

All Aboard the Network Express!



SHARE Orlando, FL 2026 Winter Technical Conference
Session 269

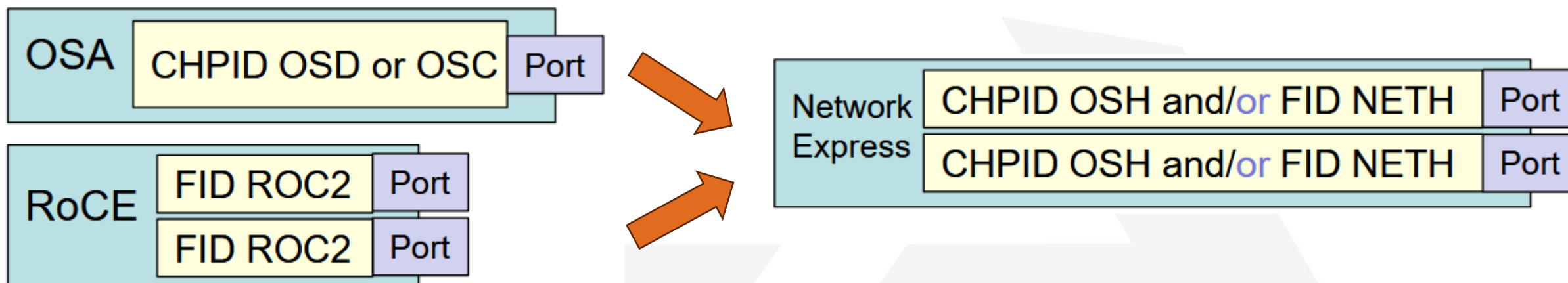
Grant Mericle – gmericle@us.ibm.com
Karthik Sundaresan – ksundar@ibm.com

Why are you here?

- Your company bought/plans to buy Network Express cards and now they expect YOU to configure them
- Your company is buying a z17 and is asking YOU whether or not to purchase Network Express cards
- You heard Salon 23 has dim lighting and are hoping to catch a nap

So what IS Network Express?

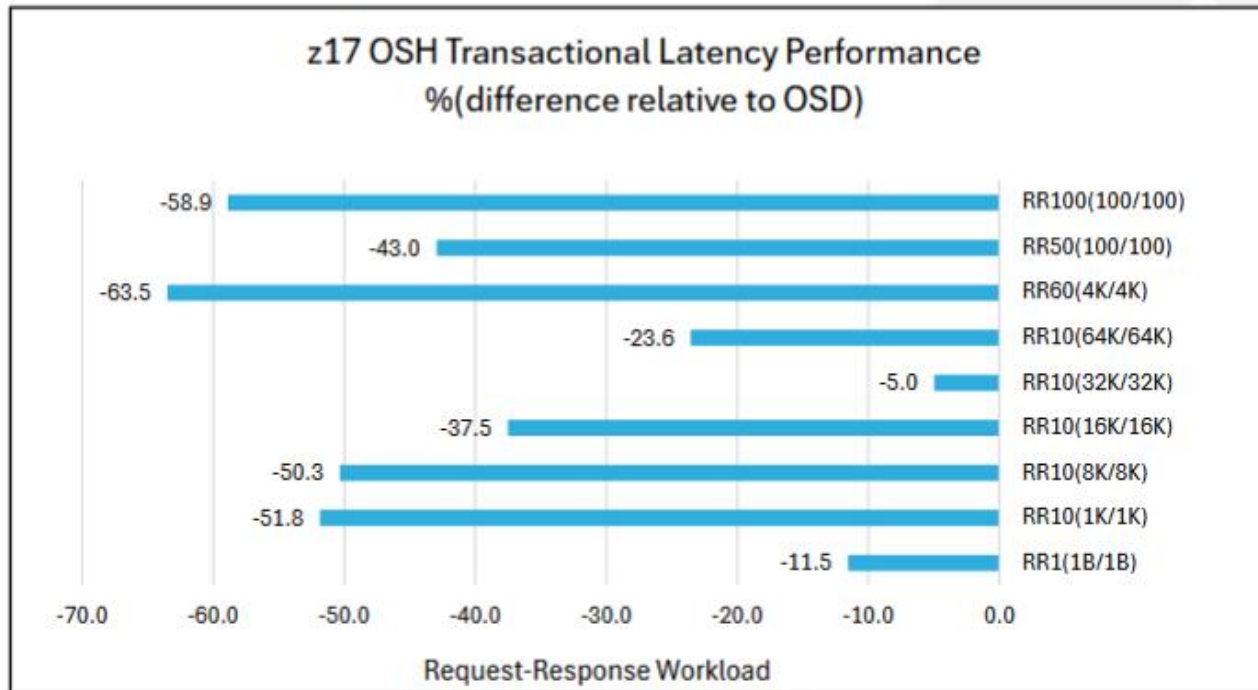
- next-gen OSA
- also the next-gen RoCE!
- both functions are now supported on a single port



What if I don't care about SMC-R (RoCE)?

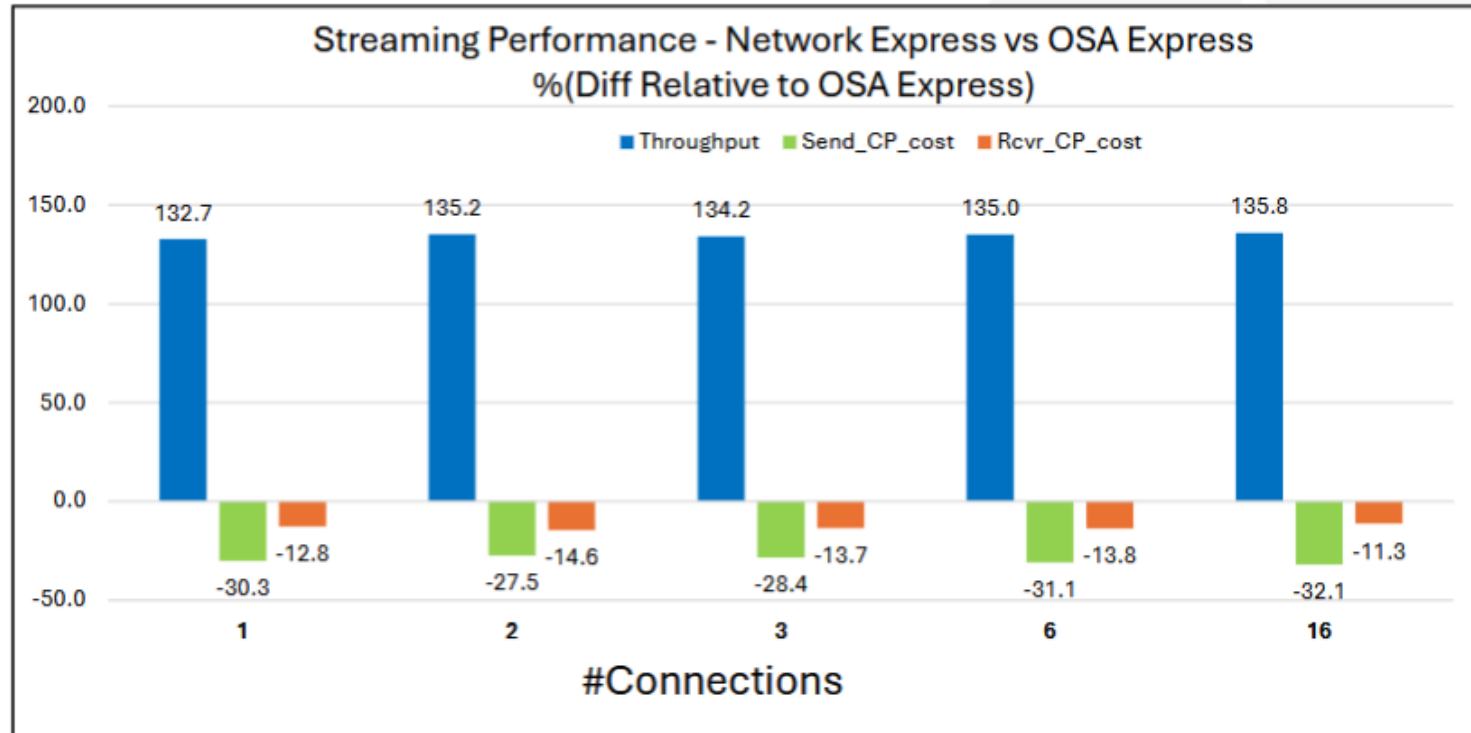
- You should at least try it
- Seriously, you have no excuse now
- In the meantime... OSH has shown some impressive performance benefits over OSD.

Interactive (RR) Performance

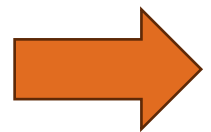


- OSH (EQDIO) compared to OSD (QDIO)
- 2CPs + 25Gb + 1500 MTU + LSO
- Significant reduction in transactional latency
 - Up to 60% improvement in network response times
- Upcoming hardware bundle updates could further improve performance

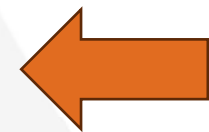
Streaming (STR) Performance



- OSH (EQDIO) compared to OSD (QDIO)
- 2CPs + 25Gb + 1500 MTU + LSO
- Streaming performance (throughput, sender & receiver CPU cost)
- Significant performance improvement
 - Up to **135%** improvement in throughput compared to OSD
 - **30%** average savings in CPU cost



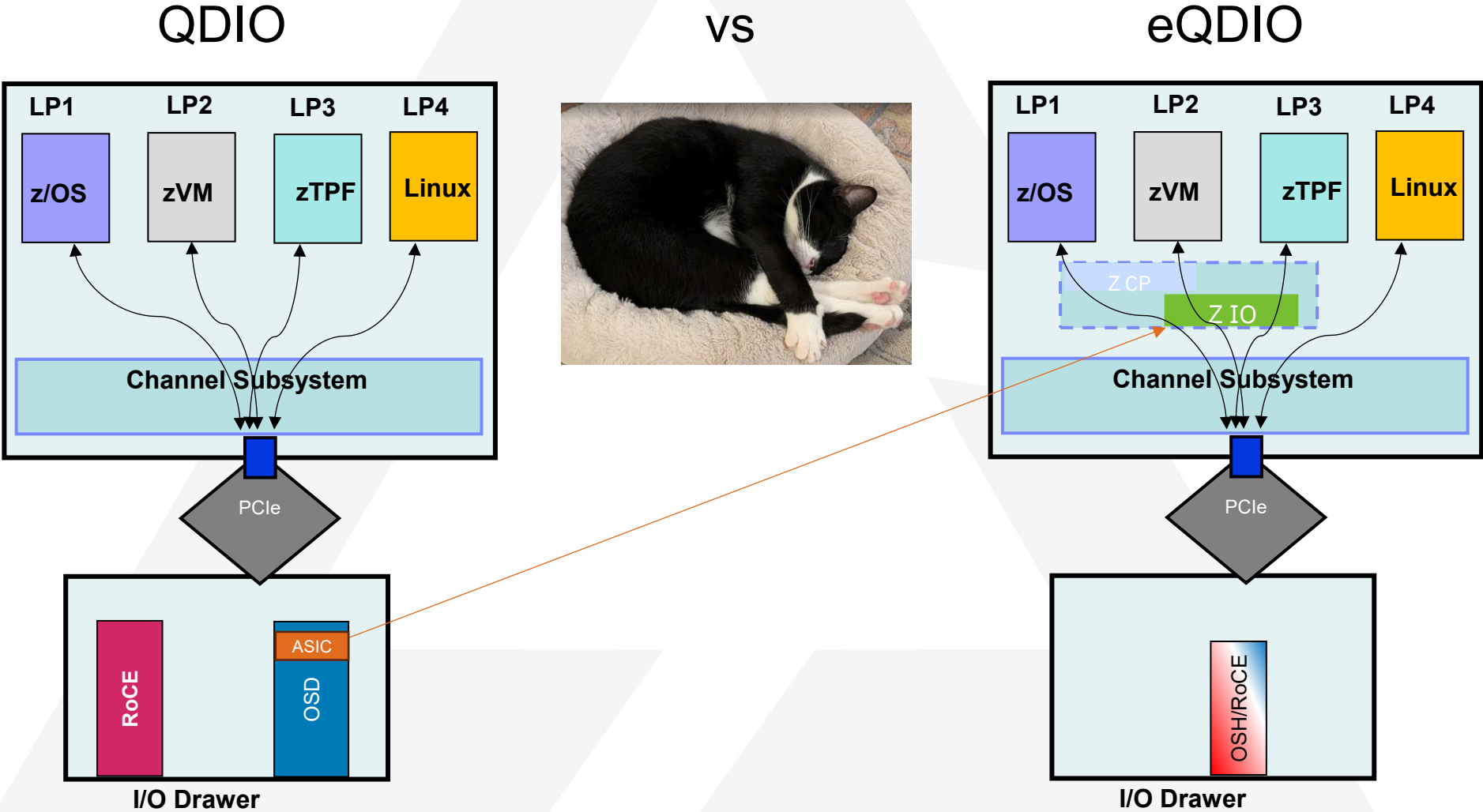
z/OS Communications Server Performance Update – 3:45 TODAY, Salon 16



How's it different from OSA Express?

- You no longer need to buy a separate OSA card and RoCE card
- 2 ports, each with its own PCHID (looks like two separate cards in HCD)
- Both 10Gb or both 25Gb — don't mix!

Anything else that's different?



Is the configuration any different than OSD?

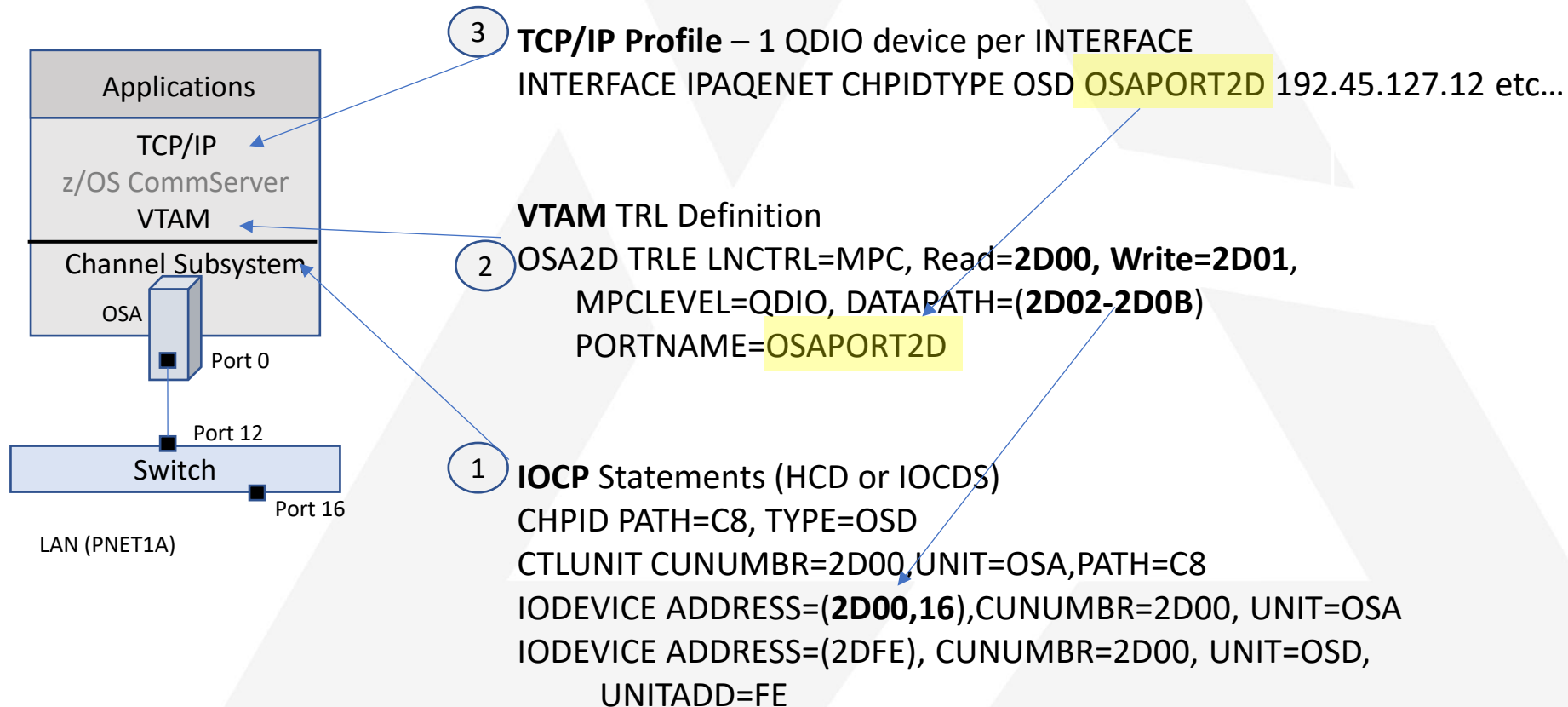
- New CHPID type (OSH)
- New INTERFACE type: EQENET
- Interface cleanup and simplification

Type	Configuration
OSD (QDIO)	IPAQENET/6 INTERFACE
OSH (EQDIO)	EQENET/6 INTERFACE

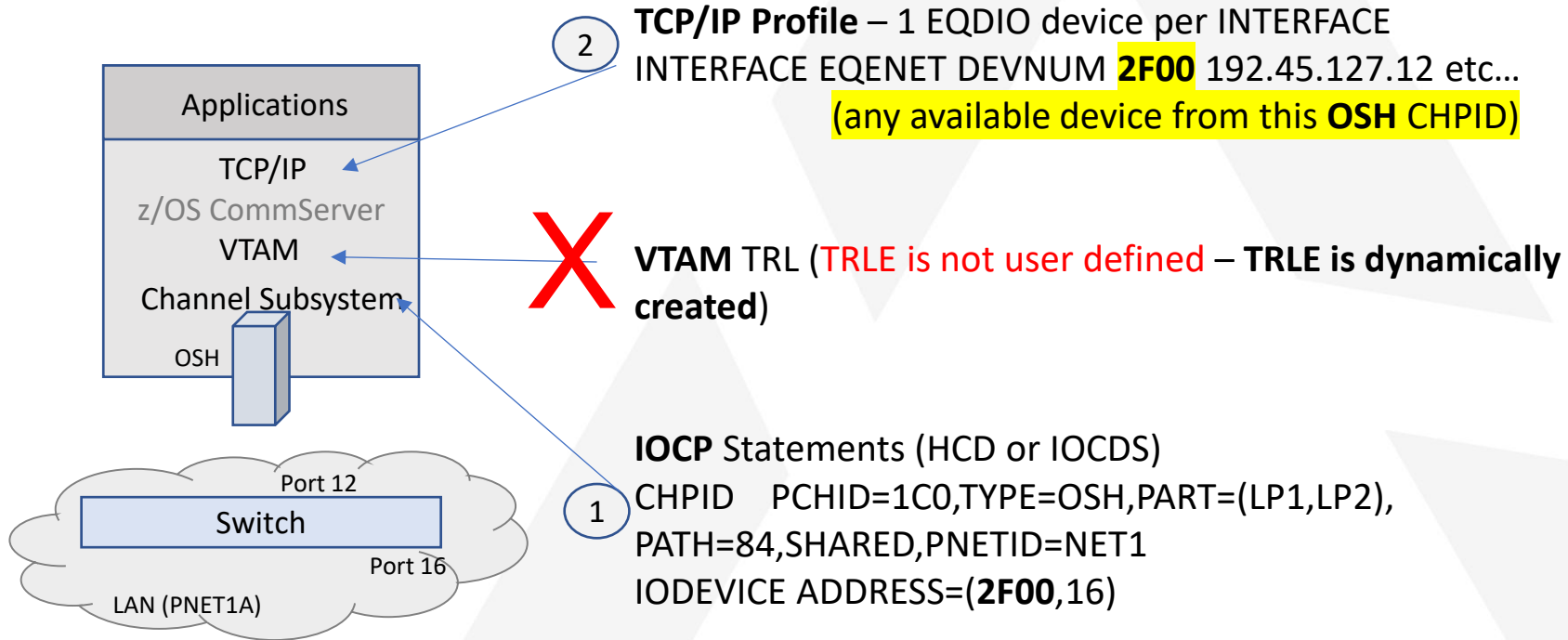
EQENET Changes

- Removed many old parameters
 - CHPIDTYPE, PORTNAME, PRIROUTER/SECROUTER, OLM, DYNVLANREG, READSTORAGE, INBPERF
- Kept core operational parameters
 - IPADDR (subnet mask now required), VLANID, VMAC, MTU, etc.
- Only new parameter: DEVNUM
- Removed user-defined TRLE requirement

Review time – OSA Express config



Network Express config



Where can I find the device addresses?

08.57.47 **D M=CHP (71)**

08.57.47 IEE174I 08.57.47 DISPLAY M 867

CHPID 09: **TYPE=35**, DESC=**OSA HYBRID**, ONLINE

DEVICE STATUS FOR CHANNEL PATH 09

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
810	+	+	+	+	.	.	.	.	.	.	.	.	.	.	.	.

SWITCH DEVICE NUMBER = NONE

***** SYMBOL EXPLANATIONS *****

+ ONLINE @ PATH NOT VALIDATED - OFFLINE . DOES NOT EXIST

* PHYSICALLY ONLINE \$ PATH NOT OPERATIONAL

DEVNUM: Insider Tip

- ALWAYS use the first address in CHPID device range
- If multiple interfaces on a stack share the CHPID, they MUST specify the same DEVNUM anyway
- Configured vs actual device may differ

Usage & Invocation – TCP/IP Configuration

EQENET definition

```
INTERFACE OSA1 DEFINE EQENET
DEVNUM 8100
IPADDR 16.12.37.160/20
VLANID 702
VMAC
MTU 1500
SMCR PFID 3140 SMCRIPADDR 16.12.37.165
```

Netstat devlinks report (netstat -d)

```

INTFNAME: OSA1          INTFTYPE: EQENET          INTFSTATUS: READY
DEVNUM: 8100          ACTDEVNUM: 8101          DEVSTATUS: READY
CHPIDTYPE: OSH        CHPID: 71              PCHID: 0139
SMCD: YES             SMCR: YES              PNETID: NETWORK7CD
ADAPTER GEN: NETWORK EXPRESS
TRLE: IUTE8101        CODE LEVEL: 3031000000001440
PORTNAME: EZAP8101
SPEED: 0000010000 (10G)
IPBROADCASTCAPABILITY: NO
VMACADDR: 42000A69C925  VMACORIGIN: OSA      VMACROUTER: ALL
CFGMTU: 9000          ACTMTU: 9000
IPADDR: 16.11.37.160/20
VLANID: 602           VLANPRIORITY: DISABLED
TOTAL READ STORAGE: 19.0M
CHECKSUMOFFLOAD: YES   SEGMENTATIONOFFLOAD: YES
SECCLASS: 255          MONSYSPLEX: YES
ISOLATE: NO
MULTICAST SPECIFIC:
MULTICAST CAPABILITY: YES
GROUP      REFCNT      SRCFLTMD
-----
224.0.0.5  0000000001  EXCLUDE
SRCADDR: NONE
224.0.0.1  0000000001  EXCLUDE
SRCADDR: NONE

```

c

- Interface type is EQENET
- Configured device and actual device
- CHPIDTYPE is OSH (EQDIO)
- OSA Generation
- TRLE and PORTNAME dynamically generated
- OSA firmware level
- ReadStorage dynamically managed (not user defined)

Netstat devlinks report (part 2)

INTERFACE STATISTICS:

```
BYTESIN                = 8042214
INBOUND PACKETS        = 81593
INBOUND PACKETS IN ERROR = 0
INBOUND PACKETS DISCARDED = 0
INBOUND PACKETS WITH NO PROTOCOL = 0
BYTESOUT               = 1025416
OUTBOUND PACKETS       = 10502
OUTBOUND PACKETS IN ERROR = 0
OUTBOUND PACKETS DISCARDED = 0
```

ASSOCIATED IQD CONVERGED INTERFACE: EZAIQCF9 IQC STATUS: READY

```
BYTESIN                = 19714
INBOUND PACKETS        = 121
BYTESOUT               = 53236
OUTBOUND PACKETS       = 464
```

SMCR CAPABILITY: V2

ASSOCIATED MULTI-SUBNET RNIC INTERFACE: EZARIUT1314E

ROCE PFID: 314E SMCRMTU: 4096

SMCRIPADDR: 16.11.39.160

UNASSOCIATED ISM INTERFACES: EZAISMU1 EZAISMU2 EZAISMU3 EZAISMU4

IPV4 LAN GROUP SUMMARY

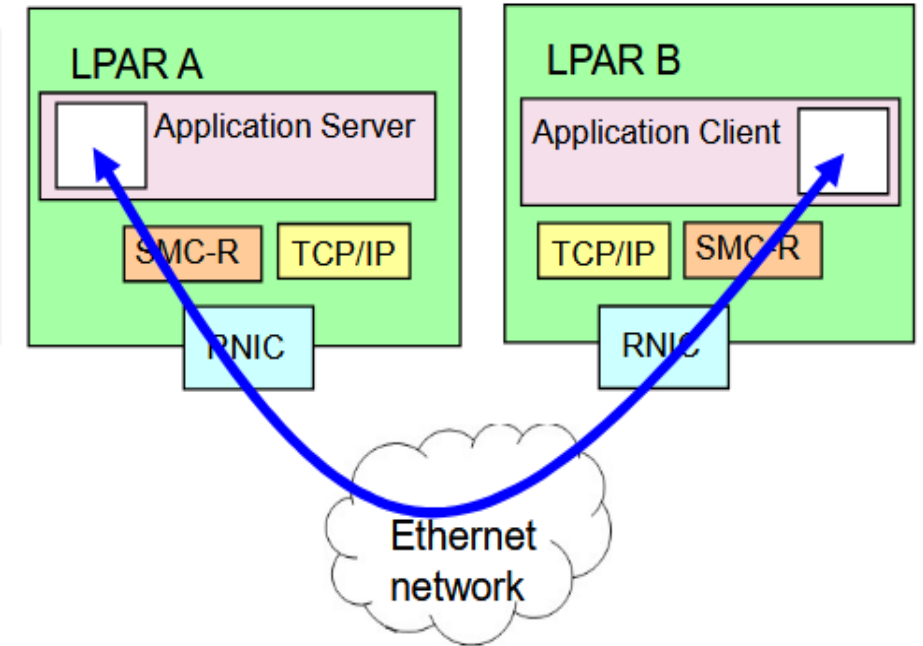
LANGROUP: 00002

NAME	STATUS	ARPOWNER	VIPAOWNER
----	-----	-----	-----
O4ETHA1	ACTIVE	O4ETHA1	NO
O4ETHA0	ACTIVE	O4ETHA0	YES

- Associated HSCI interface
- Associated SMC interfaces

What's SMC-R again?

- Shared Memory Communication over RDMA
- Before SMC-R existed, an application would have to be programmed specifically to use RDMA. SMC-R allowed for a TCP/IP application to use RDMA *transparently* – no changes required
- During a TCP handshake, both endpoints determine whether the connection is eligible for SMC-R, and if so, switch over automatically.



How do I configure SMC-R? Step 1

- In HCD, define NETH PFIDs associated with the Network Express PCHID
- Use the same PNetID that was defined for the OSH CHPID

```
09.38.57 d pcie
09.38.57 IQP022I 09.38.57 DISPLAY PCIE 767
PCIE      0010 ACTIVE
PFID      DEVICE TYPE NAME      STATUS  ASID  JOBNAME  CHID VFN
00000501  Network Express      CNFG           02FC 0001
00000502  Network Express      CNFG           02FC 0002
```

```
09.58.32 d pcie,pfid=501
09.58.32 IQP024I 09.58.32 DISPLAY PCIE 776
PCIE      0010 ACTIVE
PFID      DEVICE TYPE NAME      STATUS  ASID  JOBNAME  CHID VFN
00000501  Network Express      CNFG           02FC 0001
CLIENT ASIDS: NONE
PNetID 1: ZOSNET
```

How do I configure SMC-R? Step 2

1. Add SMCR PFID xxxx SMCRIPADDR a.b.c.d to EQENET definition
 - For OSA Express & RoCE, you could pick any PFID you wanted (separate cards)
 - For Network Express, the OSH and PFID should be for the same PCHID (i.e. port)
 - SMCRIPADDR must be on same subnet as the EQENET IPADDR

```
;
INTERFACE OSA1 DEFINE EQENET
  DEVNUM 8100
  IPADDR 16.12.37.160/20
  VLANID 702
  VMAC
  MTU 1500
  SMCR PFID 3140 SMCRIPADDR 16.12.37.165
```

How do I configure SMC-R? Step 3

Add this to the TCPIP profile:

```
GLOBALCONFIG SMCR SMCGLOBAL SMCEID TEST ENDSMCEID
```

Note: You can turn SMCR off dynamically with an obeyfile: GLOBALCONFIG NOSMCR

How do I try out SMC-R?

Start a TCP connection (SSH, FTP, etc) between the two systems where you configured SMC-R

Once established, the netstat -A (use -P <port> to filter) report will show whether the connection has switched over to SMC-R:

SMC Information:

SMCDV2Status: Inactive

SMCReason: 00005325 - Peer did not request SMC-Dv2

SMCDV1Status: Inactive

SMCReason: 0000530F - Peer did not request SMC-Dv1

SMCRV2Status: Active

SMCGroupId: C9280200

LocalSMCLinkId: C9280201

RemoteSMCLinkId: DB930101

LocalSMCRcvBuf: 64K

RemoteSMCRcvBuf: 64K

SMCEID: TEST

Usage & Invocation – Current Limitations / Guidance



- No EQDIO support on z/OS 2.4 or a VSE client
 - NETH support on z/OS 2.4 for SMC-Rv1 is available (PH66350/OA67843)
- SNMP support for OSA (OSH) will not be available at z17 GA
 - Support targeted for 1Q/2026 (zOS CS 2.5, 3.1 and 3.2) – APAR PH67742
- OSAENTA (Network Traffic Analyzer) trace support currently unavailable
 - Alternatively, a sniffer (wireshark) trace may be used
 - Under consideration
- QDIOSYNC support currently unavailable (Automated mechanism to collect SW and HW traces concurrently)
 - Under development ~2Q/2026
- NETH PFIDs (RoCE) - Guideline
 - For clients using IBM Network Express 10G or 25G with z/OS CommServer 2.5 or 3.1, it is recommended to set up a maximum of 16 NETH FIDs on the same adapter when OSH is also enabled.
- *VSwitch EQDIO Simulation for z/OS guest support*
 - Allows z/OS guests to connect to a Layer-2 Vswitch
 - APARs VM66877 (7.4 CP) and VM66878 (7.3 CP), and driver D61C Bundle #15
- “QDIO Accelerator” is supported for EQDIO but is limited to EQDIO-to-EQDIO interfaces only.

Enablement APARs

z/OS CommServer APARs (2.5 and 3.1):

- OA64896 for SNA
- PH54596 for TCP/IP

IOS APAR:

- OA63265

Next Steps (Pre-Migration)

- Move from DEVICE/LINK/HOME to INTERFACE
- Move from PRIROUTER/SECROUTER to VMAC ROUTEALL
- Add subnet mask to IPADDR – this should match what's defined in your omproute interface statement
- Do you use static routes + no subnet mask defined? → email me

Once you have Network Express

- Create your EQENET INTERFACE
- Use Auto-Migrate only if IPLing between z16 ↔ z17
- Wait. What's Auto-Migrate?

Usage & Invocation – Auto-Migrate

OSD IPAQENET Interface statement that dynamically switches to OSH EQENET when activated on z17

INTERFACE IPAQENET
PORTNAME + DEVNUM

IPLed
on z16?

IPLed
on z17?

Activate
OSD using
Portname

Activate
OSH using
DEVNUM

User defined static TRLE

TRLE
Dev 1

Dynamically created TRLE

TRLE
Dev 1

Usage & Invocation – TCP/IP Configuration

EQENET definition

```
;
INTERFACE OSA1 DEFINE EQENET
  DEVNUM 2B70
  IPADDR 16.12.37.160/20
  VLANID 702
  VMAC
  MTU 1500
  SMCR PFID 3140 SMCRIPADDR 16.12.37.165
```

“Auto-migrate” definition

```
;
INTERFACE OSA1 DEFINE IPAQENET ; Enhanced QDIO Migration OSH
  DEVNUM 2B70
  PORTNAME PORTQDIO
  IPADDR 16.12.37.160/20
  VLANID 702
  VMAC
  MTU 1500
  READSTORAGE GLOBAL
  INBPERF DYNAMIC WORKLOADQ
  SMCR PFID 3140 SMCRIPADDR 16.12.37.165
```

- You MUST code a subnet mask

Usage & Invocation – Netstat devlinks

netstat devlinks (-d) report for auto-migrated OSH EQENET IPv4 (partial report)

```
INTFNAME: 04ETHA2          INTFTYPE: EQENET          INTFSTATUS: READY
**AUTOMIGRATED**
DEVNUM: 8102              ACTDEVNUM: 8104          DEVSTATUS: READY
CHPIDTYPE: OSH            CHPID: 71              PCHID: 0139
SMCD: YES                 SMCR: YES               PNETID: NETWORK7CD
ADAPTER GEN: NETWORK EXPRESS
TRLE: IUTE8104            CODE LEVEL: 3031000000001440
PORTNAME: EZAP8104
SPEED: 0000010000 (10G)
IPBROADCASTCAPABILITY: NO
VMACADDR: 42000C69C925  VMACORIGIN: OSA        VMACROUTER: ALL
CFGMTU: 9000              ACTMTU: 9000
IPADDR: 16.12.37.160/20
VLANID: 702              VLANPRIORITY: DISABLED
TOTAL READ STORAGE: 19.0M
CHECKSUMOFFLOAD: YES      SEGMENTATIONOFFLOAD: YES
SECCLASS: 255            MONSYSPLEX: YES
ISOLATE: NO
```

- Auto-migrated interface
- Missing parms for EQENET can be auto-generated (like VMAC)
- However, subnet mask MUST be defined by user

Your feedback is important!

Submit a session evaluation for each session you attend:

www.share.org/evaluation

