IBM®

zSTSU 2004

Session 3601

z/OS Communications Server Enhancements in V1R5 and V1R6

October 7, 2004

Washington Systems Center, Gaithersburg, MD 20879
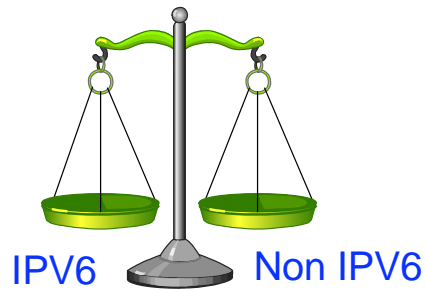Gwen Dente - gdente@us.ibm.com

@server

---

z/OS V1R5 & V1R6 Communications Server

z/OS V1R5 and V1R6 continue the effort started in z/OS V1R4 to provide IPV6 on z/OS and provide much Non-IPV6 function also!

★ IPv6 support is technology-driven in preparation for future
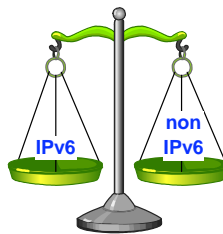
IPV6   Non IPV6

V1R5 - Continue to provide a balance of IPv6 and non-IPv6 functions

IPv6 focus in V1R5:

- ✓ Applications
- ✓ APIs
- ✓ DLCs
- ✓ Network management
- ✓ Policy
- ✓ Routing
- ✓ Security
- ✓ Enterprise Extender

non-IPv6 focus in V1R5:

- ✓ TN3270 server enhancements
- ✓ FTP enhancements
- ✓ Sysplex enhancements
- ✓ Network management
- ✓ Policy
- ✓ Security
- ✓ Msys
- ✓ SNA enhancements
- ✓ Enhancements for new hardware
- ✓ Routing enhancements
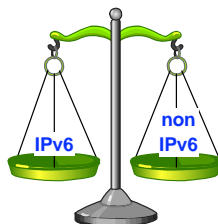- ✓ Mail

V1R6 - Continue to provide a balance of IPv6 and non-IPv6 functions

IPv6 focus in V1R6:

- ✓ Dynamic Routing
- ✓ Hipersockets
- ✓ Network management
- ✓ Sysplex

non-IPv6 focus in V1R6:

- ✓ TN3270 server enhancements
- ✓ FTP enhancements
- ✓ Sysplex enhancements
- ✓ Network management
- ✓ Policy
- ✓ Security
- ✓ SNA enhancements
- ✓ Enhancements for new hardware

Agenda

z/OS V1R5 Communications Server
- Overview
- Sysplex
- Applications
- Security
- Other TCP/IP Enhancements
- Enterprise Extender and SNA

z/OS V1R5

z/OS V1R6 Communications Server
- Overview
- Sysplex
- Applications
- Security
- Other TCP/IP Enhancements
- Enterprise Extender and SNA

z/OS V1R6

z/OS V1R5

z/OS V1R6

APPENDIX:
IPv6 Since V1R4 and Futures

The information in this presentation represents information delivered at SHARE in NY, August 2004, by the IBM Communications Server Strategy and Design group in Raleigh, NC:
Mac Devine, Sam Reynolds, Alfred Christensen.

---

z/OS Communications Server
Disclaimer

➤ Plans for the z/OS Communications server are subject to change prior to general availability

➤ Information provided in this presentation may not reflect what is actually shipped for z/OS Communications Server

➤ This presentation includes an early overview of future z/OS Communications Server enhancements

z/OS
CS

Note: Plans are subject to change!
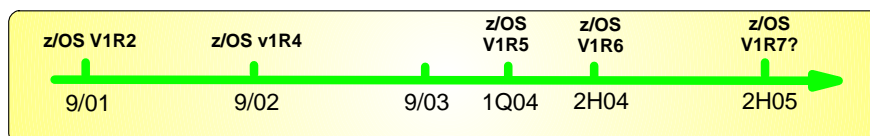
z/OS V1R2	z/OS v1R4		z/OS V1R5	z/OS V1R6	z/OS V1R7?
9/01	9/02	9/03	1Q04	2H04	2H05

Enhancements in CS for z/OS V1R5

z/OS V1R5

IBM.

Sysplex Enhancements

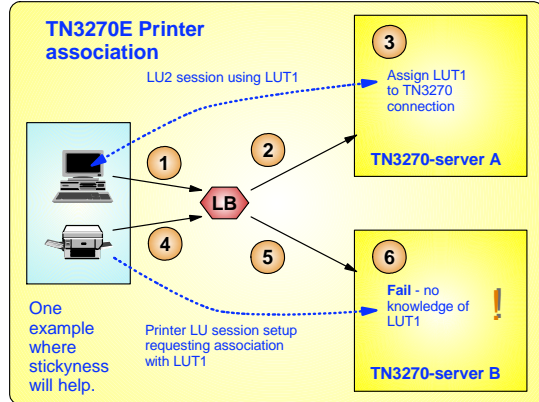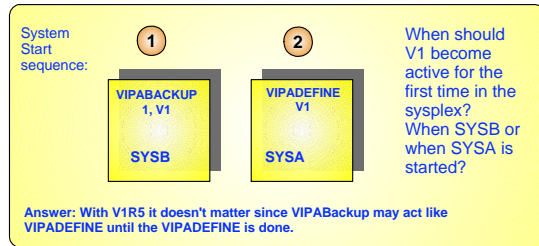z/OS V1R5

IBM.

Sysplex enhancements in z/OS V1R5

Sysplex Enhancements

- Increase ports on VIPADISTRIBUTE from 4 to 64 (PTFed back to z/OS V1R2 - APAR PQ65205)

- Dynamic port definition for VIPADISTRIBUTE dynamic VIPA when server binds to dynamic VIPA

- Increase limit of DVIPAs per stack from 256 to 1024

- Support DVIPA activation based on VIPABACKUP before VIPADEFINE ever processed

- Sysplex Distributor affinity
  - Configurable timer-based stickyness per source IP address, server DVIPA and port

- New round-robin distribution method in Sysplex Distributor (PTFed back to z/OS V1R4 - APAR PQ76866)
  - Alternative to WLM-based distribution
  - Useful where availability is more important than capacity

System Start sequence:

1 VIPABACKUP 1, V1 — SYSB

2 VIPADEFINE V1 — SYSA

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 may act like VIPADEFINE until the VIPADEFINE is done.

TN3270E Printer association

LU2 session using LUT1

3 Assign LUT1 to TN3270 connection — **TN3270-server A**

1 **2** LB

4 **5**

6 **Fail** - no knowledge of LUT1 — **TN3270-server B**

One example where stickyness will help.

Printer LU session setup requesting association with LUT1

---

zSTSU: 2004

Application Enhancements

IBM.

FTP enhancements in z/OS V1R5

z/OS V1R5

FTP File Transfer

- ➤ **Support allocation of target PDS(E) like existing source PDS(E)**

- ➤ **Enable user to specify if a PDS or a PDSE should be created**

- ➤ **Allow configuration control of passive data port range on FTP server**

- ➤ **Support use of extended passive (EPSV) and active mode (EPRT) for NAT firewall relief**

- ➤ **Provide consistent error codes from FTP client**

- ➤ **Enable MVS syslog message when batch FTP client operation fails (for automation purposes)**

- ➤ **Allow SSL/TLS login to FTP server without password when client authentication is used**

- ➤ **Deliver FTP server load module as RMODE=ANY**
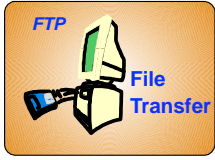
Secure FTP and Network Address Translation (NAT) firewalls in z/OS V1R5
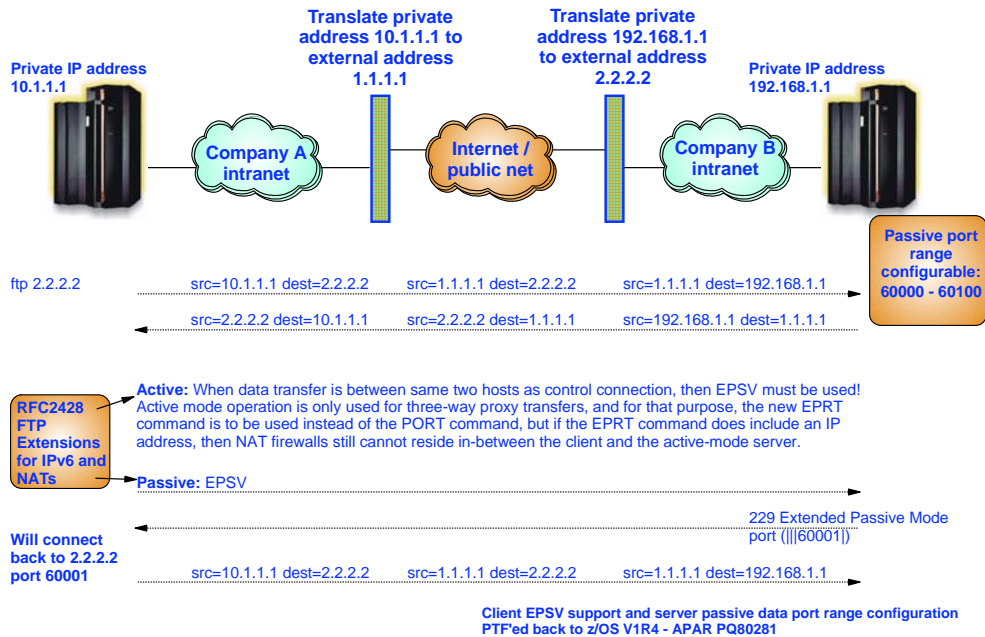
z/OS V1R5

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

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 |

RFC2428 FTP Extensions for IPv6 and NATs

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

Will connect back to 2.2.2.2 port 60001

229 Extended Passive Mode 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 |

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

TN3270 server enhancements in z/OS V1R5

➤ **Improve Tn3270/e connection recovery**
  ▸ End user need not know LUNAME to recover session

Remote terminal access

➤ **Provide SNMP network management data for tn3270/e**
  ▸ Support in new Enterprise-specific TN3270 MIB and new TN3270 subagent
  ▸ Focus on connection performance data
  ▸ Provides end-to-end stats over Tcp/Ip and SNA for connections
  ▸ Data accessible via MIB or D TCPIP,,TELNET,CONN,CONN=connid detail command

➤ **Enhance usability with the following items:**
  ▸ Support range of ipaddress on IPGroup in profile
  ▸ Better compatibility between type ahead-capable clients and applications
    – UNLOCKKEYBOARD statement in PROFILE.TCPIP

➤ **Performance enhancement for definite Rsp. sessions**
  ▸ Server turns off delay acks for clients that have negotiated it. Client can do many small sends without waiting for acks from server.

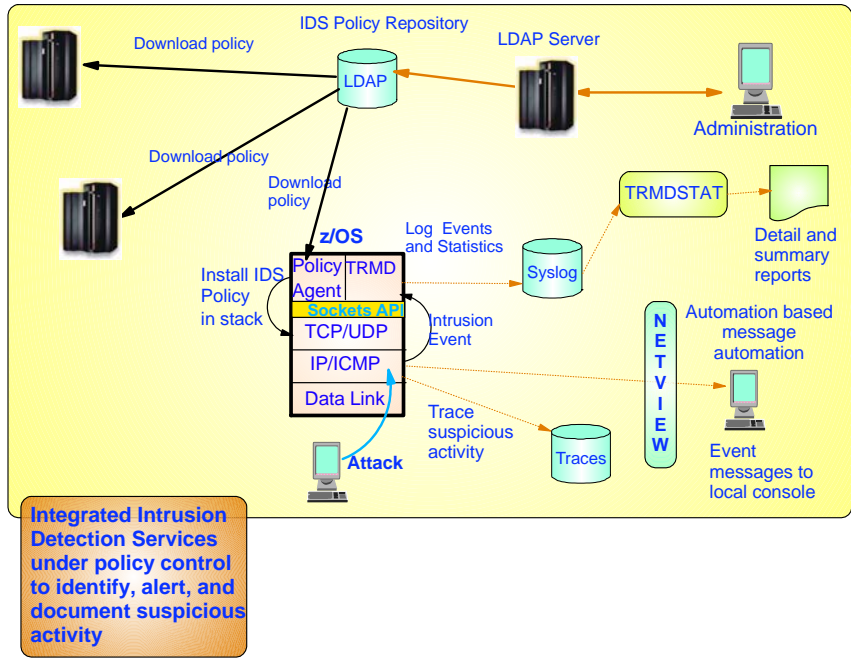---

zSTSU: 2004

Security

IBM.

Intrusion Detection Services enhancements

z/OS V1R5

Intrusion Detection Services

▸ Added new Policy-based interface flood detection
  – Detects floods of "garbage" data on a per-interface basis (ex:)
    ★ Unsupported protocol
    ★ Examine packet discards

IDS Policy Repository

Download policy

LDAP Server

LDAP

Administration

Download policy

Download policy

z/OS

Install IDS Policy in stack

Policy Agent | TRMD
Sockets API
TCP/UDP
IP/ICMP
Data Link

Log Events and Statistics

Syslog

TRMDSTAT

Detail and summary reports

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

---

# Multi-Level security enhancements

z/OS V1R5

➤ **Governments require control over information accessed through computer networks**

  ▸ Applications/users and networks are classified and assigned a level of security
    ▸ Classifications are done in RACF with SECLABEL

  ▸ MLS for Communication's Server TCPIP stack is based upon the existing NETACCESS function
    ▸ IP addresses, both INBOUND and OUTBOUND are defined to be in a security zone
    ▸ Security zone is defined to RACF using a SERVAUTH class profile
    ▸ User may be RACF PERMITted to access the SERVAUTH profile (ie. security zone)
    ▸ In an MLS environment, user and SERVAUTH profile (zone) are assigned a SECLABEL

  Key: For a user to communicate with a partner in a security zone, SECLABELS must be equivalent (user = zone)

  user = CICS SECLABEL HIGH

  ZONEB SECLABEL HIGH

  ZONEK SECLABEL TOPSEC

  9.27.100.1

  TCPIP Profile:
  NETACESS INBOUND OUTBOUND
   9.27.100.0/24 ZONEB
  ENDNETACESS
  In RACF:
  RDEFINE SERVAUTH ZONEB
  PERMIT CICS ZONEB
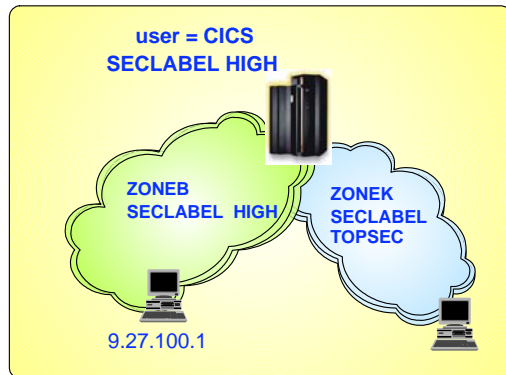
MLS
notes

N O T E S

- ➤ For z/OS users communicating in an MLS environment across a sysplex (IUTSAMEHOST and XCF links), network packets carry the SECLABEL of the sender and are checked for SECLABEL equivalence at the destination host
- ➤ Packets sent /received outside of a sysplex are assume to have equivalent SECLABELS if NETACESS checks are successful

- ➤ NETACCESS and MLS for Applications
  - ▸ FTP Poe - RACF Port of Entry (PoE) checks are done to determines if client/user is permitted to logon. Applies to MLS and nonMLS environments.
    - Determines if the client's NETACESS zone is permitted to the zone of the logon userid.
    - For MLS, client and logon userid must have equivalent SECLABELs.
  - ▸ Tn3270 server Poe - determines if the client's NETACESS zone has an equivalent SECLABEL to the Tn3270 LUname selected
    - LUGROUP gets a SECLABEL
  - ▸ Tn3270 can also determine if a client is NETACESS permitted to a Tn3270 port
    - NonMLS
    - Port gets a security zone assigned in Tn3270 profile

---

zSTSU: 2004

Other TCP/IP Enhancements

IBM.

Network Management Instrumentation Architecture

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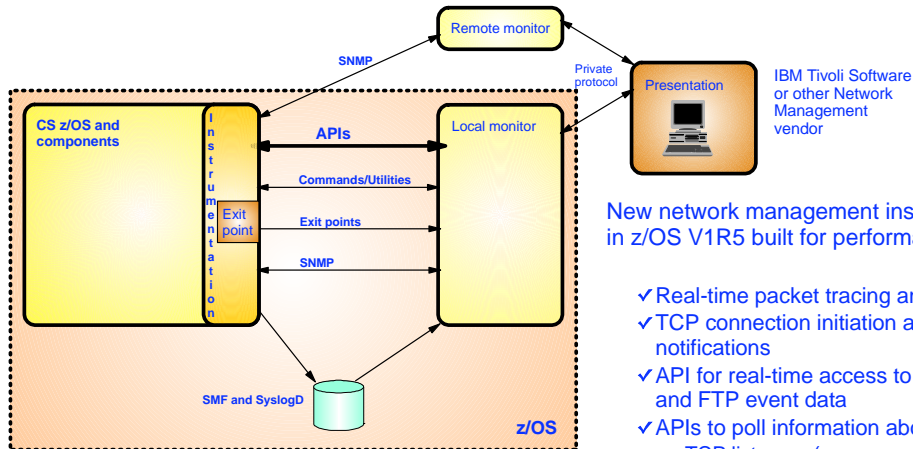
New network management instrumentation APIs in z/OS V1R5 built for performance:

- ✓ Real-time packet tracing and formatting
- ✓ TCP connection initiation and termination notifications
- ✓ API for real-time access to TN3270 server and FTP event data
- ✓ APIs to poll information about currently active
  - – 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

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.

Other TCPIP enhancements

z/OS V1R5

➤ **OMPROUTE will be changed in this release to:**
  - ‣ Improved trace performance by using CTRACE
  - ‣ Multipath enhanced to support 16 routes to same destination
  - ‣ Support advertisement for more than 254 interfaces (no upper limits)
  - ‣ Some display clean-up to show N/A where an option doesn't apply
  - ‣ Ignore undefined interfaces (don't have to define everything to OSPF)

➤ **Enhance MVS system symbol resolution**
  - ‣ In TCPIP.DATA
    - ‣ No longer requires use of EZACFSM1 preprocessor utility
  - ‣ In Resolver setup file

➤ **Increase Max. Number of allowed sockets from 2000 to 65535**
  - ‣ Macro API (EZASMI), Sockets Extended (callable EZASOKET), CICS Sockets (callable EZASOKET), and IMS Sockets

➤ **Improve performance for asynchronous stream socket receive operations**
  - ‣ LE C/C++ API
  - ‣ z/OS UNIX Assembler Callable Services Sockets API

➤ **MSys support for complete configuration of the FTP client and server**

Full VLAN support by z/OS and other OSA enhancements for z990

z/OS V1R5

- ➤ Provide Full VLAN support for OSA-Express
  - z/OS V1R2 supported only VLAN priority-tagged frames
    - ▸ supported via Policy Agent configuration statement as parameter on SetSubnetPrioTosMask
  - Extend VLAN in z/OS V1R5 to support non-null VLAN ID (Full VLAN) with OSA Express
    - ▸ Supported for OSA Express QDIO Gigabit Ethernet and Fast Ethernet with IBM Server zSeries z900 (OSA driver level 3G), and z990
    - ▸ VLAN ID is configured in TCPIP profile on the LINK (IPv4) or Interface (IPv6) Statement
    - ▸ Supports only one VLAN ID for IPv4 interface and one VLAN ID for IPv6 (IDs can be different) per OSA port
- ➤ OSA Express QDIO performance improvements for z990
  - Offload checksum to Improve performance for IPv4 packets using QDIO mode
  - Provide control of storage utilization inbound per Tcpip link/interface
  - Provide some controls over the inbound performance per link/interface - latency vs. CPU cost
- ➤ Broadcast support for Hipersockets Interface (IPv4) for z990

- ▸ IEEE standard 802.1Q defines operation of Virtual Bridged LANs
- ▸ Defines additional information in the MAC frame (tag carries VLAN-ID)
- ▸ Each port on the VLAN-aware bridge has a VLAN-ID
- ▸ VLAN-aware bridges can be configured to:
  - ✓ Admit only VLAN-tagged frames (i.e.. have both priority and non-null VLAN ID)
  - ✓ Admit all frames

★ Allows packets to be isolated to selected network interfaces and physical networks and adapters to be shared for multiple purposes, such as various security zones or customers...

---

zSTSU: 2004

Enterprise Extender and SNA Enhancements

z/OS V1R5

IBM.

NAT Compatibility with EE VRN

z/OS V1R5

hostx = 1.1.1.2
EBNx = 1.1.1.1

x.com public DNS

y.com public DNS

hosty = 2.2.2.2
EBNy = 2.2.2.1

LOCATE Reply(...CV46(...A5(192.168.2.1))

EBN
EBNx.x.com

x.com internal DNS

hostx = 10.2.1.1
EBNx = 10.1.1.1

NNx

ENx
dest=192.168.2.1 ?

hostx.x.com

Company x.com Intranet

FW

FW

EBN
EBNy.y.com

y.com internal DNS

hosty = 192.168.2.1
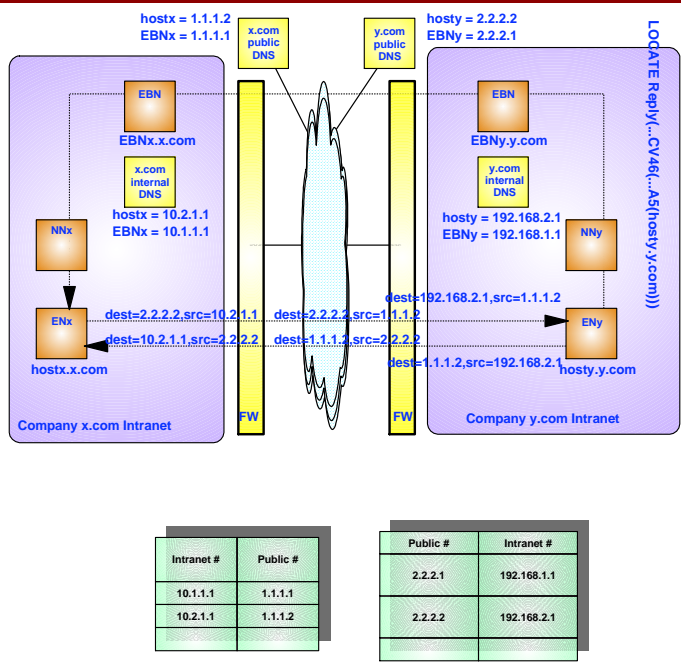EBNy = 192.168.1.1

NNy

ENy

hosty.y.com

Company y.com Intranet
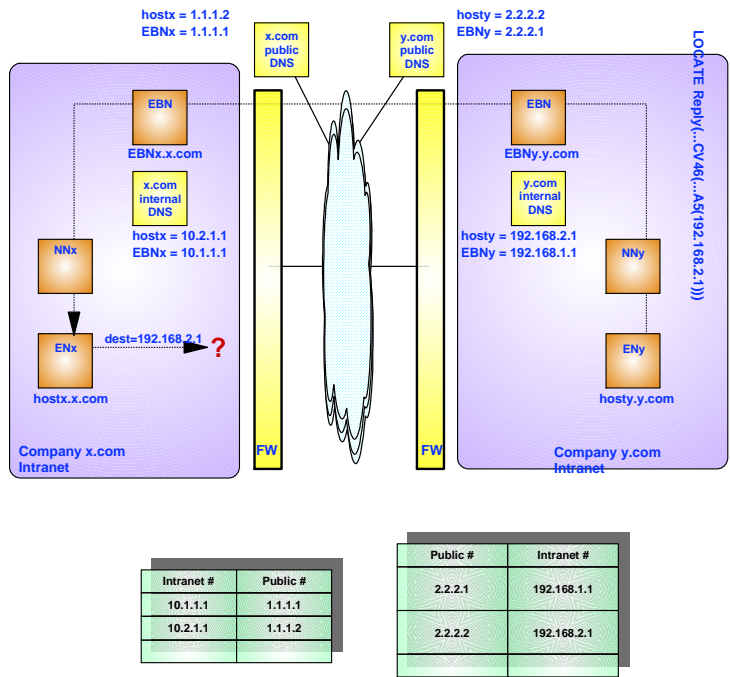
➢ The control vectors representing an EE connection network carry the IP address of the node that owns the vectors.

➢ Applications that carry IP addresses in their data stream are fundamentally incompatible with Network Address Translation (NAT).

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

Enhanced HOSTNAME Support

z/OS V1R5

hostx = 1.1.1.2
EBNx = 1.1.1.1

x.com public DNS

y.com public DNS

hosty = 2.2.2.2
EBNy = 2.2.2.1

LOCATE Reply(...CV46(...A5(hosty.y.com)))

EBN
EBNx.x.com

x.com internal DNS

hostx = 10.2.1.1
EBNx = 10.1.1.1

NNx

ENx
dest=2.2.2.2,src=10.2.1.1   dest=2.2.2.2,src=1.1.1.2
dest=10.2.1.1,src=2.2.2.2   dest=1.1.1.2,src=2.2.2.2

hostx.x.com

Company x.com Intranet

FW

FW

EBN
EBNy.y.com

y.com internal DNS

hosty = 192.168.2.1
EBNy = 192.168.1.1

NNy

ENy

dest=192.168.2.1,src=1.1.1.2
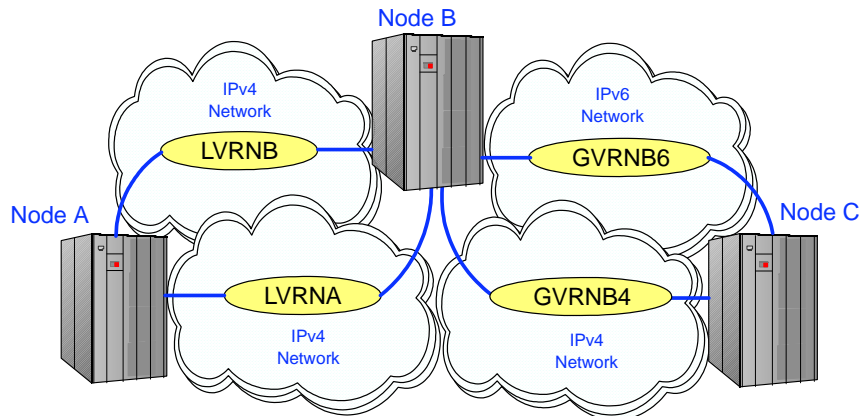dest=1.1.1.2,src=192.168.2.1

hosty.y.com

Company y.com Intranet

➢ 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.

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

Multiple VRN/VIPA Support

Node B

IPv4 Network

LVRNB

IPv6 Network

GVRNB6

Node A

Node C

LVRNA

GVRNB4

IPv4 Network

IPv4 Network

- 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

Enterprise Extender Model PUs

- Currently, unless the Configuration Services XID exit is used, EE does not allow the user to code a model for non-connection- network dynamic PUs created on the dial-in side.
  - If DYNPU=YES is specified, EE creates a "vanilla" dynamic PU.
- A new model type, DYNTYPE=EE, is added to the PU statement in the Model Major Node.
- An EE model PU allows for customizing dynamic non-connection network PUs in various manners:
  - Coding the DISCNT operand so that a disconnect delay time can be specified.
  - Specifying the DWINOP, REDIAL, and REDDELAY operands to specify whether to drive redial attempts, and how often and how many times to attempt redial, when INOPs occur for the connections using these dynamic PUs.
  - Overriding the default TG characteristics that in the past were used for these dynamic PUs.

```
MODELA1A VBUILD TYPE=MODEL
*
EEMODEL  PU      DYNTYPE=EE,
                 CAPACITY=100K,
                 COSTTIME=5,
                 CPCP=YES,
                 DISCNT=NO,
                 DWINOP=YES,
                 REDIAL=30,
                 REDDELAY=60
```

D RTPS Enhancement

- In V1R5, the DISPLAY RTPS command is enhanced to:
  - Allow the display of all RTP pipes using a particular link for the first hop. The link can be specified by either:
    - ALSNAME
    - TG number & partner CP name
  - Allow the display of all RTP pipes going to a particular Net ID (either as the destination or as the first hop). This is done by specifying *netid.* where a CP name is specified.

```
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,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
```

Other EE/SNA enhancements

➢ When dial failures occur across a connection network, the virtual routing node state is usually not affected, so the APPN topology and routing component continues to select the VRN if it is the preferred path. V1R5 adds a more specific failure message to indicate that a dial across a VRN has failed, providing an easier trigger for customer notification and automation:

```
IST1903I FAILURE OVER VRN NETA.SSCPVN TO CP NETA.SSCP2A
```

➢ V1R5 adds new messages for accounting and problem determination capability for DLUR-based sessions
  - Messages issued when a DLUR-served PU session is established or terminated with VTAM:
  - New message to include DLUR name on ACTPU and ACTLU failures
➢ Support for multiple concurrent APINGs
➢ Remove restriction that the backup NNS must be in same sysplex as the End Node for generic resource to work (allows the End Node to be searched for the resource)
➢ HPR Resequencing Optimization
  - Improves inbound processing of out of order segmented packets
➢ Modify IOBUF pool exapansion limit without recycling VTAM

Other EE/SNA Enhancements...

z/OS V1R5

Serviceability/Problem Determination
- Usability improvements to CSDUMP command (dump for sense code or message)
  - Display settings of CSDUMP command and allow for the deletion of triggers
- APPN Trace enhancement
  - New SUBTRACE option TGVC will provide TG Vectors in appropriate trace records
- Track CSM Buffers internally through components
- Enable dumping for VTAM Inoperative conditions with new Modify INOPCode command
  - Works in conjuction with existing Modify INOPDump command
- Allow for extended information to be displayed for search failures that cause sessions not to be established: DSIRFMSG start option

Performance/Storage
- Multipath Channel HPDT Packing (nonXCF) examines packet sizes for optimal outbound transmission
  - Packing turned on/off at TRLE. Requires some workload examination for tuning.

New/Enhanced Functionality
- Sift parameters from the Group on model major node for dynamic switched LUs
  - Allows for definition reduction
- SWNORDER and DLRORDER
  - Enhanced to allow greater control over PU selection during connection processing. On start option or in XCA or NCP major nodes.
- New search option, SSCPORD, provides granular control over searches for resources
  - defined on start option or in ADJSSCP table

Improving EE Performance

z/OS V1R5

- 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 to at least 0326
  - CPU utilization: Some customer environments will benefit from OA04393 (Inactivity Timer Optimization)
- Monitor the EE Informational APAR (II12223) for news on further enhancements

Enhancements in CS for z/OS V1R6

z/OS V1R6

IBM.

Sysplex Enhancements

z/OS V1R6

IBM.

Single system image (SSI) from an IP perspective in the Sysplex since V1R4

Connect to DRVIPA1

Inbound SSI

Sysplex

z/OS   z/OS   z/OS

Connect to ? from SRCVIPA1   **Connect to ? from SRCVIPA2**   **Connect to ? from SRCVIPA3**
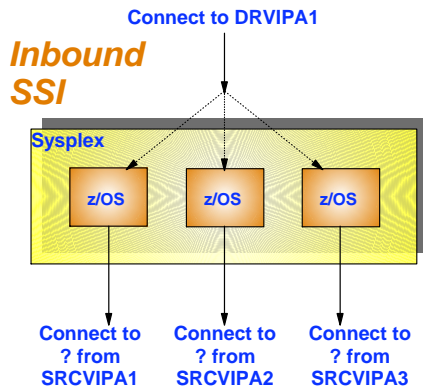
- We have single system image capability for inbound connections where a single distributed VIPA address can represent all images in the Sysplex - and remote users do not need to select a specific image when connecting to their server application.
- But if we establish outbound connections from the images 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.
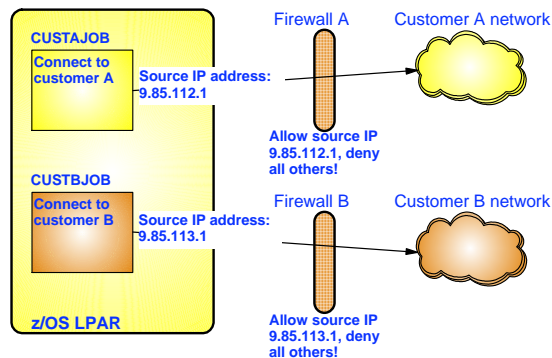
- ▸ z/OS V1R4 introduced new capabilities that allow a single Sysplex-wide source VIPA address to be used for outbound TCP connections by all images in the Sysplex - resulting in single system image capabilities for both inbound and outbound connections.
- ▸ z/OS V1R5 significantly improved the performance when using a Sysplex-wide source VIPA
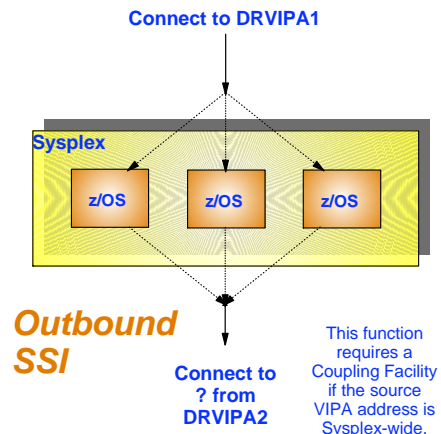
Connect to DRVIPA1

Sysplex

z/OS   z/OS   z/OS

Outbound SSI

Connect to ? from DRVIPA2

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

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

z/OS V1R6

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

```
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
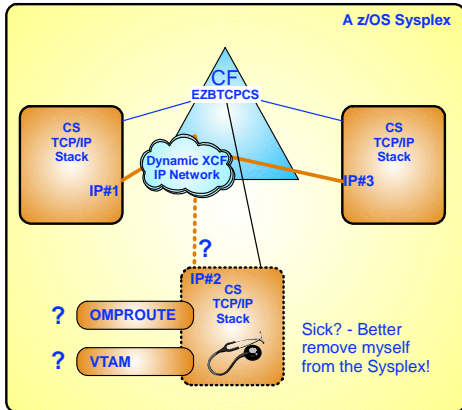   - ▸ Overrides TCPSTACKSOURCEVIPA and SOURCEVIPA specifications

TCP/IP Sysplex autonomics in z/OS V1R6 to let TCP/IP proactively handle error conditions

z/OS V1R6

A z/OS Sysplex

CF
EZBTCPCS

CS TCP/IP Stack — IP#1

CS TCP/IP Stack — IP#3

Dynamic XCF IP Network

IP#2 CS TCP/IP Stack

? OMPROUTE

? VTAM

Sick? - Better remove myself from the Sysplex!

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

- 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
  DELAYJOIN|NODELAYJOIN
```

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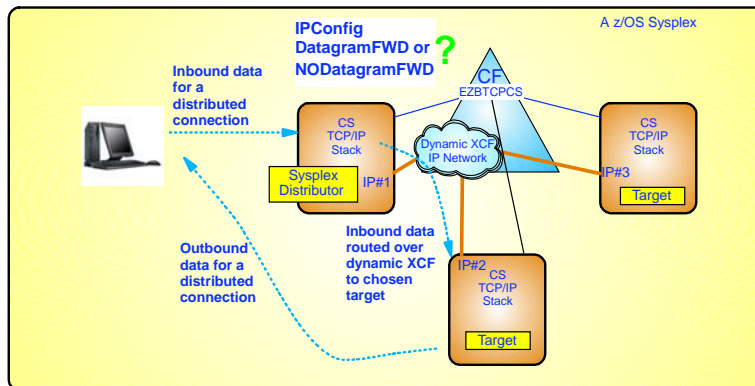- *NODELAYJOIN* - Join Sysplex immediately

---

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

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
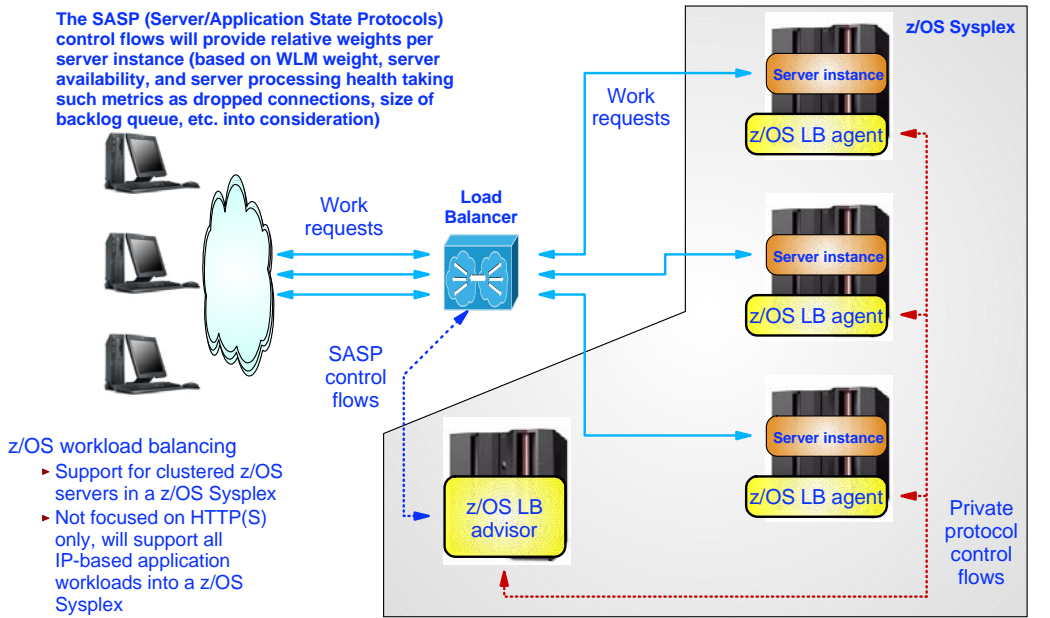

IPConfig DatagramFWD or NODatagramFWD ?

A z/OS Sysplex

Inbound data for a distributed connection

CF EZBTCPCS

CS TCP/IP Stack — Sysplex Distributor — IP#1

Dynamic XCF IP Network

CS TCP/IP Stack — IP#3 — Target

Outbound data for a distributed connection

Inbound data routed over dynamic XCF to chosen target

IP#2 CS TCP/IP Stack — Target

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

The SASP (Server/Application State Protocols) 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 Sysplex

Server instance

z/OS LB agent

Work requests

Work requests

Load Balancer

Server instance

z/OS LB agent

SASP control flows

Server instance

z/OS LB agent

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

z/OS LB advisor

Private protocol control flows

The z/OS Load Balancing Advisor technology is a new z/OS Communications Server technology that will be made ganerally available 4Q2004

---

zSTSU: 2004

Application Enhancements

z/OS V1R6

IBM.

z/OS FTP Client programming interface for improved automation and integration of z/OS file transfers

z/OS V1R6

➤ Provides an interface that allows an application to programatically 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 z/OS V1R7
- ► 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 is used. (Mappings for common programming languages are provided)
- ► Both blocking (wait for a response), and non-blocking (polling-mode) calls are supported
- ► In non-blocking mode, progress notifications can be returned to the calling application as the transfer progresses
- ► The simple commands tell the interface what to do, for example: initialize, terminate, execute an FTP client 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
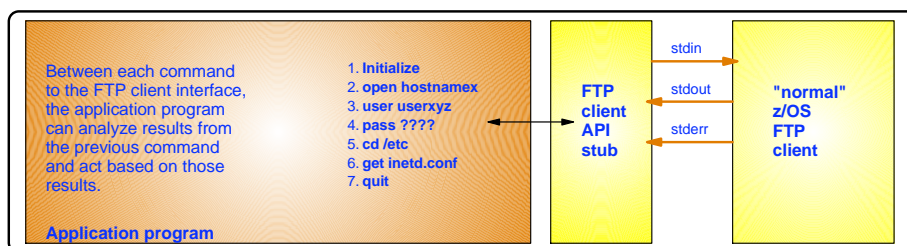
Between each command to the FTP client interface, the application program can analyze results from the previous command and act based on those results.

Application program

1. **Initialize**
2. **open hostnamex**
3. **user userxyz**
4. **pass ????**
5. **cd /etc**
6. **get inetd.conf**
7. **quit**

FTP client API stub

stdin
stdout
stderr

"normal" z/OS FTP client

I ♥ FTP

Enhanced DBCS and MBCS codepage support in FTP

z/OS V1R6

➤ Enhance 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 codepage 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

TN3270 Server in its own address space for improved management

z/OS V1R6

- ➤ Currently TN3270 Server runs as a subtask of the IBM TCPIP stack address space

- ➤ In z/OS V1R6, provide customers with a choice:
  - ▸ Run the TN3270 server as a separately started address space from TCPIP
  - ▸ Continue to run TN3270 server as a subtask of the TCPIP address space

- ➤ Reasons why an installation may want to run the TN3270 server in a separate address space:
  - ▸ Allows for prioritization of TCPIP address space vs TN3270 server
  - ▸ Much less likely for TN3270 server failure to cause a total TCPIP failure
  - ▸ Allow for easier problem diagnosis for both TCPIP 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 TCPIP
  - ▸ Commands are the same but must be directed to the intended TN3270 procedure name
  - ▸ Multiple TCPIP stacks supported
    - ‒ One server per stack (affinity)
    - ‒ One server associated with all stacks (Generic Server)
  - ▸ Must runTN3270 server with affinity for the following functions
    - ‒ TN3270 SNMP subagent (and must be only Telnet to that TCPIP)
    - ‒ WLM function
  - ▸ Requirements
    - ‒ Separate start up JCL. Sample is provided.
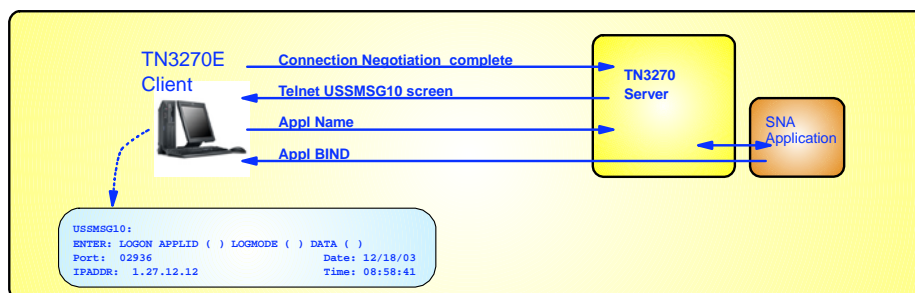
Remote terminal access
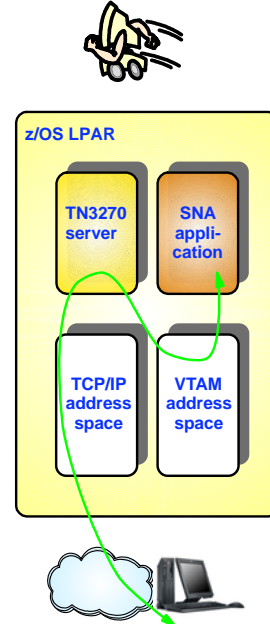
z/OS LPAR

| TN3270 server | SNA appli-cation |
| TCP/IP address space | VTAM address space |

Other TN3270 server enhancements in z/OS V1R6

z/OS V1R6

- ➤ Support for up to 128 000 connections per TN3270 server port
- ➤ Improved handling of the case where a TN3270 session is reconnected from an IP address that is different from the original IP address
  - ● Ensure new IP address is passed to VTAM for inclusion in displays and control vectors
- ➤ Support for USSMSG10 screen in SNA Character Stream (SCS) format
  - ● VTAM Unformatted System Services supports two formats of a USSMSG10 screens:
    - ▸ 3270 Data format
    - ▸ SNA Character Stream (SCS) format
  - ● The TCP/IP TN3270 server has sofar only supported USSMSG10 screens in 3270 data stream format
  - ● This release adds support for USSMSG10 screens in SCS format
    - ▸ Only TN3270E clients support SCS format
      - ‒ SCS data flows as SSCP-LU data
    - ▸ TN3270 connections continue to use 3270 Data format

TN3270E Client

Connection Negotiation complete → **TN3270 Server**

Telnet USSMSG10 screen ←

Appl Name → **SNA Application**

Appl BIND ←

```
USSMSG10:
ENTER: LOGON APPLID ( ) LOGMODE ( ) DATA ( )
Port:  02936                    Date: 12/18/03
IPADDR:  1.27.12.12             Time: 08:58:41
```

zSTSU: 2004

Security

z/OS V1R6

IBM.

Multi-Level security support by TCP/IP ...

➢ Requirement to protect resources available through the network - who can access what resources?

➢ Classify users and system resources such as networking interfaces and networking applications. Classification is defined in RACF using security labels which represent a level of security (ex: high, med, low). Use this general rule:

● Information may only be read by/written to users with equivalent security labels!
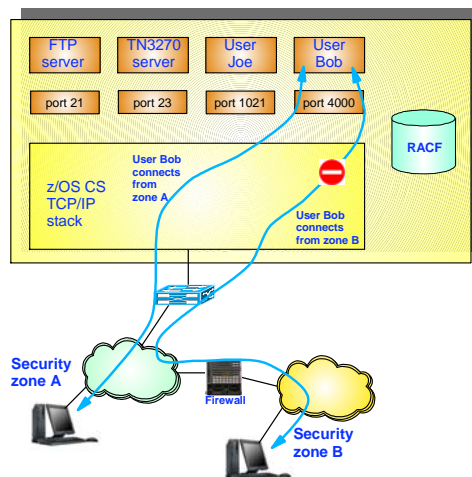
| FTP server | TN3270 server | User Joe | User Bob |
| port 21 | port 23 | port 1021 | port 4000 |

RACF

z/OS CS TCP/IP stack

User Bob connects from zone A

User Bob connects from zone B

Security zone A

Firewall

Security zone B

Multi-Level security support by TCP/IP

- ➤ 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/maynot 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 stack determines the end users have security label equivalence and may communicate (user seclabel compared to servauth zone seclabel)
    - ► Additionally, Packets sent over (XCF or SAMEHOST are labeled). If label is present, this is used to determine equivalence
  - ● Netaccess and MLS extended to Applications
    - ► Port of entry (PoE) checks for the FTP (MLS and NonMLS)
      - – Checks ensure a client may logon to a userid from FTP
    - ► Netaccess controls for Tn3270 server ports (nonMLS)
      - – Client Netaccess 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

---

# z/OS Communications Server IPSec

z/OS V1R7 ??

- ➤ 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
- ➤ Features
  - ● Configuration support
    - ► Optimized for z/OS host-to-host and z/OS host-to-gateway (z/OS gateway still supported)
    - ► *New!* 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
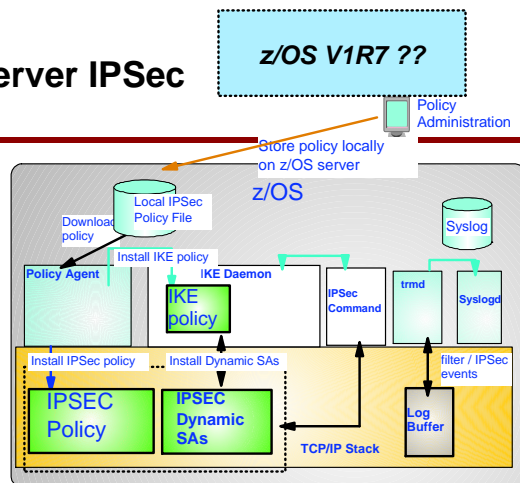
Planned for z/OS V1R7!

** Firewall Technologies still supported

zSTSU: 2004

Other TCP/IP Enhancements

z/OS V1R6

IBM.

Summary of new hardware features for z990 Stage 3

z/OS V1R6

➤ Continue Channel Subsystem Enhancements (CSS) from Stage 2

- From 2 to 4 Logical Channel Subsytems allows for 1024 channels (256*4)
  - Each image supports 256? channel programs
- Each LCSS can support up to 15 LPARS for a total of 60 LPARS (15*4) from 30
- Spanned channels
- Multiple subchannel sets

➤ OSA (Hydra-3)
- 10 GB Enet
- Ficon-3
- Support spanned channels

➤ IPv6 support for HiperSockets

➤ Greater than 16 CPUs per image

Fyi

Spanning channels allow a channel to be configured to multiple logical channel subsystems (LCSS)and to be transparently shared by any LPAR regardless of the LCSS to which the partition is configured.

z/OS V1R6 Communications Server support for z990

z/OS V1R6

➤ Expanded Channel subsystem is transparent for Communications Server

- OSA will provide support for spanning and should not affect Communications Server
- Changes made in z/OS V1R5 to support spanned channels for IQDIO for z990 Stage 2

➤ Hydra 3 (OSA Express)

- 10 Gigabit Ethernet supported with current OSA Express QDIO (IPAQENET)
  ‣ Testing will be done for TCP/IP and Enterprise Extender workloads
- Allow OSA Proxy Subagent to support OSA LCS SNMP objects so OSA can provide its own SNMP data. Support for OSA in the IBM MVS TCP/IP Enterprise-specific MIB will be deprecated.

➤ >16 CPs is supported

➤ IPv6 Hipersockets (discussed under IPv6)
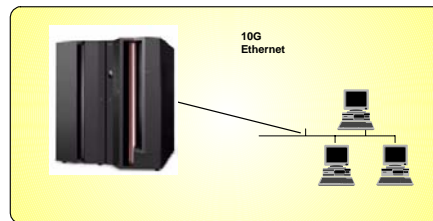
10G
Ethernet

---

zSTSU: 2004

Enterprise Extender and SNA Enhancements

z/OS V1R6

IBM.

VARY Command Enhancements for EE XCA

- V1R5 increases EE flexibility with support for multiple VRNs, multiple VIPAs, IPv6, and NAT compatibility. With the flexibility comes additional complexity, and this often leads to the need to add or change definitions.
  - However, currently those definitions cannot be changed or augmented once the XCA major node is active, without inactivating the major node, thereby **disrupting all existing Enterprise Extender connections**.
- V1R6 improves usability by allowing the UPDATE operand on the VARY command for the EE XCA major node, thereby allowing adding of GROUPs and changing of operand values, without bouncing the major node.
  - However, note that a GROUP must itself be inactive before its operands can be changed.
  - To simplify the inactivation of a GROUP prior to a change (or activation after a change), V1R6 also allows a VARY ACT (or VARY INACT) command to be issued against an EE XCA GROUP, thereby activating (or inactivating) the GROUP and all subordinate LINEs/PUs.
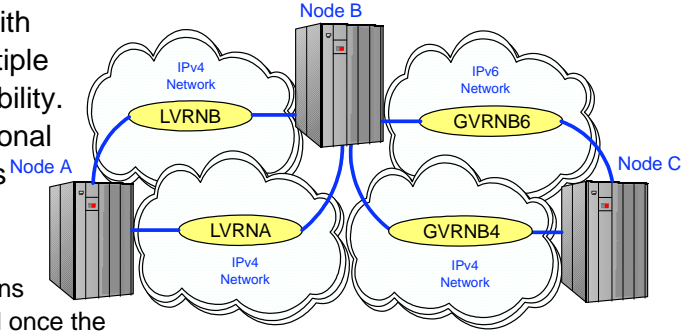

Display EE Command

- 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
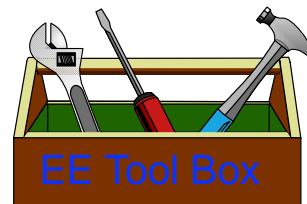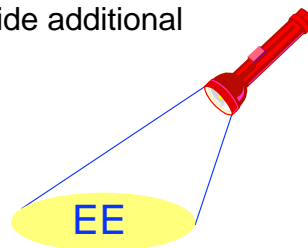
Example: D EE: General Information

```
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
```

Enhanced RTP Displays

- V1R6 adds yet more information to the detailed RTP PU output (D NET,ID=<RTP_pu_name>,HPRDIAG=YES, and also improves the organization of the output.
- 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.
  - This provides a means to correlate the RTP PU name (CNRxxxxx) on one side of an RTP connection to the corresponding RTP PU name on the other side of that same RTP connection.

```
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

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
```

EE Tool Box

EE Connection Network Failures
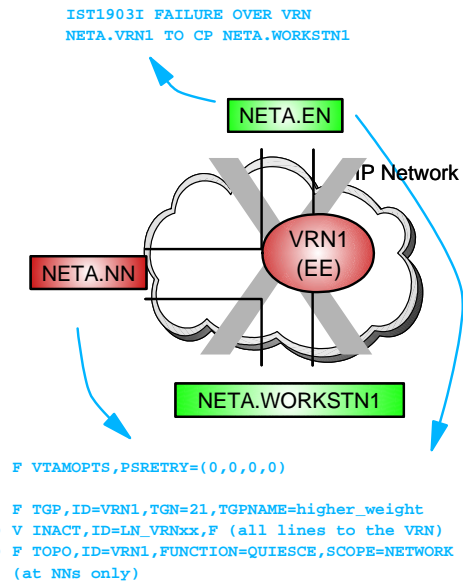
- Typical VRN Configuration
  - ENs predefine links to NNS for CP-CP
  - Dynamic CN links used EN-to-EN

- 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 Issues 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
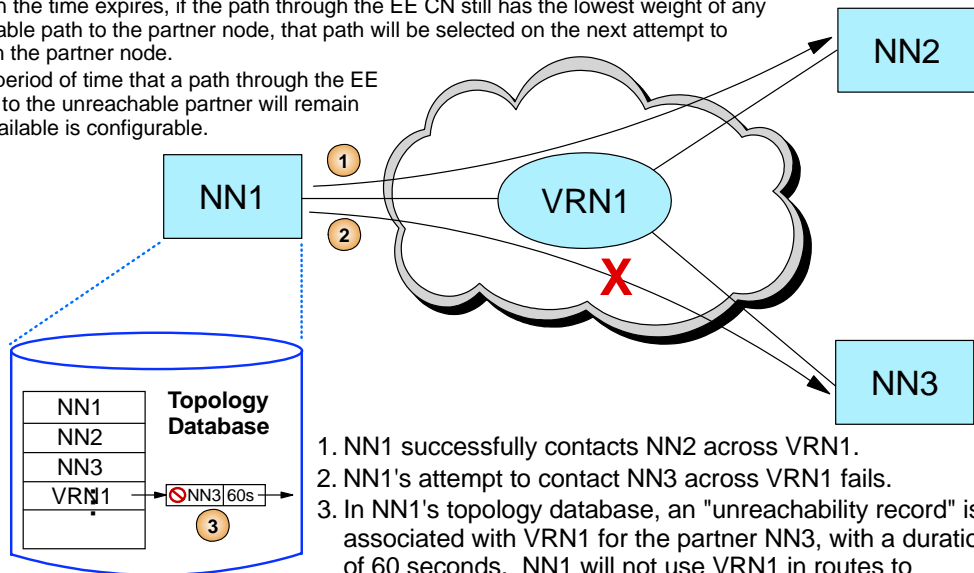 - Similar actions may be required at the partner node and/or at an EN's NNS

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

NETA.EN

IP Network

VRN1
(EE)

NETA.NN

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)
```

EE Connection Network Reachability Awareness

- EE Connection Network Reachability Awareness detects a dial failure or connection INOP for a connection over an EE connection network and prevents that specific path to the partner node from being used for a period of time, instead choosing an alternate route, if available.
- When the time expires, if the path through the EE CN still has the lowest weight of any available path to the partner node, that path will be selected on the next attempt to reach the partner node.
- The period of time that a path through the EE VRN to the unreachable partner will remain unavailable is configurable.

NN2

NN1

1

2

VRN1

X

NN3

Topology Database

| NN1 |
| NN2 |
| NN3 |
| VRN1 |

NN3 60s

3

1. NN1 successfully contacts NN2 across VRN1.
2. NN1's attempt to contact NN3 across VRN1 fails.
3. 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.

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
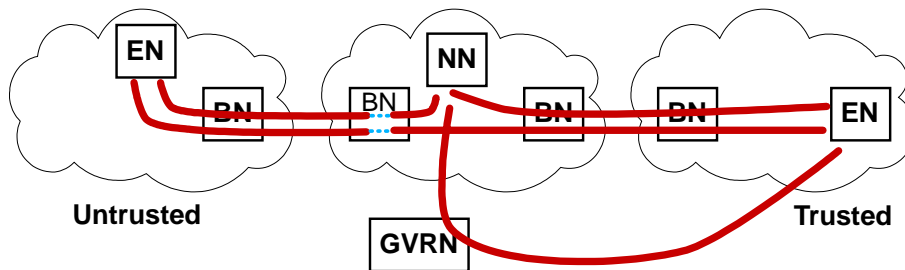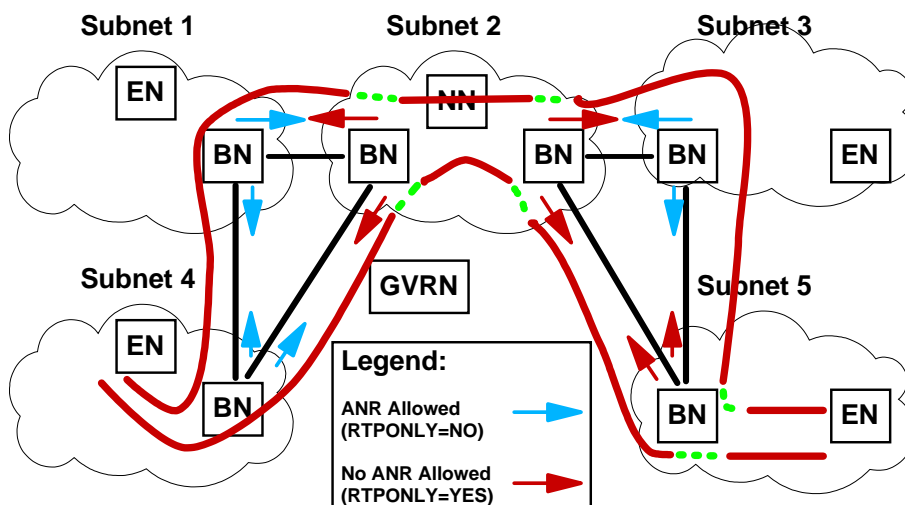

Untrusted

Trusted

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

Legend:

ANR Allowed (RTPONLY=NO) →

No ANR Allowed (RTPONLY=YES) →

Other EE/SNA Enhancements

- Move element addresses for RTP PUs and DLUR PUs to high-order address space
- MultiNode Persistent Sessions Predatory Takeover (determine when application is hung)
  - MNPS only supported a "planned takeover" where application had to close successfully
  - Application must support MNPS and support a forced takeover
    - IMS v9 will exploit this function
- Packet Trace Formatting for Enterprise Extender for TCPIP Packet Trace
- Control displaying and sending (cross domain) of TN3270 TCPIP visibility information
  - IP Characteristics (IP Address, Host name, Port) associated with a TN3270 LU in CV64
  - Can allow full visibility (local and cross-domain), restrict to just local exits and display, restrict to display only, or completely suppress
- Provide indicator to VTAM exits - logon, SME, management - that TN3270 IP characteristics could change if TKxxxxLURECON is defined in TN3270 profile
- Inactivation of Model Application Major Node and all instantiated clones via single command
- Provide extended information for APPN locate failure (LSIRFMSG)
- Improvements to APPN search cleanup (VARY TERM,SCOPE=APPN)
- Display additional information for Switched PUs - CPNAME,SAP and Mac addrs,etc..
- Display message to indicate that INACT FINAL=Yes was issued for the DLUR
- Display pending AUTOLOGONs
- Additional dump analysis enhancements for HPR RTPs
- Alert and perform extended recovery for RTP endpoint when buffers are lost

APPENDIX:
IPv6 on z/OS Communications Server:
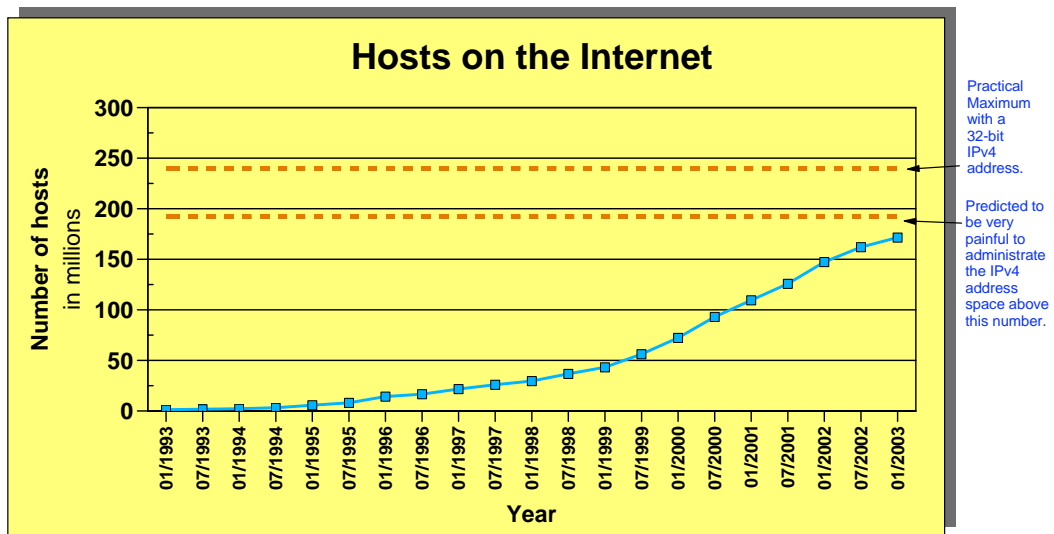# z/OS V1R4, V1R5, V1R6, and later

IBM.

We have IPv4 addresses enough - or do we?
Latest IPv4 address space usage overview

The chart shows IPv4 address allocation over time. The "blue" line is the actual allocation, the "purple" line is the smoothed allocation. This shows the IPv4 address pool being depleted in the next 5 years or so.
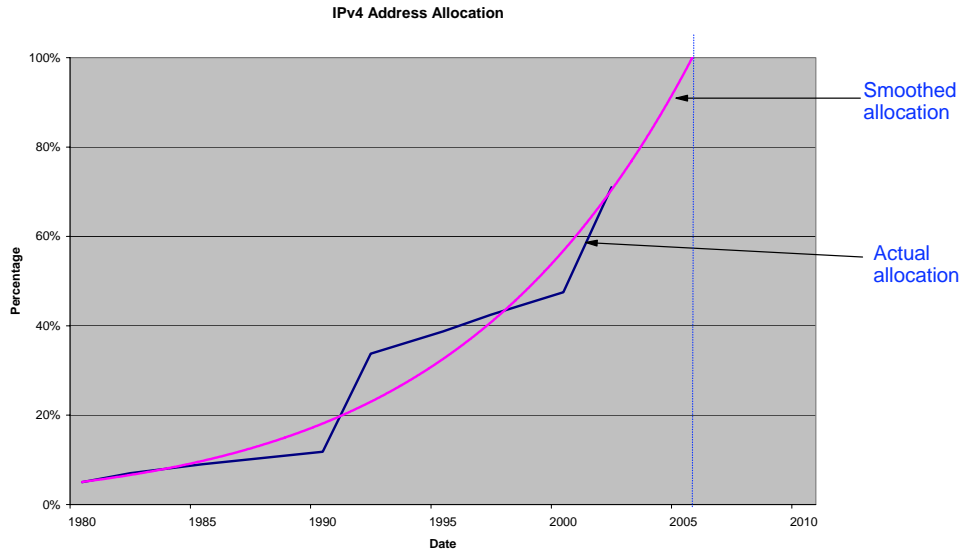
IPv4 Address Allocation

Smoothed allocation

Actual allocation

Visible IPv4 hosts on the Internet
through the last 10 years

Hosts on the Internet

Practical Maximum with a 32-bit IPv4 address.

Predicted to be very painful to administrate the IPv4 address space above this number.

Source: http://www.isc.org/ds/WWW-200301/index.html and "*An update on the H ratio*", RFC3174

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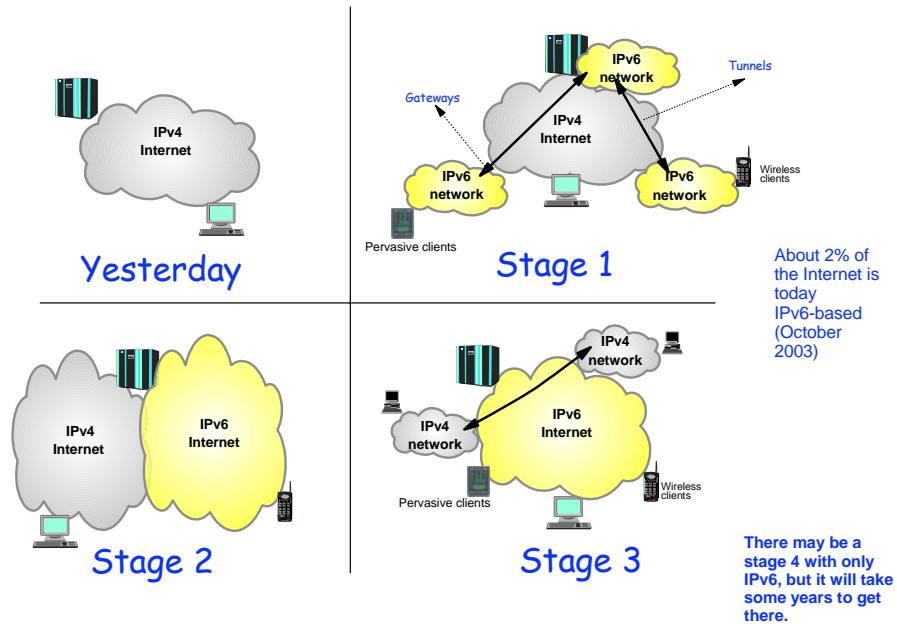
Yesterday

Stage 1

Stage 2

Stage 3

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

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

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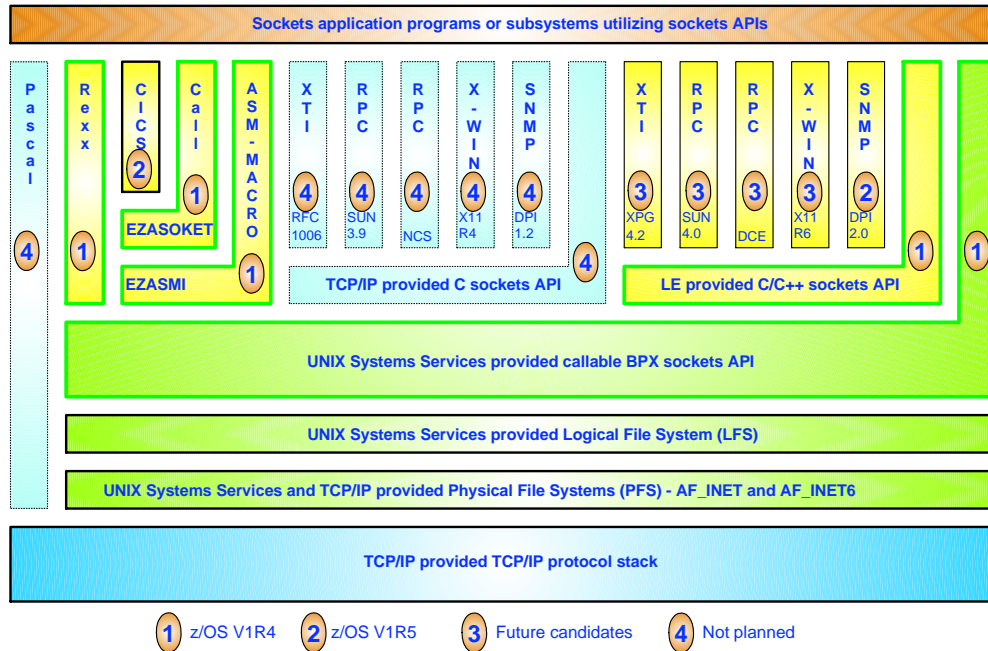 - ‣ **Additonal Network Management MIBs**
  - ‣ **HiperSockets DLC**
- **After z/OS V1R6**
  - ‣ **Integrated IPSec**
  - ‣ **Advanced Socket APIs**
  - ‣ **Extended Stats MIB, OSPFv3 MIB**
  - ‣ **Intrusion Detection Services**
  - ‣ **IPv6 mobility support**

The Internet - a worldwide digital utility.
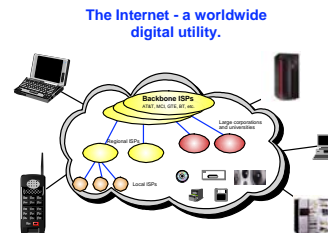
Objective is to have IPv6 production ready on the platform when you need it!

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

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

Sockets application programs or subsystems utilizing sockets APIs

Pascal	Rexx	CICS	Call	ASM-MACRO	XTI	RPC	RPC	X-WIN	SNMP	XTI	RPC	RPC	X-WIN	SNMP		

CICS **2**
Call **1**
EZASOKET
EZASMI **1**
Rexx **1**
Pascal **4**

4 RFC 1006 | **4** SUN 3.9 | **4** NCS | **4** X11 R4 | **4** DPI 1.2 — **TCP/IP provided C sockets API** **4**

3 XPG 4.2 | **3** SUN 4.0 | **3** DCE | **3** X11 R6 | **2** DPI 2.0 — **LE provided C/C++ sockets API** **1** **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

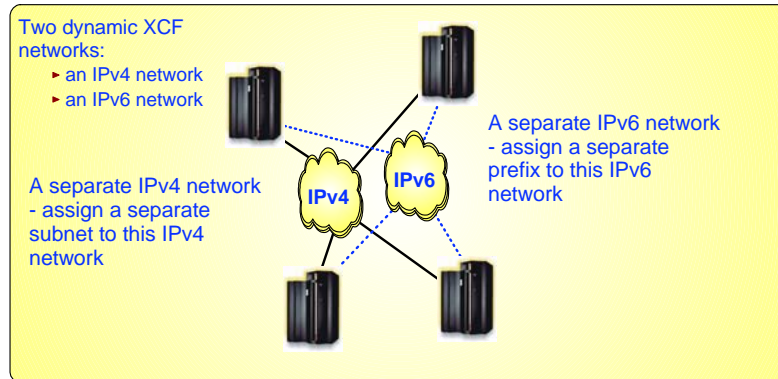---

zSTSU: 2004

IPv6 in z/OS V1R5

z/OS V1R5

IBM.

IPv6: New interface support

IPv6 MPCPTP-type Interface support - Extends connectivity options for IPV6

- ► IUTSAMEHOST to other stacks in same LPAR

- ► XCF to other stacks in same Sysplex
  - ► Both static and dynamic XCF, but not at the sametime! Must either have static xcf or dynamic.

- ► ESCON/FICON to another z/OS image (not to any known Channel-attached Routers)

Two dynamic XCF networks:
- ► an IPv4 network
- ► an IPv6 network

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

IPv4

IPv6

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

---

IPv6: Another batch of AF_INET6 enabled applications

IPv6-enable another batch of applications

- ➤ tftpd (trivial file transfer server)

- ➤ syslogd

- ➤ dcas (digital certificate access server)

- ➤ sntpd (simple network time protocol server)

- ➤ TN3270 Server: provide connectivity to SNA applications from IPv6 TN3270 clients

- ➤ Enterprise Extender: allow SNA applications to communicate over IPv6 networks

- ➤ sendmail 8.12.5 (new port of sendmail picks up IPv6 enablement too)

- ➤ MVS rshd/rexecd server
  - ► the UNIX rexecd and rshd servers were IPv6-enabled in z/OS V1R4

- ➤ TSO rsh/rexec clients
  - ► provide updated version that can be used in all z/OS environments (batch, TSO, REXX, etc.)

- ➤ New UNIX rsh client that is IPv6-enabled from start
  - ► the UNIX rexec client was IPv6-enabled in z/OS V1R4

Changes to support partners with V6 addresses:
- ◆ Changes to socket api calls and structures
- ◆ Configuration file changes

**IPv6-enable the CICS listener, sockets infrastructure, and sockets libraries**

IPv6: Add RIPng to OMPRoute and other OMPROUTE enhancements

z/OS V1R5

- ➤ **OMPROUTE has been enhanced to support IPv6 and RIPng**
  - 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
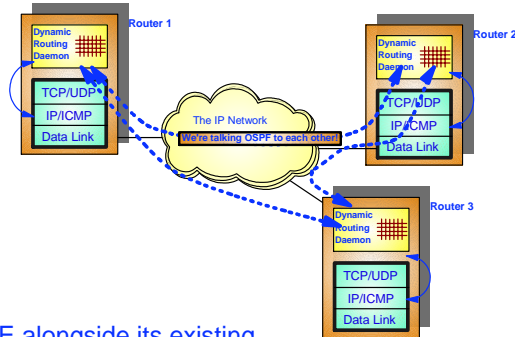
- ➤ **OMPROUTE has also been changed in this release to:**
  - Improved trace performance by using CTRACE
  - Multipath enhanced to support 16 routes to same destination
  - Support advertisement for more than 254 interfaces (no upper limits)
  - Some display clean-up to show N/A where an option doesn't apply
 - Ignore undefined interfaces (don't have to define everything to OSPF)

IPv6: Network management support for IPv6

z/OS V1R5

- ➤ **Network management SNMP support**
  - Support SNMP agent (OSNMPD)
  - DPI 2.0 enabled for AF_INET6
  - Support MVS TCPIP subagent
  - osnmp command
  - The trap forwarder daemon enabled for AF_INET6
  - 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
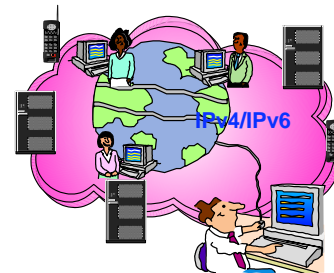
- ➤ **SMF119 records support**
  - The redesign in z/OS V1R2 did factor in IPv6 addresses, so most subtypes are already in z/OS V1R4 supporting IPv6 addresses
  - Some changes needed to selected records to capture additional IPv6-related data, such as interface records and statistics records

- ➤ **Non IPv6-specific enhancements to SMF119 records:**
  - Security information has been added to the FTP server and client file transfer completion records
  - TCP connection termination records now, for TN3270 server connections, include TN3270-specific information: LU name, SNA Application name, protocol, etc.

Must be able to monitor an IPv6 network

IPv4/IPv6

- ➤ And some netstat goodies:
  - ► add support for a hostname filter for connection-type reports
  - ► add existing IP address filter to BYETINFO report
  - ► add interface/link filter to DEVLINKS reports
  - ► add interface statistics sections to DEVLINKS report

At this time, the UDP MIB draft was very unstable and too risky to implement - will follow in a later release when RFCs have stabilized.

zSTSU: 2004

# IPv6 in z/OS V1R6

z/OS V1R6

IBM

IPv6 HiperSockets support on z990 stage 3

z/OS V1R6

➤ Enables High-Speed Connectivity for IPv6 Inside z9xx 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

HTTP/DMZ Zone **Application Zone** **Data Zone**

IDS   IDS

HTTP server or Edge Server Firewall Appli-cation Server Firewall Data Server

OSA-E

➤ **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

➤ **High Speed Connectivity to Other Operating Systems**
- Linux for zSeries
  - ▸ Can be configured as part of same internal network as z/OS
 - ▸ High speed access to enterprise data on z/OS
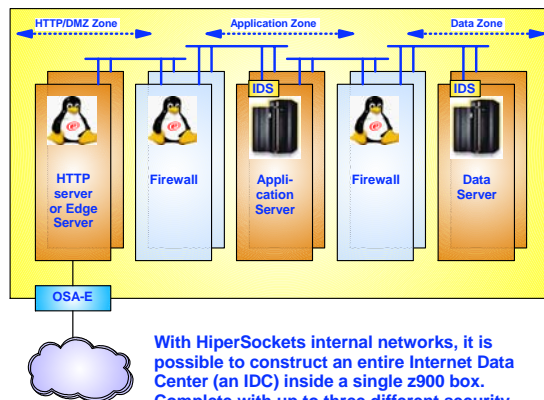
With HiperSockets internal networks, it is possible to construct an entire Internet Data Center (an IDC) inside a single z900 box. Complete with up to three different security zones and a mix of z/OS and Linux for zSeries images.

IPv6 Sysplex support

➤ 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
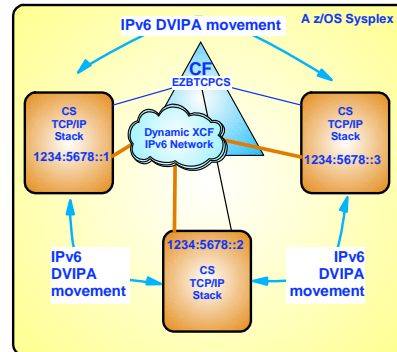
---

IPv6 OSPF support on z/OS V1R6

➤ **Dynamic routing is needed on z/OS to support sysplex functions like VIPA, Dynamic VIPA, and Dynamic VIPA takeover**
  - z/OS V1R5 added RIP for IPv6

➤ **z/OS V1R6 adds support of OSPFv3 for IPv6 (RFC2740) to OMPROUTE**
  - Single daemon will support RIP and OSPF for IPv4 and IPv6

➤ OSPF for IPV6 is algorithmically similar to IPV4, but there are differences:
  - IPV6 uses the term "link" to replace "subnet" or "network"; therefore, OSPF runs per-link
  - Addressing semantics have been removed from protocol packets except in Link State Advertisement (LSA) payloads
  - Flooding Scope for LSA's have been added - link-local, Area, and AS scope
  - Explicit support for multiple OSPF protocol instances on a single link
  - Support for Link-Local addresses - routers can't forward ip addr's with link local source
  - Authentication has been removed from OSPF and now uses IPSEC
  - Some packet format changes - Hello has no address information (uses interface id), version number, authentication fields removed, etc..
  - LSA format changes - addressing info in headers is removed, addresses inside use prefix & prefix length, new Link-LSA's introduced for link-local and prefix advertisement
  - Some changes for Stub Area support
  - All routers on a given link are now identified by Router ID (in V4 broadcast and Point to multipoint used addresses)
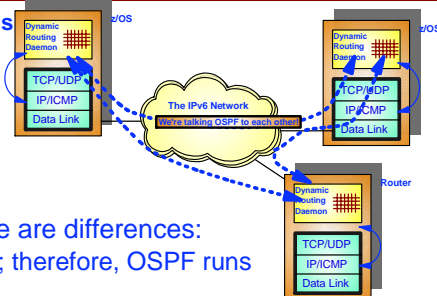
IPv6: Network management support for IPv6

Must be able to monitor an IPv6 network

➤ **Network management SNMP support in z/OS V1R5:**
- Support SNMP agent (OSNMPD)
- DPI 2.0 enabled for AF_INET6
- Support MVS TCPIP subagent
- osnmp command
- The trap forwarder daemon enabled for AF_INET6
- 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 in z/OS V1R5:**
- The redesign 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

IPv4/IPv6
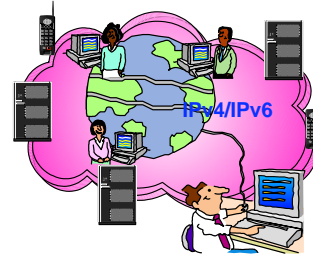
z/OS V1R6:
- ▶ **Add support for the version-neutral (both IPv4 and IPv6) MIB data in the following:**
  - ▸ Standard UDP-MIB (current draft rfc2013)
  - ▸ IBM MVS TCP/IP Enterprise-specific MIB

---

zSTSU: 2004

Getting Started with IPv6

IBM.

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

Applications

AF_INET6 PFS   **AF_INET PFS**

Common TCP and UDP Transport

IPv6 Raw Transport   **IPv4 Raw Transport**

IPv6   **IPv4**

- NeD
- MLD
- Stateless autoconfig

ICMPv6

QoS TRM IDS   **ARP**   **IGMP**   **ICMP**

Firewall Functions

Common DLC Functions

IPv6 DLCs   **IPv4 DLCs**

OSA-E QDIO

IPv4 and IPv6 packets on the same LAN
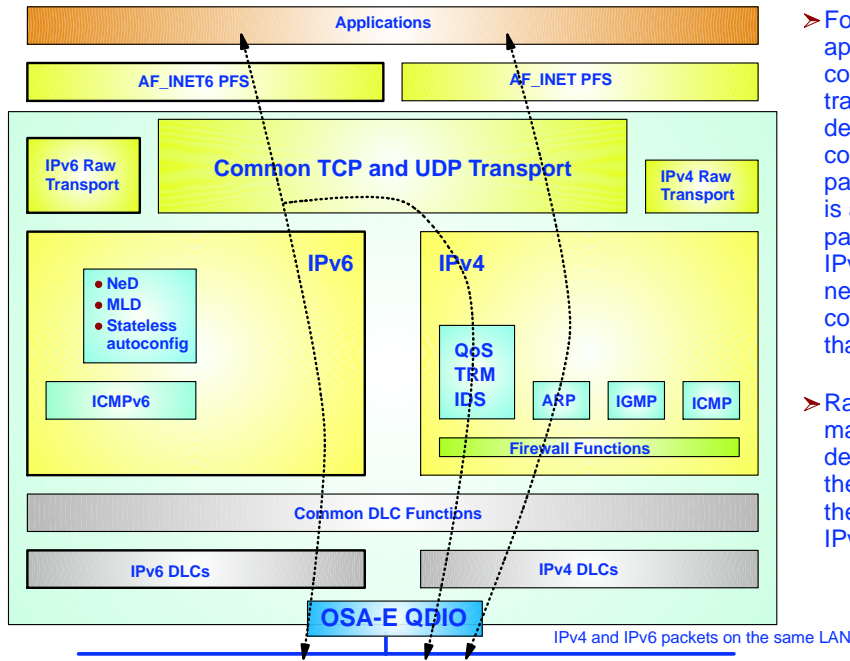
➤ For AF_INET6 applications, the common TCP or UDP transport layer determines per communication partner if the partner is an IPv4 or an IPv6 partner - and chooses IPv4 or IPv6 networking layer component based on that.

➤ Raw applications make the determination themselves when they choose IPv4 or IPv6 raw transport.
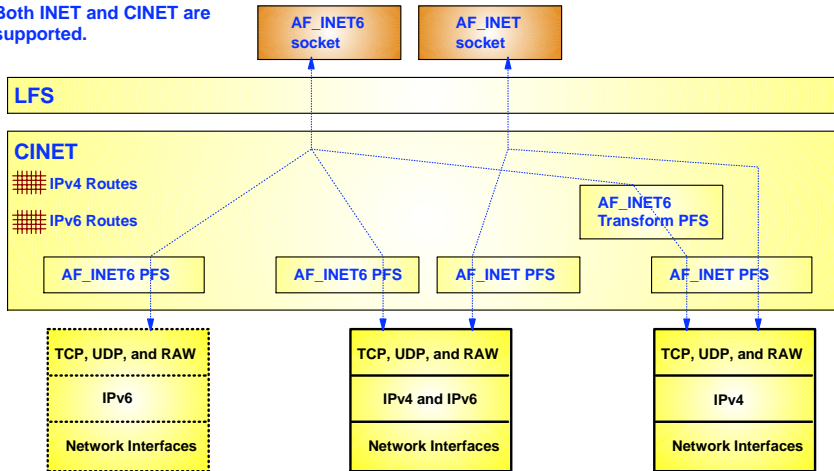
OSA-E for IPv6 requires zSeries hardware

How do we enable IPv6 support 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.

AF_INET6 socket   **AF_INET socket**

LFS

CINET

IPv4 Routes

IPv6 Routes

AF_INET6 Transform PFS

AF_INET6 PFS   **AF_INET6 PFS**   **AF_INET PFS**   **AF_INET PFS**

TCP, UDP, and RAW   **TCP, UDP, and RAW**   **TCP, UDP, and RAW**

IPv6   **IPv4 and IPv6**   **IPv4**

Network Interfaces   **Network Interfaces**   **Network Interfaces**

IPv6-only TCP/IP Stack
This will not be the case on z/OS for the forseeable 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
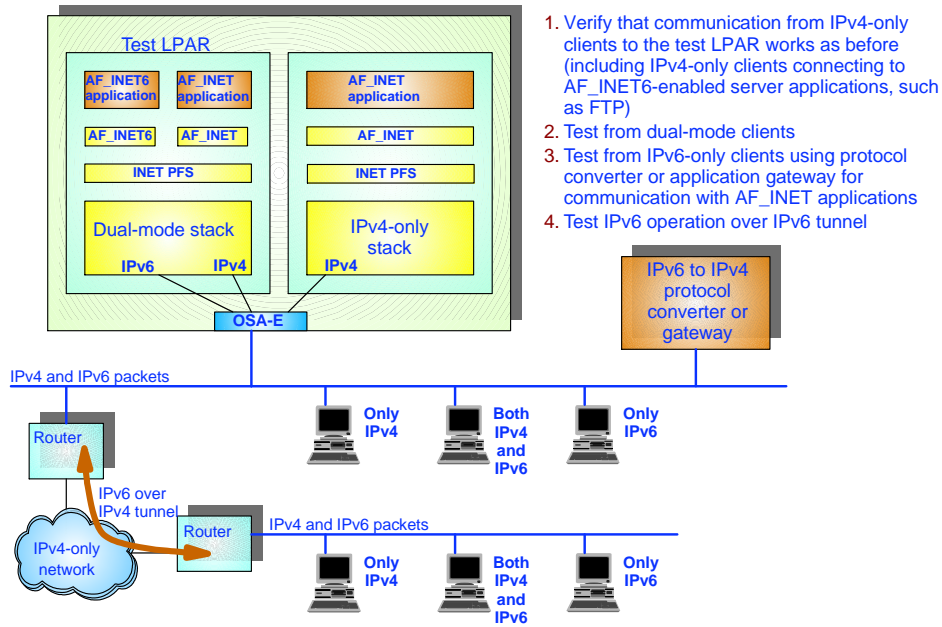
IPv4-only TCP/IP Stack
(such as AnyNet or an OEM TCP/IP stack)

▶ 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 to that they will also be able to communicate with IPv6 partners.

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!

Start testing IPv6 on z/OS

Test LPAR

| AF_INET6 application | AF_INET application |
| AF_INET6 | AF_INET |
| INET PFS |

Dual-mode stack
IPv6   IPv4

| AF_INET application |
| AF_INET |
| INET PFS |

IPv4-only stack
IPv4

OSA-E

IPv6 to IPv4 protocol converter or gateway

IPv4 and IPv6 packets

Router

Only IPv4   **Both IPv4 and IPv6**   **Only IPv6**

IPv6 over IPv4 tunnel

IPv4-only network

Router

IPv4 and IPv6 packets

Only IPv4   **Both IPv4 and IPv6**   **Only IPv6**

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

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	Networking & communications software
http://www.ibm.com/software/network/commserver	Communications Server
http://www.ibm.com/software/network/commserver/library	Communications Server white papers, product documentations, etc.
http://www.redbooks.ibm.com	ITSO redbooks
http://www.ibm.com/software/network/commserver/os390/support	Communications Server technical Support
http://www.ibm.com/support/techdocs/	Advanced technical support (flashes, presentations, white papers, etc.)
http://www.rfc-editor.org/rfcsearch.html	Request For Comments (RFC)

zSTSU: 2004

End of Topic
z/OS CS V1R5 & V1R6 Overview

IBM.