

Thomas-Krenn.AG und NexentaStor™

HERBSTWORKSHOP 2010 DER THOMAS-KRENN.AG



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



Thomas-Krenn.AG und NexentaStor™

OPEN STORAGE FOR THE 21ST CENTURY



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



Die Herausforderungen im Storage Bereich

It is all about having the right data at the right time at the right place.

Zentralen Storage Herausforderungen

- Datenwachstum
- Datenintegrität
- Flexibilität

Herausforderung: Datenwachstum

- Bekannt: Menge gespeicherter Daten wächst
- Bestehende Technologien können teilweise nicht mit dem Wachstum mithalten
 - File-System Größen
 - Backup- und Restore Zeiten
 - Fehlerraten von Raid5 (und Raid6)
 - Geschwindigkeit skaliert nicht mit Datenmenge (Mehr Daten bei weniger Spindeln, siehe)

Herausforderung: Datenintegrität

- Datenwachstum erzeugt Probleme für Datenintegrität
 - Bitfehlerwahrscheinlichkeit bei X TB Disks relevant
 - LBE (Latent Block Errors) häufiger als erwartet (CERN Studie: 1 von 1500 Dateien korrupt)
- Fehlererkennung und Fehlerkorrektur notwendig

* See <http://storagemojo.com/2007/09/19/cerns-data-corruption-research/>

* See <http://indico.cern.ch/getFile.py/access?contribId=3&sessionId=0&resId=1&materialId=paper&confId=13797>

Herausforderung: Flexibilität

- Flexible Konnektivität
 - Frontend (z.B. FCoE)
 - Backend (z.B. SAS2)
- Neue Speichertechnologien (SSD)
- Workload Mix in Virtualisierten Umgebungen
- Performance Anforderungen wachsen dynamisch

... Storage Systeme müssen **agil** sein

Open Storage by NexentaStor™

... a new way of thinking of storage.

NexentaStor™ Open Storage ?

- **Open Storage = Commodity Hardware + clevere offene Software NexentaStor™**
- Kein Vendor lock-in (offene Disk Formate)
- Entwicklungsgeschwindigkeit und TCO von Commodity Hardware
- Nutzung neuer Features durch einfaches Update von Software ohne neue Hardware möglich
- **man ist wieder Herr seiner Daten**

Hands on ...



Wie NexentaStor™ Open Storage hilft

... the secret of getting ahead is getting started.

Open Storage Lösung: Datenwachstum

- 128 Bit File-System **ZFS**
- Keine künstlichen Limits
 - Kein (praktisches) Datenlimit
 - Unbegrenzte Snapshots
 - Unbegrenzte Clones
 - (praktisch) unbegrenzte Disks
 - 256 Targets mit jeweils 16384 LUN's = 4194304
- Skalierbarkeit durch einfaches Hinzufügen von Commodity Hardware (HBA, Enclosures, Disks)

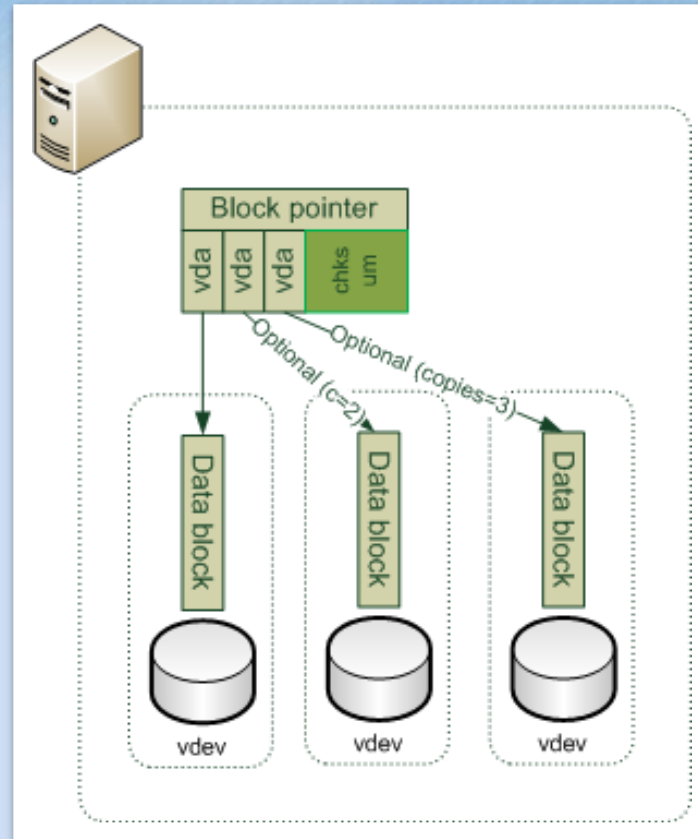
Open Storage Lösung: Datenintegrität

- End-to-End Checksumming aller Daten
- Fehlererkennung und Fehlerkorrektur
 - Auch für Fehler in passiven Daten (scrubbing)
 - Snapshots für Versionierung
- Redundante Speicherung der Daten ohne Hardware Raid
- Transaktionelle Spiegelung der Daten an zweiten Standort möglich

Beispiel

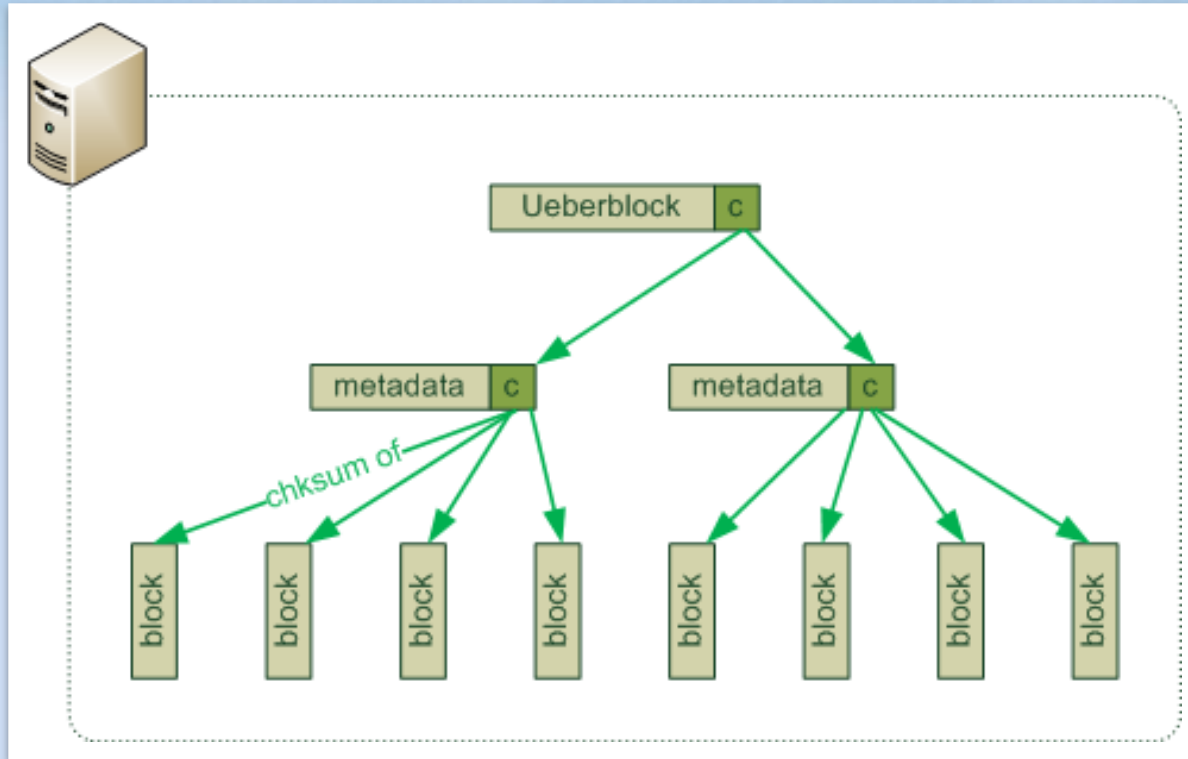
End to End Checksumming / Self Healing

Datenintegrität: Self Healing



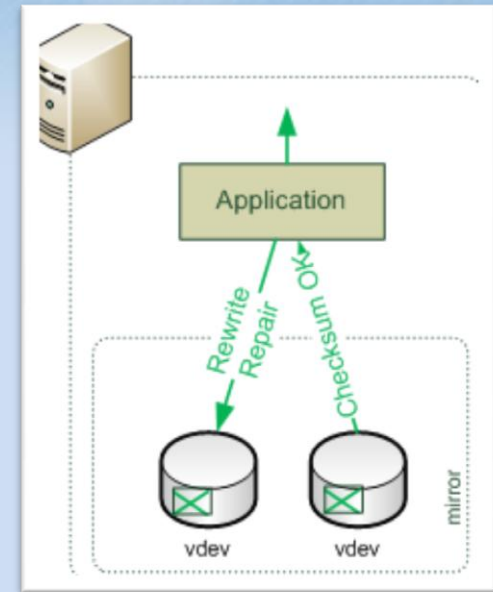
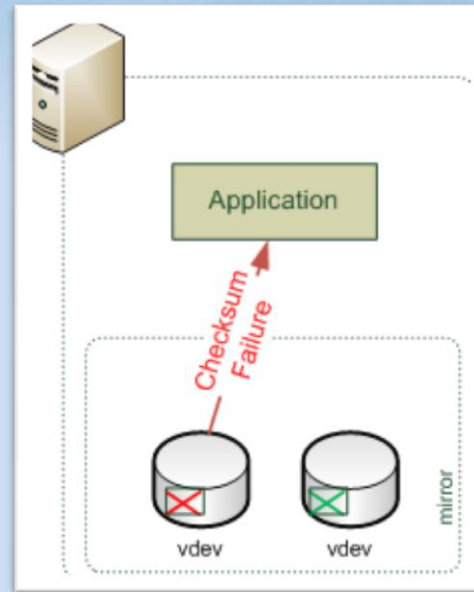
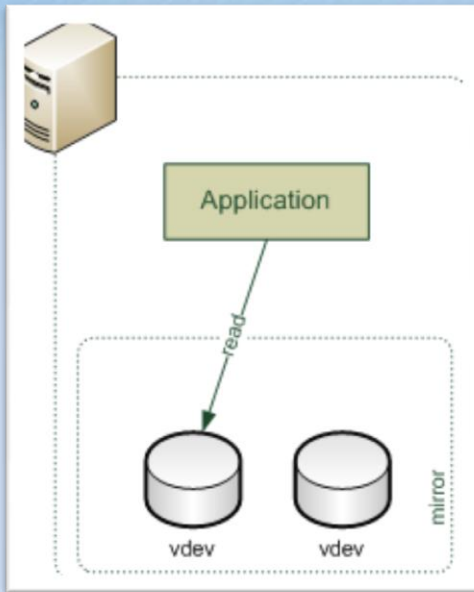
- Bis zu drei Kopien der Datenblöcke speicherbar
- Metadaten immer mindestens 2-mal vorhanden

Datenintegrität: Self Healing



- Alle Daten in ZFS sind mit Checksummen versehen

Datenintegrität: Self Healing



- Prüfung der Checksummen bei jedem Lesen
- Bei fehlerhaften Blöcken, nutzen von alternativen Blöcken und Reparatur der fehlerhaften Blöcke
- Sowohl für File- als auch für Block-Storage

Hands on ...



Open Storage Lösung: Flexibilität

- Unified: SAN-Storage und NAS-Filer
- Heterogene Technologien möglich
 - Frontend (FC, iSCSI, NAS)
 - Backend (FC, iSCSI, SAS, NAS)
- Hybrid Storage Pools: Effektives Caching durch SSD beschleunigten Speicher
- Die Lösung wächst mit den Anforderungen

Beispiel

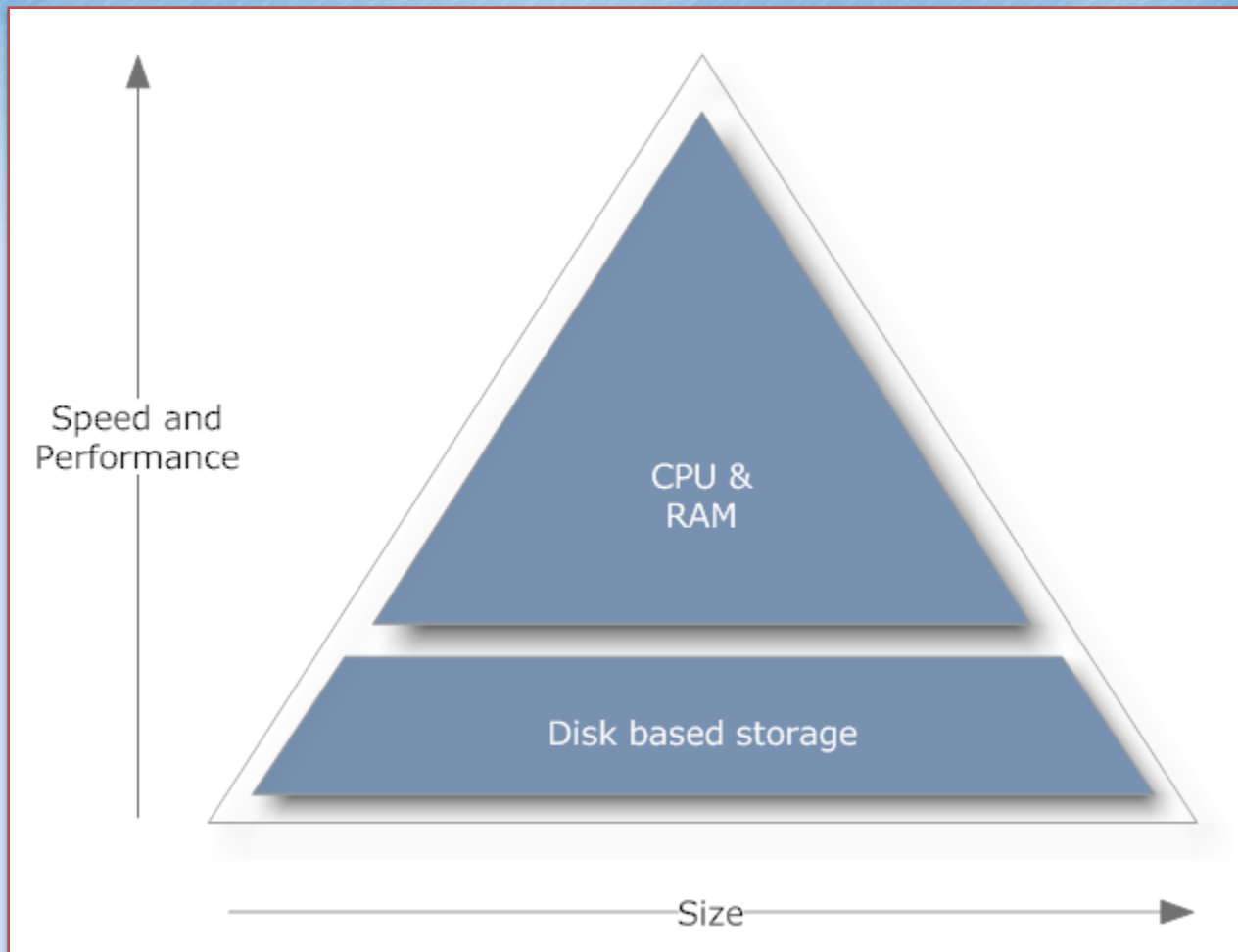
Hybrid Storage Pool

Hybrid Storage Pool: der Hintergrund

	1991 Seagate ST32550N	1997 Seagate ST19101N	2000 Seagate Cheetah X15	2007 Seagate Savio 15.1K (2,5")	2009 SSD Disk Intel X-25E
Kapazität (GB)	2,1	9	18	73	64
Durchsatz	10 MB	20 MB	40 MB	80 MB	220 MB
Zugriffszeit (RPM)	9 ms (7200)	8,2 ms (1000)	6,55 ms (15000)	3,3 ms (15000 RPM)	80 μ s
IOPS (1/Access)	110 IOPS	120 IOPS	150 IOPS	280 IOPS	3300 (W) – 35000 (R)
200 GB DB IOPS *	10476 (? €)	2666 (? €)	1666 (4600 €)	767 (700 €)	10312 - 109375 (2200 €)

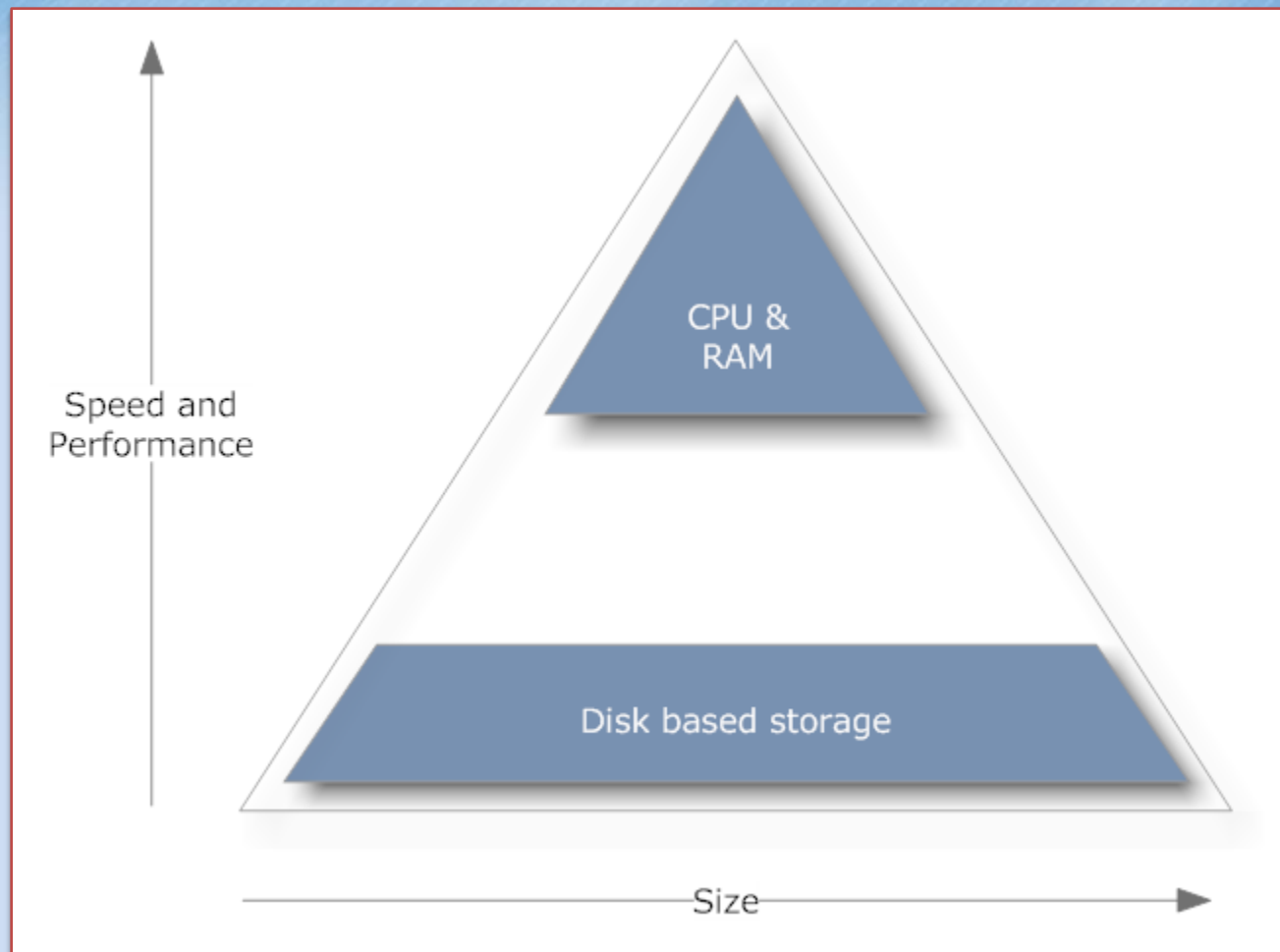
* Performance einer Datenbank mit einer Größe von 200 GB

Hybrid Storage Pool: Die 90er ...



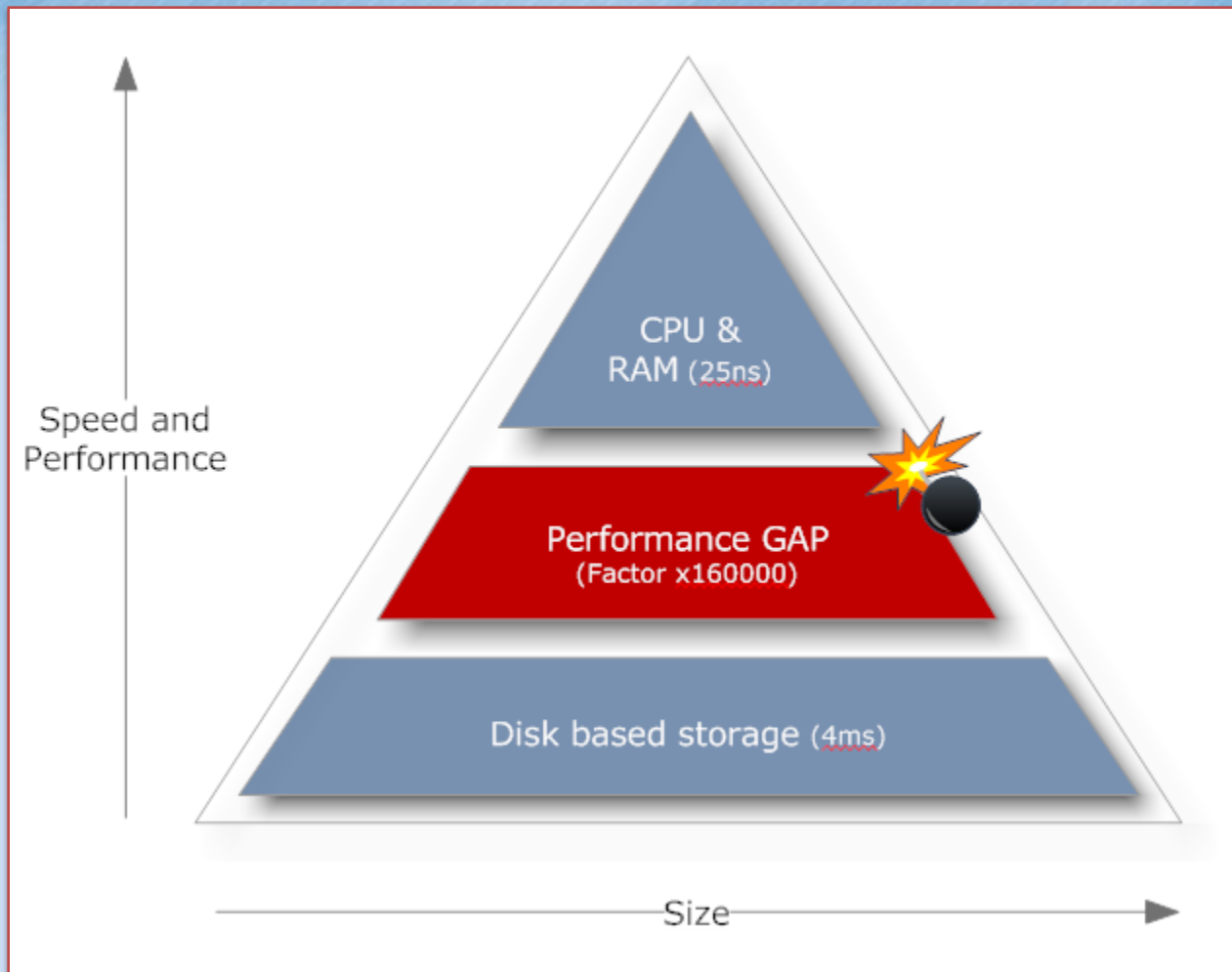
- Ausgangssituation in den 90er Jahren „Disks schnell genug“

Hybrid Storage Pool: ... seit dem



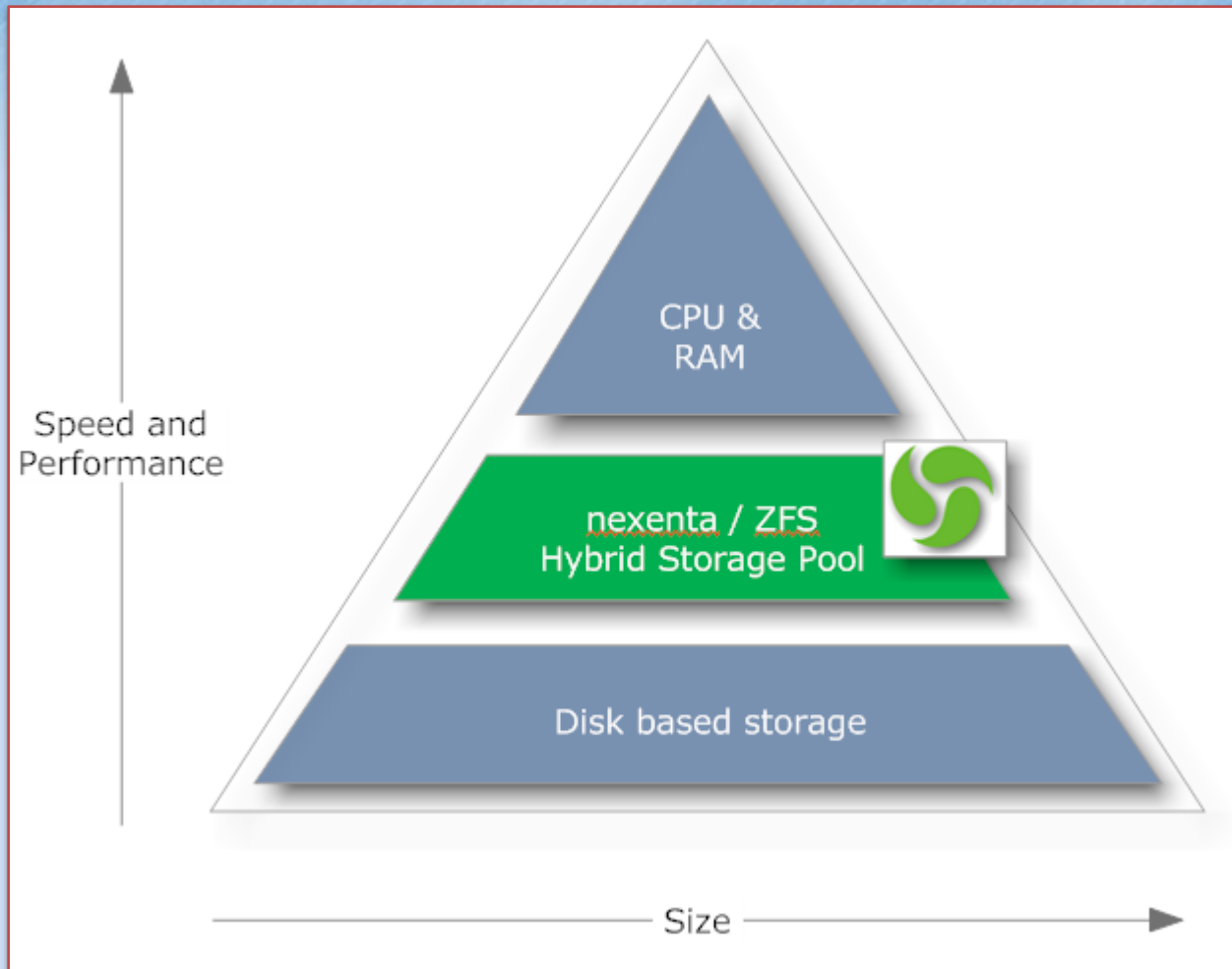
- Entwicklung seit dem: CPU Entwicklung schneller als HDD

Hybrid Storage Pool: Performance Gap



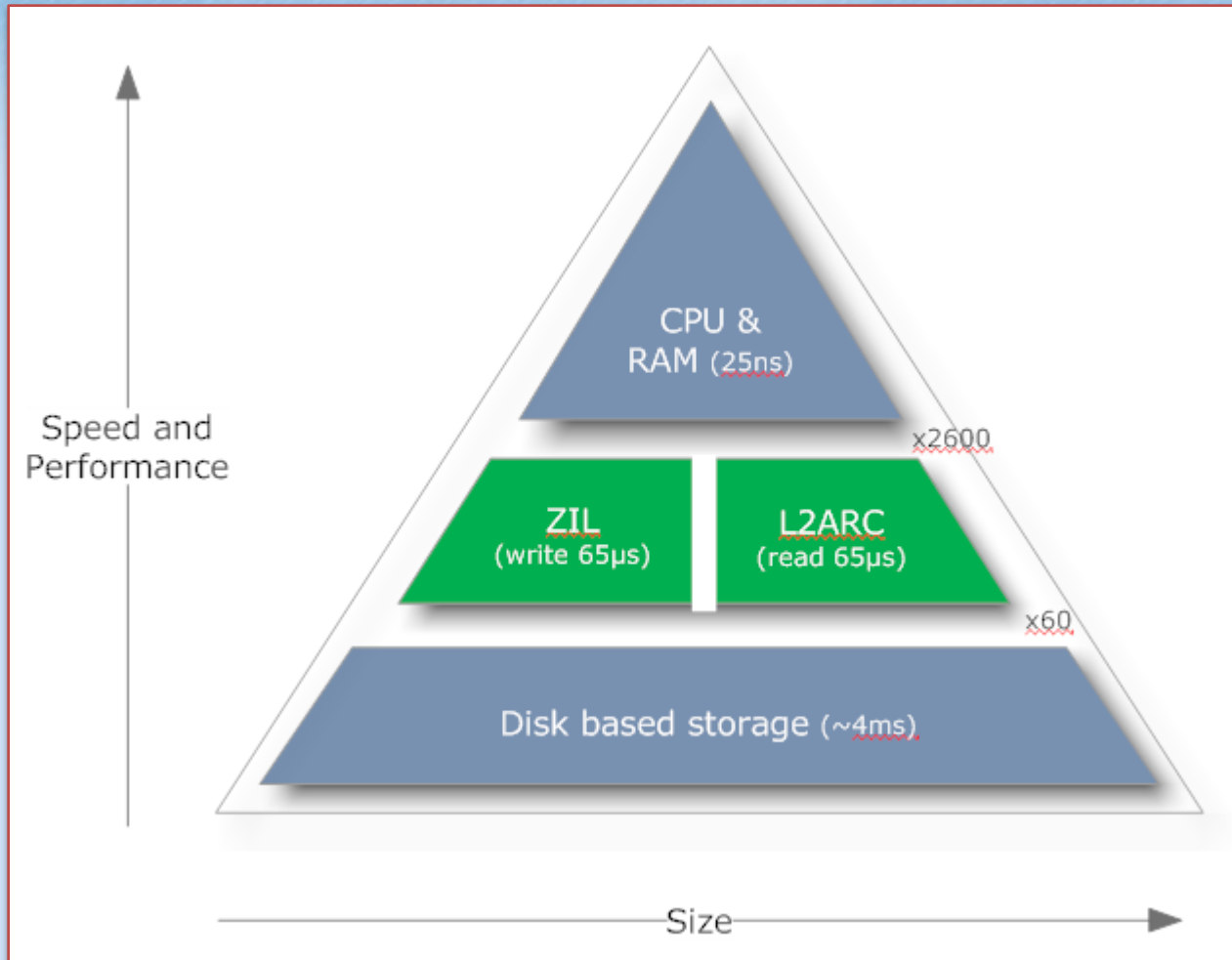
- Problem: CPU „zu Schnell“ für aktuelle HDD's

Hybrid Storage Pool: Die Lösung



- Lösung: Hybrid Storage Pool von NexentaStor

Hybrid Storage Pool: Die Lösung



- ZIL: Schreib-optimierte Speicher (RAM / SLC SSD)
- L2ARC: Lese-optimierte Speicher (MLC SSD)

Hands on ...



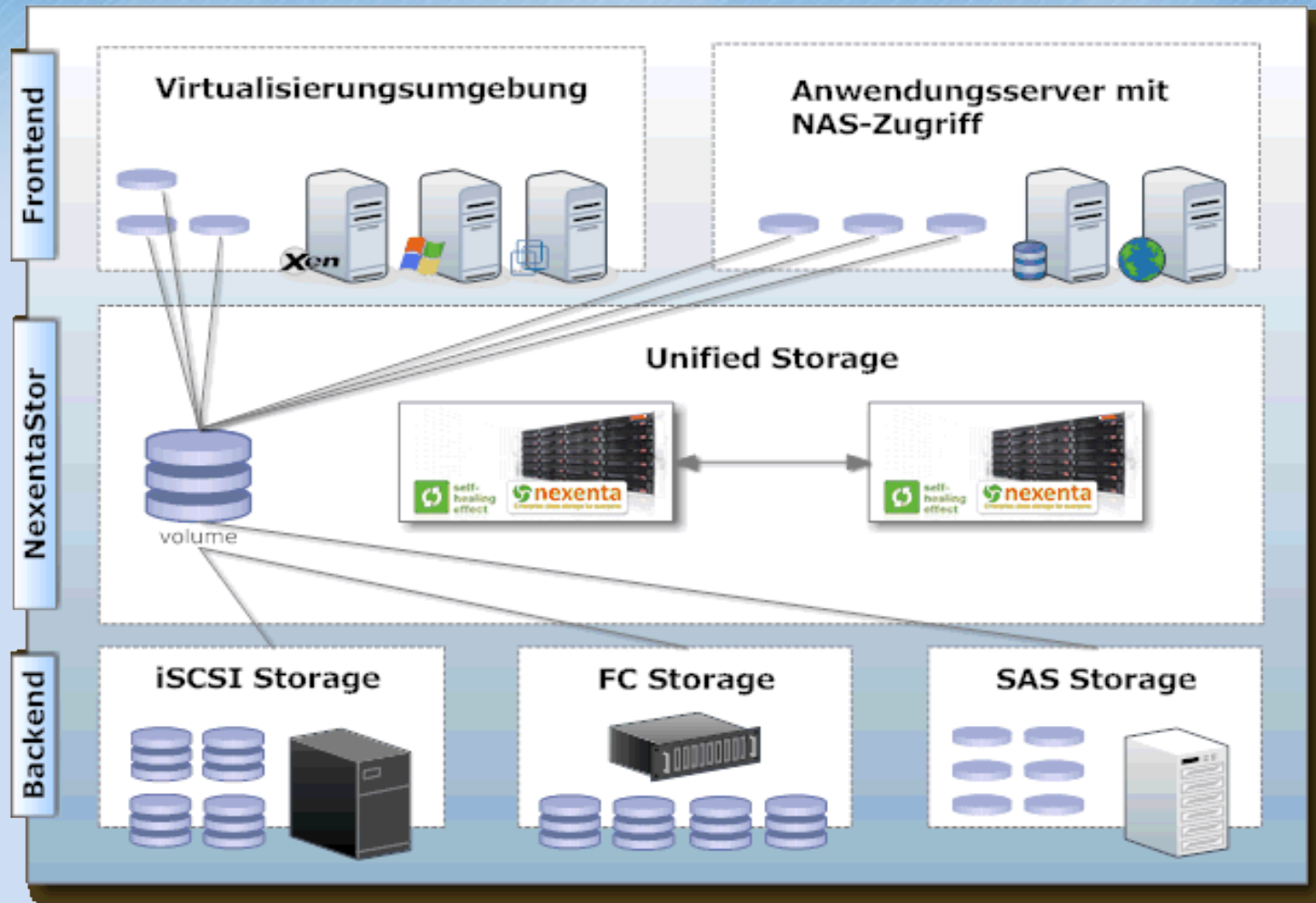
Thomas-Krenn.AG und NexentaStor™ Open Storage

... why selecting hardware vendors carefully still makes sense in an open world.

Thomas-Krenn.AG und NexentaStor™

- ZFS stellt gewisse Anforderungen um die Datenintegrität gewährleisten zu können
- NexentaStor™ Support Feedback: Mehrheit von Support Anfragen sind Hardware bezogen (> 60%)
- Qualifizierte Hardware macht Sinn!
- Thomas-Krenn.AG bietet die erste zertifizierte NexentaStor™ Lösung in D-A-CH seit März 2010 an

Thomas-Krenn.AG 846 Storage



Are you **OPEN** for **STORAGE** ?

Thank you / Danke

<http://www.nexenta.de>

<http://www.thomas-krenn.com>

Referenzen und Links

- [CERN] Study on data integrity
<http://indico.cern.ch/getFile.py/access?contribId=3&sessionId=0&resId=1&materialId=paper&confId=13797>
- Study on ZFS data integrity
www.cs.wisc.edu/wind/Publications/zfs-corruption-fast10.pdf

Thomas Krenn Herbstworkshop 2010

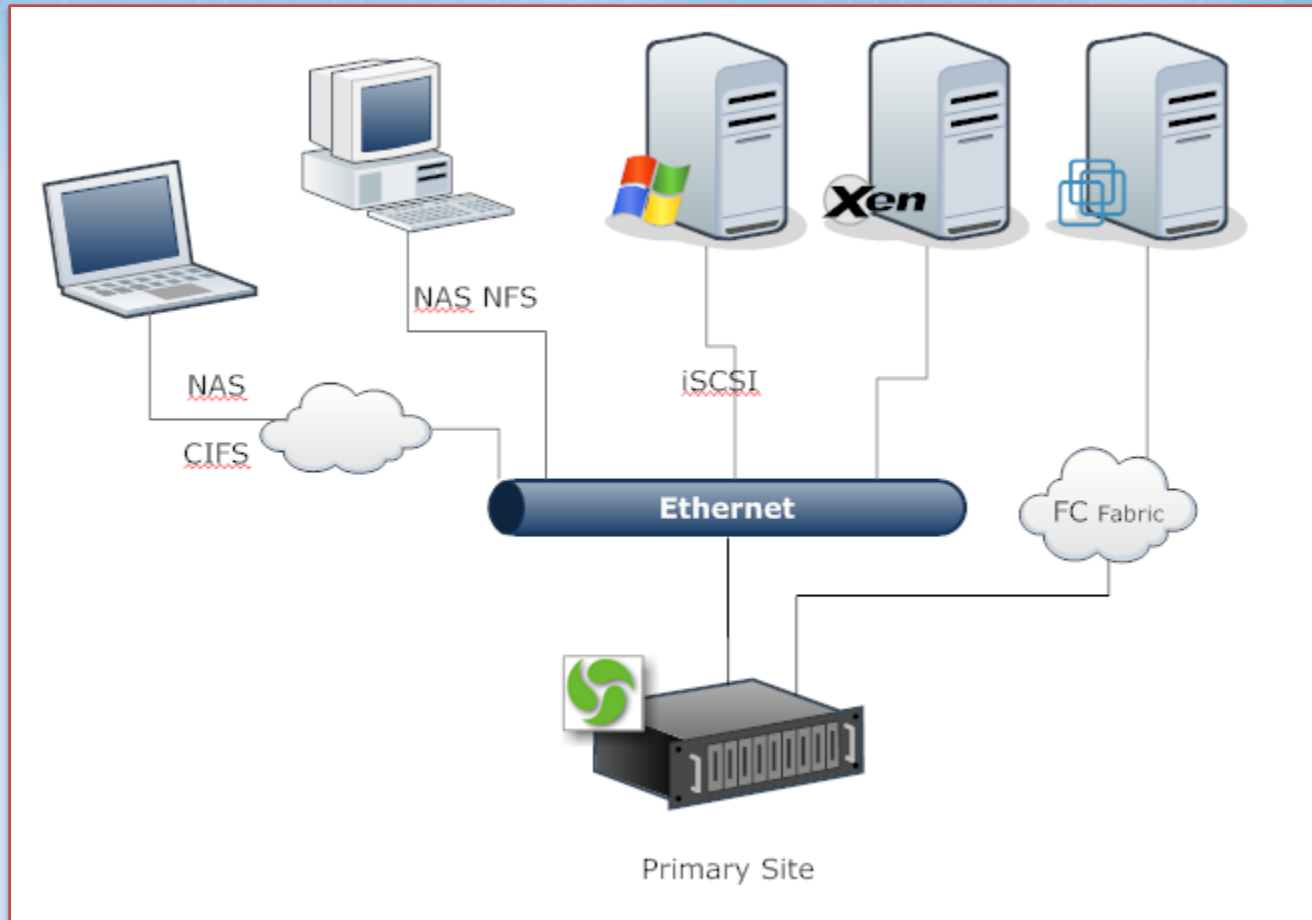
BACKUP SLIDES



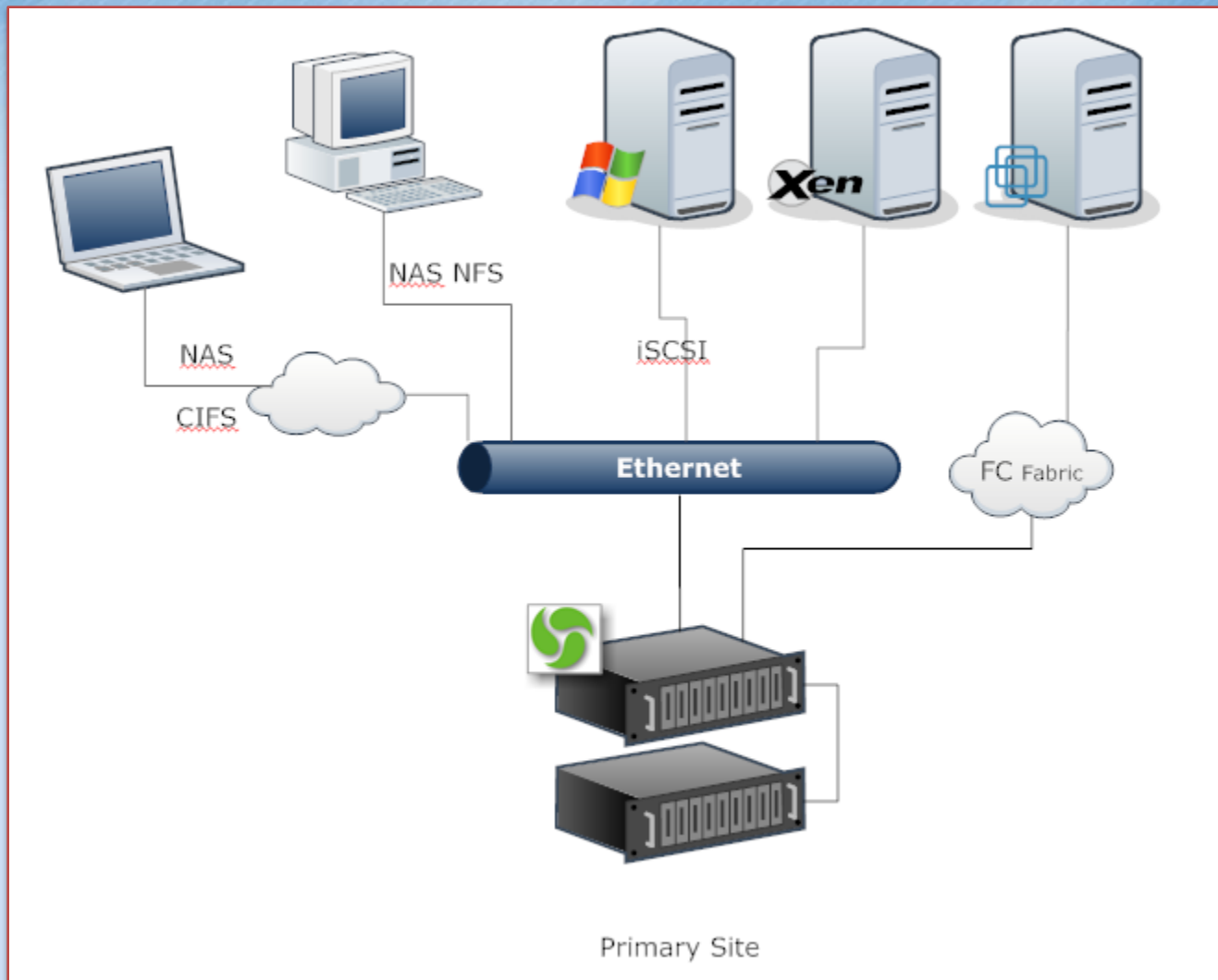
Thomas-Krenn.AG[®]
Speed is (y)our success



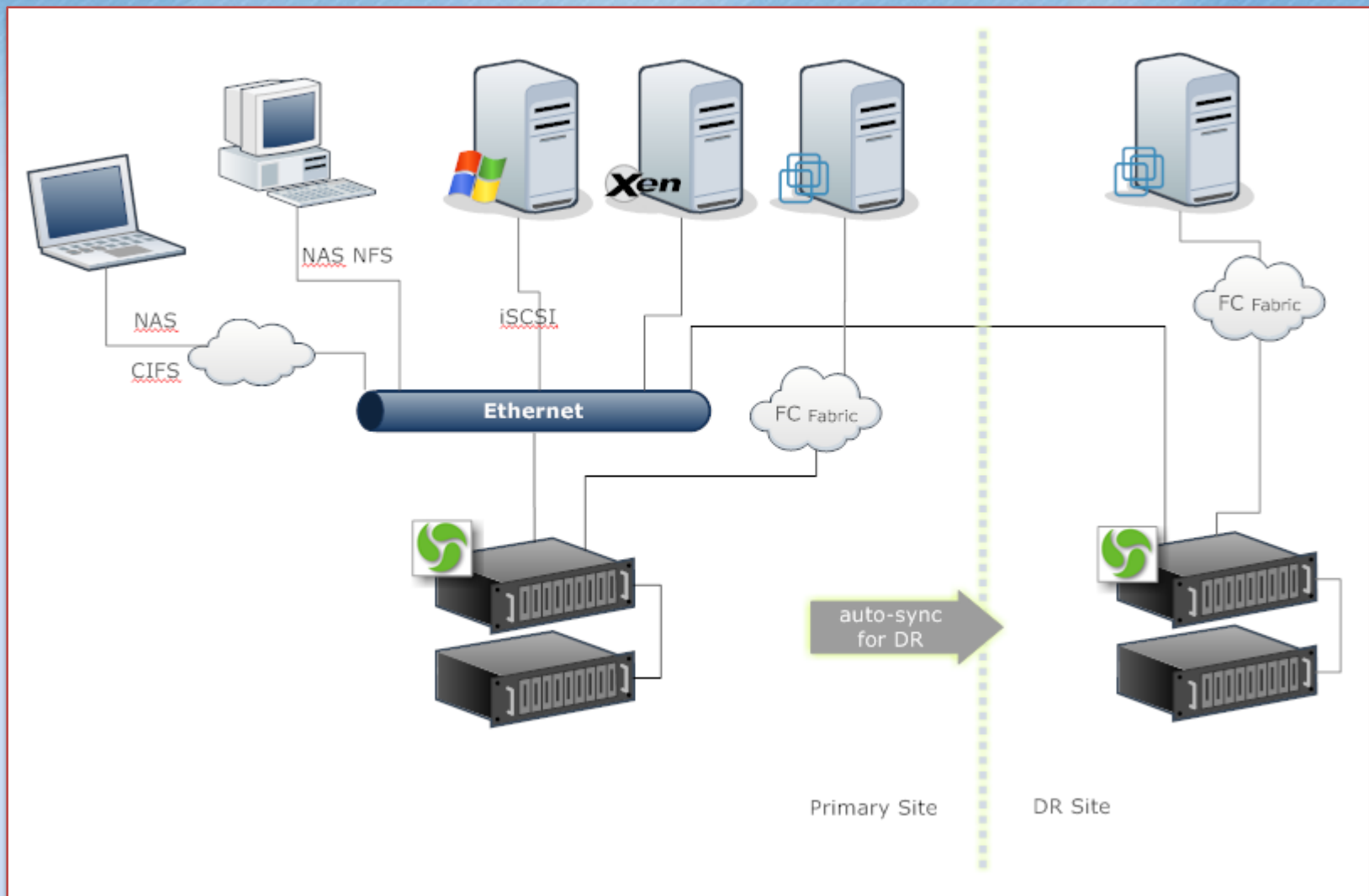
Beispielconfig: Basis



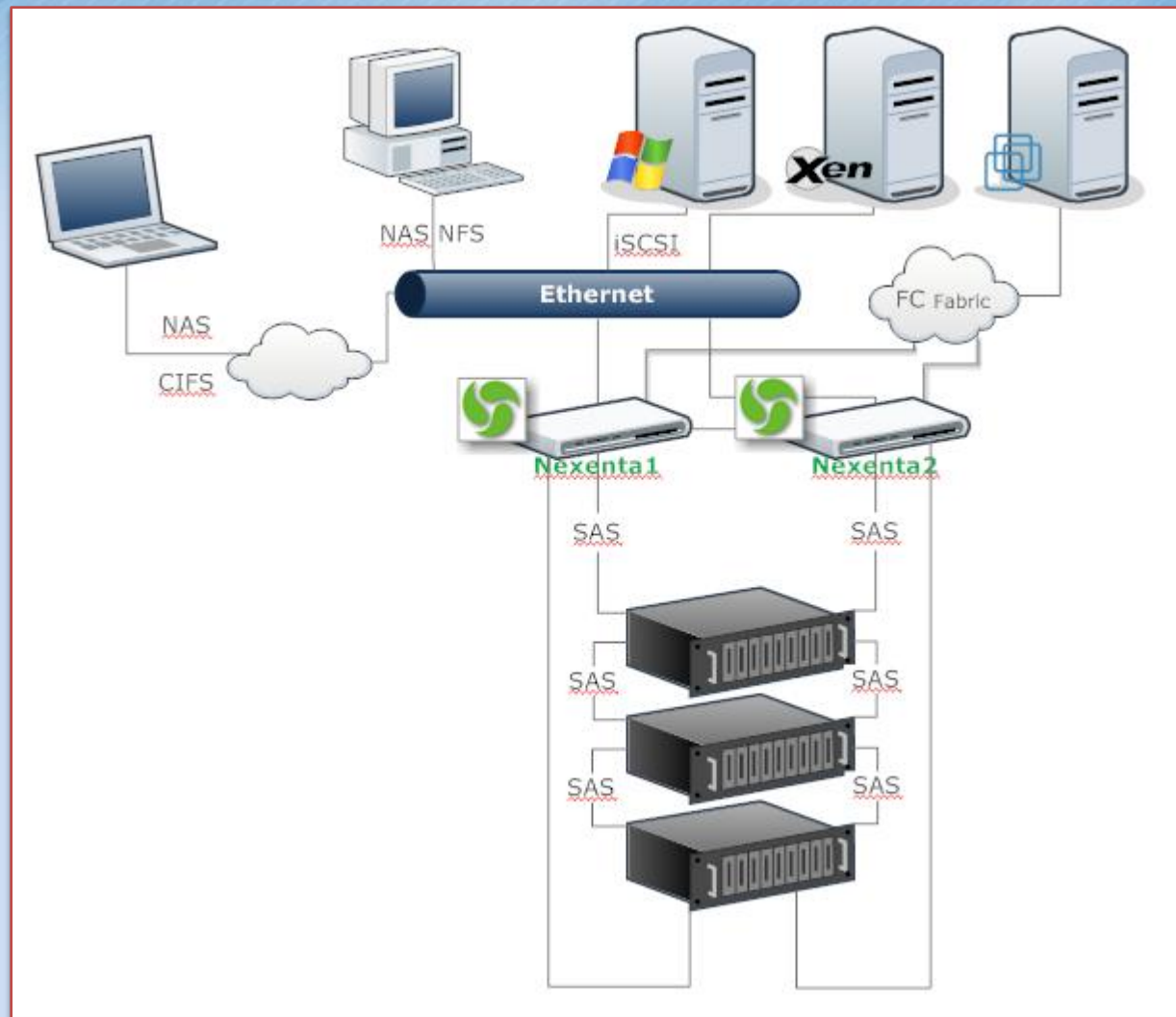
Beispielconfig: Basis Erweitert



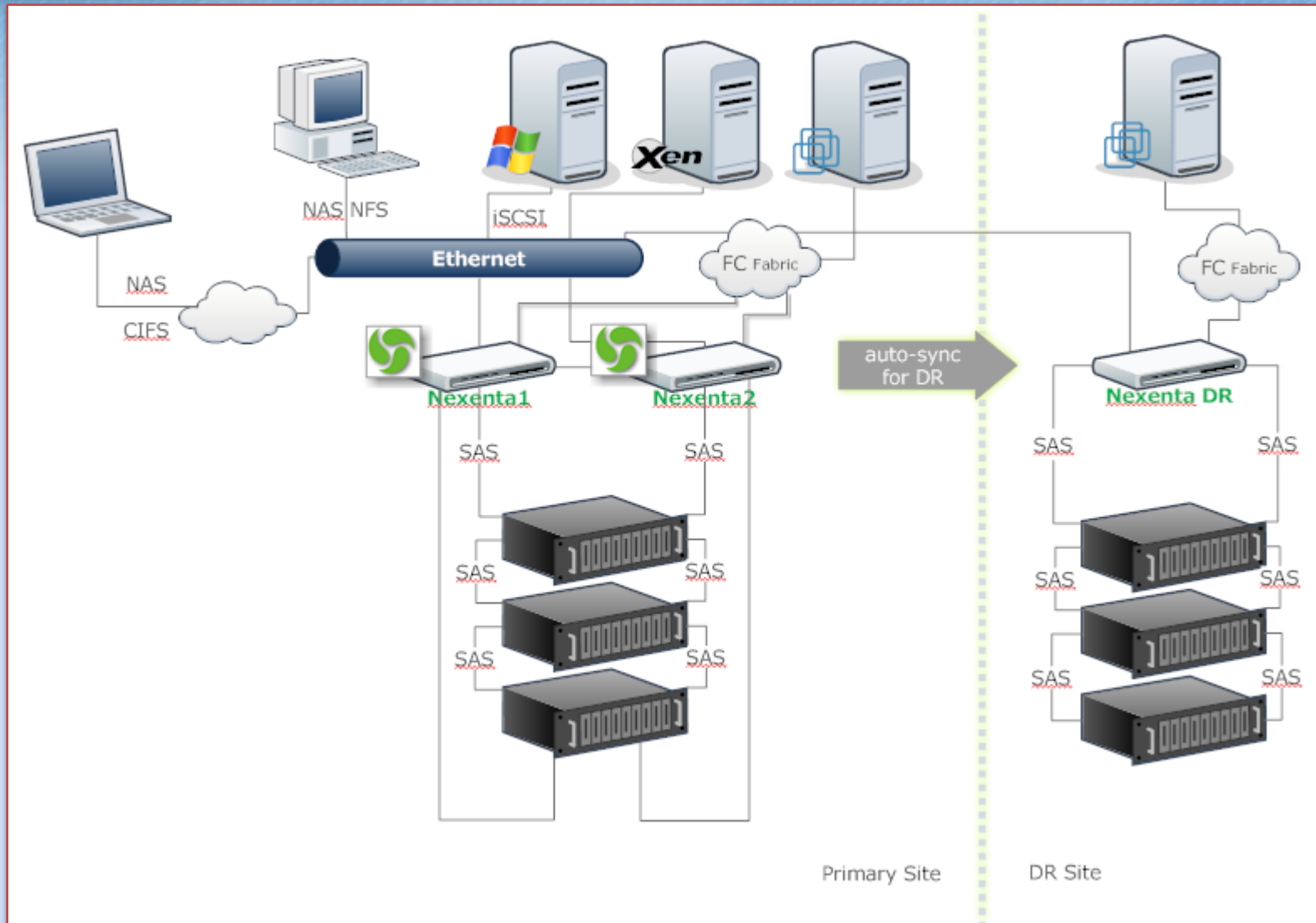
Beispielconfig: DR Setup



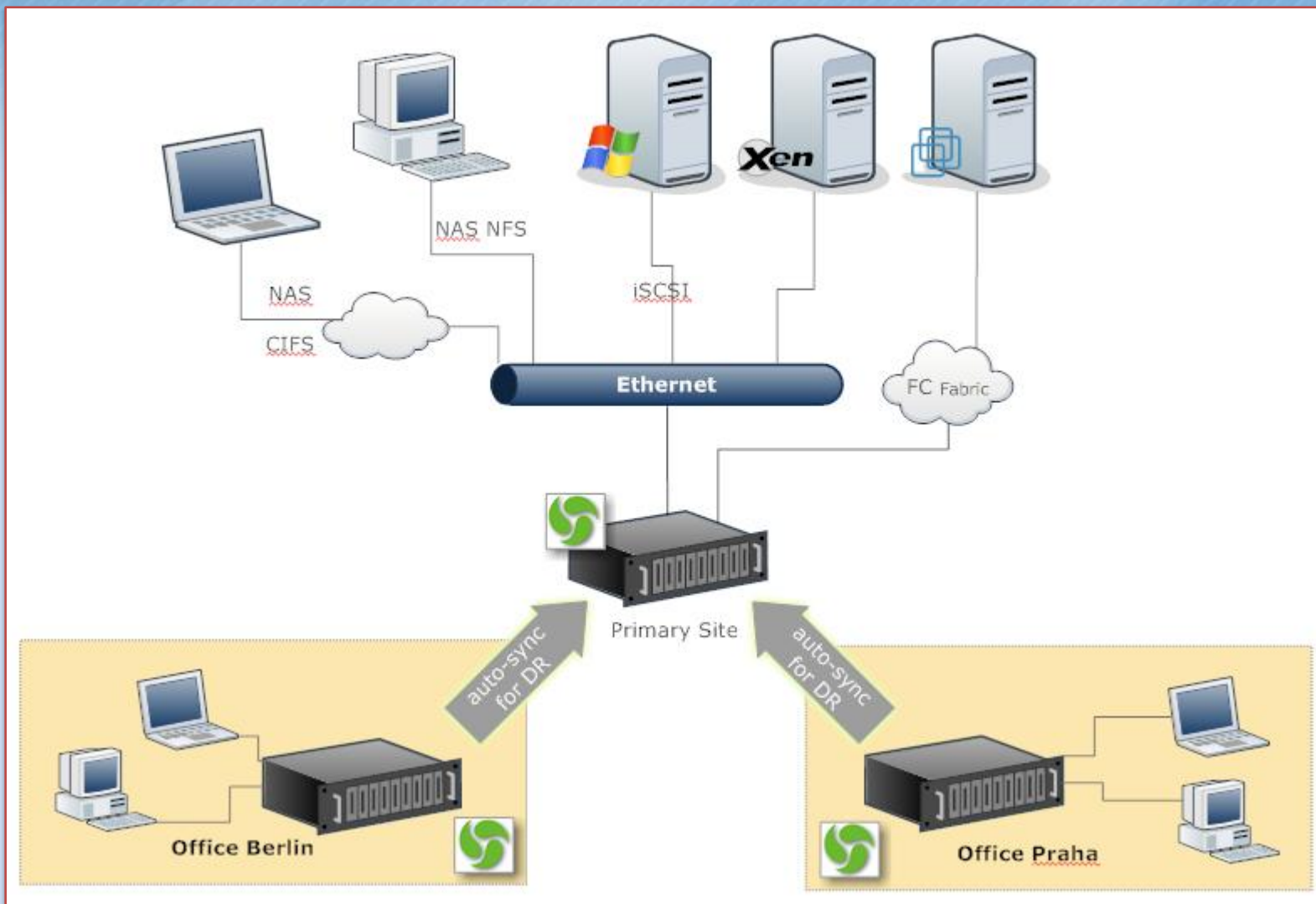
Beispielconfig: HA Cluster



Beispielconfig: Maximum Availability



Beispielconfig: Multi-Site



Datenwachstum: Datenmenge

