



Workshop

LINUX

Host Anbindung ST6x40/ST25x0

Ralf Werner

Senior Storage Consultant

Sun Microsystems GmbH

Storage Group

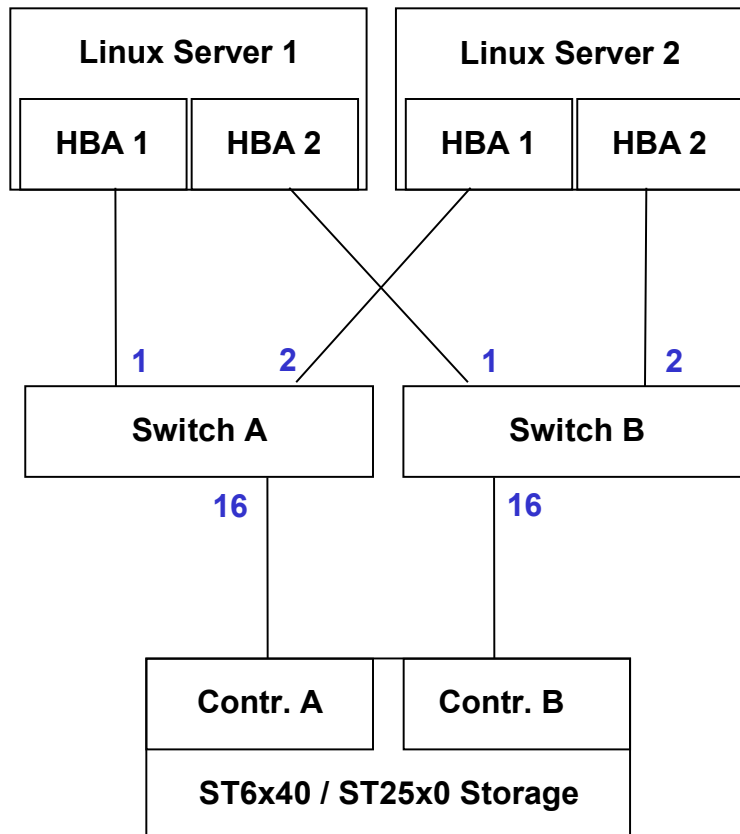
+49 173 6505175

Einführung

- **Die Konfigurationshinweise in dieser Präsentation gelten sowohl für ST6x40 als auch für ST25x0 Systeme**
 - > Die relevanten Eigenschaften sind bei beiden Systemen identisch implementiert
 - > Cache Mirroring, IO-Alignment, Stripe Sets, ...
- **Beschriebene Einstellungen basieren auf Vorgehensweisen und Werte, mit denen gute Erfahrungen gemacht wurde.**
 - > Andere Einstellungen möglich
 - > Abhängig von individueller Gesamtkonfiguration und Anforderungen

Zoning

- **WWPN- oder Port-Zoning**
 - > Switch nicht auf “Durchzug” stellen
 - > Dringend empfohlen: “Single-Initiator-Zoning”



Zonen

Switch A

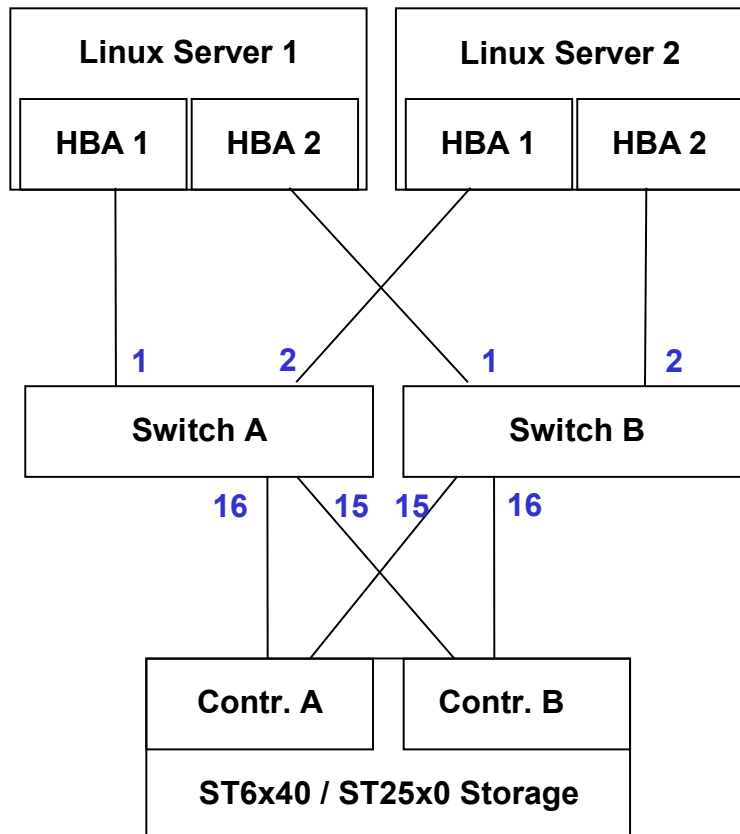
- > Zone 1: Port 1 & Port 16
- > Zone 4: Port 2 & Port 16

Switch B

- > Zone 1: Port 1 & Port 16
- > Zone 4: Port 2 & Port 16

Zoning

- **Optional: Zusätzliche “Kreuzverkabelung” zwischen Switch und Storage**
 - > Jedes Volume eines Servers über alle beide HBA's erreichbar
 - > Round-Robin Load-Balancing über HBA's zu einem Controller (mit MPP-Treiber)
 - > Kein Volume Transfer auf anderen Controller bei Switch Ausfall (mit MPP-Treiber)



Zonen

Switch A

- > Zone 1: Port 1 & Port 16
- > Zone 2: Port 2 & Port 16
- > Zone 3: Port 1 & Port 15
- > Zone 4: Port 2 & Port 15

Switch B

- > Zone 1: Port 1 & Port 16
- > Zone 2: Port 2 & Port 16
- > Zone 3: Port 1 & Port 15
- > Zone 4: Port 2 & Port 15

Multipathing

mppRDAC

- **Multi-Path-Proxy Treiber**
 - > Empfohlener Linux Multipath Treiber für ST6x40 & ST25x0 Systeme
 - > “Stateful” Treiber
 - > Load-Balancing über alle Pfade zum SELBEN Controller
 - > Failover zwischen Pfaden der beiden Controller
 - > Download vom Sun Download Center (<http://www.sun.com/download/index.jsp>)
 - > 09.00.A2.xx – Kernel 2.4 – SuSE SLES 8 und RedHat RHEL & AS 3 und AS 2.1
 - > 09.01.B2.xx – Kernel 2.6 – SuSE SLES 9 & 10 und RedHat RHEL 4
 - > 09.01.C2.xx – Kernel 2.6 – SuSE SLES 10/SP1 und RedHat RHEL 5
 - > Vor der Installation immer Readme.txt lesen
 - > Im Archiv des Treibers enthalten

mppRDAC Konfiguration

/etc/mpp.conf

- VirtualDiskProductId=StorageArray
- DebugLevel=0x0
- NotReadyWaitTime=270
- BusyWaitTime=270
- QuiescenceWaitTime=270
- InquiryWaitTime=60
- MaxLunsPerArray=256
- MaxPathsPerController=4
- ScanInterval=60
- InquiryInterval=1
- MaxArrayModules=30
- ErrorLevel=3
- SelectionTimeoutRetryCount=0
- UaRetryCount=10
- RetryCount=10
- SynchTimeout=170
- FailOverQuiescenceTime=20
- FailoverTimeout=120
- FailBackToCurrentAllowed=1
- ControllerIoWaitTime=300
- ArrayIoWaitTime=600
- **DisableLUNRebalance=0**
- IdlePathCheckingInterval=60
- RecheckFailedPathWaitTime=30
- FailedPathCheckingInterval=60
- ArrayFailoverWaitTime=300
- PrintSenseBuffer=0

Im Cluster den Auto-Failback der FC-Pfade deaktivieren, um einen "Ping-Pong-Effekt" zu verhindern. Hierfür den Parameter auf `DisableLUNRebalance=1` ändern. Anschließend einen "mppUpdate" ausführen.

Persistent Device Naming

- **Verhindert eine Änderung von Device Namen der LUN's nach einem Reboot**
- **Empfohlen für mppRDAC Multipath Treiber**

- **Utility “devlabel” für Kernel 2.4**
 - > <http://www.redhat.com/docs/manuals/enterprise/RHEL-3-Manual/sysadmin-guide/ch-devlabel.html>
 - > http://kbase.redhat.com/faq/FAQ_80_4625.shtm
- **Utility “udev” für Kernel 2.6**
 - > http://kbase.redhat.com/faq/FAQ_85_8082.shtm
 - > <http://www.redhat.com/docs/manuals/enterprise/RHEL-5-manual/Virtualization-en-US/ch-virt-lun-persistence.html>
 - > <http://www.kernel.org/pub/linux/utils/kernel/hotplug/udev.html>

Multipathing

Symantec DMP

- **Dynamic MultiPathing Treiber**
 - > Bestandteil der Storage Foundation
 - > “Stateless” Treiber
 - > Benötigt Auto-Volume-Transfer auf den ST6x40 Systemen
 - > NVSRAM Parameter
 - > In Firmware 7.10 separater Host Typ für Linux mit DMP
 - AVT aktiviert
 - > Firmware 6.xx hat nur Host Typ Linux
 - AVT nicht aktiviert
 - Host Typ muss ggf. angepasst werden (nicht supported!!!)
 - > Benötigt ASL (Array Specific Library) für Speichersystem von Symantec
 - > ASL's in der Storage Foundation enthalten
 - > Symantec Homepage auf aktualisierte Versionen überprüfen
 - > Zur Zeit nicht mit ST2540 Systemen unterstützt
 - > Sun Interop Website überprüfen
 - <https://extranet.stortek.com/interop/interop>

Linux Disks

Partitionen

- **Default: Linux legt Start Partitions auf zweiten Track (Track 1)**
 - > Track 1 kann bei Sektor 32 oder 63 beginnen
 - > Startsektor abhängig von Größe des Tracks
 - > Größe des Tracks abhängig von Größe des Volumes
 - > Dies kann zu Mis-alignment mit der Segment Size der ST6x40 LUN führen
 - > Sektor für den Start der Partition muss ggf. angepasst werden
 - > Partition sollte mindestens an einem Segment der LUN beginnen
 - > Besser noch, die Partition beginnt an einem Stripe Set
 - > Ideal für Full Stripe Writes
 - > Gut für große Blöcke
 - > Gute für Oracle ASM

Linux Disks

Beispiel: Default Partitionierung mit Mis-aligned IO's

fdisk /dev/sdc

Command (m for help) **p**
 Disk /dev/sdc: 513.5 GB, 513556348928 bytes
 255 heads, 63 sectors/track, 62436 cylinders
 Units = cylinders of 16065 * 512 = 8225280 bytes

Print Partition Table

Device	Boot	Start	End	Blocks	Id	System
/dev/sdc1		1	62436	501517138+	83	Linux

Anzeige suggeriert, das Partition am Cylinder 1 (zweiter Cylinder) beginnt.

Command (m for help) **u**
 Changing display/entry units to sectors

Anzeigeeinheiten auf Sektoren umschalten

Command (m for help): **p**

Disk /dev/sdc: 513.5 GB, 513556348928 bytes
 255 heads, 63 sectors/track, 62436 cylinders, total 1003039744 sectors
 Units = sectors of 1 * 512 = 512 bytes

Nach Umschalten der Units auf Sektoren:
 Sektor beginnt tatsächlich auf Sektor 63. Das passt zu keiner möglichen Segment Size.

Device	Boot	Start	End	Blocks	Id	System
/dev/sdc1		63	1003034339	501517138+	83	Linux

IO-Alignment

- **Partition muss an einem Segment oder Stripe Set beginnen**
- **Tabelle gibt Startsektoren für verschiedene Raid-5 Konfigurationen an**
 - > Startsektor für Segment alignment
 - > = Segment Size * 2
 - > Startsektor für Stripe Set alignment
 - > = Segment Size in kB * 2 * (Anzahl Disks je Vdisk - 1)

Anzahl Disks je Vdisk	RAID-Level	ST6x40 LUN Segment Size	Startsektor Segment aligned	Startsektor Stripe Set aligned
5	5	16kB	32	128
5	5	32kB	64	256
5	5	64kB	128	512
5	5	128kB	256	1024
5	5	256kB	512	2048
5	5	512kB	1024	4096
9	5	16kB	32	256
9	5	32kB	64	512
9	5	64kB	128	1024
9	5	128kB	256	2048
9	5	256kB	512	4096
9	5	512kB	1024	8192

Linux Disks

Partition anlegen

fdisk /dev/sdc

Command (m for help): **u**

Changing display/entry units to sectors

Command (m for help): **p**

Disk /dev/sdc: 143.4 GB, 143457779712 bytes

255 heads, 63 sectors/track, 17441 cylinders, total 280190976 sectors

Units = sectors of 1 * 512 = 512 bytes

Device	Boot	Start	End	Blocks	Id	System
--------	------	-------	-----	--------	----	--------

Command (m for help): **n**

Command action

e extended

p primary partition (1-4) **p**

Partition number (1-4): **1**

First sector (63-280190975, default 63) **8192**

Last sector or +size or +sizeM or +sizeK (32768-280190975, default 280190975):

Using default value 280190975

Command (m for help): **w**

Startsektor auf 4 MB gesetzt (8192 * 512 Byte).
Denn Wert ggf. entsprechend der vorherigen Folie anpassen.

Linux Disks

Beispiel: Partitionierung mit IO-Alignment

fdisk /dev/sdc

Command (m for help): **u**

Changing display/entry units to sectors

Command (m for help): **p**

Disk /dev/sdc: 143.4 GB, 143457779712 bytes

255 heads, 63 sectors/track, 17441 cylinders, total 280190976 sectors

Units = sectors of 1 * 512 = 512 bytes

Device	Boot	Start	End	Blocks	Id	System
/dev/sdc1		8192	280190975	140091392	83	Linux

Command (m for help): **w**

The partition table has been altered!

Calling ioctl() to re-read partition table.

Syncing disks.



Q & A

Danke!



THE MATERIAL CONTAINED WITHIN THIS PRESENTATION MAY NOT BE ALTERED OR DUPLICATED IN ANY WAY WITHOUT THE EXPRESS AUTHORISATION OF THE AUTHOR.

MENTION OF THIRD-PARTY PRODUCTS IS FOR INFORMATION PURPOSES ONLY AND SUN ACCEPTS NO LIABILITY FOR THEIR SELECTION OR PERFORMANCE.