SmarterComputing

Introduction to z/OS Communications Server

Sam Reynolds
IBM Enterprise Networking Solutions
samr@us.ibm.com
August 5, 2014

SHARE 2014 Summer Technical Conference
Session 16216

Trademarks

The following are trademarks of the International Business Machines Corporation in the United States and/or other countries.

AIX*	DB2*	HiperSockets*	MQSeries*	PowerHA*	RMF	System z*	zEnterprise*	z/VM*
BladeCenter*	DFSMS	HyperSwap	NetView*	PR/SM	Smarter Planet*	System z10*	z10	z/VSE*
CICS*	EASY Tier	IMS	OMEGAMON*	PureSystems	Storwize*	Tivoli*	z10 EC	
Cognos*	FICON*	InfiniBand*	Parallel Sysplex*	Rational*	System Storage*	WebSphere*	z/OS*	
DataPower*	GDPS*	Lotus*	POWER7*	RACF*	System x*	XIV*		

* Registered trademarks of IBM Corporation

The following are trademarks or registered trademarks of other companies.

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 therefrom.

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.

IT Infrastructure Library is a registered trademark of the Central Computer and Telecommunications Agency which is now part of the Office of Government Commerce.

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.

Java and all Java based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.

Linear Tape-Open, LTO, the LTO Logo, Ultrium, and the Ultrium logo are trademarks of HP, IBM Corp. and Quantum in the U.S. and

Linux is a registered trademark of Linus Torvalds 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.

OpenStack is a trademark of OpenStack LLC. The OpenStack trademark policy is available on the OpenStack website.

TEALEAF is a registered trademark of Tealeaf, an IBM Company.

Windows Server and the Windows logo are trademarks of the Microsoft group of countries.

Worklight is a trademark or registered trademark of Worklight, an IBM Company.

UNIX is a registered trademark of The Open Group in the United States and other countries.

* Other product and service names might be trademarks of IBM or other companies.

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.

This information provides only general descriptions of the types and portions of workloads that are eligible for execution on Specialty Engines (e.g, zIIPs, zAAPs, and IFLs) ("SEs"). IBM authorizes customers to use IBM SE only to execute the processing of Eligible Workloads of specific Programs expressly authorized by IBM as specified in the "Authorized Use Table for IBM Machines" provided at www.ibm.com/systems/support/machine_warranties/machine_code/aut.html ("AUT"). No other workload processing is authorized for execution on an SE. IBM offers SE at a lower price than General Processors/Central Processors because customers are authorized to use SEs only to process certain types and/or amounts of workloads as specified by IBM in the AUT.

Agenda

- What is Communications Server?
- Overview of SNA
- Overview of TCP/IP
- Communications Server Product Structure
- Communications Server Technology and Applications
  - Network connectivity
  - Clustering
  - Security
- Appendix

What is Communications Server?

What is Communications Server?

- The Communications Server name is used for products running on different platforms:
  - System z and zSeries, pSeries, xSeries, or OEM
  - Communications Server is currently provided for:
    - z/OS, Windows, AIX, UNIX, Linux
  - With z/OS, Communications Server is a base element (part of the operating system)
    - Consists of TCP/IP and SNA Services
    - Consists of prior products, TCP/IP and VTAM, for MVS

z/OS Communications Server History

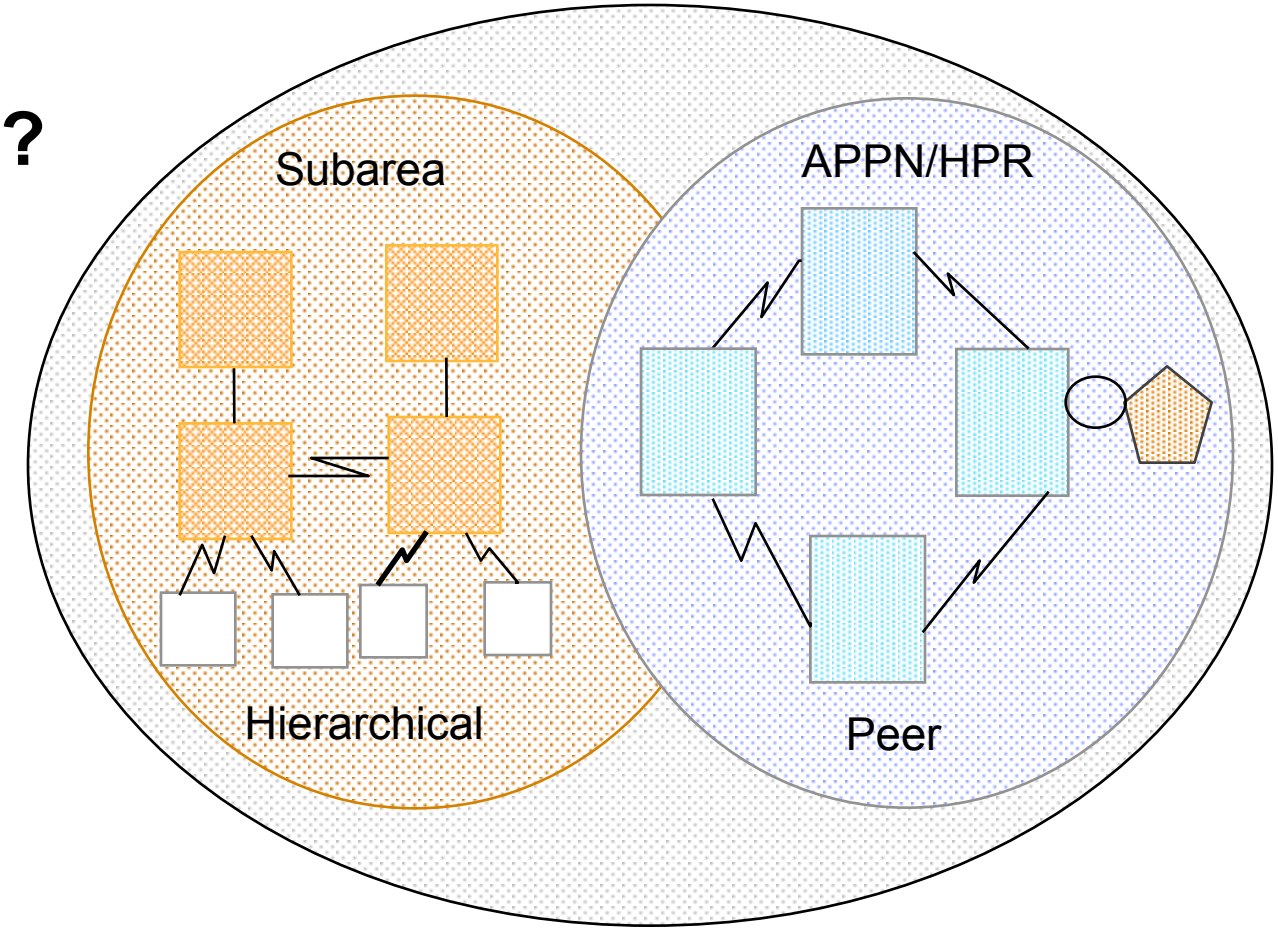
- VTAM ("Virtual Telecommunications Access Method") implements SNA and has been available as mainframe software since 1974
  - Has been continuously enhanced
- TCP/IP available as mainframe software since the 1980s
  - Has been continuously enhanced
  - Original MVS software was ported from VM
- TCP/IP and VTAM were combined into single product (Communications Server) in 1996
  - Has been continuously enhanced

Overview of SNA

SNA History

- SNA "Systems Network Architecture" developed and distributed by IBM and announced in 1974
  - Specifications have also been provided so that a large number of other vendors also provide products that implement SNA
  - SNA is an architecture defining protocols such as:
    - Link Protocols
    - Node Intercommunication Protocols
    - Application Protocols
    - VTAM is the mainframe and NCP is the front-end SNA product
  - Runs with three main operating systems
    - MVS, VM, VSE

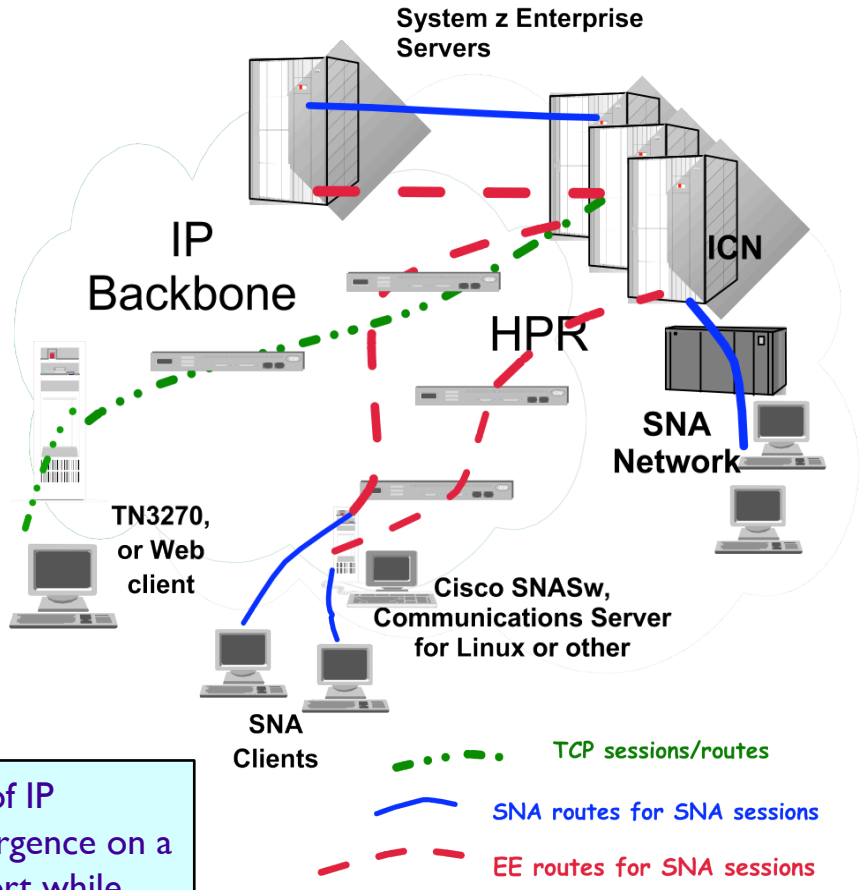
What is SNA?



- SNA originally consisted of strictly hierarchical subarea protocols with few dynamics
  - network resources predefined
- Advanced Peer to Peer networking introduced mid 1980s
  - made SNA more flat and dynamic
- High Performance Routing introduced in 1990s
- Enterprise Extender (HPR over UDP) introduced in 1999

Enterprise Extender

- Allows use of IP network for SNA sessions
  - To an SNA application, the IP network looks like a link
  - To the IP network, EE is just a UDP application
- SNA traffic is sent as UDP datagrams (on ports 12000-12004) over the IP network
- Supported by many platforms:
  - z/OS
  - CS/Linux, CS/AIX, CS/Windows
  - PComm
  - i5/OS
  - Cisco SNASw
  - Microsoft HIS



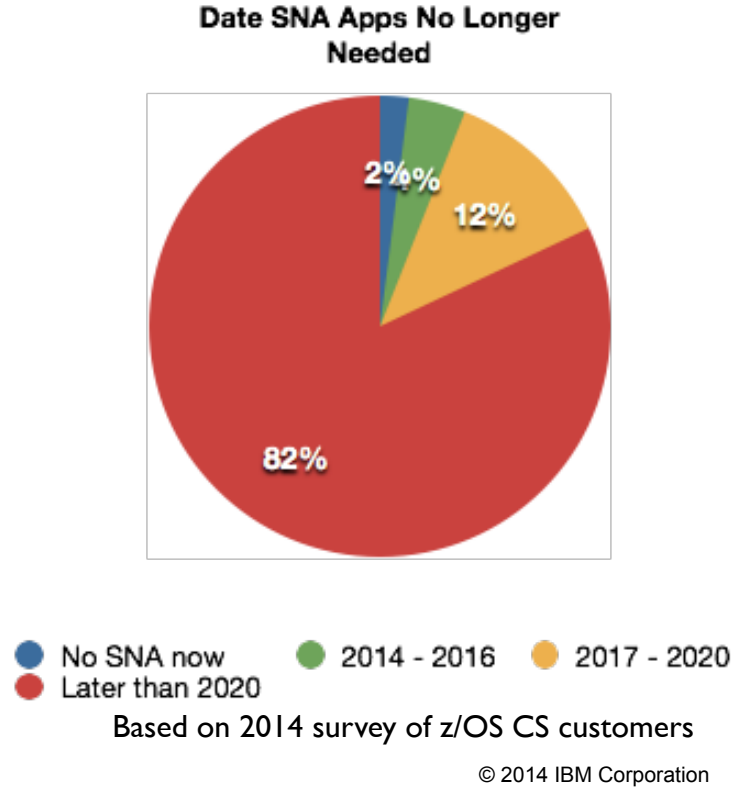
EE allows enablement of IP applications and convergence on a single network transport while preserving SNA application and endpoint investment.

15513: Enterprise Extender on z/OS Communications Server: SNA Hints and Tips
Monday, August 4, 2014: 3:00 PM-4:00 PM
Room 405 (David L. Lawrence Convention Center)
Speaker: Sam Reynolds (IBM Corporation)

State of SNA

"The report of my death is an exaggeration"
Mark Twain, 1897
SNA, 2014

- SNA is a legacy networking protocol but it's still in wide use and support is still required.
  - There are over a trillion lines of customer written application code based on CICS, IMS, and DB2
  - A large percentage of all business data is still accessed using SNA applications



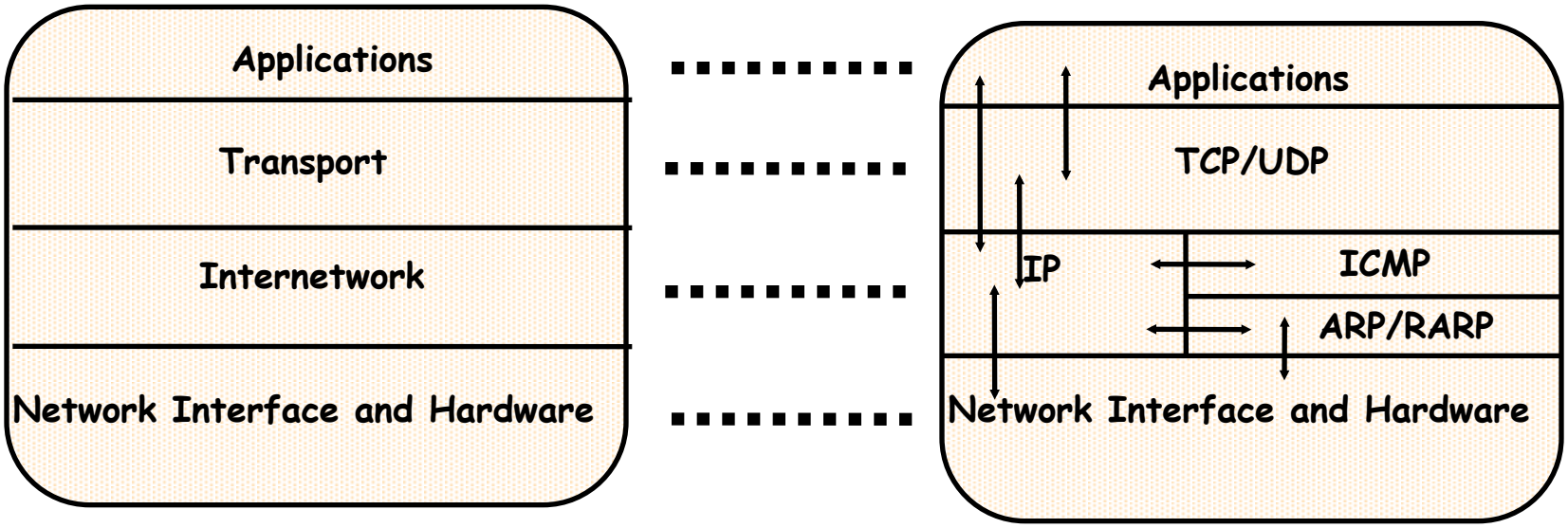
Based on 2014 survey of z/OS CS customers

Overview of TCP/IP

TCP/IP History

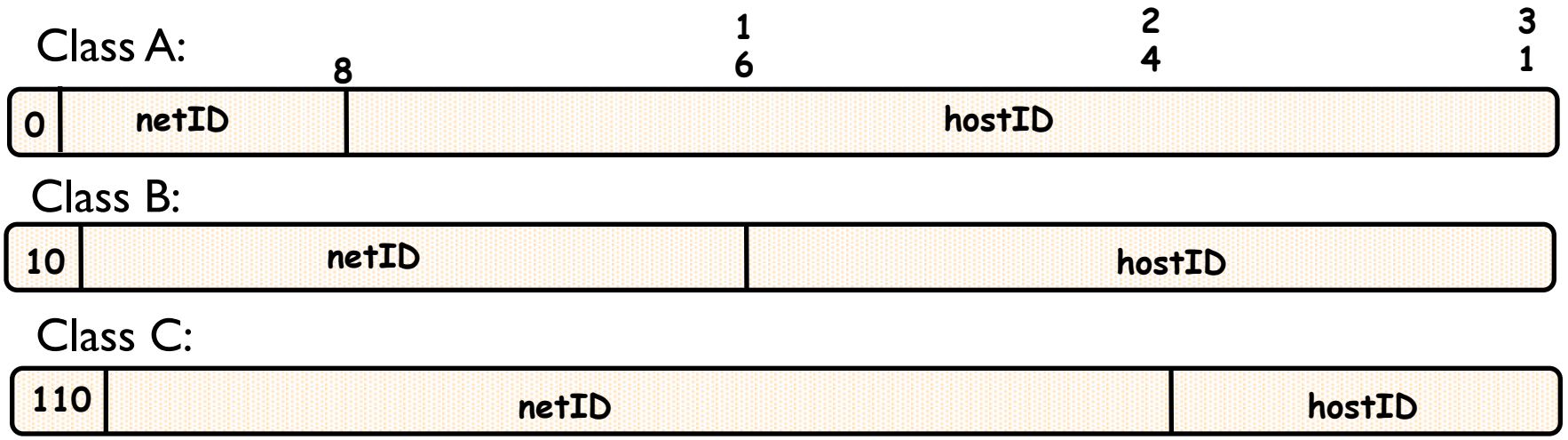
- TCP/IP really needs no introduction, but for reference…
- ARPANET procotol suite introduced concepts of layering and virtualizing in the world of networking in early 1970s
  - Funded by Defense Advanced Research Projects Agency (DARPA)
  - Later Bolt, Beranek, and Newman (BBN) developed TCP/IP protocols for Berkeley UNIX on the VAX
- Code distributed for free with University of California's Berkeley's UNIX operating system
- First release of Berkeley Software Distribution in 1983 (4.2BSD)
- Spread rapidly and several releases followed
- New WANs in US created and attached to ARPANET
  - Developed into the Internet

TCP/IP Overview



- Application Layer
  - User process cooperating with partner on same or different host, i.e. Telnet, FTP, SMTP
- Transport Layer
  - TCP (connection oriented) and UDP (connectionless)
- Internetwork Layer
  - Shields higher levels from the physical network architecture
- Network Interface Layer
  - Interface to network hardware, called link or data link layer

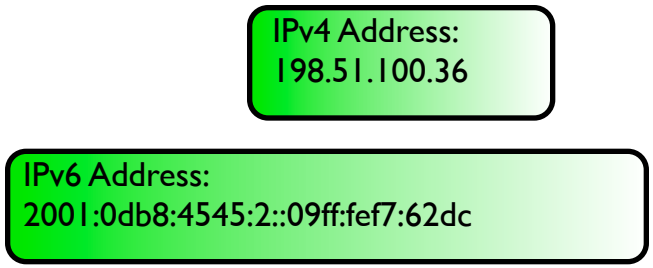
TCP/IP Addressing



- TCP/IP uses four octet addressing structure (IPv4) e.g., 198.51.100.36
  - Broken into netID and hostID portion
  - Subnetting can be used to indicate portion of host bits are to be used to indicate additional networks
- Interface between applications and transport layer is defined by port numbers and sockets
  - Each process that wants to communicate with another process identifies itself by a 16 bit port number

What is IPv6?

- IPv6 is an evolution of the "current" version of IP, which is known as IPv4



- Changes from IPv4 to IPv6
  - Expanded Routing and Addressing
    - Address space increased from 32 bits to 128 bits
    - 340 billion billion billion billion addresses (v4 = 1 billion)
    - Enough to allow every atom on the Earth to have its own IP address – 100 times over!
- Header Format Simplification
  - Reduced common-case processing cost of packet handling
  - Improved Support for Options
- New encoding allows for more efficient forwarding
- Greater flexibility for introducing new options in the future
  - Plug-and-Play Support
- Autoconfiguration of host address, default routers, MTU size, and other IP-related information

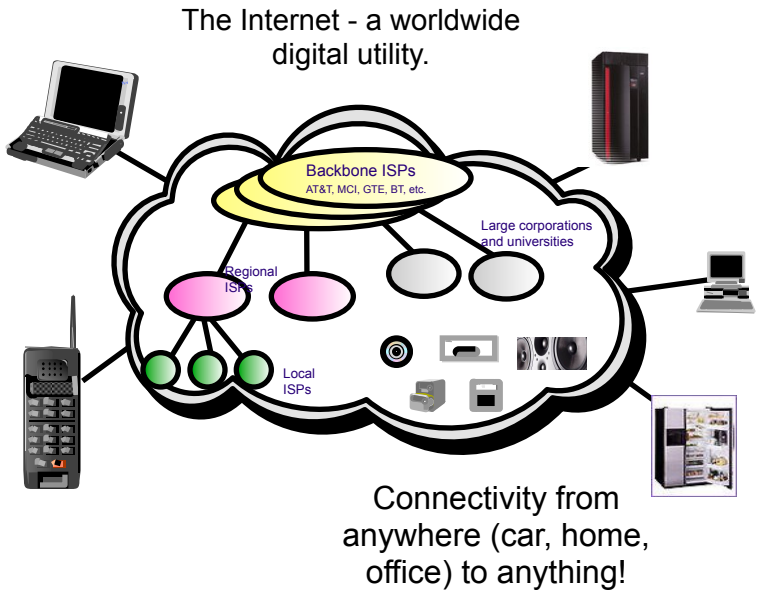
15407: IPv6: The Basics
Thursday, August 7, 2014: 11:15 AM-12:15 PM
Room 316 (David L. Lawrence Convention Center)
Speaker: Laura Knapp (AES)

Why IPv6 now?

- The effects of the IPv4 address limitations have been reduced/postponed for a while using Network Address Translation

BUT –

- A new surge of demand for IP addresses is coming from:
  - The far east where the use of the Internet is accelerating fast
  - The assignment of IP addresses to cell phones
  - The assignment of IP addresses to household appliances of various types "the internet of things"
  - Multiple addresses per person
  - Government mandates for IPv6 support

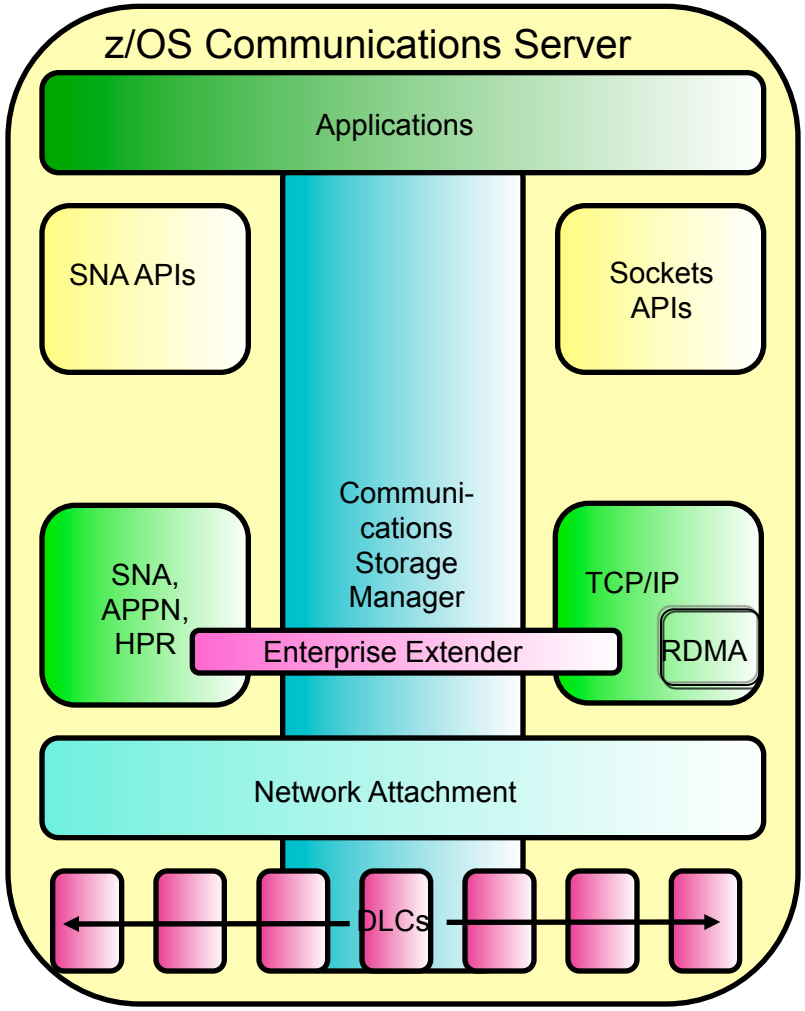


- This demand cannot be met with IPv4 addressing space and IPv6 is now required for parts of the internet
  - IPv6 support is now required to bid on IT contracts for the US government!

Communications Server Product Structure

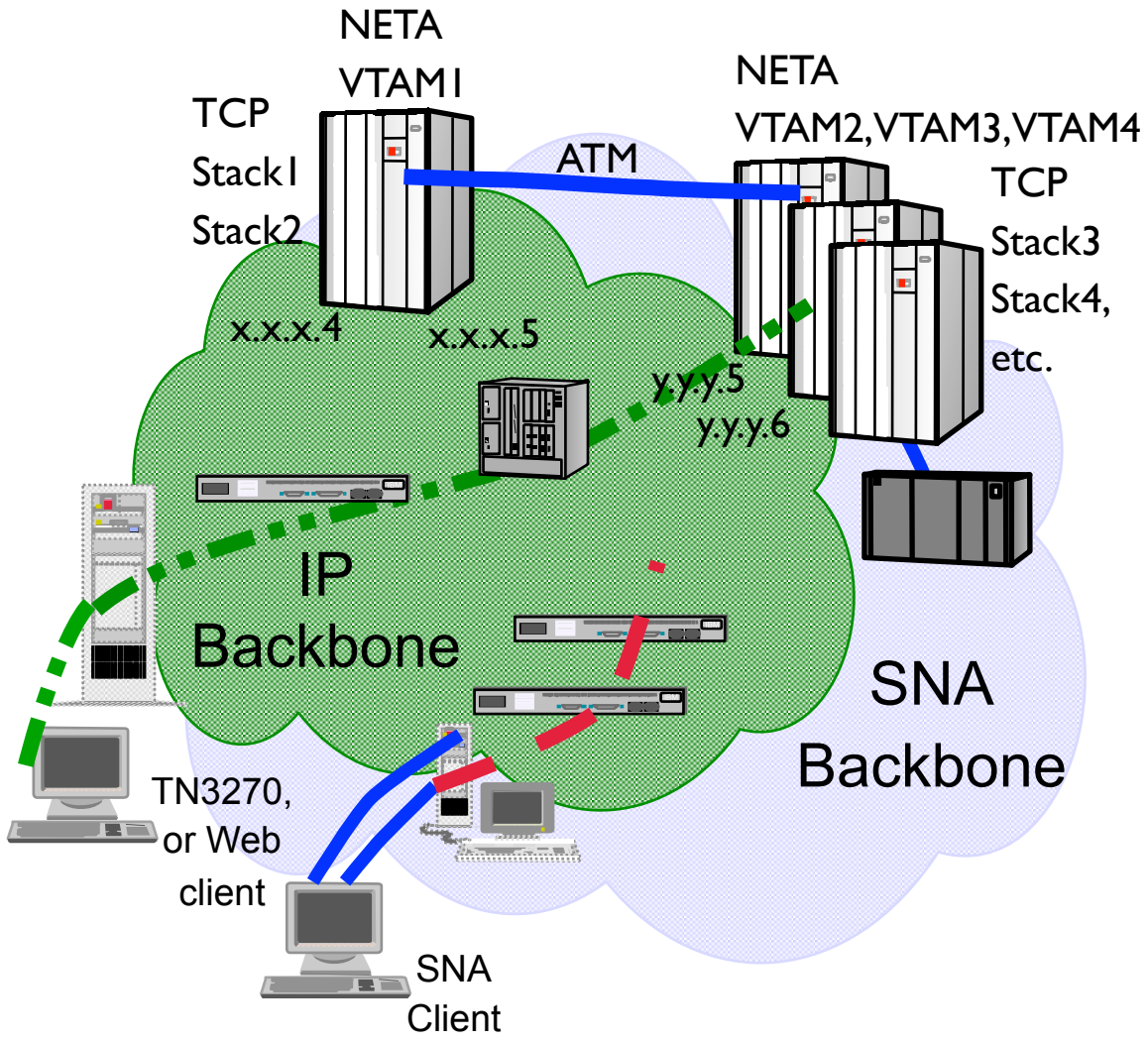
z/OS Communications Server

- Provide common services within CS
  - Network attachment
  - Storage management
- TCP/IP and SNA integration
  - TN3270
  - Network access
  - Internal optimizations
  - Enterprise Extender
- Standard TCP/IP applications
- Multi-protocol Solutions
- Sockets (TCP/IP) applications
- z/OS Unix offers z/OS users access to a wide range of UNIX-based applications over IP networks
- SNA applications
  - SNA applications are supported over native SNA or IP networks

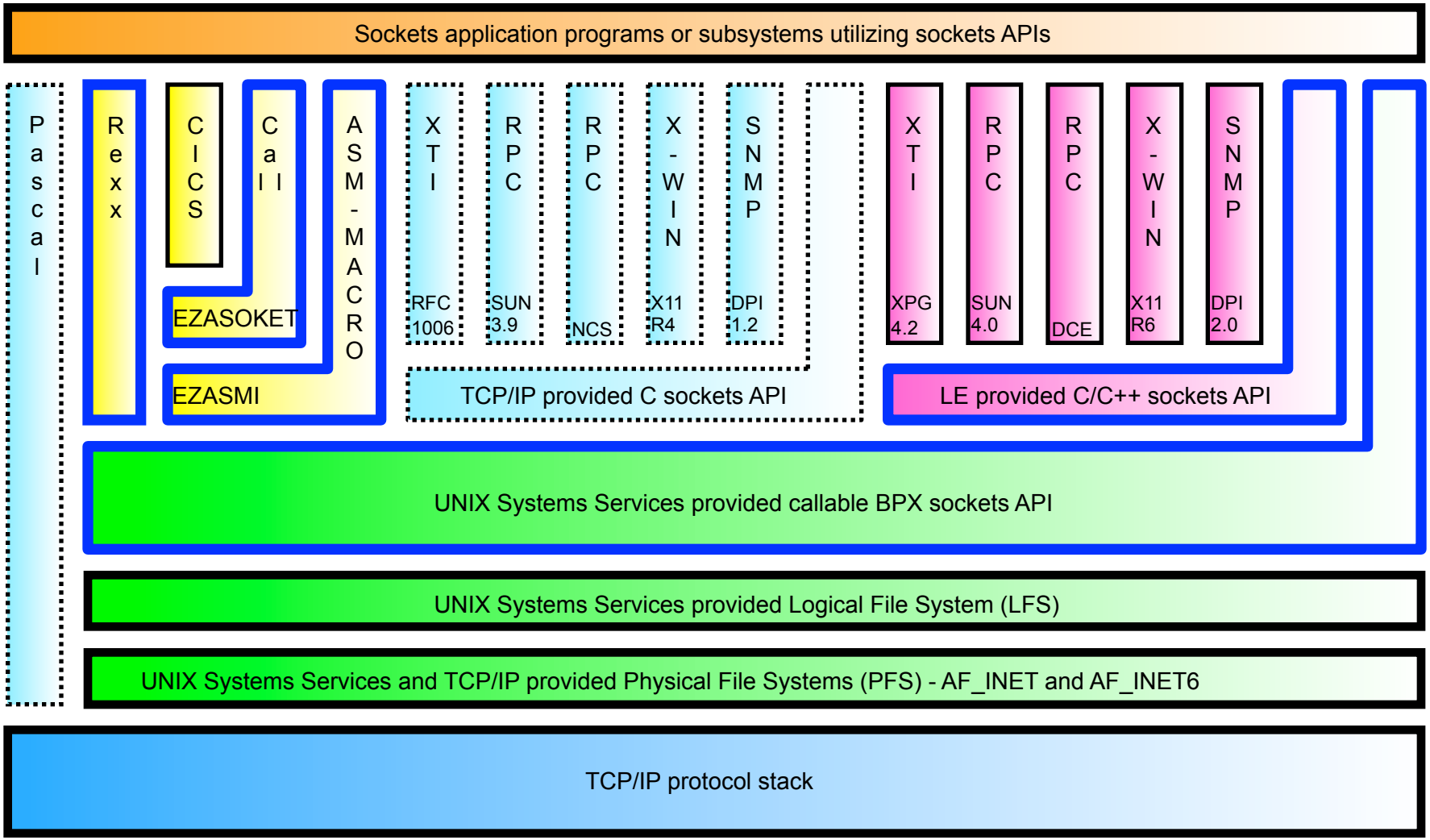


LPARs and z/OS Communications Server

- Only one VTAM per z/OS image
- May run multiple TCP/IP stacks on a single z/OS image (but generally not recommended)
- May use different physical networks for SNA and TCP/IP, or may use a single network for both



Sockets APIs? Yep, we've got a few



Some standardized application protocols

Mostly provided by CS as applications that ship with the product

Routing
OMPRoute

Remote terminal access

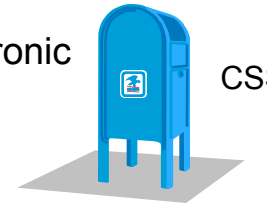
Telnet
TN3270(e)
TN5250

REXEC
RSH

Remote
Command
Execution

TCP/IP
Network

Electronic
mail



CSSMTP

NFS
DFS

Data
Sharing

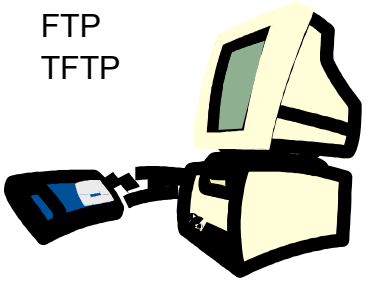
SNMP
ICMP

Network Management

HTTP
HTTPS
MIME

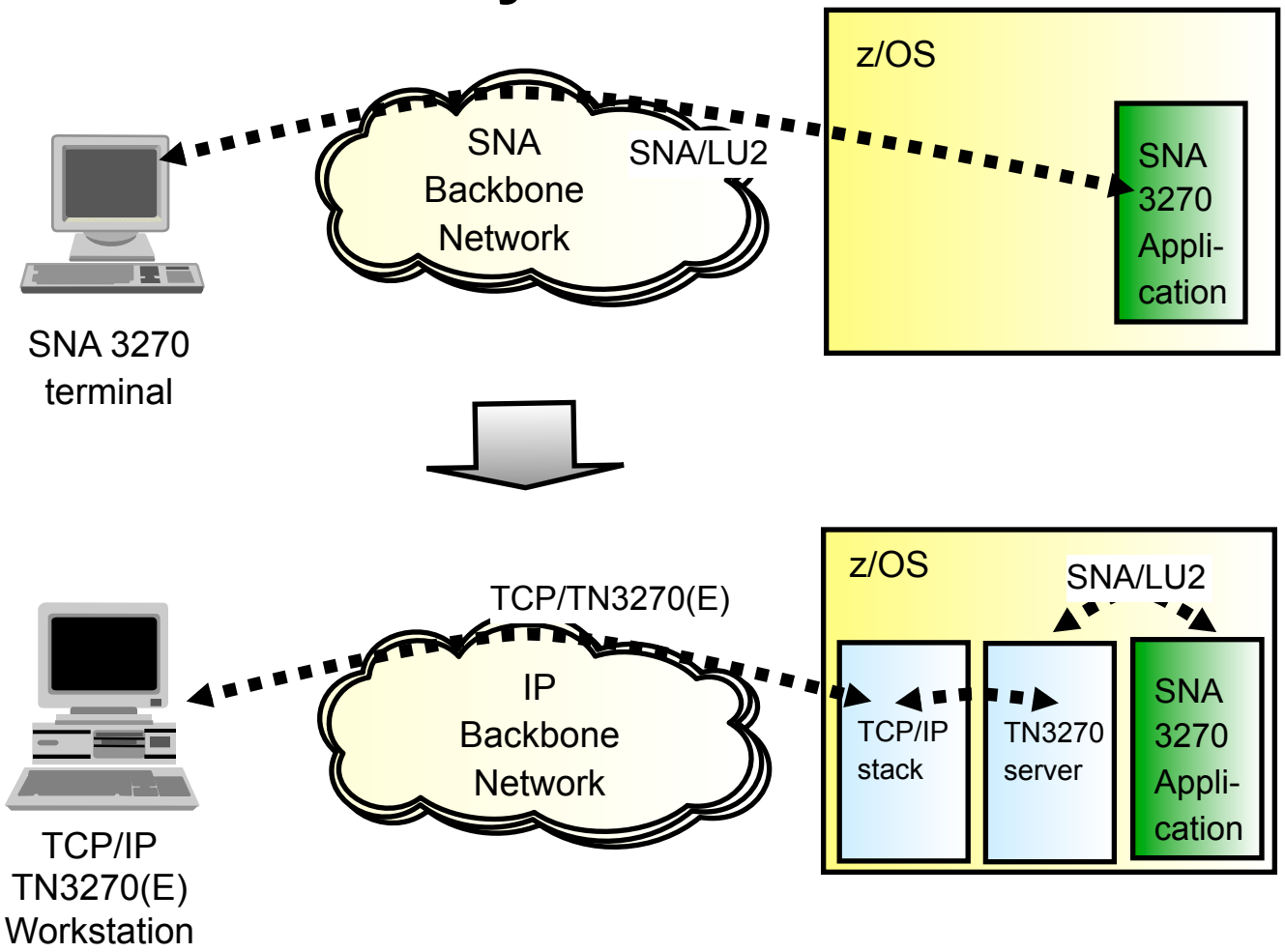
World Wide Web

FTP
TFTP



File
Transfer

Why is TN3270 Needed?



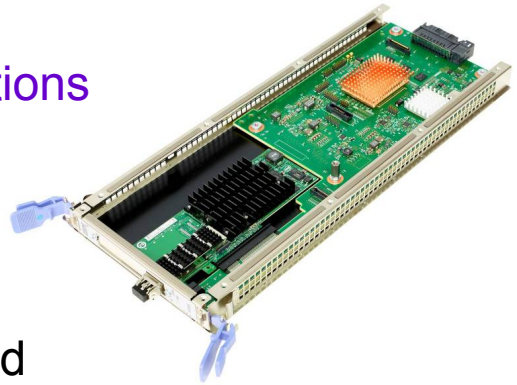
- There are hundreds of thousands of SNA LU2-based applications out there today that will probably never be migrated to native sockets.

- Many OEM and IBM client products available
- IBM products include PCOM and Rational Host On Demand

Communications Server Technology and Applications

Communications Server Network Connectivity Technologies

Open Systems Adapter (OSA)

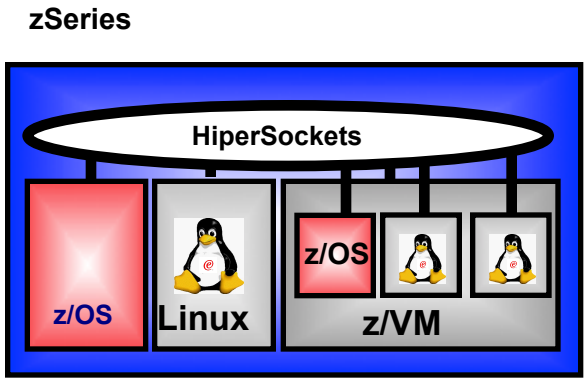


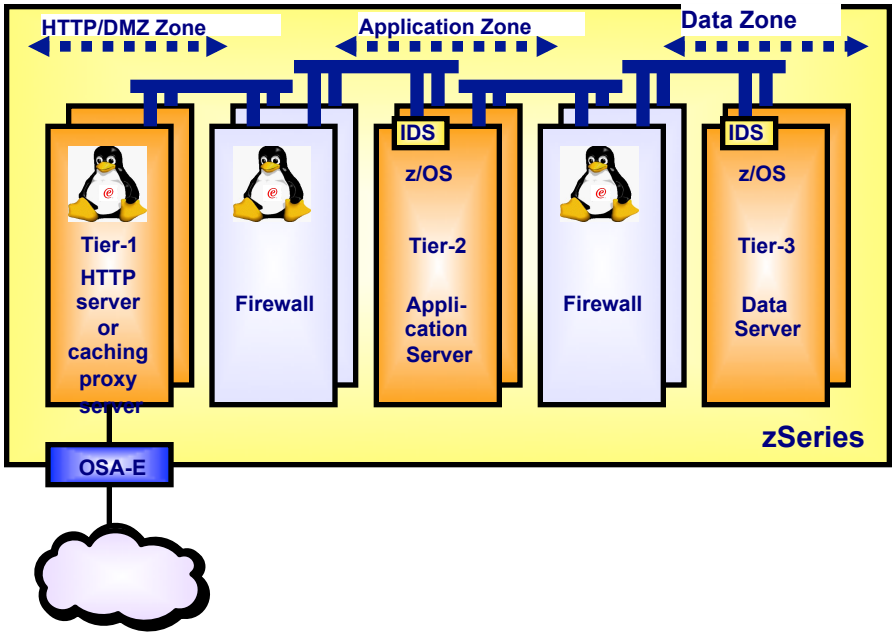
- The Open Systems Adapter is the strategic communications device for the mainframe architecture.
  - OSA integrates several hardware features and supports many networking transport protocols.
  - Powerful virtualization capabilities allow the OSA-Express adapter to be shared by all LPARs and stacks on a CEC, providing economical scalability
    - zLinux, z/VM, and z/OS LPARS can all share an OSA adapter
  - The OSA-Express adapter includes the NIC and the network controller
  - The OSA adapter can communicate with the attached host using a technology called "Queued Direct I/O" or QDIO.
  - The OSA-Express adapter (also called a feature) is a powerful network controller that can offload many processing-intensive networking tasks from the stack and the general purpose CPU
    - Offloading these tasks reduces z/OS CPU cost, increases efficiency, and can improve availability

From a previous SHARE conference:
"Getting the Most Out of Your OSA Adapter with z/OS CS"
http://www.share.org/p/do/sd/sid=9497&type=0

System z HiperSockets – The IP network within the box

- Hipersockets provides IP networking among virtual servers within a zSeries CEC
  - Up to 16 internal IP networks that interconnect z/OS, Linux on zSeries, VSE/ESA, and z/VM including z/VM guest operating systems
  - Improved response time due to very low latency
  - Highly secure (no external cables to tap into)
  - Highly available
  - Cost savings (no external switch equipment needed)
  - Flexible
  - Simple to install, operate, maintain
- A Hipersockets network looks like an internal LAN
- Hipersockets is sometimes referred to as “internal QDIO” or IQDIO





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

Shared Memory communications over RDMA (SMC-R)

- Shared Memory Communications over RDMA (SMC-R) is a protocol that allows *TCP sockets* applications to **transparently and autonomically** exploit RDMA over your datacenter Ethernet network using RoCE.
- SMC-R is a "hybrid" solution that:
  - Uses TCP connection to establish SMC-R connection
  - Each TCP end point exchanges TCP options to dynamically learn whether the connection is eligible for the SMC-R protocol
    - *If eligible, RDMA attributes are then exchanged within the TCP data stream*
  - Socket application data is exchanged via RDMA write operations
  - TCP connection remains active and controls SMC-R connection
  - This model preserves many critical operational and network management features of TCP/IP with minimal IP admin changes
    - Resilience/failover
    - Load balancing
    - Connection based security (IP filters, policy, VLAN, SSL)

SMC-R and 10GbE RoCE Express Requirements

- **Operating system requirements**
  - Requires z/OS 2.1 (GA 9/30/13) which supports the SMC-R protocol
- **Server requirements**
  - Exclusive to zEC12 (with Driver 15E) and zBC12
  - New 10 GbE RoCE Express feature for PCIe I/O drawer (FC#0411)
    - Single port enabled for use by SMC-R
    - Each feature must be dedicated to one LPAR
  - Recommended minimum configuration two features per LPAR for redundancy
    - Up to 16 features supported
  - OSA Express – either 1 GbE or 10 GbE
    - Configured in QDIO mode (OSD CHPIDs only, not OSX)
    - Does not need to be dedicated to the LPAR
  - Standard 10GbE Switch or point to point configuration supported
    - Does not need to be CEE capable
    - Switch must support and have enabled Global pause frame (a standard Ethernet switch feature for Ethernet flow control described in the IEEE 802.3x standard)

Shared Memory communications over RDMA (SMC-R) References

15508: z/OS V2R1 Communications Server: Shared Memory Communications (SMC-R), Part 1 of 2
Tuesday, August 5, 2014: 11:15 AM-12:15 PM
Room 316 (David L. Lawrence Convention Center)
Speaker: Gus Kassimis (IBM Corporation)

15509: z/OS V2R1 Communications Server: Shared Memory Communications (SMC-R), Part 2 of 2
Tuesday, August 5, 2014: 1:30 PM-2:30 PM
Room 316 (David L. Lawrence Convention Center)
Speaker: Dave Herr (IBM Corporation)

YouTube:
"Shared Memory Communications Over RDMA (SMC-R) – Overview"
http://youtu.be/8_5JviApQXw

YouTube:
"Shared Memory Communications Over RDMA (SMC-R) – Implementation"
http://youtu.be/TN0eS-I1FoE

z/OS Communications Server Web Page:
Shared Memory Communications Over RDMA Reference Information
http://www-01.ibm.com/software/network/commserver/SMCR/

Summary: TCP/IP Data Link Controls

- Discussed previously:
  - MPCIPA - to other TCP/IP hosts via QDIO or HiperSockets
  - HiperSockets -- to other TCP/IP (z/OS, Linux, and/or VM) hosts on the same zSeries CEC
  - RoCE – RDMA to other TCP/IP hosts using SMC-R
- Others:
  - CTC - channel to another IBM zSeries
  - MPCPTP - to another zSeries via channel or XCF, to another stack, or to routers/R6K via channel
  - LCS - to other TCP/IP hosts via 3172, 8232, router, OSA, etc.
  - SAMEHOST - to other address spaces or stacks in same LPAR
  - Also: CDLC, CLAW, HYPERCHANNEL, MPCOSA, ATM, X.25, SNALINK (Legacy, not recommended – These are being withdrawn in the next release of z/OS.)

SNA Data Link Controls

- CDLC
  - To 3745/3746 or 3x74 via channel
- LSA
  - To 3172 or router via channel
  - To LAN via OSA2 or OSA-Express (non-QDIO mode)
- CTC
  - To another VTAM via channel using subarea protocols
- Subarea MPC
  - To another VTAM via multiple subchannels using subarea protocols
- MPC
  - To another VTAM or router via channel (APPN)

Most common

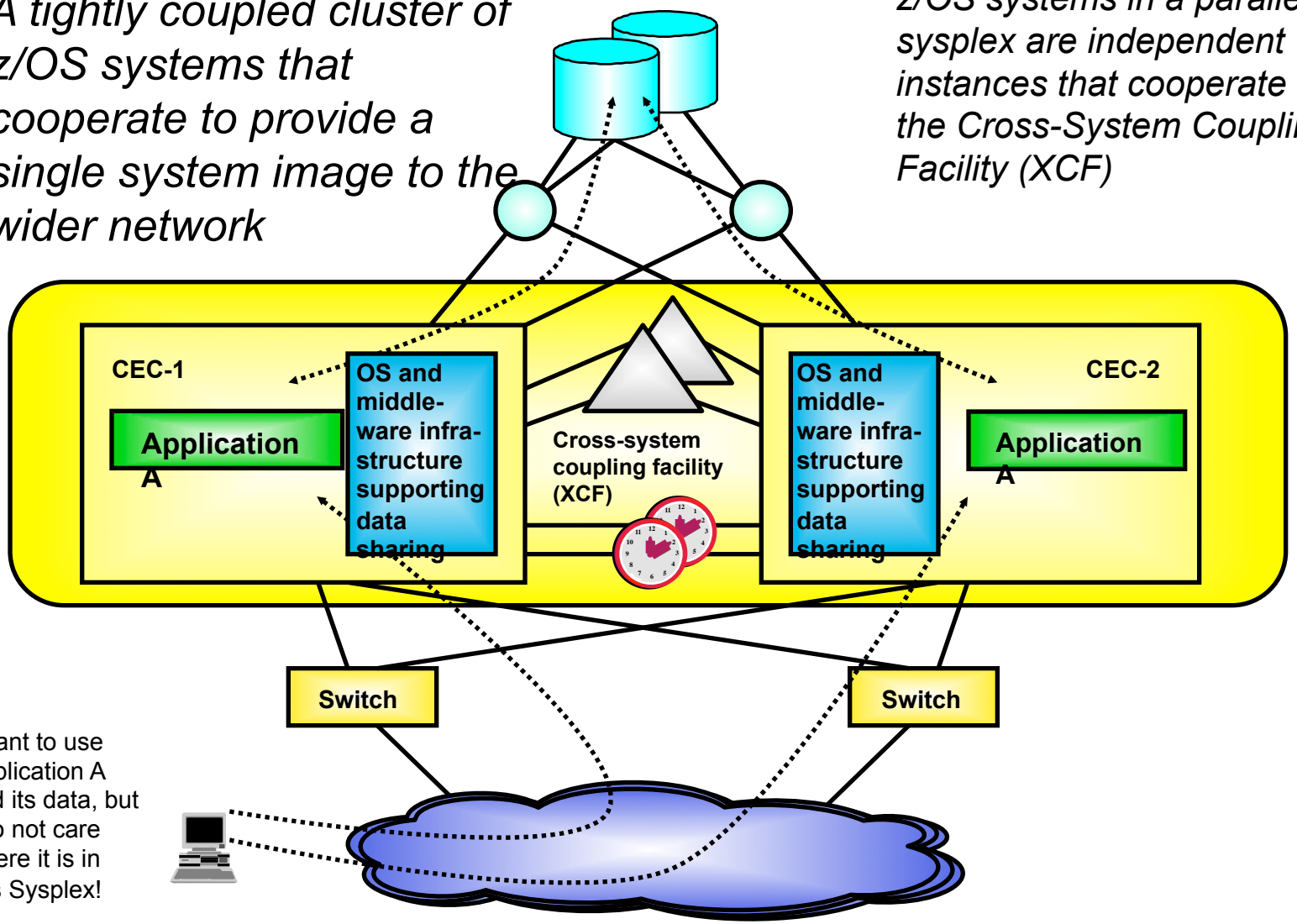
- MPC+
  - To another VTAM via channel or XCF (HPR) or to a router via channel
- EE LDLC
  - To Enterprise Extender partner (using IP network)

Communications Server Clustering Technologies

Basic definition of a parallel sysplex

A tightly coupled cluster of z/OS systems that cooperate to provide a single system image to the wider network

z/OS systems in a parallel sysplex are independent instances that cooperate using the Cross-System Coupling Facility (XCF)



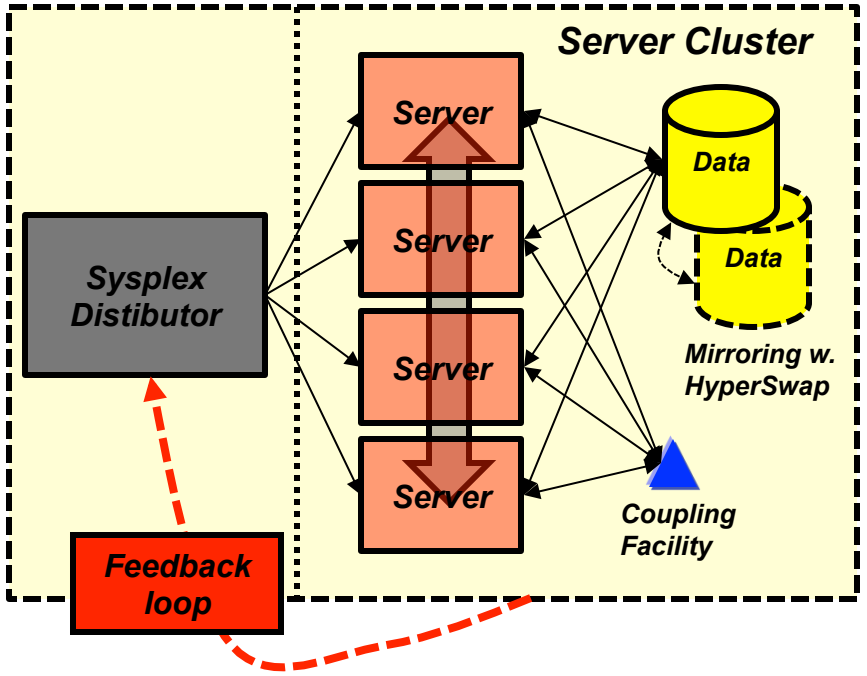
I want to use Application A and its data, but I do not care where it is in this Sysplex!

Sysplex Workload Distribution and Load Balancing

Sysplex Distributor manages load balancing and distribution within a sysplex

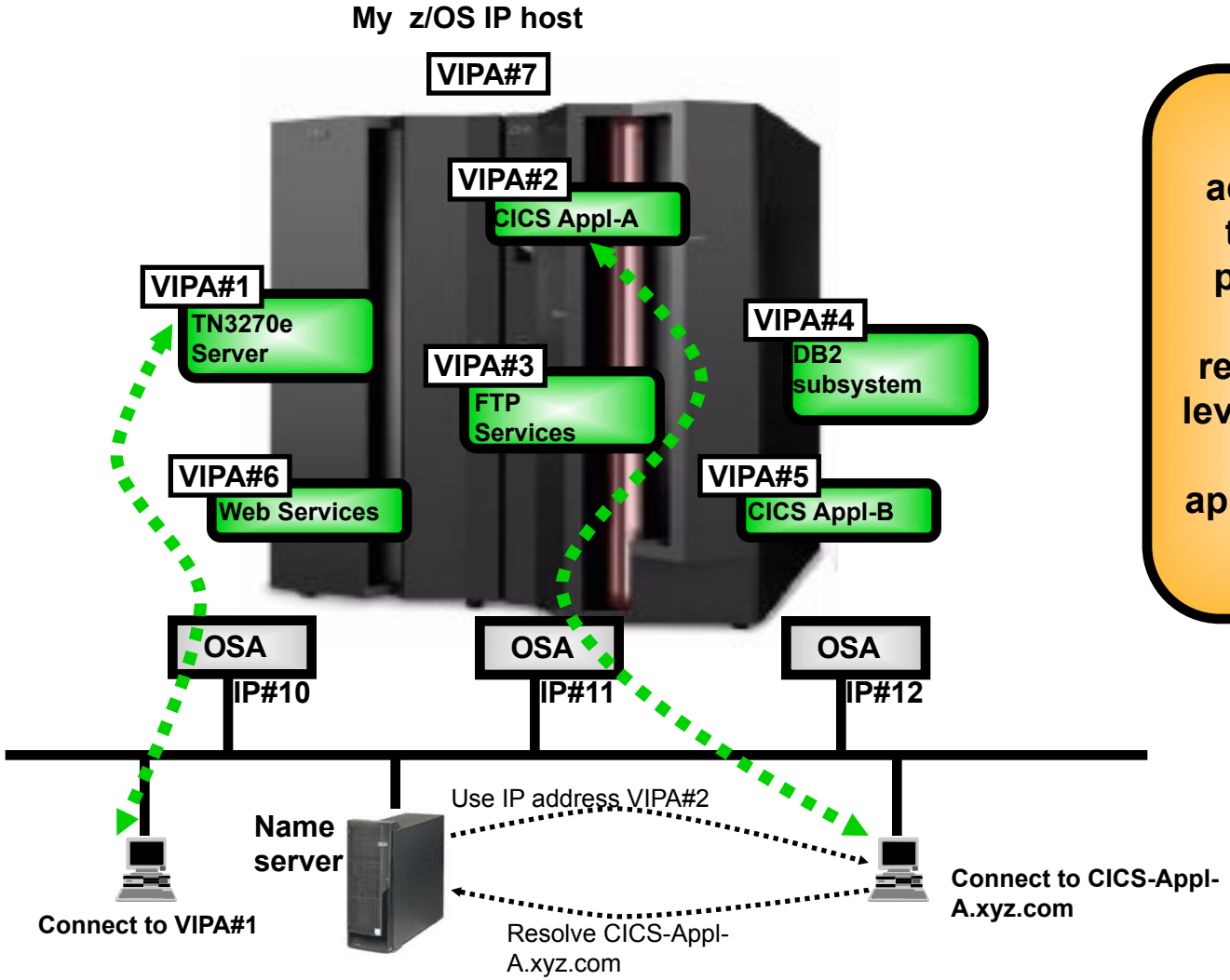
- **Performance**
  - Workload management across a cluster of server instances
  - One server instance on one hardware node may not be sufficient to handle all the workload
- **Availability**
  - As long as one server instance is up-and-running, the "service" is available
  - Individual server instances and associated hardware components may fail without impacting overall availability
- **Capacity management / horizontal growth**
  - Transparently add/remove server instances and/or hardware nodes to/from the pool of servers in the cluster
- **Single System Image**
  - Give users one target hostname to direct requests to
  - Number of and location of server instances is transparent to the user

All server instances must be able to provide the same basic service. In a z/OS Sysplex that means the applications must be Sysplex-enabled and be able to share data across all LPARs in the Sysplex.



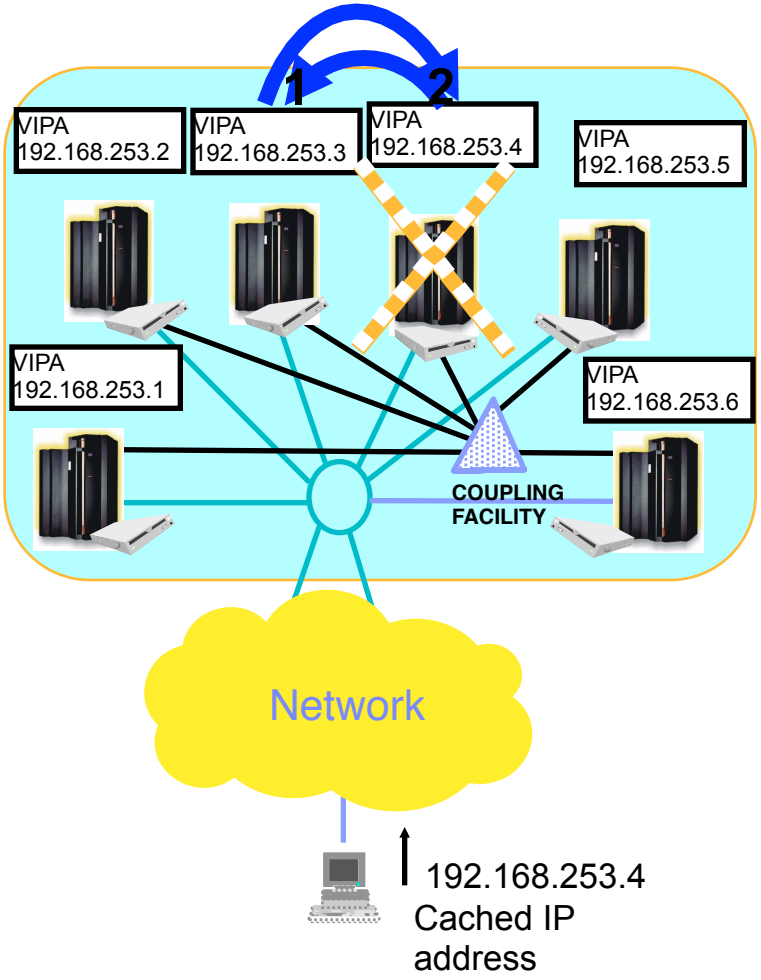
In order for the load balancing decision maker to meet those objectives, it must be capable of obtaining feedback dynamically, such as server instance availability, capacity, performance, and overall health.

Virtual IP Address (VIPA) definition



A VIPA is an IP address that is not tied to a specific physical interface and instead represents a higher level concept such as a host, an application, or a load balancer.

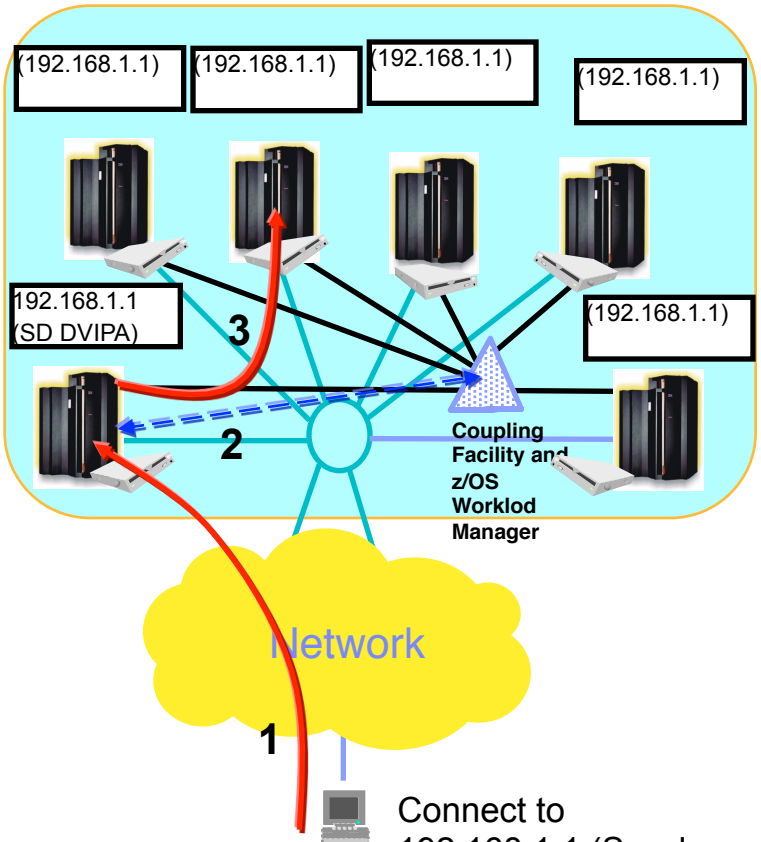
Dynamic VIPA movement – stack managed DVIPAs



Dynamic VIPA Support

- VIPAs can survive any outage by moving to another stack in Sysplex via VIPA Takeover
  - Another application instance can pick up workload or Application can be restarted on takeover stack
  - Connections broken but Reset sent to client upon takeover
    - Significantly reduces down time
- Dynamic VIPA Takeback
  - VIPA moves back to recovered primary owner
    - New Connections Handled By Primary Owner again
    - Connections Established To Backup are allowed to continue
      - ✓Data forwarded from primary owner to backup
    - Allows Movement Of Application Server Without Impacting Existing Workload
- Works in conjunction with dynamic routing
  - OSPF (implemented by the z/OS OMPROUTE daemon) recommended
- Useful for planned outages as well
  - Operator commands allow you to move Dynamic VIPAs non-disruptively

Distributed Dynamic VIPA



A Distributed Dynamic VIPA represents a Sysplex Distributor and the servers it manages

- Only the distributor node advertises the address
- Back-end servers use the address as well but do not advertise it to the network
- Clients connect to the distributor using the distributed DVIPA address and the distributor forwards the connection to the appropriate back-end server.

In this manner, distributed dynamic VIPA provides the single system image for the entire cluster managed by the Sysplex Distributor!

15507: Sysplex Networking Technologies and Considerations
Monday, August 4, 2014: 4:15 PM-5:30 PM
Room 405 (David L. Lawrence Convention Center)
Speaker: Gus Kassimis (IBM Corporation)

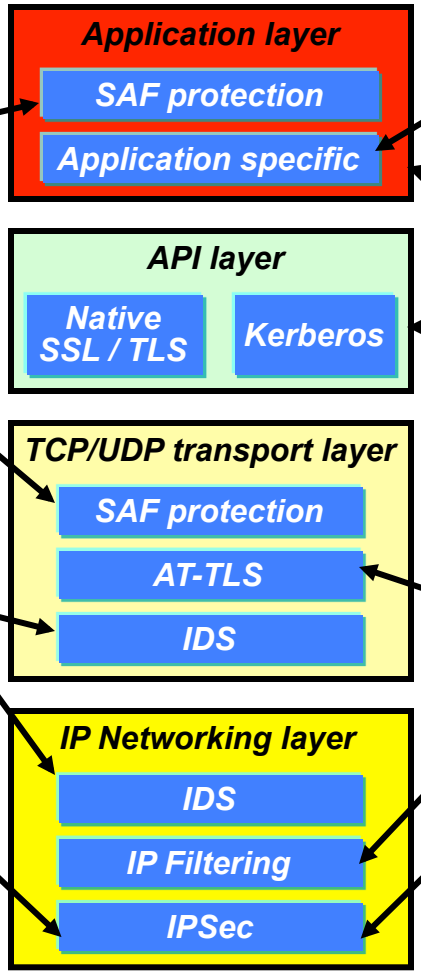
From a previous SHARE conference:
"Sysplex Networking Technology Overview"
http://www.share.org/p/do/sd/sid=527&type=0

Communications Server Security Technologies

Roles and Objectives: Communications Server security technologies

Protect the system

Protect data in the network



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

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

SSH (not part of z/OS CS) provides an umbrella of secure applications (secure shell access, secure file transfer, etc.)

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

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

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

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

IP filtering blocks out all IP traffic that this systems doesn't specifically permit. Can be configured or applied dynamically as "defensive" filters

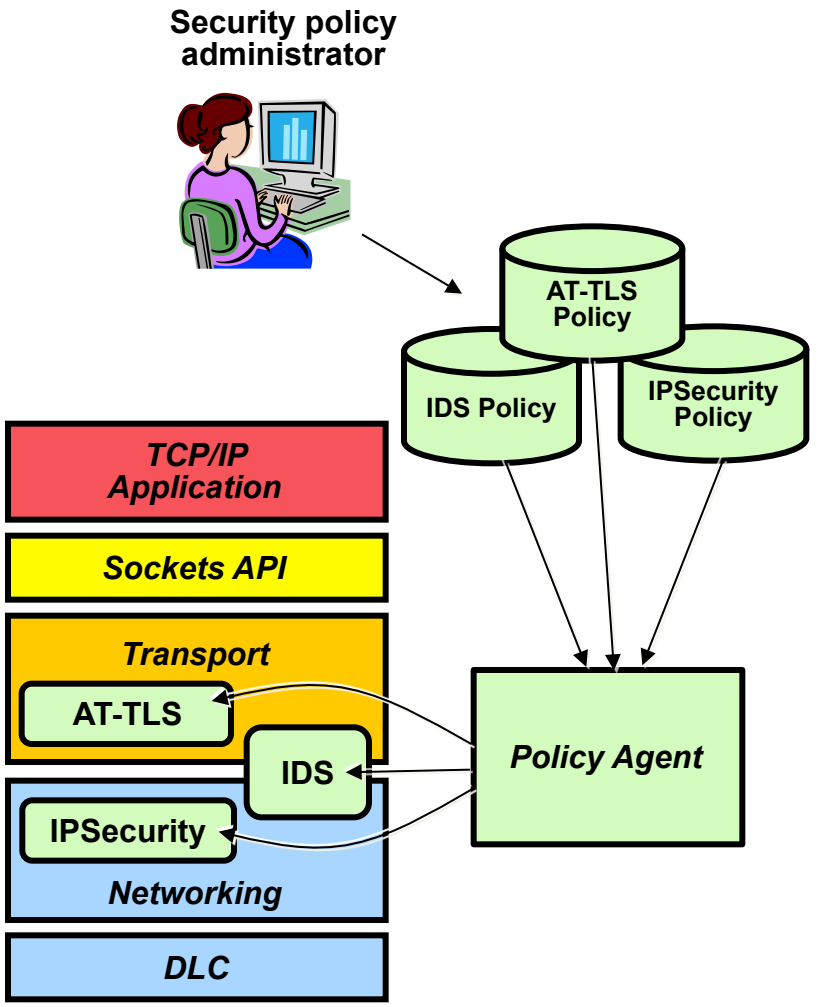
IP filters specify traffic that requires IPsec protection.

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

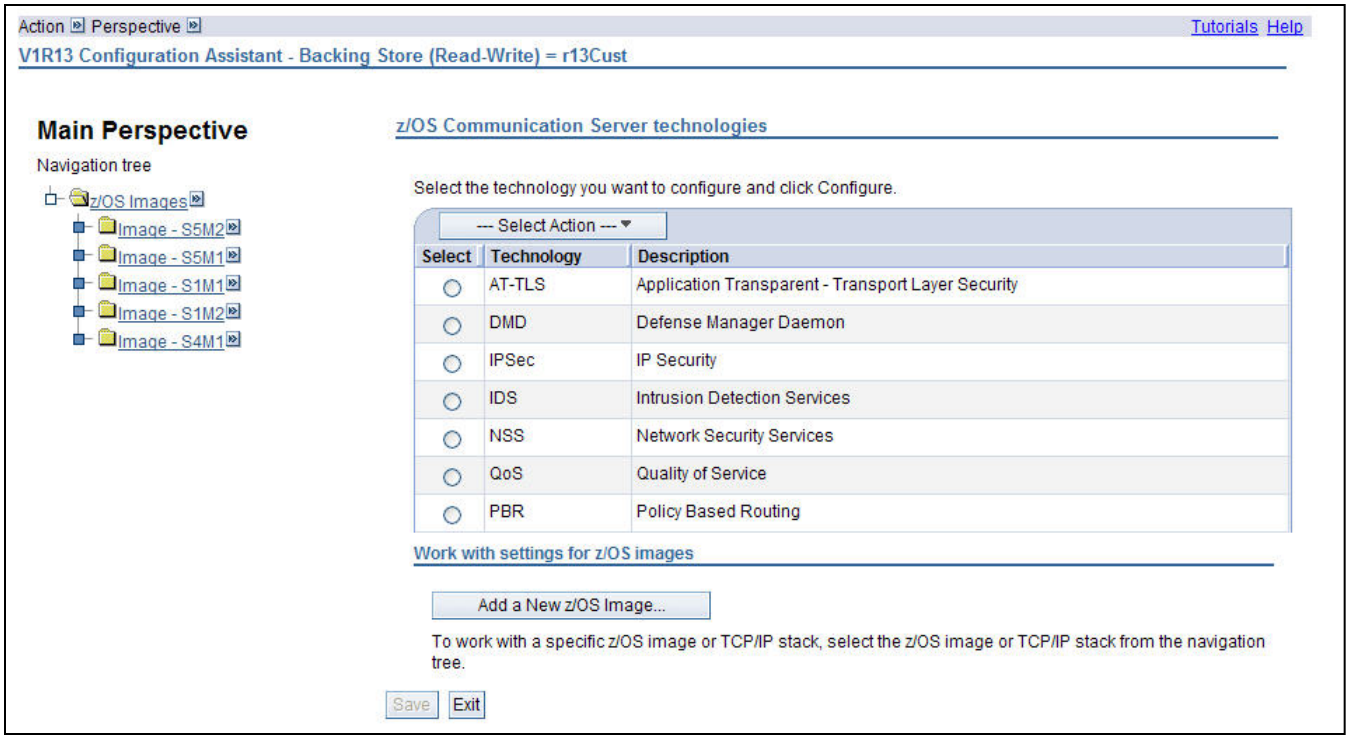
From a previous SHARE conference:
"z/OS Communications Server Network Security Overview"
http://www.share.org/p/do/sd/sid=9179&type=0

Policy-based network security on z/OS: Overview

- Policy is created through Configuration Assistant for z/OS Communications Server
  - GUI-based tool
  - Configures each security discipline (AT-TLS, IP security and IDS) using consistent model
  - Generates and uploads policy files and related content to z/OS
- Policy Agent processes and installs policies into TCP/IP stack
  - Policies are defined per TCP/IP stack
  - Separate policies for each discipline
  - Policy agent also monitors and manages the other daemons and processes needed to enforce the policies (IKED, syslogd, trmd, etc.)
- Provides network security without requiring changes to your applications
  - Security policies are enforced by TCP/IP stack
  - Different security disciplines are enforced independent of each other



Policy-based network security on z/OS: Configuration Assistant



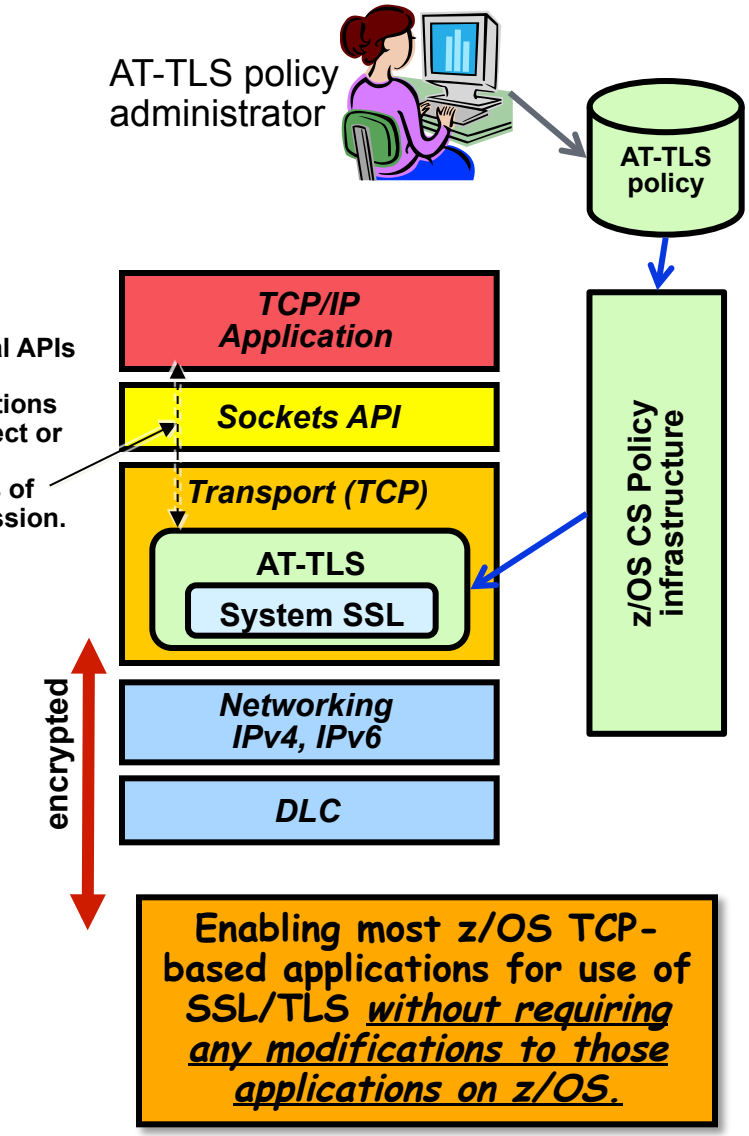
- **z/OSMF-based web interface**
- **Configures all policy disciplines**
- **Separate perspectives but consistent model for each discipline**

- **Focus on concepts, not details**
  - what traffic to protect
  - how to protect it
  - De-emphasize low-level details (though they are accessible through advanced panels)
- **Builds and maintains**
  - Policy files
  - Related configuration files
  - JCL procs and RACF directives

- **Supports import of existing policy files**
- **Supports current z/OS release plus past two**
- **Actively imports certain configuration information**

Application Transparent TLS on z/OS overview

- **Basic TCP/IP stack-based SSL/TLS**
  - SSL/TLS process performed at TCP layer (via System SSL) **without requiring any application change (transparent)**
  - AT-TLS policy specifies which TCP traffic is to be SSL/TLS protected based on a variety of criteria
    - Local address, port
    - Remote address, port
    - Connection direction
    - z/OS userid, jobname
    - Time, day, week, month
- **Application transparency**
  - Can be fully transparent to application
  - Application has option to inspect or control certain aspects of ATTLS processing
- **Available to TCP applications**
- **Supports SSLv2, v3, TLSv1.0, v1.1, v1.2**
- **Uses System SSL for SSL/TLS protocol processing**
  - Remote endpoint sees an RFC-compliant implementation
  - Interoperates with other compliant implementations

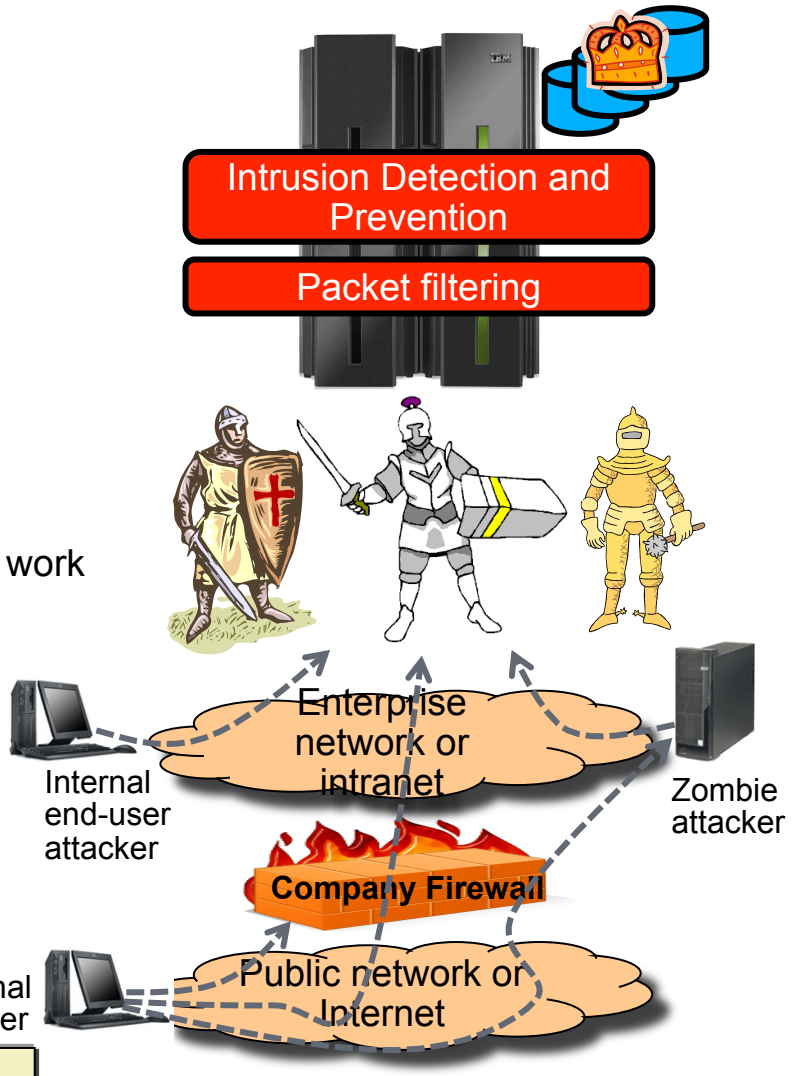


16082: Leveraging z/OS CS AT-TLS for a Lower Cost and More Rapid TLS Deployment
Thursday, August 7, 2014: 12:25 PM-1:15 PM
Room 405 (David L. Lawrence Convention Center)
Speaker: Lin Overby (IBM Corporation)

YouTube:
"AT-TLS in a Nutshell"
http://youtu.be/YKEzX70moOQ

IDS: Protecting against intentional and unintentional attacks on your system

- **What is an intrusion?**
  - Information Gathering
    - Network and system topology
    - Data location and contents
  - Eavesdropping/Impersonation/Theft
    - On the network/on the host
    - Base for further attacks on others through Amplifiers, Robots, or Zombies
  - Denial of Service - Attack on availability
    - Single packet attacks - exploits system or application vulnerability
    - Multi-packet attacks - floods systems to exclude useful work
- **Attacks can occur from Internet or intranet**
  - Company firewalls and intrusion prevention appliances can provide some level of protection from Internet
  - Perimeter security strategy alone usually not sufficient.
    - Some access is permitted from Internet – typically into a Demilitarized Zone (DMZ)
    - Trust of intranet
- **Attacks can be intentional (malicious) but often occur as a result of errors on nodes in the network (config, application, etc.)**



15516: z/OS Communications Server Intrusion Detection Services
Thursday, August 7, 2014: 1:30 PM-2:30 PM
Room 316 (David L. Lawrence Convention Center)
Speaker: Lin Overby (IBM Corporation)

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
http://www.ibm.com/systems/z/hardware/networking/index.html	IBM System z Networking
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 zSeries
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 - white papers, product documentation, etc.
http://www.redbooks.ibm.com	IBM Redbooks
http://www.ibm.com/software/network/commserver/support	IBM Communications Server Technical Support
http://www.ibm.com/support/techdocs/	Technical Support Documentation (techdocs, flashes, presentations, white papers, etc.)
http://www.rfc-editor.org/rfcsearch.html	Request For Comments (RFCs)
http://publib.boulder.ibm.com/infocenter/ieduasst/stgv1r0/index.jsp	IBM Education Assistant

YouTube:
"z/OS Communications Server Overview"
http://youtu.be/t55pyd7XuTI

Please complete your session evaluation

- Introduction to z/OS Communications Server
- Session # 16216
- QR Code:

IBM Comm Server

Find us on Facebook at
http://www.facebook.com/IBMCommserver

Follow us on Twitter at
http://www.twitter.com/IBM_Commserver

Read the z/OS Communications Server blog at
http://tinyurl.com/zoscsblog

Visit the z/OS CS YouTube channel at
http://www.youtube.com/user/zOSCommServer

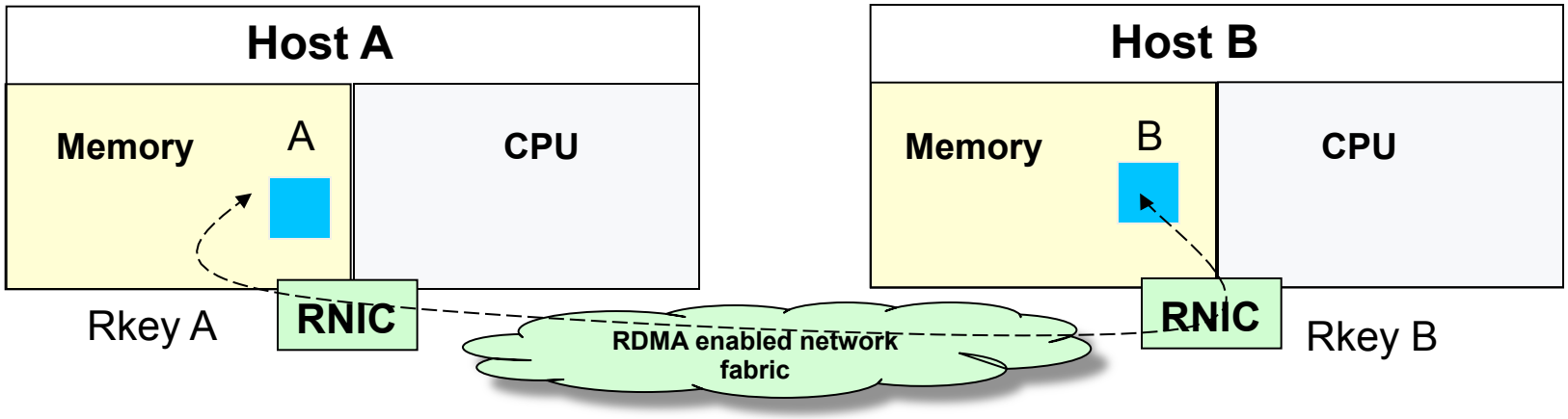
Appendix

Overview of RDMA and RoCE

RDMA (Remote Direct Memory Access) Technology Overview

Key attributes of RDMA

- Enables a host to read or write directly from/to a remote host's memory with minimal (or no) involvement from either host's CPU, by registering specific memory for RDMA partner use.
  - Requires an RDMA network interface card (RNIC)
  - Reduces networking stack overhead by bypassing TCP/IP semantics and offloading networking CPU cycles to the RNIC
    - *Data is moved directly by the RNICs, bypassing the TCP/IP stack.*



RoCE - RDMA over Converged (Enhanced) Ethernet

- RDMA based technology has been available in the industry for many years – primarily based on Infiniband (IB)
  - IB requires a completely unique network eco system (unique hardware such as host adapters, switches, host application software, system management software/firmware, security controls, etc.)
  - IB is popular in the HPC (High Performance Computing) space
- RDMA technology is now available on Ethernet – RDMA over Converged Ethernet (RoCE)
  - RoCE uses existing Ethernet fabric but requires advanced Ethernet hardware (RDMA capable NICs and RoCE capable Ethernet switches)
  - ***RoCE is a game changer!***
    - ***RDMA technology becomes more affordable and prevalent in data center networks***

Communications Server Configuration and Operation

Configuration basics – TCP/IP

- TCP/IP profile specifies configuration for the TCP/IP stack
  - IP interfaces
  - Port reservations
  - Networking options and configuration information including sysplex related
  - Details are specified in *IP Configuration Reference*
  - Accompanying *IP Configuration Guide* describes best practices and provides step by step guides to many common configuration tasks
- TCP/IP profile is stored in MVS datasets
  - Location specified in the TCP/IP started PROC (//PROFILE DD statement)
    - There are other default places we look if that's not specified, but make life easy for people reading your JCL by specifying it
  - Main profile can include sub-files
- TCP/IP configuration can be dynamically changed with VARY OBEY operator command
  - Operator specifies a dataset containing the new TCP/IP profile or a subset of a TCP profile with the new configuration information
  - E.g., VARY TCPIP,,OBEYFILE,DSN=SYS1.TCPPARMS(NEWPROF)

Operational basics – TCP/IP

- Communication server provides a rich set of displays via the NETSTAT operator command.
  - Can be run from the operator console, a TSO session or the z/OS Unix shell
  - Has options to display just about any configuration or state
    - Device state, routing tables, connection information, stack information, etc… far too many to list here
  - Documented in the *IP System Administrator's Commands.*
- TCP/IP stack is controlled with the VARY TCPIP,,*command* operator command. Values for *command* include:
  - OBEYFILE to change configuration (see previous slide)
  - SYNTAXCHECK to perform a syntax check on a TCP/IP Profile file without enacting the specified configuration
  - DATTRACE,PKTTRACE, and OSAENTA to control traces
  - START and STOP to control network resources (e.g., IP interfaces)
  - SYSPLEX to control the sysplex configuration of the stack (e.g., quiesce, leave or join sysplex groups, etc)
  - PURGECACHE to delete the ARP/Neighbor cache

Example of a TCP/IP Netstat command

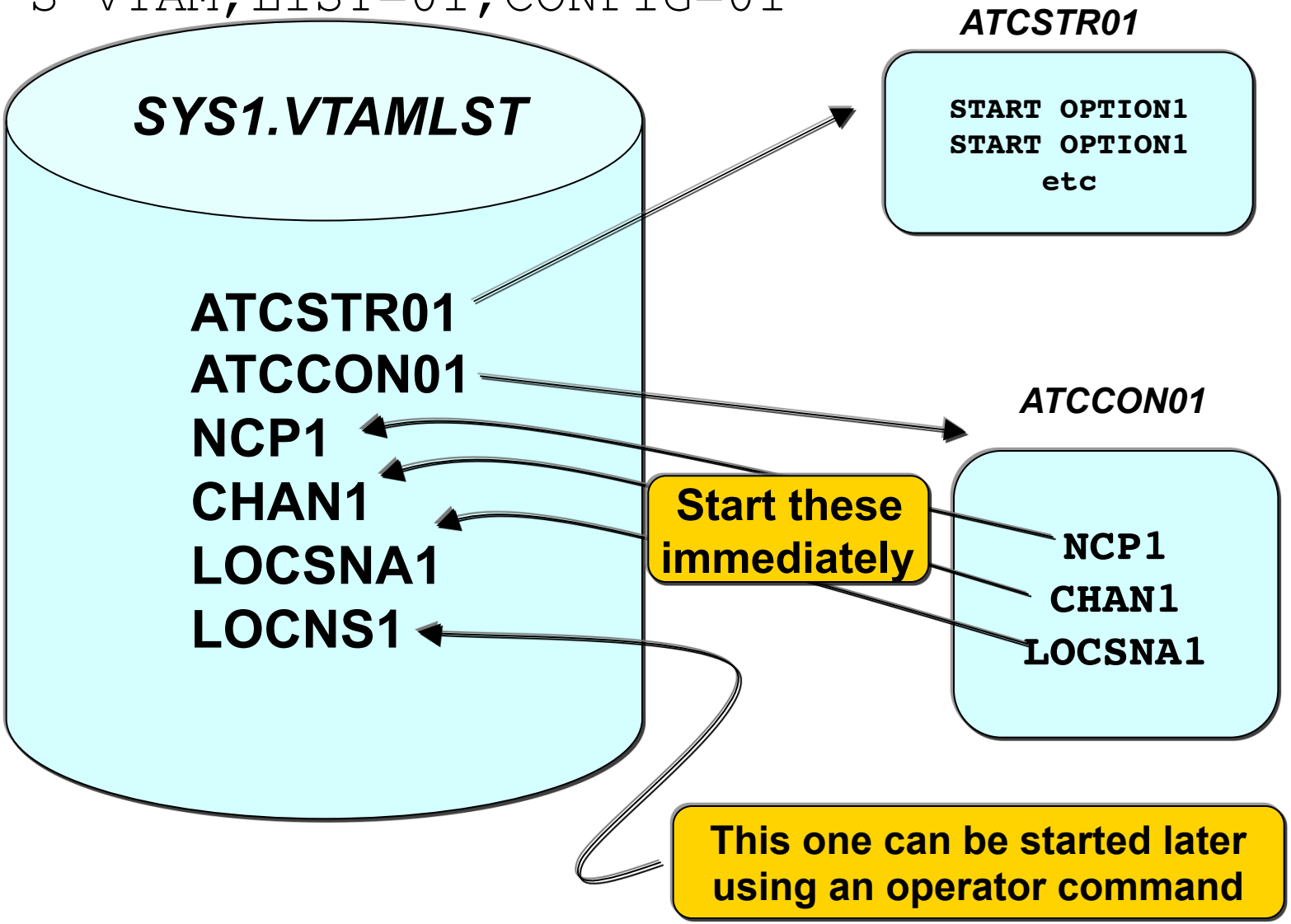
```
D TCPIP,,NETSTAT,ROUTE
EZZ2500I NETSTAT CS V2R1 TCPIP 278
 DESTINATION          GATEWAY           FLAGS     REFCNT  INTERFACE
 DEFAULT              192.168.246.1     UGS       000013  ETH1
 DEFAULT              192.168.247.129   GS        000000  OSATR104
 9.21.111.210/32      0.0.0.0           H         000000  SNA1
 9.23.246.0/24        0.0.0.0           US        000000  ETH1
 9.23.246.0/24        0.0.0.0           S         000000  ETH2
 9.23.246.16/32       0.0.0.0           UH        000000  VLINK1
 9.23.246.23/32       0.0.0.0           H         000000  ETH2
 9.23.246.24/32       0.0.0.0           UH        000000  ETH1
 9.23.247.128/25      0.0.0.0           US        000000  OSATR100
 9.23.247.128/25      0.0.0.0           S         000000  OSATR104
 9.23.247.130/32      0.0.0.0           H         000000  OSATR104
 9.23.247.135/32      0.0.0.0           UH        000000  OSATR100
 127.0.0.1/32         0.0.0.0           UH        000020  LOOPBACK
 13 OF 13 RECORDS DISPLAYED
```

Configuration basics – VTAM

- All VTAM configuration files are contained within the VTAM definition library. Its location is specified by the VTAMLST DD card in the VTAM start proc JCL (usually SYS1.VTAMLST)
- VTAM is configured using a combination of start options and configuration lists:
  - Start options control the configuration of VTAM itself and are usually specified in the ATCSTRxx member of the VTAM definition library.
    - ✓ you specify XX with the LIST=XX parameter when starting VTAM. This would be in your VTAM start proc JCL (xx can be letters or numbers or some symbols)
    - ✓ Start options can also be specified as parameters to VTAM (in your JCL)
  - Configuration lists specify the resources that VTAM will control and are usually specified in a hierarchy of files starting with the ATCCONxx member of the VTAM definition library
    - ✓ You specify XX with the CONFIG=XX parameter when starting VTAM
    - ✓ This file contains a list of member names in the VTAM definition library, each member defines a resource that VTAM will start and control. These are resources like application programs, physical and logical units, networking domains, etc.
    - ✓ Resources are defined hierarchically.
- The reference manual for this configuration is the *SNA Network Implementation Guide* and the *SNA Resource Definition Reference*

Example of VTAM configuration hierarchy

```
S VTAM,LIST=01,CONFIG=01
```



Configuration interdependency – VTAM and TCP/IP

- All DLC services in Communications Server are provided by VTAM
- This means that when configuring TCP/IP interfaces, part of the definition is in VTAM
  - The DLC port is defined in VTAM as a Transport Resource List Entry (TRLE)
  - The IP interface is defined in TCP/IP
  - The IP definition refers to the VTAM definition by TRLE PORTNAME
  - Some IP interface types dynamically create their own TRLE definitions

TCP/IP profile statement

```
INTERFACE OSAQDIO24
 DEFINE IPAQENET
 PORTNAME OSAQDIO2
 IPADDR 100.1.1.1/24
```

VTAMLST member

```
TRLHYDRA VBUILD TYPE=TRL
HYD1     TRLE   LNCTL=MPC,
                READ=(2EC0),
                WRITE=(2EC1),
                MPCLEVEL=QDIO,
                DATAPATH=(2EC2,2EC3),
                PORTNAME=(OSAQDIO2),
                PORTNUM=1
```

Operational basics – VTAM

- VTAM displays are done on the operator console with the DISPLAY NET command
  - A rich set of display commands is provided, similar to TCP/IP Netstat in scope and coverage
  - Documented in *SNA Operation*
- VTAM is controlled with the MODIFY (VTAM jobname) and VARY NET operator commands
  - In general, the MODIFY command controls configuration and resources and VARY command controls state
    - For example, VARY command is used to start and stop resources and MODIFY command is used change configuration information
    - But this rule is very loosely followed and is only a general guideline
  - Documented in *SNA Operation*

Example of a VTAM display command (Enterprise Extender information)

```
D NET,EE
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 = XCAEE2
IST2004I LIVTIME = (10,0) SRQTIME = 15 SRQRETRY = 3
IST2005I IPRESOLV = 0
IST2231I CURRENT HPR CLOCK RATE = STANDARD
IST2232I HPR CLOCK RATE LAST SET TO HIGH ON 11/14/06 AT 22:58:41
IST2233I HPR CLOCK RATE LAST EXITED HIGH ON 11/14/06 AT 22:58:45
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
-------------------------------------------------------------
IST2017I TOTAL RTP PIPES = 4 LU-LU SESSIONS = 3
IST2018I TOTAL ACTIVE PREDEFINED EE CONNECTIONS = 2
IST2019I TOTAL ACTIVE LOCAL VRN EE CONNECTIONS = 0
IST2020I TOTAL ACTIVE GLOBAL VRN EE CONNECTIONS = 0
IST2021I TOTAL ACTIVE EE CONNECTIONS = 2
IST314I END
```

Diagnostic information – TCP/IP

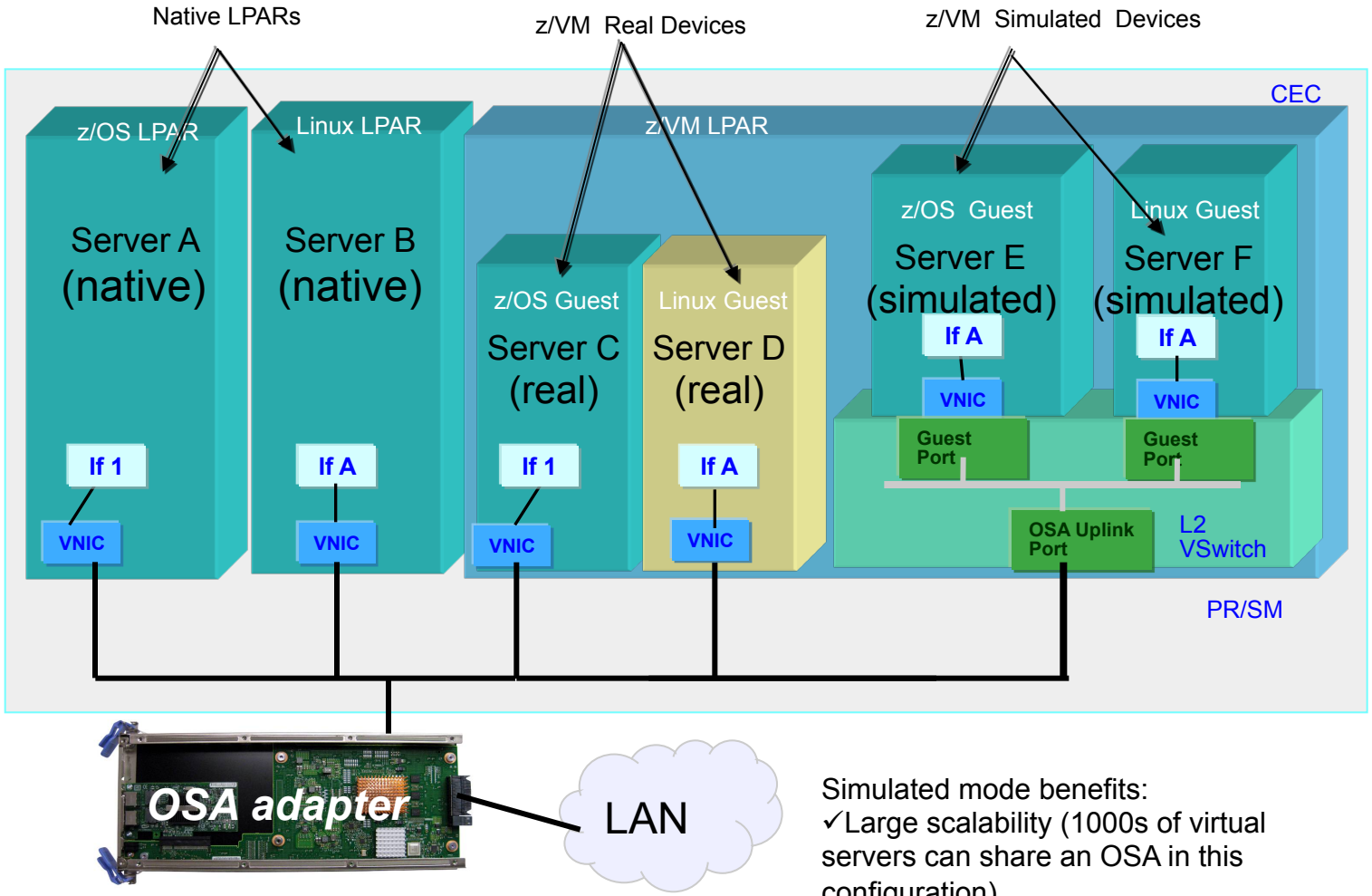
- TCP/IP makes extensive use of the MVS component trace facility and when reporting a problem you will typically be asked to provide this trace. There are multiple TCP/IP component traces including:
  - SYSTCPIP – for TCP/IP stack and Telnet internals
  - SYSTCPRT – for the OMPROUTE routing daemon
  - SYSTCPDA – packet and data trace
  - SYSTCPIK – for the IKE daemon
  - SYSTCPRE – for the resolver (interfaces with DNS to find resources in the network)
  - SYSTCPOT – OSA trace
  - And others.. Documented in *IP Diagnosis*
- TCP/IP also provides a large set of IPCS commands for analyzing TCP/IP dumps
  - Also documented in *IP Diagnosis*
  - You will usually be asked to provide an SVC dump of TCP/IP when reporting a problem
    - Be sure to provide the following:
    - TCP/IP and VTAM address spaces.
    - SDATA options RGN, CSA, LSQA, NUC, PSA, and LPA.
    - CSM data spaces. Add DSPNAME=(1.CSM*) to the DUMP command to include them in the dump.

Diagnostic information – VTAM

- VTAM has its own trace called the VTAM Internal Trace (VIT)
  - Similar to CTRACE in that what's captured is controlled by TRACE option specified in VTAMLST and/or modified by MODIFY TRACE command
  - Is in 64 bit HVCOMMON storage and is captured when VTAM is dumped if CSA is specified on the dump command.
  - If too much information than can fit in the VTAM address space is needed, MVS Generalized Trace Facility (GTF) can be used to write this trace to external datasets.
  - Format of trace records is documented in *SNA Diagnosis volume 2*
- Like TCP/IP SNA provides a large set of IPCS commands for analyzing dumps
  - Documented in *SNA Diagnosis volume 1.*
  - Also like TCPIP, you will likely be asked to provide a dump of the VTAM address space to diagnose problems
    - Should be acquired similarly to how TCP dumps are described on the previous page

Communications Server Technology and Applications

OSA sharing and connectivity summary



Simulated mode benefits:
✓Large scalability (1000s of virtual servers can share an OSA in this configuration)
✓Link aggregation (z/VM can make multiple OSAs appear as a single link to the guests, simplifying redundancy)

Linux OSA drivers are included in Red Hat Enterprise Linux and SUSE Linux Enterprise Server

z/VM Connectivity modes:

Real mode: z/VM provides direct access to the OSA adapter; works like a native LPAR

Simulated mode: z/VM provides a virtual switch and a network of virtual adapters to connect to an OSA-Express adapter. (Note that checksum and segmentation offloads are not supported in simulated mode).

OSA Offload capabilities

The OSA-Express adapter (also called a feature) is a powerful network controller that can offload many processing-intensive networking tasks from the stack and the general purpose CPU

- Address Resolution Protocol (ARP)
  - The stack registers its local IP addresses to the OSA adapter and the adapter responds to ARP requests for those addresses on the stack's behalf
  - **Key value:** If multiple OSAs are attached to the same physical LAN segment and one fails, the others detect the failure and automatically take over its ARP responsibilities (seamlessly moving adapter IP addresses from the failed adapter to the surviving ones!)
- Checksum
  - When enabled, OSA can generate checksums on outgoing packets and validate them on incoming packets, on behalf of the TCP/IP stack
  - **Key value:** reducing CPU cost
- Segmentation
  - When enabled, transfers the overhead of segmenting TCP outbound data into individual IP packets to the OSA-Express device.
  - **Key value:** significantly reduces CPU cost of TCP streaming (e.g., FTP)
- Multiple Queues (z/OS only): OSA adapters can be configured pre-sort different types of inbound traffic onto separate I/O queues to keep incompatible types of traffic (e.g., bulk data) apart. **Key value:** makes z/OS processing more efficient

Use cases for SMC-R and 10GbE RoCE Express for z/OS to z/OS communications

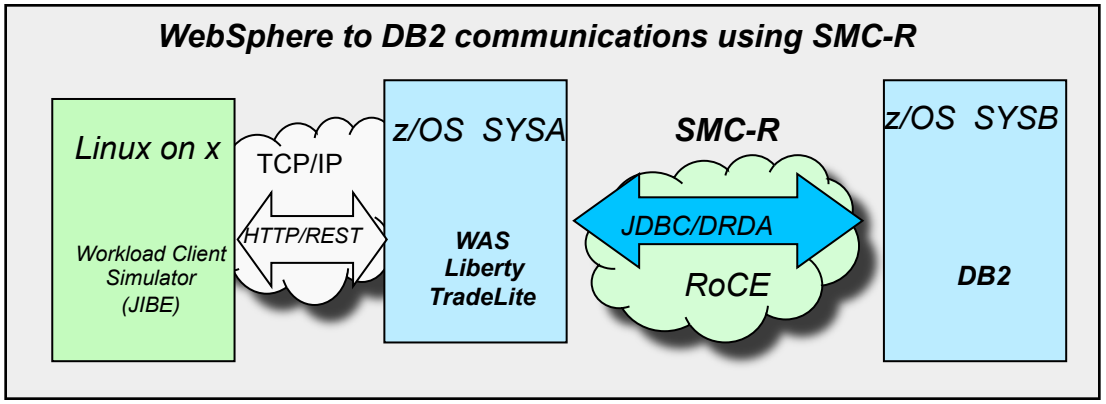
Use Cases

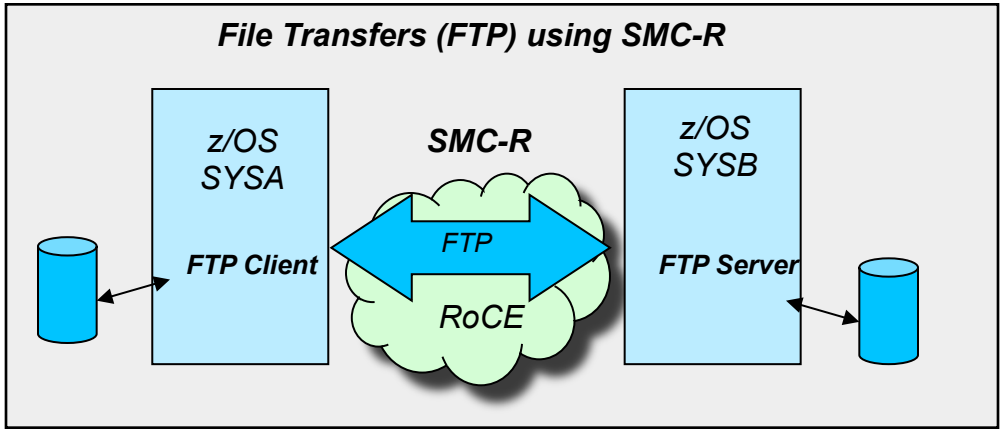
- Application servers such as the z/OS WebSphere Application Server communicating (via TCP based communications) with CICS, IMS or DB2 – particularly when the application is network intensive and transaction oriented
- Transactional workloads that exchange larger messages (e.g. web services such as WAS to DB2 or CICS) will see benefit.
- Streaming (or bulk) application workloads (e.g. FTP) communicating z/OS to z/OS TCP will see improvements in both CPU and throughput
- Applications that use z/OS to z/OS TCP based communications using Sysplex Distributor

Plus … ***Transparent to application software – no changes required!***

Performance impact of SMC-R on real z/OS workloads

40% reduction in overall transaction response time for WebSphere Application Server v8.5 Liberty profile TradeLite workload accessing z/OS DB2 in another system measured in internal benchmarks *



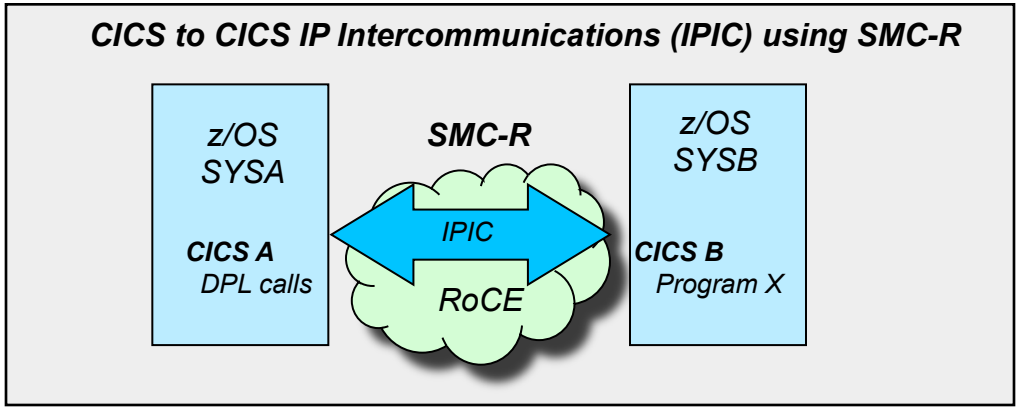


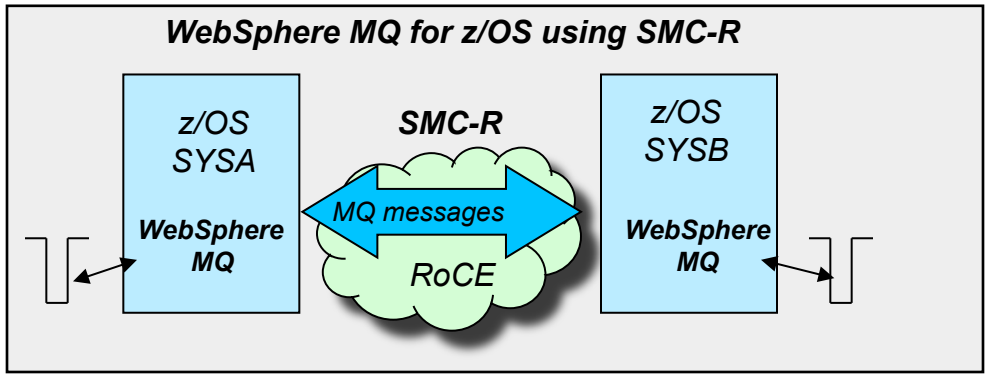
Up to ***50% CPU savings*** for FTP binary file transfers across z/OS systems when using SMC-R vs standard TCP/IP **

* Based on projections and measurements completed in a controlled environment. Results may vary by customer based on individual workload, configuration and software levels.
** Based on internal IBM benchmarks in a controlled environment using z/OS V2R1 Communications Server FTP client and FTP server, transferring a 1.2GB binary file using SMC-R (10GbE RoCE Express feature) vs standard TCP/IP (10GbE OSA Express4 feature). The actual CPU savings any user will experience may vary.

Performance impact of SMC-R on real z/OS workloads (cont)

Up to 48% reduction in response time and up to 10% CPU savings for CICS transactions using DPL (Distributed Program Link) to invoke programs in remote CICS regions in another z/OS system via CICS IP interconnectivity (IPIC) when using SMC-R vs standard TCP/IP *



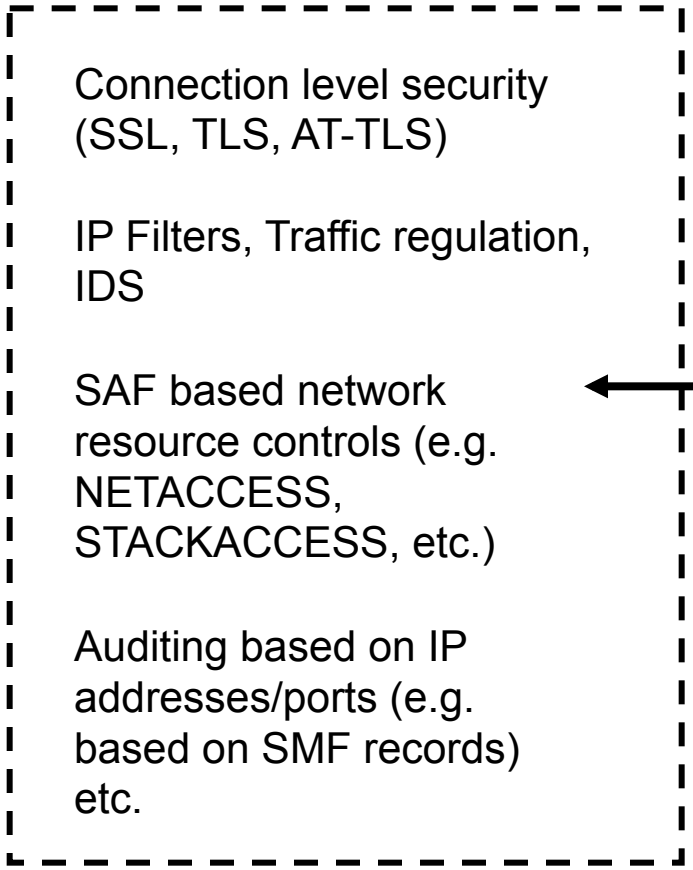


WebSphere MQ for z/OS ***realizes up to 200% increase in messages per second*** it can deliver across z/OS systems when using SMC-R vs standard TCP/IP ****

* Based on internal IBM benchmarks using a modeled CICS workload driving a CICS transaction that performs 5 DPL (Distributed Program Link) calls to a CICS region on a remote z/OS system via CICS IP interconnectivity (IPIC), using 32K input/output containers. Response times and CPU savings measured on z/OS system initiating the DPL calls. The actual response times and CPU savings any user will experience will vary.

** Based on internal IBM benchmarks using a modeled WebSphere MQ for z/OS workload driving non-persistent messages across z/OS systems in a request/response pattern. The benchmarks included various data sizes and number of channel pairs The actual throughput and CPU savings users will experience may vary based on the user workload and configuration.

SMC-R preserves existing security model



The hybrid nature of SMC-R (beginning with TCP/IP, then switching to SMC-R) allows all existing IP and TCP layer security features to automatically apply for SMC-R connections

- Without requiring any changes from a customer perspective
- And without requiring these functions to be retrofitted into a new protocol

Note: IPSec tunnels are not supported with SMC-R. When IPSec is enabled for a TCP connection then z/OS will automatically opt out of using SMC-R

Note: Security functions which require TCP packet inspection or modification are not supported because RDMA replaces TCP packets**.** These include: IPSec tunnels, filters that deny packets during an active connection, or methods that require packet tagging such as MLS.

- z/OS will automatically opt out of using SMC-R for TCP connections that require these security types.

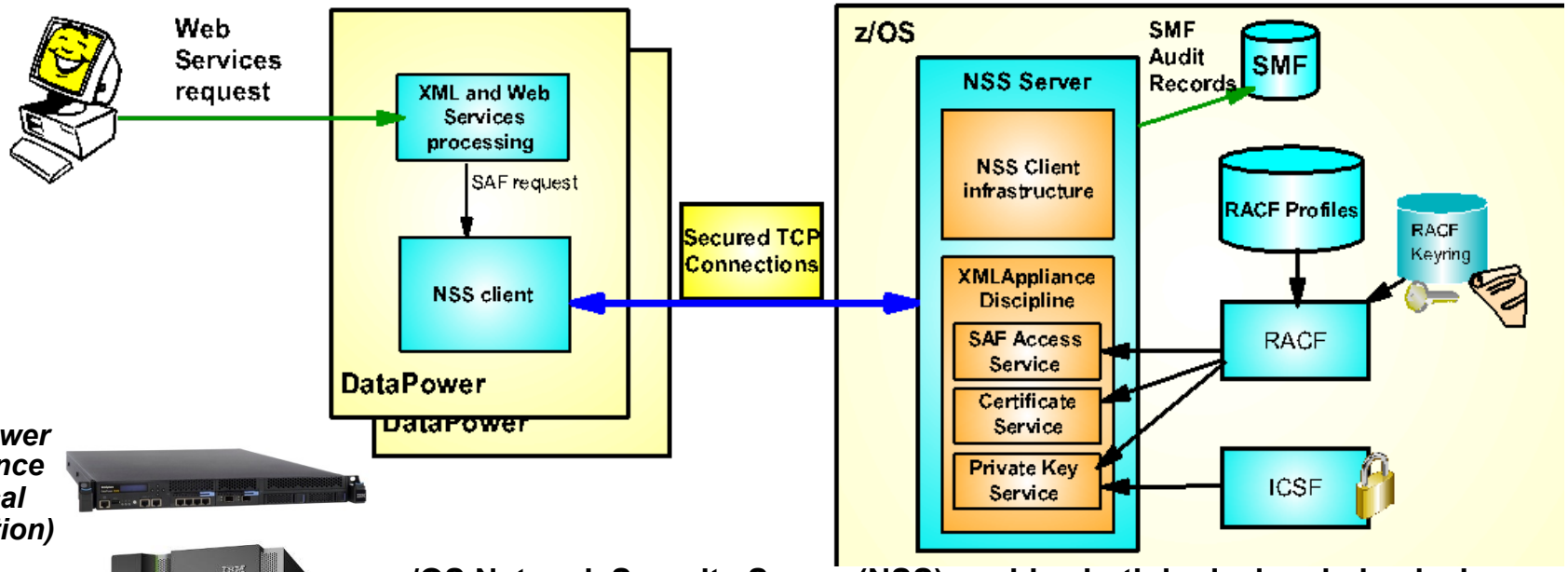
SMC-R roadmap

- First platform to support SMC-R is z/OS V2R1 (GA 9/30/13)
  - z/OS to z/OS communications within the datacenter
- IBM intends for SMC-R to be an open, pervasive protocol
  - To allow RDMA communication between any closely coupled enterprise platforms
    - E.g., zLinux, Linux on other platforms, etc. IBM intends to work with the Linux community to distribute SMC-R on Linux
  - Specification is published to the IETF to enable pervasiveness
    - http://tools.ietf.org/html/draft-fox-tcpm-shared-memory-rdma-04
- Future considerations:
  - Virtualization support for the RoCE Express feature
  - z/VM guest support for PCIe devices (statement of direction issued)
    - RoCE Express feature is a PCIe device

DataPower appliances integrate with z/OS security

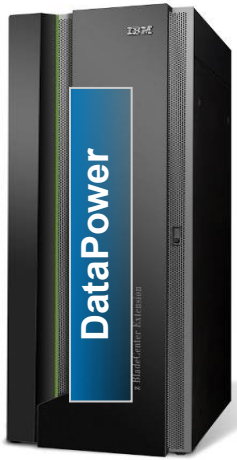
WebSphere DataPower Appliances

- Advanced WebServices operations
- Message format transformation for traditional z/OS applications
- Offload XML and WebServices security



DataPower Appliance (logical integration)

DataPower XI50z Integrated Blade (physical integration)

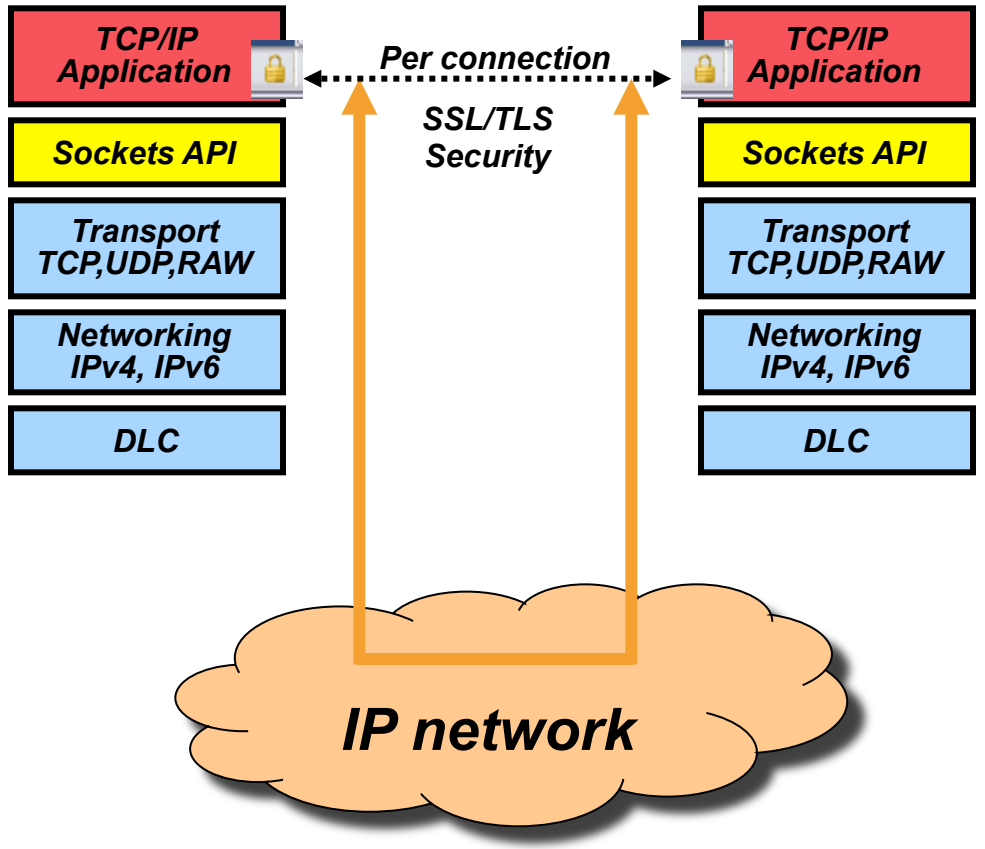


- **z/OS Network Security Server (NSS) enables both logical and physical integration between DataPower and z/OS security.**
- **DataPower accesses centralized SAF services via z/OS NSS for:**
  - **SAF access service**: Provides SAF-based authentication (of DP users) and access control (of DP resources) with SMF auditing
  - **Certificate service**: Provides retrieval of RSA certificates from a SAF keyring
  - **Private key service** provides:
    - Private RSA key retrieval (clear key only)
    - RSA signature and decryption operations (secure key only)
- **z/OS security administered through z/OS facilities like RACF (SAF), ICSF**

Transport Layer Security (TLS) overview

- Traditionally a socket layer service
- TCP only
  - Requires reliable transport
- Application-to-application protection
- Partner authentication via digital certificates
- Data protection using “cipher suites” that combine data authentication, integrity and encryption algorithms.
- z/OS implementations:
  - System SSL (part of z/OS Cryptographic Services) for use with C and C++ applications
  - Java Secure Sockets Extension (JSSE)

Source code changes are required to use either of these



Some z/OS applications that use AT-TLS

- Communications Server applications
  - TN3270 server
  - FTP client and server
  - CSSMTP
  - Load Balancing Advisor
  - IKED NSS client
  - NSS server
  - Policy Agent
- DB2 DRDA
- IMS Connect
- JES2 NJE
- Tivoli NetView applications
  - MultiSystem Manager
  - NetView Management Console
- RACF Remote Sharing Facility
- CICS Sockets applications
- 3rd party applications
- Customer applications

Advantages of using AT-TLS on z/OS

- **Reduce costs**
  - Application development
    - Cost of System SSL integration
    - Cost of application's TLS-related configuration support
  - Consistent TLS administration across z/OS applications
  - Gain access to new features with little or no incremental development cost

- **Complete and up-to-date exploitation of System SSL features**
  - AT-TLS makes the vast majority of System SSL features available to applications **– V2R1 is a perfect example!**
  - AT-TLS keeps up with System SSL enhancements – as new features are added, your applications can use them by changing AT-TLS policy, not code

- **Ongoing performance improvements**
  Focus on efficiency in use of System SSL

- **Great choice if you haven't already invested in System SSL integration**
  Even if you have, consider the long-term cost of keeping up vs. short term cost of conversion

IDS: z/OS Communications Server IDS features

IDS Events

- **Scans – attempts by remote nodes to discover information about the z/OS system**

- **Attacks – numerous types**

  - **Malformed packets**
  - **IP option and IP protocol restrictions**
  - **Specific usage ICMP**
  - **Interface and TCP SYN floods**
  - **and so forth… including several new attack types added to V1R13**

- **Traffic Regulation**
  - **TCP - limits the number of connections any given client can establish**
  - **UDP – limits the length of data on UDP queues by port**

Defensive actions

- Packet discard
- Limit connections
- Drop connections (V1R13)

Reporting

- Logging
- Console messages
- IDS packet trace
- Notifications to external event managers (like Tivoli NetView)

z/OS in-context IDS broadens overall intrusion detection coverage, complements network-based IDS:

- "In-context" because z/OS is the communications end point, not an intermediary
- Ability to evaluate inbound encrypted data - IDS applied after decryption on the target system
- Avoids overhead of per packet evaluation against table of known attacks - IDS policy checked after attack probe fires
- Detects statistical anomalies realtime - target system has stateful data / internal thresholds that generally are unavailable to external IDSs
- Policy can control prevention methods on the target, such as connection limiting and packet discard