zSeries Availability and Workload Balancing Tutorial

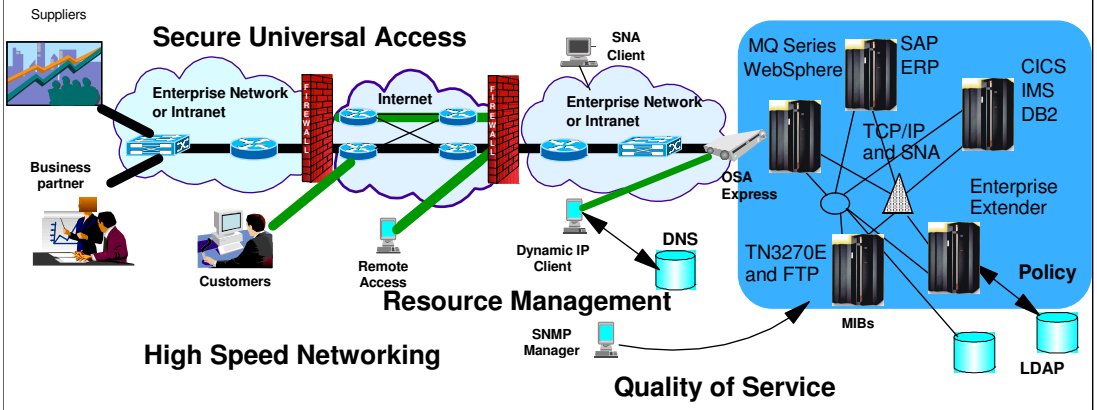
Enterprise Networking Solutions

z/OS CS is the engine that drives all application workloads for zSeries

- ✓**24X7 Access for over a trillion lines of application code**
- ✓**70% of all business data reside on zSeries**

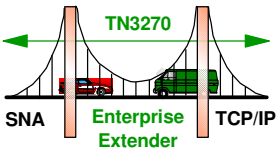
Leveraging Application Investments While Integrating New Technologies

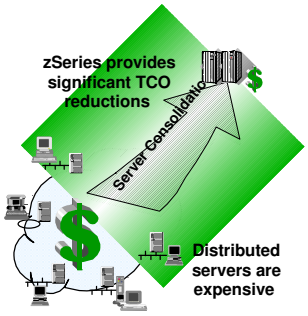
Clustering for Availability and Scalability via Sysplex



Architecture for z/OS CS

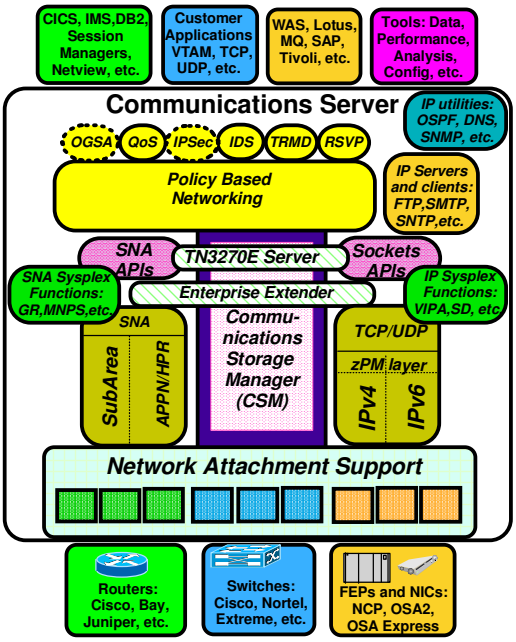
Bridge between the SNA world and the IP world



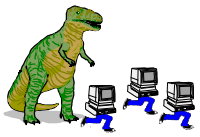


zSeries provides significant TCO reductions

Distributed servers are expensive

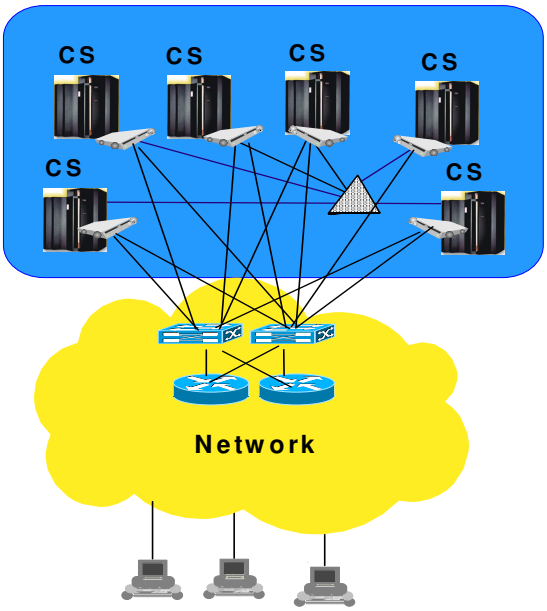


z/OS is ready for e-business on demand



Mainframes are back! (Were they ever gone?)

zSeries Sysplex

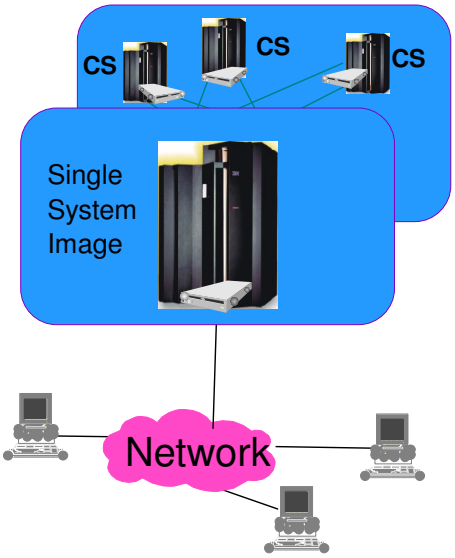


zSeries Sysplex Requirements

- Availability
  - For planned and unplanned outages
    - Network
    - Application
- Single System Image
  - End users and clients do not have to know the makeup of the data center
  - Applications on different nodes may share data
    - Via shared CF structures
      - ✓ Becomes Parallel Sysplex
- Horizontal Growth
  - Grow data center non-disruptively
- Internode Signaling
  - All systems need to be able to share information with one another
    - Via messaging or CF structures
- Multi-Protocol
  - SNA and TCP/IP

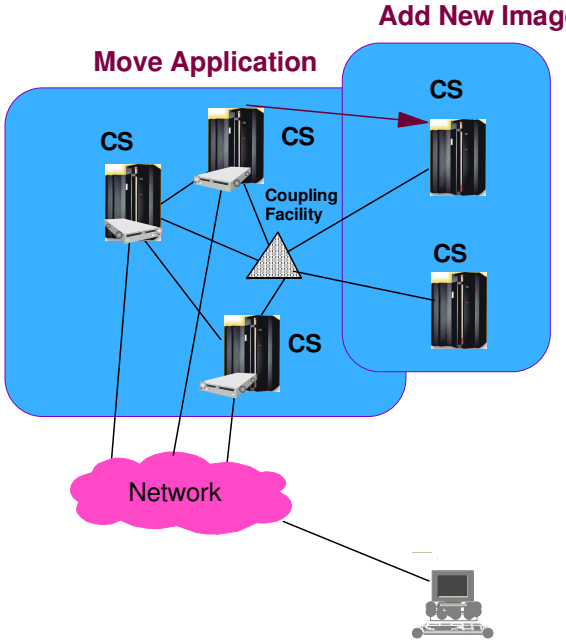
Sysplex Enables Single System Image

- **Transparent location of applications**
  - **Multiple images of same application appear as single application to end user**
    - **Balance Workload within Sysplex**
    - **Minimize Application failure impact**
    - **Freedom to move application workload to other images**
- **Network routing services**
  - **HPR Route Dynamics**
    - **Pathswitch around network failures**
    - **Congestion control**
  - **TCP/IP routing protocols**
    - **Reroute around intermediate network failures**
    - **Virtual IP Addressing (VIPA) for network attachment failures**
    - **Dynamic routing updates for movement of application workload**



Sysplex Enables Horizontal Growth

- **Connectivity Services**
  - Discovery of new sysplex members
  - Dynamic connectivity via XCF links
- **Routing Services**
  - Dynamic route calculations
  - Dynamic rerouting
- **Directory Services**
  - Dynamic registration for applications
  - Transparent location of resources
- **System Administration Services**
  - System Cloning
  - Application Cloning

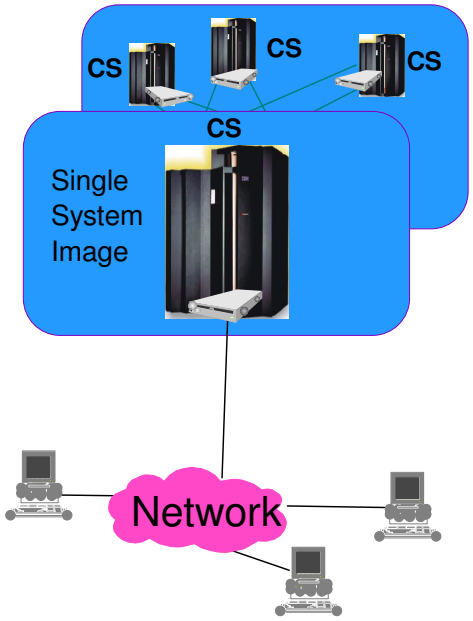


How to reduce your definitions and operator requirements so that horizontal growth and movement of application workload is more usable?

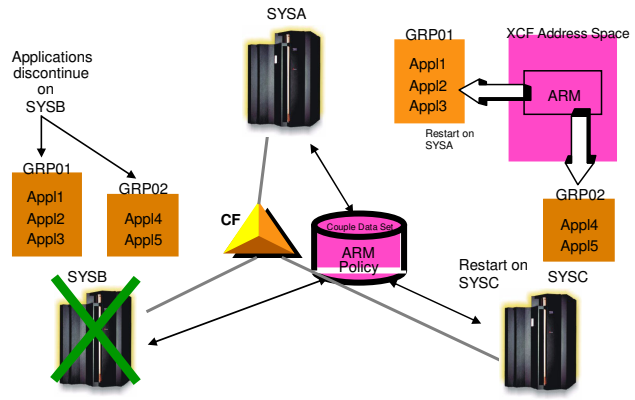
Reduced Definitions via System Symbolics

Exploit Cloning Support

- **System Cloning**
  - System Symbolics in VTAMLST members
    - Allows VTAMLST definitions to be shared among Sysplex members
  - System Symbolics in TCP Config files
    - Allows TCP Config Files to be shared among Sysplex members
- **VTAM Application Cloning**
  - Dynamic Definition of VTAM Applications
    - System Symbolics and wildcards in APPL names (e.g. APPL.&sysclone, APPL*)
  - Reduction in VTAM resources used for APPL definitions
    - Network Address allocated at OPEN ACB and released at CLOSE ACB
    - APPL definition created at OPEN ACB and deleted at CLOSE ACB
  - Allows for easy APPL relocation
- **Cloning for TN3270 Server**
  - Support includes TN3270 clients represented by VTAM APPL definitions
    - Reduction in VTAM resources used to represent TN3270 clients
    - Simplifies TN3270 Server relocation



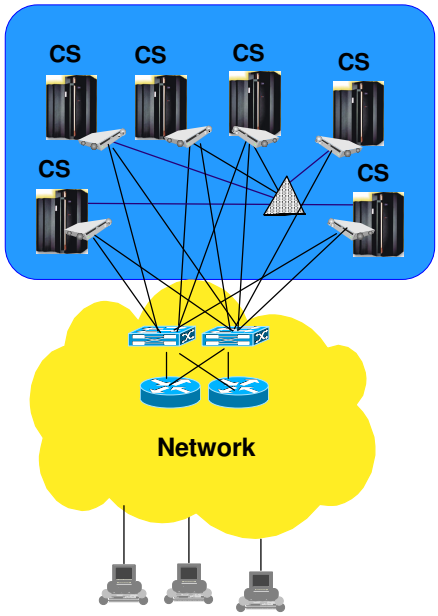
Automatic Recovery



Exploit zSeries Automatic Restart Manager (ARM)

- **Registered applications automatically restarted on failure**
  - ARM policy provides an ordered list for recovery
  - VTAM registers with ARM for restart
  - TCP/IP stack registers with ARM for in-place restart
- **ARM facility is open interface which can be exploited by any application**
  - Exploited by CICS, IMS, DB2

Network Design Decision Points

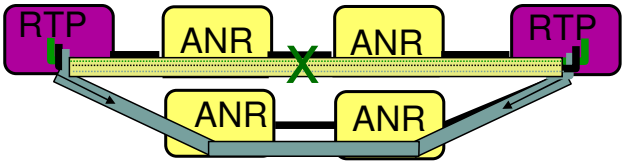


Implementation decisions
- Network Availability
  - Network access to SNA applications
    - Implement APPN/HPR?
      - ✓Data center or extend to network?
    - Implement SNA to IP integration?
      - ✓EE and/or TN3270E?
  - Network access to IP applications
    - Which network attachment choice?
      - ✓Multi-protocol or single protocol?
    - Which security choice?
      - ✓Network level and/or application level?
    - Implement VIPA?
      - ✓Static and/or Dynamic?
- Application Availability and Single System Image
  - SNA Applications
    - Implement Generic Resources?
    - Implement MNPS?
  - TCP/IP Applications
    - Which Workload Distribution choice?
    - Implement VIPA?
- Horizontal Growth and Internode Signaling
  - Implement XCF dynamics?
    - SNA and/or IP?

How to mesh connect your sysplex images so that you can promote data sharing and grow your business without impact to your customers?

How should you attach your sysplex to your WAN so that you can reduce networking costs while providing secure universal access to your SNA and IP applications?

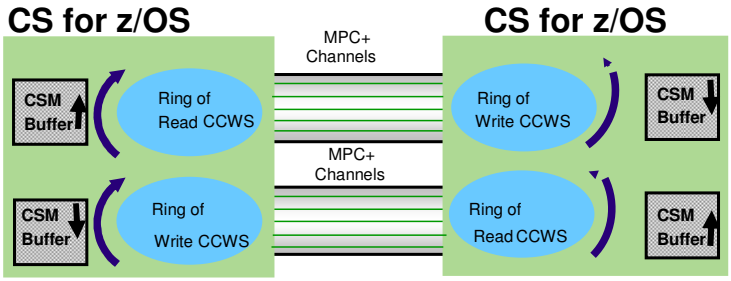
HPR Eliminates Network Outages



High Performance Routing (HPR)

- Reroutes sessions in event of a planned or unplanned node or link failure
- No impact to end user
- Support for application and terminal sessions
- Sessions rerouted to new route determined by Class of Service
- Sessions may be maintained across link "hit" without *any* switch occurring
- Discarded data retransmitted using HPR selective retransmission
- Congestion control (ARB) adapts to new route

Sysplex Connectivity Via MPC+



Improved IO Subsystem usage
- **–Significant with multiple channels**
  - **Better channel utilization**
  - **Better throughput**

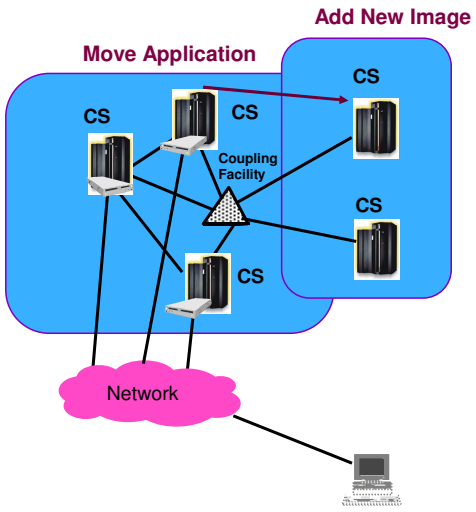
High Performance Data Transfer
- **–Significant CPU cycle savings**
  - **Reduction in data moves**
  - **Savings greater with HPDT applications**

A Multi-Protocol Channel as well as a Multi-Path Channel

XCF Dynamics Enables Horizontal Growth

XCF uses Coupling Facility Links for Data Transport

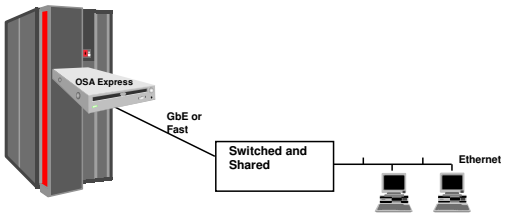
- **Eliminates requirement for ESCON definitions for SNA and IP**
- **Provides dynamic discovery and connectivity of other nodes in Sysplex**
  - **No coordinated definition required to add new images**
- **XCF messaging used to assist applications and TCP/IP stacks with data sharing**
  - **Sysplex Sockets**
  - **Dynamic VIPA**
  - **Sysplex Distributor**
- **Provides notification of new sysplex members as well as members who have failed**
  - **Dynamically updates existing members of sysplex**



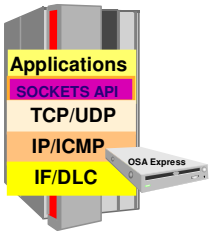
Network Attachments to Parallel Sysplex

	SNA	SNA	SNA	TCP/IP	
Interface	**Subarea**	**APPN**	**HPR**	**IP**	**Comments**
CDLC	Y	Y	Y	Y	Attached 37xx, 3174
MPC	Y	Y	Y	N	Early ESCON MPC - SNA only
MPC+	N	N	Y	Y	MPC+ ESCON, 2216. CS/NT, AIX, OEM- Bustech
Native ATM			Y	Y	390-390, NHD, OSA, OEM- Hitachi
LSA	Y	Y	Y	n/a	FDDI, TR, ENET,FNET, LANE - 2216, 3172, OSA, OEMs- Hitachi, Bustech
LCS	n/a	n/a	n/a	Y	FDDI,TR,ENET,FNET,LANE -same as LSA
CLAW	n/a	n/a	n/a	Y	AIX, OEMs
FICON	Y	Y	Y	Y	Uses MPC and MPC+, mostly transparent
GNET, FNET	N	N	N	Y	QDIO based OSA Express

Queued Direct I/O Support for OSA Express



OSA Express is a Logical Extension of TCP/IP Stack of CS for z/OS



Queued Direct I/O

- –Faster dispatch and reduced CPU utilization
  - ƒ Uses Direct Memory Access and CSM
- –Priority queuing
  - ƒ Queues correspond to TOS priorities set by App or Service Policy Agent
- –Dynamic Routing Updates to OSA Express Supported via IP Assist
  - ƒ OSA Express treated as logical extension of TCP/IP stack
    - •OSA/SF not needed for configuration and VIPA Takeover supported

Access to Data Center Via QDIO

- –SNA application access uses TN3270E (e.g.LU2) or Enterprise Extender (e.g. LU62)
- –Gigabit Ethernet And Fast Ethernet

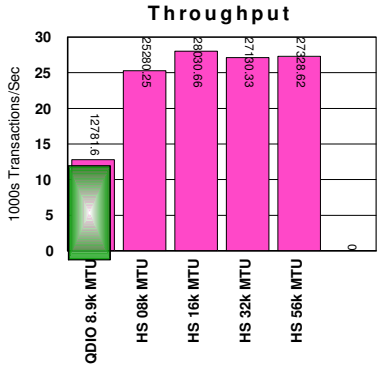
Performance and Reduced Cost

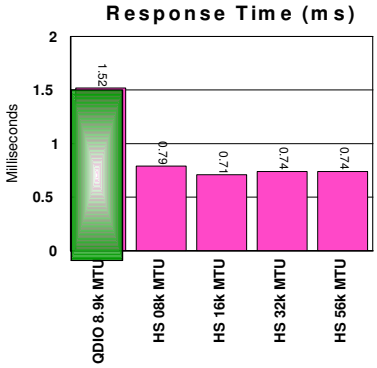
- –12X improvement over Escon for Throughput
- –2X improvement over Escon in CPU cycles
- –32X improvement over Escon in TCO

Hipersocket Support

High-Speed Connectivity Inside CEC

- **Internal TCP/IP network**
  - **LPAR to LPAR communication via shared memory**
  - **High speed, low latency, secure, dependable**
    - **Data never flows outside the CEC and no network adapter/cabling needed**
    - **Builds on Queued Direct I/O GbE architecture**
- **Provided transparently as a part of Dynamic XCF**
  - **No additional configuration**
  - **TCP/IP stack automatically selects fastest transport between stacks across LPARs**
    - **Stack bypasses normal IP over XCF when partner stack reachable over IQDIO**





TN3270E Server of CS for z/OS

Improved Usability and Management

- TN3270 Visibility
  - IP address and Port
  - Domain name
- Cloning TN3270 clients

Continuing Enhancements

- Resource Pooling
- Full Autologon Support
- TN3270 Visibility to SNA exits and applications
- Use of client certificates for SNA logon

Security

- Secure Sockets Layer
- Client authentication via certificates

Increased Capacity and Performance

- 64,000 clients tested
- 70% increase in TPUT/CPU second
- IP QoS via Service Policy Agent
- OSA Express access

Workload Management

- DNS name to LU name mapping
- Single system image
  - DNS/WLM support
  - Sysplex Distributor

Availability

- VIPA Takeover
- Fast Reconnect
- Sysplex Distributor

Key SNA Functions

- Full USSMSG support
- Response time monitoring
- Enhanced Addressing

Enterprise Extender (HPR Over UDP)

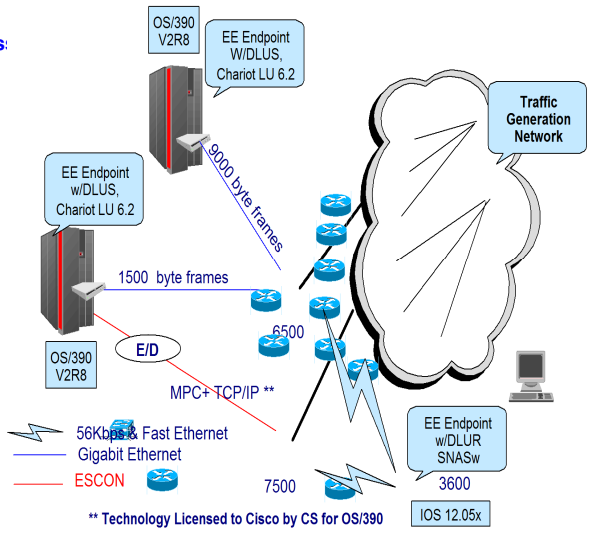
More capacity for your money

Business-critical traffic priority

24x7 Network & Application Availability

Enterprise Extender

- Enables single "wire protocol"
  - End to end SNA solution over IP transport
    - Including preserving SNA priorities
- Improves datacenter connectivity and acces
  - Exploits OSA Express and HiperSockets
- Simplifies APPN rollout
  - Significantly reduces network flows in WAN
    - Topology updates and broadcast searching
- Can replace SNI with IP technologies
  - Uses EBN connectivity
- Significant enhancements in z/OS CS V1R4
  - Dial usability enhancements
  - Performance improvements
  - Display enhancements
  - Enhanced addressing for EE PUs and Lines
- Significant enhancements in z/OS CS V1R5
  - NAT compatibility for Global VRN support
  - IPv6 support
  - Multiple VRN support
  - Multiple VIPA support



SNA Intra-Sysplex Connectivity
MPC+, XCF or EE using QDIO?

Within the sysplex, is XCF sufficient, or should I define MPC+ connections?

Is EE using QDIO an option? What about HiperSockets?

The usual answer: It depends!

- **All things being equal, MPC+ throughput should exceed XCF throughput, but using multiple XCF links can significantly increase throughput.**
- **XCF links can bypass the coupling facility, increasing throughput if the CF is being used for other functions (GR, MNPS)**
- **XCF will use more VTAM CPU cycles due to the API to the XCF facility.**

Many customers want to define both MPC+ and XCF links, but want to prefer MPC+, with XCF available for backup.

- **This can be accomplished by adding COSTBYTE=1 to the XCF TGP (in IBMTGPS) which is automatically associated with XCF TGs (assuming IBMTGPS has been activated). This makes the XCF link have a higher weight (and therefore is less desirable) than the MPC+ link for the IBM-supplied APPN Classes of Service.**

EE using QDIO is a valid option for intra-sysplex connectivity but...

- **Cross CEC traffic must go out onto the data center LAN and via 2 OSA Express cards**
- **Generally recommended for backup SNA path except for case where all SNA workload is within same CEC and HiperSockets is available**

IP Intra-Sysplex Connectivity and Network attach for Sysplex

The unusual answer: QDIO with OSA Express wins on all points

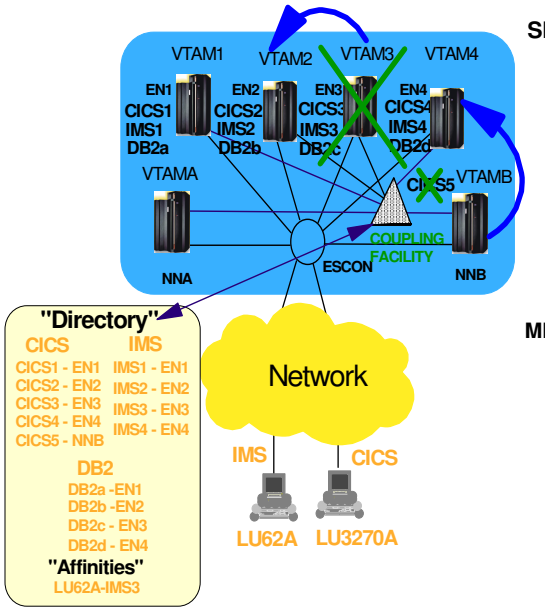
- Best for IP and EE (APPN/HPR) in terms of
  - Price
  - Throughput
  - CPU Utilization
  - Ease of implementation
- Some issues still need to be considered
  - SNA Access is only available via Enterprise Extender which requires APPN/HPR implementation
    - On z/OS and on remote Enterprise Extender capable node
      - ✓ e.g. Cisco's SNASw or IBM's Communications Servers for NT, AIX, or Linux.

Recommendations

- Begin migrating from legacy network connectivity (i.e. FEPs and ESCON) to QDIO with OSA Express by consolidating legacy traffic (i.e. non IP and non EE)
- Move TN3270E workload to zSeries to reduce networking costs, increase availability, improve manageability, and provide better scalability
- Use Enterprise Extender for non-3270 workloads to reduce networking costs, improve performance, increase availability and provide better scalability

How to balance your SNA and IP workloads so that you can maximize your availability and productivity?

Generic Resources and Multi-Node Persistent Sessions



SNA Generic Resources

- Multi- system application seen as single application to end user
  - Balances Workload within Parallel Sysplex
  - New logons not affected by application outage
- Dynamic registration performed by application on activation
  - Application is de-registered when no longer available
- Exploited by CICS,IMS,DB2, APPC/MVS and TSO/VTAM
  - Available to any SNA application

MNPS Support for SNA Sessions

- Support for both RAPI and APPC sessions
- Eliminates or reduces outage (VTAM, MVS, or hardware) impact
  - Persistence support now available for planned application workload takeover
- Requires HPR within a parallel sysplex
  - Enhanced by extending HPR outside the sysplex
- Exploited by CICS and APPC/MVS
- Exploited in IMS release Version 7
  - Known as Rapid Network Reconnect

MNPS or Generic Resources

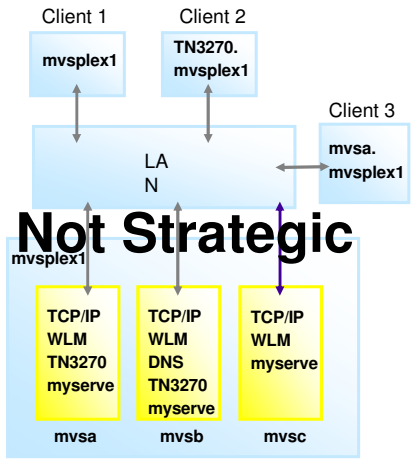
The usual answer: It depends!

- **The most important aspect to understand is whether session affinities exists for the application workloads involved**
  - **3270 workload typically mean no session affinity exists upon application outage**
    - **Generic Resources allows the end user to logon to another application instance immediately after the original application outage**
      - ✓ **Recovery is faster than MNPS and requires far less overhead during normal operations**
  - **LU62 workload typically means that a session affinity does exist upon application outage**
    - **Generic Resources is not allowed to choose another application instance upon a subsequent logon from the same end user after the original application outage**
      - ✓ **Original application must be recovered**
- **Another important aspect is to determine the importance of availability vs overall system performance**
  - **MNPS provides superior availability but does impact performance during normal operation**
    - **Storage impact, CF access, etc.**
  - **Generic Resources does not impact performance of the data path**
    - **Resolution only done during session establishment**

Recommendation:

- **Implement Generic Resources for IMS, CICS, and DB2 workloads**
- **Consider implementing MNPS if LU62 workload is critical enough to justify the extra CPU cycles**

z/OS DNS with MVS Workload Manager



What is DNS/WLM?

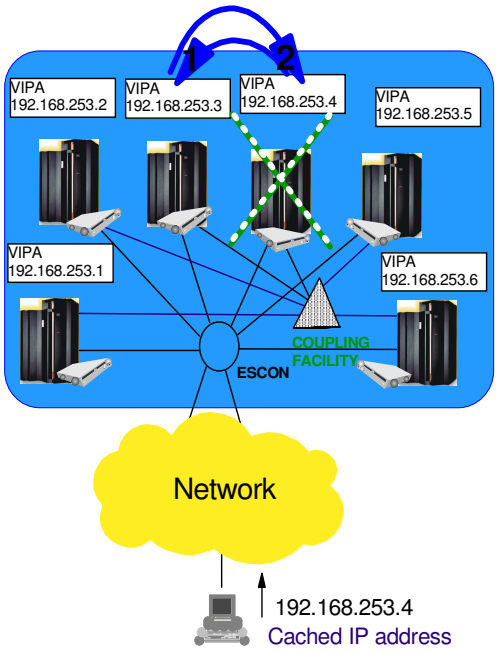
- Domain Name Service which interfaces with MVS Work Load Manager
  - Targeted for long duration connections
    - DNS resolution for every connection
  - More availability than DNS round-robin methods
    - Provided caching not done at clients or other DNS nodes
- Work load distribution on user defined goals
  - Clustered host names, server names or Weighted IP Addresses

Benefits of DNS/WLM

- Distributes connections based on current load and capacity
- Distributes load across adapters on a host
- Dynamically avoids crashed hosts and servers
  - Client can reconnect to same Server instance if required
- Dynamically avoids crashed TCP/IP stacks
  - When using sysplex name
- Highly scalable
  - New servers added without DNS administration
- Inexpensive to deploy
  - Uses existing technology

***Only available on DNS BIND 4.93 and not likely supported in the future

Virtual IP Addressing (VIPA)



Static VIPA Support
- Provides fault tolerance for network attachments
  - VIPAs can represent stack or TCP applications

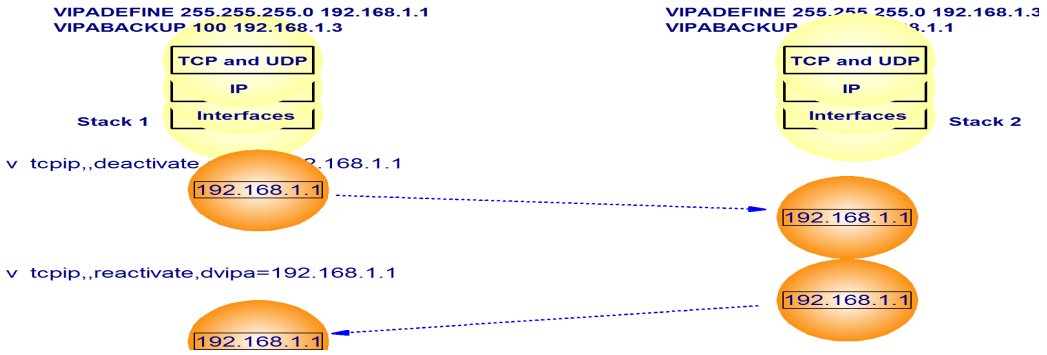
Dynamic VIPA Support
- VIPAs can survive any outage by moving to another stack in Sysplex via VIPA Takeover
  - VIPAs exchanged by TCP/IP stacks in sysplex via XCF messaging
  - Another appl instance can pick up workload or Appl can be restarted on takeover stack
  - Connections broken but Reset sent to client upon takeover
    - Significantly reduces down time

Dynamic VIPA Takeback
- VIPA moves back to recovered primary owner
  - New Connections Handled By Primary Owner again
  - Connections Established To Backup are allowed to continue
    - Data forwarded from primary owner to backup
  - Allows Movement Of Application Server Without Impacting Existing Workload
    - Allows Offload For Planned Maintenance Outages

Recent and Coming Enhancements
- z/OS CS V1R4
  - Sysplex wide SourceVIPA
- z/OS CS V1R5
  - Max number of dynamic increased from 256 to 1024
  - Allows flexible sequence of VIPA startup activation
- z/OS CS V1R6
  - Job Specific SourceVIPA
- z/OS CS V1R7
  - New operator commands to alllow for DVIPA movement
    - To designated backups or back to original owner
    - Does not require permanent configuration changes

V1R7

Sysplex Networking: Improved Operations

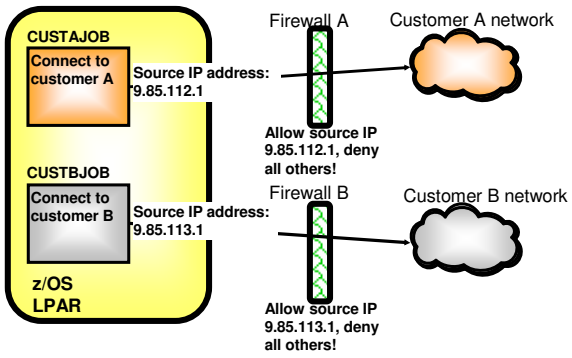


Two new commands to allow TCP/IP to relinquish or reclaim ownership of a Dynamic VIPA (DVIPA)

- Vary Tcpip,,Sysplex,Deactivate
  - DVIPA is deactivated and a configured backup stack will takeover the DVIPA
  - Backup DVIPA can be deactivated also removing eligibility as a backup
- Vary Tcpip,,Sysplex,Reactivate
  - Original owner can regain ownership
  - Can also reactivate a backup DVIPA that's been deactivated
  - Prior to these commands, Vary obey files were needed to cause a DVIPA takeover
  - These commands can't be used on a DVIPA's created from VIPARANGE with bind, ioctl(), or the Modvipa utility

V1R6

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

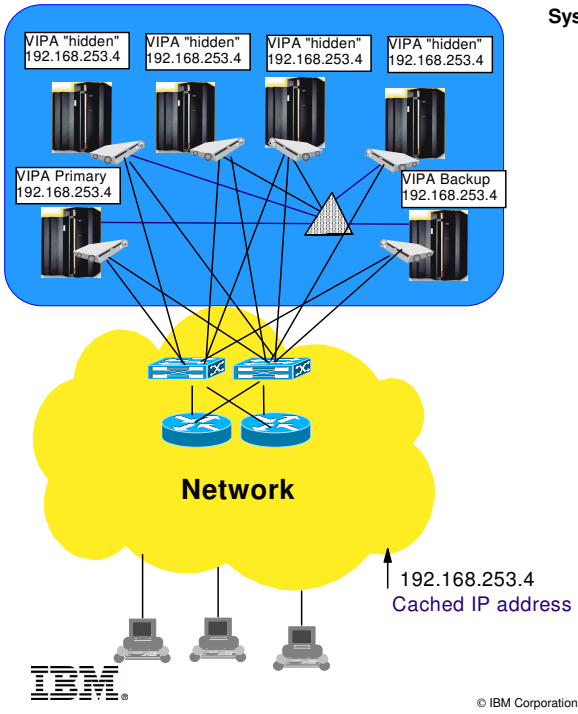



```
BEGINSRCIP
     CUSTAJOB    9.85.112.1
     CUSTBJOB    9.85.113.1
     User1*      888:555::222     ===> Wildcards allowed!
ENDSRCIP
```

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 BEGINSRCIP/ENDSRCIP allows the selection of a source IP address for outbound TCP connections by job name**
  - **Overrides TCPSTACKSOURCEVIPA and SOURCEVIPA specifications**
- ✓**Allows Single System Image view for outbound traffic initiated from z/OS Sysplex environment**

Sysplex Distributor

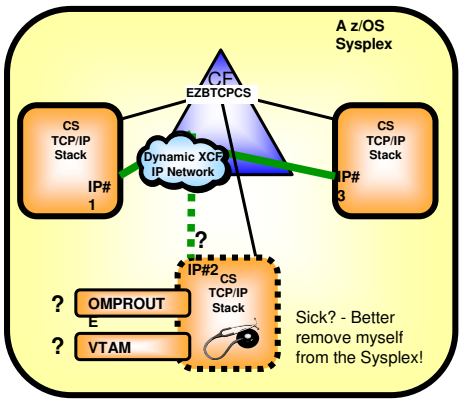


Sysplex Distributor

- **Built on Dynamic VIPA support**
  - Used primarily for non-Web workloads
    - TN3270E, FTP, DB2, CICS, etc.
- **Application centric load balancing**
  - Unavailable applications are never chosen
    - Stacks communicate application availability via XCF messaging
  - Decision based on WLM, QoS, and Policy inputs
    - Ensure SLAs are satisfied for each application as well as each client group
- **Can communicate with Cisco network**
  - Sysplex Distributor becomes MNLB service manager
    - QoS and Policy info included in criteria for best server
    - Direct path for inbound data provided by routers and switches
- **Enhancements in z/OS CS V1R4**
  - Adds distribution support for passive mode FTP
    - Already supported other modes of FTP
  - Increase distributed port limit from 4 to 64
- **Enhancements in z/OS CS V1R5**
  - Application based affinity
    - Timer based stickiness
  - Round Robin distribution
    - Can overide WLM recommendation to ensure evenly distributed workloads
- **Enhancements in z/OS CS V1R6**
  - Sysplex Autonomics
- **Enhancements in z/OS CS V1R7**
  - Improved Operations support, support for non-XCF forwarding, improved load balancing decisions

TCP/IP Sysplex autonomics

V1R6



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

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

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

- *Timersecs* - used to determine duration of the troubling condition before issuing messages or leaving the sysplex
- *Recovery* - TCPIP removes itself from the sysplex(recommended).
- *NORecovery* - TCPIP does not remove itself from the sysplex (default)
- *AutoRejoin* - allows TCP/IP to rejoin the sysplex if the problem conditions are relieved (***New*** for z/OS V1R7)

Messages are always issued to the console when these conditions are detected regardless of SYSPLEXMONITOR Recovery specification
Messages are eventual action (deleted when the action is taken or problem is resolved)

New operator command is provided to allow TCPIP to leave the sysplex (ie. EZBTCPCS xcf group)
Vary TCPIP,,SYSPLEX,LEAVEGROUP

To have TCPIP rejoin the sysplex group, a Vary Obey of the TCPIP profile with sysplex configuration statements is needed.
Severe problems may require a TCPIP stack restart

V1R7

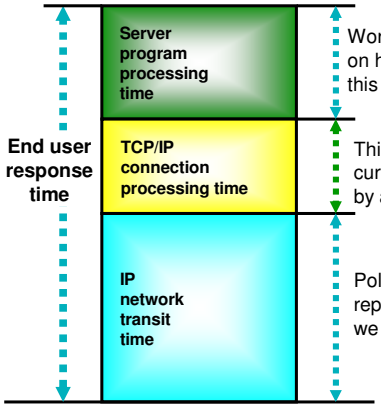
Sysplex Distributor: Improved Operations

Two new commands to gracefully quiesce a target system or application that are targets for Sysplex Distributor

- Ability quiesce a system/application prior to shutdown
  - Planned maintenance scenarios of system/application
    - Allows existing systems/applications to drain work queue prior to shutdown
  - Relieve temporary constraints of resources on target system
  - Temporary - Does not affect Sysplex Distributor's permanent configuration
  - Issued on target system being affected
- Vary Tcpip,,Sysplex,Quiesce,*option*
  - **TARGET** - Quiesces all applications on target stack.
  - **PORT=xxx -** Quiesce all applications bound to the specified port on this stack
    - **JOBNAME=jobname -** Allows quiesce of a single application in SHAREPORT group
    - **ASID=asid -** Further qualify job being quiesced (i.e. deal with duplicate jobnames)
  - No new TCP connections sent to the quiesced target (stack or application)
    - For all Distributed DVIPAs that the entity is a target for
  - Existing TCP connections are maintained (i.e. non-disruptive)
- Vary Tcpip,,Sysplex,Resume,*option*
  - **TARGET|PORT|JOBNAME|ASID**
  - Allows identified target stacks and/or applications to once again be targets for distribution

Sysplex Distributor
Improving the load balancing decision

V1R7



Workload Manager (WLM) reports on how well we meet targets for this portion

- Prior to z/OS V1R7, WLM recommendations reflected overall system capacity
- In z/OS V1R7, WLM recommendations have been enhanced to reflect how well individual servers are performing (compared to WLM policy goals)

This portion is currently not covered by any reporting

Sysplex Distributor now monitors how well target applications/systems are handling new TCP connections

- Are servers accepting new connections fast enough (i.e. is the backlog queue building?)
- Are connections being established? Is a network path back to the client available?

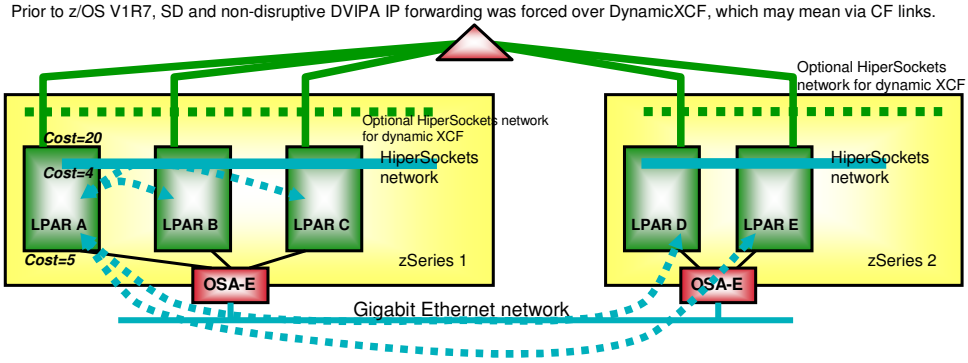
Policy Agent QoS reports on how well we do in this portion

- Ideally, when a load balancer makes decisions about which nodes and server instances to forward new workload to, it would factor in all elements that contribute to the final overall end user response time.
- Sysplex Distributor factors in server program processing time via WLM weights and network transit time via QoS fractions that are used to modify the WLM weights.
- Any delays caused by connection setup processing, such as forwarded connections that are lost due to lost network connectivity between the distributor and the target stack, or connections being dropped by the target stack due to a backlog queue full condition, are currently (prior to V1R7) not factored in by Sysplex Distributor

V1R7

Intra Sysplex IP communication improvements

Prior to z/OS V1R7, SD and non-disruptive DVIPA IP forwarding was forced over DynamicXCF, which may mean via CF links.



- In z/OS V1R7, SD and non-disruptive DVIPA IP forwarding can be directed to use any available IP network between the LPARs in a sysplex.
  - Can help alleviate constraints on Coupling Facility links
  - OSA Express over GigE provides good performance
  - Provides greater fault tolerance (any network route can be used)
  - Can be configured as needed (e.g. for inter-CEC communications but not for Intra-CEC communications)
- DynamicXCF definitions still required:
  - DynamicXCF links must still be defined in z/OS V1R7
  - Will always be used for Sysplex Wide Security Association (IPSec) packets and Multi Level Security (MLS) tagged packets

Messages are always issued to the console when these conditions are detected regardless of SYSPLEXMONITOR Recovery specification
Messages are eventual action (deleted when the action is taken or problem is resolved)

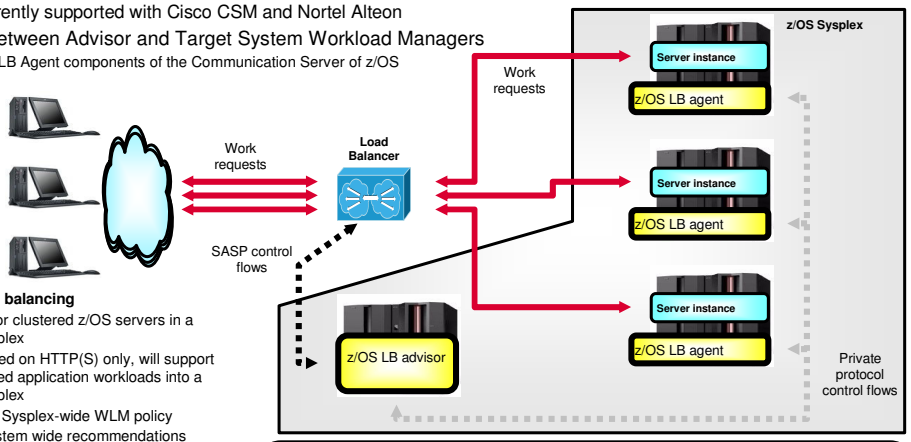
New operator command is provided to allow TCPIP to leave the sysplex (ie. EZBTCPCS xcf group)
Vary TCPIP,,SYSPLEX,LEAVEGROUP

To have TCPIP rejoin the sysplex group, a Vary Obey of the TCPIP profile with sysplex configuration statements is needed.
Severe problems may require a TCPIP stack restart

z/OS Load Balancing Advisor (LBA)

A z/OS SASP implementation compatible with EWLM specification

- Load Balancers work with both z/OS WLM and EWLM environments simultaneously
  - ✓Server/Application State Protocol (SASP) - Open protocol
  - ✓Currently supported with Cisco CSM and Nortel Alteon
- Flows between Advisor and Target System Workload Managers
  - ✓New LB Agent components of the Communication Server of z/OS



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
  - ➔ System wide recommendations
  - ➔OR Server specific recommendations (z/OS V1R7 only)
- Server health also taken into account
  - ➔Examine TCP backlog queues, UDP receive queues

The z/OS Load Balancing Advisor technology is a new z/OS Communications Server technology that was just made available on z/OS V1R4 and higher:

PQ90032 for z/OS V1R4
PQ96293 for z/OS V1R5 and V1R6
Part of z/OS V1R7 CS
For more details, see *www.ibm.com/support/docview.wss?uid=swg27005585*

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)

Which IP workload distribution model(s) do you choose?

The usual answer: It depends!

- **No single model satisfies all customer requirements**
  - Some models don't meet availability requirements while others don't meet performance requirements
  - Customer load balancing requirements will become more, rather than less, severe over time
    - Faster adjustment to cluster changes, finer degrees of granularity, SLA enforcement, etc.
- **Several configuration issues limit available choices**
  - Customers view 2216 Interactive Network dispatcher as a 'dead end'
    - Similar implementations from other vendors may not have access to WLM information or may require special networking equipment
  - AIX Interactive Network Dispatcher generally not cost acceptable
    - Forces an "extra" box in the data center
- **DNS methods are easy to implement but have availability issues**
  - DNS round-robin and DNS weighted round-robin methods have no access to WLM information
    - Knowledge of "down" servers or true server capacity is not available
  - DNS/WLM on z/OS have client dependencies for maximum availability and is not strategic
    - Clients must not cache resolved IP address
      - ✓If z/OS is secondary DNS then primary DNS must not cache either
    - Only available on DNS BIND 4.9.3
      - ✓z/OS CS will likely not ship this version of DNS BIND beyond z/OS V1R7
- **Key Considerations**
  - What functions are needed? Which solution provides them?
  - Who will administer the solution? z/OS administrators or Network administrators?
  - Performance and Availability requirements

Which IP workload distribution model(s) do you choose?

Recommendations

- **Decide on type of solution (Internal vs External Load Balancing solution)**
- **Internal Load Balancing (Sysplex Distributor)**
  - **Well suited to non-web traffic**
  - **Built on dynamic VIPA support for simple implementation and maximum availability**
  - **No new equipment required, maintains administrative controls within z/OS**
- **External Load Balancing Solutions**
  - **Typically better suited for workload distribution of Web traffic**
    - **Allows distribution to outboard caches as well as target z/OS systems**
  - **Can handle non-web traffic and non z/OS platforms**
  - **May or may not require additional equipment**
  - **Administrative functions typically performed by network administrator**
  - **Selecting a solution that supports SASP can greatly improve the quality of the load balancing decision**
    - **Using the z/OS Load Balancing Advisor significantly improves visibility into the Sysplex environment and the current workload conditions**
    - **SASP will be an open architecture so it is expected that there will be a wider range of load balancer choices in the future**

For More Information....

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

Trademarks, Copyrights and Disclaimers