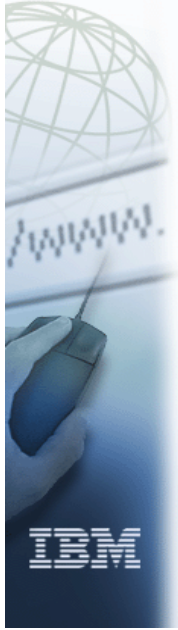


ibm.com



e-business



SNA and TCP/IP Networking Technologies in the z/OS Sysplex and for Linux on System z9 and zSeries

Introduction

Alfred B. Christensen - alfredch@us.ibm.com



Redbooks

International Technical Support Organization

© Copyright IBM Corp. 2005. All rights reserved.

Trademarks and notices

The following terms are trademarks or registered trademarks of International Business Machines Corporation in the United States or other countries or both:

- | | | | |
|----------------------|-------------------------|--------------------|-------------------|
| • AIX® | • GDDM® | • PrintWay™ | • xSeries® |
| • AnyNet® | • GDPS® | • PR/SM™ | • z/Architecture™ |
| • AS/400® | • HiperSockets™ | • pSeries® | • z/OS® |
| • Candle® | • IBM® | • RACF® | • z/VM® |
| • CICS® | • Infoprint® | • Redbooks™ | • zSeries® |
| • CICSplex® | • IMS™ | • Redbooks (logo)™ | |
| • CICS/ESA® | • IP PrintWay™ | • S/390® | |
| • DB2® | • iSeries™ | • System/390® | |
| • DB2 Connect™ | • Language Environment® | • System z9™ | |
| • DPI® | • MQSeries® | • ThinkPad® | |
| • DRDA® | • MVS™ | • Tivoli® | |
| • e business (logo)® | • MVS/ESA™ | • Tivoli (logo)® | |
| • ESCON® | • NetView® | • VM/ESA® | |
| • eServer™ | • OS/2® | • VSE/ESA™ | |
| • ECKD™ | • OS/390® | • VTAM® | |
| • FFST™ | • Parallel Sysplex® | • WebSphere® | |

- > Java and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.
- > Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.
- > Intel, Intel Inside (logos), MMX and Pentium are trademarks of Intel Corporation in the United States, other countries, or both.
- > UNIX is a registered trademark of The Open Group in the United States and other countries.
- > Linux is a trademark of Linus Torvalds in the United States, other countries, or both.
- > Red Hat is a trademark of Red Hat, Inc.
- > SUSE® LINUX Professional 9.2 from Novell®
- > Other company, product, or service names may be trademarks or service marks of others.
- > This information is for planning purposes only. The information herein is subject to change before the products described become generally available.
- > All statements regarding IBM future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

All performance data contained in this publication was obtained in the specific operating environment and under the conditions described and is presented as an illustration. Performance obtained in other operating environments may vary and customers should conduct their own testing.

Refer to www.ibm.com/legal/us for further legal information.



© Copyright IBM Corp. 2005. All rights reserved.

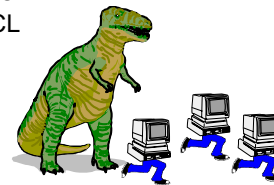
ibm.com/redbooks


Workshop objectives

➤ The overall objectives of the networking workshop day are:

- ▶ Make attendees aware, at a conceptual level, of selected new functions and capabilities in the Communications Server for z/OS V1R7.
 - Focus is on explaining concepts and where the new functions may be useful
 - Configuration principles will be covered at a conceptual level, but not in detail
 - For detailed configuration information, the attendees are referred to the product documentation

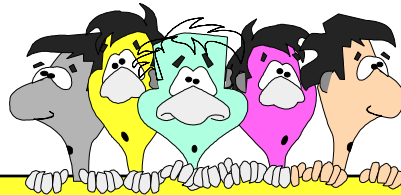
- ▶ Make attendees aware, at a conceptual level, of existing and new functions in the Communications Server for Linux on zSeries and in the Communication Controller for Linux (CCL) in zSeries
 - Provide conceptual information about the current CCL V1R1 product
 - Provide early planning information for new functions in CCL




 Redbooks © Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Workshop content



- ✓ IBM enterprise networking solutions - introduction
- ✓ CS z/OS V1R7 IP networking
 - ▶ Sysplex networking
 - ▶ zSeries and System z9
 - ▶ Applications and application enablement
 - ▶ Application transparent network security technologies
 - ▶ IPv6
 - ▶ Enterprise Extender and SNA
- ✓ SNA and IP integration in the enterprise
 - ▶ Introduction to IBM's SNA/IP integration strategy and technology
 - ▶ Distributed Communications Servers including CS Linux for zSeries and System z9 Version 6.2.1
 - ▶ Communication Controller for Linux (CCL) on zSeries and System z9 Version 1 Release 2
- ✓ A brief look at what's next

 Redbooks © Copyright IBM Corp. 2005. All rights reserved.

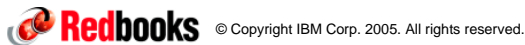
ibm.com/redbooks

Tentative time schedule for the networking day

| Time | Topic |
|---------------|--|
| 09:00 - 09:15 | Introduction |
| 09:15 - 12:00 | CS z/OS V1R7 - Sysplex, Load Balancing Advisor, Hardware exploitation, (Applications), and Integrated IP Security |
| 12:00 - 13:00 | Lunch |
| 13:00 - 15:00 | CS z/OS V1R7 - Application-Transparent SSL/TLS, (CICS Sockets), (Management), IPv6, Enterprise Extender, and SNA |
| 15:00 - 16:45 | SNA using Linux on zSeries - SNA/IP integration, Communications Server for Linux, Communication Controller for Linux |
| 16:45 - 17:00 | Wrap up |

This is a tentative schedule. Workshop-location specific requirements may change the exact timing and duration.

The handouts include many "notes" pages and some selective sections that will not all be presented. They are there for your own reading in case you need to go back and clarify something. If you see a "notes" page you want me to present, please let me know and I will do so.



ibm.com/redbooks

Practical information

A certain level of familiarity with both SNA and TCP/IP networking technologies in general and on z/OS specifically is assumed.

This is a technical update workshop.

Questions are welcome all the time.



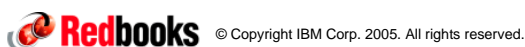
We will take frequent breaks for coffee, tea, lunch, or other personal needs.



Anything that says BEEP, BOINK, DING-DONG, or plays Beethoven's Ninth

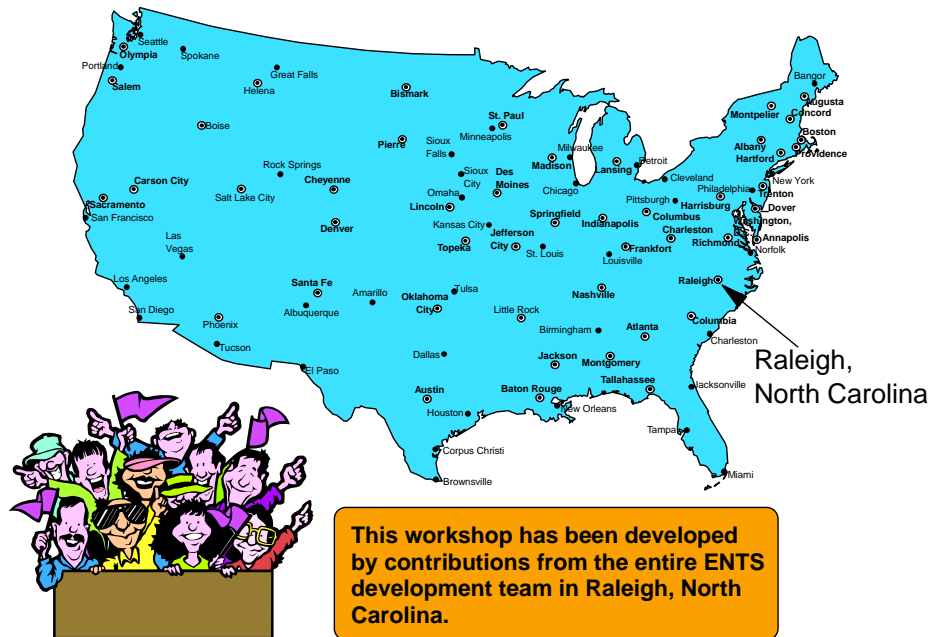



Please put phones into buzzer, vibrate, whatever non-noisy mode they support.



ibm.com/redbooks

Research Triangle Park, Raleigh, North Carolina - the home of Enterprise Networking and Transformation Solutions



 © Copyright IBM Corp. 2005. All rights reserved.


ibm.com/redbooks

Enterprise Networking and Transformation Solutions

The networking software products that are managed from a development perspective by Enterprise Networking and Transformation Solutions in Research Triangle Park, North Carolina are:

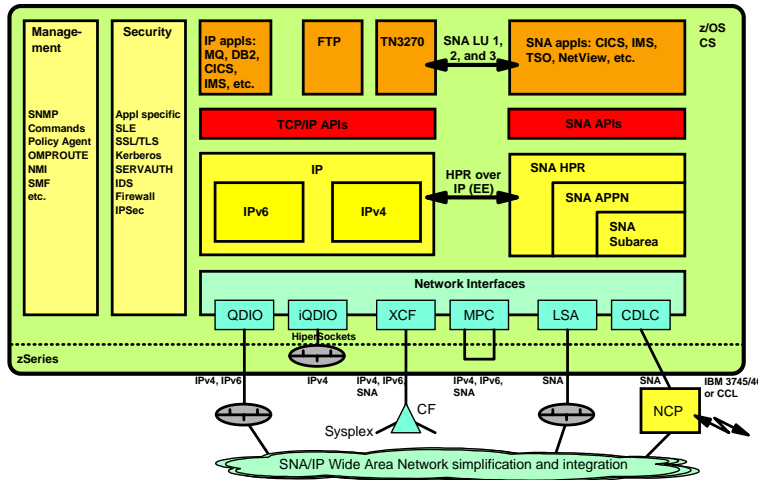
- ✓ Communications Server for z/OS - SNA and TCP/IP (CS z/OS)
- ✓ Communications Server for AIX - SNA (CS AIX)
- ✓ Communications Server for Windows - SNA (CS Windows)
- ✓ Communications Server for Linux (on Intel and Power) - SNA (CS Linux)
- ✓ Communications Server for Linux on zSeries - SNA (CSL)
- ✓ Application Workload Modeler (AWM)
- ✓ Communication Controller for Linux on zSeries and System z9 (CCL)
- ✓ WebSphere Host Access Transformation Services (HATS)
- ✓ Host On Demand (HOD)
- ✓ IBM Personal Communications (PCOM)
- ✓ Network Control Program (NCP)
- ✓ NCP Packet Switching Interface (NPSI)
- ✓ ... plus a few more ...



 © Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Communications Server on z/OS - What drives the selection of functions being added to CS z/OS?



> IPv6 support

- ▶ Staged delivery of the next generation IP network for z/OS
 - Projections are that within a few years from now, half of all nodes on Internet will be IPv6 due to a new surge in demand for IP addresses

> Middleware optimization

- ▶ Optimize the application and middleware environment on z/OS
 - Application Transparent TLS, Sysplex Load Balancing Advisor, CICS Socket Optimization, etc.

> High speed networking for zSeries

- ▶ Support for new high speed network attachment technologies
 - 10 GbE, AES Encryption, etc.

> Critical customer requirements

- ▶ Focus on critical customer requirements in the following areas:
 - Security, availability, reliability, scalability, capability and performance
 - Sysplex, EE, FTP, TN3270E, IDS, Policy, etc.

CS z/OS V1R7 - high-level overview (page 1/4)

> Sysplex networking - more options for autonomic recovery and for scalability

- ▶ Sysplex automics phase II
 - Rejoin the Sysplex group
- ▶ Deactivate/reactivate stack-managed Dynamic VIPAs (VIPADefine/VIPABackup)
- ▶ Quiesce/resume target applications for Sysplex Distributor
- ▶ Sysplex Distributor (SD) load balancing decision enhancements
- ▶ Sysplex Distributor (SD) optimized forwarding of distributed workload to target TCP/IP stacks
- ▶ z/OS Load Balancing Advisor (LBA) for outboard load balancers

> zSeries and System z9 hardware exploitation

- ▶ 10 Gigabit Ethernet Support
- ▶ QDIO OSA-Express2 Segmentation Offload (large send support)
- ▶ z9-109 IPv6 HiperSockets Support
- ▶ Dynamic VLAN registration

> CICS Sockets - revitalized for increased performance and security

- ▶ Performance enhancement - CICS Sockets tracing
- ▶ Performance enhancement - CICS monitoring
- ▶ Move TRUE to 31-bit storage
- ▶ Using the CICS Open Transaction Environment (OTE)
- ▶ SSL/TLS enabling CICS Sockets transactions through AT-TLS

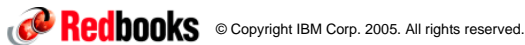
CS z/OS V1R7 - high-level overview (page 2/4)

> Standard CS z/OS server applications - new functions and enhanced security

- ▶ TN3270 server
 - Option to control use of SSL V2 or not
 - AES encryption support
- ▶ Sendmail
 - AES encryption support
- ▶ FTP
 - AES encryption support
 - Delegated RACF resource profiles for TLS FTP
 - FTP support for mixed-case RACF passwords
 - FTP JES SAPI interface changes
 - FTP client API in C programming language
 - FTP data transfer reliability feedback
 - FTP configurable end-of-line character
 - Enable/disable extended directory search

> Management - on-demand (ODI:RM) support

- ▶ Adding CIM Management in support of the on-demand Infrastructure
- ▶ SNMP UDP IPv6 MIB support
- ▶ CTRACE optimization
- ▶ Various Netstat enhancements in support of other items in this release



ibm.com/redbooks

CS z/OS V1R7 - high-level overview (page 3/4)

> Application transparent security technologies for ease of use, performance, and scalability

- ▶ Integrated IP security
 - IP filtering
 - IPSec Virtual Private Networks (VPNs)
- ▶ Application Transparent Transport Layer Security (AT-TLS)
- ▶ Policy agent support for IP security and AT-TLS policies
- ▶ z/OS network security configuration assistant GUI
- ▶ General support for mixed-case passwords

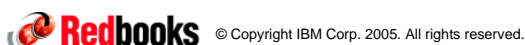
**CS z/OS
V1R7 is a
major
security
release!**

> Next generation Internet support on z/OS - we're almost done with IPv6 support on z/OS

- ▶ Advanced sockets API for IPv6 update
- ▶ Maintain 2 IPv6 routers in default list

> And let's not forget SNA - EE and SNA usability enhancements

- ▶ VTAM and Enterprise Extender display enhancements
- ▶ New DISPLAY EEDIAG command
- ▶ Model definition of VTAM cross domain resources
- ▶ Subarea VTAM XCF support
- ▶ LOGAPPL enhancements
- ▶ VTAM start option to DISPLAY RSCVs
- ▶ VTAM command to remove generic resource from a Sysplex

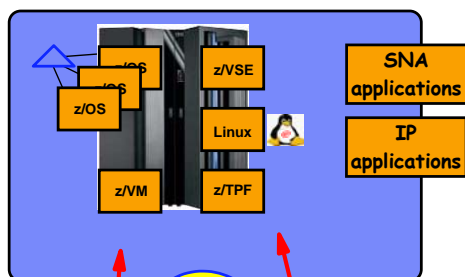


ibm.com/redbooks

CS z/OS V1R7 - high-level overview (page 4/4)

- **OROUTED is not shipped with z/OS V1R7**
 - ▶ Use OMPROUTE for both RIP V1, RIP V2, and OSPF dynamic routing protocol support
- **PAGTSNMP is not shipped with z/OS V1R7**
 - ▶ Use the new SLAPM2 MIB and supporting subagent NSLAPM2
- **z/OS V1R7 will be the last release that ships the old Firewall Technologies component**
 - ▶ Use z/OS V1R7 to migrate to the new Integrated IP Security components
- **z/OS V1R7 will be the last release that ships support for AnyNet**
 - ▶ IP over SNA: not needed any longer
 - ▶ SNA over IP: use more functionally rich and better performing technologies:
 - Enterprise Extender
 - Distributed Communications Server remote SNA API client/server technology

SNA data center access strategy



- **Protect investment in SNA applications while growing IP workloads**

- **Simplify network infrastructure**

- ▶ Consolidate the logical SNA network to the data center
- ▶ A single optimized, secure, scalable wide area network that transports IP traffic
 - Using one of many SNA over IP technologies for transporting SNA over an IP network
- ▶ A LAN environment that transports both SNA and IP traffic in the data center (and optionally in the branch if not all nodes in the branch support SNA over IP integration technologies, but require an SNA gateway)

- **Common SNA access- and application technology based on CS Linux that scales both out (to xSeries and OpenPower) and up (to System z9 and zSeries)**

- ▶ One product, one administrator interface, one skill set across the enterprise
- ▶ Developed in coordination with CS z/OS

- **Preserve the NCP capabilities through Linux on System z9 and zSeries for those who continue to rely on an NCP for remote SNA connectivity**

- ▶ SNI to business partners
- ▶ Boundary function nodes
- ▶ X.25 non-SNA access and SNA access

SNA over IP technologies that are transparent to end users or to remote SNA applications

- ▶ TN3270 (to a TN3270 server in the data center)
- ▶ HTTP/S (to a HATS server in the data center)
- ▶ Remote API (to a remote API server in the data center)
- ▶ DLSw (to a DLSw aggregation point in the data center)
- ▶ Enterprise Extender (to an EE partner in the data center)

Communications Server for Linux Version 6.2.1 - high-level overview

- Linux 2.6 kernel support
- Power platform support added
- New remote API client platforms
- Primary LU API support

Communication Controller for Linux on zSeries and System z9 (future functions) - high-level overview

- **OSN CDLC channel connectivity on System z9**
 - ▶ VTAM and TPF see the NCP as though it were channel-attached
- **NPSI X.25 support using XOT for downstream connectivity**
 - ▶ XOT Linux support to be provided by another vendor
- **Performance enhancements**
 - ▶ CCL V1R1 performance PTF
 - ▶ Additional performance enhancements
- **QDIO Layer 2 support by CCL when running on a Linux 2.6 kernel**
 - ▶ Removes requirement for OSA LCS copper ports for SNA LLC2 communication with an NCP running in a CCL environment
 - ▶ Much improved sharing of OSA resources
- **IP transmission group connectivity between two CCL NCPs**
 - ▶ Direct high-speed TCP connection between two CCL NCPs (INN or SNI traffic)



For those of
you who still
need an NCP!

CS z/OS V1R7 functions that were APARed back to earlier releases

| APAR | Back to | Description |
|---------|------------------|---|
| OA09759 | z/OS V1R4 | 10 Gigabit Ethernet support - VTAM changes |
| PQ96769 | z/OS V1R4 | 10 Gigabit Ethernet support - TCP/IP changes |
| OA11148 | z/OS V1R6 | TCP large send support - VTAM changes |
| PK02490 | z/OS V1R6 | TCP large send support - TCP/IP changes |
| OA10532 | z/OS V1R2 | OSA-Express2 support - VTAM changes |
| PQ99770 | z/OS V1R2 | OSA-Express2 support - TCP/IP changes |
| PQ90032 | z/OS V1R4 (only) | z/OS Load Balancing Advisor APAR for z/OS V1R4 |
| PQ96293 | z/OS V1R5 | z/OS Load Balancing Advisor APAR for z/OS V1R5 and V1R6 |
| | | |
| LI70764 | CCL V1R1 | CCL V1R1 support for Red Hat release 4 (RHEL4 U1) |
| LI70826 | CCL V1R1 | Performance enhancements |
| | | |

CS information in z/OS books and all CS books now unlicensed

> z/OS Migration

- ▶ Lists Communications Server function that requires you to take action to migrate to V1R7
- ▶ Approx. 15 Communications Server actions described
- ▶ This information is not provided in this format in the Communications Server library

> z/OS Summary of Messages and Interface Changes

- ▶ List all new and changed Communications Server commands, parameters, socket API changes, FTP and Telnet changes, etc.
- ▶ This information is not provided in this format in the Communications Server library

> z/OS Introduction and Release Guide

- ▶ Duplicates new function description in New Function Summary

> Five Communications Server books become unlicensed in z/OS V1R7:


- ▶ SNA Diagnosis, Vol 1: Techniques and Procedures, GC31-6850-00
- ▶ SNA Diagnosis, Vol 2: FFST Dumps and the VIT, GC31-6851-00
- ▶ SNA Customization, SC31-6854-00
- ▶ SNA Data Areas, 1, GC31-6852-00
- ▶ SNA Data Areas, 2, GC31-6853-00

z/OS Library Center

NOTES

➤ <http://www-1.ibm.com/servers/eserver/zseries/zos/bkserv/>



 © Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Some useful links to the web

NOTES

➤ APAR and ++ HOLD Documentation (containing IP and SNA info APARs) are available at the z/OS Internet library Web site at:

- <http://www.ibm.com/servers/eserver/zseries/zos/bkserv/>

➤ Both IP and SNA messages are available through the OS/390 and z/OS LookAt tool:


- <http://www.ibm.com/servers/s390/os390/bkserv/lookat/lookat.html>

➤ Where to find more information (including being able to download PDF version of all the manuals in the Communications Server library):

- <http://www.ibm.com/servers/eserver/zseries/zos/bkserv/>

➤ Communications Server home page:

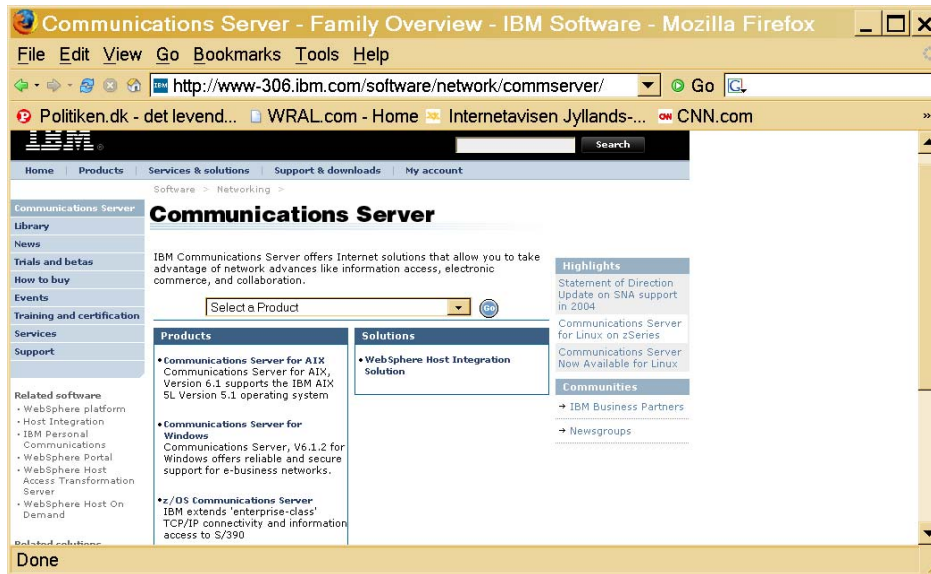
- <http://www.ibm.com/software/network/commserv>

 © Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Communications Server home page

➤ <http://www.ibm.com/software/network/commserver>



NOTES



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Communications Server for z/OS - ITSO Redbook update

- ITSO is currently running a 10-person project to create four new Communications Server for z/OS Implementation Guide Redbooks based on a CS z/OS V1R7 level.
- The objective is to replace the current 7-volume implementation guide Redbooks and the TCP/IP in a Sysplex Redbook.
- The Redbooks are targeted to be made available as drafts some time in December 2005 or early January 2006.
- The Redbook numbers are:
 - ▶ SG24-7169-00 Communications Server for z/OS Implementation, Volume 1 - Base Functions, Connectivity, and Routing
 - Brazil: Octavio Luiz Damasceno Ferreira
 - ▶ SG24-7170-00 Communications Server for z/OS Implementation, Volume 2 - Standard Applications
 - ▶ SG24-7171-00 Communications Server for z/OS Implementation, Volume 3 - High Availability, Scalability, and Performance
 - Brazil: Valirio Braga
 - ▶ SG24-7172-00 Communications Server for z/OS Implementation, Volume 4 - Security
- Note: the titles may change before the books are made generally available.



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

For more information....



**N
O
T
E
S**

| URL | Content |
|---|--|
| http://www.ibm.com/servers/eserver/zseries | IBM eServer zSeries Mainframe Servers |
| http://www.ibm.com/servers/eserver/zseries/networking | Networking: IBM zSeries Servers |
| http://www.ibm.com/servers/eserver/zseries/networking/technology.html | IBM Enterprise Servers: Networking Technologies |
| http://www.ibm.com/software/network/commserver | Communications Server product overview |
| http://www.ibm.com/software/network/commserver/zos/ | z/OS Communications Server |
| http://www.ibm.com/software/network/commserver/z_lin/ | Communications Server for Linux on zSeries |
| http://www.ibm.com/software/network/ccl | Communication Controller for Linux on zSeries |
| http://www.ibm.com/software/network/commserver/library | Communications Server products - white papers, product documentation, etc. |
| http://www.redbooks.ibm.com | ITSO Redbooks |
| http://www.ibm.com/software/network/commserver/support | Communications Server technical Support. NB: this is where you will find the downloadable configuration GUI tools. |
| http://www.ibm.com/support/techdocs/ | Technical support documentation (techdocs, flashes, presentations, white papers, etc.) |
| http://www.rfc-editor.org/rfcsearch.html | Request For Comments (RFC) |



© Copyright IBM Corp. 2005. All rights reserved.

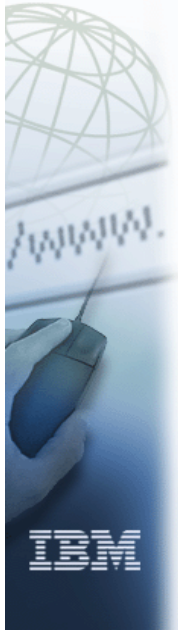
ibm.com/redbooks

This page intentionally left blank

ibm.com



e-business



Communications Server for z/OS V1R7 - Technical Update Sysplex Networking



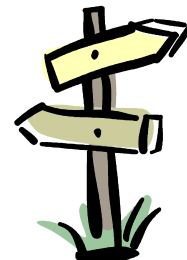
Redbooks

International Technical Support Organization

© Copyright IBM Corp. 2005. All rights reserved.

CS z/OS V1R7 Enhancements to IP Workload in a z/OS Sysplex

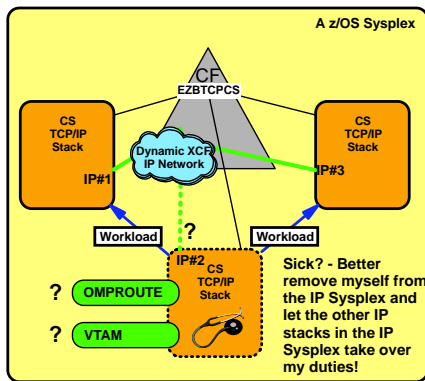
- Sysplex autonomics phase II
 - Rejoin the Sysplex group
- Deactivate/reactivate stack-managed Dynamic VIPAs (VIPADefine/VIPABackup)
- Quiesce/resume target applications for Sysplex Distributor
- Sysplex Distributor (SD) load balancing decision enhancements
- Sysplex Distributor (SD) optimized forwarding of distributed workload to target TCP/IP stacks



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Background information: TCP/IP Sysplex autonomics phase I - overview (at a z/OS V1R6 level)



- Autonomic functions to reduce single point of failure for distributed applications in a Sysplex
 - Monitor CS health indicators
 - Storage usage - CSM, TCPIP Private & ECSA
 - Monitor dependent networking functions
 - OMPROUTE availability
 - VTAM availability
 - XCF links available
 - Monitor Communications Server component-specific functions
- Monitors determine if this TCPIP stack will remove itself from the Sysplex and allow a healthy backup to take ownership of the Sysplex duties (own DVIPAs, distribute workload)
- Monitoring is always done, but configuration controls in the TCPIP Profile determine if the TCPIP stack will remove itself from the Sysplex.

The assumption is that if a TCP/IP stack determines it can no longer perform its Sysplex functions correctly, it is better for it to leave the TCP/IP XCF group and by doing so, signal the other TCP/IP stacks in the Sysplex that they are to initiate whatever recovery actions have been defined, such as moving dynamic VIPA addresses or removing application instances from distributed application groups.

```
GLOBALCONFIG SYSPLEXMONITOR TIMERSECS
seconds RECOVERY|NORECOVERY
DELAYJOIN|NODELAYJOIN
```

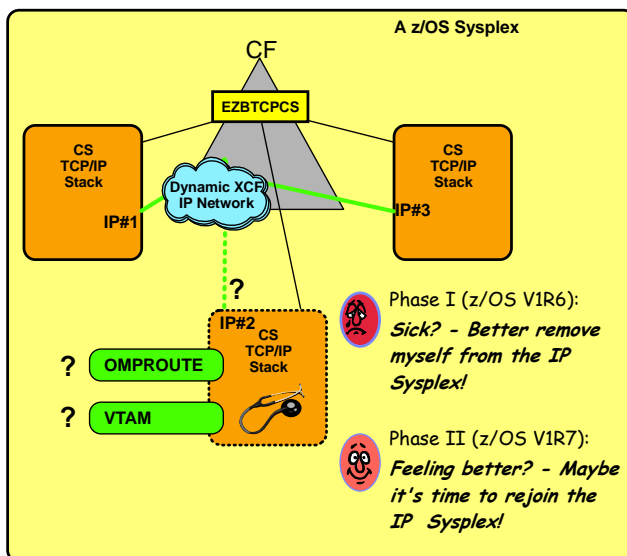
- *Timersecs* - used to determine duration of the troubling condition before issuing messages or leaving the Sysplex (if Recovery)
- *RECOVERY* - TCPIP removes itself from the Sysplex. Recommended and is the default value.
- *NORECOVERY* - TCPIP does not remove itself from the Sysplex.
- *DELAYJOIN* - Delay joining Sysplex until OMPROUTE is up
- *NODELAYJOIN* - Join Sysplex immediately



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

TCP/IP Sysplex autonomics phase II - new functions added in z/OS V1R7 to rejoin an IP Sysplex



➤ **z/OS V1R7 adds the following functions to the TCP/IP Sysplex autonomics:**

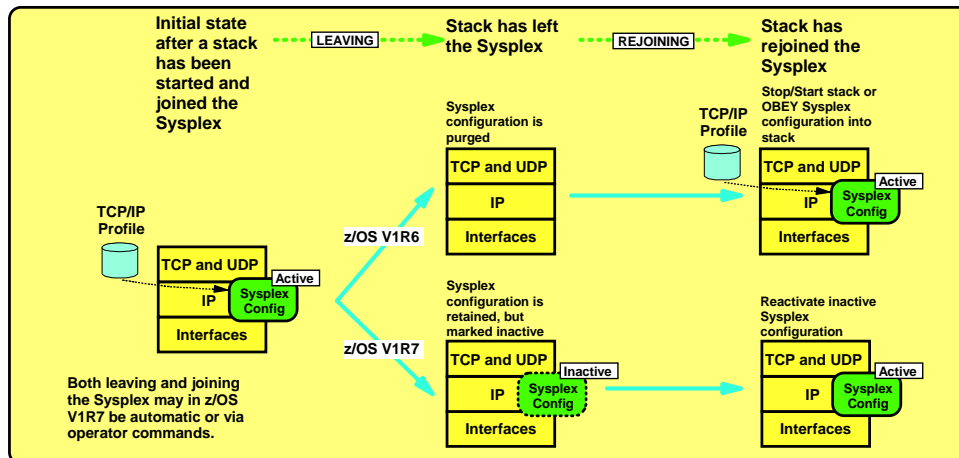
- Retain the current Sysplex configuration data in an inactive state when a stack leaves the Sysplex
- Reactivate the currently inactive Sysplex configuration when a stack rejoins the Sysplex
- New options for rejoining the Sysplex:
 - Via an operator command
 - Automatically when the error condition that caused the stack to leave the Sysplex has been cleared



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

TCP/IP stacks leaving and rejoining the Sysplex

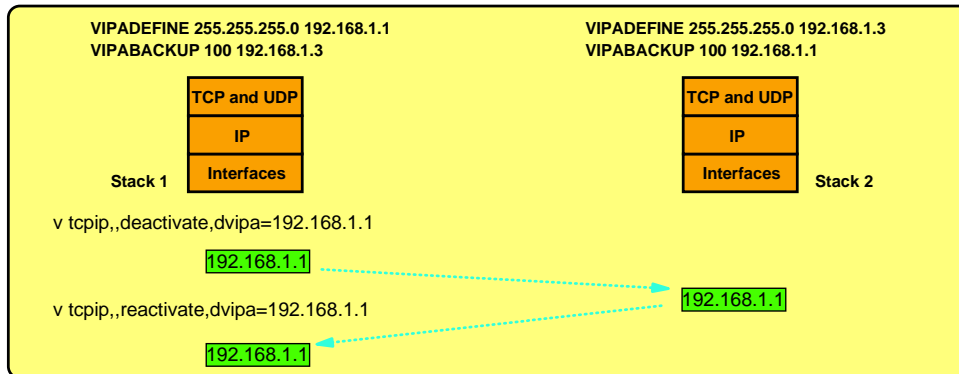


- **Leaving the Sysplex in z/OS V1R6, purges Sysplex configuration data from the stack's internal configuration blocks.**
 - To rejoin the Sysplex, the Sysplex configuration data must be reapplied to the stack's active configuration through a restart or an OBEY command
- **In z/OS V1R7, a stack's Sysplex configuration data will be retained in an inactive status when a stack leaves the Sysplex**
 - The inactive Sysplex configuration data will be shown on the NETSTAT VIPADCFG report as inactive
 - Rejoining the Sysplex will then reactivate the currently inactive Sysplex configuration data

Deactivate/reactivate stack-managed Dynamic VIPAs (VIPADefine / VIPABackup)

- **Types of dynamic VIPA addresses:**
 - Stack-managed
 - Defined through VIPADefine and VIPABackup
 - All stack-managed DVIPAs are activated/deactivated when a stack joins/leaves the Sysplex
 - Event-managed
 - Defined through VIPARANGE
 - Individual DVIPAs are activated/deactivated when an application binds to one, or a MODDVIPA command is issued against one, or an application issues an IOCTL sockets call for one
 - Distributed DVIPAs
 - Defined on distributing (owning) stack through VIPADefine/VIPABackup
 - Activated/deactivated on owning stack as other stack-managed DVIPAs
 - Defined on target stacks through VIPADISTRIBUTE statement on implicitly by distributing stack
- **Event-managed DVIPAs can be moved around the Sysplex individually based on one of the events listed above**
- **Stack-managed DVIPAs (all of them) can only be moved when a stack leaves or joins the Sysplex**
 - There is no mechanism to request movement of an individual stack-managed DVIPA except through dynamic configuration changes - OBEYFILE processing
- **z/OS V1R7 implements a new operator command to request movement of individual stack-managed DVIPAs**

Improved operations: operator-initiated movement of individual stack-managed dynamic VIPA addresses



> Deactivate

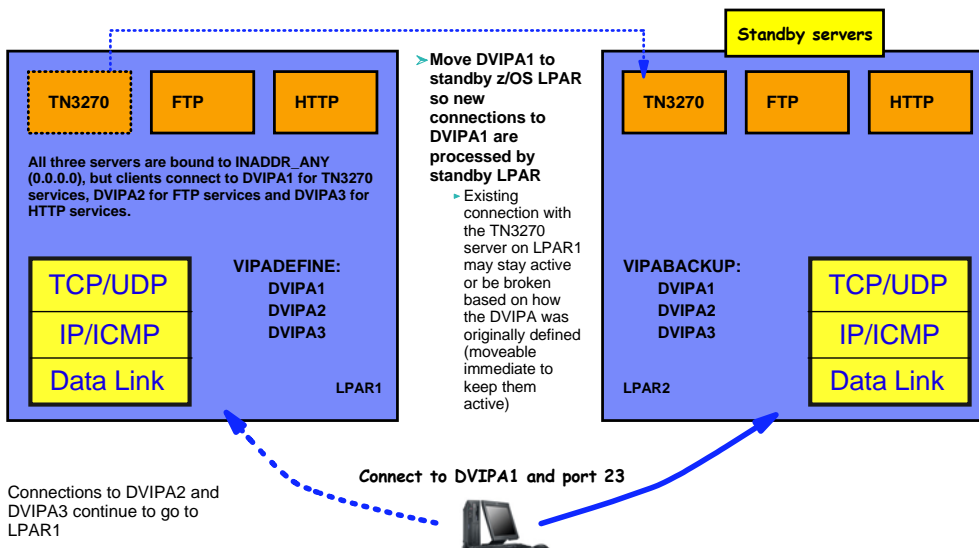
- DVIPA is deactivated and a configured backup stack will takeover the DVIPA
- Backup DVIPA can be deactivated also removing eligibility as a backup

> Reactivate

- Original owner can regain ownership
- Can also reactivate a backup DVIPA that's been deactivated
- Prior to these commands, Vary OBEY files were needed to cause a DVIPA takeover
- These commands can't be used on a DVIPA that was created from VIPARANGE with bind, ioctl(), or the MODDVIPA utility

Scenario where deactivating/reactivating stack managed DVIPAs may be useful

- > If we wanted to move all three services together to the standby LPAR, we could just remove LPAR1 from the IP Sysplex.
 - V TCPIP,,SYSPLEX,LEAVEGROUP
- > If we want to move just one of the three services, we will use the new DVIPA deactivate command to do so.
 - V TCPIP,,DEACTIVATE,DVIPA=DVIPA1



Sysplex Distributor: operator-initiated quiesce and resume of individual target server applications or full target systems

➤ Ability to quiesce a target system or an application instance prior to shutdown

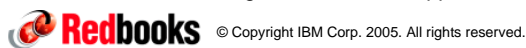
- ▶ Planned maintenance scenarios of system or application
 - Allows existing systems or applications to drain work queue prior to shutdown
- ▶ Relieve temporary constraints of resources on target system
- ▶ Temporary - Does not affect Sysplex Distributor's permanent configuration
- ▶ Issued on target system being affected
- ▶ Can also be used to control individual server applications in a SHAREPORT group
- ▶ Only way to achieve similar capability earlier was via temporary configuration changes based on OBEYFILE commands

➤ VARY TCPIP,,SYSPLEX,QUIESCE,options

- ▶ TARGET - Quiesces all applications on target stack.
- ▶ PORT=xxx - Quiesce all applications bound to the specified port on this stack
 - JOBNAME=jobname - Allows quiesce of a single application in SHAREPORT group
 - ASID=asid - Further qualify job being quiesced (such as when dealing with duplicate jobnames)
- ▶ No new TCP connections sent to the quiesced target (stack or application)
 - For all Distributed DVIPAs that the entity is a target for
- ▶ Existing TCP connections are maintained (or in other words, the process is non-disruptive)

➤ VARY TCPIP,,SYSPLEX,RESUME,options

- ▶ TARGET|PORT|JOBNAME|ASID
- ▶ Allows identified target stacks and/or applications to once again be targets for distribution



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Improved IP workload distribution quality focus in z/OS V1R7

➤ Sysplex Distributor uses server-specific WLM Interfaces to determine if target server is meeting its goal

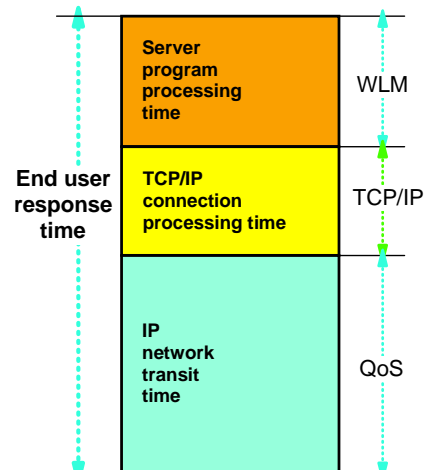
- ▶ Extracts WLM recommendations for each distributed server to determine which server(s) get new connections
- ▶ More precise than existing WLM method, which uses recommendations based upon displaceable capacity of the system
- ▶ Can also be used for SHAREPORT balancing

➤ Sysplex Distributor will detect target server unresponsiveness

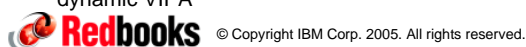
- ▶ Target stacks push key TCP/IP "health" statistics for target application(s) to distributor, such as number of connections dropped due to backlog
- ▶ When load balancing, the distributor uses these indicators along with values for WLM and QoS to determine which stack gets the connection
- ▶ Strengthens overall evaluation of a server's health

➤ Allow Sysplex Distributor to route over any available route

- ▶ Removes the need for Sysplex Distributor to route only over coupling facility links
- ▶ Allows for use of high-speed links such as Ethernet and OSA Express QDIO interfaces
- ▶ Can use any interface on the target except cannot specify a dynamic VIPA



Addresses some storm-drain scenarios, but not all.



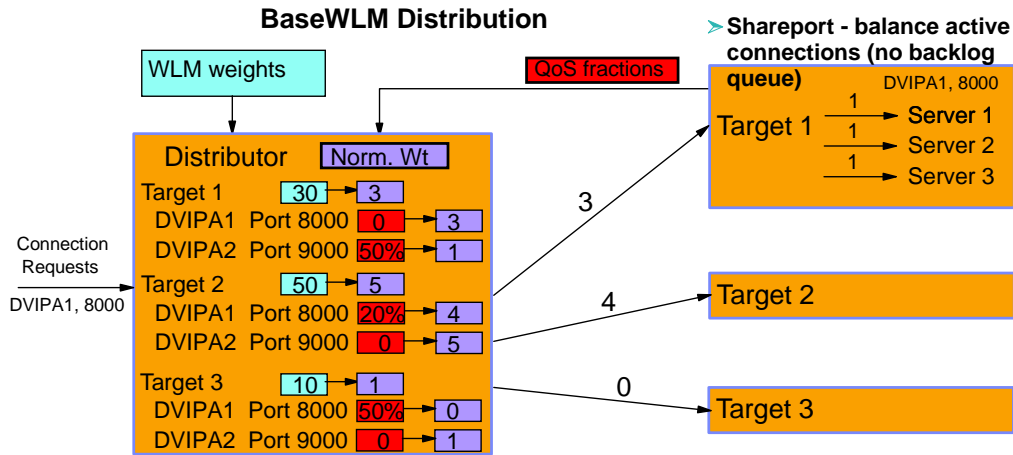
© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Calculating Sysplex Distributor weights before z/OS V1R7

NOTES

- > Using the Sysplex Distribution function, incoming connections for a DVIPA and port are distributed to multiple target stacks. Target selection is determined using
 - RoundRobin
 - BaseWLM - capacity recommendations from WLM (weights) for each system
 - Weights are normalized - optionally modified with policy information (QoS fractions) from each target



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Background notes

NOTES

- > When determining a BaseWLM weight, WLM assigns a relative weight to each system in the sysplex with the highest weight going to the system with the most available CPU capacity. The weights range between 0 & 64. If all systems in the Sysplex are running at or near 100% utilization, WLM will assign the highest weights to the systems with the largest amounts of lower importance work. In this way, new connection requests will be distributed to the systems with the highest displaceable capacity.
- > Normalizing and determining the QoS modified WLM weight
 - WLM weights are normalized - the WLM weights range in value from 1 to 64. These returned system weights are divided by the smallest system weight. For example, if BaseWLM system weights of 50, 30, and 10 are returned, the normalized weights are 5, 3, and 1.
 - A QoS Service level fraction is received from the target for each group of connections that map to a DVIPA/PORT for that service level. The fraction represents the performance of this group of connections. This is based on maximum connection limit for the service level, the target-to-client performance (ratio of retransmits and timeouts to number of packets sent, overall throughput and throughput/connection against desired values) - the lower the fraction, the better the performance.
 - The normalized WLM weight is reduced by the QoS Fraction percentage. For example if the normalized WLM weight is 5 and the QoS Fraction is 20%, the modified weight is 4 (5 - (5 * 20%).)
- > Distribution of connections to DVIPA1, Port 8000
 - Connections come in destined for DVIPA1, Port 8000.
 - Based on the QoS modified WLM weights for this DVIPA/Port and service level, as 7 connection requests are received, 3 connections are distributed to Target 1 and 4 connections to Target 2.
 - Target 1 is configured with SHAREPORT for Port 8000. Connections are evenly distributed among the servers that have no backlog queue such that each server has the same number of active connections.



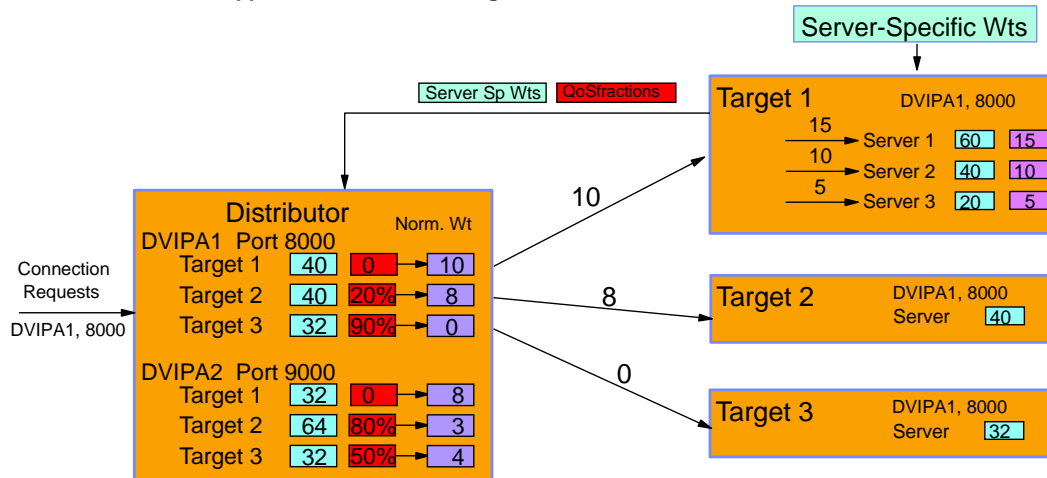
© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Enhanced workload distribution based on server-specific WLM weights

NOTES

- > WLM will assign a weight based on:
 - ▶ How well a server is meeting the goals of its service class.
 - ▶ The displaceable capacity for new work based on the importance of its service class.
- > Server-specific WLM weights are received for each server at the target.
- > The SHAREPORT distribution algorithm will also be able to use the server-specific weight.
- > QoS fractions are applied before normalizing.



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Determining the normalized WLM weight

NOTES

- > Server-specific weights are requested for Sysplex Distributor when DISTMETHOD SERVERWLM is specified.
- > Server-specific weights are requested for a SHAREPORT group when SHAREPORTWLM is specified.
- > A Server-specific weight is sent from the target to the distributor for each DVIPA/Port. In the case of multiple Shareport Servers, an average weight is sent to the distributor.
- > Determining the QoS modified WLM weight and normalizing - to preserve more of the distinctions between different weights, the QoS fraction is applied to the raw WLM weight before normalizing, and the normalization algorithm is changed. From the previous chart, the weight for Target 2's DVIPA1, Port 8000 Server is calculated as follows:
 - ▶ The QoS Service level fraction is applied against the raw WLM weight before the WLM weight is normalized. The WLM weight is 40 and the QoS fraction is 20%, so the QoS modified WLM weight is 32 ($40 - (40 * 20\%)$)
 - ▶ The normalized weight is 8 - determined by dividing by 4.
 - ▶ The exception to this would be if all of the received WLM weights associated with a DVIPA/Port were less than or equal to 16. In that case normalization is not done. After the QoS fraction is applied against the raw weight, the weights are left unchanged.
 - ▶ To change a server weight, WLM depends on the server receiving work. So even if a server weight is zero, a connection request will still be forwarded infrequently to that server to generate new WLM values.
- > Distribution of connections to DVIPA1, Port 8000
 - ▶ Connections come in destined for DVIPA1, Port 8000.
 - ▶ Based on the QoS modified WLM weights for this DVIPA/Port and service level, as 18 connection requests are received, 10 connections are distributed to Target 1 and 8 connections to Target 2.
 - ▶ Target 1 is configured with SHAREPORT for Port 8000. As 30 connections are received, 15 will be distributed to server 1, 10 to server 2, and 5 to server 3.

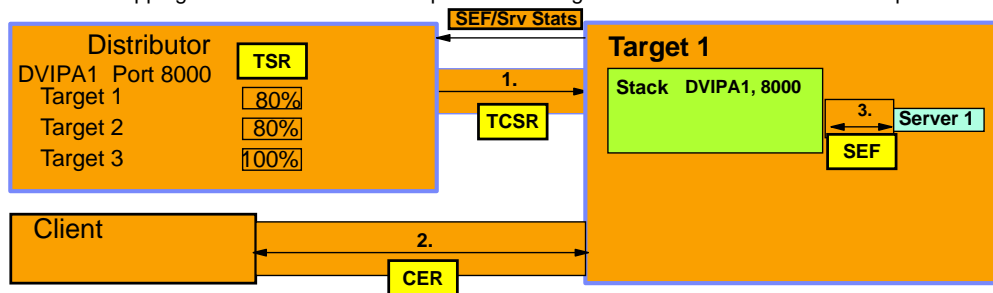


© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Sysplex autonomies health monitor for target stacks

- **TCSR - Target Connectivity Success Rate**
 - ▶ Monitoring connectivity between the distributing stack and the target stack - are the new connection requests reaching the target?
- **CER - Connection Establishment Rate**
 - ▶ Monitor network connectivity between server and client - are new connections being established?
- **SEF - Server accept Efficiency Fraction**
 - ▶ Monitor Target Server responsiveness - is the server accepting new work?
- **TSR - Target Server Responsive fraction**
 - ▶ The target sends SEF values and server statistics to the distributor which creates a Target Server Responsiveness Fraction (TSR) based on the TCSR and SEF (which includes CER).
- **All values are expected to be 100 unless there is a problem.**
 - ▶ TCSR dropping to 25 or lower will drive optimized routing function to do a new route lookup.



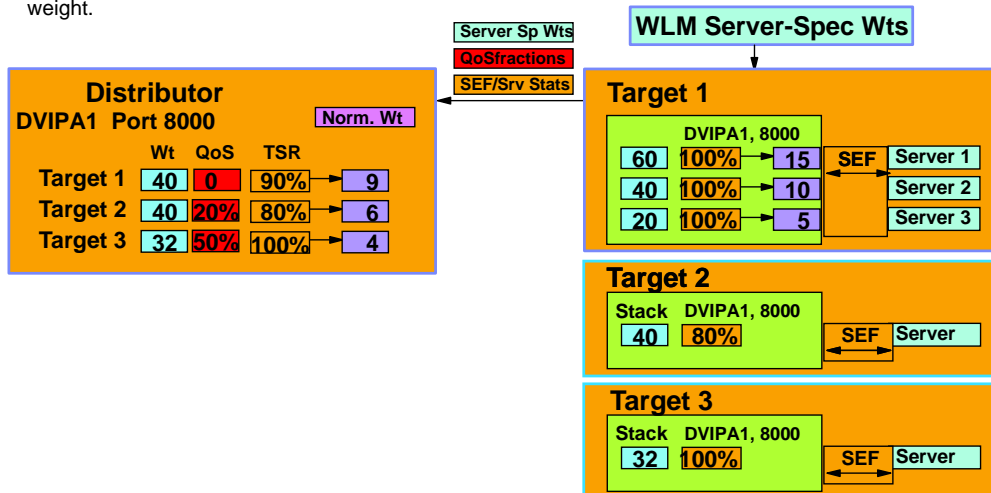
TSR calculations

NOTES

- TCSR - The distributor determines this value from the number of SYN segments it has sent to the target and the statistics returned from the target
- SEF - This value is based on whether the server is processing new connections
 - ▶ New connections are being established - Connection Establishment Rate (CER)
 - ▶ The server is accepting the new connections
- TSR - Based on SEF value (which includes CER) and TCSR value

Server-specific WLM with Sysplex autonomics health monitor for target stacks

- The distributor receives WLM server-specific weights, QoS fractions, SEF fractions, and server statistics from the targets for each server.
- The distributor calculates a normalized weight from the raw server-specific weight, QoS fraction, and TSR fraction.
- If SHAREPORTWLM is being used, the target will use the SEF fraction and apply it to the server-specific weight.



Determining the normalized WLM weight

NOTES

- From the previous chart, the weight for Target 2's DVIPA1, Port 8000 Server is calculated as follows:
 - The QoS Service level fraction is applied against the raw WLM Server weight. For example if the WLM weight is 40 and the QoS fraction is 20%, the QoS modified WLM weight is 32 ($40 * 20\%$).
 - A TSR fraction is calculated from the SEF value and server statistics that are received from the target and from information that the distributor keeps for each server. A higher fraction means a healthier server. So if the QoS modified WLM weight is 32, and the Server fraction is 80%, the new modified weight is 25 ($32 * 80\%$).
 - Finally the normalized weight of 6 is determined by dividing by 4.
- SHAREPORTWLM distribution
 - The target calculates an SEF from the statistics for each SHAREPORT server. At the target, the fractions are applied against the raw server weights and normalized (as above).
 - The average of the SEF values and statistics is sent to the distributor.
 - The average of the raw server weights is sent to the distributor.
- SHAREPORT distribution - if the existing SHAREPORT parameter is used, distribution is changed to use the SEF value alone. The SEF value is applied against an assumed raw weight of 64 (the highest weight) and a normalized weight is calculated as above. It is no longer based on balancing the number of active connections among the servers.

Netstat VIPA destination port report with the detail option

> Netstat VDPT/O DETAIL - display TCSR, CER, and SEF values

Long Format:

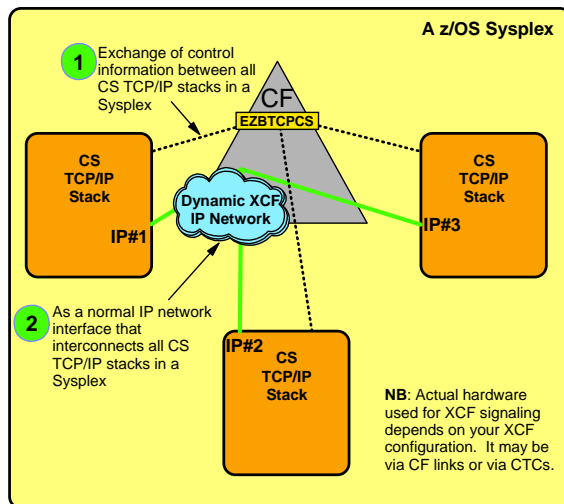
```
MVS TCP/IP NETSTAT CS V1R7          TCPIP Name: TCPCS
Dynamic VIPA Destination Port Table:
Dest:          201.2.10.11..8000
  DestXCF:     193.9.200.1
  TotalConn:  0000000050  Rdy: 001  WLM: 15  TSR: 100
  Flg: ServerWLM
  TCSR: 100 CER: 100 SEF: 100
  QoSPlcAct:  *DEFAULT*
    W/Q: 15
  QoSPlcAct:  Gold-Service
    W/Q: 10
Dest:          201.2.10.11..8000
  DestXCF:     193.9.200.2
  TotalConn:  0000000050  Rdy: 001  WLM: 15  TSR: 100
  Flg: ServerWLM
  TCSR: 100 CER: 100 SEF: 100
  QoSPlcAct:  *DEFAULT*
    W/Q: 15
  QoSPlcAct:  Gold-Service
    W/Q: 15
```

Things to think about when enabling SERVERWLM

- > **WLM server recommendations can only be used if the Sysplex Distributor and all target stacks for a distributed DVIPA port are V1R7 or later.**
 - If all targets for a DVIPA do not provide server-specific weights, then BASEWLM will be used.
- > **Target server responsiveness values can only be used if the Sysplex Distributor and all target stacks for a distributed DVIPA port are V1R7 or later.**
 - TSR values can be used with BASEWLM or SERVERWLM weights.

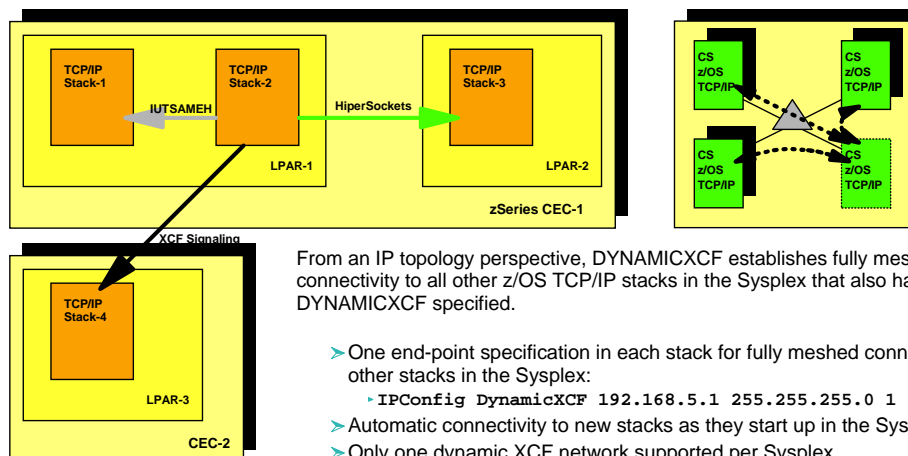
Background information: z/OS TCP/IP requires use of XCF signaling, but what is it used for?

XCF signaling is used for two purposes:



1. When a CS TCP/IP stack with the SYSPLEXROUTING option enabled in IPCONFIG starts in a Sysplex, the stack always joins a predefined XCF group (named EZBTCPCS). This group is used by all CS TCP/IP stacks in the same Sysplex to exchange control information over, such as which IP addresses each stack has in its HOME list and event notification when an IP address is added or deleted. This group is also the group that is used to keep track of which stacks are up and running, so that a stack that is defined as VIPABACKUP for a VIPA address that is active on a stack that goes down can take over the address at the point in time the first stack goes down. There are no configuration controls to enable or disable this use of XCF.
2. XCF can optionally also be used as an IP network interface over which CS TCP/IP stacks can send IP packets to each other. This use is under configuration control and can be defined using either static XCF links or allowing all stacks to join an IP XCF network dynamically (DYNAMICXCF). If one uses Sysplex Distributor or Non-disruptive Dynamic VIPA movement functions in a Sysplex, then dynamic XCF must be enabled.

Background information: is XCF signaling always used for the DYNAMICXCF IP network?



From an IP topology perspective, DYNAMICXCF establishes fully meshed IP connectivity to all other z/OS TCP/IP stacks in the Sysplex that also have DYNAMICXCF specified.

> Dynamic XCF network usage:

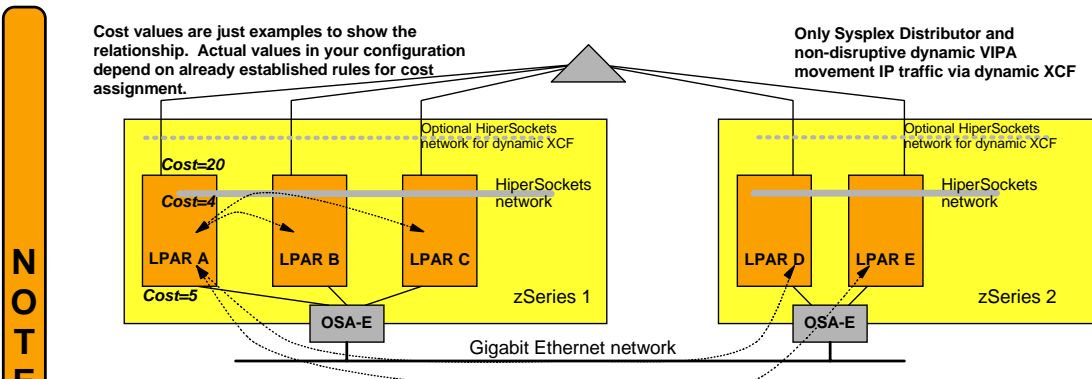
- > Sysplex Distributor and non-disruptive DVIPA movement forwarding of packets to target stack
- > General IP routing between the stacks in the Sysplex

- > One end-point specification in each stack for fully meshed connectivity to all other stacks in the Sysplex:
 - IPConfig DynamicXCF 192.168.5.1 255.255.255.0 1
- > Automatic connectivity to new stacks as they start up in the Sysplex
- > Only one dynamic XCF network supported per Sysplex

Under-the-covers DYNAMICXCF will choose one of three transport technologies depending on availability and location of partner stack:

- > Inside same LPAR: IUTSAMEH (memory-link inside a z/OS system)
- > Inside same zSeries CEC: HiperSockets (if enabled for that purpose via the IQDCHPID VTAM start option)
- > Outside CEC: XCF signaling

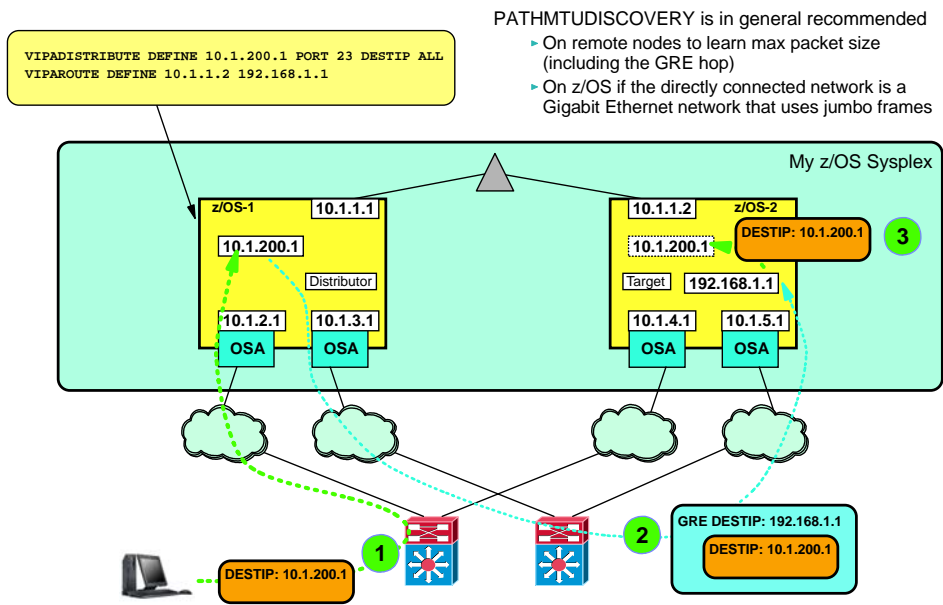
Background information: Guidelines for how to control use of the DynamicXCF IP network for general IP routing



NOTES

- > Objective:
 - > Only use dynamic XCF network for the purposes where it at this point in time is required: Sysplex Distributor and non-disruptive dynamic VIPA movement
 - > Use a HiperSockets network for IP communication between LPARs in the same CEC
 - > Use a Gigabit Ethernet infrastructure for IP communication between LPARs in different CECs
- > Define the dynamic XCF network with a rather high routing cost so it will not be used for normal IP routing unless it is the only interface that is available - or define it is a non-OSPF interface (recommended).
- > Define in each CEC a second HiperSockets network (through DEVICE/LINK definitions that interconnects all LPARs in that same CEC) - and use a low routing cost
- > Define Gigabit Ethernet connectivity from all LPARs and use a low routing cost (at least one higher than the HiperSockets network)

SD and non-disruptive DVIPA movement forwarding via non-DynamicXCF interfaces



Details on use of VIPAROUTE

NOTES

- > **Controlled by a new VIPAROUTE statement**
 - Indicates which IP address (target_ipaddr) on the target stack is to be used as the destination IP address during the route lookup selection
 - Used to select a route from a distributing stack to a target stack.
 - Used for distribution to all DVIPAs for which a matching dynamic XCF address, or ALL, was specified on a VIPADISTRIBUTE statement.
 - Used by TAKEOVER stack to forward packets for existing connections to a stack that previously owned the DVIPA (DVIPA in MOVING status).
- > **Allows optimal interface to each target, for example:**
 - IUTSAMEH within same MVS image
 - HiperSockets within same CEC
 - OSA Express Gigabit Ethernet between CECs
- > **Other considerations**
 - Multipath routing will be supported.
 - Multipath routes used on a per connection basis if the stack has been configured to use any kind of multipath (per connection or per packet).
 - If IP routing tables have changed or Target Connectivity Success Rate (TCSR) is low
 - Sysplex Distributor will perform a new route lookup to retrieve the current best route approximately every 60 seconds.
 - IPv4 uses GRE encapsulation
 - IPv6 uses an outer IP header
- > **When a connection from the client needs to be processed by Sysplex Distributor, it will determine if a matching VIPAROUTE statement has been specified or not. If it has, the best available route will be determined using the normal IP routing tables. If no matching VIPAROUTE statement exists for that target, IP packets distributed by Sysplex Distributor to that target will use Dynamic XCF interfaces.**
- > **When a VIPAROUTE statement is in effect, packets are sent from the distributor to the target encapsulated in either a GRE wrapper (IPv4) or an IPv6 header. The outer IP header will contain the VIPAROUTE target IP address as its destination IP address and the distributor's dynamic XCF address as the source IP address.**
- > **Generic Routing Encapsulation (GRE) is a standard protocol described by RFC1701. GRE allows a wrapper to be placed around a packet during transmission of the data. A receiving stack that supports GRE will remove the GRE wrapper, allowing the original packet to be processed by the receiving stack. This is often used to deliver a packet to a stack using an alternate destination IP address. For more information regarding GRE, please refer to RFC1701.**
- > **For a pre-V1R7 backup stack, the stack will not be able to process the VIPAROUTE statement.**
- > **For a pre-V1R7 target stack, all IP packets distributed from the routing stack have to be sent over the Dynamic XCF interfaces.**
- > **It is therefore strongly recommended that all TCP/IP stacks participating in VIPAROUTE distribution must be at least z/OS V1R7.**

Performance impacts of optimized Sysplex routing

> Streams workload - remote get processing (getting a file from z/OS)

| Connectivity | Trans / Second | Trans/Sec Delta % | CPU / Tran (SysDist) | CPU/Tran Delta % (Sys Dist) | CPU / Tran (Targets) | CPU/Tran Delta % (Targets) |
|--------------|----------------|-------------------|----------------------|-----------------------------|----------------------|----------------------------|
| XCF | 3.0191 | Base | 82410 | Base | 89100 | Base |
| OSAE-GbE | 2.9480 | - 2.4 % | 61190 | - 25.7 % | 75510 | - 15.3 % |
| IQDIO | 3.1650 | + 4.8 % | 71790 | - 12.9 % | 86890 | - 2.5 % |

> Streams workload - remote put processing (moving a file to z/OS)

| Connectivity | Trans / Second | Trans/Sec Delta % | CPU / Tran (SysDist) | CPU/Tran Delta % (Sys Dist) | CPU / Tran (Targets) | CPU/Tran Delta % (Targets) |
|--------------|----------------|-------------------|----------------------|-----------------------------|----------------------|----------------------------|
| XCF | 0.9108 | Base | 305700 | Base | 267000 | Base |
| OSAE-GbE | 2.6358 | + 189.4 % | 223000 | - 27.1 % | 142900 | - 46.5 % |
| IQDIO | 2.6505 | + 191.0 % | 209500 | - 31.5 % | 144700 | - 45.8 % |

Performance impacts of optimized Sysplex routing (*continued*)

➤ Transactional workload - connect, request, response, close (CRR)

| Connect-ivity | Trans / Second | Trans/Sec Delta % | CPU / Tran (SysDist) | CPU/Tran Delta % (Sys Dist) | CPU / Tran (Targets) | CPU/Tran Delta % (Targets) |
|---------------|----------------|-------------------|----------------------|-----------------------------|----------------------|----------------------------|
| XCF | 0.9108 | Base | 305700 | Base | 267000 | Base |
| OSAE-GbE | 2.6358 | + 189.4 % | 223000 | - 27.1 % | 142900 | - 46.5 % |
| IQDIO | 2.6505 | + 191.0 % | 209500 | - 31.5 % | 144700 | - 45.8 % |

Things to think about when enabling VIPAROUTE

➤ DynamicXCF must still be defined in z/OS V1R7:

- ▶ Target address for VIPADISTRIBUTE definitions is dynamic XCF IP address of target stacks
- ▶ Some workload will still be routed via DynamicXCF:
 - Sysplex Wide Security Association (IPSec) packets
 - Multi Level Security (MLS) tagged packets
 - Policy Agent QoS performance data collection
- ▶ To minimize XCF signalling, use HiperSockets for same-CEC DynamicXCF
 - Actual XCF (CF-links) will only be used for cross-CEC communication

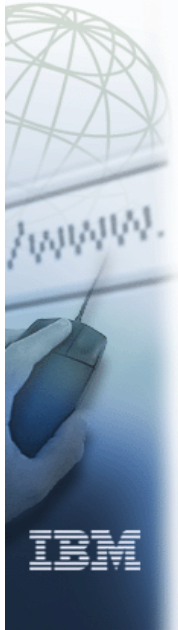
➤ Applications using the SO_CLUSTERCONNTYPE option on the GETSOCKOPT socket API

- ▶ Applications exploiting this internal indicator should continue to function properly from a communications perspective, but they may no longer optimize their processing when the destination address being used is a Dynamic VIPA or a Distributed Dynamic VIPA.
- ▶ If you have applications that exploit this socket option with Dynamic VIPAs or Distributed Dynamic VIPAs, you should consider modifying the configuration to use Static VIPAs as the destination addresses.

ibm.com



e-business



Communications Server for z/OS V1R7 - Technical Update z/OS Load Balancing Advisor



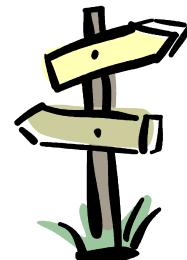
Redbooks

International Technical Support Organization

© Copyright IBM Corp. 2005. All rights reserved.

z/OS Load Balancing Advisor (LBA) - agenda

- > Background
- > The Server Application State Protocol (SASP)
- > How to implement the z/OS Load Balancing Advisor (LBA)
- > Recovery scenarios
- > Migration and coexistence
- > Cisco CSM configuration sample



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Why do we need the z/OS Load Balancing Advisor (LBA)?

➤ Let's assume that you have selected an external IP load balancing solution for your z/OS Sysplex environment

➤ Some possible reasons:

- Prefer to have a single load balancing solution across multiple platforms in your environment
- Administration of the load balancing functions belongs to network administration domain (not z/OS administrators)
- Requirements for content-based load balancing
 - Need to perform load balancing/routing decisions based on data content (inspection of URL, session IDs, cookies, etc.)
 - This is often combined with SSL offloading functions - need to decrypt data prior to inspection

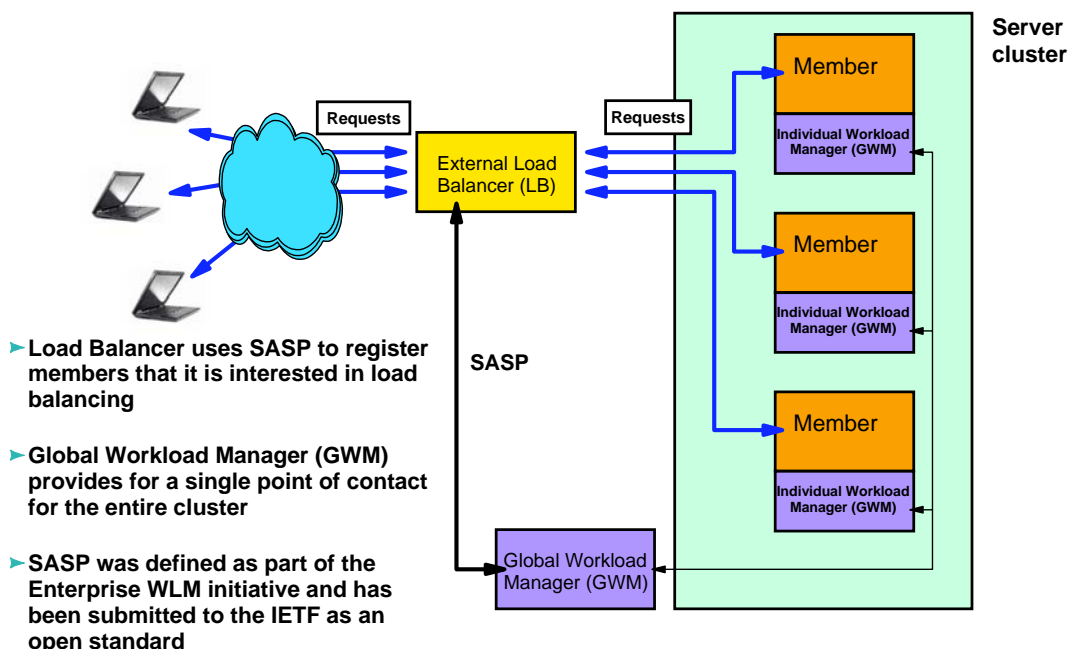
➤ But what if the external load balancing solution has no awareness of the Sysplex environment?

- Is the Sysplex treated just like any other server cluster?
- Is it aware of the current/changing workload conditions on the various systems in the Sysplex cluster?
- Is it aware of the health and status of applications and/or systems?

➤ Making the external load balancing solution "Sysplex aware" can help answer many of these questions

- The z/OS Load Balancing Advisor is a key component that allows any external load balancing solution to become "Sysplex aware".
 - The external load balancer needs to support the Server Application State Protocol (SASP) to obtain Sysplex information to use in its load balancing decisions.

Server Application State Protocol (SASP) - architecture



Load balancer registrations through SASP

➤ The load balancer may register two types of groups for which it wants weights:

- ▶ A system group
 - Represented by a list of IP addresses only.
 - IP addresses are matched to TCP/IP stacks in the Sysplex.
 - WLM weights for the LPARs are retrieved.
 - CS weight indicates if IP address is active in the Sysplex or not
 - 0 means quiesced
 - 100 means not quiesced
 - LBA displays will show a protocol value of zero for system group registrations.
- ▶ An application group
 - Represented by a list of IP addresses, protocols (TCP or UDP), and ports.
 - Server address spaces are matched to registrations.
 - WLM weights for the LPARs are retrieved.
 - CS weights are calculated factoring in how well the server instances are performing.
 - LBA displays will show protocol as TCP or UDP with the registered port numbers

➤ When an external load balancer connects to a global workload manager, it instructs the manager how it wants weights presented:

- ▶ The load balancer will poll every so often to obtain the current weights
- ▶ The load balancer requests the advisor to push weights down at certain intervals or when the weights change
 - This is how a Cisco CSM external load balancer behaves

SASP update frequency

➤ SASP supports both a "push" and a "pull" model for updating the load balancer with workload recommendations

- ▶ Support of either by the load balancer is implementation dependent
- ▶ Load balancer tells GWM which model it wants to use
- ▶ "Pull" model
 - GWM "suggests" a polling interval to the load balancer
 - z/OS Load Balancing Advisor uses the configurable update_interval value for this purpose
 - Load balancer has the option to ignore this value
 - Load balancer requests updates each polling interval
- ▶ "Push" model
 - GWM sends updated information to the load balancer on an interval basis
 - z/OS Load Balancing Advisor uses the configurable update_interval value for this purpose
 - GWM may send data more frequently than the interval period
- ▶ Load balancer determines whether it wants information about all members it registered or only changed information about its registered members

Products that support the SASP protocol (mid-2005)

> SASP Global Workload Managers (GWMs)

- ▶ EWLM (Enterprise Workload Manager)
 - Part of IBM Virtualization Engine 1.0
 - Supported platforms:
 - IBM AIX 5L Version 5.2
 - Microsoft Windows 2000 Advanced Server, 2000 Server, 2003 Enterprise Edition, 2003 Standard Edition
 - Sun Microsystems Solaris 8 (SPARC Platform Edition), 9 (SPARC Platform Edition)
 - Linux on zSeries and System z9 for next release of IBM Virtualization Engine
- ▶ z/OS Load Balancing Advisor
 - Part of z/OS Communications Server (z/OS V1R4 and higher)

> Load Balancers

- ▶ Cisco Content Switching Module (CSM) level 4.1 (2.5)
 - Uses the "Push" model
- ▶ Cisco Content Services Switch (CSS)
 - Being tested with z/OS LBA
- ▶ Other vendors likely in the future

z/OS Load Balancing Advisor (LBA) to improve quality of load balancing decisions made by outboard load balancers

The Server Application State Protocol (SASP) control flows will provide relative weights per server instance (based on WLM weight, server availability, and server processing health taking such metrics as dropped connections, size of backlog queue, etc. into consideration).

SASP is also used by the new eWLM infrastructure.



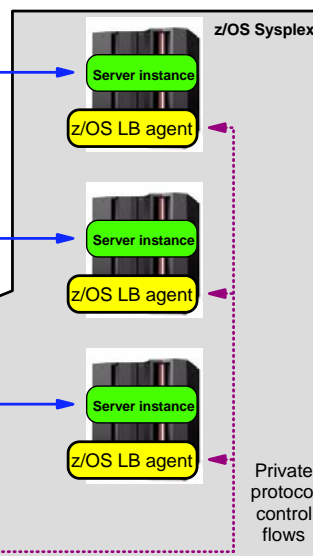
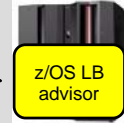
Work requests

Cisco CSM currently supports the SASP protocol.



Work requests

SASP control flows

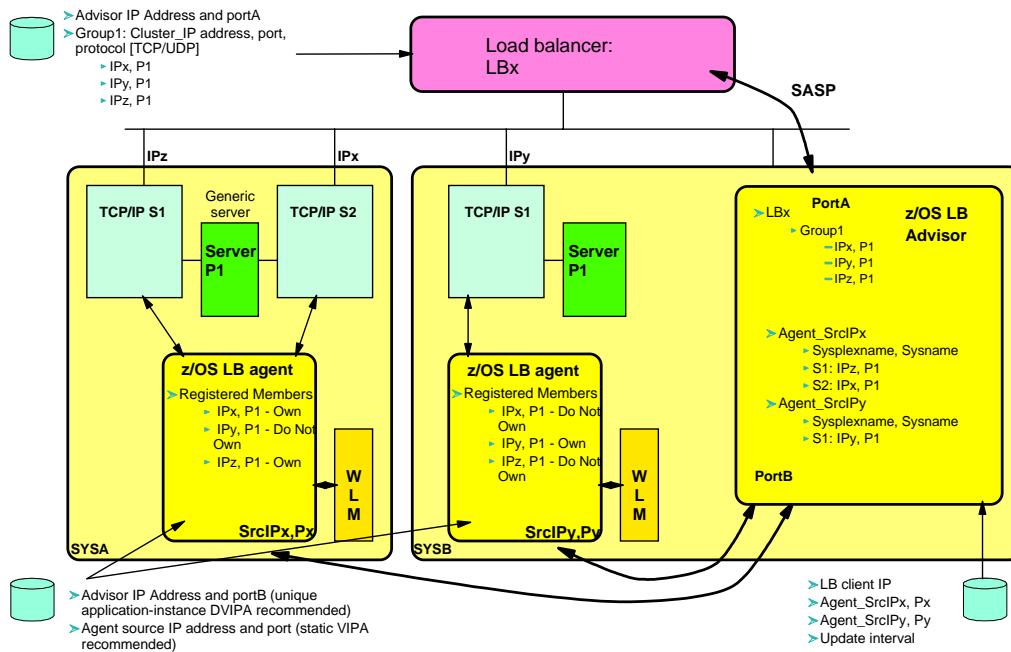


z/OS workload balancing

- ▶ Support for clustered z/OS servers in a z/OS Sysplex
- ▶ Not focused on HTTP(S) only, will support all IP-based application workloads into a z/OS Sysplex
- ▶ Based on Sysplex-wide WLM policy
- ▶ Scope is a z/OS Sysplex

The z/OS Load Balancing Advisor technology is a new z/OS Communications Server technology that was made generally available in 4Q2004 via APARs PQ90032 (V1R4) and PQ96293 (V1R5/V1R6). It is fully integrated into z/OS V1R7.

z/OS LB Advisor/Agent structure - overview



The weights that are returned to the external load balancer

> The weights are composed of two main elements:

▶ **WLM weight**

- The WLM weight that we know from other WLM-based load balancing solutions, such as Sysplex Distributor
 - ✓ A numeric value between 0 and 64

▶ **Communications Server weight**

- This weight is calculated based on the availability of the actual server instances (are they up and ready to accept workload) and how well TCP/IP and the individual server instances process the workload that is sent to them?
 - ✓ Expressed as a numeric percentage value between 0 and 100
- Purpose of calculations is to:
 - ✓ Prevent stalled server from being sent more work (accepting no new connections and new connections are being dropped due to backlog queue full condition)
 - ✓ Proactively react to server that is getting overloaded (accepting new connections, but size of backlog queue increases over time approaching the max backlog queue size)

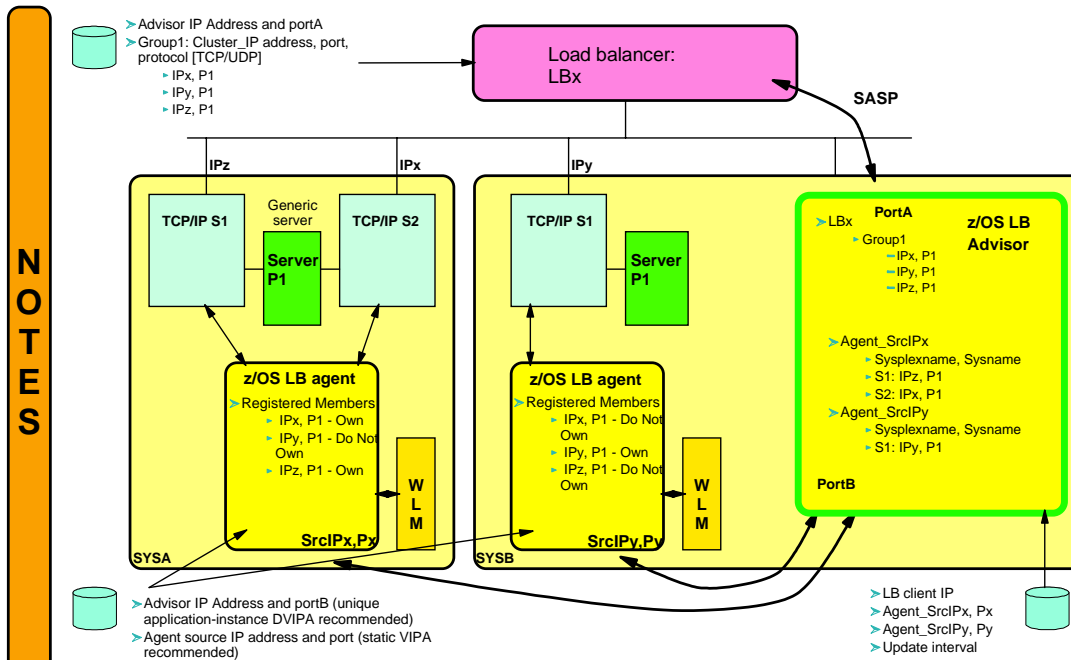
> The final weight is calculated by combining the WLM and the CS weights into a single metric

- ▶ Final weight = $WLM\ weight * CS\ weight / 100$

> Due to current external load balancer behavior when a weight of zero is returned for all members of a group, the z/OS LBA currently will never return a zero weight for all members in a group

- ▶ In the case that all members indeed do have a weight of zero, they will all be reported to the LB as having a weight of one
- ▶ Weights that are returned to the load balancer are normalized to values between 1 and 64
 - If all server instances have the same final weight (example 32), then a 1 will be returned for all server instances

z/OS LB Advisor address space - configuration overview



Redbooks © Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Advisor configuration example

NOTES

```

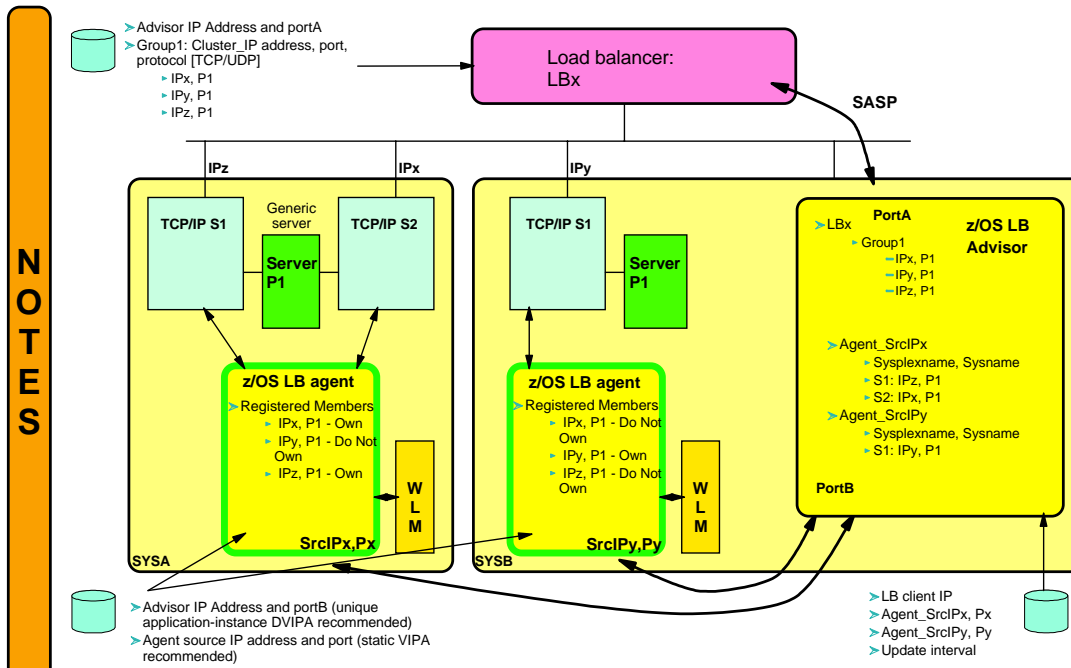
debug_level      15                # Error, Warning, Event, Info
update_interval  120               # Agent updates every 2 minutes
lb_connection_v4 10.67.5.1..3860   # DVIPA advisor listen endpoint
lb_id_list
{
  10.67.1.11                # SDBAV4
}
agent_connection_port 8100         # Agents connect to Advisor on this port
agent_id_list
{
  10.67.1.1..8000           # SD1AV4
  10.67.1.2..8000           # SD2AV4
  10.67.30.22..8000         # SD2A2V4
  10.67.1.10..8000          # SDAAV4
}

wlm serverwlm                # Request server-specific WLM weights
port_list
{
  21 wlm basewlm            # Use system WLM weights for FTP
}
  
```

Redbooks © Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

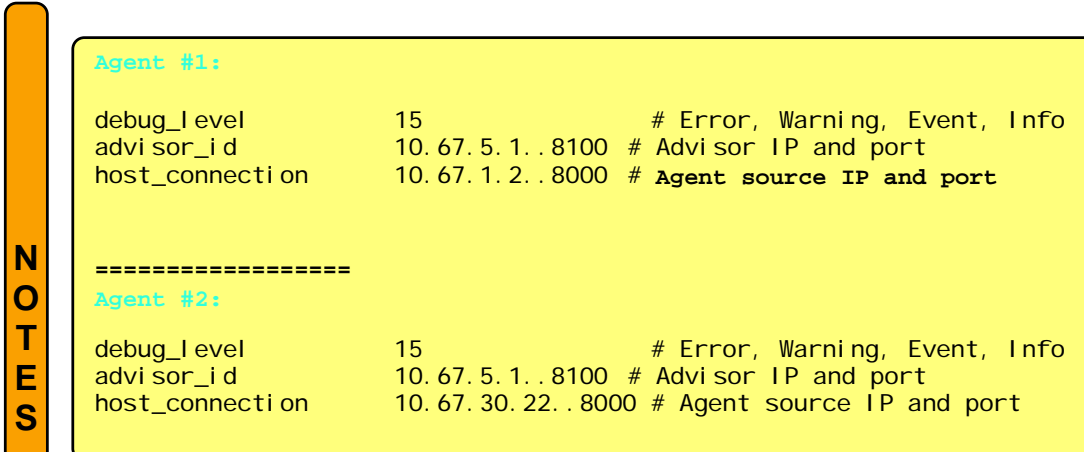
z/OS LB agent address spaces - configuration overview



Redbooks © Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Agent configuration examples



Redbooks © Copyright IBM Corp. 2005. All rights reserved.

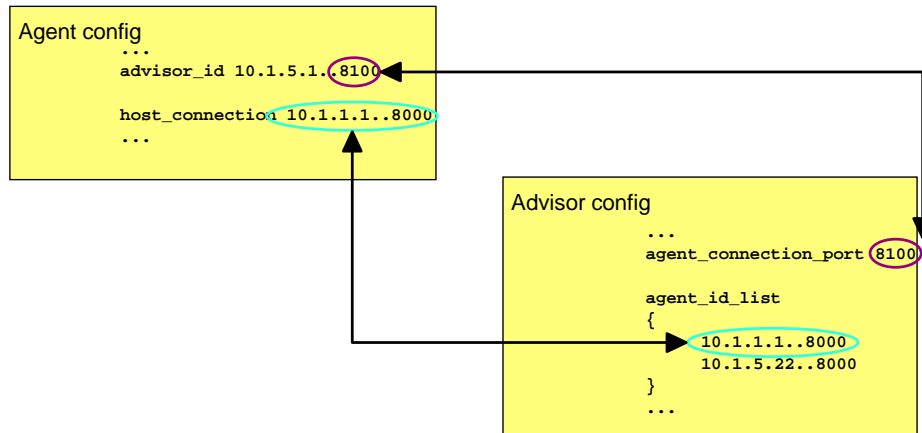
ibm.com/redbooks

Configuration hints and tips

NOTES

➤ Configuration relationships among the advisor and agent

- See connected arrows
- IP address in advisor_id statement can be any IP address belonging to the TCP/IP stack the advisor is running on, however, it is recommended this be a unique application-instance DVIPA.



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Security and control considerations

NOTES

➤ Define OMVS segments if not already defined

➤ Configure RACF

- Define user IDs to RACF, associate with OMVS segment, and assign UIDs
- Add default users to the STARTED class
- Optionally, restrict which users can start the applications
 - Only one instance of the advisor is permitted per Sysplex
 - Only one instance of the agent is permitted per MVS system
- Permit access to the BPX.WLMSEVER resource profile
 - For RACF, explicit access not required if the resource profile is not defined
 - For other security products, consult the product documentation as to whether explicit access is required for your installation
- Sample RACF definitions can be found in hlq.SEZAINST(EZARACF)
 - Look for "LBADV" and "LBAGENT"

➤ Ensure the Advisor and Agents receive the proper dispatching priority

- Verify Advisor and Agents are assigned to the WLM SYSSTC service class
- See "MVS Planning: Workload Management" for more information

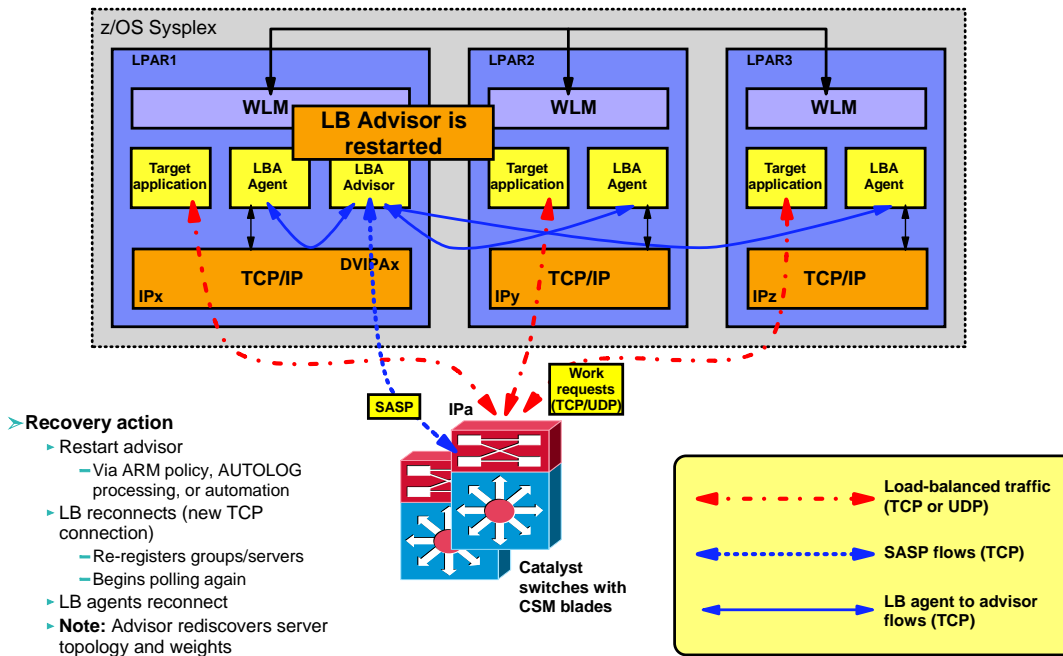
➤ Make Advisor and Agent non-swappable

- V1R7
 - Both will run non-swappable by default. No action required.
- V1R4, V1R5, and V1R6
 - Configuration of the Program Properties Table (PPT) in the SCHEDxx member is required in order to run non-swappable.

© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

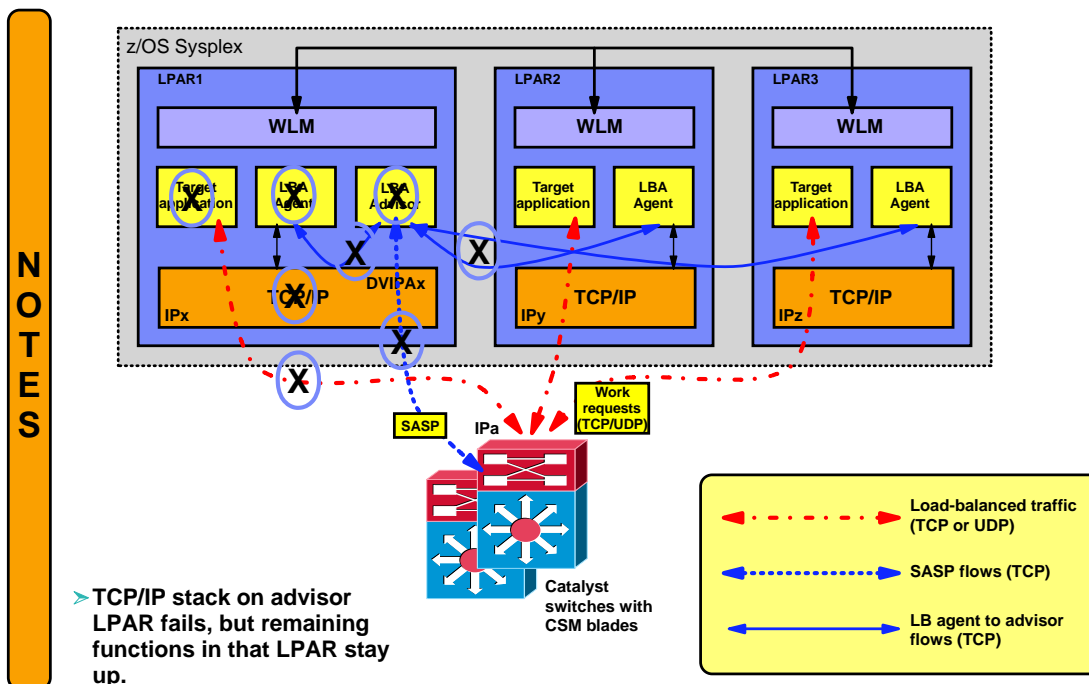
Sample setup - recovery scenario A - advisor fails (continued)



Redbooks © Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

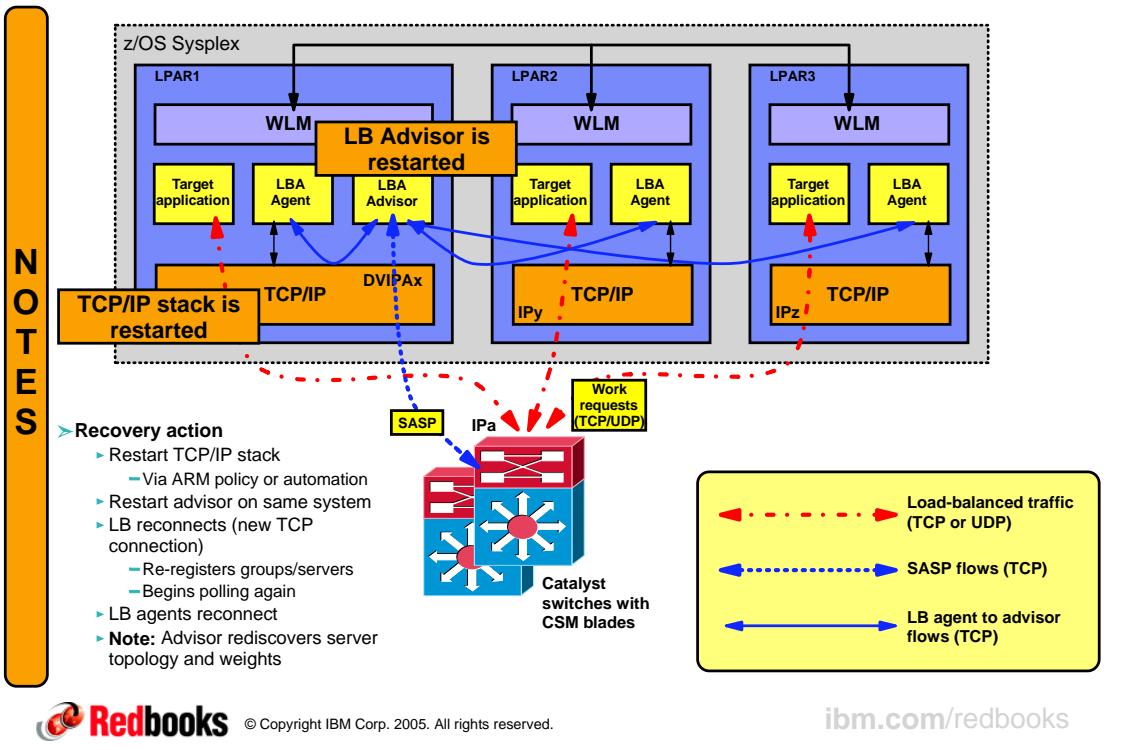
Sample setup - recovery scenario B - TCP/IP stack on advisor system fails



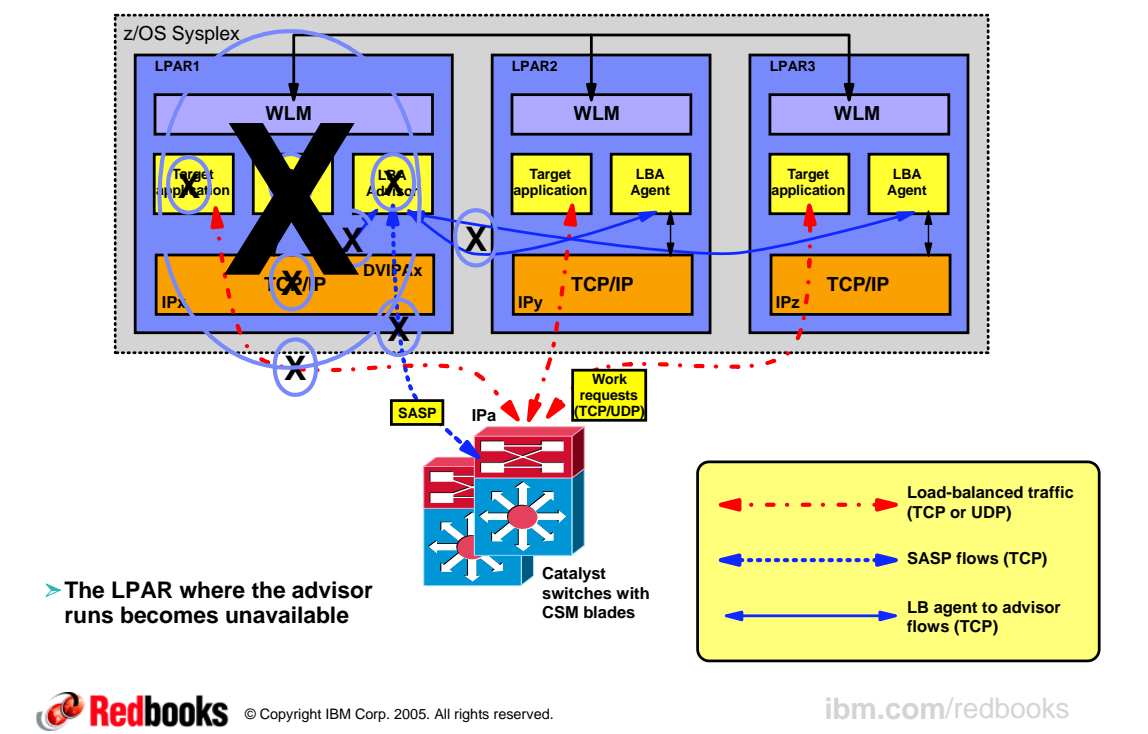
Redbooks © Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Sample setup - recovery scenario B - TCP/IP stack on advisor system fails (continued)



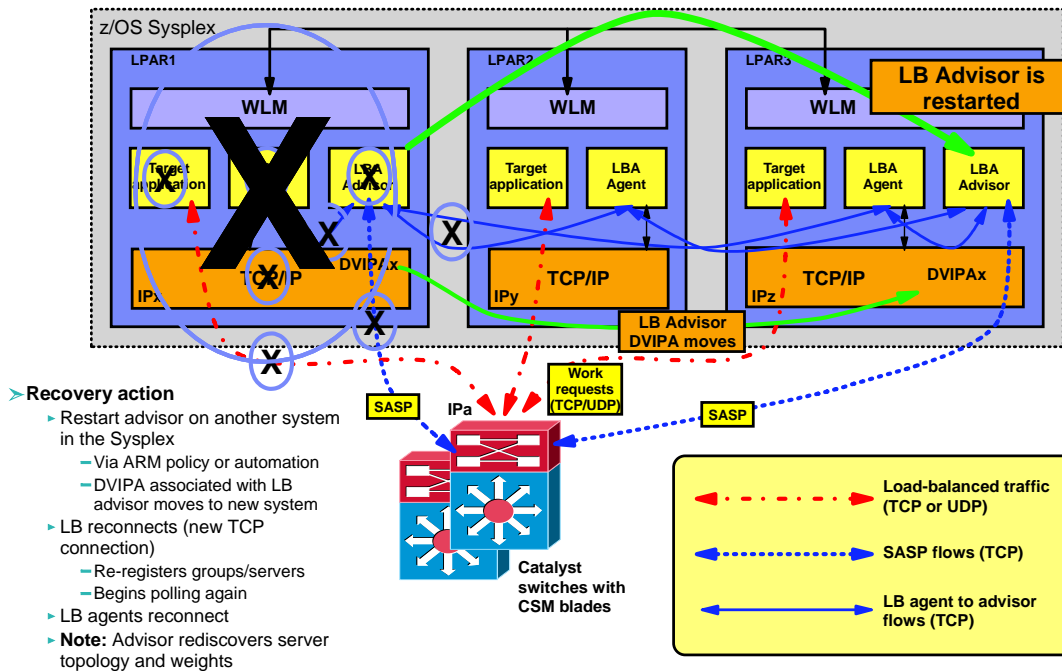
Sample setup - recovery scenario C - advisor system fails



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

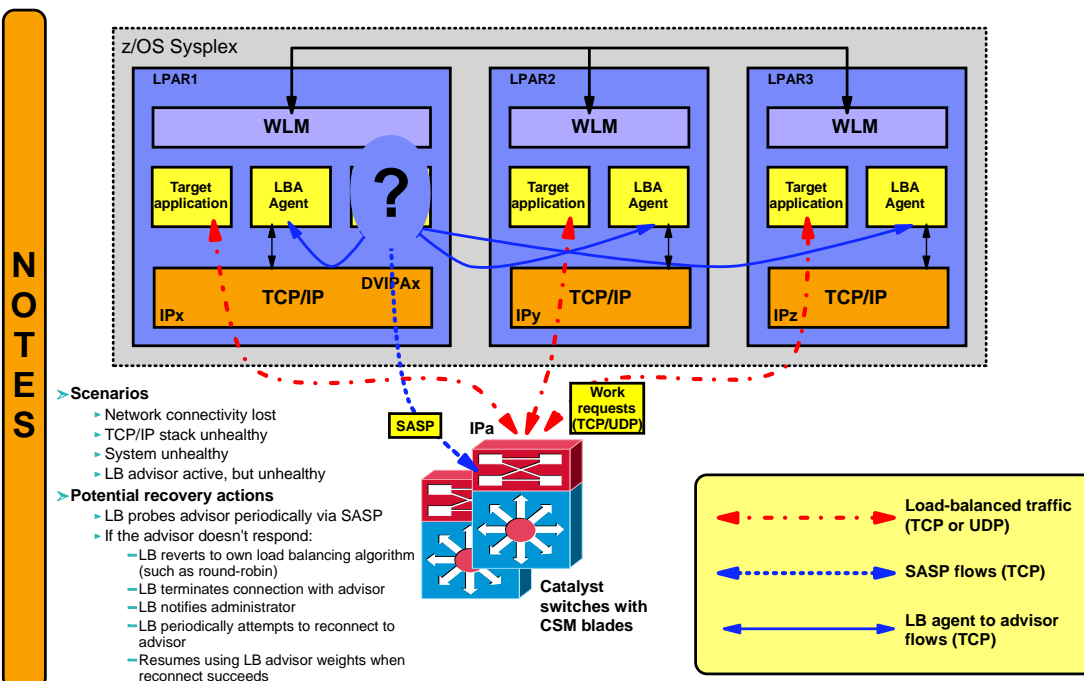
Sample setup - recovery scenario C - advisor system fails (continued)



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

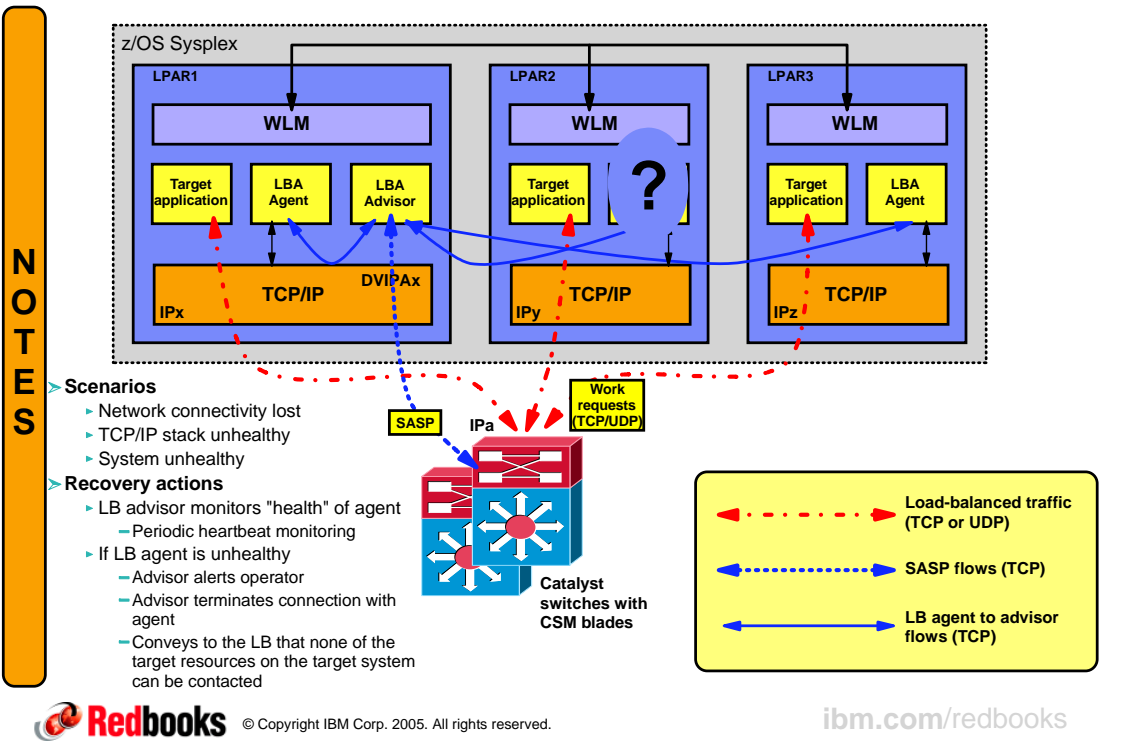
Sample setup - recovery scenario D - advisor not responding



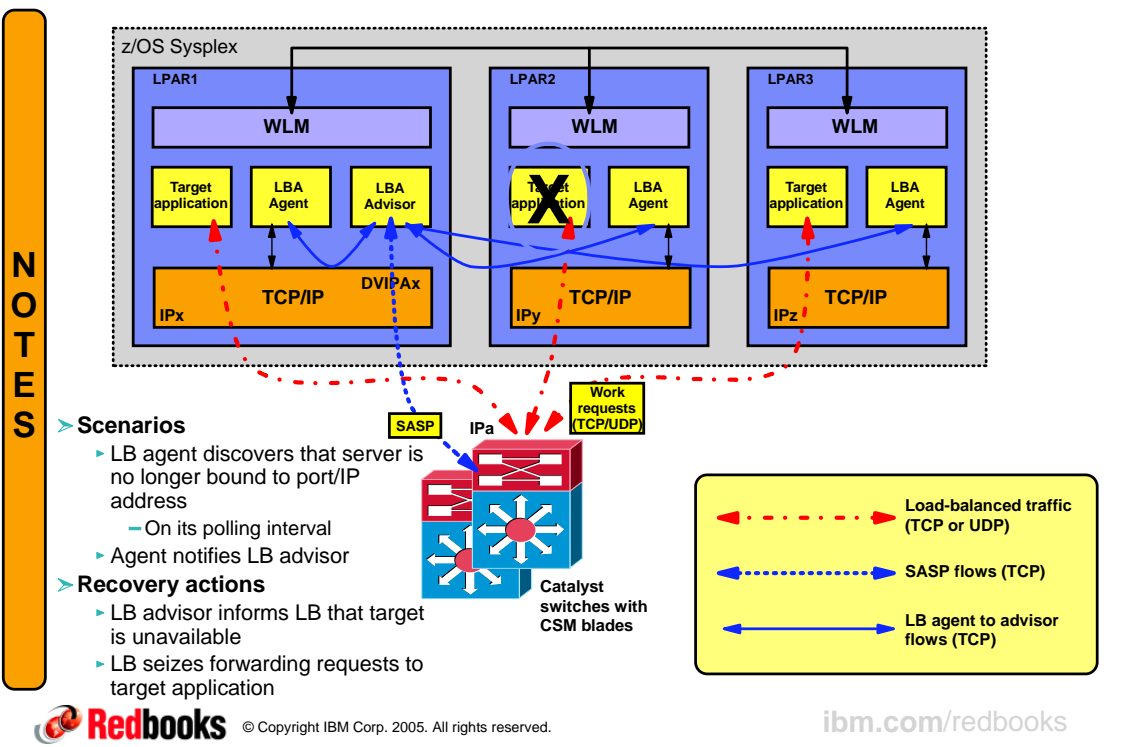
© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

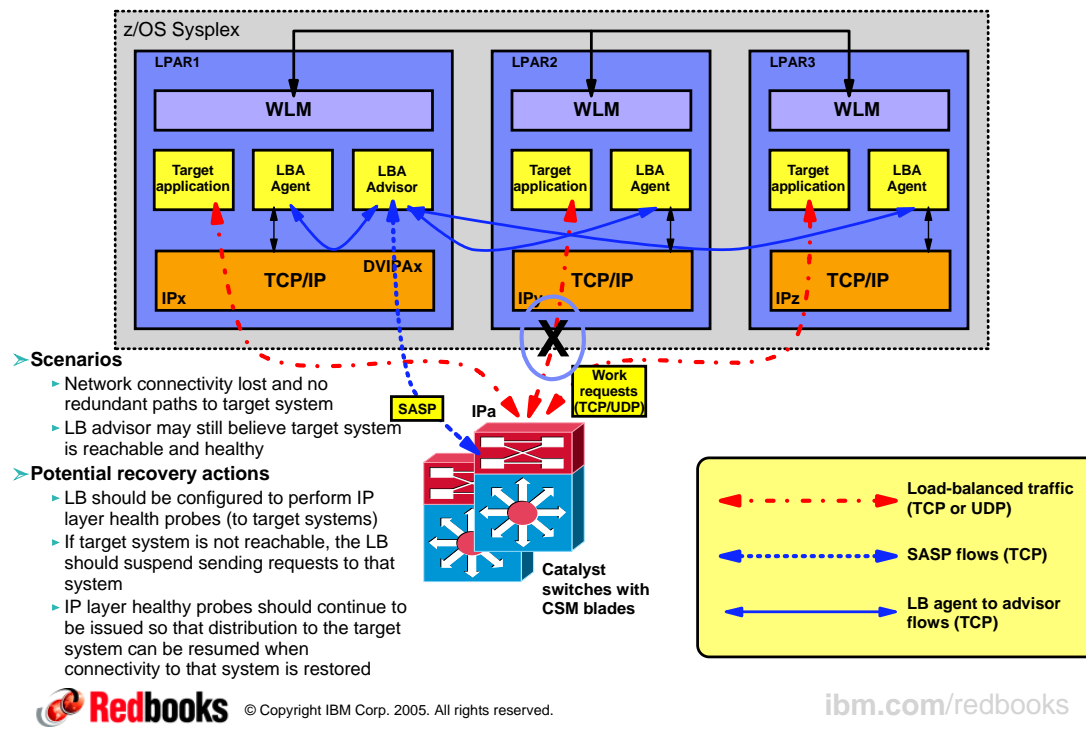
Sample setup - recovery scenario E - agent not responding



Sample setup - recovery scenario F - target application fails



Sample setup - recovery scenario G - loss of network connectivity between LB and target system or application



Managing workload distribution

> MVS operator display commands

- Display detailed information about connected load balancers
- Display detailed information about registered groups

> MVS operator Quiesce/Enable commands

- Available as MODIFY command on agents
- "Quiesce" removes member from future workload distribution eligibility
- "Enable" restores member's eligibility for workload balancing
- Only affects workload arriving through external load balancer
- Scope
 - All members owned by the MVS system of the agent
 - All members belonging to a particular TCP/IP stack
 - Port level (i.e. all members sharing a port, or all members using port on multi-stack system)
 - An individual member

> Load balancer administrator Quiesce/Enable

- SASP protocol allows for this function. Availability may be implementation dependent
- Scope
 - Implementation dependent

> MVS operator vs. load balancer administrator Quiesce/Enable interaction

- Quiesce by either makes member unavailable for load balancing - no hierarchy
- Enable by one cannot undo Quiesce by the other

Things to think about when using the z/OS LBA

➤ z/OS Load Balancing Advisor PTFed back

- ▶ PTFs only support system WLM recommendations
- ▶ PTFs do NOT support server-specific WLM recommendations
- ▶ PTFed back to V1R4, V1R5, and V1R6
 - V1R4 APAR
 - PQ90032
 - V1R5 and V1R6 APAR
 - PQ96293
- ▶ APAR publications doc at...
 - <http://www.ibm.com/support/docview.wss?rs=852&uid=swg27005585>

Things to think about when using the z/OS LBA (*continued*)

➤ Server-specific WLM recommendations

- ▶ V1R7 supports server-specific WLM recommendations
- ▶ Previous releases use system WLM recommendations exclusively
- ▶ Server-specific vs. system WLM recommendations are determined on a group basis
- ▶ Server-specific WLM recommendations will be used when all of the following are true:
 - Advisor is V1R7
 - All members of a group are owned by V1R7 agents
 - Group only contains application members (vs. system members)
- ▶ Otherwise, system WLM recommendations are used for the group
- ▶ Server-specific WLM will usually result in better workload distribution
 - e.g. HTTP
 - Exception: Servers that serve as access points to applications that run in their own address space (thus use a different WLM service class). Examples:
 - TN3270
 - FTP
 - INETD

➤ IPv6

- ▶ If using IPv6 and DVIPA for the advisor-load balancer connection or an advisor-agent connection, movement of the advisor will be limited to those z/OS releases that support IPv6 DVIPAs (V1R6 and higher)

Things to think about when using the z/OS LBA (*continued*)

> EWLM

- ▶ A group defined to a load balancer may not contain mix of members managed by EWLM and z/OS Load Balancing Advisor. All Members of a group must be managed by one or other.

> Sysplex Distributor

- ▶ May coexist with z/OS Load Balancing Advisor
- ▶ Typically you would not use both methodologies to distribute workload to the same applications

> Swappable vs. non-swappable

- ▶ V1R7 runs non-swappable by default
 - Prior to V1R7, customization is required to run non-swappable
 - See "PPT Entries to Make Non-swappable" (earlier Notes page)
 - When migrating from pre-V1R7 releases, manually added PPT entries should be removed

Enabling SASP on a Cisco CSM load balancer

> Enabling SASP on an existing CSM configuration is a simple operation. The following changes were required:

- ▶ Each load balancer connecting to a z/OS Load Balancing Advisor must have a unique ID. By default, the CSM will have the same ID; therefore, if multiple CSMs are deployed using the same z/OS Load Balancing Advisor, a unique ID needs to be configured for each (see variable SASP_CSM_UNIQUE_ID)

```
module ContentSwitchingModule 5
variable ROUTE_UNKNOWN_FLOW_PKTS 1
variable SASP_CSM_UNIQUE_ID Cisco-CSM-6509A
```

- ▶ The BINDID associates each serverfarm with the configured DFP agent. Each vserver must utilize separate serverfarms in order to register application-specific members; otherwise the CSM will only register system members.

```
serverfarm TN3270
nat server nat client
ZOS bindid 65520
real 9.42.88.9 inservice
real 9.42.88.13 inservice
real 9.42.88.1 inservice
```

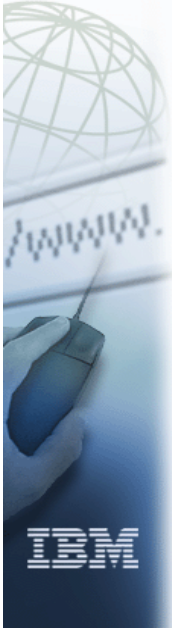
- ▶ The DFP agent is configured with the IP address and listening port of the z/OS Load Balancing Advisor along with the BINDID.

```
dfp agent 9.42.88.217 3860 65520
```

ibm.com



e-business



Communications Server for z/OS V1R7 - Technical Update System z9 and zSeries Hardware Exploitation



Redbooks

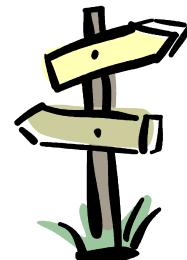
International Technical Support Organization

© Copyright IBM Corp. 2005. All rights reserved.

System z9 and zSeries hardware exploitation - agenda



- > OSA update
- > 10 Gigabit Ethernet support
- > QDIO OSA-Express2 segmentation offload
- > Support of dynamic VLAN registration
- > z9-109 IPv6 HiperSockets support



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

OSA-Express connectivity and CHPID support - overview

| Feature | Feature Name | Ports | z800 z900 | z890 z990 | z9-109 | CHPIDs | Connectors |
|---------|----------------------------|-------|--------------|--------------|--------|----------------------|------------------|
| 5201 | OSA-2 Token Ring | 2 | X | N / A | N / A | OSA | Copper, RJ-45 |
| 5202 | OSA-2 FDDI | 1 | X | N / A | N / A | OSA | Fiber, SC Duplex |
| 2362 | OSA-E 155 ATM SM | 2 | X | RPQ | N / A | OSD, OSE | Fiber, SC Duplex |
| 2363 | OSA-E 155 ATM MM | 2 | X | RPQ | N / A | OSD, OSE | Fiber, SC Duplex |
| 2364 | OSA-E GbE LX | 2 | X | C | C | OSD | Fiber, SC Duplex |
| 2365 | OSA-E GbE SX | 2 | X | C | C | OSD | Fiber, SC Duplex |
| 2366 | OSA-E Fast Ethernet | 2 | X | C | C | OSD, OSE | Copper, RJ-45 |
| 2367 | OSA-E Token Ring | 2 | X | X | N / A | OSD, OSE | Copper, RJ-45 |
| 1364 | OSA-E GbE LX | 2 | 09/04 | 06/03 | C | OSD | Fiber, LC Duplex |
| 1365 | OSA-E GbE SX | 2 | 09/04 | 06/03 | C | OSD | Fiber, LC Duplex |
| 1366 | OSA-E 1000BASE-T Ethernet | 2 | N / A | 06/03 | C | OSC, OSD, OSE | Copper, RJ-45 |
| 3364 | OSA-E2 GbE LX | 2 | N / A | 01/05 | X | OSD, OSN * | Fiber, LC Duplex |
| 3365 | OSA-E2 GbE SX | 2 | N / A | 01/05 | X | OSD, OSN * | Fiber, LC Duplex |
| 3366 | OSA-E2 1000BASE-T Ethernet | 2 | N / A | N / A | X | OSC, OSD, OSE, OSN * | Copper, RJ-45 |
| 3368 | OSA-E2 10 GbE LR | 1 | N / A | 01/05 | X | OSD | Fiber, SC Duplex |

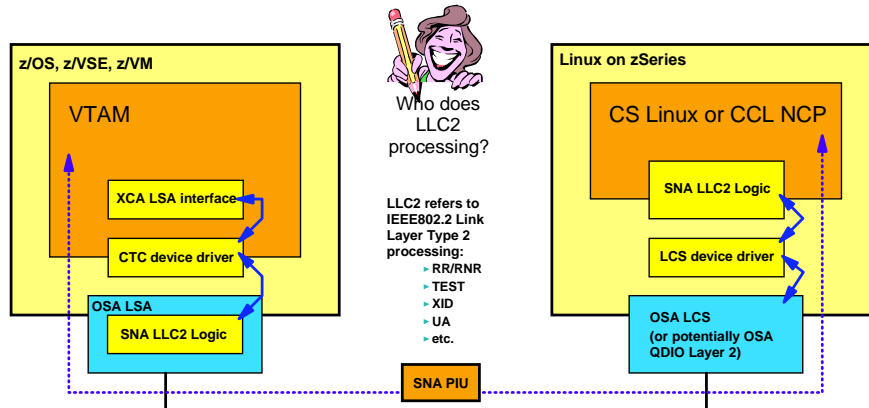
LX = Long wavelength transceiver, SX = Short wavelength transceiver, LR - Long Reach transceiver
 X = Available for ordering, C = Carry forward on an upgrade from z900 or z990
 * OSN is exclusive to z9-109. Hardware availability is 09/16/05

What are the CHPID types used for?

| CHPID type | Feature | Traffic type | | | | OSA/SF required |
|-----------------------------|---|--|--------|---------|-----|-----------------|
| | | SNA/APPN/HPR | TCP/IP | Console | NCP | |
| OSD zSeries System z9 | GbE, 10 GbE 1000BASE-T Ethernet Fast Ethernet | No (L3) Use EE or TN3270E Yes (L2) | Yes | No | No | No |
| OSE zSeries System z9 | 1000BASE-T Ethernet Fast Ethernet | Yes | Yes | No | No | Yes |
| OSC z990, z890 z9-109 | 1000BASE-T Ethernet | No | No | Yes | No | No |
| OSN z9-109 exclusive | 1000BASE-T Ethernet GbE | No | No | No | Yes | No |

- z/OS and Linux on zSeries support both IPv4 and IPv6 traffic over QDIO layer 3 interfaces.
- QDIO layer 2 mode is supported on z890, z990, and z9-109 only.
- Only Linux currently supports QDIO layer 2 mode.
 - When using QDIO layer 2 mode for IP traffic, none of the OSA QDIO layer 3 IP assist functions are available
 - ARP offload, Large send segmentation offload, checksum offload, etc.

What is the difference between VTAM's OSA LSA usage and Linux's OSA LCS usage for SNA LAN traffic?



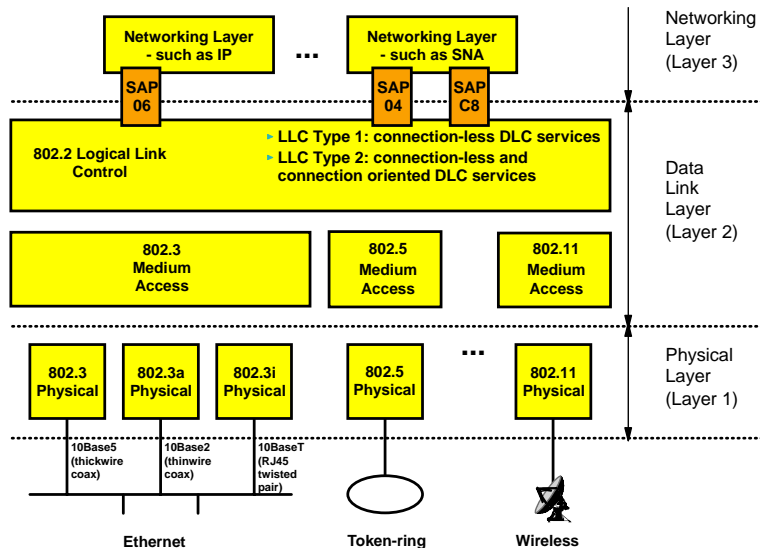
- > VTAM does not include SNA LLC2 processing logic, but has a higher-level interface defined that generally is known as LSA (Link Services Architecture).
 - ▶ VTAM and SNA-specific LLC2 code in the OSA adapter communicates with each other using the LSA primitives
 - ▶ The actual device driver between VTAM and the OSA adapter in the case of LSA is a CTC device driver
- > Linux and the SNA software running on Linux provide full SNA LLC2 logic and are able to present fully built SNA LAN frames to the OSA adapter
 - ▶ Can use the LCS device driver to interface with the OSA adapter (a LAN frame is a LAN frame!)
 - ▶ Also opens up for potentially using QDIO in layer 2 mode since the SNA solutions on Linux do not depend on SNA-specific capabilities in the OSA adapter



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

IEEE802 - the lower layers - structure



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

10 Gigabit Ethernet support - QDIO

- CS z/OS v1R7 adds support for OSA-Express2 10 Gigabit Ethernet (Gbe) LR (LR = Long Range) feature
- Requires z990, z890, or z9-109
- Configured and managed exactly like Gigabit Ethernet
- Transparent except the following will reflect the actual speed:
 1. the Speed field on the Netstat DEVLINKS/-d report output
 2. the SNMP MIB object ifHighSpeed (from the IF-MIB)

Support was
PTFed back to
z/OS V1R4
(APARs:
OA09759 and
PQ96769)

```
DevName: OTGETH1          DevType: MPCIPA
DevStatus: Ready
LnkName: LOTGETH1         LnkType: IPAQENET  LnkStatus: Ready
NetNum: n/a  QueSize: n/a  Speed: 0000010000
IpBroadcastCapability: No
.
.
DevName: OGETHD          DevType: MPCIPA
DevStatus: Ready
LnkName: LOGETHD         LnkType: IPAQENET  LnkStatus: Ready
NetNum: n/a  QueSize: n/a  Speed: 0000001000
IpBroadcastCapability: No
.
.
```



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

TCP segmentation offload - also known as "large send"

- **TCP segmenting consumes host CPU cycles in the TCP/IP stack**
 - TCP segmentation is normally done by the TCP layer to split large send buffers up into smaller TCP segments that each fit into an IP packet of a specific size (the MTU size)
 - MTU size is based on interface characteristics, route characteristics or is learned dynamically via path MTU discovery
- **Non-optimal use of Direct Memory Access for QDIO transfer**
- **CS z/OS V1R7 adds support for new OSA-Express feature (segmentation offload also referred to as 'Large Send')**
 - Offload most IPv4 TCP segmentation processing to OSA-Express in QDIO mode
 - Decrease host CPU utilization
 - Increase data transfer efficiency for IPv4 packets
- **Support automatically enabled when available in adapter**
 - Similar to existing checksum offload function
 - Checksum is offloaded whenever segmentation is offloaded
 - No configuration controls in TCP/IP
- **Applies to the OSA-Express2 features Gigabit Ethernet SX and LX, 10 Gigabit Ethernet LR**
 - Supports QDIO mode only (CHPID type OSD), and is exclusive to z990, z890, and z9-109
- **Segmentation offload support is available for z/OS V1R6.0 Communications Server.**
 - Solution was PTFed back to z/OS V1R6

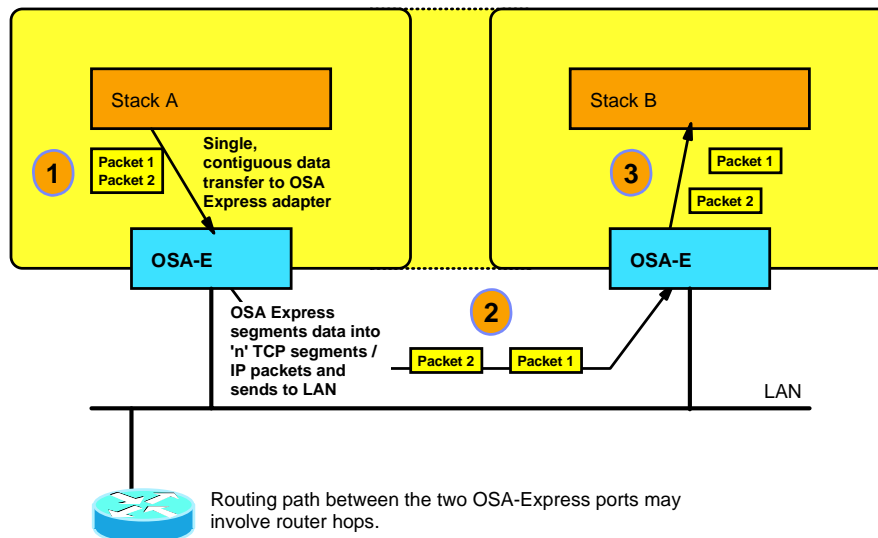


© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Segmentation offload when next hop is reached via LAN

> Segmentation can be offloaded to the OSA.



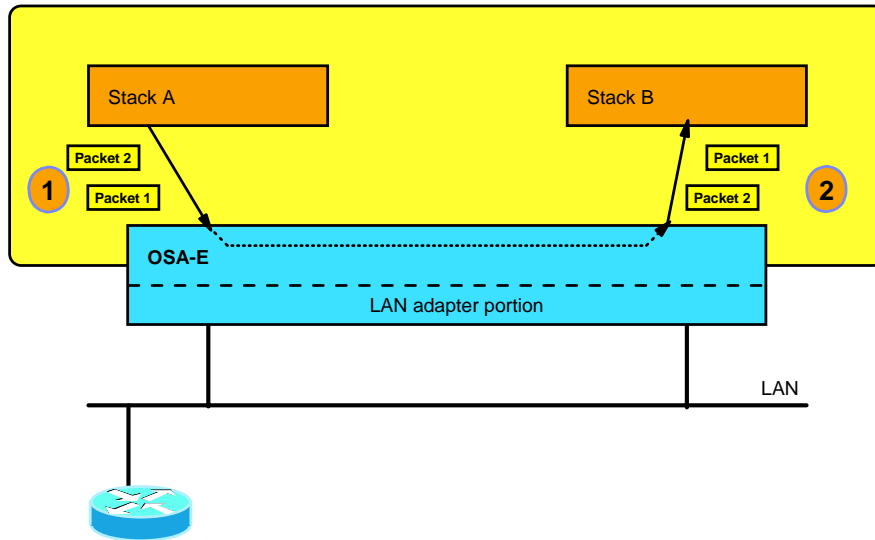
Restrictions

NOTES

- > IPv4 only
- > TCP transport only
- > Outbound packets only
- > Packets written to the LAN only (not to another stack sharing the OSA)
- > Packets larger than MSS only
- > For multipath, only when all devices in the multipath group support segoffload
- > No IPSEC packets

No segmentation offload when next hop is not reached via LAN

- Segmentation cannot be offloaded when packets loop through the adapter to another target stack that shares the adapter.



Things to think about when enabling large send support

- **Big send buffer (up to 56KB) maximizes offloading. Configure the TCP send buffer size using the following existing mechanisms...**
 - ▶ TCPSENDBfrsize on TCPCONFIG statement sets default for all applications
 - ▶ SETSOCKOPT (SO_SNDBUF) by the application overrides default
- **Send buffer size also limited by receive buffer size at other end of connection.**
 - ▶ TCPRCVBufsize on TCPCONFIG statement sets default for all applications
 - ▶ SETSOCKOPT (SO_RCVBUF) by the application overrides default
- **APARs supplied for QDIO OSA-Express2 Segmentation Offload for z/OS V1R6**
 - ▶ TCP APAR: PK02490 - PTF: UK04060 and UK04061
 - ▶ VTAM APAR: OA11148 - PTF: UA18116
- **There is no interdependency between the VTAM code and the TCP/IP code. The VTAM code can be applied without the TCP/IP code and vice versa. However, segmentation offload is not enabled unless both pieces are applied.**

Segmentation offload performance details

> OSAE-2, 1 GbE
(versus no
segmentation
offload):

| Workload | Trans/Sec Delta % | CPU/Tran Delta % |
|----------------------|----------------------|---------------------|
| RR 60 | + 1.3 % | - 0.7 % |
| CRR 9 | + 2 % | - 0.1 % |
| STR (1/20M): | | |
| 64K(send)/32K(recv) | Equal | - 28.9 % |
| 180K(send)/64K(recv) | Equal | - 36.3 % |
| 256K(send)/64K(recv) | Equal | - 39.2 % |

> OSAE-2, 10 GbE
(versus no
segmentation
offload):

| Workload | Trans/Sec Delta % | CPU/Tran Delta % |
|----------------------|----------------------|---------------------|
| RR 60 | + 1.7 % | - 2 % |
| CRR 60 | + 5.2 % | - 1 % |
| STR (1/20M): | | |
| 64K(send)/32K(recv) | + 1.1 % | - 33.4 % |
| 180K(send)/64K(recv) | + 1.5 % | - 41.5 % |
| 256K(send)/64K(recv) | + 0.4 % | - 44.9 % |

Dynamic VLAN registration - GVRP


- > OSA has recently added support for dynamic VLAN registration protocols
 - ▶ Both the OSA microcode and the switch to which the OSA port is connected must support dynamic VLAN registration
 - ▶ The protocol is referred to as GARP (Generic Attribute Registration Protocol) VLAN Registration Protocol (GVRP)
 - ▶ Simplifies management of VLAN environments
- > New TCP/IP profile keywords on the IPv4 LINK statement and the IPv6 INTERFACE statement for QDIO network interfaces are used to control if TCP/IP should request dynamic VLAN registration or not
 - ▶ Default is to not use dynamic VLAN registration

```
LINK .... NODYNVLANREG/DYNVLANREG
INTERFACE .... NODYNVLANREG/DYNVLANREG
```

- > New fields on a netstat devlinks report will indicate if the dynamic VLAN registration is supported by the OSA port and if dynamic VLAN registration is configured for the link or interface
- > CS z/OS V1R7 APAR PK05337
 - ▶ PTFs UK06129 and UK06130

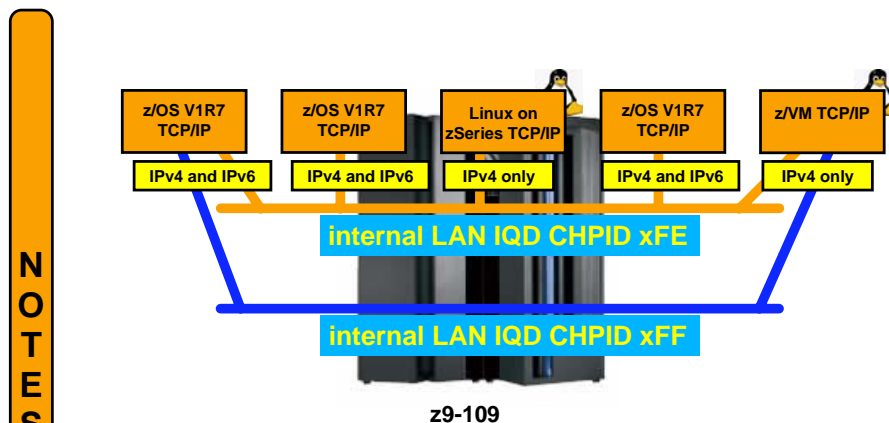
z9-109 IPv6 HiperSockets support

- z/OS V1R4 introduced IPv6, but only provided connectivity via OSA-Express QDIO
- z/OS V1R5 added MPCPTP IPv6 connectivity
- z/OS V1R5 IPCONFIG6 DYNAMICXCF support does not include HiperSockets
- z/OS V1R7 adds IPv6 connectivity over HiperSockets
 - ▶ Include HiperSockets in IPv6 automatic connectivity options under IPCONFIG6 DYNAMICXCF
 - ▶ Requires a z9-109 processor
 - ▶ No IPv6 support for HiperSockets Accelerator function
- Configure INTERFACE statement for IPAQIDIO6
 - ▶ CHPID keyword identifies the HiperSockets CHPID
 - ▶ To use HiperSockets for both IPv4 and IPv6 for the same CHPID, specify the same value both on the CHPID keyword of the INTERFACE statement and the xx suffix of the IUTIQDxx device_name on the DEVICE statement
 - ▶ Optional INTFID keyword to specify interface ID (and override value returned by the hardware)
- Similar attributes to existing IPv6 support
 - ▶ INTERFACE statement options to:
 - Add/delete/deprecate addresses
 - Specify a VIPA for SOURCEVIPA
 - ▶ Use interface name for START/STOP and on static routes (BEGINROUTES)
 - ▶ Separate START/STOP of IPv4 and IPv6
 - ▶ Separate interface counters for IPv4 and IPv6

 © Copyright IBM Corp. 2005. All rights reserved.


ibm.com/redbooks

IPv6 HiperSockets example



NOTES

- Hardware platform is a z9-109 processor
- Both Linux, z/OS, and z/VM currently support IPv6
 - ▶ But only z/OS currently supports IPv6 over HiperSockets
- A HiperSockets CHPID can be used for IPv4 and IPv6 concurrently

 © Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Things to think about when enabling IPv6 HiperSockets connectivity

- **HiperSockets IPv6 connectivity not supported to pre-V1R7 z/OS stacks**
- **Existing message EZZ4347I if hardware does not support IPv6 HiperSockets**
 - ▶ Note: This message is suppressed for XCF Dynamics.
- **Only HiperSockets on System z9 supports IPv6**
- **Only z/OS V1R7 supports IPv6 on HiperSockets**
 - ▶ Neither z/VM nor Linux on System z9 support IPv6 over HiperSockets

This page intentionally left blank

ibm.com



Communications Server for z/OS V1R7 - Technical Update Applications



Redbooks

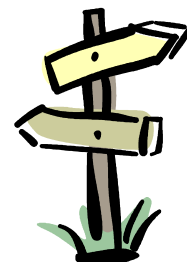
International Technical Support Organization

© Copyright IBM Corp. 2005. All rights reserved.

Applications agenda

This section or parts thereof will only be presented if requested by workshop participants.

- > **Introduction to AES**
- > **TN3270 server**
 - Option to control use of SSL V2 or not
 - AES encryption support
- > **Sendmail**
 - AES encryption support
- > **FTP**
 - AES encryption support
 - Delegated RACF resource profiles for TLS FTP
 - FTP support for mixed-case RACF passwords
 - FTP JES SAPI interface changes
 - FTP client API in C programming language
 - FTP data transfer reliability feedback
 - FTP configurable end-of-line character



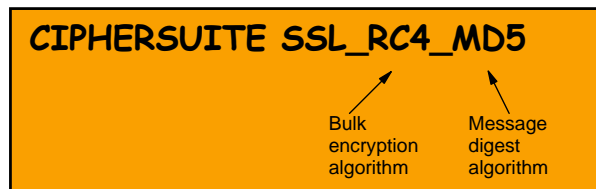
A little background information on cipher suites

➤ A cipher suite is of a collection of cryptographic algorithms:

- ▶ A key exchange algorithm
 - System SSL on z/OS uses RSA algorithms for key exchange - also known as public private key or assymmetric encryption
 - Authentication is based on digital x.509 certificates
 - SSL/TLS server always has a certificate; SSL/TLS client may optionally have a certificate also
- ▶ A bulk encryption algorithm
 - Used to encrypt/decrypt the data that is exchanged between the connection endpoints
 - Needs to be fast and efficient - symmetric encryption algorithms are used for this purpose
- ▶ A message digest algorithm
 - Used to ensure each message exchange has not been altered in transit, and that it came from the intended sender (also sometimes referred to as a digital signature)

➤ System SSL uses a 2-digit number to identify the various cipher suites it supports.

- ▶ Sendmail configuration is based on those 2-digit numbers
- ▶ TN3270 and FTP support a text string-based configuration that is then translated by TN3270 and FTP to the 2-digit numbers system SSL uses



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Full list of system SSL cipher suites in z/OS V1R7

NOTES

SSL V2 ciphers

- 1 = 128-bit RC4 encryption with MD5 message authentication (128-bit secret key)
- 2 = 128-bit RC4 export encryption with MD5 message authentication (40-bit secret key)
- 3 = 128-bit RC2 encryption with MD5 message authentication (128-bit secret key)
- 4 = 128-bit RC2 export encryption with MD5 message authentication (40-bit secret key)
- 6 = 56-bit DES encryption with MD5 message authentication (56-bit secret key)
- 7 = 168-bit Triple DES encryption with MD5 message authentication (168-bit secret key)

SSL V3 ciphers

- 00 = No encryption or message authentication and RSA key exchange
- 01 = No encryption with MD5 message authentication and RSA key exchange
- 02 = No encryption with SHA-1 message authentication and RSA key exchange
- 03 = 40-bit RC4 encryption with MD5 message authentication and RSA key exchange
- 04 = 128-bit RC4 encryption with MD5 message authentication and RSA key exchange
- 05 = 128-bit RC4 encryption with SHA-1 message authentication and RSA key exchange
- 06 = 40-bit RC2 encryption with MD5 message authentication and RSA key exchange
- 09 = 56-bit DES encryption with SHA-1 message authentication and RSA key exchange
- 0A = 168-bit Triple DES encryption with SHA-1 message authentication and RSA key exchange
- 0C = 56-bit DES encryption with SHA-1 message authentication and fixed Diffie-Hellman key exchange signed with a DSS certificate
- 0D = 168-bit Triple DES encryption with SHA-1 message authentication and fixed Diffie-Hellman key exchange signed with a DSS certificate
- 0F = 56-bit DES encryption with SHA-1 message authentication and fixed Diffie-Hellman key exchange signed with an RSA certificate
- 10 = 168-bit Triple DES encryption with SHA-1 message authentication and fixed Diffie-Hellman key exchange signed with an RSA certificate
- 12 = 56-bit DES encryption with SHA-1 message authentication and ephemeral Diffie-Hellman key exchange signed with a DSS certificate
- 13 = 168-bit Triple DES encryption with SHA-1 message authentication and ephemeral Diffie-Hellman key exchange signed with a DSS certificate
- 15 = 56-bit DES encryption with SHA-1 message authentication and ephemeral Diffie-Hellman key exchange signed with an RSA certificate
- 16 = 168-bit Triple DES encryption with SHA-1 message authentication and ephemeral Diffie-Hellman key exchange signed with an RSA certificate
- 2F = 128-bit AES encryption with SHA-1 message authentication and RSA key exchange
- 30 = 128-bit AES encryption with SHA-1 message authentication and fixed Diffie-Hellman key exchange signed with a DSS certificate
- 31 = 128-bit AES encryption with SHA-1 message authentication and fixed Diffie-Hellman key exchange signed with an RSA certificate
- 32 = 128-bit AES encryption with SHA-1 message authentication and ephemeral Diffie-Hellman key exchange signed with a DSS certificate
- 33 = 128-bit AES encryption with SHA-1 message authentication and ephemeral Diffie-Hellman key exchange signed with an RSA certificate
- 35 = 256-bit AES encryption with SHA-1 message authentication and RSA key exchange
- 36 = 256-bit AES encryption with SHA-1 message authentication and fixed Diffie-Hellman key exchange signed with a DSS certificate gsk_environment_open()
- 37 = 256-bit AES encryption with SHA-1 message authentication and fixed Diffie-Hellman key exchange signed with an RSA certificate
- 38 = 256-bit AES encryption with SHA-1 message authentication and ephemeral Diffie-Hellman key exchange signed with a DSS certificate
- 39 = 256-bit AES encryption with SHA-1 message authentication and ephemeral Diffie-Hellman key exchange signed with an RSA certificate



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

AES - Advanced Encryption Standard

➤ AES - Advanced Encryption Standard

- ▶ AES is an official U.S. Government standard. The Secretary of Commerce approved the adoption of the AES as an official government standard, effective May 26, 2002
 - Federal Information Processing Standard
 - FIPS publication 197
- ▶ AES is stronger than the Data Encryption Standard (DES) and therefore should be a popular standard both inside and outside the United States.
- ▶ AES is a bulk encryption algorithm
 - Suitable for TLS
 - More secure than DES (Data Encryption Standard)
- ▶ For more information on AES, a fact sheet is available at the following Web site:
 - <http://csrc.nist.gov/CryptoToolkit/aes/aesfact.html>

➤ Supported by SSL element of z/OS since z/OS V1R4

➤ Support being added to TN3270, FTP, and Sendmail in z/OS V1R7

- ▶ Mostly a question of adding new keywords to the configuration files.

➤ System SSL must be installed with the security level 3 feature to support AES:

- ▶ FMID JCPT341
- ▶ Not included in base element System SSL Cryptographic services

Most important cipher suites and their relationship to System SSL FMIDs

- FMID HCPT340 - Cryptographic Services System SSL
- FMID JCPT341 - Cryptographic Services Security Level 3

| Encryption type and key size | Cipher | FMID HCPT340 | FMID JCPT341 |
|------------------------------|--------|--------------|--------------|
| 512 bit keys | | ✓ | ✓ |
| 1024 bit keys | | ✓ | ✓ |
| 2048 bit keys | | ✓ | ✓ |
| SSL V2.0 RC4 US | 1 | | ✓ |
| SSL V2.0 RC4 Export | 2 | ✓ | ✓ |
| SSL V2.0 RC2 US | 3 | | ✓ |
| SSL V2.0 RC2 Export | 4 | ✓ | ✓ |
| SSL V2.0 DES 56-bit | 6 | ✓ | ✓ |
| SSL V2.0 Triple DES US | 7 | | ✓ |
| SSL V3.0 NULL MD5 | 01 | ✓ | ✓ |
| SSL V3.0 NULL SHA-1 | 02 | ✓ | ✓ |
| SSL V3.0 RC4 MD5 Export | 03 | ✓ | ✓ |
| SSL V3.0 RC4 MD5 US | 04 | | ✓ |
| SSL V3.0 RC4 SHA-1 US | 05 | | ✓ |
| SSL V3.0 RC2 MD5 Export | 06 | ✓ | ✓ |
| SSL V3.0 DES SHA-1 Export | 09 | ✓ | ✓ |
| SSL V3.0 Triple DES SHA-1 US | 0A | | ✓ |
| SSL V3.0 AES 128 Bit SHA-1 | 2F | | ✓ |
| SSL V3.0 AES 256-Bit SHA-1 | 35 | | ✓ |

TN3270 server SSL/TLS support - 128/256 bit AES encryption

> TN3270 supports more levels of SSL/TLS protocols:

- ▶ SSLv2 - generally considered to be a weak security protocol
 - The TN3270 server in z/OS V1R7 adds configuration control to disallow/allow a client to negotiate use of SSL V2 protocol levels:
 - SSLV2 or NOSSLV2 - default is NOSSLV2
 - This option can be specified on TelnetGlobals, TelnetParms, or in a ParmsGroup
- ▶ SSLv3
- ▶ TLSv1

> TN3270 server support of cipher suites:

| Cipher suite | Telnet display abbreviation |
|------------------------|-----------------------------------|
| SSL_RC4_SHA | 4S |
| SSL_RC4_MD5 | 4M |
| SSL_AES_256_SHA | A2 <== New in z/OS V1R7 |
| SSL_AES_128_SHA | A1 <== New in z/OS V1R7 |
| SSL_3DES_SHA | 3S |
| SSL_DES_SHA | DS |
| SSL_RC4_MD5_EX | 4E |
| SSL_RC2_MD5_EX | 2E |
| SSL_NULL_SHA | NS |
| SSL_NULL_MD5 | NM |
| SSL_NULL_Null | NN |

This is the default TN3270 server cipher suite list in the order of preference.

If you need to change that order or to exclude certain choices, code the ENCRYPTION / ENDECRYPTION block in TelnetGlobals, TelnetParms or in a Parmsgroup.

The two-digit telnet display abbreviation codes are not the same as the system SSL 2-digit cipher suite codes!

Sendmail SSL/TLS AES support

> Sendmail support for AES encryption:

- ▶ New AES cipher suites added to the CipherLevel parameter in the z/OS specific Sendmail configuration file:
 - For 128 bit keys, specify: cipherlevel 2F
 - For 256 bit keys, specify: cipherlevel 35

> No new diagnostics

> No migration concerns

> Default CipherLevel for Sendmail depends on the level of encryption support that is installed on the system (these are the system SSL defaults):

- ▶ Domestic encryption: 050435363738392F303132330A1613100D0915120F0C0306020100
- ▶ Otherwise: 0915120F0C0306020100

> Note: the Sendmail configuration file uses the 2-digit cipher suite codes that are assigned by system SSL

FTP server and client SSL/TLS AES support

> **FTP.DATA cipher suite parameter additions:**

- > CIPHERSUITE SSL_AES_128_SHA
- > CIPHERSUITE SSL_AES_256_SHA

> **FTP SMF 119 records will report the standard system SSL 2-digit codes corresponding to the above cipher suites:**

- > CIPHERSUITE SSL_AES_128_SHA - code 2F
- > CIPHERSUITE SSL_AES_256_SHA - code 35

> **The following FTP SMF 119 records contain information about cipher suites:**

- > FTP server transfer completion record
- > FTP server logon failure record
- > FTP client transfer completion record

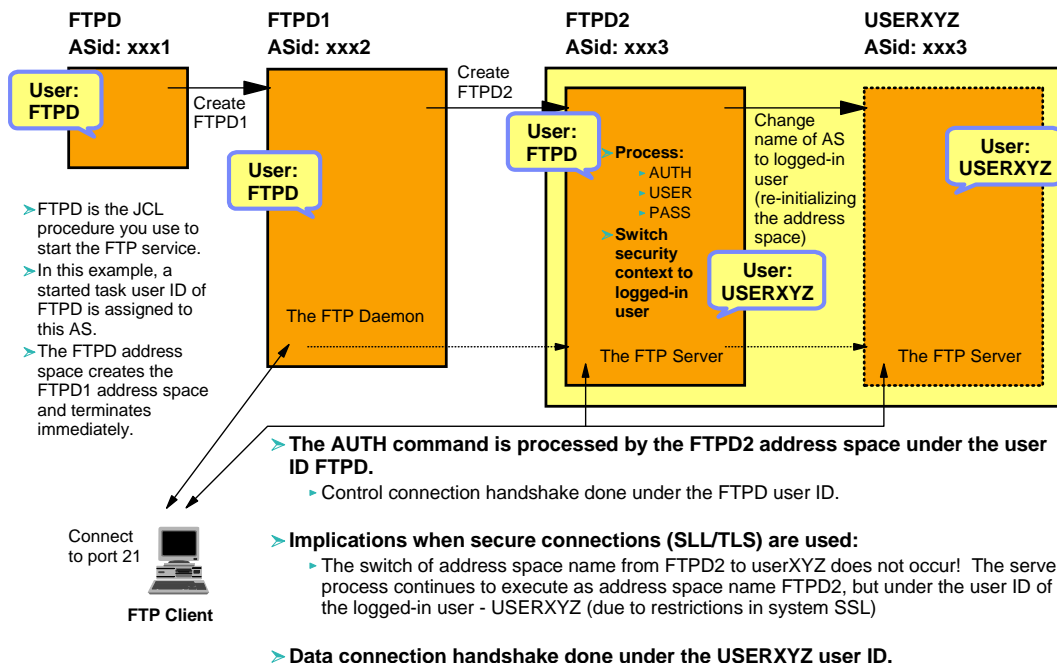
Cipher suites supported by FTP client and server on z/OS

| |
|------------------------|
| SSL_RC4_SHA |
| SSL_RC4_MD5 |
| SSL_AES_256_SHA |
| SSL_AES_128_SHA |
| SSL_3DES_SHA |
| SSL_DES_SHA |
| SSL_RC4_MD5_EX |
| SSL_RC2_MD5_EX |
| SSL_NULL_SHA |
| SSL_NULL_MD5 |
| SSL_NULL_Null |

This is the default FTP client and server cipher suite list in the order of preference.

If you need to change that order or to exclude certain choices, code the CIPHERSUITE option in the z/OS FTP client and/or server FTP.DATA

The FTP server and its related address spaces - implications when using SSL/TLS FTP



FTP's relationship to ICSF is through system SSL

- **FTP TLS hardware encryption is through ICSF services.**
 - ▶ FTP calls System SSL to do the encryption and, if hardware encryption is available, system SSL calls ICSF services to do the encryption
- **ICSF - Integrated Cryptographic Service Facility - is a software element of z/OS.**
 - ▶ ICSF provides
 - An interface to cryptographic hardware (services)
 - Storage for private cryptographic keys (CKDS and PKDS)
 - ▶ For more information on ICSF, see "z/OS ICSF Overview", SA22-7519
- **You can control access to cryptographic services and keys using an SAF-compliant security product such as RACF.**
 - ▶ CSFKEYS class
 - You can define resource profiles in the CSFKEYS class to control access to cryptographic keys.
 - ▶ CSFSERV class
 - You can define resource profiles in the CSFSERV class to control access to cryptographic services
 - The "IP Configuration Guide", Appendix B SSL/TLS Security lists resource profiles in CSFSERV class for TLS
- **For more information on using RACF to protect ICSF cryptographic keys and services, please see:**
 - ▶ "z/OS V1R6.0 ICSF Administrator's Guide", SA22-7521



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

z/OS FTP and ICSF resource access

- **When FTP sessions are secured using SSL/TLS, and hardware encryption through ICSF is used by system SSL, then the involved address space user IDs need access to the relevant CSFSERV and CSFKEYS resources**
- **All resources are important, but the most important resources to protect are in general the private keys that are used to sign and authenticate users and messages**
- **The FTP daemon user ID needs access to the FTP server's private key in order to complete the SSL/TLS handshake for the control connection and sign messages sent by the server on the control connection.**
- **The same private key is also used to perform the handshake for the data connection, but at that point in time, the address space is executing under the logged-in user's identity and no longer under the FTP daemon's identity.**
- **Originally, the only way the data connection handshake could succeed in such an environment was to permit all users, who were to use secure FTP connections, access to the FTP daemon's private key.**
 - ▶ An unacceptable security exposure to most installations.
- **For z/OS FTP client jobs that use client authentication, a user-specific private key (and certificate) is needed - and in that case that individual user needs access to that user's specific private key.**
 - ▶ Which is what we expect and not in any way a security exposure



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Intermediate solution before z/OS V1R7 and proper solution in z/OS V1R7

> z/OS FTP APAR PQ80574

- ▶ Interim fix
 - Only the FTP daemon user ID needs access to CSF resources
 - Switching security context back and forth between system SSL calls

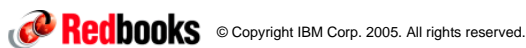
> PTF List:

- ▶ Release 120 : UQ86659
- ▶ Release 140 : UQ86660
- ▶ Release 150 : UQ86661

> APAR solution was integrated into V1R6

> FTP in z/OS V1R7 will exploit a new RACF function that is referred to as delegated resource profiles

- ▶ TLS protected sessions only
- ▶ Delegated Resource Profiles:
 - New V1R7 RACF function
 - RACF profiles that are marked 'RACF-delegated' are treated differently
- ▶ Faster and more secure than intermediate APAR solution
- ▶ Not specific to FTP
 - Will work for applications that use a similar daemon-server model as FTP does



ibm.com/redbooks

How delegated resource profiles work

Resource profile

FTP server user not permitted to resource? NO!

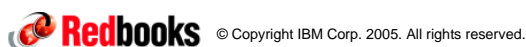
access denied!

Delegated resource profile

FTP server user not permitted to resource? NO!

FTP daemon permitted to resource? YES!

Access approved!



ibm.com/redbooks

Enabling delegated resource profiles for FTP when using SSL/TLS

> Mark resource profiles as delegated

- ▶ CSFSERV, CSFKEYS classes
- ▶ Done by specifying APPLDATA('RACF-DELEGATED') in resource definitions
 - RALTER CSFSERV CSFENC APPLDATA('RACF-DELEGATED')

> Permit FTP daemon to resource profiles

- ▶ READ access required
 - PERMIT CSFENC CLASS(CSFSERV) ID(FTPD) ACCESS(READ)

> Revoke FTP login user access to resources

- ▶ Also called client access
 - PERMIT CSFENC CLASS(CSFSERV) ID(FTPUSER) DELETE

> Refresh CSFKEYS and CSFSERV classes

Sample RACF commands are in the EZARACF sample job in hlq.SEZAINST

Things to think about if using FTP SSL/TLS with ICSF

> If you are on z/OS levels prior to V1R6 and have APAR PQ80574 (or its PTF) installed or you have z/OS V1R6 installed and -

- ▶ FTP sessions are TLS protected?
- ▶ Cryptographic hardware is in use?
- ▶ Resource profiles in CSFSERV and/or CSFKEYS classes are defined?
- ▶ Do not want to permit FTP login user IDs to CSFKEYS and CSFSERV resources?

> Then you must migrate to delegated resource profiles in z/OS V1R7!

- ▶ FTP APAR PQ80574 in z/OS V1R6 is not supported in z/OS V1R7

> Application Transparent TLS (AT-TLS) avoids this problem (new function CS z/OS v1R7)

Mixed-case RACF password support in FTP

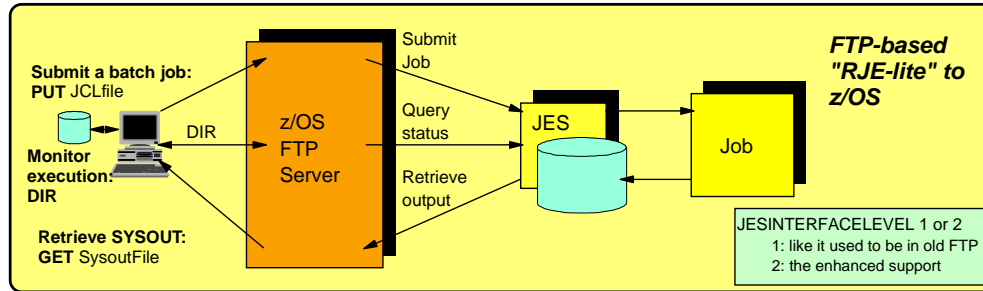
- **RACF in z/OS V1R7 supports mixed-case passwords**
 - ▶ Password possibilities increased
 - Still eight (8) characters
 - ▶ Likelihood of succeeding with a brute force attack on password combinations is reduced
- **z/OS FTP processes user-entered passwords in many scenarios and so far has always upper-case translated the passwords before handing them over the SAF interface.**
- **RACF administrator in z/OS V1R7 can toggle system-wide setting of password case support:**
 - ▶ SETROPTS PASSWORD(MIXED)
 - Passwords are used by FTP exactly as entered and handed over to SAF as entered
 - If stored password in RACF was last set when NOMIXED was active and never reset while MIXED was active, RACF will upper-case the passed value before checking for validity
 - ▶ SETROPTS PASSWORD(NOMIXED)
 - All passwords are upper-cased by FTP before handing over to SAF (as pre-V1R7)
- **FTP will in z/OS V1R7 adapt accordingly**
 - ▶ No new FTP configurations or options
- **You need to carefully evaluate FTP password usage before trying mixed-case passwords**
 - ▶ Remember the RACF options are system-wide and there are many other subsystems and applications to analyze before enabling mixed-case passwords

Things to think about for FTP if enabling mixed-cased passwords

- **Interactive FTP client**
 - ▶ Educate interactive users!
 - ▶ Reply to password prompt with correct-case password!
- **FTP client using NETRC data set**
 - ▶ Code passwords in correct-case in NETRC data set
- **FTP client API programs**
 - ▶ Code password in correct case
- **FTP client batch jobs**
 - ▶ Code passwords in INPUT file in correct case
- **REXX programs stacking FTP client commands**
 - ▶ Code passwords in correct case in REXX programs
- **FTP server FTP.DATA ANONYMOUS statement**
 - ▶ If a password is coded on this statement, make sure it is coded in the correct case
 - ▶ Same consideration if you use the ANONYMOUS keyword on the EXEC PARM field when starting the FTP daemon - must be coded in correct case

For security reasons, FTP does not provide any traces that will print the password values. If RACF fails a login request because of mixed case problems, you will not be able to diagnose that using traditional FTP debugging technologies. For FTP client jobs, you can check all the locations above.

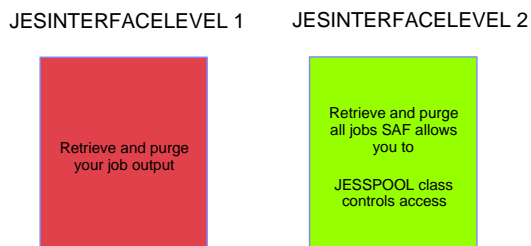
JES Interface Support in the z/OS FTP Server



Characteristics of JESINTERFACELEVEL 2 support:

- The JES interface uses the SAPI subsystem interface to JES2 and JES3
- All JES types (Jobs, started tasks, TSO, APPC)
- No jobname restrictions
- Details on DIR command output for jobs in input, active, or output status
- Filtering and access based on SAF interface - JESSPOOL and selected SDSF ISFCMD SAF profiles - allows access to jobs owned by others if proper SAF profiles are defined
- Filtering of jobs selected for DIR output is controlled via three SITE options:
 - JESJOBNAME (default <userID>*)
 - JESOWNER (default <userID>)
 - JESSTATUS (default ALL or OUTPUT or INPUT - determined by access to ISFCMD resources)
- **Use of this interface (SAPI) prior to z/OS V1R7 required UPDATE access to JESSPOOL resources**

JESSPOOL resource access control - READ access is enough, no need for UPDATE access to just browse spool data sets



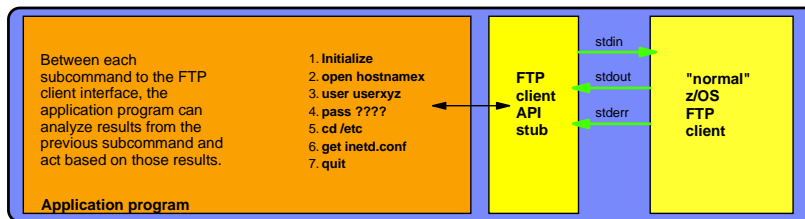
> Both SDSF and FTP have relaxed access requirement for JESSPOOL resources:

- Prior to z/OS V1R7, the user had to have:
 - UPDATE access in order to display or retrieve job
- From z/OS V1R7, the user only needs to have:
 - READ access in order to display and retrieve jobs
 - In order to purge jobs, the user still needs UPDATE access

**Remember that SAPI is a JES2-only technology.
This does not work with JES3.**

z/OS FTP client programming interface for improved automation and integration of z/OS file transfers - C API added in z/OS V1R7

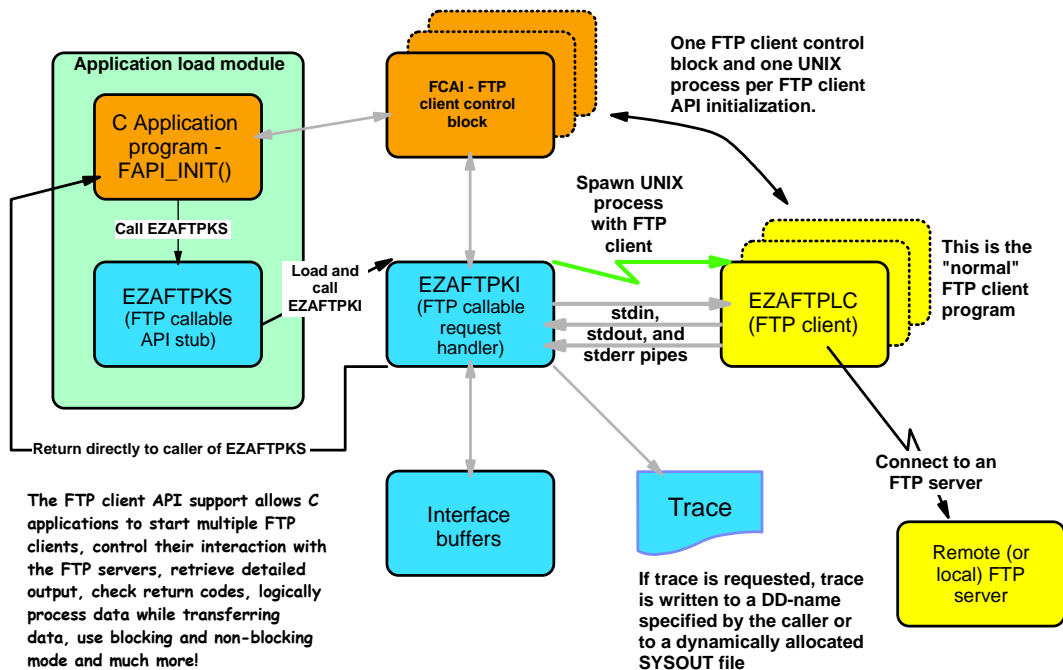
- Provides an interface that allows an application to programmatically invoke the FTP client on z/OS from common environments (UNIX shell, TSO, or MVS batch job)
- Characteristics of the interface:
 - z/OS V1R6 provides a callable interface to be used from assembler, Cobol, PL/I (or any z/OS supported programming language that supports a call interface)
 - **z/OS V1R7 adds a C API version of this programming interface**
 - Interface is reentrant and does support multiple parallel FTP client sessions by tasks within an address space
 - For communication between the program and the interface, a simple set of commands and data areas are used (Mappings for common programming languages are provided)
 - Both blocking (wait for a response), and non-blocking (polling-mode) calls are supported
 - In non-blocking mode, progress notifications can be returned to the calling application as the transfer progresses
 - The simple commands tell the interface what to do, for example: initialize, terminate, execute an FTP client subcommand, process output from the FTP client subcommand that was executed, poll for command completion
 - Results are returned as structured fields in communication area control blocks (return codes from interface and server replies or possibly local subcommand) along with free-format replies from the FTP client code
 - Debugging options are provided



C FTP client API added in z/OS V1R7

- **z/OS V1R7 extends the FTP client API with a C programming interface**
 - A C header file
 - Map the FCAI_MAP control block
 - Provide defines for the constants
 - Provide C static in-line stubs
 - A C sample file
 - The FTP C API is based on the existing FTP callable API.
 - APIs take character string commands
 - Blocking and unblocking commands are supported for flexibility
 - Return codes are grouped into categories to ease program logic
 - Posix and non-Posix environments are supported
- **Application programs that use this interface, should run in a POSIX environment.**
 - Running programs in a non-POSIX environment may provide unpredictable results.

FTP client API structure



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

FTP file transfer confidence reporting

- **The FTP protocol assumes that a data transfer is completed when the sending side of a data transfer closes the sending socket.**
 - There have been cases in the past where an erroneous close by a remote FTP client or server has made the z/OS FTP client or server believe a data transfer had successfully completed, when in fact only parts of the file or data set had been transferred to z/OS.
 - Note that if the sending site abends or the sending system crashes, such a condition will be detected by the TCP transport layer and error-return codes will be passed to the receiving program, which will abort the receive operation accordingly.
- **How can a user determine with some level of certainty that the transfer of a file in structure file and mode stream completed successfully?**
- **Calculate a confidence-of-success level for each file transfer**
 - Report the confidence level to the user
- **There are three ways in which the confidence level can be conveyed:**
 - FTP server logging
 - Requires the FTP.DATA statement FTPLOGGING to be set to TRUE
 - Uses message EZYFS86I
 - FTPOSTPR user exit
 - Indication in the parameter list to the exit routine
 - FTP client message sent to user "console" (TSO, UNIX shell, OUTPUT file, etc.)
 - Uses message EZA2108I
- **New FTP.DATA option for both z/OS FTP client and z/OS FTP server:**
 - CHKCONFIDENCE TRUE or FALSE (default is FALSE)

© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

FTP file transfer confidence levels

➤ Confidence checking has 5 levels of granularity:

- ▶ Confidence is **HIGH**
 - No errors or anomalies were detected.
- ▶ Confidence is **NOEOF**
 - An EOF marker was not found in STRUCT R or MODE B or C transfer.
- ▶ Confidence is **LOW**
 - The client did not respond following the transfer or an error was reported.
- ▶ Confidence is **UNKNOWN**
 - Outbound transfers and only set if checking is active.
 - Outbound transfers are given an UNKNOWN confidence level as the highest confidence level that they may obtain. This is due to the fact that:
 - The receiver is not necessarily a z/OS FTP client or server
 - We have no mechanism for determining with 100% accuracy that the file transfer completed successfully
 - If an error is detected on an outbound transfer, then the confidence level is lowered from UNKNOWN to LOW.
- ▶ Confidence checking is **INACTIVE**
 - Only reported to the FTPOSTPR user exit when CHKCONFIDENCE is set to FALSE.

Reporting of confidence level

➤ FTP server logging

- ▶ Confidence level is reported in syslog (such as /tmp/syslog.log)

```
Sep 22 15:20:41 MVS049 ftps[83886110]: EZYFS86I ID=FTPD10003 TRANS  
Confidence=High
```

➤ FTP client message

- ▶ Confidence level is reported to the client

```
get myfile /tmp/myfile  
EZ A1701I >>> EPSV  
229 Entering Extended Passive Mode (|||1117|)  
EZ A1701I >>> RETR myfile  
125 Sending data set USER1.MYFILE  
250 Transfer completed successfully.  
EZ A2108I Confidence=High for GET of /tmp/myfile  
EZ A1617I 154 bytes transferred in 0.010 seconds. Transfer rate 15.40 Kbytes/sec
```


FTPOSTPR exit routine interface change

>FTPOSTPR user exit

- ▶ Added confidence level
 - Pointer at offset +76 bytes (19th parameter)
 - 1 byte field
 - Expected values:
 - X'00'Confidence level is High
 - X'01'Confidence level is NoEOF
 - X'02'Confidence level is Low
 - X'03'Confidence level is Unknown
 - X'04'Confidence level checking is not active

>Users of FTPPOSTPR user exit may have a migration issue

- ▶ New parameter is always sent to the exit routine
 - New parameter added at the end of the existing parameter list, so "properly" written FTPPOSTPR exit routines would not be affected by it.

End-of-line byte sequence for FTP text transfers

>Encoding schemes for character data:

- ▶ SBCS - Single-Byte Character Set
 - 1 byte per character
- ▶ DBCS - Double-Byte Character Set
 - 2 bytes per character
- ▶ MBCS - Multiple-Byte Character Set
 - Typically 2 or more bytes per character

>EOL - End-of-line termination character

- ▶ Refers to the character(s) following a line of data that denote its end
- ▶ Exact byte value depends on encoding

>The FTP protocol as defined in RFC 959 DEMANDS that the EOL sequence be a Carriage Return character followed by a Line Feed character - <CRLF> sequence

- ▶ For SBCS ASCII that is a x'0D0A' byte sequence
- ▶ For UNICODE UCS-2 (MBCS) that is a x'000D000A' byte sequence
- ▶ Only a few customers have requested alternatives to the standard CRLF sequence

>z/OS V1R7 adds a configurable EOL termination selection for outbound transmission of ASCII data in stream mode

- ▶ SBCS support and MBCS support
- ▶ DBCS is NOT supported
- ▶ Most customers will not require this option
- ▶ Any use of this option should be planned out carefully

EOL character alternatives

➤ **There are four EOL terminators to choose from:**

- ▶ CRLF - Carriage Return Line Feed
 - Default that customers have always used in the past
- ▶ CR - Carriage Return only
- ▶ LF - Line Feed only
- ▶ NONE - No EOL terminator

➤ **SBSENDEOL**

- ▶ Used for selecting the SBCS EOL terminator

➤ **MBSSENDEOL**

- ▶ Used for selecting the MBCS EOL terminator

| | SBCS or MBCS line of data | EOL Terminator | FTP.DATA Entry | |
|--|------------------------------|-------------------|-----------------------------------|------------------------|
| Carriage Return & Line Feed (CRLF) | | x'0D' x'0A' | SBSENDEOL CRLF MBSSENDEOL CRLF | (default) (default) |
| Carriage Return Only (CR) | | x'0D' | SBSENDEOL CR MBSSENDEOL CR | |
| Line Feed Only (LF) | | x'0A' | SBSENDEOL LF MBSSENDEOL LF | |
| No EOL Terminator (NONE) | | | SBSENDEOL NONE MBSSENDEOL NONE | |

How to specify the EOL character to use for outbound transfers

➤ **There are three methods for selecting the SBCS and MBCS EOL terminator:**

- ▶ FTP.DATA statement
 - SBSENDEOL CRLF
 - MBSSENDEOL NONE
- ▶ SITE option
 - SITE SBSENDEOL=CR
 - SITE MBSSENDEOL=CRLF
- ▶ LOCSITE option
 - LOCSITE SBSENDEOL=NONE
 - LOCSITE MBSSENDEOL=LF

➤ **Control connection**

- ▶ Is NOT affected by these settings

➤ **Data connection**

- ▶ Is affected by EOL terminator selection but only for outbound
- ▶ z/OS FTP always expects the CRLF sequence for inbound text data transfers

Things to think about if using anything but the default end-of-line byte sequence

- **Stream mode restarts and SBSENDEOL values other than CRLF do not work together**
 - ▶ Stream mode restart option relies on the SIZE command
 - ▶ SIZE command relies on the EOL terminator being CRLF
- **Before setting SBSENDEOL and MBSSENDEOL to other than CRLF, ensure the receiving client or server supports the new EOL characters.**
 - ▶ Only very special-case situations should use anything but CRLF
- **The SBSENDEOL and MBSSENDEOL CRLF setting is the default and the standard line terminator defined by RFC 959.**
 - ▶ The z/OS FTP server and FTP client can receive ASCII data only in this format.
 - ▶ CRLF is the required setting for data sent to a z/OS FTP server or FTP client.

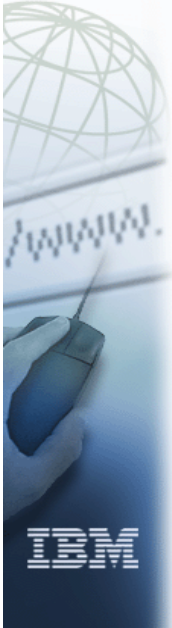
Note that FTP on z/OS cannot improve the reliability of FTP transfers - to do so would require changes to the underlying FTP protocol that would make z/OS FTP incompatible with all other FTP clients and servers. z/OS FTP can improve the reporting of how reliable a file transfer to or from z/OS appears to be.

This page intentionally left blank

ibm.com



e-business



Communications Server for z/OS V1R7 - Technical Update Integrated IP Security



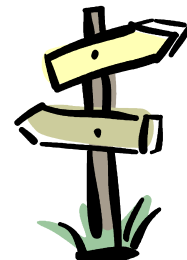
Redbooks

International Technical Support Organization

© Copyright IBM Corp. 2005. All rights reserved.

Integrated IP security agenda

- > Integrated IP security on z/OS - introduction
- > Integrated IP security - IP filtering
- > Integrated IP security - IPSec Virtual Private Networks (VPNs)
- > Configuring and enabling integrated IP Security
- > Preliminary performance data



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

What are the tools in the CS z/OS V1R7 IP security toolbox?

NOTES

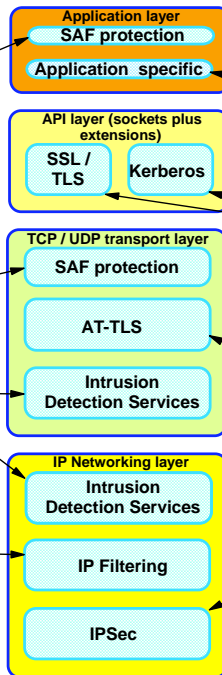
Protect the system

z/OS CS TCP/IP applications use SAF to authenticate users and prevent unauthorized access to data sets, files, and SERVAUTH protected resources.

The SAF SERVAUTH class is used to prevent unauthorized user access to TCP/IP resources (stack, ports, networks).

Intrusion detection services protect against attacks of various types on the system's legitimate (open) services. IDS protection is provided at both the IP and transport layers.

IP filtering blocks out all IP traffic that this system doesn't specifically permit.



Protect data in the network

Examples of application protocols with built-in security extensions are SNMPv3, DNS, and OSPF.

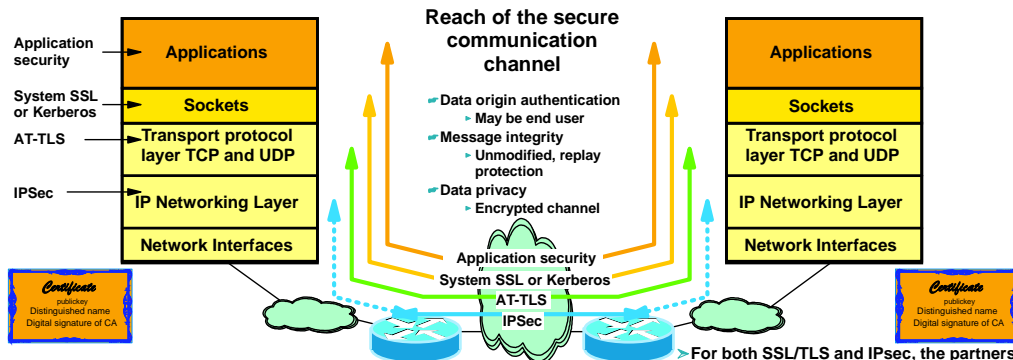
Both Kerberos and SSL/TLS are located as extensions to the sockets APIs and applications have to be modified to make use of these security functions. Both SSL/TLS and Kerberos are connection-based and only applicable to TCP (stream sockets) applications, not UDP.

AT-TLS is a TCP/IP stack service that provides SSL/TLS services at the TCP transport layer and is transparent to applications.

IPSec resides at the networking layer and is transparent to upper-layer protocols, including both transport layer protocol and application protocol.



Secure network communication between two endpoints



- > **Message level security at the application layer**
 - Secure Network Services (SNMPv3, Secure DNS) - always end-to-end
- > **Socket layer security service**
 - TLS/SSL, Kerberos - always end-to-end and TCP only
- > **Transparent application security services over IP network**
 - IPSec provides blanket protection for all IP applications - end-to-end or segment of data path between two VPN routers
 - Application transparent TLS - provides connection level protection for TCP applications - always end-to-end
 - SSL TN3270 securely extends reach of SNA applications over IP network - TN3270 client to TN3270 server data path secured
 - SNA Session Level Encryption - secures SNA session traffic transparently between two SNA application end points (LUs) of which one end could be the TN3270 server and the other the final SNA application

> For both SSL/TLS and IPSec, the partners can authenticate each other based on digital certificates, and perform a public/private key encryption handshake to generate and exchange a symmetric encryption session key to be used for encrypting and decrypting the data between the partners.

- SSL/TLS defines this as the handshake phase
- For IPSec it is part of the IKE negotiation

What is z/OS Firewall Technologies?

- **The z/OS Firewall Technologies were originally ported from a non-z/OS environment.**
 - ▶ Focus was traditional firewall capabilities.
 - ▶ Today's z/OS IP security focus is more directed towards "self protection".
- **z/OS Firewall Technologies have been available since OS/390 V2R4 and are today shipped partly with the Communications Server and partly with the Integrated Security Services component of z/OS.**
- **Most of the functions are useful both in a traditional firewall configuration and as self-protection functions on z/OS.**

| The firewall technologies functions that are shipped with z/OS | Included in Communications Server | Included in Integrated Security Services | Useful in firewall configuration | Useful as self-protection layer in z/OS |
|--|-----------------------------------|--|----------------------------------|---|
| IPv4 packet filters | ✓ | | ✓ | ✓ |
| IPv4 IPSec (VPN) | ✓ | | ✓ | ✓ |
| IPv4 Network Address Translation | ✓ | | ✓ | |
| Internet Key Exchange (IKE) | | ✓ | ✓ | ✓ |
| Command-line configuration | | ✓ | ✓ | ✓ |
| GUI configuration | | ✓ | ✓ | ✓ |
| FTP Proxy server | | ✓ | ✓ | |
| SOCKS V4 server | | ✓ | ✓ | |



© Copyright IBM Corp. 2005. All rights reserved.

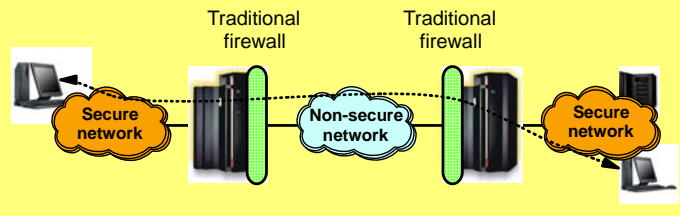
ibm.com/redbooks

Firewall technologies usage Scenarios on z/OS

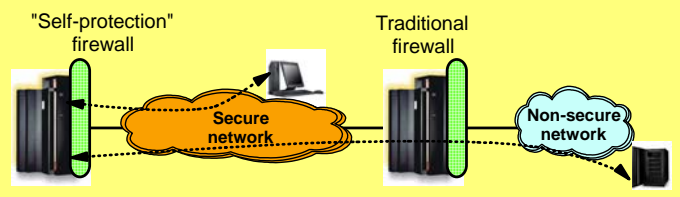
NOTES

- You can choose to use the z/OS Firewall Technologies to set up a traditional firewall structure where the firewall(s) reside in a z/OS LPAR
 - ▶ Isolates secure networks from non-secure networks
 - ▶ Provides the first line of defense from outside attacks
 - ▶ Utilizes IP Security function
 - ▶ May also utilize techniques to "hide" internal (secure) addresses from the external (non-secure) world
- You can also choose to use the z/OS Firewall Technologies on your normal z/OS LPARs to
 - ▶ Provides protection from secure network
 - ▶ Provides additional protection from non-secure network
 - ▶ Address hiding techniques are not applicable

Traditional firewall configuration



Self-protection configuration



© Copyright IBM Corp. 2005. All rights reserved.

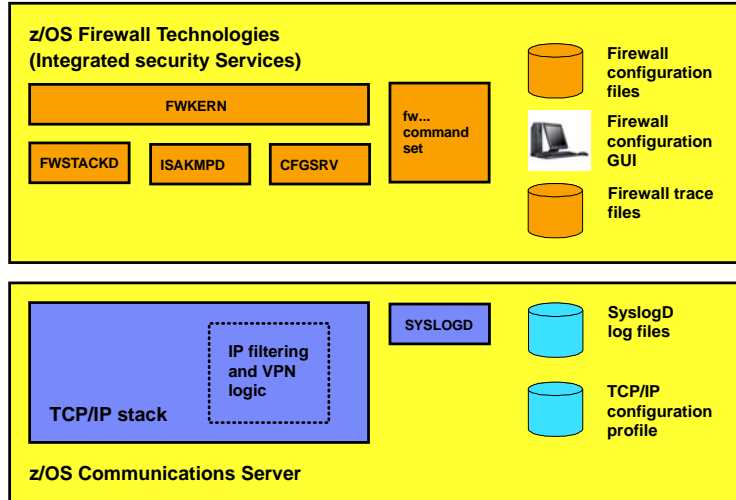
ibm.com/redbooks

Customer-identified issues with the current firewall technologies implementation

NOTES

➤ **z/OS Communications Server prior to z/OS V1R7 does not provide all the elements required for IP Security**

- ▶ z/OS Firewall Technologies must be installed and configured
- ▶ Documentation split across multiple z/OS elements
- ▶ Configuration does not exploit z/OS Communications Server configuration techniques
 - Policy Agent (Pagent)
- ▶ Firewall command set is large
- ▶ Overhead to maintain firewall servers (fwkern, fwstackd, isakmpd, and cfgrsv)
- ▶ Service ambiguity
 - Which service group is responsible for an IP Security problem
- ▶ Scalability and performance in general

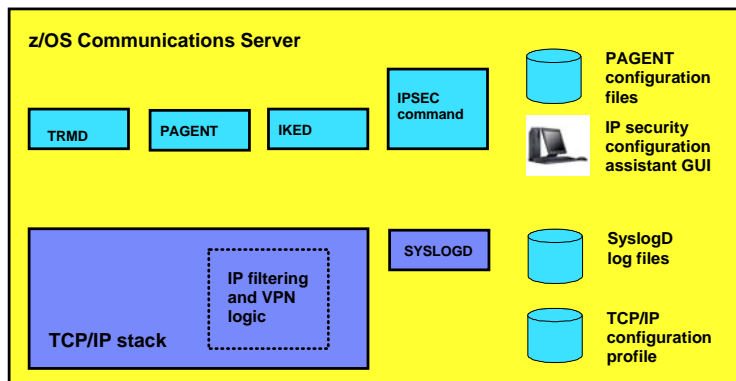


Integrated IP Security in z/OS V1R7

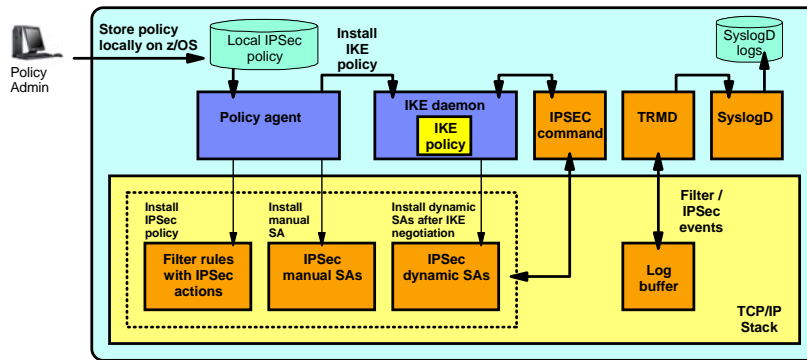
NOTES

➤ **Provide a z/OS Communications Server alternative to using the z/OS Firewall Technologies IP Security support**

- ▶ Provide a Communications Server equivalent to z/OS Firewall Technologie's ISAKMPD
- ▶ Eliminate the need to run z/OS Firewall Technologies's fwkern, fwstackd, and cfgrsv
- ▶ Provide a Pagent-based configuration file to replace the existing z/OS Firewall Technologies configuration commands.
- ▶ Provide one new UNIX System Service command to replace the multiple existing z/OS Firewall Technologies IP security management commands.
- ▶ Continue to ship the z/OS Firewall Technologies IPsec support in z/OS V1R7
 - In a future release z/OS Firewall Technologies will no longer be shipped (this includes the IP Security functions and the additional traditional firewall functions (NAT, SOCKS, and FTP proxy))
- ▶ A stack can use Integrated IPsec or z/OS Firewall Technologies IPsec, but not both
- ▶ In z/OS V1R7, Integrated IP Security is an IPv4-only solution



Integrated IP security infrastructure in z/OS V1R7



> Integrated IP security in z/OS V1R7 covers:

- ▶ IP filtering
- ▶ Virtual private networks based on IPsec

> Configuration support

- ▶ Optimized for z/OS host-to-host and z/OS host-to-gateway (z/OS gateway still supported)
- ▶ NAT IP address traversal support

> Simplified infrastructure

- ▶ Eliminates need for FW Technologies daemons
- ▶ Policy agent reads and manages IPsec and IKE policy

> Simplified configuration

- ▶ New configuration GUI for both new and expert users
- ▶ Direct file edit into local configuration file
- ▶ Reduced definition, more "wildcarding"

> Improved serviceability

- ▶ Improved messages and traces

> Default filters part of TCP profile

- ▶ More granular control before policy is loaded

> Administrative controls

- ▶ pasearch, new IPsec command



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Integrated IP security - RFC standards

NOTES

> Ability to control and protect IP traffic on one or more TCP/IP stacks

- ▶ Accomplished by:
 - IP filtering
 - Permitting or denying specific IP traffic patterns
 - Virtual Private Networks (VPNs)
 - Authenticating and/or encrypting data associated with a specific IP data pattern
- ▶ Based on RFCs defined by the IETF IPsec working group

> IPsec RFCs implemented by Integrated IP Security include

- ▶ RFC 2401: Security Architecture for the Internet Protocol
- ▶ RFC 2402: IP Authentication Header
- ▶ RFC 2403: The Use of HMAC-MD5-96 within ESP and AH
- ▶ RFC 2404: The Use of HMAC-SHA-1-96 within ESP and AH
- ▶ RFC 2406: IP Encapsulating Security Payload (ESP)
- ▶ RFC 2407: The Internet IP Security Domain of Interpretation for ISAKMP
- ▶ RFC 2408: Internet Security Association and Key Management Protocol (ISAKMP)
- ▶ RFC 2409: The Internet Key Exchange (IKE)
- ▶ RFC 2410: The NULL Encryption Algorithm and Its Use with IPsec
- ▶ RFC 2451: The ESP CBC-Mode Cipher Algorithms
- ▶ RFC 3947: Negotiation of NAT-Traversal in the IKE
- ▶ RFC 3948: UDP Encapsulation of IPsec ESP Packets



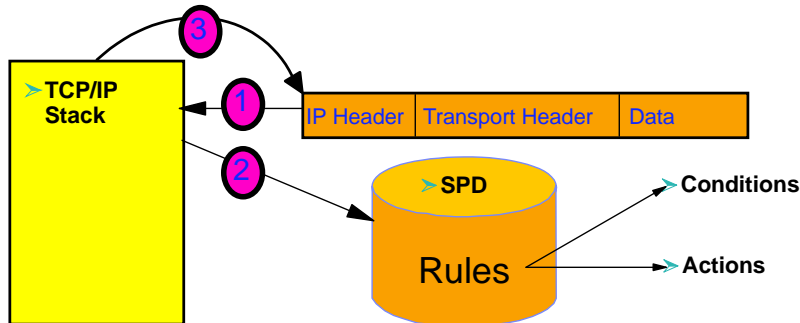
© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

IP filtering overview

NOTES

- > **Inbound or outbound IP packet arrives**
 - Can be applied for both local and routed traffic
- > **Consult rules in a Security Policy Database (SPD)**
 - Rules have conditions and actions
- > **Apply action of matching rule to packet**
 - Deny
 - Permit
 - Permit with additional processing applied



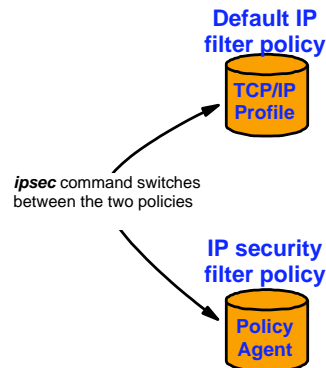
© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Integrated IP security - filter policies

> Integrated IPsec's Security Policy Database (SPD)

- Default IP filter policy
 - Intended to allow limited access while IP security filter policy is being loaded
 - Can be reverted to in an "attack" situation
 - Defined in the TCP/IP profile
 - Default is to deny all traffic
 - Provides basic filtering function
 - Permit rules only
 - No VPN support
- IP security filter policy
 - Intended to be the primary source of filter rules
 - Defined in a Policy Agent IPsec configuration file
 - Can be generated by the z/OS IP Security Configuration Assistant GUI
 - Default is to deny all traffic
- *ipsec* command is used to switch between default and IP security filter policy



> Requires the IPSECURITY option on the IPCONFIG statement

- IPSECURITY option enables use of the new integrated IP security functions
- The IPSECURITY option is mutually exclusive with the FIREWALL option
 - Separate FIREWALL and IPSECURITY stacks may coexist on one z/OS image

> Implicit filter rules

- Always present, not user-defined
 - Deny all inbound traffic
 - Deny all outbound traffic
- Appended to Default IP filter policy by the stack
- Appended to IP Security filter policy by Pagent
- If neither policies are defined, the implicit rules become the default policy (deny all)



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

A little more details on the default filter policy

NOTES

- **Provides initial protection of the stack during initialization**
 - Used until IP security filter policy is loaded
- **Generally restrictive; these user-defined rules should include**
 - Traffic needed for basic services
 - Examples
 - OMPROUTE
 - OSPF traffic
 - IGMP traffic
 - DNS queries
 - UDP traffic with a destination port of 53
 - Traffic needed to fix problems with IP security filter policy
 - Examples
 - FTP traffic from the workstation running the z/OS Network Security Configuration Assistant GUI
 - Telnet traffic from the network administrator's workstation
- **Implicit filter rules**
 - Always present, not user-defined
 - Deny all inbound traffic
 - Deny all outbound traffic
 - Appended to Default IP filter policy by the stack
 - Appended to IP Security filter policy by Pagent
 - If no policies are defined, the implicit rules become the default policy (deny all)

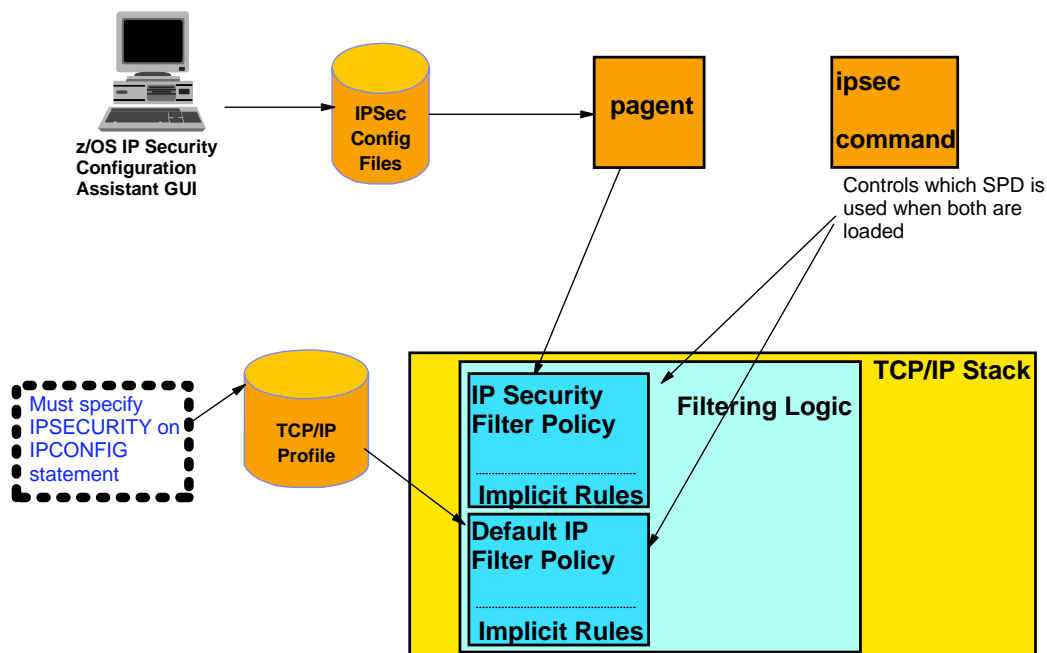
Filter rule search order

Filter rule 1
Filter rule 2
Filter rule 3

.....

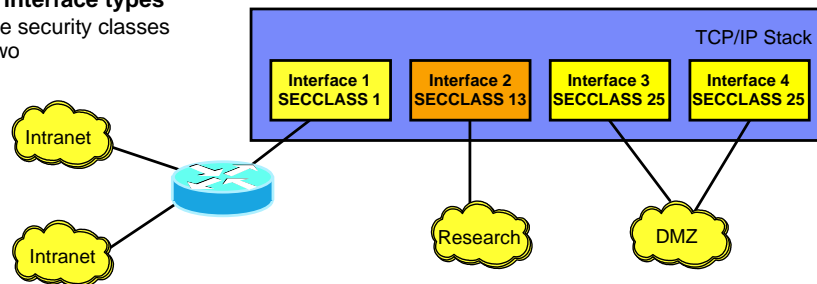
**Implicit filter rule:
Deny everything!!!**

IP filter policy on z/OS - overview



Interface security class (SECCLASS)

- Can be assigned only to non-virtual interfaces
- Defined in the TCP/IP profile
 - ▶ LINK statement (SECCLASS parameter)
 - ▶ IPCONFIG DYNAMICXCF statement (SECCLASS parameter)
- Value 1 to 255 (default is 255)
 - ▶ Value is just a classification identifier; it has no inherent meaning
 - Can be referred to in the filter rules
- Packets inherit the security class of the interface they traverse
- A more flexible and expandable mechanism than the traditional firewall's "secure" vs. "non-secure" interface types
 - ▶ 254 interface security classes instead of two



IP filter conditions - differences between the default and the security filter policy definitions

NOTES

| Criteria | Default IP Filter Policy | IP Security Filter Policy |
|-----------------|------------------------------|---|
| IP addresses | Single/Subnet | Single/ <i>Range</i> /Subnet |
| Protocol | Single/All | Single/All |
| Ports | Single/All for UDP and TCP | Single/ <i>Range</i> /All for UDP and TCP |
| Type | Single/All for ICMP and OSPF | Single/All for ICMP for OSPF |
| Code | Single/All for ICMP | Single/All for ICMP |
| Direction | Bidirectional | Bidirectional(1)/ <i>Inbound/Outbound</i> |
| Routing | Local | Local/ <i>Routed/Either</i> |
| Security Class | Single/Any | Single/Any |
| Time Conditions | Not Applicable | <i>Time Specification</i> |

Note: 1) Can control who initiates TCP connections

Text in italics above: highlights difference between the two policies

Filter actions

> Allowed actions for filter policies

| Default IP Filter Policy | IP Security Filter Policy |
|--|---|
| <ul style="list-style-type: none"> ✓ Permit | <ul style="list-style-type: none"> ✓ Permit ✓ Deny ✓ IPSec (both manual and dynamic) |

> Both policies allow filter logging to be enabled/disabled

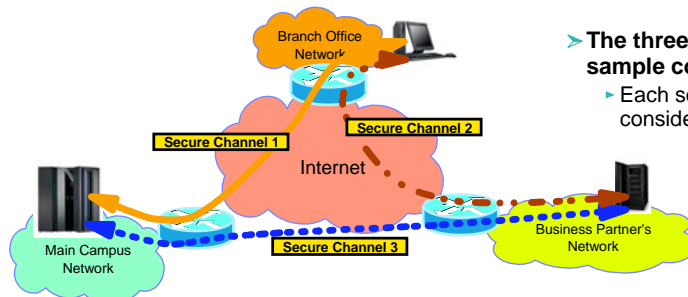
> IP security filter policies using an action of IPSec:

- ▶ Used to implement Virtual Private Networks (VPNs)
- ▶ Must be bidirectional
- ▶ Can only specify a security class of 0
 - Indicates the rule applies to all interfaces
- ▶ Require the definition of additional policy actions
 - Manual VPN actions
 - Dynamic VPN actions
- ▶ Based on Internet standards defined by the IPSec working group
 - RFC 2401 and related RFCs
- ▶ Packets matching an SPD rule with an IPSec action are modified to provide authentication and/or data encryption

IPSec Virtual Private Network (VPN) overview

> Virtual Private Network

- ▶ Logical network of connected nodes that communicate over unsecure networks using one or more secure channels



> The three secure channels in this sample configuration make up a VPN

- ▶ Each secure channel in itself can be considered a VPN

> A secure channel is commonly called an IPSec security association (SA) and uses authentication and/or encryption

- ▶ The term "tunnel" is also sometimes used in this context, but it is ambiguous and can be confused with tunnel vs. transport mode

> A secure channel provides point-to-point security

> Integrated IPSec utilizes IP security protocols defined by the IPSec working group

- ▶ RFC 2402 - IP Authentication Header (AH) protocol
 - Data authentication
 - IP header authentication
 - Data origin authentication
- ▶ RFC 2406 - IP Encapsulating Security Payload (ESP)
 - Data authentication
 - Data origin authentication
 - Data privacy

IPSec VPN concepts - encapsulation mode

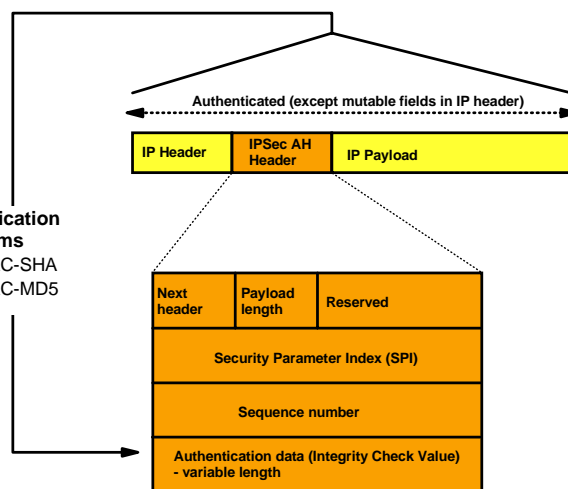
NOTES

- Indicates how to construct an IPSec packet
- Two modes
 - ▶ Transport mode
 - Inserts AH and/or ESP headers between original IP header and protected data
 - ▶ Tunnel mode
 - Creates a new IP header with an AH and/or ESP header
 - AH/ESP header followed by original IP header and protected data
- If one or both security endpoints are acting as a gateway
 - ▶ Tunnel mode must be selected
- If neither security endpoint is acting as a gateway
 - ▶ Tunnel or transport may be selected
 - ▶ Usually transport mode is used in this case
 - No need for extra cost of adding a new IP header in this case
- The counterpart to encapsulation is decapsulation

IPSec VPN concepts - Authentication Header (AH) protocol

NOTES

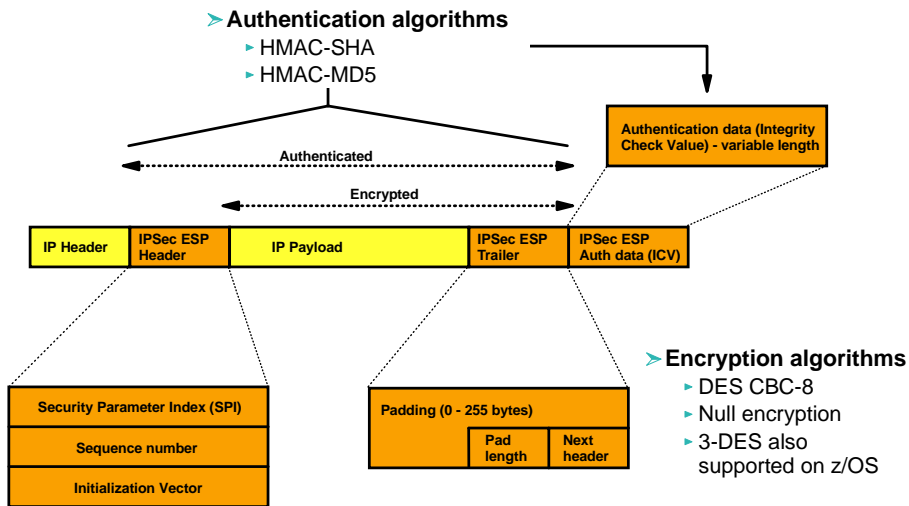
- Authentication algorithms
 - ▶ HMAC-SHA
 - ▶ HMAC-MD5



- If transport mode, then "Payload" contains the original transport header and original data
- If tunnel mode, then "Payload" contains the original IP header, original transport header, and original data

IPSec VPN concepts - Encapsulating Security Payload (ESP) protocol

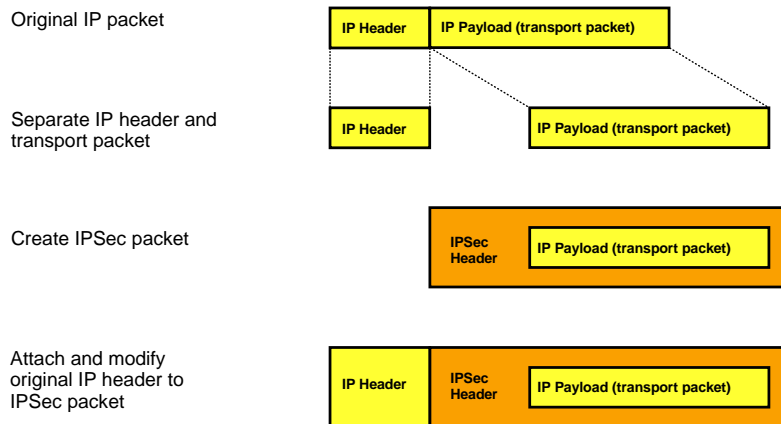
NOTES



- ▶ If transport mode, then "Payload" contains the original transport header and original data (possibly encrypted)
- ▶ If tunnel mode, then "Payload" contains original IP header, original transport header, and original data
 - ▶ "Payload" can be encrypted

IPSec VPN concepts - creating an IPSec packet using transport mode

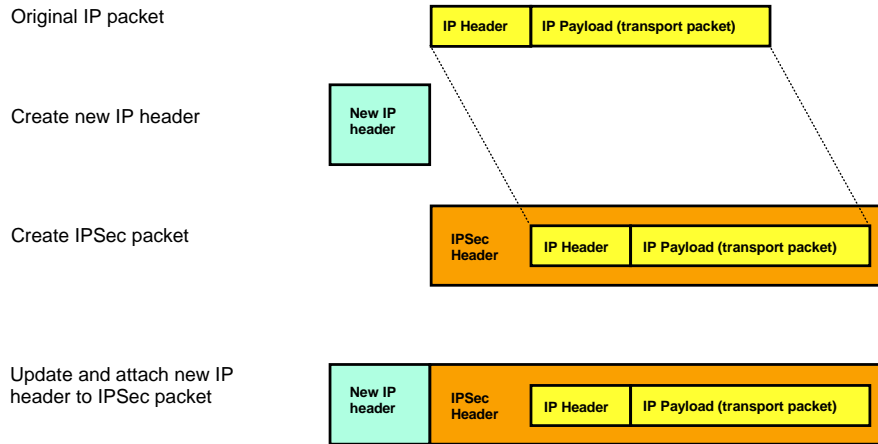
NOTES



Transport mode is typically used between two hosts that establish an IPSec VPN end-to-end between them.

IPSec VPN concepts - creating an IPSec packet using tunnel mode

NOTES



Tunnel mode is used if at least one of the two IPSec VPN endpoints is a gateway.

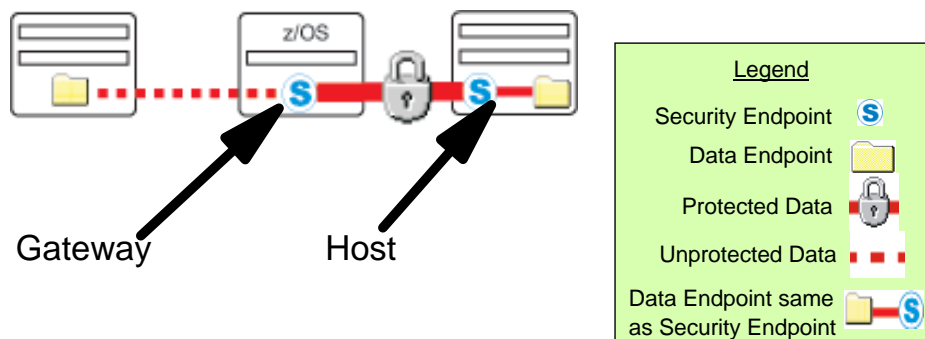
IPsec VPN concepts - security endpoint

> The endpoints of an IPSec secure channel

- Where IPSec protection is applied

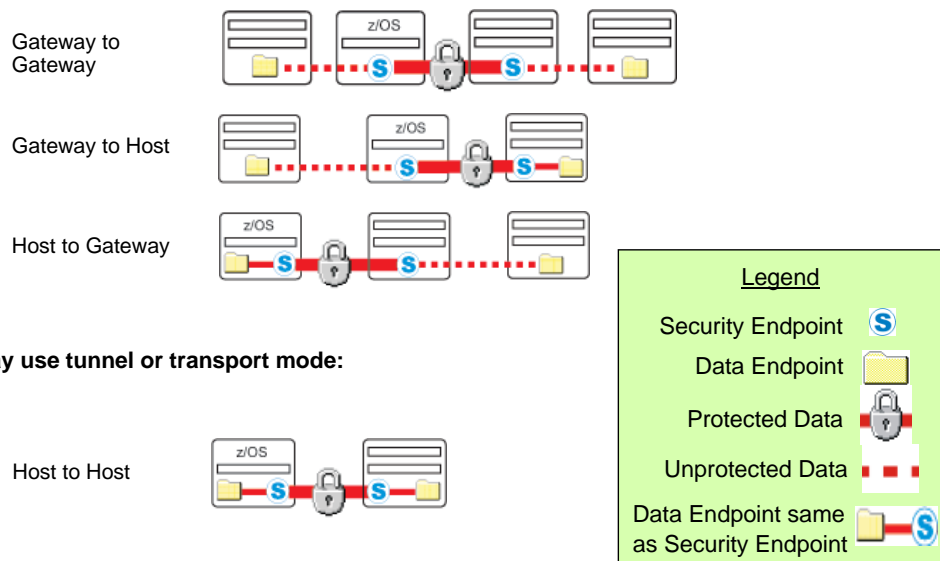
> Endpoint roles

- Host
 - ▬ Local data endpoint and secure channel endpoint are the same IP address
- Gateway (or Security Gateway)
 - ▬ Local data endpoint and secure channel endpoint are different IP addresses



IPsec VPN concepts - encapsulation mode rules

> Must use tunnel mode:



IPsec VPN concepts - security associations (SAs)

> IPsec secure channel endpoints must agree on how to protect traffic

- ▶ Security protocol
 - AH
 - ESP
- ▶ Algorithms to be used by the security protocols
 - Encryption Algorithm
 - DES or Triple DES
 - Authentication Algorithm
 - HMAC_MD5 or HMAC_SHA
- ▶ Cryptographic keys
- ▶ Encapsulation mode
 - Tunnel
 - Transport
- ▶ Lifetime/lifesize (for dynamic SAs)

> This agreement is known as a "security association" - or for short, an SA

IPSec VPN concepts - more about IPSec security associations (SAs)

NOTES

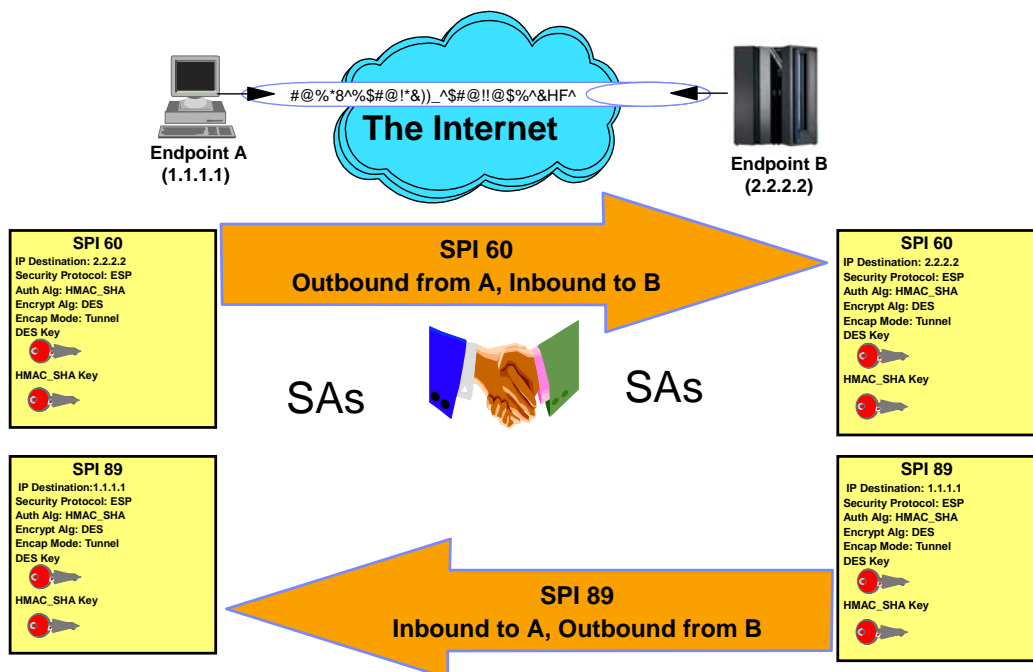
- > Used to protect IP traffic
- > Unidirectional
 - Need one for inbound and another for outbound - each IPSec secure channel endpoint consists of two SAs
 - Generally symmetrical with regards to algorithms used
 - Cryptographic keys will be different
 - A pair of matching SAs are on z/OS referred to as a "Tunnel ID" - in a sense identifying the secure channel
- > An SA is identified by:
 - A Security Parameter Index (SPI)
 - The SPI is a 32-bit value
 - SPI numbers in themselves may not be unique on a given IPSec node
 - The SPI is carried in the IPSec headers
 - IPSec protocol
 - Destination IP address information
- > Manually defined SAs
 - Statically defined in the Security Policy Database (SPD - Pagent IPSec config file)
- > Dynamically defined SAs
 - Negotiated using the Internet Key Exchange protocol
 - Acceptable values (policy) defined in the SPD (Pagent IPSec config file)
- > Security Association Database (SAD)
 - The collection of all SAs known to the stack



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

IPSec VPN concepts - IPSec security association example



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

IPSec VPN concepts - manually defined SAs

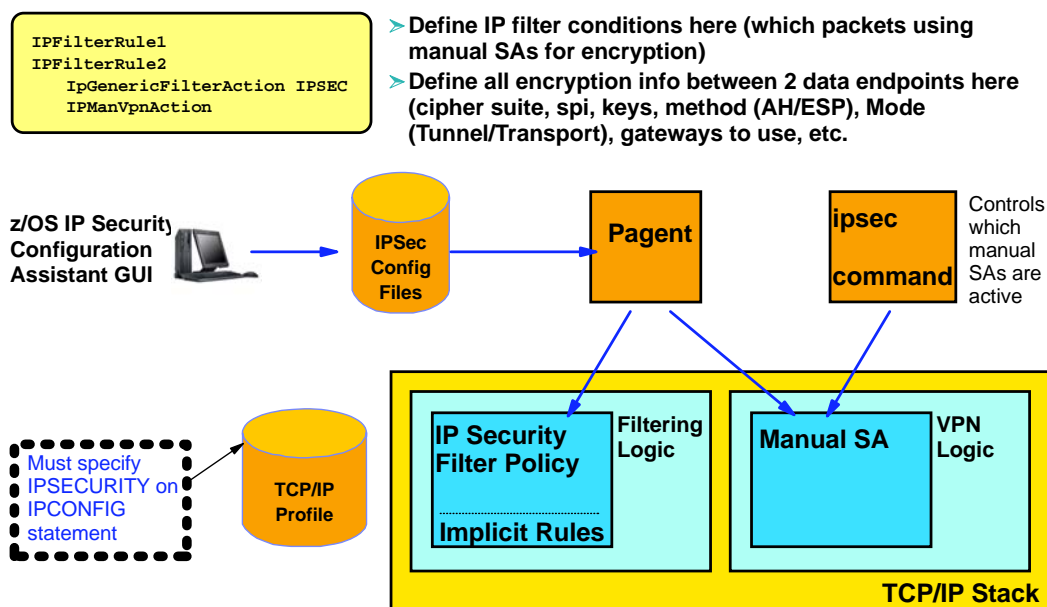
- **Not commonly used**
 - ▶ Do not provide a scalable solution
 - ▶ In the long run difficult to manage
- **Require the IPSECURITY option on the IPCONFIG statement**
 - ▶ Mutually exclusive with the FIREWALL option
- **Defined in a Pagent IPsec configuration file**
 - ▶ Cannot be used when default filter policy is in effect
 - ▶ Utilized by filter rules with an action of "ipsec"
 - ▶ SA is defined by a manual VPN action
 - Can be generated by the z/OS IP Security Configuration Assistant GUI
- **Use the ipsec command to activate/deactivate manual SAs**
 - ▶ Can also be automatically activated when policy is installed
- **Definition of SA attributes require mutual agreement between tunnel endpoint administrators**
 - ▶ Cryptographic keys and IPSec Security Protocol parameters must be mutually agreed to between tunnel endpoint administrators
 - ▶ Need to decide how to safely exchange keys (physical mail/courier service)
 - ▶ Need to decide how to refresh keys
 - Manual SAs must be deactivated and activated when refreshing keys
 - Refreshing keys must be coordinated with the remote tunnel endpoint's administrator
 - ▶ Remote endpoint may need to reactivate a manual SA if you locally deactivate the SA and then locally activate the SA.



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

IPSec VPN concepts - integrated IP Security manual SAs overview



© Copyright IBM Corp. 2005. All rights reserved.

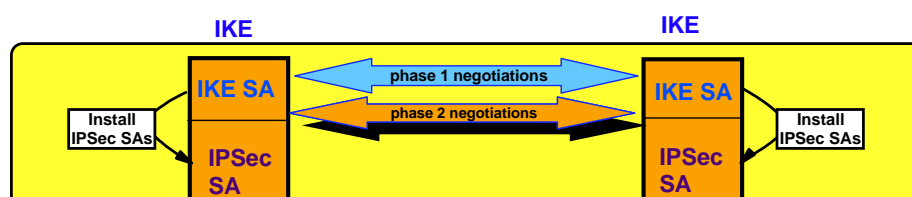
ibm.com/redbooks

IPSec VPN concepts - dynamically defined SAs

- **Currently state of the art**
 - ▶ Scalable
 - ▶ Initially requires more configuration than a manual SA
 - ▶ In the long run easier to manage
 - Set and forget it
- **Require the IPSECURITY option on the IPCONFIG statement**
 - ▶ Mutually exclusive with the FIREWALL option
- **Cannot be used when default filter policy is in effect**
- **Dynamic SAs are negotiated by the IKE daemon**
- **Dynamic IPSec VPN policy defined in a Pagent IPSec configuration file**
 - ▶ Can be generated by the z/OS IP Security Configuration Assistant GUI
 - ▶ Dynamic IPSec VPN action identifies "acceptable" SA attributes
 - Utilized by filter rules with an action of "ipsec"
 - ▶ Key exchange policy defines how to protect dynamic SA negotiations
- **The IKE daemon implements the Internet Key Exchange protocol**
 - ▶ Defined in RFC 2409
 - ▶ A two phase approach to negotiating dynamic IPSec SAs
- **The IKE daemon obtains its policy from Pagent**
 - ▶ Policy information for negotiating IPSec SAs
 - Dynamic IPSec VPN actions
 - ▶ Policy for creating a secure channel used to negotiate IPSec SAs
 - Key Exchange Policy
 - ▶ Policy for ipsec command activation and autoactivation
 - Local Dynamic IPSec VPN Policy
- **Utilizes UDP ports 500 and 4500 to communicate with remote security endpoints**
 - ▶ Negotiating SAs
 - ▶ Sending informational messages

IPSec VPN concepts - two phases of IKE negotiations

- **Phase 1 negotiation**
 - ▶ Creates a secure channel with a remote security endpoint
 - Negotiates an IKE SA
 - Generates cryptographic keys that will be used to protect Phase 2 negotiations and Informational exchanges
 - Authenticates the identity of the parties involved
 - Bidirectional, and not identified via SPIs
 - ▶ Requires processor-intensive cryptographic operations
 - ▶ Done infrequently
- **Phase 2 negotiation**
 - ▶ Negotiates a pair of IPSec SAs with a remote security endpoint
 - Generates cryptographic keys that are used to protect data
 - Authentication keys for use with AH
 - Authentication and/or encryption keys for use with ESP
 - ▶ Performed under the protection of an IKE SA
 - ▶ Done more frequently than phase 1



IPSec VPN concepts - IKE SAs

NOTES

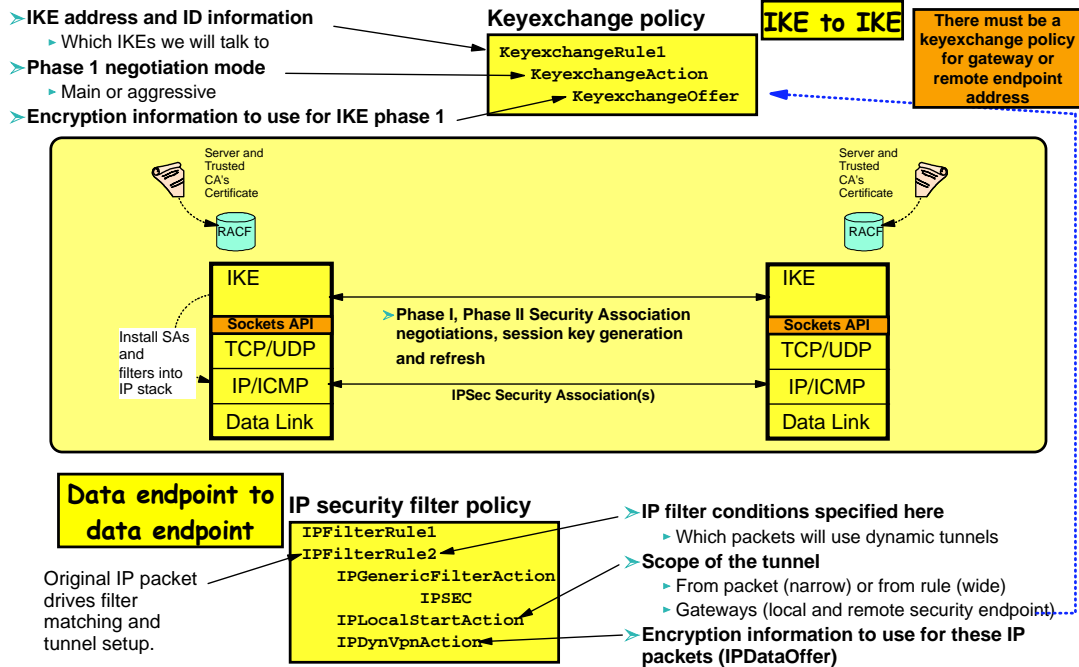
- **Used to protect Phase 2 negotiations**
- **Bidirectional**
- **Endpoints must agree on**
 - Encryption algorithm
 - DES/Triple DES
 - Hash Algorithm
 - MD5/SHA1
 - Authentication Method
 - Preshared Key
 - RSA Signature
 - Diffie-Hellman Group
 - Lifetime/Lifesize
- **Policy definition is based on identities exchanged during phase 1**
 - Key Exchange Policy
 - A set of filter rules for IKE

IPSec VPN concepts - dynamic SA activation

NOTES

- **Requires definition of local dynamic IPSec VPN policy:**
 - Command-line activation
 - ipsec -y activate command
 - Autoactivated
 - Activation attempted when a stack connects to IKED or when IP Security filter policy is reloaded
- **Does not require definition of local dynamic IPSec VPN policy:**
 - On-demand activation
 - Activation attempted when the stack receives an outbound packet requiring the protection of a new dynamic tunnel
 - Remote activation
 - A remote security endpoint initiates the negotiation of a new SA

Integrated IP security - on demand and remote activation policy highlights



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Dynamic IPsec VPN - on demand and remote activation policy highlights

NOTES

- > Key Exchange Policy - This is strictly for IKE-to-IKE flows. What IKEs we will talk to, what encryption to use to flow IKE to IKE data such as Phase I and Phase II negotiations.
 - Key Exchange Rule
 - Define IP filter conditions here for IKE; which IKE addresses and IDs will be used for Phase I negotiations - local and remote
 - Key Exchange Action
 - Whether to initiate phase I, and if so, whether to use main or aggressive mode. If responding, whether to use main or aggressive mode
 - Key Exchange Offer
 - Define what encryption information to use for Phase I negotiations
- > IPFilterrule - This is defining an encryption rule for a set of one or more data endpoints. The rule is composed of a set of filter conditions - which packets for which this rule applies, and a dynamic VPN action - what encryption to use when setting up the dynamic tunnels for this set of data endpoints.
 - IPGenericFilterAction IPFilterAction IPSEC
 - Must be entered to get dynamic VPN
 - IPLocalStartOption
 - This is where you define the scope of the Phase 2 negotiation. If you specify Packet, much of the information for the Phase 2 negotiation comes from the incoming packet. If you specify rule, it comes from the rule that matched the incoming packet
 - This is where you also specify which security endpoints to use - local and remote gateway addresses
 - IPDynVpnAction
 - IPDataOffer - here is where you specify the encryption information to use for the encryption for the data flow for this connection.



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Integrated IP security - command or autoactivated policy highlights

Command or autoactivated SAs require the existence of a LocalDynVPN policy - in addition to the security filter and keyexchange policies.

- > IKE address and ID information
 - Which IKEs we will talk to
- > Phase 1 negotiation mode
 - Main or aggressive
- > Encryption information to use for IKE phase 1

Keyexchange policy

KeyexchangeRule1
KeyexchangeAction
KeyexchangeOffer

IKE to IKE

There must be a keyexchange policy for gateway or remote endpoint address

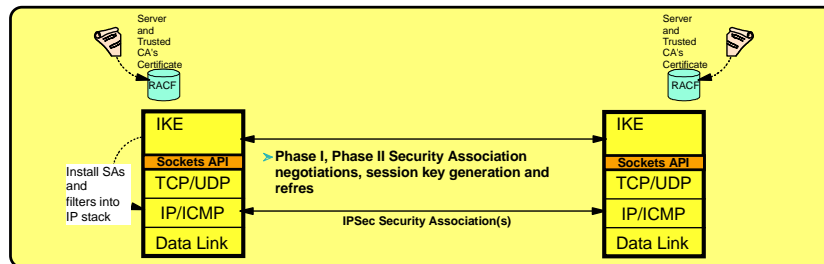
ipsec

Ipsec command controls which LocalDynVPN SAs are active

LocalDynVPN policy

LocalDynVpnRule
Autoactivate

- > Policy specification of IP addr info drives filter matching and tunnel setup



Data endpoint to data endpoint

Original IP packet drives filter matching and tunnel setup.

IP security filter policy

IPFilterRule1
IPFilterRule2
IPGenericFilterAction
IPSEC
IPLocalStartAction
IPDynVpnAction

- > IP filter conditions specified here
 - Which packets will use dynamic tunnels
- > Scope of the tunnel
 - From packet (narrow) or from rule (wide)
 - Gateways (local and remote security endpoint)
- > Encryption information to use for these IP packets (IPDataOffer)



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Dynamic IPsec VPN - command or autoactivated policy highlights

NOTES

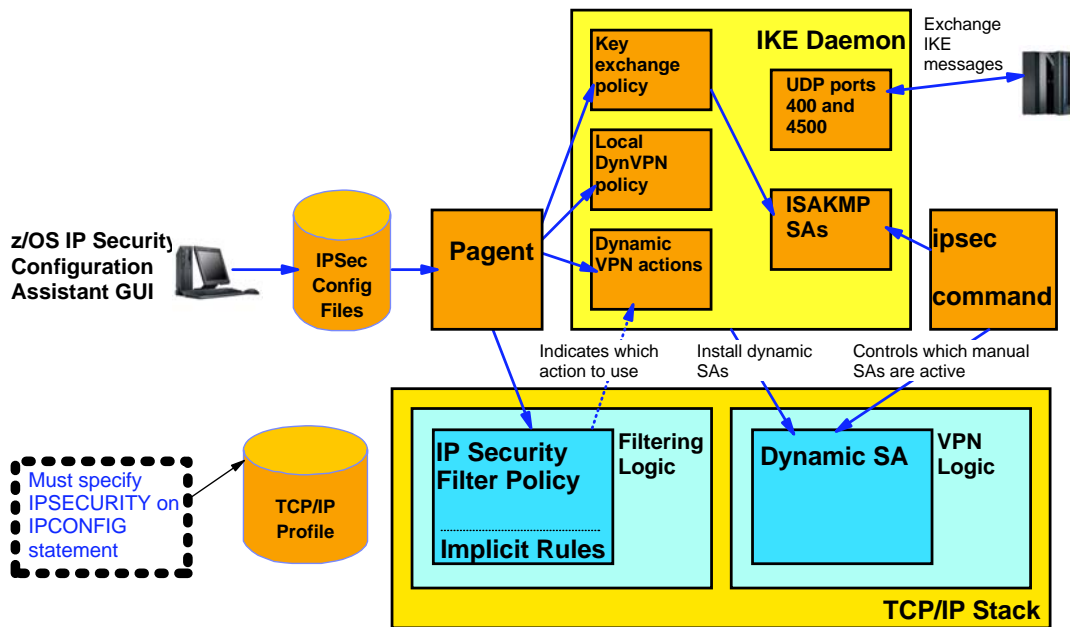
- > Key Exchange Policy - This is strictly for IKE to IKE flows. What IKEs we will talk to, what encryption to use to flow IKE to IKE data such as Phase I and Phase II negotiations.
 - Key Exchange Rule
 - Define IP filter conditions here for IKE; which IKE addresses and IDs will be used for Phase I negotiations - local and remote
 - Key Exchange Action
 - Whether to initiate phase I, and if so, whether to use main or aggressive mode. If responding, whether to use main or aggressive mode
 - Key Exchange Offer
 - Define what encryption information to use for Phase I negotiations
- > IPfilterrule - This is defining an encryption rule for a set of one or more data endpoints. The rule is composed of a set of filter conditions - which packets for which this rule applies, and a dynamic VPN action - what encryption to use when setting up the dynamic tunnels for this set of data endpoints.
 - IPGenericFilterAction IPFilterAction IPSEC
 - Must be entered to get dynamic VPN
 - IPLocalStartOption
 - This is where you define the scope of the Phase 2 negotiation. If you specify Packet, much of the information for the Phase 2 negotiation comes from the incoming packet. If you specify rule, it comes from the rule that matched the incoming packet
 - This is where you also specify which security endpoints to use - local and remote gateway addresses
 - IPDynVpnAction
 - IPDataOffer - here is where you specify the encryption information to use for the encryption for the data flow for this connection.
- > LocalDynVPNPoly - This gives the customer a way to drive Phase 1 and Phase 2 tunnel activation without a packet coming in. Effectively, this LocalDynVpnRule defines a set of addresses/ports/protocols. When the LocalDynVpnRule has the autoactivate parm, or is activated by IPSEC cmd, dynamic tunnels and IKE tunnels are created/used as though a packet with these addresses/ports/protocols was received.



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

IPSec VPN concepts - integrated IP Security dynamic SAs overview



© Copyright IBM Corp. 2005. All rights reserved.

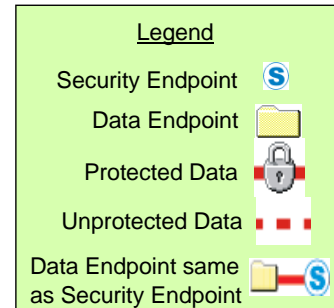
ibm.com/redbooks

IPSec VPN concepts - general NAT/NAPT restrictions

- > Only ESP is supported (AH is not allowed by RFC 3947/3948 restriction)
- > z/OS is optimized for host configuration (does not support acting as a security gateway for SAs that traverse a NAT)
- > z/OS only supports SAs that traverse a NAT, not SAs that traverse an NAPT
 - NAPT is a NAT that maps many private addresses to 1 public address by performing port translation (also known as Port Address Translation (PAT))
- > Tunnel mode with ESP (Responder only)



- > Tunnel or transport mode with ESP
 - Potential issues when interoperating with non-z/OS platforms
 - When z/OS initiates an SA for specific ports or protocol
 - When z/OS initiates data on a tunnel mode SA for all ports and protocols



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

IPSec VPN concepts - UDP encapsulation (NAT traversal)

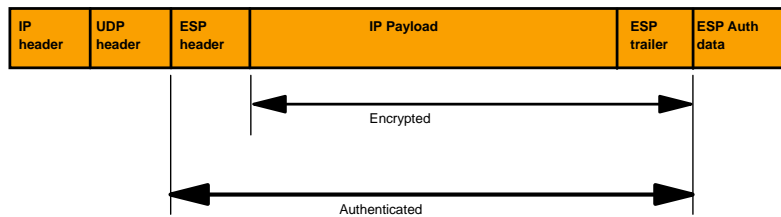
NOTES

- > **Additional encapsulation modes used when a NAT is traversed**
 - UDP-encapsulated transport
 - UDP-encapsulated tunnel
- > **Only valid with ESP packets**
 - Normal transport/tunnel mode encapsulation performed
 - Inserts an additional UDP header in front of the ESP header
- > **Allows ESP packets to traverse a NAT**
- > **On z/OS the decision to use UDP-encapsulation is made by the IKE daemon if a NAT is detected**
- > **NAT traversal support can be enabled or disabled in IP security policy**

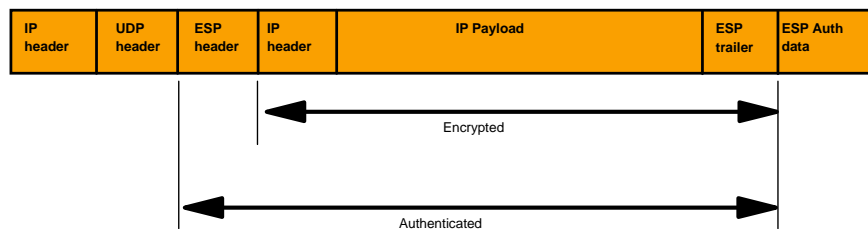
IPSec VPN concepts - UDP-encapsulated packets

NOTES

- > Below shows the format of a UDP-encapsulated transport mode packet



- > Below shows the format of a UDP-encapsulated tunnel mode packet



IPSec VPN concepts - Sysplex Wide Security Association (SWSA) considerations

- **A dynamic VIPA may be the endpoint of an SA - IPSec SAs will be distributed to target stacks of distributed dynamic VIPAs**
 - ▶ Used to distribute IPSec-protected workload
 - ▶ Used for VIPA takeover
- **Requires the DVIPSEC keyword on the IPSEC statement in the TCPIP profile**
- **Compatibility with z/OS Firewall Technologies IPSec**
 - ▶ A FIREWALL stack can be the target of an IPSECURITY stack
 - ▶ An IPSECURITY stack can be the target of a FIREWALL stack
 - ▶ A FIREWALL stack can be a backup for an IPSECURITY stack
 - ▶ An IPSECURITY stack can be a backup for a FIREWALL stack
- **Policies must be consistent on distributing and target stacks**
- **Requires the use of the Coupling Facility EZBDVIPA structure**
- **NAT traversal restrictions - SAs that traverse a NAT:**
 - ▶ Cannot be taken over if the remote host is a security gateway
 - ▶ Are not supported by z/OS Firewall Technologies IPSec (distributor, target, nor backup)

SECCLASS option on link definitions

NOTES

- **Updated to include SECCLASS**
 - ▶ Used to uniquely identify an interface or group of interfaces with similar security requirements
 - ▶ Used as an IP filtering criteria
 - Can only be specified on rules with an action of permit/deny
 - Allows broad rules to be written for all IP traffic that uses a group of interfaces without explicit knowledge of IP address
 - ▶ Can be specified for all link types except VIRTUAL
 - ▶ To modify
 - Stop the device
 - Delete the LINK statement
 - Add the LINK statement with the updated value
 - Restart the device
 - ▶ Sample syntax:

```
>> LINK Existing Link Specification | SECCLASS_255 |><  
| SECCLASS_nnn |
```


IPSEC default filter rule specification example

```
IPSEC LOGENable
; Rule      SrcAddr DstAddr  Logging Protocol  SrcPort  DestPort  Secclass
; OSPF protocol used by Omproute
IPSECRule *      *      NOLOG  PROTO OSPF
; IGMP protocol used by Omproute
IPSECRule *      *      NOLOG  PROTO 2
; DNS queries to UDP port 53
IPSECRule *      *      NOLOG  PROTO UDP  SRCPort *  DESTport 53
; Administrative access
IPSECRule *      9.1.1.1  LOG    PROTO *
ENDIPSEC
```

➤ Remember that these statements create bidirectional filters.

➤ The administrative access rule allows

- ▶ Traffic from remote IP address 9.1.1.1, any protocol over interfaces with a class of 100
- ▶ Traffic to remote IP address 9.1.1.1, any protocol over interfaces with a class of 100

Policy Agent (Pagent) configuration files

NOTES

➤ **Main configuration file**

- ▶ New CommonIPSecConfig statement
 - Identifies an IPsec configuration file containing policy definitions that are common to all stacks in a z/OS image

➤ **Image configuration file**

- ▶ New IPsecConfig statement
 - Identifies an IPsec configuration file containing policy definitions that are specific to a stack

➤ **IPsec configuration file (New)**

- IpFilterPolicy
- KeyExchangePolicy
- LocalDynVpnPolicy
- ▶ Stack-specific IPsec configuration file can be generated by the z/OS IP Security Configuration Assistant GUI or it can be edited by hand using a text editor, such as ISPF/PDF.

➤ **For additional details see**

- ▶ "Chapter 15. Policy-based networking" in the "IP Configuration Guide"
- ▶ "Chapter 18. IP Security" in the "IP Configuration Guide"
- ▶ "Chapter 21. Policy Agent and policy applications" in the "IP Configuration Reference"

pasearch command changes and a new ipsec command

NOTES

> Pasearch command additions

- ▶ New -v option
 - Displays IPsec policies
 - All IPsec policy entries (pasearch -v a)
 - IpFilter policy entries (pasearch -v f)
 - KeyExchange policy entries (pasearch -v k)
 - LocalDynVpn policy entries (pasearch -v l)
 - Can be used with other options to control output (e.g. pasearch -v a -n will just display the names of IPsec policy objects)

> New ipsec command

- ▶ Displays IP security information
 - Current filter rules
 - Manual and dynamic IPsec tunnels
 - IKE tunnels
 - Stack interface information
 - Matching filter rules for a traffic pattern
- ▶ Modifies IP security state
 - Change the filter policy the stack considers to be active
 - Default filter rules
 - IP Security filter rules
 - Activate/deactivate/refresh manual and dynamic IPsec tunnels
 - Deactivate/refresh IKE tunnels
- ▶ Runs APF authorized
- ▶ RACF profiles must be defined to use the ipsec command

ipsec command

NOTES

| Primary Command | Main functions provided |
|-----------------|---|
| ipsec -f | <ul style="list-style-type: none"> • Display information about active filter set • Display information about default IP filter rules • Display information about IP Security filter rules • Make the default IP filter rules the active filter set • Make the IP Security filter rules the active filter set |
| ipsec -m | <ul style="list-style-type: none"> • Display information about manual tunnels • Activate manual tunnels • Deactivate manual tunnels |
| ipsec -k | <ul style="list-style-type: none"> • Display information about IKE tunnels • Deactivate IKE tunnels • Refresh IKE tunnels |
| ipsec -y | <ul style="list-style-type: none"> • Display information about dynamic tunnels (stack's view) • Display information about dynamic tunnels (IKED's view) • Activate dynamic tunnels • Deactivate dynamic tunnels • Refresh dynamic tunnels |
| ipsec -i | <ul style="list-style-type: none"> • Display interface information |
| ipsec -t | <ul style="list-style-type: none"> • Locate matching filter rule |
| ipsec -o | <ul style="list-style-type: none"> • Display NATT port translation table information |
| ipsec -? | Help |

See the "IP System Administrator's Commands" for the complete syntax

ipsec command protection via SERVAUTH profiles

NOTES

➤ Command access controlled by profiles in the SERVAUTH class

| Resource name | ipsec options allowed |
|---|--|
| EZB.IPSECCMD.sysname.tcprocname.* | All of ipsec options |
| EZB.IPSECCMD.sysname.tcprocname.DISPLAY | -f display -m display -k display -y display -t -i -o |
| EZB.IPSECCMD.sysname.tcprocname.CONTROL | -f default -f reload -m activate -m deactivate -k deactivate -k refresh -y activate -y deactivate -y refresh |

Sample JCL job to define these profiles provided in SEZAINST(EZARACF)

IKE daemon

NOTES

➤ Provides for the negotiation of IPSec SAs using the Internet Key Exchange (IKE) as defined in RFC 2409, including support of RFC 3947 (Negotiation of NAT-Traversal in the IKE)

➤ APF authorized UNIX application

- ▶ Can be started from UNIX shell (iked) or started proc (sample in SEZAINST(IKED))

➤ IKED_FILE

- ▶ Specifies where to find the IKE Daemon configuration file
 - IKED_FILE=/etc/security/iked.conf
- ▶ If not specified the default is /etc/security/iked.conf

➤ IKED_CTRACE_MEMBER

- ▶ Specifies the name of a parmlib member in the form CTIIKExx that contains default CTRACE settings
 - Example: IKED_CTRACE_MEMBER=CTIIKE3A
- ▶ If not specified the default is CTIIKE00
- ▶ Must be set prior to starting IKED
 - CTRACE settings only read during IKED initialization

Enabling integrated IP security for a stack

NOTES

- **Ensure Pagent is configured and started**
 - Define IPsec policy
 - Update Pagent configuration file to contain IPsecConfig and optionally CommonIPsecConfig
- **Ensure TRMD is configured and started**
- **Ensure syslogd is configured and started**
 - TRMD will write IPsec messages to local4
- **Update TCP/IP profile**
 - Add IPSECURITY to IPCONFIG statement
 - Classify devices by adding SECCLASS (optional)
 - Define default filter rules
- **Create security definitions for ipsec command**
 - Details provided in the Integrated IPsec Externals section
 - Update the SERVAUTH class
 - Restrict access to the marker files
- **Authorize the stack to ICSF**
 - If hardware encryption is available

Enabling the IKE daemon

NOTES

- **Create the IKE daemon configuration file**
 - Sample in `/usr/lpp/tcpip/samples/iked.conf`
- **If starting from the UNIX shell**
 - `set _BPX_JOBNAME` (optional)
- **If starting from a cataloged procedure**
 - Update sample from `SEZAINST(IKED)`
- **Update the PORT statement in the TCP/IP profile to reserve UDP ports 500 and 4500**
- **Authorize the IKE daemon to the External Security Manager (ESM)**
- **Ensure syslogd is configured and started**
 - The IKE daemon will write syslog records to local4
 - For performance reasons it is recommended that IKE daemon syslog records be written to a zFS file

Enabling the IKE daemon (continued)

NOTES

- Define the location of the IKE daemon configuration file and parmlib member for CTRACE
 - IKED_FILE and IKED_CTRACE_MEMBER
- If RSA signature is being utilized, set up the IKE daemon keyring
- Performance considerations
 - Set appropriate dispatching priority at or just below TCPIP's priority
 - If running WLM should be assigned to the SYSSTC service class
- Decide how you want the IKE daemon to be started
 - Automated
 - AUTOSTART in the TCP/IP profile (use this technique if there is only 1 IP security stack running)
 - Using the COMMNDxx member of parmlib
 - From UNIX shell
 - iked
 - From operator's console
 - S IKED



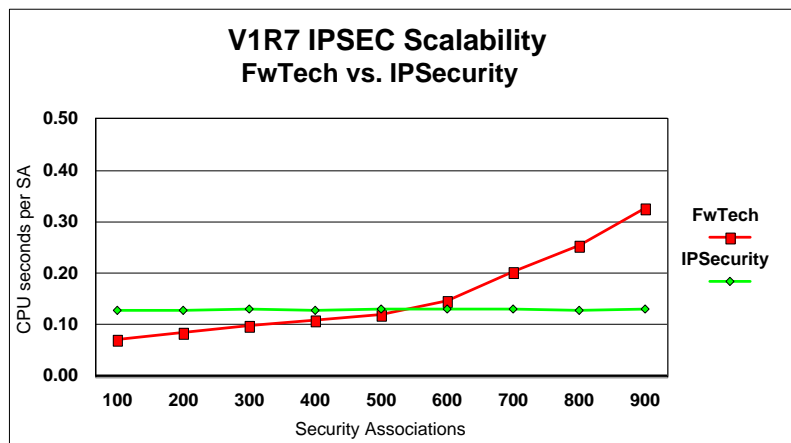
© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

CPU cost of setting up an IPSec security associations

➤ IPSec scalability

- CPU cost per SA is the same as one adds more Security Associations



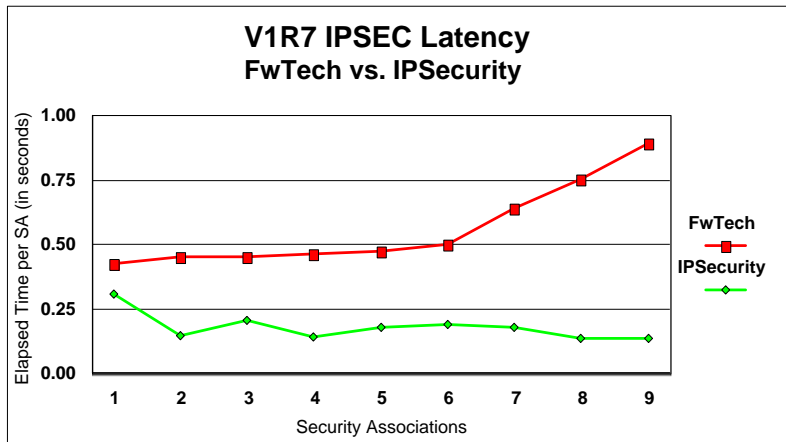
© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Time to set up an IPsec security association

> IPSEC latency

- ▶ Elapsed time it takes for an SA to be established for a connection is approximately the same as one adds more SA's



© Copyright IBM Corp. 2005. All rights reserved.

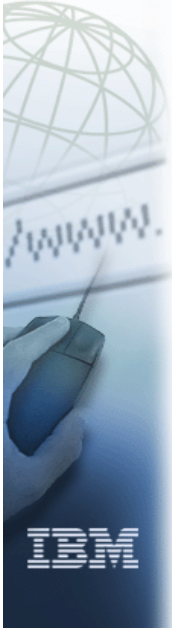
ibm.com/redbooks

This page intentionally left blank

ibm.com



e-business



Communications Server for z/OS V1R7 - Technical Update Application Transparent Transport Layer Security (AT-TLS)



Redbooks

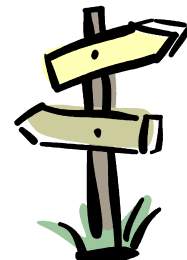
International Technical Support Organization

© Copyright IBM Corp. 2005. All rights reserved.

Application Transparent Transport Layer Security (AT-TLS) agenda



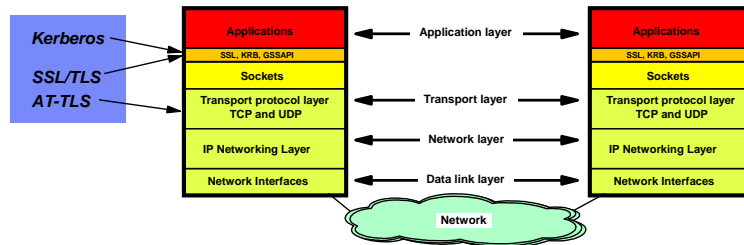
- > AT-TLS background and introduction
- > AT-TLS concepts and modes of operation
- > AT-TLS Netstat reports
- > Things to think about



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

SSL/TLS and Kerberos overview on z/OS



- **Transport Layer Security (TLS) is defined by the IETF**
 - ▶ Based on Secure Sockets Layer (SSL)
 - SSL originally defined by Netscape to protect HTTP traffic
 - ▶ TLS defines SSL as a version of TLS for compatibility
 - TLS clients and server should drop to SSL V3 based on partner's capabilities
- **Traditionally provides security services as a socket layer service**
 - ▶ Requires reliable transport layer
 - UDP applications cannot be TLS enabled
 - ▶ Application source code changes are generally needed to SSL/TLS-enable a sockets program
- **z/OS applications can be TLS enabled with System SSL**
 - ▶ System SSL is part of z/OS Integrated Security Services element
 - ▶ System SSL supports z/OS UNIX System Services C/C++ applications only
- **Kerberos support is implemented using the Kerberos and GSSAPI functions of the z/OS Security Server and provides:**
 - ▶ Third-party authentication
 - ▶ Optional message integrity
 - ▶ Optional message privacy (encryption)
- **The Kerberos environment must be set up on the z/OS system.**
 - ▶ The Kerberos support is documented in the publication "Network Authentication and Privacy Service: Administration", SC24-5926
- **Some z/OS applications that are kerberized:**
 - ▶ FTP server and client
 - ▶ UNIX Telnet daemon (OTelnetD)
 - ▶ UNIX RSH daemon (ORshD)
 - ▶ z/OS WAS Server
- **Mostly of value where a Kerberos-based infrastructure already is in place**

Symmetric encryption (secret or shared key encryption)

NOTES

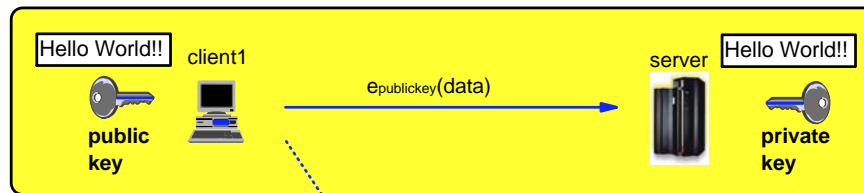


- Symmetric encryption algorithms are fast and efficient.
 - Key management is a major concern for secret key-based encryption - to whom do we send the key and how do we do that safely?
- Data that is encrypted with the secret key can only be decrypted by someone who has the same secret key (sharing the same secret).
- Someone who does not have a copy of the secret key cannot decrypt any intercepted data.

Examples of symmetric encryption algorithms are: DES and AES

RSA public/private key encryption (asymmetric encryption)

NOTES



Public and private keys match, but they are not the same value.

- > Public/private key encryption is CPU intensive
- > The server's public key can be distributed freely to anyone who needs it.

$e_{\text{publickey}}(\text{data})$

- Data encrypted by the public key can only be decrypted by the matching private key.
- Data encrypted by the private key can be decrypted by the public key.



`.*x-yz/%%&&8/dvvvv`

Someone who only has a copy of the public key cannot decrypt data that was encrypted with the same public key.

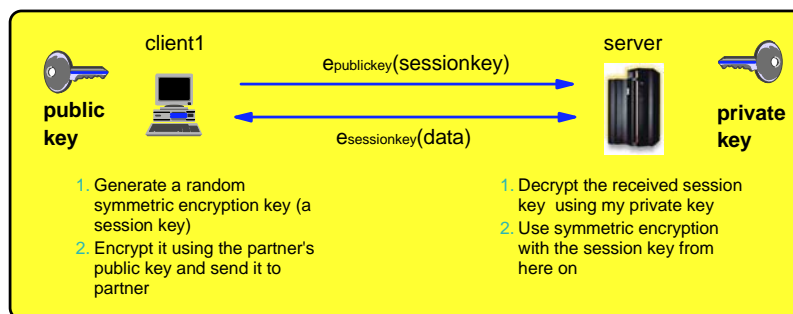


© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Combining public/private key encryption with symmetric encryption - session keys

NOTES



1. Generate a random symmetric encryption key (a session key)
2. Encrypt it using the partner's public key and send it to partner

1. Decrypt the received session key using my private key
2. Use symmetric encryption with the session key from here on

Secure Sockets Layer (SSL) uses RSA to generate symmetric session keys

1. Server has distributed a public key to client earlier.
2. Client generates a random session key, encrypts it under the server's public key, and transmits it to the server. Only the server is able to decrypt this message.
3. Server decrypts the message and the server and the client use the session key for succeeding encrypted data exchanges in this session.

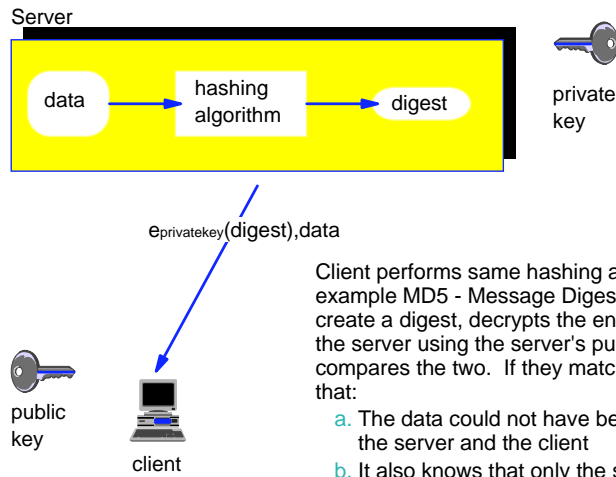


© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Digital signature / message authentication

NOTES

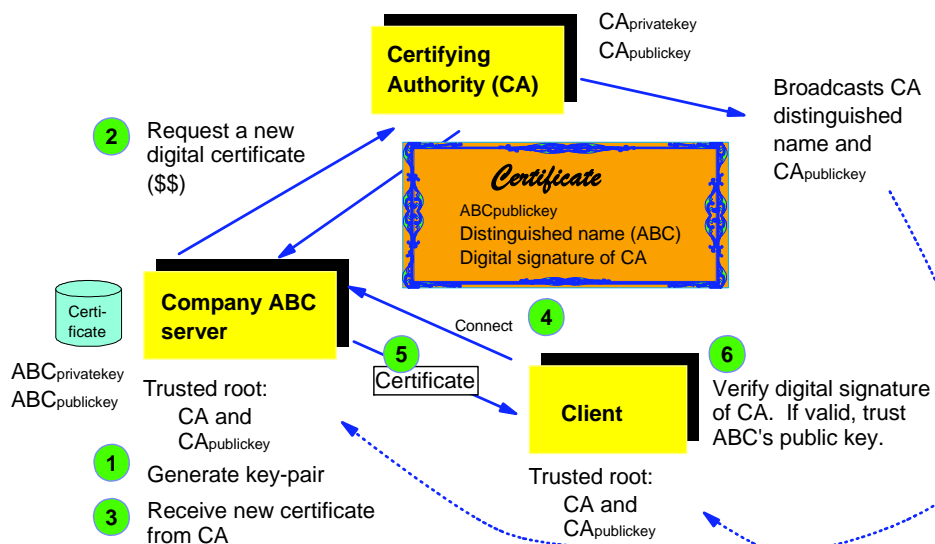


Client performs same hashing algorithm (for example MD5 - Message Digest 5) on the data to create a digest, decrypts the encrypted digest from the server using the server's public key, and compares the two. If they match, the client knows that:

- a. The data could not have been altered between the server and the client
- b. It also knows that only the server with the right private key could have created the message.

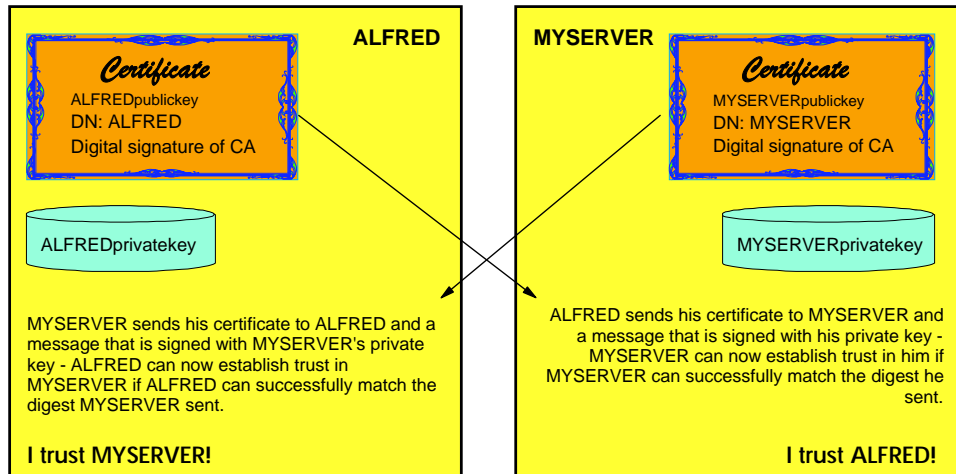
X.509 certificates - trust relationships

NOTES



Server and client certificates

NOTES

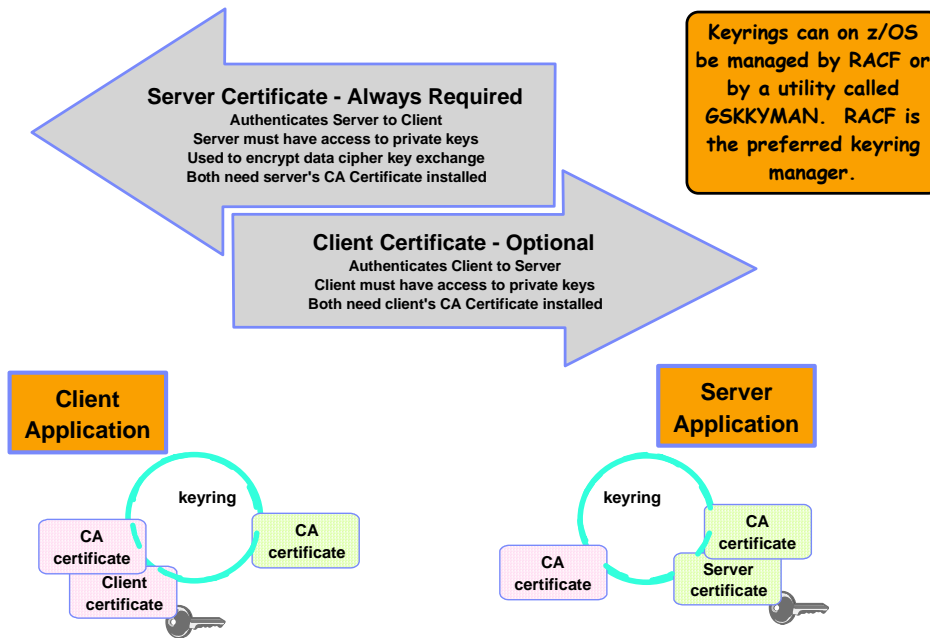


- In order to create session keys and encrypt the data stream, only the server needs to have a certificate.
- If the client has a certificate, then the server can use that to authenticate the client. On z/OS a client certificate can be mapped to a RACF user ID. Password authentication may still be required by the server application, but the password can at least now be submitted in encrypted form over the network.

SSL/TLS background information

- **SSL: Secure Sockets Layer**
 - ▶ Created by Netscape
 - ▶ Originally implemented inside Web clients and servers
 - ▶ Above sockets and below application protocol
 - ▶ SSLv1 is no longer supported
 - ▶ SSLv2 is still used in some limited cases
 - mostly public access compatibility concern
 - ▶ SSLv3.0 improved security
- **TLS: Transport Layer Security**
 - ▶ TLSv1.0 (SSLv3.1)
 - ▶ IETF RFC 2246
- **End-to-end application pipe**
 - ▶ TCP connections
 - ▶ Server authentication
 - ▶ Optional client authentication
 - ▶ Authentication
 - Public key cryptography, third-party signed certificate
 - ▶ Data privacy
 - Negotiated private key cryptography
 - SSL record protocol

Certificate and keyring background Information



Current SSL/TLS API support on z/OS is limited to C and Java

> Application layer implementation

- ▶ Development expense repeated for each application
- ▶ Toolkits available for limited programming environments
 - C and Java
 - Forking or Threaded POSIX model
 - Require application change
- ▶ Many existing z/OS applications do not fit this model
 - COBOL, assembler sockets programs
 - CICS sockets transactions
 - etc.
- ▶ Many applications purchased or otherwise not available for change
 - Source code is needed to enable them for SSL/TLS

> Application specific deployment

- ▶ Unique configuration for each application
- ▶ Different levels of SSL/TLS architecture support
- ▶ Not all toolkits support/exploit z/OS and zSeries capabilities
 - RACF keyrings
 - Certificates associated with user IDs
 - Hardware cryptography

Application Transparent Transport Layer Security (AT-TLS) introduction

> Application Transparent

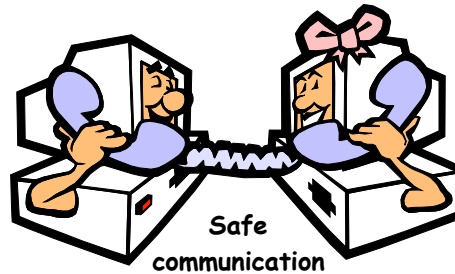
- ▶ Support existing applications without change
- ▶ Allow applications to optionally exploit/control advanced features
 - Simple ioctl
 - Extract status, certificate, and associated user ID
 - Permit cleartext negotiation prior to starting secure connection

> Transport Layer

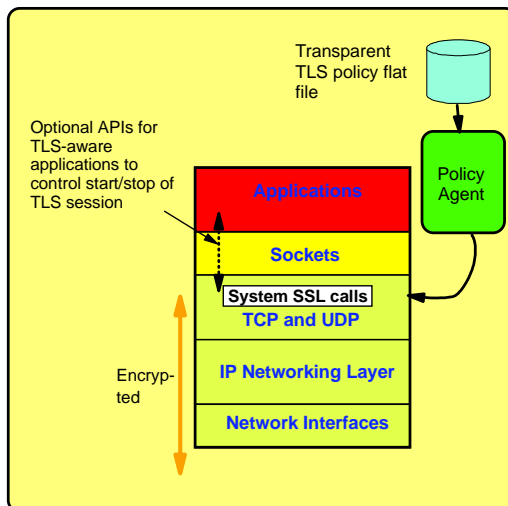
- ▶ Implement inside the TCP layer of the stack
- ▶ Common configuration through policy
- ▶ Exploit z/OS
 - RACF
 - SystemSSL
 - ICSF
 - Hardware cryptography

> Security

- ▶ Multiple Protocols
 - TLS (SSL V3.1)
 - SSL V3.0
 - SSL V2



Transparent application security: policy-controlled transparent SSL/TLS support - SSL/TLS for all z/OS sockets applications



> Basic TCP/IP stack-based TLS

- ▶ TLS process performed at TCP layer without requiring any application change (transparent)
- ▶ All connections to specified port are designated as TLS required
 - Can be further qualified by source/destination IP addresses
- ▶ Transparent TLS policies managed via Policy Agent

> Transparent TLS can be requested by application

- ▶ Application issues transparent TLS API calls to indicate that connection should start/stop using TLS

> TCP/IP stack-based TLS with client identification services for application

- ▶ Application issues TLS API calls to receive user identity information based on X.509 client certificate

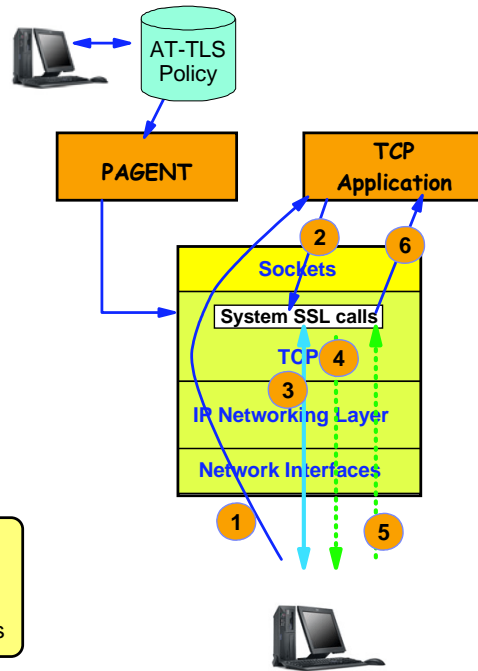
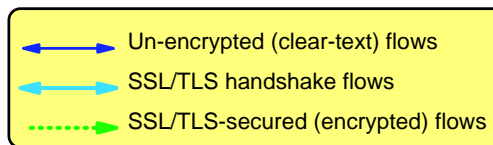
> Available to any TCP application

- ▶ CICS Sockets and JES/NJE are primary focus of this support
- ▶ All programming languages supported

AT-TLS basic principles

> Configured AT-TLS policy for the TCP application to use AT-TLS:

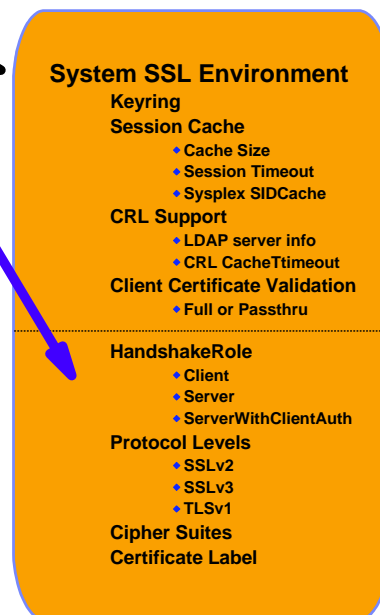
- ▶ Client connects to server and connection becomes established
- ▶ Server sends data in the clear and TCP layer queues it.
- ▶ TCP layer invokes System SSL to perform SSL handshake under identity of the server.
- ▶ TCP layer invokes System SSL to encrypt queued data and sends it to client.
- ▶ Client sends encrypted data, TCP layer invokes System SSL to decrypt.
- ▶ Server receives data in the clear.



AT-TLS environment concepts

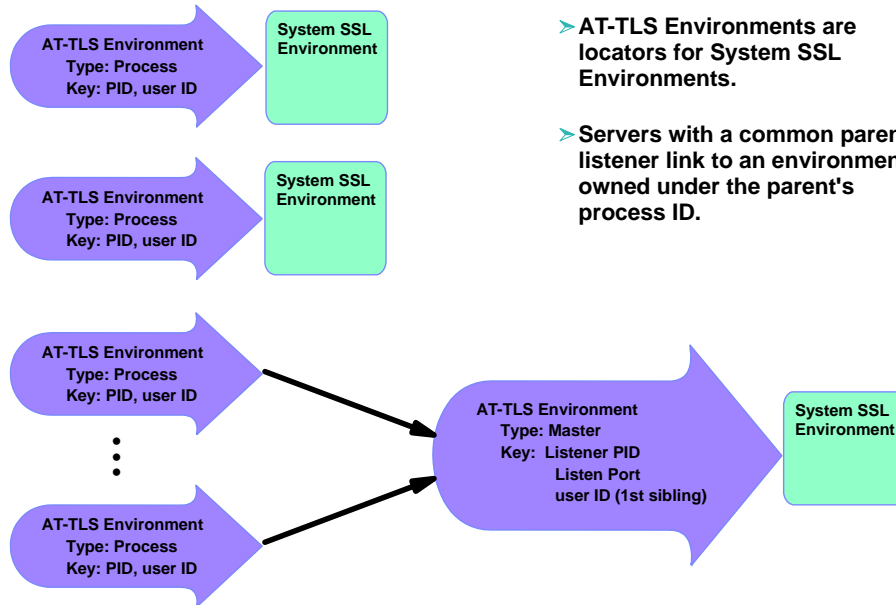
NOTES

- > **TTLSEnvironmentAction**
 - ▶ Values for System SSL Environment Attributes
- > **TTLSEnvironmentAction**
 - ▶ Values for SSL Attributes overridden at Connection
- > **System SSL Environment Sharing Rules**
 - ▶ Connection policy must reference the same instance of:
 - TTLSEnvironmentAction
 - ▶ Connection must meet one of the following criteria:
 - Same process ID and user ID or
 - Same server process family
 - HandShakeRole Server and
 - Passive connection with same parent process and port and same user ID as siblings
- > **System SSL Environment Life Cycle**
 - ▶ Dynamically created when none found to share
 - ▶ Dynamically removed when
 - TTLSEnvironmentAction is active and
 - No connections for 10-20 minutes
 - TTLSEnvironmentAction is stale and
 - No current connections



AT-TLS environments

NOTES



> AT-TLS Environments are locators for System SSL Environments.

> Servers with a common parent listener link to an environment owned under the parent's process ID.

Redbooks © Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

AT-TLS policy preview

NOTES

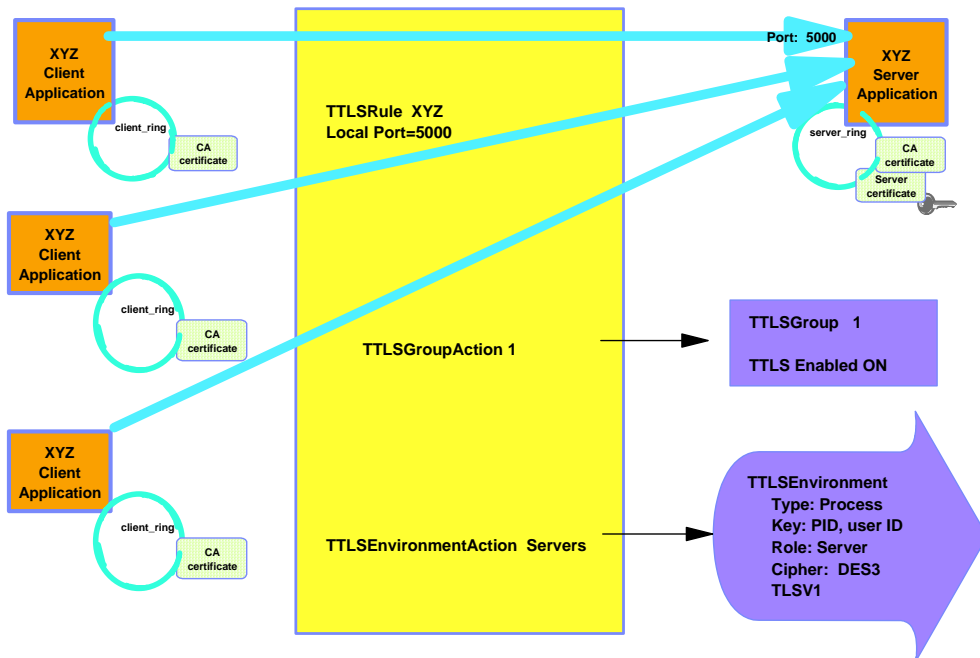
> PolicyAgent main configuration file

- CommonTTLSSConfig statement names a file containing AT-TLS objects shared across TCP/IP stacks
 - TTLSSRule statements
 - TTLSSGroupAction statements
 - TTLSSEnvironmentAction statements
 - TTLSSConnectionAction statements
- PEPInstance statement names a file containing policy for one TCP/IP stack
 - TTLSSConfig statement names a file containing AT-TLS objects for this stack
 - TTLSSRule statements
 - TTLSSGroupAction statements
 - TTLSSEnvironmentAction statements
 - TTLSSConnectionAction statements

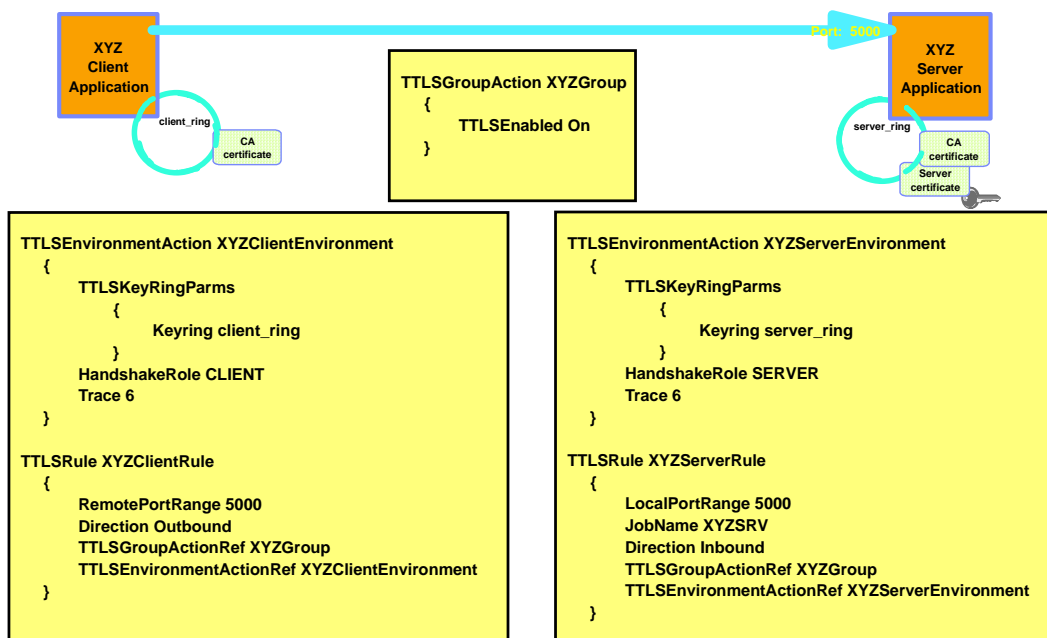
Redbooks © Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

AT-TLS basic policy overview



AT-TLS basic policy examples



AT-TLS policy mapping

NOTES

> Rule search

- ▶ One-time event for each connection
- ▶ Result persists for life of connection
- ▶ Search order is Rule Name (alphanumeric) within Priority (high to low)
- ▶ Conditions:
 - Connection direction, Local / remote IP address and port, Jobname, User ID

> Security context

- ▶ Caller's security context is "cloned" into stack at time of mapping
 - Includes: User ID, Group ID, UID and GID
- ▶ This security context is used to access keyring and certificate keys

> Mapping events

- ▶ Outbound
 - Connect
- ▶ Inbound
 - Select or poll for readable or writable
 - Any form of read or write
- ▶ SIOCTLSCTL ioctl

> Secure session auto start

- ▶ If ApplicationControlled Off, Secure connection is AutoStarted when mapped except
 - On connect, AutoStart is deferred until connection is established
 - SIOCTLSCTL ioctl never AutoStarts



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

AT-TLS policy mapping rationale

NOTES

- > A critical aspect of AT-TLS operation is the security context cloning required to access keyrings and certificates on behalf of an owning application. Analysis of several common application models indicated the need to defer security context cloning on inbound connections to sometime after accept().
- > This is most apparent in the family of servers invoked by inetd. The passive socket is created by inetd {socket(), bind(), listen()}. New connections are recognized by inetd {accept()}. Based on the port connected to, inetd creates a new server process {fork()}, optionally changes the security context in the new process {setuid()}, and then turns control over to the server program {exec()}. In many cases, the application protocol includes some form of login negotiation. The server program then changes its security context to one supplied by the client over the new connection.
- > The optimal security context to clone is the one initially used by the server process. The inetd security context does not allow enough granularity and protection of server certificates - any server invoked by inetd could be configured to use any server certificate that inetd had access to. The client security context would require all clients to have access to the server's private keys - this would be a serious breach of security.
- > In all analyzed application models, the server security context is the one presented to the stack on the first data oriented service requested over the socket.



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

AT-TLS application types

> Not enabled

- ▶ Pascal API and Web Fast Response Cache Accelerator (FRCA) not supported
- ▶ No policy or policy explicitly says Enabled OFF
 - Includes those permitted to start during InitStack window
- ▶ Application may optionally use SSL or TLS toolkit directly

> Basic

- ▶ Policy says Enabled ON
- ▶ Application unchanged and unaware of AT-TLS
- ▶ Application protocol unaffected by use of AT-TLS (consider http: versus https:)

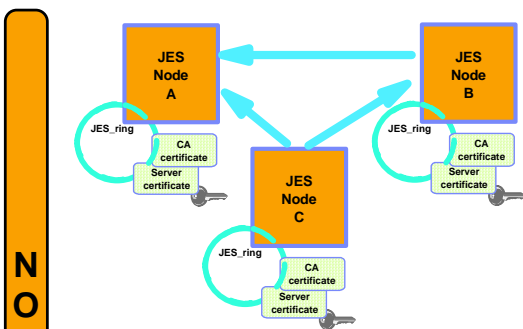
> Aware

- ▶ Policy says Enabled ON
- ▶ Application changed to use SIOCTTLSCTL ioctl to extract AT-TLS information:
 - Policy status, negotiated version and cipher, partner certificate, associated user ID

> Controlling

- ▶ Application protocol may negotiate use of TLS in cleartext prior to starting secure session
- ▶ Policy says Enabled ON and ApplicationControlled ON
- ▶ Application changed to use SIOCTTLSCTL ioctl to extract and control AT-TLS
 - Policy status, negotiated version and cipher, partner certificate, associated user ID
 - Start secure session, reset session, reset cipher

AT-TLS common policy examples (peer)



NOTES

> JES peer network

- ▶ Each node tries all known nodes, then listens for others.
- ▶ Use of TLS negotiated in the clear.
- ▶ HandshakeRole follows connect direction.

```

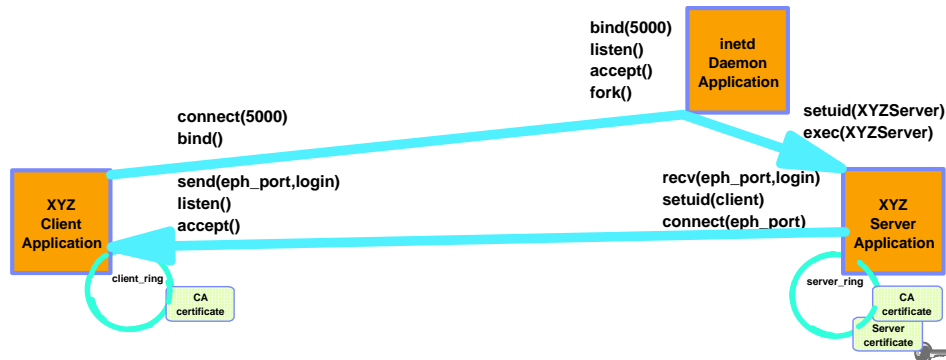
TTLSEnvironmentAction JESEnvironment
{
    HandshakeRole Server
    TLSKeyRingParms
    {
        Keyring JES_ring
    }
    Trace 6
    TTLSEnvironmentAdvancedParms
    {
        ApplicationControlled ON
    }
}

TLSConnectionAction ClientConnection
{
    HandshakeRole Client
}

TTLSEnvironmentAction JESInbound
{
    JobName JES
    Direction Inbound
    TTLSEnvironmentActionRef JESEnvironment
}

TTLSEnvironmentAction JESOutbound
{
    JobName JES
    Direction Outbound
    TTLSEnvironmentActionRef JESEnvironment
    TLSConnectionActionRef ClientConnection
}
    
```

Applications with secondary connections



- Some applications create a second connection between the client and the forked server instance. This second connection can be established in either direction.
- Some applications use two dynamically assigned ports (for example, rexec or rsh stderr connection). It can be difficult to define an effective rule for these connections.
- In other applications the server changes to a client login identity before the second connection is mapped (for example, FTP data connection). The server's security context is no longer available for accessing the server's keyring and certificate keys.

AT-TLS SecondaryMap

- **AT-TLS policy for primary connection**
 - ▶ TTLSxxxActionAdvancedParms
 - SecondaryMap ON
 - ▶ Primary connection is recorded in internal table
 - Entry created when primary connection is mapped
 - Indexed by process ID, local IP, and remote IP
 - Entry removed when primary connection closes
- **When future connections are mapped**
 - ▶ Normal policy lookup
 - ▶ Check internal table for primary connection with
 - Same process ID and same IP endpoints
 - ▶ New connection is a secondary connection if match found in internal table and
 - No policy found for new connection or
 - Found policy is lower priority than primary connection policy
 - ▶ Secondary connections
 - Share SSL Environment with primary connection
 - Use security context from primary connection
- **Application model requirements**
 - ▶ One primary connection (serially reused client, forking server)
 - ▶ May have multiple secondary connections

New Netstat filter to limit connection reports to AT-TLS connections

➤ **Netstat filter (CONNType/-X) added to limit ALLConn/-a and CONn/-c responses by connection type**

➤ **Subfilters allow for specification of connection type:**

- ▶ NOTTLSPolicy
 - Connections not mapped to AT-TLS policy
- ▶ TTLSPolicy
 - Connections mapped to AT-TLS policy
- ▶ TTLSPolicy,CURRent
 - Connections mapped to AT-TLS policy - rule and actions still available for use with new connections
- ▶ TTLSPolicy,GRoup=groupid
 - Connections using the specified AT-TLS group
- ▶ TTLSPolicy,STALE
 - Connections mapped to AT-TLS policy - rule or at least one action no longer available for use with new connections

```
--CONNType-+-NOTTLSPolicy-----+----->
      '-TTLSPolicy-+-----+-'
                +-CURRENT-----+
                +-GROUP-groupid-+
                '-STALE-----'
```

New Netstat TTLS report added

➤ **Netstat option (TTLS/-x) added to display AT-TLS data**

▶ Suboptions allow for display of specific AT-TLS data:

- GRoup
 - Summary information for AT-TLS groups.
- GRoup,DETAIL
 - Detailed information for AT-TLS groups.
- CONn=connid
 - Name of AT-TLS policy rule and names of associated actions for specified connection.
- CONn=connid,DETAIL
 - Details of AT-TLS policy rule and associated actions for specified connection.

```
.-GROUP-----+
--TTLS-+-----+----->
      +-CONN-connid-+-----+
      |             '-DETAIL-' |
      '-GROUP-+-----+
              '-DETAIL-'
```

Netstat TTLS,CONN=connid report sample

NOTES

```
D TCPIP,TCPCS3,TTLS,Conn=connid report:

D TCPIP,TCPCS3,N,TTLS,CONN=31
EZD0101I NETSTAT CS VIR7 TCPCS3 393
CONNID: 00000031
JOBNAME:      FTPD1
LOCALSOCKET:  ::FFFF:9.42.104.156..21
REMOTESOCKET:  ::FFFF:9.27.154.137..1638
SECLEVEL:     TLS VERSION 1
CIPHER:       05 TLS_RSA_WITH_RC4_128_SHA
CERTUSERID:   N/A
MAPATYPE:     PRIMARY
TTLSRULE:     FTP_SERV_21
TTLSGRPACTION: GRP_ACT1
TTLSENVACTION: ENV_ACT_SERV
1 OF 1 RECORDS DISPLAYED
END OF THE REPORT
```

AT-TLS SMF records and network management interface changes

- New SMF 119 TCP connection termination subsection if AT-TLS was used for the connection:

| Offset | Name | Length | Format | Description |
|----------|---------------------|--------|--------|--|
| 0 (x'0') | SMF119AP_TTTTLSSP | 2 | Binary | AT-TLS SSL Protocol: <ul style="list-style-type: none"> • x'0200': SSL Version 2 • x'0300': SSL Version 3 • x'0301': AT-TLS Version 1 |
| 2(x'2') | SMF119AP_TTTTLSSNC | 2 | EBCDIC | AT-TLS Negotiated Cipher |
| 4(x'4') | SMF119AP_TTTTLSSST | 1 | Binary | AT-TLS Security Type: <ul style="list-style-type: none"> • x'01': Client • x'02': Server • x'03': Server with client authentication, ClientAuthType = PassThru • x'04': Server with client authentication, ClientAuthType = Full • x'05': Server with client authentication, ClientAuthType = Required • x'06': Server with client authentication, ClientAuthType = SAFCheck |
| 5(x'5') | SMF119AP_TTTTLSSRV1 | 3 | Binary | Reserved |
| 8(x'8') | SMF119AP_TTTTLSSUID | 8 | EBCDIC | AT-TLS Partner UserID |

Dependencies and restrictions - AT-TLS

➤ z/OS Cryptographic Services System Secure Sockets Layer (System SSL)

- ▶ The PDS pdsname.SIEALNKE contains the System SSL DLLs.
- ▶ It must be in the program search order for TCPIP and Policy Agent.
- ▶ If it's not in the linklist or LPA,
 - use the STEPLIB DD statement in your TCPIP JCL
 - use the STEPLIB environment variable in the shell:
export STEPLIB=\$STEPLIB:pdsname.SIEALNKE

➤ z/OS UNIX APAR OA11339 is required.

- ▶ If it is not installed, all AT-TLS connections will fail with the message
 - syslogd:EZD1286I ... RC: 5019 Initial Handshake
 - console: EZD1287I ... RC: 5019 Initial Handshake

➤ AT-TLS does not support the following applications.

- ▶ These connections will not map to AT-TLS policy.
- ▶ They will be permitted to proceed in clear text.
 - Applications using the Pascal API to access TCP/IP
 - Line Print daemon and commands
 - LPD, LPQ, LPRM
 - Simple Mail Transfer Protocol (JES Spool Server)
 - TSO Telnet client
 - Web servers using Fast Response Cache Accelerator
 - Network administration applications permitted to EZB.INITSTACK profile
 - Connections established and mapped prior to installation of AT-TLS policy will proceed in clear text.
 - Connections established and mapped after installation of AT-TLS policy are subject to policy installed.

Things to think about when enabling AT-TLS

➤ z/OS CS ships some applications with native SSL/TLS support.

- ▶ Some may use either the native support or AT-TLS.
- ▶ Don't configure both for the same application!

➤ Digital Certificate Access Server (DCAS)

- ▶ Not currently an AT-TLS Aware application
- ▶ Do not use with AT-TLS

➤ FTP client and FTPD server

- ▶ Must specify SecondaryMap in AT-TLS policy
- ▶ If using implicit secure socket 990, see policy sample for guidance
 - /usr/lpp/tcpip/samples/pagent_TTLS.conf (/usr/lpp/tcpip/samples/IBM/EZAPAGFT)

➤ TN3270E server

- ▶ Must specify Basic Port (no security information in TN displays)
- ▶ No security parameters accepted
 - (Keyring/LDAP/Encryption/ConnType/SAFCert/ExpressLogon)

➤ Sendmail

IPSec and AT-TLS comparison - a few selected characteristics

NOTES

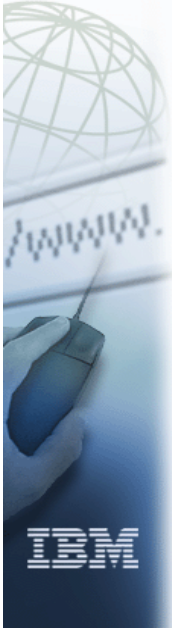
| | IPSec | AT-TLS |
|---|--|---|
| Traffic protected with data authentication and encryption | All protocols | TCP |
| End-to-end protection | Yes | Yes |
| Network segment protection | Yes | No |
| Scope of protection | <u>Security association</u> 1)all traffic 2)protocol 3)single connection | <u>TLS session</u> 1)single connection |
| How controlled | <u>IPSec policy</u> 1)z/OS responds to IKE peer 2)z/OS initiates to IKE peer based on outbound packet, IPSec command, or policy autoactivation | <u>AT-TLS policy</u> 1)For handshake role of server, responds to TLS client based on policy 2)For handshake role of client, initializes TLS based on policy 3)Advanced function applications |
| Requires application modifications | No | No, unless advanced function needed 1)Obtain client cert/userid 2)Start TLS |
| Type of security | Device to device | Application to application |
| Type of authentication | Peer-to-peer | 1)Server to client 2)Client to server (opt) |
| Authentication credentials | 1)Preshared keys 2)X.509 certificates | X.509 certificates |
| Authentication principals | Represents host | Represents user |
| Session key generation/refresh | Yes with IKE No with manual IPSec | TLS handshake |

This page intentionally left blank

ibm.com



e-business



Communications Server for z/OS V1R7 - Technical Update Networking Security Configuration Assistant GUI



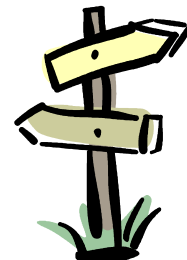
Redbooks

International Technical Support Organization

© Copyright IBM Corp. 2005. All rights reserved.

Security configuration assistant GUI agenda

- > CS z/OS configuration GUI overview
- > Network security configuration assistant



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

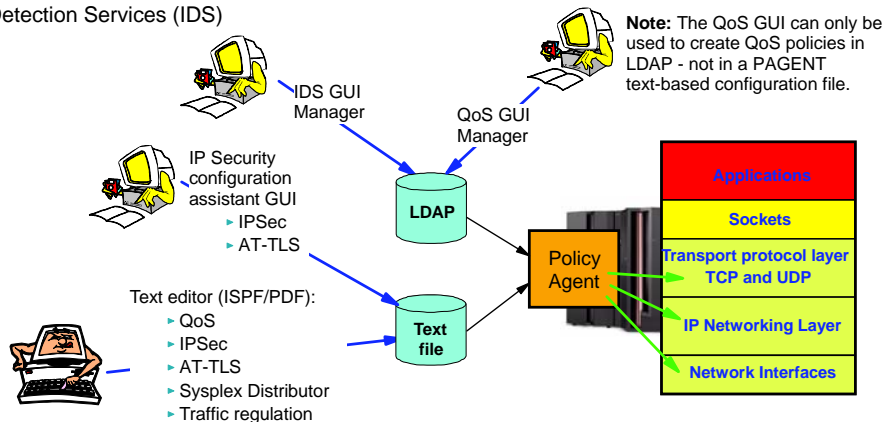
Configuring the Policy Agent

➤ **The following PAGENT policies can be stored in a flat text file format:**

- QoS policies (alternatively supported in LDAP)
- IPSec VPN policies
- IP filter policies
- AT-TLS policies
- Sysplex Distributor policies
- Traffic regulation policies

➤ **The following PAGENT policies must be stored in LDAP:**

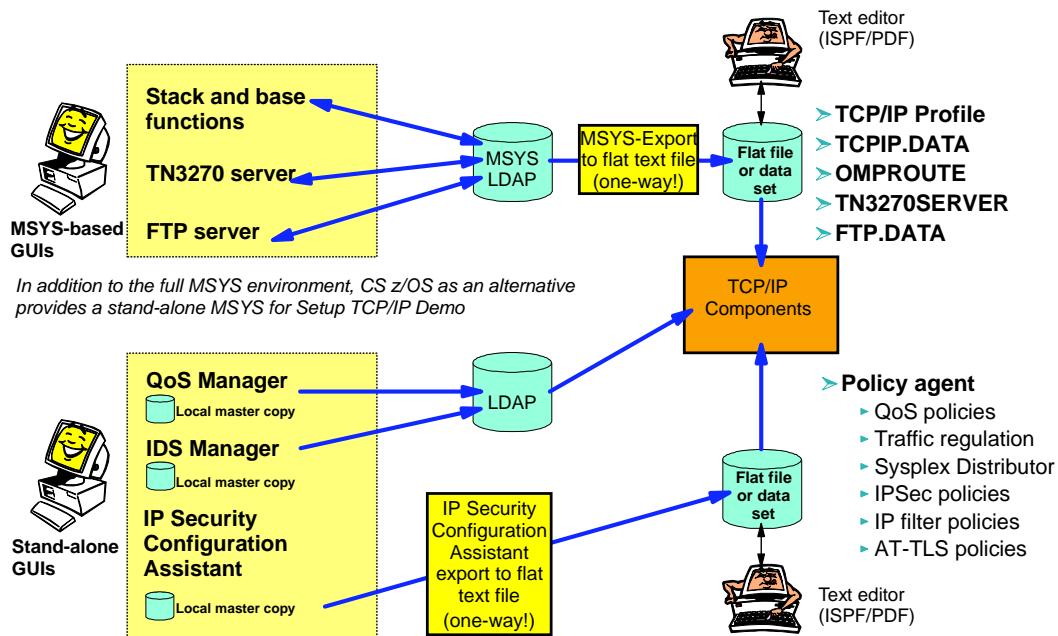
- Intrusion Detection Services (IDS)



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

GUI-assisted CS configuration overview



Note: If text editor updates are made to the flat file configuration data, those changes will not be reflected back into LDAP (for MSYS) or the local master copy for the IP security configuration assistant.



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

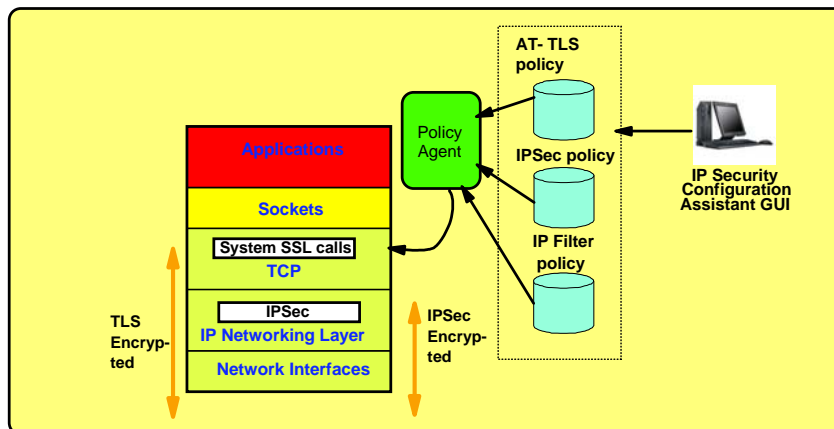
CS z/OS configuration GUIs

NOTES

- These GUIs are all available from the z/OS Communications Server support page at
 - ▶ <http://www.ibm.com/software/network/commserver/zos/support>.
- Click on the All Tools link under Download.

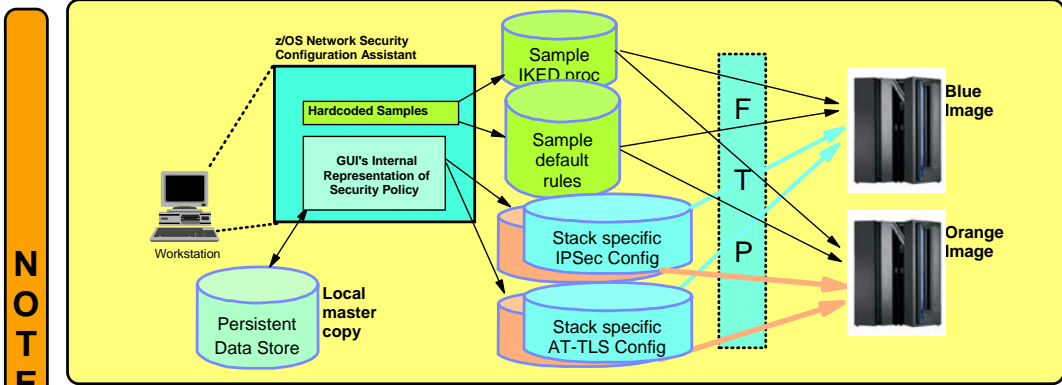
| Tool | URL |
|---|---|
| zQoS Manager | http://www.ibm.com/support/docview.wss?rs=852&uid=swg24007692 |
| zIDS Manager | http://www.ibm.com/support/docview.wss?rs=852&uid=swg24007607 |
| eServer IDS Configuration Manager | http://www.ibm.com/support/docview.wss?rs=852&uid=swg24006805 |
| z/OS Managed System Infrastructure for Setup (msys) TCP/IP Demo | http://www.ibm.com/support/docview.wss?rs=852&uid=swg24006591 |

Policy-controlled application-transparent network security



- Network security without requiring application changes
 - ▶ IPsec
 - ▶ Transparent SSL/TLS
- Configuration single administrative task
 - ▶ Higher level of abstraction
 - Focus on what traffic to protect and how to protect
 - Less focus on low-level details (though available on expert panels)

z/OS V1R7 network security configuration assistant overview



NOTES

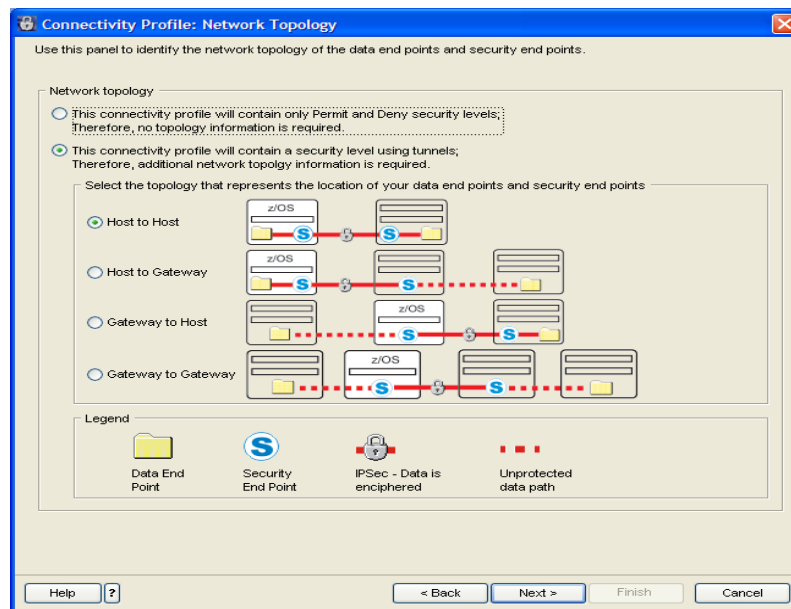
- IPsec, filtering, and AT-TLS policies can be defined by manually editing a Policy Agent configuration text file on z/OS.
- The policies can also be defined using a new downloadable policy configuration tool that runs on a workstation using a graphical user interface.
 - Policy text files that are created by the tool are transferred to z/OS using FTP
- Allows policy definition to be performed at higher level of abstraction than policy file statements
 - Define policy for both CS IPsec and AT-TLS as a single administrative task
 - ▬ Generates separate policy files for CS IPsec and AT-TLS
- Note: The uploaded policy configuration text files can be directly edited on z/OS; however policy tool persistent data store on the workstation will not have changes and are not reflected back into the tool



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

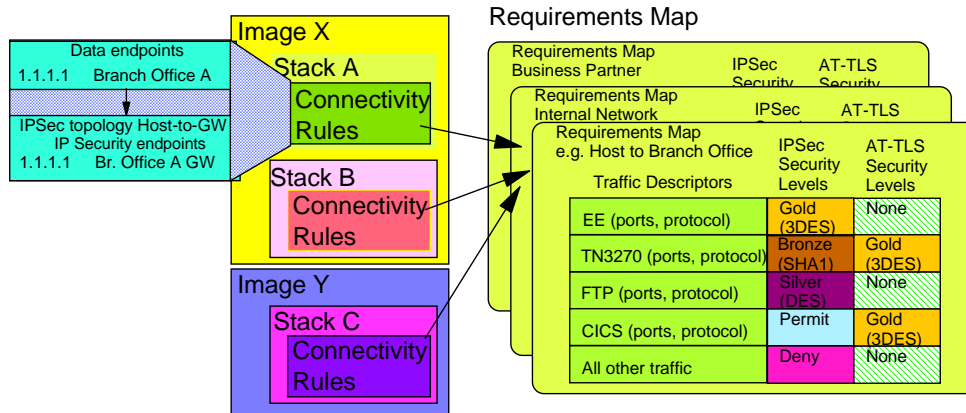
Network security configuration assistant - example



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

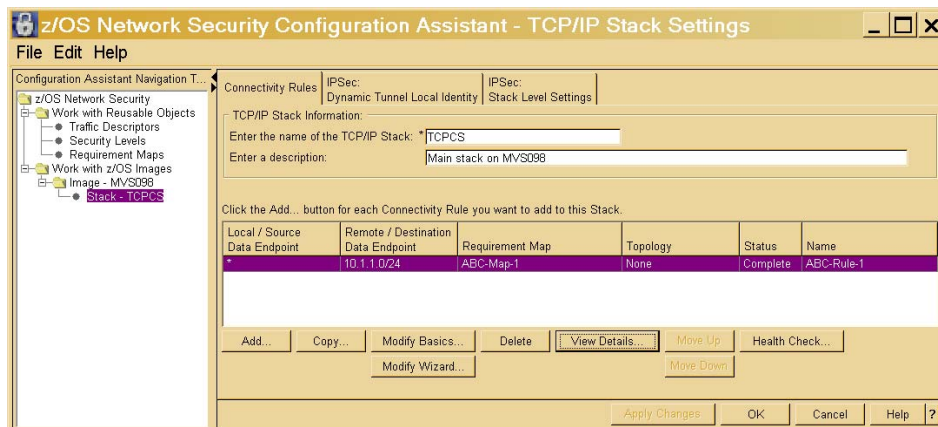
Network security configuration assistant - configuration data model



- **A system image contains one or more stacks**
 - Multiple system images may be defined
- **A stack contains a set of connectivity rules**
 - Data endpoint information
 - Security endpoint information
- **Reusable objects (can be shared across images and stacks)**
 - Requirements Map, Security Level, Traffic Descriptor

Connectivity rule example

- **A stack's connectivity rule applies a requirement map to a pair of data endpoints.**
- **The IPv4 addresses in a packet are compared with the IPv4 addresses of the data endpoints of the connectivity rules in the order that those rules appear in the table.**
- **When the IPv4 addresses match, the packet is compared with that connectivity rule's traffic descriptors in the order they appear in the requirement map; when a match is found, the corresponding security level is applied. For IPSec, each requirement map ends with an implicit rule to deny all traffic.**
- **For AT-TLS, if a packet matches no rule, it is allowed to flow with no AT-TLS protection.**



Requirement map example

> A requirement map is a collection of traffic descriptors

- ▶ You might define a requirement map named BranchOffice that provides a high level of protection for TN3270 and Web traffic but disallows (denies) all other traffic.
- ▶ You might define another requirement map named BusinessPartner that provides a high level of protection for Web traffic but disallows all other traffic.
- ▶ Then you could associate BranchOffice with the addresses of your branch offices in some connectivity rules.
- ▶ And associate BusinessPartner with the IPv4 addresses of your business partners in other connectivity rules.



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

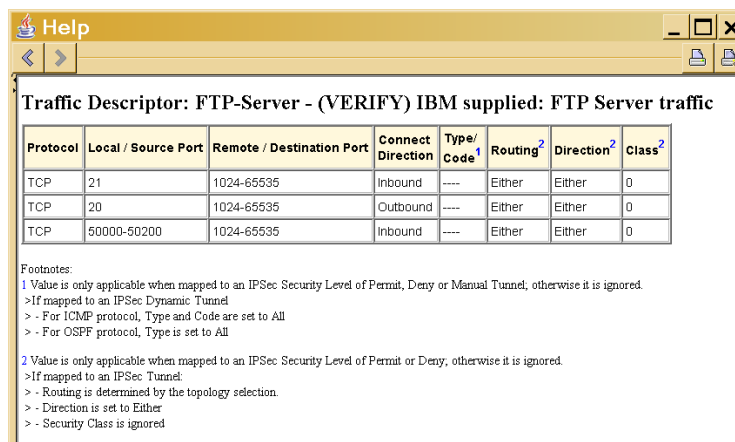
Traffic descriptor example

> The IP Security configuration assistant comes with many traffic types already defined

- ▶ They can be used as-is
- ▶ Or they can be modified to better match your local needs

> This is an example of FTP server traffic

- ▶ You may want to change the port range for passive data connections based on your local FTP server's PASSIVEDATAPORT value
 - In this example, we use the range from 50,000 to 50,200

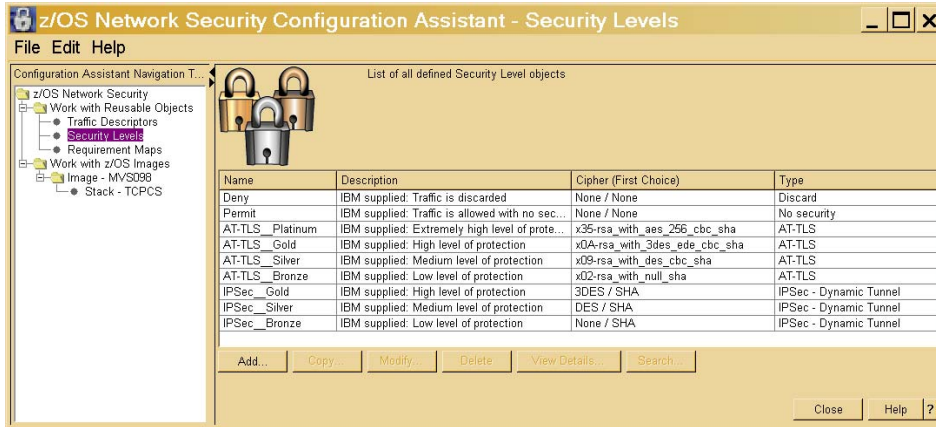


© Copyright IBM Corp. 2005. All rights reserved.

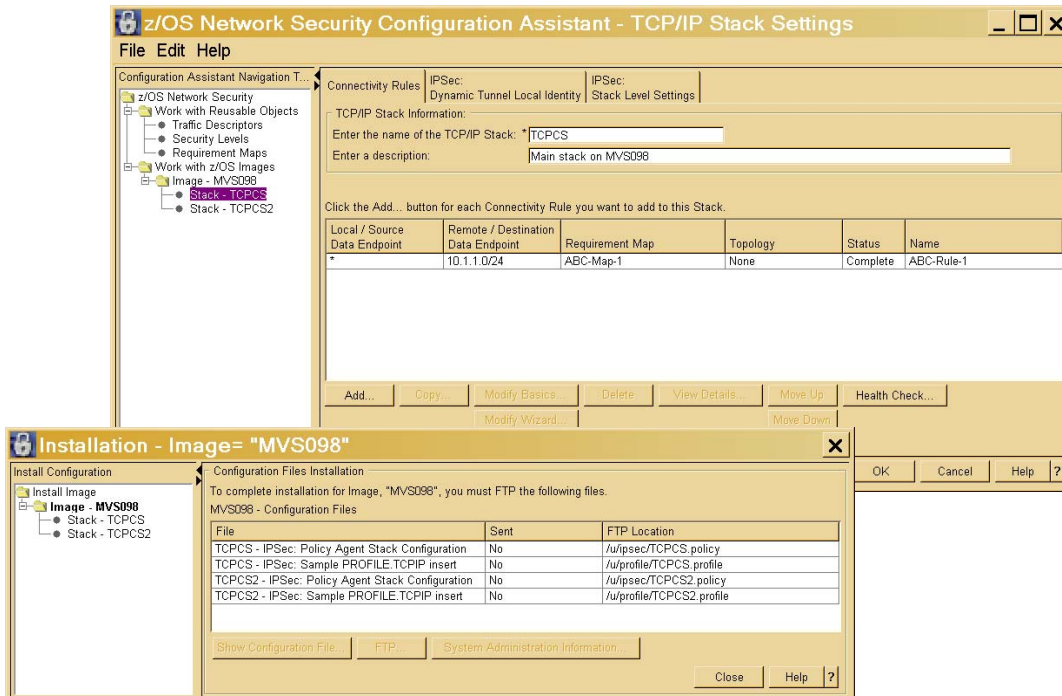
ibm.com/redbooks

Security levels

- Security levels define different ways to protect data in the network:
 - ▶ IPsec - Gold/Silver/Bronze levels
 - ▶ AT-TLS - Platinum/Gold/Silver/Bronze levels



Getting ready to FTP the policy agent configuration files to z/OS



Example policy agent configuration file for IP security and AT-TLS

```

IPSec Policy Agent Configuration File for Stack: T...
##
## IPSec Policy Agent Configuration file for:
## Image: MVS098
## Stack: TCPCS
##
## Created by the z/OS Network Security Configuration Assistant
## Date Created: Wed Aug 31 16:13:40 EDT 2005
##
## Copyright = None
##
IpGenericFilterAction      Permit-LogYes
{
  IpFilterAction           Permit
  IpFilterLogging          Yes
}

IpGenericFilterAction      Deny-LogYes
{
  IpFilterAction           Deny
  IpFilterLogging          Yes
}

IpService                  DNS
{
  Protocol                 UDP
  SourcePortRange          53
  DestinationPortRange    1024 65535
  Direction                BiDirectional
  Routing                  Either
}

IpService                  DNS-1
{
  Protocol                 UDP
  SourcePortRange          53
  DestinationPortRange    53
  Direction                BiDirectional
  Routing                  Either
}
    
```

➤ Locate or create a new Policy Agent configuration file that identifies the target stack by jobname and the location of its image file.

▸ The image file indicates the location of the policy configuration file.

➤ For example, if the stack jobname is TCPCS, then the Policy Agent configuration file `/etc/pagent.conf` contains the following statement:

▸ `TcpImage TCPCS /etc/tcps1.image`

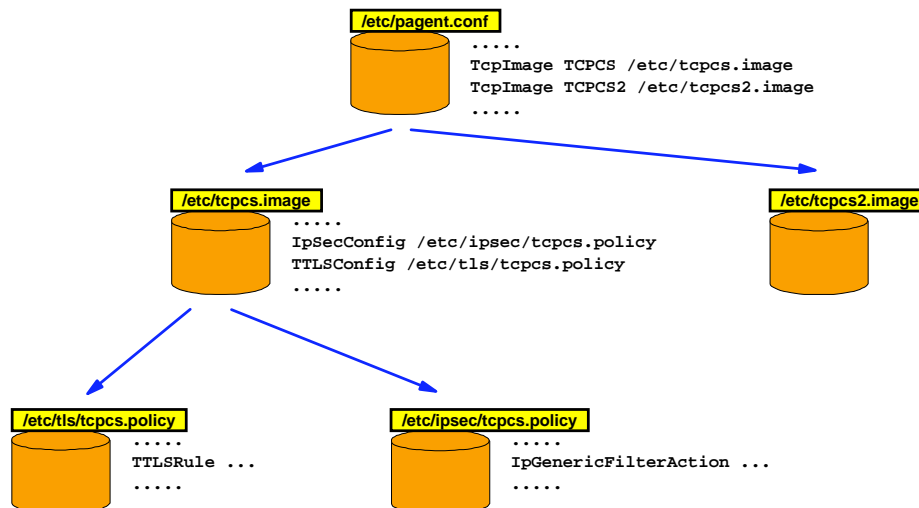
➤ And `/etc/tcps.image` contains the following statement:

▸ `IpSecConfig /etc/tcps.policy`

➤ And start Policy Agent:

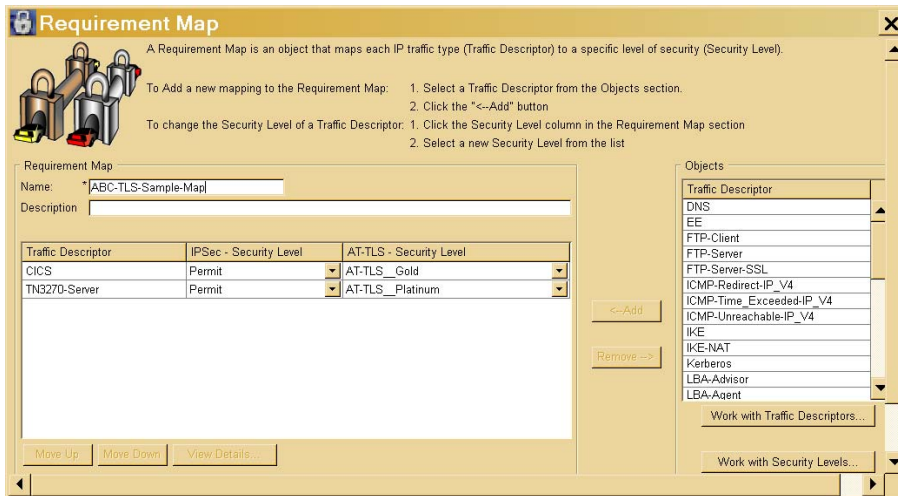
▸ `pagent -c /etc/pagent.conf`

PAGENT configuration file relationship



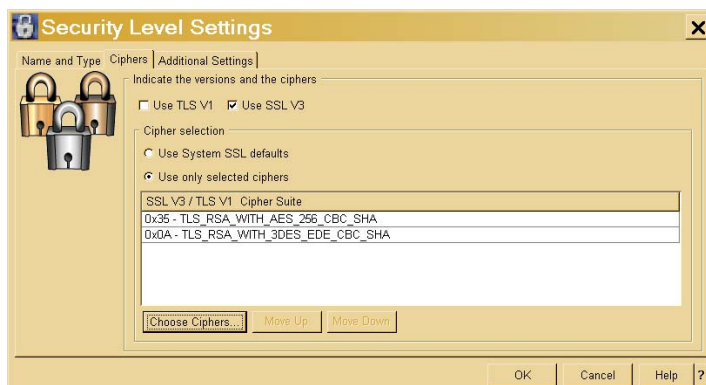
AT-TLS example for TN3270 and CICS

- **Start making a requirement map**
 - ▶ Copy the AT-TLS_Sample as a starting pint

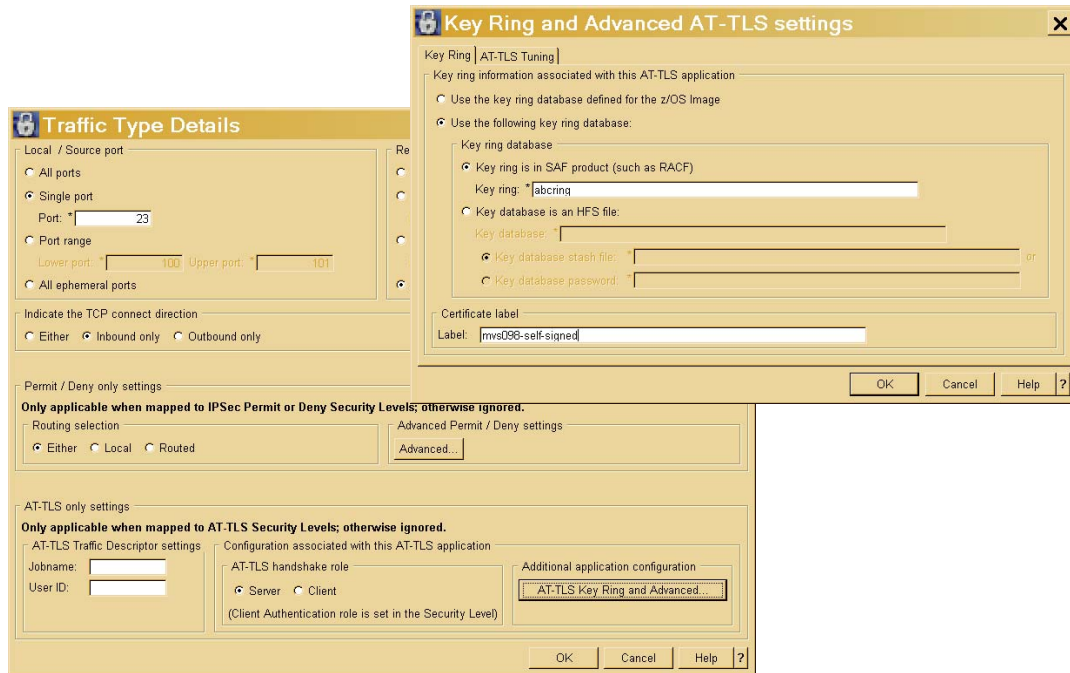


AT-TLS security level details

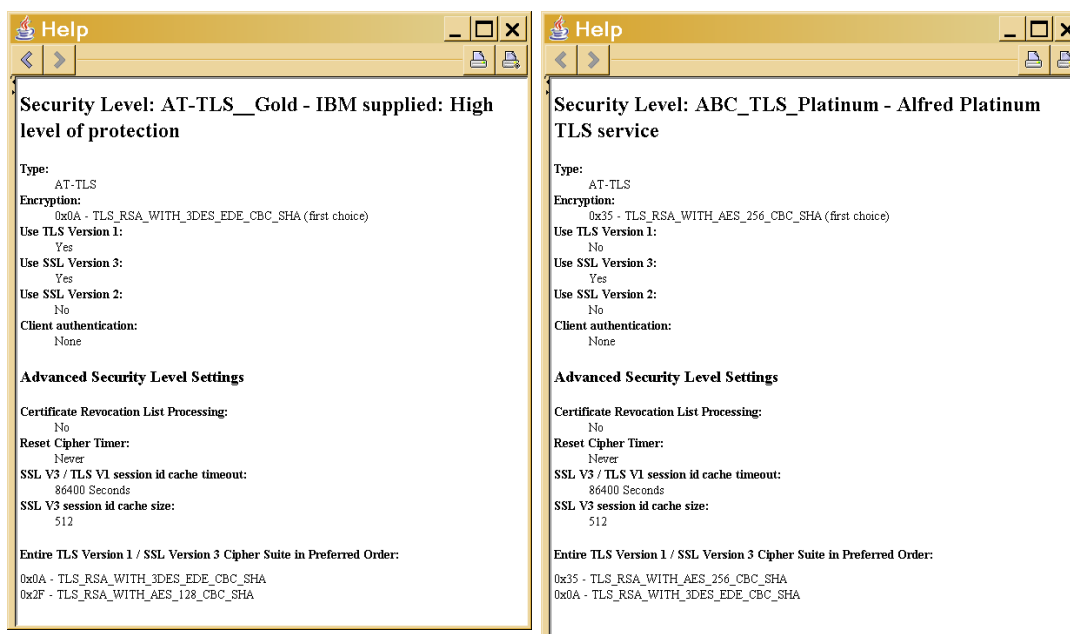
- The keyring may either be in an HFS file (managed by GSKKYMANT) or in RACF
- The keyring location can be specified at a z/OS image level or on a traffic descriptor that describes a specific application
- SSL/TLS protocol levels and ciphers can be chosen in the security level settings
- Support for checking with a Certificate Revocation List server (or multiple) is also supported



AT-TLS keyring specification in a traffic descriptor



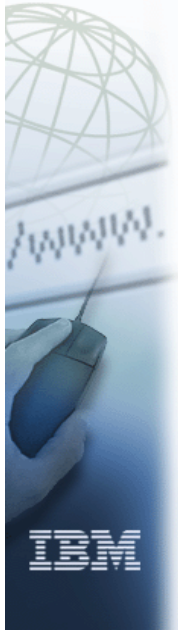
AT-TLS gold and platinum service levels



ibm.com



Communications Server for z/OS V1R7 - Technical Update CICS Sockets



Redbooks

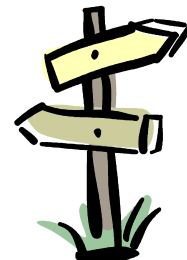
International Technical Support Organization

© Copyright IBM Corp. 2005. All rights reserved.

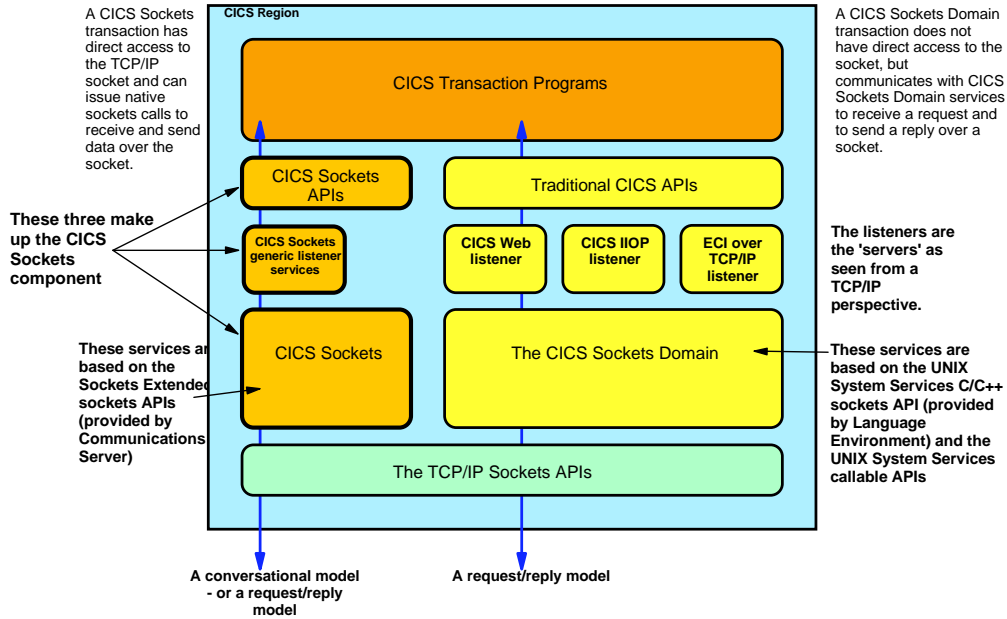
CICS Sockets

This section or parts thereof will only be presented if requested by workshop participants.

- > CICS Sockets background
- > Performance enhancement - CICS Sockets tracing
- > Performance enhancement - CICS monitoring
- > Move TRUE to 31-bit storage
- > Various general CICS Sockets enhancements
- > Using the CICS Open Transaction Environment (OTE)
- > SSL/TLS enabling CICS Sockets transactions through AT-TLS



Direct TCP/IP communication into the CICS region - an overview



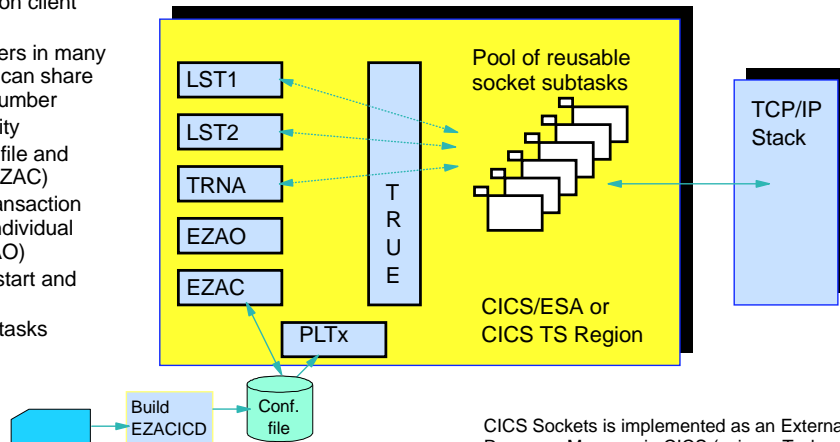
© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

CICS Sockets overview

- Multiple listeners - each instance separately configurable
- Enhanced listener has no requirements on client input data
- Multiple listeners in many CICS regions can share listener port number
- User ID security
- Configuration file and transaction (EZAC)
- Operations transaction to start/stop individual listeners (EZAO)
- PLT-enabled start and termination
- Reusable subtasks
- IPv6 support

- ▶ CICS Sockets is a component of the Communications Server for z/OS, not CICS TS itself.
- ▶ It is a general-purpose sockets programming API to be used by CICS application programmers for implementing native (low-level) sockets communication in z/OS CICS transaction programs.



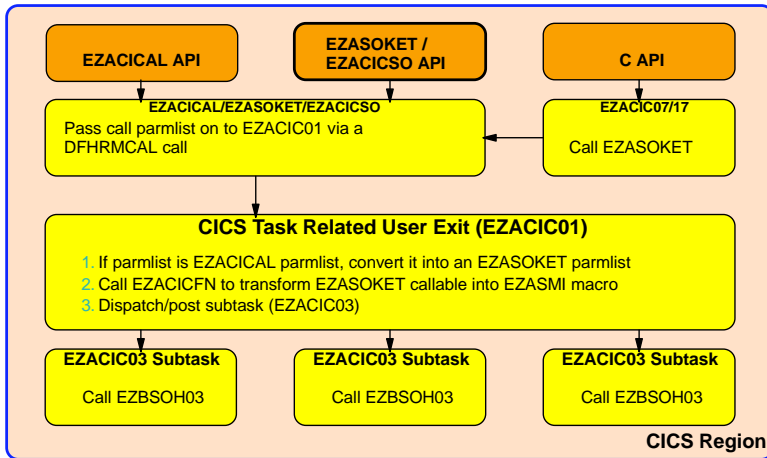
CICS Sockets is implemented as an External Resource Manager in CICS (using a Task Related User Exit - a TRUE).



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

CICS Sockets APIs



➤ If you use the new entry point name (EZACICSO) that was introduced in z/OS V1R4, then you don't have to do anything special during link-edit of your CICS Sockets program.

➤ If your program calls EZACICAL or EZASOKET, please make sure you submit link-edit input control statements as outlined below.

PQ28963 ships re-entrant version of EZACIC07, called EZACIC17.

```
CICS C-Socket Program
Linkage Edit control:

//SYSLIN DD *
INCLUDE SYSLIB(EZACIC07)
NAME MYCPGM(R)
/*
```

```
CICS Call EZACICAL and Call
EZASOKET program Linkage Edit control:

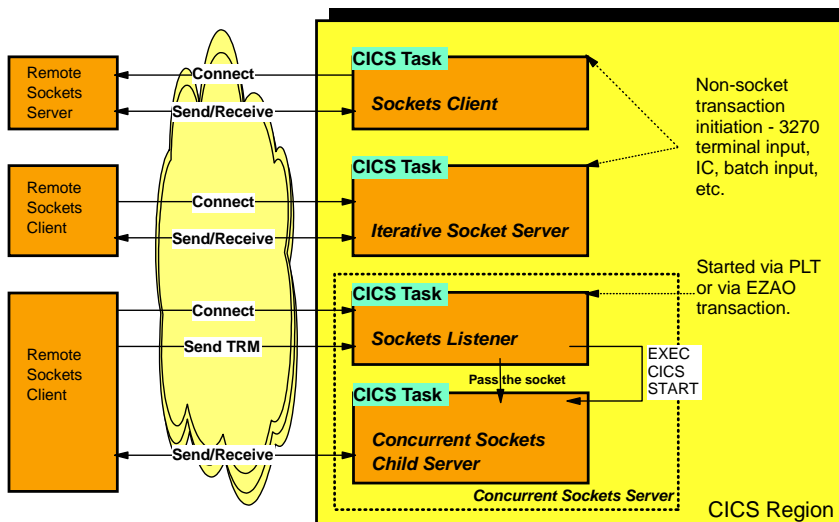
//SYSLIN DD *
INCLUDE SYSLIB(EZACICAL)
NAME MYSOKPGM(R)
/*
```



Remember this, Please!!

CICS Sockets program categories in CICS

NOTES



TRM: Transaction Request Message

CICS Sockets trace control - enable/disable

➤ **The IP CICS Sockets interface will only issue CICS Trace records when required through the use of the TRACE configuration option.**

- ▶ Associated with the IP CICS Sockets interface.
- ▶ Specified as TRACE=YES|NO with the default being YES.
- ▶ When TRACE is specified as YES then the IP CICS Sockets interface will generate CICS AP 199 or '00C7'x records for every EZASOKET call.
 - These records are generated when a CICS Auxiliary trace is active.
- ▶ Requires the CICS Master User trace flag to be on.
- ▶ Formatted by using the CICS Auxiliary Trace formatting utility program.
- ▶ Four records are written for each EZASOKET call.
- ▶ Can be specified on the EZACICD macro when building the CICS Sockets configuration file.
- ▶ Can also be specified when defining a CICS system using the EZAC transaction.
- ▶ Can be turned on/off dynamically via the EZAO START/STOP TRACE transaction

```

EDIT ---- CFGTRACE JCL A1 ----- COLUMNS 001 080
COMMAND ==>>                                SCROLL ==>> CSR
000075 CICS1A  EZACICD TYPE=CICS,           Generate configuration record      X
000076                APPLID=CICS1A,       APPLID of CICS                    X
000077                TCPADDR=TCPCS,      Address space name for TCP/IP     X
000078                CACHMIN=0,          Minimum refresh time for CACHE   X
000079                CACHMAX=20,         Maximum refresh time for CACHE   X
000080                CACHRES=5,          Maximum number of active resolvers X
000081                OTE=YES,             Use Open Transaction Environment  X
000082                TCBLIM=12,          TCB Limit                          X
000083                TRACE=NO,           No tracing needed                 X
000084                ERRORTD=TCPM         Name of TD queue for error messages
    
```



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

CICS Sockets trace performance impacts

➤ **CICS Sockets trace improvements**

- ▶ CPU reduced when using z/OS V1R7 CICS Sockets tracing (TRACE=YES)
- ▶ New option (TRACE=NO) to turn off CICS Sockets tracing
- ▶ z/OS V1R7 vs V1R6 :

| Release Trace ON/OFF | Trans / Second | Trans/Sec Delta % | CPU/Tran | CPU/Tran Delta % |
|----------------------|----------------|-------------------|----------|------------------|
| V1R6 (Trace ON) | 1552 | Base | 1142.4 | Base |
| V1R7 Trace=YES | 1682.3 | + 8.4 % | 1082.9 | - 5.2 % |
| V1R7 Trace=N0 | 1824.7 | + 17.7 % | 1020.4 | - 10.7 % |



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Things to think about when deciding whether to enable or disable CICS Sockets tracing

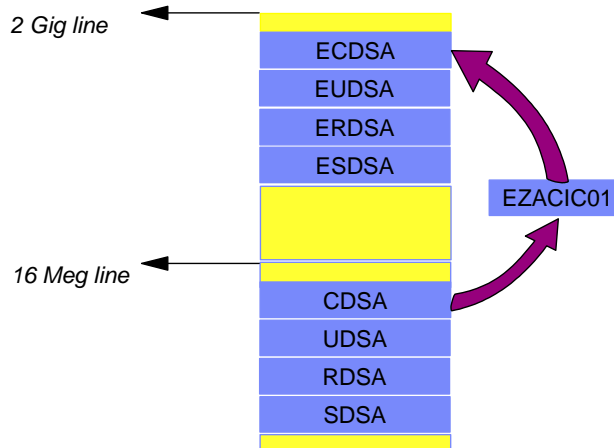
- **The data contained in the IP CICS Sockets trace records are designed to assist the IBM service team in diagnosing a problem. If at all possible, try to capture this trace data by ensuring the following:**
 - ▶ That the CICS Master User flag is on.
 - Use the CICS CETR transaction to verify this flag.
 - ▶ That the IP CICS Sockets TRACE=YES is specified.
 - Use the EZAO command to dynamically enable tracing.
 - ▶ That the CICS Auxiliary trace is active when tracing to the auxiliary trace data sets.
 - Use the CEMT SET AUX command to enable CICS Auxiliary tracing.
- **No migration changes are needed to continue IP CICS Sockets CICS tracing, since the default for TRACE is YES.**
- **If you would like to prevent IP CICS Sockets trace records from being generated, say for a production-level CICS region, then:**
 - ▶ Either add an entry on the EZACICD,TYPE=CICS macro specifying TRACE=NO
 - ▶ Or use the EZAC,ALTER,CICS transaction specifying TRACE as NO.
 - ▶ Or dynamically alter the tracing status by specifying EZAO,STOP,TRACE or by specifying EZAO,SET,CICS specifying TRACE as NO. If you dynamically change TRACE then make sure you reflect that change in your IP CICS Sockets configuration.

CICS Sockets - CICS monitoring performance enhancement by dynamically learning which EMPs are enabled

- **The IP CICS Sockets interface will create CICS Event Monitoring Point (EMP) data only when an associated entry exists in the CICS Monitor Control Table (MCT).**
 - ▶ EMPs are recorded for both the Task Related User Exit (TRUE), EZACIC01, and the Listener, EZACIC02.
 - ▶ There is no external control other than the MCT entries and the overall CICS Monitoring status.
 - ▶ A check is made by IP CICS Sockets to determine whether the EMP about to be executed has previously failed. The failure being tested is due to CICS returning a response of INVREQ whenever the EXEC CICS MONITOR command is invoked. When the EMP has failed with a response of INVREQ then all future attempts to execute that specific EMP will be disabled.
 - ▶ If the EMP has not previously failed then the EXEC CICS MONITOR command will be issued.
 - ▶ If the EMP has previously failed then the EXEC CICS MONITOR will be skipped.
 - ▶ The TRUE and Listener will steadily learn what EMPs are not specified in the MCT.
 - ▶ The use of the IP CICS Sockets MCT entries are totally optional. All or any number may be specified in the MCT.
 - ▶ IP CICS Sockets must be recycled to reset any disabled EMPs. The MCT must be updated to reflect any desired associated entries.
- **The IP CICS Sockets MCT entries are designed to give statistical information about the usage of the IP CICS Sockets interface and Listener. They are currently broken up into two distinct categories:**
 - ▶ Task Related User Exit (TRUE) - EZACIC01
 - The TRUE is invoked for each call to EZASOKET.
 - ▶ Listener - EZACIC02
 - The listener is basically an application program that calls EZASOKET.

Move CICS Sockets Task Related User Exit (TRUE) modules to 31-bit storage

- IP CICS Sockets Task Related User Exit (TRUE) is EZACIC01 and is the vehicle used to support sockets in CICS application programs. TRUEs are enabled to CICS and CICS will load these programs in the Dynamic Storage Areas based on link edit attributes.



Dynamic Storage Area's (CICS/ESA 4.1 and up):



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Enhanced CICS Sockets configuration and operator transactions

NOTES

- Both the configuration, EZAC, and operator, EZAO, transaction share the same Basic Mapping Support (BMS) map. It continues to support a 3278-2 (24x80) terminal type.
- In order to add new attributes to the enhanced Listener definition, it was required to move to a second screen.
 - ▶ The attributes on the first Listener screen are those that are common to both the standard and enhanced listener.
 - ▶ The attributes on the second Listener screen are those that are unique to the standard and enhanced Listener.

```

EZAC,DISplay,LISTENER (enhanced listener. screen 1 of 2)  APPLID = CICS1A

APPLID      ==> CICS1A          APPLID of CICS System
TRANID      ==> CSKM           Transaction Name of Listener
PORT        ==> 03011         Port Number of Listener
AF          ==> INET          Listener Address Family
IMMEDIATE   ==> YES           Immediate Startup Yes|No
BACKLOG     ==> 020          Backlog Value for Listener
NUMSOCK     ==> 050          Number of Sockets in Listener
ACCTIME     ==> 999          Timeout Value for ACCEPT
GIVTIME     ==> 999          Timeout Value for GIVESOCKET
REACTIME    ==> 999          Timeout Value for READ

Verify parameters, press PF8 to go to screen 2

PF 3 END                8 NEXT                12 CNCL
    
```

- Enhanced Listener attributes: screen 1 of 2.



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Enhanced configuration and operator transactions

NOTES

➤ Configuration and operator transaction enhancements:

- ▶ Consistent flow across all elements
- ▶ Better field placement
- ▶ Consistent look, feel, field attributes
- ▶ Protect unchangeable fields
- ▶ Enhanced error checking
- ▶ Enhanced run-time checks
- ▶ Enhanced change control


```
EZAC,DISplay,LISTENER (enhanced listener. screen 2 of 2) APPLID = CICS1A

CSTRANId   ==> CIST           Child Server Transaction Name
CSSTYPe    ==> KC             Startup Method (KC|IC|TD)
CSDELAY    ==> 000000         Delay Interval (hhmmss)
MSGLENgth  ==> 011           Message Length (0-999)
PEEKDATA   ==> YES           Enter Y|N
MSGFORMat  ==> ASCII         Enter ASCII|EBCDIC
USEREXIT   ==> CISTSE        Name of User/Security exit
GETTID     ==> NO            Get AT-TLS ID (YES|NO)
USERID     ==>              Listener User ID
WLM group 1 ==>              Workload Manager Group Name 1
WLM group 2 ==>              Workload Manager Group Name 2
WLM group 3 ==>              Workload Manager Group Name 3

Verify parameters, press PF7 to go back to screen 1
Press ENTER or PF3 to exit

PF 3 END           7 PREV           12 CNCL
```


➤ Enhanced Listener attributes: screen 2 of 2.

 © Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Various smaller enhancements to CICS Sockets

- Enhance the CICS trace data reported on the AP 199 '00C7'x CICS trace records.
- Enhance the EZACICD macro to allow it to support more than 255 listener definitions per interface.
 - ▶ Maximum is now 4095
- Enhance the listener definition to allow an appropriate user ID to be associated with the Listener task and possibly the started child server transaction.
 - ▶ If this new parameter is not specified, then the Listener task gets the user ID from either the CICS PLT user ID (if the Listener is started via the CICS PLT) or the ID of the user that invoked the EZAO transaction (if the Listener is started via the EZAO transaction).
 - ▶ If this new parameter is specified, then any user that starts the Listener (the PLT user if the Listener is started via the PLT) must have surrogate security access to this user ID. This user ID would have to be permitted to any resources the Listener accesses, such as child server transactions and programs.
- Define EZACIC06 as reentrant. Reentrant programs calling EZACIC06 will be able to establish and retain the reentrancy attribute when being link edited.
 - ▶ EZACIC06 is a utility program that may be used to translate bit-masks into character arrays and character arrays into bit-masks. This program is useful for COBOL programmers for building and interpreting bit-masks for SELECT and SELECTEX calls.
- Use the CSMT CICS Transient data queue when ERRORTD is not defined to CICS.
 - ▶ ERRORTD is the name of the CICS Transient Data queue used to deliver IP CICS Sockets messages. If the queue name specified by ERRORTD is not found during IP CICS Sockets initialization then the default CSMT queue is used.
 - ▶ Ensure that the Transient data queue is defined to CICS. A recycling of IP CICS Sockets will be necessary to reset the ERRORTD queue once it is forced to CSMT.

 © Copyright IBM Corp. 2005. All rights reserved.

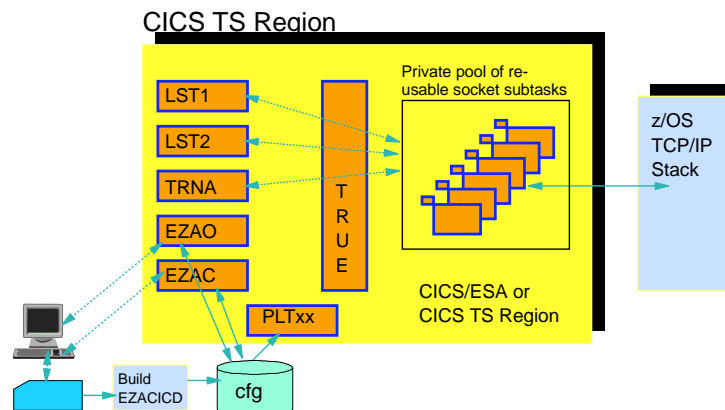
ibm.com/redbooks

What is the CICS quasi-reentrant TCB?

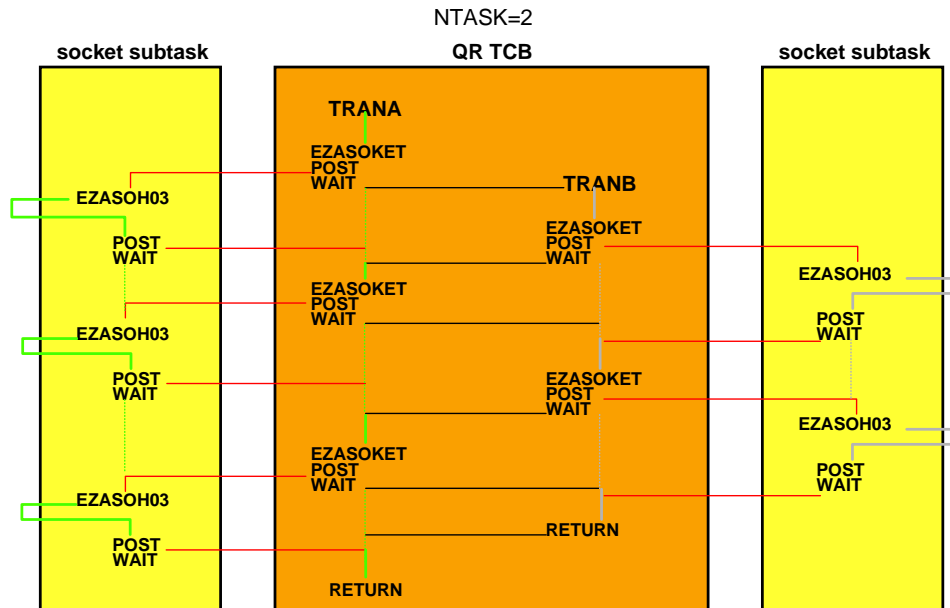
- ▶ The CICS quasi-reentrant (QR) Task Control Block (TCB) is the TCB where the customers application work is processed.
- ▶ Programs are said to be quasi-reentrant programs because they take advantage of the behavior of the CICS dispatcher and the QR TCB.
 - ▶ There is only ever one CICS task active under the QR TCB.
 - The same program can be executed by multiple CICS tasks
 - Only one of those CICS tasks is active at any given point in time
- ▶ Quasi-reentrant programs running under the QR TCB are safe in the knowledge that they are the only CICS user task running at that instance.
 - ▶ Can access shared resources such as the CICS Common Work Area (CWA)
 - ▶ Can access shared storage obtained via EXEC CICS GETMAIN SHARED
 - ▶ Running under the QR TCB guarantees serialized access to shared resources
- ▶ The QR TCB structure limits multi-processing.
 - ▶ One of the key reasons why multiple CICS regions are typically deployed for scalability in a multiprocessor environment.

CICS Sockets subtasking method - private: NTASKS

- ▶ The current subtasking method IP CICS Sockets uses is that of a privately managed pool of MVS TCBs.
 - ▶ Determined by what is coded for the NTASKS configuration option
 - ▶ Established when IP CICS Sockets is initialized
 - New TCBs are dynamically generated
 - If the subtask pool is defined too small for the IP CICS Sockets workload
 - The task is a listener
 - The same TCB will be DETACHed when the task using it has ended



EZASOKET transactions in CICS/TS 1.3



What is CICS Open Transaction Environment?

➤ CICS TS Open Transaction Environment (OTE) introduces a new class of Task Control Blocks (TCBs) called an open TCB, which can be used by applications.

- ▶ Characterized by the fact it is assigned to a CICS task for the life of the CICS task
- ▶ Multiple OTE TCBs may run concurrently in CICS

➤ There are several modes of open TCBs, used to support various functions.

- ▶ Java in CICS, for example employs a type of OTE TCB commonly referred to as "J8"
- ▶ Open API Task Related User Exits employ the "L8" TCB

➤ There is no sub-dispatching of other CICS tasks under the open TCB.

- ▶ An application executing under an open TCB can issue non CICS API requests that may involve the TCB being blocked.
- ▶ Blocking is allowed because only this TCB is halted, and not the whole of CICS
 - This is what happens if a blocking EZASOKET request is issued under the QR TCB.
 - Blocking means the TCB is halted, the TCB is not being dispatched.
- ▶ Examples of non CICS APIs would be:
 - MVS services
 - GETMAIN
 - MVS UNIX System Services POSIX functions.
 - DB2 SQL
 - MQSeries

What is CICS Open Transaction Environment? (*continued*)

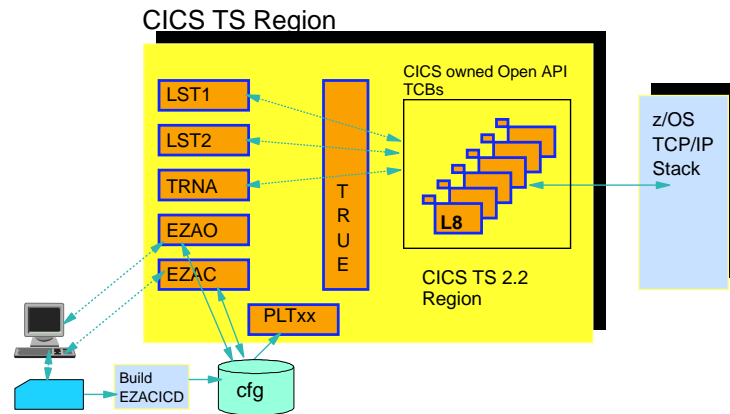
- **Since multiple tasks can potentially access shared resources simultaneously when executing under an OTE TCB, applications that access shared resources must bear the responsibility of ensuring the integrity of those resources by implementing an appropriate serialization technique.**
 - ▶ For example, a counter in the CICS common work area (CWA)
- **CICS assumes responsibility for ensuring the integrity of the resources it manages.**
 - ▶ Either the CICS TS code has been amended to run on multiple TCBs safely
 - temporary storage requests
 - ▶ Or CICS TS will ensure the code runs on the QR TCB
 - File Control requests.
- **Therefore the use of non-threadsafe CICS commands that must run on the QR TCB has a performance penalty (due to the need to switch TCBs), but there is no risk to data integrity.**
- **If the same quasi-reentrant program would run in an OTE environment, multiple instances of this program could execute at the same time.**
 - ▶ The counter value in the CWA could be changed by multiple executors at the same time and one instance would never be sure about the counter value when it stops or gets suspended.

What is threadsafe from a CICS application perspective?

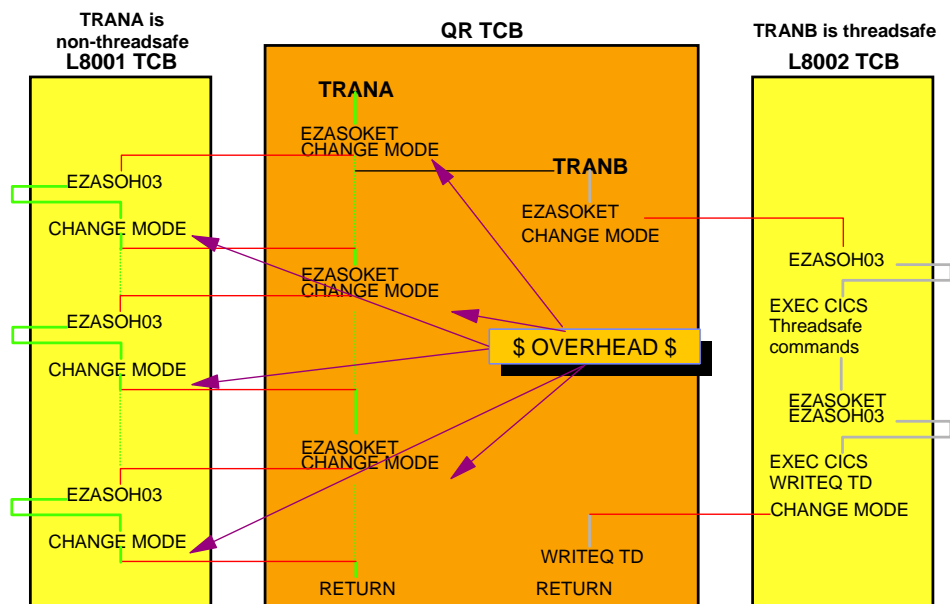
- **“Threadsafe application”**
 - ▶ A collection of application programs that employ an agreed-upon form of serialized access to shared resources.
- **A program written to “threadsafe standards” is a program that implements the agreed-upon serialization techniques.**
- **It is important to understand a single program operating without the agreed-upon serialization technique can destroy the predictability and therefore integrity of an entire system of otherwise threadsafe programs.**
- **Therefore, an application system cannot be “threadsafe” until all programs that share a common resource implement that application’s threadsafe standards.**

CICS Sockets subtasking method - Open API

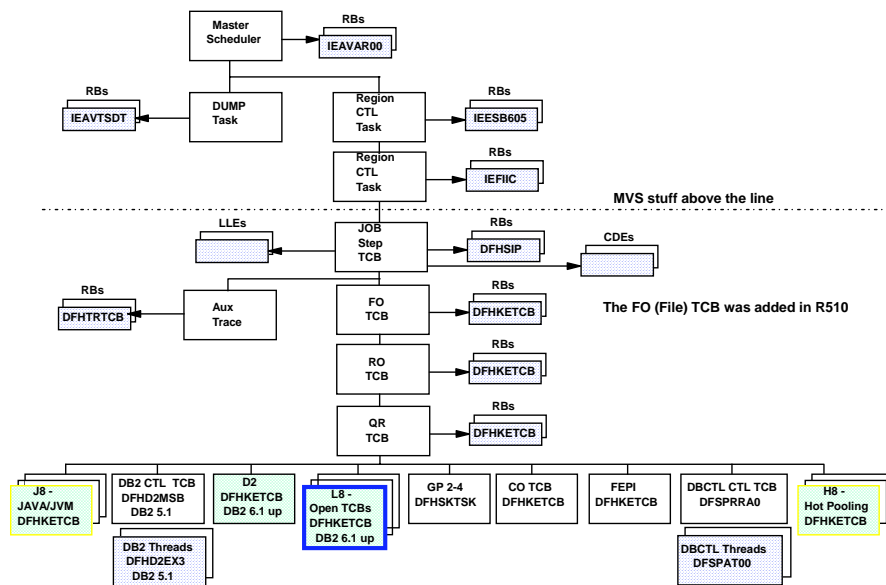
- When exploiting OTE, the IP CICS Sockets Task Related User Exit (TRUE) will be enabled as Threadsafe. When it is invoked by an EZASOKET call, CICS will switch the task from the QR TCB to an L8 TCB.
- The L8 TCBs are solely managed by CICS TS. The active L8 TCB pool size is limited by the CICS MAXOPENTCBS System Initialization parameter. The CICS ACTOPENTCBS will indicate the number of L8 TCBs in use at any instance.



EZASOKET transactions in CICS/TS 2.2



MVS TCB structure for CICS/TS R220



Update CICS configuration for OTE - MAXOPENTCBS

- CICS TS must be upgraded to at least V2R2 with any Open Transaction Environment and threadsafe PTFs applied. IP CICS Sockets will perform a runtime check to ensure this environment exists before the interface is enabled.
- MAXOPENTCBS is a CICS TS configuration option that is used to limit the size of the Open API, L8, TCB pool. Its range is from 1-2000 with a default of 12. When the number of tasks using L8 TCBS reaches MAXOPENTCBS, then any new work will be suspended by CICS TS until tasks end or MAXOPENTCBS is increased.
 - ▶ Remember, TCB storage is allocated from Local System Queue Area (LSQA). MAXOPENTCBS can be set by using the CEMT Set Dispatcher MAXOpentcbs(nnnn) command.

```

File Edit Edit_Settings Menu Utilities Compilers Test Help
-----
EDIT          DCICS.V2R3M0.SYSIN(DFH$SIPT) - 01.83          Columns 00001 00080
Command ==>>>                                         Scroll ==>> CSR
000021 MXT=260                               Set maximum tasks to 32
000022 DSALIM=16M
000023 EDSALIM=640M
000024 SPOOL=YES                               System spooling interface is required
000025 MAXOPENTCBS=260                          Limit of Open API TCB in pool
000026 FORCEQR=NO                               Do not force threadsafe pgms to QR TCB
000027 MCT=SO                                  Monitor Control Table for Sockets
000028 MN=ON                                   Monitor Control on at initialization
000029 MNEXC=ON                               Exception class monitoring is active
000030 MNPER=ON                               Performance class monitoring is active
    
```


Configure IP CICS Sockets for OTE

> OTE

- ▶ A value of YES causes the IP CICS Sockets task-related user exit to execute using the CICS Open Transaction Environment.
- ▶ A value of NO causes IP CICS Sockets to continue executing EZASOKET calls on an MVS subtask managed by the IP CICS Sockets interface.
- ▶ If OTE=YES, then the values of NTASKS, DPRTY and TERMLIM will be forced to zero if specified.

> OTE is supported on CICS/TS V2R2M0 and higher.

- ▶ If OTE=YES is specified on a pre-CICS/TS V2R2M0 system, then the IP CICS Sockets interface will fail initialization.

> When OTE=YES is specified, CICS TS will switch all calls from the QR TCB to an L8 TCB

- ▶ All EZASOKET calls
- ▶ All IP CICS C Socket functions

> IP CICS Sockets applications must be

- ▶ Coded using threadsafe programming practices as defined by CICS and
- ▶ Defined to CICS as threadsafe

```
EDIT ---- CFGOTE JCL A1 ----- COLUMNS 001 080
COMMAND ==> SCROLL ==> CSR
000075 CICS1A  EZACICD TYPE=CICS,      Generate configuration record      X
000076                APPLID=CICS1A,    APPLID of CICS                    X
000077                TCPADDR=TCPCS,    Address space name for TCP/IP     X
000078                CACHMIN=0,        Minimum refresh time for CACHE    X
000079                CACHMAX=20,       Maximum refresh time for CACHE    X
000080                CACHRES=5,        Maximum number of active resolvers X
000081                OTE=YES,          Use Open Transaction Environment   X
000082                TCBLIM=12,        TCBLIM=12                          X
000083                TRACE=NO,        No tracing needed                  X
000084                ERRORTD=TCPM       Name of TD queue for error messages
```



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Configure IP CICS Sockets for OTE - TCBLIM

> TCBLIM

- ▶ Specifies the maximum number of open TCBs that can be used by the IP CICS Sockets interface.
- ▶ Listeners will not be subject to this limitation; however, they will be subject to CICS's MAXOPENTCBS.
 - This allows listeners to be started thereby prohibiting a possible denial of service.
 - If MAXOPENTCBS is reached
 - Then no more open API TCBs are available in the CICS region and
 - The IP CICS Sockets task-related user exit cannot obtain an open TCB for its use
- ▶ If OTE=NO and TCBLIM>0, then TCBLIM will be forced to 0.
- ▶ IP CICS Sockets supports a TCB limiting mechanism to manage its use of the L8 TCBs.
- ▶ When TCBLIM is 0, no limiting factor is imposed.
 - TCBLIM=0 is the default.
- ▶ When TCBLIM is set to the same value as MAXOPENTCBS, it will never be enforced due to CICS's management of the L8 pool size.
- ▶ When a CICS region is at MAXOPENTCBS, any new work exploiting an Open API enabled TRUE will wait until an L8 TCB becomes available either through task end or by increasing MAXOPENTCBS
 - The EZAO,SET,CICS transaction can be used to change TCBLIM dynamically
- ▶ Listeners defined to the IP CICS Sockets interface are not subject to TCBLIM but are subject to MAXOPENTCBS



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

L8 TCB management for CICS Sockets

➤ The IP CICS Sockets Operator transaction can be used to inquire about the following:

- ▶ The current CICS MAXOPENTCBS setting
- ▶ The current number of L8 TCBs in use
- ▶ The current IP CICS Sockets TCBLIM setting
- ▶ The current number of L8 TCBs that are subject to TCBLIM
- ▶ The current number of tasks that are queued by TCBLIM
- ▶ The current queue depth of tasks that have been queued by TCBLIM

```
EZAO, INQUIRE, CICS                                APPLID = CICS1A

TRACE          ===> NO                            Trace CICS Sockets
MAXOPENTCBS    ===> 00260                          CICS Open API, L8, TCB Limit
ACTOPENTCBS    ===> 00000                          Active CICS Open API, L8, TCBS
TCBLIM         ===> 00000                          Open API TCB Limit
ACTTCBS        ===> 00000                          Number of Active Open API TCBS
QUEUEDEPTH     ===> 00000                          Number of Suspended Tasks
SUSPENDHWM     ===> 00000                          Suspended Tasks HWM

PF 3 END                                           12 CNCL
```



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

IP CICS Sockets components - threadsafe or not

➤ The following IP CICS Sockets programs are threadsafe:

- ▶ EZACIC01 - Task Related User Exit
- ▶ EZACIC02 - Listener
 - The Listener will incur less TCB switching if run on CICS TS V2R3
- ▶ EZACIC12 - WLM Registration/Deregistration
- ▶ EZACICME - Message module
- ▶ Sample programs
 - EZACICAC - Sample IPv4 child server
 - EZACIC6C - Sample IPv6 child server
 - EZACICSC - Sample COBOL child server
- ▶ Utility programs
 - EZACIC04, EZACIC05, EZACIC06, EZACIC08, EZACIC09, EZACIC14, EZACIC15
- ▶ Application stub
 - EZACIC17

➤ The following IP CICS Sockets programs do not need to be threadsafe:

- ▶ EZACIC00 - Operator
- ▶ EZACIC03 - MVS subtask
- ▶ EZACIC20 - PLT program
- ▶ EZACIC21 - Initialization
- ▶ EZACIC22 - Termination
- ▶ EZACIC23 - Configuration
- ▶ EZACIC25 - Domain Name Service
 - This is not marked as threadsafe as it will always incur a TCB switch due to non-threadsafe CICS commands.
- ▶ EZACIC07 is the non-reentrant application stub providing C socket support for non-reentrant C CICS programs. Application programs using the EZACIC07 application stub cannot be defined as threadsafe. Use EZACIC17 instead.

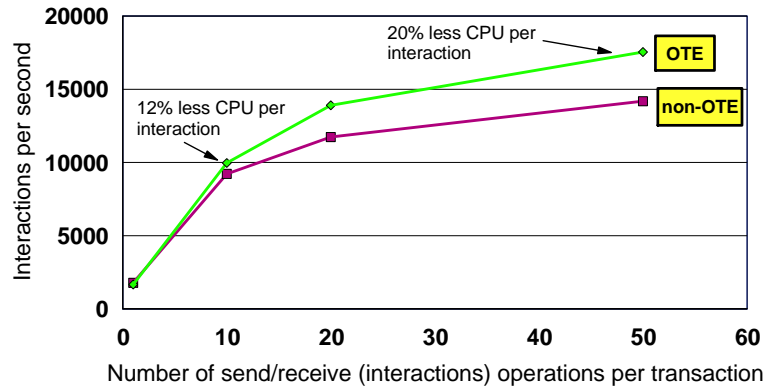


© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

CICS Sockets - OTE vs. non-OTE - performance

CICS Sockets Transactions OTE vs. non-OTE



- CICS Sockets transactions that issue many sockets calls and/or use both SQL and sockets calls will see most benefit from using OTE



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Things to think about when using OTE for CICS Sockets

- If you specify a listener user/security exit, then it must be coded to threadsafe programming standards and defined to CICS as CONCURRENCY(THREADSAFE) to prevent a TCB switch and ensure shared resource integrity.
- Child server transaction program must be coded to threadsafe programming standards and defined to CICS as CONCURRENCY(THREADSAFE) to prevent TCB switching and ensure shared resource integrity.
- Use the CICS supplied load module scanner program, DFHEIDTH, to locate non-threadsafes CICS commands in your programs

```
#####
# CICS LOAD MODULE SCANNER FILTER TABLE - THREADSAFE INHIBITORS
# This table identifies commands which "may" cause the program not to
# be threadsafe in that they allow accessibility to shared storage and
# the application must have the necessary synchronization logic in
# place to guard against concurrent update.
#####
# The extract command obtains the address and length of the global
# work area for the GLUE or TRUE.
#####
EXTRACT EXIT GASET *
GETMAIN SHARED *
ADDRESS CWA *
ASKTIME *           # Threadsafes in CICS TS V2R3
FORMATTIME *       # Threadsafes in CICS TS V2R3
SYNCPPOINT *
WRITE JOURNALNAME
WRITE FILE
WRITEQ TD
```

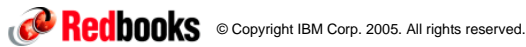


© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Things to think about when using OTE for CICS Sockets (continued)

- **Reassemble any user-written programs using any of the external IP CICS Sockets macros:**
 - ▶ EZACICA - IP CICS Sockets control blocks
 - ▶ EZACICSX - Listener security/user exit COMMAREA layout
- **The following APAR is required to exploit CICS Sockets OTE**
 - ▶ CICS TS V2R2 and V2R3
 - PQ93953 - CICS EXEC CICS SET TASK PURGE OR FORCEPURGE CMD PROCESSING FAILS
 - ▶ PTFs
 - UK01007 for CICS TS R2.2
 - UK01008 for CICS TS R2.3
- **The solution to the following APARs is recommended to enable best sockets performance in an OTE environment:**
 - ▶ APARs OA13252 and OA13278
- **Reference the following for more information on threadsafe programming practices:**
 - ▶ CICS TS documentation library
 - ▶ IBM Redbook - "*Threadsafe considerations for CICS*", SG24-6351
 - ▶ Share presentation - "What Does It Mean to be Threadsafe In CICS Transaction Server R2.2?" by Jim Grauel



© Copyright IBM Corp. 2005. All rights reserved.

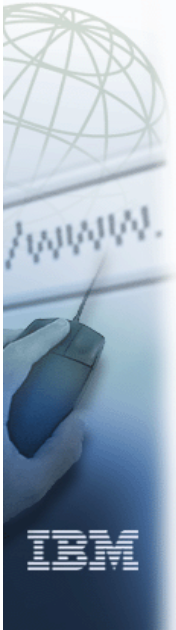
ibm.com/redbooks

This page intentionally left blank

ibm.com



Communications Server for z/OS V1R7 - Technical Update Network Management



Redbooks

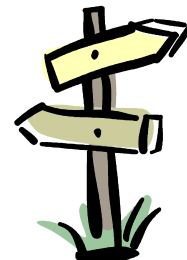
International Technical Support Organization

© Copyright IBM Corp. 2005. All rights reserved.

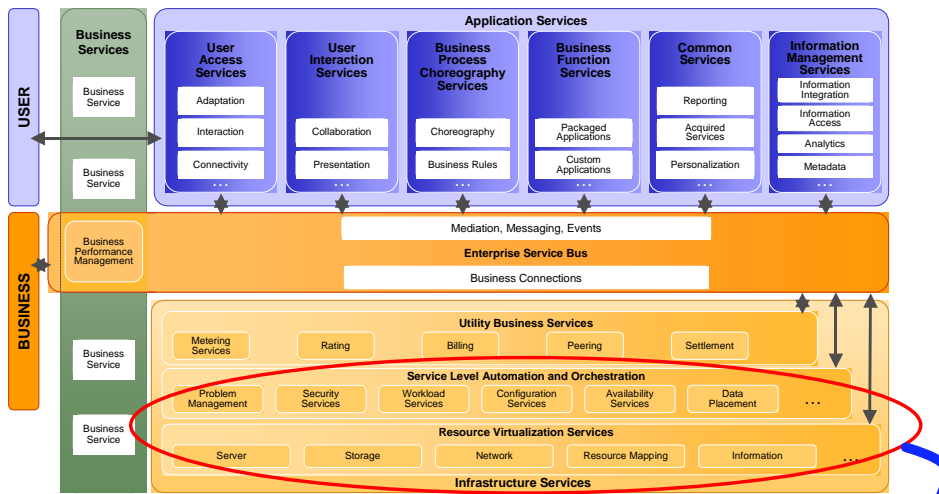
Management agenda

This section or parts thereof will only be presented if requested by workshop participants.

- > Adding CIM management in support of the On-Demand infrastructure
- > SNMP UDP IPv6 MIB support
- > Netstat changes in z/OS V1R7



On-Demand Operating Environment (ODOE)



➤ **Common Information Model (CIM) - one of the standards used for resource management!**



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Introduction to CIM

➤ Web-Based Enterprise Management (WBEM) initiative

- ▶ Key initiative of the Distributed Management Task Force (DMTF)
 - The DMTF is a nonprofit association of industry members (including IBM) dedicated to promoting enterprise and systems management and interoperability.
- ▶ Set of management and Internet standard technologies developed to unify the management of distributed computing environments.
- ▶ Defines the protocols and interfaces for CIM.

➤ Important core standards that make up WBEM

- ▶ Common Information Model (CIM) Standard
 - Provides an object-oriented data model
 - Provides a common definition of management information for systems, networks, applications, and services, and allows for platform extensions
 - CIM schema - Set of classes that define the data to be managed
- ▶ CIM-XML (one example of the WBEM protocols)
 - Defines CIM messages (operations) in XML over HTTP (for example, GetInstance, GetClass)
 - How CIM management entities communicate

➤ Related links

- ▶ DMTF - www.dmtf.org
- ▶ WBEM - www.dmtf.org/standards/wbem
- ▶ CIM schema v2.8 - www.dmtf.org/standards/cim/cim_schema_v28
- ▶ CIM client operations (for example, GetInstance) - www.dmtf.org/standards/documents/WBEM/DSP200.html



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Retrieving CIM management data

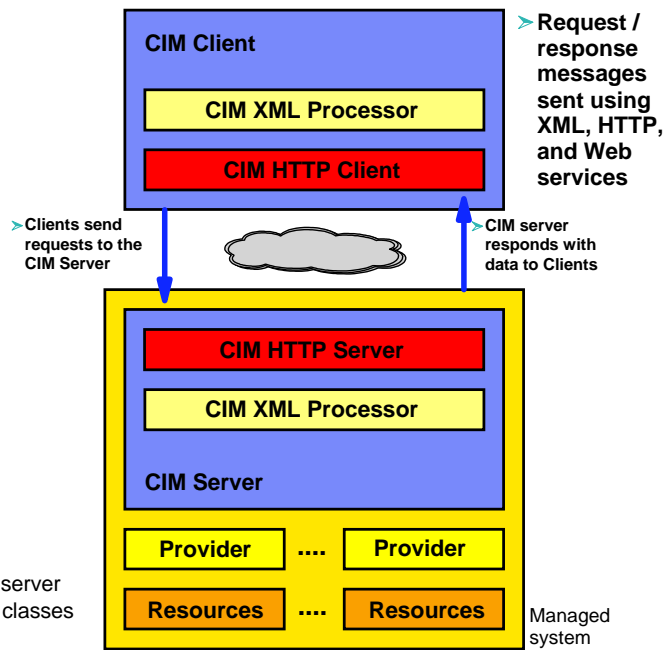
- **CIM servers are referred to as CIM Object Managers**

- **Similarities between CIM and SNMP**

- ▶ Platform extensions similar to enterprise-specific MIB modules
- ▶ CIM server similar to SNMP agent
- ▶ CIM providers similar to SNMP subagents
- ▶ Architected provider functions similar to SNMP requests. For example, CIM "`<class>GetInstance`" is similar to SNMP "get"

- **Providers**

- ▶ Accept requests for data from server
- ▶ Gather data in support of CIM classes
- ▶ Returns response with data



z/OS CIM functions

- **IBM selects the OpenPegasus CIM server from the Open Group Consortium**

- ▶ Implements the CIM/WBEM standards
- ▶ Ported to z/OS as a base element of V1R7
- ▶ Called the z/OS CIM server

- **z/OS V1R7 CIM operating system (OS) class support:**

- ▶ OS management profile instrumentation
 - ComputerSystem
 - OperatingSystem
 - Process (AS+USS)
 - Processor
 - FileSystem (USS)
 - Network
- ▶ OS Monitoring profile instrumentation
 - Performance metrics based on RMF data

- **z/OS CIM server**

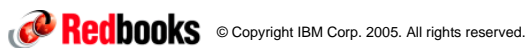
- ▶ For more information about the z/OS CIM Server, see *z/OS Common Information Model User's Guide*

- **IBM eServer CIM support**

- ▶ For information regarding the implementation of CIM management data across the eServer platforms, see *IBM eServer Common Information Model*

z/OS V1R7 CS CIM support

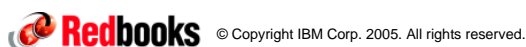
- **CS in z/OS V1R7 creates the z/OS CS CIM classes and provides the CIM network providers**
- **Support CIM schema Version 2.8**
- **Created z/OS CS platform schema extensions to the CIM classes:**
 - ▶ IBMzOS_EthernetPort - Subclass of CIM_EthernetPort
 - Supports all the IPv4 Ethernet interfaces configured to the TCP/IP stacks on the MVS image
 - Added TcpipProcName as platform-specific property
 - ▶ IBMzOS_IPProtocolEndpoint - Subclass of CIM_IPProtocolEndpoint
 - Supports all the IPv4 addresses configured to the TCP/IP stacks on the MVS image
 - Added TcpipProcName as platform-specific property
 - ▶ IBMzOS_CSNetworkPort - Subclass of CIM_SystemDevice
 - Supports the association between a computer system (an MVS image) and the network ports (network interfaces) configured to the computer system
 - z/OS CS provides data only for associations between an MVS image and the IPv4 Ethernet interfaces configured to the TCP/IP stacks on the image.
 - ▶ IBMzOS_NetworkPortImplementsIPEndpoint - Subclass of CIM_PortImplementsEndpoint
 - Supports the association between a network port (network interface) and the IP addresses configured on the interface.
 - z/OS CS provides data only for associations between an IPv4 Ethernet interface and its IP addresses.



ibm.com/redbooks

z/OS V1R7 CS CIM support (continued)

- **One CIM provider created per z/OS CIM class**
- **Provides data for IPv4 Ethernet interfaces and IP addresses**
- **Access to TCP/IP stack data controlled by security resource**
 - ▶ Resource is required if user ID associated with the client of the z/OS CIM server is not defined as a z/OS UNIX superuser.
 - ▶ Resource name is: EZB.CIMPROV.sysname.tcpname
 - ▶ Resource defined in the SERVAUTH class
 - ▶ Access is granted if the user ID associated with the client of the z/OS CIM Server is permitted for read access to the resource.
- **Providers installed in the /usr/lpp/tcpip/lib HFS directory**
- **z/OS CS CIM class definition and provider registration files installed in the /usr/lpp/tcpip/mof HFS directory**
 - ▶ Already integrated into z/OS CIM Server
 - ▶ Shipped due to service considerations
 - ▶ The class definitions can be reviewed to determine platform-specific properties
 - ▶ Platform specific properties also documented in the *IP Configuration Guide*



ibm.com/redbooks

How to enable CIM network support

- **No configuration necessary to activate the z/OS CS CIM provider support**
 - ▶ Providers automatically loaded by CIM server on first request for class data.
- **The z/OS CIM server must be configured and activated in order for the data supported by the z/OS CS CIM providers to be available to clients.**
- **Security resource must be defined for clients whose user IDs are not defined as z/OS UNIX superusers.**
 - ▶ See the *IP Configuration Guide* for more information about defining this resource.
- **z/OS CS CIM data class definitions and provider registration information (in MOF syntax) installed in new HFS directory /usr/lpp/tcpip/mof.**

Version-neutral UDP MIB

- **Enhanced the TCP/IP subagent to support the version-neutral UDP management data in the IETF internet draft version of the UDP-MIB.**
 - ▶ Data defined in UDP-MIB from draft-ietf-ipv6-rfc2013-update-03.txt (4/2004)
- **Added the following version-neutral UDP management data to the TCP/IP Enterprise-specific MIB module:**
 - ▶ `ibmTcpipMvsUdpEndpointTable` - provides counters and additional data
 - ▶ `ibmTcpipMvsUdpMcastTable` - provides data regarding UDP sockets that are receiving multicast data
- **Enhanced the following Netstat reports to display the remote IP address and port values for connected UDP sockets:**
 - ▶ ALL/-A
 - ▶ ALLCONN/-a
 - ▶ BYTEINFO/-b
 - ▶ CONN/-c
 - ▶ SOCKETS/-s

UDP management data enhancements

➤ Standard UDP-MIB data supported:

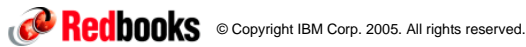
- ▶ udpHCInDatagrams/udpHCOutDatagrams
 - 64-bit UDP transport layer counters
 - Complement existing 32-bit counters
- ▶ udpEndpointTable
 - Provides local/remote IP address and port information for all UDP endpoints

➤ TCP/IP Enterprise-specific UDP data supported:

- ▶ ibmTcpiMvsUdpEndpointTable - augments the entries in the udpEndpointTable
 - 32-bit and 64-bit datagram and byte counters
 - Connection ID and resource name
 - Last activity value
 - Socket options
 - Information regarding UDP sockets that are sending multicast data
- ▶ ibmTcpiMvsUdpMcastTable - provides information regarding UDP sockets that are receiving multicast data

➤ IETF UDP-MIB internet draft shipped with product

- ▶ Because IETF internet drafts expire in six months, the version of the IETF UDP-MIB internet draft supported by V1R7 is shipped with the product and installed in the HFS in the /usr/lpp/tcpip/samples directory as file udpmib mi2.



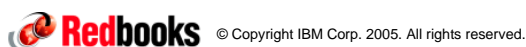
ibm.com/redbooks

Netstat enhancement in support of new UDP MIB

➤ Netstat reports display the remote IP address and port values for connected UDP sockets:

- ▶ Reports enhanced:
 - ALL/-A
 - ALLCONN/-a
 - BYTEINFO/-b
 - CONN/-c
 - SOCKETS/-s
- ▶ Previous releases' Netstat CONN/-c display of UDP Remote IP address and port for connected sockets:
 - Foreign Socket: *.*
- ▶ V1R7 Netstat CONN/-c display of UDP Remote IP address and port for connected sockets:

```
USER18 0000003D UDP
Local Socket:  ::ffff:0.0.0.0..1555      (AF_INET6 socket connected to IPV4)
Foreign Socket:  ::ffff:9.42.105.99..1555
USER18 00000039 UDP
Local Socket:   9.42.103.27..4444      (AF_INET socket)
Foreign Socket:  9.42.105.99..4444
USER18 0000003C UDP
Local Socket:   1::1..1444             (AF_INET6 socket)
Foreign Socket:  3::3..1444
```



ibm.com/redbooks

Things to think about when using new UDP MIB

➤ **Deprecated SNMP UDP management data**

- ▶ The status of 'deprecated' for MIB objects means that the objects are still supported but they will either become obsolete in the future, or they have been replaced by better objects.
 - If the objects were deprecated because they have been replaced, then management applications should plan on migrating their support to the replacement.
- ▶ Network management applications may not support deprecated management data.
- ▶ Deprecated standard UDP management data from the new UDP-MIB
 - New UDP-MIB from IETF internet draft deprecates the SNMP table, udpTable.
- ▶ Deprecated UDP management data from the TCP/IP Enterprise-specific MIB
 - ibmTcipMvsUdpTable
 - ibmTcipMvsUdpEndpMcastTable

➤ **Netstat display of remote IP address and port for connected UDP sockets**

- ▶ Automated programs that process Netstat report output may have to be updated

Netstat changes in z/OS V1R7

➤ **Starting in z/OS V1R7, all MVS console Netstat reports include "END OF THE REPORT" to clarify the end of the display report.**

➤ **Netstat ALL report changes**

- ▶ Displays an additional byte in the bit map of socket options for a UDP socket entry.
- ▶ Displays, for listening port, if the server has been quiesced for DVIPA Sysplex Distributor workload balancing.
- ▶ Displays if port sharing is being used by a listening port and if so, displays the type of port sharing (BASE or WLM).
- ▶ Displays, for a listening port, the number of active connections. Displays the server's accept efficiency fraction (SEF)
- ▶ Displays the remote IP address and port values for connected UDP sockets.

➤ **Netstat ALLCONN and CONN report changes**

- ▶ Filters the connection list display to exclude or include connections using AT-TLS Policy. For connections using AT-TLS Policy, the report can also be filtered to display only connections that are:
 - Using a current rule with current actions
 - Within one group
 - Using a stale rule or at least one stale action.
- ▶ In addition, the report was changed as follows:
 - Displays the remote IP address and port values for connected UDP sockets.

➤ **Netstat BYTEINFO report changes**

- ▶ Report changed to display the remote IP address and port values for connected UDP sockets.

Netstat changes in z/OS V1R7 (*continued*)

> Netstat CONFIG report changes

- ▶ Displays the setting of IPSECURITY on IPCONFIG and the SECCLASS setting on DYNAMICXCF on IPCONFIG.
- ▶ Displays the length of the routing prefix on the dynamic XCF IPv6 address when specified on the IPCONFIG6 DYNAMICXCF statement.
- ▶ Displays new information (AUTOREJOIN) in the Sysplex Monitor section.

> Netstat DEVLINK report changes

- ▶ Displays the appropriate SPEED for an OSA-Express 10 gigabit Ethernet adapter.
- ▶ Displays the SECCLASS values for the displayed links.
- ▶ Displays information about IPv6 HiperSockets interfaces.
- ▶ Displays whether an IPAQENET link is enabled for TCP segmentation offload.

> Netstat ND report changes

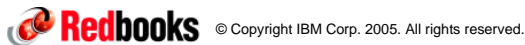
- ▶ Report is enhanced to display IPv6 addresses on HiperSockets internal LANs to which this stack has a route.

> Netstat PORTLIST report changes

- ▶ Changed to indicate if a new type of port sharing is being used (WLM).

> Netstat SOCKET report changes

- ▶ Report changed to display the remote IP address and port values for connected UDP sockets.



ibm.com/redbooks

Netstat changes in z/OS V1R7 (*continued*)

> Netstat TELNET report changes

- ▶ Report in short format is changed to display the BytesIn and BytesOut fields in two forms.

> Netstat TTLS new report

- ▶ The GROUP [,DETAIL] option lists information for AT-TLS groups.
- ▶ The CONN=connid [,DETAIL] option lists AT-TLS information for the specified connection.

> Netstat VCRT report changes

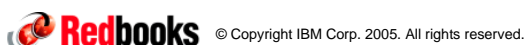
- ▶ In V1R7, the DETAIL report is changed to display additional route related information.

> Netstat VDPT report changes

- ▶ Indicates if a new distribution method ServerWLM is being used.
- ▶ Displays the Target Server Responsiveness (TSR), and, in the DETAIL display, the component fractions for the Target Connectivity Success Rate (TCSR), the Connection Establishment Rate (CER), and the Server's accept Efficiency Fraction (SEF).

> Netstat VIPADCFG report changes

- ▶ If the stack has left the sysplex group, new messages (EZZ2502I and EZZ2503I) will precede the report.
- ▶ Message EZZ2505I precedes the report to indicate if the VIPADYNAMIC configuration information cannot be displayed at this time.
- ▶ The VIPA Distribute information will contain data for backup, as well as active, dynamic VIPAs.
- ▶ Displays configured VIPAROUTE information.
- ▶ The IPAddr filter is added to provide the report for a specific IP address.
- ▶ A new subheading is added to the end of the report for deactivated dynamic VIPA information.



ibm.com/redbooks

Netstat changes in z/OS V1R7 *(continued)*

> Netstat VIPADYN report changes

- ▶ Displays the status of VIPAROUTE information.

This page intentionally left blank

ibm.com



e-business



Communications Server for z/OS V1R7 - Technical Update IPv6



Redbooks

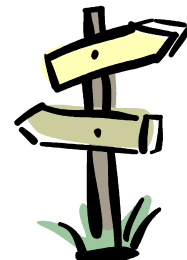
International Technical Support Organization

© Copyright IBM Corp. 2005. All rights reserved.

IPv6 agenda



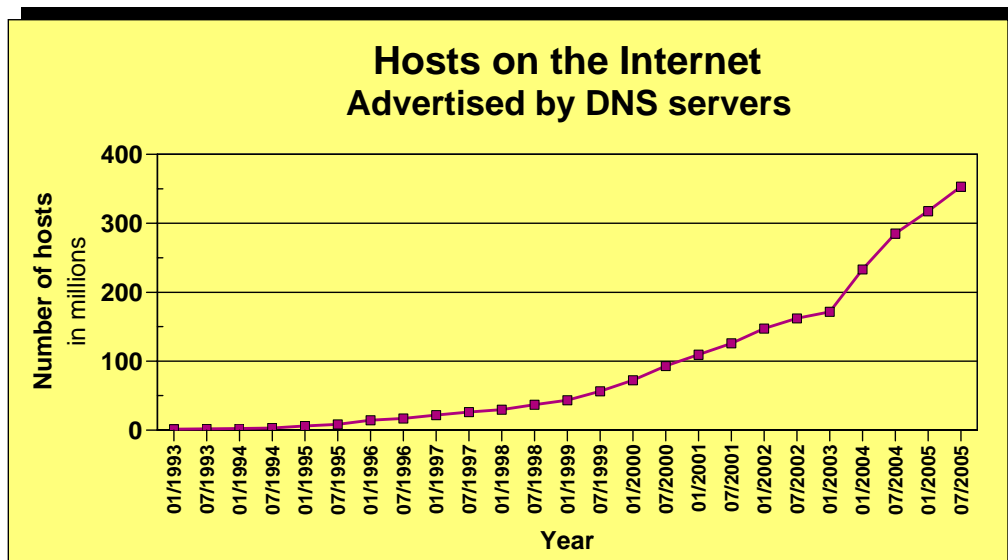
- > Introduction
- > Advanced sockets API for IPv6 update
- > Maintain 2 IPv6 routers in default list



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Visible IPv4 hosts on the Internet through the last 12 years



- What is the upper practical limit (the ultimate pain threshold) for number of assigned IPv4 addresses? Some predictions say 250,000,000 (250 million), others go up to 1,000,000,000 (one billion or one milliard).
- Source: <http://www.isc.org/index.pl?/ops/ds>

Both z/OS V1R5 and V1R6 have been certified with the IPv6 Ready logo

| Item | Content |
|--------------------------------|---|
| Logo ID | 01-000156 |
| Vendor Name | IBM Corporation |
| Country Name | US |
| Product Name (Original) | z/OS |
| Product version (Original) | V1R5 |
| Product Description (Original) | Highly secure scalable high-performance enterprise operating system |
| Product Name (Update) | |
| Product version (Update) | |
| Product Description (Update) | |
| Product Category | Host |
| Applied date | 20031217 |
| Application ID | US-20031217-000136 |
| Current Status | Approved |
| Certificated Date | 20040326 |

CS z/OS V1R7 is in the process of being certified too.

The Journey to IPv6 for z/OS Communications Server

> The first phase (z/OS V1R4)

- ▶ Stack support for IPv6 base functions - (APIs, Protocol layers)
- ▶ Resolver
- ▶ High speed attach (OSA Express QDIO))
- ▶ Service tools (Trace, Dump, etc.)
- ▶ Configuration and Netstat, ping, traceroute, SMF
- ▶ Static Routing
- ▶ FTP, otelnetd, unix rexec, unix rshd/rexecd

> The second phase (z/OS V1R5)

- ▶ Network Management
 - Applications and DPI
 - Version-neutral TCP/IP Standard MIBs
 - Additional SMF records
- ▶ Applications/Clients/APIs
 - Tn3270 server, CICS Sockets, sendmail, ntp, dcas, rxserve, rsh client
- ▶ Enterprise Extender
- ▶ Point to Point - type DLCs
- ▶ Dynamic Routing Protocol w/ OMPROUTE (only RIPng)

> The third phase (z/OS V1R6)

- ▶ Sysplex Exploitation (Dynamic VIPA, Sysplex Distributor functions)
- ▶ Dynamic Routing Protocol w/ OMPROUTE (OSPFv3)
- ▶ Additional Network Management MIBs

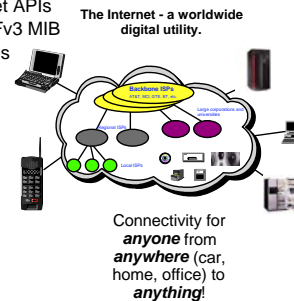
> The fourth phase (z/OS V1R7)

- ▶ SNMP UDP standard MIB (RFC2013) and IBM MVS TCP/IP Enterprise-specific MIB for UDP
- ▶ Advanced Socket API support - RFC3542
- ▶ IPv6 Two Default Routers - required for IPv6 compliance
- ▶ HiperSockets DLC (z9-109)

> After z/OS V1R7

- ▶ Integrated IPsec
- ▶ Complete Advanced Socket APIs
- ▶ Extended Stats MIB, OSPFv3 MIB
- ▶ Intrusion Detection Services
- ▶ IPv6 mobility support

Objective is to have IPv6 production ready on the platform when you need it!



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Advanced sockets API for IPv6

> Lets applications modify and receive information about packets

- ▶ Control and modify outbound packet information, such as
 - First hop address and routing headers
 - Hop options and destination options
 - Traffic class
 - Packet fragmentation
 - MTU discovery
- ▶ Receive inbound packet information such as
 - Arriving interface
 - Destination IP address
 - Hop limit
 - Source routing
 - IPv6 options (routing headers, destination options, etc.) set by the sender

> Defined first in RFC 2292, then in RFC 3542

- ▶ Defined for use by 'advanced' IPv6 applications
 - For example, ping, traceroute, and routing daemons
 - Geared more towards applications using RAW sockets
- ▶ Separate from the Basic IPv6 Socket APIs in RFC 3493



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Existing IPv6 advanced socket API support as of z/OS V1R4

NOTES

- > z/OS V1R4 introduced partial support of the IPv6 advanced socket API
- > The V1R4 implementation was based on draft RFC 2292.
- > The following options were supported by the V1R4 implementation

| Level | Option Name | Get | Set | Data type | Data path | Transports supported |
|----------------|-------------------|-----|-----|--------------|-----------|----------------------|
| IPPROTO_IPV6 | IPV6_USE_MIN_MTU | Y | Y | int | Outbound | TCP, UDP, RAW |
| | IPV6_PKTINFO | Y | Y | in6_pktinfo | Outbound | UDP, RAW |
| | IPV6_RECVHOPLIMIT | Y | Y | int | Inbound | TCP, UDP, RAW |
| | IPV6_RECVPKTINFO | Y | Y | in6_pktinfo | Inbound | UDP, RAW |
| IPPROTO_ICMPV6 | IPV6_CHECKSUM | Y | Y | int | Outbound | RAW |
| | ICMP6_FILTER | Y | Y | icmp6_filter | Inbound | RAW |

- > **Outbound options could be set**
 - ▶ 'Sticky' using setsockopt - all packets on the socket used the option
 - ▶ Or 'per packet' using sendmsg ancillary data - affects only that packet
- > **Inbound packet information is received on recvmsg() as ancillary data.**
- > **The options could be used on z/OS UNIX callable services and using LE C/C++ APIs**



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

New advanced socket API options for IPv6

- > **RFC 3542 is now implemented for UDP and RAW sockets and partially supported for TCP sockets.**
 - ▶ TCP only supports the very basic set of the IPv6 Advanced socket APIs
 - ▶ The more advanced APIs are geared towards UDP and RAW sockets
- > **Means to provide RACF authentication to allow/disallow users and applications from using the APIs are provided.**
 - ▶ The RACF authentication is granular enough to specify access restrictions for each option of the API.
- > **The options are supported only for z/OS UNIX callable services and LE C/C++ APIs**

| Level | Option Name | Data Path | Transports Supported |
|--------------|-------------------|-----------|----------------------|
| IPPROTO_IPV6 | IPV6_HOPOPTS | Outbound | UDP, RAW |
| | IPV6_RECVHOPOPTS | Inbound | UDP, RAW |
| | IPV6_RTHDR | Outbound | UDP, RAW |
| | IPV6_RECVRTHDR | Inbound | UDP, RAW |
| | IPV6_RTHDRDSTOPTS | Outbound | UDP, RAW |
| | IPV6_DSTOPTS | Outbound | UDP, RAW |
| | IPV6_RECVDSTOPTS | Inbound | UDP, RAW |
| | IPV6_RECVTCLASS | Inbound | TCP, UDP, RAW |
| | IPV6_TCLASS | Outbound | TCP, UDP, RAW |
| | IPV6_NEXTHOP | Outbound | UDP, RAW |
| | IPV6_RECVPATHMTU | Outbound | UDP, RAW |
| | IPV6_PATHMTU | Outbound | UDP, RAW |
| | IPV6_DONTFRAG | Outbound | UDP, RAW |

These are the new advanced socket API options for IPv6 that are provided in z/OS V1R7.



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

RACF protection of the advanced IPv6 socket options

> Access to the socket options is allowed under only three conditions

- ▶ Application is APF authorized - or -
- ▶ User executing the application has super user authority - or -
- ▶ Option resource name defined and the application has at least READ access to the resource.

> The resource names are:

| API option | RACF Resource Name |
|-------------------|---|
| IPV6_NEXTHOP | EZB.SOCKETP.sysname.tcpname.IPV6_NEXTHOP |
| IPV6_TCLASS | EZB.SOCKETP.sysname.tcpname.IPV6_TCLASS |
| IPV6_RTHDR | EZB.SOCKETP.sysname.tcpname.IPV6_RTHDR |
| IPV6_HOPOPTS | EZB.SOCKETP.sysname.tcpname.IPV6_HOPOPTS |
| IPV6_DSTOPTS | EZB.SOCKETP.sysname.tcpname.IPV6_DSTOPTS |
| IPV6_RTHDRDSTOPTS | EZB.SOCKETP.sysname.tcpname.IPV6_RTHDRDSTOPTS |
| IPV6_HOPLIMIT | EZB.SOCKETP.sysname.tcpname.IPV6_HOPLIMIT |
| IPV6_PKTINFO | EZB.SOCKETP.sysname.tcpname.IPV6_PKTINFO |

> Migration concerns are only for a multilevel security (MLS) environment

- ▶ IPV6_PKTINFO changed authorization in an MLS environment
- ▶ To use the options in an MLS environment, the resource name must be defined and the application must have at least READ access to the resource.

Maintain 2 IPv6 routers in default list for IPv6-Ready logo compliance

> IPv6 standards require a minimum of 2 default routers

- ▶ Required for the IPv6-Ready logo certification

> In certain situations, z/OS CS does not meet this requirement

- ▶ If default routes are being removed from the stack routing table by OMPROUTE due to lost network connectivity, the number of default IPv6 routers may go to zero

> When the last default route is deleted from the routing table

- ▶ Add the default routers back to the routing table

> No new configuration options and no migration concerns

ibm.com



e-business



Communications Server for z/OS V1R7 - Technical Update Enterprise Extender and SNA



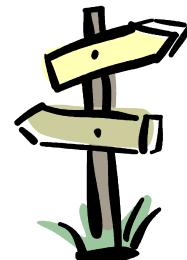
Redbooks

International Technical Support Organization

© Copyright IBM Corp. 2005. All rights reserved.

EE and SNA agenda

- > VTAM and Enterprise Extender display enhancements
- > New DISPLAY EEDIAG command
- > Model definition of VTAM cross domain resources
- > Subarea VTAM XCF support
- > LOGAPPL enhancements



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Removal of AnyNet - z/OS V1R7 last release to support AnyNet

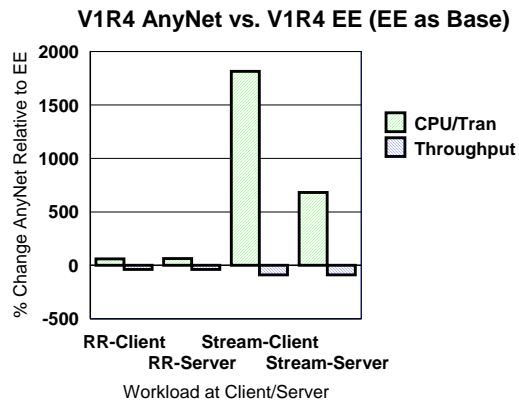
➤ Enterprise Extender, TN3270, and distributed Communications Server Remote API functions are the strategic protocols for SNA/IP integration

▸ AnyNet has not been enhanced in years

➤ EE is functionally superior, but also significantly outperforms AnyNet by all measures:

▸ AnyNet exhibits lower throughput and higher CPU utilization relative to EE:

- Interactive workloads
 - Throughput down 39%
 - CPU utilization up 63%
- Stream workloads
 - Throughput down 89%
 - CPU utilization up 682-1817%



➤ z/OS V1R7 is the last release of z/OS to include AnyNet as a component of Communications Server

Display negotiated RU sizes on the D NET,SESSIONS command

➤ D NET,SESSIONS,SID= sidvalue command enhanced

- Includes negotiated RUSIZES for this session
 - IST2064I PLU TO SLU RU SIZE = plu_to_slu_rusize SLU TO PLU RU SIZE = slu_to_plu_rusize

NOTES

```
D NET,SESSIONS,SID=EAABEEC3E5A79CCB

IST097I DISPLAY ACCEPTED
IST350I DISPLAY TYPE = SESSIONS
IST879I PLU/OLU REAL = NETA.APPL2      ALIAS = ***NA***
IST879I SLU/DLU REAL = NETA.APPL1      ALIAS = ***NA***
IST880I SETUP STATUS = ACTIV
IST933I LOGMODE=INTERACT, COS=*BLANK*
IST1635I PLU HSCB TYPE: FMCB LOCATED AT ADDRESS X'0155F5B8'
IST1635I SLU HSCB TYPE: FMCB LOCATED AT ADDRESS X'0155F720'
IST2064I PLU TO SLU RU SIZE = 65535    SLU TO PLU RU SIZE = 65535
IST1636I PACING STAGE(S) AND VALUES:
IST1637I PLU--STAGE 1--SLU
IST1638I STAGE1: PRIMARY TO SECONDARY DIRECTION - ADAPTIVE
IST1639I PRIMARY SEND: CURRENT = 7      NEXT = 8
IST1640I SECONDARY RECEIVE = 32767
IST1641I STAGE1: SECONDARY TO PRIMARY DIRECTION - ADAPTIVE
IST1642I SECONDARY SEND: CURRENT = 7    NEXT = 8
IST1643I PRIMARY RECEIVE = 32767
```

Information on queued work elements returned by the D NET,,HPRDIAG=YES command

NOTES

➤ D NET,ID=rtppuname,HPRDIAG=YES command enhanced

- ▶ Includes the current and highwater mark for the number of work elements queued for outbound transmission
- ▶ Includes a time stamp to note when the highwater mark was most recently reached.
 - IST2085I NUMBER OF NLPS ON OUTBOUND WORK QUEUE = num_nlps
 - IST2086I MAXIMUM NUMBER OF NLPS ON OUTBOUND WORK QUEUE = max_num_nlps
 - IST2087I OUTBOUND WORK QUEUE MAX REACHED ON date AT time

Example of HPRDIAG command output

NOTES

```
D NET,ID=CNR00001,HPRDIAG=YES

IST097I  DISPLAY ACCEPTED
IST075I  NAME = CNR00001, TYPE = PU_T2.1
IST1392I DISCNTIM = 00010 DEFINED AT PU FOR DISCONNECT
.
IST924I  -----
IST1973I OUTBOUND TRANSMISSION INFORMATION:
IST1974I NUMBER OF NLPS SENT = 12 (0K)
IST1975I TOTAL BYTES SENT = 1823 (1K)
IST1849I LARGEST NLP SENT = 161 BYTES
IST1980I SEQUENCE NUMBER = 372 (X'00000174')
IST1842I NUMBER OF NLPS RETRANSMITTED = 0
IST1976I BYTES RETRANSMITTED = 0 (0K)
IST1478I NUMBER OF UNACKNOWLEDGED BUFFERS = 0
IST1958I NUMBER OF ORPHANED BUFFERS = 0
IST1843I NUMBER OF NLPS ON WAITING-TO-SEND QUEUE = 0
IST1847I NUMBER OF NLPS ON WAITING-FOR-ACKNOWLEDGEMENT QUEUE = 0
IST1977I MAXIMUM NUMBER OF NLPS ON WAITING-FOR-ACK QUEUE = 6
IST1978I WAITING-FOR-ACK QUEUE MAX REACHED ON 01/08/04 AT 13:44:03
IST2085I NUMBER OF NLPS ON OUTBOUND WORK QUEUE = 0
IST2086I MAXIMUM NUMBER OF NLPS ON OUTBOUND WORK QUEUE = 5
IST2087I OUTBOUND WORK QUEUE MAX REACHED ON 01/08/04 AT 14:03:24
IST1511I MAXIMUM NETWORK LAYER PACKET SIZE = 16410 BYTES
IST924I  -----
.
```



D NET, RTPS enhancements

NOTES

> D NET RTPS command enhanced to add

- ▶ LIST=DETAIL option
 - Default
 - Displays essentially the same information as current display
 - New message IST2084I replaces message IST1454I in the output.
 - Shows number of displayed RTP pipes as well as the number of matching RTP pipes.
 - Shows the total number of RTP pipes that matched the display criteria when the output was truncated due to MAX limitations.
 - IST1454I count RTP(S) DISPLAYED
 - IST2084I count OF total MATCHING RTP PIPES DISPLAYED
- ▶ LIST=SUMMARY option
 - Provides a summary of the matching RTP pipes.

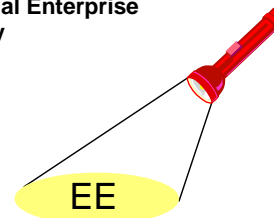
```
D NET,RTPS,FIRSTCP=NETA.SSCP2A,LIST=SUMMARY
IST097I DISPLAY ACCEPTED
IST350I DISPLAY TYPE = RTPS
IST2075I DISPLAY RTPS SUMMARY INFORMATION
IST2076I TOTAL MATCHING PIPES = 5
IST2077I CPSVCMG PIPES = 2
IST2078I RSETUP PIPES = 1
IST2079I LU-LU PIPES = 2
IST2080I PATH SWITCHING PIPES = 0
IST2081I CONGESTED PIPES = 0
IST2082I STALLED PIPES = 0
IST2083I SESSIONS = 4
IST314I END
```

Display EEDIAG command

> V1R7 provides a new operator command to provide additional Enterprise Extender problem determination and management capability

> Two basic forms:

- ▶ Display EE connections that have a packet retransmission rate that exceeds a specified threshold
- ▶ Display EE connections that have exceeded a specified number of LDLC "TEST" frame retries (SRQRETRY)



Display EE and display EEDIAG commands

➤ DISPLAY EE command available in Communications Server V1R6

- ▶ The DISPLAY EE command available in Communications Server V1R6 was provided to better manage Enterprise Extender networks. This was the first VTAM command to provide some elementary information about Enterprise Extender. The outputs from the DISPLAY EE commands are very useful, but it is cumbersome to manually perform calculations necessary to see Enterprise Extender retransmission and LDLC retry information.
- ▶ Useful, but....
- ▶ Need diagnostic data
 - Retransmission rates
 - LDLC retry information

➤ A new VTAM display command, DISPLAY EEDIAG, has been developed to extend the Enterprise Extender network management support.

- ▶ Various formats of the new display give the operator the ability to obtain:
 - Display Enterprise Extender connections that meet or exceed a specified retransmission threshold.
 - Display Enterprise Extender connections that meet or exceed a specified SRQRETRY threshold.

D NET,EEDIAG,REXMIT example

- Find all Enterprise Extender connections associated with a particular VIPA with retransmission rates that meet or exceed 5% and display the output in detail format

```
D NET,EEDIAG,REXMIT=5,IP=9::67:1:1,LIST=DETAIL
IST097I DISPLAY ACCEPTED
IST350I DISPLAY TYPE = EEDIAG
IST2065I ENTERPRISE EXTENDER CONNECTION REXMIT INFORMATION
IST2067I EEDIAG DISPLAY ISSUED ON 08/27/04 AT 13:31:05
IST924I -----
IST1680I LOCAL IP ADDRESS 9::67:1:1
IST1910I LOCAL HOSTNAME IP.SSCP1AV6
IST1680I REMOTE IP ADDRESS 9::67:1:6
IST1909I REMOTE HOSTNAME IP.SSCP2AV8
.
.
IST924I -----
IST2032I PORT PRIORITY = HIGH
IST2036I NLPS SENT = 134 ( 000K )
IST2038I NLPS RETRANSMITTED = 67 ( 000K )
IST2068I NLP RETRANSMIT RATE = 50%
IST924I -----
.
.
IST924I -----
IST2035I TOTALS FOR ALL PORT PRIORITIES
IST2036I NLPS SENT = 1948 ( 001K )
IST2038I NLPS RETRANSMITTED = 67 ( 000K )
IST2068I NLP RETRANSMIT RATE = 3%
IST2069I REXMIT COUNTERS LAST CLEARED ON 08/27/04 AT 13:20:42
IST2042I 1 OF 1 EE CONNECTIONS DISPLAYED
IST314I END
```

D NET,EEDIAG,CLEAR and SRQRETRY examples

- Clear all diagnostic counters for all Enterprise Extender connections.

```
D NET,EEDIAG,CLEAR
IST097I DISPLAY ACCEPTED
IST350I DISPLAY TYPE = EEDIAG
IST2067I EEDIAG DISPLAY ISSUED ON 08/23/04 AT 22:05:22
IST2071I ALL DIAGNOSTIC COUNTERS CLEARED FOR 3 EE CONNECTIONS
IST314I END
```

- Find all Enterprise Extender connections that experienced LDLC retries of three or more attempts. Present the output in detailed format and clear the SRQRETRY counters for all connections.

```
D NET,EEDIAG,SRQRETRY=3,LIST=DETAIL,CLEAR=SRQRETRY
IST097I DISPLAY ACCEPTED
IST350I DISPLAY TYPE = EEDIAG
IST2066I ENTERPRISE EXTENDER CONNECTION SRQRETRY INFORMATION
IST2067I EEDIAG DISPLAY ISSUED ON 08/23/04 AT 20:00:01
IST924I -----
IST1680I LOCAL IP ADDRESS 9.67.1.1
IST1910I LOCAL HOSTNAME IP.SSCP1AV6
IST1909I REMOTE HOSTNAME IP.SSCP2AV7
IST2024I CONNECTED TO SWITCHED PU CNV00006
IST2074I SUCCESSFUL SRQRETRY ATTEMPT = 0 OCCURRENCES = 98
IST2074I SUCCESSFUL SRQRETRY ATTEMPT = 1 OCCURRENCES = 5
IST2074I SUCCESSFUL SRQRETRY ATTEMPT = 2 OCCURRENCES = 1
IST2074I SUCCESSFUL SRQRETRY ATTEMPT = 3 OCCURRENCES = 1
IST2070I SRQRETRY COUNTERS LAST CLEARED ON 08/23/04 AT 18:55:15
IST2073I SRQRETRY COUNTERS CLEARED FOR 3 EE CONNECTIONS
IST2042I 1 OF 1 EE CONNECTIONS DISPLAYED
IST314I END
```

Cross domain resource basics

- Cross-domain resources are represented by RDTEs called CDRSCs.
- CDRSCs may be predefined as part of a CDRSC major node.
- CDRSCs may be dynamically defined as needed.
 - ▶ Added to ISTCDRDY.
 - ▶ Dependent upon CDRDYN start option and CDRSC parameter on CDRM statement.
 - If CDRDYN=NO, then no dynamic CDRSCs will be created at this VTAM.
 - If CDRDYN=YES, then dynamic CDRSC creation depends on CDRM's CDRSC value.
 - CDRSC=REQ means CDRSCs are required to be predefined for resources owned by the SSCP represented by the CDRM.
 - CDRSC=OPT means CDRSCs may be dynamically created for resources owned by the SSCP represented by the CDRM.
- This gives customers three options:
 - ▶ Require predefinition of all CDRSCs.
 - ▶ Allow dynamic definition of CDRSCs for any requests.
 - ▶ Determine dynamic capabilities on a CDRM by CDRM basis.
- With customers connecting to more and more networks, they are increasingly faced with the choice between security concerns and systems management concerns:
 - ▶ If they do not allow dynamic CDRSCs to be created for the new network to which they are connecting, thousands of predefined CDRSCs may have to be created.
 - ▶ If they do allow dynamic CDRSCs to be created for the new network to which they are connecting, they forfeit some measure of control over access to their network or are required to deal with it in another way.

Model CDRSC

>VTAM will now allow model CDRSCs to be defined in a CDRSC major node.

- ▶ The customer can create model CDRSC definitions, similar to the way model APPL definitions are created.

>Using the appropriate model CDRSC definitions, VTAM will dynamically create CDRSCs as needed.

- ▶ CDRSCs created from model CDRSCs are called clone CDRSCs.
- ▶ Unaffected by CDRDYN start option and CDRSC keyword.

>Benefits

- ▶ Amount of definition reduced.
- ▶ Control provided by knowledge of the new network's naming conventions.
- ▶ More parameters can be specified for model CDRSCs than dynamic CDRSCs.
 - For example, DLOGMOD
- ▶ Control provided for when clone CDRSCs are deleted.

Example of a model CDRSC major node

>Model CDRSCs may be defined by specifying '?' or '*' within the name field of the CDRSC definition statement:

- ▶ '?' represents exactly one character and can be in any position in the name.
- ▶ '*' represents zero or more characters and can be in any position but the first one.
- ▶ The CDRSCs after the NETWORK statement are considered to be real CDRSCs.
- ▶ The CDRSCs before the NETWORK statement are considered to be alias CDRSCs.
- ▶ Make names unique enough to ensure that the right model definition is used.

>New DELETE keyword for model CDRSC definitions only.

- ▶ DELETE=YES - the clone CDRSC is deleted when the sessions are ended. (Default)
- ▶ DELETE=NO - the clone CDRSC is not deleted when the sessions are ended.

```
CDRSCMOD  VBUILD  TYPE=CDRSC
?APPL     CDRSC
APPL*     CDRSC
APPL?     CDRSC
          NETWORK NETID=NETA
APPL1*    CDRSC DELETE=NO
APPL1?    CDRSC
ABCD*     CDRSC DELETE=YES
EF?G*     CDRSC DELETE=NO
```

Cloned CDRSCs

NOTES

- **Clone CDRSCs may be created:**
 - During session setup.
 - During DSRLIST processing.
 - During MODIFY ALSLIST, ACTION=CREATE processing.
- **Clone CDRSCs are built based on the active model CDRSC that is a best match:**
 - Character by character comparison, left to right.
 - Actual character is better than '?', '?' is better than '*'.
 - NQNMODE considerations.
- **Name the model used to create the following clone CDRSCs:**
 - APPL1
 - APPL2
 - APPL1A
 - APPL2A
- **In general, operator commands that work against predefined CDRSCs work the same when issued against clone CDRSCs.**
- **Most operator commands issued against model CDRSCs do not affect the already created clone CDRSCs, but they will affect any future clone CDRSCs created from the model CDRSC.**
 - An exception: when SCOPE=ALL is specified on Modify TRACE and Modify NOTRACE against a model CDRSC, current clone CDRSCs are affected as well as future clone CDRSCs.
 - Note that the DELETE parameter of the model CDRSC is a characteristic of the model CDRSC, not the clone CDRSC.

Activation/Inactivation of model CDRSCs

NOTES

- **Model CDRSCs are activated when their major node is activated if ISTATUS=ACTIVE.**
- **If a model CDRSC is inactive, it can be activated via the VARY NET,ACT command.**
- **A model CDRSC can be inactivated via the VARY NET,INACT command.**
- **A model CDRSC must be active for it to be used to create clone CDRSCs.**
- **V NET,ACT,SCOPE=ALL can be used to activate a model CDRSC and all the clone CDRSCs that have been built from that model CDRSC.**
- **V NET,INACT,SCOPE=ALL can be used to inactivate a model CDRSC and all the clone CDRSCs that have been built from that model CDRSC.**
- **Clone CDRSCs can also be specified in the V NET,ACT and V NET,INACT commands.**
 - Specifying the DELETE operand on the VARY INACT command against a clone CDRSC will override the value of the DELETE parameter specified on the model CDRSC definition.

Display of CDRSC major node

NOTES

```
D NET,ID=CDRSCMOD,E
IST097I DISPLAY ACCEPTED
IST075I NAME = CDRSCMOD, TYPE = CDRSC SEGMENT
IST486I STATUS= ACTIV, DESIRED STATE= ACTIV
IST478I CDRSCS:
IST1276I ?APPL          ACTIV      CDRM = ***NA***
IST1276I APPL*          ACTIV      CDRM = ***NA***
IST1276I APPL?          ACTIV      CDRM = ***NA***
IST1276I NETA.APPL2     ACT/S     CDRM = SSCP2A
IST1276I NETA.APPL1*    ACTIV      CDRM = ***NA***
IST1276I NETA.APPL1     ACT/S     CDRM = SSCP2A
IST1276I NETA.APPL1?    ACTIV      CDRM = ***NA***
IST1276I NETA.ABCD*     ACTIV      CDRM = ***NA***
IST1276I NETA.EF?G*     ACTIV      CDRM = ***NA***
IST1500I STATE TRACE = OFF
IST314I END
```

- Note that the clone CDRSCs are displayed following the model CDRSC with which they were created.
- Note that the model CDRSCs that were defined after the NETWORK statement have the netid in the IST1276I message.
- Note that both clone CDRSCs are displayed with their netid in IST1276I.

Display of model CDRSC

NOTES

```
D NET,ID=APPL?,E
IST097I DISPLAY ACCEPTED
IST075I NAME = APPL?, TYPE = MODEL CDRSC
IST486I STATUS= ACTIV, DESIRED STATE= ACTIV
IST599I REAL NAME = ***NA***
IST1447I REGISTRATION TYPE = NO
IST977I MDLTAB=***NA*** ASLTAB=***NA***
IST1333I ADJLIST = ***NA***
IST861I MODETAB=***NA*** USSTAB=***NA*** LOGTAB=***NA***
IST934I DLOGMOD=***NA*** USS LANGTAB=***NA***
IST597I CAPABILITY-PLU ENABLED ,SLU ENABLED ,SESSION LIMIT NONE
IST231I CDRSC MAJOR NODE = CDRSCMOD
IST2095I MODEL CDRSC DELETE = YES
IST479I CDRM NAME = ***NA***, VERIFY OWNER = NO
IST1131I DEVICE = CDRSC
IST654I I/O TRACE = OFF, BUFFER TRACE = OFF
IST1500I STATE TRACE = OFF
IST228I ENCRYPTION = NONE , TYPE = DES
IST1563I CKEYNAME = APPL? CKEY = PRIMARY CERTIFY = NO
IST1552I MAC = NONE MACTYPE = NONE
IST2088I CDRSCS DEFINED USING THIS MODEL:
IST1276I NETA.APPL2     ACT/S     CDRM = SSCP2A
IST314I END
```

Dynamic XCF support prior to z/OS V1R7

➤ Current VTAM support for dynamic XCF connectivity:

- ▶ Allows dynamically
 - Joining ISTXCF Sysplex group
 - Building APPN and TRLE definitions
 - Connecting to the other VTAM APPN nodes in the Sysplex
- ▶ Based on XCFINIT start option (YES or NO)
- ▶ For VTAM APPN nodes only

➤ Current TCP/IP support for dynamic XCF connectivity:

- ▶ Allows dynamically
 - Using VTAM XCF connections
 - Building TCP/IP devices and interfaces
 - Connecting to the other TCP/IP stacks in the Splex
- ▶ Based on IPCONFIG and IPCONFIG6 DYNAMICXCF parameters
- ▶ Requires running on a VTAM APPN node
- ▶ Requires VTAM APPN XCF connectivity



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

z/OS V1R7 new dynamic XCF support

➤ With this release, z/OS Communications Server will allow:

- ▶ TCP/IP connectivity through XCF on APPN nodes without having to first establish APPN connections
- ▶ TCP/IP connectivity through XCF from pure subarea nodes
- ▶ This allows users to utilize the full range of TCP/IP Sysplex functions without having to redefine the SNA network to use APPN communications

➤ For APPN nodes:

- ▶ A new value for the XCFINIT start parameter, DEFINE, is now allowed
- ▶ If XCFINIT = DEFINE:
 - VTAM will join the ISTXCF Sysplex group
 - VTAM will build the definitions necessary for XCF connectivity between this node and other nodes in the Sysplex
 - The XCF APPN PU and XCF TRLE definitions will be built
 - VTAM will not activate those connections
 - TCP/IP connectivity is allowed, using either static or dynamic XCF definitions
- ▶ XCFINIT=YES will remain the default for APPN nodes



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

TRLE names for dynamic XCF

NOTES

- **TRLE definitions created on a pure subarea node will use the default naming convention:**
 - ISTTxyy, where xx is this node's Sysclone value, and yy is the partner's Sysclone value
- **The first 4 characters (ISTT) can be changed using the TRLE parameter on a model XCF PU definition within the model major node deck**
 - Specify TRLE=cccc* on the model XCF PU definition, where:
 - cccc must be from 1-4 characters.
 - The first character can be alphabetical (A-Z) or the national characters @, #, or \$.
 - Any characters after the first character can be alphabetical (A-Z), numerical (0-9), or the national characters @, #, or \$.
 - The generated TRLE name will be ccccxxyy, where xx is this node's Sysclone value, and yy is the partner's Sysclone value
- **An APPN PU will not be created from the model PU definition**

Things to think about when using the new XCFINIT support

- **Since XCFINIT=DEFINE is the default for pure subarea nodes, any pure subarea nodes in the Sysplex will automatically join the ISTXCF group when they are upgraded to this release. Other nodes in the Sysplex will become aware of these nodes and may try to establish SNA XCF connectivity with them (see next bullet). If this behavior is not desired, specify XCFINIT=NO on these nodes.**
- **If a VTAM node with XCFINIT=DEFINE specified (or defaulted) is started in a Sysplex, and there are down-level VTAM APPN nodes in the Sysplex, the down-level nodes will attempt to activate an XCF PU to establish connectivity with the node specifying XCFINIT=DEFINE. Connectivity will not be established. The PU on the down-level node will remain in Pending Request Contact (PREQC) state.**
 - If the node specifying XCFINIT=DEFINE is an APPN node, connectivity can be established by issuing a Vary ACT for the XCF PU on that node.
 - If connectivity is not desired with the down-level APPN node, or the XCFINIT=DEFINE node is a subarea node, issue a Vary INACT for the XCF PU on the down-level node.
 - See the *SNA Network Implementation Guide* section on Dynamic Definition of VTAM-to-VTAM connections for more information.

New VARY NET,AUTOLOG command

> Automatic logons

- Coding LOGAPPL on an LU or by issuing a VARY LOGON (or VARY ACT,LOGON) to an LU enables the LU to do an automatic logon to a specified application when the LU becomes session capable.

> Pending autolog request

- Should the automatic session attempt fail, a pending autolog request is established in the LU host. The reallocation of a pending autolog request is attempted when a notification of the application's availability is received or when the conditions defined on the AUTOTI and AUTORTRY VTAM start option are met.

> The reallocation of pending autolog requests can now be driven with the new VARY AUTOLOG operator command.

> The new VARY AUTOLOG command will allow customers to immediately drive pending autologon requests into session if the controlling application is located and is session capable.

> The VARY AUTOLOG command has an option of acting upon a selected PLU name or all PLUs for which there is a pending autolog request.

```
>>_VARY NET,AUTOLOG _____,ID=*_<
      |_____|
      |_____,ID=controlling_appl_____|<
```



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

New pending autolog request

> Effective in z/OS V1R7 a pending autolog request is established when an LU normally terminates its session with the controlling (automatic logon) application.

> V1R7 includes message IST2100I in the DISPLAY AUTOLOGON command to show the pending autolog requests that were created as a result of a normal termination of an LU with its controlling application.

```
D net,autolog,scope=all

IST350I DISPLAY TYPE = AUTOLOG
IST1990I PENDING AUTOLOGON REQUESTS FOR:
IST1992I NETA.APPL1 - WAITING FOR AUTOTI TIMER
IST1997I NETA.LU1
IST2100I NETA.APPL1 - NORMALLY LOGGED OFF LUS
IST1997I NET1.LU4
IST314I END
```

Note: The pending autolog request for NET.LU4 was created when an LU-to-LU session between NETA.APPL1 and NETA.LU4 terminated normally.

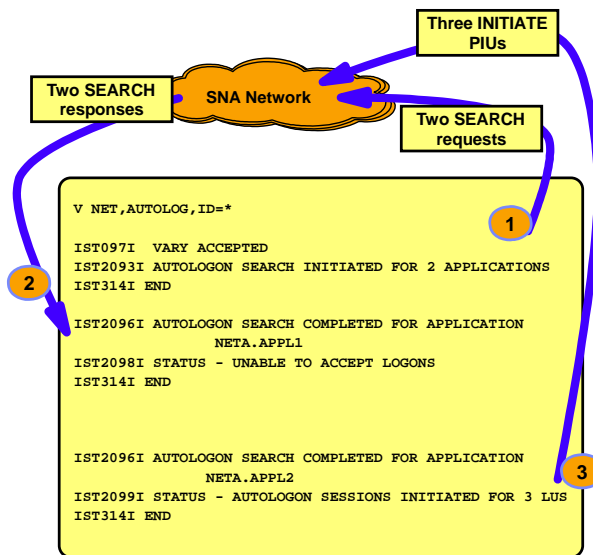


© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

V NET,AUTOLOG,ID=* command processing example

- Upon processing the VARY AUTOLOG operator command, external search PIUs are sent into the network to locate applications NETA.APPL1 and NETA.APPL2
- The search response indicates whether the application could be located. If the application is located the status of the application is included.
 - Each search response generates an IST2096I message group.
 - **NOTE:** In this example both NETA.APPL1 and NETA.APPL2 were located. The status of NETA.APPL1 shows that it is not session capable. The status of NETA.APPL2 shows that it is session capable.
- Since NETA.APPL2 is capable of LU-LU sessions, the SLU host generates an initiation for each LU with a pending autologon request for application NETA.APPL2.



Normal termination pending autologon requests

- The pending autologon requests generated by the normal termination of a session between an LU and its controlling application (as displayed in message IST2100I) are only acted upon by the VARY AUTOLOG command.
 - They are not driven by the AUTOTI or AUTORTRY events, nor by PLU notification.
- Once the VARY AUTOLOG command is issued:
 - If the application is session capable
 - A session will be initiated for the LU of the pending autolog request
 - If the application is not session capable or not found
 - The pending autologon request type is changed to a request that can also be driven by the setting of the AUTORTRY and AUTOTI start options or by PLU notification.

```

D NET,AUTOLOG,SCOPE=ALL
IST097I DISPLAY ACCEPTED
IST350I DISPLAY TYPE = AUTOLOG
IST1990I PENDING AUTOLOGON REQUESTS FOR:
IST2100I NETB.APPL1 - NORMALLY LOGGED OFF LUS
IST1997I NETA.LU1
IST314I END
    
```

```

V NET,AUTOLOG,ID=*
IST097I VARY ACCEPTED
IST2093I AUTOLOGON SEARCH INITIATED FOR 1 APPLICATIONS
IST314I END
IST2096I AUTOLOGON SEARCH COMPLETED FOR APPLICATION
      NETB.APPL1
IST2098I STATUS = UNABLE TO ACCEPT LOGONS
IST314I END
    
```

```

D NET,AUTOLOG,SCOPE=ALL
IST097I DISPLAY ACCEPTED
IST350I DISPLAY TYPE = AUTOLOG
IST1990I PENDING AUTOLOGON REQUESTS FOR:
IST1991I NETB.APPL1 - WAITING FOR PLU NOTIFICATION
IST1997I NETA.LU1
IST314I END
    
```


ibm.com



e-business



SNA Networking using Linux on zSeries

SNA/IP Integration



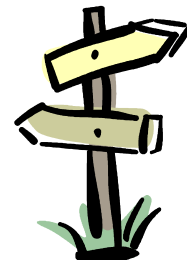
Redbooks

International Technical Support Organization

© Copyright IBM Corp. 2005. All rights reserved.

SNA/IP integration - agenda

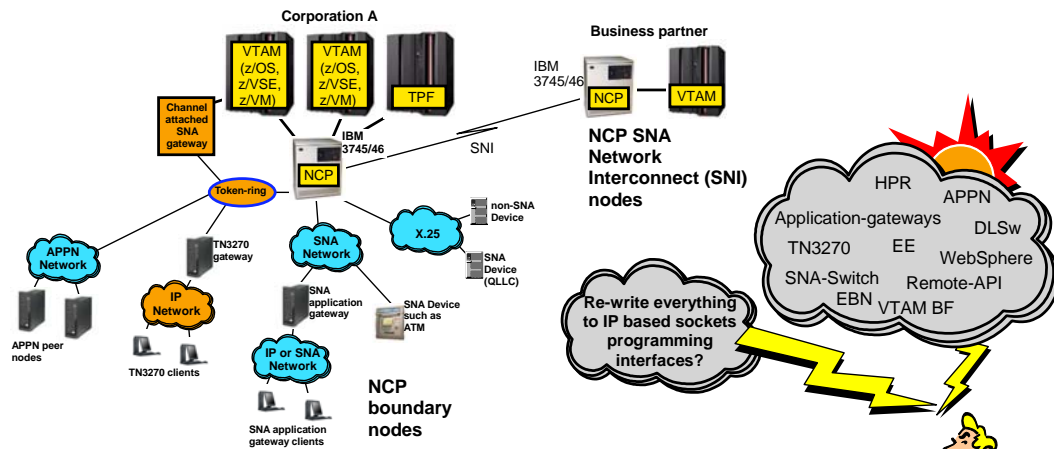
- > Introduction to SNA/IP integration
- > Role of the mainframe
- > SNA environment scenarios



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

A traditional SNA network infrastructure



CCL is recommended for installations that continue to require traditional SNA connectivity, such as SNA Network Interconnection (SNI) to business partners.

Many SNA installations have over the last few years managed to move ahead to more modern SNA functions based on APPN, HPR, and Enterprise Extender - or managed migrating SNA access to one of many forms of IP-based connectivity. Such installations should continue their efforts to deploy those newer technologies.

How do I maintain reliable and cost-efficient access to mainframe SNA business applications and business partners with an aging SNA networking infrastructure?

- ▶ IBM 3745/46 Communication Controller
- ▶ Token-ring technology
- ▶ ESCON channel-attached SNA controllers in general
- ▶ IBM 2216 Nways Multi-access Connector
- ▶ AnyNet

What's the problem with the SNA network?

- **SNA networking infrastructure hardware components are being withdrawn and no longer marketed by networking equipment vendors:**
 - ▶ New replacement or expansion units difficult to locate and acquire
 - ▶ Spare part inventory for repairs is of limited size and shrinking
 - ▶ Technical support by vendors has already or will cease within the next few years
- **SNA networking infrastructure end-to-end design, definition, operation, management, and problem determination skills are diminishing and not being replenished:**
 - ▶ No new graduates have knowledge of SNA networking architecture
 - ▶ Existing SNA support staff is approaching retirement
- **Only a limited set of today's software vendors provide new or enhanced SNA software technologies:**
 - ▶ SNA software can hardly be considered a growth area
 - ▶ New SNA software solutions aim at preserving existing SNA workload, not growing it
 - ▶ Only a very limited set of vendors have an interest in preserving SNA workload

How can I base my business on an SNA network infrastructure that cannot be repaired, expanded, managed, or modernized?

What's the problem with the SNA applications?

➤ **Most mainframe installations still have a mainframe SNA application portfolio to support and to provide access to:**

- ▶ An SNA application either uses SNA programming interfaces directly or relies on its underlying middleware, such as CICS or IMS, to use SNA programming interfaces on its behalf
- ▶ More recent application subsystems are either network programming interface agnostic or are based on IP network programming interfaces - such as sockets
- ▶ Most mainframe installations today have a mix of application subsystems where some require SNA networking technologies and some require IP networking technologies

➤ **The branch and regional offices are increasingly moving towards IP-based client platforms, of which the ultimate is the universal client interface: a WEB browser:**

- ▶ The IP-based client platforms need access to both newer IP-based services in the data center and to older SNA-based services
- ▶ To access SNA-based mainframe applications from such IP-based client interfaces, one or more gateways that transform between different networking- and application protocols are needed somewhere in-between the client platform and the mainframe
- ▶ That's where SNA/IP integration technologies come into the picture!

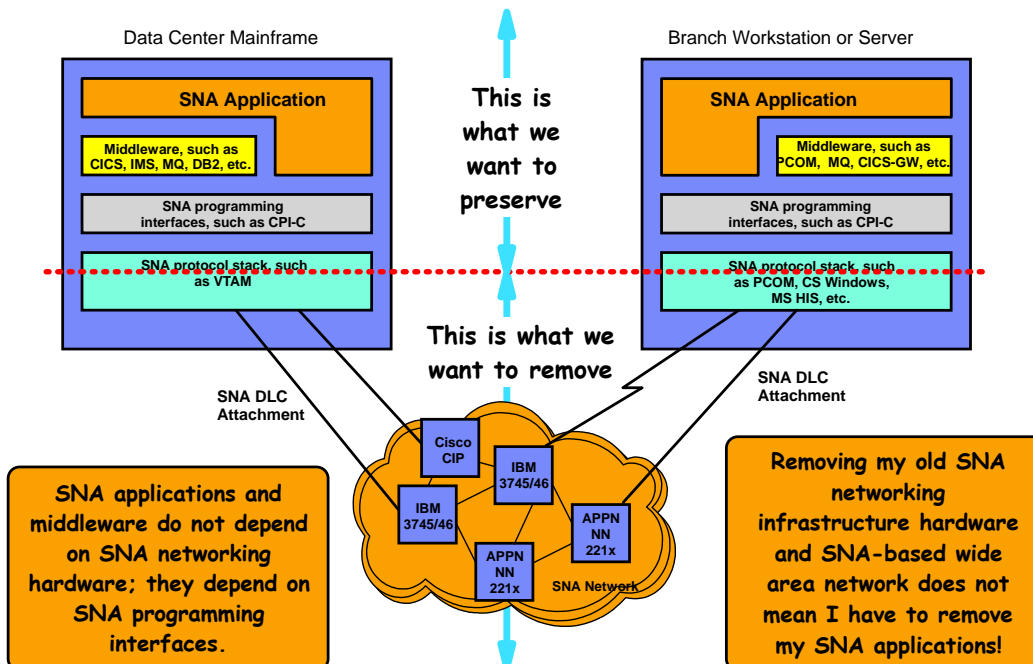
How can I continue to provide access to my SNA-based mainframe applications, while at the same time expanding my IP-based mainframe application portfolio?



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

SNA: what is worth preserving and what isn't?



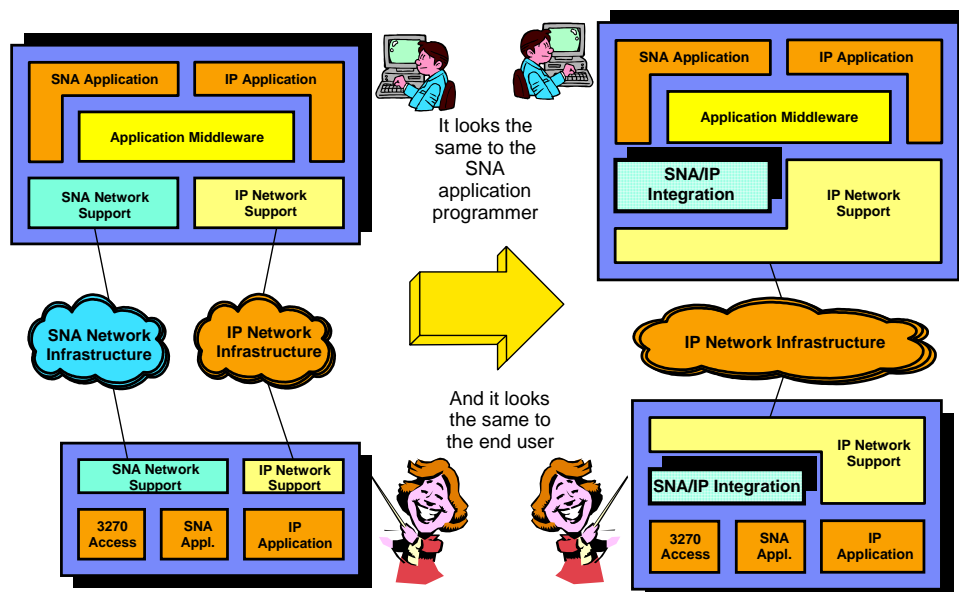
© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Objectives of SNA/IP and host integration

- **Allow installations to preserve SNA applications as long as they have business value:**
 - ▶ There is no need for the "big bang" application migration development project - just because we no longer can buy a new IBM 3745/46
- **Assist installations in modernizing and simplifying their end-user platforms using thin-client technologies, but at the same time preserve SNA server applications on the mainframe**
- **Help remove dependency on an outdated SNA networking hardware infrastructure:**
 - ▶ IBM 3745/46
 - ▶ IBM 2216
 - ▶ OEM ESCON channel-attached SNA gateways, such as Cisco CIP
 - ▶ Token-Ring hardware
 - ▶ Etc.
- **Assist in reducing the need for SNA skills in the enterprise:**
 - ▶ Remove the need for SNA wide area networking skills
 - ▶ Some of the data center related SNA skills may still, to some extent, be needed
- **Reduce the complexity of the overall enterprise networking infrastructure by using a single high-capacity, scalable, reliable, and secure IP-based transport network to provide enterprise-wide connectivity for both SNA-based and IP-based application services:**
 - ▶ One network
 - ▶ One skill set
 - ▶ One set of management tools and procedures

The basics of SNA/IP integration



Examples of SNA/IP integration technologies

> Telnet 3270 (TN3270):

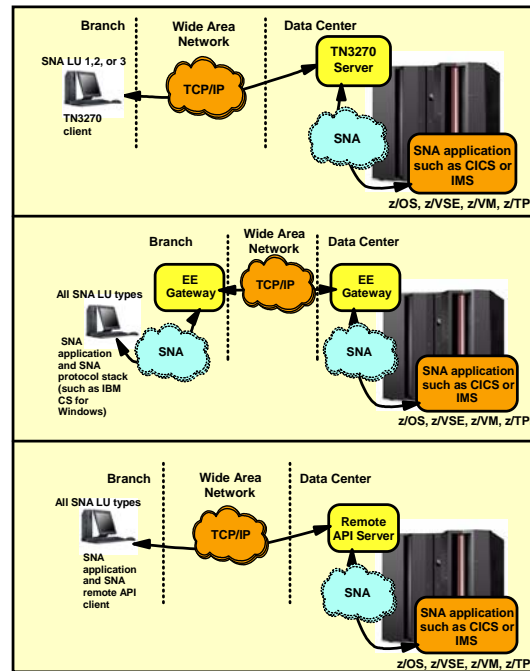
- Client user interface is a completely "normal" IBM 3270 terminal interface
- Support for both SNA terminals and SNA printers
- Communication from client workstation to a TN3270 server is IP-based
- TN3270 server acts as a gateway between the TN3270 TCP connection and an SNA secondary LU
- The TN3270 server secondary LU establishes a session with the SNA application
- The SNA application sees no difference from a traditional real SNA 3270 terminal

> Enterprise Extender (SNA switch, or HPR over IP):

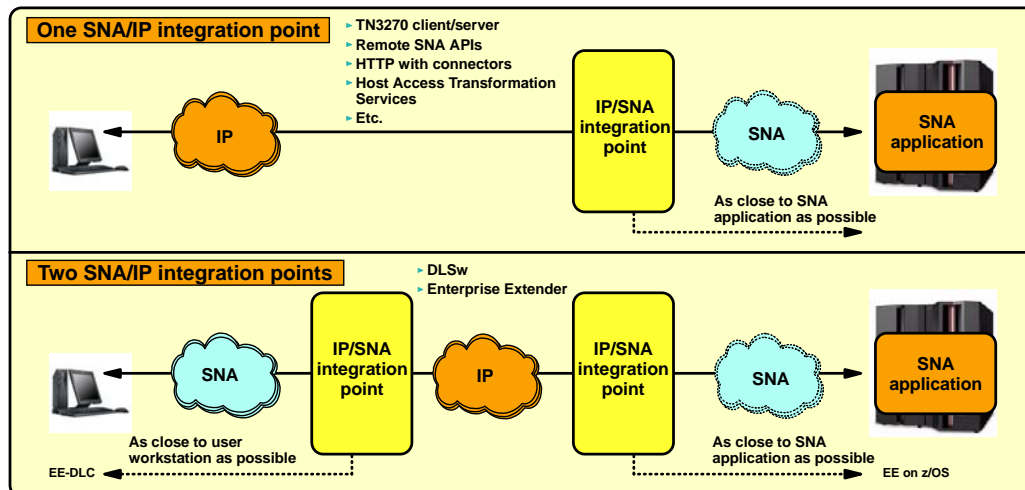
- Uses an IP network as an APPN/HPR link
- Link endpoints can be on gateways (as this example) or, if the platforms support it, directly on the nodes where the SNA applications reside
- EE is an integral part of an SNA APPN/HPR network topology
- Transparent to SNA applications at both application endpoints
- Typically used where SNA applications (LU0 or LU6.2) are located in the branch and need SNA application access to the mainframe

> Remote API (split-stack):

- No SNA protocol stack needed on remote workstation - only a shim layer that is referred to as the remote API client component
- Remote API client component processes all SNA API calls from local SNA application and ships each call over a TCP connection to a remote API server
- The remote API server has a full SNA protocol stack and executes the SNA API call functions on behalf of the remote SNA application
- Transparent to both SNA applications at both application endpoints



Where should the SNA/IP integration points be placed?



> Remember that on one side of the integration point, there has to be some form of an SNA network - so you don't want that side to face your wide area network:

- Place TN3270 servers in the data center - as close to the mainframe as possible (optimal: on the mainframe)

> Where your chosen technology uses two integration points, you want to place those two nodes as close to the SNA application nodes as possible:

- In the branch on the end-user work station or on a branch server
- In the data center as close to the mainframe as possible (optimal: on the mainframe)

Two basic models for SNA/IP integration point on System z9 and zSeries

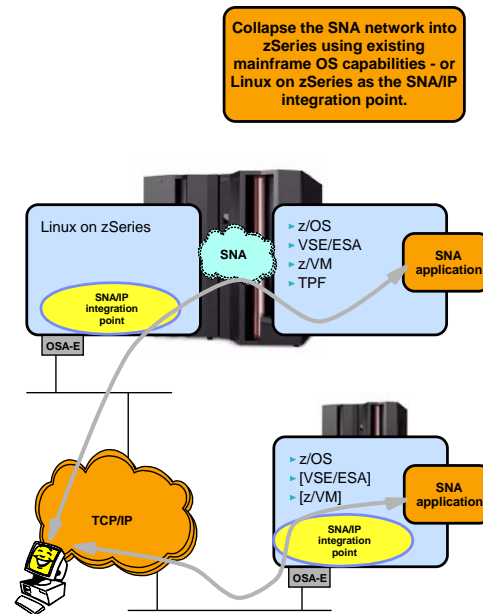
➤ You may end up mixing and matching - using some SNA/IP integration technologies inside your existing System z9 and zSeries operating system environment and others inside Linux on System z9 and zSeries

➤ The Linux on zSeries operating system image may be on the same zSeries as your existing zSeries operating system, or it may be on another zSeries that has some kind of SNA network connectivity:

- ▶ CTC/MPC
- ▶ Shared SNA LAN
- ▶ HiperSockets (EE)

➤ Advantages of implementing SNA/IP integration in Linux:

- ▶ One SNA/IP integration technology may serve all the traditional zSeries operating system environments:
 - z/OS, z/VSE, z/VM, (TPF)
- ▶ Minimal changes to the traditional operating systems' SNA definitions
- ▶ The exact same SNA/IP integration technology may be implemented on zSeries and on other platforms: Intel and power
 - One skill set to work with SNA/IP integration in the branch and in the data center



Premier SNA/IP integration technology: Enterprise Extender

➤ EE offers a one-stop solution for SNA/IP integration that supports both branch and business partner communication and offers the opportunity for use of IP network flows end-to-end

➤ Use of EE requires no changes to SNA applications

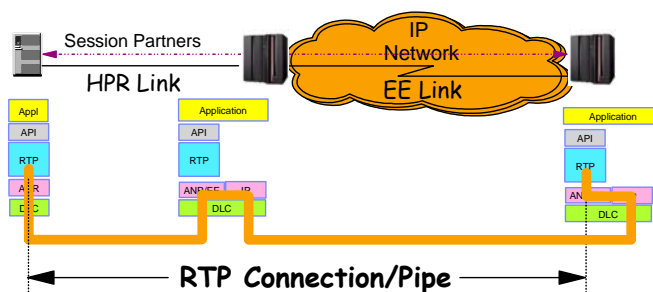
➤ Network infrastructure is native IP, which allows the router infrastructure to maximize router efficiency - no need for routers to perform functions beyond native IP routing

➤ Use of EE can reduce the APPN network complexity by collapsing the APPN Network Node (NN) topology into the data center

- ▶ Minimizes the effect of APPN network searches

➤ EE can be used to implement business partner communication based on the APPN Extended Border Node (EBN) function

- ▶ SNA over IP end-to-end with z/OS business partners



So - why don't all z/OS installations use EE on z/OS?

- Requires APPN and HPR enablement of the z/OS environment
- EE uses UDP packets and that causes problems for firewall administrators
- EE requires coordinated actions by both endpoints
 - Issue for business-to-business communication
 - Not all business partners are able to make similar changes

Positioning CSL, CCL, and HATS

> Communications Server for Linux (CS Linux) on zSeries and System z9

- ▶ SNA/IP integration technologies on zSeries and System z9
 - TN3270 server
 - SNA gateway
 - Enterprise Extender gateway
 - Split stack - remote SNA API services (main migration technology for Anynet)
- ▶ SNA APPN node functions (NN, EN, BX)
 - Migrate IBM 3746 MAE and NNP functions to CS Linux

> Communication Controller for Linux (CCL) on zSeries and System z9

- ▶ Preserving selected NCP functions
 - Migrate the IBM 3745 NCP to the CCL platform
- ▶ Preserving existing SNA subarea connectivity if desirable
 - Including SNI
 - Including traditional SNA boundary functions
- ▶ Alternative for those who for various reasons cannot make the move to an APPN/HPR and Enterprise Extender environment

> The WebSphere Application Server environment - the platform for the Services Oriented Architecture environment

- ▶ Host Access Transformation Services
 - Thin client access to SNA 3270 based applications - from a Web browser
 - Option to transform IBM 3270 dialog to a modernized HTTP/HTML-based dialog without changing the mainframe 3270 applications
- ▶ Connectors
 - Standard and home-written to act as gateways between the WebSphere Application Server environment and existing mainframe SNA applications



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Is there a recommended order of implementing SNA/IP integration?

A multi-step approach:

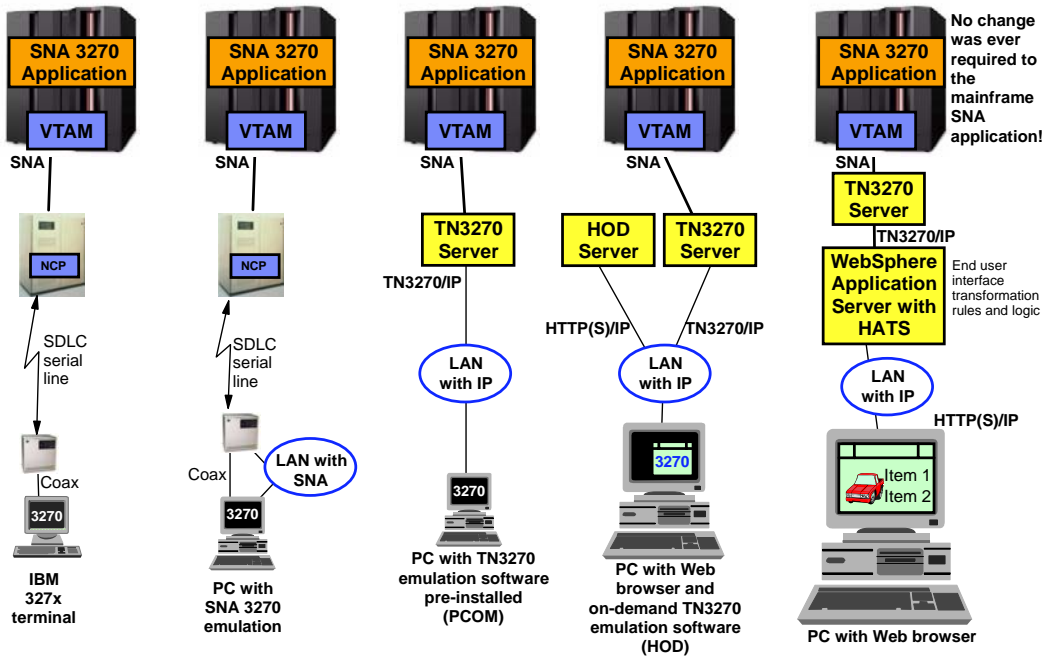
1. Consolidate intranet **SNA 3270** traffic from programmable workstations (LU1/SCS, LU2, LU3/DSC) into the data center:
 - Using TN3270 client software (PCOM, HOD, OEM) on the user workstation connecting to a TN3270 server in the data center, which could be the Communications Server on z/OS or Communications Server for Linux on zSeries.
 - Using standard Web browser on the user workstation connecting to WebSphere Application Server Host Access Transformation Services on a server node in the data center, which could be z/OS or Linux on zSeries.
2. Move **middleware** communication off SNA where applicable:
 - DB2 DRDA, MQ, and many file transfer applications can be migrated to native IP communication without impact on end-user applications.
3. For **SNA Client/Server applications** in the branches or remote locations (LU0, LU6.2), use one of the following technologies to transport the SNA data over an IP network:
 - Use Enterprise Extender to transport native SNA flows over an IP WAN network from the branch and into the data center.
 - Use a remote SNA API technology to ship SNA application calls over an IP network to an SNA API server running on CS Linux in the data center.
 - Use a DLSw infrastructure and continue using an NCP to perform the SNA boundary functions, but move the NCP to the Communication Controller for Linux on zSeries environment.
4. For **business partner** SNA communication, use one of the following technologies:
 - Use Enterprise Extender and the APPN Extended Border Node (EBN) capabilities to migrate from SNI to APPN Multiple Network Connectivity.
 - Continue using SNI, but move the SNI NCP to the Communication Controller for Linux on zSeries environment.



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

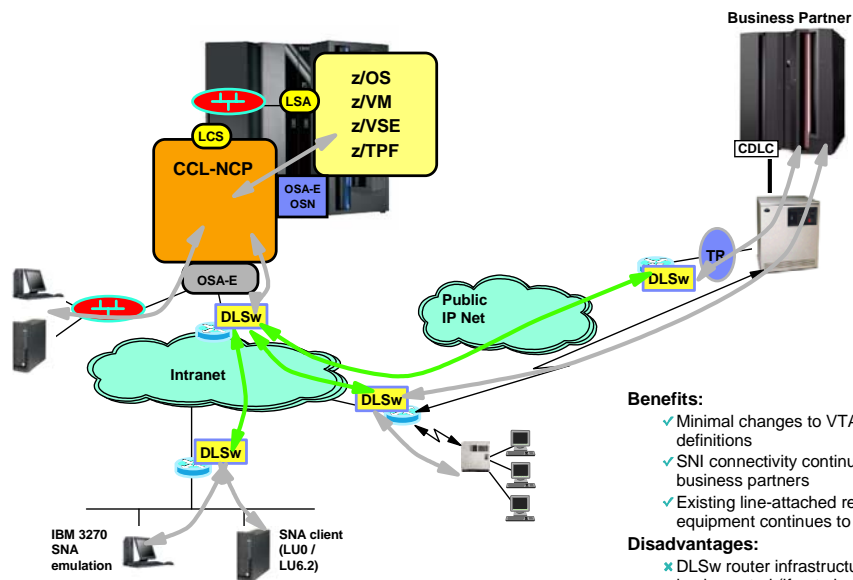
IBM 3270 SNA application access evolution



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Scenario 1: Retain SNA subarea environment as unchanged as possible



Benefits:

- ✓ Minimal changes to VTAM and NCP definitions
- ✓ SNI connectivity continues unchanged to business partners
- ✓ Existing line-attached remote SNA equipment continues to operate as today

Disadvantages:

- ✗ DLSw router infrastructure needs to be implemented (if not already there)
- ✗ No exploitation of modern SNA capabilities as available in APPN/HPR



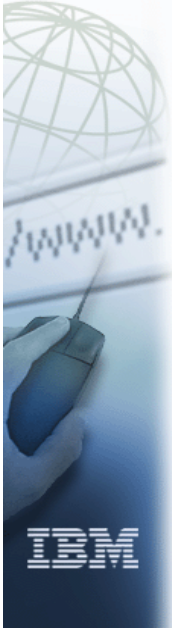
© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

ibm.com



e-business



SNA Networking using Linux on zSeries Communications Server for Linux



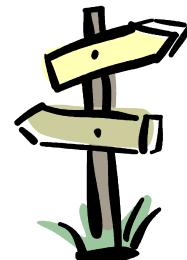
Redbooks

International Technical Support Organization

© Copyright IBM Corp. 2005. All rights reserved.

Communications Server for Linux - agenda

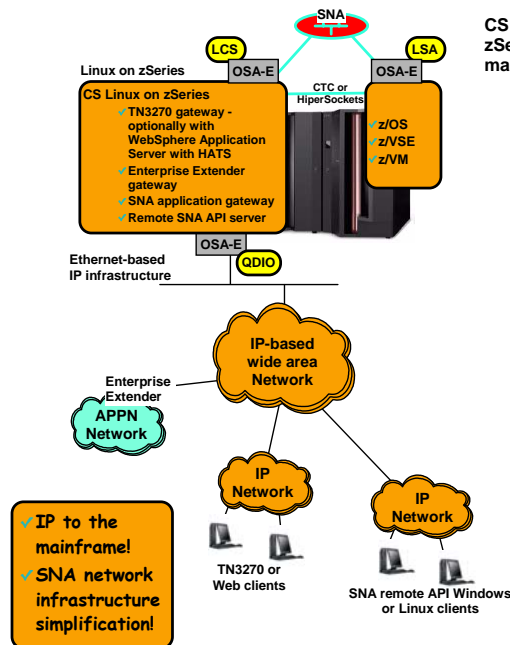
- > CS Linux Version 6.2.1
- > CS Linux on zSeries - connectivity
- > CS Linux on zSeries - selected functions



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Communications Server for Linux on zSeries - zSeries resident traditional SNA/IP integration - an overview



CS Linux on zSeries offers SNA/IP integration technologies on zSeries, but with no or minimal changes to the traditional mainframe OS environment.

1. Enterprise Extender same-NETID gateway functions
 - Using APPN/ISR routing to/from VTAM and EE downstream
 - EE gateway to z/OS, z/VSE, or z/VM VTAM
2. TN3270 server on zSeries
 - Supports TN3270 access to z/OS, z/VSE, and z/VM
 - Can be combined with WebSphere Application Server and Host Access Transformation Services
 - IP all the way to zSeries
 - No or minimal change to VTAM definitions if consolidating existing distributed TN3270 servers
3. TN3270 SSL offload - using the TN3270 redirector
4. APPN Network Node or Branch Extender node in an APPN network infrastructure
5. SNA gateway for consolidation of multiple downstream SNA PUs
6. SNA application platform for Web-based access to SNA applications
7. Remote API services for remote SNA application access without having SNA protocol stacks on distributed Windows, Linux, and AIX nodes



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

IBM Communications Server for Linux on zSeries Version 6.2.1

- **Advanced Peer-to-Peer Networking (APPN) support**
 - APPN End Node (EN) or APPN Network Node (NN) support
 - Uses Dependent LU Requester (DLUR) for dependent LU access over an APPN network
 - Supports connection networks
- **High Performance Routing (HPR) including Enterprise Extender (EE, also known as HPR over IP)**
- **Branch Extender (BX) support**
 - Allows for APPN network topology simplification
- **SNA API support**
 - CPI-C and APPC APIs for both dependent and independent LU6.2 - including extensions for both Java and C
 - Java Host Access APIs
 - LUA APIs (Request Unit Interface (RUI) and Session Level Interface (SLI)) for dependent LU functions (LU types 0, 1, 2, and 3)
 - Primary LU 0 support for the LUA APIs
 - Remote SNA client/server APIs
 - ◆ Client support on Windows, AIX (32 and 64 bit), Linux (Intel i686 and x86_64, Power ppc64, zSeries s390 and s390x)
 - APPC application suite (AFTP, APING, AREXEC, ATELL, ACOPI, and ANAME)
- **TN3270E server**
 - Including SSL with client authentication and Express Logon support
 - Telnet redirector - allows Telnet port mapping and/or Telnet passthru from SSL to non-SSL
- **Administration**
 - Motif-based administration (GUI interface)
 - Network Operator Facility (NOF) APIs for programmed administration
 - Internationalization
 - 31-bit and 64-bit support
 - Runs on both Red Hat and SUSE (both 2.4 and 2.6 kernel levels)
- **Network attachments for SNA**
 - Enterprise Extender (HPR over IP)
 - (V)CTC using MPC channel protocols (Linux as a PUT2.1 - APPN/ISR routing)
 - Native SNA (SNA LLC2) over shared LAN (Ethernet or Token-Ring)



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

The distributed IBM Communications Server family

> The distributed IBM Communications Servers are supported on both Windows, AIX, and Linux:

- ▶ AIX
 - IBM Communications Server for AIX Version 6.1
 - We do plan a follow-on release for Communications Server for AIX 6.1.
 - It is our intention to continue supporting CS/AIX 6.1 to provide reasonable migration time for our customers.
- ▶ Windows 2000, Windows 2003, or Windows XP:
 - IBM Communications Server for Windows Version 6.1.2
- ▶ Linux on Intel and Power platforms
 - IBM Communications Server for Linux Version 6.2.1
- ▶ Linux on IBM zSeries and System z9
 - IBM Communications Server for Linux on zSeries Version 6.2.1

A comprehensive set of consistent, operating system- and hardware platform-independent SNA/IP integration solutions to be deployed in the enterprise across branches, regional offices, and the data center.

The IBM Communications Servers for Linux Version 6.2.1 product set:

| Architecture | Platform | RHAS 2.1 2.4 kernel | SLES 8 2.4 kernel | RHEL 3 2.4 kernel | SLES 9 2.6 kernel | RHEL 4 2.6 kernel |
|--------------|------------------------------|------------------------|----------------------|----------------------|----------------------|----------------------|
| i686 | Intel, 32-bit | ✓ | ✓ | ✓ | ✓ | ✓ |
| ppc64 | OpenPower or Power 5, 64-bit | | | | ✓ | ✓ |
| s390 | zSeries, 31-bit | | ✓ | ✓ | ✓(1) | ✓(1) |
| s390x | zSeries, 64-bit | | ✓ | ✓ | ✓ | ✓ |

Note (1): Indicates that support may be deprecated in future Linux releases.



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

What is new in CS Linux Version 6.2.1?

> 2.6 Linux Kernel support

- ▶ SLES 9, RHEL 4

> AIX Remote API Client

- ▶ Client support for AIX V5 (5.x), CICS Transaction Server needs

> Linux on Power

- ▶ OpenPower and Power 5 platforms
 - SLES 9 and RHEL 4 only
- ▶ ppc64 kernel on server
- ▶ ppc32, ppc64 on client

> Primary RUI interface

- ▶ Interface is documented in "What's New in this Release" document.

> Remote API Clients

- ▶ APARS included for latest maintenance release

> CS Linux on zSeries

- ▶ Support for SNA LLC2 access over QDIO Layer-2 OSA-Express ports

> Communications Server for Linux (5724-I33):

- ▶ Communications Server (Linux, i686, Intel - SLES 8, SLES 9, RHAS 2.1, RHEL 3, RHEL 4)
- ▶ Communications Server (Linux on Power, ppc64 - SLES 9, RHEL 4)
- ▶ Remote API Clients (Windows, Linux (Intel), Linux on Power, Linux on zSeries, AIX)

> Communications Server for Linux on zSeries and System z9 (5724-I34):

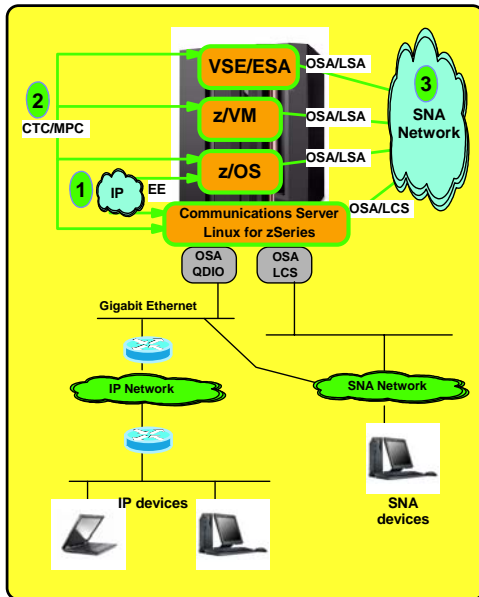
- ▶ Communications Server (Linux on zSeries and System z9, s390, s390x - SLES 8, SLES 9, RHEL 3, RHEL 4)
- ▶ Remote API Clients (Windows, Linux (Intel), Linux on Power, Linux on zSeries, AIX)



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

SNA connectivity options



➤ SNA between Linux for zSeries and other zSeries operating systems:

1. Enterprise Extender (to z/OS only).

- z/OS needs to be EE-enabled in order for this option to work.
- Physical connectivity may be via HiperSockets
- z/OS is the only mainframe operating system that supports EE (besides Linux). z/VM, VSE/ESA, and TPF do not support EE connectivity.

2. An MPC (Multi Path Channel) CTC (Channel to Channel) channel driver supports APPN Node-to-Node Communication over a CTC (virtual, EMIF, or real channel-to-channel).

- Use of this option requires both endpoints of the channel to be PU type 2.1 nodes, which means existing mainframe environments that have not enabled APPN support in VTAM will have to do some APPN enablement to communicate with CS Linux for zSeries using this option.

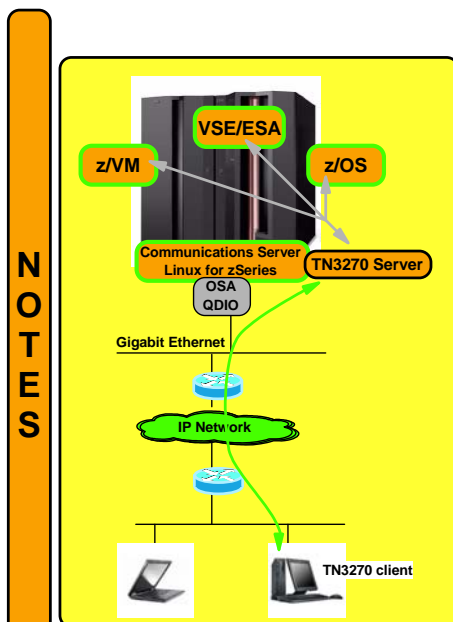
3. Use of Linux for zSeries LCS device driver with a configurable SAP on an OSA adapter to exchange SNA LLC2 frames with the underlying LAN (Ethernet or token-ring).

- This option depends on OSA-express microcode upgrades (z800, z900: 3.50 - z990: 5.50). There are no plans to ship this support for pre-zSeries models.
- Works for OSA-2 as-is (in non-shared TCP/IP passthru mode, not in any shared mode). Works for pre-zSeries models also.
- Also works for QDIO Layer-2 mode OSA-Express ports.
- Use of this option allows the mainframe operating system to view CS Linux as one of three SNA node types:
 - Peripheral node (a plain PU type 2.0)
 - LEN node (a PU type 2.1)
 - APPN node (a PU type 2.1)

© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Consolidating existing distributed TN3270 Servers into Linux for zSeries



➤ Minimal or no changes to VTAM definitions of TN3270 server PUs and LUs

- ▶ Continue to look like a PU type 2.1 (or 2.0) with dependent LUs of type 1, 2, and 3
- ▶ USS table handling continues to be performed by the VTAM SSCP
- ▶ Default application logon continues to be handled via existing VTAM definitions

➤ Configuration concepts for TN3270 servers remain similar to how they were for the distributed TN3270 servers

➤ Connectivity to zSeries via Gigabit Ethernet and QDIO

➤ SNA connectivity between Linux for zSeries and:

- ▶ z/OS: EE (HiperSockets), CTC/MPC, or shared LAN
- ▶ z/VM and VSE/ESA: CTC/MPC or shared LAN

➤ SNA collapsed into the data center

➤ In most configurations, the LU element addresses will come out of VTAM's high-order address pool

➤ Reduced dependency on IBM3745/46, CIP, or token-ring hardware

➤ We do not recommend customers moving from the z/OS TN3270 to the CS Linux TN3270 server

© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Example of functional comparison between two alternatives

NOTES

| Functional area of interest | CS Linux on zSeries | CS z/OS |
|---|--|---|
| Multiple TN3270 server ports per OS image | Yes, multiple ports can be defined | Yes, multiple ports (255) per server instance (8 instances per z/OS LPAR) |
| LU assignment rules per port | Shared among all ports | Can be shared inside a server instance or separate per port |
| LU name assignment based on client IP address | Yes | Yes |
| LU name assignment based on client host name | Yes | Yes |
| LU name assignment based on server IP address | No | Yes |
| LU name assignment based on server link name over which connection was received | No | Yes |
| LU name assignment based on user ID (if SSL/TLS with client authentication is used) | No | Yes |
| Secure connections | Yes (SSL connections) | Yes (SSL or TLS connections) |
| Secure connections with client authentication | Yes (signature verification of certificate signer) | Yes (signature verification with optional SAF authentication and port protection) |
| Support for ELF (Express Logon Feature) | Yes (via z/OS DCAS server) | Yes |

Example of functional comparison between two alternatives (continued)

NOTES

| Functional area of interest | CS Linux on zSeries | CS z/OS |
|---|--|---|
| TN3270E support | Yes | Yes |
| TN3270E contention support | Yes | Yes |
| Printer association | Yes | Yes |
| Specific LU name request - both specific LU and specific LU group | Yes | Yes |
| Support for user exit routine to assign LU names | No | Yes |
| ANS=CONT support | Yes | No |
| Capacity per server instance | Tests done with up to 20,000 connections | Tests done with up to 128,000 connections |
| Built-in response time monitoring | No | Yes - SNA, IP, and full round-trip response time |
| SNA session re-connect support | No | Yes - for both generic and specific LU name assignments |
| Telnet redirector support | Yes - including SSL offload to redirector with non-SSL redirection | No |

Example of functional comparison between two alternatives (continued)

NOTES

| Functional area of interest | CS Linux on zSeries | CS z/OS |
|--|---|--|
| How is initial SNA application chosen? | N/A - done via traditional VTAM definitions (LOGAPPL) | Can be assigned based on client IP address, host name, server IP address, link name, or user ID |
| Can TN3270 server perform access authorization to SNA application? | No | Yes - via assignment rules based on certificate-derived user ID, or via user ID derived through use of built-in network solicitor function |
| USS table support | N/A - uses standard VTAM SSCP USS table processing | Controlled by TN3270 server - VTAM USS table can be used as-is, or TN3270 server-specific versions can be used |
| Connection load-balancing support | Traditional load balancing | Sysplex Distributor or traditional load balancing with SASP support |
| Server identity take-over | Manual or automated operations to move IP address to another Linux OS image | Sysplex dynamic VIPA policies to move IP address to another z/OS image in the Sysplex |
| Accounting data for charge back | | Yes - SMF records |

Great for "smaller" TN3270 server environments!

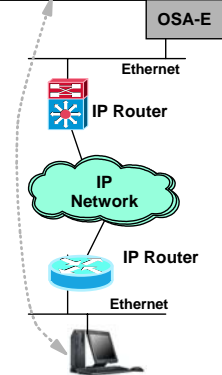
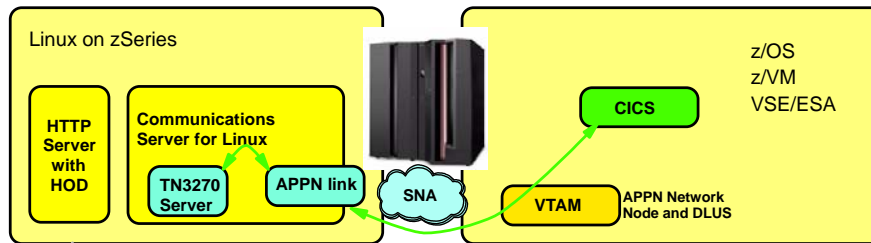
Remains the preferred choice for large TN3270 server environments!

© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

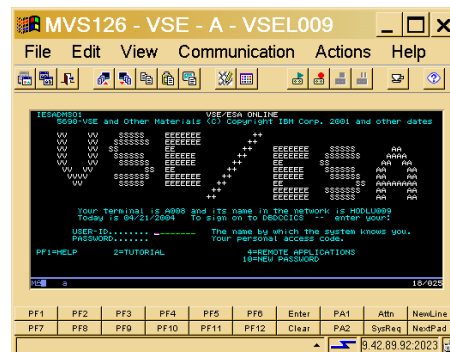
Combining CS Linux on zSeries with Host On Demand

NOTES



To avoid pre-installing 3270 emulator software on the workstation, an HTTP server serving a HOD client can be deployed on Linux and combined with the TN3270 server of CS Linux

- Web browser used to access a HOD client
- HOD client connects over TN3270/IP to TN3270 server in Linux
- TN3270 server in Linux uses DLUR to connect to DLUS in VTAM in z/OS, z/VM, or VSE/ESA



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

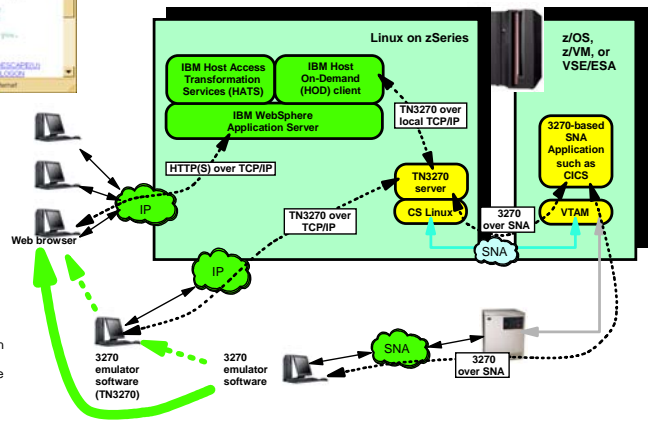
IBM 3270 access: one step further - CS for Linux on zSeries and IBM's Host Access Transformation Services

NOTES



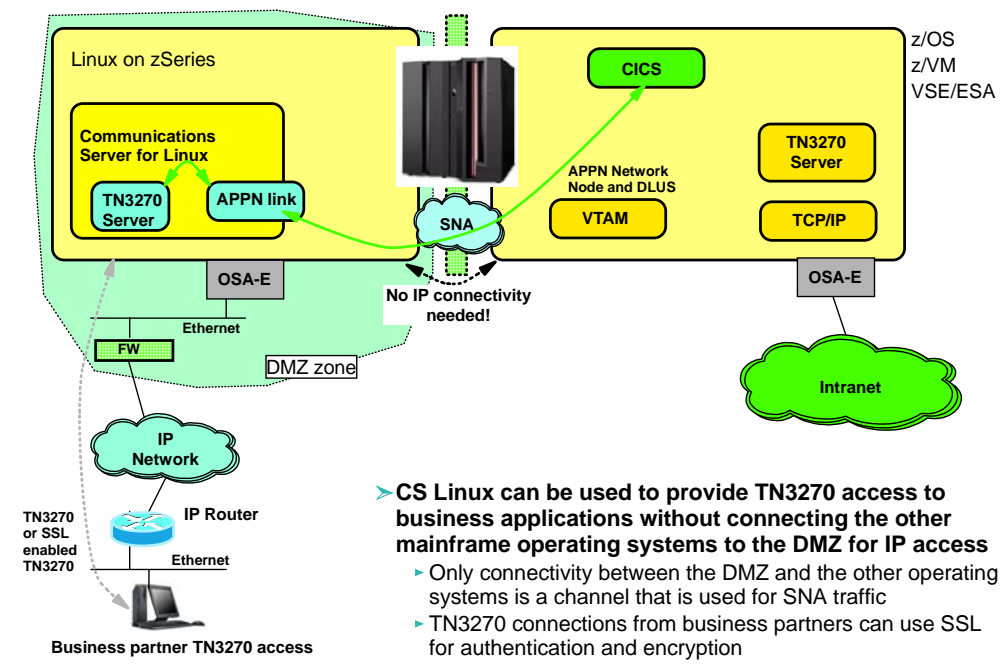
- > Universal workstation client: Web browser
 - Basically all platforms with a Web browser are supported
- > No 3270 emulator software (fat client or down-loaded) on workstation
- > Only HTTP/HTTPS over IP between workstation and WebSphere Application Server
 - Simplifies firewall setup
- > No changes to existing mainframe SNA 3270 applications
- > User interface can remain 3270-like, or it can be transformed using the WebSphere Studio tooling

- Network infrastructure simplification**
 - IP network access from end user to mainframe
 - SNA network collapsed into zSeries
- Scalability**
 - Vertical scaling through zSeries 64-bit storage support and powerful parallel CPU engines
 - Horizontal scaling through z/VM technologies
- Availability**
 - zSeries hardware availability
 - Multiple parallel virtual environments can be deployed
- Security**
 - Security-rich internal network connectivity between Linux and the mainframe operating systems
 - Encryption/decryption of HTTPS connections done with zSeries IFL engines and hardware crypto support



Secure business partner TN3270 access via CS Linux

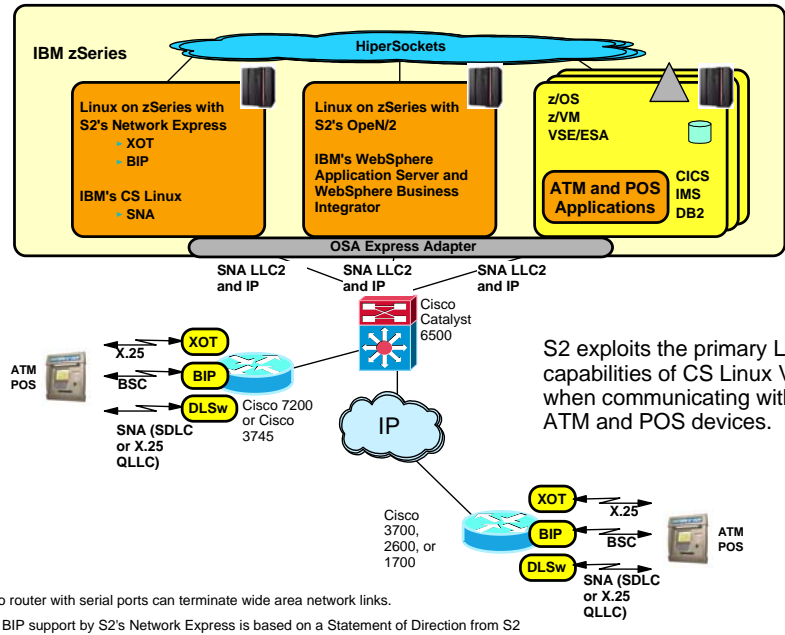
NOTES



- > CS Linux can be used to provide TN3270 access to business applications without connecting the other mainframe operating systems to the DMZ for IP access
 - Only connectivity between the DMZ and the other operating systems is a channel that is used for SNA traffic
 - TN3270 connections from business partners can use SSL for authentication and encryption

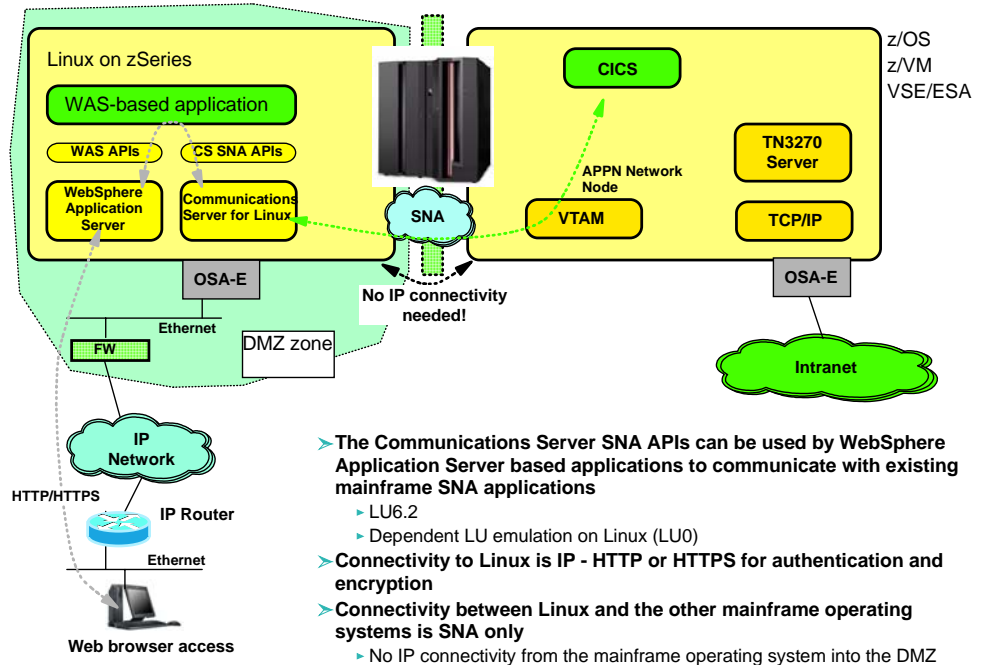
ATM and POS access scenario - Utilizing S2's Network Express and OpenN/2

NOTES



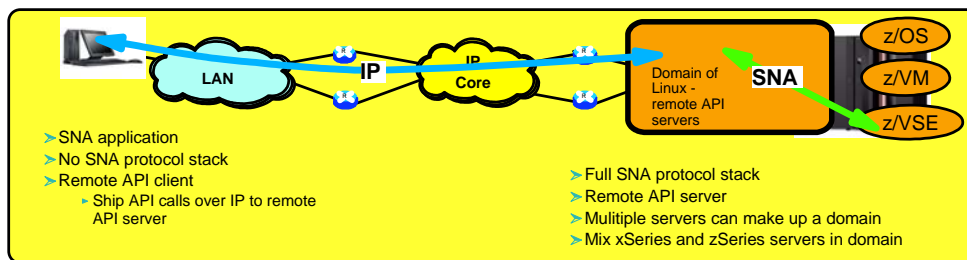
Business partner Web access to mainframe SNA applications through the CS Linux SNA APIs

NOTES



Remote API overview

- The remote SNA API support allows SNA application programs to reside on nodes that don't implement a full SNA protocol stack.
- The SNA API calls are intercepted by a shim layer that ships the calls over a TCP connection to a remote API server where the actual SNA API calls are executed.
- This technology provides a solution for SNA application programs that must remain in remote locations - without requiring SNA protocol stacks on those remote nodes.
 - Removing the need for SNA stack configuration skills, management, and operations procedures outside the data center where the remote SNA API servers may be collapsed
- This technology also provides built-in availability and load-balancing to a pool (domain) of remote API servers
 - A remote API client is not limited to use a single remote API server
 - LUs and pools of LUs can be shared across servers on a remote API server domain.
 - Servers can be configured to back up each other
- There is no charge for installing the remote API client - usage is covered by per-user server charge
- Support Windows (XP, 2000, 2003), AIX, Linux, Linux on Power and Linux on zSeries clients

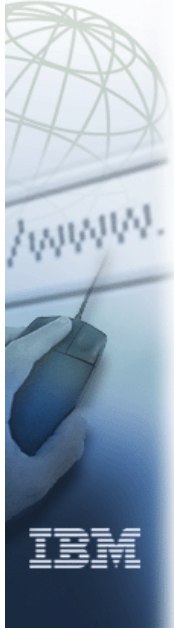


This page intentionally left blank

ibm.com



e-business



SNA Networking using Linux on zSeries

Communication Controller for Linux on System z9 and zSeries



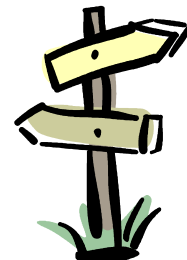
Redbooks

International Technical Support Organization

© Copyright IBM Corp. 2005. All rights reserved.

Communication Controller for Linux (CCL) agenda

- > CCL V1R1 overview
- > CCL implementation and NCP migration planning
- > CCL use of OSA copper ports in LCS mode
- > MAC addressing when using SNA over a LAN
- > Installation and customization overview
- > Introducing CCL V1R2
- > Selected CCL solution scenario:
 - CDLC channel connectivity to CCL
- > Selected CCL R2 solution scenario:
 - QDIO Layer 2 access from native Linux LPAR

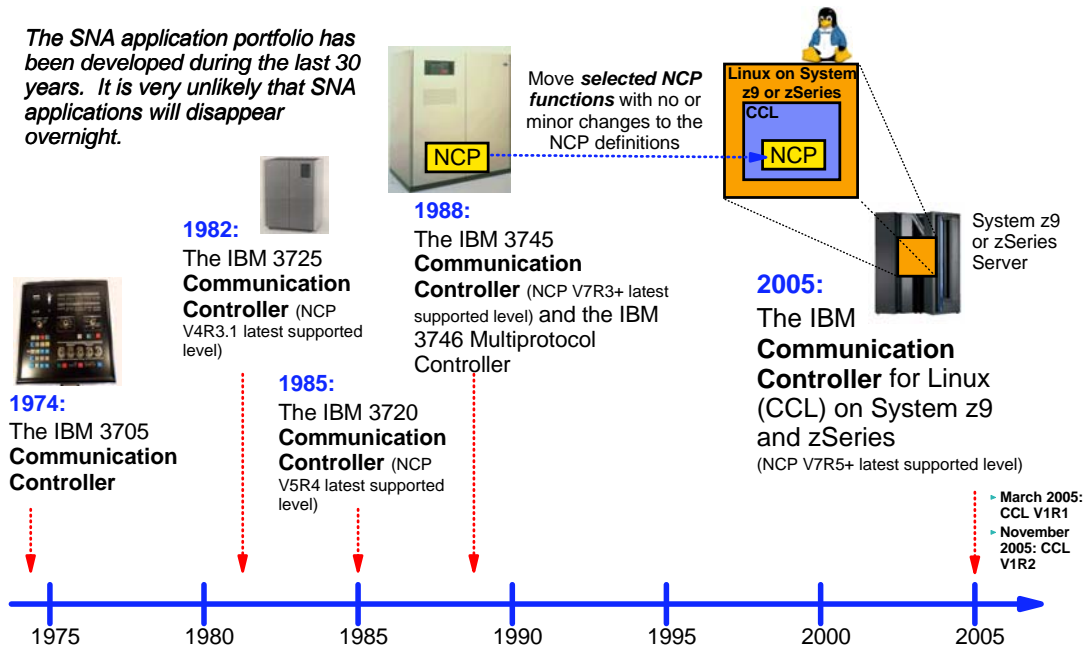


© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

IBM Communication Controllers - the foundation of SNA application access to the IBM mainframe since 1974

The SNA application portfolio has been developed during the last 30 years. It is very unlikely that SNA applications will disappear overnight.



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

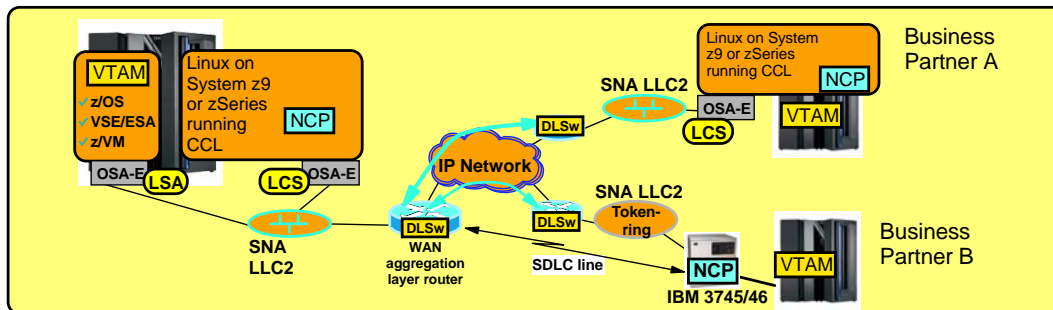
CCL V1R1 overview



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

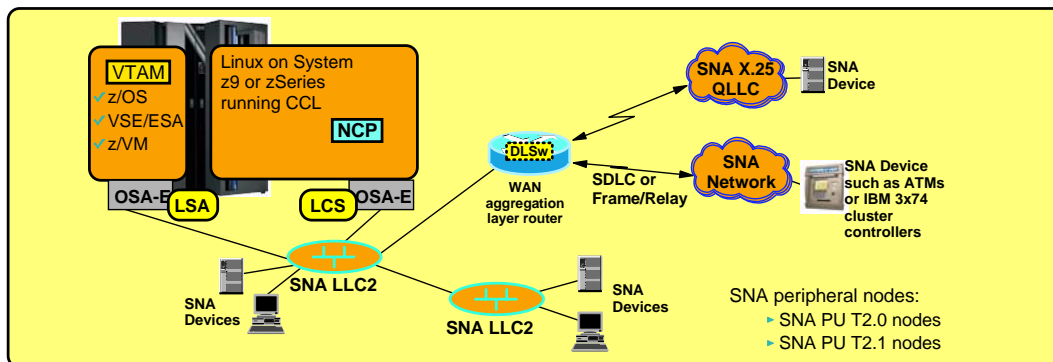
CCL V1R1 recap - primary objective: preserve SNI connectivity to business partners (SNI/INN traffic)



> SNI NCP gateway functions move to Linux on System z9 or zSeries.

- ▶ Business partners may continue to use IBM 3745/46 technology or move to a CCL implementation.
- ▶ In CCL V1R1 all SNA traffic leaves/enters CCL as SNA LAN traffic (SNA LLC2) over an OSA port operating in LCS mode.
 - OSA ports operating in LCS or LSA mode are limited to copper-based cabling (Cat5 and RJ45).
- ▶ VTAM sees the NCP as a LAN-attached remote NCP over its OSA port operating in LSA mode.
- ▶ SNA traffic to/from the business partner location can be tunneled (typically DLSw) over an IP network.
- ▶ An SDLC line from the business partner's IBM 3745/46 can be terminated in a local wide area network aggregation layer router (a router with WAN network interfaces).
 - Aggregation layer routers do not support Multi-Link Transmission Groups (MLTG) in such a setup.
- ▶ Has no impact on existing SNI topology.
- ▶ Has minimal impact on existing SNA network management procedures and disciplines.

CCL V1R1 recap - secondary objective: preserve selected NCP boundary functions (BNN traffic)

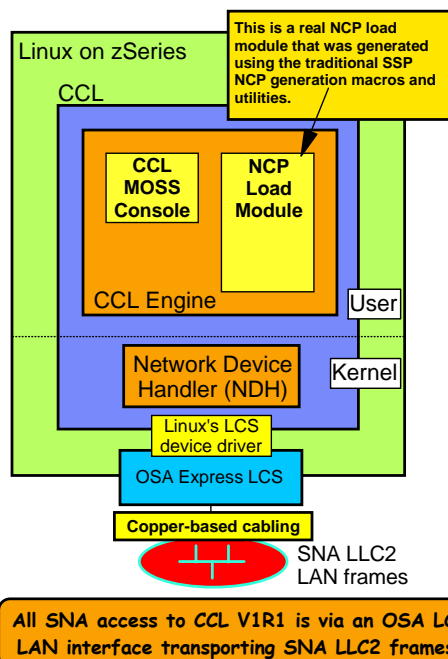


- > Also for boundary functions, all SNA traffic in CCL V1R1 leaves/enter the CCL as SNA LAN traffic (SNA LLC2) over an OSA port operating in LCS mode.
- > SNA wide area network links to which SNA peripheral nodes are attached must physically be moved from the IBM 3745/46 to a wide area network aggregation layer router that switches the SNA frames between the serial line interfaces and the local LAN:
 - ▶ NCP SNA connectivity via SDLC, Frame Relay, and SNA X.25 QLLC links are supported
 - The X.25 SNA support (QLLC) does not imply NPSI support by CCL V1R1
 - ▶ NCP boundary function support includes standard availability functions such as SSCP takeover, support for duplicate MAC addressing, and XRF
 - ▶ NPA-LU, NtuneMON, and NRF are also supported
 - ▶ Dial-in to the aggregation layer router will work, but NCP-initiated dial-out is not supported
- > Remote SNA LANs can be connected to the data center LAN by bridges or DLSw technology

CCL V1R1 was NOT a complete replacement for the IBM 3745/46 Communication Controller!

| CCL Functional Overview Matrix | CCL V1R1 supports | CCL V1R1 support of serial lines via an aggregation layer router | CCL V1R1 does not support |
|------------------------------------|--|--|--|
| Software | NCP (V7R5 and above) and compatible levels of NRF SSP, NTuneMON, NetView, and NPM continue to work as they have in the past | | Other IBM 3745 software products: NPSI, XI/NSF, EP, NTO, NSI, MERVA, and TPNS Functions provided by the IBM 3746 MAE or NNP NCP-based IP routing |
| Physical network interfaces | OSA token-ring and Ethernet LAN (uses an LCS interface that is only supported by certain, copper-based, OSA cards) Though NCP only supports SNA over token-ring, CCL transparently converts Ethernet frames to token-ring for the NCP | SDLC, Frame Relay, X.25 QLLC, and ISDN serial line interfaces are not supported directly by CCL, but are supported via an aggregation layer router | Channel, BSC, ALC, Start/Stop, and X.25 non-SNA lines |

CCL structure and components



> CCL consists of both user-space and kernel-space functions:

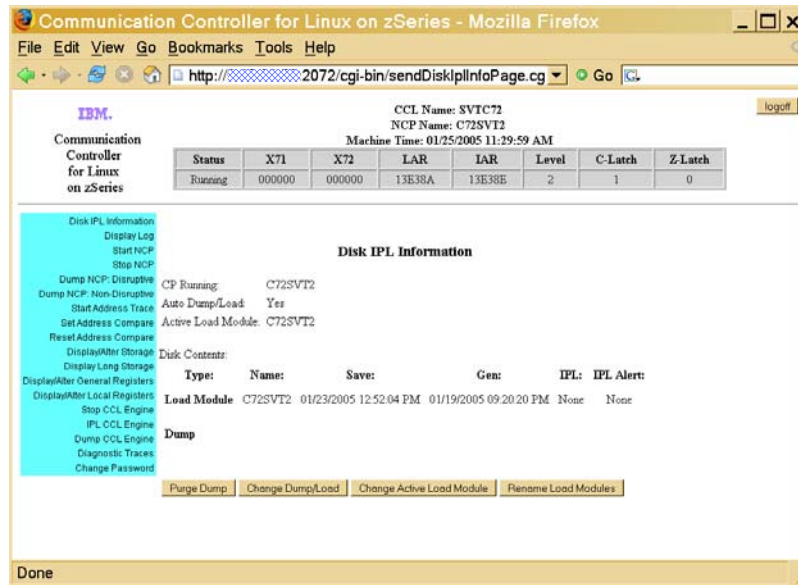
- > **CCL engine** emulates an IBM 3745-31A with 16 MB memory supporting an NCP load module and a MOSS console interface.
- > The **MOSS console** is accessed through a standard Web browser.
- > **Network Device Handler (NDH)** is a kernel extension that acts as the interface between an OSA port operating in LCS mode and the NCP Token-Ring Interface (NTRI).
 - The only supported network interface from an NCP perspective is an SNA TIC interface.
 - The actual LAN to which the OSA port is connected may be either token-ring or IEEE802.3 Ethernet (NDH will transform between the frame formats).

> **NDH components**

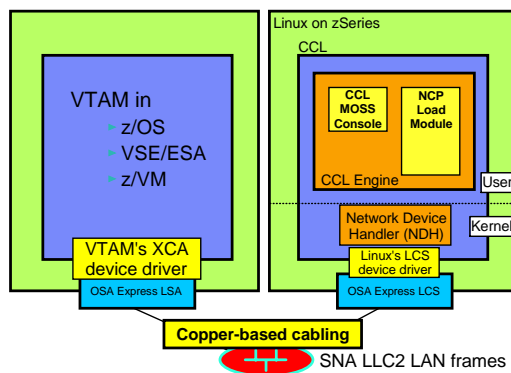
- > NDH itself consists of two components:
 - A small source code isolation module that is built during installation of CCL
 - An object code only NDH module
- > Both are dynamically loaded into kernel space. No kernel rebuild/reboot is required.

CCL and the MOSS console interface

➤ The CCL MOSS console functions are accessed via a Web browser.



CCL V1R1 and VTAM communication



➤ VTAM sees the CCL NCP as a LAN-attached remote NCP through its XCA (eXternal Communications Adapter) interface, such as an OSA LSA interface

- VTAM may either be the owning host or a data host to the CCL NCP
- For VTAM to be the owning host, a VTAM PTF will be made available for VTAM to activate and own a CCL NCP through an XCA network interface:
 - New keyword on the XCA PU statement to allow VTAM to activate and own CCL NCP resources over an XCA interface
 - ALLOWACT=NO/YES

- VTAM maintenance details:
- OS/390 and z/OS VTAM: APAR OA10425
 - VSE/ESA VTAM: APAR DY46311
 - z/VM VTAM: APAR VM63677

➤ No SNA subarea topology changes - VTAM is still a PU Type 5 and the NCP is a PU Type 4

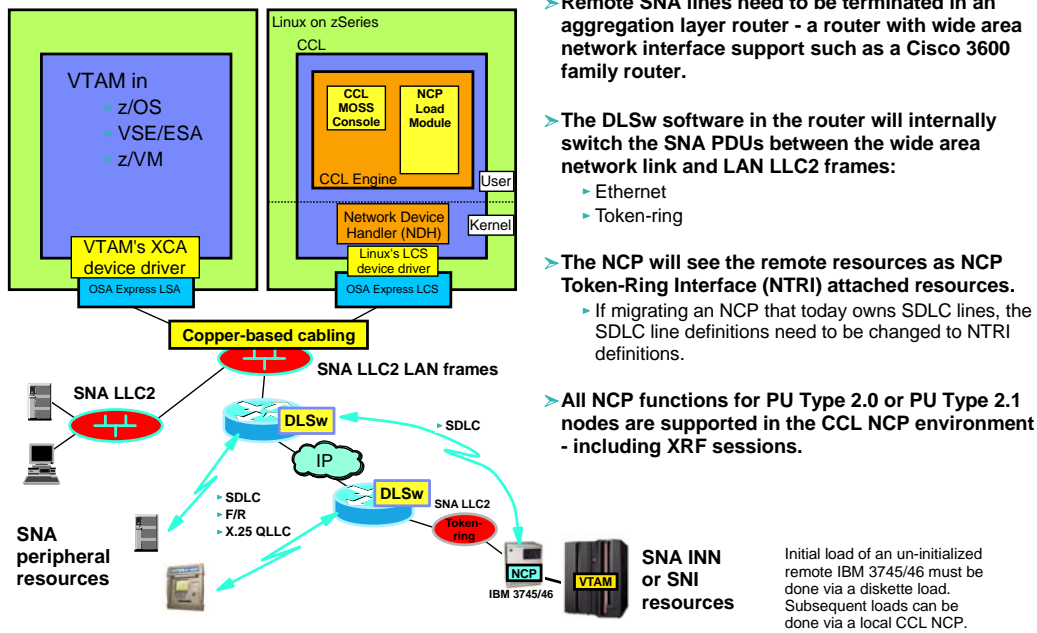
- In most cases no changes to SNA subarea pathing definitions

➤ None or minor changes to VTAM definitions and operations

➤ In most cases no changes to NetView definitions and operations

The only supported connectivity option between VTAM and CCL V1R1 is via a shared LAN to which VTAM is attached over an OSA LSA port, and CCL over an OSA LCS port.

CCL V1R1 and SNA connectivity - overview



> Remote SNA lines need to be terminated in an aggregation layer router - a router with wide area network interface support such as a Cisco 3600 family router.

> The DLSw software in the router will internally switch the SNA PDUs between the wide area network link and LAN LLC2 frames:

- ▶ Ethernet
- ▶ Token-ring

> The NCP will see the remote resources as NCP Token-Ring Interface (NTRI) attached resources.

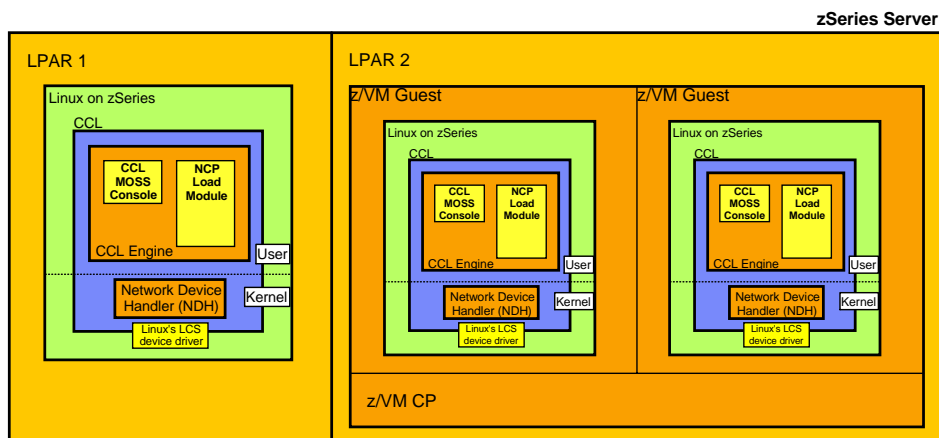
- ▶ If migrating an NCP that today owns SDLC lines, the SDLC line definitions need to be changed to NTRI definitions.

> All NCP functions for PU Type 2.0 or PU Type 2.1 nodes are supported in the CCL NCP environment - including XRF sessions.

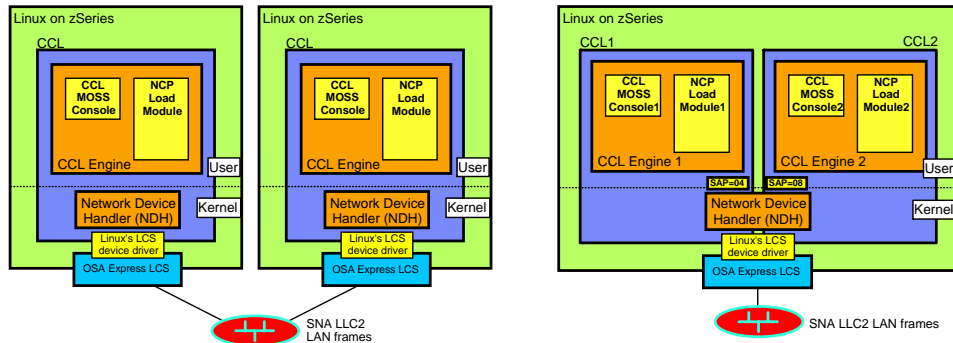
Linux deployment model

> Linux deployment:

- ▶ LPAR mode - one Linux image in an LPAR (no requirement for z/VM)
 - If you need a single or two CCL NCPs, and have no z/VM skills - this may be the option to choose
- ▶ As a z/VM guest
 - If you need more CCL NCPs and/or have z/VM skills, this may be a better option to choose



Deployment models for multiple NCPs



- **One CCL NCP per Linux image**
 - ▶ Each NCP operates completely independent of other NCPs
 - ▶ Each NCP may run in Linux images that are guests under z/VM or individual LPARs
- **Multiple CCLs per Linux image**
 - ▶ One NCP per CCL
 - ▶ Each CCL has its own MOSS console
 - ▶ Multiple CCLs share one NDH instance
- **No major difference in throughput**
- **Two Linux images may require more overhead (DASD, Memory, CPU) than one**
- **One Linux image may become a single point of failure**

CCL use of OSA copper ports (OSE - LCS mode)

- **CCL V1R1 exchanges SNA network flows with the network over a Linux LCS device driver interface only:**
 - ▶ For NCP to VTAM communication (VTAM attached to shared LAN via an OSA LSA port)
 - ▶ For downstream communication where aggregation layer routers switch SNA PDUs to/from wide area network connections or over IP networks (DLSw)
- **Only OSA copper-based interface ports can be configured as LCS ports - not fiber-based ports**
- **OSA/SF is needed for locally administered MAC addresses on OSA ports and for maintenance of the OSA Address Table (OAT) when sharing OSA LCS ports between multiple Linux images**
 - ▶ Locally administered MAC addresses can alternatively be set via the Hardware Management Console (HMC)

| Processor type | Required MCL level | Ethernet (OSA-Express FCs) | Token-Ring (OSA-2 and OSA-Express FCs) |
|----------------|--------------------|--|--|
| G5/G6 | | OSA-Express FC 2340 Fast Ethernet (10/100 Mbps) - 1 port/feature | OSA2 ENTR card FC 5201 (4/16 Mbps) - 2 ports/feature (each port can also be configured as a 10 Mbps Ethernet port) |
| z/800 or z/900 | 3.5 | OSA-Express FC 2366 Fast Ethernet (10/100 Mbps) - 2 ports/feature | OSA-Express FC 2367 (4/16/100 Mbps) - 2 ports/feature |
| z/890 or z/990 | 5.50 | OSA-Express FC 1366 (upgraded) 1000BaseT (10/100/1000 Mbps) - 2 ports/feature | OSA-Express FC 2367 (4/16/100 Mbps) - 2 ports/feature |
| z9-109 | | OSA-Express FC 2366 (carried forward) Fast Ethernet (10/100 Mbps) - 2 ports/feature OSA-Express FC 1366 (carried forward) 1000BaseT (10/100/1000 Mbps) - 2 ports/feature OSA-Express2 FC 3366 1000BaseT (10/100/1000 Mbps) - 2 ports/feature | Token-ring connectivity is not supported on a z9-109 |

OSA-Express copper connectivity overview

| Feature | Feature Name | Ports | z800 z900 | z900 z990 | z9-109 | CHPIDs | Connectors |
|---------|-----------------------------------|-------|--------------|--------------|--------|----------------------|------------------|
| 5201 | OSA-2 Token Ring | 2 | X | N/A | N/A | OSA | Copper, RJ-45 |
| 5202 | OSA-2 FDDI | 1 | X | N/A | N/A | OSA | Fiber, SC Duplex |
| 2362 | OSA-E 155 ATM SM | 2 | X | RPQ | N/A | OSD, OSE | Fiber, SC Duplex |
| 2363 | OSA-E 155 ATM MM | 2 | X | RPQ | N/A | OSD, OSE | Fiber, SC Duplex |
| 2364 | OSA-E GbE LX | 2 | X | C | C | OSD | Fiber, SC Duplex |
| 2365 | OSA-E GbE SX | 2 | X | C | C | OSD | Fiber, SC Duplex |
| 2366 | OSA-E Fast Ethernet | 2 | X | C | C | OSD, OSE | Copper, RJ-45 |
| 2367 | OSA-E Token Ring | 2 | X | X | N/A | OSD, OSE | Copper, RJ-45 |
| 1364 | OSA-E GbE LX | 2 | 09/04 | 06/03 | C | OSD | Fiber, LC Duplex |
| 1365 | OSA-E GbE SX | 2 | 09/04 | 06/03 | C | OSD | Fiber, LC Duplex |
| 1366 | OSA-E 1000BASE-T Ethernet | 2 | N/A | 06/03 | C | OSC, OSD, OSE | Copper, RJ-45 |
| 3364 | OSA-E2 GbE LX | 2 | N/A | 01/05 | X | OSD, OSN * | Fiber, LC Duplex |
| 3365 | OSA-E2 GbE SX | 2 | N/A | 01/05 | X | OSD, OSN * | Fiber, LC Duplex |
| 3366 | OSA-E2 1000BASE-T Ethernet | 2 | N/A | N/A | X | OSC, OSD, OSE, OSN * | Copper, RJ-45 |
| 3368 | OSA-E2 10 GbE LR | 1 | N/A | 01/05 | X | OSD | Fiber, SC Duplex |

LX = Long wavelength transceiver, SX = Short wavelength transceiver, LR = Long Reach transceiver
 X = Available for ordering, C = Carry forward on an upgrade from z900 or z990
 * OSN is exclusive to z9-109. Hardware availability is 09/16/05

CCL implementation and NCP migration planning

CCL project outline

➤ Make a physical inventory of your current Communication Controller environment

- ▶ Communication Controller model, size, features, line interfaces, LAN interfaces, etc.

➤ Make a logical and functional inventory

- ▶ NCP related functions
 - Boundary function lines, INN lines, SNI lines
 - Use of duplicate TIC MAC addressing for availability and scalability
 - XRF, NRF
 - NTuneMON, NPA-LU
- ▶ Functions that are not supported by CCL V1R1, such as NTO, XI, NSI, and NSF
- ▶ Network Node Processor functions (3746-900 or 3746-950)

Refer to *IBM Communication Controller Migration Guide, SG24-6298* appendix A and B for inventory worksheets.

➤ Reconcile and optimize

- ▶ Identify hardware and software components that are no longer used
- ▶ Remove hardware components that are no longer used (can reduce both maintenance cost and NCP Tier pricing)
- ▶ Clean up NCP definitions accordingly

➤ Controller consolidation and migration strategy planning

- ▶ Define high availability strategy - levels of redundancy and switchover capabilities
- ▶ Identify workloads that could be moved off SNA wide area networking via SNA/IP integration technologies
- ▶ Which NCPs to move to CCL and in which order
- ▶ Which NCPs to consolidate when moving to CCL NCPs and in which order
- ▶ Which NNP or MAE functions to migrate to CSL (Communications Server for Linux on zSeries)
- ▶ Which remaining functions to consolidate into fewer Communication Controller footprints

CCL project outline (continued)

➤ Overall CCL environment design for high availability

- ▶ Number of Linux images and number of CCL NCPs required to support strategy
- ▶ Linux and CCL NCP deployment from a data center and CEC perspective - LPARs or z/VM, which CCL NCP goes where
- ▶ Linux and CCL availability design - management and recovery procedures and tools
- ▶ Network availability and load balancing through duplicate MAC support for token-ring or Ethernet LAN connectivity to CCL NCPs (Ethernet LAN requires additional design of DLSw components)
- ▶ Wide area network connectivity through aggregation layer routers - how many, what type of WAN interfaces, how to provide redundancy for WAN termination if required
 - Consider optimization opportunities by terminating WAN lines in remote locations that today are already connected through an IP backbone to the data center - using DLSw technology over the IP backbone
- ▶ LAN infrastructure changes - token-ring and/or Ethernet, how to interconnect
- ▶ zSeries hardware requirements: IFLs, memory, DASD, OSA ports
- ▶ Physical LAN cabling between OSA ports, switches, and aggregation layer routers
- ▶ Define any changes business partners may have to implement (depends on migration strategy)
- ▶ Define any changes to peripheral SNA link stations (depends on migration strategy)

CCL provides improved options for redundancy design: you don't need to buy an extra IBM 3745 to deploy a stand-by NCP.

CCL project outline (continued)

➤ Establish a test environment

- ▶ One or two Linux on zSeries systems with CCL and test NCPs
- ▶ Experiment, test, learn - make sure to include recovery scenarios in the test activities

➤ CCL detailed migration planning per CCL NCP

- ▶ NCP migration strategy
 - Deactivate old NCP subarea and activate new NCP with same subarea in CCL (this requires the least amount of NCP changes and allows reuse of existing TIC MAC addresses by OSA ports)
 - Keep old NCP subarea active, activate new NCP with new subarea in CCL, and migrate resources over time to new NCP (requires changes to SNA subarea path definitions and may prevent you from reusing existing TIC MAC addresses in the new environment)
- ▶ WAN connectivity migration strategy
 - If the existing IBM 3745/46 has TIC interfaces, a migration of WAN lines to aggregation layer routers could be considered before moving the NCP to CCL (simplifies the move to CCL)
 - Otherwise the move of WAN lines to aggregation layer routers cannot start until the NCP has been activated in the CCL
- ▶ Fall back planning for each planned step

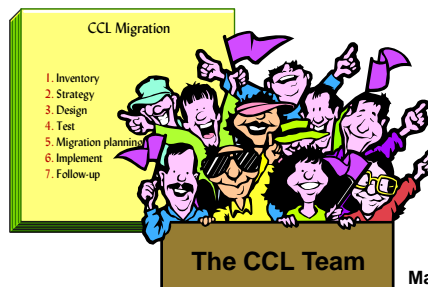
CCL project outline (continued)

➤ CCL implementation

- ▶ Establish planned infrastructure (Linux images, CCLs, OSA ports, cabling, switches, etc.)
- ▶ Migrate one NCP at a time according to detailed plan

➤ Consolidation of remaining IBM 3745/46 resources

- ▶ Functions not supported by CCL should be consolidated into fewer and smaller IBM 3745/46s
- ▶ Clean up NCPs and associated licenses for old environment



Make sure you have a plan before you start!

CCL use of OSA copper ports in LCS mode

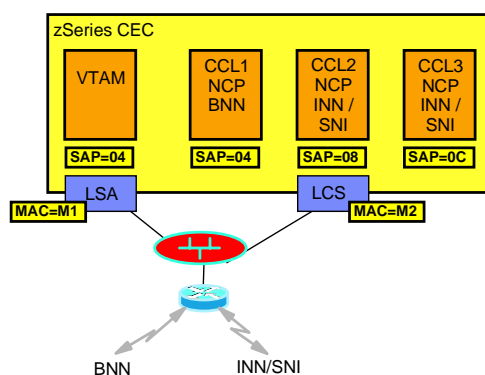
CCL V1R1 use of OSA ports and MAC addresses

- **This section focuses on CCL usage of OSA copper ports operating in LCS mode, which is the only SNA LAN connectivity option CCL V1R1 supports.**
- **The next release of CCL will support use of OSA Express fiber ports operating in QDIO layer-2 mode.**
 - ▶ That support will significantly improve the sharing capabilities of OSA ports, and it will reduce (and in some cases remove) the requirements for copper-based cabling
 - ▶ QDIO layer-2 mode is supported on z890, z990, and z9-109 only.
 - ▶ If you implement on earlier hardware platforms, you still need to use OSA copper-based ports and cabling.
- **Even with the next release, there will be scenarios where use of OSA copper port in LCS mode continues to be required - such as when implementing CCL on a z800 or a z900 that do not support QDIO Layer 2 mode**
- **VTAM will also continue to require an OSA copper port operating in LSA mode if VTAM does not reside on the same System z9 CEC as where CCL resides.**

OSA port usage by CCL V1R1 - summary of the rules

- Only OSA copper-based ports can be used by CCL R1 for SNA traffic - configured in LCS mode
- A Linux image can use the same OSA LCS port for SNA and IP access
- The physical LAN may be either token-ring or Ethernet IEEE802.3
- Two VTAMs can share an LSA port as long as they use two unique local SAP numbers (for example SAP 04 and SAP 08)
- VTAM and CCL cannot share an OSA port for communication between them - VTAM's LSA port cannot be the same as CCL's LCS port
- Two CCL NCPs cannot share an OSA LCS port for BNN traffic
 - The NCP uses local SAP 04 and SAP C8 (HPR) for peripheral node communication. These SAP numbers cannot be overridden in the NCP definitions. BNN traffic to/from a CCL NCP must go to/come from SAP 04 and SAP C8 for HPR
- One BNN NCP and one or more INN/SNI NCPs can share an LCS port using different local SAPs for the INN/SNI traffic
 - The NCP allows overriding the local SAP number for INN/SNI traffic
 - To share an LCS port between two Linux images, OSA/SF must be used to create an OAT (OSA Address Table) that indicates which Linux image a SAP number belongs to.
 - If the OSA LCS port is also used for IP access to Linux, remember also to add the home IP addresses of each Linux image to the OAT
- Two INN/SNI NCPs can establish a subarea link between them using a shared OSA port as long as they code UNIQUE=NO on the corresponding subarea PU

CCL V1R1 - how many OSA ports? One VTAM, one CCL BNN NCP, and two CCL INN/SNI NCP



➤ Outline:

- One VTAM
- One CCL BNN NCP
- Two CCL INN/SNI NCP

➤ Rules:

- Two INN/SNI NCPs can share an LCS port using different SAPs
 - For subarea links, Source SAP can be coded on the subarea PU statement
- Two INN/SNI NCPs can establish a subarea link between them using a shared OSA port as long as they code UNIQUE=NO on the corresponding subarea PU
 - CCL1, CCL2, and CCL3 in this configuration can have INN links defined between them if they code UNIQUE=NO on their corresponding subarea PU statements
 - otherwise they need separate OSA ports

Two OSA copper ports

Note:

The example shows all three CCLs in the same zSeries CEC. Separating them into separate zSeries CECs would mean each of them would need its own OSA LCS port.

For availability reasons, you need to consider doubling the number of ports for an implementation without single points of failure.

MAC addressing when using SNA over a LAN

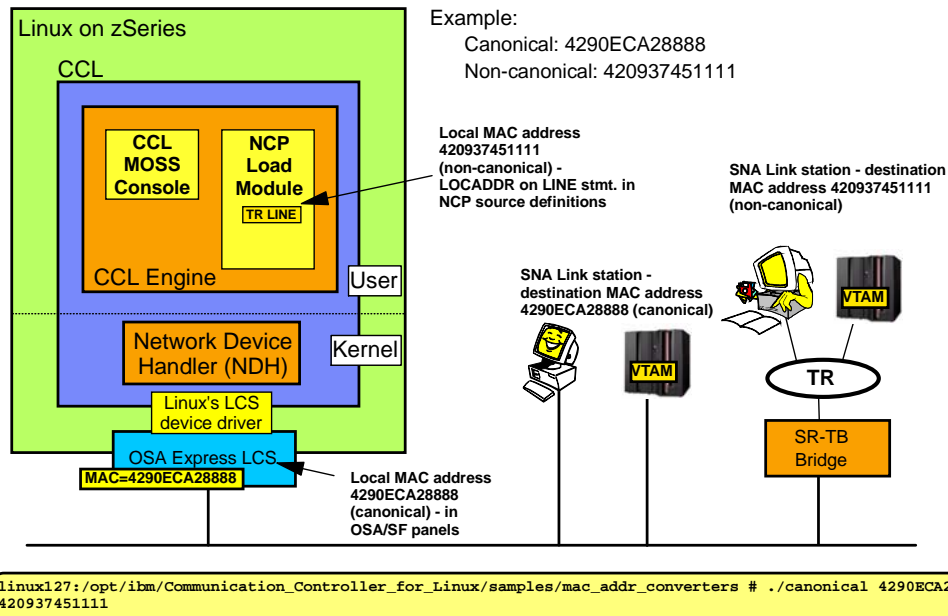
MAC address formats - token-ring (non-canonical) and Ethernet (canonical)

> The NCP sees all LAN interfaces as being token-ring

- ▶ A token-ring MAC address is in the non-canonical form and this form is what must be coded in the NCP generation deck.
- ▶ The NCP requires locally administered MAC addresses
 - MAC addresses starting with B'x1xx xxxx'
- ▶ If the OSA port is token-ring, then the MAC address in the NCP and in OSA/SF for the OSA port match
- ▶ If the OSA port is Ethernet, then the MAC address in the NCP must be the non-canonical form of the Ethernet canonical MAC address as specified in OSA/SF
- ▶ Canonical is little-endian, while non-canonical is big-endian
- ▶ A utility is provided with CCL to assist in the conversion
 - Canonical
 - Canonical.cmd (REXX version)

| | | | | | | |
|------------------------------------|----------|----------|----------|----------|----------|----------|
| Canonical address (Ethernet) | 08 | 00 | 3f | e1 | 4d | a8 |
| Binary | 00001000 | 00000000 | 00111111 | 11100001 | 01001101 | 10101000 |
| Reverse bits in each byte | 00010000 | 00000000 | 11111100 | 10000111 | 10110010 | 00010101 |
| Non-canonical version (token-ring) | 10 | 00 | fc | 87 | b2 | 15 |

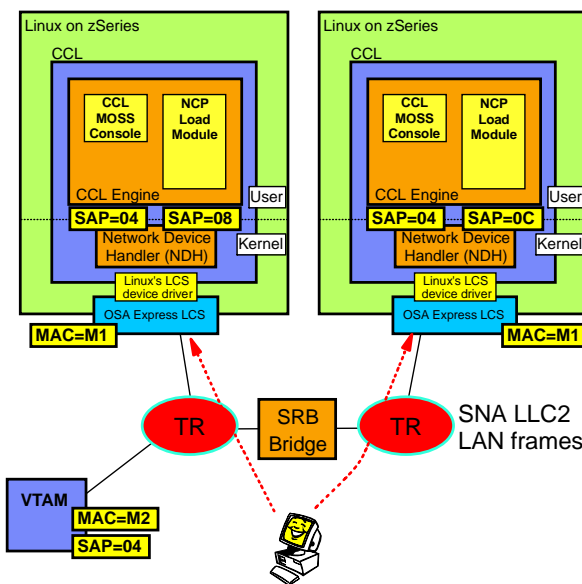
Ethernet considerations - canonical or non-canonical MAC address



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Duplicate token-ring MAC addresses in two CCL NCPs for NCP availability and load balancing



➤ Load balancing of remote SNA link station access when both CCL instances are up and running

- Each NCP needs unique SAP for VTAM links - VTAM needs to know them as separate NCPs (SAP 08 and 0C in this setup)
- The two NCPs need the same SAP for downstream links (SAP 04 in this setup)

➤ Availability

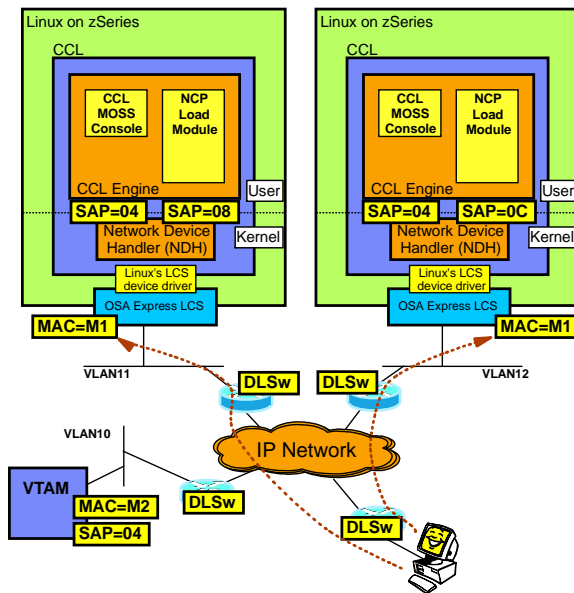
- If an LCS port, a Linux image, an NCP, or a CCL engine goes down, remote link stations can recover over the other CCL instance
 - As usual in an SNA network, such a switch is disruptive to SNA sessions (subarea and APPN)
 - SNA sessions over HPR will survive such a switch
- Traditional availability aspects would direct one towards two Linux images on two different zSeries CECs in two different data centers for maximum availability



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

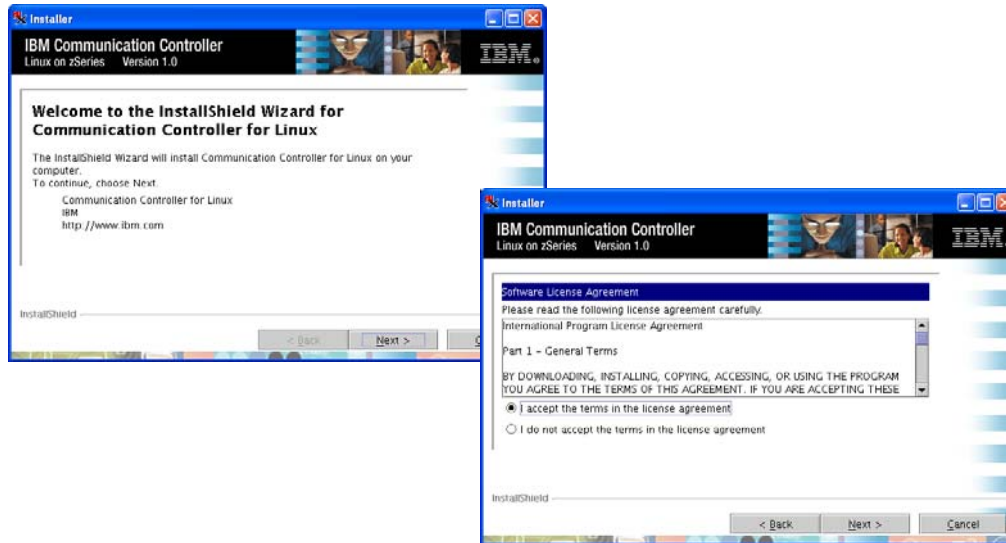
Duplicate Ethernet MAC addresses in two CCL NCPs for NCP availability and load balancing



- **Ethernet segments with duplicate MAC addresses cannot be bridged together**
 - Ethernet bridging technology doesn't support that
- **Elements of the DLSw technology can instead be used to connect the Ethernet segments with duplicate MAC addresses together and with other LAN segments**
 - DLSw can be configured to use round-robin towards the two Ethernet segments with duplicate MAC addresses - allowing an even spread of SNA link station connections when both are available
- **Same considerations for VTAM links from the NCPs**
 - Communication between the NCPs and VTAM also need to use DLSw (typically local)
 - VLAN10 cannot be bridged to VLAN11 and VLAN12
 - If each CCL instance has a separate OSA port for VTAM connectivity, DLSw isn't needed for the VTAM communication

Installation of CCL

CCL InstallShield



The InstallShield takes you through a number of questions and eventually performs the installation.

Preparing to build the NDH isolation module

- **Example based on SUSE SLES8 (2.4.21-273 kernel level).**
- **Enter the following commands in a root user shell:**
 - ▶ `cd /usr/src/linux`
 - Position on the current Linux kernel level files
 - ▶ `make clean`
 - Purges any old build from the build directory
 - ▶ `make mrproper`
 - Removes any old build files
 - ▶ `make cloneconfig`
 - Creates a current `/usr/src/linux.config` file
 - ▶ `make dep`
 - Builds all kernel dependencies
 - This step may take several minutes
 - This is a *go-get-yourself-a-cup-of-coffee* command
- **All the make commands will produce messages being sent to your shell session. They can generally be ignored.**



Build the NDH isolation module and load the NDH modules

```
login as: root
Sent username "root"
root@linux127.tcp.raleigh.ibm.com's password:
Last login: Tue Jan 25 14:14:42 2005 from sig-9-49-154-253.mts.ibm.com
linux127:/usr/src/linux # uname -a
Linux linux127 2.4.21-273-default #1 SMP Mon Jan 17 13:17:09 UTC 2005 s390 unknown
linux127:/usr/src/linux # cd /opt/ibm/Communication_Controller_for_Linux
linux127:/opt/ibm/Communication_Controller_for_Linux # cd drivers

linux127:/opt/ibm/Communication_Controller_for_Linux/drivers # ./build.sh
No Linux Kernel source directory specified.
Using defaulting to directory: /usr/src/linux

cc -DLCSSSEQ1 -D__KERNEL__ -DMODULE -Wall -O2 -I/usr/src/linux/include -c -o cclndh.o cclndh.c

linux127:/opt/ibm/Communication_Controller_for_Linux/drivers # cd ..

linux127:/opt/ibm/Communication_Controller_for_Linux # ./load_ndh.sh
NDH kernel modules loaded. You are now able to run the cclengine
linux127:/opt/ibm/Communication_Controller_for_Linux #
```

Build the NDH
isolation module

Load the NDH modules
into the kernel

- This has changed a little since the original shipment
 - ▶ You now only need to run the load_ndh.sh script - it determine if NDH has been built and if not, it will automatically invoke the build.sh script before proceeding.

Directory structure for CCL engines

- CCL installs by default into:
 - ▶ /opt/ibm/Communication_Controller_for_Linux
 - Suggestion: consider changing that to something shorter during the InstallShield dialog - such as /opt/ibm/CCL
 - ▶ This is the directory where the **cclengine** load module resides and this is the directory you must position to before issuing the **cclengine** command.
 - ▶ A CCL engine is named when issuing the **cclengine** command. Each CCL engine name must exist as a subdirectory under the directory where the **cclengine** command resides.
 - ▶ Assuming we want to start two CCL engines - CCL1 and CCL2:
 - /opt/ibm/Communication_Controller_for_Linux/CCL1
 - /opt/ibm/Communication_Controller_for_Linux/CCL2
 - ▶ The NCP load modules must reside in the CCL1 and CCL2 directories.
 - Should be stored in uppercase name
 - ▶ Starting the CCL1 engine with an NCP module (NIA760G) located in the CCL1 subdirectory:
 - ./cclengine -m NIA760G -p 4000 CCL1 &

How to load an NCP load module into CCL over a LAN

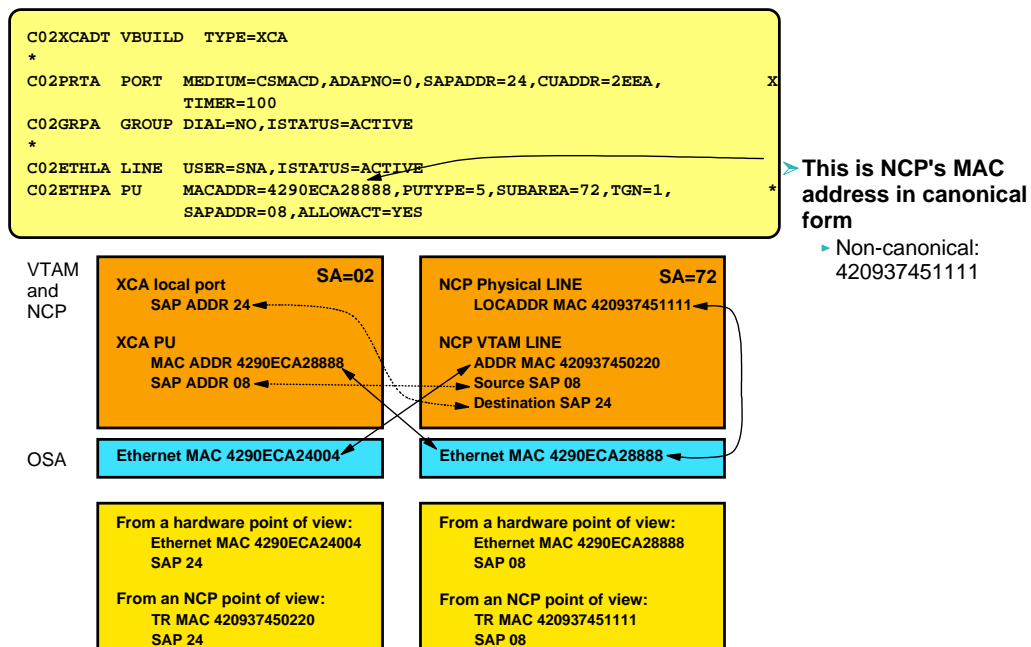
- The very first NCP load module must be manually transferred to Linux and loaded into the CCL via a shell command interface
 - ./cclengine -m<NCP load mod name> -p<MOSS port> <ccl engine name>
 - This process can be automated to be performed during IPL of Linux
- VTAM's XCA definitions need to be activated:
 - VARY net,ACT,ID=XCA_pu
- The NCP can then be activated from VTAM using a normal V NET,ACT,ID=<NCP name> command
 - VARY net,ACT,ID=NCPname
- The LOADFROM=HOST option is not supported by CCL over a LAN, but will be by CCL V1R2 when connecting to a CCL NCP over an OSA for NCP (OSN) CHPID
- The LOADFROM=EXTERNAL option is not supported for a CCL that is directly adjacent to VTAM
- NCP load modules on the MOSS disk can from then on be refreshed using the existing VTAM MODIFY LOAD commands to save a new NCP load module to the MOSS disk (a Linux file), and to schedule a timed IPL of the newly transferred NCP load module:
 - MODIFY net,LOAD,ID=NCPname,ACTION=ADD/REPLACE,LOADMOD=loadmod,IPLTIME=
 - MODIFY net,LOAD,ID=NCPname,ACTION=SETTIME,LOADMOD=loadmod,IPLTIME=



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Sample VTAM XCA major node for Ethernet connectivity



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Sample NCP definitions for Ethernet connectivity

```

C72PTRG1 GROUP ECLTYPE=(PHY,ANY),ADAPTER=TIC2,ANS=CONT,MAXTSL=16732, X
RCVBUFC=32000,USSTAB=AUSSTAB,ISTATUS=ACTIVE,XID=NO, X
RETRIES=(20,5,5)
*-----*
* Physical Ethernet LINE - BNN and INN
*-----*
C72TR88 LINE ADDRESS=(1088,FULL),TRSPEED=16,PORTADD=88, X
LOCADD=420937451111
C72PU88A PU
*****
* NTRI BNN LOGICAL LINES FOR TOKEN RING PORT 1088 *
*****
C72BNNG1 GROUP ECLTYPE=LOGICAL,ANS=CONTINUE,AUTOGEN=200,CALL=INOUT, X
ISTATUS=ACTIVE,PHYSRSC=NONE, X
USSTAB=AUSSTAB,RETRIES=(10,10,10,20),XMITDLY=NONE, X
MODETAB=AMODETAB
*****
* NTRI INN LOGICAL LINES FOR TOKEN RING PORT 1088 *
*****
C72INNG1 GROUP ECLTYPE=(LOGICAL,SUBAREA),ANS=CONT,PHYSRSC=C72PU88A, X
LOCALTO=13.5,REMOTTO=18.2,T2TIMER=(0.2,0.2,3), X
ISTATUS=ACTIVE,SOLCST=(C72PRI,C72SEC),MONLINK=CONT
*-----*
* Linkstation to VTAM NETC.C02N
*-----*
C72LG2A LINE TGN=1,TGCONF=SINGLE
C72PG2A PU ADDR=18420937450220,SSAP=(08,W)

```

NCP name in this example is C72DUPE

> This is the NCP MAC address in non-canonical form

▶ Canonical: 4290ECA28888

> The first two digits is VTAM's SAP number in hexadecimal

▶ 'X'18' = SAP 24

> The remaining digits is the non-canonical form of VTAM's MAC address

▶ Canonical: 4290ECA24004

> For the subarea link to VTAM, we will use local SAP 08



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Sample VTAM activation and display commands

```

v net,act,id=c02xcadt,all
IST093I C02XCADT ACTIVE
IST464I LINK STATION C02ETHPA HAS CONTACTED C72DUPE SA 72
IST093I C02ETHPA ACTIVE

v net,act,id=c72dupe,all
IST093I C72DUPE ACTIVE
IST093I C72PU88A ACTIVE
IST464I LINK STATION C72PG2A HAS CONTACTED C02NPU SA 2
IST093I C72PG2A ACTIVE

D NET,ID=C02ETHLA,E
IST097I DISPLAY ACCEPTED
IST075I NAME = C02ETHLA, TYPE = LINE 592
IST486I STATUS= ACTIV---E, DESIRED STATE= ACTIV
IST087I TYPE = LEASED, CONTROL = SDLC, HPDT = *NA*
IST134I GROUP = C02GRPA, MAJOR NODE = C02XCADT
IST1500I STATE TRACE = OFF
IST1656I VTAMTOPO = REPORT, NODE REPORTED - YES
IST1657I MAJOR NODE VTAMTOPO = REPORT
IST396I LNKSTA STATUS CTG GTG ADJNODE ADJSA NETID ADJLS
IST397I C02ETHPA ACTIV--W-E 1 1 C72DUPE 72 NETC
IST314I END

```

The 'W' flag indicates that ALLOWACT=YES was coded on this PU in VTAM's XCA major node.



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Introducing CCL V1R2

Where did CCL V1R1 work well?

➤ **SNA installations that had managed to move most SNA traffic off the IBM 3745/46 environment, but had a few SNI partners left that were unable to move away from SNI.**

- ▶ Example is USI, a bank in Italy: *"USI considers CCL V1R1 to be a good program innovation as it enables us to replace our old IBM 37xx SNI connections easily as no changes are required to our business partners' network. CCL has also reduced the number of NCP definitions and complexity of our NCP configuration."*

➤ **SNA installations that wanted to move away from old hardware, such as IBM 3745/46, token-ring infrastructure, and older channel-attached SNA technologies.**

- ▶ Example is National City, a bank in the US: *"Our token-ring SNI connections are the last remnants of the token-ring environment at National City and the last remaining connections to our IBM 3745/46s. Moving to CCL will save the the cost of IBM 3745/46 hardware and the maintenance cost of the token-ring switch hardware. In addition, there are power, floor space, and rack space savings."*
- ▶ Other examples are various installations who are consolidating connections that are not supported by CCL, such as EP-based connections, onto a single IBM 3745 and migrating other connections that are supported by CCL to a CCL environment in order to eliminate the majority of their IBM 3745/46s.



Where did CCL V1R1 work well? (continued)

➤ SNA installations setting up disaster recovery sites.

- ▶ An example is an SNA installation on the US west coast who is looking at setting up a DR site on the east coast. They had originally assumed they needed to buy additional IBM 3745/46 equipment, but are now looking at using CCL as an alternative.

➤ SNA installations with floor space issues.

- ▶ Migrating NCPs from IBM 3745/46s to CCL can save both floor space and power consumption.



Where did CCL V1R1 not work so well?

➤ SNA installations that had multiple types of connections that were not supported by CCL V1R1.

- ▶ There is still a large amount of non-SNA X.25 connections in use. This is especially the case in selected geographies, where X.25 networks were promoted by the telecom industry.

➤ TPF installations. TPF cannot use an OSA LSA port to communicate with a CCL NCP. TPF only supports a channel-attached NCP.

- ▶ For some SNA configuration scenarios, TPF requires an NCP for communication to VTAM.

➤ SNA installations that had many, highly utilized IBM 3745s.

- ▶ CCL V1R1 CPU requirements were in some of those cases prohibitive. If the IBM 3745 CCU was between 70% and 90% utilized, each migrated NCP would in general require one z990 IFL engine.
- ▶ This issue was somewhat addressed by a CCL V1R1 performance PTF August 2005.

➤ SNA installations that use many IBM 3745/46 TIC adapters.

- ▶ This was especially the case for large boundary function installations. In order to migrate those NCPs to CCL without requiring changes to the SNA link station definitions in the peripheral SNA nodes, there had to be a one-to-one mapping of LAN adapters (MAC addresses).
- ▶ OSA LCS interfaces do not provide enough virtualization capabilities to do so in an efficient manner, and in many cases those installations would need more physical OSA copper ports than what they had slots available for.



Highlights of CCL V1R2

➤ Provide CDLC-based channel connectivity between CCL and same-CEC SNA operating systems on System z9

- ▶ z/OS, z/VSE, and z/VM VTAM to CCL NCP
- ▶ TPF to CCL NCP

➤ Enable CCL usage of OSA-Express fiber optic ports instead of OSA copper-based ports

- ▶ IP-based encapsulation of SNA traffic (utilizing IP QDIO layer-3 network interfaces to Linux):
 - IP-Transmission Group (IP-TG) for direct IP connectivity (TCP connection) between two CCL NCPs
 - INN and SNI traffic between two CCL NCPs
- ▶ Native SNA LLC2 traffic over an OSA-Express fiber infrastructure:
 - OSA-Express QDIO layer-2 support for native SNA LLC2 traffic to/from a CCL NCP
 - INN, SNI, and boundary function traffic
 - MAC addresses are virtual with QDIO layer-2 support
 - Much improved sharing capabilities of one OSA port by many CCL NCPs

➤ Reduce CCL CPU requirements

- ▶ Some of these enhancements were PTFed back to CCL V1R1 in August, but more are included in CCL V1R2
 - Compiler optimization
 - More functions re-written to native zSeries and System z9 assembler

➤ Add support for NPSI non-SNA X.25 connectivity

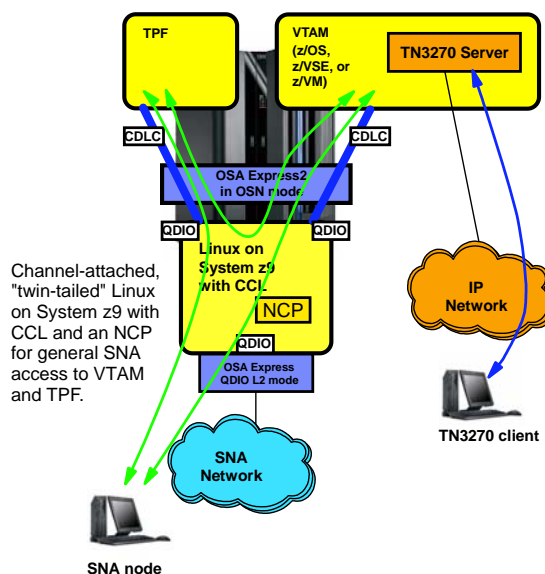
- ▶ Part of the NPSI X.25 connectivity solution is provided by a vendor product that is not included in CCL V1R2



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

CDLC channel connectivity between VTAM or TPF and CCL V1R2 on System z9



Channel-attached, "twin-tailed" Linux on System z9 with CCL and an NCP for general SNA access to VTAM and TPF.

➤ OSA Express2 on System z9 implements a new OSA CHPID type - known as OSA for NCP (OSN).

- ▶ OSA-Express2 1000Base-T and 1 Gigabit Ethernet features on z9-109

➤ TPF and VTAM see the OSA Express OSN port as a channel-attached IBM 3745 to which they communicate using the usual CDLC channel protocol.

- ▶ OSA CHPID defined as OSN
- ▶ TPF and VTAM device number defined as IBM 3745
- ▶ Linux device numbers (three in a set) defined as OSN (accessed through QDIO from Linux)
- ▶ Same OSN port can be shared among more SNA host systems and more CCL NCPs

➤ The OSA microcode relays the CDLC data over a QDIO interface to CCL, which presents it to the NCP as though it had arrived over an IBM 3746 channel interface.

➤ The fact that the NCP runs in CCL instead of an IBM 3745/46 is transparent to TPF and VTAM.

- ▶ Existing configuration definitions are used unchanged
- ▶ Existing activation and management flows continue to work as before

➤ The normal Load/Dump functions over a channel are fully supported

- ▶ No need to FTP an NCP load module to the Linux file system
- ▶ Loads very fast (less than 10 seconds)

➤ TPF or VTAM must reside on the same System z9 CEC as where CCL resides

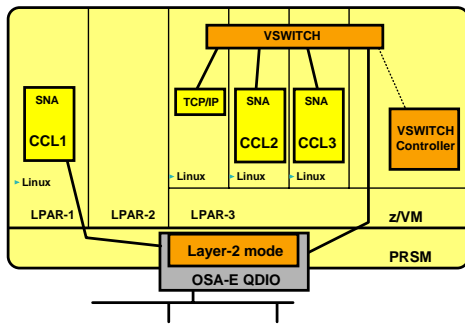
- ▶ This is a same-CEC connectivity technology



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

QDIO layer-2 for native SNA LLC2 traffic over QDIO interfaces



Fast Ethernet, 1000BASE-T Ethernet, Gigabit Ethernet, and 10 Gigabit Ethernet features on OSA-Express and OSA-Express2 on z890, z990, and z9-109

➤ QDIO layer-2 mode advantages from a Linux SNA perspective:

- ▶ Retain SNA LLC2 communication over an OSA port
- ▶ Supports fiber-optic Gigabit or 10 Gigabit network connectivity
 - Reuse existing fiber optic cabling and switch infrastructure
- ▶ No shortage on MACs for SNA access - they are virtual
 - Multiplex many SNA link stations over one physical network interface
- ▶ With QDIO layer-2 support, the CCL V1R1 restrictions for sharing an OSA port across multiple boundary function NCPs are all gone.

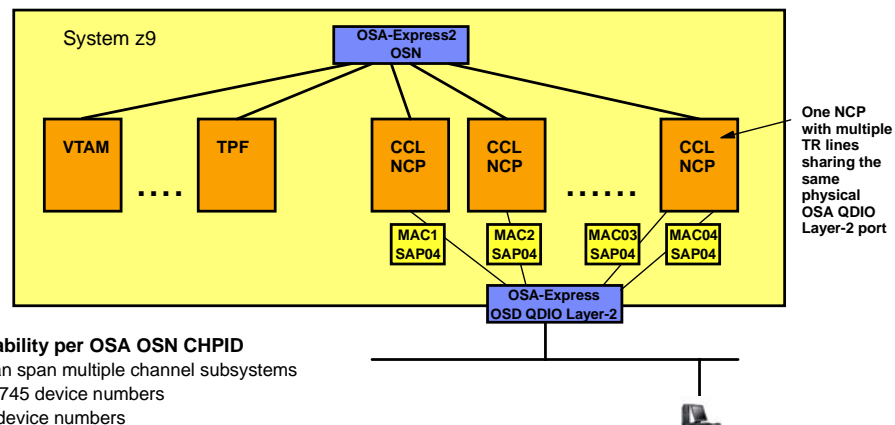
➤ QDIO in layer-2 mode is network protocol (layer-3) agnostic, which allows it to handle traffic for any network protocol - such as NetBIOS, SNA, IPX, IPv4, and IPv6.

➤ Each endpoint is identified by a Media Access Control (MAC) address, which in this case is a virtual MAC address that is assigned by the QDIO device driver in operating systems that support QDIO layer-2 mode (which at this time is Linux on System z9 or zSeries only as well as the z/VM virtual switch).

➤ QDIO layer-2 mode is supported both in native LPAR and under z/VM.

- ▶ QDIO layer-2 mode is supported by CCL V1R2 when running on a Linux 2.6 kernel only.

OSA for NCP and QDIO Layer-2 scalability



➤ OSA for NCP scalability per OSA OSN CHPID

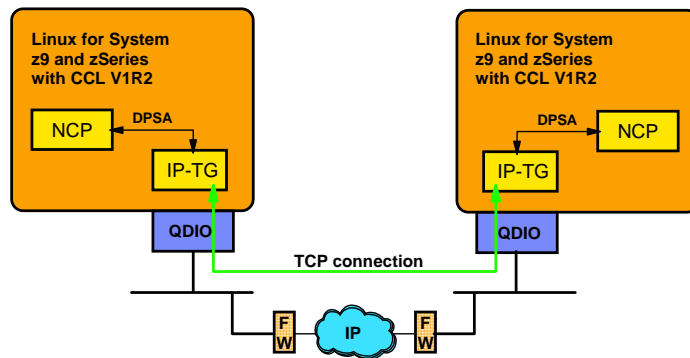
- ▶ OSN CHPIDs can span multiple channel subsystems
- ▶ Up to 180 IBM 3745 device numbers
- ▶ Up to 480 OSN device numbers
 - Accessed from Linux as QDIO device groups
 - QDIO device groups are linked to IBM 3745 device numbers via a concept known as Channel Connection Identifiers (CCID)
 - Multiple IBM 3745 device numbers may be mapped to a single QDIO device group when the NCP is connected to multiple VTAMs or TPF systems

➤ OSA Layer-2 scalability per OSA OSD CHPID

- ▶ Up to 2048 virtual MAC addresses
- ▶ Up to 1920 QDIO device numbers (Each QDIO device group is an NCP LAN interfaces - all potentially using the standard SNA service access point (SAP) number of 04)

➤ Two LPARs can share an OSD CHPID where one LPAR uses the port in traditional QDIO Layer-3 mode and the other uses it in QDIO Layer-2 mode - but they cannot communicate with each other over that shared OSA port

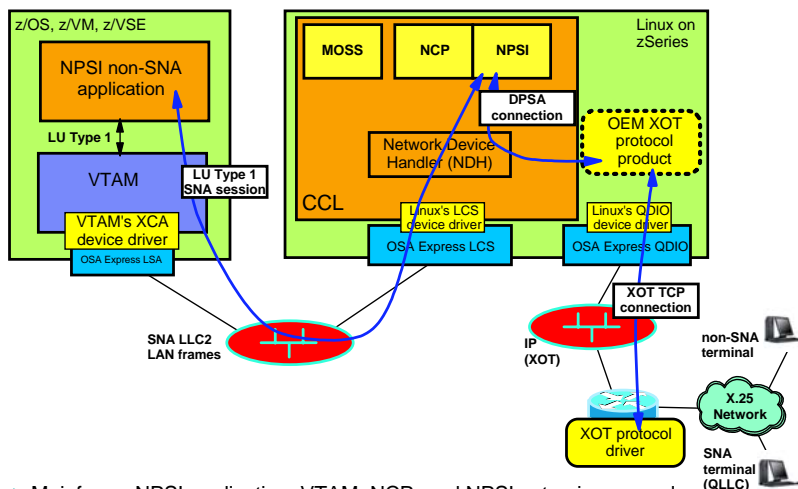
IP transmission group for INN/SNI traffic between two CCL V1R2 NCPs



IP-TG in combination with the CDLC connectivity to VTAM provides up to 6 times better throughput than two IBM 3745 INN/SNI NCPs connected via token-ring.

- > IP transmission group exchanges INN/SNI traffic between two CCL V1R2 NCPs over a TCP connection.
- > The NCP sees the IP-TG endpoint as a TIC3 token-ring adapter
 - TIC3 adapters normally reside in the IBM 3746 frame and are attached to Token-Ring Processors (TRP)
 - A TRP in a real IBM 3746 does all the SNA LLC2 processing on behalf of the NCP
 - No LLC2 overhead in the NCP
 - The NCP interfaces to the TRP using the Dynamic Parameter Status Area (DPSA) programming interface
 - Used by the NCP for all line and channel resources that are located in an IBM 3746
- > Because there is no real LLC2 processing when using IP-TG, IP-TG performs very well for INN/SNI traffic between two CCL V1R2 NCPs.
- > The IP-TG TCP connection can optionally be secured (SSL) using the STUNNEL technology of Linux.
- > Configuration options allow for control of port numbers at both endpoints for easier firewall configuration between business partners.

CCL V1R2 - NPSI non-SNA X.25 support overview



XOT is an open standard and defined in RFC 1613 "Cisco Systems X.25 over TCP (XOT)"

The XOT protocol support on Linux on System z9 and zSeries will not be provided by IBM as part of the CCL offering itself. It will be provided by other vendors as a separately priced feature.

- > Mainframe NPSI application, VTAM, NCP, and NPSI setup is as usual.
- > NPSI processing remains offloaded from the mainframe OS environment.
- > Physical connectivity to X.25 network is via an aggregation layer router.
- > Connectivity between aggregation layer router and NPSI is via an X.25 Over TCP/IP (XOT) TCP connection (IP network flows).
- > Interface between NPSI and local XOT protocol component is the same as NPSI uses today when communicating over X.25 adapters in an IBM 3746 unit - the Dynamic Parameter Status Area (DPSA) interface.

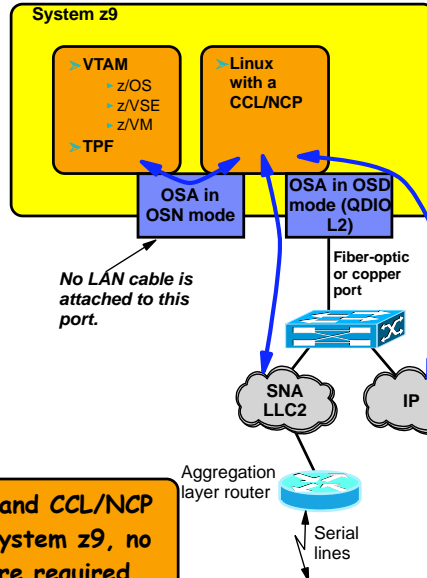
CCL V1R2 physical connectivity on System z9

> VTAM or TPF connectivity to a same-CEC CCL/NCP

- Shared OSA port configured in OSN mode
- VTAM, TPF and NCP see the connectivity as a traditional ESCON CDLC channel

> VTAM connectivity to a CCL/NCP in another CEC:

- VTAM OSA LSA copper port to shared LAN



> CCL/NCP downstream connectivity via OSA port configured as OSD and operated in QDIO layer-2 mode:

- SNA LLC2
 - Continues to also be supported over an OSA port configured in LCS mode
- SNA over IP:
 - IP-TG
 - XOT

> Multiple SNA linkstations per physical OSA port:

- Each linkstation configured using a separate virtual MAC address

> Alternative IP connectivity:

- Via another OSA port configured as OSD and operated in QDIO layer-3 mode

When VTAM or TPF and CCL/NCP reside in the same System z9, no OSA copper ports are required.

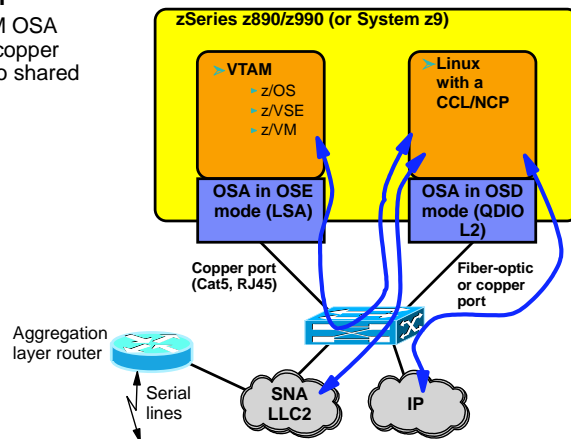
Redbooks © Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

CCL V1R2 physical connectivity on zSeries z890/z990

> VTAM connectivity to CCL/NCP

- VTAM OSA LSA copper port to shared LAN



> CCL/NCP connectivity via OSA port configured as OSD and operated in QDIO layer-2 mode:

- SNA LLC2
 - Continues to also be supported over an OSA port configured in LCS mode
- SNA over IP:
 - IP-TG
 - XOT

> Multiple SNA linkstations per physical OSA port:

- Each linkstation configured using a separate virtual MAC address

> Alternative IP connectivity:

- Via another OSA port configured as OSD and operated in QDIO layer-3 mode

VTAM needs one OSA copper port. CCL/NCP does not require any OSA copper ports.

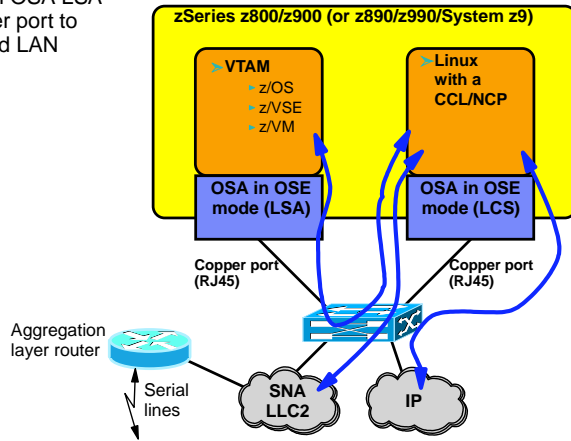
Redbooks © Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

CCL V1R2 physical connectivity on zSeries z800/z900

> VTAM connectivity to CCL/NCP

- ▶ VTAM OSA LSA copper port to shared LAN



> CCL/NCP connectivity via OSA copper port configured as OSE and operated in LCS mode:

- ▶ SNA LLC2
- ▶ SNA over IP:
 - IP-TG
 - XOT

> Multiple SNA linkstations per physical OSA port:

- ▶ Each linkstation must be configured using the same MAC address and separate SAP numbers

> Alternative IP connectivity:

- ▶ Via an OSA port configured as OSD and operated in QDIO layer-3 mode

**VTAM needs one OSA copper port.
CCL/NCP requires at least one OSA copper port.**

CCL V1R2 - connectivity summary

> IP traffic

- ▶ XOT
- ▶ IP-TG

> SNA LLC2

- ▶ Native SNA over the LAN

- SNA PU Type 4 NCP (running in a CCL)

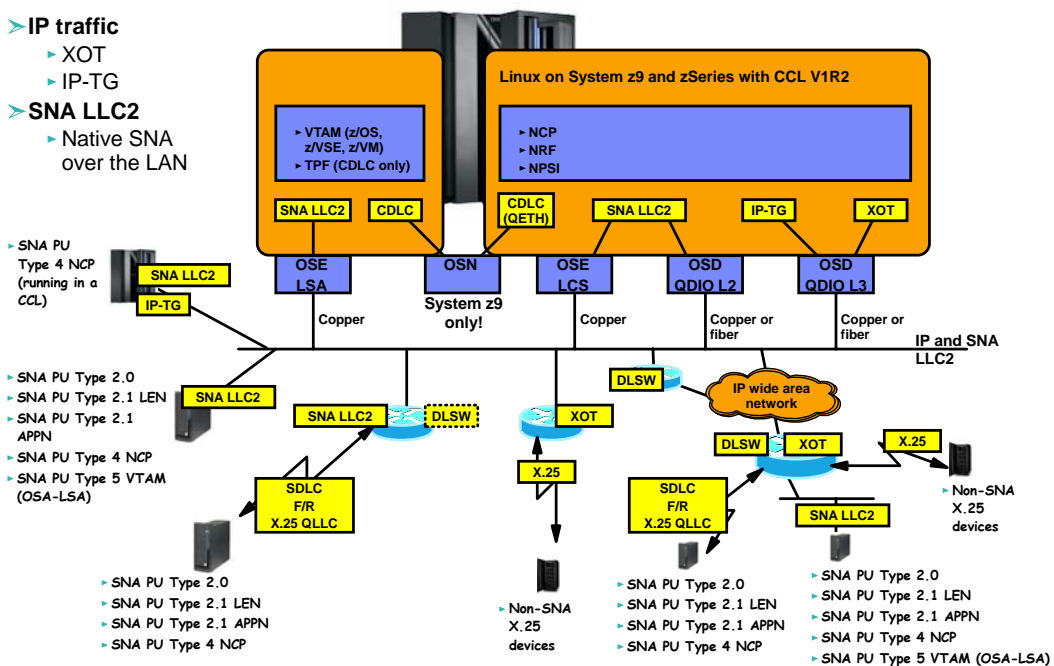
- SNA PU Type 2.0
- SNA PU Type 2.1 LEN
- SNA PU Type 2.1 APPN
- SNA PU Type 4 NCP
- SNA PU Type 5 VTAM (OSA-LSA)

- SNA PU Type 2.0
- SNA PU Type 2.1 LEN
- SNA PU Type 2.1 APPN
- SNA PU Type 4 NCP

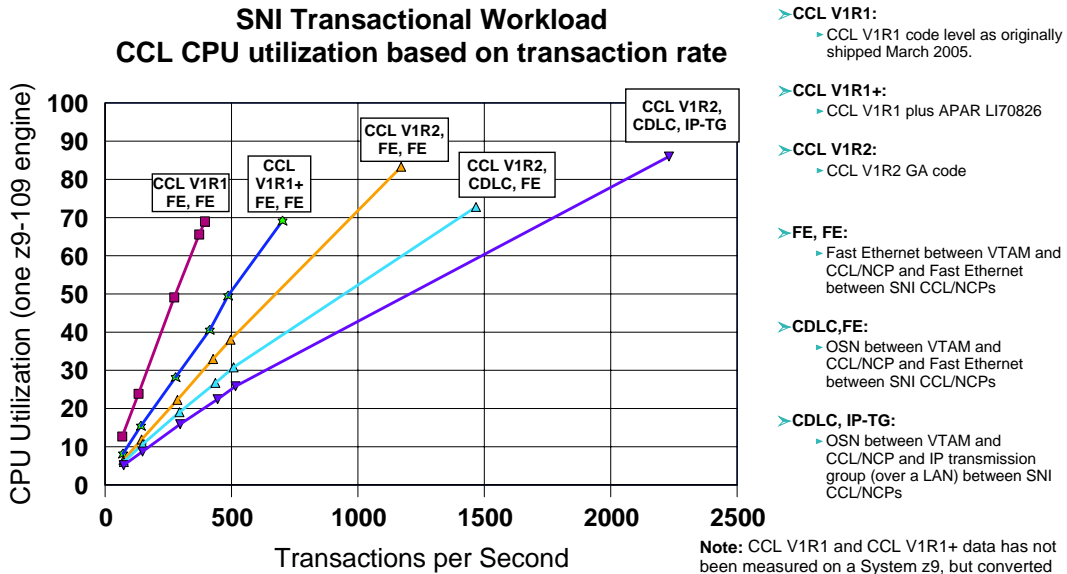
- Non-SNA X.25 devices

- SNA PU Type 2.0
- SNA PU Type 2.1 LEN
- SNA PU Type 2.1 APPN
- SNA PU Type 4 NCP

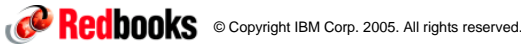
- SNA PU Type 2.0
- SNA PU Type 2.1 LEN
- SNA PU Type 2.1 APPN
- SNA PU Type 4 NCP
- SNA PU Type 5 VTAM (OSA-LSA)



Preliminary SNI performance data for CCL V1R2

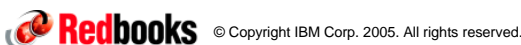
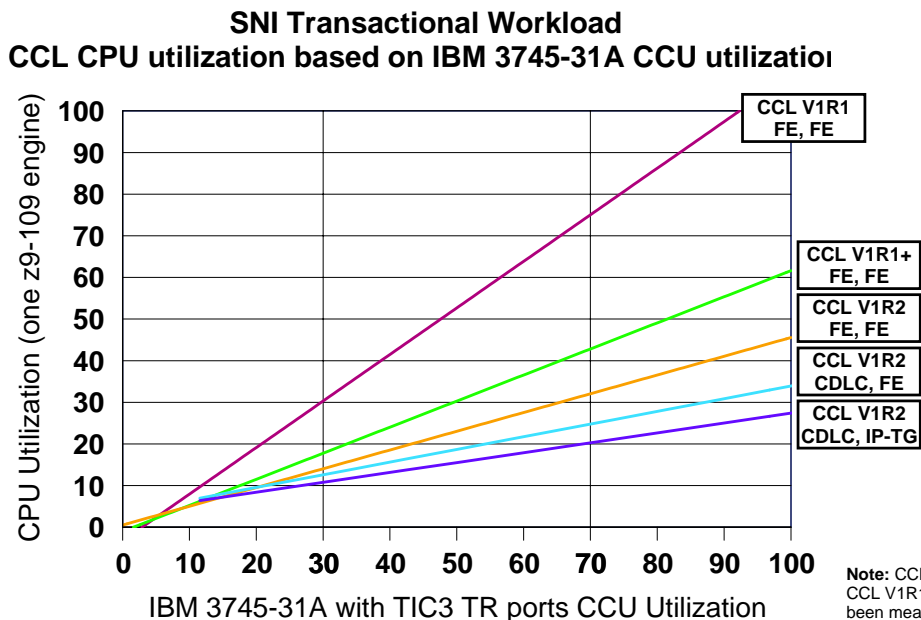


An IBM 3745-31A sample configuration with TIC3 ports maxes out around 380 transactions per second



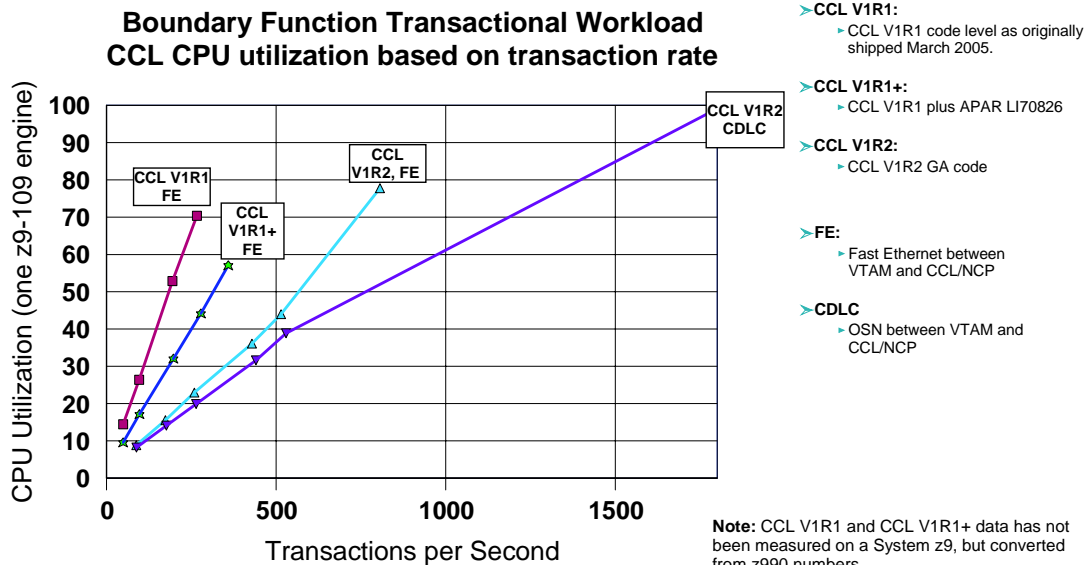
ibm.com/redbooks

Preliminary SNI workload CPU capacity planning data for CCL V1R2 running on System z9



ibm.com/redbooks

Preliminary boundary function performance data for CCL V1R2



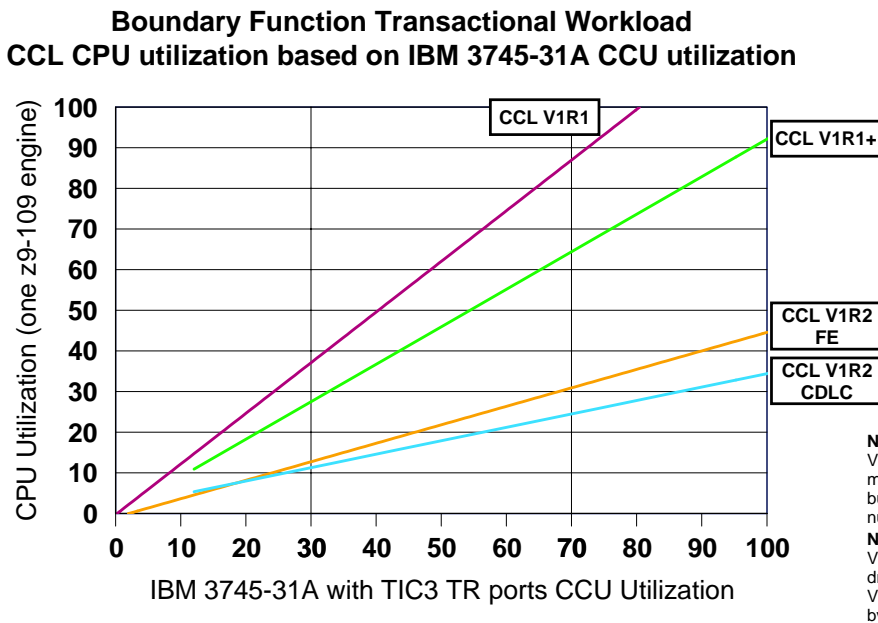
An IBM 3745-31A sample configuration with TIC3 ports maxes out around 268 transactions per second



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Preliminary boundary function workload CPU capacity planning data for CCL V1R2 running on System z9

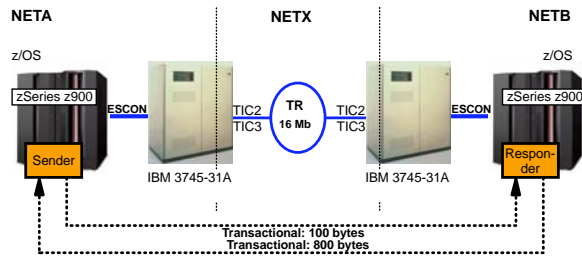


© Copyright IBM Corp. 2005. All rights reserved.

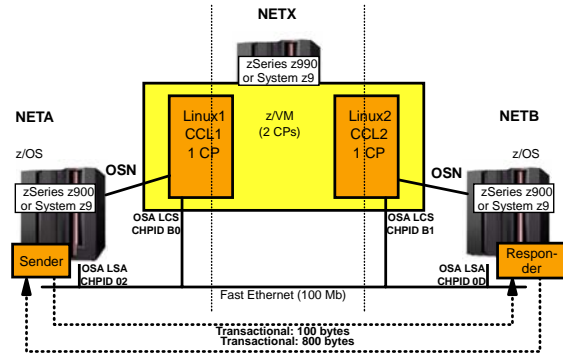
ibm.com/redbooks

CCL CPU usage - SNI transactional workload - setup notes

NOTES



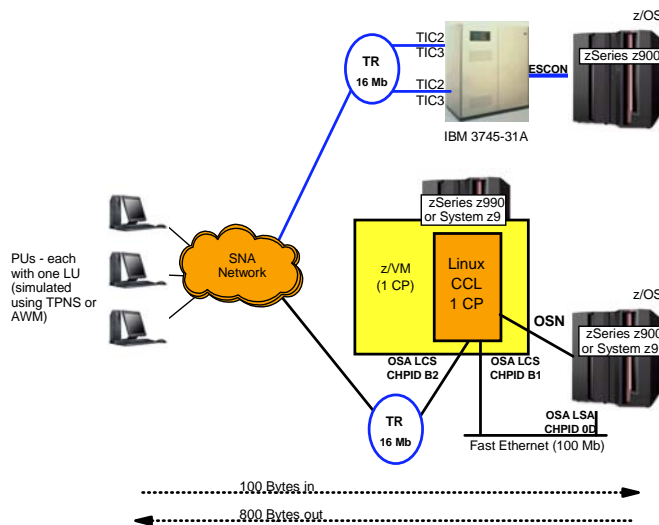
- > **IBM 3745-61A**
 - ▶ Tests were done with an IBM 3745-61A - divided into two IBM 3745-31A units
 - ▶ CCU capacity equals an IBM 3745-31A
 - ▶ One TIC2 or one TIC3 adapter
- > **CCL as z/VM guest**
 - ▶ 2 dedicated z990 or z9 CPs to z/VM
 - ▶ 1 virtual CP to each CCL Linux guest
 - ▶ "Client" side CCL working set 105 MB (as reported by z/VM)
 - ▶ "Server" side CCL working set 340 MB (as reported by z/VM)



- > **z990 or z9-109 hardware for CCL**
- > **CCL V1R2: SUSE SLES9 SP2 - Linux 2.6 kernel**
- > **CCL V1R1 and V1R1+: SUSE SLES8 SP4 - Linux 2.4 kernel**
- > **Transactional workload characteristics**
 - ▶ Up to 175 LU 6.2 sessions used to generate traffic
 - ▶ 100 bytes in per transaction
 - ▶ 800 bytes out per transaction
 - ▶ Thinktime 330 milliseconds between transactions per SNA session
 - ▶ VTAM IOBUF size=932
 - ▶ IBM 3745 MAXBFRU=20

CCL CPU usage - boundary function transactional workload - setup notes

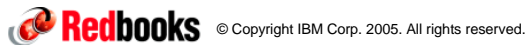
NOTES



- > **IBM 3745-31A**
 - ▶ CCU capacity equals an IBM 3745-31A
 - ▶ Two TIC2 or two TIC3 adapters
- > **CCL as z/VM guest**
 - ▶ 2 dedicated z990 or z9 CPs to z/VM
 - ▶ 1 virtual CP to the CCL Linux guest
- > **z990 or z9-109 hardware for CCL**
- > **CCL V1R2: SUSE SLES9 SP2 - Linux 2.6 kernel**
- > **CCL V1R1 and V1R1+: SUSE SLES8 SP4 - Linux 2.4 kernel**
- > **Transactional workload characteristics**
 - ▶ BF devices were emulated as one LU per PU SNA devices
 - ▶ Each transaction consisted of a 100-byte request in and a 800-byte response
 - ▶ A thinktime of 330 milli seconds was used between each transaction
 - ▶ VTAM IOBUF size was set to 932 in all test runs.
 - ▶ NCP MAXBFRU was set to 20 in all test runs.
 - ▶ CCL NCP MAXOUT was set to 7 for CCL to VTAM.
 - ▶ CCL V1R1 and V1R1+ workloads were driven by TPNS.
 - ▶ CCL V1R2 workloads were driven by AWM

CCL requirement for zSeries or System z9 hardware

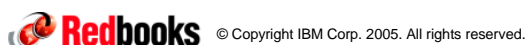
- **Processor support**
 - ▶ G5/G6, z800/z900, z890/z990, or z9-109
- **CP requirements (can be IFL engines on zSeries and System z9)**
 - ▶ Depends on workload and connectivity options
 - ▶ In general it is possible to migrate two heavily used (CCU utilization over 70% each) IBM 3745 SNI NCPs to one z990 IFL engine and up to three IBM 3745 SNI NCPs to one System z9 IFL engine
- **OSA port requirement**
 - ▶ Copper-based ports for SNA LLC2 (LCS) - can be used on all hardware levels
 - ▶ Fiber optic or copper ports for SNA LLC2 (QDIO layer-2) - z890, z990, z9-109 only
 - ▶ Fiber optic or copper ports for SNA over IP such as IP-TG or XOT (QDIO layer-3, QDIO layer-2, or LCS)
 - ▶ OSN port for CDLC connectivity - z9-109 only
- **Memory requirements**
 - ▶ Memory per CCL engine: 20 MB
 - ▶ Usual memory requirements for Linux on zSeries
 - Memory: 256 - 512 MB memory (depending on distribution, packages, and kernel level)
- **DASD requirements**
 - ▶ DASD for CCL: 50 MB
 - ▶ DASD for CCL traces, dumps, logs, NCP load modules: 80 - 100 MB per CCL engine
 - ▶ Usual DASD requirements for Linux on zSeries
 - Approximate DASD space equivalent to two 3390-3 DASD volumes
 - Use the Linux Logical Volume Manager (LVM) to group the volumes together



ibm.com/redbooks

CCL V1R2 requirements for Linux on System z9 and zSeries

- **Minimum Linux requirements for CCL V1R2**
 - ▶ SUSE LINUX Enterprise Server 8 for IBM zSeries and IBM S/390 (SLES8), kernel 2.4.21
 - Minimum level supported: Service Pack 4 (SLES8 + SP4)
 - ▶ SUSE LINUX Enterprise Server 9 for IBM zSeries and IBM S/390 (SLES9), kernel 2.6.5
 - Minimum level supported: Service Pack 1 (SLES9 + SP1)
 - ▶ Red Hat Enterprise Linux AS 4 (RHEL4), kernel 2.6.9
 - Minimum level supported: Update 1 (RHEL4 + Update1)
 - Note: IBM shipped an enabling PTF to CCL V1R1 (LI70764) on July 21, 2005, for CCL V1R1 to run on RHEL 4 Update 1
 - Note: LI70764 will be integrated into GA-level CCL V1R2
 - ▶ Both 31-bit and 64-bit distributions are supported
- **Minimum Linux requirements for CCL V1R2 communication via QDIO layer 2:**
 - ▶ Processors: IBM System z9 or IBM eServer zSeries z890, z990
 - ▶ Linux support is available (only for kernel 2.6) as source code patch on developerWorks:
 - <http://www.ibm.com/developerworks/linux/linux390/linux-2.6.5-s390-27-april2004.html>
 - ▶ IBM is working with its Linux distribution partners to ensure that this function will be provided in future kernel 2.6 distribution releases or service updates.
 - Note: CCL V1R1 also supports communication via OSA Layer 2
- **Minimum Linux requirements for CCL V1R2 communication via CDLC:**
 - ▶ Processors: IBM System z9
 - ▶ Linux support is available (only for kernel 2.6) as source code patch on developerWorks:
 - <http://www.ibm.com/developerworks/linux/linux390/linux-2.6.5-s390-29-april2004.html>
 - ▶ IBM is working with its Linux distribution partners to ensure that this function will be provided in future kernel 2.6 distribution releases or service updates.
- **For availability of further distributions supporting CCL V1R2 functions and specific package requirements on top of available distributions refer to:**
 - ▶ <http://www.ibm.com/software/network/ccl>



ibm.com/redbooks

CCL V1R2 is still NOT a complete replacement for the IBM 3745/46 Communication Controller, but it gets closer!

| CCL Functional Overview Matrix | CCL V1R2 supports | CCL V1R2 support of serial lines via an aggregation layer router | CCL V1R2 does not support |
|------------------------------------|--|--|--|
| Software | NCP (V7R5 and above) and compatible levels of NRF SSP, NTuneMON, NetView, and NPM continue to work as they have in the past <i>NCP Packet Switching Interface (NPSI)</i> | | Other IBM 3745 software products: XI/NSF, EP, NTO, NSI, MERVA, and TPNS Functions provided by the IBM 3746 MAE or NNP NCP-based IP routing |
| Physical network interfaces | OSA token-ring and Ethernet LAN (uses an LCS interface that is only supported by certain, copper-based, OSA cards) <i>CDLC channel connectivity through OSA on System z9</i> <i>OSA fiber optic connectivity QDIO layer-2 for SNA LLC2 traffic</i> <i>IP-TG for direct IP connectivity between two CCL NCPs</i> | SDLC, Frame Relay, X.25 QLLC, and ISDN serial line interfaces are not supported directly by CCL, but are supported via an aggregation layer router <i>X.25 circuits are not supported directly by CCL, but are via an aggregation layer router that uses the XOT protocol to transport the X.25 packets to/from NPSI running in CCL</i> | BSC, ALC, Start/Stop |



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

CCL V1R1 update reference

- **Service included by APAR LI70732:**
 - The DISK IPL info on the MOSS HTTP console displays wrong Gen date/time.
- **Service included by APAR LI70798:**
 - Error processing IMAGENAME longer than 10 characters.
- **Service included by APAR LI70764:**
 - New function has been added to CCL in support of RedHat Release 4 (RHEL4 U1).
- **Service included by APAR LI70826:**
 - New function - CCL Engine Performance Enhancement

For more info on latest maintenance and PTF download:
<http://www-1.ibm.com/support/docview.wss?rs=2192&uid=swg24009855>



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Contact information

- CCL home page:
 - ▶ <http://www.ibm.com/software/network/ccl>
- For more information, contact:
 - ▶ EMEA: Peter Redman - Peter_Redman@uk.ibm.com
 - ▶ Americas: Erika Lewis - erika@us.ibm.com
 - ▶ AP: Chuck Gardiner - cgardine@us.ibm.com
- For planning and installation services, contact:
 - ▶ April Singer in IBM Software Services for Websphere, Enterprise Transformation Services - singeraf@us.ibm.com
- For technical assistance in the Americas, IBMers can submit a TechExpress through w3.ibm.com or a question through WWQ&A




For further technical assistance:

US:

- ▶ Access installation and technical support information via the WWQA database
 - IBMers can access via the WWQA database via QASearch on <http://w3.viewblue.ibm.com>
 - Customers can access installation and technical support information from [IBMLink/ServiceLink](#).
- ▶ Please research questions through all available resources before submitting a question to the Q&A database.


EMEA

- ▶ Techline and local Field Technical Support Specialists provide technical pre-sales assistance. Additional technical support is available through worldwide Question & Answer (WWQA), QASearch function on ViewBlue or EHONE. For some brands/products, authoring of questions is only available via Techline.

 © Copyright IBM Corp. 2005. All rights reserved.

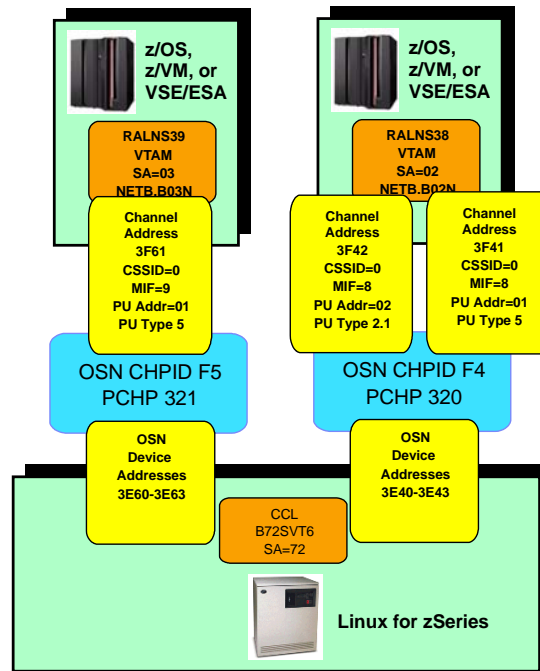
ibm.com/redbooks

Appendix 1: CDLC channel connectivity to CCL

 © Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

CDLC channel connectivity to CCL: OSN configuration scenario



IOCP definitions for OSN

```

RESOURCE PART=((CSS(0),(RANS38,8),(RANS39,9),(RALNS27,D)))
*
CHPID PCHID=320,PATH=(CSS(0,1),F4),TYPE=OSN,SHARED
CHPID PCHID=321,PATH=(CSS(0,1),F5),TYPE=OSN,SHARED
*
CNTLUNIT CUNUMBR=3E40,PATH=((CSS(0),F4),(CSS(1),F4)),UNIT=OSN
CNTLUNIT CUNUMBR=3E60,PATH=((CSS(0),F5),(CSS(1),F5)),UNIT=OSN
*
IODEVICE ADDRESS=(3E40,32),CUNUMBR=3E40,UNIT=OSN,
UNITADD=20
IODEVICE ADDRESS=(3F41,10),CUNUMBR=3E40,UNIT=3745,
UNITADD=01
IODEVICE ADDRESS=(3F4E,1),CUNUMBR=3E40,UNIT=OSAD,
UNITADD=FE
*
IODEVICE ADDRESS=(3E60,32),CUNUMBR=3E60,UNIT=OSN,
UNITADD=20
IODEVICE ADDRESS=(3F61,10),CUNUMBR=3E60,UNIT=3745,
UNITADD=01
IODEVICE ADDRESS=(3F6E,1),CUNUMBR=3E60,UNIT=OSAD,
UNITADD=FE
    
```

← OSN CHPIDs

← OSN Linux devices

← IBM 3745 VTAM/TPF devices

Defining OSN devices to Linux

- Create file `hwcfg-qeth-bus-ccw-0.0.3e60` in `/etc/sysconfig/hardware`

```
#-----  
# hwcfg-qeth-bus-ccw-0.0.3e60  
#  
# Hardware configuration for a qeth device at 0.0.3e60  
# Automatically generated by netsetup  
#-----  
  
STARTMODE="auto"  
MODULE="qeth"  
MODULE_OPTIONS=""  
MODULE_UNLOAD="yes"  
  
# Scripts to be called for the various events.  
SCRIPTUP="hwup-ccw"  
SCRIPTUP_ccw="hwup-ccw"  
SCRIPTUP_ccwgroup="hwup-qeth"  
SCRIPTDOWN="hwdown-ccw"  
# CCW_CHAN_IDS sets the channel IDs for this device  
# The first ID will be used as the group ID  
CCW_CHAN_IDS="0.0.3e60 0.0.3e61 0.0.3e62"  
  
# CCW_CHAN_NUM set the number of channels for this device  
# Always 3 for an qeth device  
CCW_CHAN_NUM=3  
  
# CCW_CHAN_MODE sets the port name for an OSA-Express device  
CCW_CHAN_MODE="GIGE3E60"
```

Similar configurations will need to be made for other OSN devices - such as in this sample setup: 3e40



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Defining OSN devices to Linux (continued)

- Create file `ifcfg-qeth-bus-ccw-0.0.3e60` in `/etc/sysconfig/network`

```
BOOTPROTO="static"  
UNIQUE=""  
STARTMODE="onboot"
```

- The previous definition will allow the OSN device to become active at startup.
- In order for the new OSN device to be recognized at startup, you must modify the `/etc/sysconfig/hardware/scripts/hwup-ccw` as follows:
 - ▶ Find Line: 1731/01|1731/05
 - ▶ Change to: 1731/01|1731/05|1731/06
- There is an alternative way of defining the OSN devices to Linux by echo'ing information into `/sys/bus/ccwgroup/drivers/qeth/group`
 - ▶ `echo 0.0.3e60,0.0.3e61,0.0.3e62 > /sys/bus/ccwgroup/drivers/qeth/group`
 - ▶ `echo 1 > /sys/bus/ccwgroup/drivers/qeth/0.0.3E60/online`
 - ▶ `ifconfig -e`
 - ▶ `ifconfig osn0 up`
- If using this method, one needs to repeat those echo commands after each IPL of Linux



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Establishing controls for load/dump of an NCP over the CDLC channel

- You can use the CCL load/dump program (cclcldp) to load a specified NCP into the CCL engine using a CDLC connection.
- You must define an `iplportdefs` configuration file for each CCL Engine that will be loaded from VTAM over a CDLC connection.
 - File name is `iplportdefs` and it must reside in the CCL engine directory

```
IPLPORTDEFS
*
*-----*
*      NCP DEFINITION:  ADDRESS=2112, HOSTLINK=9, ADDR=1
*      OSN DEFINITION:  CSS_ID=X'0' MIF_ID=X'09' UNITADD=X'01'
*                      CCID=X'00090001'
*                      DEVICE=X'3E60'
*-----*
*
*      ADDRESS      2112
*      HOSTLINK     9
*      ADDR         01
*      DEVICE       3e60
```

Optionally customize how to correlate NCP definitions with OSA CDLC resources

- The ESCON logical PU definitions in the NCP gen must be mapped to the OSA CDLC resources so that when a given NCP PU is activated, that PU connects the NCP to the desired VTAM (or TPF) link station (the specific 3745 device that is defined to the SNA host).
- The CDLC support in CCL is designed so you can encode this information in the NCP gen. During ESCON resource activation, NCP passes this information to CCL, which passes it to the (correct) OSA, which establishes the CDLC connection. The values are encoded as follows:
 - The physical LINE ADDRESS value maps to SNA host's CSS ID
 - The logical LINE HOSTLINK value maps to SNA host's MIF ID
 - The logical PU ADDR value maps to UNITADD value of 3745 device
 - The last four characters of logical PU name identifies the QETH device
- If this mapping doesn't match, you need to create a `CCLDEFS` file in which you code the correlation between the NCP definitions and the OSA CDLC resources.
 - File name is `ncp_load_module_name.CCLDEFS` and it must reside in the CCL engine directory

CCL NCP definitions

```

*****
*                CCL CDLC PHYSICAL LINE 2112                *
*****
B72GRP  GROUP  LNCTL=CA,ANS=CONT
B72C2112 LINE  ADDRESS=2112,ANS=CONT,SRT=(32765,32765),
          XMONLNK=YES,SPEED=18000000
B72P2112 PU    PUTYPE=1
*
*****
* Logical Group for Physical Line 2112                      *
*****
B72CALG1 GROUP LNCTL=CA,PHYSRSC=B72P2112,MAXPU=32,NPACOLL=NO,ANS=CONT,
          TIMEOUT=180,DELAY=0.0,CASDL=10,SRT=(32765,32765)
*
*****
* CONNECTION TO RALNS39 --- CSS ID = 0: MIF = 9: PUADDR=01
*****
B72LL03 LINE  ADDRESS=NONE,HOSTLINK=09,SPEED=18000000,MONLINK=YES
C3P13E60 PU    PUTYPE=5,ADDR=01,TRANSFR=140,TGN=1,MONLINK=YES
*
*****
* CONNECTION TO RALNS38 --- CSS ID = 0: MIF = 8: PUADDR=01
*****
B72LL02 LINE  ADDRESS=NONE,HOSTLINK=08,SPEED=18000000,MONLINK=YES
C2P13E40 PU    PUTYPE=5,ADDR=01,TRANSFR=140,TGN=1,MONLINK=YES
C2P23E40 PU    PUTYPE=2,ADDR=02

```

➤ TYPEGEN=NCP
instead of
TYPEGEN=NCP-R
in the BUILD
macro

➤ The VERSION
keyword on the
BUILD macro
must have the "F"
extension for
ODLC lines

The last 4 characters of
the PU Name (3E60)
equates to the READ
device address of the
OSN CHPID defined for
the Linux host

The last 4 characters of
the PU Name (3E40)
equates to the READ
device address of the
OSN CHPID defined for
the Linux host

Subarea 03 VTAM definitions

➤ Create a channel-attached major node

```

B03CA  VBUILD  TYPE=CA
B03GRP  GROUP  LNCTL=NCP
*
*****
* C3P13E60 PU ADDR = 01: CSS ID = 0: MIF = 9:          *****
*****
*
B03CALN LINE  ADDRESS=3F61,MAXBFRU=36
B03PU   PU    CHANCON=COND,MAXDATA=32768,TGN=1

```

Subarea 02 VTAM definitions

> Create a channel-attached major node

```
B02CA  VBUILD  TYPE=CA
B02GRP GROUP  LNCTL=NCP
*
*****
*   C2P13E40 PU ADDR = 01: CSS ID = 0: MIF = 8:           *****
*****
*
B02CALN LINE    ADDRESS=3F41,MAXBFRU=36
B02PU   PU      CHANCON=COND,MAXDATA=32768,TGN=1
```

> Create a local major node (for APPN activation)

```
B02LCL  VBUILD TYPE=LOCAL
*
*****
*   C2P23E40 PU ADDR = 02: CSS ID = 0: MIF = 8:           *****
*****
*
B02LCLP1 PU      PUTYPE=2,CUADDR=3F42,ISTATUS=ACTIVE,XID=YES,
                VPACING=0,SSCPFM=USSSCS,MAXBFRU=255,DYNLU=YES,
                CONNTYPE=APPN,CPCP=YES
```

Activating the NCP over the CDLC channel

> Start the CCL engine with the NCP load module already loaded (in this example: B72SVT6):

```
nohup ./cclengine -mB72SVT6 -p2072 SVTB72 &
```

> From NETB.B03N, activate the local CA major node

```
V NET,ACT,ID=B03CA,ALL
IST097I VARY ACCEPTED
IST093I B03CA ACTIVE
IST464I LINK STATION B03PU HAS CONTACTED B72SVT6 SA 72
IST093I B03PU ACTIVE
```

> From NETB.B03N, activate the NCP major node

```
V NET,ACT,ID=B72SVT6,ALL
IST097I VARY ACCEPTED
IST093I B72SVT6 ACTIVE
IST093I B72NPPU ACTIVE
IST093I B72P2112 ACTIVE
IST464I LINK STATION C2P13E40 HAS CONTACTED B02N SA 2
IST093I C2P13E40 ACTIVE
IST464I LINK STATION C3P13E60 HAS CONTACTED ISTPUS SA 3
IST093I C3P13E60 ACTIVE
```

Activating the NCP over the CDLC channel (continued)

- To contact the NCP from VTAM subarea 02 (NETB.B02N) via the channel major node, activate the local CA major node

```
V NET,ACT,ID=B02CA,ALL
IST097I VARY ACCEPTED
IST093I B02CA ACTIVE
IST464I LINK STATION B02PU HAS CONTACTED B72SVT6 SA 72
IST093I B02PU ACTIVE
```

- To contact the NCP from VTAM subarea 02 (NETB.B02N) via the local major node (as an APPN node), activate the Local major node

```
V NET,ACT,ID=B02LCL,ALL
IST097I VARY ACCEPTED
IST093I B02LCL ACTIVE
IST1086I APPN CONNECTION FOR NETB.B03N IS ACTIVE - TGN = 21
IST093I B02LCLP1 ACTIVE
IST1096I CP-CP SESSIONS WITH NETB.B03N ACTIVATED
```

Activating and loading the NCP over the CDLC channel

- Start the CCL engine - but use a load module name of cclcldp to instruct the CCL engine that the load will come from the VTAM activation command:

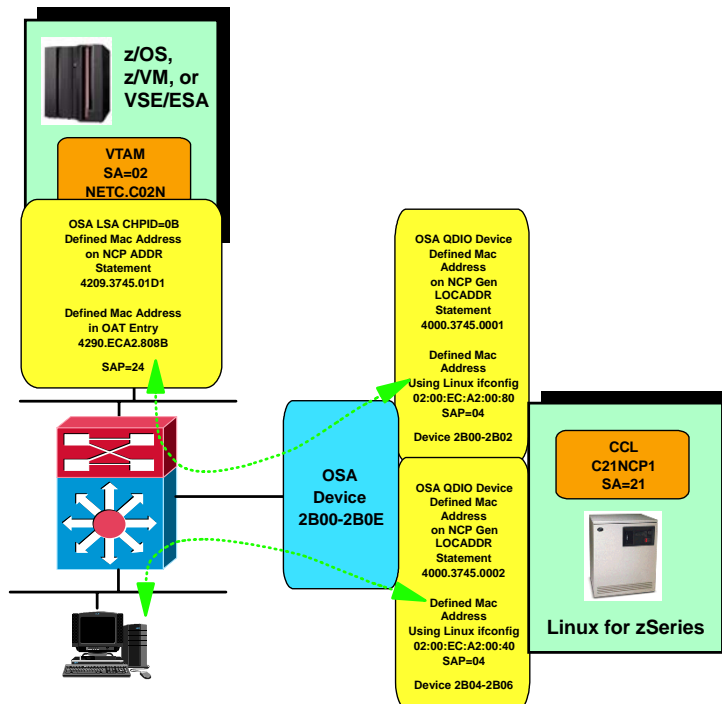
```
nohup ./cclengine -mcclcldp -p2072 SVTB72 &
```

- From NETB.B03N, load and activate the NCP Major Node

```
V NET,ACT,ID=B72SVT6,ALL,LOAD=YES,U=3F61
IST097I VARY ACCEPTED
IST461I ACTIVATE FOR U/RNAME ENTRY ID = 3F61-S STARTED
IST897I LOAD OF B72SVT6 STARTED
IST270I LOAD OF B72SVT6 COMPLETE - LOAD MODULE = B72SVT6
IST464I LINK STATION 3F61-S HAS CONTACTED B72SVT6 SA 72
IST093I B72SVT6 ACTIVE IST093I B72NPPU ACTIVE
IST093I B72P2112 ACTIVE
IST464I LINK STATION C2P13E40 HAS CONTACTED B02N SA 2
IST093I C2P13E40 ACTIVE
IST464I LINK STATION C3P13E60 HAS CONTACTED ISTPUS SA 3
IST093I C3P13E60 ACTIVE
```

Appendix 2: QDIO Layer 2 access from native Linux LPAR

QDIO layer 2 mode: configuration overview



Define QDIO layer 2 devices 2B00-2B02 to Linux - hardware

- Create the script `hwcfg-qeth-bus-ccw-0.0.2b00` in the `/etc/sysconfig/hardware` directory

```
#!/bin/sh
#
STARTMODE='auto'
MODULE='qeth'
MODULE_OPTIONS=''
MODULE_UNLOAD='yes'

SCRIPTUP='hwup-ccw'
SCRIPTUP_ccw='hwup-ccw'
SCRIPTUP_ccwgroup='hwup-qeth'
SCRIPTDOWN='hwdown-ccw'

# CCW_CHAN_IDS are the device addresses
CCW_CHAN_IDS='0.0.2b00 0.0.2b01 0.0.2b02'

# CCW_CHAN_NUM set the number of channels for this device
CCW_CHAN_NUM='3'

# CCW_CHAN_MODE sets the port name for an OSA-Express device
CCW_CHAN_MODE='GIGE2B00'

# QETH_LAYER2_SUPPORT enables Layer2 support for this device.
QETH_LAYER2_SUPPORT=1
```

QDIO device addresses 2B00, 2B01, and 2B02

This is where you specify to the Linux QDIO device driver that you want to operate this QDIO port in layer 2 mode.



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Define QDIO layer 2 devices 2B00-2B02 to Linux - network

- Create this script `ifcfg-qeth-bus-ccw-0.0.2b00` in the `/etc/sysconfig/network` directory

```
LLADDR='02:00:ec:a2:00:80'
BOOTPROTO='none'
STARTMODE='onboot'
UNIQUE=''
```

- By using `LLADDR`, we can set the MAC address to any value necessary. This keyword may be different in Red Hat releases.
- The local MAC address can also be set using an `ifconfig` command:
`ifconfig eth0 hw ether 02:00:ec:a2:00:80 up`
- MAC Address defined on the NCP `LOCADDR` statement is the non-canonical version of this address – 4000.3745.0001



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Define QDIO layer 2 devices 2B04-2B06 to Linux - hardware

- Create the script `hwcfg-qeth-bus-ccw-0.0.2b04` in the `/etc/sysconfig/hardware` directory

```
#!/bin/sh
#
STARTMODE='auto'
MODULE='qeth'
MODULE_OPTIONS=''
MODULE_UNLOAD='yes'

SCRIPTUP='hwup-ccw'
SCRIPTUP_ccw='hwup-ccw'
SCRIPTUP_ccwgroup='hwup-qeth'
SCRIPTDOWN='hwdown-ccw'

# CCW_CHAN_IDS are the device addresses
CCW_CHAN_IDS='0.0.2b04 0.0.2b05 0.0.2b06'

# CCW_CHAN_NUM set the number of channels for this device
CCW_CHAN_NUM='3'

# CCW_CHAN_MODE sets the port name for an OSA-Express device
CCW_CHAN_MODE='GIGE2B00'

# QETH_LAYER2_SUPPORT enables Layer2 support for this device.
QETH_LAYER2_SUPPORT=1
```

QDIO device addresses 2B04, 2B05, and 2B06

This is where you specify to the Linux QDIO device driver that you want to operate this QDIO port in layer 2 mode.



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Define QDIO layer 2 devices 2B04-2B06 to Linux - network

- Create this script `ifcfg-qeth-bus-ccw-0.0.2b04` in the `/etc/sysconfig/network` directory

```
LLADDR='02:00:ec:a2:00:40'
BOOTPROTO='none'
STARTMODE='onboot'
UNIQUE=''
```

- By using `LLADDR`, we can set the MAC address to any value necessary. This keyword may be different in Red Hat releases.
- The local MAC address can also be set using an `ifconfig` command:
`ifconfig eth0 hw ether 02:00:ec:a2:00:40 up`
- MAC Address defined on the NCP `LOCADDR` statement is the non-canonical version of this address – 4000.3745.0002



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

VTAM channel-attached major node definitions

- Create a channel-attached major node in VTAM on subarea 02:

```
C02XCA  VBUILD  TYPE=XCA
*
C02ETHPT PORT  MEDIUM=CSMACD,ADAPNO=0,SAPADDR=24,CUADDR=2EBA,      X
          TIMER=100
C02ETHGP GROUP DIAL=NO,ISTATUS=ACTIVE
*
C02ETHL2 LINE  USER=SNA,ISTATUS=ACTIVE
C02ETHP2 PU    MACADDR=0200ECA20080,PUTYPE=5,SUBAREA=21,TGN=1,      X
          SAPADDR=04,ALLOWACT=YES
```

- In this sample setup, VTAM communicates with the NCP over QDIO devices 2B00-2B02 - using the default SAP of 04.

NCP physical line interface definitions (NTRI)

- Define the NCP physical line interfaces:

```
*****
* Physical NTRI Lines
*****
*
C21PTRG1 GROUP ECLTYPE=(PHY,ANY),ADAPTER=TIC2,ANS=CONT,MAXTSL=16732,  X
          RCVBUFC=32000,USSTAB=AUSSTAB,ISTATUS=ACTIVE,XID=NO,      X
          RETRIES=(20,5,5),NPACOLL=(YES,EXTENDED)
*
C21TR88  LINE  ADDRESS=(1088,FULL),TRSPEED=16,PORTADD=88,          X
          LOCADD=400037450001,NPACOLL=YES
C21PU88A PU
*
C21TR89  LINE  ADDRESS=(1089,FULL),TRSPEED=16,PORTADD=89,          X
          LOCADD=400037450002,NPACOLL=YES
C21PU89A PU
```

Uses QDIO layer 2 devices 2B04-2B06

Uses QDIO layer 2 devices 2B00-2B02

NCP logical line interface definitions

> Define the NCP logical line interfaces:

```
*****
* LOGICAL BNN Lines *
*****
*
C21BNNG1 GROUP ECLTYPE=LOGICAL,ANS=CONTINUE,AUTOGEN=250,CALL=INOUT, X
           ISTATUS=ACTIVE,PHYSRSC=C21PU89A, X
           USSTAB=AUSSTAB,RETRIES=(10,10,10,20),XMITDLY=NONE, X
           MODETAB=AMODETAB,NPACOLL=YES
*****
* NTRI INN LOGICAL LINES FOR TOKEN RING PORT 1088 *
*****
*
C21INNG1 GROUP ECLTYPE=(LOGICAL,SUBAREA),ANS=CONT,PHYSRSC=C21PU88A, X
           LOCALTO=13.5,REMOTTO=18.2,T2TIMER=(0.2,0.2,3), X
           ISTATUS=ACTIVE,SDLCST=(C21PRI,C21SEC),NPACOLL=YES, X
           MONLINK=CONT
*
C21LG2A LINE TGN=1,TGCONF=SINGLE
C21PG2A PU ADDR=184209374501D1,SSAP=(04,H)
```

> You start the CCL engine and loads the NCP using the usual cclengine command:
nohup ./cclengine -mC21NCP1 -p2021 SVTC21 &



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

Activating the NCP from VTAM

> From NETC.C02N activate the XCA major node

```
V NET,ACT,ALL,ID=C02XCA
IST097I VARY ACCEPTED
IST093I C02XCA ACTIVE
IST464I LINK STATION C02ETHP2 HAS CONTACTED C21NCP1 SA 21
IST093I C02ETHP2 ACTIVE
```

> From NETC.C02N activate the NCP

```
V NET,ACT,ID=C21NCP1,ALL
IST097I VARY ACCEPTED
IST093I C21NCP1 ACTIVE
IST093I C21PU88A ACTIVE
IST093I C21PU89A ACTIVE
IST093I C21NPPU ACTIVE
IST464I LINK STATION C21PG2A HAS CONTACTED C02NPU SA 2
IST093I C21PG2A ACTIVE
```



© Copyright IBM Corp. 2005. All rights reserved.

ibm.com/redbooks

ibm.com

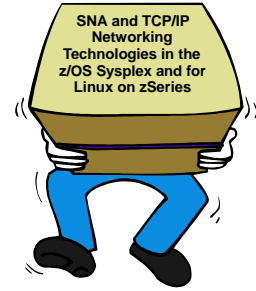


e-business



SNA and TCP/IP Networking Technologies in the z/OS Sysplex and for Linux on System z9 and zSeries

The End



Redbooks

International Technical Support Organization

© Copyright IBM Corp. 2005. All rights reserved.