

SMC-D @ BJB

Shared Memory Communications using Direct Memory Access

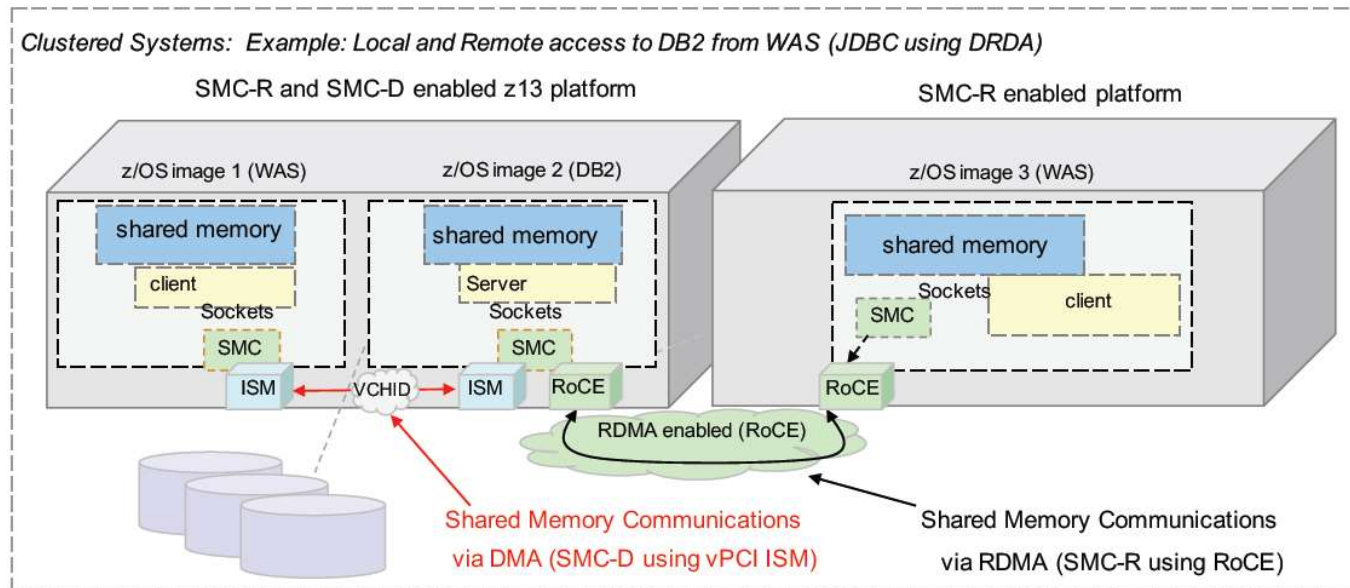
Goals

- **Reduce MSU**
- Performance (Bandwidth)
- Security

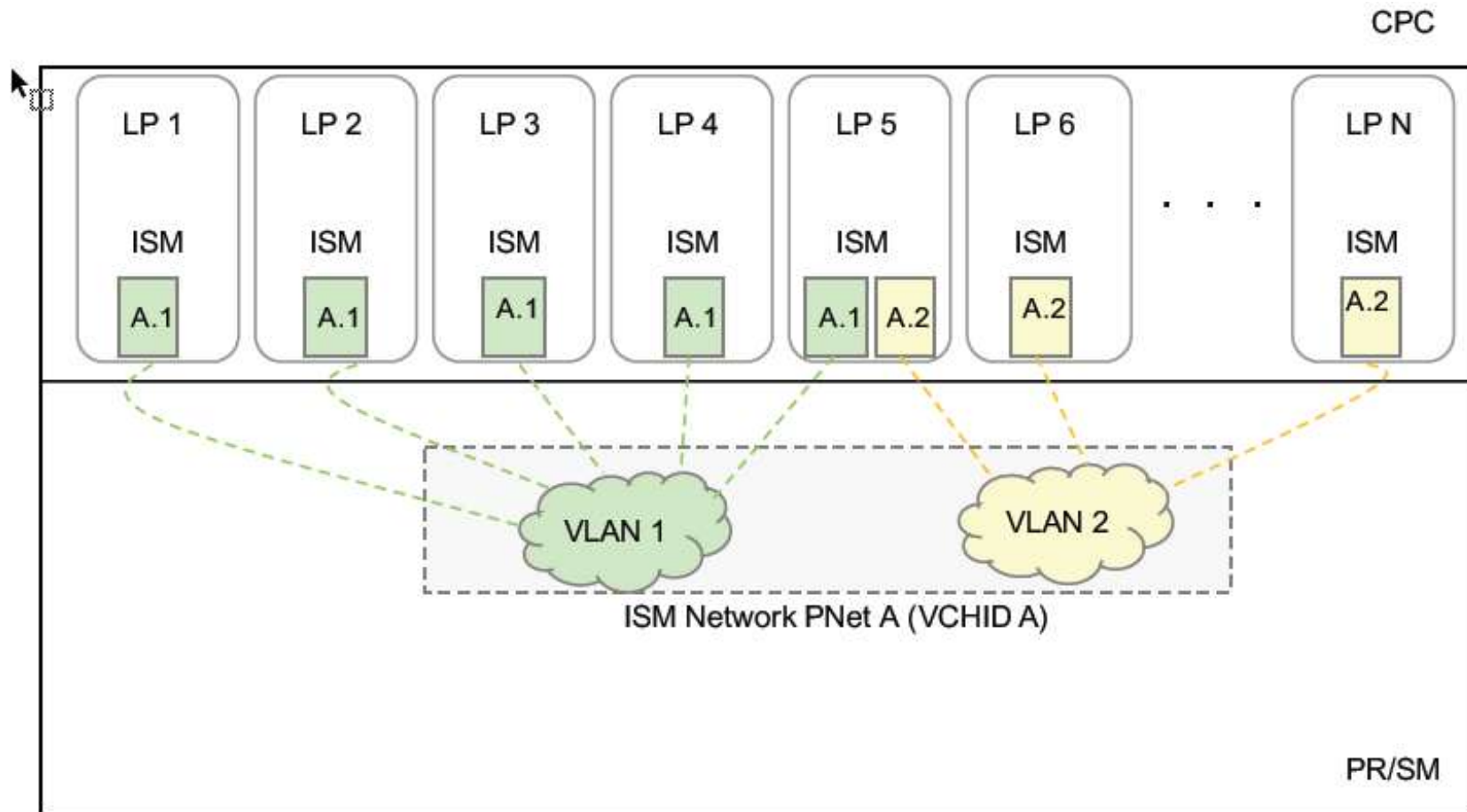
Overview



Shared Memory Communications within the enterprise data center (RoCE) and within System z (ISM)

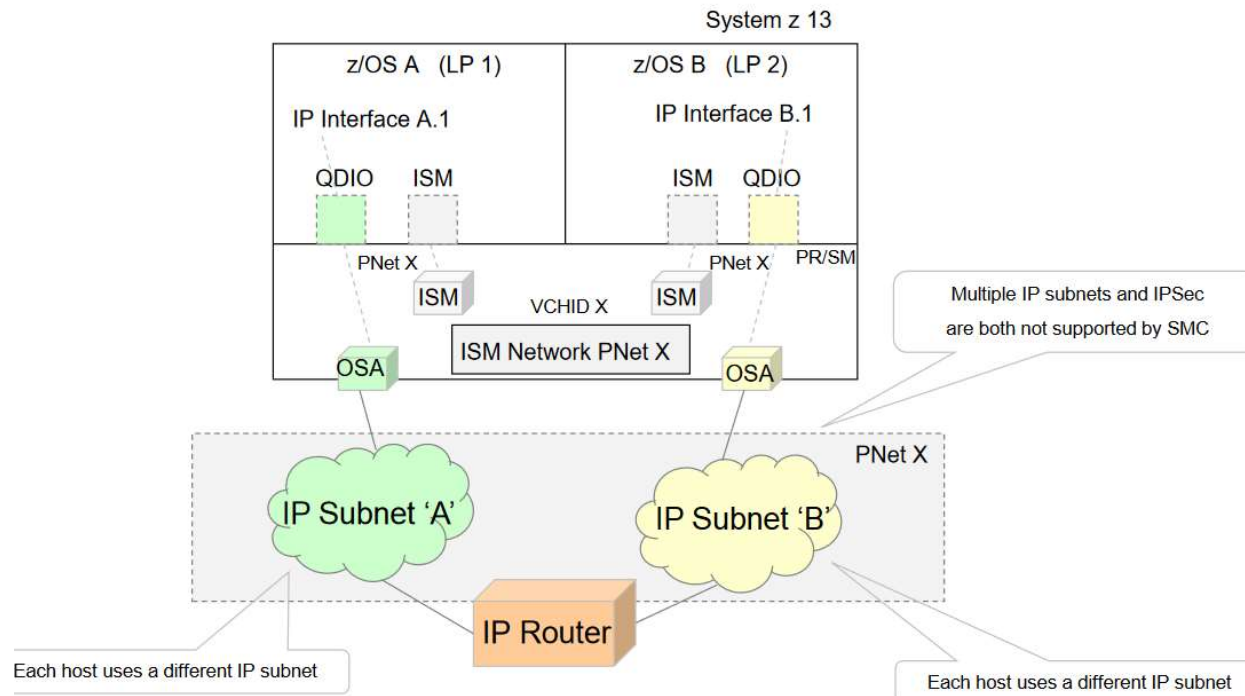


Config possibilities

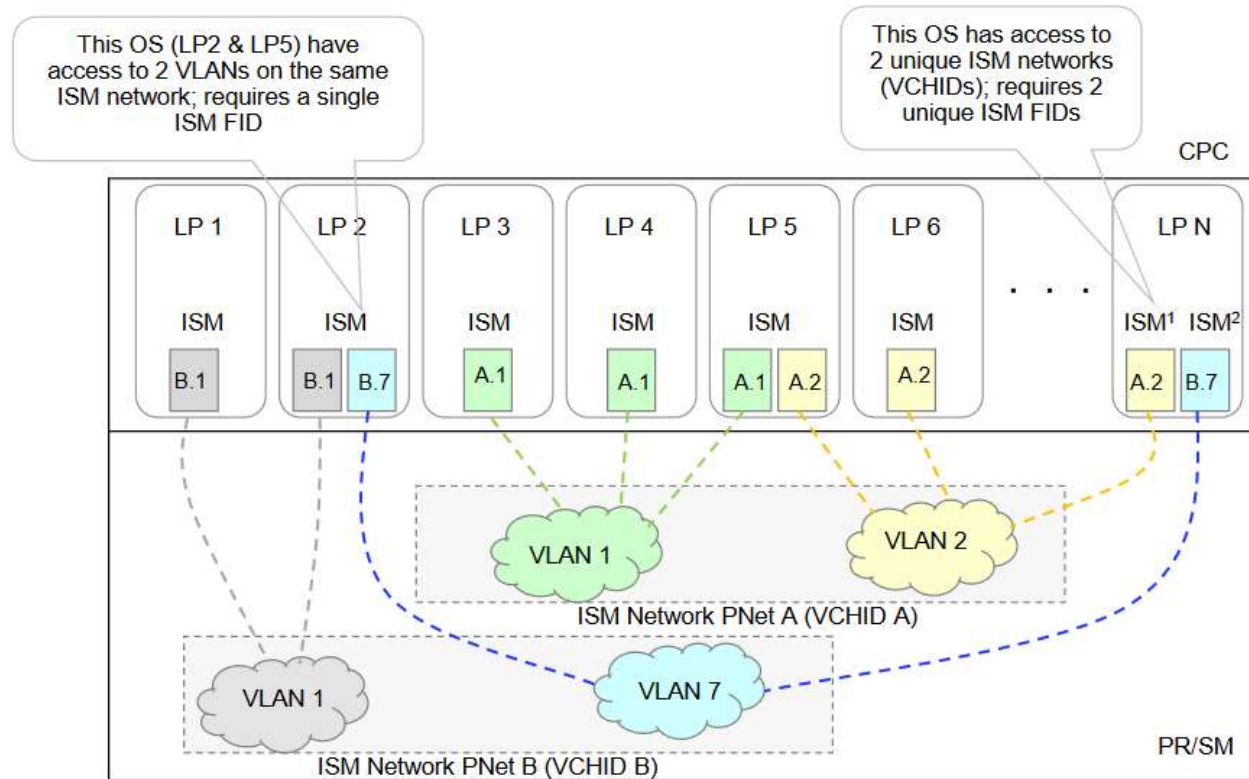


Multiple IP Subnets → Not supported

Peers must have direct connectivity over the same IP subnet to exploit SMC-R or SMC-D

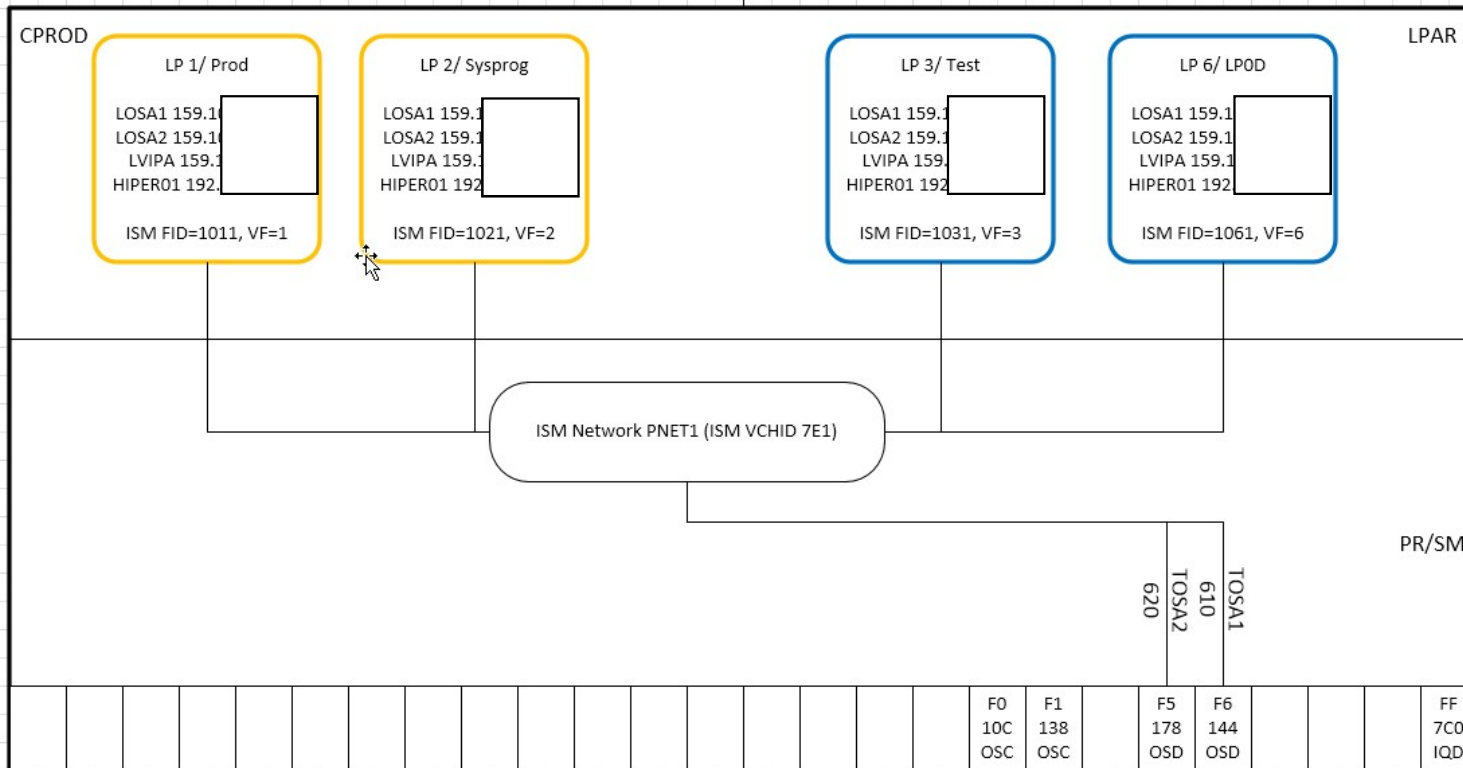


Isolating workloads



BJB Config

IP Network



HCD

Processor Support Levels

```
TIES2_2480.ecf - [24x80]
Datei Bearbeiten Ansicht Verbindung Dateiübertragung Einstellungen Fenster Hilfe
Change Processor Definition
TSTB
Available Support Levels
TSTB Command ==>>> _____ Row 1 of 3 More: >
Select the processor support level which provides the processor
capabilities you want to use.
Support Level
  2965 support, ISM, RCE, UID
- 2965 support, ISM, RCE
  2965 support, ISM, RCE, UID, CL5
***** Bottom of data *****
Local system name . . . . . CPROD
*CBDPPR1 ZSDSFOP2 -CMD      ISR@PRIM
4B C 13/05
Verbunden mit ties2.juliusbaer.com WAADGL00 (WAADGL00) UB
```

Functions List

```
TIES2_2480.ef - [24 x 80]
Datei Bearbeiten Ansicht Verbindung Dateiübertragung Einstellungen Fenster Hilfe
Goto Filter Backup Query Help
-----
TSTB                               PCIe Function List           Row 1 of 8 More: >
Command ==> _____ Scroll ==> CSR

Select one or more PCIe functions, then press Enter. To add, use F11.

Processor ID . . . . : CPR0D          DC2 z13s N10

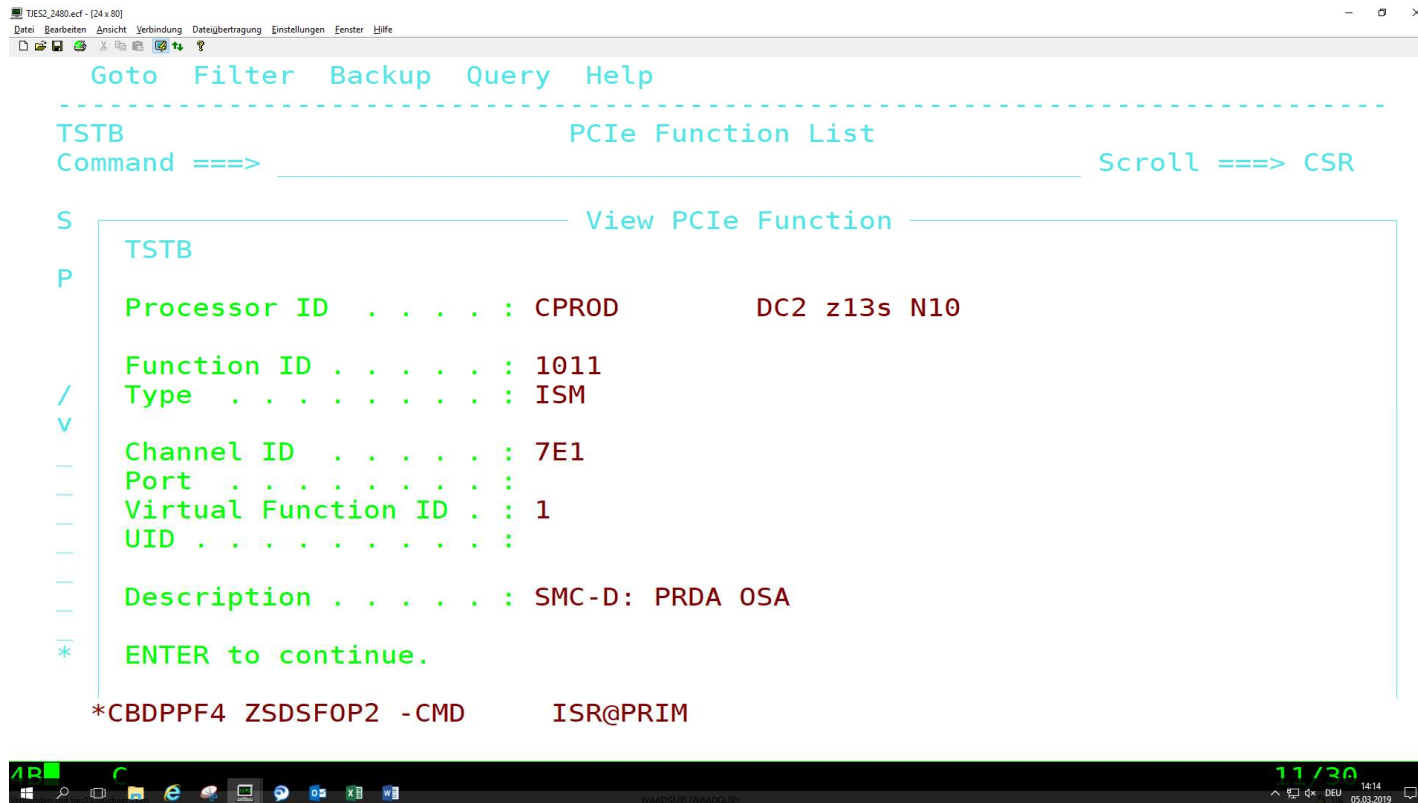
/ FID  CHID+  P+  VF+  Type+  UID  Description
- 1011  7E1     -   1   ISM    _____  SMC-D: PRDA OSA
- 1012  7E2     -   1   ISM    _____  SMC-D: PRDA HiperSocket
- 1021  7E1     -   2   ISM    _____  SMC-D: PRDC OSA
- 1022  7E2     -   2   ISM    _____  SMC-D: PRDC HiperSocket
- 1031  7E1     -   3   ISM    _____  SMC-D: TSTB OSA
- 1032  7E2     -   3   ISM    _____  SMC-D: TSTB HiperSocket
- 1061  7E1     -   6   ISM    _____  SMC-D: TSTD OSA
- 1062  7E2     -   6   ISM    _____  SMC-D: TSTD HiperSocket
***** Bottom of data *****

*CBDPPFF ZSDSFOP2 -CMD      ISR@PRIM
```



PCIe Function

```
TIES2_2480.ecf - [24x80]
Datei Bearbeiten Ansicht Verbindung Dateiübertragung Einstellungen Fenster Hilfe
Goto Filter Backup Query Help
-----
TSTB                               PCIe Function List
Command ==> _____ Scroll ==> CSR
S                                     View PCIe Function
P
TSTB
Processor ID . . . . . : CPROD          DC2 z13s N10
Function ID . . . . . : 1011
Type . . . . . : ISM
/
v
Channel ID . . . . . : 7E1
Port . . . . . :
Virtual Function ID . : 1
UID . . . . . :
Description . . . . . : SMC-D: PRDA OSA
*
ENTER to continue.
*CBDPPF4 ZSDSFOP2 -CMD          ISR@PRIM
```



ISM: Mapping PNET

```
TIES2_2480.ecf - [24x80]
Datei Bearbeiten Ansicht Verbindung Dateiübertragung Einstellungen Fenster Hilfe
Goto Filter Backup Query Help
- View Physical Network IDs
T
C
S
P
Physical network ID 1 . : PNET1
Physical network ID 2 . :
Physical network ID 3 . :
Physical network ID 4 . :

/
v 1011 7E1 - 1 ISM SMC-D: PRDA OSA
  1012 7E2 - 1 ISM SMC-D: PRDA HiperSocket
- 1021 7E1 - 2 ISM SMC-D: PRDC OSA
- 1022 7E2 - 2 ISM SMC-D: PRDC HiperSocket
- 1031 7E1 - 3 ISM SMC-D: TSTB OSA
- 1032 7E2 - 3 ISM SMC-D: TSTB HiperSocket
- 1061 7E1 - 6 ISM SMC-D: TSTD OSA
- 1062 7E2 - 6 ISM SMC-D: TSTD HiperSocket
***** Bottom of data *****

*CBDPPF4 ZSDSFOP2 -CMD ISR@PRIM

4B C 05/33
Verbunden mit ties2.juliusbaer.com WAADGL00 (WAADGL00) UB
```

OSA: Mapping PNET

```
TIES2_2480.ecf - [24x80]
Datei Bearbeiten Ansicht Verbindung Dateiübertragung Einstellungen Fenster Hilfe
Goto Filter Backup Query Help
- View Physical Network IDs
T
C
Physical network ID 1 . : PNET1
S
Physical network ID 2 . :
Physical network ID 3 . :
P
Physical network ID 4 . :
C
C

/ CHPID AID/P Type+ Mode+ Sw+ Sw Port Con Mng Description
v F5 178 OSD SHR ___ ___ ___ No OSA Express4S GbE LX
_ F6 144 OSD SHR ___ ___ ___ No OSA Express4S GbE LX
_ FF 7C0 IQD SHR ___ ___ ___ No Hipersocket
***** Bottom of data *****

* CBDPPF4 ZSDSFOP2 -CMD ISR@PRIM

4B C 05/33
Verbunden mit ties2.juliusbaer.com WAADGL00 (WAADGL00) UB
```

TCPIP Configuration

```
-----  
GLOBALCONFIG  
-----  
GLOBALCONFIG  
      SMCD  
; enable SMCD support - could be activated via obeyfile ***  
;                   - could be deactivated via obeyfile ***
```

Commands/ Debugging

-D PCIE

IQP022I 15.57.15 DISPLAY PCIE 924

PCIE 0011 ACTIVE

PFID	DEVICE	TYPE	NAME	STATUS	ASID	JOBNAME	CHID	VFN
00001031	ISM			ALLC	00B0	NET	07E1	0003
00001032	ISM			ALLC	00B0	NET	07E2	0003

-D PCIE,PFID=1031

IQP024I 16.03.21 DISPLAY PCIE 134

PCIE 0011 ACTIVE

PFID	DEVICE	TYPE	NAME	STATUS	ASID	JOBNAME	CHID	VFN	PN
00001031	ISM			ALLC	00B0	NET	07E1	0003	

CLIENT ASIDS: 26

PNetID 1: PNET1

Commands/ Debugging

```
-D TCPIP, ,N,DEV,PNETID=*
EZZ2500I NETSTAT CS V2R3 TCPIP 397
PNETID: PNET1
  INTFNAME: LOSA1           INTFTYPE: IPAQENET
  INTFNAME: LOSA2           INTFTYPE: IPAQENET
  INTFNAME: EZAISM02        INTFTYPE: ISM           ASSOCIATED: YES
PNETID: PNET2
  INTFNAME: HIPER01         INTFTYPE: IPAQIDIO
  INTFNAME: EZAISM01        INTFTYPE: ISM           ASSOCIATED: YES
5 OF 5 RECORDS DISPLAYED
```


OSA und PNET

```
-D M=CHP(F5)
IEE174I 16.05.05 DISPLAY M 617
CHPID F5: TYPE=11, DESC=OSA DIRECT EXPRESS, ONLINE
DEVICE STATUS FOR CHANNEL PATH F5
    0 1 2 3 4 5 6 7 8 9 A B C D E F
0062 + + + + + + + + + + + + + + +
SWITCH DEVICE NUMBER = NONE
PHYSICAL CHANNEL ID = 0178
PNETID 1 = PNET1
PNETID 2 =
PNETID 3 =
PNETID 4 =
```

OSA, ISM relation

INTFNAME: LOSA1 INTFTYPE: IPAQENET INTFSTATUS: READY
 PORTNAME: TOSA1 DATAPATH: 0612 DATAPATHSTATUS: READY
 CHPIDTYPE: OSD SMCR: DISABLED (GLOBALCONFIG NOSMCR)
 PNETID: PNET1 SMCD: YES
 TRLE: TRL0610

ASSOCIATED ISM INTERFACE: EZAISM02

INTFNAME: EZAISM02 INTFTYPE: ISM INTFSTATUS: READY
 PFID: 1031 TRLE: IUT01031 PFIDSTATUS: READY
 PNETID: PNET1
 GIDADDR: 28008483FF1E2965

INTERFACE STATISTICS:

BYTESIN	= 1278954854841
INBOUND OPERATIONS	= 8488263
BYTESOUT	= 69200915
OUTBOUND OPERATIONS	= 6552332
SMC LINKS	= 2
TCP CONNECTIONS	= 4
INTF RECEIVE BUFFER INUSE	= 256K
DEVICE INTERRUPTS	= 8476860

Performance

Scenario

FTP Transfer von 4GB Daten

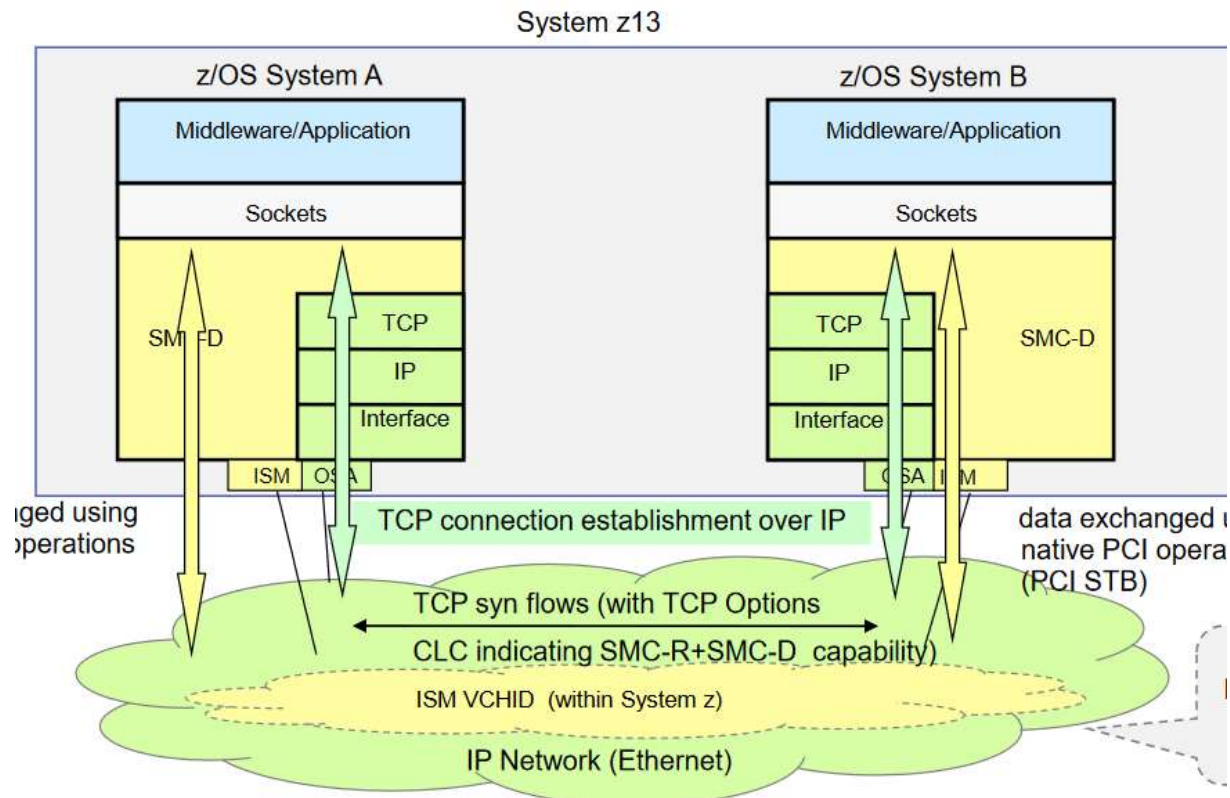
Transfer medium	Bandbreite (MB/s)	CPUs (User + TCPIP)
OSA	51	10.0 (4.44 + 5.56)
Hipersockets	123	10.4 (4.05 + 6.37)

Transfer medium	Bandbreite (MB/s)	CPUs (User + TCPIP)
SMC-D	155	3.36 (3.28 + 0.08)

Performance (2)

- Parallele FTP bleibt die Bandbreite erhalten
- 50% Reduzierung der täglichen FTP-Batch Zeiten!
- Limitation ist der Disk IO

CPU Reduction



Goals achieved

- Performance (Bandwidth)
 - About 15%. Bottleneck: Disk
- **Reduce CPU**
 - **98% CPU reduction in TCPIP**
- Security
- Umstellung auf 'Interface' Statements



Abkürzungen

- ISM Internal Shared Memory
- SMC-D Shared Memory Communications - Direct Memory Access

Links

- [ftp://public.dhe.ibm.com/software/systemz/pdf/IBM System z13 and z13s SMC-D and ISM Introduction with zOS Overview 06-30-16.pdf](ftp://public.dhe.ibm.com/software/systemz/pdf/IBM_System_z13_and_z13s_SMC-D_and_ISM_Introduction_with_zOS_Overview_06-30-16.pdf)