What's New in z/OS Communications Server

Sam Reynolds – samr@us.ibm.com
Alfred B Christensen – alfredch@us.ibm.com
IBM Raleigh, NC, USA

Session: 9159
Monday, February 28, 2011: 11:00 AM-12:00 PM

What's New in z/OS Communications Server

Session number:	9159
Date and time:	Monday, February 28, 2011: 11:00 AM-12:00 PM
Location:	Room 212A (Anaheim Convention Center)
Program:	Communications Infrastructure
Project:	Communications Server
Track:	
Classification:	Technical
Speaker:	Sam Reynolds, IBM Alfred B Christensen, IBM
Abstract:	The z/OS Communications Server combines TCP/IP and SNA support to better address the needs of today's complex networks. This session introduces new functions and capabilities for z/OS Communications Server, with a focus on the z/OS V1R12 CS release.

Trademarks, notices, and disclaimers

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

- Advanced Peer-to-Peer Networking®
- AIX®
- alphaWorks®
- AnyNet®
- AS/400®
- BladeCenter®
- Candle®
- CICS®
- DataPower®
- DB2 Connect
- DB2®
- DRDA®
- e-business on demand®
- e-business (logo)
- e business(logo)®
- ESCON®
- FICON®
- GDDM®
- GDPS®
- Geographically Dispersed Parallel Sysplex
- HiperSockets
- HPR Channel Connectivity
- HyperSwap
- i5/OS (logo)
- i5/OS®
- IBM eServer
- IBM (logo)®
- IBM®
- IBM zEnterprise™ System
- IMS
- InfiniBand ®
- IP PrintWay
- IPDS
- iSeries
- LANDP®
- Language Environment®
- MQSeries®
- MVS
- NetView®
- OMEGAMON®
- Open Power
- OpenPower
- Operating System/2®
- Operating System/400®
- OS/2®
- OS/390®
- OS/400®
- Parallel Sysplex®
- POWER®
- POWER7®
- PowerVM
- PR/SM
- pSeries®
- RACF®
- Rational Suite®
- Rational®
- Redbooks
- Redbooks (logo)
- Sysplex Timer®
- System i5
- System p5
- System x®
- System z®
- System z9®
- System z10
- Tivoli (logo)®
- Tivoli®
- VTAM®
- WebSphere®
- xSeries®
- z9®
- z10 BC
- z10 EC
- zEnterprise
- zSeries®
- z/Architecture
- z/OS®
- z/VM®
- z/VSE

* All other products may be trademarks or registered trademarks of their respective companies.

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

- Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.
- Cell Broadband Engine is a trademark of Sony Computer Entertainment, Inc. in the United States, other countries, or both and is used under license there from.
- Java and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.
- Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.
- InfiniBand is a trademark and service mark of the InfiniBand Trade Association.
- Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.
- UNIX is a registered trademark of The Open Group in the United States and other countries.
- Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.
- ITIL is a registered trademark, and a registered community trademark of the Office of Government Commerce, and is registered in the U.S. Patent and Trademark Office.
- IT Infrastructure Library is a registered trademark of the Central Computer and Telecommunications Agency, which is now part of the Office of Government Commerce.

Notes:

- Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput improvements equivalent to the performance ratios stated here.
- IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.
- All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.
- This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.
- All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.
- Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the performance, compatibility, or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.
- Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.

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

Agenda

- Introduction
- Application Integration / Data Consolidation and Standards
- Availability and business resilience
- Scalability / Performance / Constraint Relief and Accelerators
- Security
- System Management and Monitoring
- SNA and EE

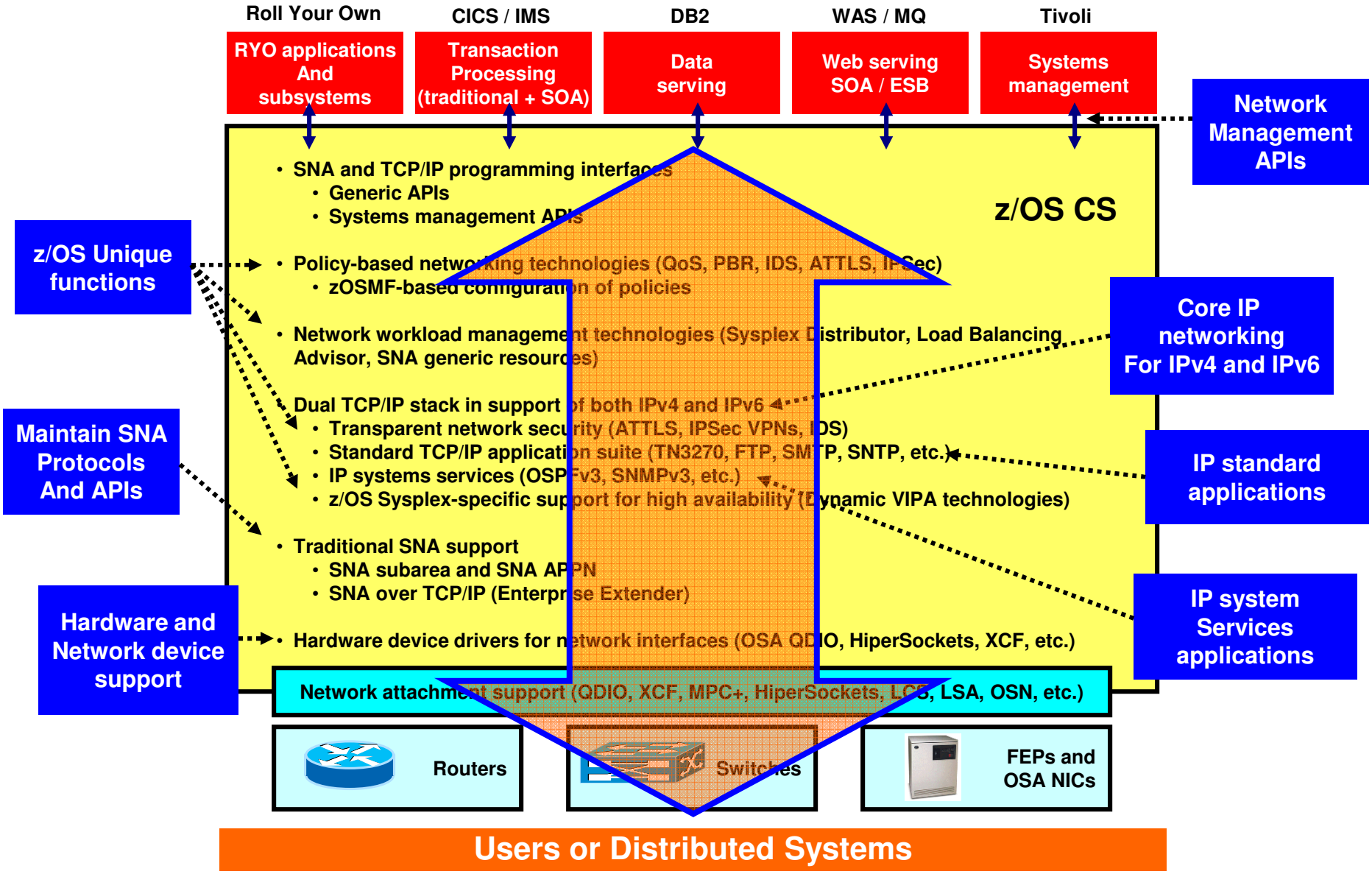
z/OS V1R12 was made available September 2010

Disclaimer: All statements regarding IBM future direction or intent, including current product plans, are subject to change or withdrawal without notice and represent goals and objectives only. All information is provided for informational purposes only, on an "as is" basis, without warranty of any kind.

z/OS V1R12 Communications Server – Technical Update

Introduction

z/OS Communications Server functional overview



Help shape future z/OS Communications Server functions!

- We need your feedback on the directions we are taking and would greatly appreciate your participation in that process!
  - We develop a function across several months in an iterative fashion.
  - You can provide feedback during development iterations. No commitment is required.
  - A non-disclosure agreement is required, and if not already in place, we will work with you to set one up
- If you are interested in providing feedback on z/OS Communications Server content, please check your areas of interest, and provide the contact information below. Thank you!

- ❑ **Security**
- ❑ **Sysplex and High Availability**
- ❑ **Connectivity**
- ❑ **Configuration and Usability**
- ❑ **Applications / APIs**
- ❑ **Problem Diagnosis**
- ❑ **General**

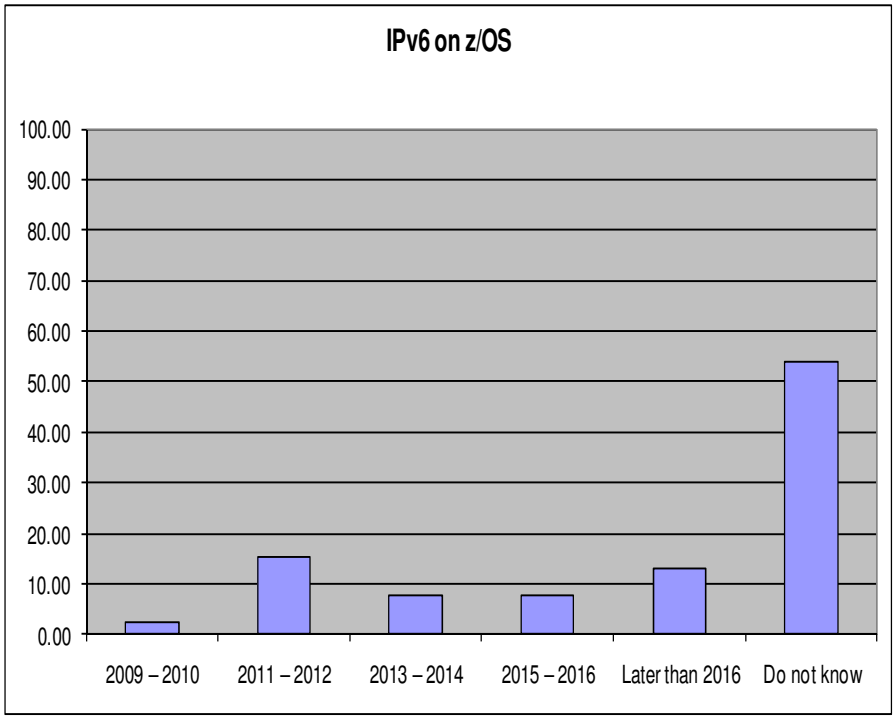
Name: _____ Company: _____

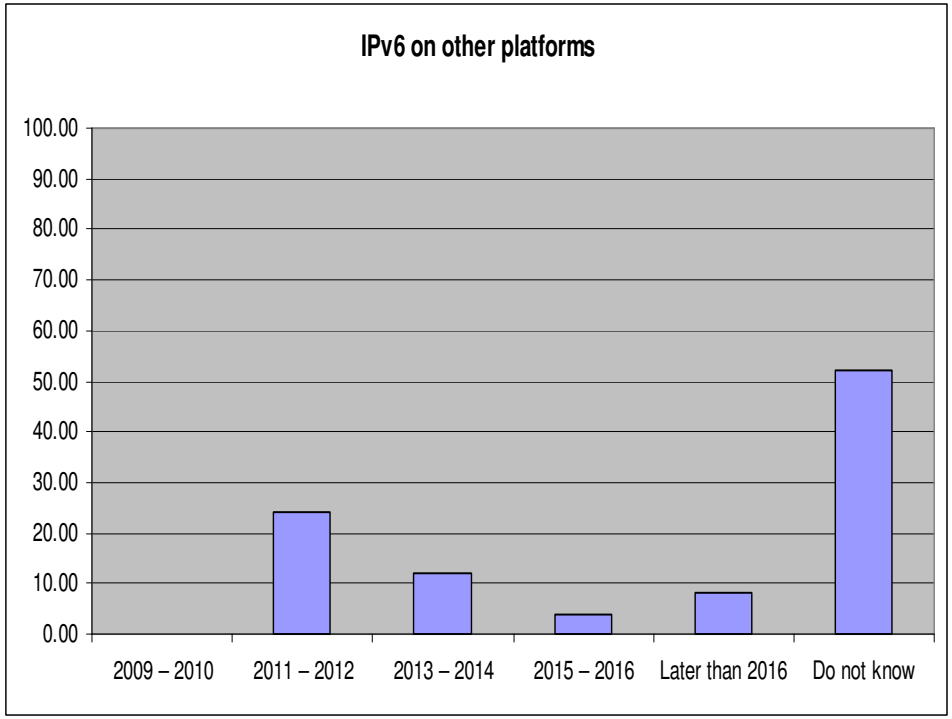
Email: _____ Phone: _____

Application Integration / Data Consolidation and Standards

When do our z/OS customers believe they will need IPv6?

- The majority of z/OS customers do not know
  - Expectations are that it will be needed slightly earlier on other platforms than z/OS
- It is time to start thinking, learning, and preparing ***now*** !

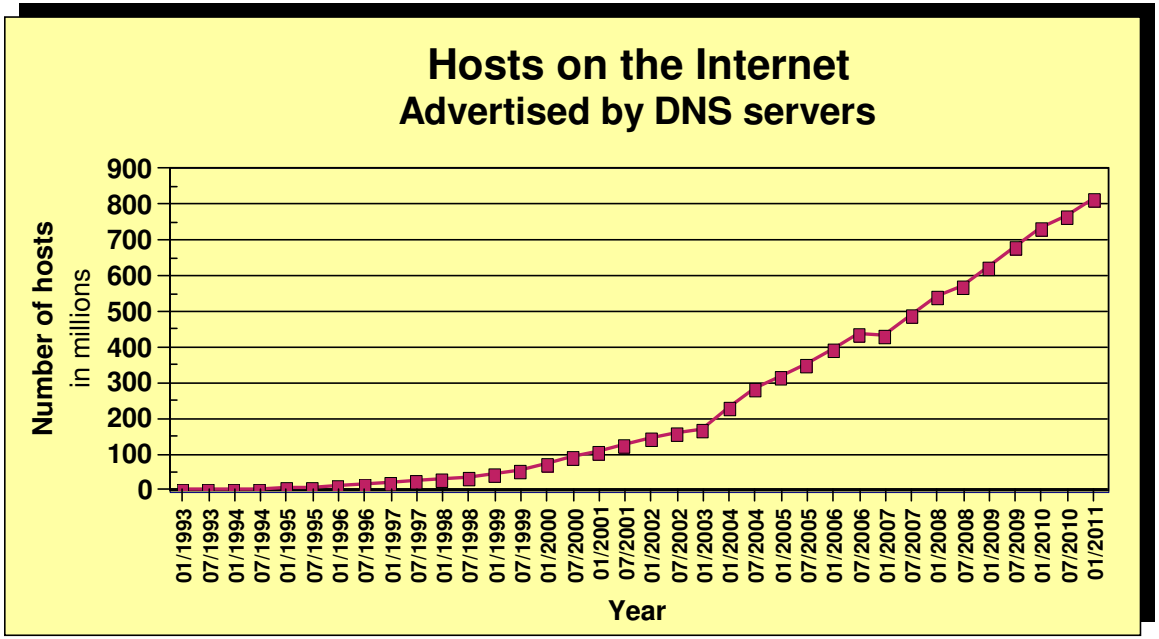




Source: Survey conducted by ENS early 2009 among a selected set of customers (39 responses to this question)

IPv4 address usage since early 1993

- Projected Internet Assigned Numbers Authority (IANA) Unallocated Address Pool Exhaustion
  - **Happened Feb 1, 2011**
- Projected Regional Internet Registries (RIR) Unallocated Address Pool Exhaustion
  - **August 2011**
- z/OS Communications Server continues to focus on IPv6 standards currency
  - US DoD/NIST
  - IPv6 Forum



- **What is the upper practical limit (the ultimate pain threshold) for number of assigned IPv4 addresses? Some predictions said 250,000,000 (250 million), others go up to 1,000,000,000 (one billion or one milliard).**
- **Source: https://www.isc.org/solutions/survey**
- **Source: http://www.potaroo.net/tools/ipv4/index.html**
- **Source: http://penrose.uk6x.com/**

If you want to stay in business after 2011/2012, you'd better start paying attention!
Do not worry too much; the sky isn't falling – IPv4 and IPv6 will coexist for many years to come.
Your applications need to be able to use both. If you write directly to the TCP/IP sockets layer, you need to start changing those applications.

Is Doomsday approaching?

http://www.potaroo.net/tools/ipv4/index.html

IPv4 Address Report

This report is auto-generated by a daily script. The report you are seeing here was generated at 26-Feb-2011 07:58 UTC.

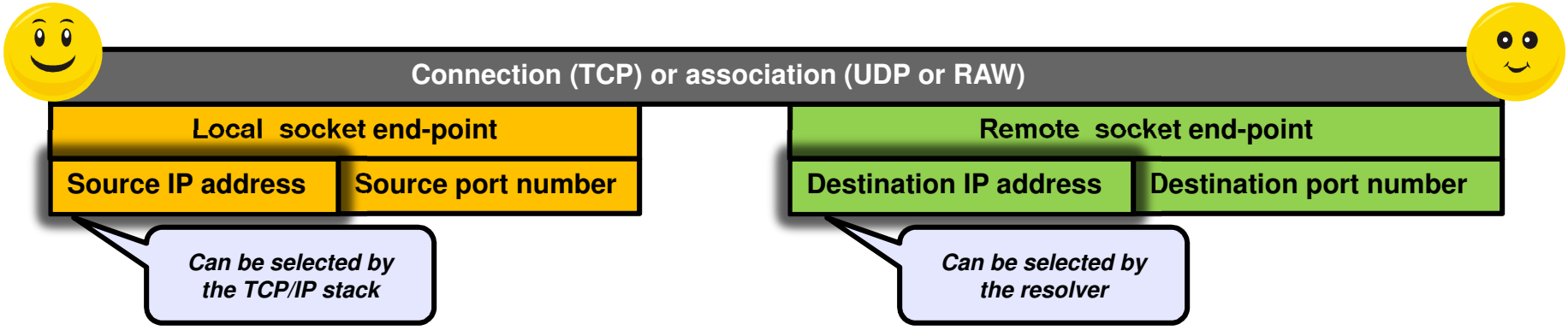
IANA Unallocated Address Pool Exhaustion: 01-Feb-2011

Projected RIR Unallocated Address Pool Exhaustion: 09-Aug-2011

This is less than six months from now!!!!

z/OS Communications Server keeps the pace, adding required new IPv6 support…

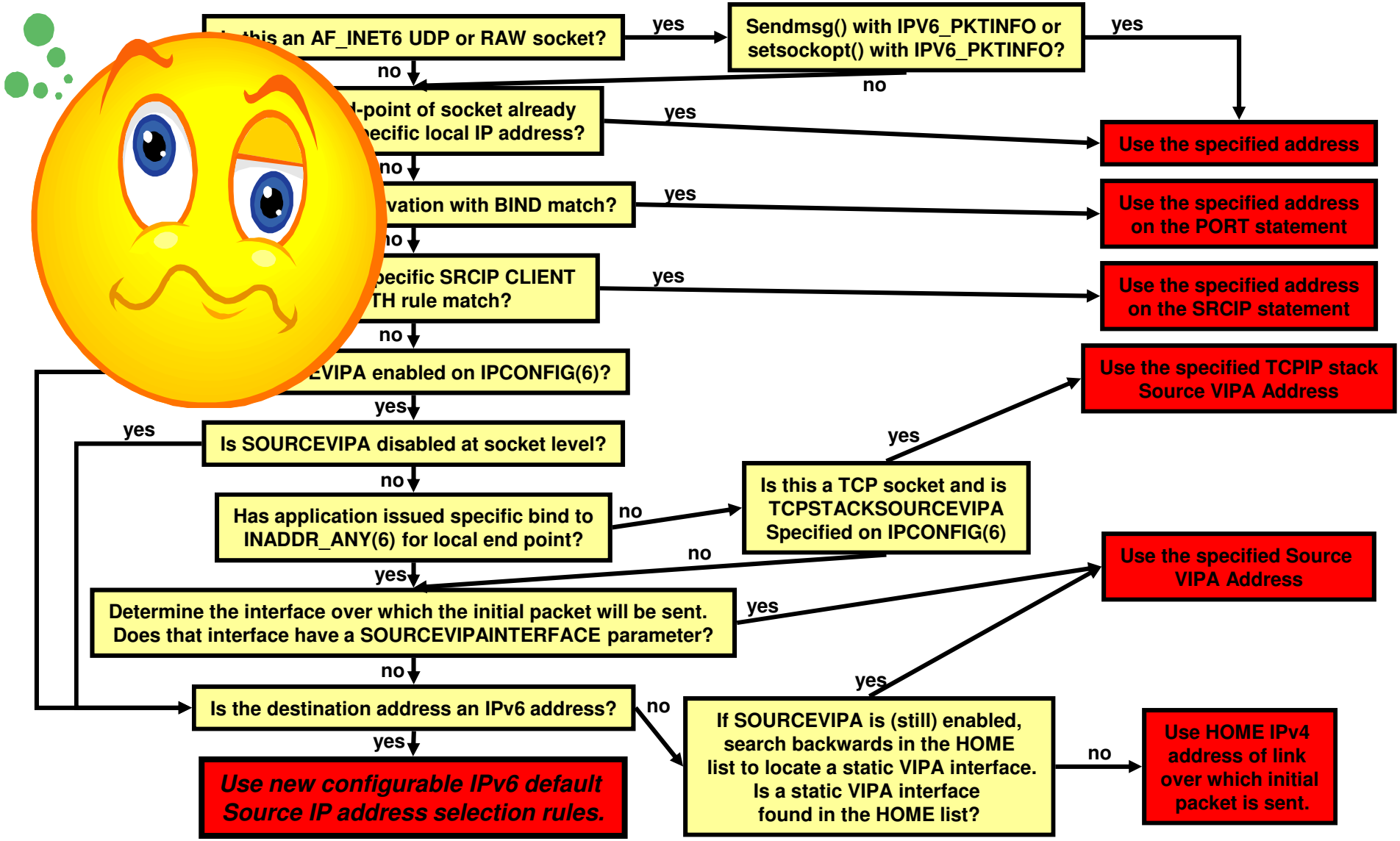
z/OS V1R12 adds support for RFC 3484 “Default Address Selection for Internet Protocol version 6 (IPv6)“



- **Source address selection**
  - No impact when destination is an IPv4 address
  - For IPv6 destinations, the new configurable rules kick in if neither SOURCEVIPA nor SRCIP selects a source IP address
  - New rules configurable via new TCP/IP profile statements

- **Destination address selection**
  - Governs the order in which IP addresses are returned by the getaddrinfo() resolver call
  - No changes for gethostbyname()
  - No changes if IPv6 is not enabled
  - SORTLIST continues to govern order of IPv4 addresses
  - New configurable rules may be used to alter preference for IPv6 over IPv4 addresses to the opposite, but otherwise no impact to IPv4 destinations

Y'all remember this: z/OS TCP/IP source IP address selection logic (simplified!)

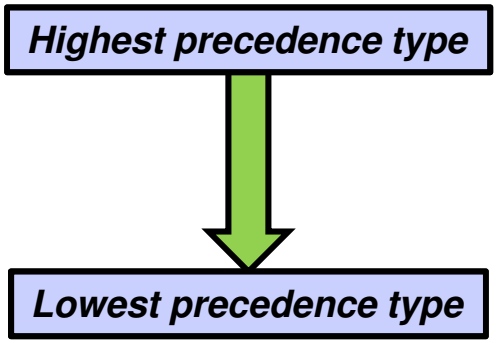


Route precedence

- Which route is installed in the routing table when routes to the same destination are received from multiple sources?

  1. Non-replaceable static routes
  2. OSPF routes
  3. RIP routes
  4. Router advertisement routes (IPv6)
  5. Replaceable static routes

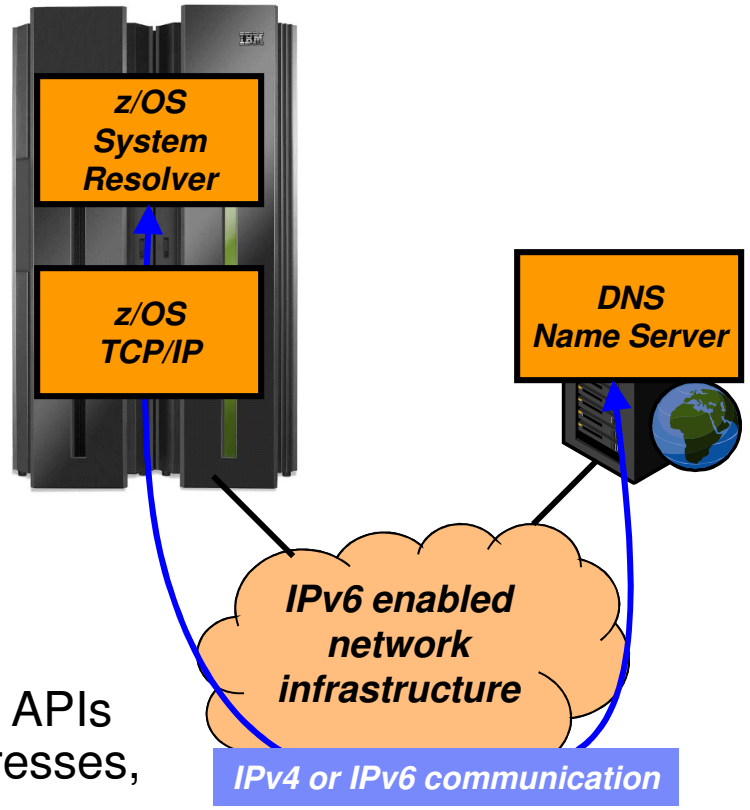
These may now also have precedence among themselves



- Managed by the TCP/IP stack and OMPROUTE in combination
- IPv6 default router advertisements have been expanded with metric
  - Router advertisement routes may now have a precedence associated
  - Allows for differentiation among multiple routers that all provide a default route
  - All router advertisements are kept by TCP/IP in case a higher precedence routes goes away
    - These kept, but currently unused router advertisements can now be displayed by netstat
- IPv6 router advertisement has also been expanded with the ability for a router to inform about off-link destinations (network prefixes) that can be reached through the router
  - These are also associated with precedence information

Resolver support for IPv6 connections to DNS name servers

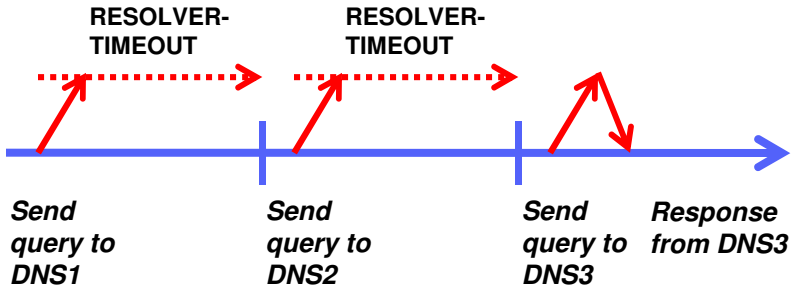
- Allows the system resolver to send requests to DNS name servers using IPv6 communication
  - Specify IPv6 addresses on the NSINTERADDR and NAMESERVER configuration statements
  - Resolver sends queries using IPv4, IPv6 or both based on the configuration



- Applications cannot manipulate IPv6 addresses using low-level resolver API calls, such as res_query and res_search
  - Only IPv4 addresses are supported on these APIs
  - The entire list, containing IPv4 and IPv6 addresses, is used for searching
    - Unless the application modifies the list, in which case only the returned IPv4 addresses are used

- The type of address returned (IPv4/IPv6) is not tied to the transport between the resolver and the name server. IPv6 addresses can be returned before z/OS V1R12

Improved resolver reaction to unresponsive name servers



Assume:

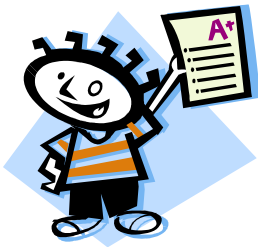
- ***3 name servers in TCPIP.DATA***
- ***2 first are un-responsive***
- ***RESOLVERTIMEOUT 30 seconds***

It takes 60+ seconds to get a response, and it will do so for every query made to the resolver

- Un-responsive name servers can impact performance significantly
  - Based on the setting of number of name servers, timeout, and retry limit in TCPIP.DATA
    - Beware that default RESOLVERTIMEOUT used to be 30 seconds – should be lowered to seconds or sub-seconds!
    - Default changed to 5 seconds in z/OS V1R12
- Before z/OS V1R12, no warning messages have been issued when name servers repetitively time out
- z/OS V1R12 adds messages to the console when name servers are un-responsive
- Configurable un-responsiveness threshold: percentage of failed queries over a 5-minute period
  - Default 25%
- A message will also be issued when a name server is deemed to have become responsive again

```
EZZ9308E UNRESPONSIVE NAME SERVER DETECTED AT IP ADDRESS 9.43.25.200
EZZ9310I NAME SERVER 9.43.25.200
           TOTAL NUMBER OF QUERIES SENT          6000
           TOTAL NUMBER OF FAILURES              2100
           PERCENTAGE                              35%
```

IPv6 – State of z/OS and z/OS Communications Server



A few applications and add-on functions still need IPv6-enablement: Intrusion Detection Services, remote commands, IPSec NAT traversal, and some less frequently used applications and functions.

z/OS Communications Server applications and z/OS-unique functions are not defined in any compliance criteria, but many are already IPv6 enabled:

- ***High-availability functions IPv6-enabled: DVIPA, Sysplex, etc.***
- ***Add-ons such as IP Security, AT-TLS, etc.***
- ***Applications (TN3270, EE, FTP, CSSMTP, etc.)***
- ***Management functions (SNMP, SMF records, NMI, OSPF, etc.)***
- ***Subsystems are picking up (WAS, CICS, MQ, etc.)***

Good for real, full-function, reliable "production" use

Good for US government use

US Government compliance criteria

1. ***Department of Defense (DoD)***
2. ***All other agencies via NIST (National Institute of Standards and Technology)***

Good for "commercial" use

IPv6 Ready Logo compliance based on "Tahi" test

IPv6 Base RFC compliance based on standards bodies specifications

Important z/OS applications and subsystems are already IPv6 enabled

z/OS V1R10 CS certified by DoD in 2008
z/OS V1R12 CS certified By USGv6 in 2010

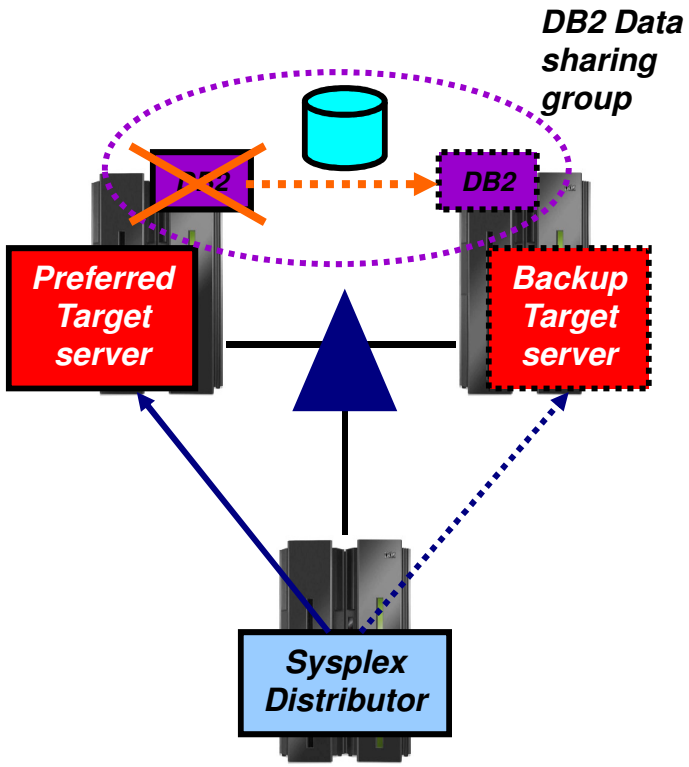
z/OS V1R8 and V1R11 CS certified as IPv6 Phase 2 Ready

Started in z/OS V1R4 CS – continually updating

Availability and business resilience

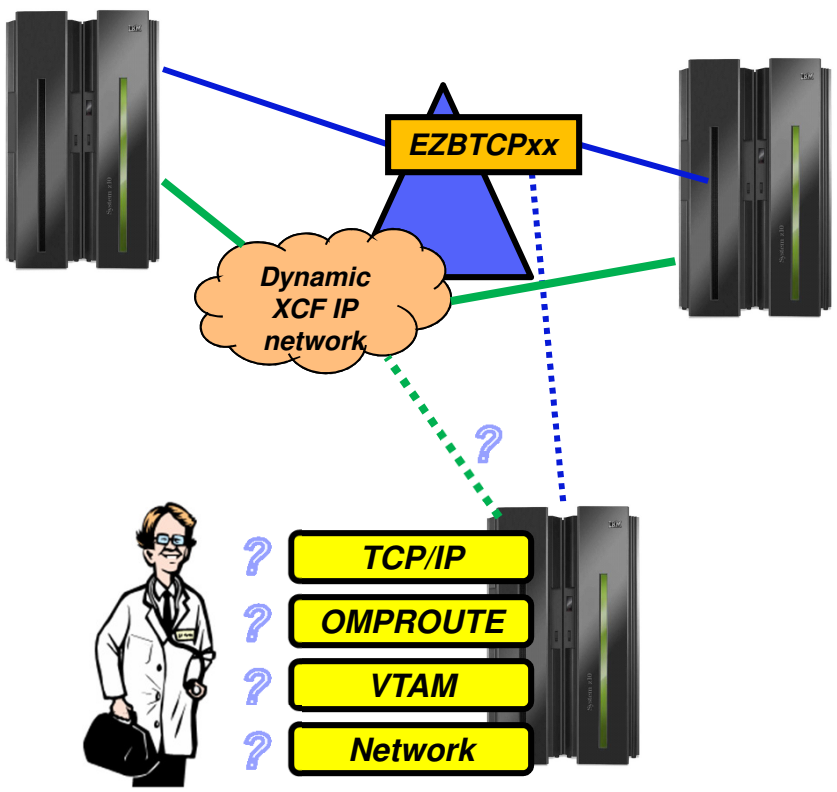
Sysplex Distributor hot standby support

- Have a single target server to receive all new connection requests
  - While other target servers are active but not receiving any new connection requests
  - Automatically route traffic to a backup target server when the active target server is not available
- Enable using a new HOTSTANDBY distribution method
  - One preferred target
    - AUTOSWITCHBACK option - switch to the preferred target if it becomes available
    - No auto switch back if reason for original switch was health problems
    - Use a V TCPIP Quiesce and Resume sequence
  - One or more backup targets ranked in order of preference
  - A target is not available when:
    - Not ready OR
    - Route to target is inactive OR
    - If HEALTHSWITCH option configured – target is not healthy when
      - TSR = 0% OR
      - Abnormal terminations = 1000 OR
      - Server reported Health = 0%




```
VIPADEFINE DVIPA1
VIPADISTRIBUTE DISTMETHOD HOTSTANDBY
  AUTOSWITCHBACK HEALTHSWITCH
  DVIPA1 PORT nnnn
  DESTIP  XCF1 PREFERRED
  DESTIP  XCF2 BACKUP 50
  DESTIP  XCF3 BACKUP 100
```

Sysplex autonomics extended with internal TCP/IP component abend monitoring



Sick? Better remove myself from the IP Sysplex!

Feeling better? Maybe it's time to rejoin the IP Sysplex

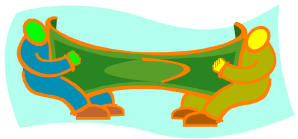
- Monitoring:
  - Monitor CS health indicators
    - Storage usage critical (>90%) - CSM, TCPIP Private & ECSA
      - For more than TIMERSECS seconds
  - Monitor dependent networking functions
    - OMPROUTE availability
    - VTAM availability
    - XCF links available
  - Monitor for abends in Sysplex-related stack components
    - Selected internal components that are vital to Sysplex processing
      - Does not include "all" components
  - Selected network interface availability and routing
  - ***Monitor for repetitive internal abends in non-Sysplex related stack components***
    - ***5 times in less than 1 minute***

New in z/OS V1R12

- Actions:
  - Remove the stack from the IP Sysplex (manual or automatic)
    - Retain the current Sysplex configuration data in an inactive state when a stack leaves the Sysplex
  - Reactivate the currently inactive Sysplex configuration when a stack rejoins the Sysplex (manual or automatic)

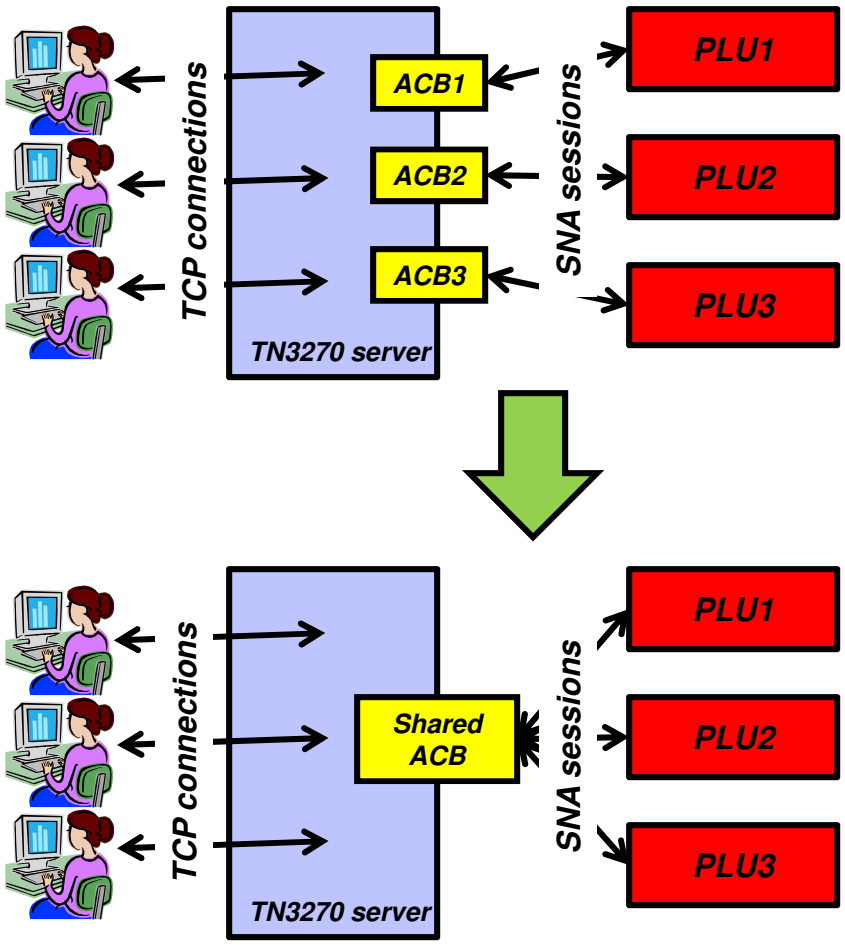
z/OS V1R12 Communications Server – Technical Update

Scalability / Performance / Constraint Relief and Accelerators



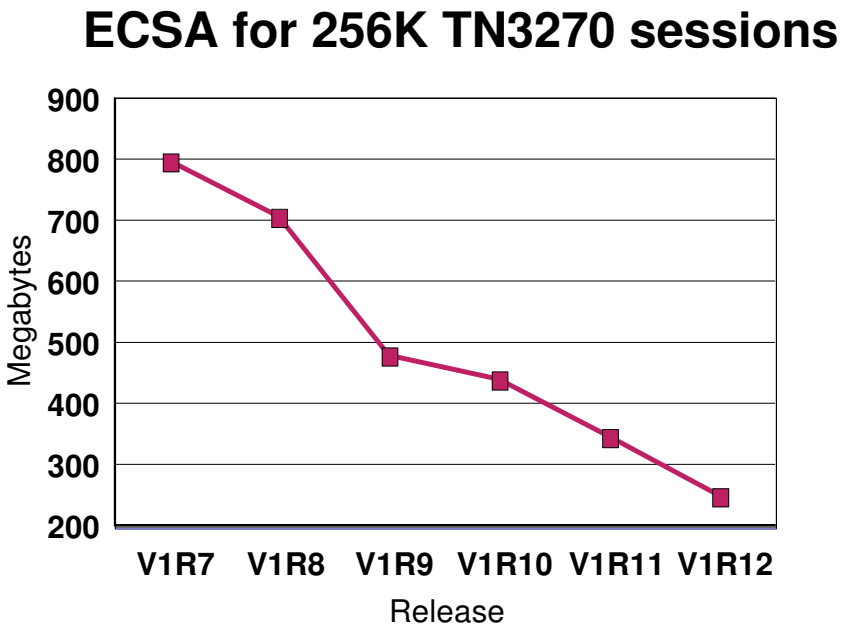
TN3270 server improvements – shared ACB support for improved performance and reduced ECSA storage use

- Telnet shared ACB support can be turned on or off with a simple statement in TELNETGLOBALS section
- VTAM model statements must be used to define the Telnet LUs
- Shared ACBs remain open until the Telnet server is ended.
  - Improve path length for client logon by using an ACB which is already open
  - Improve path length for client logoff by avoiding CLOSE ACB
  - Improve path length for Telnet termination by having fewer ACBs to close
  - Reduce the likelihood of Telnet hangs due to CLOSE ACB
  - Reduce TN3270 server ECSA usage
- No change to VTAM display commands



TN3270 server ECSA usage improvement up to and including z/OS V1R12 Communications Server

Release	ECSA for 256K TN3270 sessions
V1R7	798M
V1R8	708M
V1R9	480M
V1R10	440M
V1R11	347M
V1R12 (1)	249M



The numbers are configuration dependent, but they should give you an idea of the magnitude of the savings achieved in the recent releases.

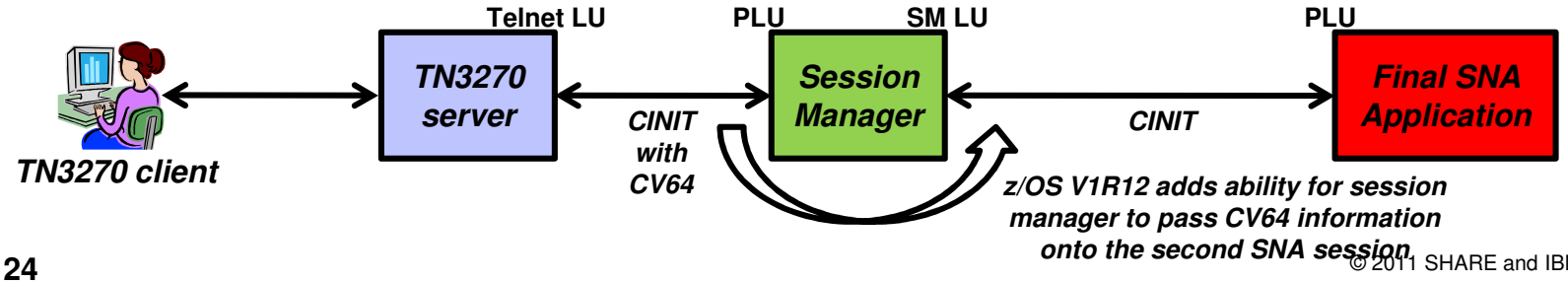
Note (1): The V1R12 number is a preliminary number based on use of shared ACBs.

TN3270 server improvements – IP management information through a relay-mode session manager

- TN3270 server passes selected IP management information to the SNA side via a control vector known as a "CV64"
  - CV64 includes client IP address, port, optionally host name and secure connection status
  - A VTAM display of the Telnet LU includes some IP information

```
IST1727I DNS NAME: CRUSET60P.RALEIGH.IBM.COM
IST1669I IPADDR..PORT 9.27.40.41..3907
```

  - The CV64 is also passed to the SNA primary logical unit (PLU) via its logon exit
- When the SNA PLU is a session manager that relays the SNA session over another LU to the final SNA application PLU, the CV64 information is lost on that second session
  - The session manager has no SNA APIs available to propagate the CV64 information
- z/OS V1R12 adds such an API, allowing an enabled session manager to pass the CV64 information to the final SNA application



Additional TN3270 server usability enhancements

- OMVS can be shutdown and restarted without re-IPLing z/OS
  - F OMVS,Shutdown
  - F OMVS,Restart
- Following the shutdown of OMVS, you are supposed to manually stop telnet
  - If Telnet stays up after OMVS is restarted, Telnet behavior is unpredictable.
- In z/OS V1R12 Telnet server address spaces register with OMVS and get notified when OMVS is being shut down
  - Telnet will shut down with OMVS
    - OMVS shutdown is delayed until Telnet has shut down
  - Must be restarted after OMVS has been restarted

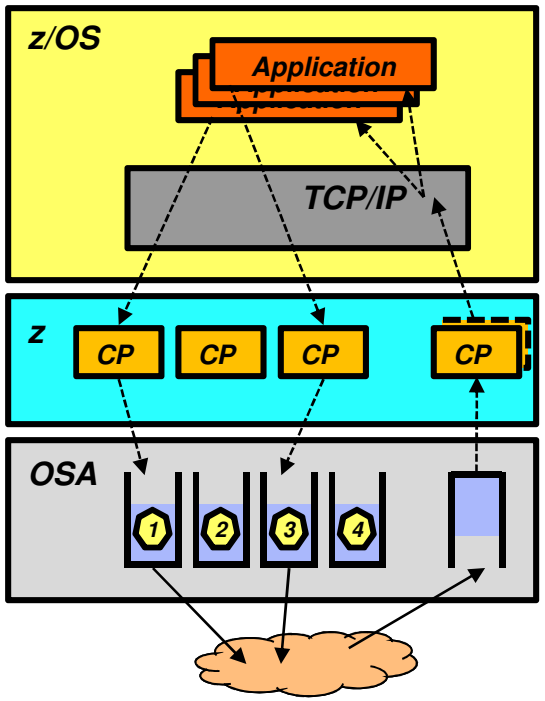
```
F OMVS,SHUTDOWN
BPXI055I OMVS SHUTDOWN REQUEST ACCEPTED
EZZ6008I TELNET STOPPING
EZZ6028I TELNET TRANSFORM HAS ENDED
EZZ6010I TELNET SERVER ENDED FOR PORT  3023
EZZ6010I TELNET SERVER ENDED FOR PORT  2023
EZZ6010I TELNET SERVER ENDED FOR PORT  1024
EZZ6010I TELNET SERVER ENDED FOR PORT  1023
EZZ6009I TELNET SERVER STOPPED
```

- A new option is passed in the CV64 control vector to an SNA primary LU on the CINIT flow
  - The option informs the SNA application if the TN3270 connection is a secure connection or not
  - Can be used by the SNA application to determine requirements for additional security
- To prevent a change of TN3270 connection attributes during a takeover process, a new configuration option is added to the takeover definitions:
  - TKOGENLURECON and TKOSPECLURECON – SAMECONNTYPE
- TN3270 server messages will now indicate the name of the TN3270 server address space instead of just saying ‘TELNET’

Pre V1R12 OSA inbound/outbound processing overview

- Queued Direct IO (QDIO) uses multiple write queues for outbound traffic separation
  - Outbound traffic is separated by priority (policy or WLM)
  - Multiple CPs can be used to manage the write queues

- QDIO uses only one read queue
  - All inbound traffic is received on the single read queue
  - Multiple CPs are used only when data is accumulating on the queue
    - During bursts of inbound data
  - Single process for initial interrupt and read buffer packaging
  - TCP/IP stack performs inbound data separation
    - Sysplex distributor traffic
    - Bulk inbound, such as FTP
    - IPv4/IPv6
    - EE traffic
    - Etc.
  - z/OS Communications Server is becoming the bottleneck as OSA nears 10GbE line speed
    - Inject latency
    - Increase processor utilization
    - Impede scalability

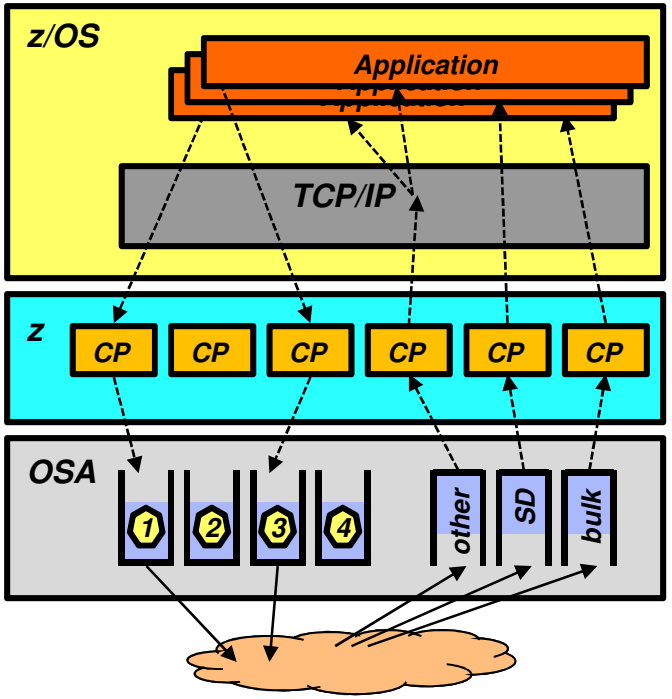


Performance problems observed for bulk inbound traffic:

- ***Multiple processes run when data is accumulating on the read queue***
- ***Inbound data for a single TCP connection can arrive at the TCP layer out of order***
- ***TCP transmits a duplicate ACK every time it sees out of order data***
- ***Sending side enters fast retransmit recovery***

OSA multiple inbound queue support: improved bulk transfer and Sysplex Distributor connection routing performance

- Allow inbound QDIO traffic separation by supporting multiple read queues
  - "Register" with OSA which traffic goes to which queue
  - OSA-Express Data Router function routes to the correct queue
- Each input queue can be serviced by a separate process
  - Primary input queue for general traffic
  - One or more ancillary input queues (AIQs) for specific traffic types
- Supported traffic types
  - Bulk data traffic queue
    - Serviced from a single process - eliminates the out of order delivery issue
  - Sysplex distributor traffic queue
    - SD traffic efficiently accelerated or presented to target application
  - All other traffic not backed up behind bulk data or SD traffic
- Dynamic LAN idle timer updated per queue
- Early performance data (**Note**: your mileage will vary)
  - Request/Response transaction rate improvements (up to 55%)
  - Streaming workload throughput improvements (up to 40%)

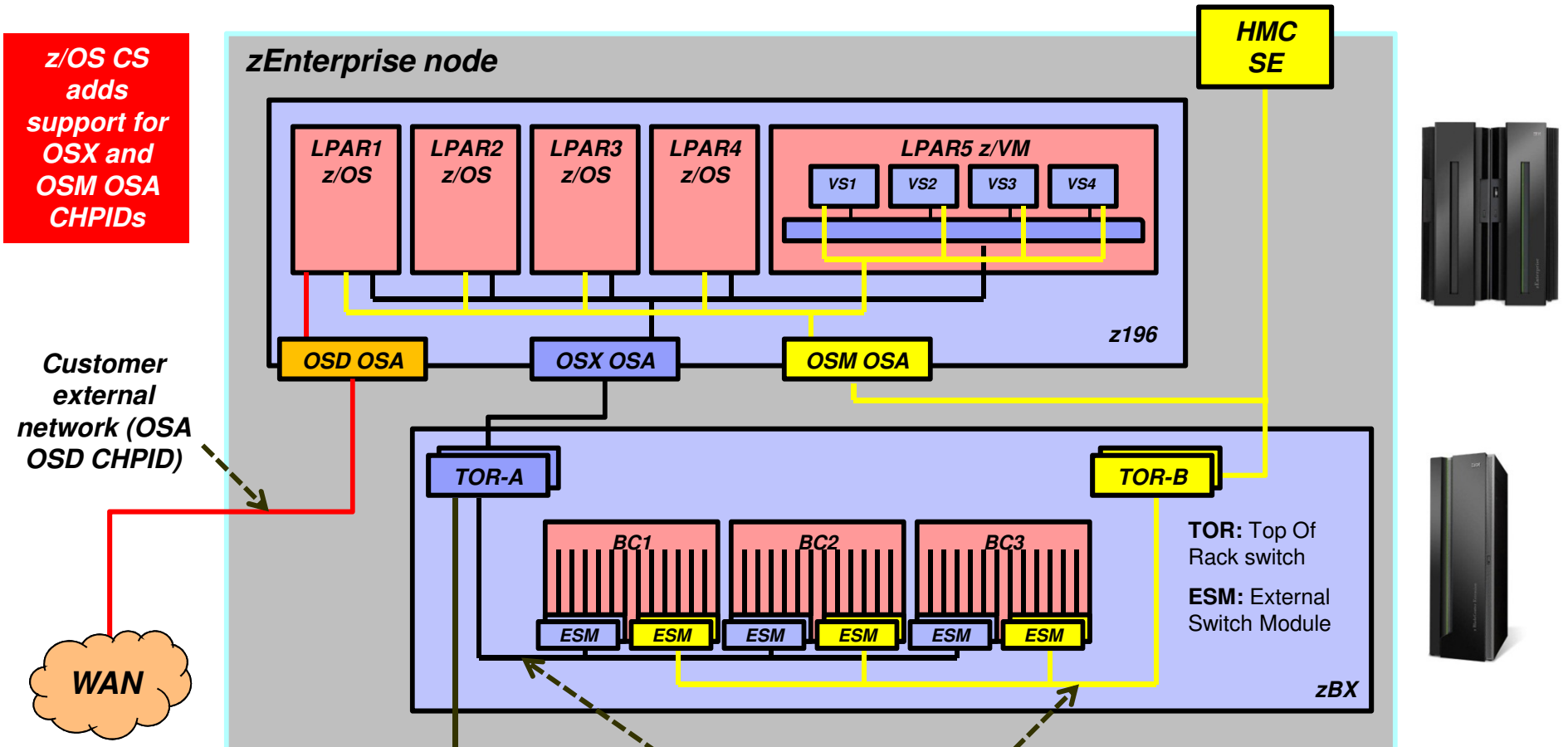


TCP/IP defines and assigns traffic to queues dynamically based on local IP address and port

Bulk traffic
- ***Application sets send or receive buffer to at least 180K***
- ***Registered per connection (5-tuple)***

SD traffic
- ***Based on active VIPADISTRIBUTE definitions***
- ***Registered on DVIPA address***

Support in z/OS Communications Server for internal networks in an IBM zEnterprise System



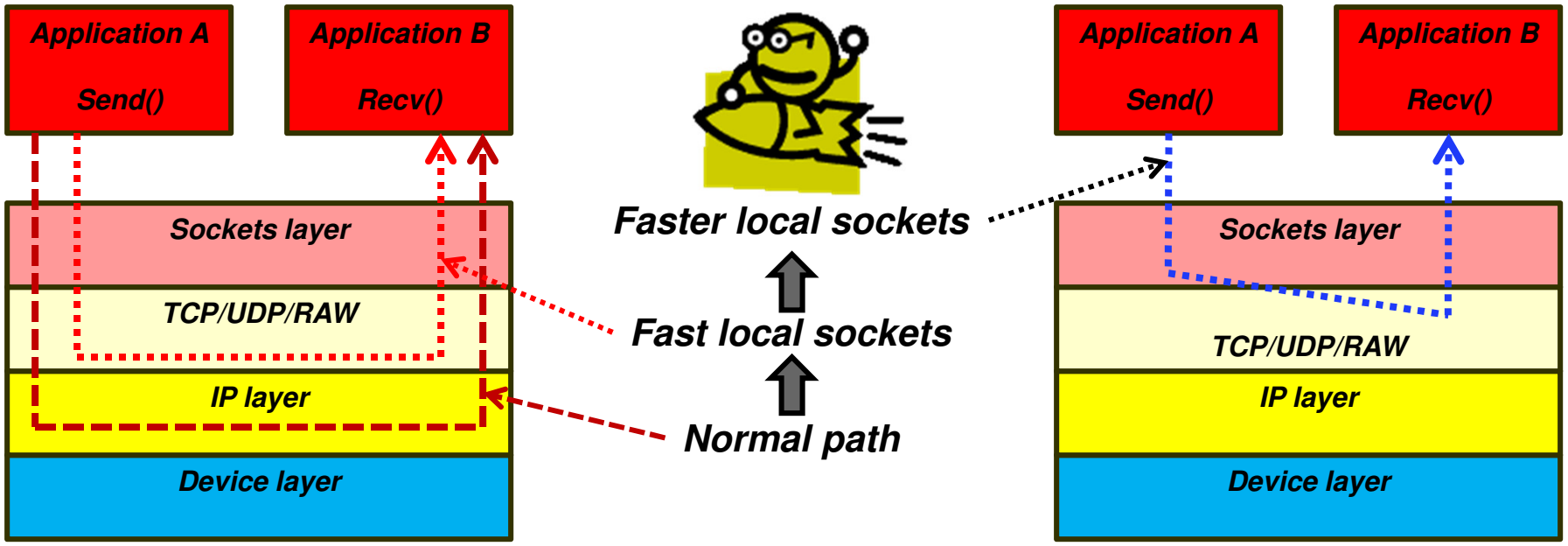
Blades may be:

- IBM Smart Analytics Optimizer for DB2 for z/OS,
- IBM POWER7
- IBM System x blade [1]
- WebSphere DataPower Appliance [1]

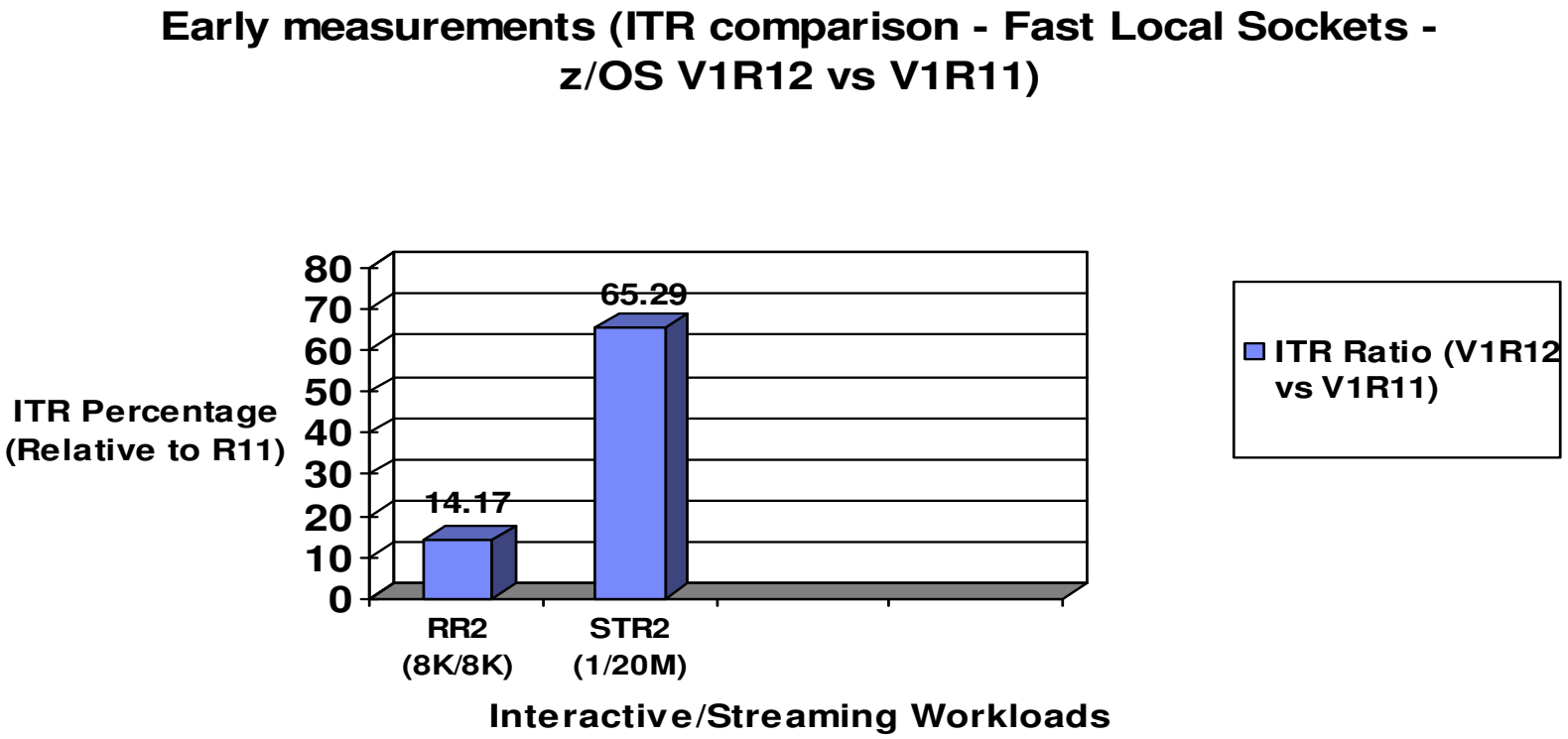
Note 1: All statements regarding IBM future direction and intent are subject to change or withdrawal without notice, and represents goals and objectives only.

Performance improvements for fast local sockets

- Fast local sockets (FLS)
  - Optimized path through TCP/IP
  - Bypassing the IP layer
    - Data placed on TCP send queue
    - Data is then moved to TCP receive queue
    - ACKs built and sent from receive side
  - Used when socket end-points are on same stack
  - Dynamic; no configuration required
- Faster local sockets (Turbo FLS)
  - Bypasses processing on both sending and receiving side
    - Data no longer placed on TCP send queue
    - Data is placed directly onto receive queue bypassing TCP inbound processing
    - Data no longer ACKed
  - Enabled automatically; no configuration changes
    - Reverts to fast local sockets if packet trace or AT-TLS is enabled
    - No impact for data trace



Performance improvements for fast local sockets…



- Faster local sockets (FLS) - Summary
  - Exploiting the co-location pattern of applications using sockets
  - Leveraging the co-location to provide substantial performance benefits (Cross-memory mode, etc).
    - And doing so transparently (to both applications and system administrators)

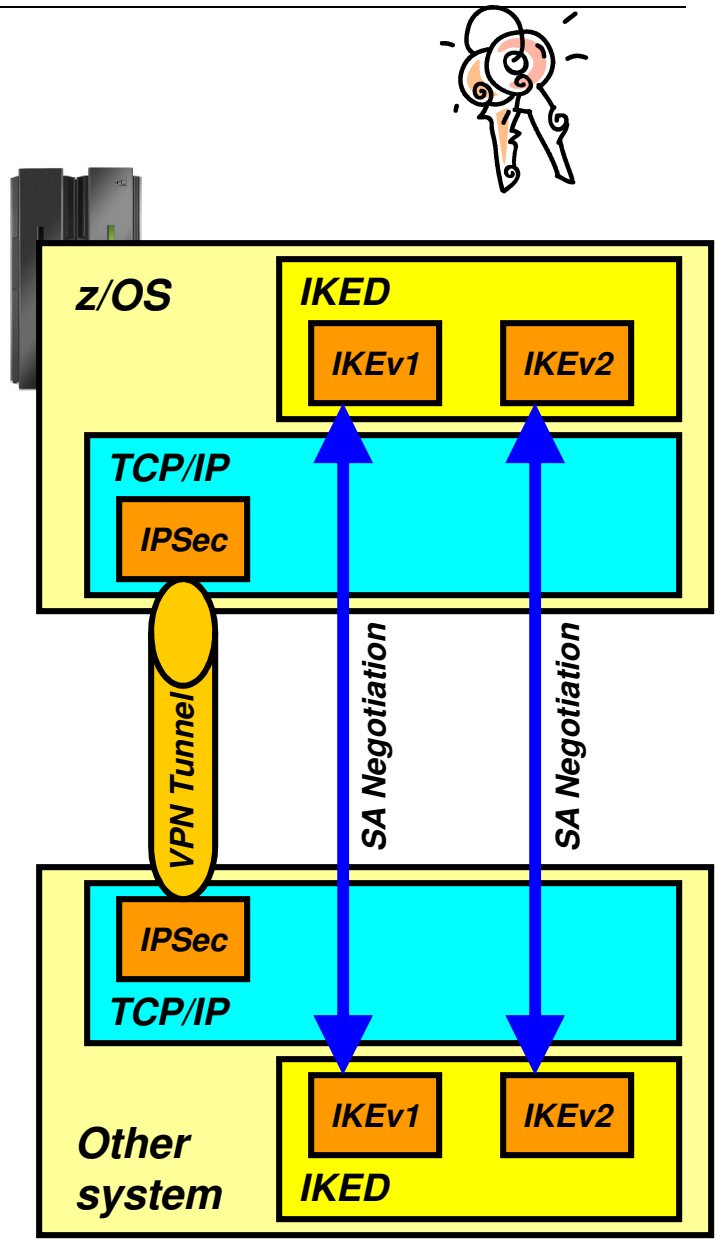
Note: The performance measurements discussed in this presentation are preliminary z/OS V1R12 Communications Server numbers and were collected using a dedicated system environment. The results obtained in other configurations or operating system environments may vary.

z/OS V1R12 Communications Server – Technical Update

Security

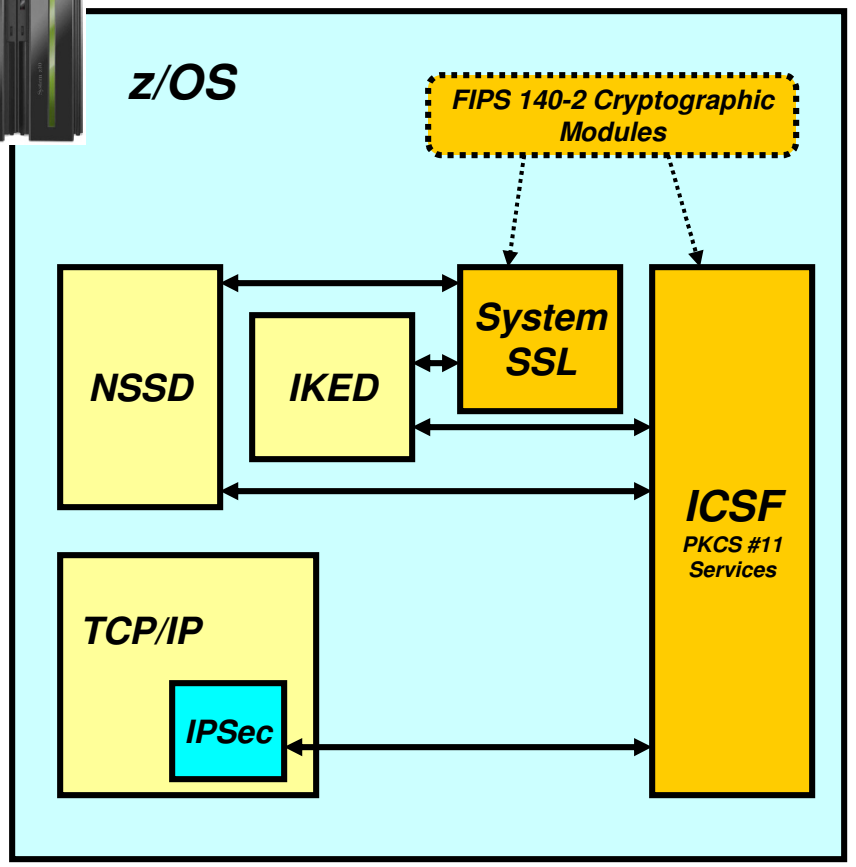
IKEv2 support

- The Internet Key Exchange (IKE) protocol provides automated management of cryptography keys and security associations used by IPSec
  - Either a portion of the data path or the entire data path can be secured
- IKEv2 is the newest version of the IKE protocol
  - Designed to replace the current version, IKEv1
  - IKEv2 is a rewrite of IKEv1 and almost wholly incompatible with IKEv1
  - However, both protocol versions need to be supported into the foreseeable future
- The existing IKE daemon will support both IKEv1 and IKEv2
  - Both protocols may be used at the same time using a single IKE daemon



IKE, IPSec, and NSS FIPS 140 mode

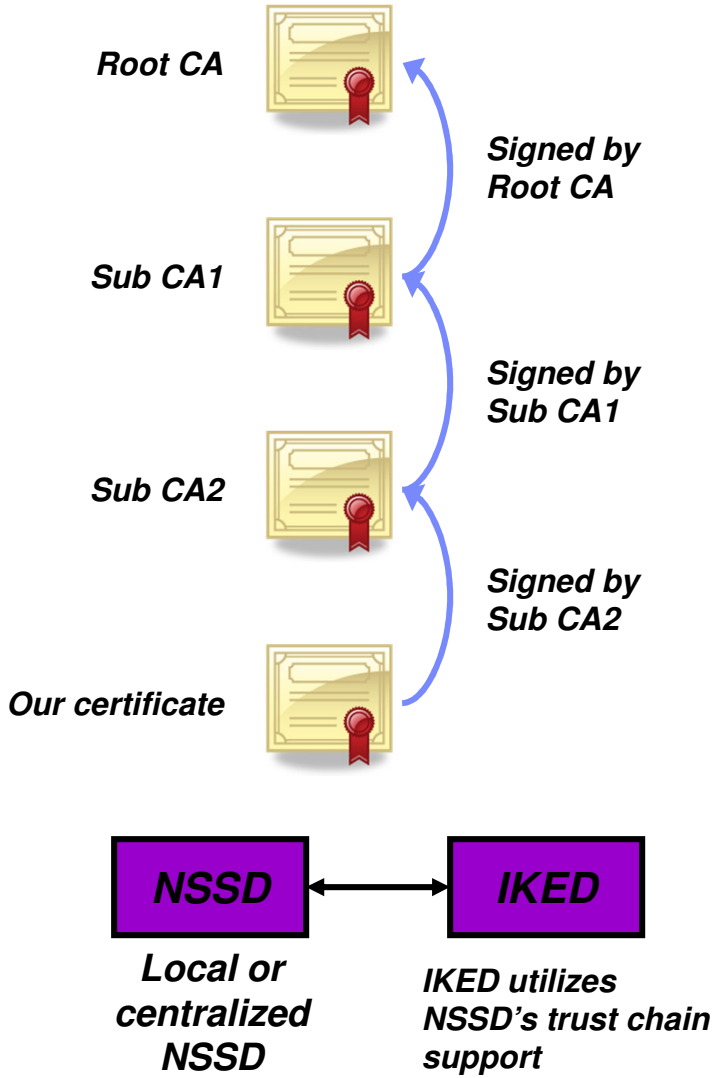
- FIPS 140 defines requirements and standards for cryptographic modules used within the US Government and elsewhere
  - Applies to cryptographic modules – not systems or applications
  - On z/OS, both System SSL and ICSF's PKCS #11 services are designed to address FIPS 140-2 requirements
- IKE, IPSec and NSS offer an optional FIPS 140 mode
  - When enabled, all IKE, IPSec and NSS IPSec-related crypto operations are performed through FIPS 140 mode System SSL or ICSF calls
  - TCP/IP stacks are individually enabled
  - IKED must be configured for FIPS 140 mode if any TCP/IP stack is enabled for FIPS 140 mode
- FIPS 140 mode reflected via the Network Management Interface



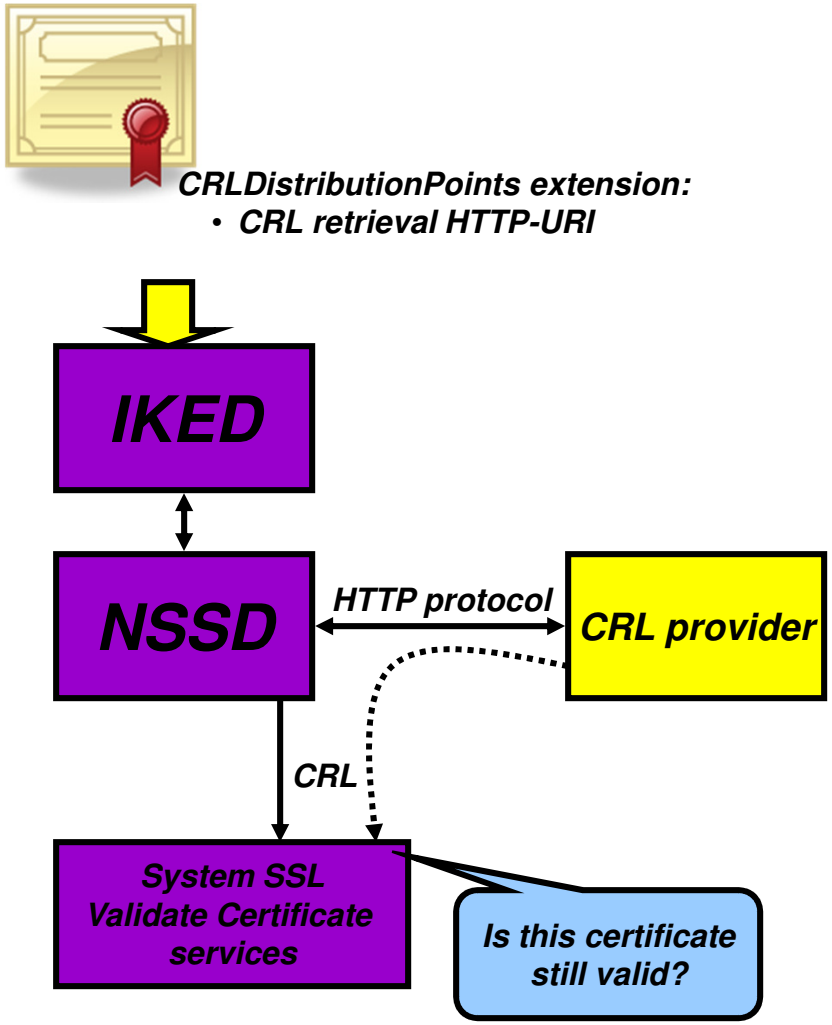
Note: AT-TLS added support to address FIPS 140-2 requirements in z/OS V1R11

IPSec support for certificate trust chains

- RFC 4306 requires support for trust chains.
  - NSSD is updated to provide support for trust chains.
  - The maximum number of certificates supported in a trust chain is 32.
- IKED is updated to exploit NSSD's trust chain support.
  - IKED's local certificate processing will not be updated to support trust chains.
  - As a result, trust chain support in IKED will only be available to stacks that are configured as a network security client.
  - When a stack is configured as a network security client, IKED will utilize trust chain support for both IKEv1 and IKEv2 exchanges.



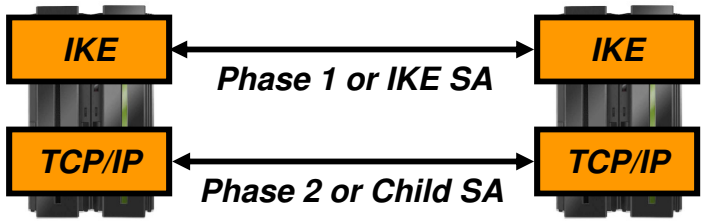
IPSec support for certificate revocation lists (CRLs)



- When IPSec authenticates a digital signature, it needs to ensure the signing certificate is still valid
- NSSD will retrieve CRLs using information in the CRLDistributionPoints extension in a certificate
  - HTTP-URIs only
- NSSD will pass CRLs to System SSL
- System SSL will validate the certificate against the CRL
  - To ensure the certificate is still valid
    - Has not expired or been revoked
- NSSD will not support retrieval of CRLs from LDAP servers
- For IKEv2, IKED depends on NSSD for this function

IPSec algorithm support

IKEv1 Phase 1 and IKEv2 IKE SA			IKEv1 Phase 2 and IKEv2 Child SA		
Purpose	**Existing**	**New**	**Purpose**	**Existing**	**New**
Encryption algorithm	DES, 3DES, AES_CBC KeyLength 128	AES_CBC Keylength 256	Encryption algorithm	DES, 3DES, AES_CBC KeyLength 128	AES_CBC KeyLength 256, AES_GCM_16 KeyLength 128 | 256
Diffie-Hellman group	Group1, Group2, Group5, Group14	Group19, Group20, Group21, Group24	Authentication algorithm	HMAC_MD5, HMAC_SHA1	AES_GMAC_128 | 256, AES128_XCBC_96, HMAC_SHA2_256_128, HMAC_SHA2_384_192, HMAC_SHA2_512_256
IKEv1 hash algorithm	MD5, SHA1	SHA2_256, SHA2_384, SHA2_512	Perfect forward secrecy group	Group1, Group2, Group5, Group14	Group19, Group20, Group21, Group24
Partner authentication	PreSharedKey, RSASignature	ECDSA-256, ECDSA-384, ECDSA-521 (these are only for IKEv2)			
IKEv2 message verification algorithm	N/A	HMAC_MD5_96, HMAC_SHA1_96 AES128_XCBC_96, HMAC_SHA2_256_128, HMAC_SHA2_384_192, HMAC_SHA2_512_256			
IKEv2 pseudo random function	N/A	HMAC_MD5, HMAC_SHA1 AES128_XCBC, HMAC_SHA2_256, HMAC_SHA2_384, HMAC_SHA2_512			



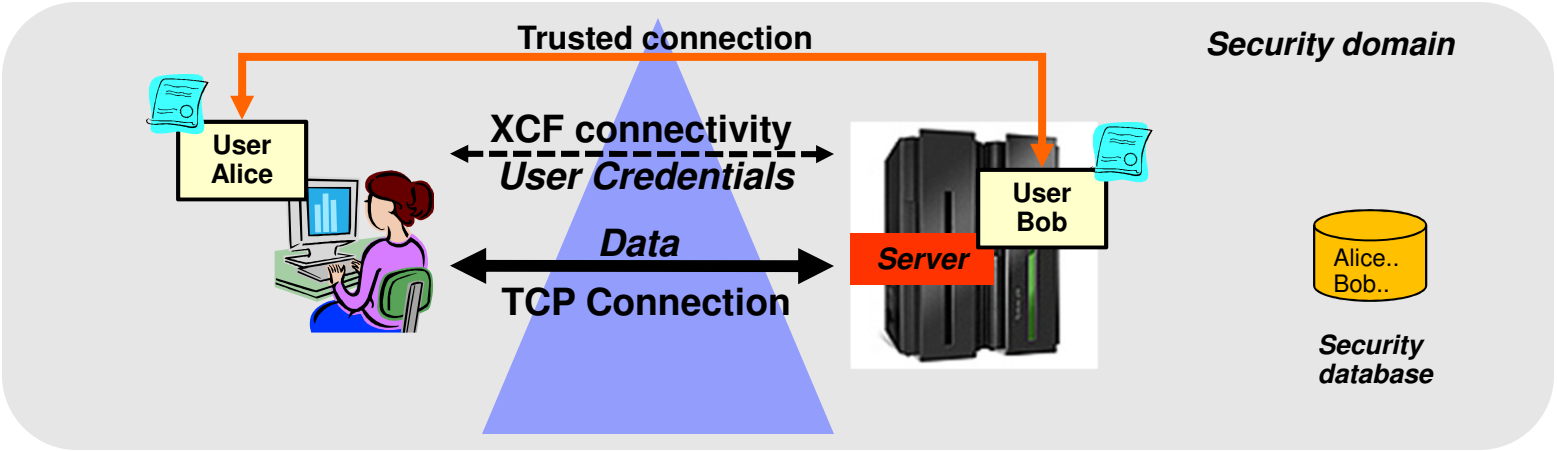
SA: Security Association aka. the tunnel

z/OS V1R12 IPSec-related RFC status - overview

RFC	Title
3566	The AES-XCBC-MAC-96 Algorithm and Its Use With IPsec
3948	UDP Encapsulation of IPsec ESP Packets
4106	The Use of Galois/Counter Mode (GCM) in IPsec Encapsulating Security Payload (ESP)
4109	Algorithms for Internet Key Exchange version 1 (IKEv1)
4301	Security Architecture for the Internet Protocol
4302	IP Authentication Header
4303	IP Encapsulating Security Payload (ESP)
4304	Extended Sequence Number (ESN) Addendum to IPsec Domain of Interpretation (DOI) for Internet Security Association and Key Management Protocol (ISAKMP)
4306	Internet Key Exchange (IKEv2) Protocol
4307	Cryptographic Algorithms for Use in the Internet Key Exchange Version 2 (IKEv2)
4308	Cryptographic suites for IPSec
4434	The AES-XCBC-PRF-128 Algorithm for the Internet Key Exchange Protocol (IKE)
4718	IKEv2 Clarifications and Implementation Guidelines
4753	ECP Groups For IKE and IKEv2
4754	IKE and IKEv2 Authentication Using the Elliptic Curve Digital Signature Algorithm (ECDSA)
4809	Requirements for an IPsec Certificate Management Profile
4835	Cryptographic Algorithm Implementation Requirements for Encapsulating Security Payload (ESP) and Authentication Header (AH)
4868	Using HMAC-SHA-256, HMAC-SHA-384, and HMAC-SHA-512 with IPsec
4869	Suite B Cryptographic suites for IPSec
4945	The Internet IP Security PKI Profile of IKEv1/ISAKMP, IKEv2, and PKIX
5282	Using Authenticated Encryption Algorithms with the Encrypted Payload of the Internet Key Exchange version 2 (IKEv2) Protocol

Trusted TCP connections within a z/OS Sysplex or Subplex

- Allow TCP connection endpoints within a Sysplex to establish a trust relationship
  - Exchanges security credentials that identify the security context of the other endpoint
    - Without the overhead and CPU-related costs of SSL/TLS with client authentication
  - Requires no application protocol changes
    - Simple API call to the TCP/IP stack
    - Transparent to the client application
  - Security credentials exchanged using secure XCF messaging
    - Application traffic may take any network path between the client and server
- Support these new socket API options for C/C++ (LE), Unix System Services Callable (BPXxxxx), and JAVA



z/OS V1R12 Communications Server – Technical Update

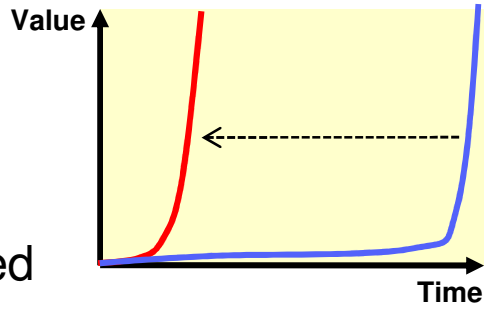
System Management and Monitoring

IBM

Focus on Consumability, Simplification and Time to Value

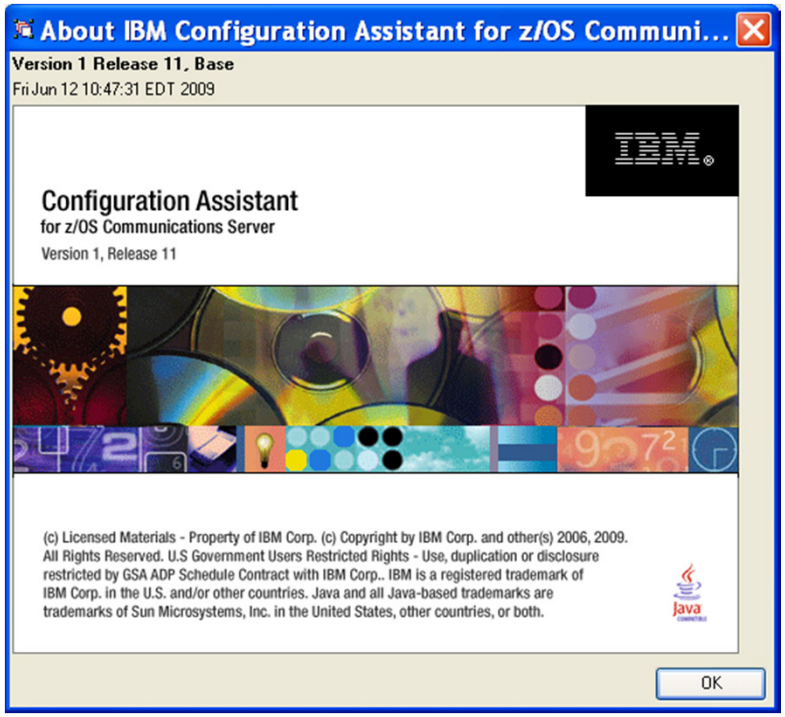
- Consumability and Simplification – what is it all about?

- Is this about using Graphical User Interfaces to configure networking functions?
  - Yes, the z/OS Communication Server Configuration Assistant is certainly a key step towards that direction
  - Continually improved since its introduction
    - Goal: Simplify tasks required to configure policy based networking functions (IPSec, IP Filters, IDS, AT-TLS, Defense Manager, NSS, QoS, Policy Based Routing, etc.)



- But it does ***not*** end there
  - Consumability and Simplification is really about "***Time to value***"
  - Delivering key features and solutions where benefits can be realized ***very quickly!***
    - By introducing functions that require minimal or no configuration on your part
    - Selecting reasonable defaults for automatically enabled functions
    - Building "***autonomic***" capabilities into our software that minimize requirements for users detecting and correcting abnormal conditions
    - Revisiting existing functions/features over time when adoption inhibitors are identified

IBM Configuration Assistant for z/OS Communications Server



The Configuration Assistant for z/OS Communications Server is a z/OSMF system management task that provides assistance in configuring TCP/IP networking policies and can help dramatically reduce the amount of time required to create network configuration files.

Use it to create configuration files for any number of z/OS images, any number of TCP/IP stacks, for the following:

- Application Transparent - Transport Layer Security
- IP Security
- Intrusion Detection Services
- Network Security Services
- Quality of Service
- Policy Based Routing
- Defense Manager Daemon

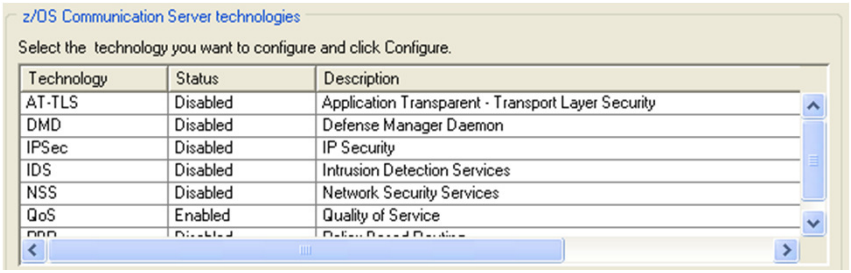
Visit the z/OSMF web page at:

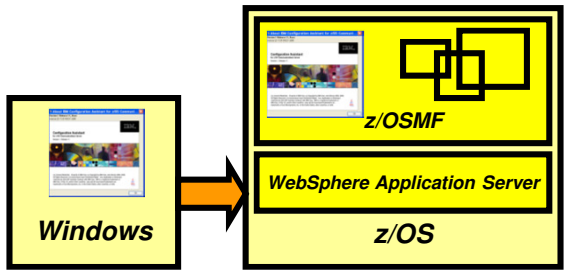
http://www-03.ibm.com/systems/z/os/zos/zosmf/

Learn even more about z/OSMF by visiting the IBM Education Assistant:

http://bit.ly/chZkQM

The Configuration Assistant for z/OS Communications Server is also available as an as-is, non-warranted, Windows-based tool that is downloadable from the Web. New functions and enhancements for the Configuration Assistant will be integrated into z/OS Management Facility, but may not be provided in the Windows-based Configuration Assistant.
Download URL: http://tinyurl.com/cgoqsa



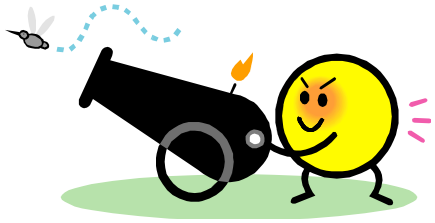


IBM Health Checker for z/OS OMPROUTE checks

- Large routing table (2000 or more routes) in the TCP/IP stack can potentially cause high CPU utilization for route changes (adds and deletes)
- Noticeable performance degradation in OMPROUTE, OMVS, and the TCP/IP stack as number of routes increase
  - Even worse with tracing enabled
- The time to process route changes may exceed OMPROUTE's Dead Router Interval for OSPF routes
  - Adjacencies with neighbors may be lost
  - Network connectivity problems may occur
- Most customer sites typically use 50-500 unique routes.
  - IP Configuration Guide documents that routing table size should be kept to a minimum:
    - OSPF: Use stub areas, route summarization, or use filters
    - RIP or Static: Use sub-netting or super-netting for route summarization or use filters
- New health checks are implemented in z/OS V1R12 to monitor the number of indirect routes in a TCP/IP stack
  - Warnings to be issued if number of indirect IPV4 or IPv6 routes exceed configurable limit (default is 2000)

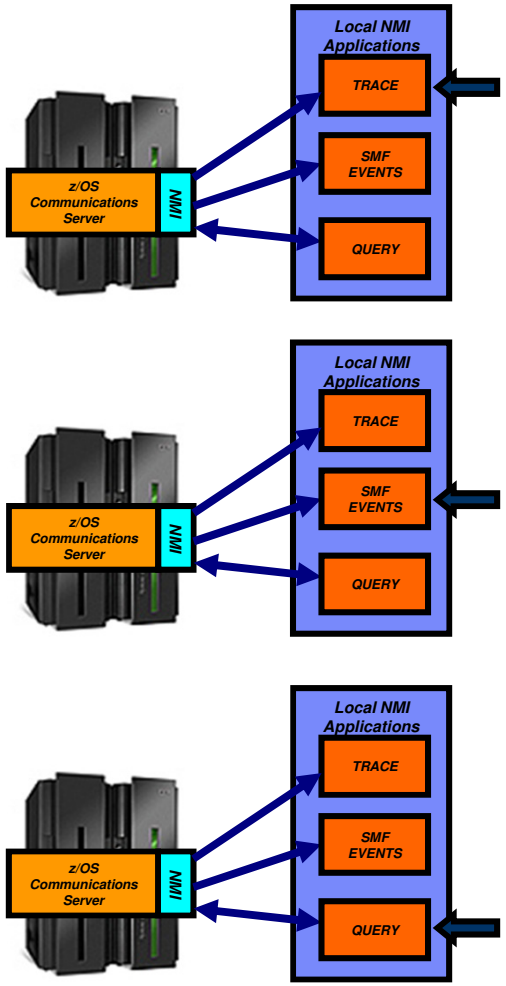
Command to drop all connections for a server

- V TCPIP,,DROP command or netstat drop command
  - Used to drop (reset) a TCP or UDP connection.
  - Must specify the connection id of the connection to be dropped.
  - Need to issue D TCPIP,,NETSTAT,CONN to find the connection id
- Can be a cumbersome process if all connections with a given server need to be dropped
  - Many display and many drop commands
- z/OS V1R12 extends the V TCPIP,,DROP command to support new parameters:
  - VARY TCPIP,,DROP,PORT=portnum,[JOBNAME=jobname,ASID=asid]
  - VARY TCPIP,,DROP,JOBNAME=jobname,[ASID=asid]
- The extended command will:
  - Scan the TCP connection table for listeners matching the filters.
  - If found, scan the table again for all child connections pointing back to listener.
  - Issue RESET for each such connection found



Overview of Network Management Interface enhancements in z/OS V1R12 Communications Server

- **Trace NMI**
  - New Data Trace records to indicate start and end of a “data flow”
    - Start record written on the first socket read or write operation
    - End record written when the socket is closed
    - Start/End records are created by default. No changes to VARY TCPIP,,DATTRACE command
  - Apply Packet Trace filters to Sysplex Distributor VIPAROUTE traffic
    - Filtering on both GRE header and encapsulated header
  - Next hop address now included in packet trace records
- **SMF event NMI**
  - Sysplex events
    - Provides support for NMI events with information similar to the earlier sysplex-related SNMP traps
  - CSSMTP events
    - Both NMI events and new SMF support
- **Query NMI**
  - Network interface and device information and TCP/IP global statistics
    - Allows applications to obtain TCP/IP interface attributes and statistics, and TCP/IP global stack statistics using the TCP/IP query NMI



Enhancements to TCP/IP storage command

- D TCPIP,,STOR

```
TCPCS     STORAGE           CURRENT    MAXIMUM      LIMIT
TCPCS     ECSA                9645K     10087K    NOLIMIT
TCPCS     POOL               14017K     14171K    NOLIMIT
TCPCS     64-BIT COMMON          1M         1M    NOLIMIT
DISPLAY TCPIP STOR COMPLETED SUCCESSFULLY
```

- Common (ECSA) usage information includes the size of the TCP/IP load modules loaded into common by dynamic LPA
  - Load module size is a stable value
  - Might be a large percentage of common usage value
  - Might mask workload related fluctuations/growth in common storage usage
- In z/OS V1R12, ECSA usage for load modules moved to separate line of the display.
- Similar changes made to the storage callable NMI interface.

```
TCPCS     STORAGE           CURRENT    MAXIMUM      LIMIT
TCPCS     ECSA                2822K     2935K    NOLIMIT
TCPCS     POOL               14194K    14194K    NOLIMIT
TCPCS     64-BIT COMMON          1M         1M    NOLIMIT
TCPCS     CSA MODULES         7419K      7419K    NOLIMIT
DISPLAY TCPIP STOR COMPLETED SUCCESSFULLY
```

Operator command to query and display OSA information

- OSA/SF has been used for years to configure OSA and display the configuration. OSA/SF has played a more central role for OSE devices (pre-QDIO) than for today's OSD devices (QDIO).
- OSD devices exclusively use IPA signals exchanged with the host to enable and configure features and register IP addresses to OSA.
- However, there has so far been no mechanism to display the information directly from OSA without OSA/SF.
- z/OS V1R12 implements a new D TCPIP,,OSAINFO command for use with OSA Express3:
  - Base OSA information
  - OSA address table information
  - Information related to the new multiple inbound queues
  - Etc.

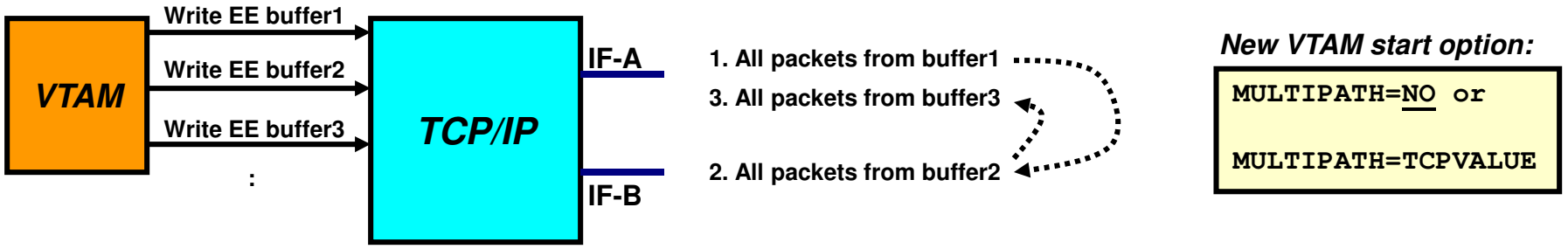
```
D TCPIP,,OSAINFO,INTFN=V6O3ETHG0,MAX=100

EZZ0053I COMMAND DISPLAY TCPIP,,OSAINFO COMPLETED SUCCESSFULLY
EZD0031I TCP/IP CS V1R12  TCPIP Name: TCPSVT    15:39:52
Display OSAINFO results for Interface: V6O3ETHG0
PortName: O3ETHG0P  PortNum: 00  DevAddr: 2D64    RealAddr: 0004
PCHID: 0270         CHPID: D6    CHPID Type: OSD  OSA code level: 5D76
Gen: OSA-E3         Active speed/mode: 10 gigabit full duplex
Media: Singlemode Fiber        Jumbo frames: Yes  Isolate: No
PhysicalMACAddr: 001A643B887C  LocallyCfgMACAddr: 000000000000
Queues defined Out: 4  In: 3   Ancillary queues in use: 2
Connection Mode: Layer 3       IPv4: No   IPv6: Yes
SAPSup: 00010293
…
```

z/OS V1R12 Communications Server – Technical Update

SNA and EE

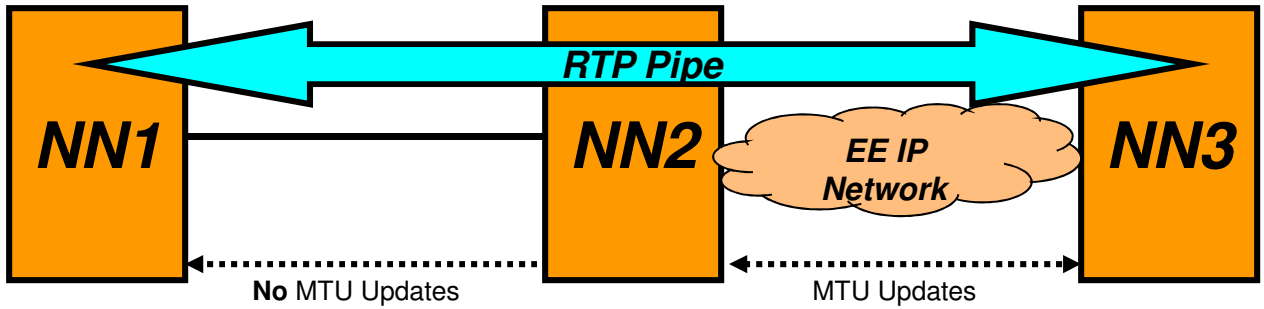
Multipath control for Enterprise Extender

- With multipath enabled in TCP/IP, all packets in one EE write buffer will be sent over one interface, and all packets in the next EE write buffer will be sent over another interface
  - A modified per-packet algorithm – really a per-EE-buffer algorithm
- Same behavior independent of PERCONNECTION / PERPACKET setting in TCP/IP
- EE traffic may incur performance issues if the different paths are not truly equal in terms of bandwidth and delay
- Per-connection multipath is generally beneficial for other TCP/IP traffic
- New support to allow TCP/IP to specify use of Multipath, but disable it by default for EE traffic



Improved recovery from RTP pipe stalls

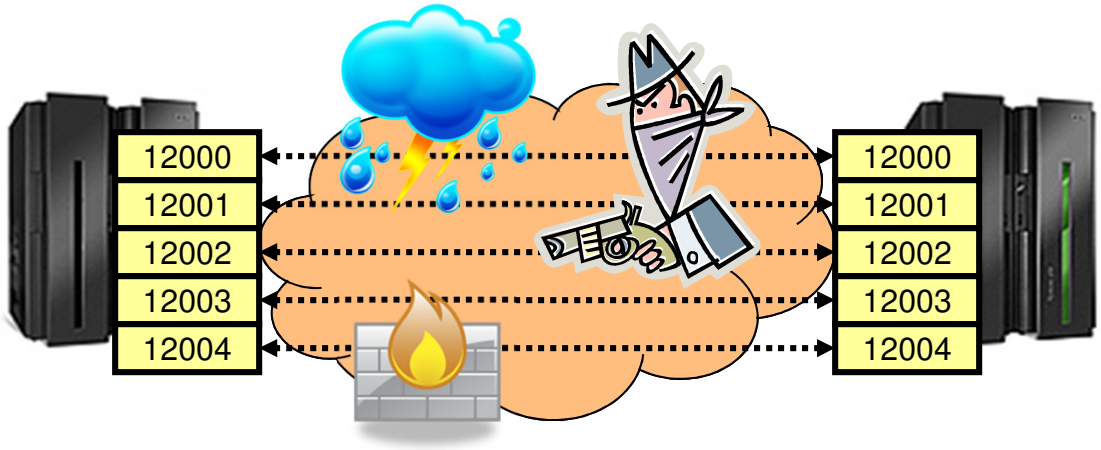
- z/OS V1R10 provided a version of Path MTU Discovery (PMTU) for Enterprise Extender.
  - However, MTU size changes over an active EE link are only communicated to the two nodes that act as the endpoint of that EE link (NN2 and NN3 below)



- If an existing RTP pipe begins on a node other than the EE link endpoint, it will not learn the PMTU-discovered MTU size, and will continue to send packets at a non-optimal size, potentially resulting in packet loss and transmission stalls.
- z/OS V1R12 adds logic for VTAM to drive the path switch logic if multiple retransmissions occur (stall detection)
  - Thereby letting NN1 above learn the new current MTU size and adapt

```
IST2335I PATH SWITCH REASON: XMIT STALL RECOVERY
```

Enterprise Extender connection health verification



- Questions:
  - Are all five EE ports reachable at EE connection initialization point in time?
  - Do all five EE ports remain reachable?
- Apart from something not working correctly, you really do not know!
- z/OS V1R12 adds optional probing logic during EE connection initialization and during the lifetime of the EE connection to verify the health of the EE connection.
  - EEVERIFY=NEVER
    - Do not send any probes
  - EEVERIFY=ACTIVATE
    - Probe during connection initialization
  - EEVERIFY=timer-interval
    - Probe during initialization and periodically at the specified timer-interval

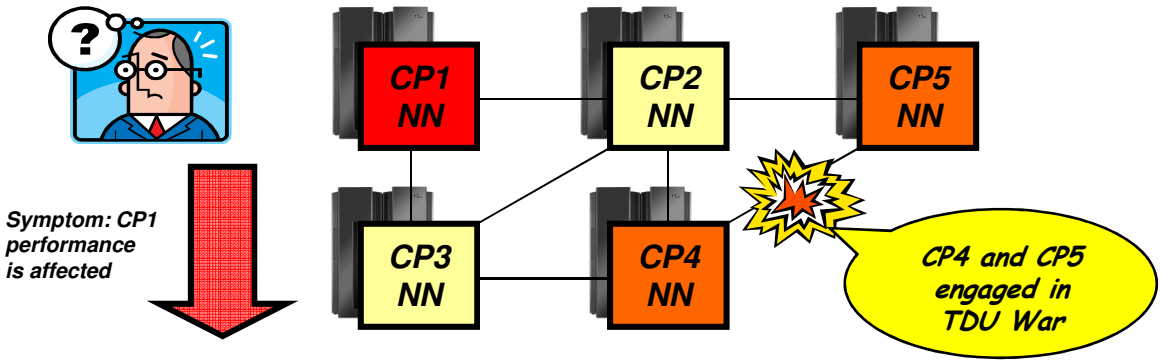
Enterprise Extender connection health verification - example

- To see all failed connections, issue the following command:

```
d net,ee,list=eeverify
 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 = XCAIP
 IST2004I LIVTIME = (10,0)      SRQTIME =    15  SRQRETRY =     3
 IST2005I IPRESOLV =    0
 IST2231I CURRENT HPR CLOCK RATE = STANDARD
 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 -------------------------------------------------------
 IST2324I EE HEALTH VERIFICATION: FAILED CONNECTION INFORMATION
 IST2325I LINE LNIP1 PU SWIP2A1 ON 12/21/09 AT 15:56:39
 IST2326I EE HEALTH VERIFICATION TOTAL CONNECTION FAILURES = 1
 IST2017I TOTAL RTP PIPES =          1     LU-LU SESSIONS =          2
 IST2018I TOTAL ACTIVE PREDEFINED EE CONNECTIONS          =          1
 IST2019I TOTAL ACTIVE LOCAL  VRN EE CONNECTIONS          =          0
 IST2020I TOTAL ACTIVE GLOBAL VRN EE CONNECTIONS          =          0
 IST2021I TOTAL ACTIVE EE CONNECTIONS                     =          1
 IST314I END
```

Enhancements to topology database diagnostics

- Enhancements in V1R11 defined a new control vector for TDU flows
  - Topology Resource Sequence Number Update (x'4E') control vector to identify node that set the RSN



- TDUDIAG start option available to control frequency of when new control vector is included
- Still required dumps and traces to diagnose TDU war
- z/OS V1R12 enhances various commands to improve the ability to better diagnose the TDU war scenario:
  - Enhance existing DISPLAY TOPO,LIST=TDUINFO output
  - New DISPLAY TOPO,LIST=TDUDIAG summary command
  - Diagnostic information from the Topology RSN Update control vector added in V1R11 is saved
  - New displays of diagnostic information from the x'4E' control vector
    - DISPLAY TOPO,LIST=TDUDIAG command for a TG
    - DISPLAY TOPO,LIST=TDUDIAG command for a node

z/OS V1R12 Communications Server – Technical Update

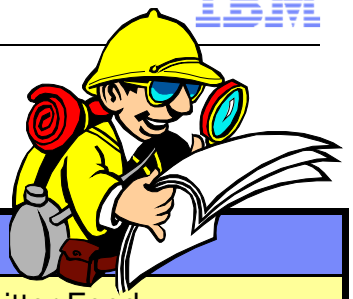
Tool

The IBM z/OS Communications Server Network Utility Assistant

- There is a new tool available for download from the z/OS Communications Server web pages:
  - http://www-01.ibm.com/support/docview.wss?uid=swg24029203
- The IBM z/OS Communications Server Network Utility Assistant tool is a TSO/ISPF front-end to the z/OS Communications Server TSO NETSTAT line-mode command.

```
*--------------------- z/OS V1R12 CS TCP/IP NETSTAT -------------------------*
Command ===>

Select a report option by number or name ==>

     1 ALL          2 ALLConn      3 ARp          4 BYTEInfo     5 CLients
     6 CONFig       7 CONN         8 DEVlinks     9 Gate        10 HOme
    11 PORTList    12 ROUTe       13 SOCKets     14 TELnet      15 UP
    16 CACHinfo    17 SLAP        18 VIPADYn     19 VIPADCFG    20 VCRT
    21 VDPT        22 IDS         23 STATS       24 ND          25 SRCIP
    26 DROP        27 TTLS        28 RESCache    29 DEFADDRT
    90 TN3270      91 CICSsock    92 FTP         93 CICSTS

Enter optional command modifiers and selection filters:

 Do you want to specify optional command modifiers   ==> N  (Y/N)
 Do you want to specify optional selection filters   ==> N  (Y/N)

Enter optional TCP/IP stack name and general options:

 Stack name    ==> TCPCS      Leave blank for default stack
 Interval      ==> 5          Seconds for interval display
 Report format ==> LONG       (Short/Long) Leave blank for stack-default
 Excl. TN3270  ==> N          (Y/N) Reply Y to exclude TN3270 connections
 Netstat debug ==> N          (Y/N) Reply Y to see debug messages from Netstat
 EZANS   debug ==> N          (Y/N) Reply Y to see debug messages from EZANS

Enter required arguments for ARP and DROP commands:

 ARP address  ==> ALL              ARP (specify an IPv4 address or ALL)
 Conn id      ==>                  DROP (Specify connection ID to drop)
```

For more information



URL	Content
http://www.twitter.com/IBM_Commserver	IBM Communications Server Twitter Feed
http://www.facebook.com/IBMCommserver	IBM Communications Server Facebook Fan Page
http://www.ibm.com/systems/z/	IBM System z in general
http://www.ibm.com/systems/z/hardware/networking/	IBM Mainframe System z networking
http://www.ibm.com/software/network/commserver/	IBM Software Communications Server products
http://www.ibm.com/software/network/commserver/zos/	IBM z/OS Communications Server
http://www.ibm.com/software/network/commserver/z_lin/	IBM Communications Server for Linux on System z
http://www.ibm.com/software/network/ccl/	IBM Communication Controller for Linux on System z
http://www.ibm.com/software/network/commserver/library/	IBM Communications Server library
http://www.redbooks.ibm.com	ITSO Redbooks
http://www.ibm.com/software/network/commserver/zos/support/	IBM z/OS Communications Server technical Support – including TechNotes from service
http://www.ibm.com/support/techdocs/atsmastr.nsf/Web/TechDocs	Technical support documentation from Washington Systems Center (techdocs, flashes, presentations, white papers, etc.)
http://www.rfc-editor.org/rfcsearch.html	Request For Comments (RFC)
http://www.ibm.com/systems/z/os/zos/bkserv/	IBM z/OS Internet library – PDF files of all z/OS manuals including Communications Server

For pleasant reading ….