SYMMETRIX DMX-4
mit Flash Laufwerken und
MC V.5773

SYMMETRIX DMX4

EMC² where information lives®

~ 19.000 Systeme

~ 4.600 Kunden

Direct Matrix

- Bis zu 256 direkte Pfade
- Bis zu 256 GB/s Datenbandbreite
- Bis zu 12,8 GB/s Kontrollbandbreite

Dynamischer Cache

- Bis zu 1 TB Cache netto
- Intelligentes “Adaptive Pre-fetch”
- QoS, LPAR und vieles mehr
- 3,2 Mio. I/Os vom und zum Cache

MC V.5873 “Enginuity”

- Aktuell funktionsreichstes Betriebssystem – nicht nur auf Speichergeräten
- Entworfen für kontinuierlichen Betrieb, “24x365¼”
- Selbstoptimierend
- Fortschrittliche erwaltung
- 100% SMI-S kompatibel

Prozessoren

- Bis zu 196 Dual-Core PowerPCs
- Vier Prozessoren pro Direktor
- Über 830.000 Dhrystone MIPS

Flexibles Backend

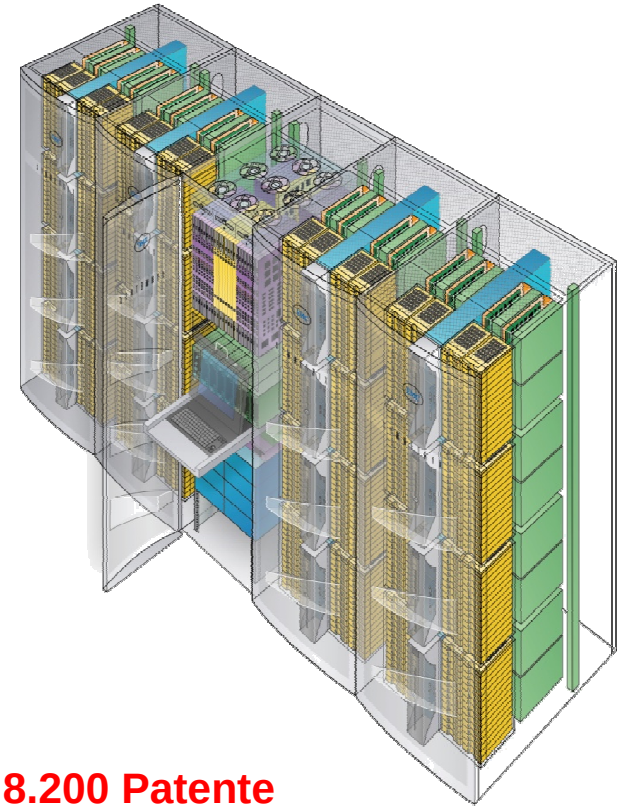
- Bis zu 128 Fibre Channel Pfade
- RAID 0, 1, 1 + 0, 5, 6
- 73, 146 GB SSD
- 146, 300, 400, 500 GB 10k U/min
- 146, 300, 450 GB 15k U/min
- 1000 GB S-ATA

Fehlertolerantes Design

- “NDU everything”
- Volle dreifache Redundanz aller Komponenten
- Jede Platte mit zwei FC-Ports
- Globale Hot Spare Platten
- Dreifach redundante Stromversorgung und eigene Batterieversorgung
- Fernunterstützung und Call-home-support

Homogene Umgebungen

- 42 Betriebssysteme
- 105 Formatierungsarten



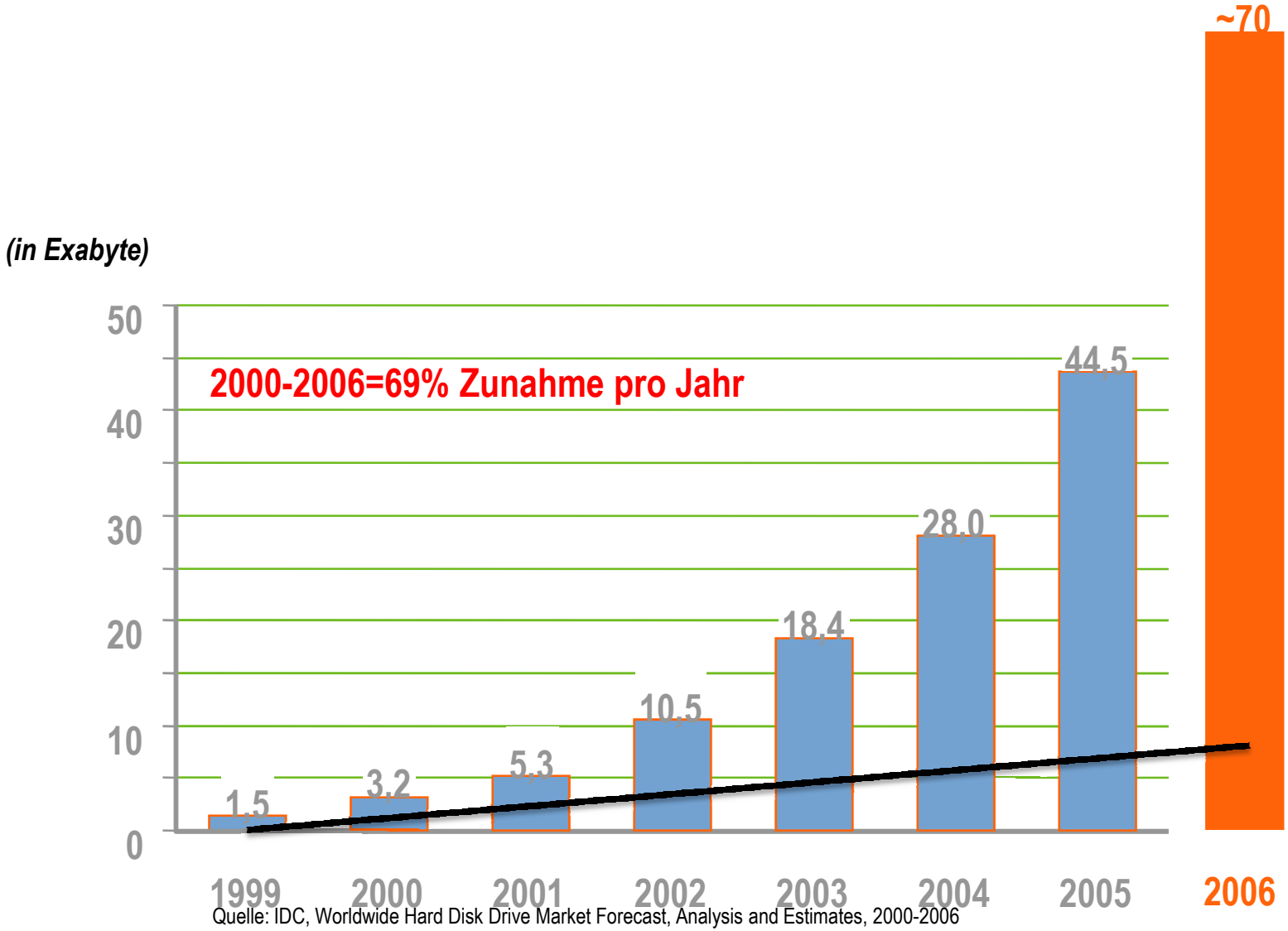
~ 8.200 Patente

In 190 Ländern auf allen Kontinenten

Aber:

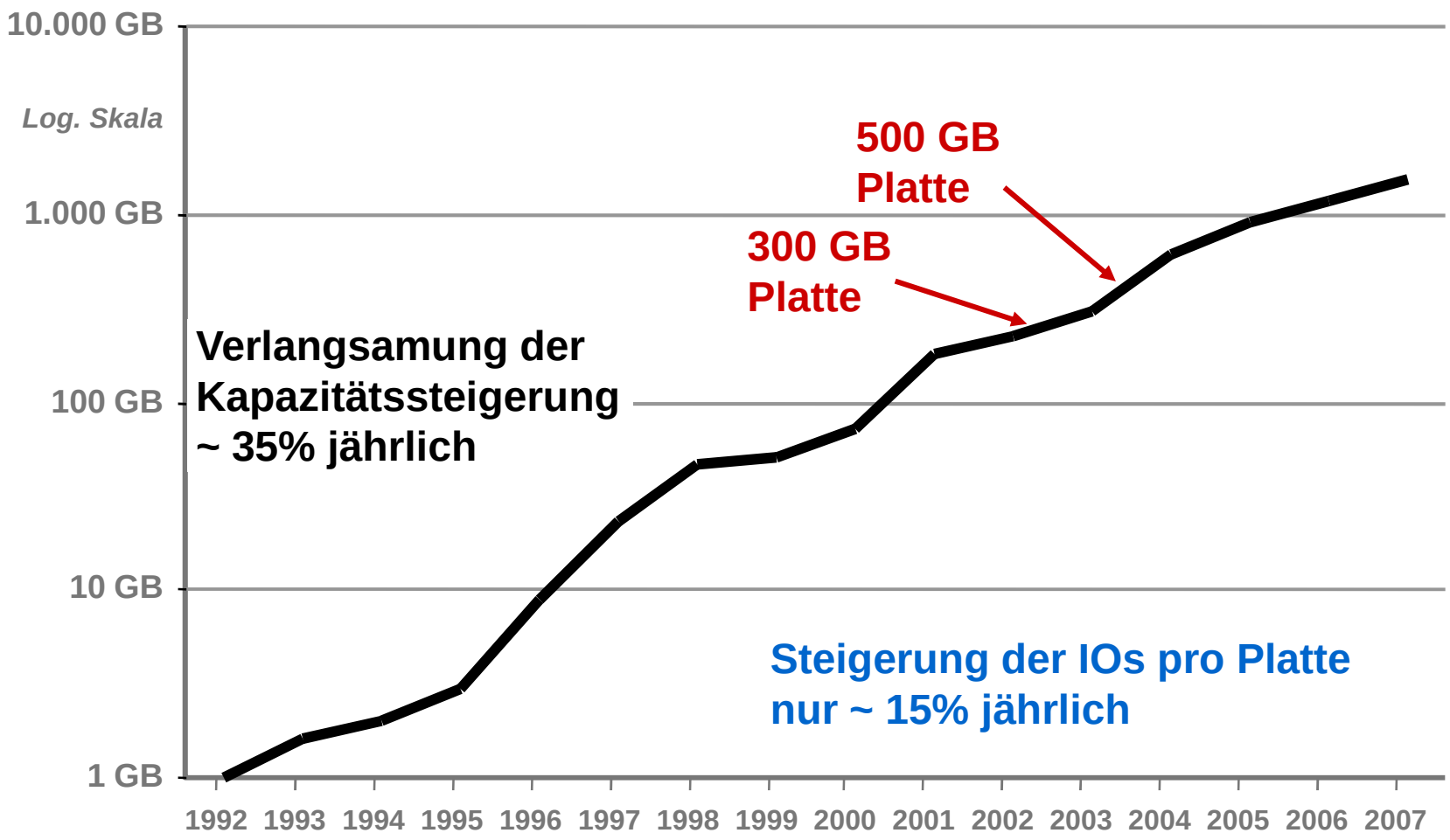
Magnetische Aufzeichnungsverfahren können mit der Menge an Informationen nicht Schritt halten

Installierter Plattenspeicher

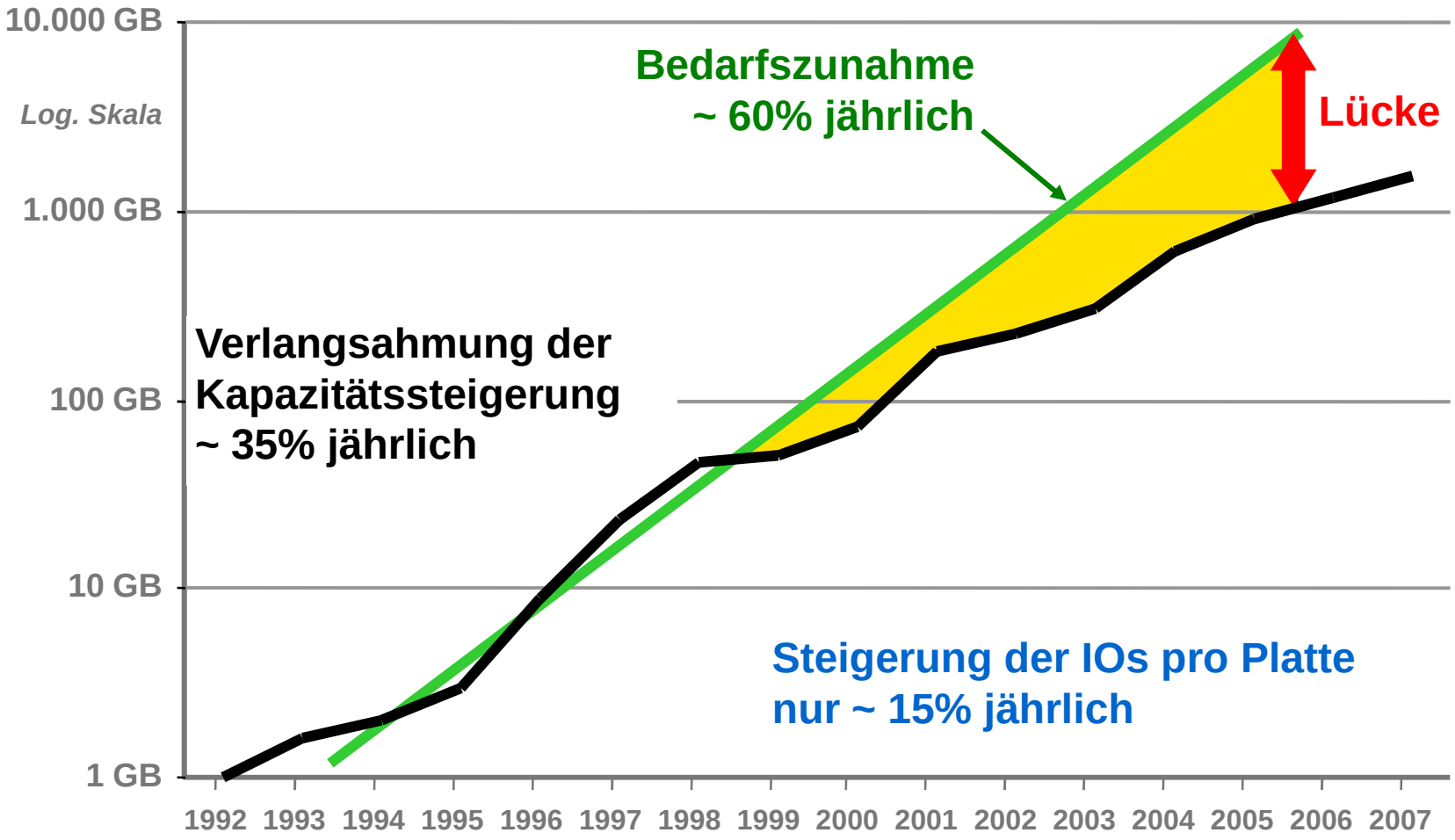


Quelle: IDC, Worldwide Hard Disk Drive Market Forecast, Analysis and Estimates, 2000-2006

Kapazitäten von Magnetplatten



Kapazitäten von Magnetplatten



Consumer vs. Enterprise Flash Technologie

EMC² where information lives®

Consumer Flash SSD

- Multi-Level-Cell (MLC), NAND oder NOR
- Speichert mehr als ein Bit in einer Zelle
- Typischerweise in leseintensiven Umfeldern genutzt
- Optimiert für Leseleistung, generell schlechte Schreibleistung
- Selten mit Ersatzzellen ausgerüstet
- Kürzere Lebendauer
- Beispiele: Memory Sticks, MP3-Player

Enterprise Flash SSD

- Single-Level-Cell (SLC), NAND
- Speichert ein Bit pro Zelle
- Duale Laufwerksschnittstelle
- Parallele I/Os über beide Kanäle
- Höhere Transferraten
- Längere Lebensdauer (typisch 200.000 Schreib-I/Os pro Zelle)
- Beispiel: EMC DMX-4 Flash Laufwerke

Verbesserung der Lebensdauer

EMC² where information lives®

Write Leveling

- Zellen haben eine begrenzte Anzahl von Schreibzyklen
- Write leveling erlaubt eine gleichmäßige Nutzung der Zellen, auch wenn der Rechner auf dieselben Adressen schreibt

Zusatzkapazität

- Jedes Flash Laufwerk verfügt über zusätzliche Kapazitäten
  - Verbessert die Leistung des Write Leveling Algorithmus
  - Ersetzt verbrauchte Zellen

SYMMETRIX DMX-4 Flash Laufwerke

EMC² where information lives®

Gleicher Formfaktor wie 3,5” HH Laufwerke

- Nutzt die gleichen Gehäuse wie Festplatten
- Erste Kapazitäten 73 GByte und 146 GByte
- FC-Schnittstelle mit zwei Kanälen und 4 Gbit/s

Erfordert SYMMETRIX DMX-4 und MC V.5773

- Verfügbar für alle SYMMETRIX DMX-4 Systeme

SYMMETRIX DMX-4 Flash Laufwerke

EMC² where information lives®

Was bringen Flash Laufwerke?

Sehr kurze Zugriffszeiten

- Passend für Hochleistungsanwendungen

Hohe I/O-Raten aus relativ geringer Kapazität

- Passend für Anwendungen mit hoher Zugriffsdichte

Überlegungen

Erste Versionen haben eine 2 Gbit/s-Schnittstelle

- Quadranten mit Flash-Laufwerken haben durchgehend 2 Gbit/s

Unterstützt einige EMC-spezifische Kommandos nicht

- Keine XOR-Errechnung auf den Platten
- Kein Secure Erase

Flash Laufwerke in Inline-Kommandos

EMC² where information lives®

```
SYSTEM PHYSICAL DRIVES INFORMATION

DIR I/F  VENDOR  CODE  L#   L1    CAPACITY USED     SERIAL #              BLKSZ PART #
-----------------------------------------------------------------------------------------
15d D/02 T146155 5507   5   3E M1 11177323 01ADE1A0 3LN2SNG9               0200 072000000
15d D/04 T146155 5507   9   CA m3 11177323 021C0A00 3LN2NX08               0200 072000000
15d D/06 T146155 5507   8   DA m3 11177323 021BD520 3LN2STWY               0200 072000000
15d D/08 T146155 5507   9   D2 m3 11177323 021C0A00 3LN2SZB7               0200 072000000
15d D/0A T146155 5507   9   C2 m3 11177323 025F5200 3LN10EHH               0200 072000000
15d D/0C T146155 5507   8   E1 m3 11177323 021BD520 3LN1GVH8               0200 072000000
16d C/00 T146155 5507   5   30 M1 11177323 01ADE1A0 3LN2RMER               0200 072000000
16d C/02 T146155 5507   5   32 M1 11177323 01ADE1A0 3LN2PGXB               0200 072000000
16d C/04 STC0146 5CC1  32   40 m3 088BB991 086F8A20 0x00000B7C              0200 N/A
16d C/06 T146155 5507   9   D0 m2 11177323 021C0A00 3LN2S8YT               0200 072000000
16d C/08 T146155 5507   9   C8 m2 11177323 021C0A00 3LN2SG73               0200 072000000
16d C/0A T146155 5507   8   D8 m2 11177323 021BD520 3LN2QCMK               0200 072000000
16d C/0C T146155 5507   9   C0 m2 11177323 025F5200 3LN11AN8               0200 072000000
16d C/0E T146155 5507  32      HS 11177323 110B2100 3LN16CPY               0200 072000000
16d D/01 T14
16d D/03 T14
16d D/05 T14
16d D/07 T14
DAF-  1b : W
```

```
UTILITY D2 -- Display PH DV RECORD of target N  : TIME: JAN/08/08 01:27:36
-------------------------------------------------

DEVICE #  40; FC Loop D; Loop ID # 04; DEVICE TYPE: STC0146  FBA  DEVICE
-------------------------------------------------------------------
VENDOR INFORMATION : -- STEC    ZeusIOPS        5CC1 SN: 0x00000BC6
ALPA: E1  WWN: 5000A72000000BC6  Present: Yes      Bypassed: No
FC Login State: [02]_....._ConnFail_ConnComp_ConnSnd  Port Speed: 2G
FC Login State Time: JAN/06/08 23:32:07  DISK SPEED(RPM): 200
CAPACITY_BLKS/MB: 088BB991 / 70007  FLAGS : 00  BLCK IO : 0000
DEFECTS PRIMARY/GROWN : - 0 / 0  IO/FATAL ERRORS : 1491 / 0
RECOVER/UNRECOVER ERRORS : - 0 / 0  PERMANENT FATAL ERRORS : 0 JOBS/QUEUED : 0 / 0
XOR FAILURES :  0   SKIP MASK REJECTIONS : 0  LONG I/O: 00/2c (26b11795)
EXPECTED BLOCK SIZE : 200     ACTUAL BLOCK SIZE : 200   MAX Q DEPTH : 128

Colors:  NORMAL  Physical Gap  Physical Overlap
LV#0000   40 5m1  BLKS: 00000100-00436AFF | LV#0001   41 5m2  BLKS: 00437DE0-0086E7DF
LV#0002   42 5m3  BLKS: 0086FAC0-00CA64BF | LV#0003   43 5m0  BLKS: 00CA77A0-010DE19F
LV#0004   44 5m1  BLKS: 010DF480-01515E7F | LV#0005   45 5m2  BLKS: 01517160-0194DB5F
LV#0006   46 5m3  BLKS: 0194EE40-01D8583F | LV#0007   47 5m0  BLKS: 01D86B20-021BD51F
LV#0008   48 5m1  BLKS: 021BE800-025F51FF | LV#0009   49 5m2  BLKS: 025F64E0-02A2CEDF
```

Flash Laufwerke in SymmWin-Fenstern

EMC² where information lives®

ID	DA 1				DA 2			
	DA 16				DA 15			
	A0	A1	B0	B1	A0	A1	B0	B1
0	A	A	A	A	A	A	A	A
1	A	A	A	A	A	A	A	A
2	A	A	A	A	A	A	A	A
3	A	A	A	A	A	A	A	A
4	B	B			B	B		
5	C		B	B			B	B
6								

ID	Name	Blocks	Fmt	Spare	Group	type	speed	BaSpeed	count
A	T146155	11177323	512	NO	0	fiber	15K	4G	32
B	STC0073	88BB991	512	NO	1	flash	90K	2G	8
C	STC0073	88BB991	512	YES	0	flash	90K	2G	1

DMX-4 950 mit neun Flash Laufwerken

ID	FDA-Link		
dirs	Dual Initiator	0	1
01a	10	2G	2G
01b	20	2G	2G
01c	30	2G	2G
01d	40	2G	2G
02a	0F	4G	4G
02b	1F	4G	4G
02c	2F	4G	4G
02d	3F	4G	4G
15a	02	4G	4G
15b	12	4G	4G
15c	22	4G	4G
15d	32	4G	4G
16a	01	2G	2G
16b	11	2G	2G
16c	21	2G	2G
16d	31	2G	2G

Setzt die Loop auf 2 Gbit/s in DirEdit

ID	DA 1							
	DA 16							
	A0	A1	B0	B1	C0	C1	D0	D1
0	A	A	A	A	A	A	A	A
1	A	A	A	A	A	A	A	A
2	A	A	A	A	A	A	A	A
3	A	A	A	A	A	A	A	A
4	C	B		B		B	B	
5								

ID	Name	Blocks	Fmt	Spare	Group	type	speed	BaSpeed	count
A	T146155	11177323	512	NO	0	fiber	15K	4G	32
B	STC0073	88BB991	512	NO	1	flash	90K	2G	4
C	STC0073	88BB991	512	YES	0	flash	90K	2G	1

DMX-4 1500 mit fünf Flash Laufwerken

Flash Laufwerke im Solutions Enabler

EMC² where information lives®

Flash drive

```
SEGapps5:/bench/work # symdisk list -disk_group 1 -sid 466

Symmetrix ID                    : 000190103466
Disks Selected                  : 240

                                                                  Capacity(MB)
Ident  Symb Int TID Vendor     Type       Hypers   Total    Free    Actual
------ ---- --- --- ---------- ---------- ------ -------- -------- --------
DF-16A 16A    D    4 STEC       STC0146        32    70007      899    70007
DF-1B  01B    D    4 STEC       STC0146        32    70007      899    70007
DF-16C 16C    D    4 STEC       STC0146        32    70007      899    70007
DF-16D 16D    C    4 STEC       STC0146        32    70007      899    70007
SEGapps5:/bench/work #
```

HDD

```
SEGapps5:/bench/work # symdisk list -disk_group 2 -sid 466

Symmetrix ID                    : 000190103466
Disks Selected                  : 240

                                                                  Capacity(MB)
Ident  Symb Int TID Vendor     Type       Hypers   Total    Free    Actual
------ ---- --- --- ---------- ---------- ------ -------- -------- --------
DF-15A 15A    C    4 SEAGATE    T146155        32   140014    70906   140014
DF-15B 15B    D    4 SEAGATE    T146155        32   140014    70906   140014
DF-2C  02C    D    4 SEAGATE    T146155        32   140014    70906   140014
DF-2D  02D    C    4 SEAGATE    T146155        32   140014    70906   140014
SEGapps5:/bench/work #
```

Leistung: Flash vs. HDD

EMC² where information lives®

Read Miss I/Os

- Flash Laufwerk: 1 ms Lieferzeit
- 15k HDD: 6 ms Lieferzeit
- 10k HDD: 9 ms Lieferzeit

Sequential Read I/Os

- Nutzt prefetch und den Cache, Transport mit DRAM-Geschwindigkeit
- Flash Laufwerk und HDD mit gleicher Leistung bei Cache Hits

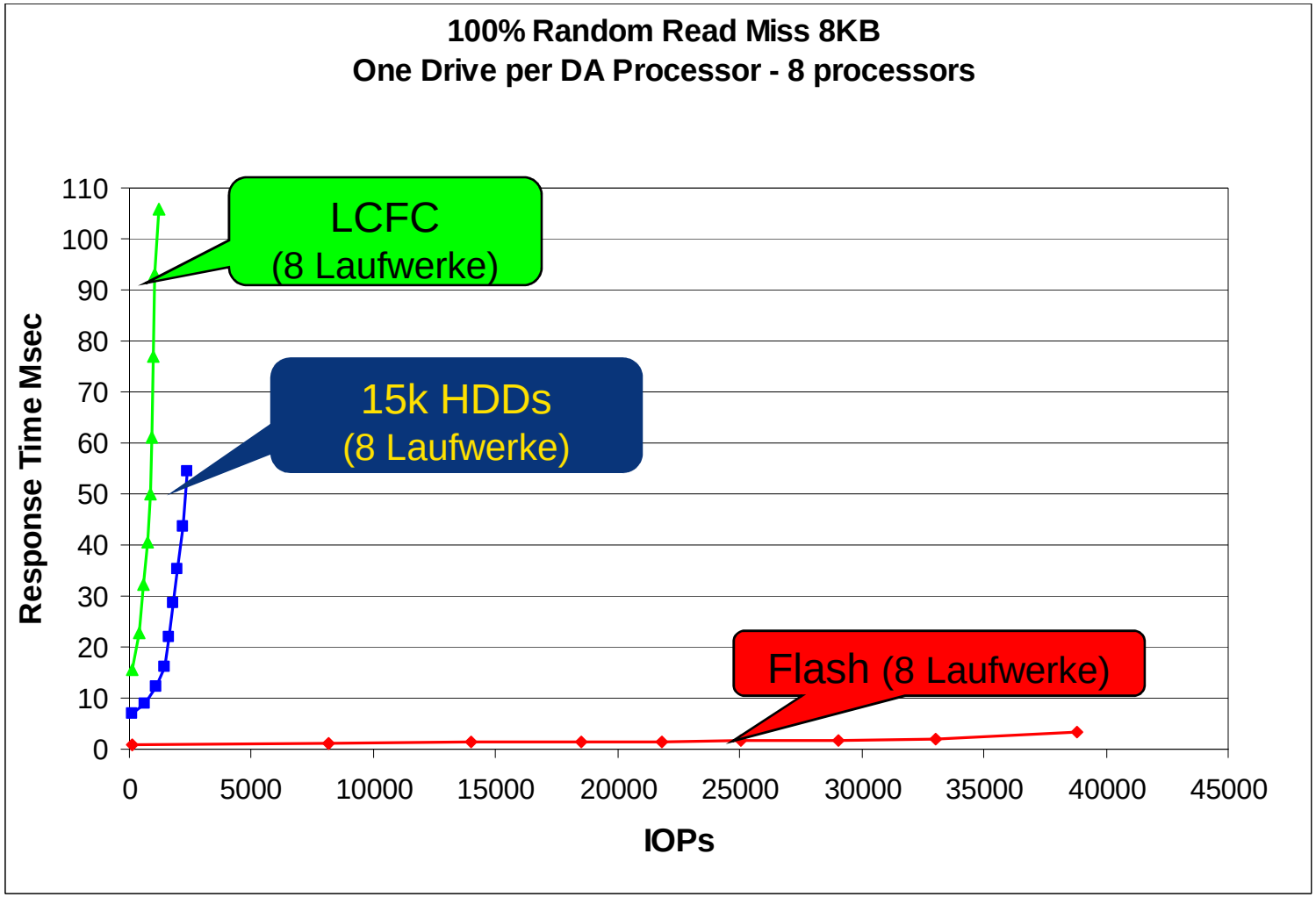
Write I/Os

- Nutzt den Cache, Transport mit DRAM-Geschwindigkeit
- Flash Laufwerk und HDD mit gleicher Leistung
- Bei “Disk bottlenecked heavy writes”: Flash Laufwerke sind besser
- Bei “DA bottlenecked heavy writes”: HDDs sind besser

Leistung: Flash vs. HDD

100% Random Read Miss 8 kByte

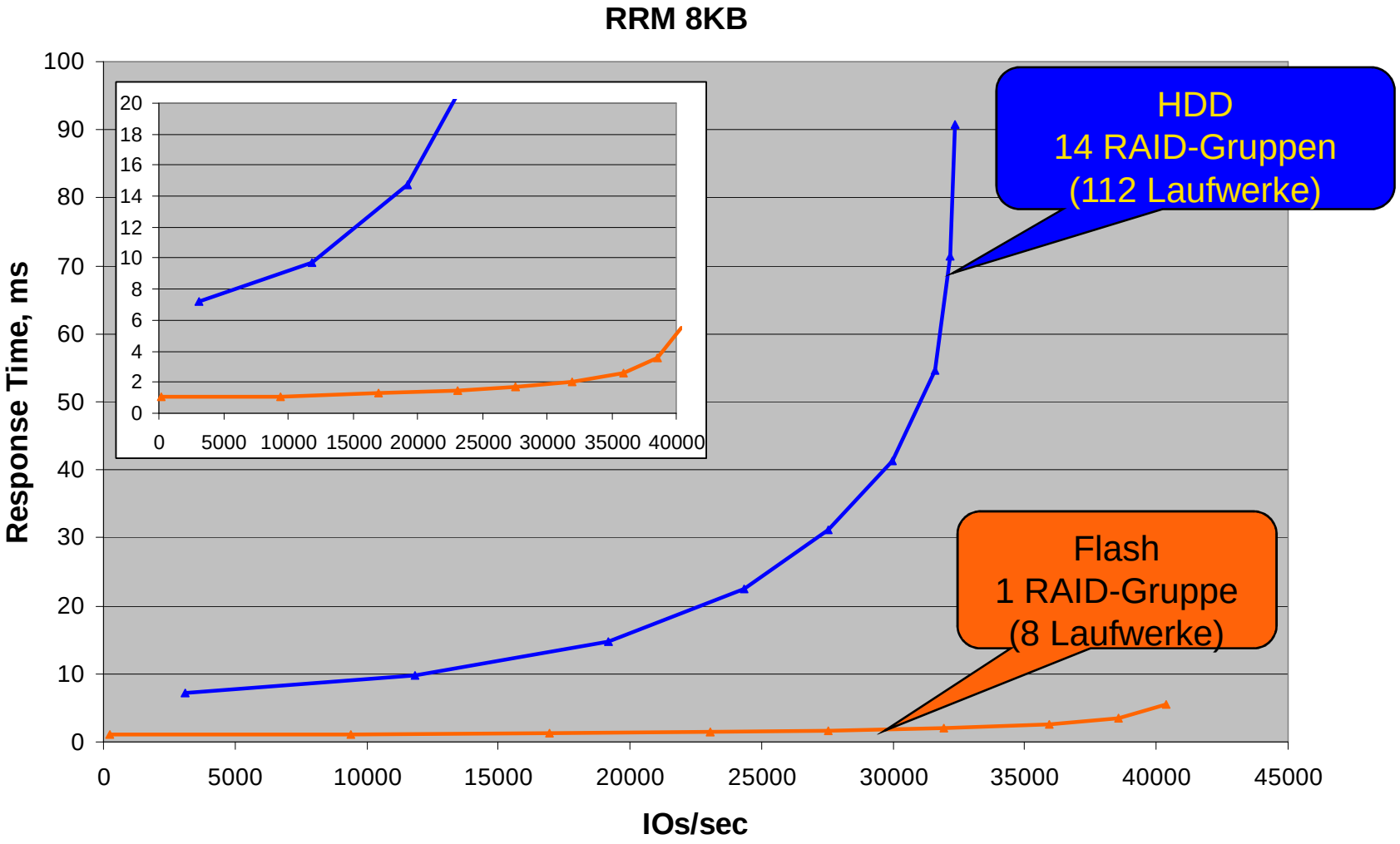
Eine RAID-Gruppe (8 Laufwerke) pro Test



Leistung: Flash vs. HDD

100% Random Read Miss 8 kByte

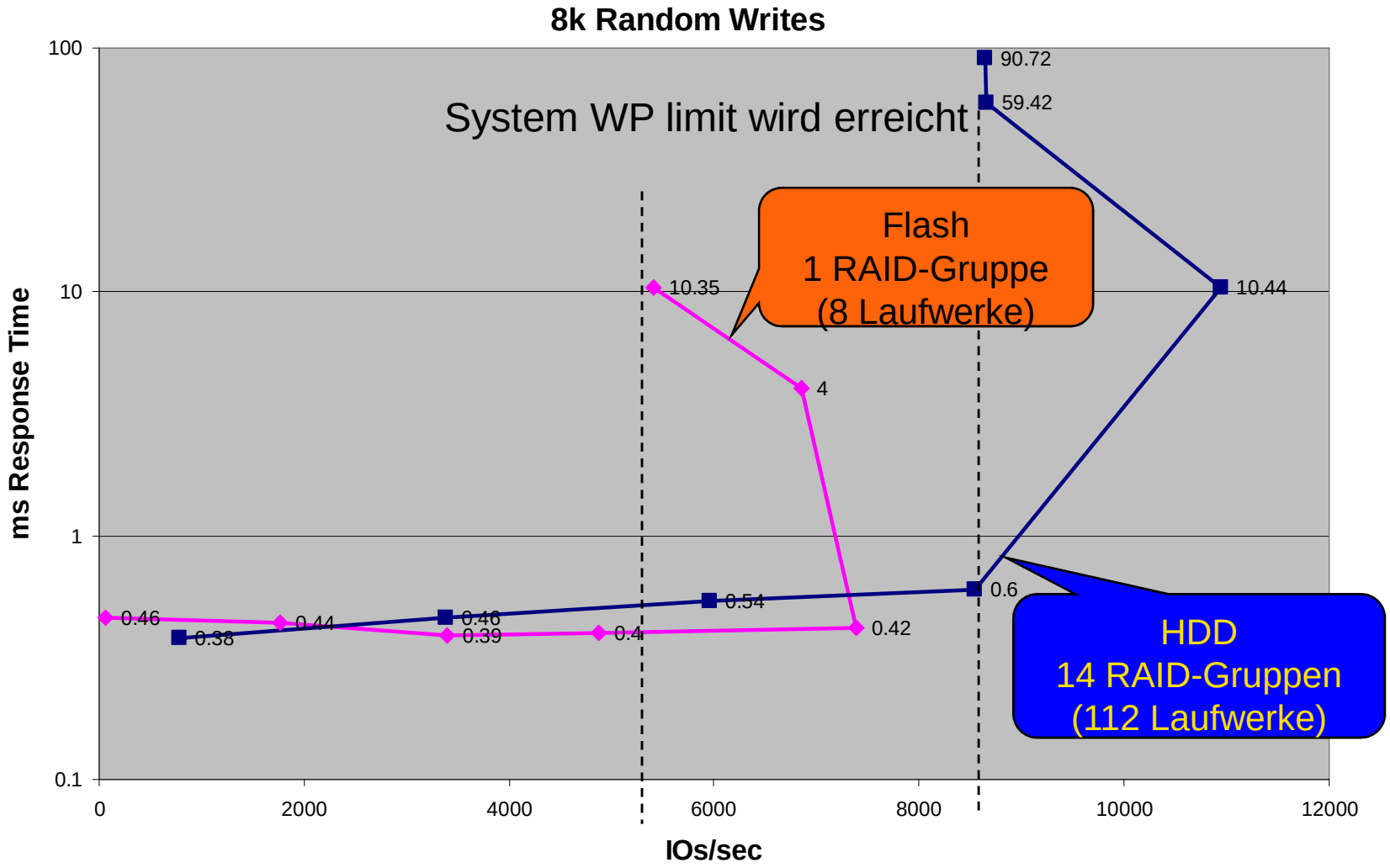
Eine RAID-Gruppe Flash und 14 RAID-Gruppen HDDs



Leistung: Flash vs. HDD

100% Random Write 8 kByte

Eine RAID-Gruppe Flash und 14 RAID-Gruppen HDDs

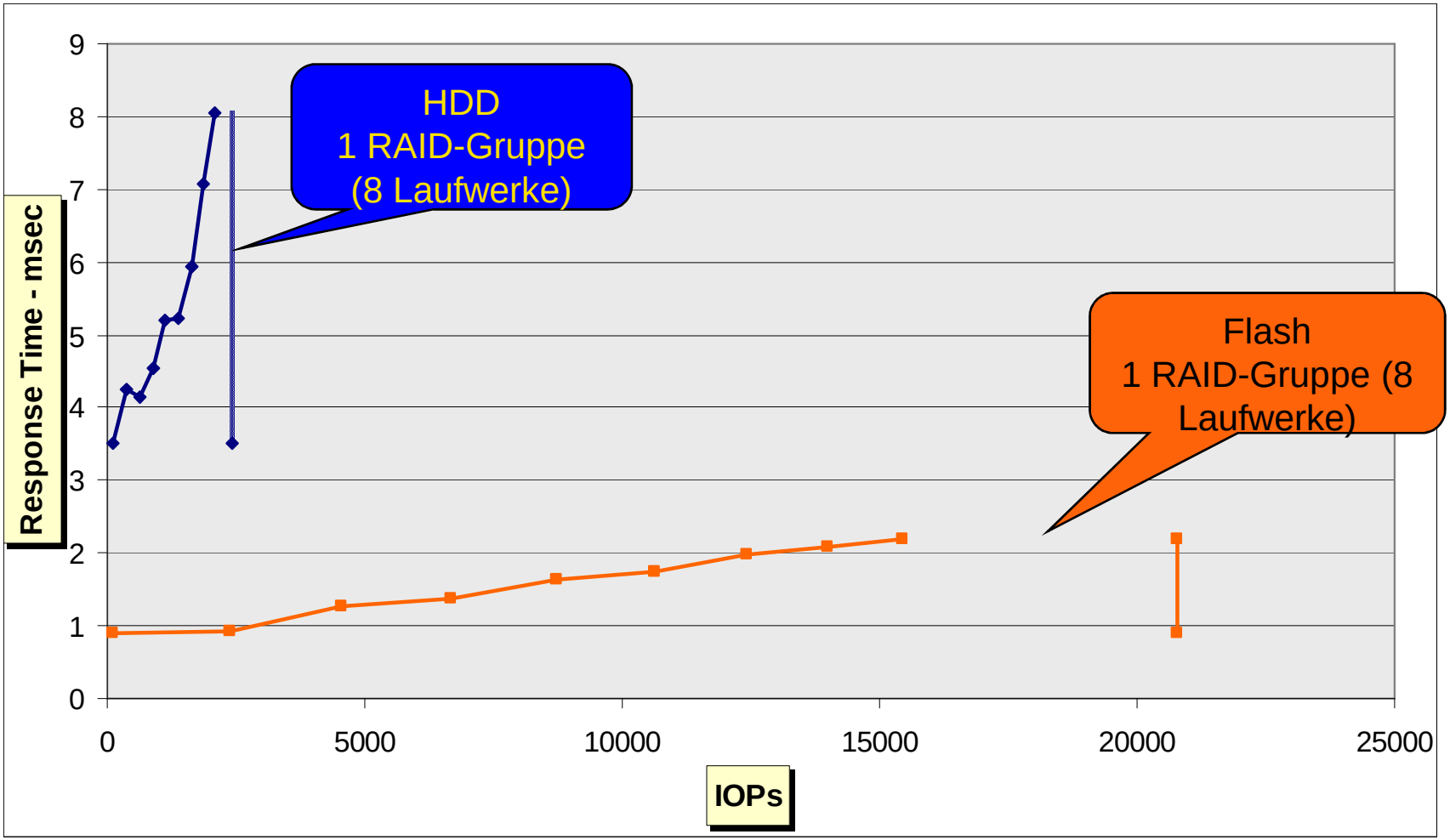


Leistung: Flash vs. HDD

EMC² where information lives®

Gemischtes I/O Profil: OLTP2

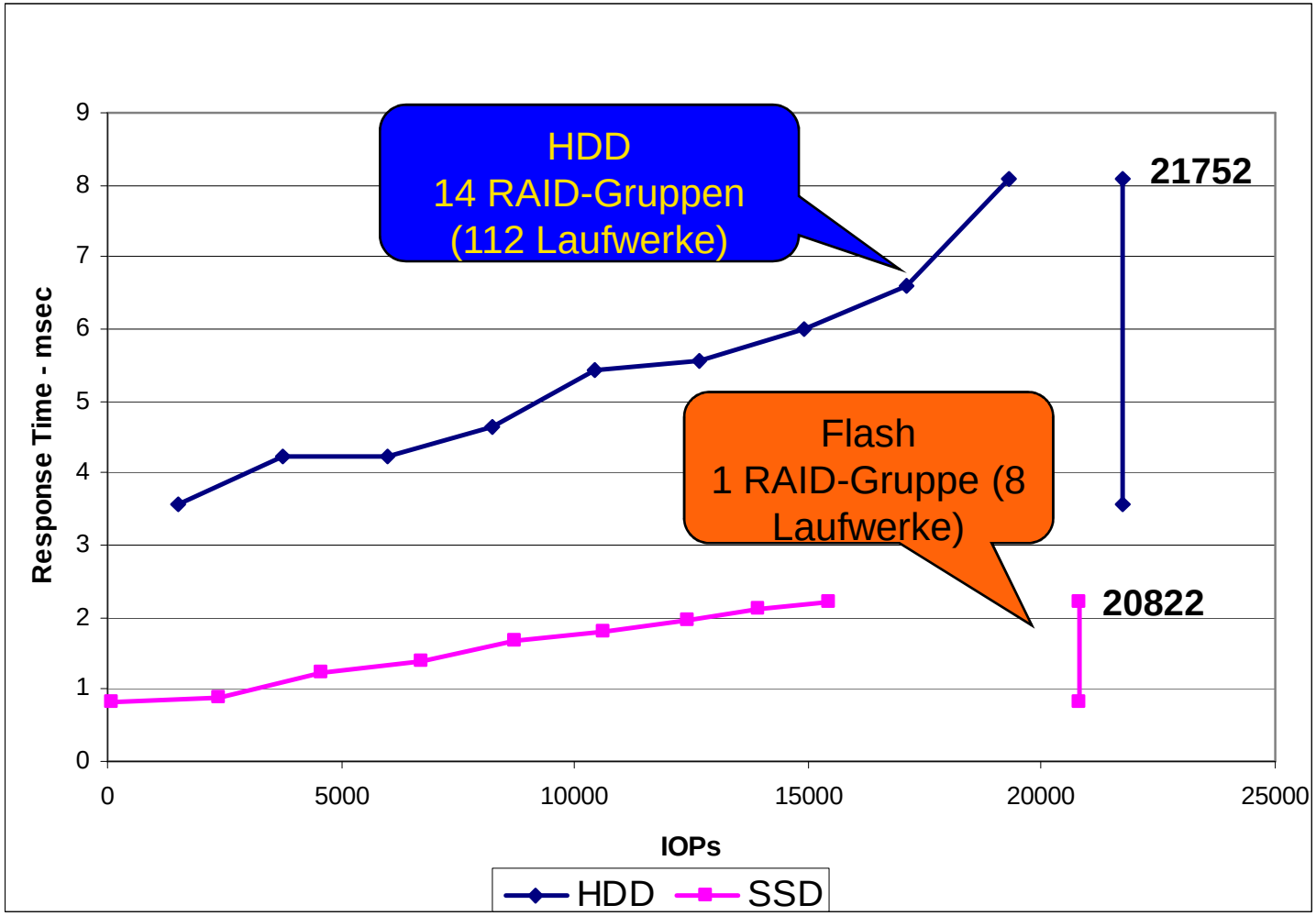
Eine RAID-Gruppe (8 Laufwerke) pro Test



Leistung: Flash vs. HDD

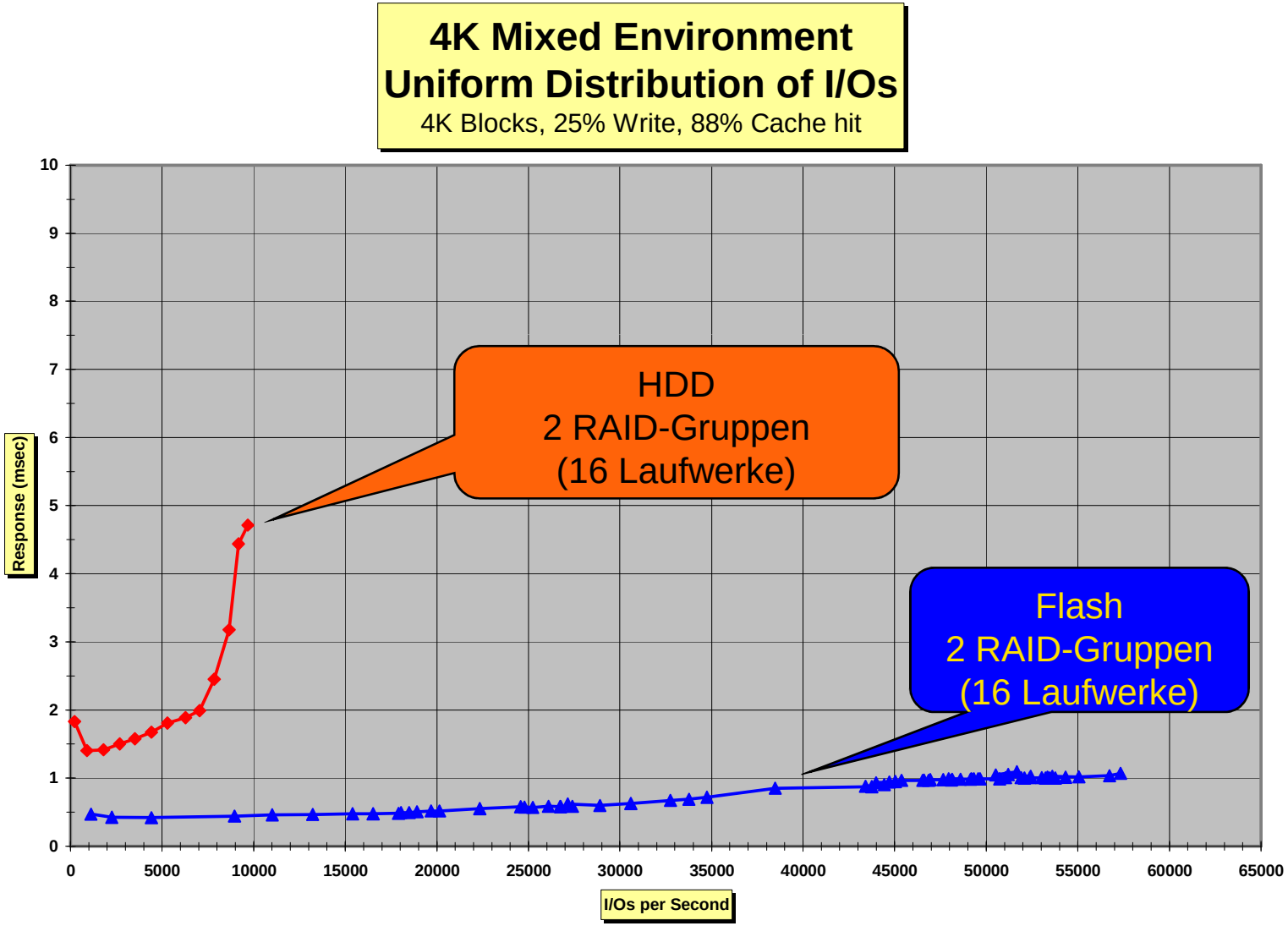
Gemischtes I/O Profil: OLTP2

Eine RAID-Gruppe Flash und 14 RAID-Gruppen HDDs



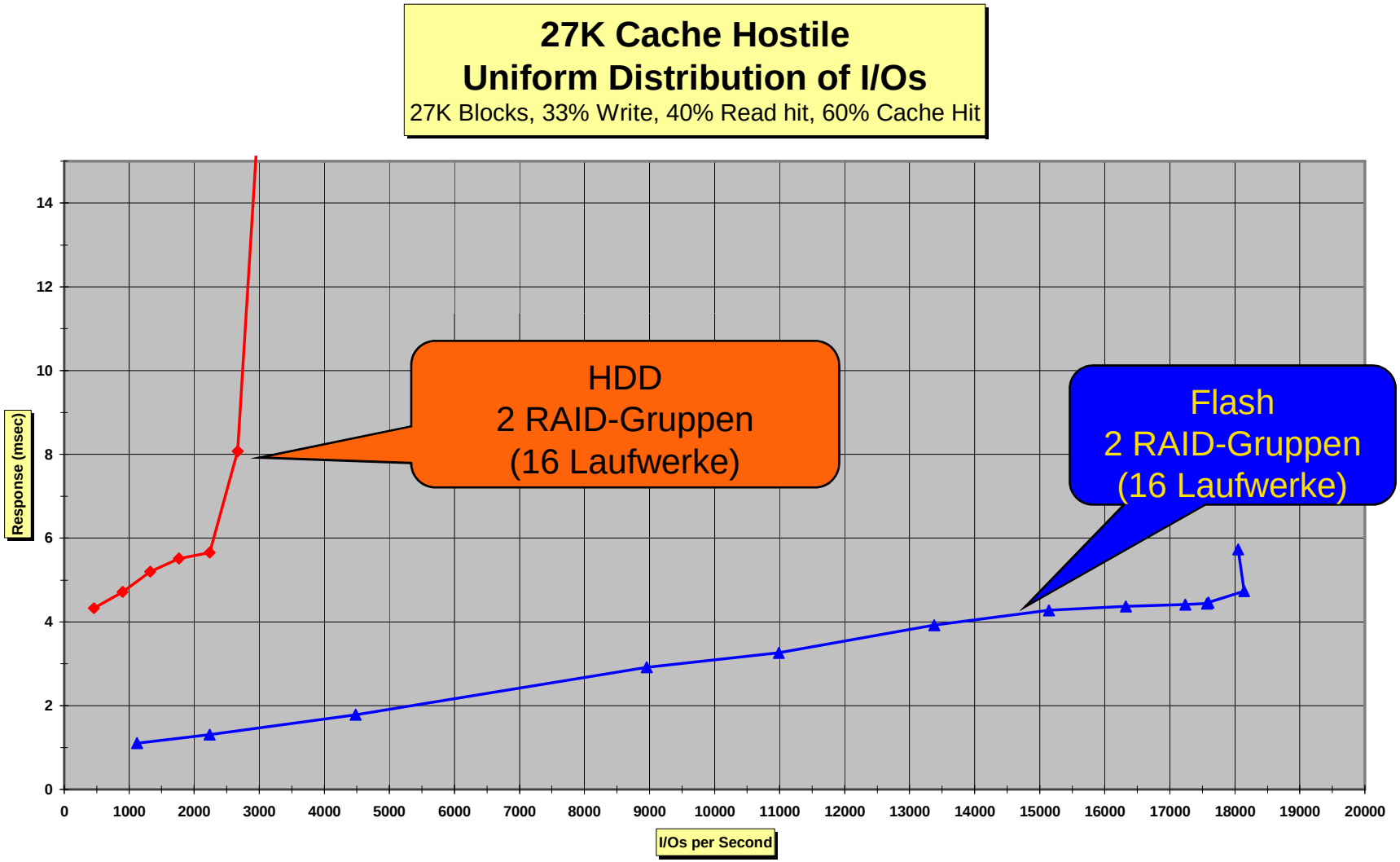
Leistung: Flash vs. HDD

Mainframe gemischter, cache-freundlicher I/O



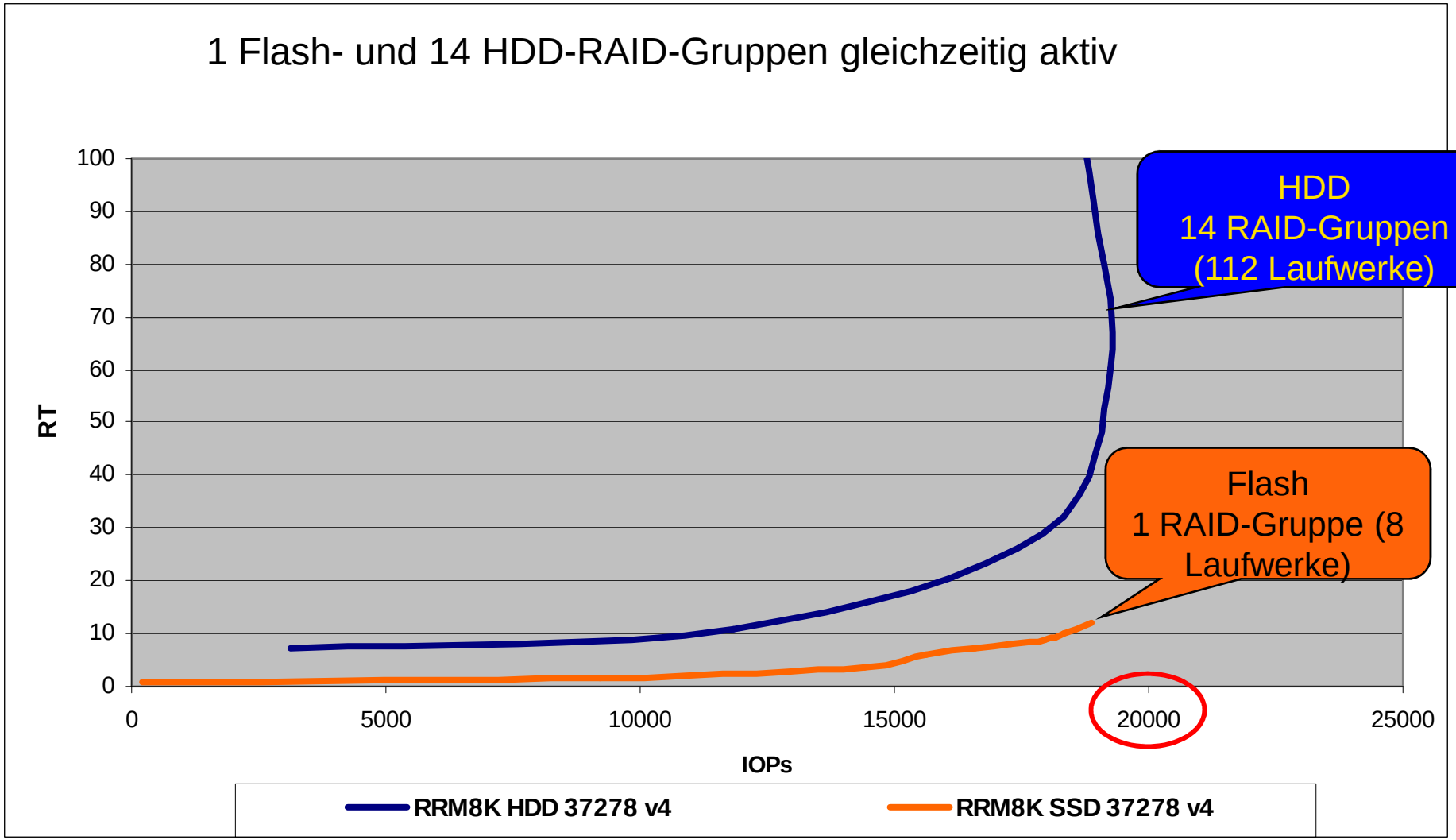
Leistung: Flash vs. HDD

Mainframe gemischter, cache-feindlicher I/O



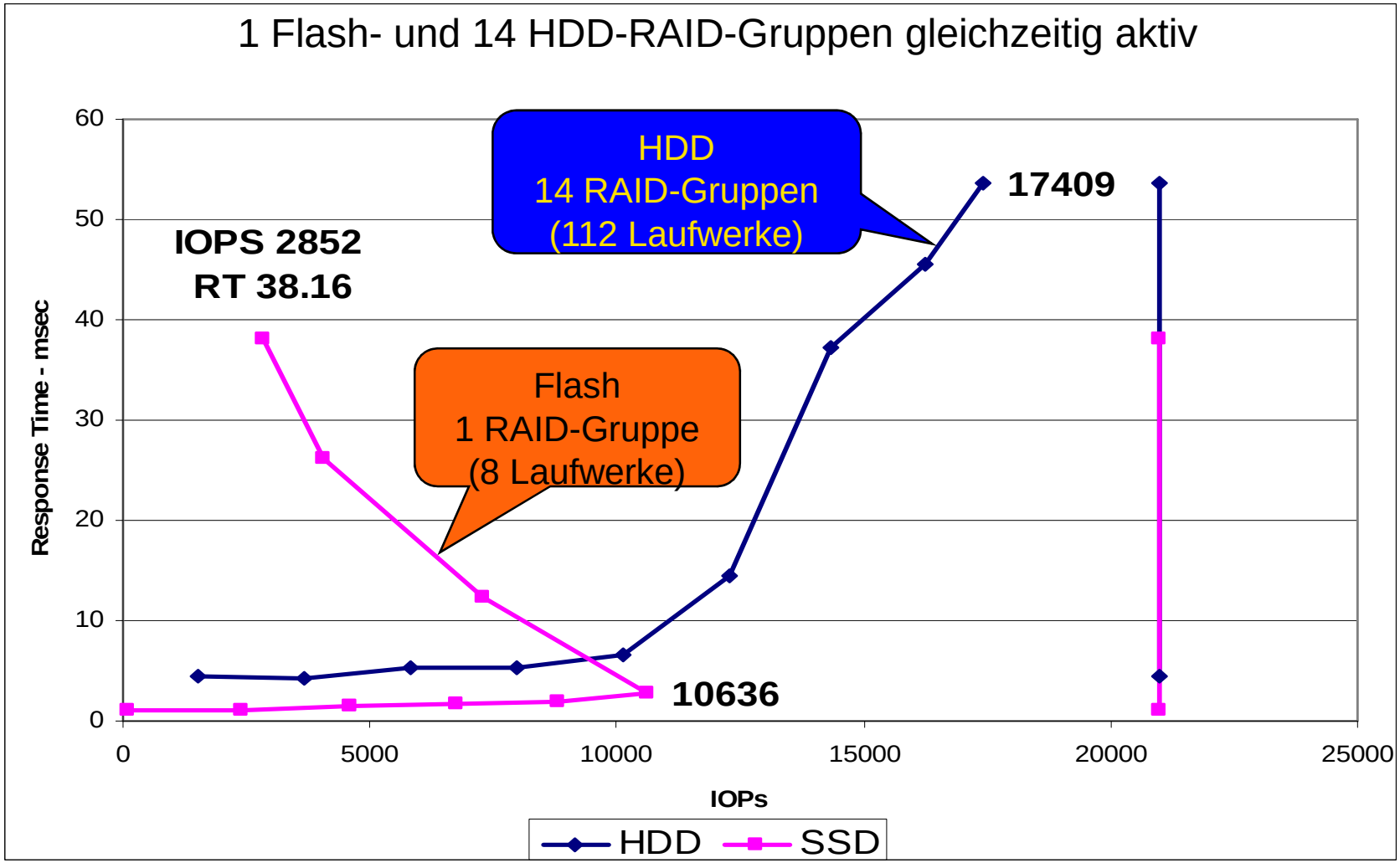
Leistung: Flash vs. HDD

100% Random Read Miss 8 kByte



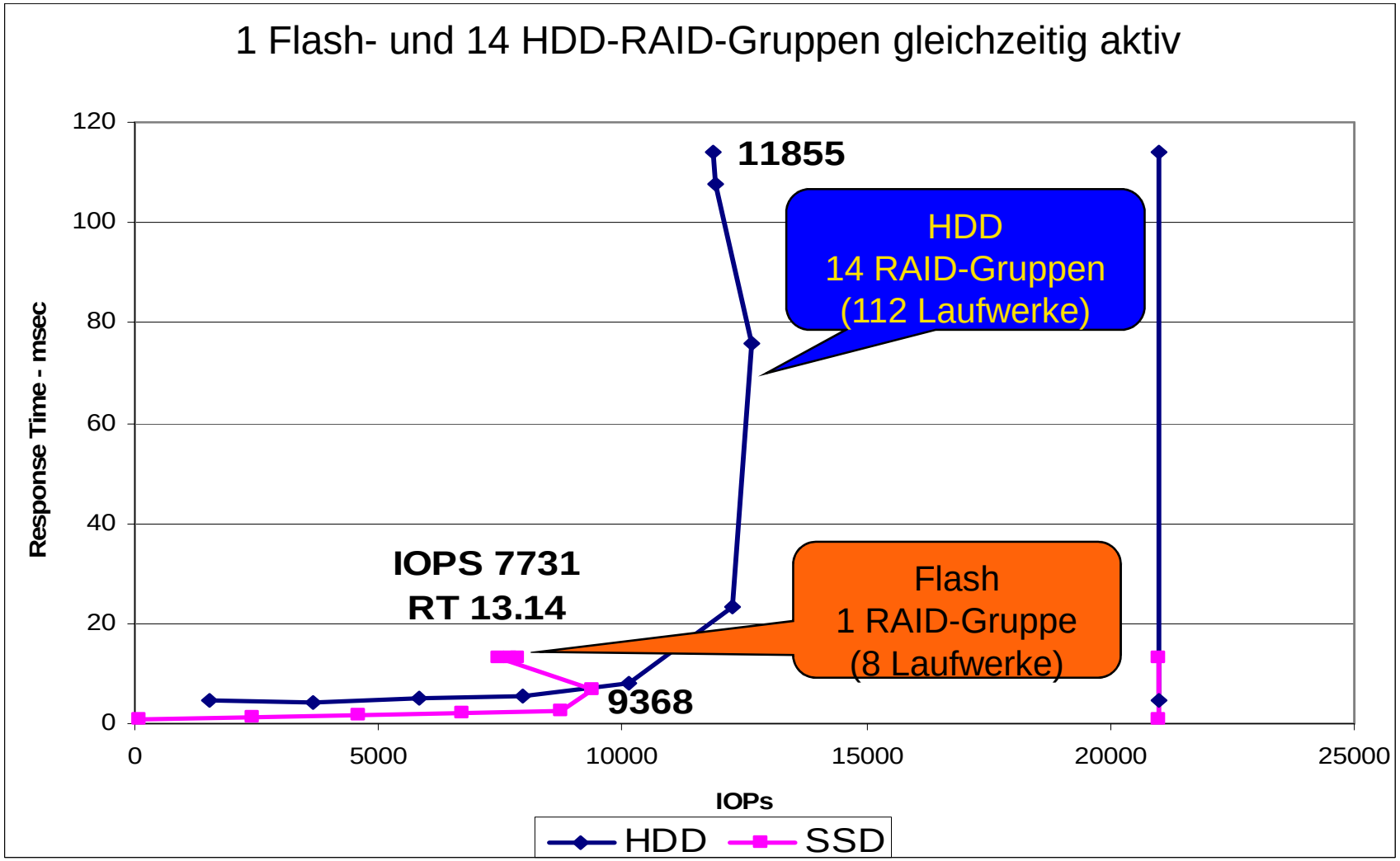
Leistung: Flash vs. HDD

OLTP2 Arbeitslast

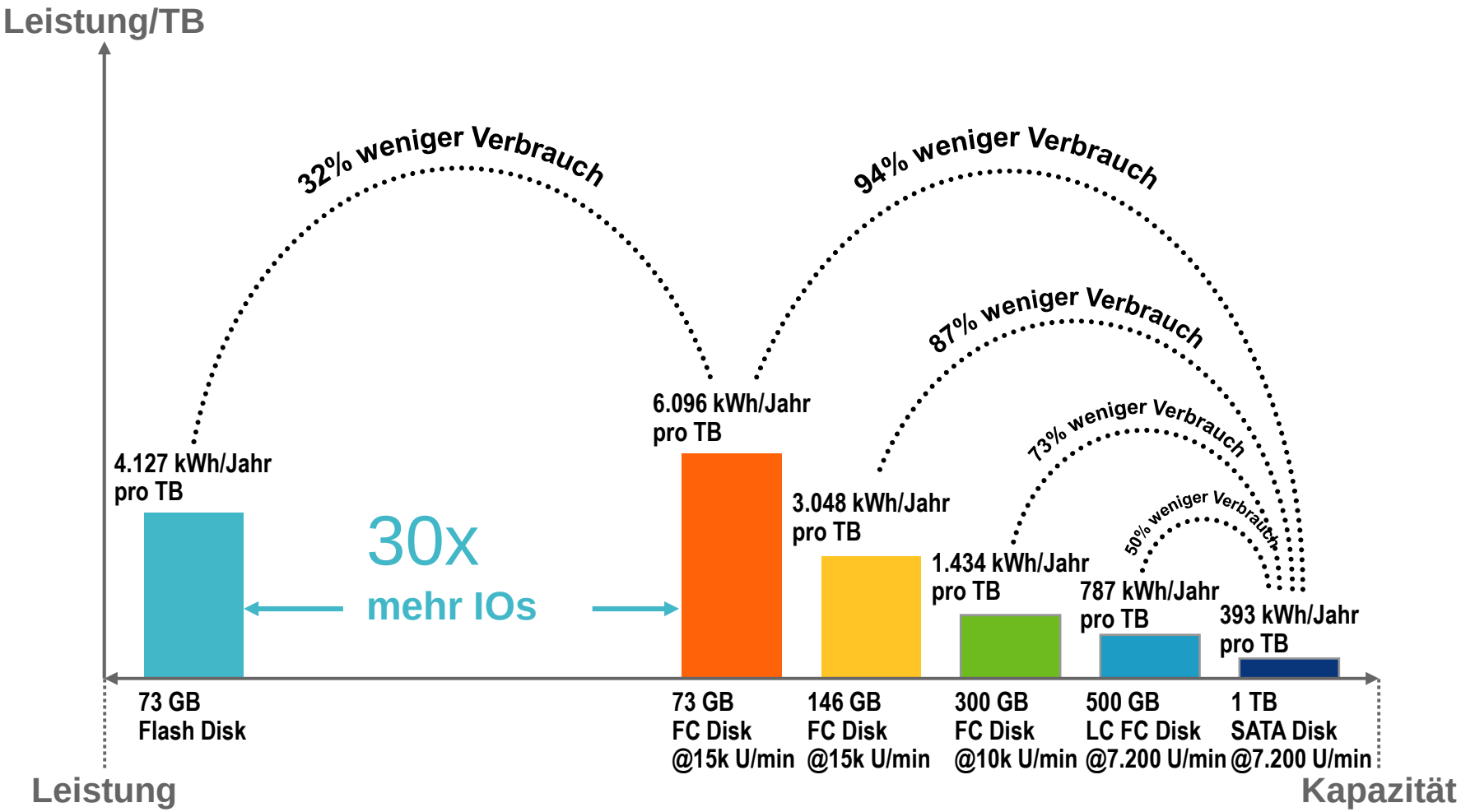


Leistung: Flash vs. HDD

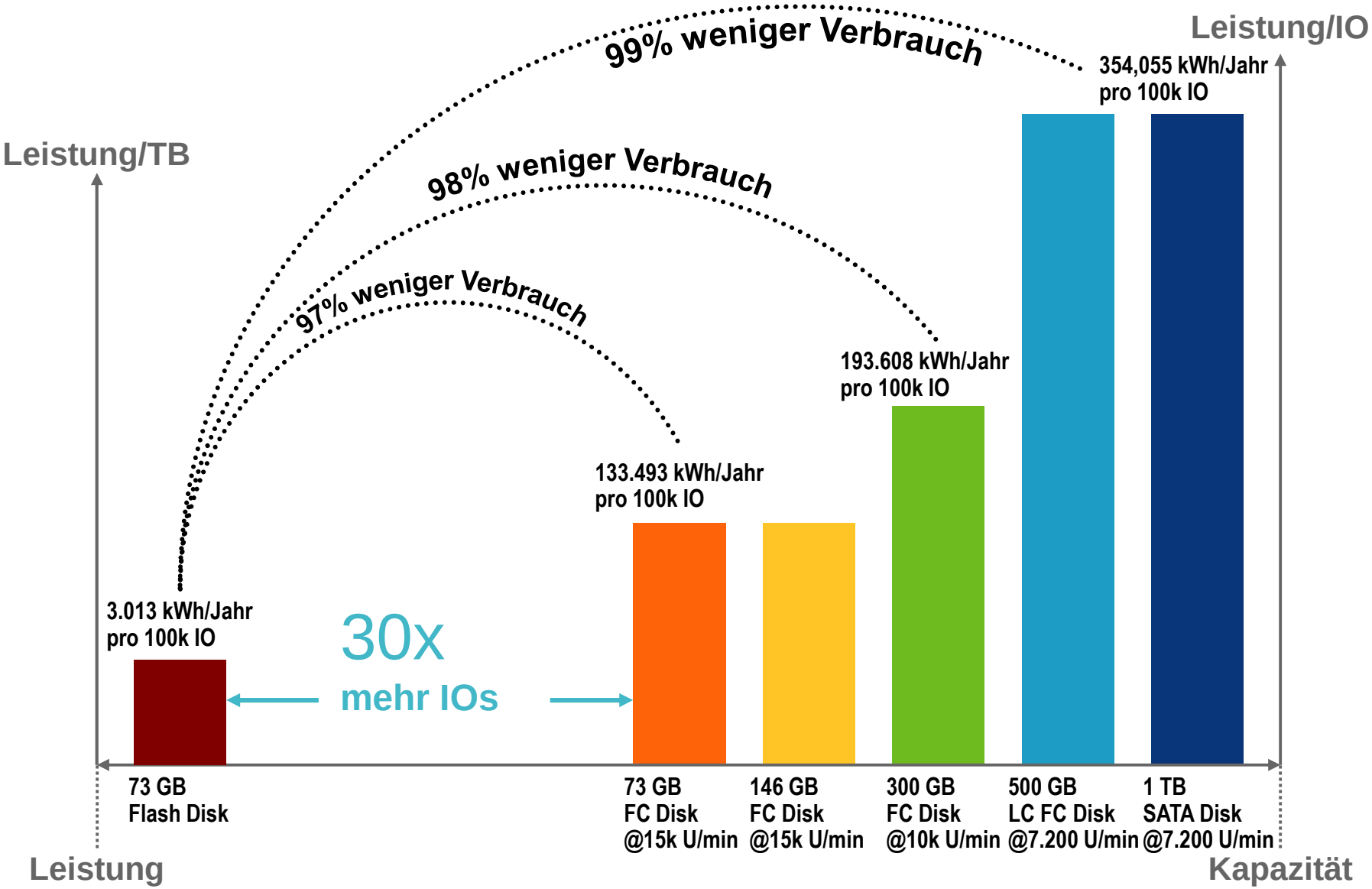
OLTP2 Arbeitslast – mit dynamischem Cache-Partitioning



Energieeffizienz: Flash vs. Festplatten



Energieeffizienz: Flash vs. Festplatten



Ideale Umgebungen für Flash-Laufwerke

EMC²
where information lives®

Anwendungen mit niedrigen Antwortzeiten

“Tier-0” Anwendungen

Alle Anwendungen mit vielen Read Misses

Beispiele:

- Algorithmic Trading
- Currency Exchange & Arbitrage
- Trade Optimization
- Real-Time Data/Feed Processing
- Contextual Web Advertising
- Andere Real-Time Transaction Systeme
- Data modeling

Wo Flash-Laufwerke nicht hingehören

EMC² where information lives®

Lasten wenigen oder ohne Read Misses

- Lasten mit hohem Read Hit Anteil
- Lasten mit hohem sequentiellen Leseanteil
  - Data Warehousing
  - Video Streaming
- Lasten mit hohem Schreibanteil

Anwendungen mit synchronem SRDF

- Read Miss Lieferzeiten werden verbessert, aber je höher die Schreibanteile werden, desto geringer wird der Gewinn durch Flas-Laufwerke

Konfigurationsdetails

EMC² where information lives®

Flash-Laufwerke können für Open Systems und Mainframe eingesetzt werden

RAID-5 (3+1 und 7+1) ist der einzig sinnvolle Schutz

HDDs und Flash-Laufwerke können in derselben Loop konfiguriert werden

Alle Mitglieder einer RAID-Gruppe müssen entweder Flash oder HDD sein

LUNs auf Flash können

- TF/Clone Quelle,
- TF/Clone Ziel oder
- TF/Snap Quelle sein

LUNs auf Flash-Laufwerken können über SRDF gekoppelt werden

Konfigurationsdetails

EMC² where information lives®

Symmetrix Optimizer Virtual LUN Migration funktioniert mit LUNs auf Flash Laufwerken

- Migration von/nach Flash zu/von HDDs

Virtual Provisioning Data LUNs können auf Flash Laufwerken abgelegt werden

Dynamic Cache Partitioning kann mit Flash Laufwerken genutzt werden

Symmetrix Priority Controls kann mit Flash Laufwerken genutzt werden

Regeln für Ersatzlaufwerke gelten genauso für Flash Laufwerke (1 + (1 für jede angefangenen 50))

Konfigurationsdetails

EMC² where information lives®

PowerVault Volumes gehören auf HDDs

Einige interne Volumen (SFS, DRV) gehören auf HDDs

DSE Pool und TF/Snap Pool Volumen gehören auf HDDs

SRDF/Star und SRDF/CR sind nur für Volumen auf HDDs erlaubt

Optimizer verlegt LUNs nicht auf Flash Laufwerke

Max. 4 Flash Laufwerke pro Slice (mehr per RPQ)

LUNs auf Flash Laufwerken können keine ORS Quellen oder Ziele sein

FlashCopy, PPRC und XRC sind nicht mit Flash Laufwerken erlaubt

TPF und i5 *noch* nicht unterstützt

Fragen?
Anregungen?
Erfahrungen?
Kommentare?

Vielen Dank
für Ihre Aufmerksamkeit!

EMC²®