



IBM eServer zSeries

# G10 – SCSI mit VSE und Linux - Konzepte und Implementierung



# Trademarks

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

CICS*	IBM*	Virtual Image
DB2*	IBM logo*	Facility
DB2 Connect	IMS	VM/ESA*
DB2 Universal Database	Intelligent Miner	VSE/ESA
e-business logo*	Multiprise*	VisualAge*
Enterprise Storage Server	MQSeries*	WebSphere*
HiperSockets	OS/390*	xSeries
	S/390*	z/Architecture
	SNAP/SHOT	z/VM
	*	z/VSE
		zSeries

\* Registered trademarks of IBM Corporation

The following are trademarks or registered trademarks of other companies.

LINUX is a registered trademark of Linus Torvalds

Tivoli is a trademark of Tivoli Systems Inc.

Java and all Java-related trademarks and logos are trademarks of Sun Microsystems, Inc., in the United States and other countries

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

Microsoft, Windows and Windows NT are registered trademarks of Microsoft Corporation.

SET and Secure Electronic Transaction are trademarks owned by SET Secure Electronic Transaction LLC.

Intel is a registered trademark of Intel Corporation.

# Agenda

## § **SCSI Grundlagen**

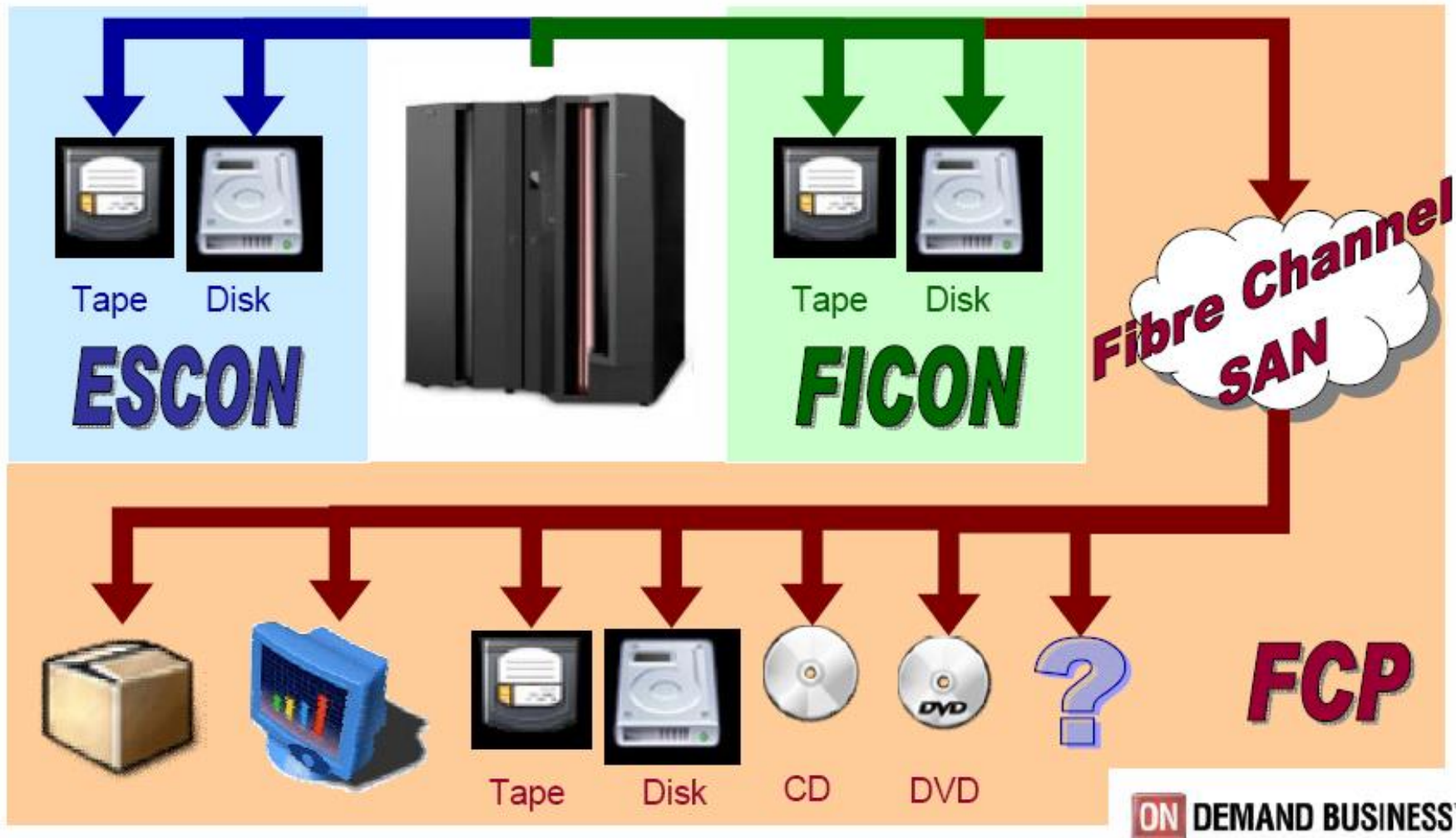
- Fibre Channel SAN
- SCSI Adressierung
- IPL von SCSI

## § **VSE und SCSI**

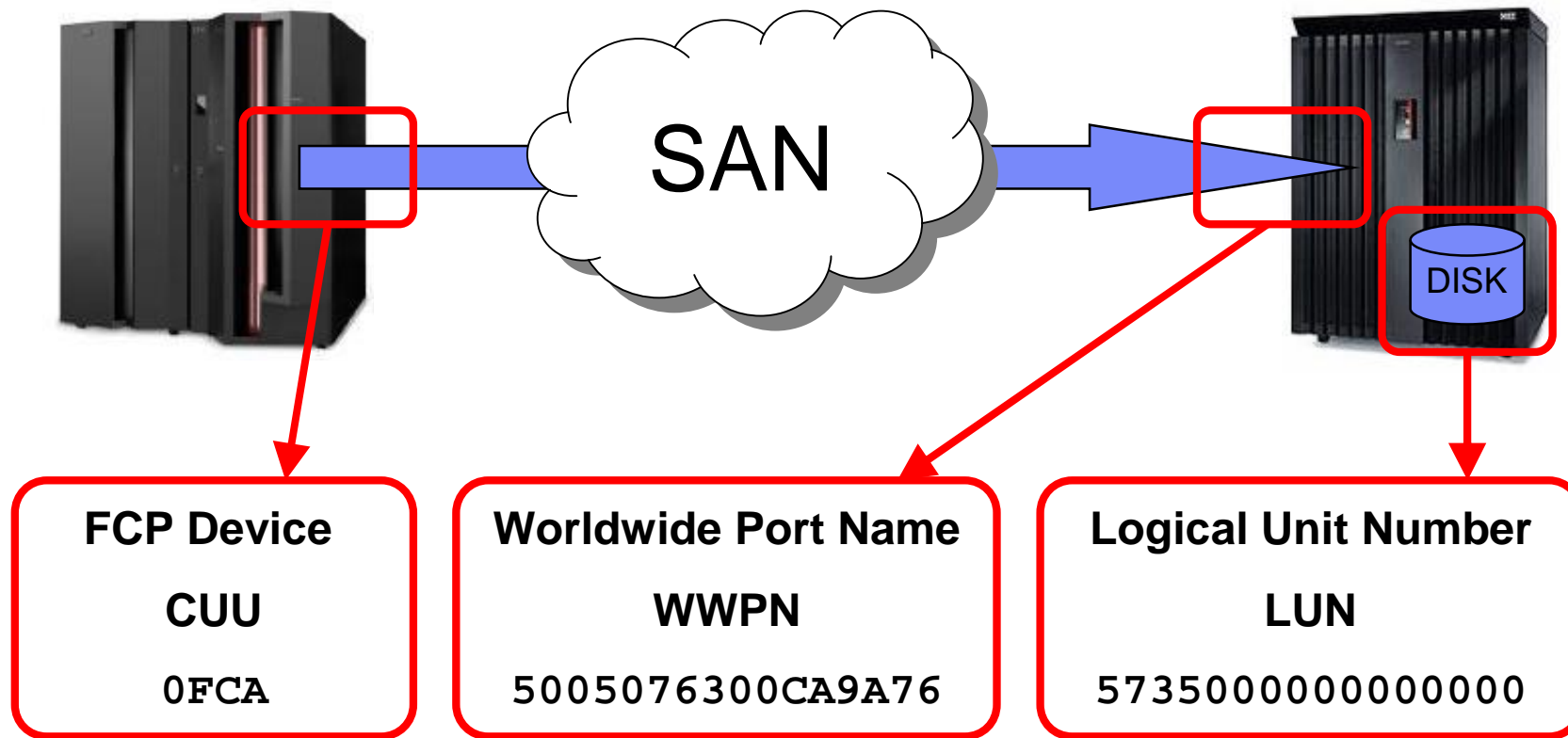
## § **Linux und SCSI**

## § **Zusammenfassung und Ausblick**

# Fibre Channel Storage Area Network (SAN)



# SCSI Adressierung



## SCSI Multipathing



### § 2 oder mehr Pfade zur selben Disk

- Bei Problemen mit dem verwendeten Pfad
  - Failover zu einem anderen Pfad
  - Failback sobald der erste Pfad wieder funktioniert
- Load-Balancing

## Sharing von SCSI Disks



- § Eine bestimmte LUN kann nur einmal pro FCP CHPID verwendet werden
- § ACT legt den „Owner“ der LUN fest (pro CHPID)
  - Zugriffe von anderen Systemen über die selbe CHPID sind nicht möglich
  - ACT kann derzeit nur über ein Linux-Tool definiert werden

	Exklusive LUN	Shared LUN
ohne ACT	„First come, first serve“	Nicht unterstützt
mit ACT	Deterministisch	Read-Only (erfordert Support von Operating System)

## IPL von SCSI

**Load**

CPC: P000F12B  
Image: ZFCP4

Load type:  Normal  Clear  SCSCI  SCSCI dump

Store status

Load address: 5C00

Load parameter:

Time-out value: 060 (60 to 600 seconds)

World wide port name: 5005076300CE93A7

Logical unit number: 5732000000000000

Boot program selector: 0

Boot record logical block address: 0000000000000000

OS specific load parameters:

OK Reset Cancel Help

### § Benützt den „Machine Loader“

- Software-Plattform unabhängiges Hardware-Tool

### § Native oder LPAR

- Load über Hardware Management Console (HMC)
  - Load Address = FCP Device
  - WWPN
  - LUN

### § Unter z/VM

- SET LOADDEV PORTNAME 50050763 00C29A76 LUN 56010000 00000000
- IPL cuu (FCP Device)



## Agenda - SCSI mit VSE

- § **Voraussetzungen**
- § **FBA Emulation**
- § **Setup + Definitionen**
- § **Kommandos + Dialoge + Messages**
- § **Installation auf SCSI**
- § **Migration von ECKD zu SCSI**

## Hardware und Software Voraussetzungen

§ **IBM eServer zSeries 800, 900, 890 oder 990**

§ **IBM zSeries FCP Adapter**

– Microcode Level:

- z800 und z900: J11233.015 oder höher
- z890 und z990: J13471.004 oder höher

§ **FCP Switch (z.B. IBM 2109)**

§ **IBM TotalStorage Enterprise Storage Server (ESS)**

– Microcode Level: 2.3.1 oder höher

§ **IBM TotalStorage DS6000 oder DS8000**

§ **z/VSE 3.1 oder höher**

§ **z/VM 4.4. oder höher (nur bei VSE unter VM)**

## Hardware und Software Voraussetzungen (2)

### § Für IPL von SCSI

- CPU Feature Code 9904
- z800 und z900:
  - Microcode Level EC-Number J12811 oder höher
- z890 und z990:
  - Microcode Level EC-Number J12221 oder höher

### § Für IPL von SCSI unter z/VM 4.4

- z/VM Service Level:
  - UM31181 (Englisch)
  - UM31180 (German)
  - UM31179 (Kanji)

### § Emulated FBA Disks:

- z/VM 5.1

## Charakteristiken einer SCSI Disk

### § Größe einer SCSI Disk

- Minimum 8 MB
- Maximum ca. 24 GB (begrenzt durch internes Model)
- 4 MB werden von z/VSE intern verwendet
- Nutzbare Größe = Größe – 4 MB
- VSAM kann nur die ersten 16 GB nutzen

### § Model

- SCSI Disks werden als FBA Devices definiert
  - 9336 Model 20

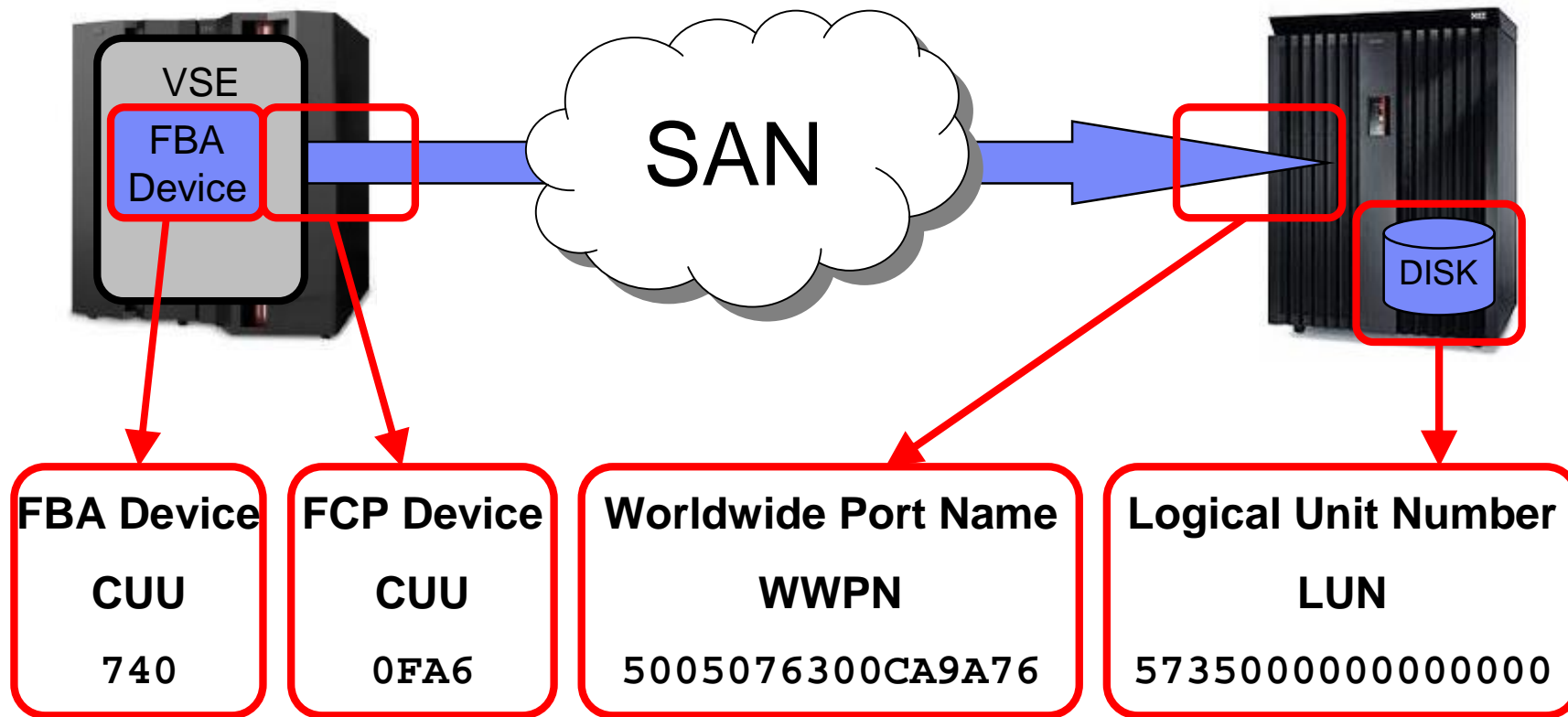
### § Block Größen

- z/VSE unterstützt nur SCSI Disks mit einer Block Größe von 512 Bytes

### § Standards

- SCSI Disks müssen ANSI SCSI Version 3 unterstützen

# SCSI Adressierung - VSE



## SCSI Setup für z/VSE



### § FCP Devices:

- ADD 4A7:4A9,FCP

### § FBA Devices:

- ADD 608:61B,FBA

### § Define SCSI:

- DEF SCSI,FBA=608,FCP=4A7,WWPN=5005076300CA9A76,LUN=5710
- DEF SCSI,FBA=609,FCP=4A7,WWPN=5005076300CA9A76,LUN=5711
- ...

### § IPL von SCSI (VM)

- Minimum 32M Memory
- SET LOADDEV PORT 50050763 00CE9A76 LUN 57350000 00000000
- IPL 4B8 (IPL von FCP device)

# SCSI Commands

## §SCSI Definition beim IPL:

```
DEF SCSI,FBA=cuu,FCP=cuu,WWPN=nnnnnnnnnnnnnnnnnn,LUN=nnnn
```

## §SCSI Definition im laufenden Betrieb:

```
SYSDEF SCSI,FBA=cuu,FCP=cuu,WWPN=nnnnnnnnnnnnnnnnnn,LUN=nnnn
```

- FBA Device und FCP Device müssen schon wärem IPL definiert worden sein

## §SCSI Definitionen löschen:

```
SYSDEF SCSI,DELETE,FBA=cuu,FCP=cuu,WWPN=nnnnnnnnnnnnnnnnnn,LUN=nnnn
```

## §Anzeigen der SCSI Definitionen:

```
QUERY SCSI
```

AR	FBA-CUU	FCP-CUU	WORLDWIDE PORTNAME	LOGICAL UNIT NUMBER
AR 0015	608	4A7	5005076300CA9A76	5710000000000000
AR 0015	609	4A7	5005076300CA9A76	5711000000000000
AR 0015	60A	4A7	5005076300CA9A76	5712000000000000
AR 0015	60B	4A7	5005076300CA9A76	5713000000000000
AR 0015	60C	4A7	5005076300CA9A76	5714000000000000
AR 0015	60D	4A8	5005076300CA9A76	5715000000000000

```
QUERY SCSI,608 (FBA Device)
```

AR	FBA-CUU	FCP-CUU	WORLDWIDE PORTNAME	LOGICAL UNIT NUMBER
AR 0015	608	4A7	5005076300CA9A76	5710000000000000

## Interactive Interface Dialoge

```

ADM$DSK2          HARDWARE CONFIGURATION: DISK LIST

Options:  2 = Alter device type code/mode              5 = Delete a disk
          3 = Specify Shared and/or Device Down by an 'X' in the appr. column
          8 = Specify DEF SCSI command

   OPT      ADDR   DEVICE  DEVICE-TYPE  DEVICE SPEC  SHARED  DEVICE
                   CODE      MODE              DOWN

   8         DA1   FBA-SCSI  FBA                -         -
   -
   -
   -
   -
   -
   -
   -
   -
   -
   -
   -
   -
   -

PF 1=HELP      2=REDISPLAY  3=END              5=PROCESS
    
```



## Interactive Interface Dialoge

```
TAS$ICME          HARDWARE CONFIGURATION AND IPL: DEF SCSI

Enter the required data and press ENTER.

FBA .....        DA1          cuu of the FBA-SCSI device
FCP .....        FA0          cuu of the FCP device
WWPN .....        5005076300CA9A76  World wide port name of the
                                remote controller
LUN .....        5600         Logical unit number of the SCSI

PF1=HELP          2=REDISPLAY  3=END
```

## Interactive Interface Dialoge

```
TAS$ICMD          HARDWARE CONFIGURATION AND IPL: DEF SCSI

Enter the required data and press ENTER.

OPTIONS: 1 = ADD          2 = ALTER
         5 = DELETE

  OPT   FBA      FCP      WWPN          LUN
  ---   ---     ---     ---             ---
  -     233     C01     5005076300C693CB  5176
  -     DA1     FA0     5005076300CA9A76  5600
  -
  -
  -
  -
  -
  -
  -
  -
  -
  -

PF1=HELP          2=REDISPLAY  3=END          5=PROCESS
```

## SCSI Messages

- § **0S45I SCSI DEVICE 618 CONSISTS OF 03906304 BLOCKS,  
03897432 BLOCKS ARE AVAILABLE, 680 BLOCKS ARE UNUSED**
  - Vielfaches von 777 Blocks
    - Internes Model: 1 „Cylinder“ = 777 Blocks
- § **0S40I SCSI PROCESSING EVENT: REASON=0060  
FUNCTION=INIT-SCSI FBA=609 FCP=4A7 WWPN=5005076300CA9A76  
LUN=5711000000000000**
  - Message-Beschreibung erklärt den Grund anhand des Reason Codes
- § **0S46I I/O ERROR ON FBA=600 FCP=C00 RC=01, REASON==052500**
  - SCSI I/O Error wurde auf eine FBA Error gemapped.
  - Siehe Message Beschreibung für den genauen Grund anhand des Reason Codes
- § **Generell werden SCSI I/O Fehler in S/390 I/O Fehler gemapped, die an das User Programm zurückgegeben werde (z.B. Unit Check)**
  - Zusätzlich kommt noch die 0S46I Message, um die genaue Ursache anzugeben

# SCSI Multipathing

## § Ein oder mehrere alternative Pfade zur selben SCSI Disk

- Erhöht die Verfügbarkeit
- Nicht: Workload-Balancing

## § Jeder Pfad muss über einen anderen FCP Adapter definiert sein

- Eine FCP Karte kann mehrere FCP Adapter (CHPID) haben
- Um die Ausfallsicherheit zu erhöhen, sollte man FCP Adapter auf verschiedenen physikalischen FCP Karten verwenden

## § Am besten auch über unterschiedliche Switches und/oder Ports

## § Beispiel:

```
DEF SCSI , FBA=DA1 , FCP=FA0 , WWPN=5005076300CA9A76 , LUN=5600
```

```
DEF SCSI , FBA=DA1 , FCP=FB0 , WWPN=5005076300C29A76 , LUN=5600
```

## § QUERY SCSI

```
AR 0015 FBA-CUU FCP-CUU WORLDWIDE PORTNAME LOGICAL UNIT NUMBER
```

```
AR 0015 DA1      FA0      5005076300CA9A76  5600000000000000
```

```
AR 0015 DA1      FB0      5005076300C29A76  5600000000000000
```

- Der zuerst gezeigte Pfad wird derzeit verwendet um auf die SCSI Platte zuzugreifen

## Sharing SCSI Disks

### § ADD cuu,FBA,SHR (FBA Device)

### § Lockfile wird benutzt (Siehe DLF)

- Per Reserve/Release SCSI Command (intern)
  - Reserve basierend auf dem FCP Adapter
  - Release muss von selbem FCP Adapter erfolgen
- Mögliches Problem:
  - Hardwait tritt auf, während die Disk reserved ist
    - Disk bleibt reserved
  - Es wird versucht die Disk im Hardwait-Processing zu releasen, das kann aber schief gehen

### § Restriktionen

- Lockfile darf nicht auf DOSRES/SYSWK1 liegen (bei SCSI)
- Kein Multipathing möglich für Lockfile-Disks
- Jedes VSE System, das auf die Lockfile zugreift, muss das über eine eigene FCP CHPID machen.

### § Empfehlung

- Separate Disk für das Lockfile

## Installation auf SCSI

### § Nur Base Installation möglich

- FSU von ECKD- auf SCSI-Disks nicht möglich

### § Automatische Installation

- IPL von Tape

```
BG 0000 SI70D IF YOU WANT TO USE SCSI DEVICES SPECIFY YES, ELSE NO
0 YES
BG 0000 SI75I ENTER SCSI COMMAND FOR DOSRES
BG 0000 SA80D SCSI,FBA=CUU,FCP=CUU,WWPN=PORTNAME,LUN=LUN
0 SCSI,FBA=608,FCP=C00,WWPN=5005076300C69A76,LUN=5745
AR 0033 0S45I SCSI DEVICE 608 CONSISTS OF 09765632 AVAILABLE, 651 BLOCKS ARE UNUSED
BG 0000 SA76I ENTER SCSI COMMAND FOR SYSWK1
BG 0000 SA80D SCSI,FBA=CUU,FCP=CUU,WWPN=PORTNAME,LUN=LUN
0 SCSI,FBA=609,FCP=D00,WWPN=5005076300C29A76,LUN=5746
AR 0033 0S45I SCSI DEVICE 609 CONSISTS OF 09765632 AVAILABLE, 651 BLOCKS ARE UNUSED
BG 0000 SI08I DOSRES IS 608, DEVICE TYPE FBA
BG 0000 SI09I SYSWK1 IS 609, DEVICE TYPE FBA
```

### § Hardware Configuration Dialog

- PF5 um die IPLPROC zu katalogisieren
  - Sonst wird beim nächsten IPL die SYSWK1 nicht gefunden

## Migration von ECKD nach SCSI

### § FSU von ECKD zu SCSI nicht möglich

- Base Installation nötig

### § Datei-Allokationen müssen angepasst werden

- Umrechnen von Tracks/Cylinder zu Blocks (für 3390):
  - 1 Track = ca. 112 Blocks
  - 1 Cylinder = ca. 1680 Blocks
- VSAM Space
  - Tipp: Cluster-Größe in RECORDS angeben, nicht in Tracks
- Sequentielle Dateien
- VSE Libraries
  - Tipp: 1 LIBR Block = 1024 Bytes = 2 SCSI Blocks

### § Programme müssen mit FBA Disks umgehen können

- Am besten Device-unabhängig implementiert

# Fragen ?

