

# *Storage Performance für virtualisierte Umgebungen*

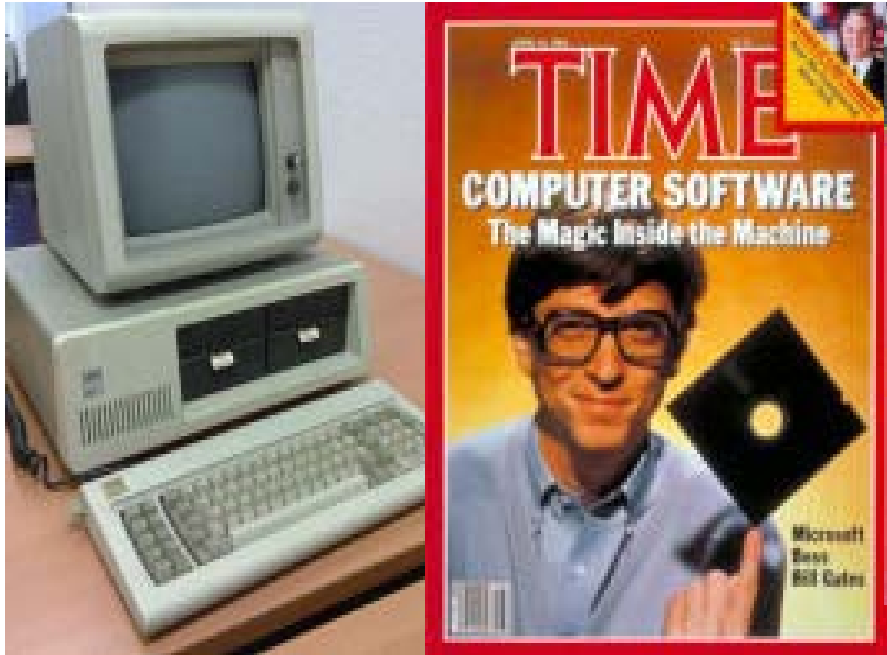
*Die Architektur für die nächste Dekade*



**Ralf Wendt**  
**HP Storage - Deutschland**

# Alt hergebrachte Architekturen haben Limitationen

Die meisten Speicherkonzepte wurde für Bedürfnisse entwickelt, wie sie vor 10- 20 Jahren bestanden



“Ich muss sagen, als ich 1981 diese Entscheidungen getroffen hatte, hatte ich das Gefühl, dass ich genug Freiräume für die nächsten 10 Jahre bereitstellen würde.

Der Schritt von 64kB auf 640kB fühlte sich an, als ob es für einen gehörigen Zeitraum reichen würde. Na ja, das tat es nicht – es dauerte nur 6 Jahre, bis man erkannte, dass es ein richtiges Problem wurde.” \*

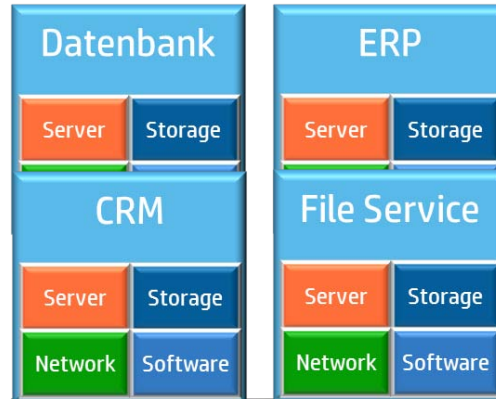
# Alt hergebrachte Architekturen haben Limitationen

Die meisten Speicherkonzepte wurde für Bedürfnisse entwickelt, wie sie vor 10- 20 Jahren bestanden

## Damalige Rechenzentren

- Wenige Applikationen
- Geringe Datenmengen
- Planbares Wachstum

## Vor 10-20 Jahren: Silo Konzepte

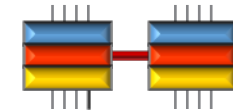


## Architektur Design der 90er Jahre

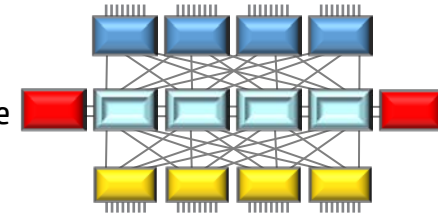
- Hardwarenahe Softwarearchitektur
- 1 zu 1 Beziehung Server/Storage
- Dedizierte Speichercontroller
- Geringe Kapazitäten

## Traditioneller Modular 2-Cntrl./monolithischer Storage

Server Verbindung



Platten Verbindung



# Die Welt hat sich gewandelt...

Die nächsten 10 Jahre ist Storage durch die „Ära der Unvorhersagbarkeit“ gekennzeichnet

## Flexible, skalierbare IT-Umgebung mit virtuellen & physikalischen Servern



## Kapazitäts- / & Performanceanforderungen treiben die Kosten in die Höhe



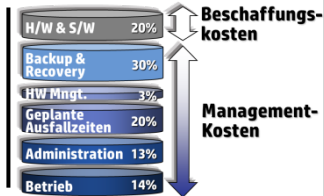
## Steigende Komplexität mit gebundenen Management- / & Kapazitätsressourcen



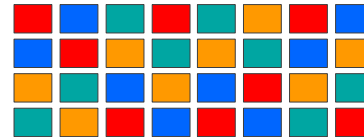
30%



15%



## Explosives Wachstum von Energie, Stellfläche & Ineffizienz



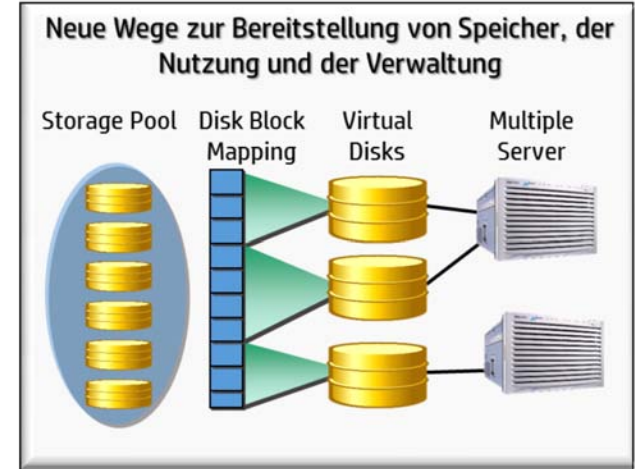
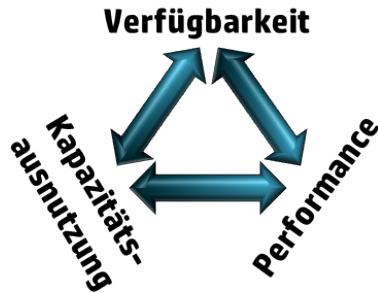
# Warum Virtualisierung bei Storage?

## Warum?

- erlaubt ein einfaches und schnelles Management
- überwindet physikalische Grenzen alter Architekturen

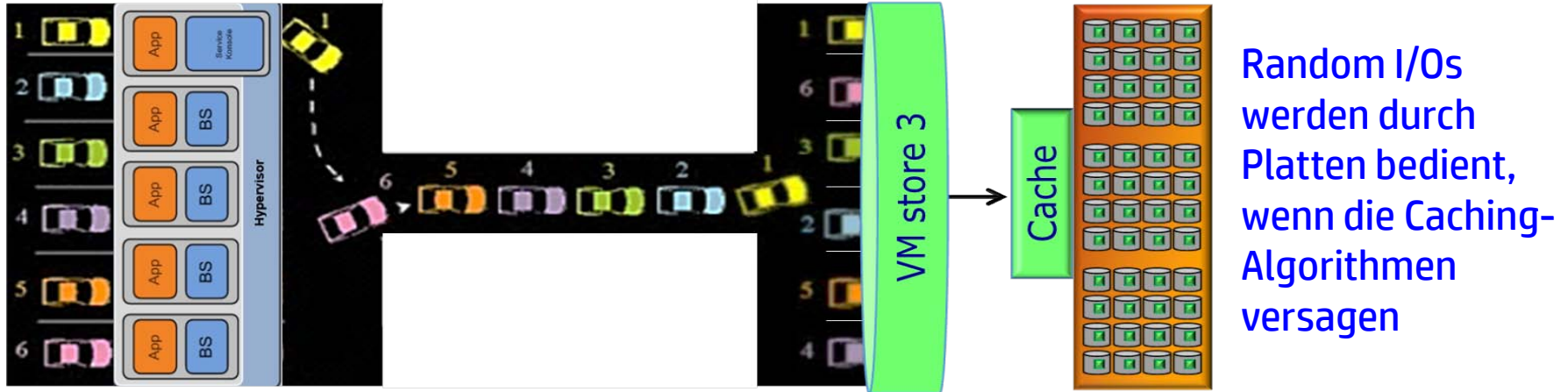
## Wie?

- Management geschieht auf der logischen Ebene, unabhängig von der physischen Realität
- Die Array-Virtualisierung übernimmt die optimale physischen Abstimmung  
→ Optimal bzgl. ...



# Herausforderungen an moderne Speicherinfrastrukturen

Technisches Beispiel: VMs & Applikationen haben grundsätzlich einen „random“ I/O



Die Performance einer VM Serverumgebung hängt sehr eng mit dem VM-Speicher und der Anzahl der physischen Festplatten seiner LUN zusammen  
→ je mehr parallele Plattenzugriffe, desto höher die Performance

# Überlegungen zur Leistungsfähigkeit von Speichersystemen



Michael Ettengruber, Business Development Manager EVA  
P6000 & FC

Hewlett-Packard GmbH



# Festplatten-Grundlagen – Performance?

Festplatten haben eine Zugriffszeit von 5-7 (bis zu 15ms bei MDL)

Millisekunden – **unerheblich**  
**ob 2,5 oder 3,5 Zoll**

Auf Cache-Daten kann mit einer Verzögerung von Mikrosekunden zugegriffen werden (1000 mal schneller)

Das tatsächliche Lesen eines 8k Blocks dauert ca. 1/20000tel Sekunde (0.05ms) (bei 160MB/sec Sustained Transfer Rate der Disk)

Zugriffszeit:  
Pos-Zeit + Rot.Latenz  
typ: 10k: 6-7ms  
typ: 15k: 5-6ms





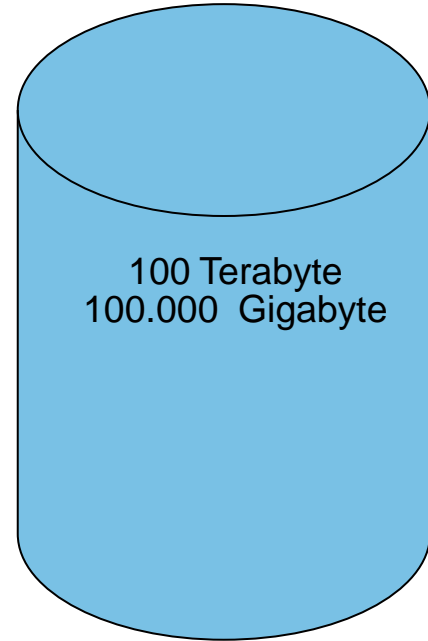
# Cache-Größe – Gigabyte vs. Terabyte

Verhältnis Cache- und Disk-Kapazität ist immer mind. Faktor 1000 – oder 1 Promille.

Trefferwahrscheinlichkeit?  
1:1000 oder schlicht  
wirkungslos ohne „Intelligenz“

Nicht die Größe des Cache ist ein Kriterium für die Qualität sondern der Umgang damit!

4/8/16/32/64 GB





# Überlegungen zur “HBA Queue Depth”

# Leistungsfähigkeit von Speichersystemen

## Überlegungen zur HBA Queue Depth

- Host HBA Queue Depth legt die Anzahl an ausstehenden I/Os des Servers zum Speichersystem pro LUN fest
- Server “sieht” 8 LUNs über den HBA und die HBA Queue Depth ist “16 “ dann kann der Server 16 I/Os auf jede der 8 LUNs ausstehend haben
- **Ergibt gesamt:  $8 * 16 = 128$  I/Os für einen Server**
- Diese ausstehenden I/Os werden benutzt um Daten asynchron auf´s Speichersystem zu schreiben



# Leistungsfähigkeit von Speichersystemen

## HBA Queue Depth Bedeutung

- Hohe HBA Queue Depth Werte können schnell zum Problem werden:
- Wenn zu viele Server auf mehrere LUNs des Speichersystem zugreifen und zu hohe QD-Werte eingestellt haben, dann wird die SCSI-Warteschlange des Speichersystem-Hostports zu gross :
  - Bei “grossen” I/Os (>32/64kB) I/Os wird einen hohe MB/sek-Last auf dem System erzeugt, die die zeitgerechte Ausführung “kleiner” I/Os verhindert
  - Bei Hosts mit “grossen” I/Os muss unbedingt die Queue Depth angepasst werden
  - Standard-Wert 32-256 I/Os – guter Mittelwert 16, bei “grossen” I/Os evtl.  $\leq 8$
  - Ausprobieren!



# Leistungsfähigkeit von Speichersystemen

## Beispiele zur HBA Queue Depth

– Hohe HBA Queue Depth Werte können **wirklich** schnell zum Problem werden:

### – Beispiel 1:

- 8 LUNs
- HBA queue Depth = 16
- I/O Grösse = 8kB
- Max. Daten, die “in einem Rutsch” am Speichersystem ankommen kann:  $8 * 16 * 8k = 1MB$

### Beispiel 2:

- 8 LUNs
- HBA queue Depth = 64 (häufiger Standard-Wert!!!)
- I/O Grösse = 256kB
- Max. Daten, die “in einem Rutsch” am Speichersystem ankommen kann :  $8 * 64 * 256k = 128MB$
- Ein einzelner Server mit 256kB I/Os und einer QD von 64 mit 8 LUNs erzeugt eine vergleichbare Last (MB/s) wie **128 (!)** Server mit 8kB I/Os und einer QD von 16
- → Was würde bei dieser Last als “Random Writes” auf RAID 5 passieren?





**“Device Busy” und “Queue Full”**

# Leistungsfähigkeit von Speichersystemen

## “Device Busy” und “Queue Full”

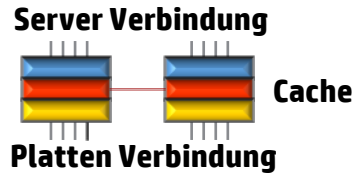
- Alle Speichersysteme brauchen einen Mechanismus um mit “Überlastung” umzugehen
  - Umsetzung: Entweder “SCSI\_Queue\_Full” oder “SCSI\_Device\_Busy” Status als Antwort auf eine I/O-Anfrage vom Server
  - Jeweils “selbsterklärend”:
    - “SCSI\_Queue\_Full” ist die Antwort bei “voller” Host Port Queue
    - “SCSI\_Device\_Busy” ist die Antwort bei vollem Schreib-Cache wenn das Speichersystem Zeit benötigt, um die Daten auf die Disks zu schreiben
- In beiden Fällen wird vom Server erwartet, eine zufällig bestimmte Zeit zu warten und dann den nächsten Versuch zu starten.
- Anmerkung: Vmware unterstützt das nicht bei allen Arrays!
  - ESX muss “Adaptive Queue Depth Throttling” für das spezifische Speichersystem unterstützen





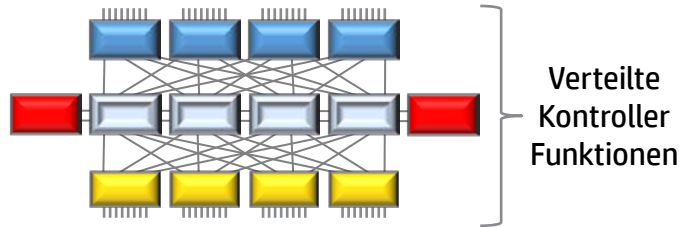
# Traditionelle vs. HP 3PAR Hardware Architektur

## Traditioneller Modularer Storage



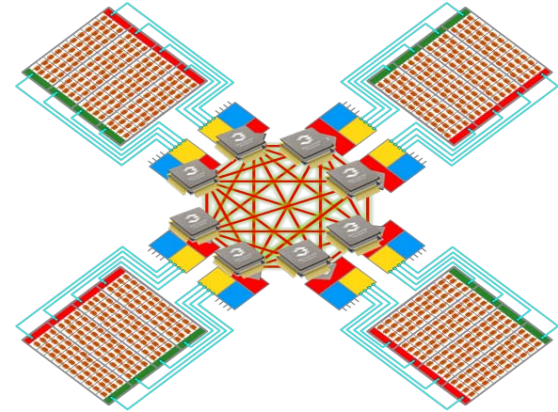
**Dual-Kontroller Design:** kosteneffizient, aber limitierte Skalierbarkeit / Belastbarkeit (Silo Architektur)

## Traditioneller Monolithischer Storage



**Monolithisches/Multi Cntrl. Design:**  
Komplex, Skalierbar / belastbar, kostenintensiv,  
keine hohe Flexibilität (Silo Architektur)

## HP 3PAR vernetzt und aktiv/aktiv



**Cluster GRID Design:** kosteneffiziente, skalierbare / belastbare Architektur. Trifft die Anforderung virtueller IT Umgebungen in Bezug auf:

- Effektivität
- Mandantenfähigkeit
- Hochverfügbarkeit
- Automatisiertes Management

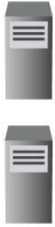
# Gemischte Lastprofile: Bereit für jeden Einsatz

## I/O Processing : Traditionelle Speicherarchitektur mit **SW** RAID

Hosts

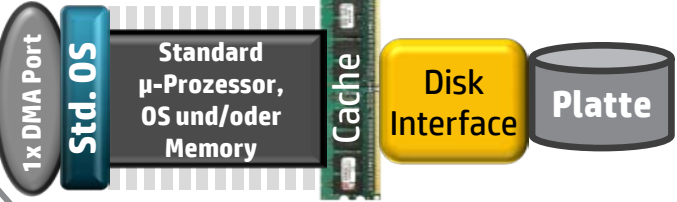
Hohes, Durchsatz orientiertes  
Lastprofil zur Verarbeitung

Hohes transaktionales  
Lastprofil zur Verarbeitung



Host  
Interface

kleine IOs warten auf Abarbeitung großer IOs

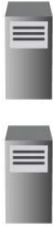


— = Kontrollinformation (Metadaten)  
— = Daten

## I/O Processing : 3PAR Controller Node mit **HW** RAID

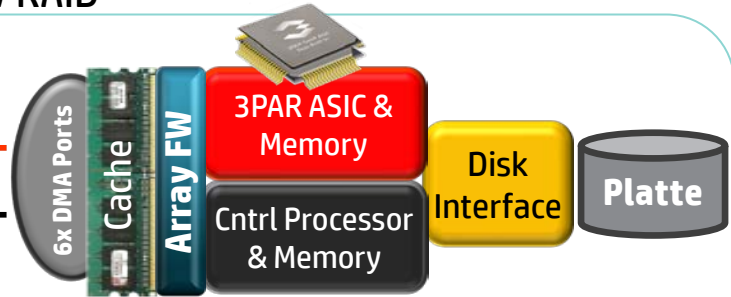
Hohes, Durchsatz orientiertes  
Lastprofil zur Verarbeitung

Hohes transaktionales  
Lastprofil zur Verarbeitung



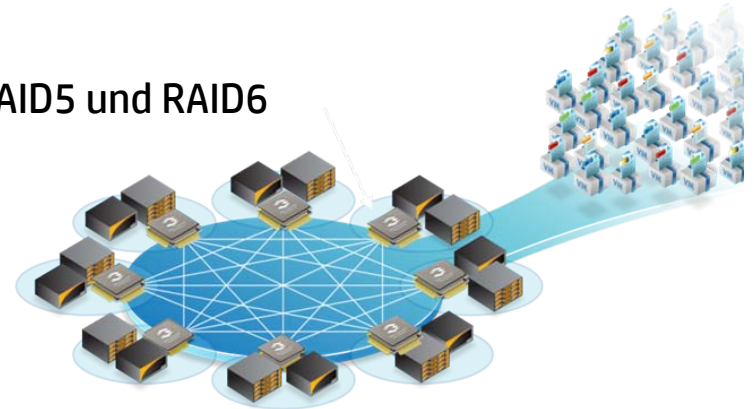
Host  
Interface

Kontrollinformation & Daten werden getrennt und separat verarbeitet  
(Konzept der „Out of Band“ Virtualisierung)



# Vorteile des 3PAR ASICs (64bit)

- Meshed **Active/Active Cluster** mit dynamischem Loadbalancing
- **Hardware** RAID Kalkulation
- **HW** unterstütztes **Rebuild** der Datenblöcke, insbes. für RAID5 und RAID6
- **Cache Kohärenz**
- **Zero Detection**
- Parallele Verarbeitung von Daten und Instruktionen (**Mixed Workload**) mit **multiplen DMAs**.  
Effiziente Rohkapazitätsnutzung mit nativer Thin Technology Implementierung.



# „Echtes“ Thin Provisioning: Merkmale

## Thin Technologien

- **Reservierungsfreie Implementierung** (Thin Provisioning)
- **inline Konvertierung** (Thin Conversion)
- **automatische Rückgewinnung** (Thin Reclamation)
- **Vermeidung „unnützer“ Daten** (Zero Detection)

## Skalierbare Performance

- Dynamische Performanceanpassung durch „**Wide Striping**“
- kein Performanceverlust zu Gunsten von Kapazitätseinsparungen

## Kleine Schreib-Datenblöcke (16kB)

- **Hohe Granularität** vermeidet unnütze Kapazitätsallokation

## Autonome Provisionierung

- Für Kapazitätserweiterungen bedarf es keines Administrators

## Für alle Datendienste

- z.B. Snapshots
- Remote Copies und damit bis zu 80% geringeren Datentransfer zwischen Fat & Thin Standorten

Start Thin



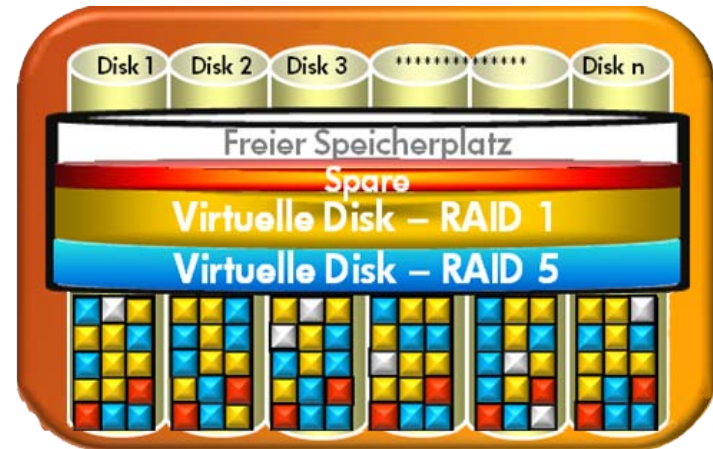
# Dynamische Optimierung - 1/2

Welche Anforderungen muss eine neue Architektur erfüllen?

Array basierte Virtualisierung bedeutet:  
Automatisierung und Aggregation von Systemaufgaben und Prozessen zur  
Verringerung der Komplexität und effizienten Ressourcennutzung

## Wide Striping

- Virtuelle Platten werden aus physikalischen Blöcken aller verfügbaren Platten erzeugt
- Online Erweiterbarkeit
- Dynamisches Load Balancing
- Keine „idle Spare Disk“

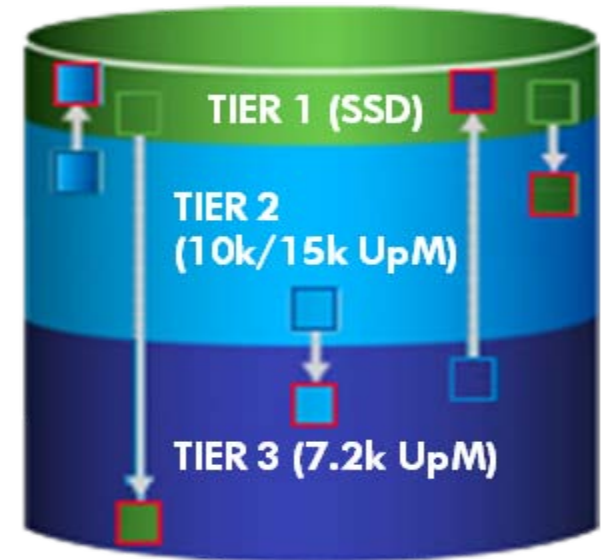


# Dynamische Optimierung - 2/2

Welche Anforderungen muss eine neue Architektur erfüllen?

## Storage Tiering

- Auf benutzerdefinierte Regeln basierendes automatisches Verschieben von kleinen Datenblöcken zwischen **allen** Speicherklassen innerhalb eines Speichersystems
- Freier Wechsel zwischen RAID Leveln
- Automatisches, freies und optimiertes platzieren der Daten auf der Platte
- Online erweiterbare Speicherklassen

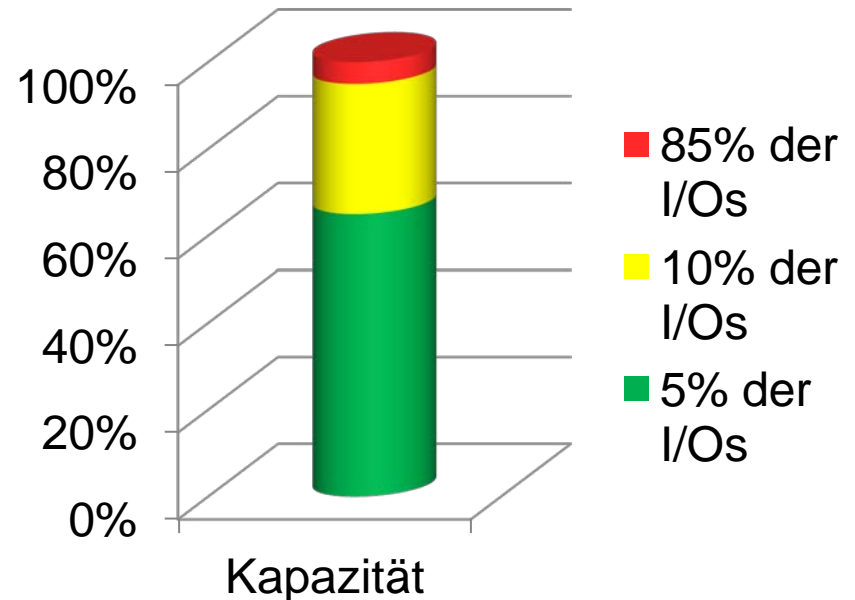


**Multi-Tier Volume**

# Adaptive Optimization

Autonomes Verteilen der Daten auf unterschiedliche Plattentechnologien

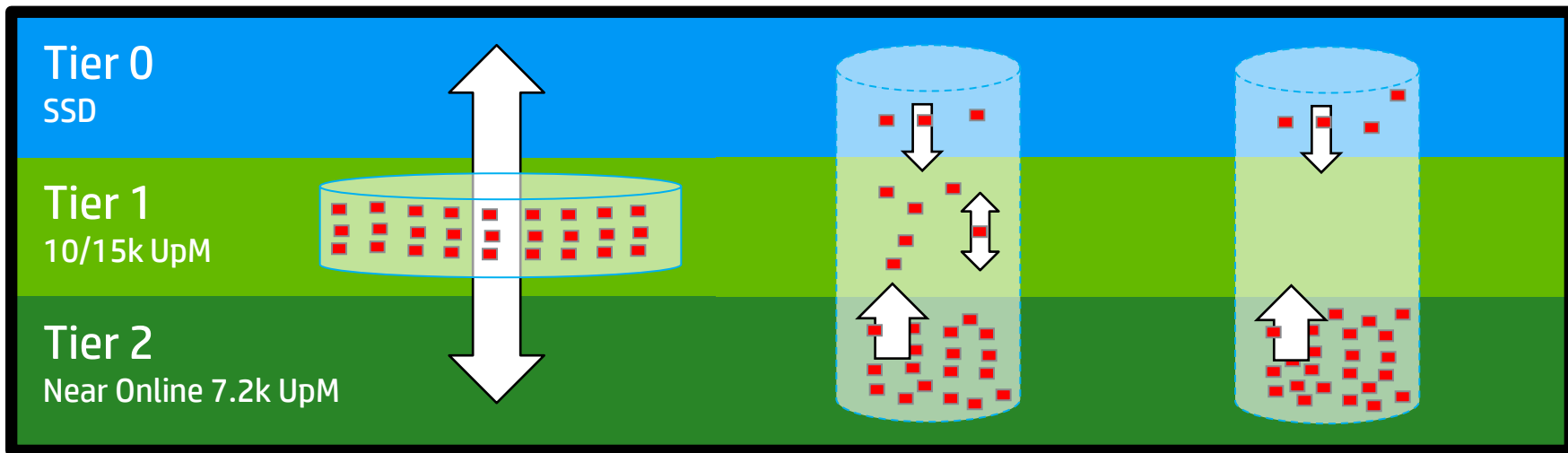
- Bei den meisten Applikationen erfordert ein kleiner Teil des Datenbestands den grössten Teil der Leistung (Bewegungsdaten)
- Ein typisches Verhältnis besteht in 95% der Zugriffe auf 5% des Datenbestands
- Aber: Applikationen und Umgebungen haben unterschiedliche “regionale Dichte”





# Die Lösung: 3PAR Adaptive LUN Optimierung (Tiering)

Automatisches Verschieben von Daten zwischen Speicherklassen – in 16kB Blöcken



■ = Datenblöcke

Dynamische  
Optimierung  
vollständiger LUNs

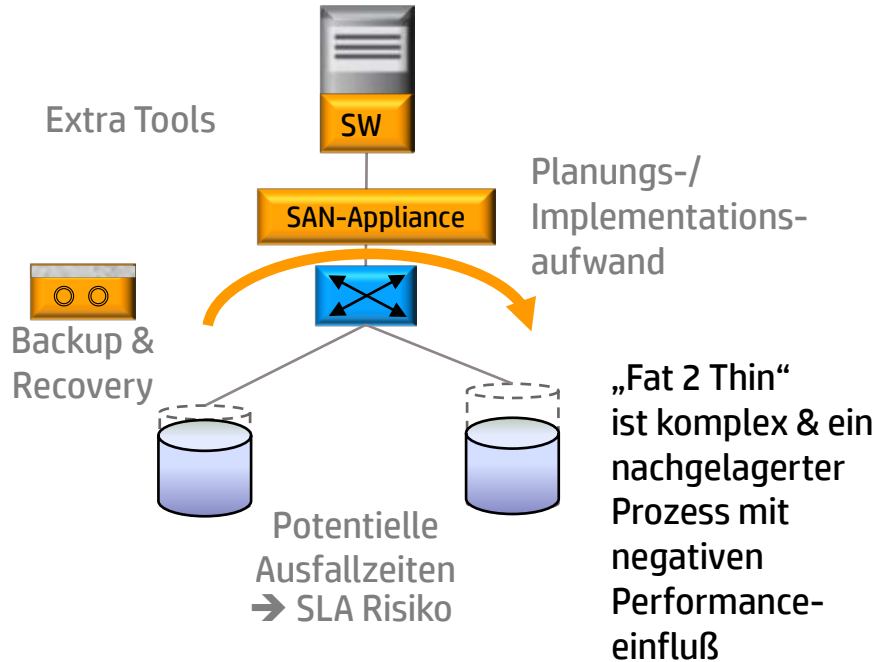
Optimierung auf Sub-LUN  
Level

Speicherklassen  
Optimierung mit  
Nearline & SSD

# Konvergente Migration — HP 3PAR Peer Motion

## Traditionelle Blockdatenmigration

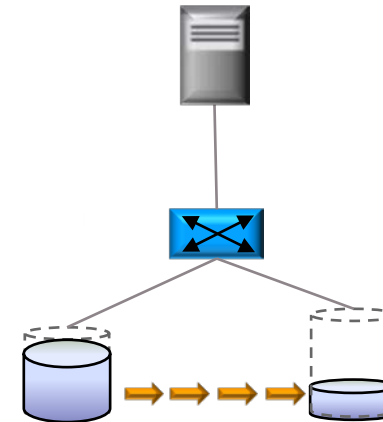
Komplex, zeitaufwendig, riskant



 = Block Migration Ansätze

## HP 3PAR Peer Motion

1ste unterbrechungsfreie „Knopfdruck“ Migration für Enterprise SAN



- einfach
- narrensicher
- online
- unterbrechungsfrei
- auf Array Ebene
- Inline Thin Conversion

„Peer Motion“ erlaubt:

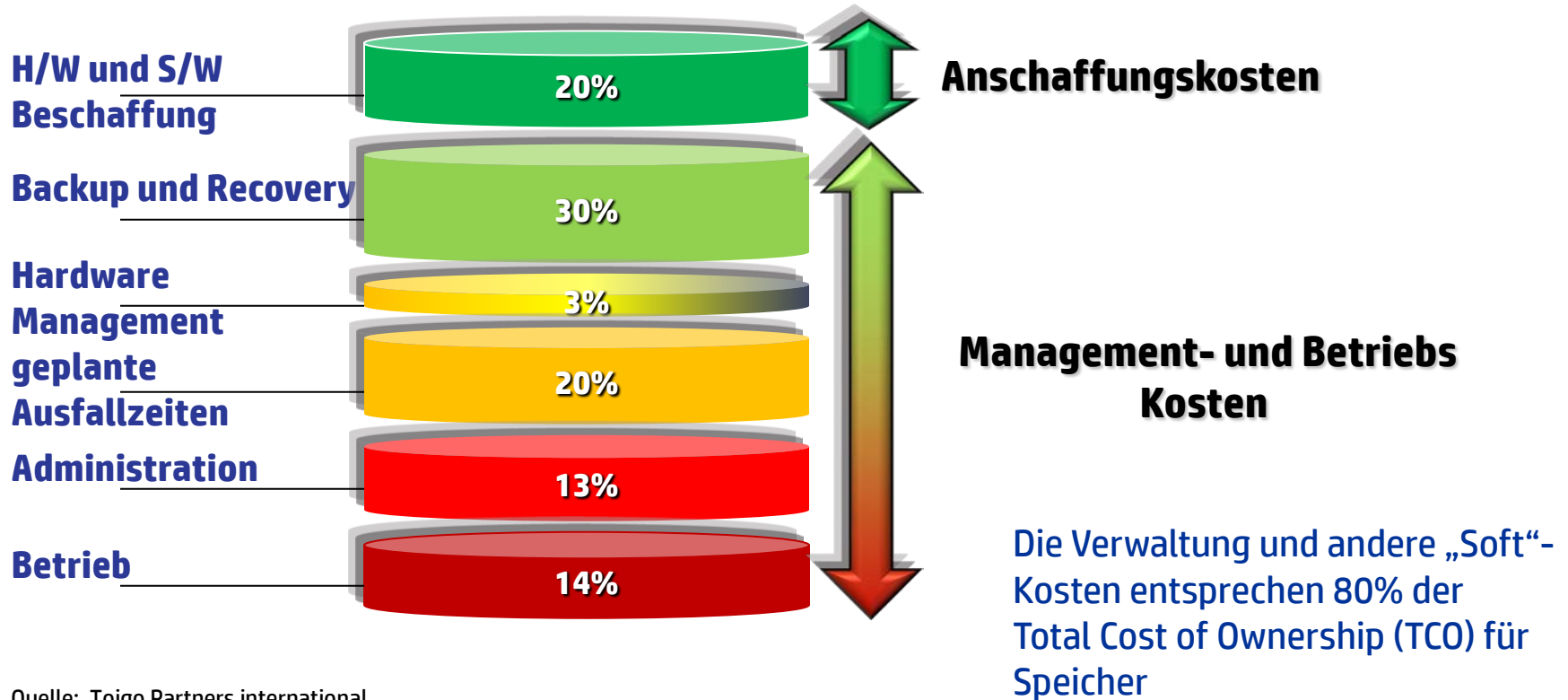
- „Load Balancing“ auf Knopfdruck
- nahtlose Technologieerneuerung
- Kosten optimiertes Array Lebenszyklusmngt
- Investitionskostenenkungen (Thin-landing)

# Unified Managementsystem

- Einheitliche Behandlung für virtuellen/physischen Servern, Storage und Netzwerk
  - Überwachung und Eventhandling
  - Inventarisierung
  - Konfigurationsschnappschüsse zum Vergleich für einzelne und mehrere Systemgruppen
- Einheitliche Betrachtung der physischen und virtuellen IT Umgebung
  - Provisionierung der gesamten Umgebung
- Integration in Management der Hypervisor
- Einheitliche Berichterstellung
  - Einzel- & multiple Systemberichte
  - Vordefinierte & individuelle Berichte
  - Wartungs- und Serviceverwaltung, z.B. Laufzeiten
  - Lizenzreporting und Überwachung



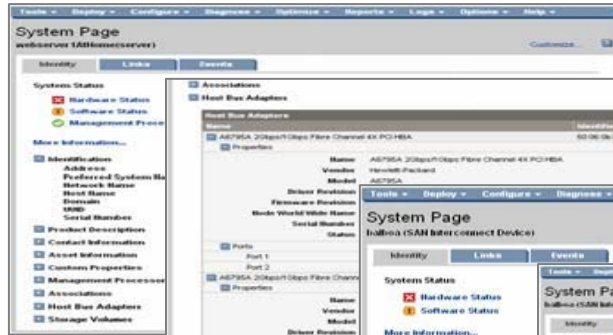
# Die Hardwarekosten sind nur die Spitze des Eisbergs...



Quelle: Toigo Partners international

# Einheitliches Server- und Speichermanagement

HP Insight Control und HP SIM – machen das Leben leichter



Server (Host) – virtuell oder physisch

Host Bus Adapter / FC-HBA

SAN Interconnect – FC Switch

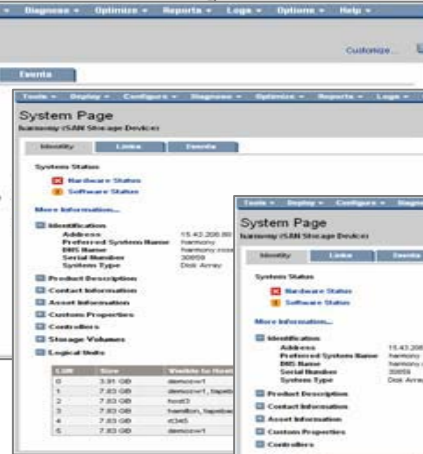
FC Switch Port

Speichersystem

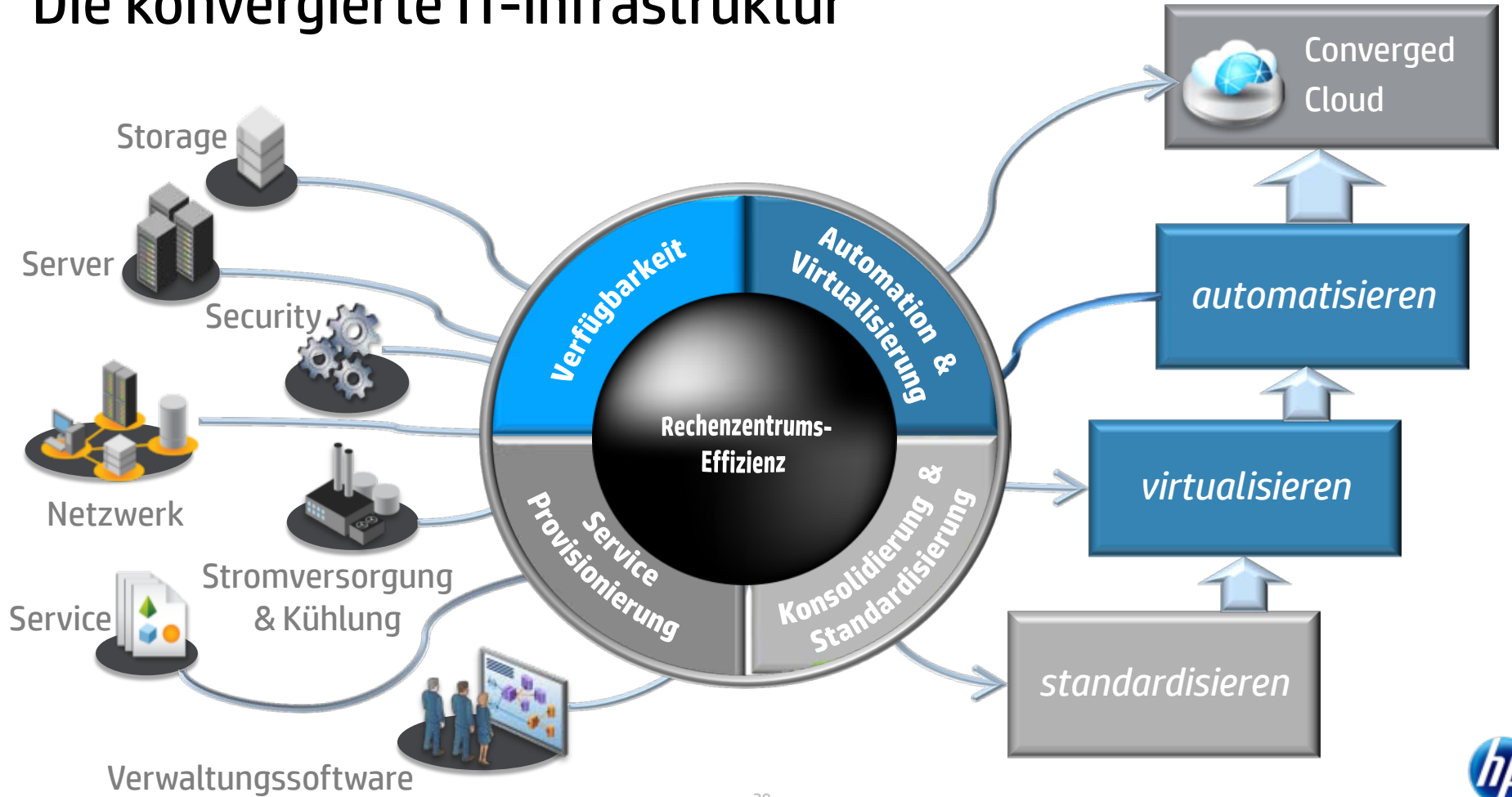
Speichersystem-  
Controller

LUN/vDisk

Identifikation und  
logischer Zusammenhang  
„out of the box“ – vom  
Server bis zur LUN



# Die konvergierte IT-Infrastruktur



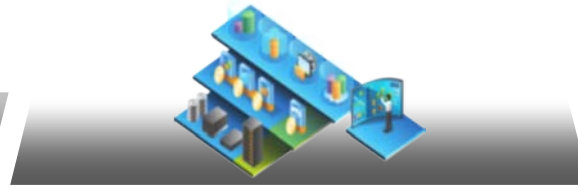
# Die HP Storage Strategie

Die Ära der IT Silos endet

## Konventioneller Speicher



## Converged Storage



## Converged Systems



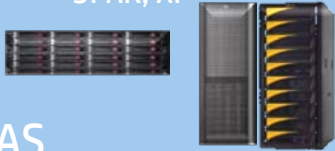
RZ Profil	<p><b><u>Gewachsene RZ Architektur</u></b> mit moderatem IT-Wachstum; relativ statische Server/Storage Beziehung; <b>IT optimieren &amp; Investitionen schützen</b></p>	<p><b><u>„Virtuelle“ RZ Architektur</u></b> moderates/starkes IT- Wachstum; IT muss flexibel &amp; skalierbar sein; hohe Virtualisierungsanforderungen <b>IT Grenzen überwinden</b></p>	<p><b><u>„automatisierte“ RZ Architektur</u></b> starkes IT- Wachstum; IT muss flexibel &amp; skalierbar &amp; automatisiert sein für das Bereitstellen von <b>IT &amp; Applikationen auf SLA Basis</b></p>
-----------	---	---	---



# HP Storage Portfolio – alles, was Sie brauchen

## Store

- **SAN:** MSA, EVA, LeftHand, 3PAR, XP



- **NAS**  
X1000, X3000, X5000, IBRIX



- **Dediziert**  
MSA, JBODs, Internal Server Storage



- **Archiv**  
HP/iCAS, DB Archiving,

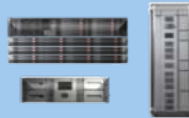


## Protect

- **Sicherheit**  
Secure Key Mgr  
Tape/Disk Encryption



- **Backup Systeme**  
StoreOnce D2D, ESL, EML, MSL



- **Datendienste**  
SnapShots, Remote Copy, 3-Site Recovery

- Backup Software  
**Data Protector**



## Optimize

- **Virtualisierung**  
LeftHand, EVA, 3PAR

- **Datenreduktion**  
D2D w/ StoreOnce single & multi node

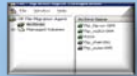


- **Thin Provisioning**  
3PAR, LeftHand, XP, EVA & EVA Dynamic Capacity Manager



## Manage

- **MOE**  
Matrix Operating Environment  
- Insight Control  
- HP SIM



- **SRM**  
Storage Resource Mngt.  
Storage Essentials



- **System Mgmt**  
Command View, IMC



- **Dokumentation & Record Mgmt**  
HP TRIM



Storage Services:

Entwerfen

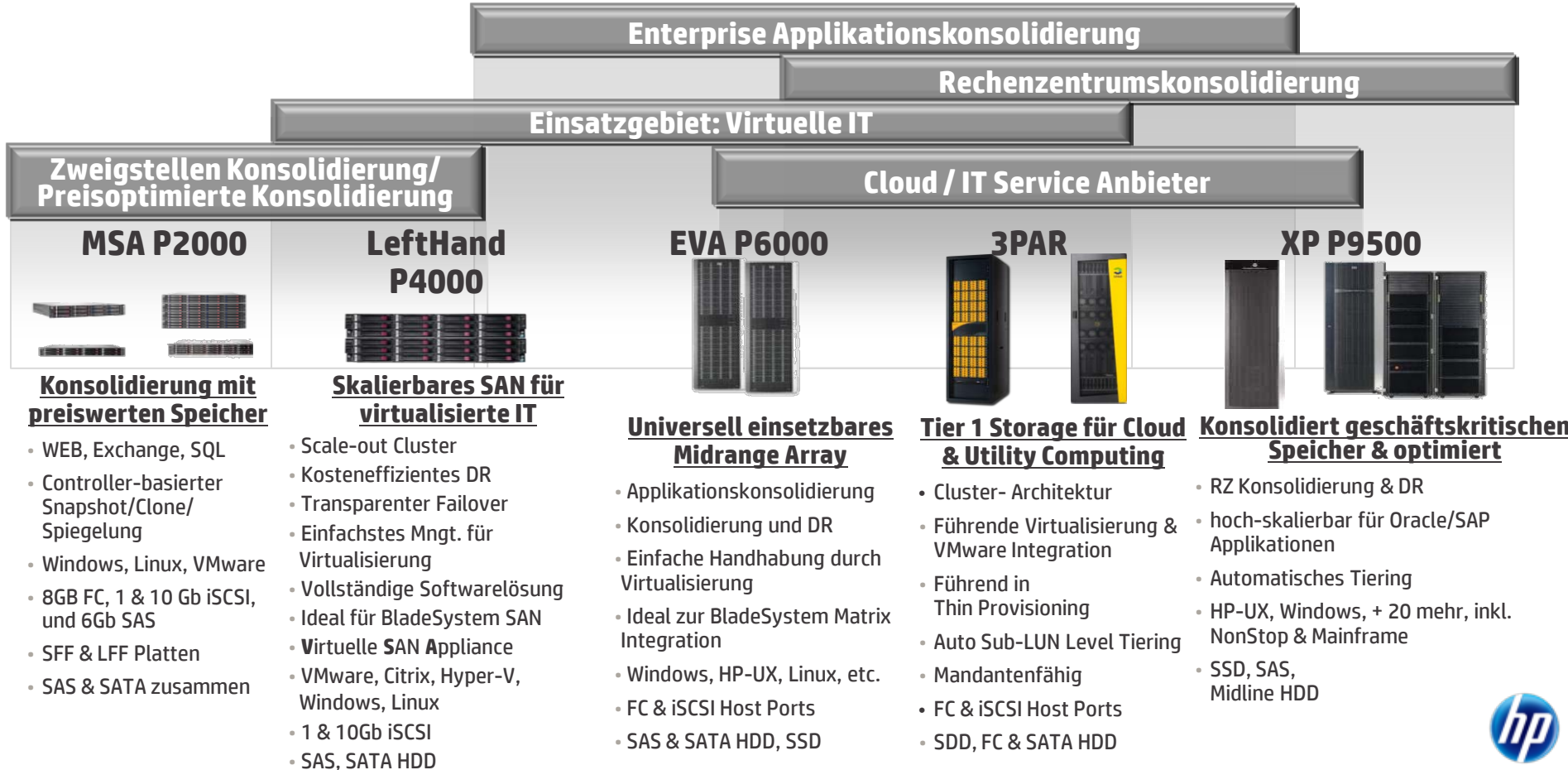
Implementieren

Integrieren

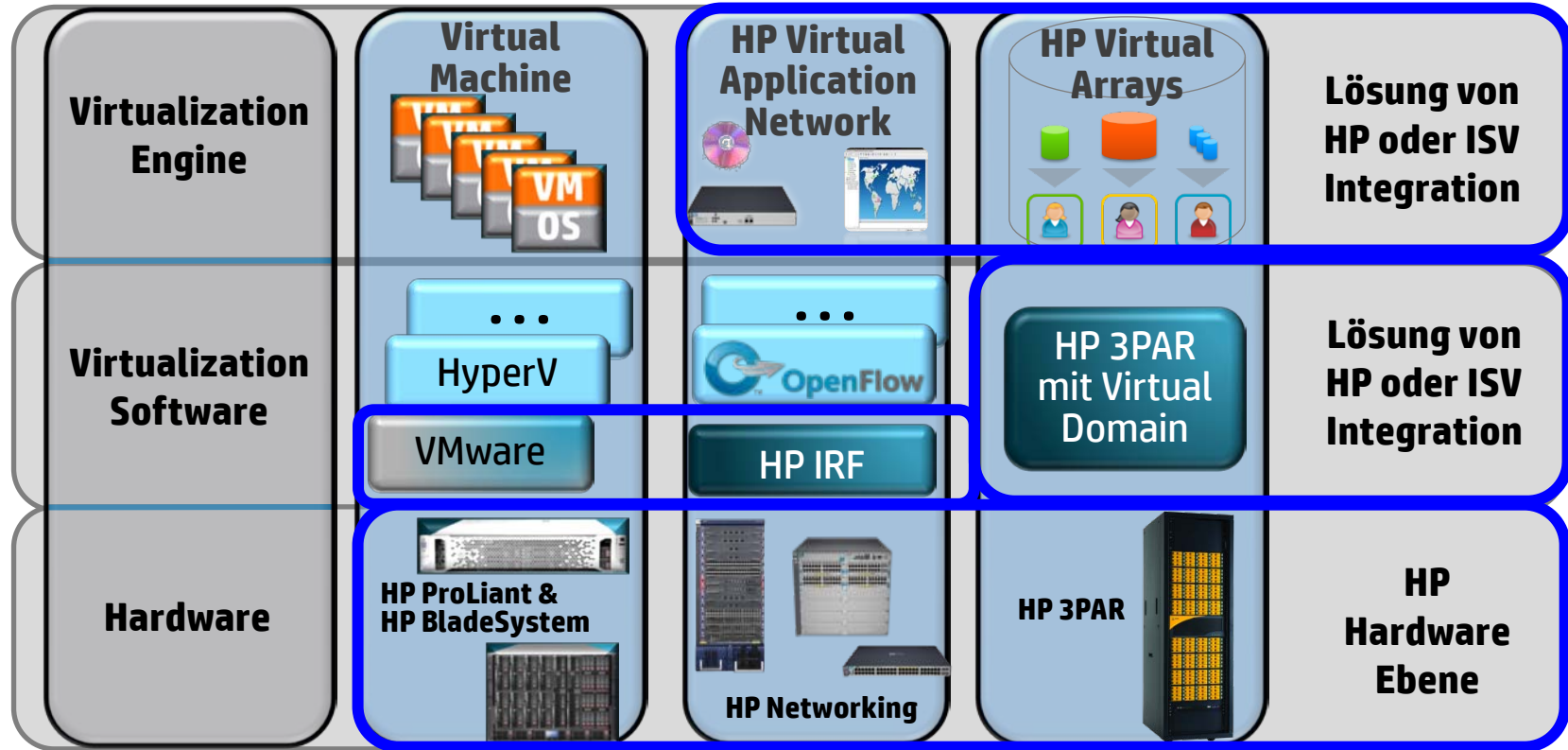
Verwalten

Weiterentwickeln

# HP Storage Store Portfolio – SAN Storage



# Virtuelle IT heute und in der Zukunft



*Vielen Dank*

