

Storagetechnologien im Vergleich

Werner Fischer
Technology Specialist Thomas-Krenn.AG
wfisher@thomas-krenn.com

Thomas-Krenn Trainings-Road-Show November 2009
Berlin – Köln – Frankfurt – Stuttgart – Wien



Thomas-Krenn.AG[®]
Speed is (y)our success



Agenda



- 1) SAN oder DAS (Direct Attached Storage)?**
- 2) Performanceanforderungen**
- 3) Storage in virtualisierten Umgebungen**
- 4) Neue Storage Technologien**
 - SAS/SATA Expander Backplane
 - Adaptec ZMCP (Zero-Maintenance Cache Protection)
 - Solid State Disks
 - Adaptec MaxIQ SSD Cache



1) SAN oder DAS?



- **DAS (Direct Attached Storage) Schnittstellen**

- (IDE/ATA)

- (SCSI)

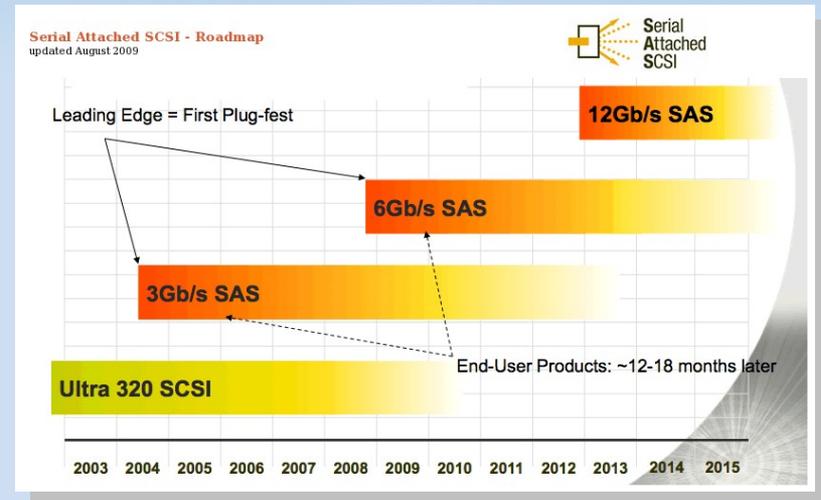
- SATA

- Serial ATA 1.5 Gbit/s
- Serial ATA 3.0 Gbit/s
- Serial ATA 6.0 Gbit/s

- SAS

- 3 Gbit/s SAS
- 6 Gbit/s SAS (derzeit in Entwicklung)
- 12 Gbit/s SAS (voraussichtlich ab ca. 2013)

- Fibre Channel



Quelle: http://www.scsita.org/aboutscsi/sas/SAS_roadmap2004.html



1) SAN oder DAS?



- **SAN (Storage Area Network) Technologien**
 - Fibre Channel (FC)
 - iSCSI
 - teilweise SAS (z.B. beim Modular Server)
- **Hauptunterschiede zu DAS**
 - zentraler Datenspeicher von mehreren Servern
 - effizientere Ausnutzung der Speicherkapazität
 - ermöglicht Cluster-Systeme ohne Datenreplizierung
 - zuverlässige Storage-Anbindung erfordert Multipathing:
 - Datenpfad geht im Gegensatz zu DAS über Server-externe Verbindungen,
 - oft sind zusätzliche Geräte (Switches) im Datenpfad
 - Anmerkung: HDDs selbst werden über die gleichen Protokolle wie bei DAS angesprochen



1) SAN oder DAS?



- **Fibre Channel**

- typische Datenübertragungsraten:

- 1 Gbit/s
- 2 Gbit/s
- 4 Gbit/s
- 8 Gbit/s

- Server-seitig HBAs (Host Bus Adapter)

- **dual fabric Konfiguration**

- nutzt zwei FC Switche
(diese sind NICHT miteinander verbunden)

- single HBA Zoning



QLogic HBA QLE2562 (dual port)



QLogic FC Switch 5802V



1) SAN oder DAS?



- **iSCSI**

- iSCSI Initiator

- Hardware Initiator (z.B. Qlogic QLA4052C)
 - Software Initiator (verwendet normale NICs)

- iSCSI Target

- Hardware Target (z.B. EMC AX4-5i oder IBM DS3300)
 - Software Target (z.B. Linux iSCSI Enterprise Target)



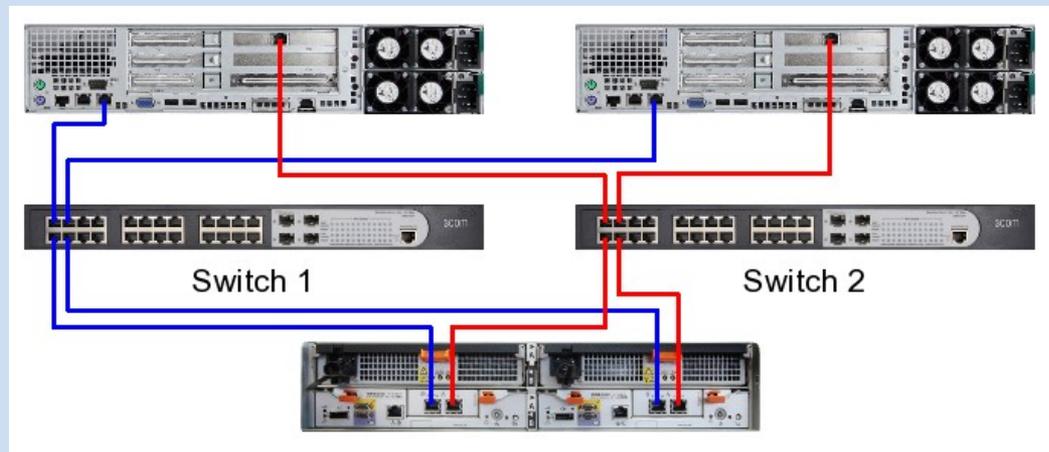
iSCSI Hardware Target EMC AX4-5i dual controller

1) SAN oder DAS?



- **iSCSI Multipathing**

- über allgemeine Multipathing-Mechanismen (z.B. Linux Device Mapper, Vmware NMP – Native Multipathing Plug-in)
- über iSCSI MC/S (Multiple Connections per Session) wird nicht von jedem Gerät unterstützt



iSCSI Multipath Setup mit 2 unterschiedlichen IP Netzen



1) SAN oder DAS?



- **Vergleich Fibre Channel / iSCSI**

Merkmale	Fibre Channel	iSCSI
Kosten	hoch (HBAs, FC-Switches)	mittel bis niedrig
Know-How	eigenes FC Know-How notwendig	Netzwerk Know-How anwendbar
Performance	sehr hoch (geringer Protokolloverhead)	mittel (höherer Protokolloverhead)
Beispiel 1 GBit/s	bringt 100 MB/sec	bringt ca. 75 MB/sec
aktuell typische Schnittstellen	4 GBit/sec	1 GBit/sec
Distanz	niedrig bis mittel	hoch (IP Netz)



2) Performanceanforderungen



- **random access Zugriffe: mittlere Zugriffszeit**
 - setzt sich zusammen aus mittlerer:
 - Spurwechselzeit (seek time), bei 2,5“ geringer als bei 3,5“
 - Latenzzeit (latency), Dauer einer halben Umdrehung
 - Kommando-Latenz (controller overhead)

HDD	RPM	seek time*	latency	controller overhead**	mittlere Zugriffszeit
WD 500 GB WD5002ABYS 3,5“	7.200	8,9 ms	4,2 ms	0,2 ms	13,3 ms
Seagate 400 GB ST3400755SS 3,5“	10.000	4,2 ms	3,0 ms	0,2 ms	7,4 ms
Fujitsu 300 GB MBA3300RC 3,5“	15.000	3,6 ms	2,0 ms	0,2 ms	5,8 ms

*) write seek time, Angabe laut Hersteller

***) Beispielannahme



2) Performanceanforderungen



- **RAID Levels**

Merkmale	RAID 0	RAID 1	RAID 5	RAID 6	RAID 10*
Mindestanzahl an HDDs	2	2	3	4	4
Datensicherheit	keine	Ausfall einer HDD	Ausfall einer HDD	Ausfall v. 2 HDDs	Ausfall v. 1 HDD pro Subarray
Kapazitätsausnutzung	100%	50%	67-94%	50-88%	50%
Typische Anwendung	High Perf.	OS DB	Fileserver	Fileserver Backup	DB

*) RAID 0 über mehrere RAID 1



3) Storage in virtualisierten Umgebungen



- **Beispiel XenServer**

- Begriffe

- Storage Repositories (SRs)
 - Virtual Disk Images (VDIs)

- Mögliche Zuordnungen von physischen Storage <-> VDIs

- Datei-basierte VHD (Virtual Hard Disk) in einem lokalen Filesystem (ext3) oder shared NFS Filesystem
 - Logical Volume based VHD auf einem LUN (lokales LVM, shared Fibre Channel, iSCSI, SAS)
 - LUN per VDI (mit Array-spezifischen Plugins, z.B. bei NetApp, Equallogic, StorageLink)



3) Storage in virtualisierten Umgebungen



- **Beispiel VMware:**
 - direktes Durchreichen von Blockdevices (Raw Device Mapping)
 - virtuelle Festplatte in einer Datei
 - thick provisioned disk
 - thin provisioned disk (neu mit vSphere 4)
 - Storagetypen
 - Fibre Channel (VMFS)
 - iSCSI (VMFS)
 - NFS
- **allgemein zusätzliche Möglichkeiten**
 - Nutzung eines iSCSI Software Initiators in einer virtuellen Maschine



4) NT: SAS/SATA Expander Backplane



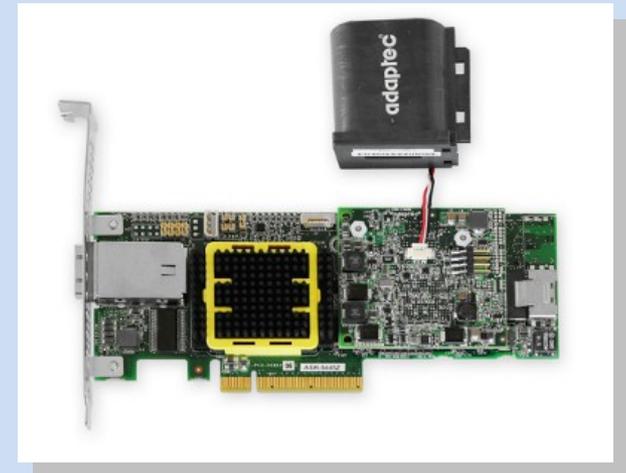
- **erlaubt den Anschluss von**
 - bis zu 24 Festplatten (SC846) an einem 4-Port-Controller
 - bis zu 256 Festplatten an einem 8-Port-Controller (4 intern, 4 extern) bei Verwendung von zusätzlichen JBODs
- **Beispiel SC846: ein internes Kabel reicht für den Anschluss aller 24 Festplatten**
- **Vorteile:**
 - kostengünstiger
 - geringere Fehleranfälligkeit
 - bessere Kühleigenschaften (Luftstrom)
 - einfache Erweiterbarkeit



4) NT: Adaptec Serie 5Z RAID Controller



- **gleicher Funktionsumfang wie Serie 5**
 - RAID 0, 1, 1E, 5, 5EE, 6, 10, 50, 60, JBOD
- **ZMCP (Zero-Maintenance Cache Protection)**
 - Doppelschicht-Kondensatoren
 - 4 GB Flash Speicher
- **Vorteile:**
 - wartungsfrei (kein Erneuern der BBU)
 - dauerhafter Cache-Schutz (BBU Akkus halten Daten max. 72 Stunden)
 - geringere Abwärme im Vergleich zum Akku einer BBU



4) NT: Intel und ATP SSD's



- **Varianten:**

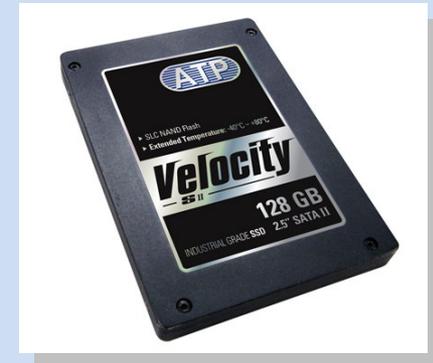
- SLC (single level cell): eine Flash-Zelle speichert ein Bit
- MLC (multi level cell): eine Flash-Zelle speichert zwei oder vier Bit, etwas geringere Performance / Anzahl Schreibzyklen, aber höhere Kapazitäten

- **Intel**

- X25-E (SLC), **32/64 GB**
- X18-M/X25-M (MLC), **80/160 GB**

- **ATP**

- Velocity SII (SLC), **32/64/128 GB**
- Velocity MII (MLC), **16/32/64/256 GB**



4) NT: Adaptec MaxiQ SSD Cache



- **32 GB Intel SLC SSD als Lese-Cache**
- **Learned-Path Algorithmus erkennt häufig gelesene Daten und kopiert diese auf die SSD**
- **drastische Verbesserung der Lese-Performance**
- **Cache bleibt auch bei Reboot bzw. Power-Off/Power-On erhalten**
- **mit folgenden RAID-Controllern nutzbar (FW 17517)**
 - Adaptec Serie 2
 - Adaptec Serie 5
 - Adaptec Serie 5Z
- **(derzeit bei Thomas Krenn noch im Test)**



Fazit



- laufend viele neue Technologien
- über 20 Techniker beschäftigen sich bei Thomas Krenn mit diesen Innovationen
- wir stellen Ihnen unsere Erfahrungen direkt zur Verfügung:

thomas-krenn.com/wiki

The screenshot shows the main page of the Thomas Krenn Wiki. The browser window title is 'Wiki Thomas-Krenn.AG - Mozilla Firefox'. The address bar shows 'http://www.thomas-krenn.com/de/wiki/Hauptseite'. The page features a navigation menu on the left with sections like 'navigation', 'Kategorien', 'Suche', and 'werkzeuge'. The main content area is titled 'Willkommen beim Thomas Krenn Wiki' and contains a grid of server-related categories such as 'Server-Hardware', 'Server-Software', 'Virtualisierung', 'Storage', and 'Netzwerk+Zubehör'. The right sidebar contains sections for 'Neueste Artikel' and 'Beliebteste Artikel'.

