ibm.com

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

Redbooks

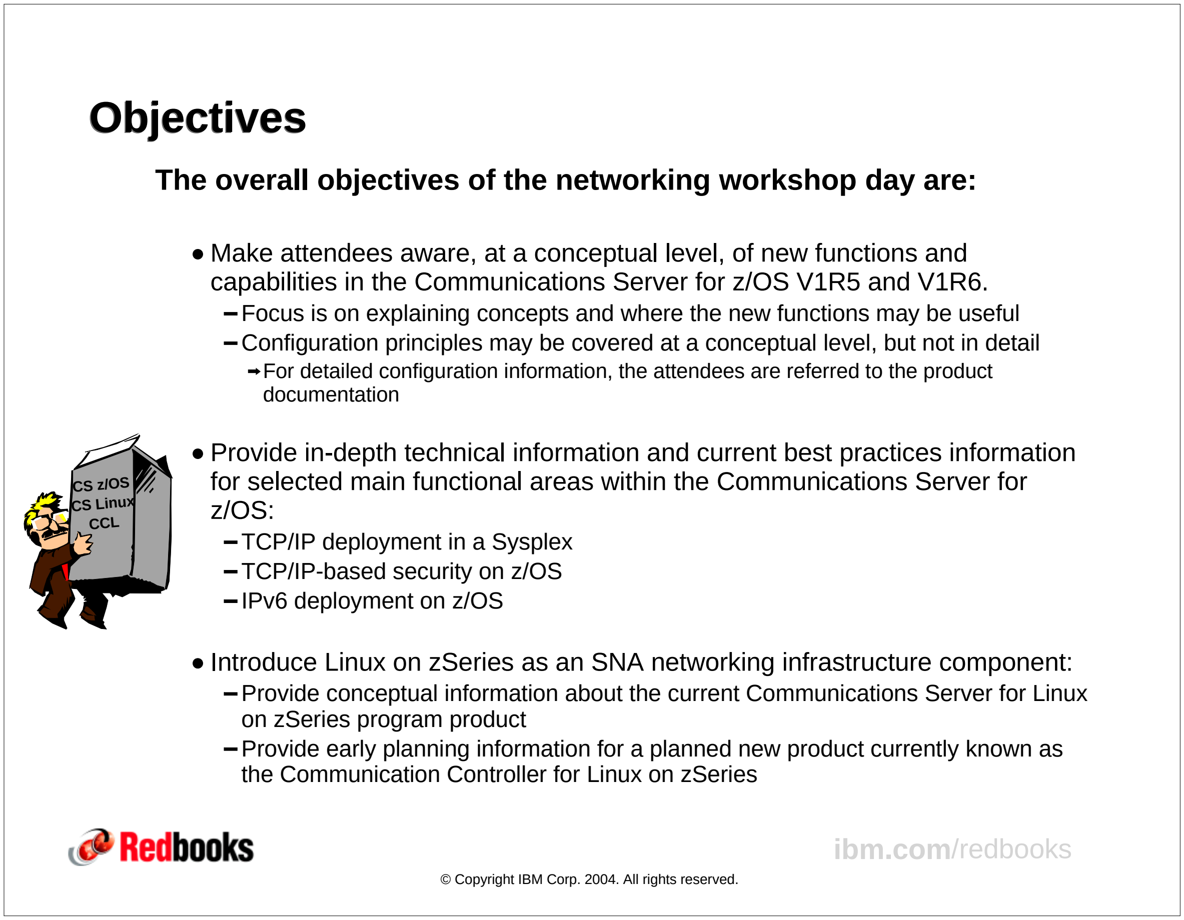
International Technical Support Organization

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

Redbooks

ibm.com/redbooks



Objectives

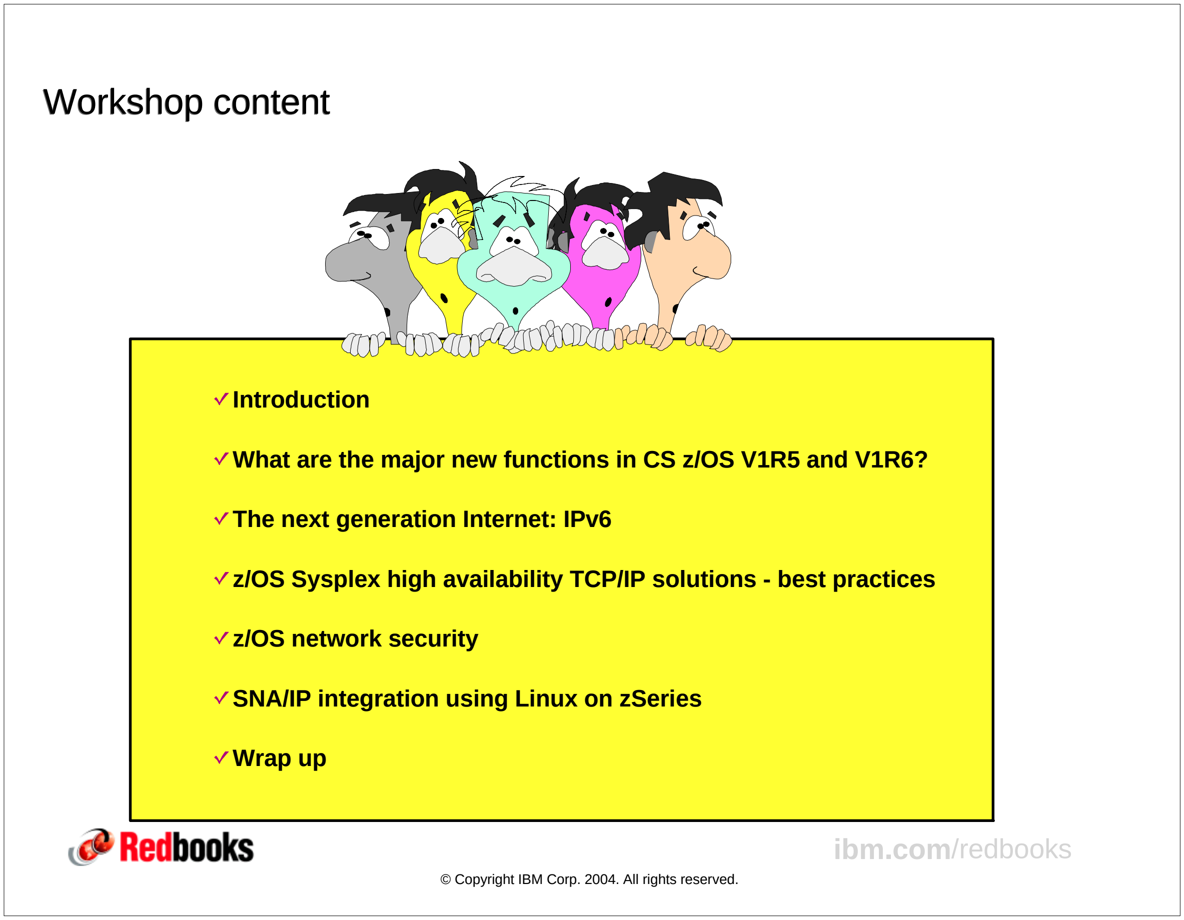
The overall objectives of the networking workshop day are:

- Make attendees aware, at a conceptual level, of new functions and capabilities in the Communications Server for z/OS V1R5 and V1R6.
  - Focus is on explaining concepts and where the new functions may be useful
  - Configuration principles may be covered at a conceptual level, but not in detail
    - ➡For detailed configuration information, the attendees are referred to the product documentation

- Provide in-depth technical information and current best practices information for selected main functional areas within the Communications Server for z/OS:
  - TCP/IP deployment in a Sysplex
  - TCP/IP-based security on z/OS
  - IPv6 deployment on z/OS
- Introduce Linux on zSeries as an SNA networking infrastructure component:
  - Provide conceptual information about the current Communications Server for Linux on zSeries program product
  - Provide early planning information for a planned new product currently known as the Communication Controller for Linux on zSeries

Redbooks

ibm.com/redbooks



Workshop content

- ✓ **Introduction**
- ✓ **What are the major new functions in CS z/OS V1R5 and V1R6?**
- ✓ **The next generation Internet: IPv6**
- ✓ **z/OS Sysplex high availability TCP/IP solutions - best practices**
- ✓ **z/OS network security**
- ✓ **SNA/IP integration using Linux on zSeries**
- ✓ **Wrap up**

Redbooks

ibm.com/redbooks

Tentative time schedule

Time	Topic
09:00 - 09:30	Introduction
09:30 - 11:30	What are the major new functions in CS z/OS V1R5 and V1R6?
11:30 - 12:30	The next generation Internet: IPv6
12:30 - 13:30	Lunch
13:30 - 15:00	z/OS Sysplex high availability TCP/IP solutions - best practices
15:00 - 15:45	z/OS network security
15:45 - 16:45	SNA/IP integration using Linux on zSeries
16:45 - 17:00	Wrap up

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

Practical information

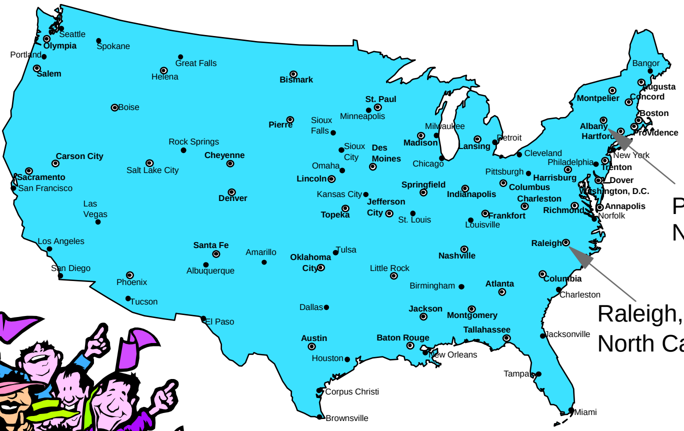
A certain level of familiarity with both SNA and TCP/IP networking technologies in general and on z/OS specifically is assumed.

Questions are welcome all the time.

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

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

Research Triangle Park, Raleigh, North Carolina - the home of Communications Servers



Poughkeepsie, New York

Raleigh, North Carolina

This workshop has been developed by contributions from the entire Communications Server development team in Raleigh, North Carolina - in cooperation with ITSO in Poughkeepsie, New York.

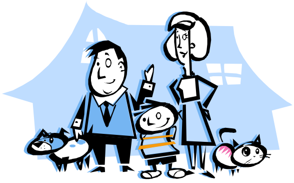
Redbooks

ibm.com/redbooks

Enterprise Networking Solutions: Communications Server Product Family

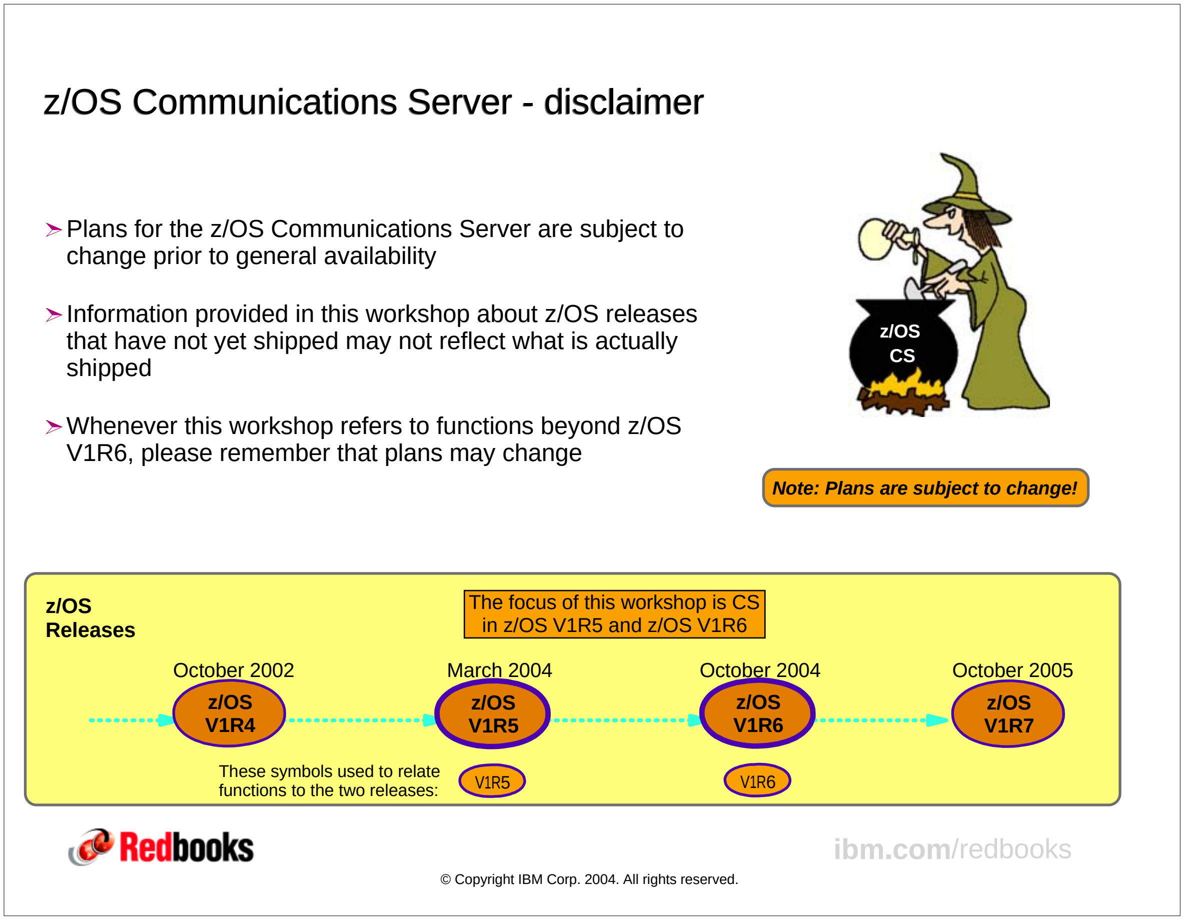
The networking family of products that are managed from a development perspective by Enterprise Networking Solutions in Research Triangle Park, North Carolina are:

- ✓ **Communications Server for z/OS - SNA and TCP/IP**
- ✓ **Communications Server for AIX - SNA**
- ✓ **Communications Server for Windows - SNA**
- ✓ **Communications Server for Linux (on Intel) - SNA**
- ✓ **Communications Server for Linux on zSeries - SNA**
- ✓ **Application Workload Modeler**
- ✓ **Communication Controller for Linux (planned 2005)**
- ✓ **ACF/SSP**
- ✓ **ACF/NCP**
- ✓ **ATunemon**
- ✓ **NPSI**



Redbooks

ibm.com/redbooks



z/OS Communications Server - disclaimer

- Plans for the z/OS Communications Server are subject to change prior to general availability
- Information provided in this workshop about z/OS releases that have not yet shipped may not reflect what is actually shipped
- Whenever this workshop refers to functions beyond z/OS V1R6, please remember that plans may change

Note: Plans are subject to change!

z/OS Releases

The focus of this workshop is CS in z/OS V1R5 and z/OS V1R6

October 2002 z/OS V1R4

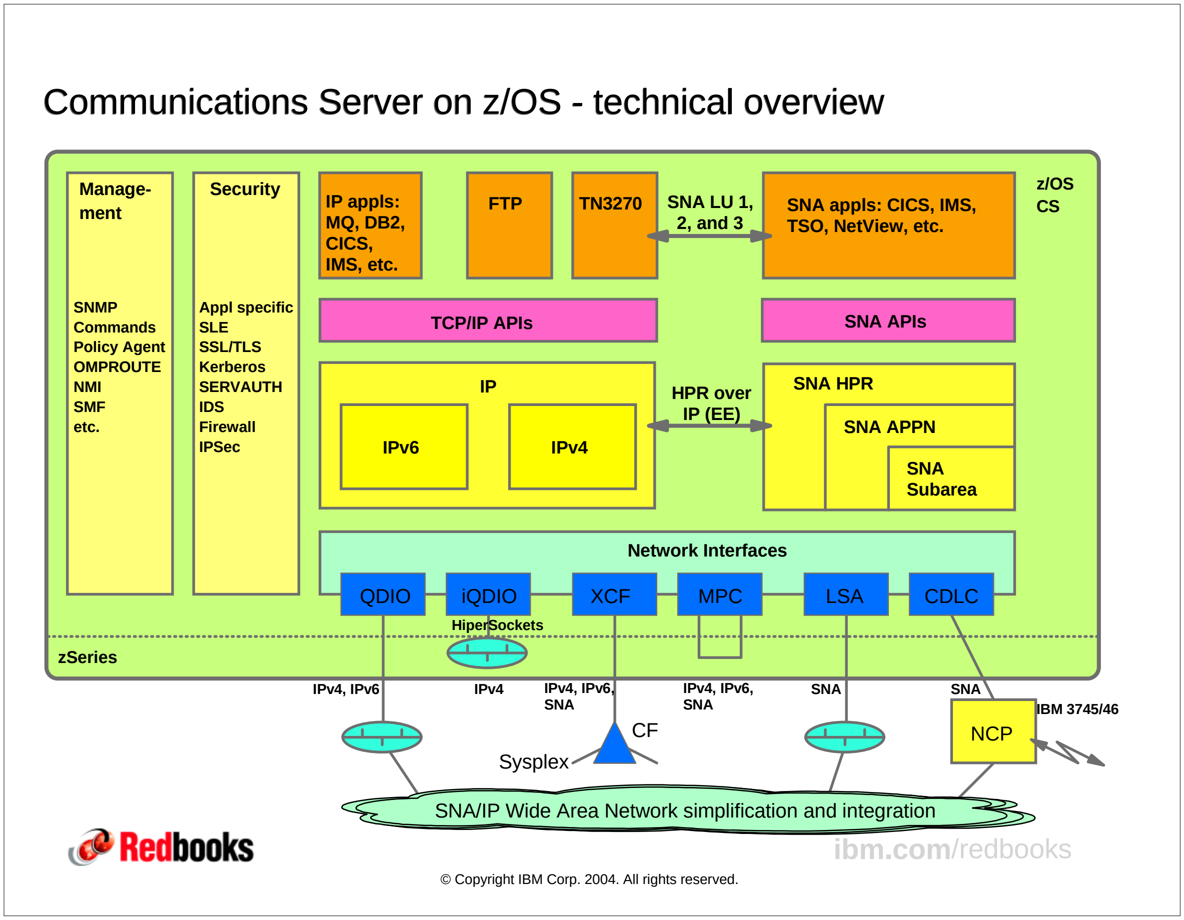
March 2004 z/OS V1R5

October 2004 z/OS V1R6

October 2005 z/OS V1R7

These symbols used to relate functions to the two releases: V1R5 V1R6

ibm.com/redbooks



Communications Server on z/OS - technical overview

z/OS CS

Manage-ment

SNMP
Commands
Policy Agent
OMPROUTE
NMI
SMF
etc.

Security

Appl specific
SLE
SSL/TLS
Kerberos
SERVAUTH
IDS
Firewall
IPSec

IP appls: MQ, DB2, CICS, IMS, etc.

FTP

TN3270

SNA LU 1, 2, and 3

SNA appls: CICS, IMS, TSO, NetView, etc.

TCP/IP APIs

SNA APIs

IP

IPv6

IPv4

HPR over IP (EE)

SNA HPR

SNA APPN

SNA Subarea

Network Interfaces

QDIO iQDIO XCF MPC LSA CDLC

HiperSockets

zSeries

IPv4, IPv6

IPv4

IPv4, IPv6, SNA

IPv4, IPv6, SNA

SNA

SNA

IBM 3745/46

NCP

CF

Sysplex

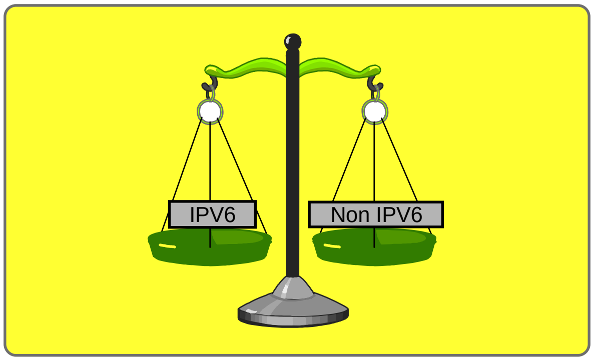
SNA/IP Wide Area Network simplification and integration

ibm.com/redbooks

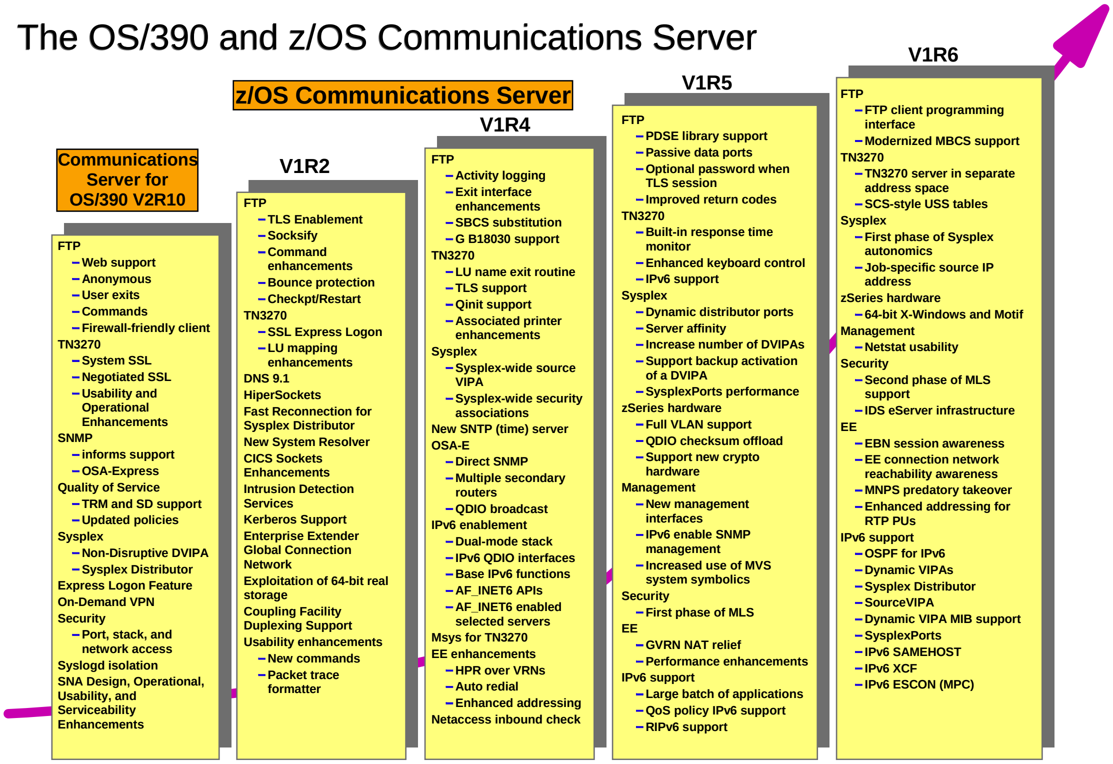
z/OS V1R5 and V1R6 Communications Server

z/OS V1R5 and V1R6 continue the effort started in z/OS V1R4 to provide IPV6 support on z/OS, but both releases also provide a lot of non-IPV6 functions!

- IPv6 support is technology driven in preparation for future



The OS/390 and z/OS Communications Server



Communications Server for OS/390 V2R10

FTP
- Web support
- Anonymous
- User exits
- Commands
- Firewall-friendly client

TN3270
- System SSL
- Negotiated SSL
- Usability and Operational Enhancements

SNMP
- informs support
- OSA-Express

Quality of Service
- TRM and SD support
- Updated policies

Sysplex
- Non-Disruptive DVIPA
- Sysplex Distributor

Express Logon Feature

On-Demand VPN

Security
- Port, stack, and network access

Syslogd isolation

SNA Design, Operational, Usability, and Serviceability Enhancements

z/OS Communications Server

V1R2

FTP
- TLS Enablement
- Socksify
- Command enhancements
- Bounce protection
- Checkpt/Restart

TN3270
- SSL Express Logon
- LU mapping enhancements

DNS 9.1

HiperSockets

Fast Reconnection for Sysplex Distributor

New System Resolver

CICS Sockets Enhancements

Intrusion Detection Services

Kerberos Support

Enterprise Extender Global Connection Network

Exploitation of 64-bit real storage

Coupling Facility Duplexing Support

Usability enhancements
- New commands
- Packet trace formatter

V1R4

FTP
- Activity logging
- Exit interface enhancements
- SBCS substitution
- G B18030 support

TN3270
- LU name exit routine
- TLS support
- Qinit support
- Associated printer enhancements

Sysplex
- Sysplex-wide source VIPA
- Sysplex-wide security associations

New SNTP (time) server

OSA-E
- Direct SNMP
- Multiple secondary routers
- QDIO broadcast

IPv6 enablement
- Dual-mode stack
- IPv6 QDIO interfaces
- Base IPv6 functions
- AF_INET6 APIs
- AF_INET6 enabled selected servers

Msys for TN3270

EE enhancements
- HPR over VRNs
- Auto redial
- Enhanced addressing

Netaccess inbound check

V1R5

FTP
- PDSE library support
- Passive data ports
- Optional password when TLS session
- Improved return codes

TN3270
- Built-in response time monitor
- Enhanced keyboard control
- IPv6 support

Sysplex
- Dynamic distributor ports
- Server affinity
- Increase number of DVIPAs
- Support backup activation of a DVIPA
- SysplexPorts performance

zSeries hardware
- Full VLAN support
- QDIO checksum offload
- Support new crypto hardware

Management
- New management interfaces
- IPv6 enable SNMP management
- Increased use of MVS system symbolics

Security
- First phase of MLS

EE
- GVRN NAT relief
- Performance enhancements

IPv6 support
- Large batch of applications
- QoS policy IPv6 support
- RIPv6 support

V1R6

FTP
- FTP client programming interface
- Modernized MBCS support

TN3270
- TN3270 server in separate address space
- SCS-style USS tables

Sysplex
- First phase of Sysplex autonomics
- Job-specific source IP address

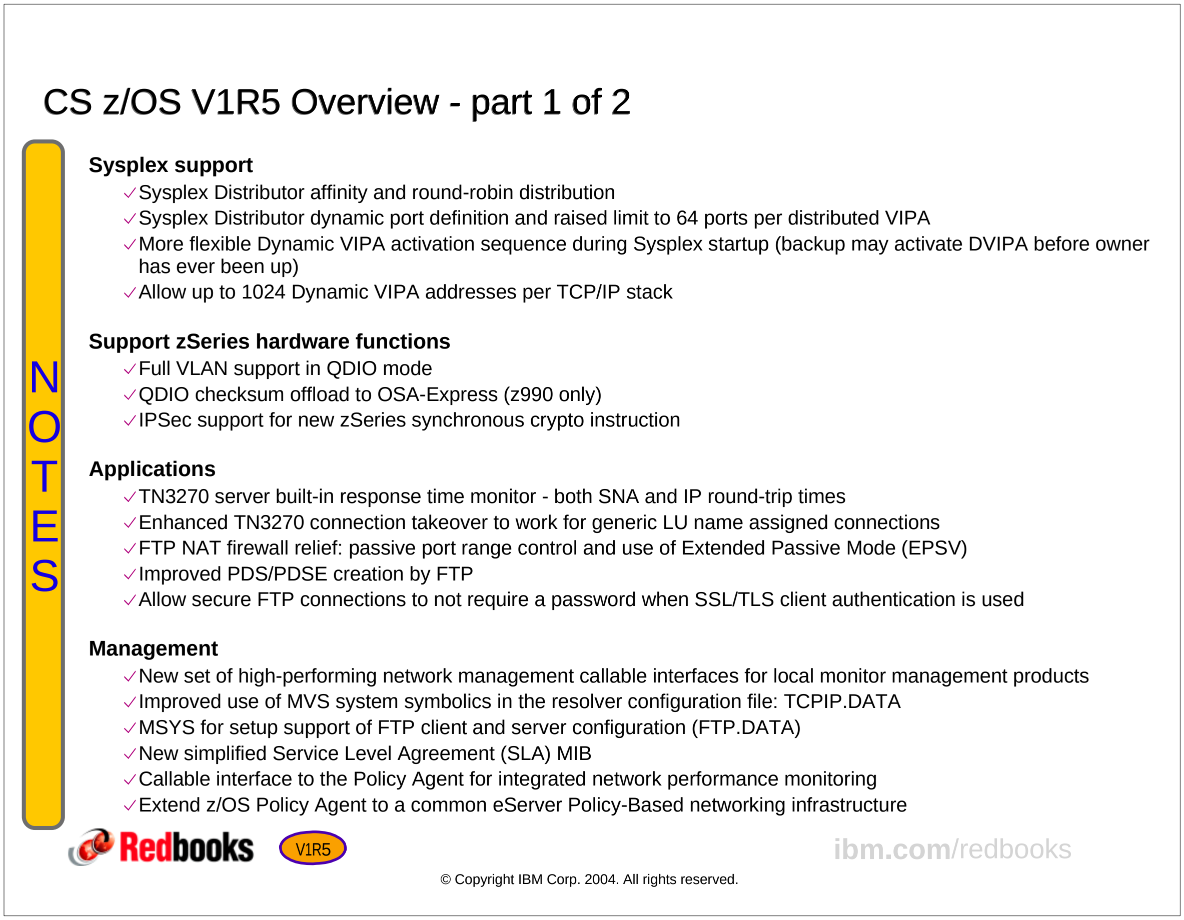
zSeries hardware
- 64-bit X-Windows and Motif

Management
- Netstat usability

Security
- Second phase of MLS support
- IDS eServer infrastructure

EE
- EBN session awareness
- EE connection network reachability awareness
- MNPS predatory takeover
- Enhanced addressing for RTP PUs

IPv6 support
- OSPF for IPv6
- Dynamic VIPAs
- Sysplex Distributor
- SourceVIPA
- Dynamic VIPA MIB support
- SysplexPorts
- IPv6 SAMEHOST
- IPv6 XCF
- IPv6 ESCON (MPC)



CS z/OS V1R5 Overview - part 1 of 2

NOTES

Sysplex support

- ✓ Sysplex Distributor affinity and round-robin distribution
- ✓ Sysplex Distributor dynamic port definition and raised limit to 64 ports per distributed VIPA
- ✓ More flexible Dynamic VIPA activation sequence during Sysplex startup (backup may activate DVIPA before owner has ever been up)
- ✓ Allow up to 1024 Dynamic VIPA addresses per TCP/IP stack

Support zSeries hardware functions

- ✓ Full VLAN support in QDIO mode
- ✓ QDIO checksum offload to OSA-Express (z990 only)
- ✓ IPSec support for new zSeries synchronous crypto instruction

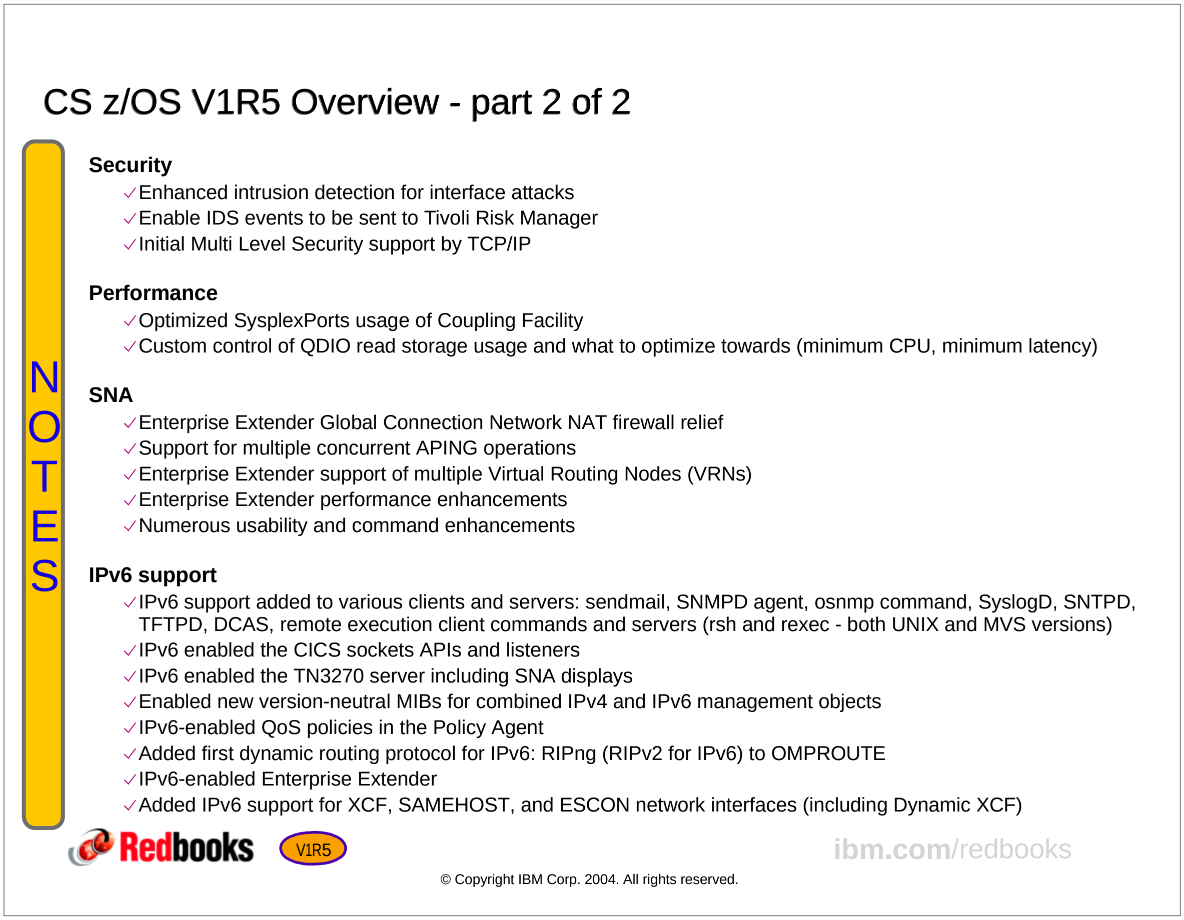
Applications

- ✓ TN3270 server built-in response time monitor - both SNA and IP round-trip times
- ✓ Enhanced TN3270 connection takeover to work for generic LU name assigned connections
- ✓ FTP NAT firewall relief: passive port range control and use of Extended Passive Mode (EPSV)
- ✓ Improved PDS/PDSE creation by FTP
- ✓ Allow secure FTP connections to not require a password when SSL/TLS client authentication is used

Management

- ✓ New set of high-performing network management callable interfaces for local monitor management products
- ✓ Improved use of MVS system symbolics in the resolver configuration file: TCPIP.DATA
- ✓ MSYS for setup support of FTP client and server configuration (FTP.DATA)
- ✓ New simplified Service Level Agreement (SLA) MIB
- ✓ Callable interface to the Policy Agent for integrated network performance monitoring
- ✓ Extend z/OS Policy Agent to a common eServer Policy-Based networking infrastructure

Redbooks V1R5 ibm.com/redbooks



CS z/OS V1R5 Overview - part 2 of 2

NOTES

Security

- ✓ Enhanced intrusion detection for interface attacks
- ✓ Enable IDS events to be sent to Tivoli Risk Manager
- ✓ Initial Multi Level Security support by TCP/IP

Performance

- ✓ Optimized SysplexPorts usage of Coupling Facility
- ✓ Custom control of QDIO read storage usage and what to optimize towards (minimum CPU, minimum latency)

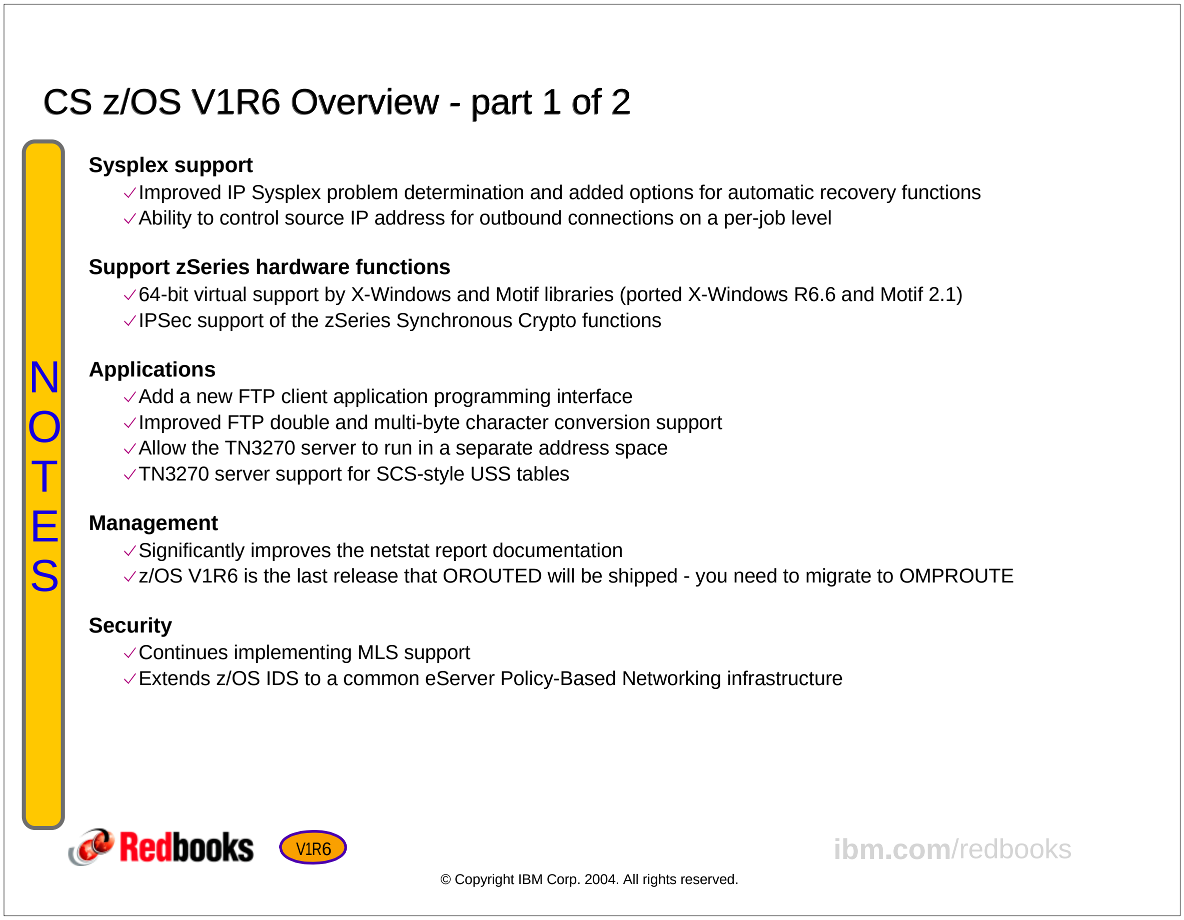
SNA

- ✓ Enterprise Extender Global Connection Network NAT firewall relief
- ✓ Support for multiple concurrent APING operations
- ✓ Enterprise Extender support of multiple Virtual Routing Nodes (VRNs)
- ✓ Enterprise Extender performance enhancements
- ✓ Numerous usability and command enhancements

IPv6 support

- ✓ IPv6 support added to various clients and servers: sendmail, SNMPD agent, osnmp command, SyslogD, SNTPD, TFTPD, DCAS, remote execution client commands and servers (rsh and rexec - both UNIX and MVS versions)
- ✓ IPv6 enabled the CICS sockets APIs and listeners
- ✓ IPv6 enabled the TN3270 server including SNA displays
- ✓ Enabled new version-neutral MIBs for combined IPv4 and IPv6 management objects
- ✓ IPv6-enabled QoS policies in the Policy Agent
- ✓ Added first dynamic routing protocol for IPv6: RIPng (RIPv2 for IPv6) to OMPROUTE
- ✓ IPv6-enabled Enterprise Extender
- ✓ Added IPv6 support for XCF, SAMEHOST, and ESCON network interfaces (including Dynamic XCF)

Redbooks V1R5 ibm.com/redbooks



CS z/OS V1R6 Overview - part 1 of 2

NOTES

Sysplex support

- ✓Improved IP Sysplex problem determination and added options for automatic recovery functions
- ✓Ability to control source IP address for outbound connections on a per-job level

Support zSeries hardware functions

- ✓64-bit virtual support by X-Windows and Motif libraries (ported X-Windows R6.6 and Motif 2.1)
- ✓IPSec support of the zSeries Synchronous Crypto functions

Applications

- ✓Add a new FTP client application programming interface
- ✓Improved FTP double and multi-byte character conversion support
- ✓Allow the TN3270 server to run in a separate address space
- ✓TN3270 server support for SCS-style USS tables

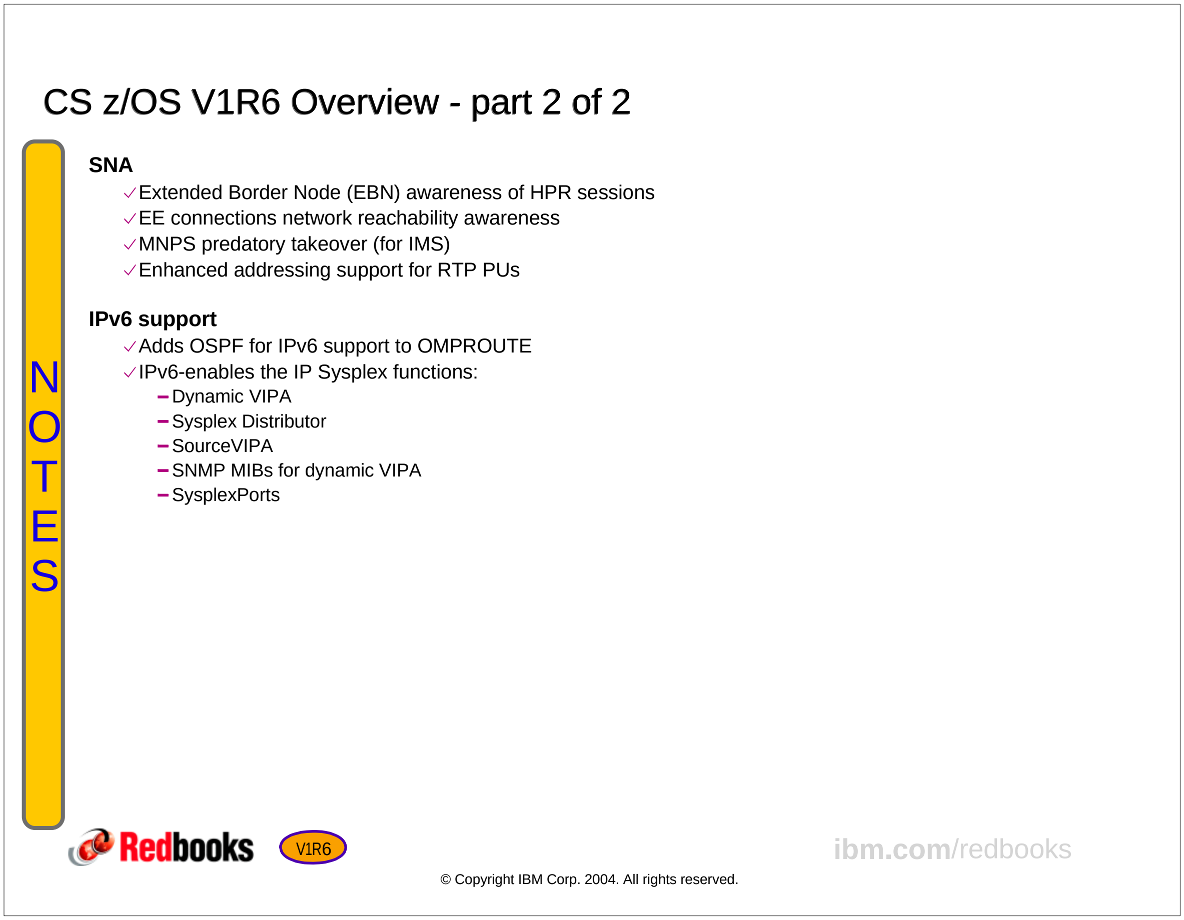
Management

- ✓Significantly improves the netstat report documentation
- ✓z/OS V1R6 is the last release that OROUTED will be shipped - you need to migrate to OMPROUTE

Security

- ✓Continues implementing MLS support
- ✓Extends z/OS IDS to a common eServer Policy-Based Networking infrastructure

Redbooks V1R6 ibm.com/redbooks



CS z/OS V1R6 Overview - part 2 of 2

NOTES

SNA

- ✓Extended Border Node (EBN) awareness of HPR sessions
- ✓EE connections network reachability awareness
- ✓MNPS predatory takeover (for IMS)
- ✓Enhanced addressing support for RTP PUs

IPv6 support

- ✓Adds OSPF for IPv6 support to OMPROUTE
- ✓IPv6-enables the IP Sysplex functions:
  - Dynamic VIPA
  - Sysplex Distributor
  - SourceVIPA
  - SNMP MIBs for dynamic VIPA
  - SysplexPorts

Redbooks V1R6 ibm.com/redbooks

ibm.com

What are the Major New Functions in CS z/OS V1R5 and V1R6?

Redbooks

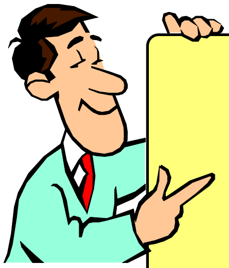
International Technical Support Organization

Objectives

The objectives of this session are to introduce the major functional enhancements that are delivered as part of the Communications Server in z/OS V1R5 and z/OS V1R6.

- In this session we will concentrate on the main enhancements that are not covered in any of the later sessions today.
- Not every functional enhancement will be discussed.
- The focus will be on the concepts and not the detailed configuration of the functions.
  - What are the functions that have been added, why have they been added, and when would you use those functions?

Agenda



1. Sysplex support
2. zSeries hardware support
3. Applications - TN3270 and FTP
4. Management
5. Security
6. SNA - APPN, HPR, Enterprise Extender
7. IPv6 support

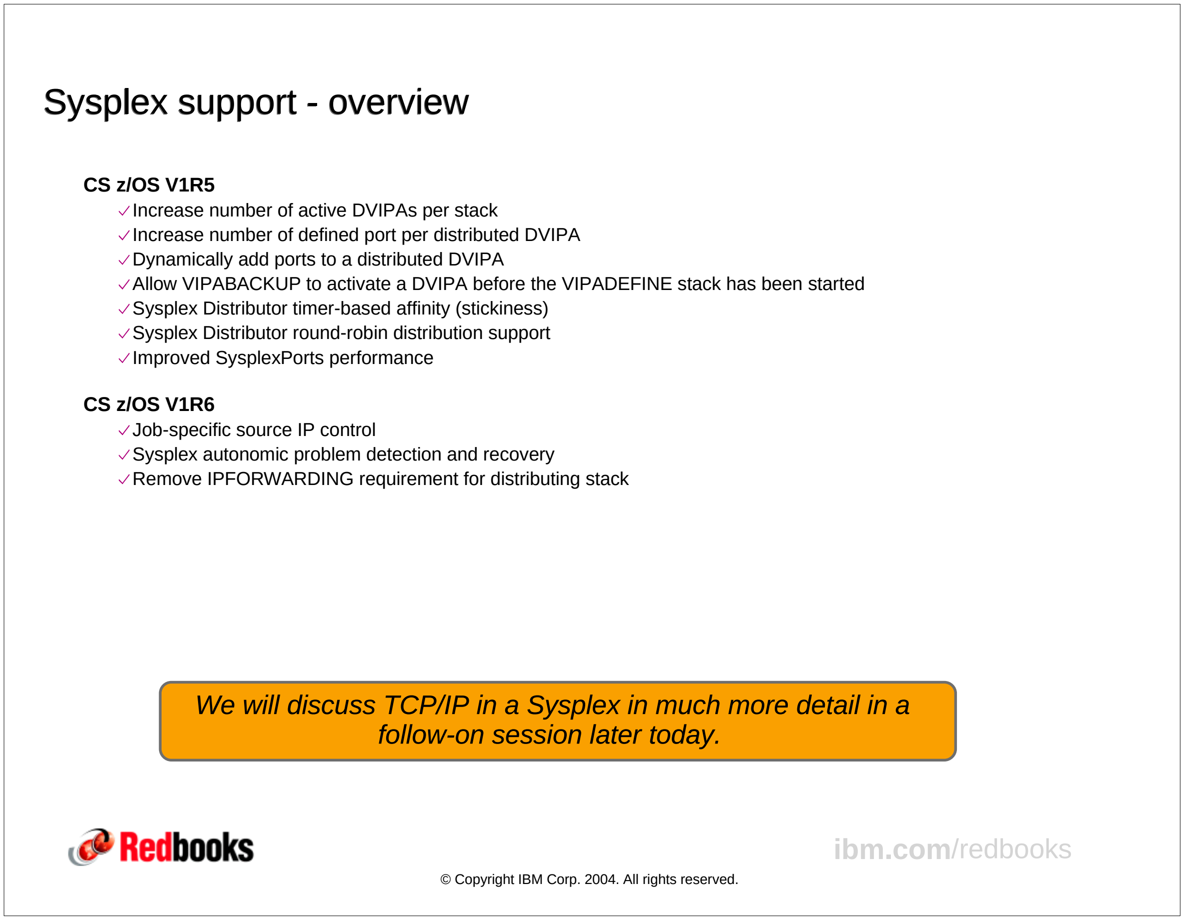
IBM®

What are the Major New Functions in CS z/OS V1R5 and V1R6?

-

Sysplex Support

International Technical Support Organization



Sysplex support - overview

CS z/OS V1R5

- ✓ Increase number of active DVIPAs per stack
- ✓ Increase number of defined port per distributed DVIPA
- ✓ Dynamically add ports to a distributed DVIPA
- ✓ Allow VIPABACKUP to activate a DVIPA before the VIPADEFINE stack has been started
- ✓ Sysplex Distributor timer-based affinity (stickiness)
- ✓ Sysplex Distributor round-robin distribution support
- ✓ Improved SysplexPorts performance

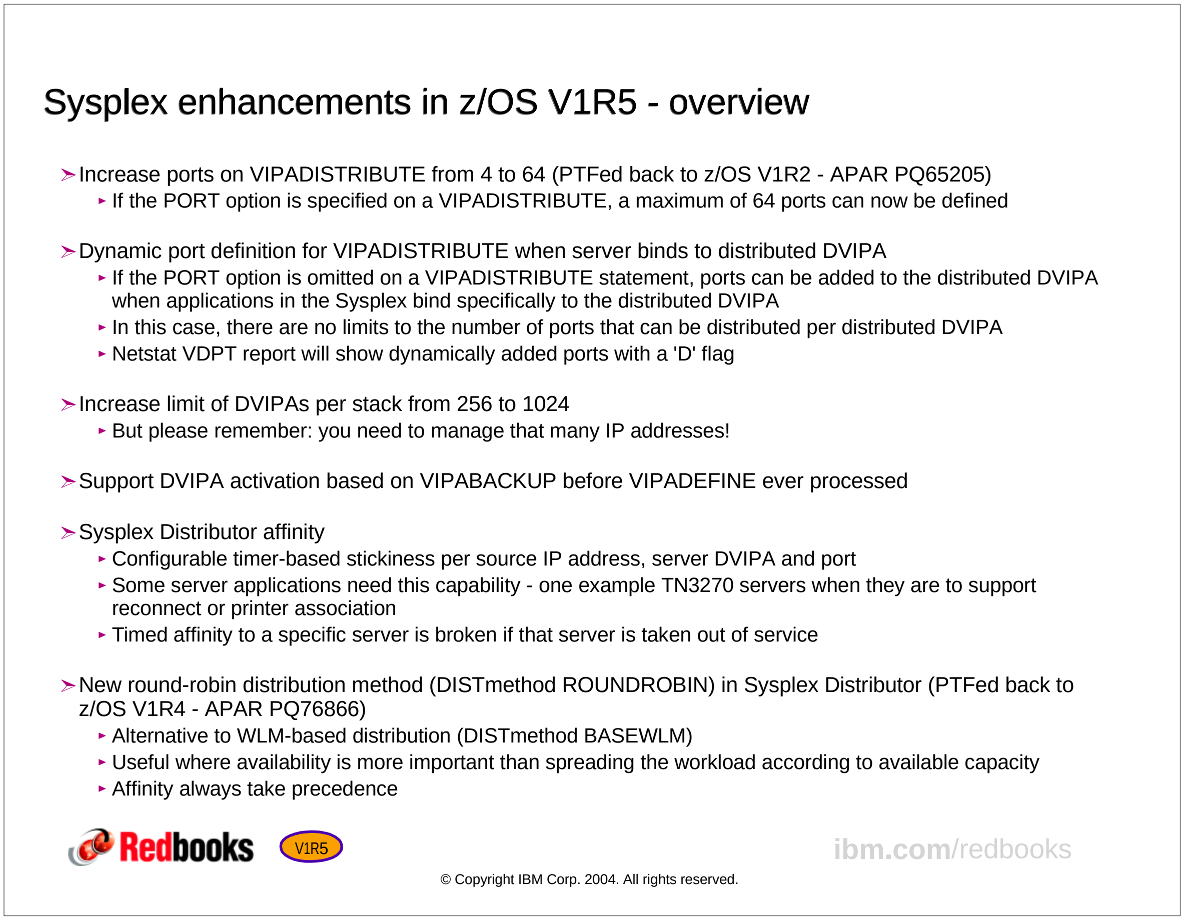
CS z/OS V1R6

- ✓ Job-specific source IP control
- ✓ Sysplex autonomic problem detection and recovery
- ✓ Remove IPFORWARDING requirement for distributing stack

We will discuss TCP/IP in a Sysplex in much more detail in a follow-on session later today.

Redbooks

ibm.com/redbooks



Sysplex enhancements in z/OS V1R5 - overview

- Increase ports on VIPADISTRIBUTE from 4 to 64 (PTFed back to z/OS V1R2 - APAR PQ65205)
  - If the PORT option is specified on a VIPADISTRIBUTE, a maximum of 64 ports can now be defined

- Dynamic port definition for VIPADISTRIBUTE when server binds to distributed DVIPA
  - If the PORT option is omitted on a VIPADISTRIBUTE statement, ports can be added to the distributed DVIPA when applications in the Sysplex bind specifically to the distributed DVIPA
  - In this case, there are no limits to the number of ports that can be distributed per distributed DVIPA
  - Netstat VDPT report will show dynamically added ports with a 'D' flag

- Increase limit of DVIPAs per stack from 256 to 1024
  - But please remember: you need to manage that many IP addresses!

- Support DVIPA activation based on VIPABACKUP before VIPADEFINE ever processed

- Sysplex Distributor affinity
  - Configurable timer-based stickiness per source IP address, server DVIPA and port
  - Some server applications need this capability - one example TN3270 servers when they are to support reconnect or printer association
  - Timed affinity to a specific server is broken if that server is taken out of service

- New round-robin distribution method (DISTmethod ROUNDROBIN) in Sysplex Distributor (PTFed back to z/OS V1R4 - APAR PQ76866)
  - Alternative to WLM-based distribution (DISTmethod BASEWLM)
  - Useful where availability is more important than spreading the workload according to available capacity
  - Affinity always take precedence

Redbooks V1R5

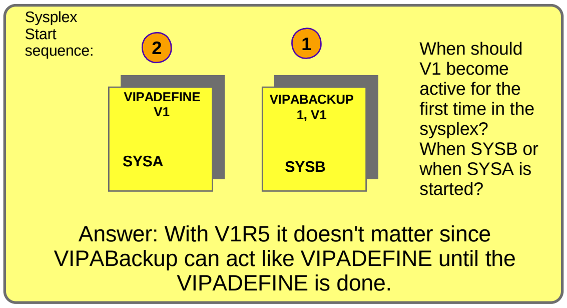
ibm.com/redbooks

VIPABACKUP before VIPADEFINE

- The original design for Dynamic VIPAs envisioned that the stack with the VIPADEFINE would be activated first, but this is not always the case.

- The VIPADEFINE profile statement contains definitions (MOVEABLE state and subnet mask) needed to activate the DVIPA. Currently, these are not present on the VIPABACKUP statement.

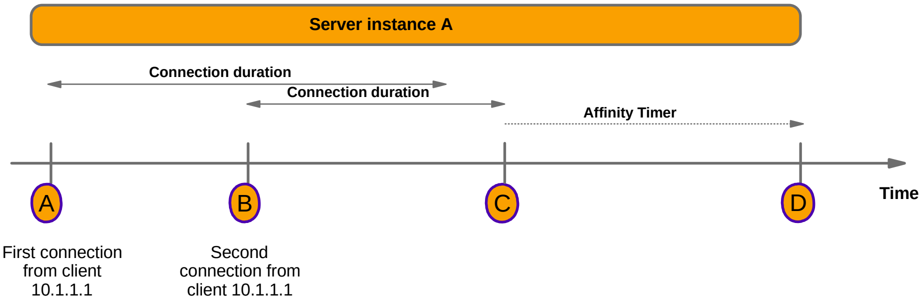
- z/OS V1R5 adds MOVEABLE, SERVICEMGR, and the subnet mask as optional parameters on the VIPABACKUP statement
  - The presence of MOVEABLE and a subnet mask designates that the DVIPA may be activated on this stack if it is not active elsewhere in the Sysplex
  - MOVEABLE, a subnet mask, and optionally SERVICEMGR, provide the required information for activation of the DVIPA.
  - If the DVIPA is already active elsewhere in the sysplex, these parameters are ignored and the DVIPA is configured as a backup DVIPA on this stack.



When should V1 become active for the first time in the sysplex? When SYSB or when SYSA is started?

Answer: With V1R5 it doesn't matter since VIPABackup can act like VIPADEFINE until the VIPADEFINE is done.

Sysplex Distributor timed affinity - when does the timer start?

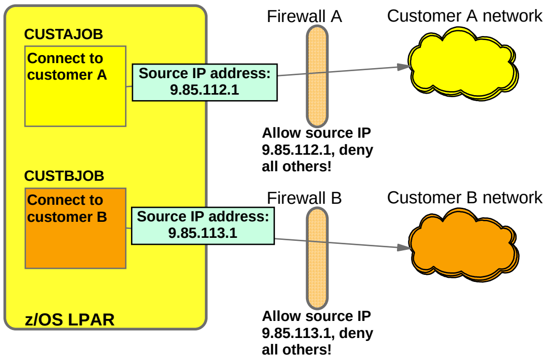


- If a client already has a connection with a timed affinity server instance, new connections from that same client will go to that server instance.
  - At time B, a second connection from the client at 10.1.1.1 arrives. Since that client already has a connection with server instance A, this second connection goes to the same server instance A.

- The affinity timer runs from the point in time the last connection from a given client ended.
  - At time C, the last connection from client 10.1.1.1 ends. The time between C and D is the affinity timer. If any new connections arrive from the client before time D, they will be sent to server instance A.

- Pay attention to nodes in your network that may make it look as if all client connections come from one (or a few) source IP addresses:
  - SOCKS servers, HTTP(S) proxy servers, NAT firewalls that perform client NATing

Job-specific source IP address control added in V1R6 for easier firewall filter rule administration


CUSTAJOB
Connect to customer A
Source IP address: 9.85.112.1
Firewall A
Customer A network
Allow source IP 9.85.112.1, deny all others!
CUSTBJOB
Connect to customer B
Source IP address: 9.85.113.1
Firewall B
Customer B network
Allow source IP 9.85.113.1, deny all others!
z/OS LPAR

```
SRCIP
     CUSTAJOB   9.85.112.1
     CUSTBJOB   9.85.113.1
     User1*     888:555::222    ===> Wildcards allowed!
ENDSRCIP
```

Extending configuration control over which local IP address to use for outbound connections from z/OS

- ✓ Outbound connections can use same IP addresses as inbound connections to same application without application change:
  - Easier for accounting and management
  - Easier for security (firewall admin)
  - Permits source IP address selection controls for applications even when application doesn't provide for this programmatically (most don't, but some do!)
- ✓ Introducing job-specific source IP addressing
  - A new TCPIP.Profile statement SRCIP/ENDSRCIP allows the selection of a source IP address for outbound TCP connections by job name
  - Overrides TCPSTACKSOURCEVIPA and SOURCEVIPA specifications

Redbooks V1R6 ibm.com/redbooks

IBM®

What are the Major New Functions in CS z/OS V1R5 and V1R6? - zSeries Hardware Support

Redbooks

International Technical Support Organization

zSeries hardware support - overview

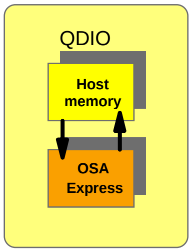
CS z/OS V1R5

- ✓OSA-Express performance control options
- ✓QDIO checksum offload to OSA-Express (z890 and z990 only)
- ✓Full VLAN support in QDIO mode

CS z/OS V1R6

- ✓64-bit virtual support by X-Windows and Motif libraries (ported X-Windows R6.6 and Motif 2.1)
- ✓IPSec support of the zSeries Synchronous Crypto functions
- ✓OSA-E direct SNMP support extended to also include LCS interfaces

Queued Direct IO via OSA-Express on z890/z990 - overview



QDIO characteristics

- ➢IP Only - use Enterprise Extender for QDIO advantages to SNA
- ➢IP-Assist to handle MAC addressing, ARP processing, some filtering
- ➢Dynamically maintains the OSA Address Table in a shared environment
- ➢Supports high-speed LPAR-to-LPAR communication
- ➢Direct Memory Access (DMA) protocol reduces IO interrupts
- ➢Supports outbound priority queuing (four priority queues - control via Policy Agent QoS)
- ➢QDIO devices support multicast and from z/OS V1R4 also broadcast
- ➢Only QDIO is capable of delivering media speed on Gigabit Ethernet
- ➢z/OS V1R5 provides full VLAN support (switch port in trunk mode)
- ➢z/OS V1R5 provides improved performance controls and use of checksum offload

z890/990 OSA-Express adapter port connectivity	IP LCS	SNA LSA	IP QDIO	IP QDIO Check-sum offload support	IP QDIO Jumbo frame support
Token-ring - 4 Mbit	X	X	X		
Token-ring - 16 Mbit	X	X	X		
Token-ring - 100 Mbit	X	X	X		
Ethernet - 10 Mbit - Copper	X	X	X		
Ethernet - 100 Mbit (fast ethernet) - Copper	X	X	X		
Ethernet - 1000 Mbit (1000BASE-T) - Copper	X	X	X	X	X
Gigabit Ethernet - Fiber			X	X	X

QDIO and iQDIO (HiperSockets) read storage use control

- Each OSA-Express QDIO device and HiperSockets device requires fixed storage for read processing
- VTAM provides start options to configure the amount of read storage but these settings apply globally
  - QDIOSTG applies to all OSA-Express QDIO devices
  - IQDIOSTG applies to all HiperSockets devices
- z/OS V1R5 provides new keywords in the TCP/IP profile to override the global VTAM read storage setting for a specific OSA-Express QDIO or HiperSockets device
  - The Communications Server uses CSM (Communications Storage Manager) dataspace buffers backed by 64-bit real storage for this read storage.
- Can specify one of the following values:
  - **GLOBAL**
    - The amount of storage is determined by the QDIOSTG or IQDIOSTG VTAM start option. This is the default.
  - **MAX**
    - You expect a heavy inbound workload over this adapter
    - QDIO: 4 MB, iQDIO: 7.8 MB
  - **AVG**
    - You expect a medium inbound workload over this adapter
    - QDIO: 2 MB, iQDIO: 6 MB
  - **MIN**
    - You expect a light inbound workload over this adapter
    - QDIO: 1 MB, iQDIO: 4 MB
- For HiperSockets, the settings only affect devices with 64K frame size

V1R5 ibm.com/redbooks

QDIO inbound performance

- The performance of an OSA-Express QDIO device is impacted by how frequently the OSA interrupts the host to process inbound packets
  - More frequent interruptions lead to minimized latency but increased CPU consumption
  - Less frequent interruptions lead to decreased CPU consumption but increased latency
- New options are provided in the TCP/IP profile to specify the desired inbound performance behavior from an OSA-Express in QDIO mode
  - PTFed back to z/OS V1R4 via APAR PQ92262
- Can specify one of the following values:
  - **MINCPU** - instructs the adapter to minimize host interrupts, thereby minimizing host CPU consumption. This mode of operation may result in minor queuing delays for packets into the host, and is not recommended for workloads with demanding latency requirements.
  - **MINLATENCY** - instructs the adapter to minimize latency, by immediately presenting received packets to the host. This mode of operation will generally result in higher CPU consumption than the other two settings, and is recommended only for workloads with demanding latency requirements. This setting should only be used if host CPU consumption is not an issue.
  - **BALANCED** (default) - instructs the adapter to strike a balance between MINCPU and MINLATENCY

Processor	Microcode level
G5/G6	4.28
zSeries 2064 GA2	2.29
zSeries 2064 GA3	3.23

V1R5 ibm.com/redbooks

QDIO checksum offload

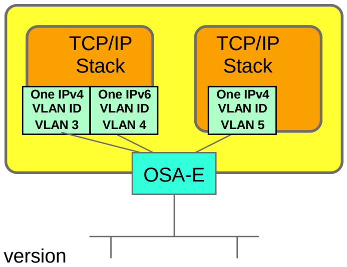
- Offloads most IPv4 checksum processing to OSA-Express in QDIO mode
- Provides improved performance for IPv4 traffic
- Enabled by the TCP/IP stack via control flows to the OSA-Express adapter (added in z/OS V1R5)
- Supported on the following OSA-Express features (which require an IBM eServer zSeries 890 or 990):
  - Feature # 1364   GbE LX
  - Feature # 1365   GbE SX
  - Feature # 1366   1000BASE-T Ethernet when configured to operate at 1 Gbps
- Only applies to IPv4 packets
- Only applies to packets that go onto the LAN
  - Not to packets that are passed back into the zSeries to an LPAR sharing the OSA-Express LAN port
- Applies to TCP, UDP, and IP header checksums
- Applies to both inbound and outbound
- Exceptions
  - Fragmentation/reassembly
  - IPSec
  - Packets between 2 stacks sharing the OSA
  - Outbound multicast and broadcast
  - Some outbound TCP control packets (e.g. SYN, RST)

V1R5

ibm.com/redbooks

Full QDIO VLAN support in z/OS V1R5

- Provides full VLAN support for OSA-Express QDIO by allowing VLAN ID to be configured
- Supported on these OSA-Express features (which require an IBM eServer zSeries 890 or 990):
  - Gigabit Ethernet (feature #s 2364, 2365, 1364, 1365)
  - Fast Ethernet (feature # 2366)
  - 1000BASE-T Ethernet (feature # 1366)
- Also supported on the following OSA-Express features (at system driver level 3G) on an IBM eServer zSeries 800 or 900:
  - Gigabit Ethernet
  - Fast Ethernet
- Conforms to the IEEE 802.1Q standard
  - VLAN tag in MAC header contains VLAN ID
- Allows VLAN priority tagging to be used with VLAN ID
- A global VLAN ID must be configured in the TCP/IP profile
  - Each stack can only specify one VLAN ID for each OSA-Express port per IP version
    - A stack can specify separate VLAN IDs for IPv4 and IPv6 for the same OSA-Express port
  - Each stack sharing an OSA-Express port may specify a different VLAN ID
- z/OS V1R4 PTF coming (APAR PQ86508)
- Multiple OSA PRIRouters supported (PRIRouter per global VLANID support)



V1R5

ibm.com/redbooks

VLAN tagging basics

Two types of frames in a VLAN environment:

- Untagged frame
  - No tag header following the source MAC address
- Tagged frame
  - Priority-tagged frame
    - Tag header includes only VLAN priority information, but no VLAN ID (VLAN ID is zero and is referred to as a null-tagged frame)
  - VLAN-tagged frame
    - Tag header includes both VLAN priority information and VLAN ID

Dest MAC address	**Source MAC address**	**Type/Length**

Ethernet layer-2 Header, untagged

Dest MAC address	**Source MAC address**	**Tag Control info**	**Type/Length**

Ethernet layer-2 Header, tagged

VLAN Tag x'8100'	**3-bit Priority**	**1-bit Canonical Always zero**	**12-bit VLAN ID**

Tag Control Information

Redbooks V1R5

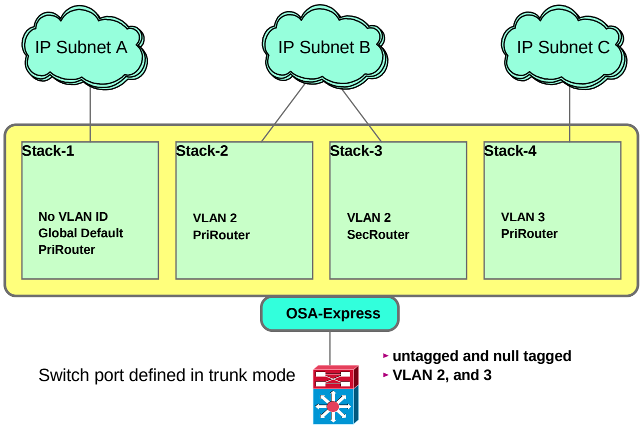
VLANs and OSA PRI/SEC router support

- The VLANID parameter of the LINK and INTERFACE statements interacts with the PRIRouter and SECRouter parameters on the DEVICE and INTERFACE statements.
  - If you configure both a VLANID and either PRIRouther or SECRouter, then this TCP/IP instance will act as a router for this VLAN only. Frames that are received at this device for an unknown IP address will only be routed to this TCP/IP instance if they are VLAN tagged with this VLAN ID.
  - If you do not configure a VLAN ID, but do configure PRIRouter or SECRouter, then this TCP/IP instance will act as a "Global default Pri/SecRouter" for all inbound unicast frames with an unknown destination IP address

✓ Stack-1 acts as PriRouter for:
- Untagged frames
- Frames tagged with a null VLAN ID
- Frames tagged with an unregistered (anything but VLAN 2 or 3) VLAN ID
- Frames tagged with a registered VLAN ID, an unknown destination IP address but no VLAN Pri/SecRrouter

✓ Stack-2 acts as PriRouter for frames tagged with VLAN 2 while Stack-3 acts as SecRouter for that same VLAN.

✓ Stack-4 acts as PriRouter for frames tagged with VLAN 3



Redbooks V1R5

General VLAN network design considerations

1. When using VLAN IDs in any TCP/IP stack sharing an OSA port, the switch port to which the OSA port is attached should be configured in trunk mode.

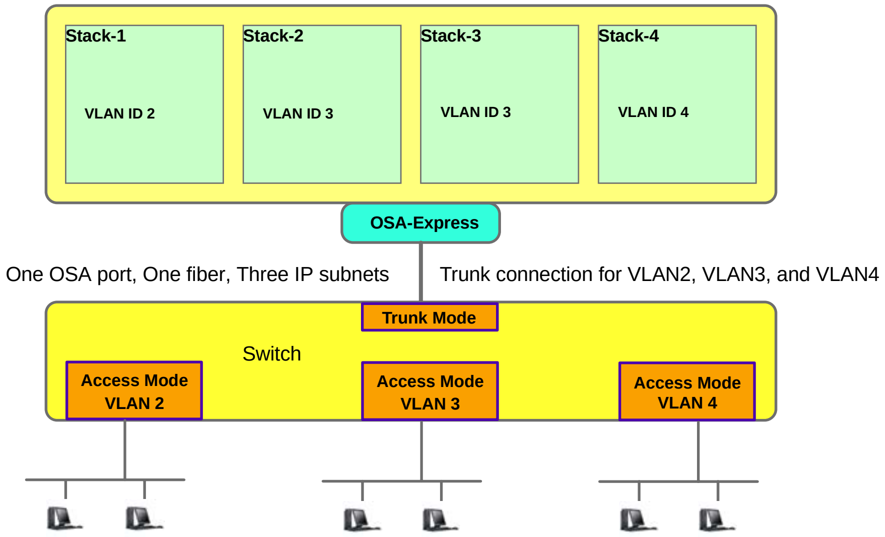
2. When not using VLAN IDs in any TCP/IP stack sharing an OSA port, the switch port to which the OSA port is attached should be configured in access mode.

3. When a TCP/IP stack uses multiple OSA ports all to the same LAN, and a VLAN ID is used on one of those ports, VLAN IDs should be used on all ports to that same LAN.

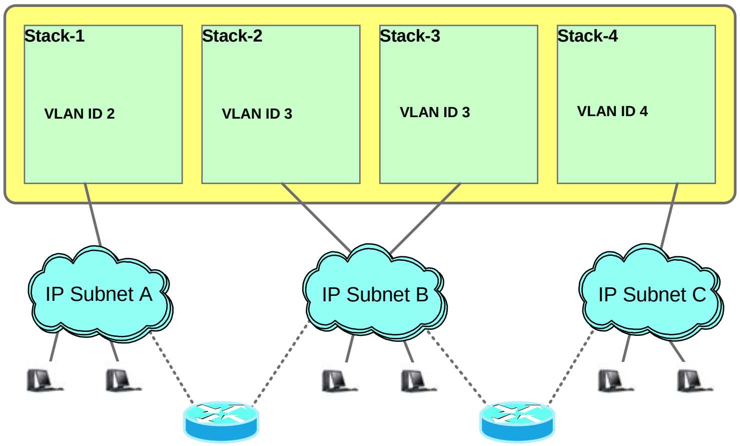
4. Some switch vendors use VLAN ID 1 for special purposes. VLAN ID 1 should be avoided when designing VLAN-based networks.

... and please make sure your networking staff creates and maintains both a physical network diagram and a logical network diagram - they look very different when you work with VLAN configurations.

Physical network connectivity diagram - Multiple stacks, separate subnets, single OSA port



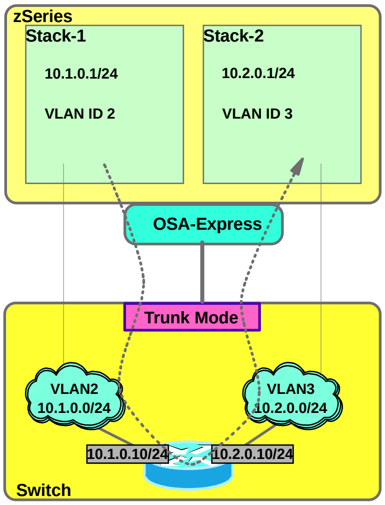
Logical network diagram - Multiple stacks, separate subnets, single OSA port



- Depending on switch configuration, the switch may interconnect the VLANs using a layer-3 IP router function.
- The subnets may belong to different routing domains or OSPF areas.
  - Test, production, demo
- The subnets may belong to different security zones.
  - Intranet, DMZ

Redbooks V1R5 ibm.com/redbooks

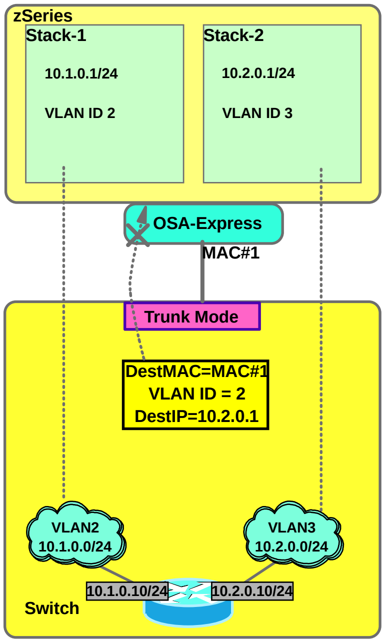
Sharing OSA ports between stacks belonging to different VLAN IDs



- Stack-1 has a global VLAN ID of 2 configured on its IPv4 LINK statement.
- Stack-2 has a global VLAN ID of 3 configured on its IPv4 LINK statement.
- Stack-1 has a routing table entry that points to 10.1.0.10 for forwarding to the 10.2.0.0/24 (VLAN ID 3) subnet.
- Stack-1 sending an IP packet with a destination IP address in the IP header of 10.2.0.1 - tagged with VLAN ID 2 - with next hop IP address 10.1.0.10 (the router interface on VLAN ID 2).
- Which path will this packet take?
  - To the OSA adapter, then on to the router at 10.1.0.10, then routed over the router's other interface to VLAN ID 3, back up to the OSA adapter, and then to Stack-2.
- This is what you want! The reasons for using VLANs is often security - separating different subnets via a router that applies IP filters before forwarding packets.

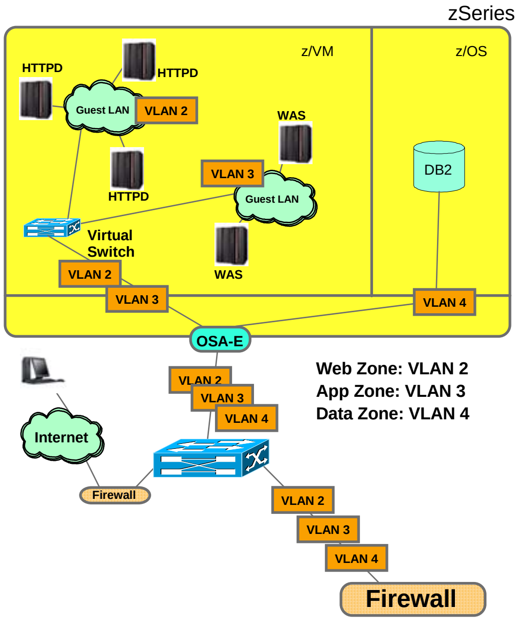
Redbooks V1R5 ibm.com/redbooks

Inbound processing - unicast frames



- Stack-1 has a global VLAN ID of 2 configured on its IPv4 LINK statement.
- Stack-2 has a global VLAN ID of 3 configured on its IPv4 LINK statement.
- Assume that a unicast frame with an IP packet arrives at MAC#1 with the following characteristics (please note that this could only be the case if a rogue node was attached to the switch using a trunk interface and the node manipulated the frame contents - also known as a hacker!):
  - Destination IP address 10.2.0.1
  - VLAN ID tag of VLAN 2
- What will the OSA adapter do with that IP packet?
  - Discard it since the VLAN ID of the registered IP address (10.2.0.1 - VLAN ID 3) doesn't match the VLAN tag in the frame (VLAN ID 2)

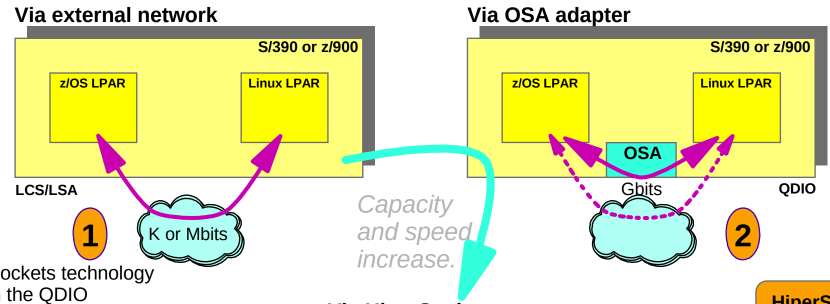
Combining outside firewalls, z/VM virtual switch, and VLAN technology

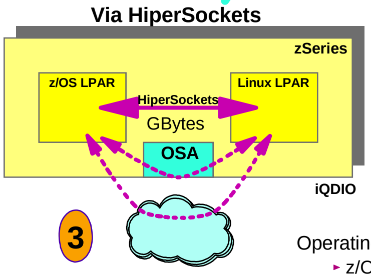


Web Zone: VLAN 2
App Zone: VLAN 3
Data Zone: VLAN 4

- z/VM's virtual switch is VLAN-aware and supports switching different VLAN IDs to different guest LANs.
- A single OSA adapter connected to a VLAN-aware external switch using a trunk mode connection can now serve all three security zones
- z/VM's virtual switch can be attached to up to three physical OSA adapters for scaling purposes if capacity requirements exceed those of a single network interface
- z/VM's virtual switch supports Ethernet IPv4 (QDIO)
  - Currently no support for other networking protocols, such as SNA, NETBIOS, or IPv6
- The virtual switch does depend on z/VM TCP/IP for initial QDIO control flows (known as the controller), but z/VM TCP/IP is not involved in data transfer via the virtual switch
- VLAN technology in combination with z/VM's virtual switch offers significantly simplified network connectivity options for a complex zSeries server environment

LPAR to LPAR IP communication - HiperSockets





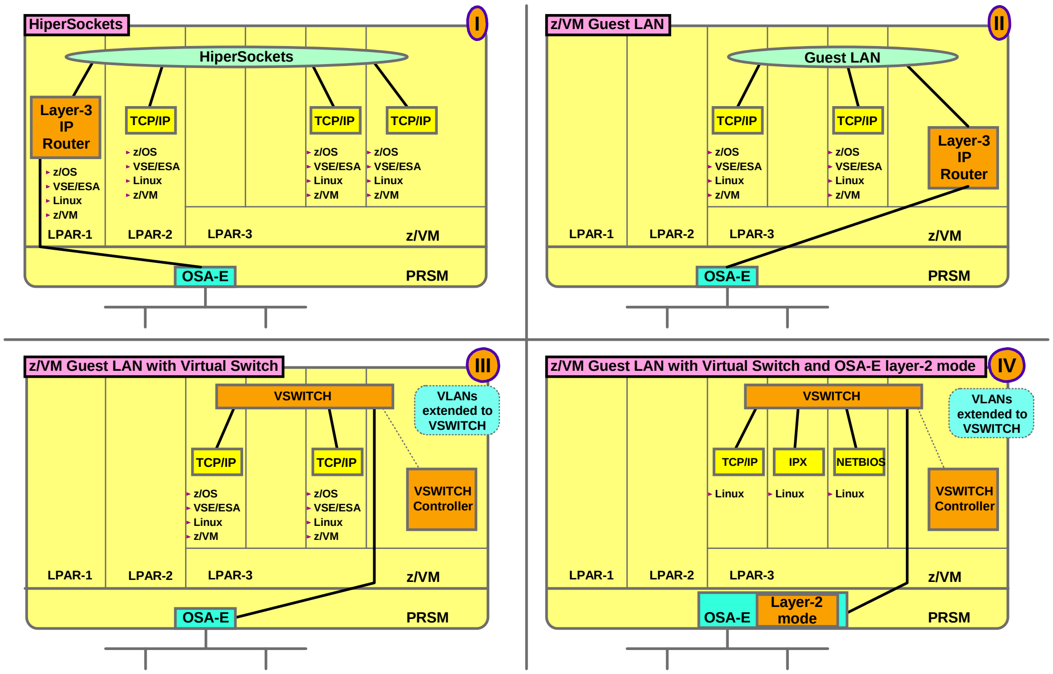
- The HiperSockets technology is based on the QDIO technology from the OSA Express adapter. For HiperSockets, it is referred to as internal QDIO - or for short iQDIO.
- A HiperSockets network looks like an internal LAN
- A z800 or z900 supports up to four independent HiperSockets networks (internal LANs), while a z890 or z990 supports up to 16 such internal networks.
- Broadcast support was added for HiperSockets in z/OS V1R5

HiperSockets on zSeries allows operating system images on the same processor complex to exchange IP traffic virtually at memory speed.

Operating system support for HiperSockets:

- z/OS 1.2 and higher,
- Linux kernel 2.4 or later,
- z/VM 4.2 or later,
- VSE/ESA 2.7 or later

Overview of internal LAN technologies on zSeries



Comparison matrix - zSeries internal LAN options

	HiperSockets	z/VM Guest LAN	z/VM with VSWITCH (z/VM 4.4)	z/VM with VSWITCH and OSA layer-2 mode
zSeries OS support	z/VM z/OS VSE/ESA Linux	z/VM z/OS VSE/ESA Linux	z/VM z/OS VSE/ESA Linux	Linux
Network protocol support	IPv4	IP4 and IPv6	IPv4	All protocols: IPv4, IPv6, IPX, NETBIOS, SNA, etc.
Device driver support	iQDIO	iQDIO or QDIO	QDIO	QDIO (For an OS to support other protocols than IP, its QDIO device driver must be able to handle protocols other than IP)
Need for mainframe layer-3 router to outside network	Required (z/OS and Linux support accelarated routing to/from HiperSockets)	Required	None (z/VM TCP/IP stack required for control, not layer-3 routing)	None (z/VM TCP/IP stack required for control, not layer-3 routing)
LPAR or z/VM guest support	Both	z/VM guests in a z/VM LPAR	z/VM guests in a z/VM LPAR	z/VM guests in a z/VM LPAR
Extend IEEE802.1q VLAN	No	No	Yes (for IPv4 frames)	Yes (for all protocol frames)

Redbooks

ibm.com/redbooks

IBM®

What are the Major New Functions in CS z/OS V1R5 and V1R6?

-

Communications Server for z/OS Provided Applications

Redbooks

International Technical Support Organization

Applications - overview

CS z/OS V1R5

- ✓ TN3270 server built-in response time monitor - both SNA and IP round-trip times
- ✓ Enhanced TN3270 connection takeover to work for generic LU name assigned connections
- ✓ TN3270 server has been IPv6 enabled
- ✓ Improved configuration controls for when TN3270 server should send "unlock keyboard" to client
- ✓ Improved TN3270 server configuration: IP range specification - use a range instead of a subnet mask
- ✓ Added support for network access control of TN3270 data flows
- ✓ FTP NAT firewall relief: passive port range control and use of Extended Passive Mode (EPSV)
- ✓ Improved PDS/PDSE creation by FTP
- ✓ Allow secure FTP connections to not require a password when SSL/TLS client authentication is used
- ✓ Cleaned up batch FTP client return codes
- ✓ MSYS for setup for FTP.DATA
- ✓ FTP server load modules now shipped as RMODE=ANY
  - (may impact your FTP server security exit routines)

CS z/OS V1R6

- ✓ Adds a new FTP client application programming interface
- ✓ Improved FTP double and multi-byte character conversion support
- ✓ Allow the TN3270 server to run in a separate address space
- ✓ TN3270 server support for SCS-style USS tables
- ✓ Control if connection takeover is allowed when client IP address changes

Redbooks

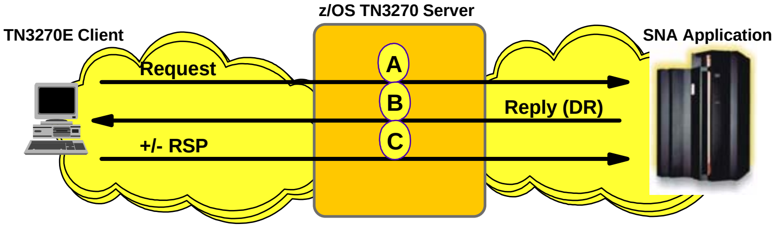
ibm.com/redbooks

TN3270 response time monitoring

- The TN3270 server in z/OS V1R5 implements a built-in TN3270 round-trip response time monitor function.
  - **Time A** - when request is sent to the SNA application
  - **Time B** - when reply is sent to the TN3270E client
  - **Time C** - when response is received from TN3270E client

- Data maintained:
  - Life of connection averages
  - Sliding window averages
  - Round-trip response times by time buckets
- Data reported via:
  - SNMP subagent
  - D TCPIP,,TELNET commands

```
Round-trip response time = Time C - Time A
IP response time         = Time C - Time B
SNA response time        = Round trip response time - IP response time
```



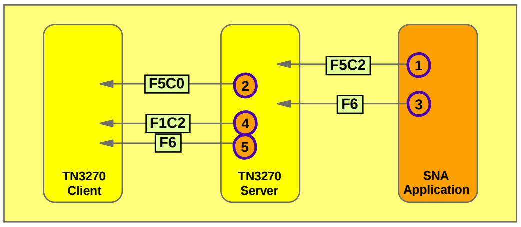
Redbooks V1R5

ibm.com/redbooks

"Stubborn" clients that don't know when to unlock the keyboard

Current TN3270 server process that works for (almost) all TN3270(E) clients:

1. F5C2 (Only in chain, no CDI, no EB). Application sends unlock keyboard, but no CDI or EB, so TN3270 server turns off unlock keyboard and remembers that the application sent it.
2. F5C0 (Unlock keyboard turned off).
3. F6 (Only in chain, CDI). Read buffer, read modified, or read partition query from SNA application.
4. TN3270 server first sends an unlock keyboard to properly finish the previously sent data
5. TN3270 server then sends the read modified



- In some cases, type-ahead has buffered data and the unlock keyboard (4) allows the client to send the data to Telnet which forwards it to the application. The application assumes this is the response to the read modified and issues errors. In this case, the unlock keyboard should have been sent AFTER the read modified was sent.
- Some applications do not send data to the client first. They expect data from the client after sending the BIND.
  - TN3270E connections send the Bind to the client alerting it to unlock the keyboard.
  - Telnet can not send a BIND to the client over a TN3270 connection. In this case the client keyboard remains locked and the end user can not send data to the application. Because the application is waiting for data, the session is essentially hung.

Two new sets of options in z/OS V1R5:

1. Send unlock keyboard before or after a READ command has been received from the SNA application
   - READ Buffer, READ Modified, READ partition query
   - Does not apply if SNA extensions have been negotiated on a TN3270E connection (the TN3270E commands for keyboard restore (KRI) and start data indicator (SDI) are used instead in that case)
2. Send unlock keyboard after BIND
   - Only applies to TN3270 connections (not TN3270E)

If you have problems with specific clients that occasionally hang, you may want to experiment with these options in a PARMSGROUP that is mapped to those specific client identifiers. APARs for OS/390 V2R10 - PQ63027, z/OS V1R2 - PQ63027 & PQ67798, and z/OS V1R4 - PQ63027 & PQ67798

```
UNLOCKKEYBOARD
    BEFOREREAD|AFTERREAD
    TN3270BIND|NOTN3270BIND
```

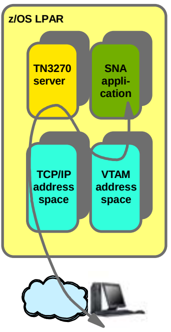
V1R5

ibm.com/redbooks

TN3270 server in its own address space in z/OS V1R6 for improved management

Remote terminal access

- z/OS V1R6 allows you to choose where to run the TN3270 server:
  - Run the TN3270 server as a separately started address space from TCP/IP
  - Continue to run TN3270 server as a subtask of the TCP/IP address space
- Reasons why an installation may want to run the TN3270 server in a separate address space:
  - Allows for prioritzation of TCP/IP address space vs TN3270 server
  - Much less likely for TN3270 server failure to cause a total TCP/IP failure
  - Allow for easier problem diagnosis for both TCP/IP and TN3270
  - Easier controls for starting and stopping the server
- Considerations
  - Profile statements are the same (minor considerations) and must be in a file separate from TCP/IP
  - Commands are the same but must be directed to the intended TN3270 procedure name
  - Multiple TCP/IP stacks supported
    - One server per stack (affinity)
    - One server associated with all stacks (Generic Server)
  - Multiple TN3270 server address spaces supported
    - Max 8 TN3270 server address spaces per LPAR
    - Only one can activate the TN3270 response time SNMP subagent in a stack
      - Must have stack affinity to that stack
      - The first one started with stack affinity and TNSACONFIG enabled activates the SNMP subagent
  - Must run TN3270 server with affinity for the following functions
    - TN3270 response time SNMP subagent
    - WLM function
  - Requirements
    - Separate startup JCL. Sample is provided.



V1R6

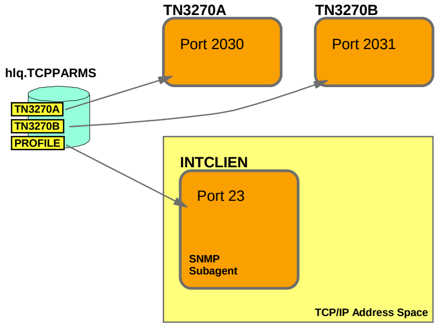
ibm.com/redbooks

Multiple TN3270 server address spaces

Per TN3270 server address space:

- Define JCL procedure - sample in hlq.SEZAINST(EZBTNPRC)
- Define started task RACF profile (assign started task user ID)
  - Started task user ID must be UID 0 or permitted to BPX.SUPERUSER and have OMVS segment
- Define TN3270 server definitions
  - Same as prior to z/OS V1R6
  - One new global option: TCPIPJOBNAME for stack affinity
  - Each server has its own set of definitions
    - Different key ring files per server address space

```
//TN3270A  PROC PARMS='CTRACE(CTIEZBTN)'
//TN3270   EXEC PGM=EZBTNINI,REGION=0M,PARM='&PARMS'
//*
//SYSPRINT DD SYSOUT=*,DCB=(RECFM=VB,LRECL=132,BLKSIZE=136)
//SYSOUT   DD SYSOUT=*,DCB=(RECFM=VB,LRECL=132,BLKSIZE=136)
//CEEDUMP  DD SYSOUT=*,DCB=(RECFM=VB,LRECL=132,BLKSIZE=136)
//*
//*TNDBCSCN DD DISP=SHR,DSN=TCPIP.SEZAINST(TNDBCSCN)
//*TNDBCSXL DD DISP=SHR,DSN=TCPIP.SEZAXLD2
//*TNDBCSER DD SYSOUT=*
//*
//PROFILE  DD DSN=USER1.TCPCS.TCPPARMS(TN3270A),DISP=SHR
//SYSTCPD  DD DSN=USER1.TCPCS.TCPPARMS(TCPDATA),DISP=SHR
```

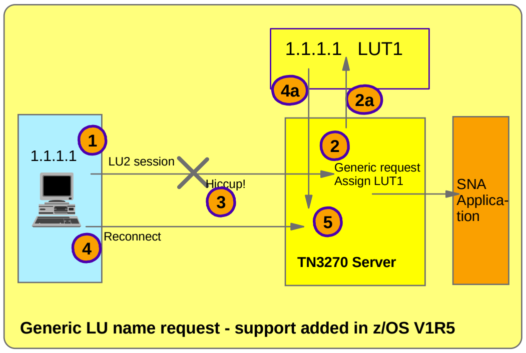


Normal MVS console TN3270 display commands - directing them to selected server address space name:

```
D TCPIP,TN3270A,T,CONN,CONN=24D
EZZ6065I TELNET CONNECTION DISPLAY 531
  CONNECTED: 12:31:28  06/28/2004  STATUS: SESSION PENDING
  CLIENT IDENTIFIER FOR CONN: 0000024D   SECLABEL: **N/A**
    CLIENTAUTH USERID: **N/A**
    HOSTNAME: NO HOSTNAME
    CLNTIP..PORT: ::FFFF:9.65.211.157..1484
    DESTIP..PORT: ::FFFF:9.42.105.45..2030
    LINKNAME: QDIO4
  PORT:  2030 QUAL: NONE
    AFFINITY: TCPCS
    STATUS: ACTIVE  BASIC         ACCESS: NON-SECURE
  PROTOCOL: TN3270E           DEVICETYPE: IBM-3278-3-E
    TYPE: TERMINAL GENERIC
    OPTIONS: ETET----   3270E FUNCTIONS: BSR----
                       NEWENV FUNCTIONS: --
  LUNAME: TCPABC50
[lines deleted]
```

Redbooks V1R6 ibm.com/redbooks

TN3270 server reconnect - generic LU requests support added



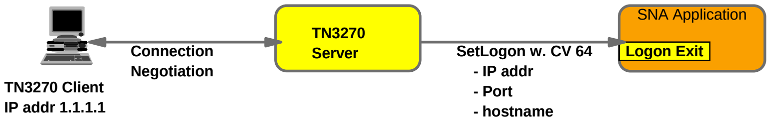
Generic LU name request - support added in z/OS V1R5

1. Initiate connection without specifying an LU name.
2. TN3270E server assigns LU name LUT1 and user connects to SNA application.
   - 2a. Telnet adds Client ID/LU name to master table.
3. Network has a hiccup and client to server IP connection is broken
4. Client sends a new connection (reconnect) request
   - 4a. Request from IP addr 1.1.1.1, so Telnet assigns LUT1 to connection request to simulate a specific LU request.
5. Reconnect logic kicks in and reconnects the user with the SNA application.
   - 5a. Telnet suspends new request
   - 5b. Sends TimeMark to original client
   - 5c. If response, rejects new and next available Generic LU is assigned.
   - 5d. If no response, new request continues setup with original LU by copying all relevent session info.
6. At disconnect time, Telnet removes the Client ID - LuName entry from the table.

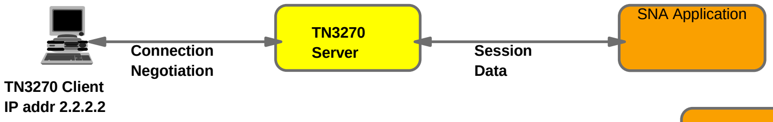
Redbooks V1R5 ibm.com/redbooks

Connection takeover when client IP address changes

- Telnet session initiation uses SetLogon to pass client information to VTAM and the APPL
  - VTAM displays include the client's IP address.
  - Some applications use the client's IP address for billing or other uses.



- Connection Takeover can be from a different IP address.



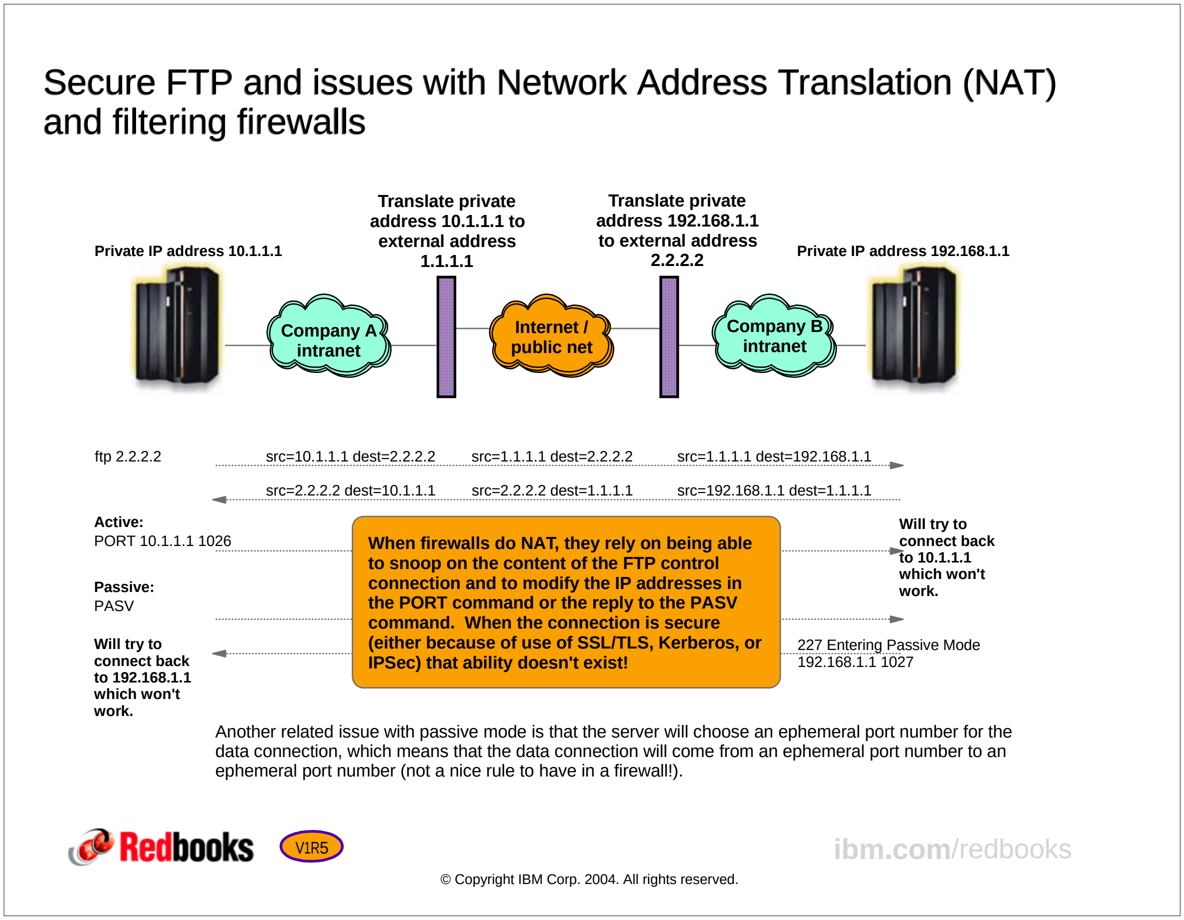
TKOSPECLURECON and TKOGENLURECON with new [NO]SAMEIPADDR keyword to control if reconnect from different IP address is allowed or not.

- Prior to z/OS V1R6 VTAM and the Application Logon exit can accept only one SetLogon request.
  - VTAM display is wrong.
  - Application has wrong IP address.
- TN3270 server will in z/OS V1R6 send SETLogon to VTAM both during initial session setup and during reconnect processing
  - VTAM displays will always be correct (show current client IP address)
  - Application logon exit still only driven once
  - New bits (CV64 subvector 81 flag X'80') in initial logon exit interface to indicate if IP address might change and application logon exit routine can then reject or accept this capability

V1R6

FTP protocol extensions for secure FTP - used for both SSL/TLS and Kerberos security for FTP

- Secure FTP refers to using the standard FTP protocol with one or more security extensions to improve security of data transfers
- "*FTP Security Extensions*", RFC2228 - defines a set of new FTP protocol commands and replies for negotiating secure FTP sessions.
- This RFC defines a framework for securing FTP. SSL/TLS and Kerberos are just two of many potential security mechanisms that could be used with FTP. Others will likely be defined in the future.
- The commands and replies are generic and are used to implement both Kerberos-based and SSL/TLS-based secure FTP sessions.
- Please note that secure FTP (sometimes referred to as ftps) as discussed in this presentation has nothing to do with what is known as ***sftp***
  - sftp is a file transfer protocol under the umbrella of SSH (Secure Shell)
    - SSH is available in an officially supported version for z/OS V1R4+ since May 2004
    - The sftp protocol in SSH on z/OS supports HFS file transfers, not MVS data sets
    - You can use sftp for transfer of HFS files to/from z/OS in an environment that has chosen to standardize on use of SSH
    - You need an SSH sftp client to exchange files with an SSH sftp server
  - sftp has absolutely nothing to do with the normal FTP standards as defined in RFC 959
  - The sftp protocol is its own protocol and sftp under SSH does in no way interoperate with normal FTP



Secure FTP and issues with Network Address Translation (NAT) and filtering firewalls

Private IP address 10.1.1.1

Translate private address 10.1.1.1 to external address 1.1.1.1

Translate private address 192.168.1.1 to external address 2.2.2.2

Private IP address 192.168.1.1

Company A intranet

Internet / public net

Company B intranet

ftp 2.2.2.2

src=10.1.1.1 dest=2.2.2.2 src=1.1.1.1 dest=2.2.2.2 src=1.1.1.1 dest=192.168.1.1

src=2.2.2.2 dest=10.1.1.1 src=2.2.2.2 dest=1.1.1.1 src=192.168.1.1 dest=1.1.1.1

Active:
PORT 10.1.1.1 1026

Passive:
PASV

Will try to connect back to 192.168.1.1 which won't work.

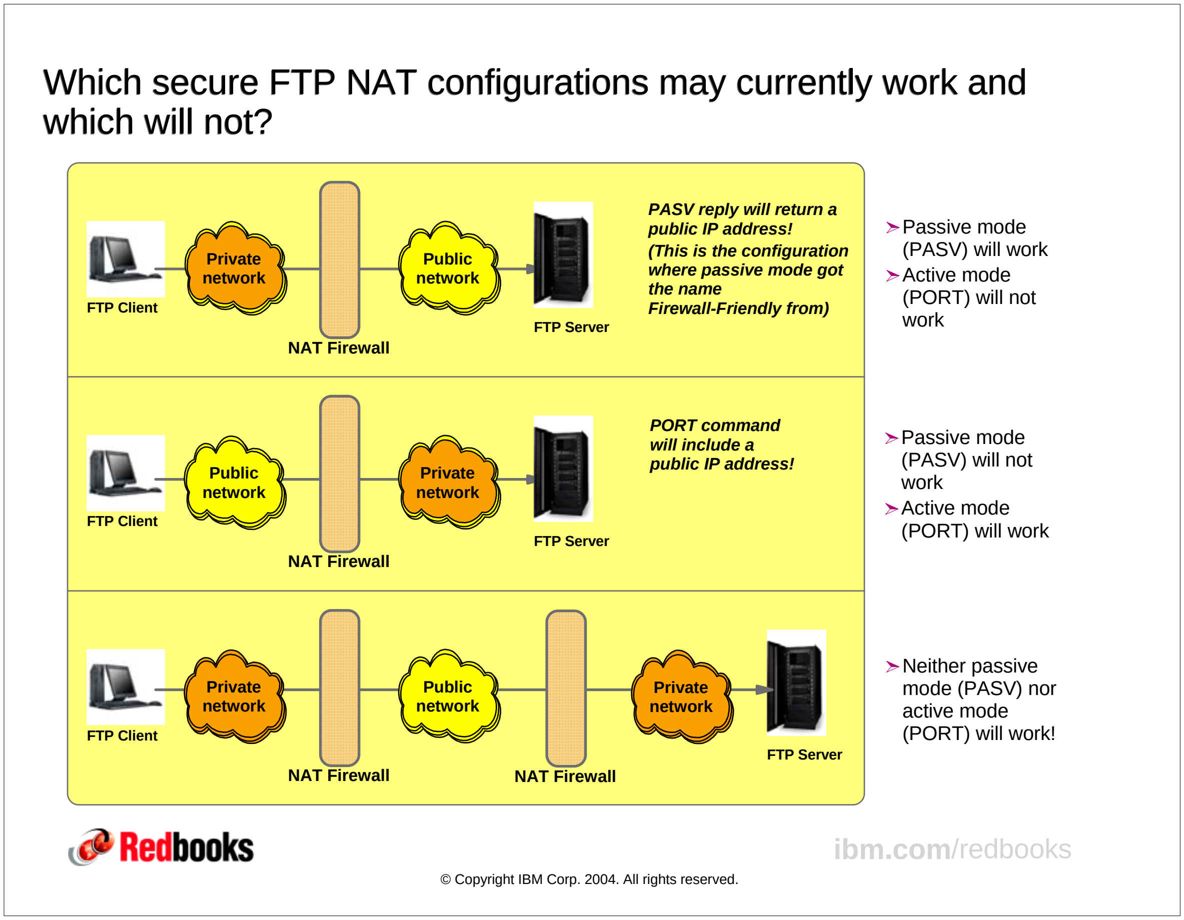
When firewalls do NAT, they rely on being able to snoop on the content of the FTP control connection and to modify the IP addresses in the PORT command or the reply to the PASV command. When the connection is secure (either because of use of SSL/TLS, Kerberos, or IPSec) that ability doesn't exist!

Will try to connect back to 10.1.1.1 which won't work.

227 Entering Passive Mode
192.168.1.1 1027

Another related issue with passive mode is that the server will choose an ephemeral port number for the data connection, which means that the data connection will come from an ephemeral port number to an ephemeral port number (not a nice rule to have in a firewall!).

Redbooks V1R5 ibm.com/redbooks



Which secure FTP NAT configurations may currently work and which will not?

FTP Client — Private network — NAT Firewall — Public network — FTP Server

PASV reply will return a public IP address! (This is the configuration where passive mode got the name Firewall-Friendly from)

- Passive mode (PASV) will work
- Active mode (PORT) will not work

FTP Client — Public network — NAT Firewall — Private network — FTP Server

PORT command will include a public IP address!

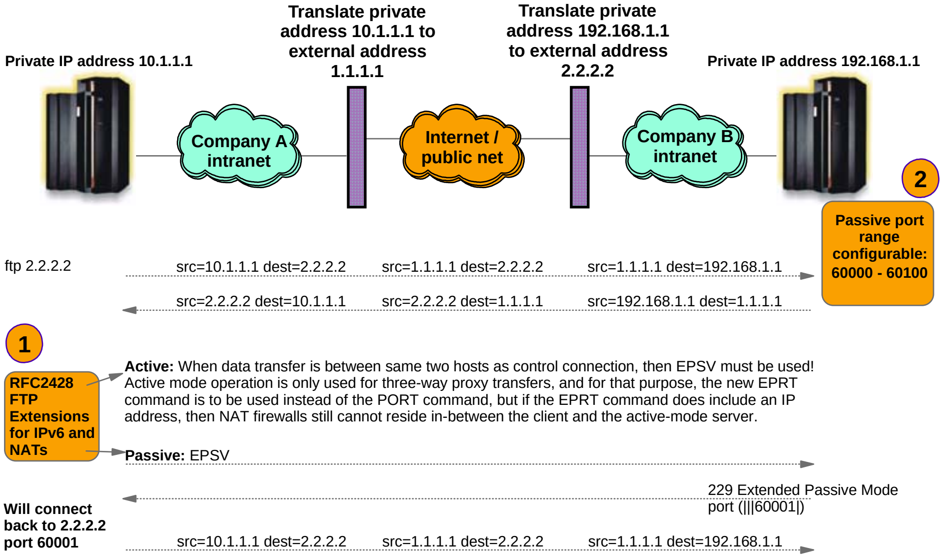
- Passive mode (PASV) will not work
- Active mode (PORT) will work

FTP Client — Private network — NAT Firewall — Public network — NAT Firewall — Private network — FTP Server

- Neither passive mode (PASV) nor active mode (PORT) will work!

Redbooks ibm.com/redbooks

Secure FTP, NAT, and filtering firewalls - addressed in z/OS V1R5


Translate private address 10.1.1.1 to external address 1.1.1.1
Translate private address 192.168.1.1 to external address 2.2.2.2
Private IP address 10.1.1.1
Private IP address 192.168.1.1
Company A intranet
Internet / public net
Company B intranet
2
Passive port range configurable: 60000 - 60100
ftp 2.2.2.2
src=10.1.1.1 dest=2.2.2.2
src=1.1.1.1 dest=2.2.2.2
src=1.1.1.1 dest=192.168.1.1
src=2.2.2.2 dest=10.1.1.1
src=2.2.2.2 dest=1.1.1.1
src=192.168.1.1 dest=1.1.1.1
1
RFC2428 FTP Extensions for IPv6 and NATs
229 Extended Passive Mode port (|||60001|)
Will connect back to 2.2.2.2 port 60001
src=10.1.1.1 dest=2.2.2.2
src=1.1.1.1 dest=2.2.2.2
src=1.1.1.1 dest=192.168.1.1

Active: When data transfer is between same two hosts as control connection, then EPSV must be used! Active mode operation is only used for three-way proxy transfers, and for that purpose, the new EPRT command is to be used instead of the PORT command, but if the EPRT command does include an IP address, then NAT firewalls still cannot reside in-between the client and the active-mode server.

Passive: EPSV

Do not expect that all secure FTP problems with firewalls have been solved - there may be more, we just haven't seen yet.

Client EPSV support and server passive data port range configuration PTF'ed back to z/OS V1R4 - APAR PQ80281

Redbooks V1R5 ibm.com/redbooks

Log in to z/OS V1R5 FTP server using SSL/TLS without a password

```
SECURE_LOGIN      VERIFY_USER        ; Authorization level indicator
                                     ; NO_CLIENT_AUTH (D)
                                     ; REQUIRED
                                     ; VERIFY_USER
SECURE_PASSWORD   OPTIONAL           ; W. clientuath is PW required?
                                     ; OPTIONAL
                                     ; REQUIRED (D)
```

Verify User is required.
New SECURE_PASSWORD option instructs if a password is required or not.
If user verification based on certificate doesn't succeed, user will be prompted for a password.

```
//ALFREDA JOB 1,ALFRED,CLASS=A,MSGCLASS=X,NOTIFY=USER1
//*
//* test of client authentication
//*
//FTP   EXEC PGM=FTP,PARM='-a TLS'
//SYSTCPD   DD DSN=USER1.TCPCS.TCPPARMS(TCPDATA),DISP=SHR
//SYSFTPD   DD DSN=USER1.TCPCS.TCPPARMS(FTPUSER1),DISP=SHR
//SYSPRINT  DD SYSOUT=*
//INPUT     DD *
;
; Test of client authentication
;
mvs098.tcp.raleigh.ibm.com 2021 (exit
user1
cd 'user1.alfred.cntl'
dir
quit
//OUTPUT    DD SYSOUT=*
```

No password in batch FTP input stream - only the user ID.

No prompt for a password from the server.

```
220-FTPSEC1 IBM FTP CS V1R5 at MVS098.tcp.raleigh.ibm.com, 11:35:39 on 2004-02-04.
220-*
220-* Welcome to the FTP server on MVS098.tcp.raleigh.ibm.com
220-* This system is used by Alfred for testing purposes.
220-* Any issues should be reported to alfredch@us.ibm.com
220-* Your host name is mvs098cs6.tcp.raleigh.ibm.com
220-*
220 Connection will not timeout.
EZA1701I >>> AUTH TLS
234 Security environment established - ready for negotiation
EZA2895I Authentication negotiation succeeded
EZA1701I >>> PBSZ 0
200 Protection buffer size accepted
EZA1701I >>> PROT P
200 Data connection protection set to private
EZA2906I Data connection protection is private
EZA1459I NAME (mvs098.tcp.raleigh.ibm.com:USER1):
EZA1701I >>> USER user1
230-*
230-* USER1 - welcome to the FTP server on MVS098.tcp.raleigh.ibm.com
230-* Login time and date is Wed Feb  4 11:35:41 2004
230-* The current working directory is /u/user1
230-*
230-User USER1 is an authorized user
230 USER1 is logged on.  Working directory is "/u/user1".
EZA1460I Command:
EZA1736I cd 'user1.alfred.cntl'
```

Redbooks V1R5 ibm.com/redbooks

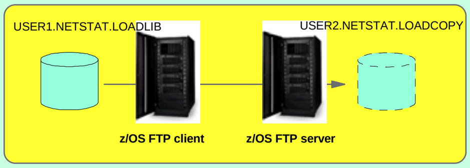
z/OS V1R5 improves FTP handling of PDS and PDSE data sets

- z/OS V1R5 improves handling of PDS/PDSE libraries by both the FTP client and the FTP server:
  - New FTP.DATA and SITE/LOCSITE options to specify whether an MKDIR or LMKDIR in MVS data set mode results in allocating a PDS or a PDSE:
    - PDSTYPE [PDS | PDSE]
  - New option on the MKDIR and LMKDIR commands to specify a 'like' partitioned data set from where to copy allocation attributes - avoding specifying SITE/LOCSITE commands before creating the new partitioned data set:
    - mkdir remote_directory (like local_directory
    - lmkdir local_directory (like remote_directory

Example: creating a new partitioned data set on the server with the attributes of a data set on the client - and transfer all members

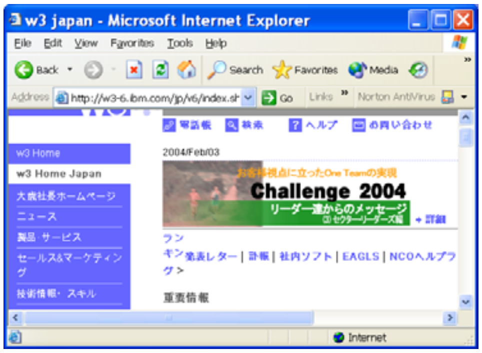
```
lcd 'user1.netstat'
cd 'user2.netstat'
;
; Make a new library on server called 'USER2.NETSTAT.LOADCOPY'
; based on how local 'USER1.NETSTAT.LOADLIB' looks like
;
mkdir loadcopy (like loadlib
lcd loadlib
cd loadcopy
mput *
dir
quit
```

Both client and server need to be at a z/OS V1R5 or higher level



Enhanced DBCS and MBCS code page support in FTP

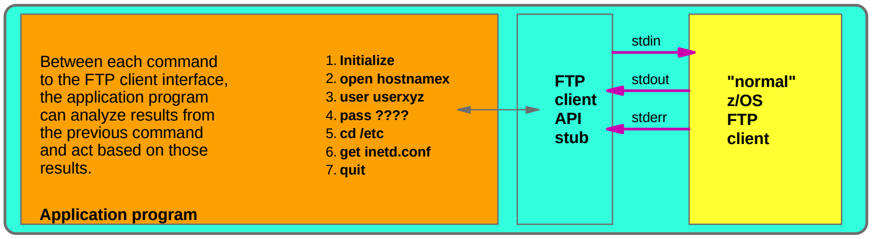
- Enhanced Multi-Byte Character Set (MBCS) - primarily support for Asian languages
- Current FTP support for Double Byte Character Set (DBCS) is based on an imbedded support in TCP/IP for selected conversions and is not ready for the latest z/OS character conversion technology
- z/OS V1R4 provides MBCS encoding support only for Chinese code standard GB18030
- z/OS V1R6 enhances MBCS to include the DBCS code pages currently supported by the existing old imbedded support
  - Some conversion parameters are not supported with the new method (they aren't standard)
- The new support is based on use of the standard FTP protocol (type ASCII) and use of SITE commands that are compatible with single byte (SBCS) conversion:
  - ENCODING SBCS/MBCS and
  - SB/MBDATACONN=(file_system_code_page , network_code_page)
- Original code page support for DBCS using LOADDBCSTABLES is still supported, but we recommend moving to the new support if at all possible
- Objective is to make FTP independent of any specific code page - as long as the underlying z/OS conversion supports a code page conversion - so will FTP
- Currently FTP uses iconv() conversion services, but will eventually move to the Unicode Conversion Services


w3 japan - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites Media
Address http://w3-6.ibm.com/jp/v6/index.sh
Go Links Norton AntiVirus
w3 Home
w3 Home Japan
2004/Feb/03
Challenge 2004
Internet

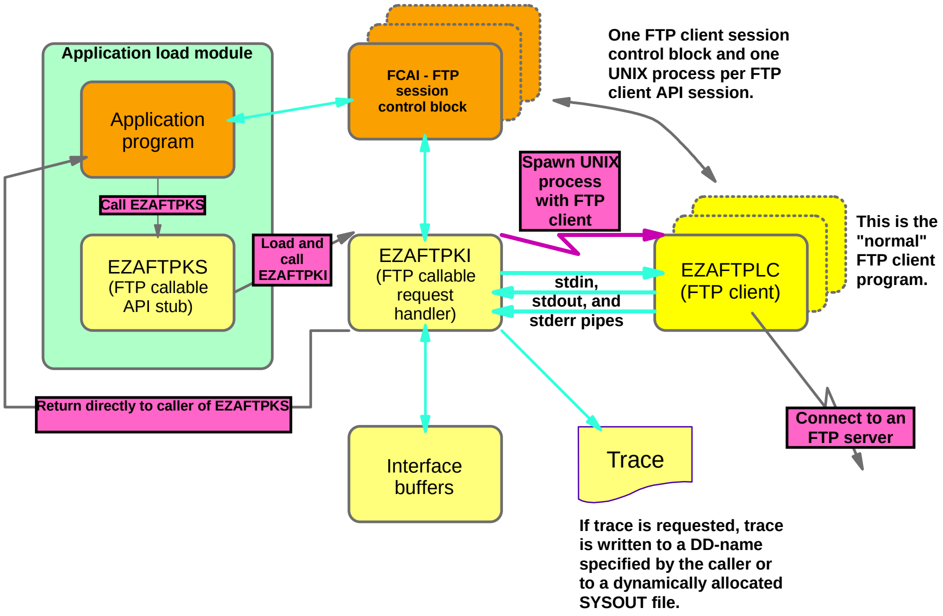
z/OS FTP client programming interface for improved automation and integration of z/OS file transfers in z/OS V1R6

- 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) - plans to add C and REXX APIs in a later release
  - 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 replies 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 command, process output from the FTP client command 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 command) along with free-format replies from the FTP client code
  - Debugging options are provided



Structure of the FTP client API implementation



Application load module
Application program
Call EZAFTPKS
EZAFTPKS (FTP callable API stub)
Load and call EZAFTPKI
Return directly to caller of EZAFTPKS
FCAI - FTP session control block
One FTP client session control block and one UNIX process per FTP client API session.
Spawn UNIX process with FTP client
EZAFTPKI (FTP callable request handler)
stdin, stdout, and stderr pipes
EZAFTPLC (FTP client)
This is the "normal" FTP client program.
Connect to an FTP server
Interface buffers
Trace
If trace is requested, trace is written to a DD-name specified by the caller or to a dynamically allocated SYSOUT file.

IBM®

What are the Major New Functions in CS z/OS V1R5 and V1R6?

-

Management

Redbooks

International Technical Support Organization

Management - overview

CS z/OS V1R5

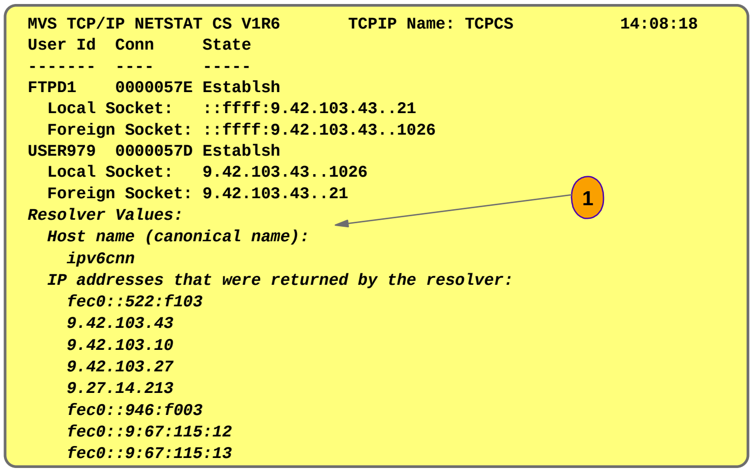
- Various enhancements to netstat reports
- New set of high-performing network management callable interfaces for local monitor management products
- Improved use of MVS system symbolics in the resolver configuration file: TCPIP.DATA
- MSYS for setup support of FTP client and server configuration (FTP.DATA)
- New simplified Service Level Agreement (SLA) MIB
- Callable interface to the Policy Agent for integrated network performance monitoring
- Extend z/OS Policy Agent to a common eServer Policy-Based networking infrastructure

CS z/OS V1R6

- Significantly improved the netstat report documentation
- z/OS V1R6 is the last release OROUTED will be shipped - migrate to OMPROUTE

Netstat report using the host name filter

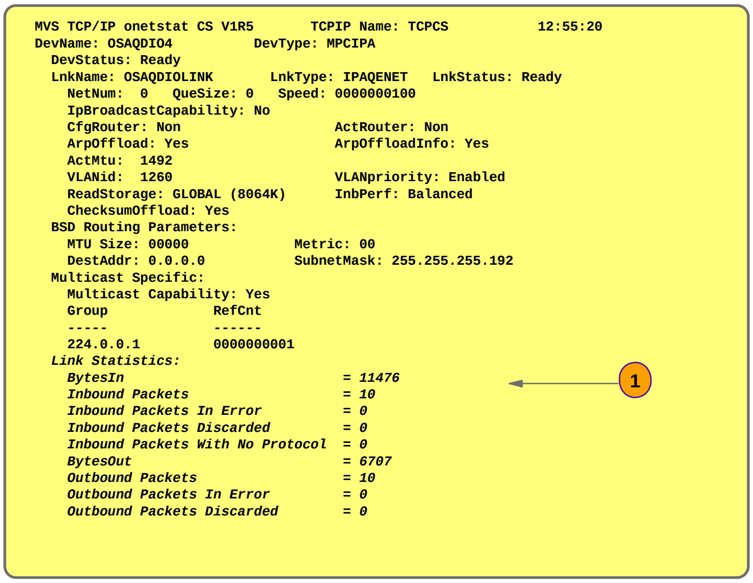
- Filter the netstat ALLCONN/-a report on HOSTName/-H *ipv6cnn*



```
MVS TCP/IP NETSTAT CS V1R6       TCPIP Name: TCPCS           14:08:18
User Id  Conn     State
-------  ----     -----
FTPD1    0000057E Establsh
  Local Socket:   ::ffff:9.42.103.43..21
  Foreign Socket: ::ffff:9.42.103.43..1026
USER979  0000057D Establsh
  Local Socket:   9.42.103.43..1026
  Foreign Socket: 9.42.103.43..21
Resolver Values:
  Host name (canonical name):
    ipv6cnn
  IP addresses that were returned by the resolver:
    fec0::522:f103
    9.42.103.43
    9.42.103.10
    9.42.103.27
    9.27.14.213
    fec0::946:f003
    fec0::9:67:115:12
    fec0::9:67:115:13
```

Netstat DEVLINKS report with new statistics

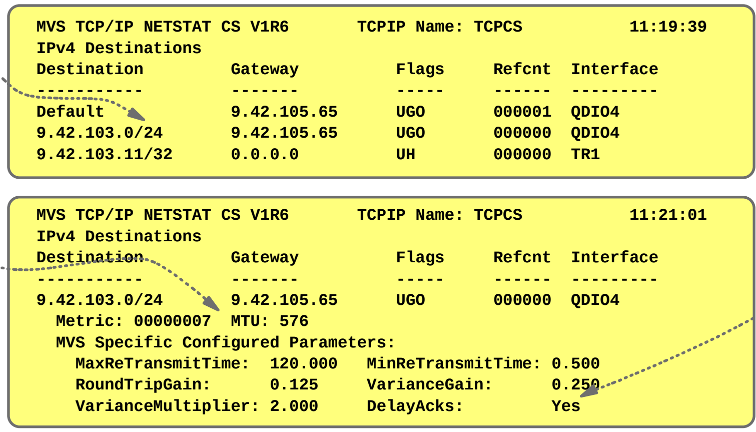
- Netstat DEVLINKS/-d



```
MVS TCP/IP onetstat CS V1R5       TCPIP Name: TCPCS           12:55:20
DevName: OSAQDIO4          DevType: MPCIPA
  DevStatus: Ready
  LnkName: OSAQDIOLINK       LnkType: IPAQENET   LnkStatus: Ready
    NetNum:  0   QueSize: 0   Speed: 0000000100
    IpBroadcastCapability: No
    CfgRouter: Non                  ActRouter: Non
    ArpOffload: Yes                 ArpOffloadInfo: Yes
    ActMtu:  1492
    VLANid:  1260                   VLANpriority: Enabled
    ReadStorage: GLOBAL (8064K)     InbPerf: Balanced
    ChecksumOffload: Yes
  BSD Routing Parameters:
    MTU Size: 00000            Metric: 00
    DestAddr: 0.0.0.0          SubnetMask: 255.255.255.192
  Multicast Specific:
    Multicast Capability: Yes
    Group             RefCnt
    -----             ------
    224.0.0.1         0000000001
  Link Statistics:
    BytesIn                           = 11476
    Inbound Packets                   = 10
    Inbound Packets In Error          = 0
    Inbound Packets Discarded         = 0
    Inbound Packets With No Protocol  = 0
    BytesOut                          = 6707
    Outbound Packets                  = 10
    Outbound Packets In Error         = 0
    Outbound Packets Discarded        = 0
```

Netstat route report changes

- Changes to the netstat ROUTE/-r report:
  - Add the configured DELAYACKS/NODELAYACKS setting for a route to the netstat ROUTE/-r report when the DETAIL modifier is specified.
  - Add the MTU size for IPv4 routes to both LONG and SHORT formats of the netstat ROUTE/-r report when the DETAIL modifier is specified.
  - Add the prefix length information for IPv4 routes to the SHORT format of the netstat ROUTE/-r report to show the subnet mask information.



```
MVS TCP/IP NETSTAT CS V1R6       TCPIP Name: TCPCS          11:19:39
IPv4 Destinations
Destination         Gateway          Flags     Refcnt  Interface
-----------         -------          -----     ------  ---------
Default             9.42.105.65      UGO       000001  QDIO4
9.42.103.0/24       9.42.105.65      UGO       000000  QDIO4
9.42.103.11/32      0.0.0.0          UH        000000  TR1
```

```
MVS TCP/IP NETSTAT CS V1R6       TCPIP Name: TCPCS          11:21:01
IPv4 Destinations
Destination         Gateway          Flags     Refcnt  Interface
-----------         -------          -----     ------  ---------
9.42.103.0/24       9.42.105.65      UGO       000000  QDIO4
  Metric: 00000007  MTU: 576
  MVS Specific Configured Parameters:
    MaxReTransmitTime:  120.000   MinReTransmitTime: 0.500
    RoundTripGain:      0.125     VarianceGain:      0.250
    VarianceMultiplier: 2.000     DelayAcks:         Yes
```

V1R6

Netstat documentation

- The z/OS netstat command has been enhanced almost every single release. Up to now, it has 26 report/functional options, 19 modifiers (additional keywords for report options), and 9 filters (select-strings for report options), so that netstat can provide 54 different reports to display the network status of the local host, including information about TCP/IP connections, network clients, gateways, and devices, etc.

- The existing netstat documentation in the *IP System Administrator's Commands* book is mainly divided into two sections:

  - The TSO NETSTAT command and its options
  - The z/OS UNIX onetstat/netstat command and its options

- Since netstat has so many different options and filters which cover almost 200 pages of documentation, it is difficult for customers to find individual option information in the TSO and UNIX sections.

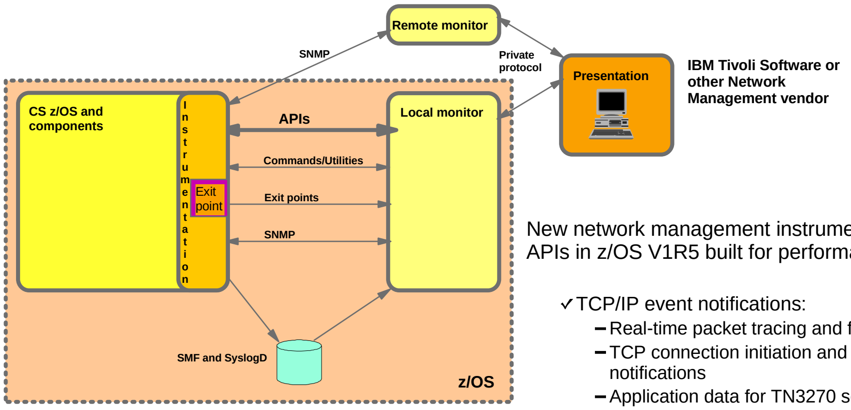
- Most of the netstat options are supported for both TSO and UNIX environments, so the existing documentation format causes the same information (in Parameters and Examples sections) to be repeated in two places.

- Netstat reports produce a high amount of detailed information to the user. But since we have a number of customers who are not that familiar with TCP/IP concepts, they get confused and are often unable to interpret the information that is provided in the netstat report.

- For some of the netstat reports, we do not explain all of the fields displayed on the report.

V1R6

Network Management Instrumentation (NMI) overview



Tivoli's IBM Tivoli Monitor / Network Performance (ITM/NP) use these new APIs

The APIs and their documentation will be available as part of CS z/OS for use by network management vendors and customer network management applications. APIs shipped for CS z/OS V1R4 as a PTF (UQ81245).

See info APAR II13699 for details.

New network management instrumentation APIs in z/OS V1R5 built for performance:

- ✓ TCP/IP event notifications:
  - Real-time packet tracing and formatting
  - TCP connection initiation and termination notifications
  - Application data for TN3270 server and FTP event data
- ✓ APIs to poll information about currently active TCP/IP activity
  - TCP listeners (server processes)
  - TCP connections (detailed information about individual connections)
  - UDP endpoints
  - CS storage usage
- ✓ API to receive and poll for Enterprise Extender management data

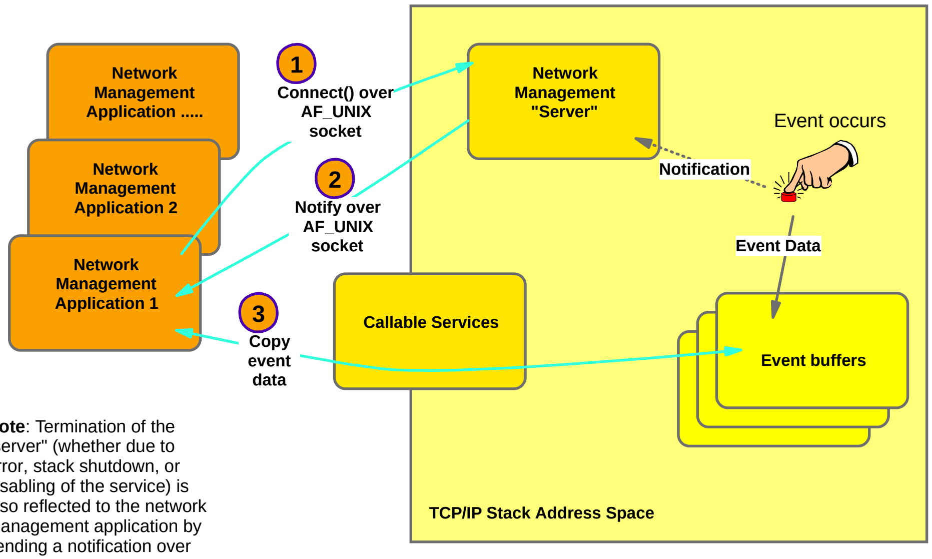
Event notification interfaces - configuration

- The event notification support must be enabled before use. This is accomplished using the NETMONitor statement in the TCP/IP PROFILE:

```
NETMONitor OFF
           | ON
           | [NOPKTTRCService|PKTTRCService]
             [NOTCPCONNService|TCPCONNService]
             [NOSMFService|SMFService]
```

- The status of this configuration option may be seen using the netstat CONFIG/-f command.
- A parameter of OFF indicates all services should be inactive, while ON indicates that all services should be made active.
- If no options are specified on the NETMONitor statement, then NETMONitor OFF is assumed.
- If NETMONitor appears in a VARY TCPIP,,OBEYFILE, then the parameters will change the existing settings, e.g.:
  - NETMONITOR OFF - turn all functions off
  - NETMONITOR ON - turn all inactive functions on
  - NETMONITOR NOSMFS - turn off SMF service if active
  - NETMONITOR TCPCONNS - turn on TCP connection service if not active

TCP/IP event notification to network management application


Network Management Application .....
Network Management Application 2
Network Management Application 1
1
Connect() over AF_UNIX socket
2
Notify over AF_UNIX socket
3
Copy event data
Network Management "Server"
Notification
Event occurs
Event Data
Callable Services
Event buffers
TCP/IP Stack Address Space

Note: Termination of the "server" (whether due to error, stack shutdown, or disabling of the service) is also reflected to the network management application by sending a notification over the AF_UNIX connection to the client.

Network management services can be protected through SERVAUTH profiles:

- EZB.NETMGMT.sysname.stackname.SYSTCPxx

V1R5

Basics of the new polling type of network management interfaces

- Network management applications which run locally on a z/OS system where they monitor TCP/IP activity and status need a high-speed, low-overhead interface to access data about TCP connections and UDP endpoints.
  - SNMP protocols can be used to access such information, but adds processing overhead that in some situations has proven to be unacceptable, especially when obtaining information for a large set of endpoints (e.g. walking the TCP connection table).
  - Some applications have attempted to parse the output from netstat to access the necessary information. This is error-prone and inefficient, and can put a severe strain on TCP/IP stack resources when done frequently.
  - An application that receives asynchronous events may need to use a polling interface to query (sample at selected intervals) the current status of one or more existing TCP or UDP connections to monitor their progress.
- The new network management callable API is a high-performance and low overhead polling interface that can return the following types of information at a given point in time:
  - **GetTCPListeners** - Information about all, or selected, active TCP endpoints that listen for incoming connections ("servers").
  - **GetConnectionDetail** - Information about all, or selected, active non-listener TCP connections.
  - **GetUDPTable** - Information about all, or selected, active UDP endpoints.
  - **GetStorageStatistics** - Information about TCP/IP stack usage of common and private storage.

V1R5

Filtering support by polling function

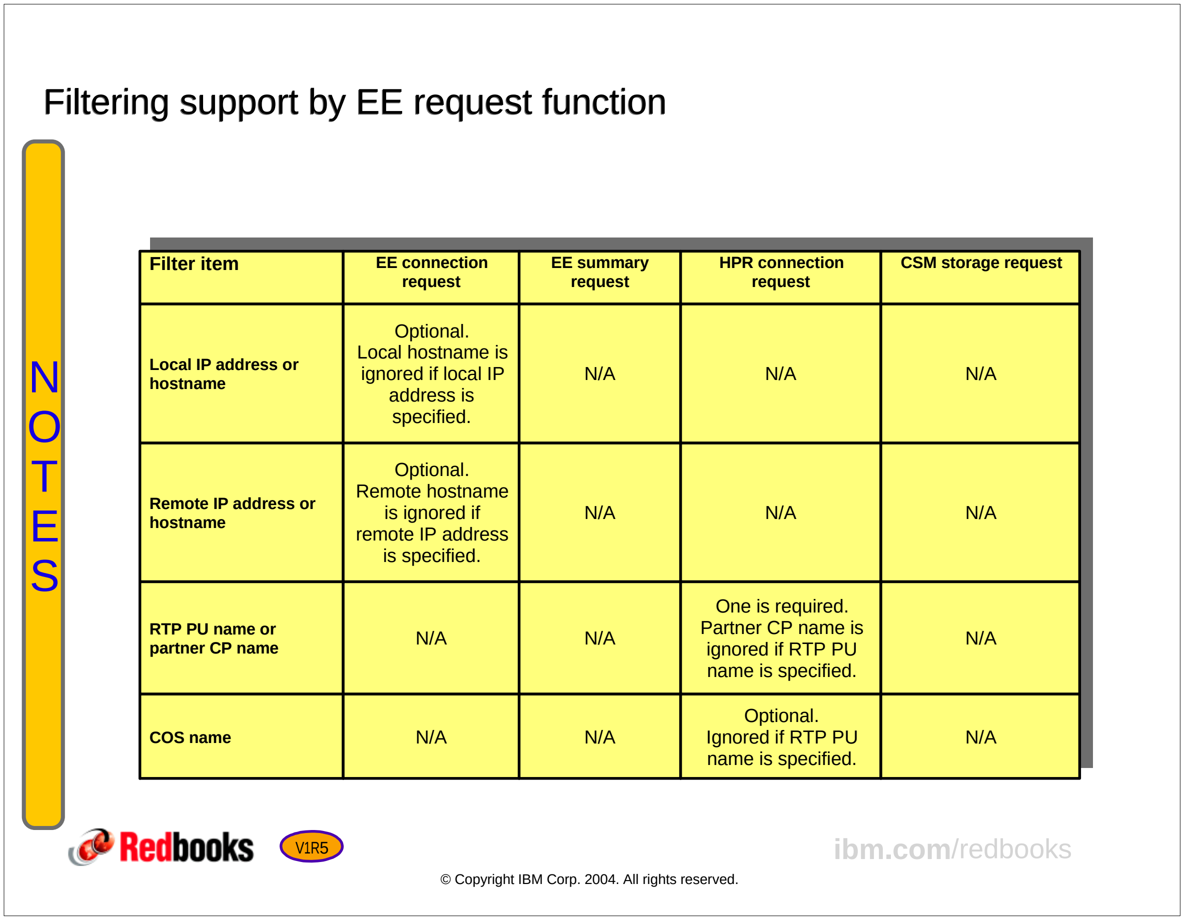
Filter item	GetTCPListeners	GetUDPTable	GetConnectionDetail	GetStorageStatistics
ASID	yes	yes	yes	no
Resource name (Job name)	yes	yes	yes	no
Resource ID (conn ID)	yes	yes	yes	no
Server resource ID (conn ID of server)	no	no	yes	no
Local IP address	yes	yes	yes	no
Local IP address prefix	yes	yes	yes	no
Local port	yes	yes	yes	no
Remote IP address	no	no	yes	no
Remote IP address prefix	no	no	yes	no
Remote port	no	no	yes	no

Redbooks V1R5 ibm.com/redbooks

Basics of new Enterprise Extender management interfaces

- Network management information for Enterprise Extender (EE) is not very extensive
  - While there are a number of network management tools available for monitoring and for problem determination with TCP/IP and SNA, there is nothing available specifically to assist with EE network management.
- Network management information for High Performance Routing (HPR) exists, but is not easily processed
  - The DISPLAY NET,ID=*rtpname*,HPRDIAG=YES command does provide some HPR statistics for a given HPR connection (added in z/OS V1R4).
  - However, operator must still issue command periodically, and someone must parse and sort the output, to get useful monitoring information.
- Network management information for Common Storage Management (CSM) exists, but is not easily processed
  - The DISPLAY NET,CSM command provides data currently.
  - Performance Monitor Interface (PMI) processing provides application interface to collect the data, but requires monitoring application to open an ACB to operate.
- Rather than require operators or customers to collate the existing data on their own, and rather than require ACB overhead, a new interface is provided for acquisition of information about EE, HPR, and/or CSM.
- The management application can request the pertinent information at regular intervals and thus get a more complete picture of system usage.
- This new SNA Network Monitor Interface (NMI) is a polling interface, based on the AF_UNIX socket interface, for requesting information about Enterprise Extender, High Performance Routing, and Common Storage Management

Redbooks V1R5 ibm.com/redbooks

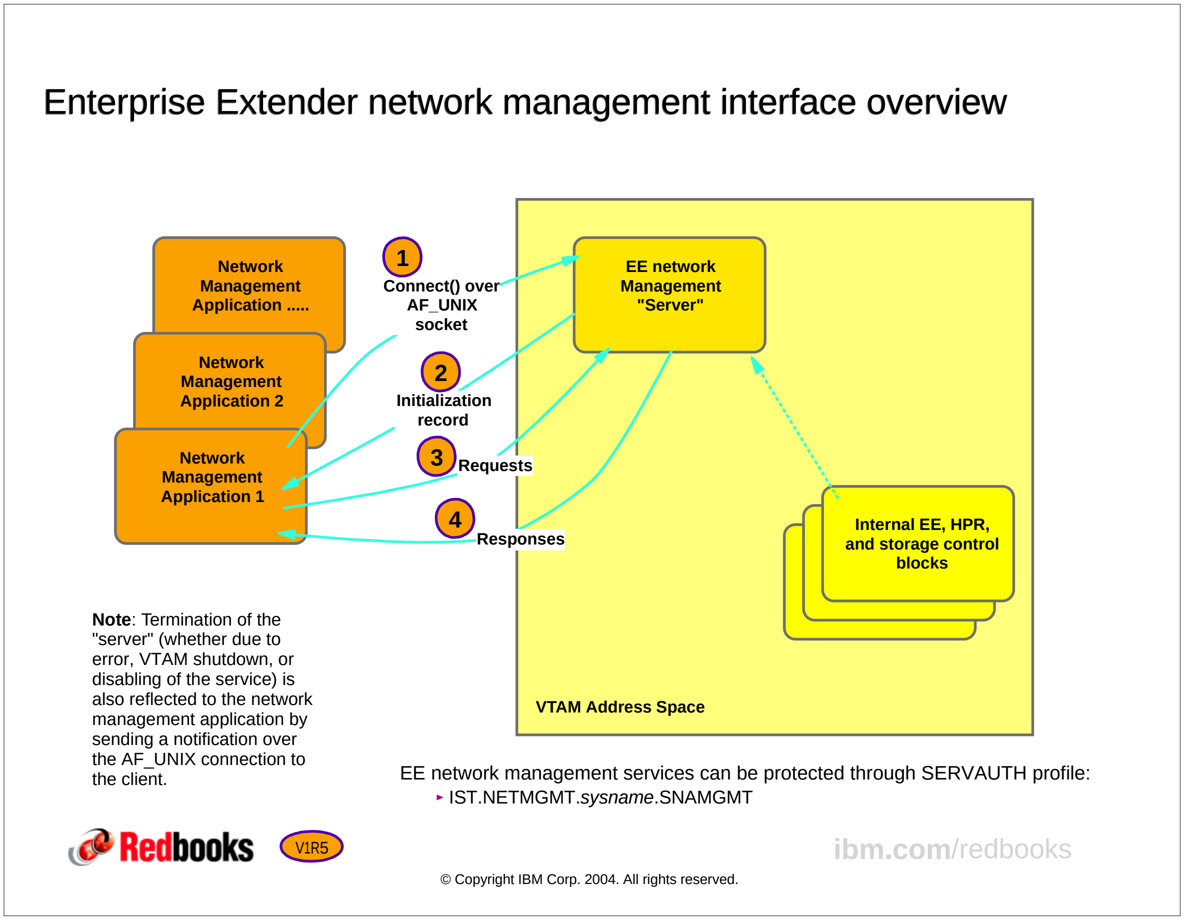


Filtering support by EE request function

NOTES

Filter item	EE connection request	EE summary request	HPR connection request	CSM storage request
Local IP address or hostname	Optional. Local hostname is ignored if local IP address is specified.	N/A	N/A	N/A
Remote IP address or hostname	Optional. Remote hostname is ignored if remote IP address is specified.	N/A	N/A	N/A
RTP PU name or partner CP name	N/A	N/A	One is required. Partner CP name is ignored if RTP PU name is specified.	N/A
COS name	N/A	N/A	Optional. Ignored if RTP PU name is specified.	N/A

Redbooks V1R5 ibm.com/redbooks



Enterprise Extender network management interface overview

Network Management Application .....
Network Management Application 2
Network Management Application 1
1
Connect() over AF_UNIX socket
2
Initialization record
3
Requests
4
Responses
EE network Management "Server"
Internal EE, HPR, and storage control blocks
VTAM Address Space

Note: Termination of the "server" (whether due to error, VTAM shutdown, or disabling of the service) is also reflected to the network management application by sending a notification over the AF_UNIX connection to the client.

EE network management services can be protected through SERVAUTH profile:

- IST.NETMGMT.*sysname*.SNAMGMT

Redbooks V1R5 ibm.com/redbooks

NMI documentation

NOTES

Note: download the documentation at these URLs:

- z/OS V1R4: ftp://ftp.software.ibm.com/s390/zos/commserver/V1R4NMUG.pdf
- z/OS V1R5: ftp://ftp.software.ibm.com/s390/zos/commserver/V1R5NMUG.pdf

Redbooks

ibm.com/redbooks

IBM®

What are the Major New Functions in CS z/OS V1R5 and V1R6?

-

Security

Redbooks

International Technical Support Organization

Security - overview

CS z/OS V1R5

- ✓Enhanced intrusion detection for interface attacks
- ✓Enable IDS events to be sent to Tivoli Risk Manager
- ✓Initial Multi Level Security support by TCP/IP

CS z/OS V1R6

- ✓Continues implementing MLS support
- ✓Extend z/OS IDS to a common eServer Policy-Based Networking infrastructure

We will discuss TCP/IP security in much more detail in a follow-on session later today.

IBM®

What are the Major New Functions in CS z/OS V1R5 and V1R6?

-

SNA - APPN, HPR, and Enterprise Extender

Redbooks

International Technical Support Organization

SNA - overview

CS z/OS V1R5

- ✓ Enterprise Extender Global Connection Network NAT firewall relief
- ✓ Support for multiple concurrent APING operations
- ✓ Enterprise Extender support of multiple Virtual Routing Nodes (VRNs)
- ✓ Enterprise Extender performance enhancements
- ✓ Numerous usability and command enhancements

CS z/OS V1R6

- ✓ Extended Border Node (EBN) awareness of HPR sessions
- ✓ EE connections network reachability awareness
- ✓ MNPS predatory takeover (for IMS)
- ✓ Enhanced addressing support for RTP PUs

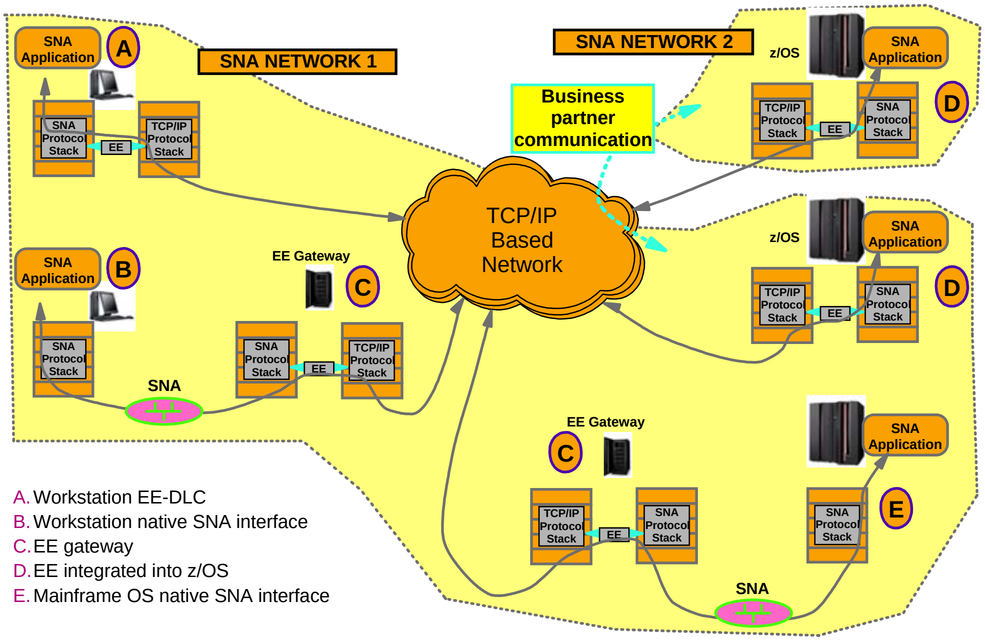
Key points about Enterprise Extender

NOTES

- ➢ Enterprise Extender is based on the latest SNA architecture - High Performance Routing (HPR)
  - ▸ Includes all the dynamics of Advanced Peer to Peer Networking (APPN)
    - – Dynamic SNA resource and topology discovery
  - ▸ Includes all the performance and availability aspects of HPR
    - – Non-disruptive SNA path switches - SNA sessions survive HPR link failures if an alternate path is available or when the original link recovers
  - ▸ EE is HPR over IP and looks at the underlying IP network as a single HPR link
    - – Benefits of IP network high-availability features:
      - ● Redundant network equipment and paths
      - ● Dynamic recovery from IP topology changes
      - ● QoS based routing with IP QoS derived from SNA Class Of Service (COS)
  - ▸ To be an EE node, that node also needs to have both APPN and HPR enabled
- ➢ EE may be deployed in an internal network using two base topologies
  - ▸ EE pushed to the edge of the network - IP end-to-end
    - – On z/OS - allowing IP traffic to enter/leave z/OS over OSA-Express adapters using QDIO or using HiperSockets between LPARs on a zSeries
    - – On the workstation using EE-DLC (Data Link Control) features of products such as IBM Personal Communications or Communications Server for Windows or Linux
  - ▸ EE deployed on gateways in the network - simple to implement
    - – EE gateways can be deployed in the branch on CS Windows, CS Linux, CS AIX, or SNA Switch
    - – EE gateways can be deployed in the data center on the same platforms as in the branch or directly on zSeries using either z/OS or CS Linux on zSeries acting as gateways to other zSeries operating systems
      - ● An EE gateway on zSeries allows network flows between the zSeries and the network to be IP based

Enterprise Extender - topology principles

NOTES



A. Workstation EE-DLC
B. Workstation native SNA interface
C. EE gateway
D. EE integrated into z/OS
E. Mainframe OS native SNA interface

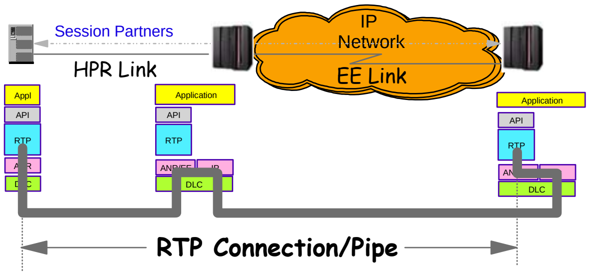
Redbooks

ibm.com/redbooks

Key points about Enterprise Extender - continued

NOTES

- 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
- 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



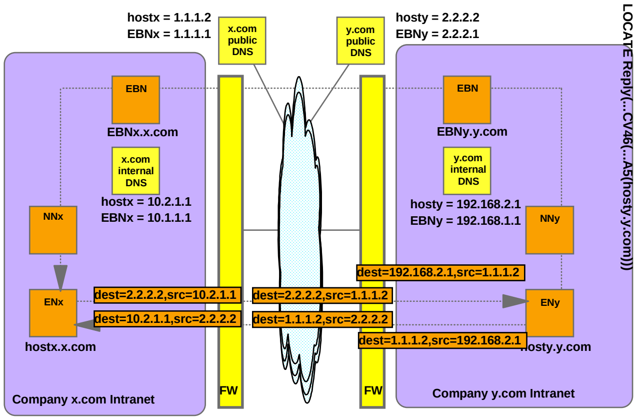
Why don't all installations use EE on z/OS?

- 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

Redbooks

ibm.com/redbooks

EE connection network firewall relief and IPv6 support

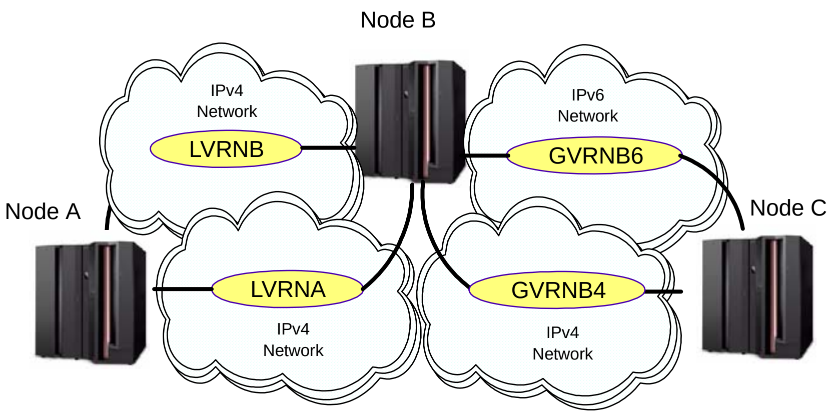


Intranet #	Public #
10.1.1.1	1.1.1.1
10.2.1.1	1.1.1.2

Public #	Intranet #
2.2.2.1	192.168.1.1
2.2.2.2	192.168.2.1

- ➢ EE architecture has been updated to allow the EE connection network control vectors to carry the hostname corresponding to the EE VIPA.
- ➢ IPv4 connection network control vectors will continue to carry IP address as well as hostname to ensure compatibility with downlevel nodes.
- ➢ Administrative requirement of coordinating NAT tables and public DNS entries is a known administrative procedure to installations that use NAT.
- ➢ There will be recommended maximum fully qualified hostname lengths due to limited space in the route selection control vector.
- ➢ If hostname resolves to an IPv6 address, then EE will use IPv6

Multiple VRN/VIPA support



- ➢ V1R5 will allow the specification of multiple local and/or multiple global EE connection networks.
  - ▸ In the diagram above, Node B defines 2 local VRNs (both IPv4) and 2 global VRNs (one IPv4 and one IPv6)
- ➢ EE will allow multiple (static) VIPAs, defined on a GROUP basis in the EE XCA major node.
  - ▸ All EE VIPAs must still belong to a single TCP/IP stack

D RTPS enhancement

- D NET,RTPS displays HPR pipes that match the input criteria. Prior to z/OS V1R5, several keywords (APPNCOS, CPNAME, CONGEST, SWITCH, and ID) could be used to determine the matching criteria.
- To address the need to display the HPR pipes traversing a particular first hop, we added the following keywords to the D NET,RTPS command: FIRSTTG, FIRSTCP, and ALSNAME. They can be used alone, in conjunction with each other, or with the other already existing keywords.
- To address the need to display HPR pipes going to a particular network, we added the capability of specifying the CPNAME keyword with a network qualified name, where the netid is specified, but the name is an asterisk (*). This will indicate that the netid portion of the CP name must match, but the name portion of the CP name will not be a matching criteria. For example, if CPNAME=NETB.* is specified, then IST1697I will only be displayed for HPR pipes with a destination CP in the network known as NETB. No other format other than netid.* will be allowed, when using the *.
- The new FIRSTCP operand can use the asterisk in the same way as CPNAME.
- The customer can use the new keywords individually, together, and in conjunction with the current keywords:
  - D NET,RTPS,FIRSTTG=21
  - D NET,RTPS,FIRSTCP=SSCP1A
  - D NET,RTPS,FIRSTCP=NETA.SSCP1A
  - D NET,RTPS,ALSNAME=AHHCPU1
  - D NET,RTPS,FIRSTTG=21,FIRSTCP=SSCP1A
  - D NET,RTPS,ALSNAME=AHHCPU1,CONGEST=YES
- The customer can use the netid.* format for the CPNAME and FIRSTCP keywords:
  - D NET,RTPS,CPNAME=NETA.*
  - D NET,RTPS,FIRSTCP=NETB.*

Redbooks V1R5

D RTPS examples

NOTES

```
d net,rtps,firsttg=21,firstcp=sscp2a
IST097I DISPLAY ACCEPTED
IST350I DISPLAY TYPE = RTPS
IST1695I PU NAME       CP NAME      COS NAME SWITCH CONGEST  SESSIONS
IST1696I CNR00005 NETA.SSCP2A       #INTER     NO      NO           1
IST1696I CNR00004 NETA.SSCP2A       #BATCH     NO      NO           1
IST1454I 2 RTP(S) DISPLAYED
IST314I END

d net,rtps,alsname=ahhcpu1
IST097I DISPLAY ACCEPTED
IST350I DISPLAY TYPE = RTPS
IST1695I PU NAME       CP NAME      COS NAME SWITCH CONGEST  SESSIONS
IST1696I CNR00001 NETA.SSCP1A       CPSVCMG    NO      NO           1
IST1454I 1 RTP(S) DISPLAYED
IST314I END

d net,rtps,cpname=neta.*
IST097I DISPLAY ACCEPTED
IST350I DISPLAY TYPE = RTPS
IST1695I PU NAME       CP NAME      COS NAME SWITCH CONGEST  SESSIONS
IST1696I CNR00005 NETA.SSCP2A       #INTER     NO      NO           1
IST1696I CNR00004 NETA.SSCP2A       #BATCH     NO      NO           1
IST1696I CNR00001 NETA.SSCP1A       CPSVCMG    NO      NO           1
IST1454I 3 RTP(S) DISPLAYED
IST314I END
```

Redbooks V1R5

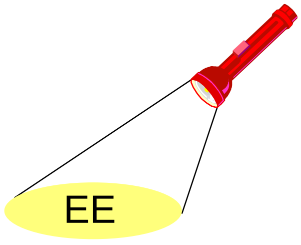
Multiple concurrent APING support

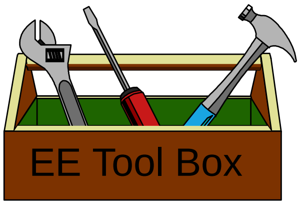
- The DISPLAY APING function is implemented using two Transaction Programs (TPs). They are ISTAPING and APINGDTP.
  - The ISTAPING TP is more commonly referred to as APINGTP. This is the TP that initiates the transaction. When a VTAM operator issues a DISPLAY APING command, the APINGTP will issue an allocate to start a conversation over an LU 6.2 session.
  - The APINGD TP behaves like a server. It receives the transaction from the ISTAPING TP. It completes allocation of the conversation and responds by sending reply data.
    - The MODIFY APINGDTP command allows you to change the number of APINGD transaction programs permitted to run concurrently to respond to APING requests from other nodes.
    - The DISPLAY APINGDTP command displays the number of APINGD transaction programs permitted to run concurrently to respond to APING request from other nodes.
- Customers would like to have the ability to issue multiple DISPLAY APING commands simultaneously.
- Customers want to use DISPLAY APING under programmatic control to gather performance data. Being limited to a single DISPLAY APING at a time greatly increases the time to gather such information for the entire network.
- Many customers are primarily concerned about error scenarios, where the DISPLAY APING target is entered incorrectly, or is not findable. In those cases, the operator must wait for the searching to complete and the previous DISPLAY APING to fail before being able to try again.
- In a case where the DISPLAY APING inadvertently connects to an APPC application which does not understand the APING transaction protocols, the target application might not respond and the conversation can hang until cancelled by the operator.

Redbooks V1R5 ibm.com/redbooks

Display EE command

- V1R6 provides a new operator command to provide additional details about Enterprise Extender connectivity
- Three basic forms:
  - General information
    - Basic XCA settings
    - Local IP addresses and/or hostnames
    - RTP pipe and LU-LU session counts
    - Connection counts
  - Specific connection information
    - Local IP address and/or hostname
    - PU information
    - LDLC information
    - Data transfer statistics
  - Aggregate connection information
    - Local IP address and/or hostname
    - Connection counts
    - Aggregate data transfer statistics





Redbooks V1R6 ibm.com/redbooks

D EE: General information

NOTES

```
D NET,EE,LIST=DETAIL

IST097I DISPLAY ACCEPTED
IST350I DISPLAY TYPE = EE
IST2000I ENTERPRISE EXTENDER GENERAL INFORMATION
IST1685I TCP/IP JOB NAME = TCPCS
IST2003I ENTERPRISE EXTENDER XCA MAJOR NODE NAME = XCAIP1A
IST2004I LIVTIME =    10  SRQTIME =    15  SRQRETRY =     3
IST2005I IPRESOLV =    0
IST924I -------------------------------------------------------------
IST2006I PORT PRIORITY =  SIGNAL    NETWORK     HIGH   MEDIUM     LOW
IST2007I IPPORT NUMBER =   12000      12001    12002    12003   12004
IST2008I IPTOS VALUE   =      C0         C0       80       40      20
IST924I -------------------------------------------------------------
IST1680I LOCAL IP ADDRESS 9.67.1.5
IST2009I RTP PIPES =          2      LU-LU SESSIONS      =          1
IST2010I INOPS DUE TO SRQRETRY EXPIRATION                =          0
IST1324I VNNAME = IP.GVRN5           VNGROUP = GPIP5     (GLOBAL)
IST2011I          AVAILABLE LINES FOR THIS EE VRN        =          0
IST2012I          ACTIVE CONNECTIONS USING THIS EE VRN   =          1
IST2013I AVAILABLE LINES FOR PREDEFINED EE CONNECTIONS   =          0
IST2014I ACTIVE PREDEFINED EE CONNECTIONS                =          0
IST2015I ACTIVE LOCAL  VRN EE CONNECTIONS                =          0
IST2016I ACTIVE GLOBAL VRN EE CONNECTIONS                =          1
IST924I -------------------------------------------------------------
IST2017I TOTAL RTP PIPES =          6     LU-LU SESSIONS =          3
IST2018I TOTAL ACTIVE PREDEFINED EE CONNECTIONS          =          0
IST2019I TOTAL ACTIVE LOCAL  VRN EE CONNECTIONS          =          0
IST2020I TOTAL ACTIVE GLOBAL VRN EE CONNECTIONS          =          3
IST2021I TOTAL ACTIVE EE CONNECTIONS                     =          3
IST314I END
```

DISPLAY RTPS by TCID

- When attempting to diagnose a possible problem with a specific RTP connection, customers often find it necessary to monitor the status and/or performance of a given RTP connection from "both ends".
- There is currently no easy way for customers to correlate the RTP (CNRxxxxx) PU name on one side of an RTP connection to the corresponding RTP PU name on the other side of that same RTP connection.
  - The only information associated with an RTP connection that can be used to correlate the local RTP PU name to the corresponding remote RTP PU name on the other side of the RTP connection is the "Remote TCID" (Transport Connection IDentifier).
  - The Local and Remote TCIDs for an RTP connection are provided in the output of the DISPLAY ID=CNRxxxxx command (on message IST1476I), but this requires that the RTP PU name already be known.
- The DISPLAY RTPS command is expanded to include the new TCID= operand, which allows an RTP PU to be found and displayed by its Local TCID.

DISPLAY RTPS by TCID...

NOTES

1. From the local host (NETA.SSCPAA), issue the DISPLAY ID=CNRxxxxx command and remember the value of the remote partner CP (from message IST1481I) and the Remote TCID (from message IST1476I).

```
d net,id=cnr00006
IST097I DISPLAY ACCEPTED
IST075I NAME = CNR00006, TYPE = PU_T2.1
    :
IST1963I APPNCOS = #INTER - PRIORITY = HIGH
IST1476I TCID X'14AB34050001001F' - REMOTE TCID X'14AB300100010020'
IST1481I DESTINATION CP NETB.SSCPBA - NCE X'D000000000000000'
IST1587I ORIGIN NCE X'D000000000000000'
    :
IST1480I RTP END TO END ROUTE - RSCV PATH
IST1460I TGN  CPNAME              TG TYPE      HPR
IST1461I  21  NETA.VRNA           APPN         RTP
IST1461I  21  NETB.SSCPBA         APPN         RTP
    :
IST314I END
```

2. From the remote host (NETB.SSCPBA), use DISPLAY RTPS,TCID= to display the corresponding RTP PU. (If desired, the TEST=YES operand can also be included on this command.)

```
d net,rtps,tcid=14AB300100010020
IST097I DISPLAY ACCEPTED
IST350I DISPLAY TYPE = RTPS
IST1695I PU NAME       CP NAME      COSNAME SWITCH CONGEST STALL SESS
IST1960I CNR00007 NETA.SSCPAA       #INTER    NO      NO     NO      1
IST1454I 1 RTP(S) DISPLAYED
IST314I END
```

Redbooks V1R6 ibm.com/redbooks

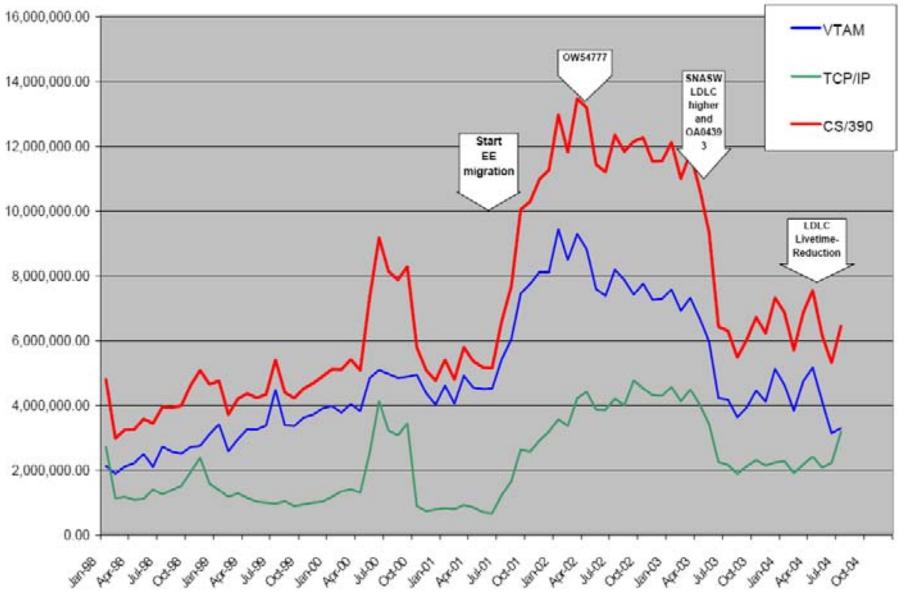
EE performance overview

- There has been significant effort put into EE performance enhancements
  - Increased throughput
  - Reduced CPU Utilization
- Most of the enhancements have been delivered via PTFs
  - Throughput: Apply OA02213 & PQ69398, and upgrade OSA Microcode Level to 3.26 (z/Series 2064 GA3) or 4.28 (G5/G6)
  - CPU utilization: Some customer environments will benefit from OA04393 (Inactivity Timer Optimization)
- Monitor the EE Informational APAR (II12223) for news on further enhancements

APAR	Purpose	PTF	Notes
OW53393	ARB Enhancements	UW94491	V1R2 only, base in V1R4
OW56896	LAN Idle	UA00067	V1R2 only, base in V1R4
OW52291	EE Packing & QDIOSTG Option	UA00131 - V1R2 UA00132 - V1R4	
OW53978	EE Outbound Data Ordering	UA00131 - V1R2 UA00132 - V1R4	Coreq: PQ69398
PQ69398	Fast UDP Outbound Ordering	UQ73923	V1R2 only, base in V1R4
OW57459 OW56893	HPR Resequencing	UA00131 - V1R2 UA00132 - V1R4	
OA02213	Send SRB Optimization	UA01999 - V1R2 UA02000 - V1R4	Will prereq all VTAM APARs in table

Redbooks V1R5 ibm.com/redbooks

CPU consumption during roll-out of EE

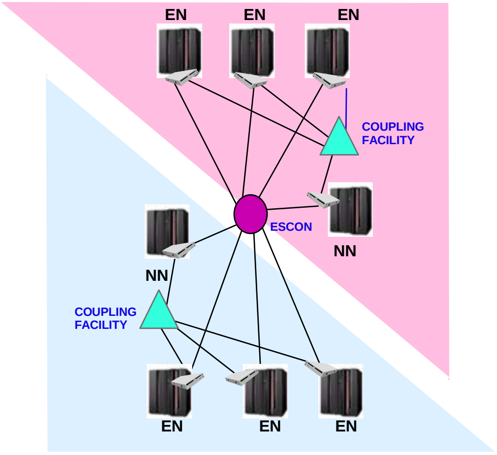


Redbooks

ibm.com/redbooks

Non-SYSPLEX Network Node Server for generic resources end nodes

- Currently the network node server for end nodes running generic resource applications must be connected to the same coupling facility structure as the served end nodes. To avoid a single point of failure this requires two network node servers in each sysplex configuration.
- Customers would like to have the flexibility of having a backup network node server that is not connected to the same sysplex as the served end nodes but continues to support the generic resource function, including session level load balancing.



Redbooks V1R5

ibm.com/redbooks

Non-SYSPLEX Network Node Server for generic resources end nodes - how to specify

- For the generic resource function to work when the network node is not connected to the generic resource structure, the end node must allow searching for unknown resources. This will allow the LOCATE, with the generic name, to be forwarded to the end node so the resolution of the generic name can be done on the end node.

- APPN architecture does not allow for end nodes to reply to LOCATEs with a different owning CP name. Therefore only the end node that owns the real instance can return the positive found reply to the LOCATE. Generic name resolution will be done by the first end node to receive the search and access the generic resource structure. This will create an affinity for that session set up. But if the resolved real instance does not reside on this end node, the affinity will be used by the owning end node to confirm the resolved name (affinity) and return the positive reply.

- Customers will control this function by coding the new ENBCAST operand on the network node server list major node.
  - The default value is ENBCAST=NO.
  - Coding ENBCAST=YES for the backup network node server in the network node server list allows a backup network node server (outside the Sysplex) to search the end node for the generic resource name (which is considered an unknown resource) as part of the end node broadcast search phase.
  - This function relies on the backup network node server to find the real instance during the domain broadcast and redirect the search correctly. This means that if any of the generic resource end nodes are using the backup network node server, then all of the other generic resource end nodes must be using that same backup network node server as well.
  - Due to the possible performance implications, we do not recommend that customers run their Generic Resources configuration in this "backup mode" of operation for a long time. The "backup mode" should only be used during a temporary outage of the primary network node server...which should still be attached to the same Sysplex as the GR ENs.

Redbooks V1R5 ibm.com/redbooks

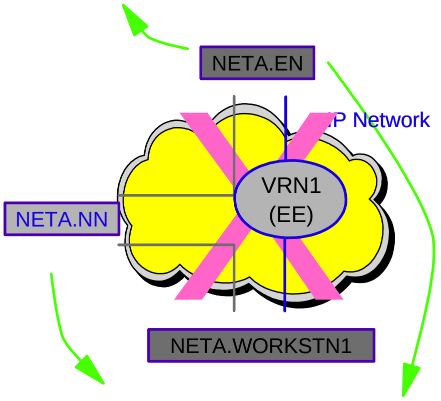
EE connection network reachability awareness of connection network failures

- Typical VRN configuration
  - ENs predefine links to NNS for CP-CP
  - Dynamic CN links used EN-to-EN

- What happens if IP network fails?
  - Affected RTPs begin path switching
  - VRN path is chosen again
    - Topology component is *not* notified of the failure
    - VRN path still has lowest weight
  - Path switch fails even though a functional alternate path exists (through NN)

- V1R5 will issue new message
  - For VRN dial failures and VRN link INOPs
    - IST1903I - Identifies the VRN and partner node
  - Allows network operator to take action
    - Disable PSRETRY
    - Prevent VRN from being used for new sessions
    - Re-enable both after IP network is restored
  - Similar actions may be required at the partner node and/or at an EN's NNS

```
IST1903I FAILURE OVER VRN
NETA.VRN1 TO CP NETA.WORKSTN1
```



```
     F VTAMOPTS,PSRETRY=(0,0,0,0)

     F TGP,ID=VRN1,TGN=21,TGPNAME=higher_weight
(or) V INACT,ID=LN_VRNxx,F (all lines to the VRN)
(or) F TOPO,ID=VRN1,FUNCTION=QUIESCE,SCOPE=NETWORK
     (at NNs only)
```

Redbooks V1R6 ibm.com/redbooks

EE connection network reachability awareness

NOTES

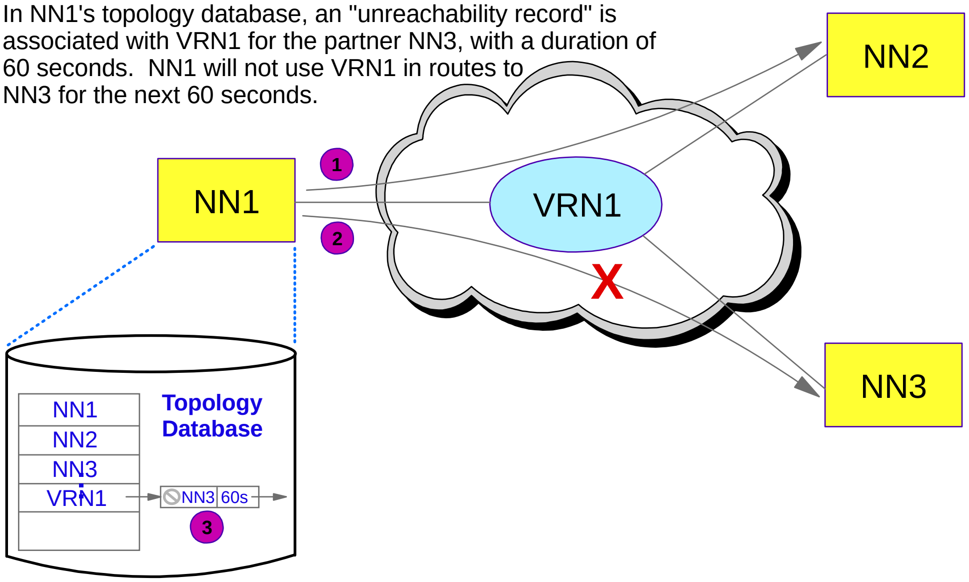
- EE Connection Network Reachability Awareness detects a dial failure or connection INOP for a connection over an Enterprise Extender connection network and prevents that specific path to the partner node from being used for a period of time. If alternate paths are available, APPN Topology and Routing services will select the optimal alternate session path for session establishment or an HPR path switch.
- When the time expires, if the path through the EE virtual routing node (VRN) still has the lowest weight of any available path to the partner node, the path over this particular VRN will be selected on the next attempt to redial the partner node.
- The period of time that a path through the EE VRN to the unreachable partner will remain unavailable is configurable.
- Unreachable partner information is maintained in the Topology Database and is associated with an EE VRN or with an end node that is on the origin side of the VRN.
- Unreachable partner information is sent to an end node's NNS or broadcast to a network node's adjacent network nodes in Topology Database Updates (TDUs).
- Once the unreachability period expires, a subsequent attempt to use the VRN will result in failure if the underlying problem with the connection has not been corrected. The new dial failure will again prevent selection of the path through this VRN to the unreachable partner node for the same period of time. This will continue until the problem with the connection path is corrected.
- It is possible for a path between two nodes, through the VRN, to be usable for route selection in one direction but not the other. This function can detect that distinction and will allow routing in the direction that is usable while preventing the path in the direction that is not usable from being selected when new session are established and HPR path switches occur.
- A new control vector has been architected to carry unreachable partner information on TDUs.

Redbooks V1R6

ibm.com/redbooks

EE connection network reachability awareness

- NN1 successfully contacts NN2 across VRN1.
- NN1's attempt to contact NN3 across VRN1 fails.
- In NN1's topology database, an "unreachability record" is associated with VRN1 for the partner NN3, with a duration of 60 seconds. NN1 will not use VRN1 in routes to NN3 for the next 60 seconds.



Redbooks V1R6

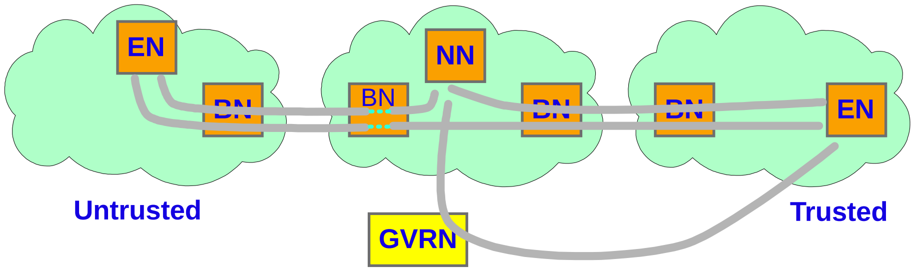
ibm.com/redbooks

EBN session awareness

- When HPR is used across subnet boundaries, the border nodes (EBNs) lose session awareness. This complicates authorization (DSME required instead of SME) and prevents accounting (at the EBN).

- Some EE/HPR exploiters (primarily service providers) require a method to:
  - Control and monitor which sessions are using their EBNs as a transport
  - Control which business partners can exploit Global VRNs (into or around their network)

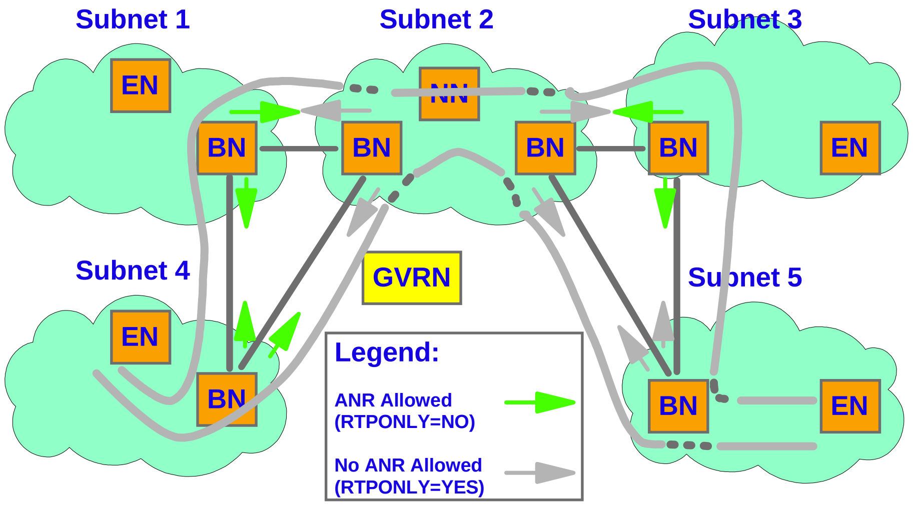
- V1R6 provides controls to force back-to-back RTP pipes at the EBN, thereby restoring session awareness on the border node



Redbooks V1R6 ibm.com/redbooks

EBN session awareness...

- New RTPONLY operand on ADJCP definitions allows control of session awareness at APPN subnetwork boundaries
  - Controls where EBNs are allowed to perform ANR routing and/or exploit GVRNs



Redbooks V1R6 ibm.com/redbooks

IBM®

What are the Major New Functions in CS z/OS V1R5 and V1R6?
-
IPv6

Redbooks

International Technical Support Organization

IPv6 - overview

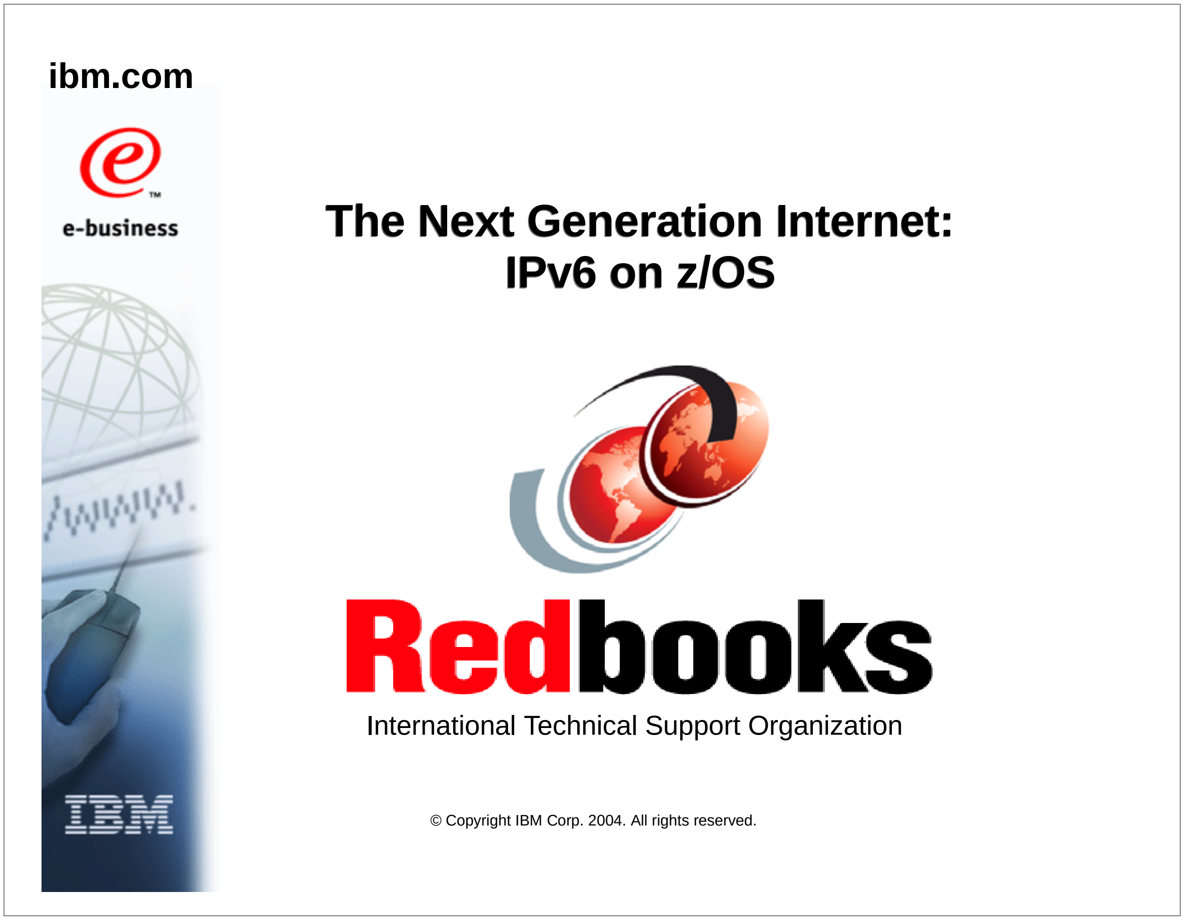
CS z/OS V1R5

- ✓IPv6 support added to various clients and servers: sendmail, SNMPD agent, osnmp command, SyslogD, SNTPD, TFTPD, DCAS, remote execution client commands and servers (rsh and rexec - both UNIX and MVS versions)
- ✓IPv6 enabled the CICS sockets APIs and listeners
- ✓IPv6 enabled the TN3270 server including SNA displays
- ✓Enabled new version-neutral MIBs for combined IPv4 and IPv6 management objects
- ✓IPv6-enabled QoS policies in the Policy Agent
- ✓Added first dynamic routing protocol for IPv6: RIPng (RIPv2 for IPv6) to OMPROUTE
- ✓IPv6-enabled Enterprise Extender
- ✓Added IPv6 support for XCF, SAMEHOST, and ESCON network interfaces (including Dynamic XCF)

CS z/OS v1R6

- ✓Adds OSPF for IPv6 support to OMPROUTE
- ✓IPv6-enable the IP Sysplex functions:
  - Dynamic VIPA
  - Sysplex Distributor
  - SourceVIPA
  - SNMP MIBs for dynamic VIPA
  - SysplexPorts

We will discuss IPv6 on z/OS in much more detail in a follow-on session later today.



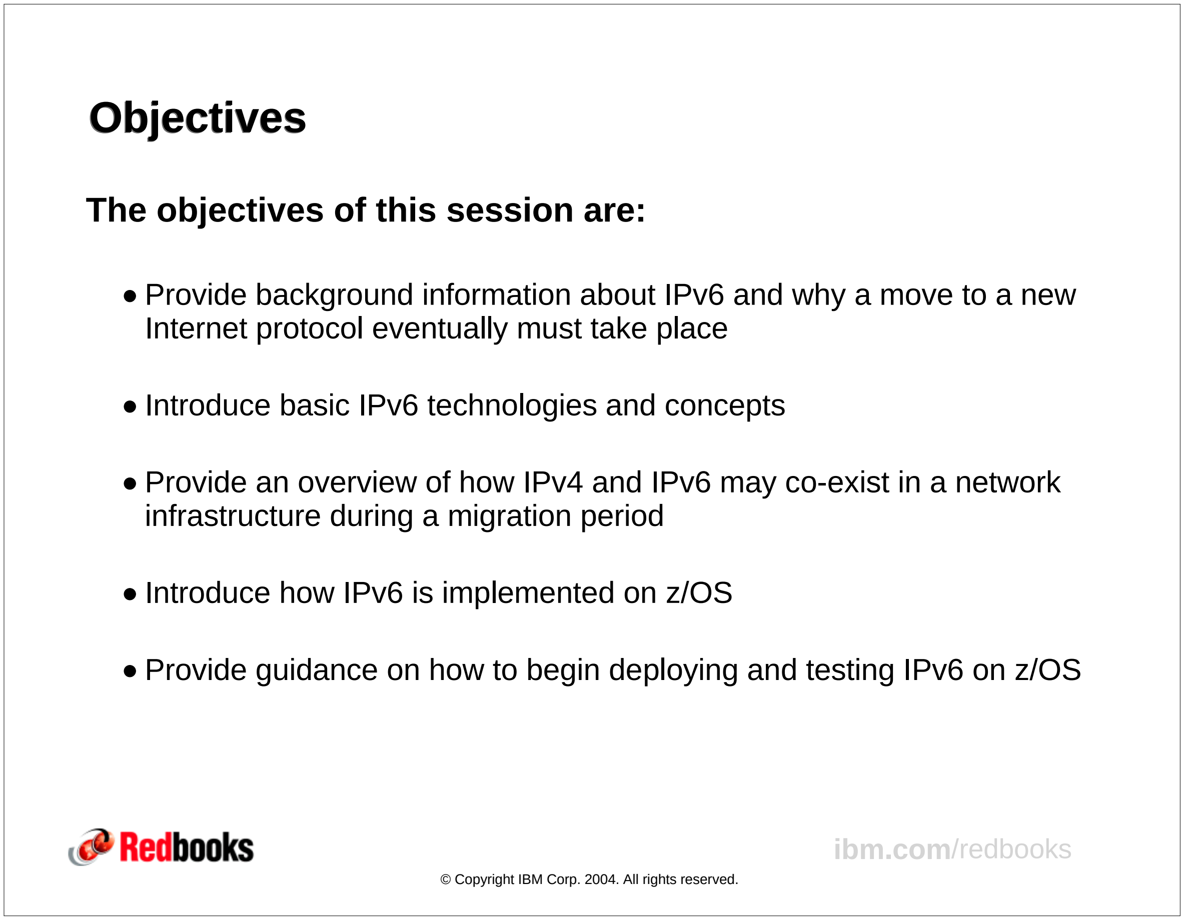
ibm.com

e-business

The Next Generation Internet: IPv6 on z/OS

Redbooks

International Technical Support Organization



Objectives

The objectives of this session are:

- Provide background information about IPv6 and why a move to a new Internet protocol eventually must take place
- Introduce basic IPv6 technologies and concepts
- Provide an overview of how IPv4 and IPv6 may co-exist in a network infrastructure during a migration period
- Introduce how IPv6 is implemented on z/OS
- Provide guidance on how to begin deploying and testing IPv6 on z/OS

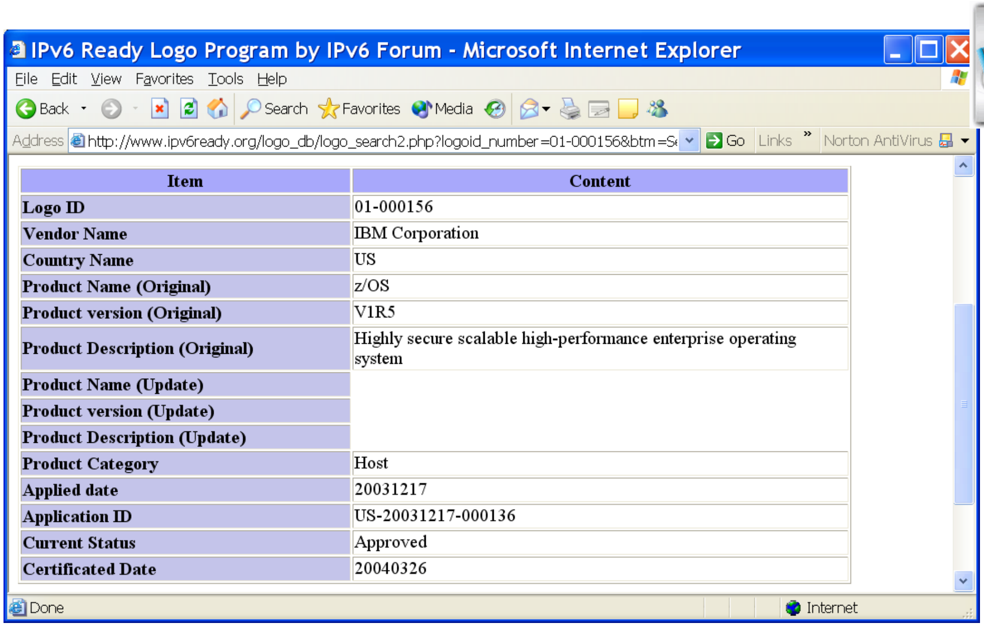
Redbooks

ibm.com/redbooks

Agenda

1. Why do we need a new Internet Protocol?
2. Brief technical introduction to IPv6
3. IPv4 / IPv6 coexistence and migration
4. IPv6 on z/OS
   - System structure
   - Status
   - What's to come

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


IPv6 Ready Logo Program by IPv6 Forum - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back
Search
Favorites
Media
Address http://www.ipv6ready.org/logo_db/logo_search2.php?logoid_number=01-000156&btm=S
Go
Links
Norton AntiVirus
Done
Internet

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

IBM®

The Next Generation Internet: IPv6 on z/OS

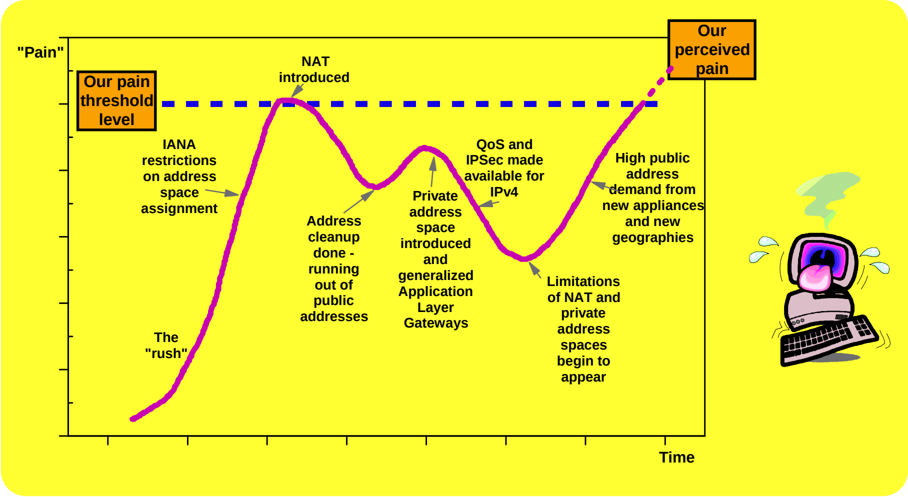
-

Why do we need a new Internet protocol?

Redbooks

International Technical Support Organization

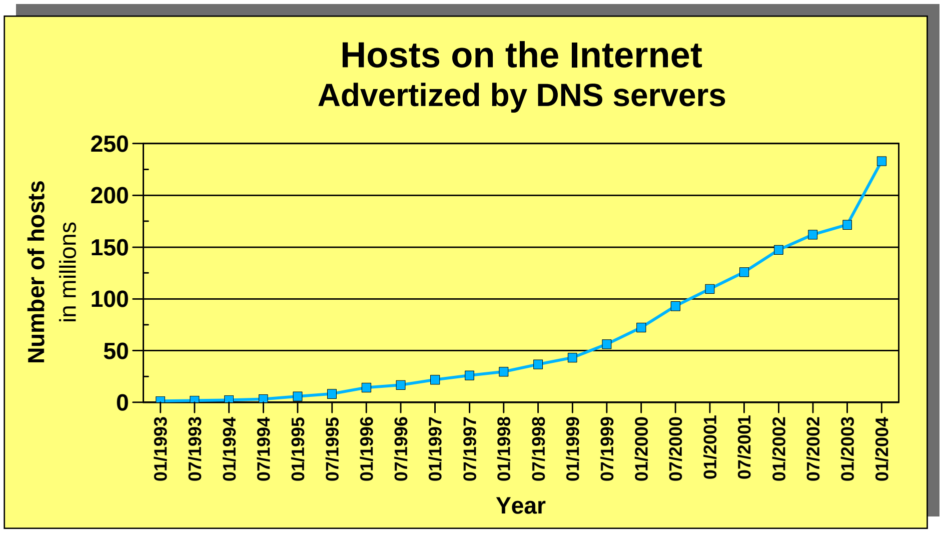
The "pain" curve
Managing the IPv4 address space



Redbooks

ibm.com/redbooks

Visible IPv4 hosts on the Internet through the last 10 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).

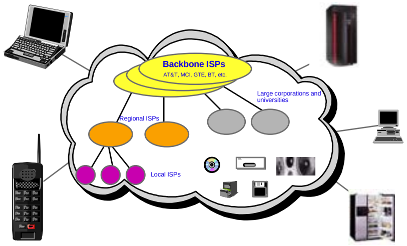
Redbooks

ibm.com/redbooks

Secure IP connectivity for anyone from anywhere to anything!

- **Growing mobility of users**
  - Internet access from anywhere (car, home, office)
  - Multiple addresses per person
  - Pervasive Computing
- **Continued rapid growth of the Internet**
  - China plans to roll out 1 billion Internet nodes, starting with a 320 million student educational network
  - Asia/Pacific, and to a lesser extent Europe, missed out on the early IPv4 address allocations
- **Government support**
  - Wide-scale IPv6 promotion underway in Japan, Korea, and Taiwan
  - European Commission (EC) encourages IPv6 research, education, and adoption in member countries
  - Government agencies beginning to mandate IPv6 capable technology
- **Convergence of voice, video and data on IP**
  - Need for reliable and scalable architecture
  - "Always-on connections"
- **New application opportunities**
  - Potentially unlimited number of IP nodes (vehicles, devices, components, individual parts, etc.)
- **Security becomes more and more important**
  - Various optional security features have been patched on top of IPv4
  - IPv6 has security features defined as part of the base protocol

The Internet - a worldwide digital utility.



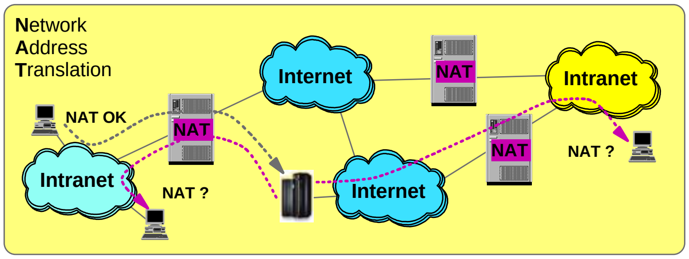
Connectivity for ***anyone*** from ***anywhere*** (car, plane, home, office) to ***anything***!

IPv6 promises true end-to-end connectivity for peer-based collaborative solutions.

Redbooks

ibm.com/redbooks

Couldn't we just add more Network Address Translating (NAT) firewalls to deal with the limited number of IPv4 addresses?



NATs work best in small end-sites, client-only

- All connections originate from clients (outbound only)
- Only a subset of clients need Internet access at any point in time
- A pool of public addresses matching the number of clients who need concurrent Internet access need to be available even when NAT is used
- Little configuration/administration is needed
- Limited applicability

When clients are servers (inbound connections):

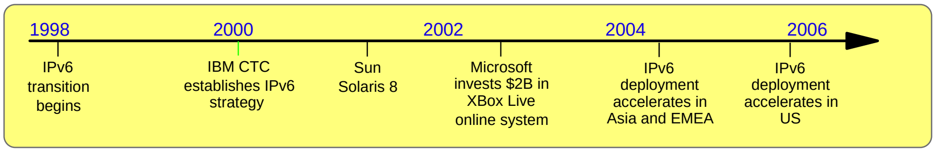
- Static NATing (manual configuration on NAT device)
- If most/all clients are servers, NAT multiplexing premise fails

- Shortage of IPv4 addresses has led to extensive use of private (not globally reachable) addresses
  - Requires Network Address Translators (NATs) or application layer gateways at intranet/Internet boundaries
  - Every NAT node between a private and a public network needs a pool of public IP addresses - adding more NAT nodes requires more public IP addresses
- NATs are a pain to design around and are generally a severe barrier to continued Internet scaling
  - NATs break protocols (FTP, IPSec, DRDA, EE, etc.) that rely on globally unique addresses
  - NATs are very often sensitive to application data being encrypted (SSL/TLS, IPSec, Kerberos)
  - NATs have operational and administrative scaling problems
  - Always-on devices need permanent, global addresses (NATs prevent this)
  - Barrier to deployment of new types of applications (true peer-to-peer)
  - Convergence of voice, video, and data over IP
- IPv6 alleviates these problems and removes barriers to continued Internet expansion

IPv6 market drivers

Investment in improved infrastructure capabilities reduces costs and enables new uses and revenues

- **IPv4 operational constraints and resulting increased operating costs**
  - Limited addressing
  - Network routing inefficiencies
- **Fortune 1000 - 70% have had to deploy NATs**
- **IPv6 improvements bring decreased operating costs**
  - Addressing constraints removed
  - Simplified header architecture and protocol
  - "Built-in" security
  - Autoconfiguration



1998 — IPv6 transition begins

2000 — IBM CTC establishes IPv6 strategy

Sun Solaris 8

2002 — Microsoft invests $2B in XBox Live online system

2004 — IPv6 deployment accelerates in Asia and EMEA

2006 — IPv6 deployment accelerates in US

- **Telecommunications**
  - Unified communication services
  - 3G wireless
  - Civilian wireless networks predicted at 1B users in 2006
  - 420M new mobile terminals sold in 2002
  - Billing applications
  - Network management services
- **Government**
  - Homeland Security
  - Department of Defense (telematic ID of individual pieces of equipment)
- **Financial Services**
  - Online banking via wireless
- **Entertainment**
  - Online gaming predicted to grow from $210M (2002) to $1.8B (2005)
  - 114M online gaming users predicted by 2006
- **Distribution**
  - Retail Services
- **Media**
  - Real-time video
  - Video on demand
  - Interactive media
- **Education**
  - Distance Learning

Source: "IPv6 - An Internet Evolution", www.IPv6.org

Why has deployment been slow so far?

- **Economic slowdown has slowed growth and spending**
  - Network infrastructure vendors are not introducing new products quickly
  - Service providers are not upgrading and expanding networks
- **IPv6 upgrades to network infrastructure are expensive**
  - IPv6 routing performance requires hardware upgrades
  - New technology requires staff training
  - New code/additional complexity will cause added support burdens
  - No current revenue stream to justify the costs
- **Major technology markets are comfortable with IPv4**
  - US and Europe have (relatively) many IPv4 addresses
  - Address shortages have been mitigated by the use of NAT
- **Benefits of IPv6 are not widely understood or not compelling**
  - Desire that it solves more problems (e.g., multihoming)
- **Need critical mass of IPv6 peers for tangible benefits**
  - Chicken and egg problem; limited incentive for legacy IPv4 sites
  - Deployments of new devices and associated new infrastructure do not have these constraints
  - ISPs will not move until pressured to do so by customers
- **Potential for rapid adoption when critical mass is reached**
  - Applications + Middleware + Infrastructure (OS, routers)
  - A few big customers will show the way

- IPv0 to IPv3 were early research and development versions.
- The name IPv5 was used for something else (the Internet Stream Protocol - RFC1819)
- IPng: IP next generation - an early name for what became IPv6

Redbooks

ibm.com/redbooks

IBM®

The Next Generation Internet: IPv6 on z/OS

-

Brief technical introduction to IPv6

Redbooks

International Technical Support Organization

What is IPv6?

- **IPv6 is an evolution of the current version of IP, which is known as IPv4**
  - Work on new IETF standard started in early 90's under the name IPng (IP next generation)
  - Not backward compatible, but migration techniques defined
- **Today's IPv4 has 32-bit addresses**
  - Theoretical limit is 4,294,967,295 addresses
  - Practical limit is significantly less - predictions range from around 250,000,000 to 1,000,000,000
- **IPv6 provides almost unlimited number of addresses**
  - IPv6 addresses are 128 bits
  - No practical limit on global addressability
  - Enough address space to meet all imaginable needs for the whole world and for generations to come
  - More addresses *cannot* be retrofitted into IPv4
- **Other important improvements:**
  - Facilities for automatic configuration
  - Improved support for site renumbering
  - End-to-end IP security
  - Mobility with route optimization (important for wireless)
  - Miscellaneous minor improvements

IPv4 Address:
9.67.122.66

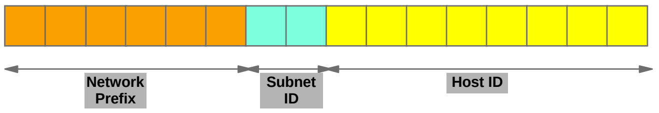
IPv6 Address:
2001:0DB8:4545:2::09ff:fef7:62dc

An added advantage of IPv6:

- You don't have to try and remember those IP addresses any longer - IPv6 addresses are plain impossible to remember!
- A DNS infrastructure is an absolute requirement in an IPv6 environment.

Expanded routing and addressing

- Expanded size of IP address space
  - Address space increased to 128 bits
    - Provides 340,282,366,920,938,463,463,374,607,431,768,211,456 addresses



Prefix-bits designate address scope:

- **link-local scope**
- **site-local scope**
- **global scope**

- As important as the expanded address space is the use of hierarchical address formats:
  - Allocation architecture
    - Network portion of address allocated by ISPs
    - Subnet portion of address is allocated by customer
    - Host Identifier is derived from the MAC address of the interface adapter
  - Facilitates efficient routing architectures
    - IPv6 uses CIDR (Classless InterDomain Routing), first introduced in IPv4
  - IPv6 hierarchical routing likely only viable method for keeping the size of the backbone router tables under control
    - Even with hierarchical routing, the current IPv4 Internet backbone maintains 90,000 or more routes

IPv6 address types and scopes

Unicast:

- Assigned to one interface. Packets destined for a unicast address are sent to only one node.
- Can be link-local scope, site-local scope or global scope

Multicast:

- Provides a means for a source to communicate with a group

Anycast:

- Allows the source to communicate with the closest member of a group

The first few bits of the address (format prefix) identify the address type and scope:

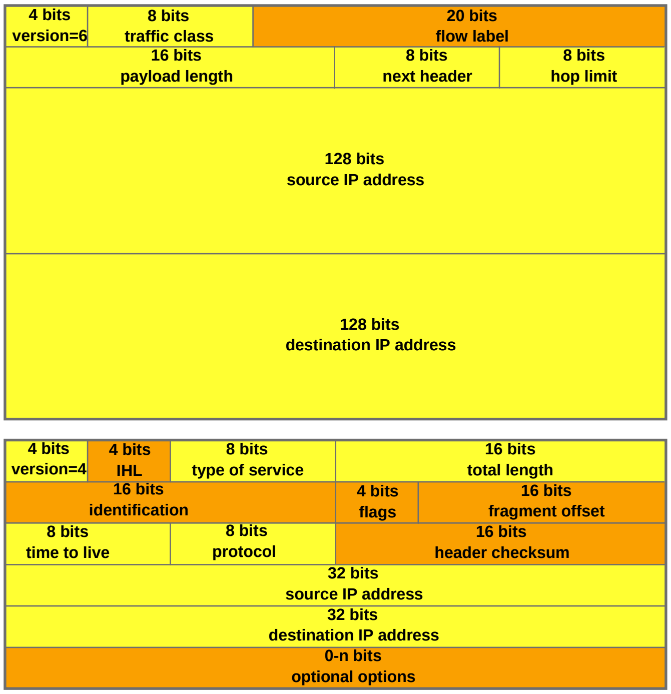
Address type	Binary prefix	IPv6 notation
Unspecified	00...0 (128 bits	::/128
Loopback	00...1 (128 bits)	::1/128
Multicast	11111111	FF00::/8
Link-local unicast	1111111010	FE80::/10
~~Site-local unicast~~	~~1111111011~~	~~FEC0::/10~~
Global unicast	(everything else)	

Site-local addresses have been deprecated by the IETF, but they may still be part of some implementations.

Multicast addresses (groups) also have scopes

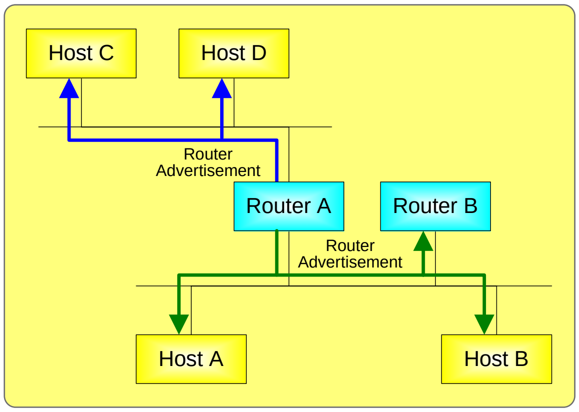
IPv6 header format

- Changes to the IP header
  - Header is fixed length
    - Optional headers daisy-chained
  - Address space quadrupled to 16 bytes
  - Fragmentation fields moved out of base header
  - Header checksum eliminated
    - Performed by the link layer
  - No hop-by-hop fragmentation
    - Path MTU discovery determines MTU size
  - Flow label added
  - Some fields renamed
    - Time to Live -> Hop Limit
    - Protocol -> Next Header
    - Type of Service -> Traffic Class



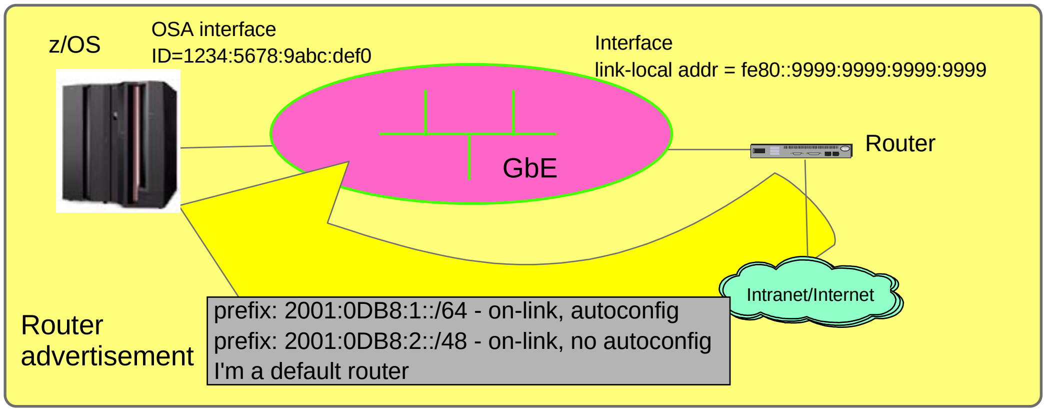
Stateless address autoconfiguration

- **Address configuration without separate DHCP server**
  - Router *is* the server, advertising key address configuration information
- **Address formed by combining routing prefix with Interface ID (IID)**
- **Link-local address configured when an interface is enabled**
  - Allows immediate communication with devices on the local link
  - Primarily used for bootstrapping and discovery
  - Well-known prefix combined with locally generated 64-bit IID
- **Other addresses configured via router advertisements**
  - RA advertises 64-bit prefixes (e.g., on-link, form an address)
  - Public (e.g., server) addresses formed from IID



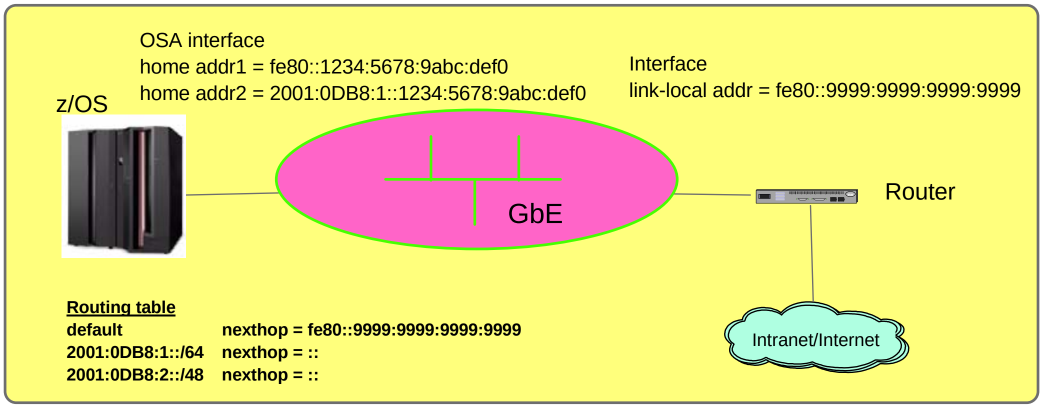
Router discovery - example

- With Router Discovery, a router informs a host of what prefixes are assigned to the medium
  - A prefix is a concept that's analogous to an IPv4 subnet, but with some key differences that will be discussed later
- A router can also use Router Discovery to advertise itself as a default router
- CS for z/OS supports receiving, but not sending, router advertisements
  - This support is in the TCP/IP stack
  - No routing daemon necessary to exploit


z/OS
OSA interface
ID=1234:5678:9abc:def0
Interface
link-local addr = fe80::9999:9999:9999:9999
GbE
Router
Intranet/Internet
Router advertisement
prefix: 2001:0DB8:1::/64 - on-link, autoconfig
prefix: 2001:0DB8:2::/48 - on-link, no autoconfig
I'm a default router

Stateless address autoconfiguration - example

- Stateless Address Autoconfiguration builds on the Router Discovery function described on the previous chart
  - A link-local address is always configured by adding the link-local prefix (FE80::/64) to the 64-bit interface ID
  - If the autoconfig flag in the router advertisement is set, other IP addresses are configured by adding an advertised prefix to the 64-bit interface ID
  - Routes are added to the routing table for prefixes that have the on-link flag set
  - On-link and autoconfiguration can both be set, or be independent
- The graphic below shows the result of autoconfiguration and contents of the routing table, after the router advertisement in the previous slide is processed.


OSA interface
home addr1 = fe80::1234:5678:9abc:def0
home addr2 = 2001:0DB8:1::1234:5678:9abc:def0
z/OS
Interface
link-local addr = fe80::9999:9999:9999:9999
GbE
Router
Intranet/Internet

Routing table

default	**nexthop = fe80::9999:9999:9999:9999**
2001:0DB8:1::/64	**nexthop = ::**
2001:0DB8:2::/48	**nexthop = ::**

IBM®

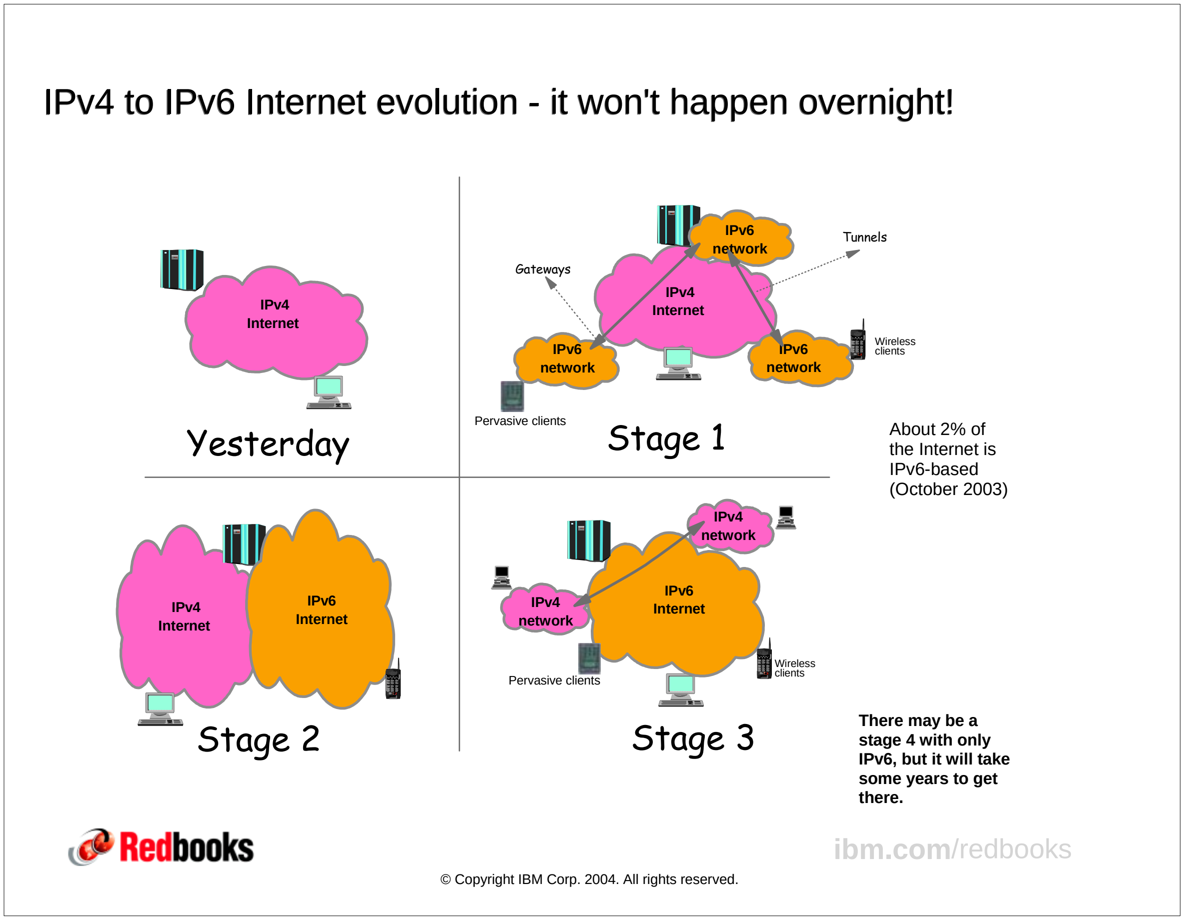
The Next Generation Internet: IPv6 on z/OS

-

IPv4 / IPv6 coexistence and migration

Redbooks

International Technical Support Organization



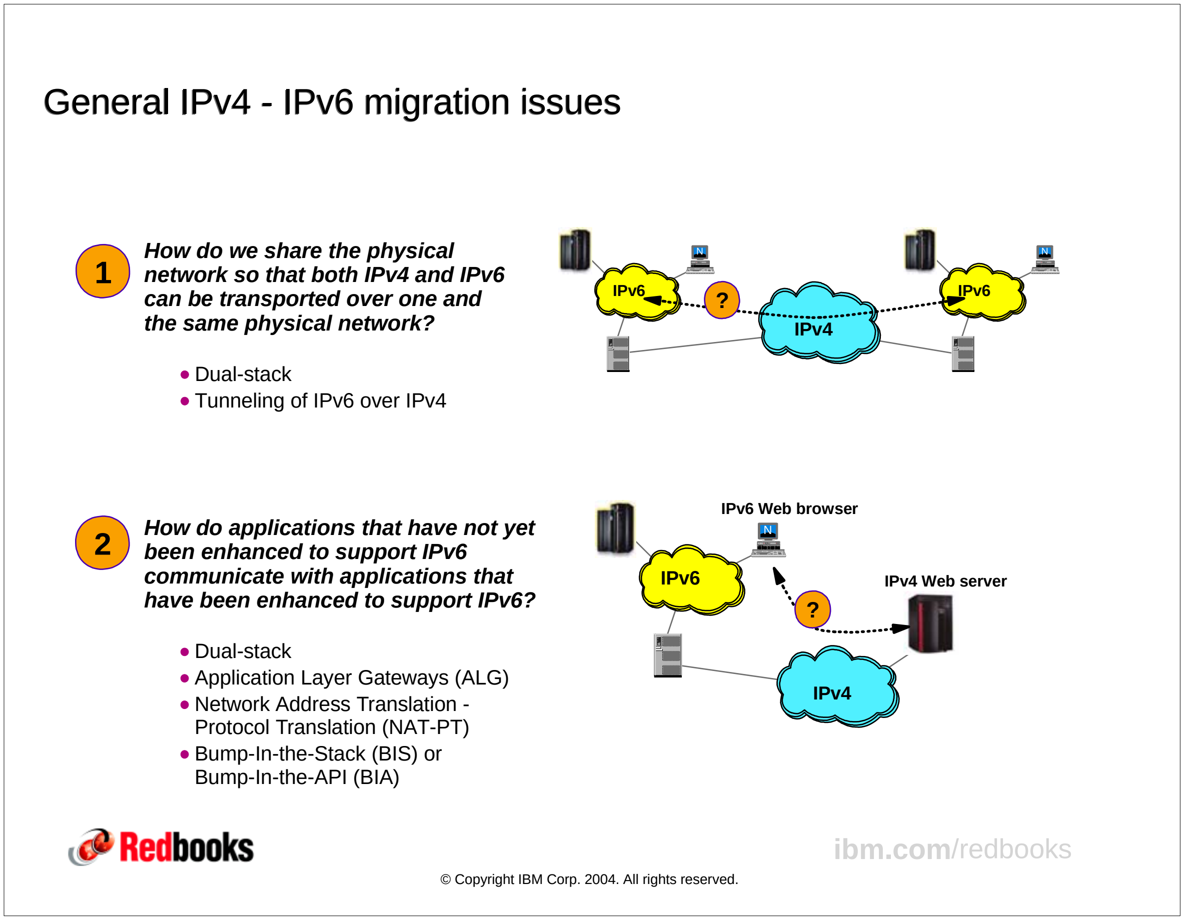
IPv4 to IPv6 Internet evolution - it won't happen overnight!

About 2% of the Internet is IPv6-based (October 2003)

There may be a stage 4 with only IPv6, but it will take some years to get there.

Redbooks

ibm.com/redbooks



General IPv4 - IPv6 migration issues

1 ***How do we share the physical network so that both IPv4 and IPv6 can be transported over one and the same physical network?***

- Dual-stack
- Tunneling of IPv6 over IPv4

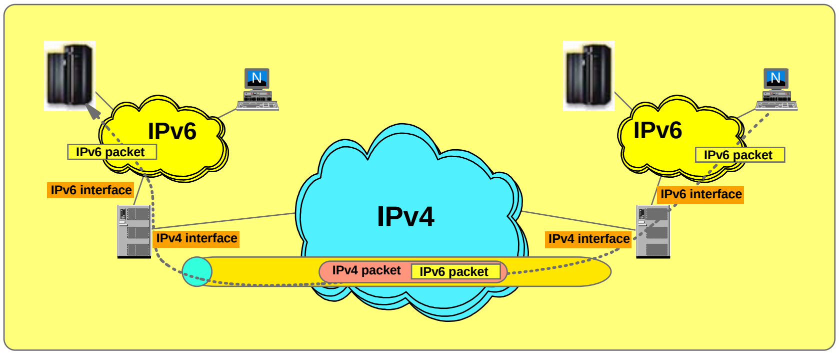
2 ***How do applications that have not yet been enhanced to support IPv6 communicate with applications that have been enhanced to support IPv6?***

- Dual-stack
- Application Layer Gateways (ALG)
- Network Address Translation - Protocol Translation (NAT-PT)
- Bump-In-the-Stack (BIS) or Bump-In-the-API (BIA)

Redbooks

ibm.com/redbooks

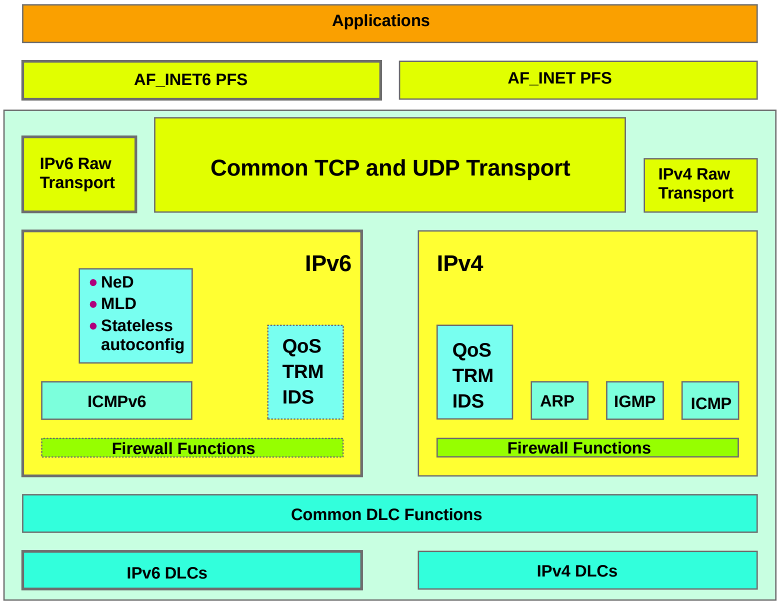
IPv6 over IPv4 network tunneling overview



- Tunneling: encapsulating an IPv6 packet in an IPv4 packet and sending the IPv4 packet to the other tunnel endpoint IPv4 address.
- Requires applications on both endpoints to use AF_INET6 sockets
- Tunnel endpoints can be in hosts or routers
  - The tunnel endpoint may be an intermediate node, the final endpoint, or a mixture of the two
- The tunnel endpoint placement depends on connectivity needs
  - Placing endpoints in routers allows entire sites to be connected over an IPv4 network
  - Placing endpoints in hosts allows access to remote IPv6 networks without requiring updates to the routing infrastructure
- z/OS does not support being a tunnel endpoint
  - Tunnel endpoints must be implemented on downstream routers connected to z/OS via a QDIO IPv6 interface

Redbooks ibm.com/redbooks

Generalized dual-mode TCP/IP structure



Applications

AF_INET6 PFS

AF_INET PFS

IPv6 Raw Transport

Common TCP and UDP Transport

IPv4 Raw Transport

IPv6

- NeD
- MLD
- Stateless autoconfig

QoS
TRM
IDS

ICMPv6

Firewall Functions

IPv4

QoS
TRM
IDS

ARP

IGMP

ICMP

Firewall Functions

Common DLC Functions

IPv6 DLCs

IPv4 DLCs

A dual-mode (or dual-stack) TCP/IP implementation supports both IPv4 and IPv6 interfaces - and both AF_INET and AF_INET6 applications.

The dual-mode TCP/IP implementation is a key technology for IPv4 and IPv6 coexistence in an internet.

z/OS uses the dual-mode implementation - a single TCP/IP address space that handles both IPv4 and IPv6.

Redbooks ibm.com/redbooks

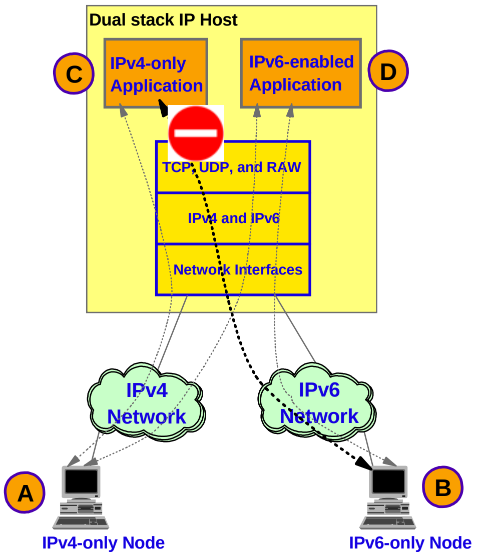
Sockets API considerations when moving to AF_INET6

NOTES

- IPv6 addresses are 128-bit in size as compared to 32 bits for IPv4
  - Data structures which store IP addresses must be modified to handle the larger size
- DNS Resolver library changes
  - New DNS calls replace gethostbyname() and gethostbyaddr()
- Textual representation of the IP address has changed
  - IPv4 addresses use dotted-decimal format
  - IPv6 addresses use colon-hex notation
- IPv6 has several scopes for IP addresses
  - An address is only unique within its given scope
  - On multihomed hosts, an IP address alone may be insufficient to select the interface over which to route
    - True for link-local addresses, may also be true for site-local addresses
  - Most applications will not care about this, but it is possible that some may
- IP addresses should not be assumed to be permanent
  - Long-term use of an address is discouraged due to renumbering
  - Applications should rely on DNS resolvers to cache the appropriate IP addresses
- sockaddr_in6
  - analogous to sockaddr_in, but larger
  - Holds 128-bit IPv6 address, port numbers, plus Flow Label and Interface Identifier
  - Two versions of sockaddr_in6 are available, converging the 4.3 and 4.4 BSD variants.
- in6_addr
  - analogous to 32-bit in_addr
  - holds a 128 bit address
- Socket calls to investigate for possible changes
  - socket(), bind(), connect(), sendmsg(), sendto(), accept(), recvfrom(), recvmsg(), getpeername(), getsockname()
- New calls
  - inet_pton(), inet_ntop()

Redbooks ibm.com/redbooks

IPv4 to IPv6 application end-to-end communication

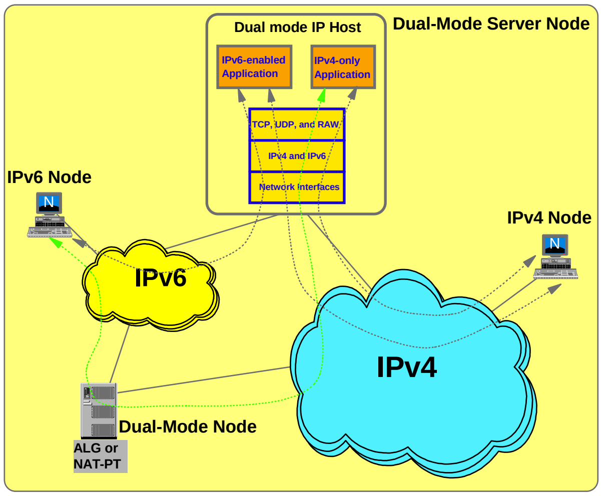


- An IPv6 application on a dual-mode stack can communicate with both IPv4 and IPv6 applications residing on IPv4-only or dual-mode stacks. D can communicate with A and B.
- An IPv6 application on an IPv6-only stack cannot communicate with an IPv4 application - even if it executes on a dual-mode stack. C cannot communicate with B.
- So from a migration/coexistence point of view, it is the communication between B and C that is of concern.
- Initially, there will be applications that are IPv4-only applications - even when they execute on a new dual-mode stack. These applications cannot be used from partner applications that reside on IPv6-only nodes.
  - IPv6-only nodes are not expected to become common right away. It is expected that most implementations of IPv6 will be as dual-mode stacks where IPv4 applications can communicate via the IPv4 side of the dual-mode stacks.
- Various technologies to address this application incompatibility issue have been identified.

Redbooks ibm.com/redbooks

Communication between IPv6 nodes and IPv4 nodes or applications

- Tools that enable communication between IPv6 nodes and IPv4 nodes or applications typically involve some form of translation
- This translation can be performed at the IP, transport, or application layer
  - At the IP layer, Simple IP/ICMP Translator (SIIT) may be used
    - Network Address Translator-Protocol Translator NAT-PT is built on top of SIIT
  - At the transport layer, SOCKS has been updated to allow IPv6/IPv4 relaying
    - The TCP or UDP connections are terminated at the boundary of the IPv6 domain and relayed to the IPv4 domain
  - At the application layer, proxies (sometimes referred to as Application Layer Gateways or ALGs) can be run on dual mode stacks



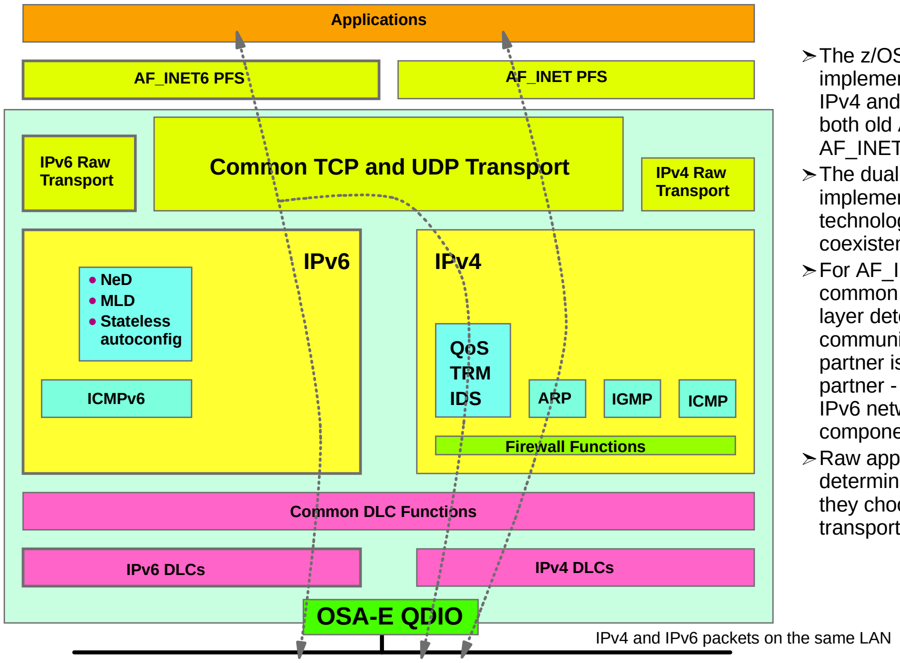
ibm.com/redbooks

IBM®

The Next Generation Internet: IPv6 on z/OS
-
IPv6 on z/OS

International Technical Support Organization

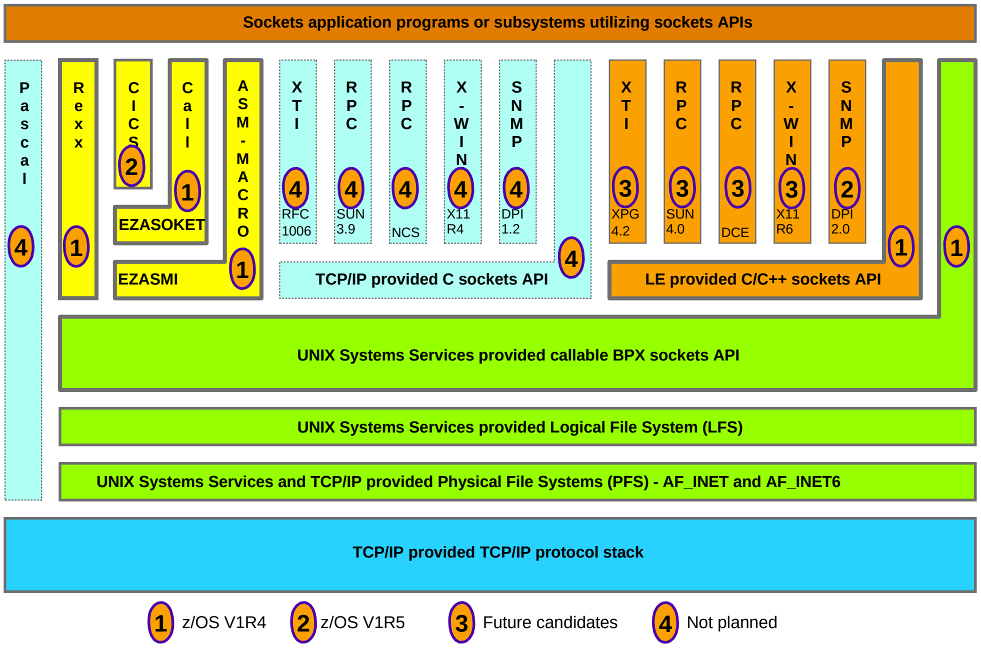
Flow between LAN and applications on an IPv6-enabled z/OS LPAR



- The z/OS dual-mode TCP/IP implementation supports both IPv4 and IPv6 interfaces - and both old AF_INET and new AF_INET6 applications.
- The dual-mode TCP/IP implementation is a key technology for IPv4 and IPv6 coexistence in an internet.
- For AF_INET6 applications, the common TCP or UDP transport layer determines per communication partner if the partner is an IPv4 or an IPv6 partner - and chooses IPv4 or IPv6 networking layer component based on that.
- Raw applications make the determination themselves when they choose IPv4 or IPv6 raw transport.

OSA-E for IPv6 requires zSeries hardware

IPv6: Sockets-related API AF_INET6 enablement overview, status, and plans


Sockets application programs or subsystems utilizing sockets APIs
Pascal 4
Rexx 1
CICS 2
Call 1
EZASOKET
ASM-MACRO 1
EZASMI
XTI 4 RFC 1006
RPC 4 SUN 3.9
RPC 4 NCS
X-WIN 4 X11 R4
SNMP 4 DPI 1.2
4
TCP/IP provided C sockets API
XTI 3 XPG 4.2
RPC 3 SUN 4.0
RPC 3 DCE
X-WIN 3 X11 R6
SNMP 2 DPI 2.0
1
LE provided C/C++ sockets API
1
UNIX Systems Services provided callable BPX sockets API
UNIX Systems Services provided Logical File System (LFS)
UNIX Systems Services and TCP/IP provided Physical File Systems (PFS) - AF_INET and AF_INET6
TCP/IP provided TCP/IP protocol stack
1 z/OS V1R4
2 z/OS V1R5
3 Future candidates
4 Not planned

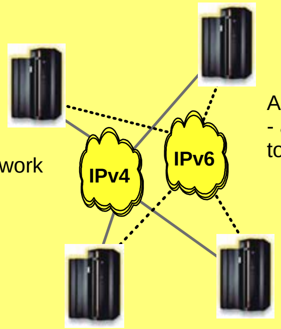
IPv6 network interface support

IPv6 network interfaces supported by CS in z/OS V1R5:

- ✓ IPv6 Loopback interface
- ✓ OSA-Express QDIO interface (zSeries hardware)
  - Gigabit Ethernet
  - Fast Ethernet
- ✓ IUTSAMEHOST to other stacks in same LPAR
- ✓ XCF to other stacks in same Sysplex
- ✓ ESCON/FICON (MPCPTP) to another z/OS image (not to any known channel-attached routers)

Two logical networks:

- an IPv4 network
- an IPv6 network



A separate IPv4 network - assign a separate subnet to this IPv4 network

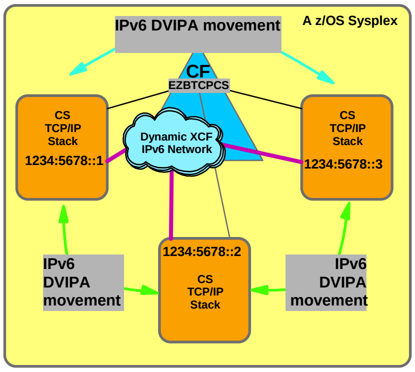
A separate IPv6 network - assign a separate prefix to this IPv6 network

The two logical networks may share the same OSA-Express adapters and the same physical network infrastructure (cabling, switches, etc.)

Redbooks V1R5 ibm.com/redbooks

IPv6 Sysplex support in z/OS V1R6

- **IPv6 Dynamic VIPA support**
  - Up to 1024 IPv6 Dynamic VIPA addressed per stack
- **IPv6 support for stack managed Dynamic VIPA addresses (VIPADEFINE/VIPABACKUP)**
- **IPv6 support for application-specific Dynamic VIPA addresses (VIPARANGE)**
- **IPv6 support for distributed Dynamic VIPA addresses and distribution of IPv6 workload by Sysplex Distributor (VIPADISTRIBUTE):**
  - WLM-based distribution
  - Round-robin distribution
  - Server affinity
  - Passive mode FTP support
  - Fast connection reset support
- **IPv6 support for Sysplex sockets**
- **IPv6 support for source VIPA address use:**
  - Interface-based selection of source VIPA
  - Sysplex-wide source VIPA addresses
  - Job-specific source VIPA
- **SNMP MIB support for IPv6 dynamic VIPA addresses**



A z/OS Sysplex

IPv6 DVIPA movement

CF
EZBTCPCS

CS TCP/IP Stack 1234:5678::1

Dynamic XCF IPv6 Network

CS TCP/IP Stack 1234:5678::3

1234:5678::2 CS TCP/IP Stack

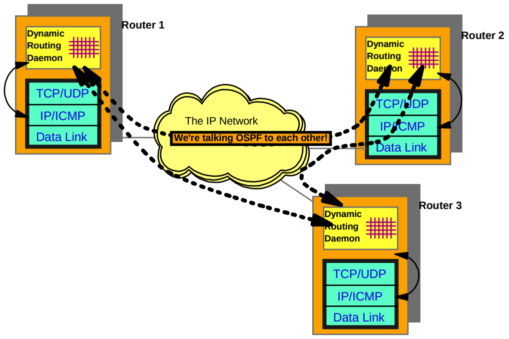
IPv6 DVIPA movement

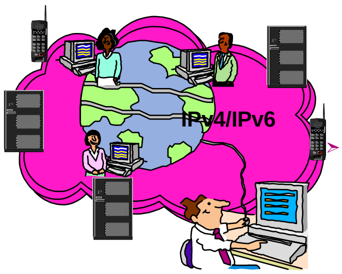
IPv6 DVIPA movement

Redbooks V1R6 ibm.com/redbooks

Dynamic routing and network management for IPv6

- **OMPROUTE extended with support for dynamic routing for IPv6: RIPng and OSPFv3**
  - Like IPv4, there is support for the routing protocol, plus support for basic IPv6 routing concepts
    - generic interfaces
    - static routes
    - direct routes
    - prefix and router advertisement routes
  - This new support has been added to OMPROUTE alongside its existing IPv4 dynamic routing support
  - You use new sets of IPv6 configuration statements and display commands to activate and monitor this new support
  - RIPng added in z/OS V1R5
  - OSPFv3 added in z/OS V1R6





- **Network management SNMP support**
  - Support SNMP agent (OSNMPD)
  - IPv6 MIB support - new RFC drafts have been published that define IP version-neutral objects
    - RFC2011 (IP and ICMP)
    - RFC2012 (TCP)
    - RFC2096 (IP routes)
    - RFC2233 (Interfaces) - this one is not version neutral
- **SMF119 records support**
  - The redesign of SMF records from SMF118 to SMF119 in z/OS V1R2 did factor in IPv6 addresses, so most subtypes are already in z/OS V1R4 supporting IPv6 addresses
  - Some changes needed to selected records to capture additional IPv6-related data, such as interface records and statistics records

V1R5

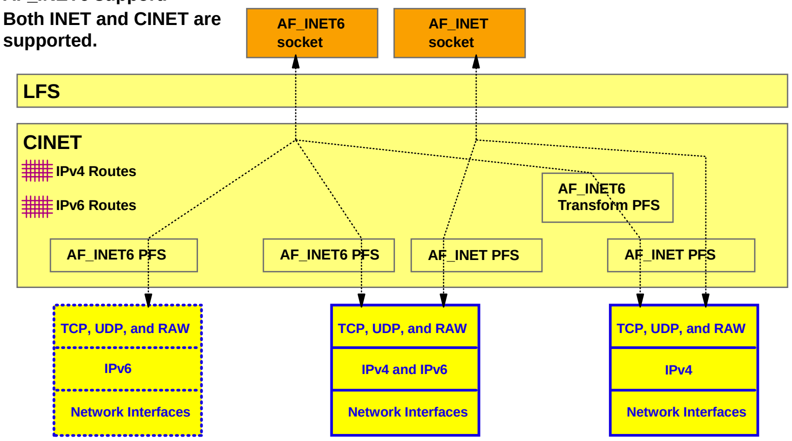
ibm.com/redbooks

How do we enable IPv6 support on z/OS and what are the consequences?

IPv6 is enabled at an LPAR level via an option in BPXPRMxx to enable AF_INET6 support.

Both INET and CINET are supported.

When IPv6 is enabled, a z/OS V1R4 TCP/IP stack will always have an IPv6 Loopback interface. You can define real IPv6 interfaces in addition to the loopback interface.



IPv6-only TCP/IP Stack

This will not be the case on z/OS for the foreseeable future! An AF_INET6 stack is required to also support AF_INET!

Dual Mode TCP/IP Stack

A z/OS V1R4 TCP/IP stack will always come up as dual-mode if AF_INET6 is enabled in BPXPRMxx

IPv4-only TCP/IP Stack

(such as AnyNet or an OEM TCP/IP stack)

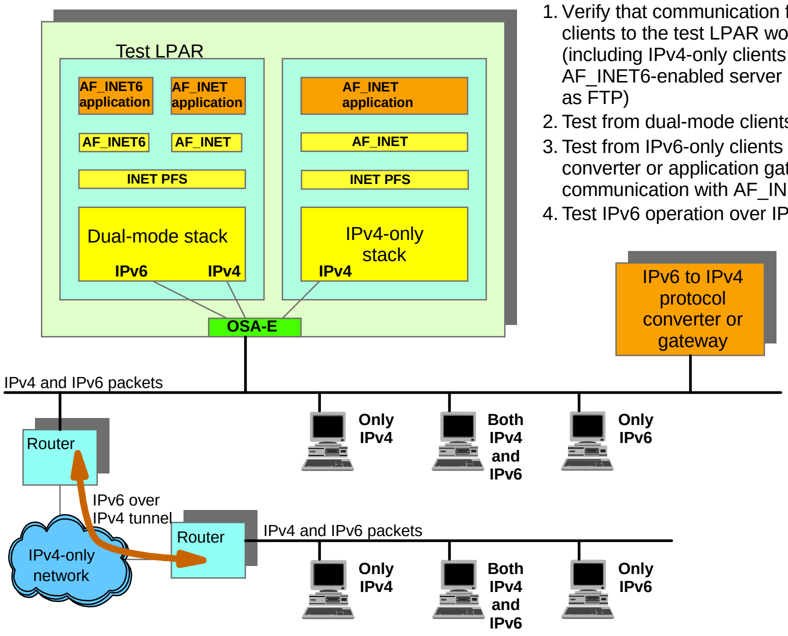
- Existing AF_INET sockets programs will continue to work as they always did - no difference in behavior or support.
- AF_INET6 enabled sockets programs will be able to communicate with IPv4 partners (just as before they were changed to support IPv6), but in addition they will also be able to communicate with IPv6 partners.

Note: When IPv6 is enabled, most netstat reports will look different because of the potential for long IPv6 addresses.

Make sure you have modified any netstat screen-scraping REXX programs you might have developed in the past!

ibm.com/redbooks

Start testing IPv6 on z/OS



1. Verify that communication from IPv4-only clients to the test LPAR works as before (including IPv4-only clients connecting to AF_INET6-enabled server applications, such as FTP)
2. Test from dual-mode clients
3. Test from IPv6-only clients using protocol converter or application gateway for communication with AF_INET applications
4. Test IPv6 operation over IPv6 tunnel

Steps for moving to an IPv6 environment

1. **Network access**
   - A LAN can carry both IPv4 and IPv6 packets over the same media
   - A single OSA-E port can be used for both IPv4 and IPv6
     - Optionally assign different VLAN IDs to IPv4 and IPv6 networks on the same OSA-E port
   - Update TCP/IP Profile to include the INTERFACE statement(s) for any IPv6 interfaces
   - LPAR to LPAR IPv6 communication can be done using shared LAN, MPCPTP6 links, or XCF for IPv6

2. **IPv6 address selection**
   - Obtain an address block from your ISP, or use one of your IPv4 addresses to create a 6to4 prefix:
     - 6to4 prefix is 2002::/16 - append a globally unique IPv4 address and a host identifier to form a globally unique IPv6 address
     - Example: chosen IPv4 address 9.1.1.1 - 6to4 prefix 2002:0901:0101::/48
   - For test purposes, site-local IPv6 addresses is sufficient, but avoid using them in production
   - IPv6 addresses can be assigned to the IPv6 Interfaces and static VIPAs
   - Addresses can be manually configured on the INTERFACE statement in the TCP/IP Profile or autoconfigured using Neighbor Discovery Stateless Autoconfiguration (VIPA addresses must be manually configured)

3. **DNS setup**
   - Any DNS BIND 9 Name Server can be used for both IPv4 and IPv6 resources
     - Host name to IPv6 address information can also be managed locally using the /etc/ipnodes file
   - Continue to use the existing host name for IPv4 connectivity to avoid possible disruption in network connectivity and IPv4-only applications on an IPv6-enabled stack
   - Create a new host name to be used for IPv6 and IPv4 connectivity
   - Optionally, a third host name which may be used only for IPv6 can be configured
   - If using stateless autoconfiguration to define IPv6 addresses, static VIPA addresses should be stored in DNS since the autoconfigured addresses will change over time and no Dynamic DNS support is available on z/OS

Steps for moving to an IPv6 environment

4. **INET or Common INET**
   - Both are supported for IPv6, but INET is simpler
   - AF_INET6 NETWORK statement must be coded in BPXPRMxx before starting IPv6-enabled stacks
     - When the AF_INET6 NETWORK statement is present in BPXPRMxx, then all CS z/OS TCP/IP stcaks that are started in that LPAR will be dual-mode stacks. If you do not define any IPv6 configuration options, such a stack will have an IPv6 loopback interface as its only IPv6 interface

5. **Selection and placement of IPv6 to IPv4 protocol converter or application gateway**
   - z/OS does not implement any functions that will allow IPv6-only nodes to communicate with z/OS-resident AF_INET applications, so an outboard protocol converter or application-layer gateway component may be needed
   - This component will only be needed if the test configuration includes IPv6-only platforms (which is not expected to become the norm for quite some years)
   - Various technologies are being made available by various vendors; SOCKS64 seems the simplest technology right now

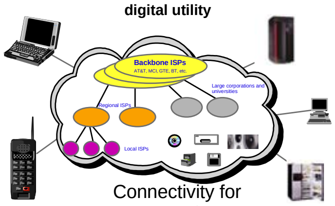
6. **Connectivity to non-local IPv6 locations**
   - Tunneling may be needed between a router connected to the LAN that z/OS is connected to, and a router at another location where IPv6 test equipment is located

The journey to IPv6 for z/OS Communications Server

IPv6 deployment phases

- 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)**
  - **Additonal Network Management MIBs**
- After z/OS V1R6
  - **Integrated IPSec**
  - **HiperSockets DLC**
  - **Advanced Socket APIs**
  - **Extended Stats MIB, OSPFv3 MIB**
  - **Intrusion Detection Services**
  - **IPv6 mobility support**

The Internet - a worldwide digital utility



Connectivity for ***anyone*** from ***anywhere*** (car, plane, home, office) to ***anything***!

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

ibm.com

z/OS Sysplex High Availability TCP/IP Solutions - Best Practices

Redbooks

International Technical Support Organization

Objectives

The objectives of this session are:

- Refresh the main design points when designing a highly available TCP/IP environment in a z/OS Sysplex
- Provide main best practices from a recent high-availability and load balancing white paper project that was run in cooperation between IBM and Cisco
- Introduce a new z/OS Sysplex Load Balancing Advisor technology that will provide WLM weights and server application health information to external load balancers

Agenda

1. TCP/IP in a Sysplex - overview
2. Network access to the z/OS Sysplex
3. Using Virtual IP Addressing for TCP/IP Application and Network Interface High Availability
4. Improved TCP/IP Sysplex autonomic behavior in z/OS V1R6
5. TCP/IP workload balancing into a Sysplex
6. Providing z/OS Sysplex WLM weights and Server Application Health Information to External Load Balancers
7. Appendix: z/OS Subsystem-specific Load Balancing Guidelines

Redbooks

ibm.com/redbooks

IBM®

z/OS Sysplex High Availability TCP/IP Solutions - Best Practices

-

TCP/IP in a Sysplex - overview

Redbooks

International Technical Support Organization

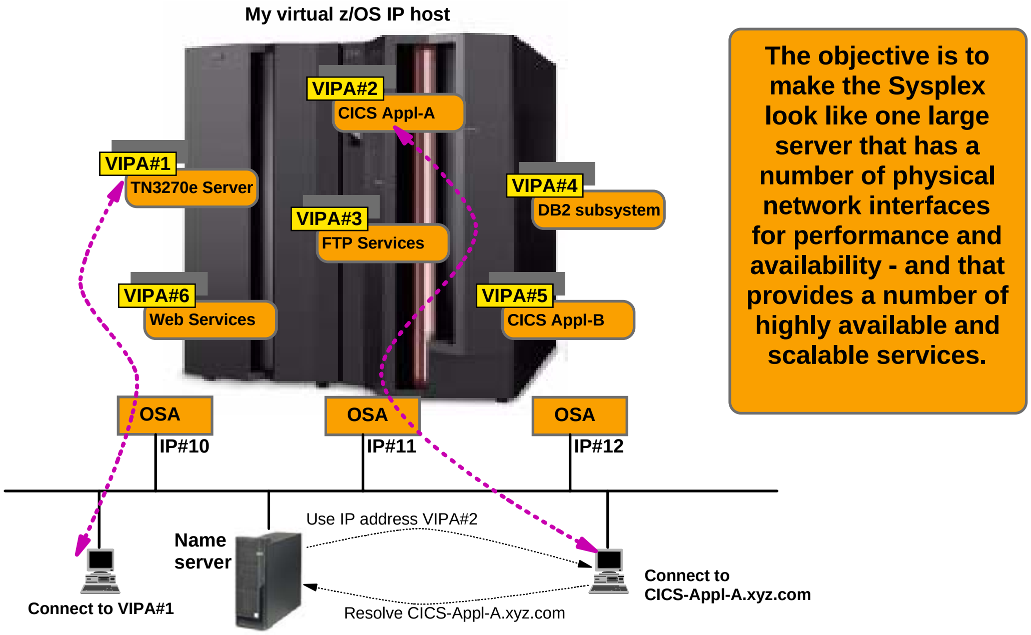
What is it we want to achieve for our IP applications in the sysplex?

- ✓ **Availability** - There must be no interruption of service for planned or unplanned outages of LPARs, network interfaces, application instances, switches, or routers. Users must perceive 100% availability of selected applications.
- ✓ **Scalability** - It must be possible to add capacity on demand for all network components, including LPARs, network interfaces, applications, switches, and routers.
- ✓ **Performance** - It must be possible to control use of network resources so critical applications and users are given premium service.
- ✓ **Load Balancing** - Network traffic must be distributed nearly equally over equal-cost network components. Application workload must be distributed among the multiple images of selected applications.
- ✓ **Single System Image** - It must be possible to reach an application based on its identity independently of which LPAR in the sysplex it executes on. Multiple instances of an application must be reachable individually and as a single entity.
- ✓ **Automated Recovery** - In case of failure, recovery plans must not rely on operator intervention, although automated operator commands can be used to aid the recovery.
- ✓ **Security** - Extend platform security to the IP environment. Focus on self-protection to reduce risc of security incidents introduced by/from the network.

Redbooks

ibm.com/redbooks

The view of a typical (large) server - many network interfaces, many services



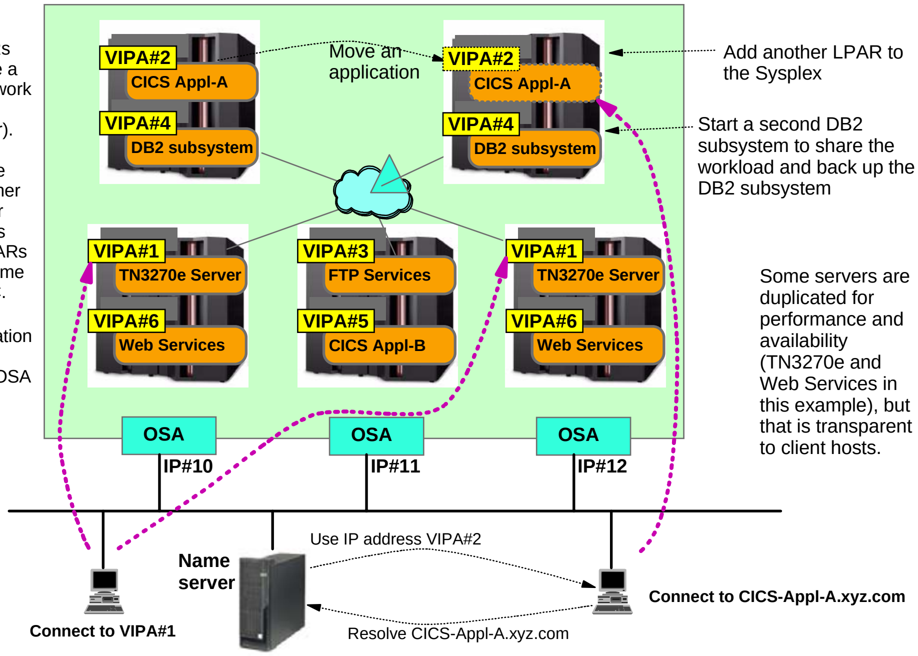
The objective is to make the Sysplex look like one large server that has a number of physical network interfaces for performance and availability - and that provides a number of highly available and scalable services.

Redbooks

ibm.com/redbooks

A typical (large) server = my Sysplex

Not all LPARs need to have a physical network interface (an OSA adapter). LPARs can communicate with each other using XCF or HiperSockets between LPARs inside the same zSeries CEC. The general recommendation is to have all LPARs use OSA adapters for performance reasons.



Add another LPAR to the Sysplex

Start a second DB2 subsystem to share the workload and back up the first DB2 subsystem

Some servers are duplicated for performance and availability (TN3270e and Web Services in this example), but that is transparent to client hosts.

Redbooks

ibm.com/redbooks

IBM®

z/OS Sysplex High Availability TCP/IP Solutions - Best Practices

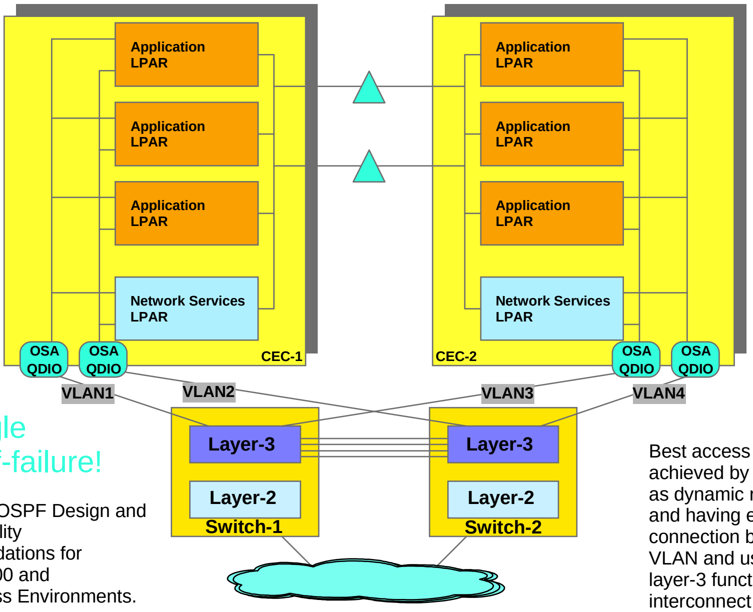
-

Network Access to the z/OS Sysplex

Redbooks

International Technical Support Organization

Downstream network connectivity to the Sysplex: OSA-Express with QDIO



No single point-of-failure!

Reference: OSPF Design and Interoperability Recommendations for Catalyst 6500 and OSA-Express Environments.

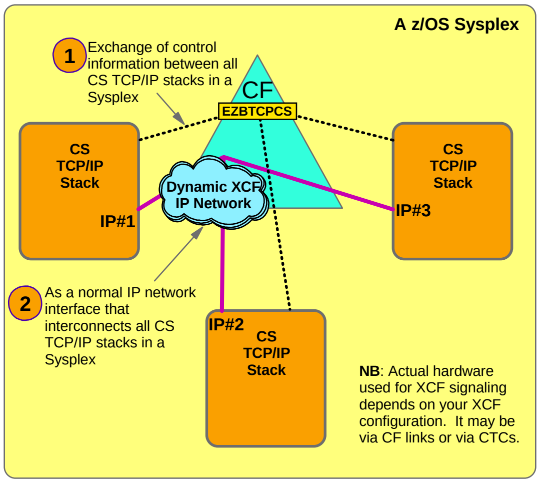
http://www-1.ibm.com/servers/eserver/zseries/networking/pdf/ospf_design.pdf

Best access availability is achieved by using OSPF as dynamic routing protocol and having each switch connection be its own VLAN and use the switch's layer-3 functions to interconnect them.

Redbooks

ibm.com/redbooks

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



A z/OS Sysplex
1
Exchange of control information between all CS TCP/IP stacks in a Sysplex
CF
EZBTCPCS
CS TCP/IP Stack
IP#1
Dynamic XCF IP Network
CS TCP/IP Stack
IP#3
2
As a normal IP network interface that interconnects all CS TCP/IP stacks in a Sysplex
IP#2
CS TCP/IP Stack
NB: Actual hardware used for XCF signaling depends on your XCF configuration. It may be via CF links or via CTCs.

XCF signaling is used for two purposes:

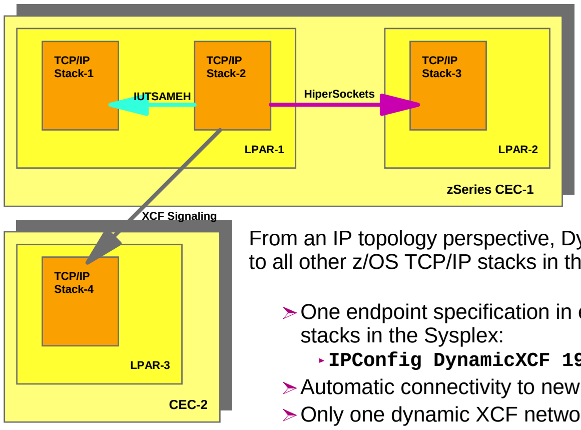
1. When a CS TCP/IP stack starts in a Sysplex, it always joins a predefined XCF group. This group is used by all CS TCP/IP stacks in the same Sysplex to exchange control information, such as which IP addresses each stack has in its home list and event notification when an IP address is added or deleted. 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.

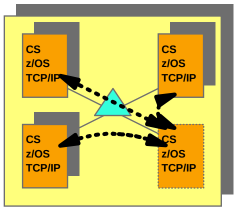
2. XCF can optionally also be used as an IP network interface over which CS TCP/IP stacks can send IP packets to each other. This use is under configuration control and can be defined using either static XCF links or allowing all stacks to join an IP XCF network dynamically (DYNAMICXCF). If one uses Sysplex Distributor or Non-disruptive Dynamic VIPA movement functions in a Sysplex, then dynamic XCF must be enabled.

Redbooks

ibm.com/redbooks

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.

- One endpoint specification in each stack for fully meshed connecitivity 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

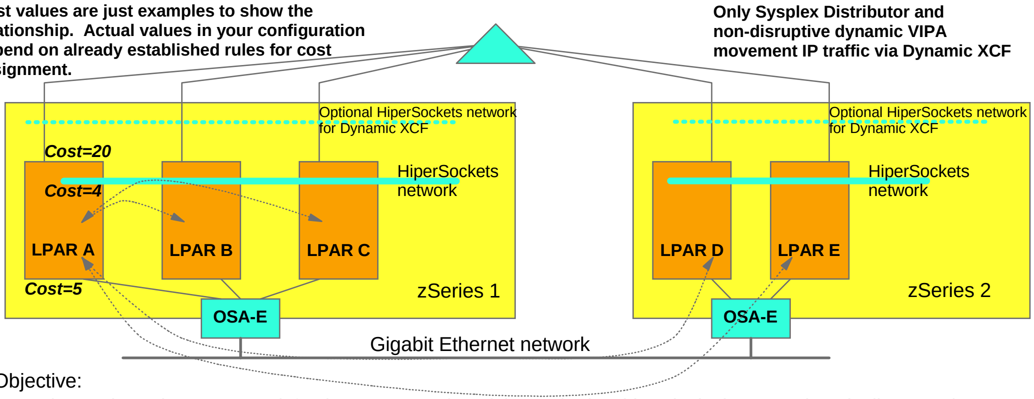
Redbooks

ibm.com/redbooks

Guidelines for controlling use of the DynamicXCF IP network for general IP routing

Cost values are just examples to show the relationship. Actual values in your configuration depend on already established rules for cost assignment.

Only Sysplex Distributor and non-disruptive dynamic VIPA movement IP traffic via Dynamic XCF



- Objective:
  - Only use dynamic XCF network for the purposes where it is required at this point in time: 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.
- Define in each CEC a second HiperSockets network (through DEVICE/LINK definitions that interconnect 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)

Redbooks

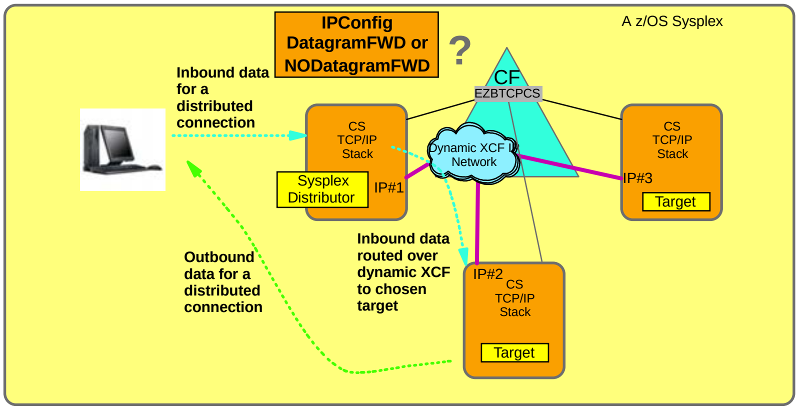
ibm.com/redbooks

General IP forwarding no longer required for Sysplex Distributor in z/OS V1R6

The distributing TCP/IP stack needs to forward both connection setup and inbound connection data over a dynamic XCF IP network to the chosen TCP/IP target stack in the sysplex

- Previous to z/OS V1R6 it was a requirement that the distributing stack had to have DATAGRAMFWD enabled
  - This option means that the TCP/IP stack is allowed to route IP packets in general from any interface to any interface (only way to limit this general routing capability was via firewall filters on z/OS)

- In z/OS V1R6, use of Sysplex Distributor does not require DATAGRAMFWD to be enabled
  - Sysplex Distributor can now be deployed without any risk of using a z/OS stack as a general intermediate routing node



V1R6

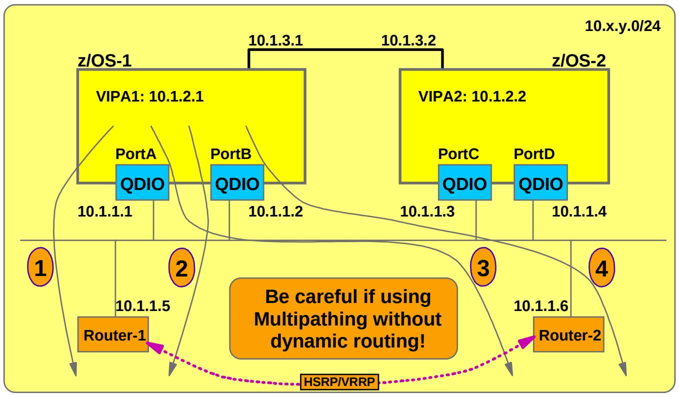
Load-balancing outbound IP packets over multiple first-hop routers (MULTIPATH)

IPCONFIG MultiPath
[PerConnection or PerPacket]

z/OS-1's IP Routing Table (extract)

Destination	Via
10.1.1.0/24	Direct delivery
Default	10.1.1.5 / PortA
Default	10.1.1.5 / PortB
Default	10.1.1.6 / Port A
Default	10.1.1.6 / Port B

z/OS V1R5 raises number of dynamic multipath routes from 4 to 16.



Static route definitions on z/OS:

- If an adapter fails in such a way that z/OS TCP/IP gets informed, it will skip over the corresponding entries from the routing table
- If one of the first-hop routers loses its connection to the backbone network or if it "dies" - z/OS TCP/IP doesn't know anything about it since it doesn't participate in dynamic routing updates - and it will continue to attempt to use the corresponding routing table entries - connections will time out, UDP packets will be lost, etc.
- If the two routers deploy VRRP or HSRP between them on the interfaces towards the z/OS systems, then the fact that one of them turns into a black hole can be hidden from z/OS - z/OS continues to send packets to both first-hop addresses, they are just both serviced by the one surviving router

Dynamic routing updates:

- z/OS TCP/IP will know both if the adapter itself fails or if the first-hop router fails - and will dynamically update the routing table entries and recover from the router outage..

IBM®

z/OS Sysplex High Availability TCP/IP Solutions - Best Practices

-

Using Virtual IP Addressing for TCP/IP Application and Network Interface High Availability

Redbooks

International Technical Support Organization

Why do I need virtual IP addresses (VIPA)?

What does the virtual IP addressing (VIPA) technology promise?

Interface resilience:

- Communication with a server host is unaffected by server physical network interface failures. As long as even a single physical network interface is available and operational on a server host, communication with applications on the server host will persist.

Application access independent of network topology:

- Separates network topology from server application topology - a VIPA address can be used to identify a server application instead of a physical network interface.
- Allows network administrators to renumber physical network topology
  - No impact to end-user accessing server applications by IP address
  - No changes needed in DNS or hosts file configuration
  - No impact on firewall filtering rules

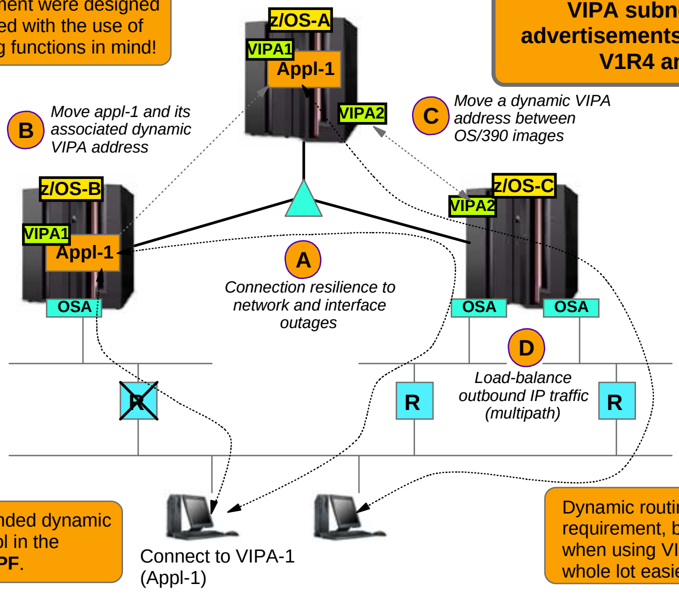
Single system image:

- Allows the Sysplex to be perceived as a single large server node, where VIPA addresses identify applications independent of which images in the Sysplex the server applications execute on.
- Applications retain their identity when moved between images in a Sysplex.
- Multiple instances of a server application can be accessed as one server.

Are dynamic routing protocols required on z/OS in order to use VIPA?

Base IP recovery as well as VIPA address movement were designed and implemented with the use of dynamic routing functions in mind!

APAR PQ82792 will prevent OMPROUTE from including the VIPA subnets in OSPF advertisements! (PTFs for z/OS V1R4 and V1R5)


z/OS-A
VIPA1
Appl-1
VIPA2
B
Move appl-1 and its associated dynamic VIPA address
C
Move a dynamic VIPA address between OS/390 images
z/OS-B
VIPA1
Appl-1
OSA
A
Connection resilience to network and interface outages
z/OS-C
VIPA2
OSA
OSA
D
Load-balance outbound IP traffic (multipath)
R
R
R
Connect to VIPA-1 (Appl-1)

Always remember that a z/OS Sysplex is not a host; it is an IP network in itself and as any IP network it needs the capability to react to topology changes in its own network and in the adjacent networks.

The recommended dynamic routing protocol in the Sysplex is **OSPF**.

Dynamic routing is not an absolute requirement, but it is ***highly recommended*** when using VIPA addresses (it makes life a whole lot easier)!

Redbooks

Dynamic VIPA usage - overview

A dynamic VIPA address has all the attributes of a static VIPA address. In addition, it has the ability to move between TCP/IP stacks in a Sysplex based on certain events - without operator intervention in terms of configuration changes.

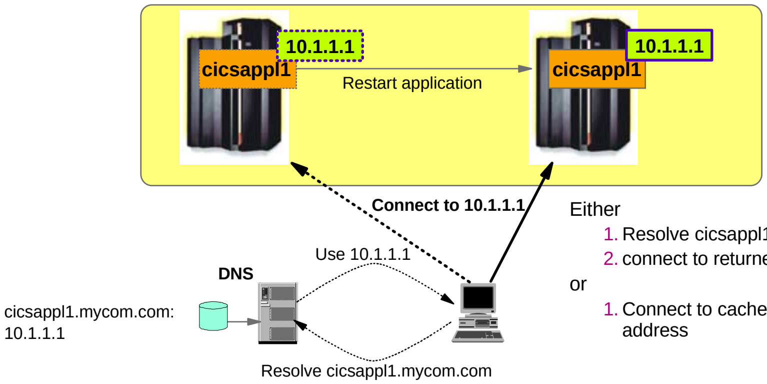
When do you want the dynamic VIPA to move?	What's the type of DVIPA to use?	How do you define it?	Application requirements	Typical use
Move to a backup stack, when the currently owning stack goes down or is taken down.	**A stack- managed DVIPA**	VIPADEFINE on primary owner - VIPABACKUP on potential backup stacks.	Applications bind to INADDR_ANY.	Multiple instances of server run on multiple stacks and can back up each other.
Move along with a specific server application that binds its listening socket to the dynamic VIPA address.	**An application-specific DVIPA**	VIPARANGE	Applications must bind to the specific dynamic VIPA address (alternatively use BIND specific on port reservation)	Single instance application that is moved between stacks - planned or unplanned.
Move when instructed to do so by executing a utility (moddvipa) or by an authorized application (using an ioctl call)	**A command-activated DVIPA**	VIPARANGE	No special requirements, but typically application binds to INADDR_ANY.	Single instance applications that cannot be controlled via bind specific functions.

Redbooks

Basic principles for recovery of single-instance IP application in a Sysplex

- Single-instance applications are applications that run in only one instance in the Sysplex, either because the application needs exclusive access to certain resources, or because there is no need to start it in more than one instance.

- Availability from an IP perspective then becomes an issue of being able to restart the application on the same LPAR or on another LPAR with as little impact to end users as possible.
  - Speed of movement - ARM or automated operations procedures
  - Retain identity from a network perspective (its IP address) - application-specific DVIPAs



Application-specific dynamic VIPA addresses come in very handy for this purpose.

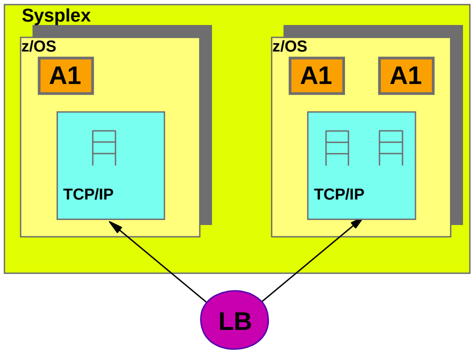
Either

1. Resolve cicsappl1.mycom.com
2. connect to returned address

or

1. Connect to cached (or hardcoded!) address

Workload balancing: a question of performance, availability, and scalability - multi-instance applications (data sharing)



Connection load balancing technologies:

Between z/OS images:

a. DNS
b. NAT - CSM, CSS, many others
c. Dispatchers - ND, CSM, Sysplex Distributor
d. Contents-based - CSM, CSS, BigIP, etc.

Inside single z/OS TCP/IP stack:

a. Port sharing

Application characteristics:

- Multiple instances of the server are able to provide the exact same services to clients (will typically require data sharing)
- No state preserved at server between two connections (application protocol has to include support for such behavior or store state data in shared storage)

Benefits of intelligent load balancing:

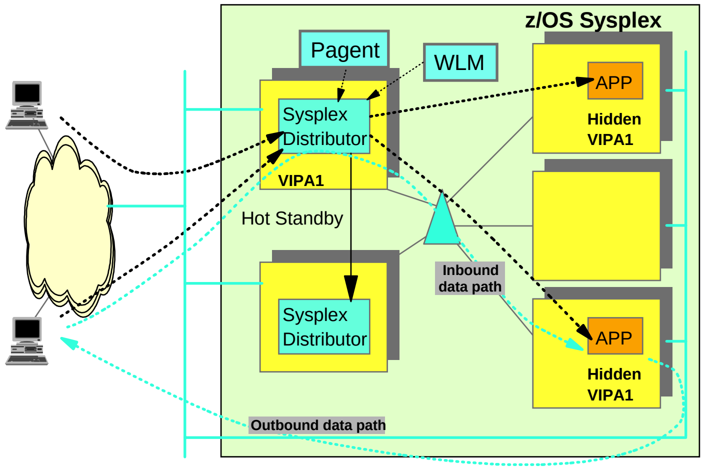
- ***Performance*** - improving response time
- ***Availability*** - If one instance goes down, connections with it break, but new connections can be established with remaining instance(s)
- ***Scalability*** - more server instances can be added on demand (horizontal growth)

Examples:

- Web server
- TN3270 server
- Some CICS applications
- FTP server
- DB2
- MQ
- WAS
- LDAP
- RYO

Sysplex Distributor: z/OS-integrated intra-Sysplex workload balancing - overview

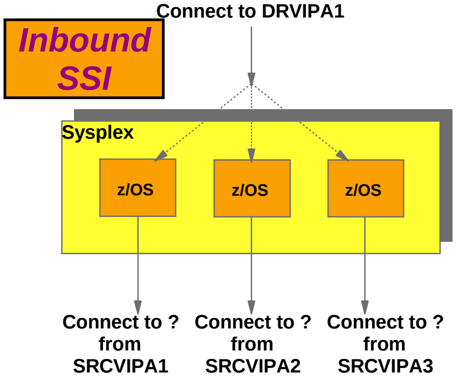
- Independent of network attachment technology. Will work with both direct (including OSA Express) and channel-attached router network connections.
- All z/OS images communicate via XCF. Each TCP/IP stack has full knowledge of IP addresses and server availability in all stacks.
- A network-connected stack owns a given VIPA address and acts as the distributor of new connection requests to that VIPA address.
- In z/OS V1R5, Sysplex Distributor supports timer-based source IP address affinity, and in addition to WLM-based balancing it also now supports a round-robin based balancing among available servers.



- Distribution of new connection requests is based on real-time consultation with WLM (if available, else round-robin), z/OS QoS policy agent and target stack for application availability.
- Connection information is stored in Sysplex distributor stack for routing of IP packets belonging to an existing connection to the appropriate target stack. Routing is based on the connection to which IP packets belong (connection-based routing).
- The hot standby stack is free to do other work, but will take over ownership of the VIPA address and the distributor function in case the primary stack fails. This takeover is non-disruptive to existing connections.

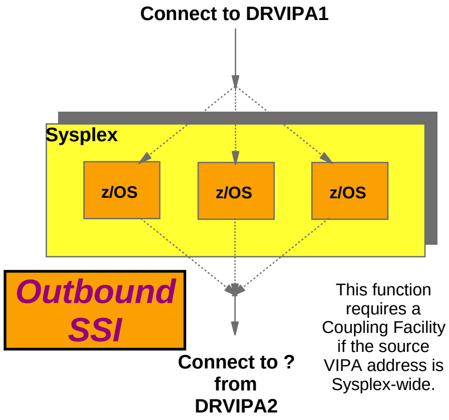
ibm.com/redbooks

Single system image (SSI) from an IP perspective in the Sysplex


Inbound SSI
Connect to DRVIPA1
Sysplex
z/OS
z/OS
z/OS
Connect to ? from SRCVIPA1
Connect to ? from SRCVIPA2
Connect to ? from SRCVIPA3

- We have single system image capability for inbound connections where a single distributed VIPA address can represent all images in the Sysplex - and remote users do not need to select a specific image when connecting to their server application.
- But if we establish outbound connections from the images in the Sysplex, each image has its own source VIPA address - so there is no single system image from an outbound connection perspective - which has implications in firewall filter setup, etc.

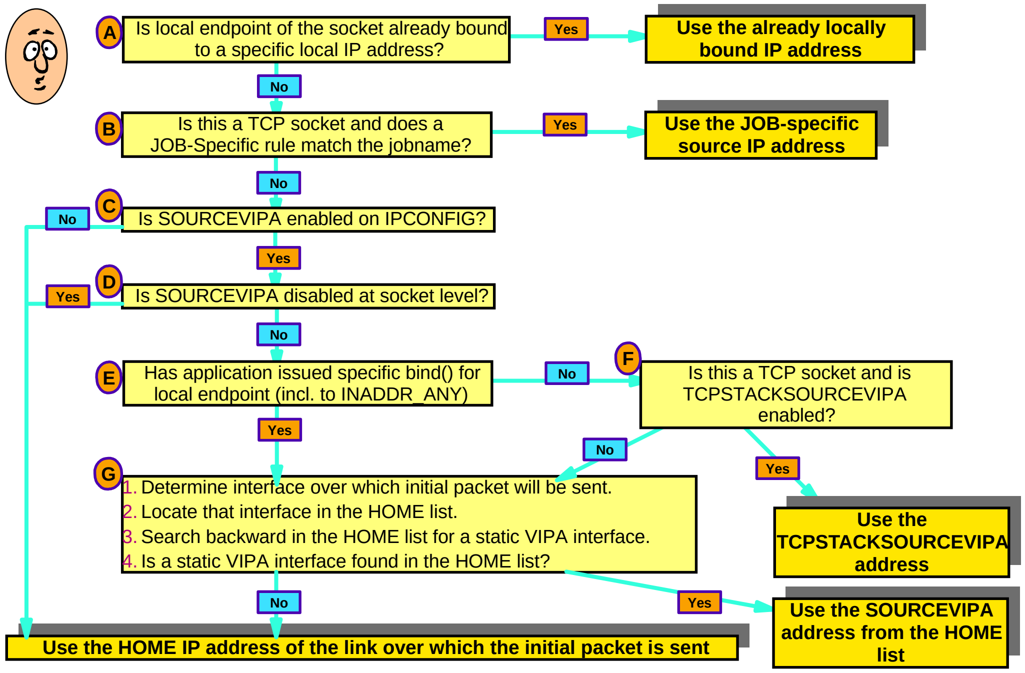
- z/OS V1R4 introduced new SysplexPorts capabilities that allow a single Sysplex-wide source VIPA address to be used for outbound TCP connections by all images in the Sysplex - resulting in single system image capabilities for both inbound and outbound connections.
- z/OS V1R5 significantly improved the performance of SysplexPorts by allocating ephemeral ports in the CF in blocks of 64.



This function requires a Coupling Facility if the source VIPA address is Sysplex-wide.

ibm.com/redbooks

Selecting source IP address for IPv4 in CS z/OS V1R6



A Is local endpoint of the socket already bound to a specific local IP address?

Yes → Use the already locally bound IP address

No

B Is this a TCP socket and does a JOB-Specific rule match the jobname?

Yes → Use the JOB-specific source IP address

No

C Is SOURCEVIPA enabled on IPCONFIG?

No

Yes

D Is SOURCEVIPA disabled at socket level?

Yes

No

E Has application issued specific bind() for local endpoint (incl. to INADDR_ANY)

No

F Is this a TCP socket and is TCPSTACKSOURCEVIPA enabled?

Yes → Use the TCPSTACKSOURCEVIPA address

No

Yes

G
1. Determine interface over which initial packet will be sent.
2. Locate that interface in the HOME list.
3. Search backward in the HOME list for a static VIPA interface.
4. Is a static VIPA interface found in the HOME list?

Yes → Use the SOURCEVIPA address from the HOME list

No

Use the HOME IP address of the link over which the initial packet is sent

V1R6

IBM®

z/OS Sysplex High Availability TCP/IP Solutions - Best Practices

-

Improved TCP/IP Sysplex Autonomic Behavior in z/OS V1R6

Redbooks

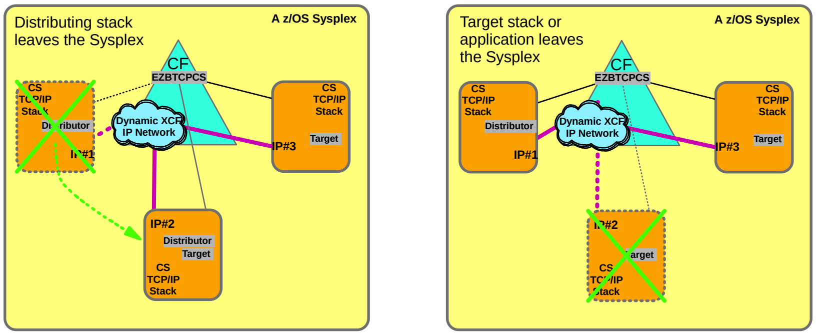
International Technical Support Organization

Sysplex autonomics - overview

- **TCP/IP sysplex technologies were originally introduced in OS/390 V2R8**
  - Functional enhancements have been shipped in all OS/390 and z/OS releases since then
  - Customers use the functions as the foundation for building highly available and scalable TCP/IP solution environments in their z/OS Sysplexes
  - Customers have adapted the use of these functions rapidly and they are today in widespread use for business-critical workloads in installations of all sizes - ranging from small to very large
- **As these technologies themselves mature and use of them increases, it is time to take a critical look at the operational aspects and identify areas where proactive improvements can be made**
  - A large experience base to harvest ideas from for operational improvements
    - Reduce overall system impact of malfunctioning dependent components such as XCF signaling, or dynamic routing daemon (OMPROUTE) availability, etc.
    - Improve operator interaction so operators can control the system without modifying the active configuration
  - Apply autonomic principles to let the technology itself detect and react to a range of error conditions - without operator intervention
- **Sysplex Distributor as a z/OS Sysplex resident load balancer, provides the opportunity to include very detailed, real-time metrics in its decision logic**
  - As the z/OS infrastructure improves in the areas of server-specific performance reporting, apply those improved metrics to the TCP/IP load balancing decision algorithms
  - Include connection processing time in the decision algorithms

TCP/IP sysplex recovery functions - before z/OS V1R6

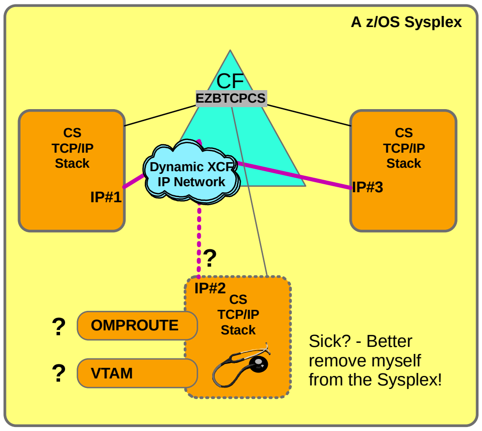
- **TCP/IP sysplex recovery functions to protect against major hardware and software failures are triggered when a TCP/IP stack leaves the TCP/IP XCF group (terminates)**
  - If the leaving stack was a DVIPA owner, a backup stack will take over the DVIPA along with any associated Sysplex Distributor responsibilities - and new workload will continue to be processed by the Sysplex
  - If the leaving stack was a target stack for distributed workload, the distributing stack will remove it from its list of candidate target stacks - stop sending more connections to it
- **If a TCP/IP stack doesn't terminate, but enters an "unresponsive" condition, recovery functions are not triggered**
  - If the unresponsive stack is a Sysplex Distributor stack, no new connections to the distributed application will be processed and routing of inbound data through the distributing stack to target stacks for existing connections will cease
  - If the unresponsive stack is a target stack, the distributing stack will continue to send new connections to it and since WLM may see the target stack as lightly loaded, that stack may even be seen as a preferred stack for new workload - sending even more workload down the drain



What can cause an unresponsive condition of TCP/IP?

- There are a few known error-conditions that can cause TCP/IP to become unresponsive or appear to be hanging - without actually terminating:
  - The downstream network lost visibility of the distributing stack due to an OMPROUTE outage or malfunction and the network routers do not know how to reach the destination DVIPA addresses
  - VTAM is malfunctioning, data link control services are not working properly, and IP packets cannot be received or sent
  - TCP/IP is in a critical storage constraint situation
  - XCF IP network connectivity (DynamicXCF) between the distributing stack and the target stacks is not functioning
  - Abends/errors in the TCP/IP sysplex code components

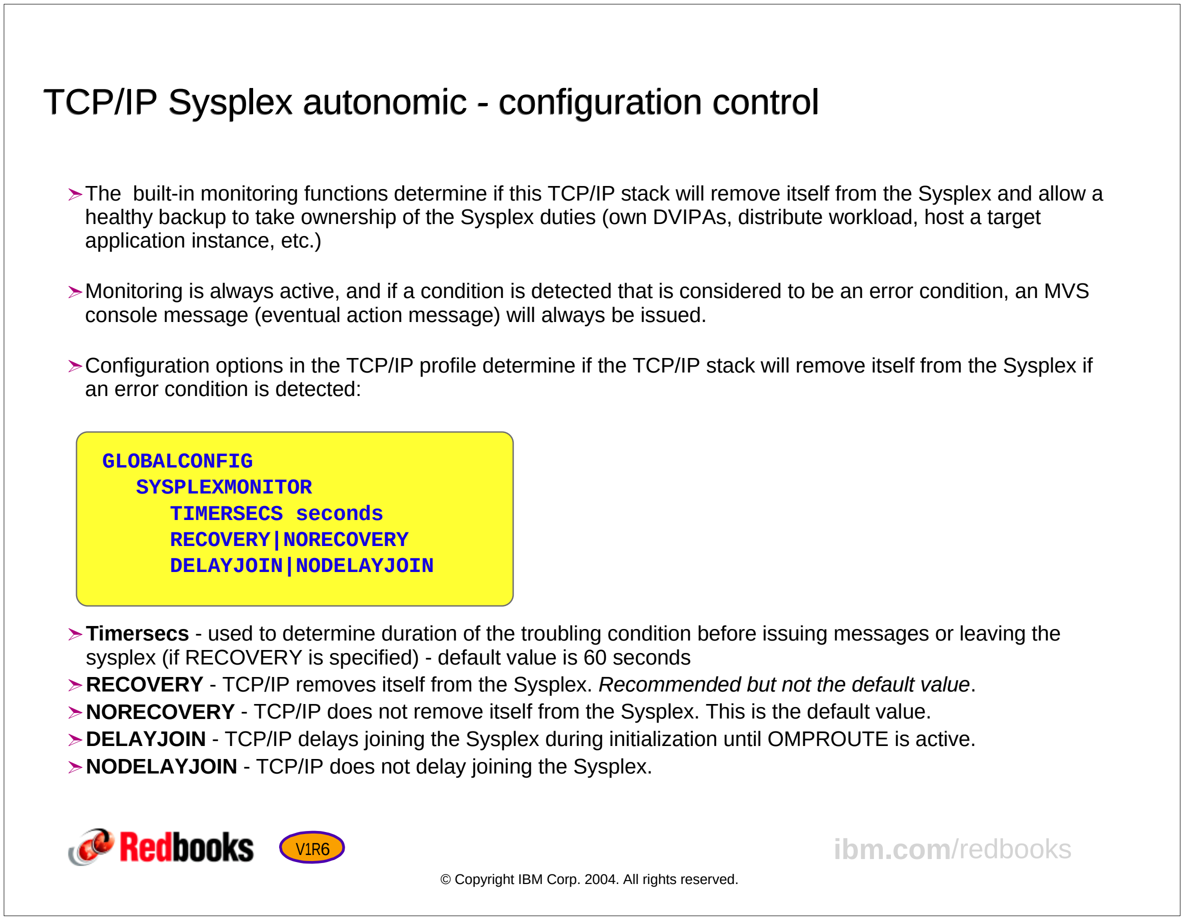
TCP/IP sysplex autonomic problem determination and recovery in z/OS V1R6 - overview



Sick? - Better remove myself from the Sysplex!

The base 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.

- Add autonomic functions to monitor system health:
  - Monitor storage usage (if critical for a *period of time*, initiate recovery action)
    - CSM, TCP/IP private and ECSA
  - Monitor dependent networking functions (if unavailable for a *period of time*, initiate recovery action)
    - OMPROUTE availability and health
    - VTAM availability
    - Dynamic XCF link availability
  - Monitor for abends in critical Sysplex-related functions
  - Monitor that specific internal functions are executed in a timely fashion (if unresponsive for a *period of time*, initiate recovery action)
    - Using independent XCF status monitor component
- Optionally, delay joining a Sysplex during initialization until OMPROUTE is active (already done for VTAM)
- Provide an MVS operator command to request a TCP/IP stack to leave the Sysplex



TCP/IP Sysplex autonomic - configuration control

- The built-in monitoring functions determine if this TCP/IP stack will remove itself from the Sysplex and allow a healthy backup to take ownership of the Sysplex duties (own DVIPAs, distribute workload, host a target application instance, etc.)

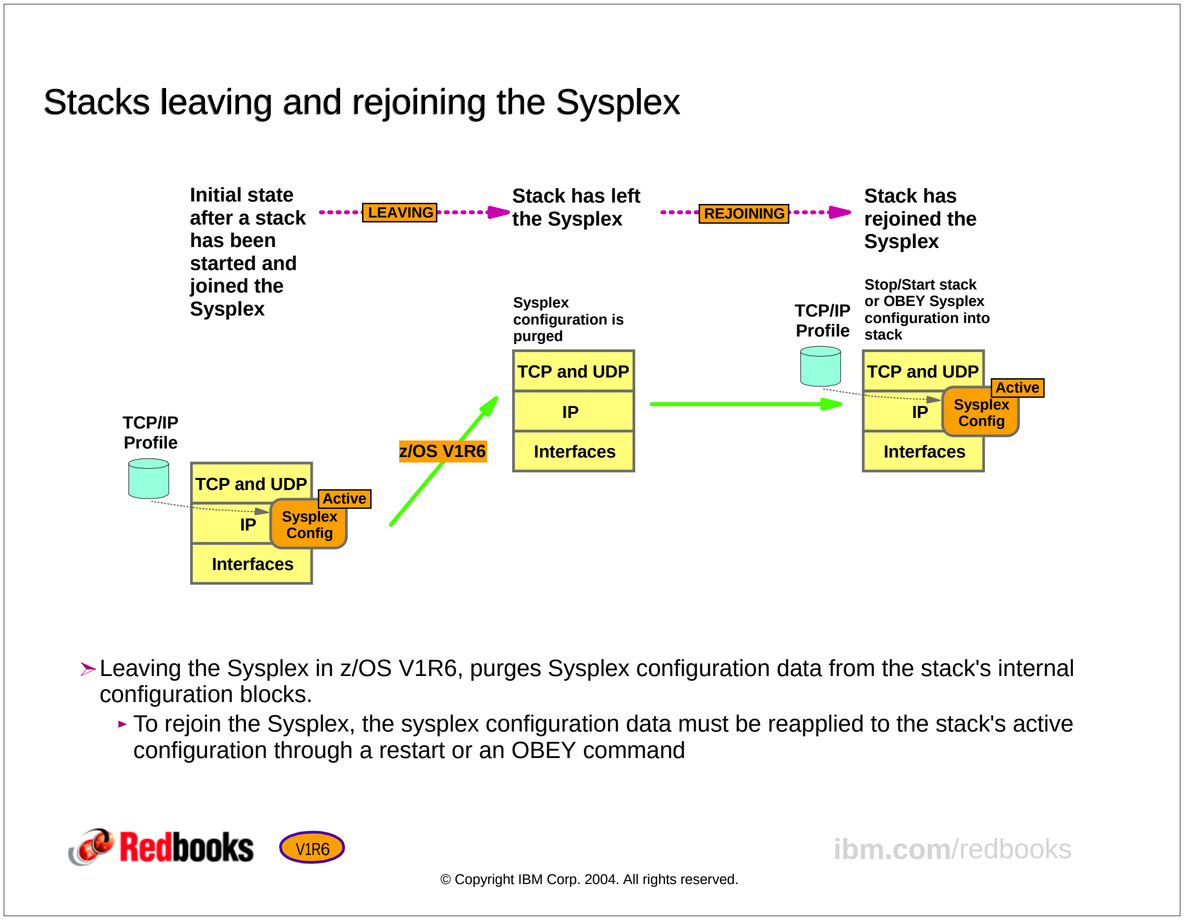
- Monitoring is always active, and if a condition is detected that is considered to be an error condition, an MVS console message (eventual action message) will always be issued.

- Configuration options in the TCP/IP profile determine if the TCP/IP stack will remove itself from the Sysplex if an error condition is detected:

```
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 is specified) - default value is 60 seconds
- **RECOVERY** - TCP/IP removes itself from the Sysplex. *Recommended but not the default value*.
- **NORECOVERY** - TCP/IP does not remove itself from the Sysplex. This is the default value.
- **DELAYJOIN** - TCP/IP delays joining the Sysplex during initialization until OMPROUTE is active.
- **NODELAYJOIN** - TCP/IP does not delay joining the Sysplex.

Redbooks V1R6 ibm.com/redbooks



Stacks leaving and rejoining the Sysplex

Initial state after a stack has been started and joined the Sysplex

LEAVING

Stack has left the Sysplex

REJOINING

Stack has rejoined the Sysplex

TCP/IP Profile

TCP and UDP

IP

Active

Sysplex Config

Interfaces

z/OS V1R6

Sysplex configuration is purged

TCP and UDP

IP

Interfaces

TCP/IP Profile

Stop/Start stack or OBEY Sysplex configuration into stack

TCP and UDP

IP

Active

Sysplex Config

Interfaces

- 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

Redbooks V1R6 ibm.com/redbooks

Leaving and rejoining the Sysplex

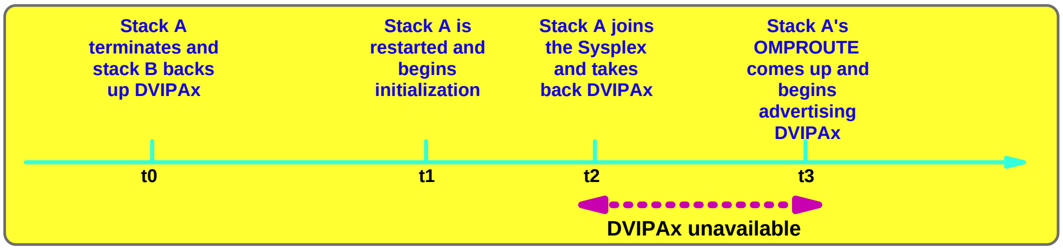
NOTES

- **When a TCP/IP stack leaves the Sysplex (disconnects from the EZBTCPCS XCF group), it will in z/OS V1R6 remove all its internal Sysplex related configuration data:**
  - The content of the VIPADYNAMIC block
- **DynamicXCF definitions will not be removed and DynamicXCF IP connectivity may in fact remain active even after a stack has left the Sysplex**
- **To rejoin the Sysplex, one of two methods can be used:**
  - Stop the stack and start it again
    - During normal initialization, all Sysplex related configuration options in the TCP/IP profile will be processed and normal sysplex initialization will join the stack into the Sysplex
  - Vary OBEY a profile that includes a DynamicXCF statement and/or a VIPADYNAMIC block
    - If the error condition was temporary and only affected Sysplex-related functions, while local functions continued to operate normally, a vary OBEY can be used after the Sysplex-related error condition was cleared to have the stack rejoin the Sysplex
    - If a stack does not have a VIPADYNAMIC block (could be the case for a target-only stack), a vary OBEY of an IPConfig statement with the DynamicXCF definitions will make the stack rejoin the Sysplex
    - When the stack rejoins the Sysplex, it will initiate takeback of any DVIPAs for which it is the primary owner (VIPADEFINEd)

Redbooks V1R6 ibm.com/redbooks

Delay joining the Sysplex until the TCP/IP environment is fully ready to do real work

- **Delay joining the Sysplex during stack initialization:**
  - The case that is being addressed by this function is where a primary stack is restarted and attempts to take back the dynamic VIPA addresses for which it is the primary owner.
  - If it attempts to do so before its OMPROUTE is up and active, a window may occur where the dynamic VIPA address have been taken back from the backup stack, but hasn't yet been advertised by the restarted stack - resulting in a time window where that address isn't available
    - That window would normally be very short (a few seconds), but could be longer depending on local operations procedures
  - The base idea is that it is better to leave the address with the backup stack until the restarting stack is fully ready to take over its responsibilities



Stack A terminates and stack B backs up DVIPAx

Stack A is restarted and begins initialization

Stack A joins the Sysplex and takes back DVIPAx

Stack A's OMPROUTE comes up and begins advertising DVIPAx

DVIPAx unavailable

Redbooks V1R6 ibm.com/redbooks

Operator command to request a TCP/IP stack to leave the Sysplex

- **Operator command to leave the Sysplex**
  - Operator command must be issued on the system where the stack that is to leave the Sysplex is running
  - Allows an operator initiated recovery of an error condition
    - Monitoring issues messages, but NORECOVERY was configured
    - Recovery from other unforeseen error condition
  - For the stack to rejoin the Sysplex, a restart of the stack or an OBEYFILE command is required to redefine all the Sysplex-related resources (DVIPAs, Distributor rules, etc.)

```
VARY TCPIP,[stackname],SYSPLEX,LEAVEGROUP
.....
EZZ0053I COMMAND SYSPLEX,LEAVEGROUP COMPLETED SUCCESSFULLY
```

Redbooks V1R6

ibm.com/redbooks

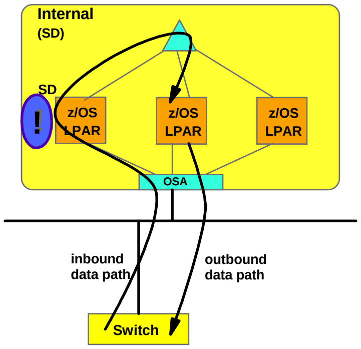
IBM®

z/OS Sysplex High Availability TCP/IP Solutions - Best Practices
-
TCP/IP Workload Balancing in a Sysplex

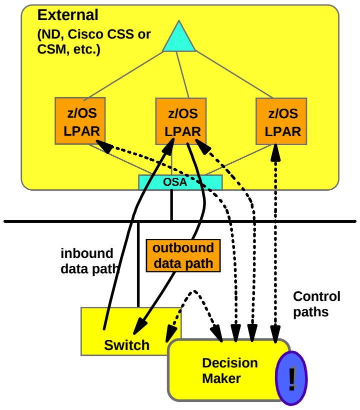
Redbooks

International Technical Support Organization

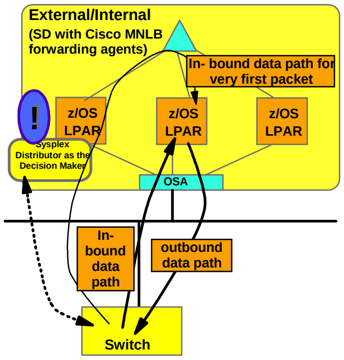
Sysplex internal vs. external workload balancing - which technology is best for me?



- Has realtime information available...
  - more timely capacity information
  - QoS from Service Policy Agent
  - application-independent server availability
- No problems with shared OSA adapters; no intermediate routers
- Uses expensive Sysplex resources for inbound traffic
  - Inbound traffic funneled through single point
  - Routing stack uses zSeries MIPs for inbound routing
  - Inbound traffic routed over XCF



- Specialized routing hardware may be more cost-effective
- May be configured for no single point of traffic flow
- So far no general Sysplex-capacity feedback technology (some specialized attempts have been made)
- Requires application-specific health probes ("application ping")
- Problems with shared OSA adapters or other intermediate routers


External/Internal
(SD with Cisco MNLB forwarding agents)
In- bound data path for very first packet
!
z/OS LPAR
z/OS LPAR
z/OS LPAR
Sysplex Distributor as the Decision Maker
OSA
In- bound data path
outbound data path
Switch

- Combines the advantages of internal decision making with routing efficiency of external control point
- Shared OSA adapters require use of Generic Routing Encapsulation by the switch.
- Preferred technology especially when workload includes large amounts of inbound data flows.

! Denotes decision point

Redbooks

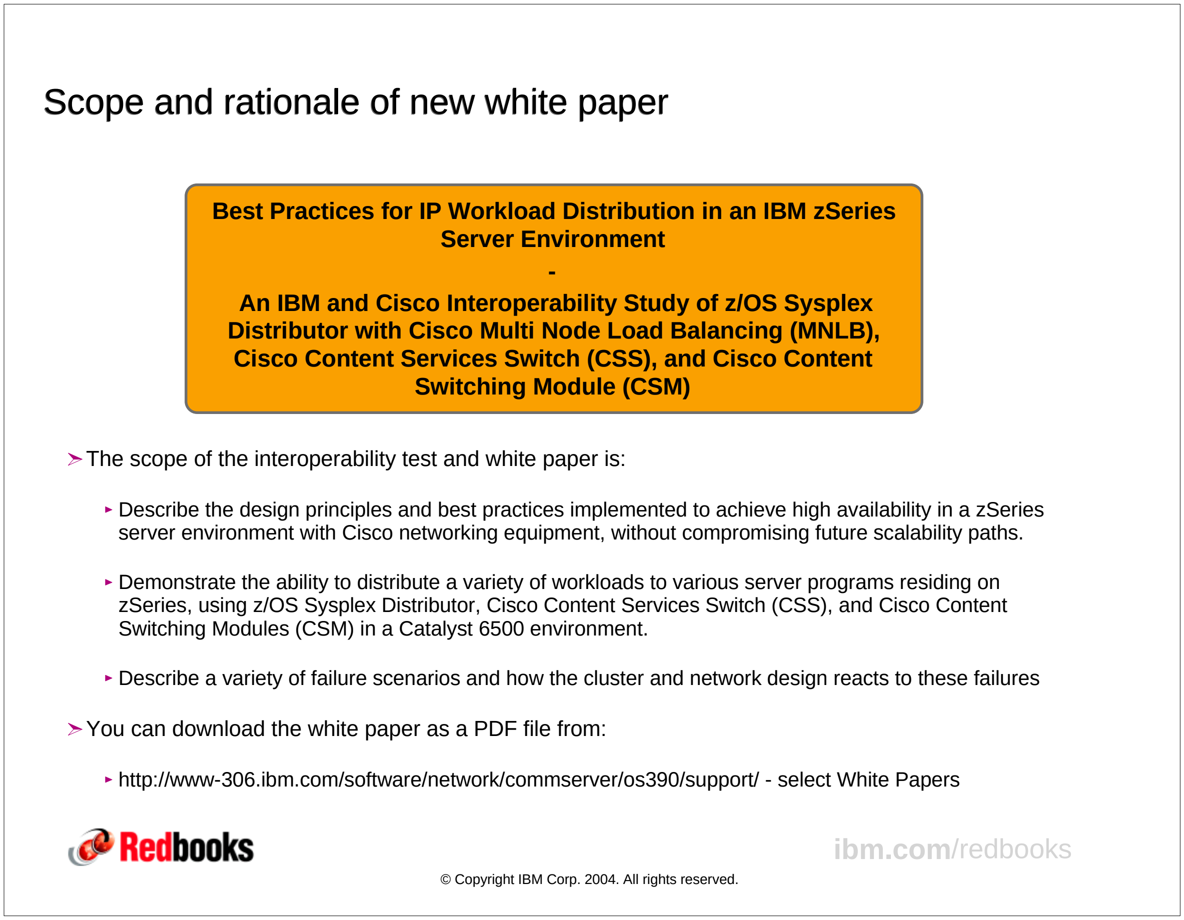
ibm.com/redbooks

Function comparison load balancing - DNS, SD, outboard LBs

	DNS	Sysplex Distributor	Outboard Load Balancers
When is server instance decision made?	Name resolution	Connection setup (in-line SYN segment)	Connection setup (in-line SYN segment / HTTP GET request arrival)
TCP connections or UDP associations	Both TCP and UDP (at time of name resolution) [UDP: EE as an example]	TCP only	Both TCP and UDP (at time of packet arrival)
Real-time application availability information available	DNS/WLM: Yes otherwise: No (at time of resolution)	Yes	Based on a polling interval. (z/OS LBA: feedback)
LPAR WLM information available	DNS/WLM: Yes otherwise: No (at time of resolution)	Yes	No (z/OS LBA: yes)
Network QoS performance available	No	Yes	No
Extra network flows	Each new contact preceded by DNS flows	Inbound via distributing stack. Virtually none if combined with Cisco MNLB forwarding agents	No (some minimal extra paths inside the switch environment)
Potential issues	DNS and resolver caches. Clients hardcoding IP addresses.	TCP only - no UDP support	No real time Sysplex information available (z/OS LBA: will be). Need to understand NATing implications.

Redbooks

ibm.com/redbooks



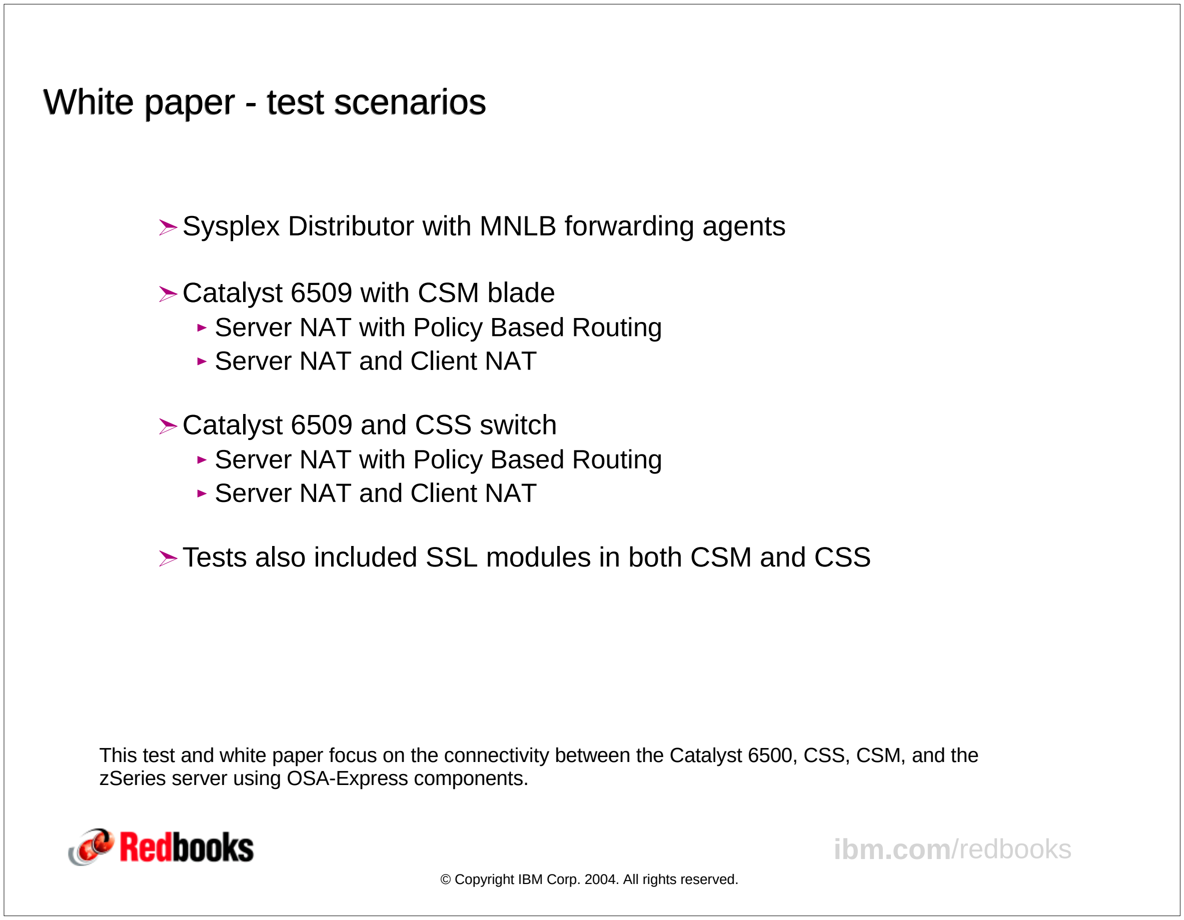
Scope and rationale of new white paper

Best Practices for IP Workload Distribution in an IBM zSeries Server Environment

-

An IBM and Cisco Interoperability Study of z/OS Sysplex Distributor with Cisco Multi Node Load Balancing (MNLB), Cisco Content Services Switch (CSS), and Cisco Content Switching Module (CSM)

- The scope of the interoperability test and white paper is:
  - Describe the design principles and best practices implemented to achieve high availability in a zSeries server environment with Cisco networking equipment, without compromising future scalability paths.
  - Demonstrate the ability to distribute a variety of workloads to various server programs residing on zSeries, using z/OS Sysplex Distributor, Cisco Content Services Switch (CSS), and Cisco Content Switching Modules (CSM) in a Catalyst 6500 environment.
  - Describe a variety of failure scenarios and how the cluster and network design reacts to these failures
- You can download the white paper as a PDF file from:
  - http://www-306.ibm.com/software/network/commserver/os390/support/ - select White Papers

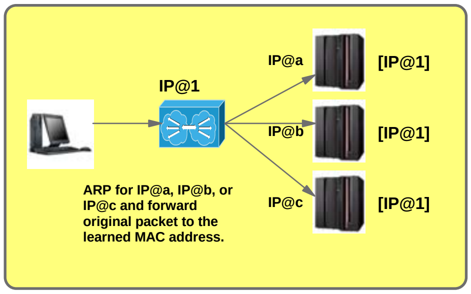


White paper - test scenarios

- Sysplex Distributor with MNLB forwarding agents
- Catalyst 6509 with CSM blade
  - Server NAT with Policy Based Routing
  - Server NAT and Client NAT
- Catalyst 6509 and CSS switch
  - Server NAT with Policy Based Routing
  - Server NAT and Client NAT
- Tests also included SSL modules in both CSM and CSS

This test and white paper focus on the connectivity between the Catalyst 6500, CSS, CSM, and the zSeries server using OSA-Express components.

Sysplex Distributor/MNLB using dispatch mode forwarding - basics



ARP for IP@a, IP@b, or IP@c and forward original packet to the learned MAC address.

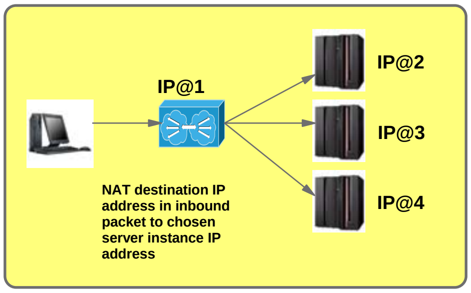
Sysplex Distributor is a disptach mode technology

- A distributed VIPA is transparently replicated to all LPARs in the Sysplex, but only the distributing LPAR advertises it to the network

Dispatch mode is also known as MAC-level forwarding.

- The server cluster IP address (IP@1) is from a networking point of view owned by the load-balancing node
- The load-balancer forwards inbound IP packets by selecting an appropriate next-hop IP address to route the packet to - there will be one configured next-hop IP address per server instance in the cluster
  - If servers are directly connected that means ARPing for the next-hop address and forwarding the original packet (including its original destination IP address) towards that discovered MAC address
- Only layer-4 load-balancing supported (TCP connections)
- If servers are more than one-hop away from the load balancer, then some encapsulation technology is needed to forward the packets (such as Generic Routing Encapsulation - GRE)
- No restrictions on outbound path from server instances (no NATing of either inbound or outbound traffic)

External load balancer using Directed Mode - basics



NAT destination IP address in inbound packet to chosen server instance IP address

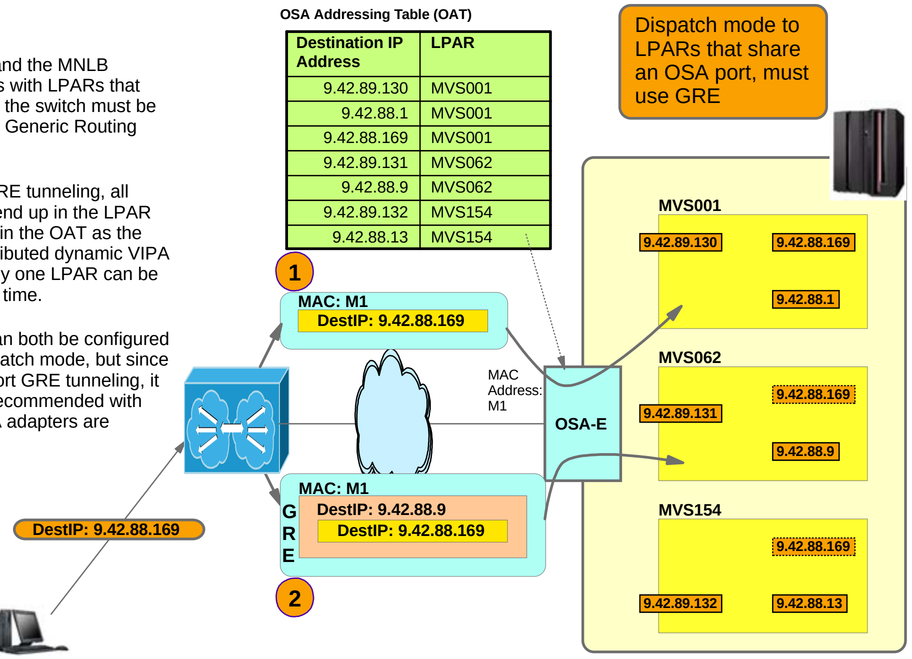
CSM and CSS are almost always configured to operate in directed mode

- They can be configured to operate in dispatch mode, but normally do not operate in that mode with a z/OS Sysplex due to issues with shared OSAs:
  - None of them are able to use GRE, which is required if OSA ports are shared
  - If OSA ports are not shared, then both CSM and CSS can be configured to operate correctly using dispatch mode

- The server cluster IP address (IP@1) is owned by the load-balancing node
- The load balancer forwards inbound IP packets by selecting an appropriate server instance specific IP address and changes the destination IP address in the inbound packets to that of the chosen server instance (performs server NAT)
- Outbound packets must be routed back via the load balancer for it to change the server IP address to the cluster IP address
- Layer-4 to layer-7 load balancing is supported
- No restrictions on network topology between the load balancer and the server nodes

Sysplex Distributor / MNLB - Dispatch Mode use of GRE tunneling

- When using SD and the MNLB forwarding agents with LPARs that share OSA ports, the switch must be configured to use Generic Routing Encapsulation.
- Without use of GRE tunneling, all connections will end up in the LPAR that is registered in the OAT as the owner of the distributed dynamic VIPA address - and only one LPAR can be so at any point in time.
- CSM and CSS can both be configured to operate in dispatch mode, but since they do not support GRE tunneling, it is generally not recommended with z/OS unless OSA adapters are non-shared.



OSA Addressing Table (OAT)

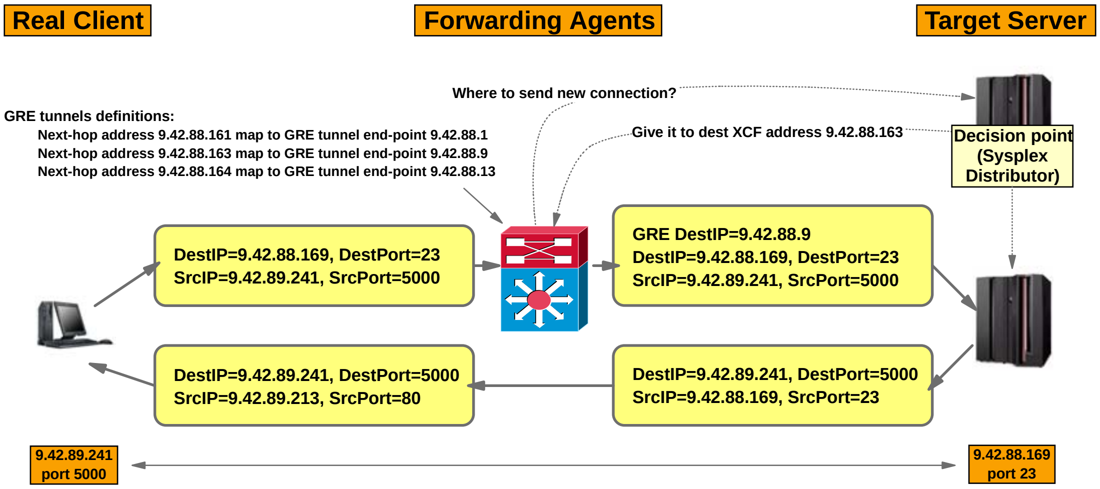
Destination IP Address	LPAR
9.42.89.130	MVS001
9.42.88.1	MVS001
9.42.88.169	MVS001
9.42.89.131	MVS062
9.42.88.9	MVS062
9.42.89.132	MVS154
9.42.88.13	MVS154

Dispatch mode to LPARs that share an OSA port, must use GRE

Redbooks

ibm.com/redbooks

Sysplex Distributor / MNLB - Dispatch mode forwarding - flow basics



Real Client
Forwarding Agents
Target Server
Where to send new connection?
GRE tunnels definitions:
Next-hop address 9.42.88.161 map to GRE tunnel end-point 9.42.88.1
Next-hop address 9.42.88.163 map to GRE tunnel end-point 9.42.88.9
Next-hop address 9.42.88.164 map to GRE tunnel end-point 9.42.88.13
Give it to dest XCF address 9.42.88.163
Decision point (Sysplex Distributor)
DestIP=9.42.88.169, DestPort=23
SrcIP=9.42.89.241, SrcPort=5000
GRE DestIP=9.42.88.9
DestIP=9.42.88.169, DestPort=23
SrcIP=9.42.89.241, SrcPort=5000
DestIP=9.42.89.241, DestPort=5000
SrcIP=9.42.89.213, SrcPort=80
DestIP=9.42.89.241, DestPort=5000
SrcIP=9.42.88.169, SrcPort=23
9.42.89.241
port 5000
9.42.88.169
port 23

Redbooks

ibm.com/redbooks

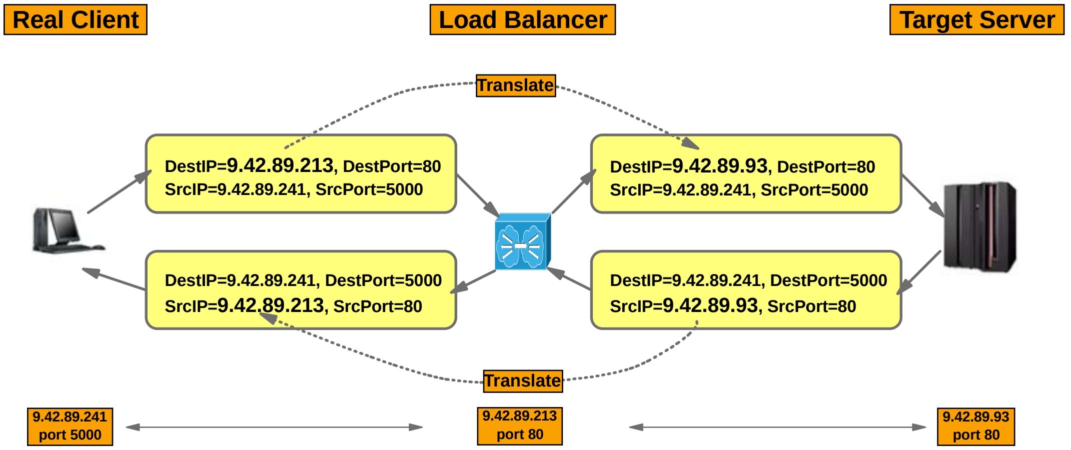
Sysplex Distributor / MNLB
Dispatch mode forwarding - key points

- Is based on dispatch mode forwarding – also known as MAC-level forwarding
- Load balancing decision point is inside the z/OS Sysplex and can take real-time server availability and LPAR capacity into consideration
- First SYN packet of a new inbound connection is routed via the Sysplex Distributor node and XCF to the target node. Succeeding packets per connection are routed directly from the 6509s to the target nodes
- Server IP address and client IP address are never NATed - there is no requirement for outbound packets to be routed via any specific path
- Dispatch mode forwarding must be combined with use of Generic Routing Encapsulation to overcome one-hop away and shared OSA limitations

Redbooks

ibm.com/redbooks

Cisco CSS/CSM - Directed mode - Server NAT and Policy Based Routing - flow basics



Real Client

Load Balancer

Target Server

Translate

DestIP=9.42.89.213, DestPort=80
SrcIP=9.42.89.241, SrcPort=5000

DestIP=9.42.89.93, DestPort=80
SrcIP=9.42.89.241, SrcPort=5000

DestIP=9.42.89.241, DestPort=5000
SrcIP=9.42.89.213, SrcPort=80

DestIP=9.42.89.241, DestPort=5000
SrcIP=9.42.89.93, SrcPort=80

Translate

9.42.89.241
port 5000

9.42.89.213
port 80

9.42.89.93
port 80

Redbooks

ibm.com/redbooks

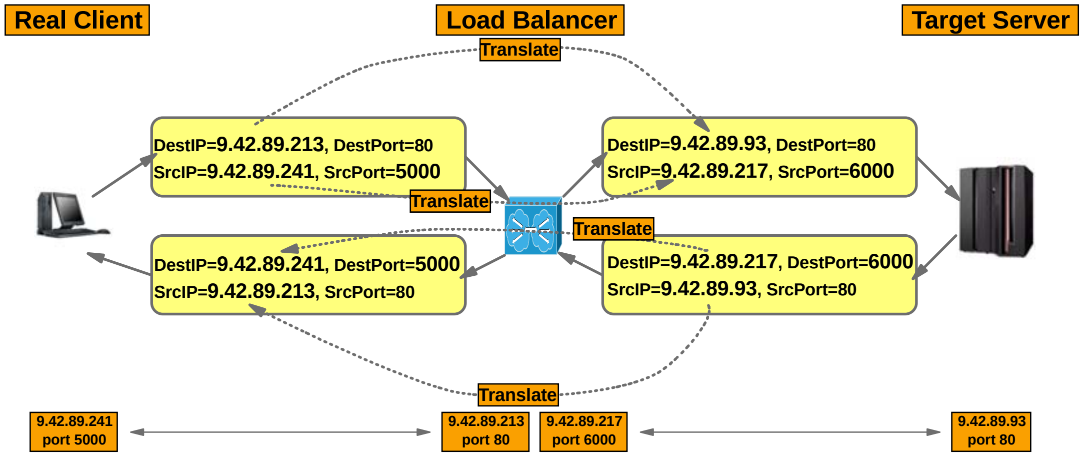
Cisco CSS/CSM - Directed mode - Server NAT and Policy Based Routing - key points

- Only the server IP address is NATed (destination IP address on inbound and source IP address on outbound)
- Policy-based routing in routing infrastructure re-directs outbound IP packets from target servers to the load-balancer so it can NAT the source IP address in outbound packets
- Outbound packets that do not need NATing of the source IP address are routed using normal IP routing table processing
- Real client IP address information is available to target servers

Redbooks

ibm.com/redbooks

Cisco CSS/CSM - Directed mode - Server NAT and Client NAT - flow basics



Real Client

Load Balancer

Target Server

Translate

DestIP=9.42.89.213, DestPort=80
SrcIP=9.42.89.241, SrcPort=5000

DestIP=9.42.89.93, DestPort=80
SrcIP=9.42.89.217, SrcPort=6000

Translate

Translate

DestIP=9.42.89.241, DestPort=5000
SrcIP=9.42.89.213, SrcPort=80

DestIP=9.42.89.217, DestPort=6000
SrcIP=9.42.89.93, SrcPort=80

Translate

9.42.89.241
port 5000

9.42.89.213
port 80

9.42.89.217
port 6000

9.42.89.93
port 80

Redbooks

ibm.com/redbooks

Cisco CSS/CSM - Directed mode - Server NAT and Client NAT - key points

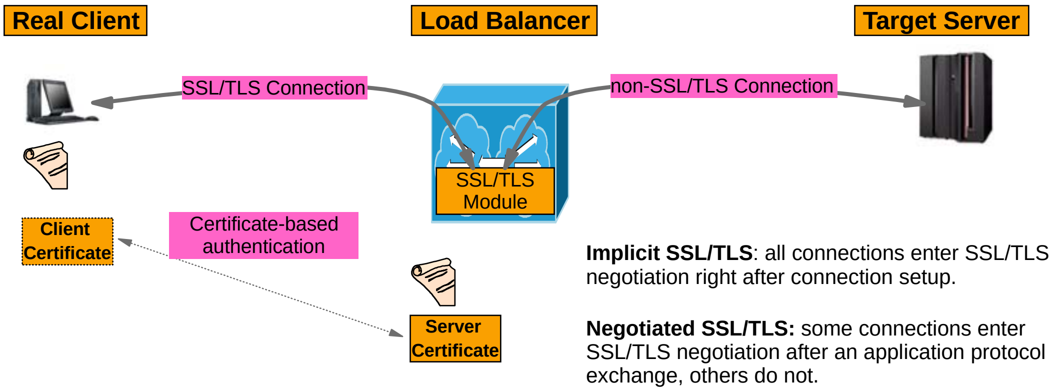
- Both server IP address and client IP address are NATed by the load balancer - no need for use of Policy Based Routing since outbound packets from target servers are destined for a load balancer IP address

- Client IP address seen by target servers is an IP address on the load balancer and not the real client IP address

- Should be used with care if any of the following functions on server nodes are in use:
  - Networking policy conditions based on client IP address information
  - NETACCESS rules for access control and/or MLS label assignment
  - Server configuration options based on source IP address information, such as TN3270 server LU name assignment

- May also complicate diagnosing certain error cases where real client IP address and port number are unknown on the server node

SSL/TLS offloading to CSS or CSM



Implicit SSL/TLS: all connections enter SSL/TLS negotiation right after connection setup.

Negotiated SSL/TLS: some connections enter SSL/TLS negotiation after an application protocol exchange, others do not.

- Implicit mode SSL/TLS workload can be offloaded (specific SSL/TLS server port number) - negotiated SSL/TLS can in general not be offloaded (shared port number)
  - FTP is negotiated
  - TN3270 can be either
  - HTTPS is implicit

- SSL/TLS processing cannot be offloaded if client authentication is required on target server

FTP workload key points

- **Both CSM and CSS stay out of the loop for FTP data connections**
  - Direct data connection between real client IP address and real server IP address
    - The case for both server NAT with PBR and Server NAT with Client NAT
  - SSL/TLS FTP (not the same as SFTP in SSH) with the SSL/TLS end-point on z/OS does work through load balancer
    - The CSM and CSS do not need to investigate the PORT command or the PASV reply
  - The new extended passive mode (EPSV) is not supported when using CSM or CSS in directed mode
    - This could be an issue if SSL/TLS FTP sessions are needed and those sessions traverse a traditional NAT firewall somewhere in the network before the connections hit the CSM/CSS load balancers - such an access path will not work
- **Sysplex Distributor / MNLB works for all combinations of FTP workload**
  - non-SSL/TLS active mode, passive mode, and extended passive mode
  - SSL/TLS active mode, passive mode, extended passive mode
    - Note: If SSL/TLS sessions pass through NAT firewalls on the path towards the Sysplex, then extended passive mode must be used

TN3270 workload load balancing key points

- Ensure TN3270 server instances do not assign the same LU names
  - Separate generic LU name pools
  - Same LU group names, but with different LU names
  - Stickiness!
- If printer association is used, printer connection request must go to same TN3270 server instance as LU2 connection request was processed by
  - Stickiness!
- For TN3270 server reconnect, reconnect request must go to same TN3270 server instance as original connection request
  - Stickiness!
- If client IP address or hostname is used in mapping rules, ensure client NATing is not used by load balancer

HTTP(S) workload load balancing key points

- Both CSM and CSS allow load-balancing based on content of the HTTP request
  - Sysplex Distributor/MNLB does not
- HTTPS connections must be terminated on the load balancer if content-based load balancing is required
  - SSL/TLS client authentication based on client's digital certificate is not possible on the real HTTP servers
- CSM and CSS can parse all HTTP header fields – including affinity cookies
  - Can direct connections with affinity to specific server instance

IBM®

z/OS Sysplex High Availability TCP/IP Solutions - Best Practices

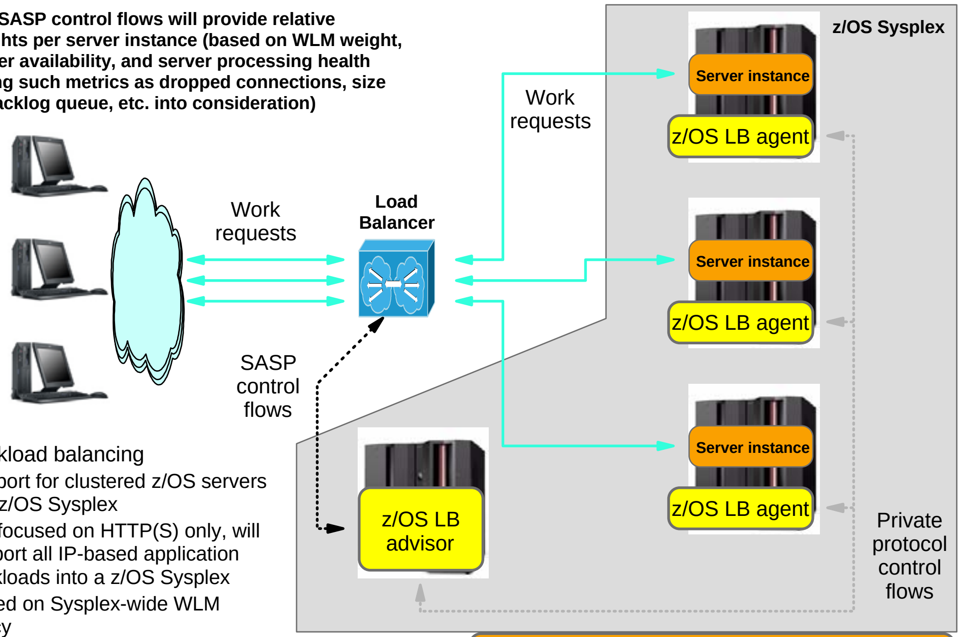
-

Providing z/OS Sysplex WLM Weights and Server Application Health Information to External Load Balancers

International Technical Support Organization

z/OS Load Balancing Advisor (LBA) for outboard load balancers

The 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)



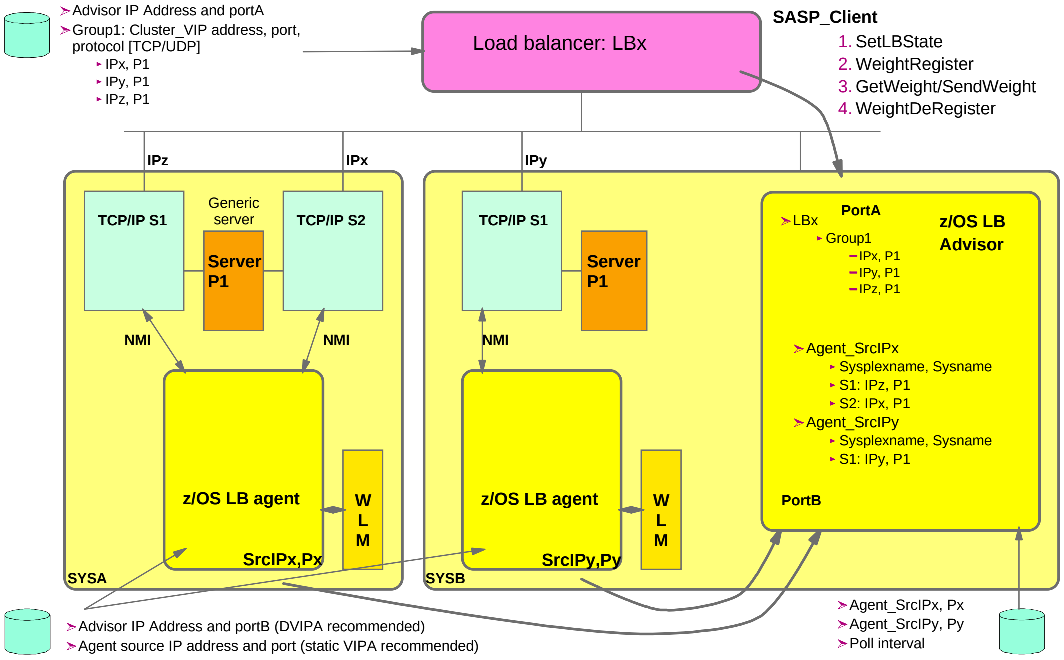
z/OS workload balancing

- 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 is planned for general availability 4Q2004

Redbooks V1R6

z/OS LB Advisor/Agent structure - overview



Redbooks V1R6

ibm.com/redbooks

The weights

- The weights are composed of two main elements:
  - ***WLM weight***
    - The WLM weight based on displaceable LPAR capacity as 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, the z/OS LBA currently will never return a zero weight - the lowest weight it will return is a weight of 1
  - 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

Redbooks V1R6 ibm.com/redbooks

Load balancer registrations

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 addressed are matched to TCP/IP stacks in the Sysplex.
  - WLM weights for the LPARs are retrieved.
  - CS weight indicate if IP address is active in the Sysplex or not (0 or 100).
  - LBA displays will show a protocol value of zero for system group registrations.
- ❑An application group
  - Represented by a list of IP address, Protocol (TCP or UDP), and port.
  - 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 the z/OS load balancing advisor, it instructs the advisor 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

Redbooks V1R6 ibm.com/redbooks

Some strategies for workload request balancing into a z/OS Sysplex

A. **DNS/WLM as workload balancing should not be used any longer**

B. **Where HTTP workload is to be balanced based on content of HTTP requests, an outboard load balancer that supports contents inspection must be deployed**

- If HTTPS workload is to be included, the load balancing node must be accompanied by an SSL/TLS offload technology
- Can be combined with a cache appliance for improved performance

C. **UDP workload balancing must be deployed using an outboard load balancer - SD does not support UDP balancing**

D. **Remaining TCP connection balancing can be deployed using either SD or an outboard load balancer:**

- SD has more real-time information available than outboard load balancers - even with outboard load balancers using the SASP protocol
- Who is to apply management control over the workload balancing function will be a major factor in deciding which solution to use
- If installation cannot combine use of SD with Cisco MNLB forwarding agent support, an outboard LB may from a performance point of view be preferable to SD

IBM®

z/OS Sysplex High Availability TCP/IP Solutions - Best Practices

-

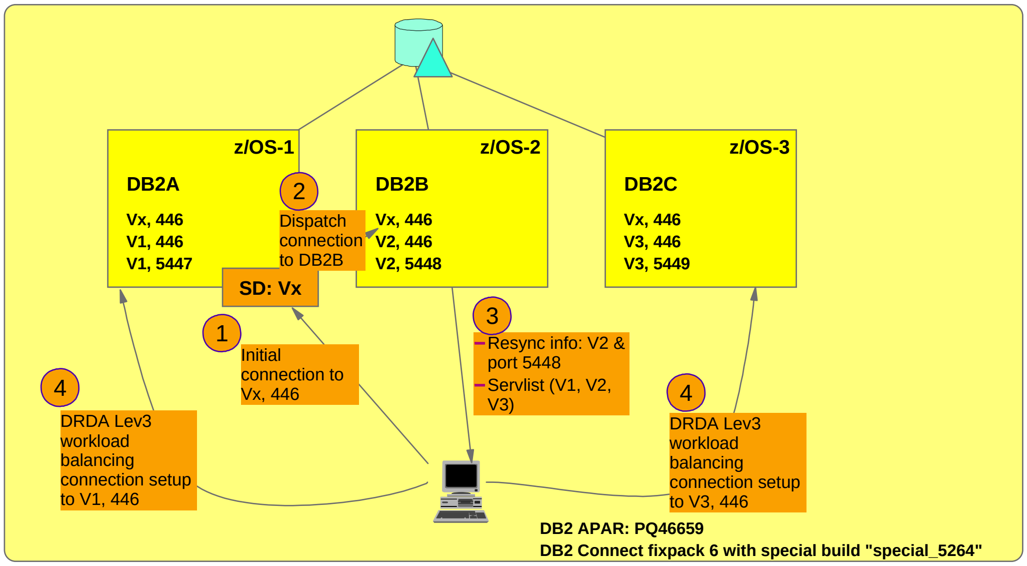
Appendix:

z/OS Subsystem-specific Load Balancing Guidelines

Redbooks

International Technical Support Organization

Dynamic VIPAs and sysplex Distributor


z/OS-1
DB2A
Vx, 446
V1, 446
V1, 5447
z/OS-2
DB2B
Vx, 446
V2, 446
V2, 5448
z/OS-3
DB2C
Vx, 446
V3, 446
V3, 5449
2
Dispatch connection to DB2B
SD: Vx
1
Initial connection to Vx, 446
3
Resync info: V2 & port 5448
Servlist (V1, V2, V3)
4
DRDA Lev3 workload balancing connection setup to V1, 446
4
DRDA Lev3 workload balancing connection setup to V3, 446

DB2 APAR: PQ46659
DB2 Connect fixpack 6 with special build "special_5264"

With the above listed APAR solution for DB2 V6 and V7, DB2 can use application-specific dynamic VIPA addresses and Sysplex Distributor for connection balancing. This allows for restart of a failed DB2 instance on another LPAR without manual movement of IP addresses.

Details of TCP/IP definitions for DB2 data sharing

z/OS-2

```
port
  446 tcp db2adist shareport bind Vx
 5447 tcp db2adist bind V1
  446 tcp db2bdist shareport bind Vx
 5448 tcp db2bdist bind V2
  446 tcp db2cdist shareport bind Vx
 5449 tcp db2cdist bind V3
VipaDynamic
 VipaRange Define 255.255.255.255 V1
 VipaRange Define 255.255.255.255 V2
 VipaRange Define 255.255.255.255 V3
 VipaBackup 1 Vx
EndVipaDynamic
```

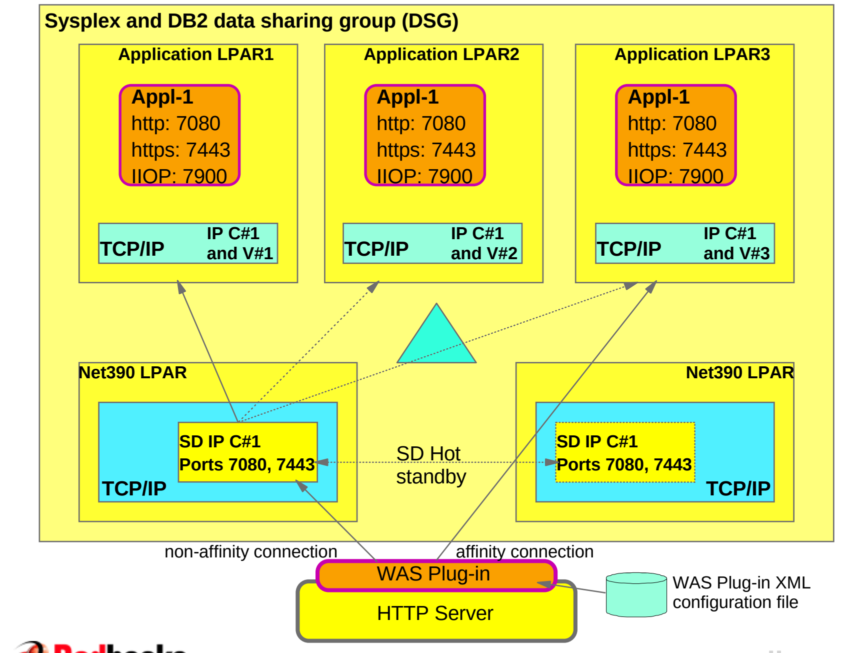
z/OS-1

```
port
  446 tcp db2adist shareport bind Vx
 5447 tcp db2adist bind V1
  446 tcp db2bdist shareport bind Vx
 5448 tcp db2bdist bind V2
  446 tcp db2cdist shareport bind Vx
 5449 tcp db2cdist bind V3
VipaDynamic
 VipaRange Define 255.255.255.255 V1
 VipaRange Define 255.255.255.255 V2
 VipaRange Define 255.255.255.255 V3
 VipaDefine 255.255.255.255 Vx
 VipaDistribute Define Vx Port 446 DestIP all
EndVipaDynamic
```

z/OS-3

```
port
  446 tcp db2adist shareport bind Vx
 5447 tcp db2adist bind V1
  446 tcp db2bdist shareport bind Vx
 5448 tcp db2bdist bind V2
  446 tcp db2cdist shareport bind Vx
 5449 tcp db2cdist bind V3
VipaDynamic
 VipaRange Define 255.255.255.255 V1
 VipaRange Define 255.255.255.255 V2
 VipaRange Define 255.255.255.255 V3
 VipaBackup 2 Vx
EndVipaDynamic
```

Sysplex Distributor as WAS application connection balancer in a Sysplex

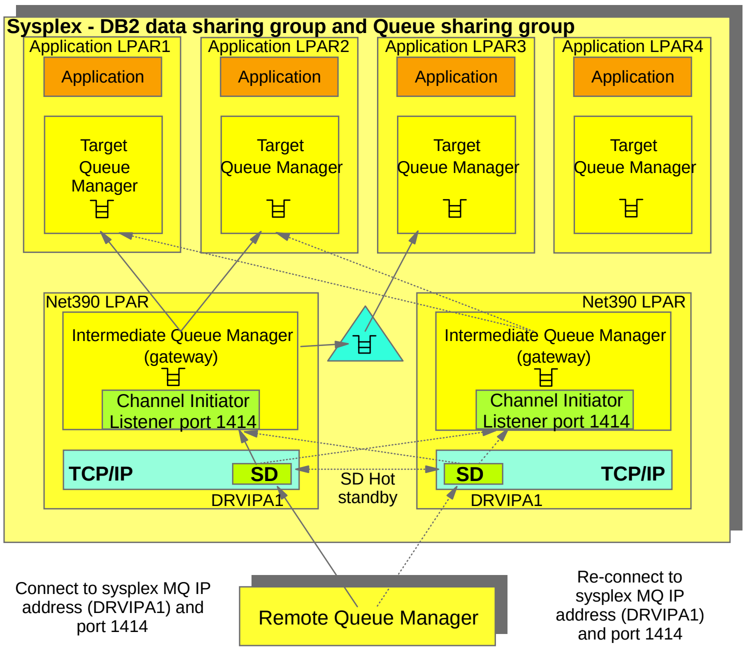


The application servers must not be bind-specific - they must be able to respond to connections that arrive for both the cluster IP address (C#1) and server-specific IP address (V#1, V#2, or V#3).

Redbooks

ibm.com/redbooks

Queue Manager gateway in a Sysplex using Net390 Nodes - one large QSG



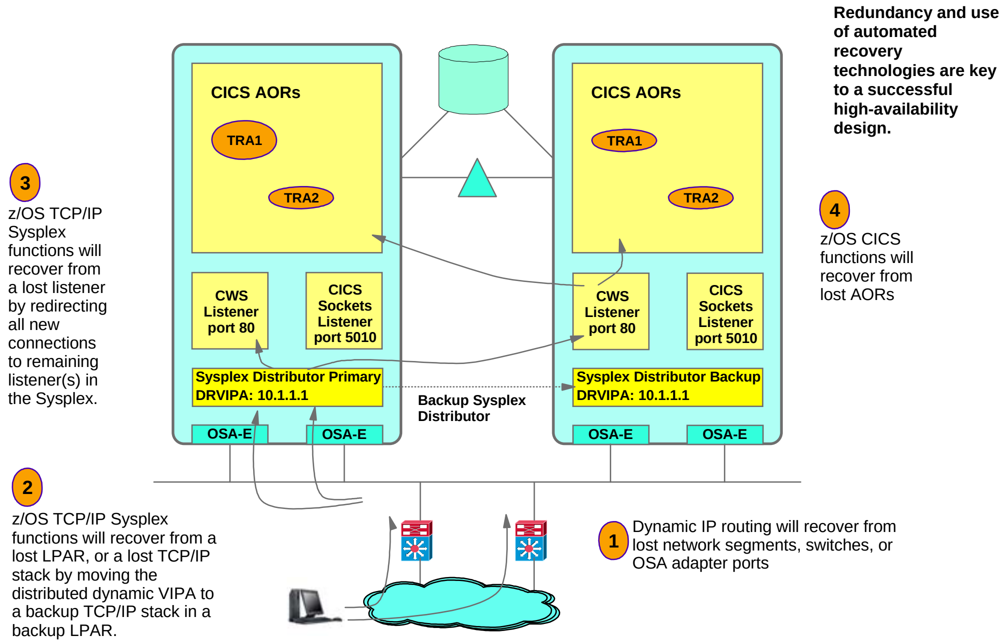
Connect to sysplex MQ IP address (DRVIPA1) and port 1414

Re-connect to sysplex MQ IP address (DRVIPA1) and port 1414

Redbooks

ibm.com/redbooks

High-availability design for TCP/IP workload into the CICS environment



Redundancy and use of automated recovery technologies are key to a successful high-availability design.

CICS AORs
TRA1
TRA2
CWS Listener port 80
CICS Sockets Listener port 5010
Sysplex Distributor Primary DRVIPA: 10.1.1.1
OSA-E
OSA-E
Backup Sysplex Distributor
CICS AORs
TRA1
TRA2
CWS Listener port 80
CICS Sockets Listener port 5010
Sysplex Distributor Backup DRVIPA: 10.1.1.1
OSA-E
OSA-E
3
4
2
1

3 z/OS TCP/IP Sysplex functions will recover from a lost listener by redirecting all new connections to remaining listener(s) in the Sysplex.

4 z/OS CICS functions will recover from lost AORs

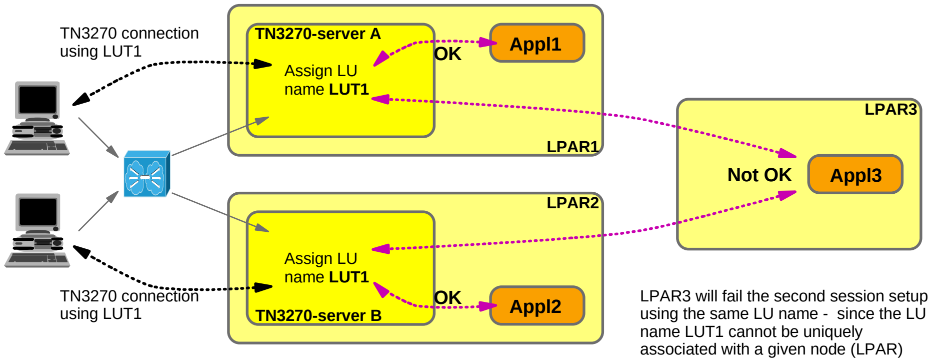
2 z/OS TCP/IP Sysplex functions will recover from a lost LPAR, or a lost TCP/IP stack by moving the distributed dynamic VIPA to a backup TCP/IP stack in a backup LPAR.

1 Dynamic IP routing will recover from lost network segments, switches, or OSA adapter ports

Assigning LU names when connection balancing TN3270 connections

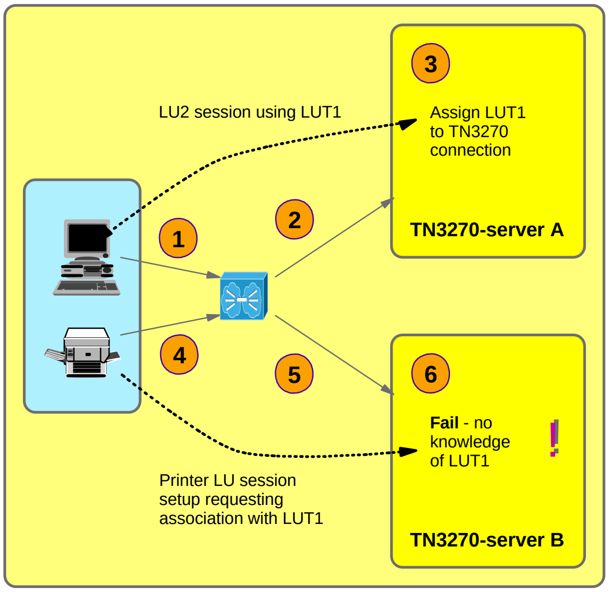
Three main ways LU names can be assigned:

1. **Generic request** - server decides entirely which LU name to assign
   - Each TN3270 server must have its own group of LU names (generic pools)
2. **Specific request with pool name** - server validates and decides which LU name in the named pool to assign
   - Pool names can be the same in more TN3270 servers, but each server must have its own group of LU names assigned to those pool names
3. **Specific request with LU name** - server validates and assigns requested LU name
   - Generally TN3270 servers can allow assigning the same LU names since theoretically a client workstation will only request an LU name once



LPAR3 will fail the second session setup using the same LU name - since the LU name LUT1 cannot be uniquely associated with a given node (LPAR)

TN3270 connection balancing and printer association issues



1. Initiate connection for establishing an LU2 session.
2. This connection is sent to TN3270-server A.
3. TN3270-server A selects an LU2 LU name - for the purpose of this example it doesn't matter if it is a generic request or a specific request. As an example an LU named LUT1 is chosen.
4. The workstation user now starts a printer emulator and a request for a new connection with the TN3270 server is initiated.
5. This connection is sent to TN3270-server B.
6. The workstation emulator now requests a printer LU name that is associated with the LU2 LU name from the first session LUT1 - and TN3270-server B is going to reject that since it doesn't know anything about the LU name that was assigned by TN3270-server A.

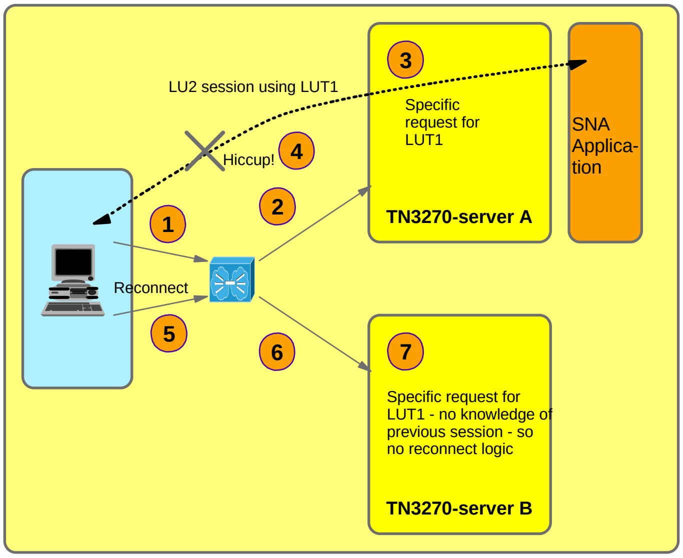
Affinity (or stickiness) needed by load balancer to handle this case so all connections from the same client go to the same server instance.

Sysplex Distributor supports affinity from z/OS V1R5.

Redbooks

ibm.com/redbooks

TN3270 connection balancing and reconnects



1. Initiate connection for establishing an LU2 session.
2. This connection is sent to TN3270-server A.
3. TN3270-server A allows LU name LUT1 and user connects to SNA application.
4. Network has hiccups and client believes connection is broken.
5. Client sends a new connection (reconnect) request.
6. Reconnect request is sent to TN3270-server B.
7. If this connection had ended up in server A, reconnect logic would have kicked in and reconnected the user with the SNA application. But server B doesn't know anything about the previous session and treats it as a new connection and allows LU name LUT1, which now is active on both LPARs.

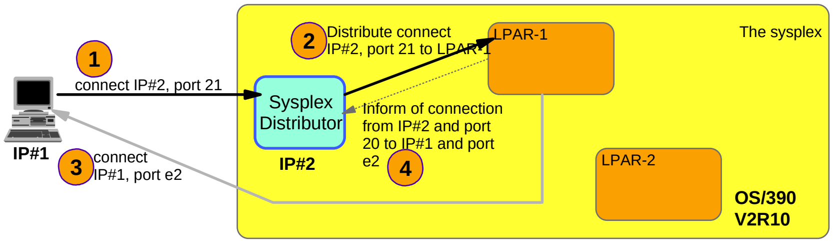
Affinity (or stickiness) needed by load balancer to handle this case so all connections from the same client go to the same server instance.

Sysplex Distributor supports affinity from z/OS V1R5.

Redbooks

ibm.com/redbooks

FTP workload balancing - Active mode FTP and Sysplex Distributor

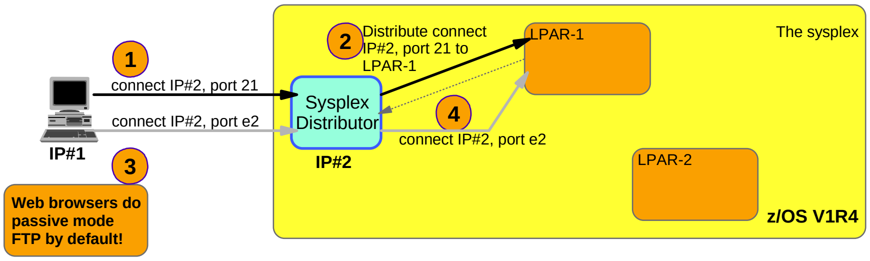


1. Connection to the cluster address (IP#2) and the FTP control connection port number (port 21) is intercepted by the Sysplex Distributor LPAR
2. Sysplex Distributor distributes the connection to LPAR-1 - and we now have a connection between IP#1, port e1 and IP#2, port 21 - Sysplex Distributor remembers this 4-tuple and forwards all inbound IP packets from IP#1 port e1 towards IP#2 port 21 to LPAR-1 hereafter
3. When the data connection is to be established, the client opens a new listening socket and sends a PORT command to the FTP server over the control connection instructing it which IP address and port number to connect to in order to set up the data connection. The FTP server will do so from a new socket that is locally bound to port 20 - and we will now have a connection between IP#2 (the cluster IP address) port 20 and IP#1 port e2.
4. Because the inbound IP packets for this data connection will come to the cluster IP address they will show up at the Sysplex Distributor node and it needs to know where to send them for this particular data connection. That information is conveyed to the Sysplex Distributor node from LPAR-1 if the source port number (port 20) is included in the list of port numbers for which the Sysplex Distributor is supposed to distribute.

Active mode FTP connections can be distributed using Sysplex Distributor since OS/390 V2R10 as long as both the control connection and data connection server port numbers are included in the VIPADISTRIBUTE statement (port 21 and port 20)

ibm.com/redbooks

Passive mode FTP and Sysplex Distributor supported in z/OS V1R4



1. Connection to the cluster address (IP#2) and the FTP control connection port number (port 21) is intercepted by the Sysplex Distributor LPAR
2. Sysplex Distributor distributes the connection to LPAR-1 - and we now have a connection between IP#1, port e1 and IP#2, port 21 - Sysplex Distributor remembers this 4-tuple and forwards all inbound IP packets from IP#1 port e1 towards IP#2 port 21 to LPAR-1 hereafter
3. When the data connection is to be established the server now opens a new socket and binds it to some ephemeral port number (e2) and the IP address to which the control connection was established (IP#1) and sends this information to the client over the control connection as the PASV reply. The client now sends a connect request to IP#1 and port e2 and this request shows up at the Sysplex Distributor node. Since the destination port is an ephemeral port we cannot pre-configure it in the VIPADISTRIBUTE statement so the distributing node doesn't know that this connection really should go to LPAR-1 - and will in releases up until z/OS V1R4 not be able to distribute such passive mode FTP sessions.
4. In z/OS V1R4 support has been added so that when the FTP server binds its data socket to the cluster IP address (IP#1) and an ephemeral port number, that information will be sent to the distributing node and the distribution control tables will be dynamically extended with this ephemeral port number and an affinity with the LPAR from where it originated. So when the data connection request arrives in the distributing node, it knows it has to go to LPAR-1.

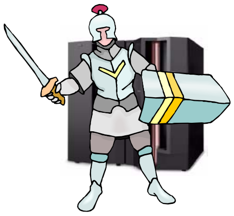
Passive mode FTP connections can be distributed using Sysplex Distributor from z/OS V1R4. Please note that this support requires SYSPLEXPORTS to be specified on the Distributed DVIPA that is used for FTP workload.

ibm.com/redbooks

ibm.com

e-business

z/OS Network Security



Redbooks

International Technical Support Organization

Objectives

The objectives of this session are:

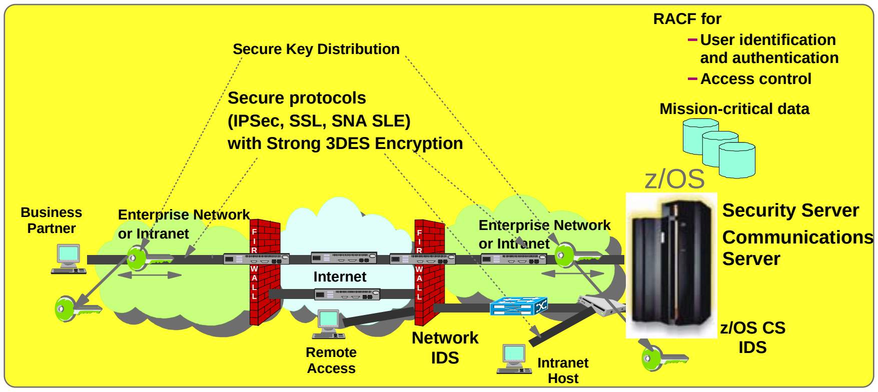
- Introduce the security roles of the Communications Server on z/OS
- Understand how CS z/OS can be used to extend external network security functions by providing a layer of self-protective functions within the z/OS operating system
- Define what system and resource protection is - integrated z/OS firewall functions, SAF-based protection, etc.
- Define z/OS's role in network security - IPSec, SSL/TLS, etc.
- Enable participants to understand the various security technologies, in order to select the proper technology for their specific needs.

Agenda

- ✓ **z/OS Communications Server security roles and objectives**
- ✓ **System and resource protection**
  - Firewall Technologies on z/OS
  - Intrusion Detection Services
  - Syslogd Protection
  - SAF SERVAUTH class protection of TCP resources
  - Multilevel Security
- ✓ **Network security**
  - Transparent Application Security
    - IPSec and VPN
    - TN3270 SSL & Express Logon
  - Built-in Application Security
    - Secure Network Services
    - TLS support for FTP

z/OS CS security roles and objectives

- ✓ **Secure access to both TCP/IP and SNA applications**
- ✓ **Focus on end-to-end security and self-protection**
- ✓ **Exploits strengths of S/390 and zSeries hardware and software**



- **Protect data and other resources on the system**
  - ***System availability***
    - Protect system against unwanted access and denial of service attacks from network
  - ***Identification and authentication***
    - Verify identity of users
  - ***Access control***
    - Protect data and other system resources from unauthorized access
- **Protect data in the network using cryptographic security protocols**
  - ***Data Origin Authentication***
    - Verify that data was originated by claimed sender
  - ***Message Integrity***
    - Verify contents were unchanged in transit
  - ***Data Privacy***
    - Conceal cleartext using encryption

IP-based security technology overview and introduction

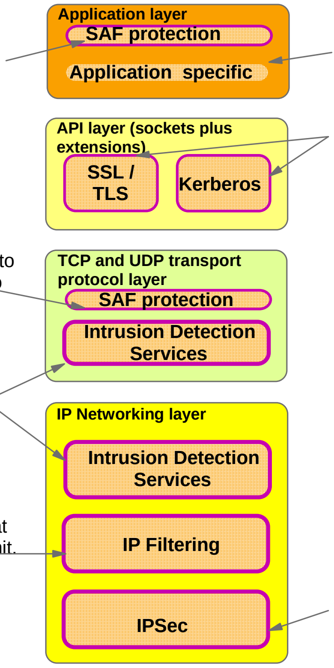
Protect the system

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.


Application layer
SAF protection
Application specific
API layer (sockets plus extensions)
SSL / TLS
Kerberos
TCP and UDP transport protocol layer
SAF protection
Intrusion Detection Services
IP Networking layer
Intrusion Detection Services
IP Filtering
IPSec

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.

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

Redbooks

ibm.com/redbooks

IBM®

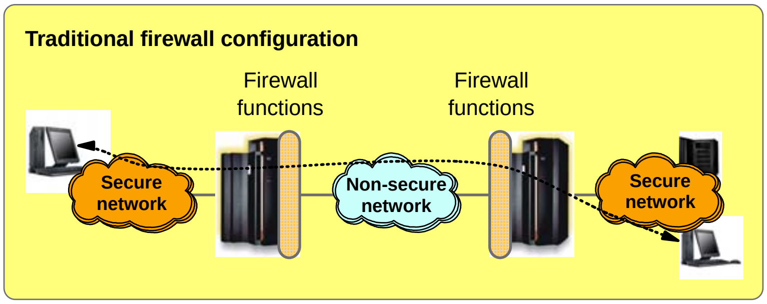
z/OS Network Security
-
z/OS Communications Server Resource Protection

Redbooks

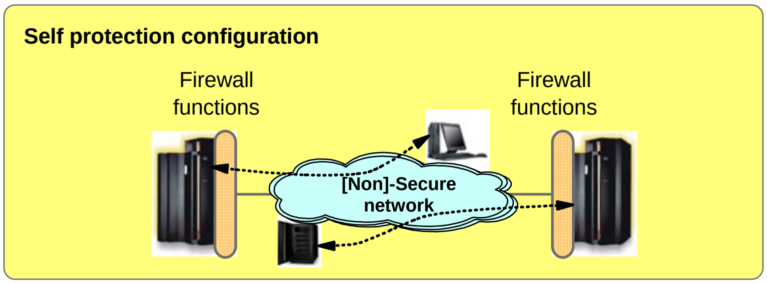
International Technical Support Organization

Firewall technologies usage scenarios on z/OS

- 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


Traditional firewall configuration

- You can also choose to use the z/OS firewall technologies on your normal z/OS LPARs to
  - add an extra layer of network access protection through IP filtering
  - support VPN end-points on z/OS


Self protection configuration

ibm.com/redbooks

z/OS firewall technologies

The firewall technologies functions that are shipped with z/OS	Included in Communications Server	Included in Security Server Free	Included in Security Server Non-free	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			✓	✓	

- z/OS firewall technologies have been available since OS/390 V2R4 and are today shipped partly with the Communications Server and partly with the Security Server on z/OS.
- Most of the functions are useful both in a traditional firewall configuration and as self-protection functions on z/OS.

ibm.com/redbooks

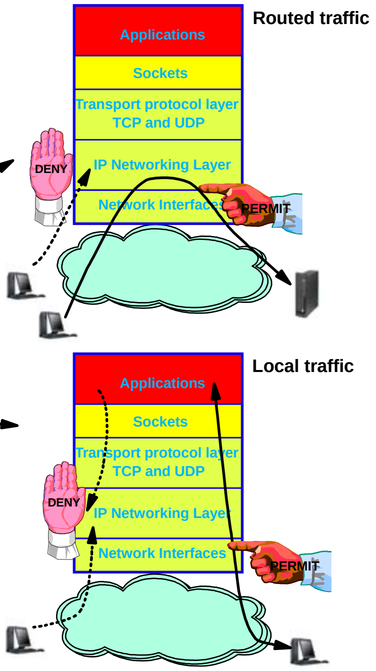
IP packet filtering on z/OS

Packet filtering at IP Layer

- Filter rules defined to deny or permit packets based on:
  - IP source/dest address
  - Protocol (TCP, TCP with ACK, UDP, ICMP, etc.)
  - Source/dest port
  - Direction of flow
  - Local or routed traffic
  - Time
  - Network interface
- Used to control
  - Traffic being routed
  - Access at destination host (local)
- When IP filtering is active, a default rule will deny all packets that are not specifically permitted

Packaging (firewall technologies)

- Security Server
  - Configuration through z/OS UNIX command line interface or Configuration GUI
- Communications Server
  - Runtime IP packet filtering



ibm.com/redbooks

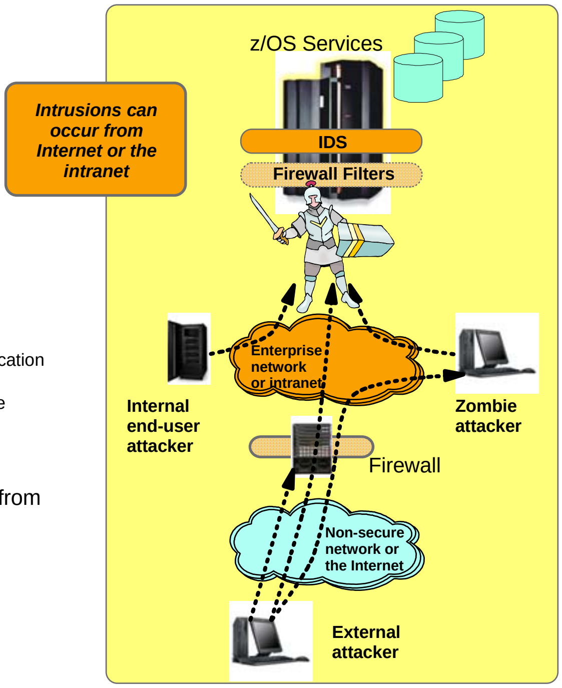
Intrusion threat - protecting against attacks on your system or your legitimate (open) services

What is an intrusion?

- Information Gathering
  - Network and system topology
  - Data location and contents
- Eavesdropping/Impersonation/Theft
  - On the network/on the host
  - Base for further attacks on others
    - Amplifiers
    - Robot or zombie
- Denial of Service
  - Attack on availability
    - Single Packet attacks - exploits system or application vulnerability
    - Multi-Packet attacks - floods systems to exclude useful work

Attacks can occur from Internet or intranet

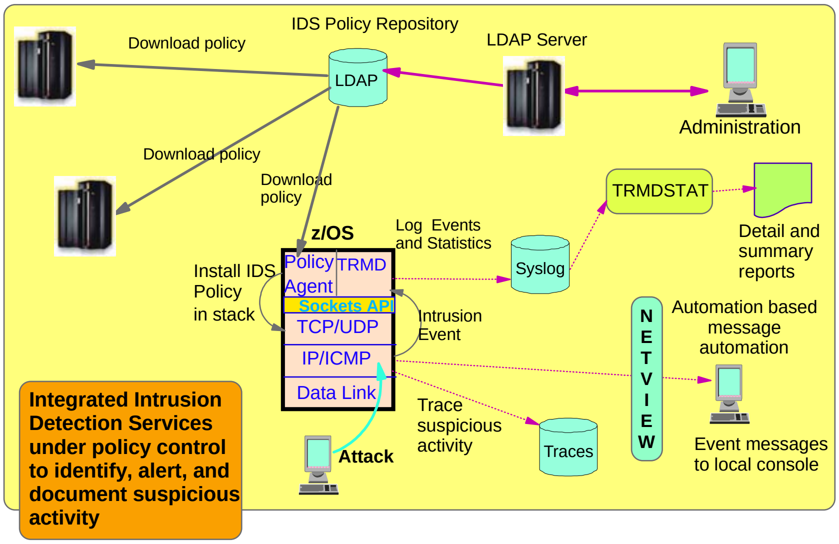
- Firewall can provide some level of protection from Internet
- Perimeter Security Strategy *alone* may not be sufficient.
  - Considerations:
    - Access permitted from Internet
    - Trust of intranet



Intrusions can occur from Internet or the intranet
z/OS Services
IDS
Firewall Filters
Enterprise network or intranet
Internal end-user attacker
Zombie attacker
Firewall
Non-secure network or the Internet
External attacker

ibm.com/redbooks

Intrusion Detection Services overview



IDS Policy Repository

Download policy

LDAP

LDAP Server

Administration

Download policy

Download policy

TRMDSTAT

Detail and summary reports

z/OS

Log Events and Statistics

Syslog

Install IDS Policy in stack

Policy Agent | TRMD

Sockets API

TCP/UDP

IP/ICMP

Data Link

Intrusion Event

NETVIEW

Automation based message automation

Trace suspicious activity

Traces

Event messages to local console

Attack

Integrated Intrusion Detection Services under policy control to identify, alert, and document suspicious activity

Events detected
- Scans, attacks against stack, flooding (both TCP and UDP)

Defensive methods
- Packet discard, limit connections

Reporting
- Logging, event messages to local console, IDS packet trace
- Notifications to NetView

Security policy stored in LDAP

z/OS IDS broadens intrusion detection coverage:
- Ability to evaluate inbound encrypted data - IDS applied after decryption on the target system
- Avoids overhead of per packet evaluation against table of known attacks - IDS policy checked after attack detected
- Detects statistical anomalies real-time - target system has stateful data / internal threshholds unavailable to external IDSs
- Policy can control prevention methods on the target, such as connection limiting and packet discard

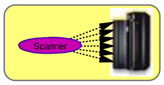
Redbooks

ibm.com/redbooks

IDS Event types

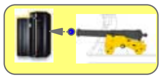
- **Scan detection and reporting**
  - Intent of scanning is to map the target of the attack (Subnet structure, addresses, masks, addresses in-use, system type, op-sys, application ports available, release levels)
    - TCP port scans
    - UDP port scans
    - ICMP scans
      - Sensitivity levels for all scans can be adjusted to control number of false positives recorded.

Scan



- **Attack detection, reporting, and prevention**
  - Intent is to crash or hang the system (Single or multiple packet)
    - Malformed packet events
    - Inbound fragment restrictions
    - IP option restrictions
    - IP protocol restrictions
    - ICMP redirect restrictions
    - Flooding events (SYN flood detections, physical interface flood detection added in z/OS V1R5)
    - Outbound raw restrictions
    - UDP perpetual echo

Attack



- **Traffic regulation for TCP connections and UDP receive queues**
  - Could be intended to flood system OR could be an unexpected peak in valid requests
    - UDP backlog management by port
      - Packets discard
    - TCP total connection and source percentage management by port
      - Connection limiting

Flooding

Redbooks

ibm.com/redbooks

IDS actions and message automation

- **Options**
  - Event logging
    - Syslogd - Number of events per attack subtype recorded in a five minute interval is limited
    - Local Console - Recording suppression provided if quantity of IDS console messages reach policy-specified thresholds
  - Statistics
    - Syslogd - Normal, Exception
  - IDS packet trace
    - Activated after attack detected
      - Number of packets traced for multi-packet events are limited
      - Amount of data trace is configurable (header, full, byte count)

- **All IDS events recorded in syslog and console messages, and packet trace records have probeid and correlator**
  - Probeid identifies the specific event detected
  - Correlator allows events to be matched with corresponding packet trace records

- **Console message can drive message automation**
  - MPF message suppression can suppress message output to system console
  - Example automation actions:
    - Route message to NetView console(s)
    - Email notification to security administrator
    - Run trmdstat and attach output to email
  - Selectors
    - Automate based on message number, other message content, such as event type or probe ID

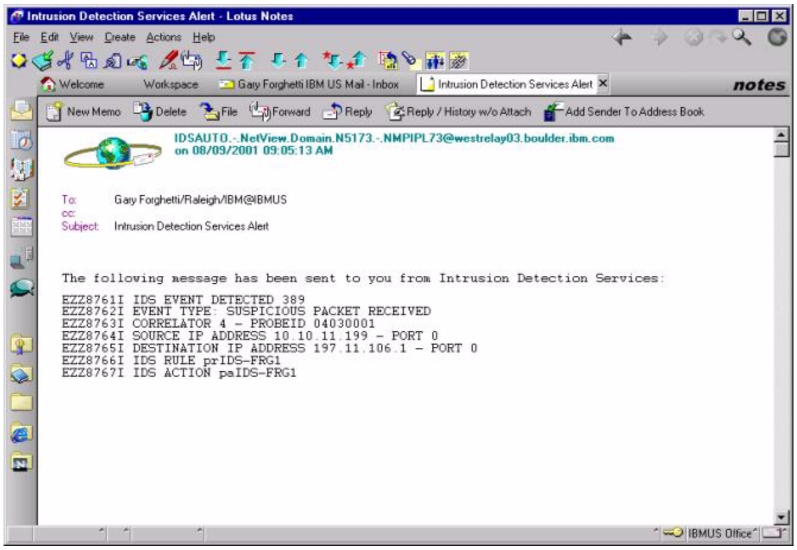
NetView clists: http://www.ibm.com/support/all_download_drivers.html
Search: idsauto

Redbooks

ibm.com/redbooks

IDSAUTO in NetView

Example of an email sent by the NetView IDSAUTO solution as the result of an Intrusion event being detected by the IDS component of z/OS TCP/IP:


Intrusion Detection Services Alert - Lotus Notes
File Edit View Create Actions Help
Welcome
Workspace
Gary Forghetti IBM US Mail - Inbox
Intrusion Detection Services Alert
notes
New Memo
Delete
File
Forward
Reply
Reply / History w/o Attach
Add Sender To Address Book
IBMUS Office

IDSAUTO.-.NetView.Domain.N5173.-.NMPIPL73@westrelay03.boulder.ibm.com
on 08/09/2001 09:05:13 AM

To: Gary Forghetti/Raleigh/IBM@IBMUS
cc:
Subject: Intrusion Detection Services Alert

```
The following message has been sent to you from Intrusion Detection Services:

EZZ8761I IDS EVENT DETECTED 389
EZZ8762I EVENT TYPE: SUSPICIOUS PACKET RECEIVED
EZZ8763I CORRELATOR 4 - PROBEID 04030001
EZZ8764I SOURCE IP ADDRESS 10.10.11.199 - PORT 0
EZZ8765I DESTINATION IP ADDRESS 197.11.106.1 - PORT 0
EZZ8766I IDS RULE prIDS-FRG1
EZZ8767I IDS ACTION paIDS-FRG1
```

Redbooks V1R5

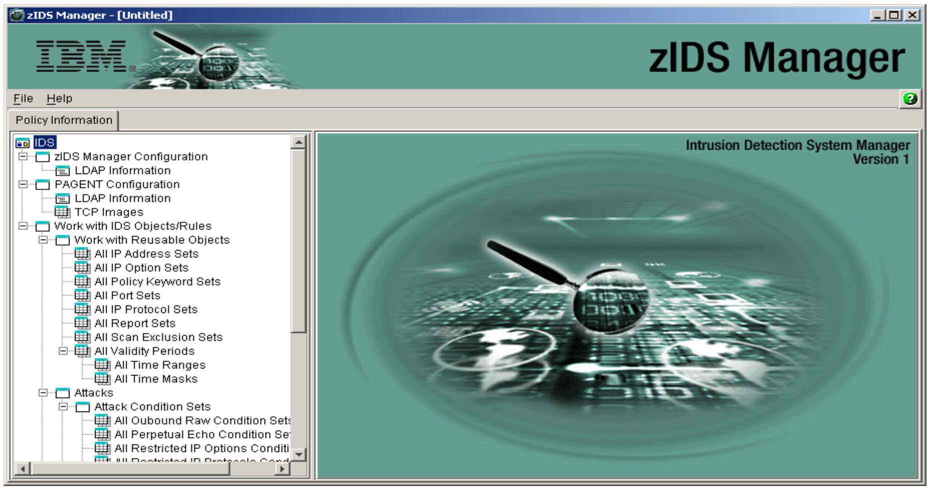
ibm.com/redbooks

Enhanced flood detection

- Prior to z/OS V1R5 individual events (for example, malformed packet, requests to an unbound port, queue full conditions, etc.) are detected and handled.
- IDS can provide notification about the individual instances but the discarded packets may be a symptom of a larger problem such as a flood.
- The only flood type currently detected by IDS Attack support is a SYN flood.
- z/OS V1R5 adds interface flood detection support as part of IDS flood detection:
  - A high percentage of discarded packets on a physical interface may indicate the interface is under attack.
  - Using the information already detected by IDS, track the discard rate by physical interface to determine if there is a potential attack
  - Notify the customer that a possible interface flood condition is occurring if the discard rate exceeds a specified limit.
  - Provide information to help determine the potential cause of the interface flood
  - Allow the customer to specify policy criteria used to detect an interface flood
  - Provide detection support without a large performance impact

Redbooks V1R5 ibm.com/redbooks

Defining IDS policy for z/OS - using Windows or Linux with the zIDS Manager



zIDS Manager - [Untitled]

zIDS Manager

File Help

Policy Information

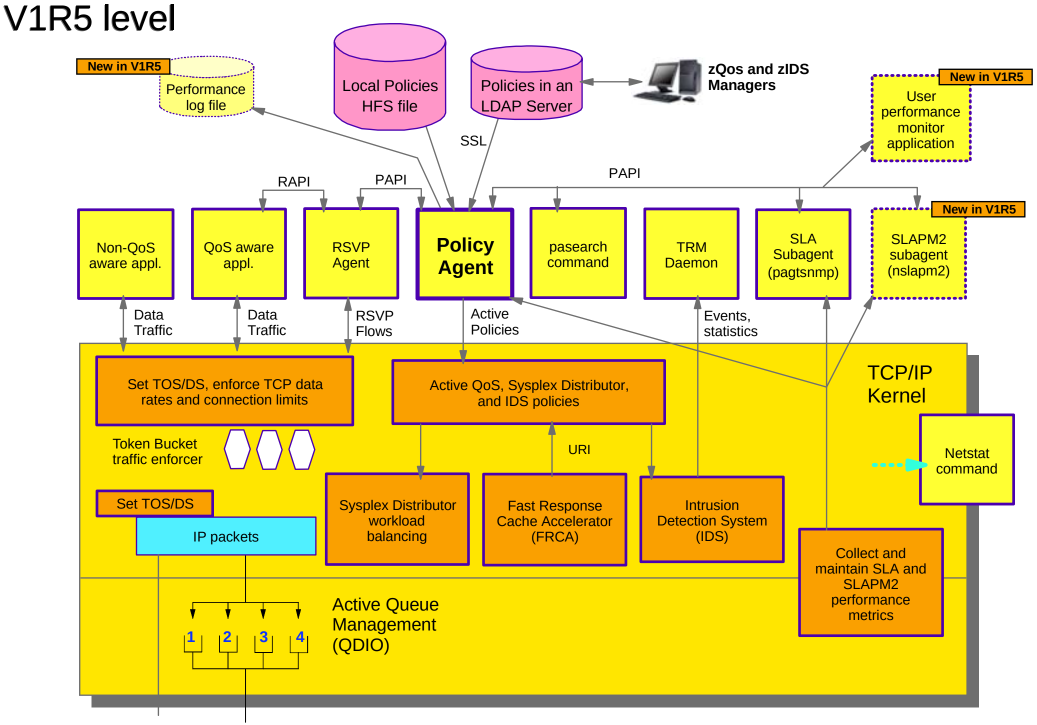
- IDS
  - zIDS Manager Configuration
    - LDAP Information
  - PAGENT Configuration
    - LDAP Information
    - TCP Images
  - Work with IDS Objects/Rules
    - Work with Reusable Objects
      - All IP Address Sets
      - All IP Option Sets
      - All Policy Keyword Sets
      - All Port Sets
      - All IP Protocol Sets
      - All Report Sets
      - All Scan Exclusion Sets
      - All Validity Periods
        - All Time Ranges
        - All Time Masks
    - Attacks
      - Attack Condition Sets
        - All Oubound Raw Condition Sets
        - All Perpetual Echo Condition Se
        - All Restricted IP Options Conditi
        - All Restricted IP Protocols Cond

Intrusion Detection System Manager
Version 1

zIDSManager "as is" Web Tool: to implement LDAP policies.
Available at: http://www-3.ibm.com/software/network/commserver/downloads
Needs Java Runtime 1.3 or 1.4.
Available at: http://java.sun.com/j2se/1.4.1/download.html

Redbooks ibm.com/redbooks

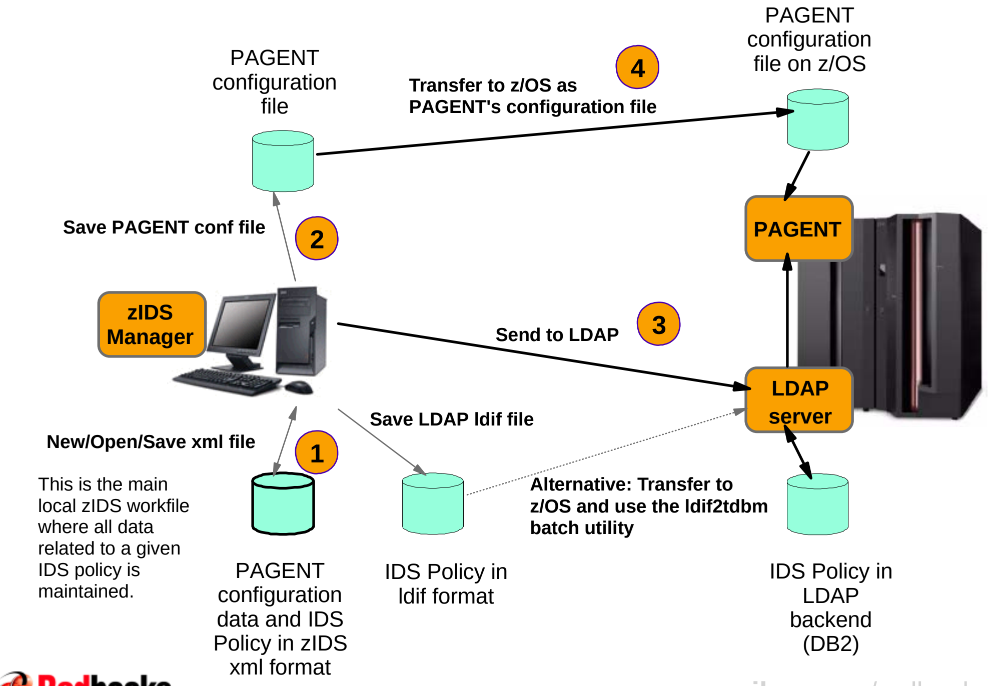
Policy-based IP networking on z/OS - component overview at z/OS V1R5 level


New in V1R5
Performance log file
Local Policies HFS file
Policies in an LDAP Server
zQos and zIDS Managers
New in V1R5
User performance monitor application
SSL
RAPI
PAPI
PAPI
Non-QoS aware appl.
QoS aware appl.
RSVP Agent
Policy Agent
pasearch command
TRM Daemon
SLA Subagent (pagtsnmp)
New in V1R5
SLAPM2 subagent (nslapm2)
Data Traffic
Data Traffic
RSVP Flows
Active Policies
Events, statistics
TCP/IP Kernel
Set TOS/DS, enforce TCP data rates and connection limits
Active QoS, Sysplex Distributor, and IDS policies
Token Bucket traffic enforcer
URI
Netstat command
Set TOS/DS
IP packets
Sysplex Distributor workload balancing
Fast Response Cache Accelerator (FRCA)
Intrusion Detection System (IDS)
Collect and maintain SLA and SLAPM2 performance metrics
1
2
3
4
Active Queue Management (QDIO)

Redbooks

ibm.com/redbooks

zIDS Manager and data locations


PAGENT configuration file
Transfer to z/OS as PAGENT's configuration file
4
PAGENT configuration file on z/OS
Save PAGENT conf file
2
PAGENT
zIDS Manager
Send to LDAP
3
LDAP server
Save LDAP ldif file
New/Open/Save xml file
1
Alternative: Transfer to z/OS and use the ldif2tdbm batch utility
PAGENT configuration data and IDS Policy in zIDS xml format
IDS Policy in ldif format
IDS Policy in LDAP backend (DB2)

This is the main local zIDS workfile where all data related to a given IDS policy is maintained.

Redbooks

ibm.com/redbooks

Must the LDAP server reside on z/OS?

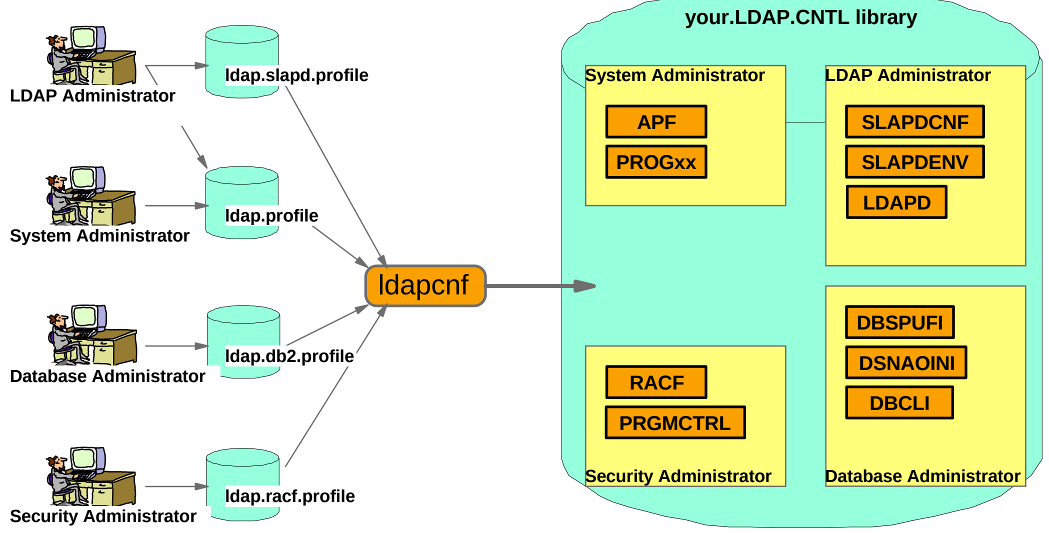
- The LDAP server does not have to reside on z/OS - PAGENT accesses the LDAP server via a TCP connection and identifies the LDAP server host via a hostname or IP address and port number.
- The LDAP server on z/OS requires a DB2 system to be available - all the LDAP data on z/OS is stored physically in DB2.
  - You do not need a separate DB2 system for LDAP - an existing DB2 system can be used. The LDAP server needs some table spaces and tables and a few other normal DB2-application type definitions to work with an existing DB2 system.
  - If you have a DB2 system already, then it is a simple and fast job to set up an LDAP server on z/OS - it won't take you more than about half an hour to collect the information you need, edit the setup files, and run the set up scripts and batch jobs.
  - If you do not have a DB2 system, but need to have a new DB2 system defined, you'd better team up with your friendly DB2 co-workers - setting up a new DB2 system isn't an easy task if you've never worked with DB2 before. (On the other hand if you know DB2, setting up another DB2 system isn't a big thing either)
- The LDAP server you use for your IDS policies (and QoS policies) can be shared with other LDAP applications and used for other LDAP purposes. You do not need a dedicated LDAP server for PAGENT's use

The easy way to set up an LDAP server on z/OS - Use the ldapcnf utility

1 Gather information

2 Generate customization jobs

3 Review and execute customization jobs



LDAP Administrator

System Administrator

Database Administrator

Security Administrator

ldap.slapd.profile

ldap.profile

ldap.db2.profile

ldap.racf.profile

ldapcnf

your.LDAP.CNTL library

System Administrator: APF, PROGxx

LDAP Administrator: SLAPDCNF, SLAPDENV, LDAPD

Security Administrator: RACF, PRGMCTRL

Database Administrator: DBSPUFI, DSNAOINI, DBCLI

Refer to "*z/OS LDAP Server Administration and Use*", SC24-5923

Protecting TCP/IP-related resources on your system through the SAF interface

- All the "traditional" SAF protection of data sets, authorized functions, etc. on a z/OS system applies to TCP/IP workload just as it applies to all other types of workload (Be careful with anonymous services such as anonymous FTP or TFTP services)

- The SERVAUTH resource class is used to specifically define and protect a number of TCP/IP unique resources

- General SERVAUTH profile format:

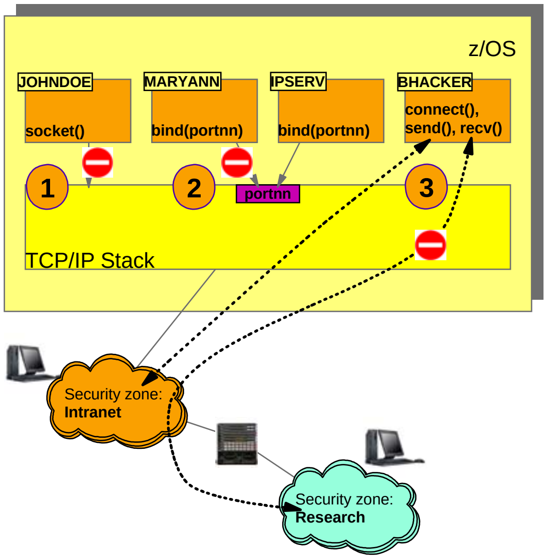
  - **`EZB.resource_category.system_name.jobname.resource_name`**
    - EZB designates that this is a TCP/IP profile
    - resource_category is capability area to be controlled e.g. TN3270, Stack Access, etc.
    - system_name is the name of the system - can be wildcarded
    - jobname is the jobname associated with the resource access request - can be wildcarded
    - optional resource_name - one or more qualifiers to indicate name of resource to be protected - can be wildcarded

- To protect one of the supported TCP/IP resources, you define a SERVAUTH profile with universal access NONE and you then permit users to have READ access to the resources

- If using OEM security packages, beware of the differences between defined/not defined resource actions

Redbooks ibm.com/redbooks

Protecting TCP/IP services, server port numbers, and network segments



SAF resources in the SERVAUTH resource class

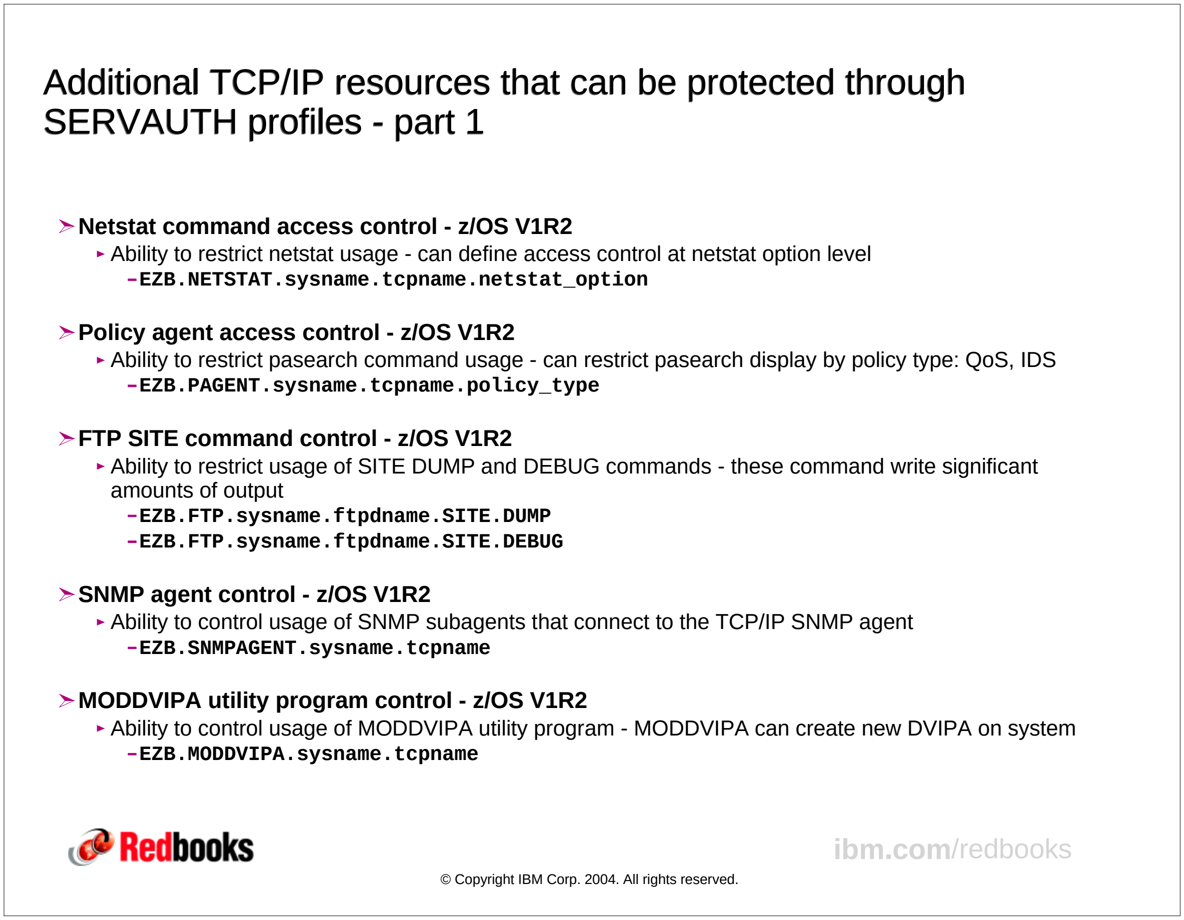
```
EZB.STACKACCESS.sysname.tcpname
EZB.PORTACCESS.sysname.tcpname.SAFkeyword
EZB.NETACCESS.sysname.tcpname.zonename
```

1. **Stack Access Control**
   - Controls user ability to open socket (use of TCP/IP services)
     - z/OS CS TCP/IP stack is considered a resource
     - Granularity is per stack in a multi-stack environment
   - Access to stack via sockets allowed if user permitted to SAF resource (SERVAUTH class: STACKACCESS)

2. **Local Port Access Control**
   - Controls user access to a local TCP or UDP port
     - Port is considered a resource
   - Function enabled
     - Via SAF Keyword on PORT or PORTRANGE
   - Access to port allowed if user permitted to SAF resource (SERVAUTH class: PORTACCESS)
   - Access to port not permitted for any user
     - Via RESERVED Keyword On PORT Or PORTRANGE

3. **Network Access Control**
   - Controls local user access to network resources
     - Network considered a resource - Network/Subnet/Specific host
   - Allows management of security zones
     - Via NETACCESS statement In TCP/IP Profile - NETACCESS statement allows grouping of network resources
   - Access to security zone allowed if user permitted to SAF resource (SERVAUTH class: NETACCESS)

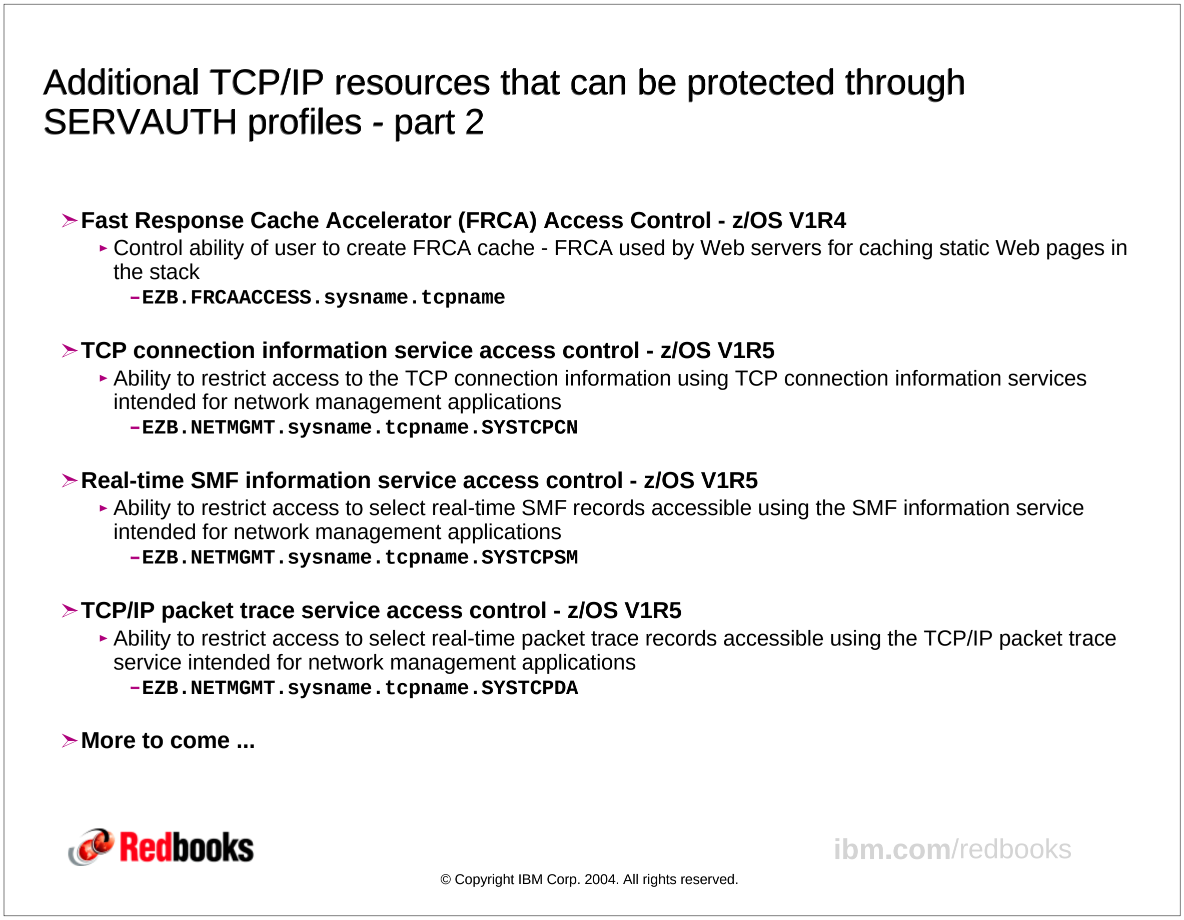
Redbooks ibm.com/redbooks



Additional TCP/IP resources that can be protected through SERVAUTH profiles - part 1

- **Netstat command access control - z/OS V1R2**
  - Ability to restrict netstat usage - can define access control at netstat option level
    - **EZB.NETSTAT.sysname.tcpname.netstat_option**
- **Policy agent access control - z/OS V1R2**
  - Ability to restrict pasearch command usage - can restrict pasearch display by policy type: QoS, IDS
    - **EZB.PAGENT.sysname.tcpname.policy_type**
- **FTP SITE command control - z/OS V1R2**
  - Ability to restrict usage of SITE DUMP and DEBUG commands - these command write significant amounts of output
    - **EZB.FTP.sysname.ftpdname.SITE.DUMP**
    - **EZB.FTP.sysname.ftpdname.SITE.DEBUG**
- **SNMP agent control - z/OS V1R2**
  - Ability to control usage of SNMP subagents that connect to the TCP/IP SNMP agent
    - **EZB.SNMPAGENT.sysname.tcpname**
- **MODDVIPA utility program control - z/OS V1R2**
  - Ability to control usage of MODDVIPA utility program - MODDVIPA can create new DVIPA on system
    - **EZB.MODDVIPA.sysname.tcpname**

ibm.com/redbooks



Additional TCP/IP resources that can be protected through SERVAUTH profiles - part 2

- **Fast Response Cache Accelerator (FRCA) Access Control - z/OS V1R4**
  - Control ability of user to create FRCA cache - FRCA used by Web servers for caching static Web pages in the stack
    - **EZB.FRCAACCESS.sysname.tcpname**
- **TCP connection information service access control - z/OS V1R5**
  - Ability to restrict access to the TCP connection information using TCP connection information services intended for network management applications
    - **EZB.NETMGMT.sysname.tcpname.SYSTCPCN**
- **Real-time SMF information service access control - z/OS V1R5**
  - Ability to restrict access to select real-time SMF records accessible using the SMF information service intended for network management applications
    - **EZB.NETMGMT.sysname.tcpname.SYSTCPSM**
- **TCP/IP packet trace service access control - z/OS V1R5**
  - Ability to restrict access to select real-time packet trace records accessible using the TCP/IP packet trace service intended for network management applications
    - **EZB.NETMGMT.sysname.tcpname.SYSTCPDA**
- **More to come ...**

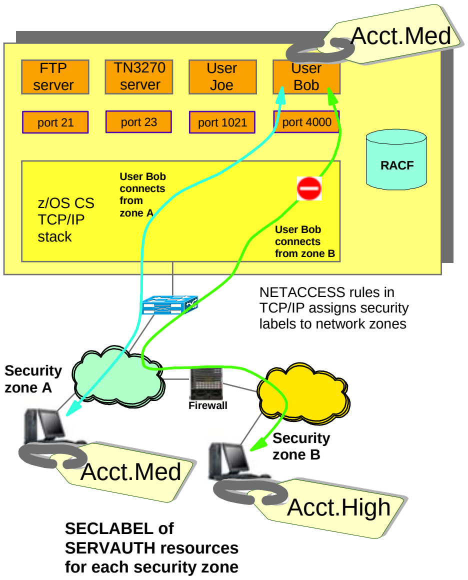
ibm.com/redbooks

Multilevel security

- Multilevel security is an enhanced security environment that can be configured on z/OS
  - Extends the B1 security support
  - IBM's MLS for z/OS has access control implications for the entire system.
    - See *z/OS Planning for Multilevel Security* GA22-7509 for system-wide MLS information
  - z/OS Communications Server TCP/IP is one element of a multilevel secure z/OS system
- Goal of MLS is to prevent declassification of data
  - All data and other resources are classified
  - All users are classified
- Classification is accomplished with security labels which combine
  - Security levels (hierarchical)
    - e.g. Top Secret, Internal Use Only, Unclassified
  - Security categories (non-hierarchical)
    - e.g. Accounting, Sales
- MLS adds a security policy check, Mandatory Access Control (MAC), to the usual Discretionary Access Control (DAC)
  - MAC ensures that data of a certain classification is accessed by a user with authority to access that classification
    - With MAC, the security administrator using RACF classifies the sensitivity and type of each resource using a security label and controls each user's access to the resource by assigning a security label to the user.
  - DAC ensures that data can be accessed only by a user permitted to access the data
    - With DAC, user-based permission to access resources
- MAC check is made prior to DAC check

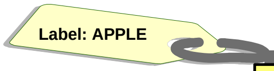
CS z/OS TCP/IP support of MLS

- **Phase 1**
  - Network Access Control outbound (OS/390 V2R10) and inbound (z/OS V1R4)
    - RACF SERVAUTH profiles define IP security zones that local users may/may not be permitted to send to/receive from over a socket (IP addresses are defined within a zone)
- **Phase 2 (z/OS V1R5)**
  - Extend Network Access Control to an MLS environment
    - When data is sent or received, the TCP/IP stack determines if the end user's security label matches the network zone security label (user SECLABEL compared to SERVAUTH zone SECLABEL)
    - Additionally, IP packets that are sent over XCF or SAMEHOST are labeled. If a label is present in an IP packet, that label is used to check against the user's label.
  - Netacess and MLS extended to applications
    - Port of entry (PoE) checks for the FTP server (MLS and NonMLS)
      - Checks ensure a client may log on to FTP
    - Netacess controls for TN3270 server ports (nonMLS)
      - Client Netacess permitted to TN3270 port zone
    - MLS controls determine if client has equivalent SECLABEL to the TN3270 LUname selected
- **Phase 3 (z/OS V1R6)**
  - Examine networking applications to determine if they will run successfully in an MLS environment.
  - Restrict IPv4 Setsockopt() source routing options
  - Documentation of MLS issues for applications/networking


Acct.Med
FTP server
TN3270 server
User Joe
User Bob
port 21
port 23
port 1021
port 4000
RACF
z/OS CS TCP/IP stack
User Bob connects from zone A
User Bob connects from zone B
NETACCESS rules in TCP/IP assigns security labels to network zones
Security zone A
Firewall
Security zone B
Acct.Med
Acct.High
SECLABEL of SERVAUTH resources for each security zone

Understanding basic MLS concepts

- **Security Level**
  - This term deals with the sensitivity of information and a person's clearance to it. Information is classified according to its sensitivity, such as CONFIDENTIAL or SECRET or TOPSECRET, etc.
  - Users are classified by their clearances.
- **Category**
  - This term is used to designate the department or type of information.
  - There might be a category for accounting, another for logistics and another for cryptographic methodology. There might be categories created for certain products or projects.
  - Categories are used to enforce broad "need to know" policies.
- **Security Label**
  - A security label (SECLABEL) is an eight character name. It represents a particular security level and a set of categories that are defined in the security server.
  - Port of Entry may be used to set or limit session security label at login.
  - Job card parameter SECLABEL= may be used to set job security label.
  - Both resources and users are assigned security labels.



Security Level	Category List
CONFIDENTIAL	ACCOUNTING, DEPT35

MLS security label - checking

- **Equal, Equivalent, Dominant and Disjoint SECLABELs**
  - These are the SECLABEL checks done to ensure that access to information is authorized.
    - **Disjoint:** If both SECLABELs have one or more categories that are not present in the other SECLABEL, they are said to be disjoint. No access is allowed.
    - **Dominate:** One SECLABEL is said to dominate another one, when its level is equal or higher and its categories are equal or a proper superset of the other.
    - **Equivalent:** Two SECLABELs are equivalent when their names are defined to have the same level and identical categories. Equivalent SECLABELs dominate each other.
    - **Equal:** Two SECLABELs are equal when they have the same name. Equal SECLABELs may be considered equivalent without asking the security server.
  - To READ data the user's SECLABEL must dominate the data SECLABEL.
  - To WRITE data the the user's SECLABEL must be dominated by the data SECLABEL.
  - To both READ and WRITE data, the user and data SECLABELs must be equivalent.
- **There are some predefined SECLABELs with special meanings:**
  - SYSLOW
    - This label is dominated by all other labels. It can be read by anyone. Distributed software is SYSLOW.
  - SYSHIGH
    - This label dominates all other labels. It can be written by anyone. System console and syslogd files are SYSHIGH. So are dumps and traces.
  - SYSNONE
    - This label is immune from SECLABEL checking. Useful for system catalog. (Intended for resources only, not users.)
  - SYSMULTI
    - This label is equivalent to all other labels. Given to authorized servers that run work securely on behalf of multiple users.

Data and user security labels (simplified!)

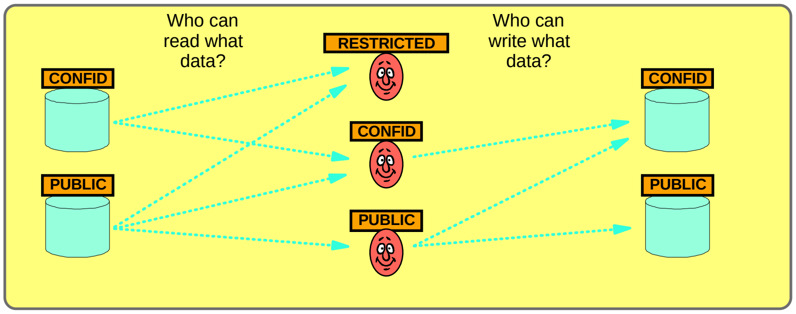
- A user can read a file that has a lower or equivalent security label
  - User security label dominates data security label

- A user can write (only) to a file that has an equivalent or higher security label
  - Data security label dominates user security label

- A user can read and write a file that has an equivalent security label

- If a user could write to a file with a lower security label, that user could in fact declassify data by reading data classified at that user's level and rewriting it at a lower level for other lower-level users to read it.

- Data can "flow" upward in the classification hierarchy; it cannot flow downward.
  - A CONFID user can read PUBLIC data and write it as CONFID or higher
  - A CONFID user can READ CONFID data, but cannot write it is as PUBLIC data



Redbooks V1R5 ibm.com/redbooks

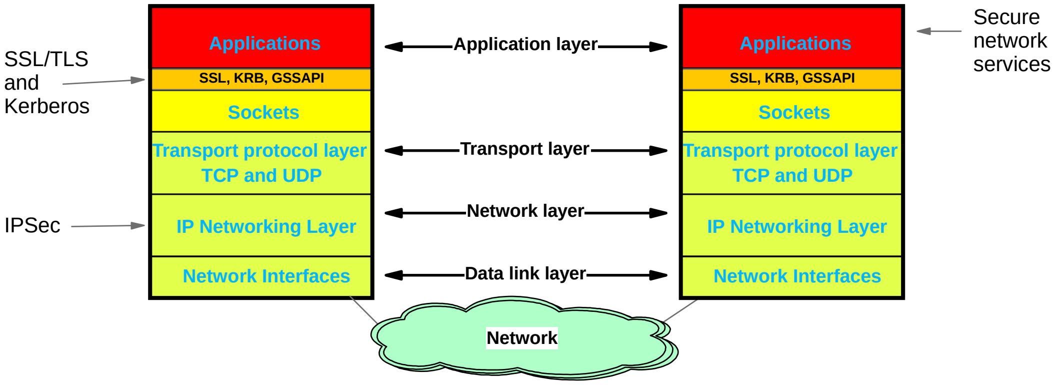
IBM®

z/OS Network Security
-
z/OS Communications Server Network Security

Redbooks

International Technical Support Organization

Workload-based security deployment



Transparent security for applications over an IP network

- IPSec provides blanket protection for all IP applications
  - End-to-end or segment of data path
- SSL/TLS TN3270 securely extends reach of SNA applications over an IP network
  - TN3270 client to TN3270 server data path secured
- SNA session level encryption secures SNA session end-to-end

Built-in application security

- Build security into application using sockets layer services (SSL/TLS, Kerberos)
- Build security into application at the message level (secure network services)

Redbooks

Encryption support by application

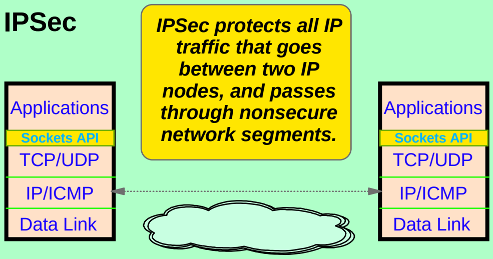
	IPSec	SSL/TLS	Kerberos	SNA Session level encryption	Secure network services
Where implemented:	Implemented in TCP/IP	Implemented in TCP application using socket layer services	Implemented in TCP application using socket layer services	Implemented in VTAM	Implemented in TCP or UDP application
SNA over IP using TN3270	**Yes, IP portion of data path**	**Yes, IP portion of data path (Note 1)**	No	**Yes, SNA portion of data path**	N/A
SNA over Enterprise Extender	**Yes, IP portion of data path**	N/A (EE uses UDP)	N/A (EE uses UDP)	**Yes, SNA portion of data path**	N/A
FTP	**Yes**	**Yes (Note 1)**	**Yes**	N/A	N/A
telnetd (UNIX)	**Yes**	No	**Yes**	N/A	N/A
rshd	**Yes**	No	**Yes**	N/A	N/A
Policy Agent	**Yes**	**Yes (LDAP)**	No	N/A	N/A
DCAS server	**Yes**	**Yes (Note 1)**	No	N/A	N/A
Sendmail	**Yes**	**Yes (new in z/OS V1R5)**	No	N/A	N/A
DNS	**Yes**	No	No	N/A	**Secure DNS**
SNMP	**Yes**	N/A	N/A	N/A	**SNMPv3**
OSPF	**Yes**	N/A	N/A	N/A	**OSPF MD5 authentication**
All other z/OS CS IP applications	**Yes**	No	No	N/A	N/A

Note 1: Required for advanced authentication and access control functions based on client's digital certificate

Redbooks

Network security - IPSec (VPNs) and SSL/TLS

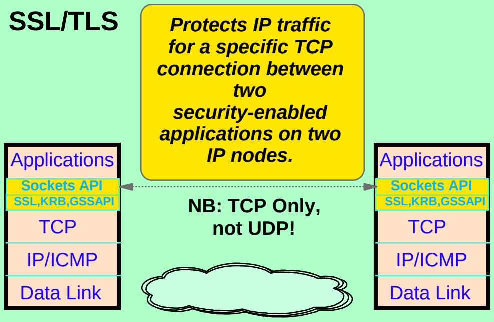
IPSec



IPSec protects all IP traffic that goes between two IP nodes, and passes through nonsecure network segments.

- **Provides authentication, integrity, and data privacy at IP layer**
  - **AH protocol provides authentication and integrity**
  - **ESP protocol provides data privacy (auth/integrity optional).**
  - **IKE protocol includes key exchange using public key cryptography and negotiation of security parameters**
  - **Management of crypto keys and SAs can also be manual**
- **IP node authentication, not user authentication**
- **Use of IPSec is transparent to upper layers including application**
  - **Blanket level protection for upper layer protocols**

SSL/TLS

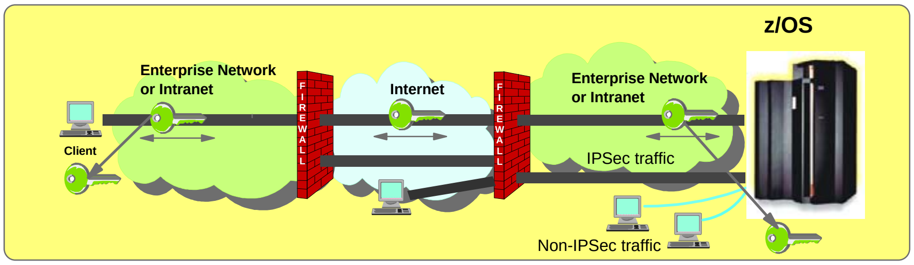


Protects IP traffic for a specific TCP connection between two security-enabled applications on two IP nodes.

NB: TCP Only, not UDP!

- **Provides authentication, integrity, and data privacy above TCP layer.**
  - **SSL handshake protocol includes key exchange using public key cryptography and negotiation of security parameters**
- **User authentication if client certificates used**
- **Applications must be changed to use SSL APIs**
  - **UDP applications cannot be SSL-enabled**
- **z/OS provides a System SSL library for z/OS applications**
  - **Available to LE C/C++ sockets programs only**

z/OS IPSec and VPN support



- Packaging (part of firewall technologies)
  - IKE daemon and configuration (Security Server)
  - TCP/IP IPSec and IKE support (Communications Server)
- Supports latest IETF standards
  - RFCs 2401-2406, 2409, and 2410
    - Maintains interoperability with previous IPSec RFC levels (1825-1829)
- Strong Crypto
  - Triple DES encryption
    - Exploits hardware S/390 cryptographic coprocessor
    - Exploits z990 CP assist instruction (new in z/OS V1R5)
  - Improved authentication algorithms (HMAC-MD5, HMAC-SHA)
- IKE Support
  - Authentication methods
    - Pre-shared key
    - RSA signature (uses X.509 certificates for host-based authentication)

Kerberos overview

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)

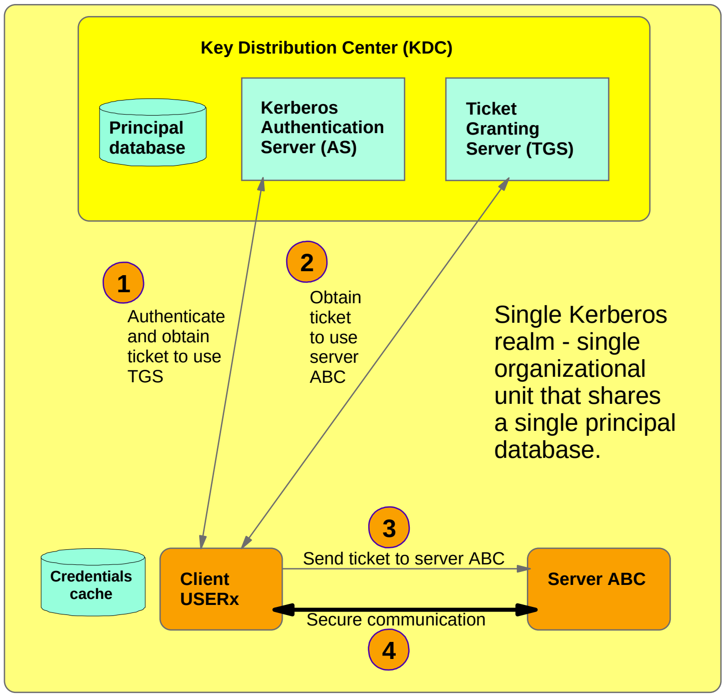
The Kerberos environment must be set up on the z/OS system.

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 infra structure already is in place


Key Distribution Center (KDC)
Principal database
Kerberos Authentication Server (AS)
Ticket Granting Server (TGS)
1
Authenticate and obtain ticket to use TGS
2
Obtain ticket to use server ABC
Credentials cache
Client USERx
3
Send ticket to server ABC
Server ABC
Secure communication
4

Single Kerberos realm - single organizational unit that shares a single principal database.

Redbooks

ibm.com/redbooks

Securing TN3270 access to z/OS

Protection of data in the network

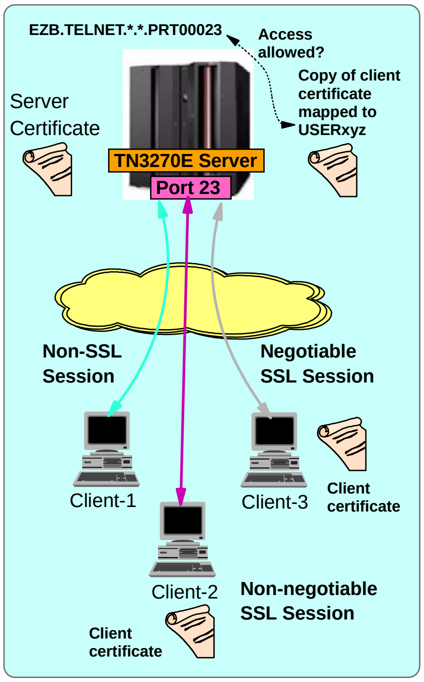
- Protect IP portion of data path - from client into z/OS TN3270 server

Optional levels of authentication based on TN3270 client certificate

- Authentication of client certificate
- SAF verification of userid associated with client certificate authorized to access TN3270 port
  - Uses SERVAUTH class - Profile EZB.TN3270.sysname.tcpname.portxxxx
- Use certificate for log on to SNA application without requiring userid/password
  - Express Logon Feature

Server configuration controls security policy for SSL session:

- The PARMSGROUP and PARMSMAP statements controls which clients must use SSL, which clients may use SSL, and which clients are not allowed to use SSL at all.
- Whether uncondtional or negotiable (or both) methods of TN3270 SSL are allowed
- Controls which ciphersuites are acceptable to TN3270 server.



OS/390 V2R6: SSL support with server certificates
OS/390 V2R8: Added support for client certificates
OS/390 V2R10: Added support for negotiable SSL
z/OS V1R4: Added TLS support

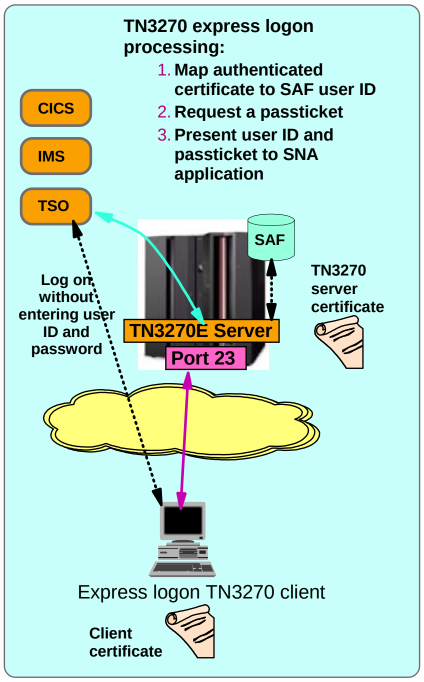
Redbooks

ibm.com/redbooks

Express logon to SNA application through TN3270 server

- **No change to SNA application required**
  - TN3270 Server uses z/OS passticket support for application transparency
    - Passticket one-time password
- **Uses Host On-Demand V5 or PCOMM V5.5 for login screen recognition and identification**
  - Administrator uses emulator's macro recording facility
    - Captures the interaction sequence and location on the panels of the userid and password
  - Emulator inserts symbolics in the inbound data stream to mark locations of userid and passticket
- **TN3270 Server gets user ID (based on client certificate and application ID) and passticket (based on user ID and application ID) using SAF**
  - TN3270 inserts user ID and passticket into inbound datastream in place of symbolics inserted during emulator's macro playback
- **A single authenticated client certificate can be used for user identification to multiple SNA applications**
  - TN3270 Server application-qualifies certificate during certificate to user ID mapping
    - Application ID provided by TN3270 client

Simplifies user ID and password administration

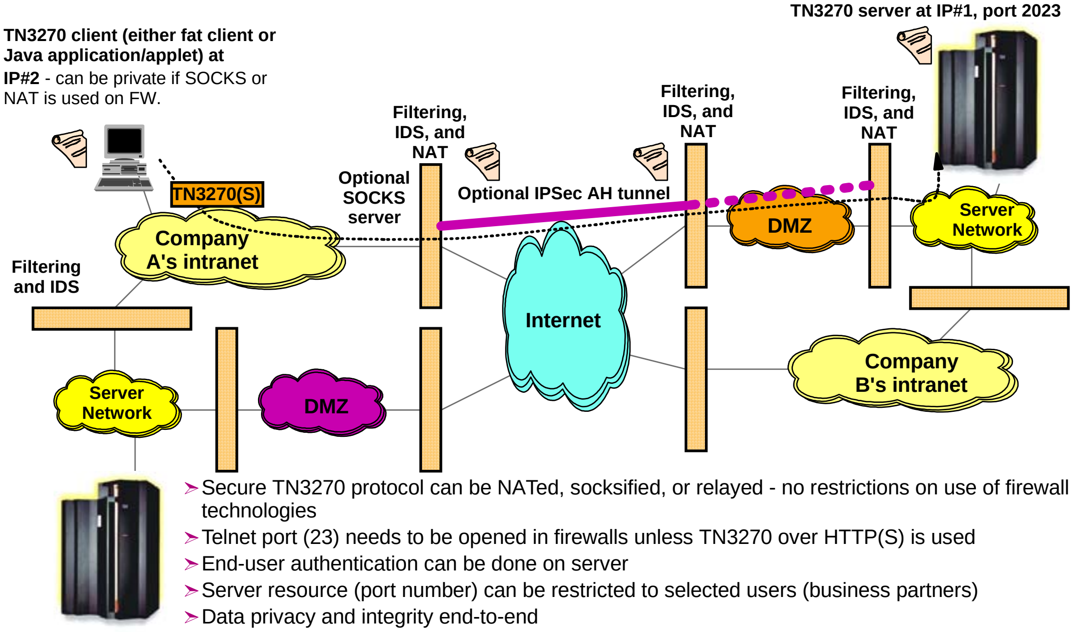


TN3270 express logon processing:

1. Map authenticated certificate to SAF user ID
2. Request a passticket
3. Present user ID and passticket to SNA application

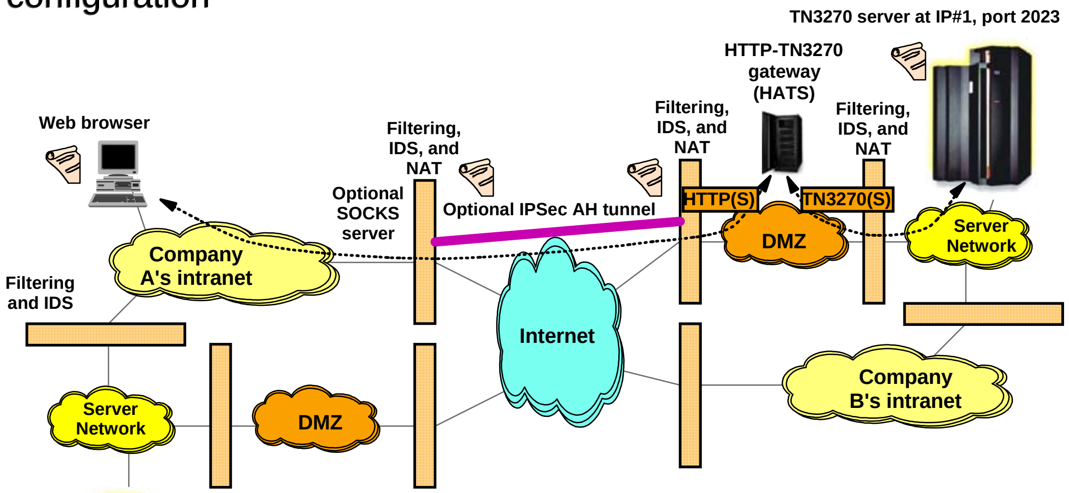
CICS
IMS
TSO
SAF
Log on without entering user ID and password
TN3270 server certificate
TN3270E Server
Port 23
Express logon TN3270 client
Client certificate

Business partners - TN3270 access sample configuration



- Secure TN3270 protocol can be NATed, socksified, or relayed - no restrictions on use of firewall technologies
- Telnet port (23) needs to be opened in firewalls unless TN3270 over HTTP(S) is used
- End-user authentication can be done on server
- Server resource (port number) can be restricted to selected users (business partners)
- Data privacy and integrity end-to-end
- IPSec AH tunnel between firewalls allows FWs to authenticate with each other and reject traffic to TN3270 server from unauthorized partners (even if they fake their source IP addresses)

Business partners - TN3270 over HTTP(S) access sample configuration

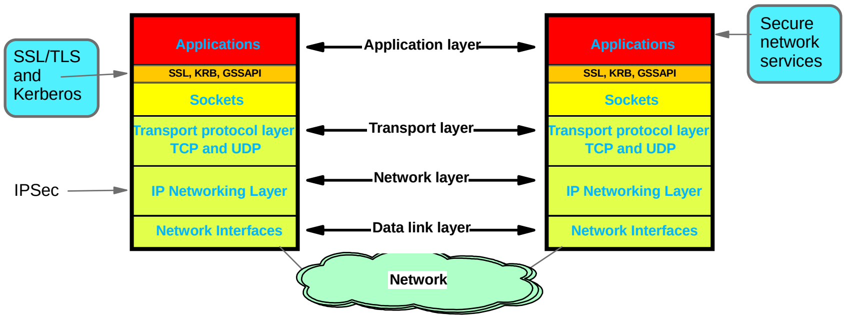


- If firewalls cannot be opened for any other protocol than HTTP and HTTPS, there are today solutions that allow a Web browser to communicate with a Web server that acts as a gateway between HTTP(S) and TN3270(S)
- IBM's Host Access Transformation Server (HATS) provides one such solution

Redbooks

ibm.com/redbooks

Built-in application security



Sockets-based security (SSL/TLS and Kerberos)

- TLS Enabled FTP server and client
  - Secure file transfer by providing encryption, authentication, and message integrity for the FTP control and data connections
    - Strong authentication using X.509 Certificates (including client authentication)
- Application kerberization
  - FTP server and client, UNIX telnet daemon, UNIX rsh daemon
    - Strong 3rd party authentication for client/server applications using secret key cryptography, encrypted data flows

Secure network services

- SNMPv3
  - Provides authentication, data integrity and privacy services for SNMP messages
  - Provides access control (read/write) for MIB objects
- Secure DNS
  - Ensures DNS query replies are authentic
- OSPF MD5 authentication
  - Ensures routing table integrity
  - Uses MD5 authentication for routing messages (RFC 2328)

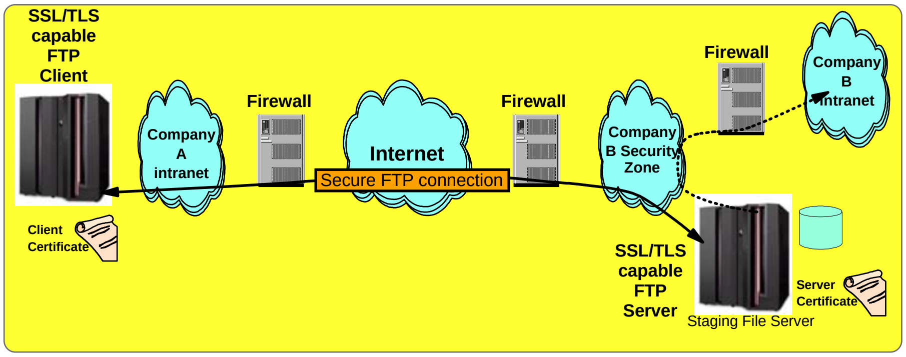
Redbooks

ibm.com/redbooks

Securing file transfers to/from z/OS with File Transfer Protocol (FTP)

Key features

- Protection of data in the network
  - Data origin authentication, data integrity, and privacy
- Server and client configuration provided to control security policy for TLS session
  - TLS negotiation will find an agreeable policy or TLS setup will not proceed.
- PKI based authentication of end users added. Options are:
  - Basic - userid/password
  - TLS authentication of client certificate
  - Cross-checking of userid entered with userid associated with client certificate
  - Use certificate for logon without requiring userid/password. New in z/OS V1R5
- Firewall Coexistence
  - TLS NAT Traversal support added in z/OS V1R5 (PTFed back to z/OS V1R4)



SSL/TLS for FTP-based file transfer operations

- **Please note that secure FTP based on SSL/TLS (sometimes referred to as *ftps*) has nothing to do with what is known as *sftp***
  - sftp is a file transfer protocol under the umbrella of SSH (Secure Shell)
  - sftp has absolutely nothing to do with the normal FTP standards as defined in RFC 959
  - The sftp protocol is its own protocol and sftp under SSH in no way interoperates with normal FTP
    - An sftp client cannot exchange files with a normal FTP server
- **SSL/TLS enabled FTP is based on IETF standards and draft standards**
  - z/OS FTP client and server do interoperate with other vendors' secure FTP implementations
- **FTP sessions can be made secure using one of two modes:**
  - Implicit mode
    - A separate FTP port number that is designated as being secure (control connection port number 989)
    - Implicit mode is not recommended for FTP according to the latest IETF standards, but is widely supported
  - Negotiated mode
    - Uses the standard FTP port number (control connection port number 21)
    - Use of SSL/TLS is negotiated per FTP session using a recent FTP protocol extension tha is known as the AUTH command
      - The AUTH command is used to negotiate both SSL/TLS-based and Kerberos-based security for FTP sessions
    - z/OS FTP server options instruct the server what to accept from the client
      - SECURE_FTP is either required or allowed by the server
    - z/OS FTP client is instructed which type of security to request through user interaction or client configuration options
- **When using SSL/TLS with FTP, the control connection is always secured**
  - Security for the data connection is optional and negotiated by the client
    - The z/OS FTP server options instruct the server what to accept from the client
      - SECURE_DATACONN is either never, clear, or private

z/OS CS security aspects summary

Protecting system resources and data from the network

- **Integrated Intrusion Detections Services**
  - Detects, records, and defends against scans, stack attacks, flooding
- **Protect system availability**
  - Built-in protection against Denial of Service attacks
  - IP packet filtering
  - Syslogd integrity and availability
  - Sysplex Wide Security Associations
- **SAF protection of z/OS resources**
  - z/OS CS application access to data sets and files
  - SERVAUTH class protection
    - Ex: Local user access to TCP/IP system, TCP and UDP ports
  - Multilevel security

Protecting mission-critical data in the network

- **Strong encryption with triple DES**
  - Using hardware assist from crypto coprocessor and CP assist instruction
- **Transparent application security**
  - IPSec for TCP/IP applications
  - Internet-ready access to SNA applications with TN3270 SSL
  - Express Logon Feature
- **Built-in application security**
  - SSL-enabled FTP, Kerberized FTP, rsh, telnet
- **Secure network services**
  - SNMPv3, Secure OSPF Authentication, Secure DNS

This page intentionally left blank

ibm.com

SNA/IP Integration Using Linux on zSeries

Redbooks

International Technical Support Organization

Objectives

The objectives of this session are:

- Review the general concepts of SNA/IP integration technologies
- Position Linux on zSeries as an element in an SNA/IP integration strategy
- Introduce the Communications Server for Linux on zSeries product
- Preview the Communication Controller for Linux plans

Agenda

1. Introducing general SNA/IP integration concepts and technologies
2. Communications Server for Linux on zSeries
3. Preview of Communication Controller for Linux on zSeries

IBM Statement of Direction update on SNA support in 2004

It is IBM's intent to support VTAM in z/OS Communications Server for the foreseeable future. Customers have a substantial investment in 3270 and SNA applications. We continue to support and enhance VTAM's capabilities while integrating it with new technologies. IBM has no plans at this time to discontinue SNA support in z/OS Communications Server. As of June 2004, customers can, for selected SNA workloads, use Communications Server products for Linux, Linux on IBM eServer zSeries, Microsoft Windows, and AIX to replace some of the old SNA infrastructure components, such as the IBM 3745/46 or other channel-attached SNA controllers. z/OS Communications Server can replace some SNA Network Interconnect (SNI) workloads using Enterprise Extender and Extended Border Node functions.

It is IBM's intent to introduce an additional solution in 2005 that uses NCP (Network Control Program) software running within Linux on zSeries. The intent is to provide a migration path for customers who use traditional SNA (including SNA Network Interconnect (SNI)) to communicate with their business partners. This solution can allow them to continue using traditional SNA without a dependency on IBM 3745 and 3746 Communication Controller hardware.

IBM®

SNA/IP Integration Using Linux on zSeries - Introducing General SNA/IP Integration Concepts and Technologies

Redbooks

International Technical Support Organization

SNA/IP integration strategy objectives

- ✓ **Preserve investment in the SNA application portfolio for the "natural" lifetime of those SNA applications.**
  - ➢ Preserve the ability to access those SNA applications using traditional end-user technologies such as an IBM 3270 terminal interface or SNA-based client/server program-to-program communication.
  - ➢ Enable reuse of those same SNA applications from an emerging e-business environment through various forms of Web-enabling technologies.
    - User interface transformation through technologies such as Host Access Transformation Services.
    - e-business application integration through various forms of WebSphere Application Server connector technologies.
- ✓ **Help reduce cost of owning and operating an enterprise networking infrastructure.**
  - ➢ Remove business dependency on SNA networking technology that is no longer strategic or is nearing end of life.
    - IBM 3745/46 Communication Controller (no longer marketed by IBM)
    - Token-ring technology (products rapidly being withdrawn in general)
    - ESCON channel-attached SNA controllers of various types (ESCON channel chips no longer manufactured)
    - AnyNet technology (z/OS V1R7 is last z/OS release to support AnyNet on z/OS)
    - OS/2 (End of Service announced for 2006)
  - ➢ Help reduce software licenses and maintenance costs associated with multiprotocol wide area networking and related management software.
  - ➢ Reduce dependency on SNA wide area network technology skills.
  - ➢ Be able to move forward with SNA to IP migration without dependency on business partner progress or lack thereof with respect to SNA to IP migration activities.
- ✓ **Help improve return on investments in the enterprise networking infrastructure by consolidating all wide area network traffic to an IP-based networking technology.**
  - ➢ Focus on establishing a highly available, scalable, and secure IP-based networking infrastructure.
  - ➢ Reduce overall networking infrastructure complexity.
  - ➢ Consolidate the SNA networking environment into the zSeries box(es) or as a minimum into the data center.

Redbooks

ibm.com/redbooks

SNA/IP integration elements - a multi-step approach

CS Linux can help

- Consolidate intranet **SNA 3270** traffic (LU1/SCS, LU2, LU3/DSC) into the data center:
  - Using TN3270 client software (PCOMM, HOD, OEM) on the user workstation connecting to a TN3270 server in the data center, which could be z/OS or 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
- Move **middleware** communication off SNA where applicable. DB2 DRDA, MQ, etc. can be migrated to native IP communication without impact on database or messaging applications.

CS Linux can help

- For remaining **SNA client/server applications** in the branches/remote locations (LU0, LU6.2), use one of the following technologies to transport the SNA data over an IP network:
  - Enterprise Extender to transport native SNA flows over an IP network from the branch and into the data center. EE can in the branch be deployed on the workstation or on an EE gateway in the branch. In the data center, EE can be deployed on an EE gateway, such as CS Linux, or on z/OS itself.
  - Use a remote SNA API technology to ship Windows and Linux SNA application calls over an IP network to an SNA API server running on CS Linux in the data center.

CCL can help

- Move IBM 3745/46-based **business partner communication** to Enterprise Extender technology or to the new IBM Communication Controller for Linux on zSeries (SOD: 2005).

CS Linux can help

- Migrate OS/2-based SNA branch server applications to Linux and use either the EE technology or the remote API technology to traverse the IP wide area network.

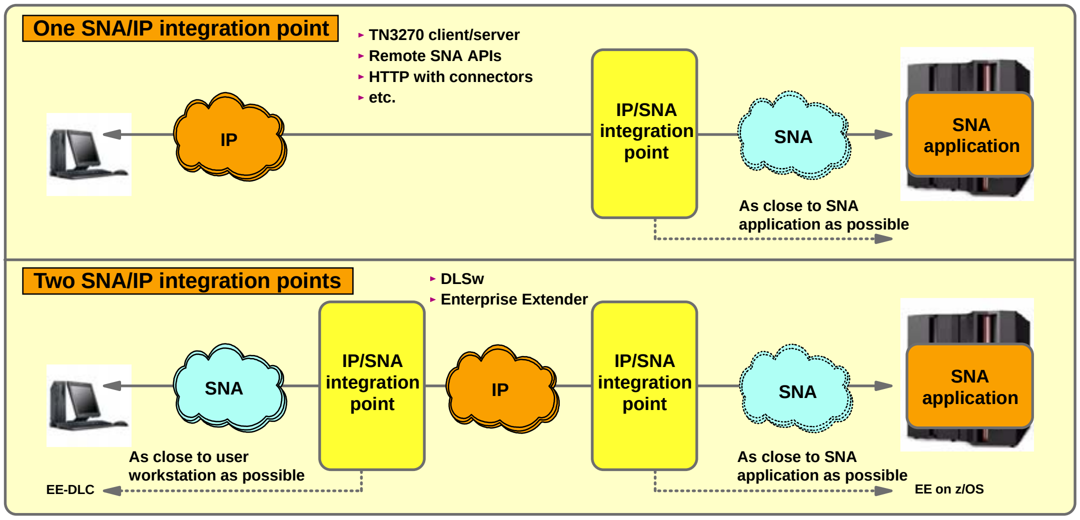
CCL can help

- To the extent supported, move remaining native SNA communication coming in through an IBM 3745/46 to the new IBM Communication Controller for Linux on zSeries using the SNA boundary functions.

Redbooks

ibm.com/redbooks

SNA/IP integration without changing SNA applications - a transformation is needed somewhere!



One SNA/IP integration point

- TN3270 client/server
- Remote SNA APIs
- HTTP with connectors
- etc.

IP

IP/SNA integration point

SNA

SNA application

As close to SNA application as possible

Two SNA/IP integration points

- DLSw
- Enterprise Extender

SNA

IP/SNA integration point

IP

IP/SNA integration point

SNA

SNA application

As close to user workstation as possible

As close to SNA application as possible

EE-DLC

EE on z/OS

Redbooks

ibm.com/redbooks

Two basic models for SNA/IP integration point on zSeries

Linux on zSeries SNA/IP integration technologies:

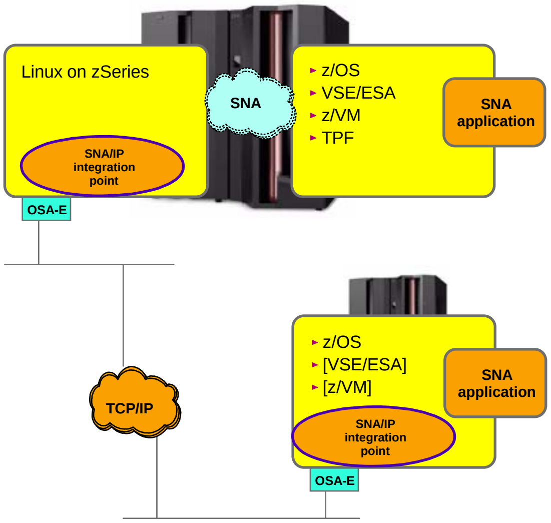
- ✓ Communications Server for Linux (CS Linux) on zSeries ***(available now!)***
- ✓ Communication Controller for Linux (CCL) on zSeries ***(planned availability 2005!)***
- ✓ The WebSphere Application Server environment ***(available now!)***
  - Host Access Transformation Services
  - Connectors

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

Collapse the SNA network into zSeries using existing mainframe OS capabilities - or Linux on zSeries as the SNA/IP integration point.



Redbooks

ibm.com/redbooks

The z/OS environment

Why place the SNA/IP integration point on z/OS?

- ✓ **Functions**
  - Functionally most rich TN3270 server on the market
  - Supports both SNA subarea-based (SNI) and EE-based (EBN) business partner connectivity
- ✓ **Performance**
  - No extra "hops"
  - IP traffic over high-speed OSA-Express interfaces
  - High-performance same-OS interfaces between the IP and SNA side of the integration point
- ✓ **Capacity/Scalability/Availability**
  - Both the SNA and TCP/IP environments on z/OS are based on and use the z/OS Sysplex technologies
  - A single TN3270 server can service up to 128,000 TN3270 clients
- ✓ **Simplification**
  - All functions are incorporated into one single operating system environment
  - The SNA network is collapsed to reside inside the z/OS Sysplex environment or the z/OS systems in the data center(s)
- ✓ **Disaster recovery**
  - Disaster recovery planning from a network connectivity point of view is significantly simplified if all network access to z/OS is based on TCP/IP - dynamic movement of IP addresses to move functions from one site to another site
- ✓ **Security**
  - IP to z/OS allows use of IP-based security functions that will secure network data all the way into z/OS

If you have a choice, choose z/OS!

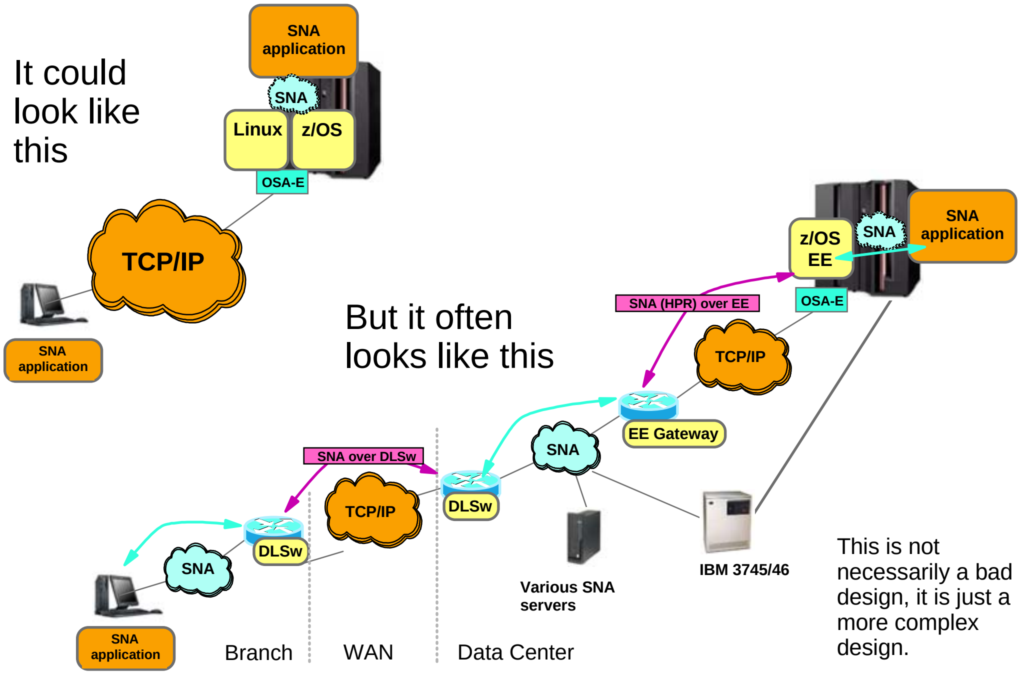
Why not place the SNA/IP integration point on z/OS?

- ✗ **Lack of z/OS SNA skills**
  - EE requires z/OS to be APPN-enabled
  - TN3270 server on z/OS requires different z/OS SNA configuration definitions than non-z/OS resident TN3270 servers do
  - EE business partner connectivity requires APPN multiple network connectivity (EBN and session management)
- ✗ **Lack of z/OS TCP/IP skills**
  - An SNA/IP integration point on z/OS requires TCP/IP to be functional on z/OS
- ✗ **Cost of z/OS MIPS (CP and software charges) versus IFL engines**

Redbooks

ibm.com/redbooks

Potential for real infrastructure simplification


It could look like this

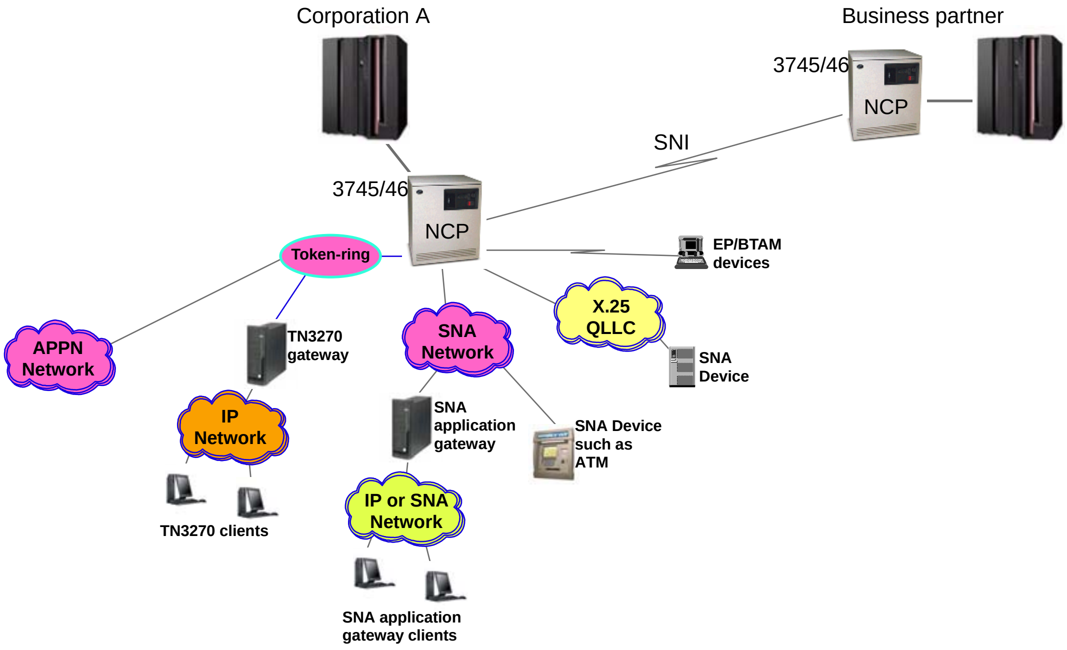
But it often looks like this

This is not necessarily a bad design, it is just a more complex design.

Redbooks

ibm.com/redbooks

Traditional SNA wide area networking infrastructure



Redbooks

ibm.com/redbooks

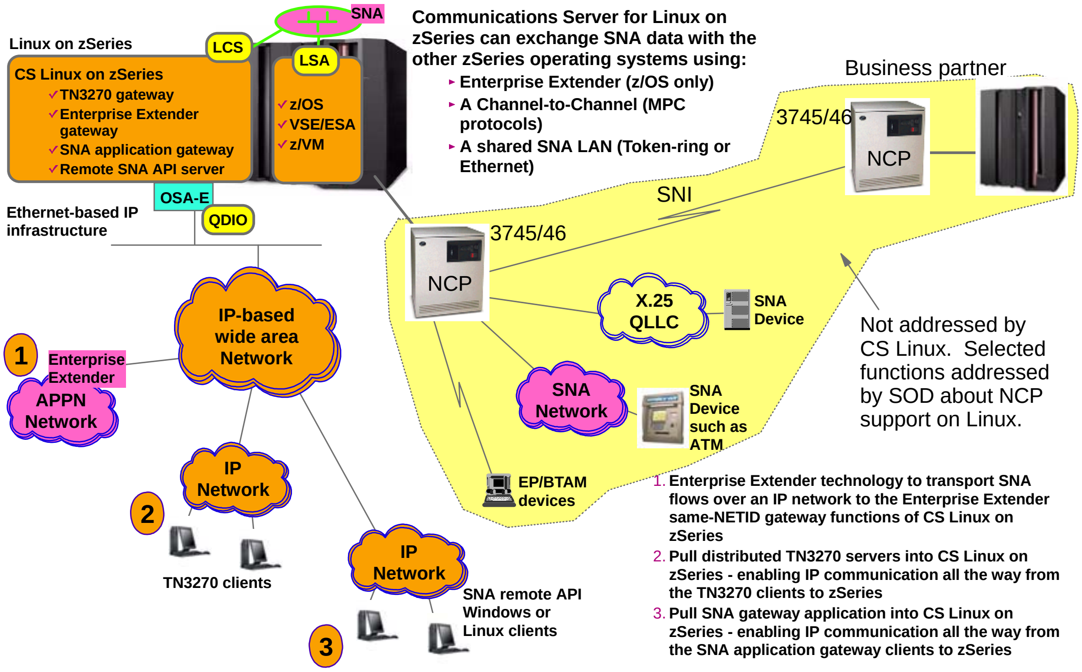
IBM®

SNA/IP Integration Using Linux on zSeries
-
Communications Server for Linux on zSeries

Redbooks

International Technical Support Organization

(May 2004): Communications Server for Linux on zSeries - available now - reduces IBM 3745/46 dependency



Communications Server for Linux on zSeries can exchange SNA data with the other zSeries operating systems using:

- **Enterprise Extender (z/OS only)**
- **A Channel-to-Channel (MPC protocols)**
- **A shared SNA LAN (Token-ring or Ethernet)**

Not addressed by CS Linux. Selected functions addressed by SOD about NCP support on Linux.

1. **Enterprise Extender technology to transport SNA flows over an IP network to the Enterprise Extender same-NETID gateway functions of CS Linux on zSeries**
2. **Pull distributed TN3270 servers into CS Linux on zSeries - enabling IP communication all the way from the TN3270 clients to zSeries**
3. **Pull SNA gateway application into CS Linux on zSeries - enabling IP communication all the way from the SNA application gateway clients to zSeries**

Redbooks

ibm.com/redbooks

IBM Communications Server for Linux on zSeries - 5734-I34

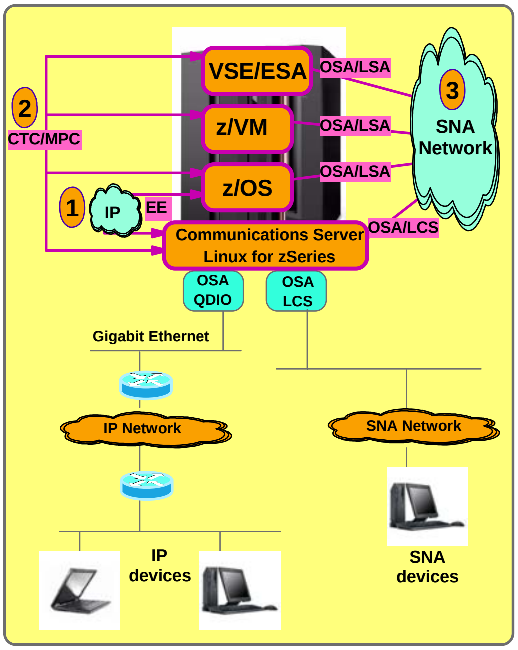
- **Advanced Peer-to-Peer Networking (APPN) support**
  - APPN End Node (EN) or APPN Network Node (NN) support
  - Uses Dependent LU Requester (DLUR) for dependent LU access over an APPN network
- **High Performance Routing (HPR) including Enterprise Extender (EE - HPR over IP)**
- **Branch Extender (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)
  - Remote SNA client/server APIs
  - APPC application suite (AFTP, APING, AREXEC, ATELL, ACOPY, 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
  - Currently runs on the 2.4 Linux kernel
- **Network attachments for SNA**
  - Enterprise Extender (HPR over IP)
  - (V)CTC using MPC channel protocols (Linux as a PUT2.1)
  - Native SNA over shared LAN (Ethernet or Token-Ring)

Available since May 2004

CS Linux is also offered in an Intel Linux version: 5724-I33.

Most of the CS Linux functions are also offered in the other distributed Communications Servers from IBM: CS/Windows and CS/AIX

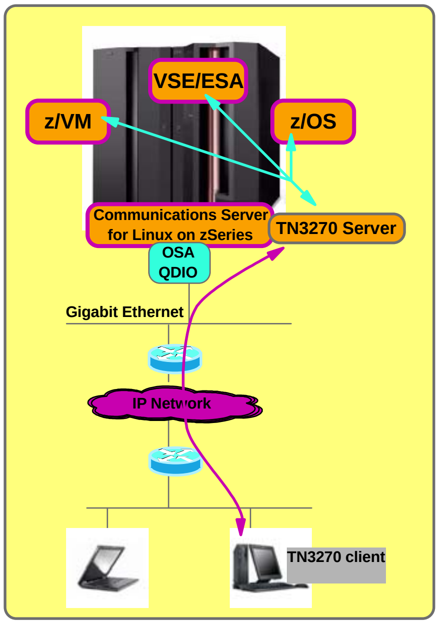
CS Linux on zSeries - overview of options for SNA connectivity to z/OS, z/VM, and VSE/ESA



- SNA between Linux on zSeries and other zSeries operating systems:

1. **Enterprise Extender (HPR over IP)**
   - For upstream to z/OS only
   - Dependent LUs: Linux DLUR - z/OS DLUS
   - Can use any IP-based connectivity between Linux and z/OS - including HiperSockets
2. **APPN Host to Host (AHHC/ANNC over MPC)**
   - Connectivity: CTC MPC channel
   - For upstream to z/OS, z/VM, and VSE/ESA
   - Both endpoints must be defined as PU Type 2.1 nodes - may mean APPN-enabling z/OS, z/VM, and VSE/ESA, if not already done (z/VM and VSE/ESA as APPN NNs)
   - Dependent LUs: Linux DLUR - z/OS, z/VM, and VSE/ESA DLUS
3. **SNA LAN (APPN, LEN, or Peripheral)**
   - Connectivity: Linux OSA LCS via shared LAN to OSA LSA
   - For upstream to z/OS, z/VM, and VSE/ESA
   - Linux attachment via LCS device driver and enhanced OSA Express microcode (zSeries only)
   - z/OS, z/VM, and VSE/ESA attachment via standard OSA LSA device driver
   - PUs may be PU Type 2.0 or 2.1

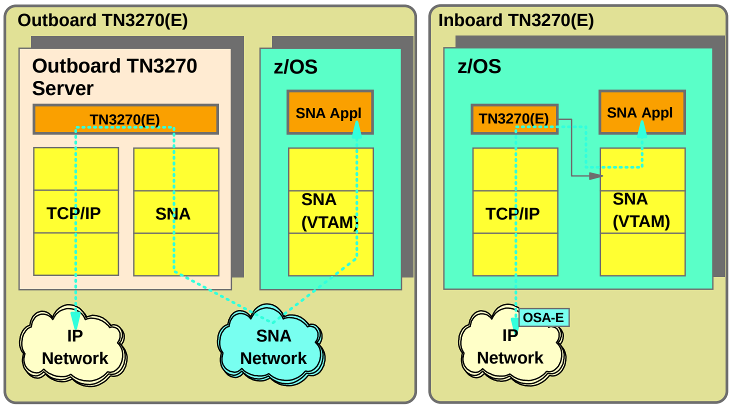
Consolidating existing distributed TN3270 Servers into Linux on zSeries



- 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 on 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
  - The TN3270 servers in the two products are not functionally equivalent

Redbooks ibm.com/redbooks

Does it really matter where I place my TN3270 server?



Application Workload Modeler (5655-J62) is an IBM product that can be used to perform in-house detailed benchmarks of both IP and SNA-based solutions.

AWM provides an echo SNA application (awmecho) that can be used as the SNA application when doing TN3270 benchmarks.

✓ Performance
✓ Scalability
✓ Functions
✓ Availability
✓ Manageability
✓ Cost
✓ Security

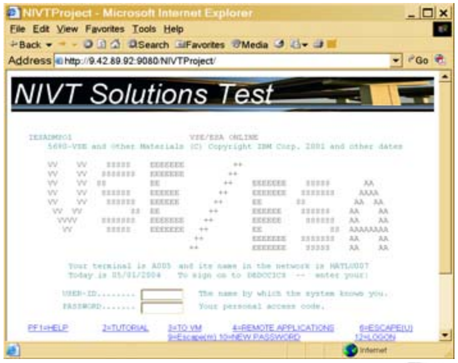
- Application Workload Modeler can be used to simulate TN3270(E) client activity
  - Capable of simulating a large number of TN3270(E) clients/activity
    - Used internally to benchmark 60,000 client sessions to a single z/OS TN3270 server
  - Can be used to benchmark outboard vs. inboard TN3270(E) server solution
    - Determine cost/performance/scalability characteristics of each solution - including use of SSL/TLS
  - Aids in decision-making process for TN3270(E) server placement
    - Functions, availability characteristics, management capabilities

Redbooks ibm.com/redbooks

TN3270 server on CS Linux on zSeries or on z/OS?

Area of interest	TN3270 server in CS Linux on zSeries	TN3270 server in z/OS
Secure TN3270 support	Yes (SSL only) - including client authentication (signature verification)	Yes (SSL and TLS) - including client authentication (signature verification) with optional SAF authentication and port protection (SERVAUTH)
Express logon	Yes (TCP-SSL connection with a z/OS DCAS server)	Yes (direct SAF interaction)
Support for RFC2355E contention resolution (important for both HOD and PCOMM)	Yes	Yes
zSeries hardware crypto exploitation	No	Yes
LU name assignment (LU name nailing)	Client IP address, client host name	Client IP address (including ranges), client host name, MVS user ID, server IP address, server interface name
Real or placeholder LU name assigned	Placeholder LU name (the locally defined name)	The real LU name
Printer association support	Yes	Yes
Specific LU requests	Yes	Yes
ANS=CONT support	Yes	No
USS table support	N/A (dependent LUs - done by the VTAM SSCP and controlled via VTAMLST definitions)	Yes - TN3270 server reuses VTAM USS table definitions (z/OS V1R6 adds support for SCS mode USS)
Selecting SNA application	N/A (dependent LUs - done by the VTAM SSCP and controlled via VTAMLST definitions)	Yes - LOGAPPL and QINIT support
Definitions	LU definitions on Linux and in VTAMLST (one PU per 255 LUs)	LU definitions in z/OS TN3270 server and VTAMLST (ACBs - cloning supported)
Capacity	Testing with 20,000 concurrent sessions has been done.	60,000 concurrent sessions tested - theoretical limit of 128,000 sessions
Load balancing	Traditional connection balancing	Traditional connection balancing. Sysplex Distributor adds value in terms of real-time LPAR capacity and server availability.

3270: one step further - CS for Linux on zSeries and IBM's Host Access Transformation Services


NIVTProject - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites Media
Address http://9.42.89.92:9080/NIVTProject/
Go
NIVT Solutions Test
Internet

```
IESADMP01                    VSE/ESA ONLINE
   5690-VSE and Other Materials (C) Copyright IBM Corp. 2001 and other dates

      Your terminal is A005 and its name in the network is NATLU007
      Today is 05/01/2004   To sign on to DBDCCICS -- enter your:

     USER-ID........         The name by which the system knows you.
     PASSWORD.......         Your personal access code.

PF1=HELP    2=TUTORIAL   3=TO VM    4=REMOTE APPLICATIONS    6=ESCAPE(U)
                         9=Escape(m) 10=NEW PASSWORD          12=LOGON
```

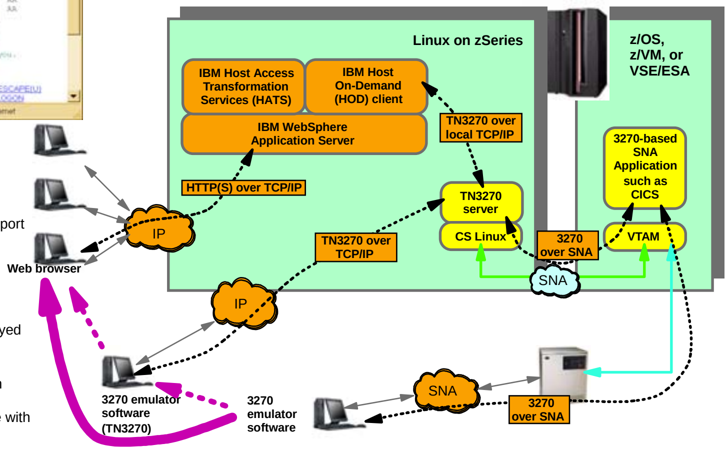
- ➢Universal workstation client: Web browser
  - Basically all platforms with a Web browser are supported
- ➢No 3270 emulator software (fat client or down loaded) on workstation
- ➢Only HTTP/HTTPS over IP between workstation and WebSphere Application Server
  - Simplifies firewall setup
- ➢No changes 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

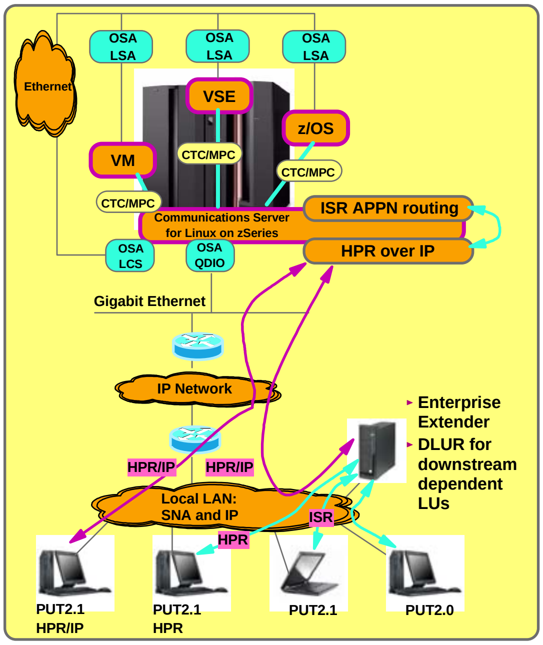
Scalability
- Vertical scaling through zSeries 64-bit storage support and powerful parallel CPU engines
- Horizontal scaling through z/VM technologies

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



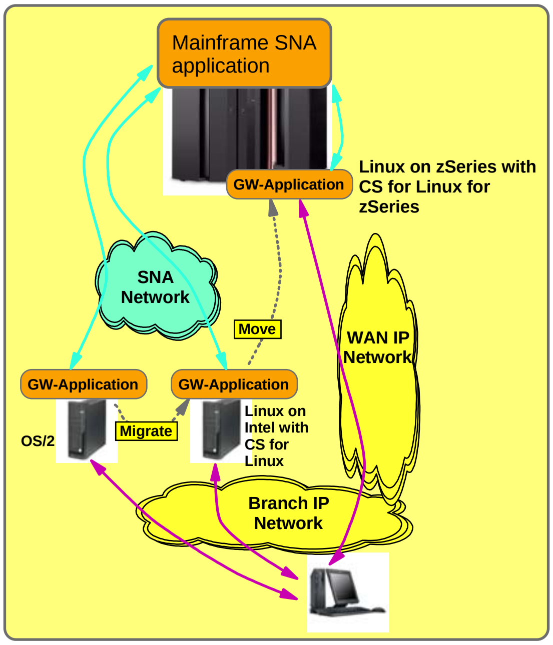
Linux on zSeries as same NETID Enterprise Extender gateway to z/OS, z/VM, and VSE/ESA



- Linux on zSeries can act as an Enterprise Extender gateway to other zSeries operating systems that do not support, or are not configured for EE
- SNA traffic reduced to branch LAN and zSeries
  - Reduced dependency on IBM 3745/46, CIP, or token-ring hardware
  - Common WAN IP infrastructure
- Network connectivity to zSeries via Gigabit Ethernet and QDIO
- For full set of functions, zSeries operating systems must be APPN enabled
  - ISR APPN routing over:
    - A CTC MPC link
    - OSA LCS via shared LAN to OSA LSA
  - HPR routing over IP (to z/OS only)
- LEN connectivity over a shared LAN provides some limited capabilities (predefined resources)
- All immediate downstream and upstream Network Nodes must be within the same SNA NETID
  - CS Linux does not support APPN boundary functions or session services extensions

Customer-written SNA application gateway programs on OS/2

- Customers migrating off OS/2 can take advantage of the CS Linux SNA API capabilities


Mainframe SNA application
GW-Application
Linux on zSeries with CS for Linux for zSeries
SNA Network
Move
WAN IP Network
GW-Application
GW-Application
OS/2
Migrate
Linux on Intel with CS for Linux
Branch IP Network

Example:

A customer-specific application that runs on OS/2 and acts as a gateway between users in a branch and SNA applications on the mainframe:

- Upstream: SNA LU0 or LU6.2
- Downstream: TCP/IP

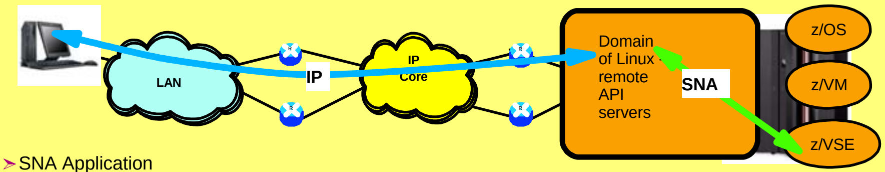
The first step is to rewrite the application to run on Linux instead of OS/2 - potentially deploying the gateway application initially on Linux for Intel in the branch. The SNA APIs provided by CS Linux are the same APIs as provided by CS OS/2.

- CS Linux provides multiple SNA programming interfaces, including:
  - CPI-C for LU6.2
  - APPC for LU6.2
  - LUA (for LU0, 1, 2, and 3)

If initially deploying on Linux on Intel, a second step is to consolidate the gateway application into Linux on zSeries, collapsing the SNA network segment to be within the zSeries environment.

Remote SNA API Client/Server technology

- 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 using a single Remote API server
  - Pools of LUs can be shared across servers on a 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
- Supports Windows XP, 2000, 2003 clients and Linux clients

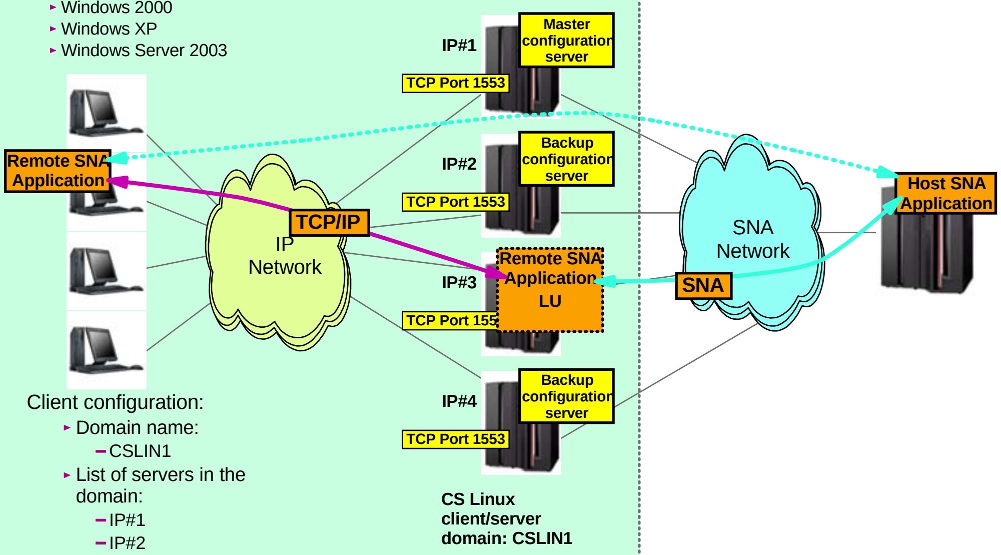


- SNA Application
- No SNA protocol stack
- Remote API client
  - Ship API calls over IP to Remote API server

- Full SNA protocol stack
- Remote API server
- Mulitiple servers can make up a domain
- Mix xSeries and zSeries servers in domain

CS Linux client/server domain overview

- Linux
- Windows 2000
- Windows XP
- Windows Server 2003


IP#1
Master configuration server
TCP Port 1553
IP#2
Backup configuration server
TCP Port 1553
Remote SNA Application
TCP/IP
IP Network
IP#3
Remote SNA Application LU
TCP Port 155
SNA
SNA Network
Host SNA Application
IP#4
Backup configuration server
TCP Port 1553
CS Linux client/server domain: CSLIN1

Client configuration:

- Domain name:
  - CSLIN1
- List of servers in the domain:
  - IP#1
  - IP#2
  - IP#3
  - IP#4

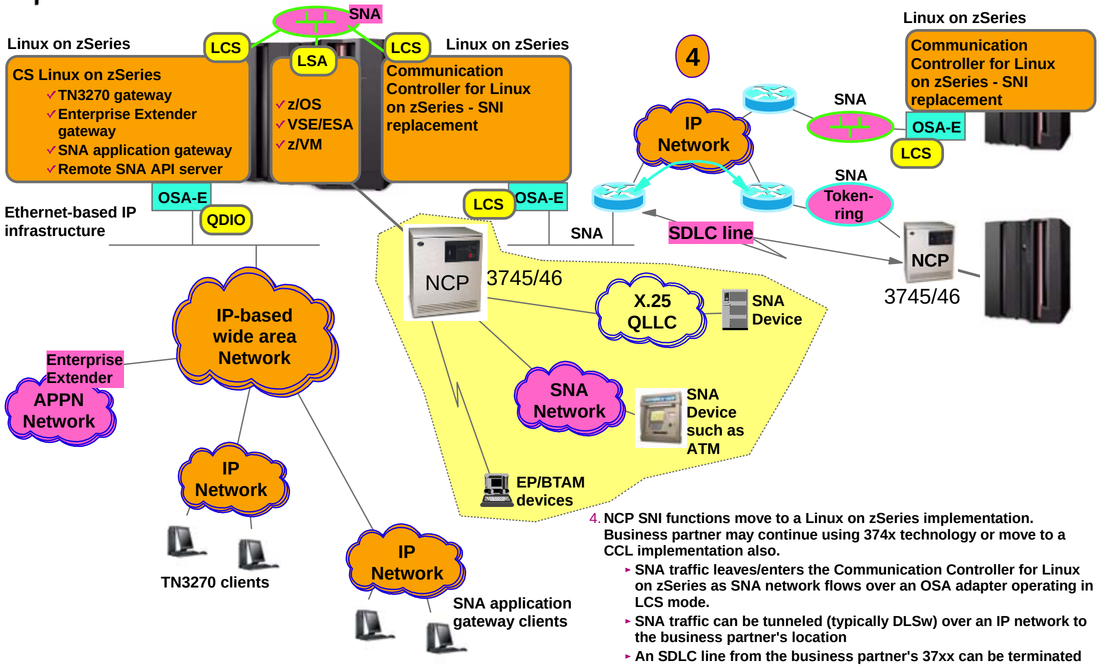
IBM®

SNA/IP Integration Using Linux on zSeries
-
Preview of Communication Controller for Linux on zSeries

Redbooks

International Technical Support Organization

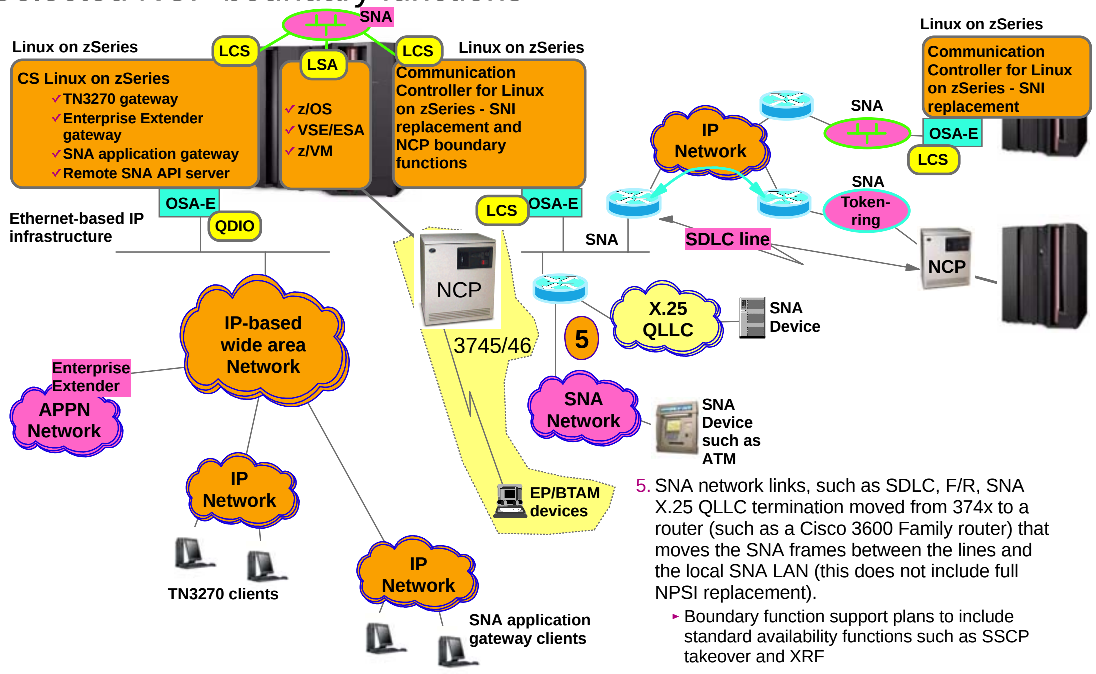
(Planned 2005): Communication Controller for Linux Phase 1 - SNI replacement



4. NCP SNI functions move to a Linux on zSeries implementation. Business partner may continue using 374x technology or move to a CCL implementation also.
   - SNA traffic leaves/enters the Communication Controller for Linux on zSeries as SNA network flows over an OSA adapter operating in LCS mode.
   - SNA traffic can be tunneled (typically DLSw) over an IP network to the business partner's location
   - An SDLC line from the business partner's 37xx can be terminated in a local router

Redbooks

(Planned 2005): Communication Controller for Linux Phase 1 - Selected NCP boundary functions

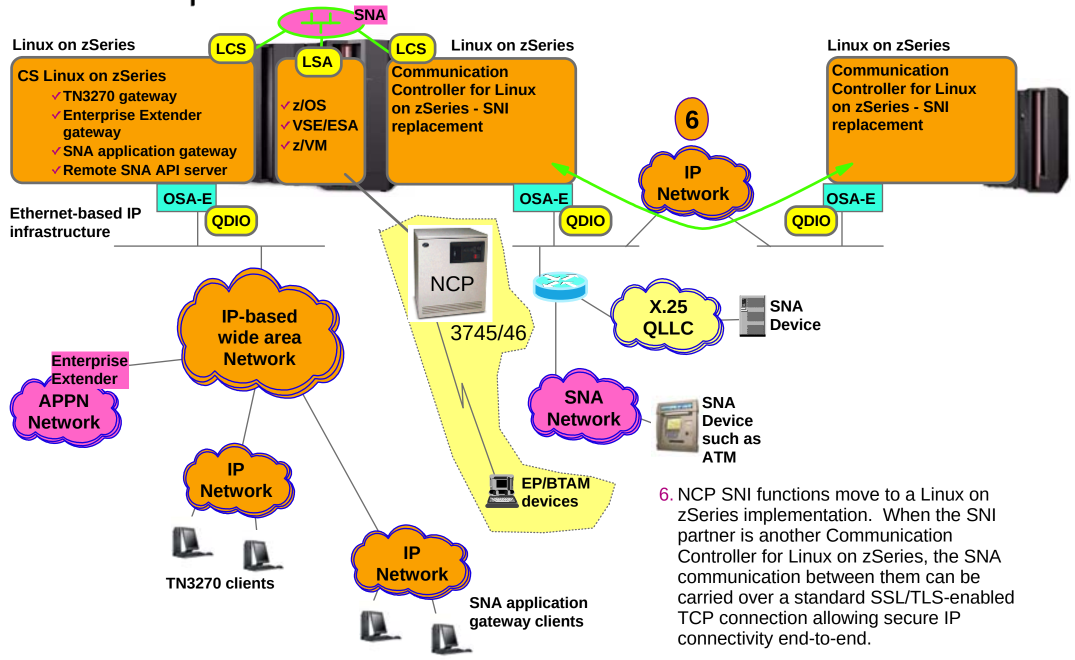


5. SNA network links, such as SDLC, F/R, SNA X.25 QLLC termination moved from 374x to a router (such as a Cisco 3600 Family router) that moves the SNA frames between the lines and the local SNA LAN (this does not include full NPSI replacement).
   - Boundary function support plans to include standard availability functions such as SSCP takeover and XRF

Redbooks

ibm.com/redbooks

(Planned 2005): Communication Controller for Linux (CCL) Phase 2 - SNI to CCL partner over IP

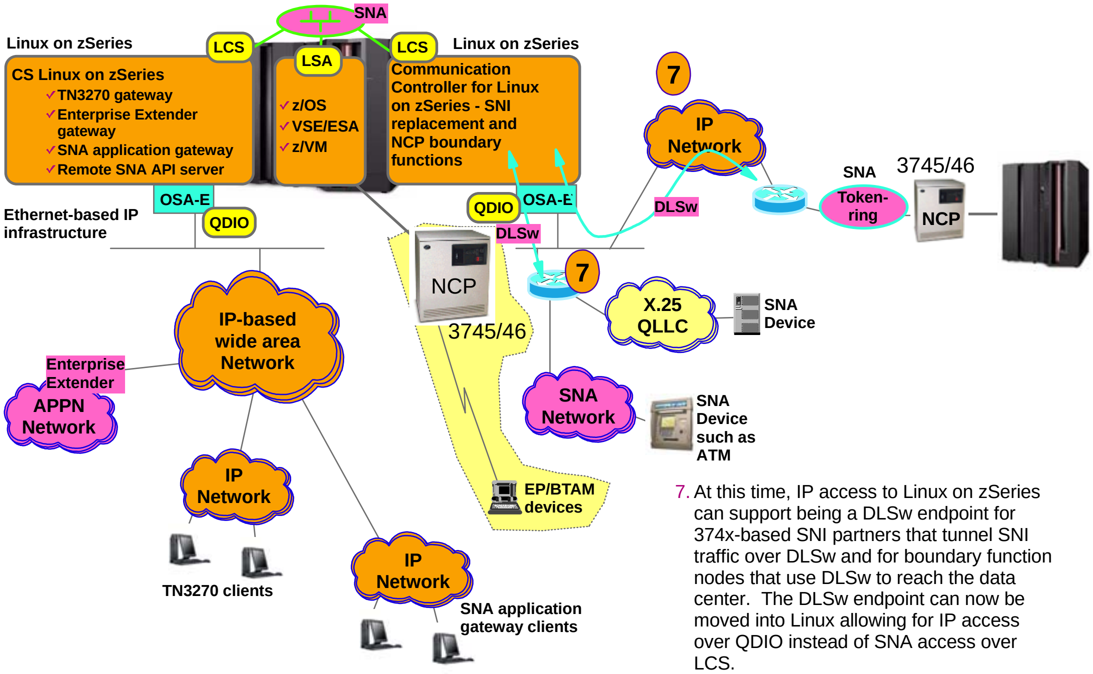


6. NCP SNI functions move to a Linux on zSeries implementation. When the SNI partner is another Communication Controller for Linux on zSeries, the SNA communication between them can be carried over a standard SSL/TLS-enabled TCP connection allowing secure IP connectivity end-to-end.

Redbooks

ibm.com/redbooks

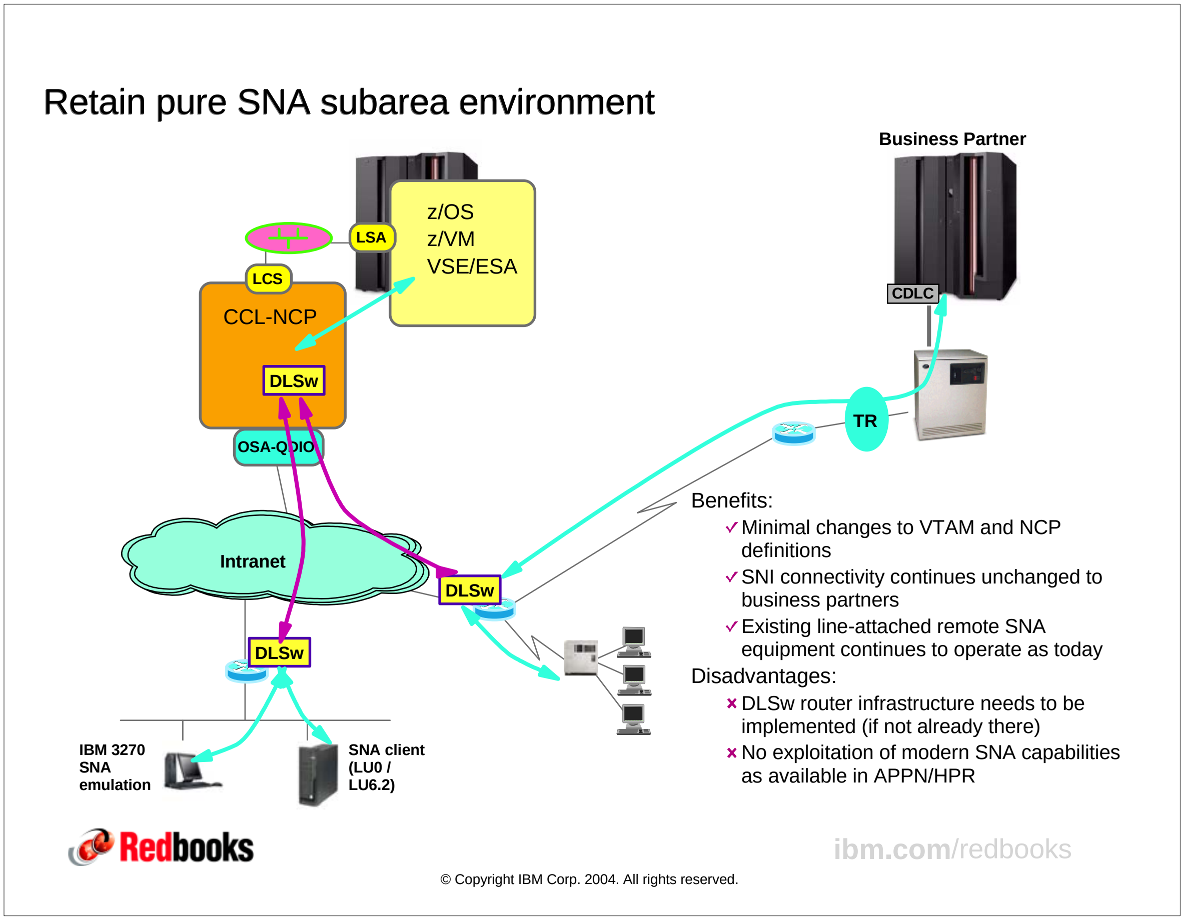
(Planned 2005): Communication Controller for Linux Phase 2 - DLSw endpoint support



7. At this time, IP access to Linux on zSeries can support being a DLSw endpoint for 374x-based SNI partners that tunnel SNI traffic over DLSw and for boundary function nodes that use DLSw to reach the data center. The DLSw endpoint can now be moved into Linux allowing for IP access over QDIO instead of SNA access over LCS.

Solutions overview matrix

Impact on existing mainframe OS SNA configuration definitions	**Virtually none - Retain pure SNA subarea environment**	**A few - Retain SNA subarea environment in mainframe OS**	**Some including APPN enablement - Exploit SNA APPN/HPR technologies**
IBM 3270 terminal access (display and printer)	SNA LLC via DLSw from branch to DC, CCL-NCP, SNA subarea connectivity to z/OS, z/VM, or VSE/ESA	TN3270 client software in branch, TN3270 server on CS Linux in DC, SNA subarea or APPN connectivity to z/OS, z/VM, or VSE/ESA	TN3270 client software in branch, TN3270 server on CS Linux in DC or on z/OS, z/VM, or VSE/ESA
SNA client/server access (LU0, LU6.2)	SNA LLC via DLSw from branch to DC, CCL-NCP, SNA subarea connectivity to z/OS, z/VM, or VSE/ESA	Remote SNA API to CS Linux in DC, SNA subarea or APPN connectivity to z/OS, z/VM, or VSE/ESA	EE from branch or SNA LLC via DLSw from branch to DC EE gateway, EE directly to z/OS or EE to CS Linux with APPN connectivity to z/VM or VSE/ESA
Business partner SNA connectivity (SNI)	CCL-NCP SNI replacement, SNA subarea connectivity to z/OS, z/VM, or VSE/ESA	CCL-NCP SNI replacement, SNA subarea connectivity to z/OS, z/VM, or VSE/ESA	EE EBN directly to business partner or to AT&T



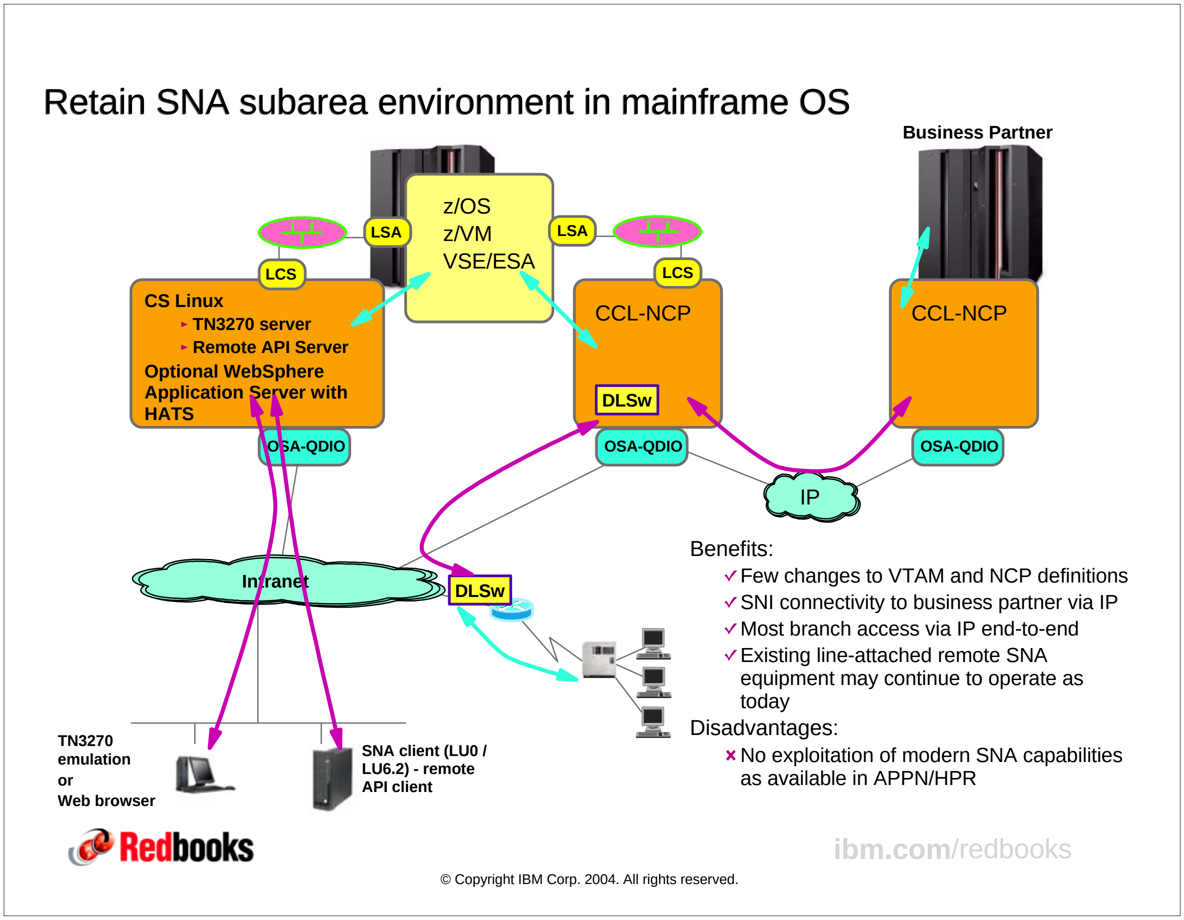
Retain pure SNA subarea environment

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



Retain SNA subarea environment in mainframe OS

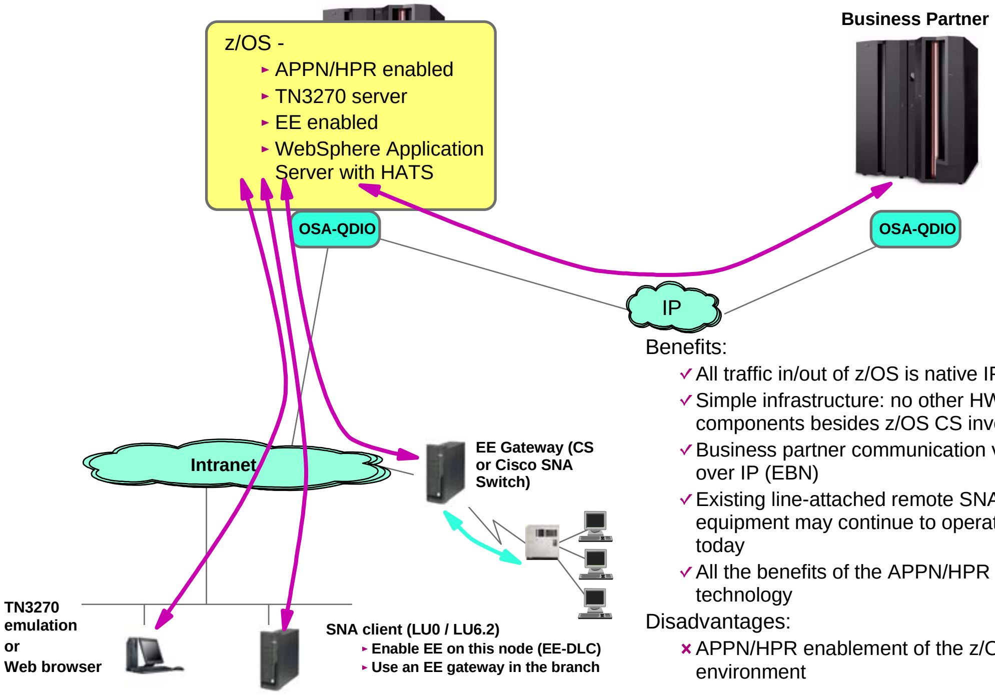
Benefits:

- ✓ Few changes to VTAM and NCP definitions
- ✓ SNI connectivity to business partner via IP
- ✓ Most branch access via IP end-to-end
- ✓ Existing line-attached remote SNA equipment may continue to operate as today

Disadvantages:

- × No exploitation of modern SNA capabilities as available in APPN/HPR

Exploit SNA APPN/HPR technologies in the z/OS environment



z/OS -

- APPN/HPR enabled
- TN3270 server
- EE enabled
- WebSphere Application Server with HATS

Business Partner

EE Gateway (CS or Cisco SNA Switch)

TN3270 emulation or Web browser

SNA client (LU0 / LU6.2)

- Enable EE on this node (EE-DLC)
- Use an EE gateway in the branch

Benefits:

- ✓ All traffic in/out of z/OS is native IP
- ✓ Simple infrastructure: no other HW/SW components besides z/OS CS involved
- ✓ Business partner communication via HPR over IP (EBN)
- ✓ Existing line-attached remote SNA equipment may continue to operate as today
- ✓ All the benefits of the APPN/HPR technology

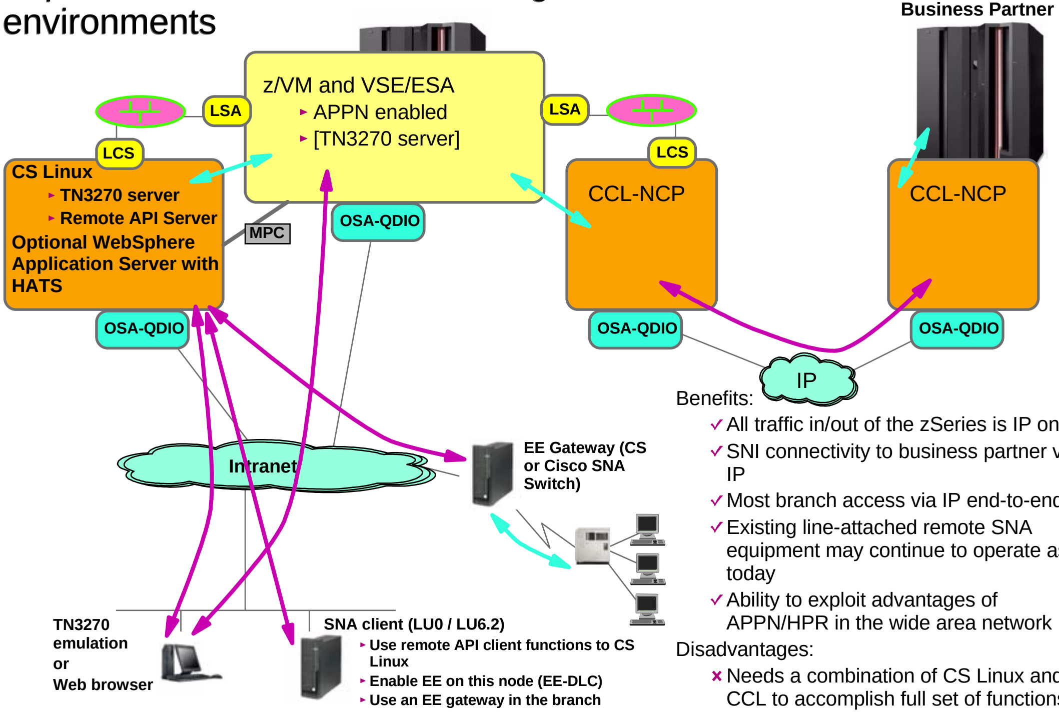
Disadvantages:

- × APPN/HPR enablement of the z/OS environment

Redbooks

ibm.com/redbooks

Exploit SNA APPN/HPR technologies for the z/VM and VSE/ESA environments



Business Partner

z/VM and VSE/ESA

- APPN enabled
- [TN3270 server]

CS Linux

- TN3270 server
- Remote API Server

Optional WebSphere Application Server with HATS

EE Gateway (CS or Cisco SNA Switch)

TN3270 emulation or Web browser

SNA client (LU0 / LU6.2)

- Use remote API client functions to CS Linux
- Enable EE on this node (EE-DLC)
- Use an EE gateway in the branch

Benefits:

- ✓ All traffic in/out of the zSeries is IP only
- ✓ SNI connectivity to business partner via IP
- ✓ Most branch access via IP end-to-end
- ✓ Existing line-attached remote SNA equipment may continue to operate as today
- ✓ Ability to exploit advantages of APPN/HPR in the wide area network

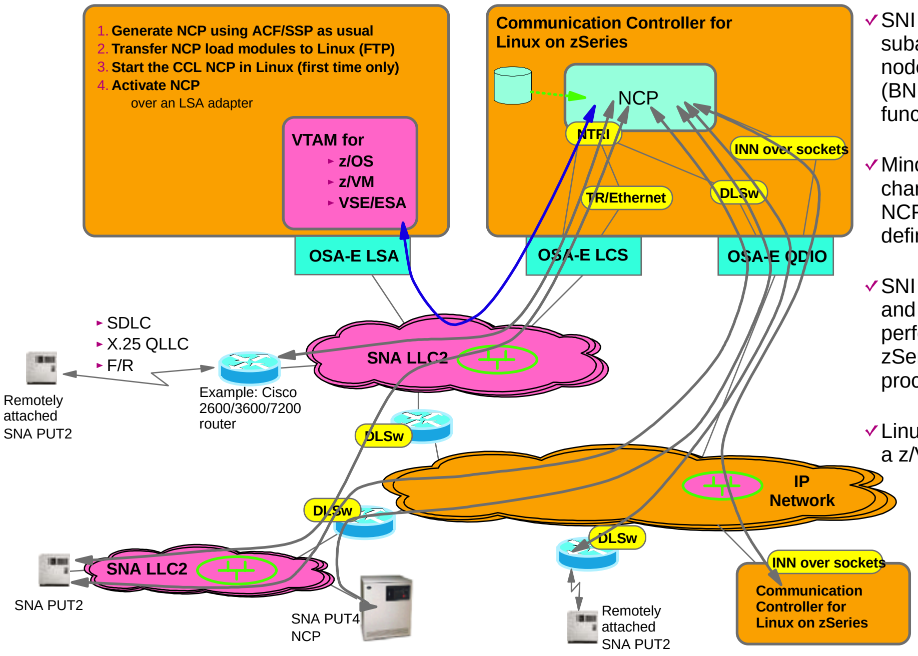
Disadvantages:

- × Needs a combination of CS Linux and CCL to accomplish full set of functions

Redbooks

ibm.com/redbooks

Subarea SNA connectivity overview for Communication Controller for Linux on zSeries - Phase 2



1. Generate NCP using ACF/SSP as usual
2. Transfer NCP load modules to Linux (FTP)
3. Start the CCL NCP in Linux (first time only)
4. Activate NCP
   over an LSA adapter

- ✓ SNI and SNA subarea boundary node functions (BNN) remain NCP functions
- ✓ Minor configuration changes to existing NCP and VTAM definitions
- ✓ SNI gateway NCP and BNN functions performed using zSeries IFL processors
- ✓ Linux in LPAR or as a z/VM guest

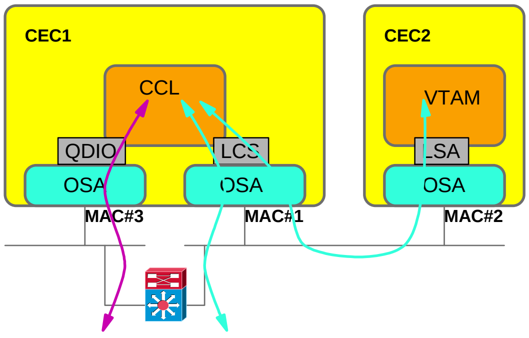
Redbooks

ibm.com/redbooks

OSA port usage - Sharing ports between CCL and VTAM

- OSA has special support that allows IP packets between two LPARs that share an LCS or a QDIO adapter to loop back to the destination LPAR without going out over the actual LAN infrastructure.
- That support does not exist for LSA (SNA).
  - OSA LSA token-ring and ATM allow packets to same MAC as where they came from (looping the frame out on the network and then back in)
  - OSA LSA Ethernet doesn't allow that
- When the LAN type is Ethernet then VTAM and CCL cannot share the same OSA port if they are to communicate with each other
  - VTAM must have one LSA port and CCL another LCS port onto the same LAN
    - Can be two separate adapters or the two ports on a two-port OSA-express adapter

The CCL LCS OSA adapter can be used for both SNA and IP communication in/out of the CCL Linux environment, but typically IP would be going over a QDIO adapter.



The VTAM owning OS and the CCL owning Linux OS may both reside in a single CEC or be split between two CECs as long as there is LAN connectivity between them.

Redbooks

ibm.com/redbooks

Availability

VTAM availability

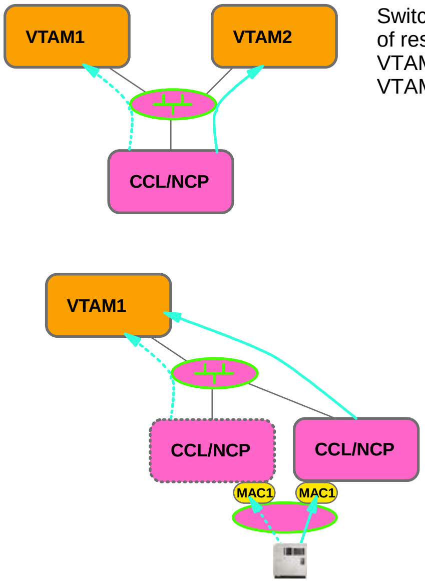
- SSCP takeover functions will work as today

Susbsystem availability

- Extended Recovery Facility (XRF) will work as today

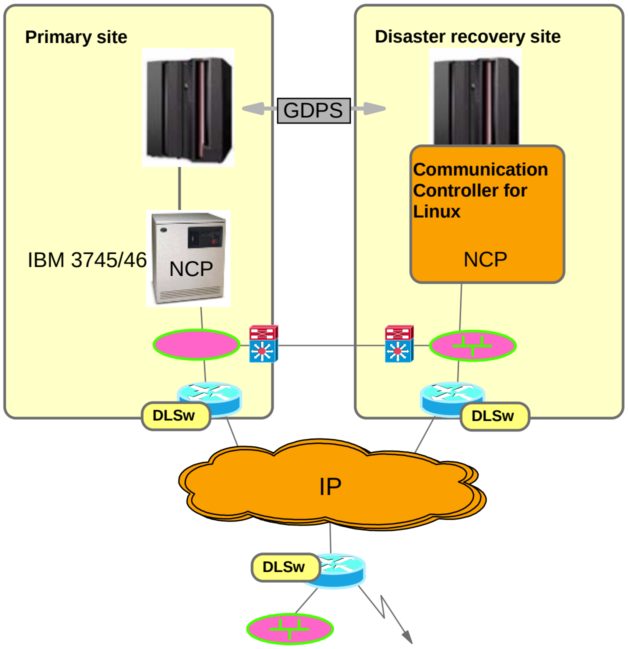
CCL/NCP availability

- Redundant CCL/NCPs with duplicate TR MAC addresses
- Similar capabilities can be deployed for Ethernet by combining VLAN technology and Cisco's DLSW+ technology



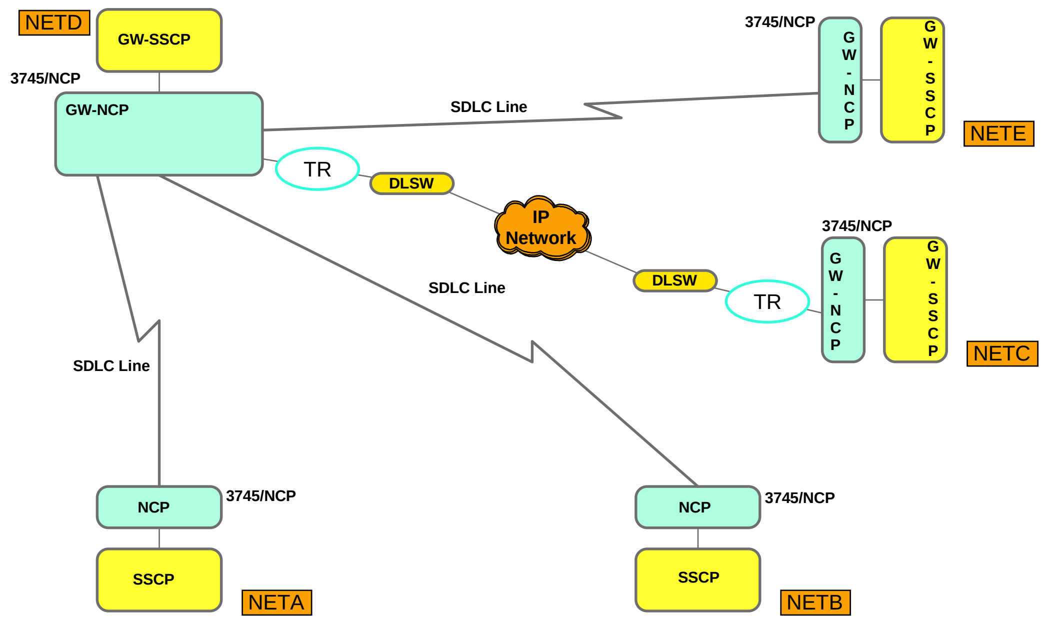
Switch ownership
of resources from
VTAM1 to
VTAM2

Disaster recovery site establishment

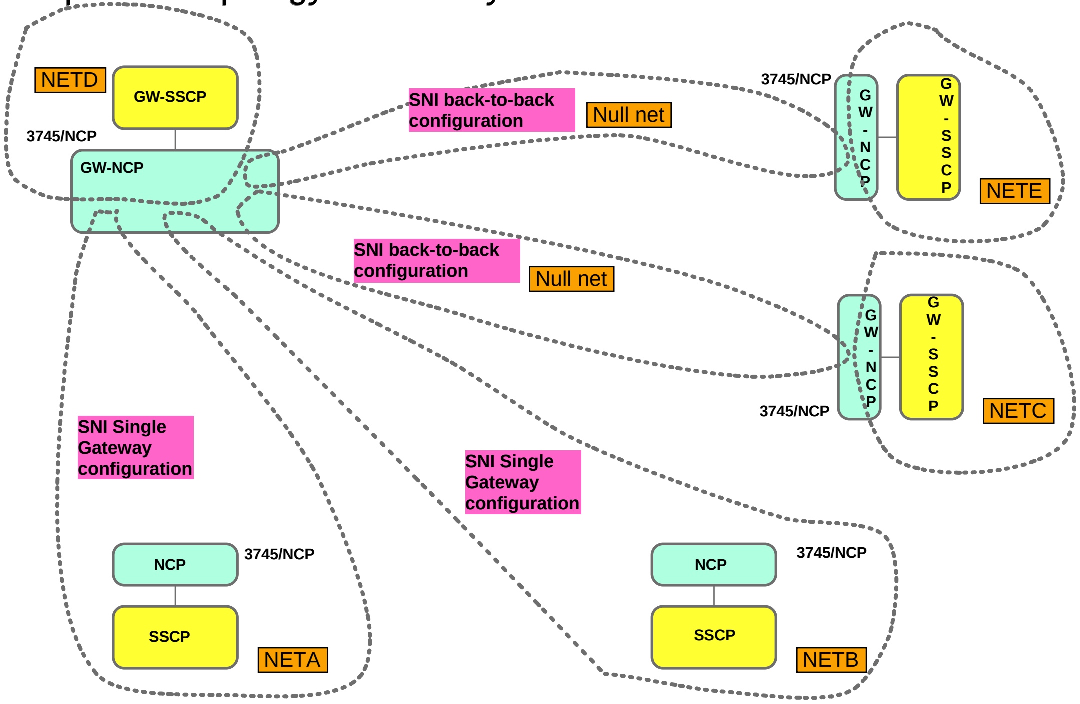


- For installations currently investing in DR sites, IBM 3745/46 redundancy poses some challenges.
- An alternative in many cases to installing spare IBM 3745/46 hardware will be to use Communication Controller for Linux running the DR NCP in the DR site.
- For LAN-attached connections, a switch to the DR site can be done using layer-2 bridging of SNA flows, or using DLSw to redirect the traffic to the DR site.
- If physical serial lines are attached to the IBM 3745/46, they need as usual to be manually switched to the DR site where they then can be terminated in a DLSw router

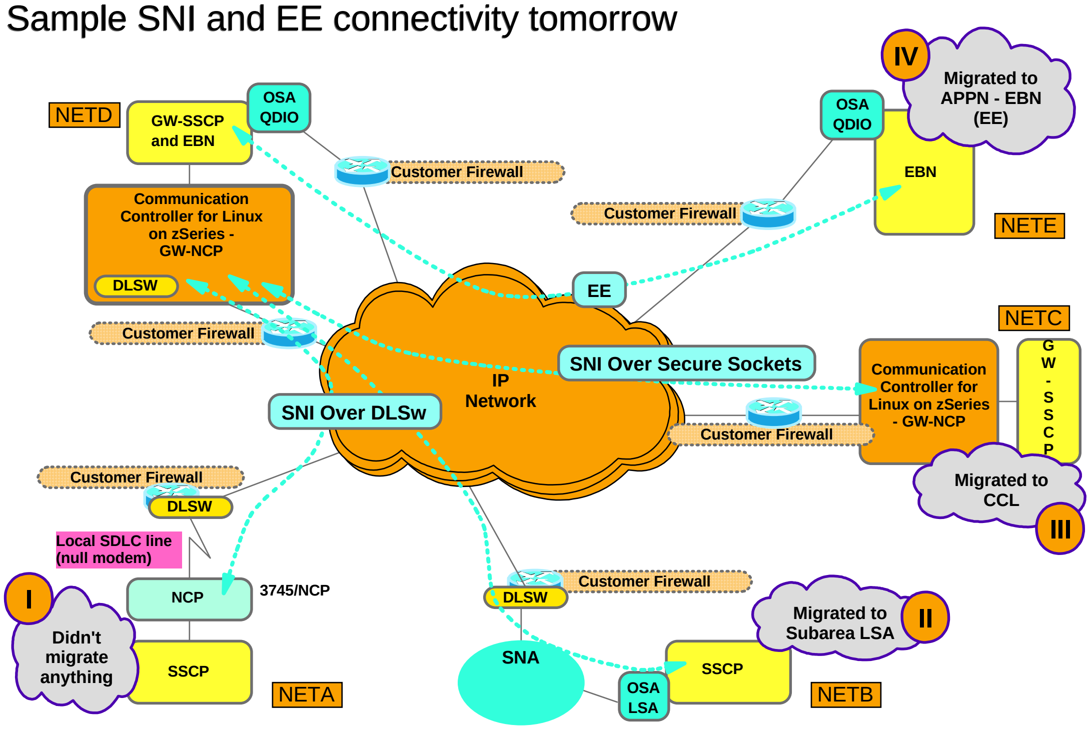
Sample SNI connectivity view today



Sample SNI topology view today

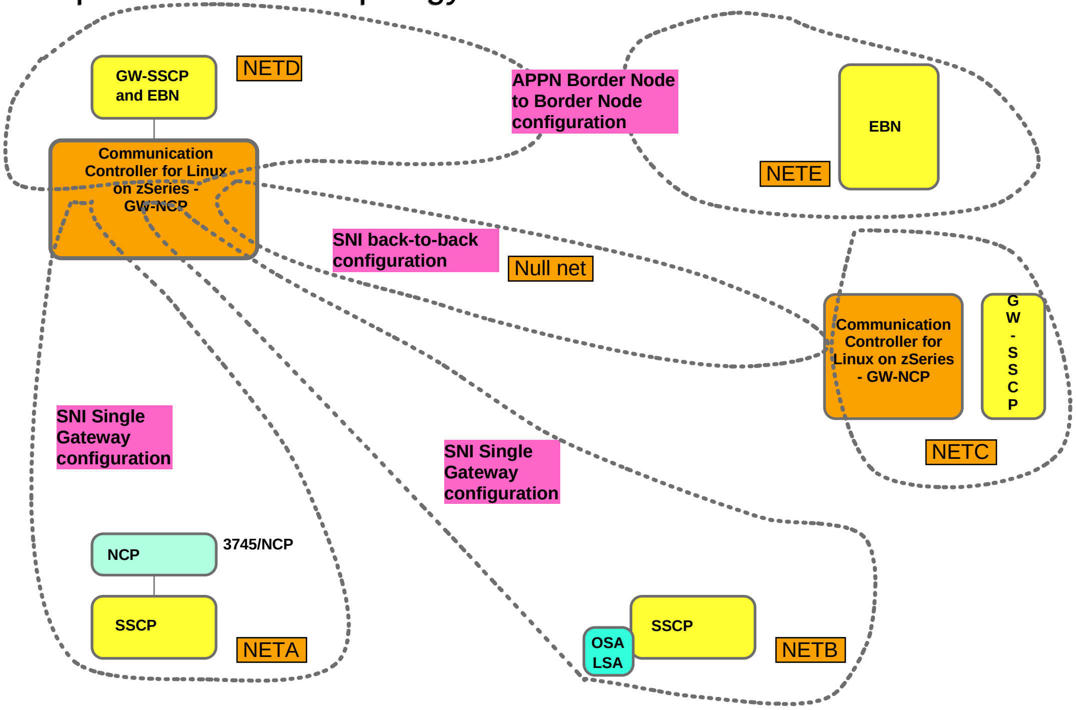


Sample SNI and EE connectivity tomorrow



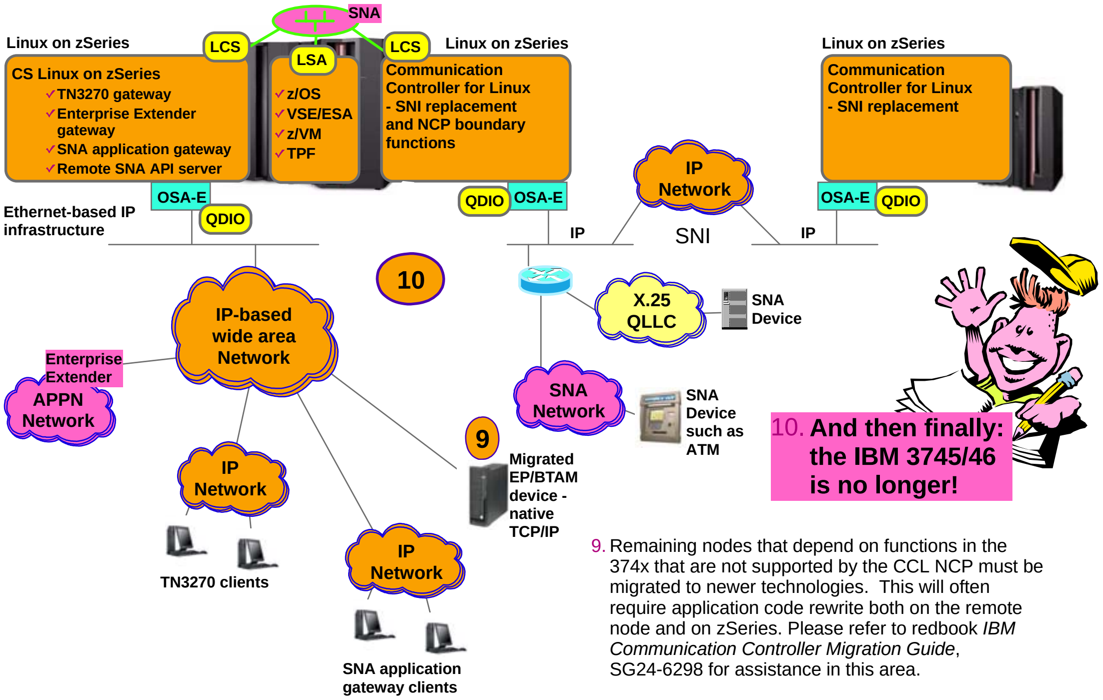
ibm.com/redbooks

Sample SNI and EE topology view tomorrow


GW-SSCP and EBN
NETD
APPN Border Node to Border Node configuration
EBN
NETE
Communication Controller for Linux on zSeries - GW-NCP
SNI back-to-back configuration
Null net
Communication Controller for Linux on zSeries - GW-NCP
G W - S S C P
NETC
SNI Single Gateway configuration
SNI Single Gateway configuration
NCP
3745/NCP
SSCP
NETA
SSCP
OSA LSA
NETB

ibm.com/redbooks

Migrate remaining 374x-dependent nodes to newer technologies



10. **And then finally: the IBM 3745/46 is no longer!**

9. Remaining nodes that depend on functions in the 374x that are not supported by the CCL NCP must be migrated to newer technologies. This will often require application code rewrite both on the remote node and on zSeries. Please refer to redbook *IBM Communication Controller Migration Guide*, SG24-6298 for assistance in this area.

Summary

- ✓ **Preserve use of existing SNA applications**
  - IBM 3270 access
  - SNA Client/Server
  - SNA subarea business partner communication (SNI)
- ✓ **Replacement technology for selected IBM 3745/46 NCP functions will be provided by IBM**
  - No need to migrate off SNA subarea technology
- ✓ **Linux on zSeries is an important component in an SNA to IP migration strategy**
  - Skills in Linux on zSeries need to be established
- ✓ **With the existing and planned IBM provided technologies, CS for Linux on zSeries and Communication Controller for Linux on zSeries, it will be possible to define an SNA to IP network migration plan that can support:**
  - Collapsing the physical SNA network to the zSeries or the data center
  - Achieving full independence of SNA wide area network hardware and software components
  - Removing need for maintaining an SNA wide area network component skills base

ibm.com

Wrap Up

Redbooks

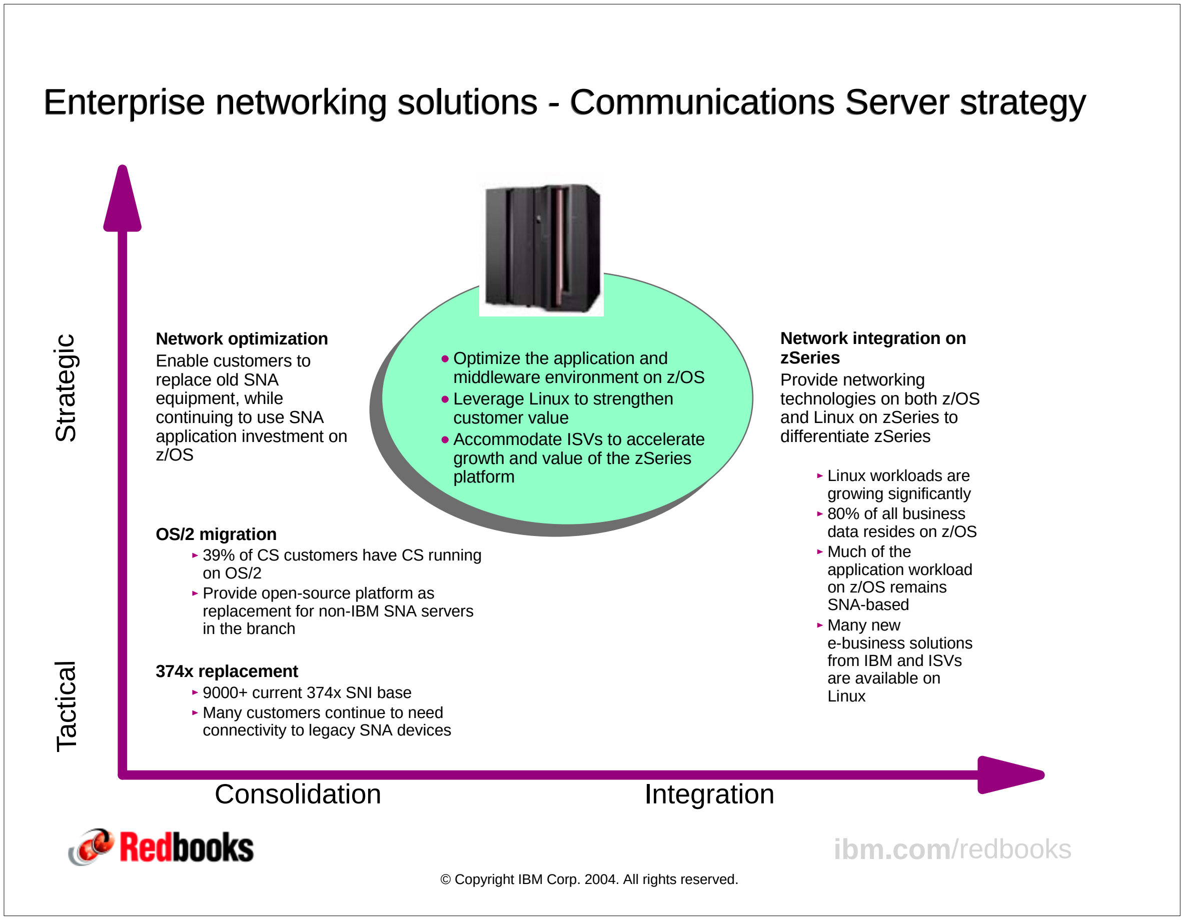
International Technical Support Organization

Enterprise Networking Solutions - overall strategy

- **Enterprise**
  - Solutions that address the enterprise networking needs
  - Solutions that are useful for both z/OS, z/VM, VSE/ESA, TPF, and Linux on zSeries
- **Networking**
  - SNA technologies in the enterprise
    - Distributed Communications Servers - Windows, AIX, Linux in general
    - NCP-related technologies
    - VTAM on z/OS, z/VM, and VSE/ESA
  - TCP/IP technologies in the enterprise
    - Primarily z/OS
    - Some IP-based technologies likely to be extended to Linux on zSeries
- **Solutions**
  - Products, documentations
  - Best practices and implementation experiences
  - Services - engagement in customer projects

ENS as the Center of Competence for enterprise networking technologies

Enterprise networking solutions - Communications Server strategy



Strategic

Tactical

Network optimization
Enable customers to replace old SNA equipment, while continuing to use SNA application investment on z/OS

- Optimize the application and middleware environment on z/OS
- Leverage Linux to strengthen customer value
- Accommodate ISVs to accelerate growth and value of the zSeries platform

Network integration on zSeries
Provide networking technologies on both z/OS and Linux on zSeries to differentiate zSeries

- Linux workloads are growing significantly
- 80% of all business data resides on z/OS
- Much of the application workload on z/OS remains SNA-based
- Many new e-business solutions from IBM and ISVs are available on Linux

OS/2 migration

- 39% of CS customers have CS running on OS/2
- Provide open-source platform as replacement for non-IBM SNA servers in the branch

374x replacement

- 9000+ current 374x SNI base
- Many customers continue to need connectivity to legacy SNA devices

Consolidation

Integration

ibm.com/redbooks

Linux-based SNA solutions - what's to come

- **Communications Server for Linux**
  - Move to a Linux 2.6 kernel base
  - Add enhanced SNA integration technology for the Web services application environment
  - Add specialized (and limited) SNA Primary LU capabilities
- **Communication Controller for Linux**
  - Phase 1 - planned at 1H2005
    - SNI replacement
    - Selected boundary function support
    - In phase 1, all SNA network attachment is via OSA LCS and SNA Link Layer Control frames
    - NCP sees all resources as NCP Token-Ring (NTRI) attached
  - Phase 2 - planned later 2005
    - Focus on IP (QDIO) connectivity to CCL
    - SNI over direct sockets (TCP connection to SNI partner CCL)
    - Datalink switching (DLSw) endpoint support in CCL

ibm.com/redbooks

CS z/OS - continue providing TCP/IP Sysplex autonomic capabilities

Self-optimizing:

- Sysplex Distributor will distribute incoming traffic to target stacks within a Sysplex using optimal available IP routes. This allows the use of high-speed interfaces such as OSA Express Gigabit Ethernet. In addition, it removes a restriction: Sysplex Distributor need no longer use only dynamic XCF interfaces for packet forwarding.
- Sysplex Distributor will exploit new z/OS WLM support to help optimize workload balancing for TCP/IP servers in a Sysplex. Sysplex Distributor will use server-specific recommendations from WLM that reflect how well target servers meet their service class goals.

Self-healing:

- Sysplex Distributor will use key performance indicators such as connection backlog queues to supplement existing measurements and WLM recommendations. This will help improve load balancing.

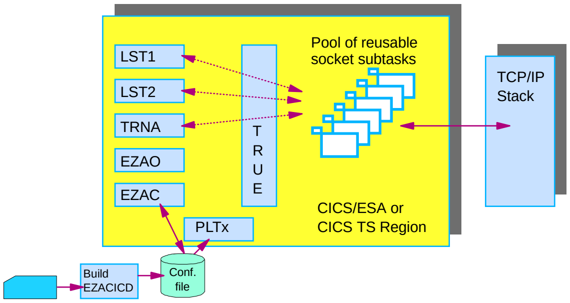
Self-configuring:

- TCP/IP will be able to rejoin a Sysplex once problems that have triggered an automatic takeover have been resolved. When a stack rejoins a Sysplex, it can automatically restore its original configuration and resume ownership of any DVIPAs for which it is the primary owner. This expands on the TCP/IP Automatic Takeover function introduced in z/OS V1.6.

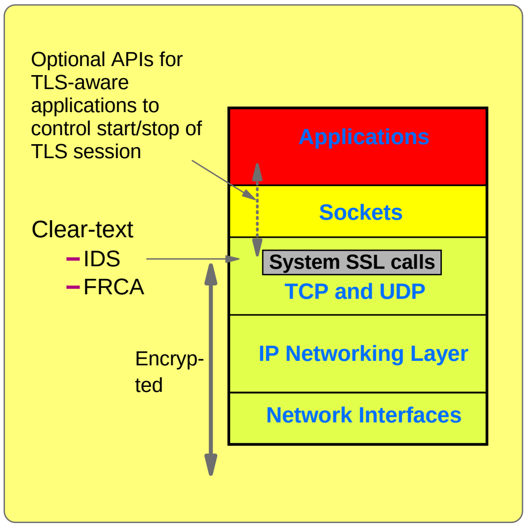
CS z/OS - focus on enhancing CICS sockets

CICS sockets enhancements are planned to improve application performance by:

- Allowing CICS sockets to use the CICS Open Transaction Environment (OTE). This is designed to reduce task switching in CICS environments.
- Helping to reduce the overhead of CICS sockets tracing and monitoring processing when these facilities are not activated.
- Allowing the IP CICS sockets Task-Related User Exit (TRUE) to be loaded above the 16MB line, providing virtual storage constraint relief.



CS z/OS - Application-transparent Transport Layer Security (TLS)

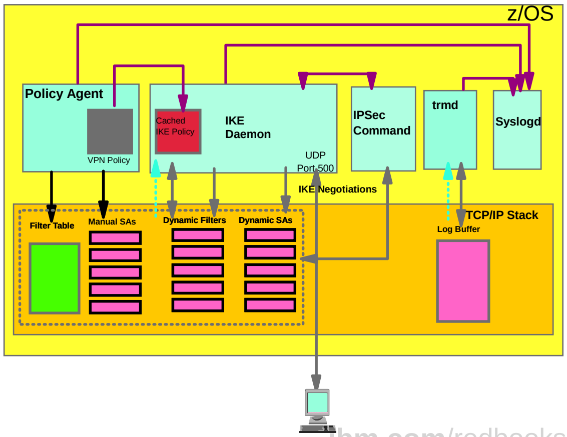


- **Basic TCP/IP stack-based TLS**
  - TLS process performed at TCP layer without requiring any application change (transparent)
  - All connections to specified port are designated as TLS required
  - Transparent TLS policies managed via Policy Agent
- **Transparent TLS can be requested by aplication**
  - Application issue 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

CS z/OS - integrated IPSec/VPN support including NAT traversal support

- Features
  - Configuration support
    - Optimized for z/OS host-to-host and z/OS host-to-gateway (z/OS gateway still supported)
    - NAT traversal support
  - Simplified infrastructure
    - Eliminates need for FW technologies daemons
  - 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

- Complete IPSec, filtering, and IKE solution part of z/OS Communications Server
  - Alternative to firewall technologies
    - New IKE daemon and configuration
- Makes use of existing Communications Server Infrastructure
  - TCP/IP stack - IPSec and IP filtering
  - Policy agent - reads and manages IPSec and IKE policy
  - trmd - monitors TCP/IP stacks for log messages

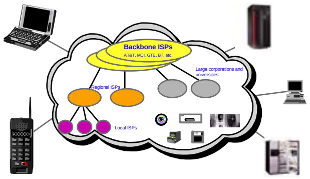


CS z/OS - the journey to IPv6 continues

IPv6 deployment phases

- 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)
  - Additonal Network Management MIBs
- After z/OS V1R6
  - Integrated IPSec
  - HiperSockets IPv6
  - Advanced Socket APIs
  - Extended Stats MIB, OSPFv3 MIB
  - Intrusion Detection Services
  - IPv6 mobility support

The Internet - a worldwide digital utility.



Connectivity for ***anyone*** from ***anywhere*** (car, home, office) to ***anything***!

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

Redbooks

ibm.com/redbooks

CS z/OS V1R5 and V1R6 functions that were APARed back to earlier releases

APAR	Back to	Description
PQ84185	z/OS V1R4	FTP secure password function (SSL/TLS login without a password)
PQ86508	z/OS V1R4	VLAN support
PQ76172	OS/390 V2R10	OMPROUTE tracing via CTRACE
PQ73161	z/OS V1R2	SMTP/NJE IPMAILER as a host name
PQ80281	z/OS V1R4	FTP EPSV and PASSIVEDATAPORT support
PQ65205	z/OS V1R2	Sysplex Distributor more than 4 ports
PQ76866	z/OS V1R4	Sysplex Distributor Round-Robin distribution method
UQ81245	z/OS V1R4	Network Management Interface support
PQ63027 and PQ67798	OS/390 V2R10	TN3270 server improved keyboard unlock control
PQ92262	z/OS V1R4	OSA Express inbound performance control

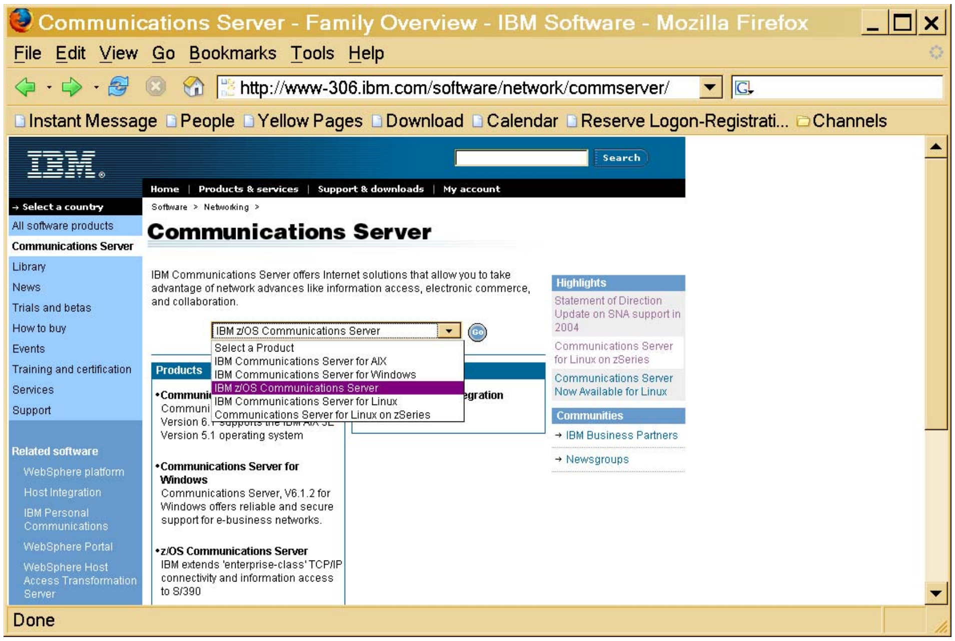
Redbooks

ibm.com/redbooks

Some useful links to the Web

- 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 publications in the Communications Server library):
  - `http://www.ibm.com/servers/eserver/zseries/zos/bkserv/`
- Communications Server home page:
  - `http://www.ibm.com/software/network/commserver`
- Communications Server support page:
  - `http://www.ibm.com/software/network/commserver/os390/support`

Communications Server home page


Communications Server - Family Overview - IBM Software - Mozilla Firefox
File Edit View Go Bookmarks Tools Help
http://www-306.ibm.com/software/network/commserver/
Instant Message People Yellow Pages Download Calendar Reserve Logon-Registrati... Channels
IBM
Search
Home | Products & services | Support & downloads | My account
→ Select a country
All software products
Communications Server
Library
News
Trials and betas
How to buy
Events
Training and certification
Services
Support
Related software
WebSphere platform
Host Integration
IBM Personal Communications
WebSphere Portal
WebSphere Host Access Transformation Server
Software > Networking >
Highlights
Statement of Direction Update on SNA support in 2004
Communications Server for Linux on zSeries
Communications Server Now Available for Linux
Communities
→ IBM Business Partners
→ Newsgroups
Products
Go
Done

Communications Server

IBM Communications Server offers Internet solutions that allow you to take advantage of network advances like information access, electronic commerce, and collaboration.

IBM z/OS Communications Server

- Select a Product
- IBM Communications Server for AIX
- IBM Communications Server for Windows
- IBM z/OS Communications Server
- IBM Communications Server for Linux
- Communications Server for Linux on zSeries

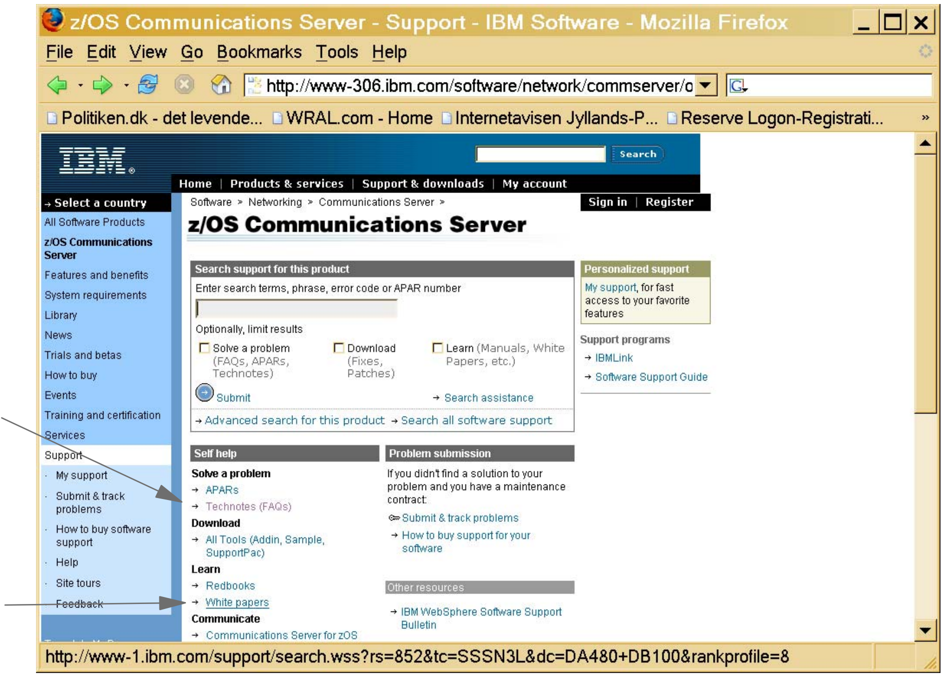
•Communi... Communi... Version 6.1 supports the IBM AIX 5L Version 5.1 operating system

...egration

•**Communications Server for Windows**
Communications Server, V6.1.2 for Windows offers reliable and secure support for e-business networks.

•**z/OS Communications Server**
IBM extends 'enterprise-class' TCP/IP connectivity and information access to S/390

z/OS Communications Server support page


z/OS Communications Server - Support - IBM Software - Mozilla Firefox
File Edit View Go Bookmarks Tools Help
http://www-306.ibm.com/software/network/commserver/o
Politiken.dk - det levende... WRAL.com - Home Internetavisen Jyllands-P... Reserve Logon-Registrati...
IBM
Search
Home | Products & services | Support & downloads | My account
→ Select a country
Sign in | Register

Software > Networking > Communications Server >

z/OS Communications Server

All Software Products
z/OS Communications Server
Features and benefits
System requirements
Library
News
Trials and betas
How to buy
Events
Training and certification
Services
Support
· My support
· Submit & track problems
· How to buy software support
· Help
· Site tours
· Feedback

Search support for this product

Enter search terms, phrase, error code or APAR number

Optionally, limit results

□ Solve a problem (FAQs, APARs, Technotes) □ Download (Fixes, Patches) □ Learn (Manuals, White Papers, etc.)

Submit → Search assistance

→ Advanced search for this product → Search all software support

Personalized support

My support, for fast access to your favorite features

Support programs
→ IBMLink
→ Software Support Guide

Self help

Solve a problem
→ APARs
→ Technotes (FAQs)

Download
→ All Tools (Addin, Sample, SupportPac)

Learn
→ Redbooks
→ White papers

Communicate
→ Communications Server for zOS

Problem submission

If you didn't find a solution to your problem and you have a maintenance contract:

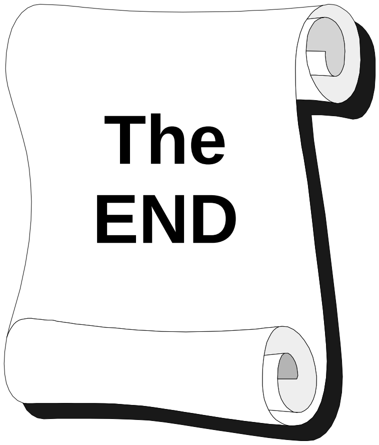
Submit & track problems
→ How to buy support for your software

Other resources
→ IBM WebSphere Software Support Bulletin

http://www-1.ibm.com/support/search.wss?rs=852&tc=SSSN3L&dc=DA480+DB100&rankprofile=8

Technotes is a collection of descriptions of typical problems and their solutions

Numerous white papers of interest



Contact Information:

Marketing or sales information:
Haechul Shin (haechul@us.ibm.com)

Technical information:
Alfred B Christensen (alfredch@us.ibm.com)

Services engagements:
April Singer (singeraf@us.ibm.com)