

White Paper

FUJITSU Server PRIMEQUEST

Performance Report PRIMEQUEST 2800B2

In diesem Dokument sind alle Benchmarks zusammengefasst, die für den FUJITSU Server PRIMEQUEST 2800B2 durchgeführt wurden.

Ferner werden die Leistungsdaten der PRIMEQUEST 2800B2 mit denen anderer PRIMEQUEST Modelle verglichen und diskutiert. Neben den Benchmark-Ergebnissen als solchen wird jeder Benchmark und die Umgebung, in der der Benchmark durchgeführt wurde, kurz erläutert.

Version

1.1

2015-11-18



Inhalt

Dokumenthistorie.....	2
Technische Daten.....	3
SPECcpu2006.....	5
Disk-I/O: Performance von RAID-Controllern.....	9
OLTP-2.....	14
vServCon.....	18
VMmark V2.....	23
STREAM.....	27
Literatur.....	29
Kontakt.....	30

Dokumenthistorie

Version 1.0 (2015-09-15)

Neu:

- Technische Daten
- SPECcpu2006
Messungen mit Intel® Xeon® Processor E7-8800 v3 Product Family
- Disk-I/O: Performance von RAID-Controllern
Messungen mit „PRAID EP420i“ Controller
- vServCon
Ergebnisse für Intel® Xeon® Processor E7 v3 Family
- VMmark V2
„Performance Only“ Messungen mit Xeon E7-8890 v3
- STREAM
Messungen mit Intel® Xeon® Processor E7-8800 v3 Product Family

Version 1.1 (2015-11-18)

Neu:

- OLTP-2
Ergebnisse für Intel® Xeon® Processor E7 v3 Family

Technische Daten

PRIMEQUEST 2800B2



In diesem White Paper werden bei Maßeinheiten Dezimalpräfixe nach SI-Standard verwendet (z.B. 1 GB = 10^9 Byte). Abweichend hiervon sind bei Kapazitäten von Caches und Speichermodulen diese Präfixe als Binärpräfixe (z.B. 1 GB = 2^{30} Byte) zu interpretieren. Im Falle weiterer Ausnahmen wird an entsprechender Stelle gesondert darauf hingewiesen.

Modell	PRIMEQUEST 2800B2
Formfaktor	Rack Server
Anzahl Systemboards	1 – 4
Anzahl I/O-Einheiten	1 – 4
Anzahl Disk-Einheiten	0 – 2
Pro Systemboard:	
Chipsatz	Intel® C602 Chipset
Anzahl Sockel	2
Anzahl bestellbarer Prozessoren	2
Prozessortyp	Intel® Xeon® Processor E7-8800 v3 Product Family
Anzahl Speichersteckplätze	48 (24 pro Prozessor)
Maximaler Speicherausbau	3 TB
Max. Anzahl interner Festplatten	0
Pro I/O-Einheit:	
Onboard LAN-Controller	I/O Unit L (1GbE, 2xbaseTports) PQ2800E: 2 x 1 Gbit/s I/O Unit F (10GbE, 2xbaseTports) PQ2800E: 2 x 10 Gbit/s
PCI-Steckplätze	I/O Unit L (1GbE, 2xbaseTports) PQ2800E: 4 x PCI-Express 3.0 x8 I/O Unit F (10GbE, 2xbaseTports) PQ2800E: 1 x PCI-Express 3.0 x8 2 x PCI-Express 3.0 x16
Pro Disk-Einheit:	
Max. Anzahl interner Festplatten	4

Prozessoren (seit System-Release)								
Prozessor	Cores	Threads	Cache	QPI-Speed	Nominal-frequenz	Max. Turbo-frequenz	Max. Speicher-frequenz ¹⁾	TDP
			[MB]	[GT/s]	[Ghz]	[Ghz]	[MHz]	[Watt]
Xeon E7-8893 v3	4	8	45	9.60	3.20	3.50	1600	140
Xeon E7-8891 v3	10	20	45	9.60	2.80	3.50	1600	165
Xeon E7-8860 v3	16	32	40	9.60	2.20	3.20	1600	140
Xeon E7-8867 v3	16	32	45	9.60	2.50	3.30	1600	165
Xeon E7-8870 v3	18	36	45	9.60	2.10	2.90	1600	140
Xeon E7-8880 v3	18	36	45	9.60	2.30	3.10	1600	150
Xeon E7-8890 v3	18	36	45	9.60	2.50	3.30	1600	165

1) BIOS-Einstellung: *Memory Operation Mode = Performance Mode*

Alle mit der PRIMEQUEST 2800B2 bestellbaren Prozessoren unterstützen Intel® Turbo Boost Technology 2.0. Diese Technologie ermöglicht den Betrieb des Prozessors mit höheren Frequenzen als der Nominalfrequenz. In der Prozessortabelle steht „Max. Turbofrequenz“ für das theoretische Frequenzmaximum bei nur einem aktiven Core pro Prozessor. Die tatsächlich erreichbare Maximalfrequenz ist abhängig von der Anzahl aktiver Cores, dem Stromverbrauch, der elektrischen Leistungsaufnahme und der Temperatur des Prozessors.

Das Erreichen der maximalen Turbofrequenz wird von Intel grundsätzlich nicht garantiert. Dies hängt mit Fertigungstoleranzen zusammen, aus denen eine Varianz bezüglich der Performance verschiedener Exemplare eines Prozessormodells folgt. Das Spektrum der Varianz überdeckt den gesamten Bereich zwischen der Nominalfrequenz und der maximalen Turbofrequenz.

Die Turbo-Funktionalität ist per BIOS-Option einstellbar. Grundsätzlich empfiehlt Fujitsu die „Turbo Mode“-Option auf der Standardeinstellung „Enabled“ zu belassen, denn durch die höheren Frequenzen wird die Performance deutlich gesteigert. Da die höheren Frequenzen jedoch abhängig von Randbedingungen und nicht immer garantiert sind, kann es für Anwendungsszenarien mit intensiver Verwendung von AVX-Instruktionen und hoher Anzahl Instruktionen pro Takteinheit, aber auch solchen, in denen eine konstante Performance oder eine niedrige elektrische Leistungsaufnahme gefordert ist, von Vorteil sein die „Turbo Mode“-Option auszuschalten.

Speichermodule (seit System-Release)									
Speichermodul	Kapazität [GB]	Ranks	Bitbreite der Speicherchips	Frequenz [MHz]	Low voltage	Load reduced	Registered	ECC	
16GB (2x8GB) 1Rx4 DDR4-2133 R ECC	16	1	4	2133			✓	✓	
32GB (2x16GB) 2Rx4 DDR4-2133 R ECC	32	2	4	2133			✓	✓	
64GB (2x32GB) 4Rx4 DDR4-2133 LR ECC	64	4	4	2133		✓	✓	✓	
128GB (2x64GB) 4Rx4 DDR4-2133 LR ECC	128	4	4	2133		✓	✓	✓	

Netzteile (seit System-Release)	max. Anzahl
Power supply 2.880W silver	6
Power Supply 2.880W platinum hp	6

Einige Komponenten sind möglicherweise nicht in allen Ländern/Vertriebsregionen verfügbar.

Detaillierte technische Informationen finden Sie im [Datenblatt PRIMEQUEST 2800B2](#).

SPECcpu2006

Benchmark-Beschreibung

SPECcpu2006 ist ein Benchmark, der die Systemeffizienz bei Integer- und Fließkomma-Operationen misst. Er besteht aus einer Integer-Testsuite (SPECint2006), die 12 Applikationen enthält, und einer Fließkomma-Testsuite (SPECfp2006), die 17 Applikationen enthält. Beide Testsuiten sind extrem rechenintensiv und konzentrieren sich auf die CPU und den Speicher. Andere Komponenten, wie Disk-I/O und Netzwerk, werden von diesem Benchmark nicht vermessen.

SPECcpu2006 ist nicht an ein spezielles Betriebssystem gebunden. Der Benchmark ist als Source-Code verfügbar und wird vor der eigentlichen Messung kompiliert. Daher beeinflussen auch die verwendete Compiler-Version und deren Optimierungseinstellungen das Messergebnis.

SPECcpu2006 beinhaltet zwei verschiedene Methoden der Performance-Messung: Die erste Methode (SPECint2006 bzw. SPECfp2006) ermittelt die Zeit, die für die Bearbeitung einer einzelnen Aufgabe benötigt wird. Die zweite Methode (SPECint_rate2006 bzw. SPECfp_rate2006) ermittelt den Durchsatz, d.h. wie viele Aufgaben parallel erledigt werden können. Beide Methoden werden zusätzlich noch in zwei Messläufe unterteilt, „base“ und „peak“, die sich in der Verwendung der Compiler-Optimierung unterscheiden. Bei der Publikation von Ergebnissen werden immer „base“-Werte verwendet, „peak“-Werte sind optional.

Benchmark	Arithmetik	Typ	Compiler-Optimierung	Messergebnis	Anwendung
SPECint2006	Integer	peak	aggressiv	Geschwindigkeit	Singlethreaded
SPECint_base2006	Integer	base	konservativ		
SPECint_rate2006	Integer	peak	aggressiv	Durchsatz	Multithreaded
SPECint_rate_base2006	Integer	base	konservativ		
SPECfp2006	Fließkomma	peak	aggressiv	Geschwindigkeit	Singlethreaded
SPECfp_base2006	Fließkomma	base	konservativ		
SPECfp_rate2006	Fließkomma	peak	aggressiv	Durchsatz	Multithreaded
SPECfp_rate_base2006	Fließkomma	base	konservativ		

Bei den Messergebnissen handelt es sich um das geometrische Mittel aus normalisierten Verhältniswerten, die für die Einzel-Benchmarks ermittelt wurden. Das geometrische Mittel führt gegenüber dem arithmetischen Mittel dazu, dass bei unterschiedlich hohen Einzelergebnissen eine Gewichtung zugunsten der niedrigeren Einzelergebnisse erfolgt. Normalisiert heißt, dass gemessen wird, wie schnell das Testsystem verglichen mit einem Referenzsystem ist. Der Wert „1“ wurde für die SPECint_base2006-, SPECint_rate_base2006, SPECfp_base2006 und SPECfp_rate_base2006-Ergebnisse des Referenzsystems festgelegt. So bedeutet beispielsweise ein SPECint_base2006-Wert von 2, dass das Messsystem diesen Benchmark etwa doppelt so schnell wie das Referenzsystem bewältigt hat. Ein SPECfp_rate_base2006-Wert von 4 bedeutet, dass das Messsystem diesen Benchmark etwa 4/[# base copies] mal so schnell wie das Referenzsystem bewältigt hat. „# base copies“ gibt hierbei an, wie viele parallele Instanzen des Benchmarks ausgeführt worden sind.

Nicht alle SPECcpu2006-Messungen werden von uns zur Veröffentlichung bei SPEC eingereicht. Daher erscheinen auch nicht alle Ergebnisse auf den Web-Seiten von SPEC. Da wir für alle Messungen die Protokolldateien archivieren, können wir jederzeit den Nachweis für die korrekte Durchführung der Messungen erbringen.

Benchmark-Umgebung

System Under Test (SUT)	
Hardware	
Modell	PRIMEQUEST 2800B2
Prozessor	Intel® Xeon® Processor E7-8800 v3 Product Family
Speicher	4 Sockel: 32 x 32GB (2x16GB) 2Rx4 DDR4-2133 R ECC 8 Sockel: 64 x 32GB (2x16GB) 2Rx4 DDR4-2133 R ECC
Software	
BIOS-Einstellungen	Energy Performance = Performance
Betriebssystem	SPECint_rate_base2006, SPECint_rate2006: Red Hat Enterprise Linux Server release 6.6 SPECfp_rate_base2006, SPECfp_rate2006: Red Hat Enterprise Linux Server release 7.1
Betriebssystem-einstellungen	SPECint_rate_base2006, SPECint_rate2006: echo always > /sys/kernel/mm/redhat_transparent_hugepage/enabled
Compiler	SPECint_rate_base2006, SPECint_rate2006: Version 14.0.0.080 of Intel C++ Studio XE for Linux SPECfp_rate_base2006, SPECfp_rate2006: C/C++: Version 15.0.0.090 of Intel C++ Studio XE for Linux Fortran: Version 15.0.0.090 of Intel Fortran Studio XE for Linux

Einige Komponenten sind möglicherweise nicht in allen Ländern / Vertriebsregionen verfügbar.

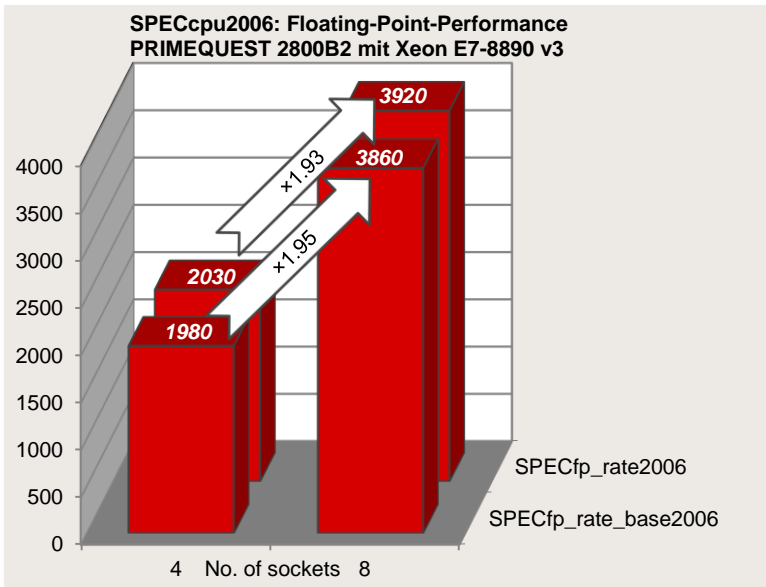
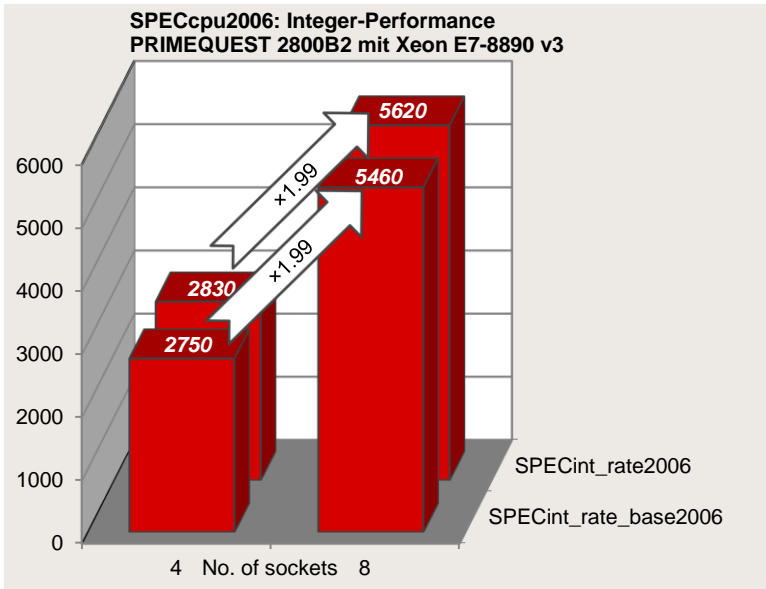
Benchmark-Ergebnisse

Das Benchmark-Ergebnis hängt prozessorseitig in erster Linie von der Größe des Prozessor-Caches, der Unterstützung von Hyper-Threading, der Anzahl Prozessorkerne und der Prozessorfrequenz ab. Die Anzahl Cores, die durch den Benchmark belastet werden, bestimmt die maximal erreichbare Prozessorfrequenz.

Prozessor	Anzahl Prozessoren	SPECint_rate_base2006	SPECint_rate2006	Anzahl Prozessoren	SPECint_rate_base2006	SPECint_rate2006
Xeon E7-8893 v3	4	850	881	8	1700	1760
Xeon E7-8891 v3	4			8	3780	3910
Xeon E7-8860 v3	4			8	4740	4870
Xeon E7-8867 v3	4			8	4920	5070
Xeon E7-8870 v3	4			8	5070	5220
Xeon E7-8880 v3	4			8	5250	5410
Xeon E7-8890 v3	4	2750	2830	8	5460	5620

Prozessor	Anzahl Prozessoren	SPECfp_rate_base2006	SPECfp_rate2006	Anzahl Prozessoren	SPECfp_rate_base2006	SPECfp_rate2006
Xeon E7-8893 v3	4	793	824	8	1580	1610
Xeon E7-8891 v3	4			8	3040	3090
Xeon E7-8860 v3	4			8	3520	3620
Xeon E7-8867 v3	4			8	3690	3750
Xeon E7-8870 v3	4			8	3670	3740
Xeon E7-8880 v3	4			8	3700	3770
Xeon E7-8890 v3	4	1980	2030	8	3860	3920

Die beiden folgenden Grafiken geben wieder, wie die Performance der PRIMEQUEST 2800B2 bei Verwendung des Xeon E7-8890 v3 von vier auf acht Prozessoren skaliert.



Disk-I/O: Performance von RAID-Controllern

Benchmark-Beschreibung

Performance-Messungen von Disk-Subsystemen bei PRIMEQUEST Servern dienen dazu, deren Leistungsfähigkeit zu beurteilen und einen Vergleich der verschiedenen Storage-Anbindungen bei PRIMEQUEST Servern zu ermöglichen. Standardmäßig werden diese Performance-Messungen mit einem definierten Messverfahren durchgeführt, das die Zugriffe realer Anwendungsszenarien anhand von Kenndaten modelliert.

Die wesentlichen Kenndaten sind:

- Anteil von wahlfreien Zugriffen / sequentiellen Zugriffen
- Anteil der Zugriffsarten Lesen / Schreiben
- Blockgröße (kB)
- Anzahl paralleler Zugriffe (# of Outstanding I/Os)

Eine gegebene Wertekombination dieser Kenndaten heißt „Lastprofil“. Die folgenden fünf Standardlastprofile lassen sich typischen Anwendungsszenarien zuordnen:

Standard-lastprofil	Zugriff	Zugriffsart		Blockgröße [kB]	Anwendung
		read	write		
File copy	wahlfrei	50%	50%	64	Kopieren von Dateien
File server	wahlfrei	67%	33%	64	File-Server
Database	wahlfrei	67%	33%	8	Datenbank (Datentransfer) Mail Server
Streaming	sequentiell	100%	0%	64	Datenbank (Log-File), Datensicherung; Video Streaming (teilweise)
Restore	sequentiell	0%	100%	64	Wiederherstellen von Dateien

Zur Modellierung parallel zugreifender Anwendungen mit unterschiedlicher Belastungsintensität wird die „# of Outstanding I/Os“ mit 1, 3, 8 beginnend bis 512 gesteigert (ab 8 in Zweierpotenzschritten).

Die Messungen des vorliegenden Dokumentes beruhen auf diesen Standardlastprofilen.

Die wichtigsten Ergebnisse einer Messung sind:

- Throughput [MB/s] Datendurchsatz in Megabytes pro Sekunde
- Transactions [IO/s] Transaktionsrate in I/O-Operationen pro Sekunde
- Latency [ms] mittlere Antwortzeit in ms

Für sequentielle Lastprofile hat sich der Datendurchsatz als übliche Messgröße durchgesetzt, während bei den wahlfreien Lastprofilen mit ihren kleinen Blockgrößen meist die Messgröße „Transaktionsrate“ verwendet wird. Datendurchsatz und Transaktionsrate sind direkt proportional zueinander und lassen sich nach der Formel

<i>Datendurchsatz [MB/s]</i>	$= \text{Transaktionsrate [IO/s]} \times \text{Blockgröße [MB]}$
<i>Transaktionsrate [IO/s]</i>	$= \text{Datendurchsatz [MB/s]} / \text{Blockgröße [MB]}$

ineinander überführen.

In diesem Kapitel sind Kapazitäten von Speichermedien durchgängig zur Basis 10 angegeben (1 TB = 10^{12} Bytes), während alle anderen Kapazitäten, Dateigrößen, Blockgrößen und Durchsätze zur Basis 2 angegeben sind (1 MB/s = 2^{20} Bytes/s).

Alle Details des Messverfahrens und Grundlagen zur Disk-I/O-Performance sind im White Paper „[Grundlagen Disk-I/O-Performance](#)“ beschrieben.

Benchmark-Umgebung

Alle in diesem Kapitel diskutierten Messergebnisse wurden mit den im Folgenden aufgelisteten Hardware- und Software-Komponenten ermittelt:

System Under Test (SUT)	
Hardware	
Controller	1 x „PRAID EP420i“
Festplatte	4 x 2.5" SAS SSD Toshiba PX02SMF040 4 x 2.5" SAS HDD HGST HUC156045CSS204
Software	
BIOS-Einstellungen	Intel Virtualization Technology = Disabled VT-d = Disabled Energy Performance = Performance Utilization Profile = Unbalanced CPU C6 Report = Disabled
Betriebssystem	Microsoft Windows Server 2012 Standard
Betriebssystem-Einstellungen	Choose or customize a power plan: High performance Für die Disk-I/O-erzeugenden Prozesse: Setzen der AFFINITY auf den CPU-Node, an den der PCIe-Slot des RAID-Controllers angeschlossen ist
Verwaltungssoftware	ServerView RAID Manager 6.1.4
Initialisierung von RAID-Verbänden	RAID-Verbände werden vor der Messung mit einer elementaren Blockgröße von 64 kB („Stripe Size“) initialisiert
Dateisystem	NTFS
Messwerkzeug	Iometer 2006.07.27
Messdaten	Messdateien von 32 GB bei 1 – 8 Festplatten; 64 GB bei 9 – 16 Festplatten; 128 GB bei 17 oder mehr Festplatten

Einige Komponenten sind möglicherweise nicht in allen Ländern / Vertriebsregionen verfügbar.

Benchmark-Ergebnisse

Die hier vorgestellten Ergebnisse sollen dabei helfen, aus den verschiedenen Konfigurationsmöglichkeiten der PRIMEQUEST 2800B2 die passende Lösung unter dem Gesichtspunkt der Disk-I/O-Performance auszuwählen. Im Folgenden sollen verschiedene Kombinationen von RAID-Controllern und Datenträgern untersucht werden.

Festplatten

Die erste wesentliche Komponente sind die Festplatten. Wenn im Folgenden von „Festplatten“ die Rede ist, so ist dies als Oberbegriff gemeint für HDDs („hard disk drives“, also konventionelle Festplatten) und SSDs („solid state drives“, also nichtflüchtige elektronische Speichermedien).

Modellvarianten

Die maximale Anzahl von Festplatten im System hängt von der Systemkonfiguration ab. Die PRIMEQUEST 2800B2 erlaubt bis zu zwei Disk Units (DU). Die Disk Units werden im Folgenden auch mit dem Oberbegriff „Untereinheit“ bezeichnet.

Die folgende Tabelle stellt die wesentlichen Fälle zusammen. Die zwei Konfigurationsvarianten der Disk Unit werden folgendermaßen abgekürzt: "Disk Unit (1C)" ist eine Disk Unit mit einem Controller, und "Disk Unit (2C)" ist eine Disk Unit mit zwei Controllern.

Für alle in diesem Kapitel behandelten Schnittstellen wird nur deren höchste unterstützte Version genannt.

Untereinheit	Formfaktor	Schnittstelle	Anzahl PCIe-Controller	Maximalzahl Festplatten
Disk Unit (1C)	2.5"	SAS 12G	1	4
Disk Unit (2C)	2.5"	SAS 12G	2	2 x 2

Durch die modulare Architektur des Systems reicht es aus, die Disk-I/O-Performance pro Controller zu betrachten. Die mögliche Gesamt-Performance des Systems ergibt sich als Summe der Performance-Maxima aller darin enthaltenen Controller.

RAID-Controller

Neben den Festplatten ist der RAID-Controller die zweite Performance-bestimmende Schlüsselkomponente.

Die folgende Tabelle fasst die wichtigsten Eigenschaften der verfügbaren RAID-Controller der PRIMEQUEST 2800B2 zusammen. Pro Controller ist hierin ein kurzer Alias angegeben, der bei der anschließenden Zusammenstellung der Performance-Werte verwendet wird.

Controller-Name / Einbauplatz	Alias	Cache	Unterstützte Interfaces		Max. # Disks in der Untereinheit	RAID Levels in der Untereinheit	BBU/FBU
PRAID EP420i in Disk Unit (1C)	PRAID EP420i (DU-1C)	2 GB	SAS 12G	PCIe 3.0 x8	4 x 2.5"	0, 1, 1E, 5, 6, 10	-/√
PRAID EP420i in Disk Unit (2C)	PRAID EP420i (DU-2C)	2 GB	SAS 12G	PCIe 3.0 x8	2 x 2.5"	0, 1	-/√

Systemspezifische Schnittstellen

Die Schnittstellen eines Controllers zum System Board (gilt auch für die Schnittstelle Disk Unit – System Board) und zu den Festplatten haben jeweils spezifische Grenzen für den Datendurchsatz. Diese Grenzen sind in der folgenden Tabelle zusammengestellt. Das Minimum dieser beiden Werte ist eine prinzipielle Grenze, die nicht überschritten werden kann. Dieser Wert ist in der folgenden Tabelle mit Fettdruck hervorgehoben.

Controller-Alias	Effektiv in der Konfiguration					Anschluss über Expander
	# Disk-Kanäle	Grenze für Durchsatz Disk-Interface	PCIe-Version	PCIe-Breite	Grenze für Durchsatz PCIe-Interface	
PRAID EP420i (DU-1C)	4 x SAS 12G	4120 MB/s	2.0	x4	1716 MB/s	-
PRAID EP420i (DU-2C)	2 x SAS 12G	2060 MB/s	2.0	x4	1716 MB/s	-

Weitere Details zu diesen RAID-Controllern finden sich im White Paper „[RAID-Controller-Performance](#)“.

Einstellungen

Der Cache von HDDs hat in den meisten Fällen einen großen Einfluss auf die Disk-I/O-Performance. Er wird häufig als Sicherheitsproblem bei Stromausfall angesehen und daher abgeschaltet. Dennoch wurde er von den Festplattenherstellern aus gutem Grund zur Steigerung der Schreib-Performance integriert. Aus Performance-Gründen ist es daher empfehlenswert den Festplatten-Cache einzuschalten. Um Datenverlusten bei Stromausfall vorzubeugen, empfiehlt es sich das System mit einer USV auszustatten.

Bei Controllern mit Cache gibt es mehrere einstellbare Parameter. Die jeweils optimalen Einstellungen können vom RAID-Level, vom Anwendungsszenario und vom Datenträgertyp abhängen. Besonders bei den RAID-Levels 5 und 6 (und den davon abgeleiteten komplexeren RAID-Levels 50 und 60) ist bei Anwendungsszenarien mit Schreibanteil das Einschalten des Controller-Caches Pflicht. Bei aktiviertem Controller-Cache sollten die darin temporär gespeicherten Daten gegen Verlust bei Stromausfall gesichert werden. Hierfür ist geeignetes Zubehör verfügbar (beispielsweise eine BBU bzw. FBU).

Zwecks einfacher und sicherer Handhabung der Einstellungen von RAID-Controller und Festplatten empfiehlt sich die mit dem Server gelieferte RAID-Manager-Software „ServerView RAID“. Üblicherweise wird man – spezifisch für den Anwendungsfall – mittels der vordefinierten Modi „Performance“ oder „Data Protection“ die kompletten Cache-Einstellungen für Controller und Festplatten en bloc vornehmen. Der Modus „Performance“ gewährleistet für die Mehrzahl der Anwendungsszenarien Performance-optimale Einstellungen.

Nähere Informationen zu den Einstellungsmöglichkeiten beim Controller-Cache sind im White Paper „[RAID-Controller-Performance](#)“ zu finden.

Performance-Werte

Generell hängt die Disk-I/O-Performance eines RAID-Verbandes von Festplattentyp und –anzahl, vom RAID-Level und vom RAID-Controller ab, sofern die Limitierungen der [systemspezifischen Schnittstellen](#) nicht überschritten werden. Daher gelten auch alle Performance-Aussagen des Dokumentes „[RAID-Controller-Performance](#)“ für die PRIMEQUEST 2800B2, soweit die dort vermessenen Konfigurationen auch von diesem System unterstützt werden.

Die Performance-Werte der PRIMEQUEST 2800B2 werden im Folgenden tabellarisch zusammengestellt, jeweils spezifisch für verschiedene RAID-Level, Zugriffsarten und Blockgrößen. Wesentlich verschiedene Konfigurationsvarianten werden getrennt behandelt. Hierbei werden die etablierten Messgrößen, wie sie schon im Unterkapitel [Benchmark-Beschreibung](#) erwähnt wurden, verwendet. Bei den wahlfreien Zugriffen wird also die Transaktionsrate angegeben, und bei den sequentiellen Zugriffen der Datendurchsatz. Um Verwechslungen der Maßeinheiten zu vermeiden, sind die Tabellen für die beiden Arten von Zugriffen getrennt.

In den Tabellenzellen sind die maximal erreichbaren Werte eingetragen. Das bedeutet dreierlei: Zum einen wurden Festplatten mit optimaler Performance verwendet (die Komponenten sind im Unterkapitel [Benchmark-Umgebung](#) näher beschrieben). Des Weiteren sind Cache-Einstellungen von Controllern und Festplatten zugrunde gelegt, die für das jeweilige Zugriffsszenario und den RAID-Level optimal sind. Und schließlich ist jeder Wert das Maximum über den gesamten Bereich von Belastungsintensitäten (# of Outstanding I/Os).

Zwecks zusätzlicher Visualisierung der Zahlenwerte ist jede Tabellenzelle mit einem waagerechten Balken hinterlegt, dessen Länge proportional zum Zahlenwert in der Tabellenzelle ist. Alle Balken, die im gleichen Längenmaßstab dargestellt sind, haben die gleiche Farbe. Es können also nur die Tabellenzellen mit gleichfarbigen Balken sinnvoll visuell miteinander verglichen werden.

Da die waagerechten Balken in den Tabellenzellen die maximal erreichbaren Performance-Werte veranschaulichen, sind sie als von links nach rechts heller werdende Farbverläufe dargestellt. Der helle Farbton am rechten Balkenende drückt aus, dass der Wert das Maximum ist und nur bei optimalen Voraussetzungen erreicht werden kann. Je dunkler dann der Farbton nach links hin wird, umso häufiger wird der entsprechende Wert in der Praxis erreichbar sein.

2.5" - Wahlfreie Zugriffe (maximale Performance-Werte in IO/s):

Base Unit PQ2800B2							
Konfigurationsvariante			RAID-Level	HDDs wahlfrei 8 kB Blöcke 67% read [IO/s]	HDDs wahlfrei 64 kB Blöcke 67% read [IO/s]	SSDs wahlfrei 8 kB Blöcke 67% read [IO/s]	SSDs wahlfrei 64 kB Blöcke 67% read [IO/s]
RAID-Controller	Festplatten-typ	# Disks					
PRAID EP420i (DU-1C)	HUC156045CSS204 SAS HDD PX02SMF040 SAS SSD	2	1	1949	1085	77312	12141
		4	10	3479	1445	106524	20473
		4	0	3939	1871	128697	32507
		4	5	2151	936	36808	13110
PRAID EP420i (DU-2C)	HUC156045CSS204 SAS HDD PX02SMF040 SAS SSD	2	1	1949	1085	77312	12141
		2	0	1950	970	105105	15537

2.5" - Sequentielle Zugriffe (maximale Performance-Werte in MB/s):

Base Unit PQ2800B2							
Konfigurationsvariante			RAID-Level	HDDs sequentiell 64 kB Blöcke 100% read [MB/s]	HDDs sequentiell 64 kB Blöcke 100% write [MB/s]	SSDs sequentiell 64 kB Blöcke 100% read [MB/s]	SSDs sequentiell 64 kB Blöcke 100% write [MB/s]
RAID-Controller	Festplatten-typ	# Disks					
PRAID EP420i (DU-1C)	HUC156045CSS204 SAS HDD PX02SMF040 SAS SSD	2	1	424	272	1489	418
		4	10	586	467	1571	736
		4	0	968	901	1570	1469
		4	5	726	669	1556	1248
PRAID EP420i (DU-2C)	HUC156045CSS204 SAS HDD PX02SMF040 SAS SSD	2	1	424	272	1489	418
		2	0	504	492	1489	816

Fazit

Die PRIMEQUEST 2800B2 erreicht also mit einem Controller im Maximalausbau mit leistungsfähigen Festplatten einen Durchsatz von bis zu 1571 MB/s bei sequentiellen Lastprofilen und eine Transaktionsrate von bis zu 128697 IO/s bei typischen wahlfreien Anwendungsszenarien.

Beim maximalen Ausbau des Systems mit zwei Disk Units mit je zwei Controllern, d.h. in Summe vier Controllern, könnte man maximal acht Festplatten betreiben. In diesem Maximalausbau würde das System bei Verwendung von leistungsfähigen Festplatten rechnerisch einen Gesamtdurchsatz von bis zu 5956 MB/s bei sequentiellen Lastprofilen und eine Gesamttransaktionsrate von bis zu 420420 IO/s bei typischen wahlfreien Anwendungsszenarien erreichen.

OLTP-2

Benchmark-Beschreibung

OLTP steht für Online Transaction Processing. Dem OLTP-2-Benchmark liegt das typische Anwendungsszenario einer Datenbanklösung zugrunde. Es werden bei OLTP-2 Zugriffe auf eine Datenbank simuliert und die Anzahl erreichter Transaktionen pro Sekunde (tps) als Maß für die Leistungsfähigkeit des vermessenen Systems ermittelt.

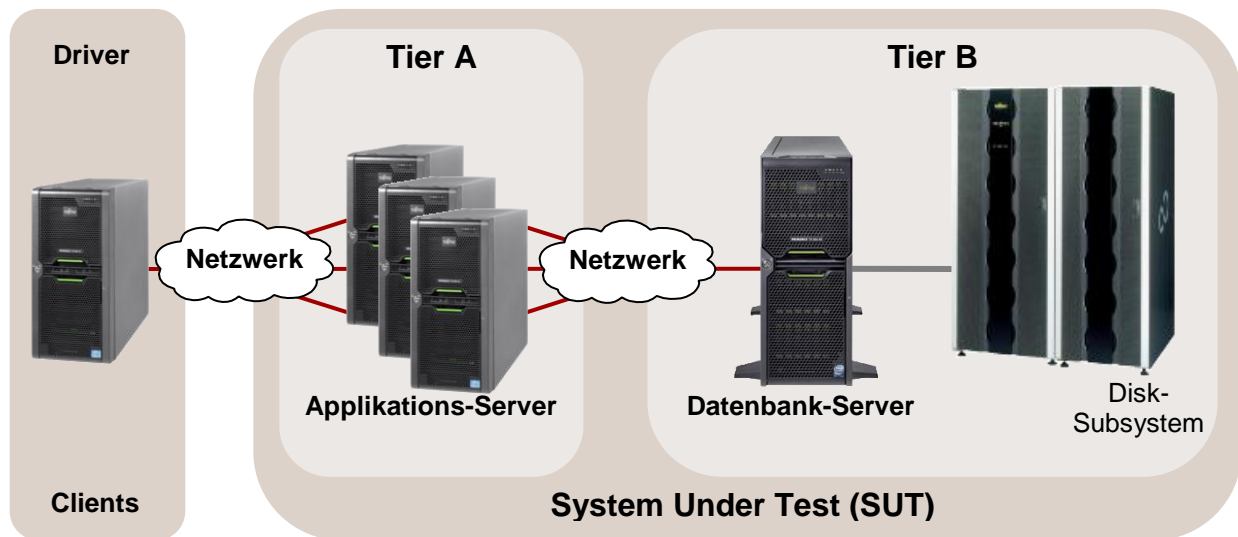
Im Gegensatz zu Benchmarks, wie beispielsweise SPECint und TPC-E, die von unabhängigen Gremien standardisiert wurden und bei denen die Einhaltung des jeweiligen Reglements überwacht wird, ist OLTP-2 ein interner Benchmark von Fujitsu. OLTP-2 basiert auf dem bekannten Datenbank-Benchmark TPC-E. OLTP-2 wurde so gestaltet, dass eine Vielzahl von Konfigurationen messbar sind, um die Skalierung eines Systems hinsichtlich CPU- und Speicherausbau darstellen zu können.

Auch wenn die beiden Benchmarks OLTP-2 und TPC-E ähnliche Anwendungsszenarien simulieren und die gleichen Lastprofile verwenden, so sind die Ergebnisse nicht vergleichbar oder gar gleichzusetzen, da die beiden Benchmarks unterschiedliche Methoden zur Simulation der Benutzerlast verwenden. Typischerweise sind OLTP-2-Werte TPC-E-Werten ähnlich. Ein direkter Vergleich oder gar die Bezeichnung des OLTP-2-Ergebnisses als TPC-E-Ergebnis ist nicht zulässig, da insbesondere kein Preis-Leistungswert ermittelt wird.

Weitere Informationen können dem Dokument [Benchmark-Überblick OLTP-2](#) entnommen werden.

Benchmark-Umgebung

Der Messaufbau wird symbolisch durch folgende Grafik veranschaulicht:



Alle Ergebnisse wurden exemplarisch auf einer PRIMEQUEST 2800E2 ermittelt.

Datenbank-Server (Tier B)	
Hardware	
Modell	PRIMEQUEST 2800E2
Prozessor	Intel® Xeon® Processor E7 v3 Family
Speicher	4096 GB: 64 x 64GB (2x32GB) 4Rx4 DDR4-2133 LR ECC 2048 GB: 32 x 64GB (2x32GB) 4Rx4 DDR4-2133 LR ECC 1024 GB: 16 x 64GB (2x32GB) 4Rx4 DDR4-2133 LR ECC
Netzwerk-Interface	2 x onboard LAN 10 Gb/s

Disk-Subsystem	PRIMEQUEST 2800E2: 1 x PRAID EP420i 2 x 300 GB 10k rpm SAS Drives, RAID1 (OS) 15 x PRAID EP420e 14 x JX40: Je 15 x 400 GB SSD Drive, RAID5 (Daten) 1 x JX40: 10 x 900 GB 10k rpm SAS Drives, RAID10 (LOG)
Software	
BIOS	Version BB15068
Betriebssystem	Microsoft Windows Server 2012 R2 Standard
Datenbank	Microsoft SQL Server 2014 Enterprise

Applikations-Server (Tier A)

Hardware	
Modell	2 x PRIMERGY RX2530 M1
Prozessor	2 x Xeon E5-2697 v3
Speicher	64 GB, 2133 MHz registered ECC DDR4
Netzwerk-Interface	2 x onboard LAN 10 Gb/s
Disk-Subsystem	2 x 300 GB 15k rpm SAS Drive
Software	
Betriebssystem	Microsoft Windows Server 2012 R2 Standard

Client (Lastgenerator)

Hardware	
Modell	1 x PRIMERGY RX300 S8
Prozessor	2 x Xeon E5-2667 v2
Speicher	64 GB, 1600 MHz registered ECC DDR3
Netzwerk-Interface	2 x onboard LAN 1 Gb/s 1 x Dual Port LAN 1 Gb/s
Disk-Subsystem	1 x 300 GB 10k rpm SAS Drive
Software	
Betriebssystem	Microsoft Windows Server 2012 R2 Standard
Benchmark	OLTP-2 Software EGen version 1.14.0

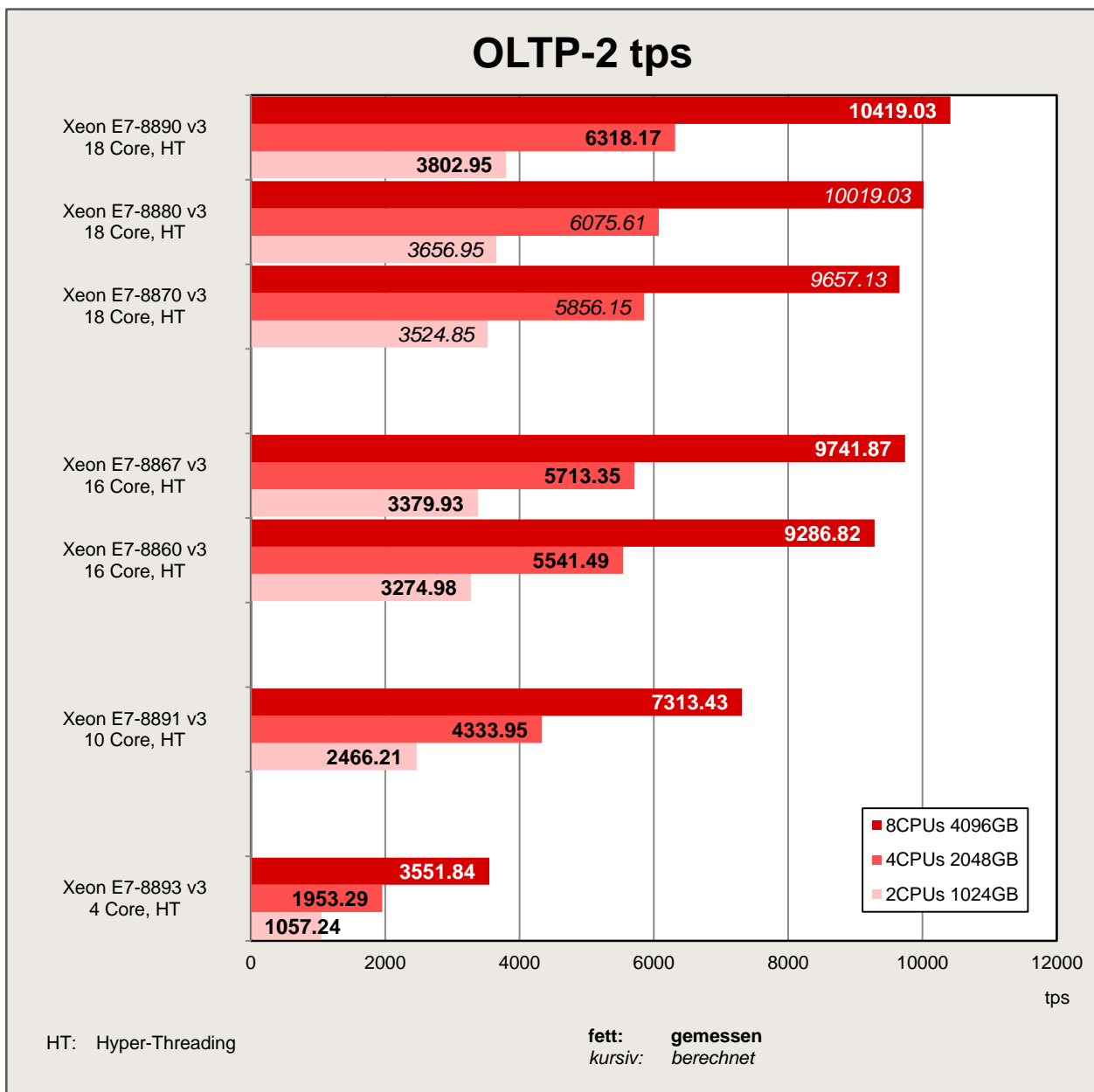
Einige Komponenten sind möglicherweise nicht in allen Ländern / Vertriebsregionen verfügbar.

Benchmark-Ergebnisse

Die Datenbank-Performance ist in hohem Maße abhängig von den Ausbaumöglichkeiten mit CPU, Speicher und den Anschlussmöglichkeiten eines für die Datenbank angemessenen Disk-Subsystems. Bei den folgenden Skalierungsbetrachtungen der Prozessoren gehen wir davon aus, dass sowohl der Speicher als auch das Disk-Subsystem adäquat gewählt ist und keinen Engpass darstellt.

Als Richtlinie für die Auswahl von Arbeitsspeicher gilt im Datenbankumfeld, dass eine ausreichende Menge wichtiger ist als die Geschwindigkeit der Speicherzugriffe. Daher wurde bei den Messungen mit zwei Prozessoren eine Bestückung mit insgesamt 1024 GB Speicher, bei den Messungen mit vier Prozessoren eine Bestückung mit 2048 GB Speicher und bei den Messungen mit acht Prozessoren eine Bestückung mit insgesamt 4096 GB Speicher betrachtet. Die Speicherbestückungen hatten einen Memory-Zugriff von 1600 MHz. Weitere Informationen über Speicherperformance sind in dem White Paper [Speicher-Performance Xeon E7 v3 \(Haswell-EX\) basierter Systeme](#) zu finden.

Die nachfolgende Grafik zeigt die OLTP-2 Transaktionsraten, die mit zwei, vier und acht Prozessoren der Intel® Xeon® Processor E7 v3 Family erreicht werden können.



Es wird deutlich, dass durch die Vielzahl an freigegebenen Prozessoren eine große Leistungsbandbreite abgedeckt wird. Vergleicht man den OLTP-2-Wert des leistungsschwächsten Prozessors Xeon E7-8893 v3 mit dem des leistungstärksten Prozessors Xeon E7-8890 v3, so ergibt sich eine Leistungssteigerung um den Faktor 2.9.

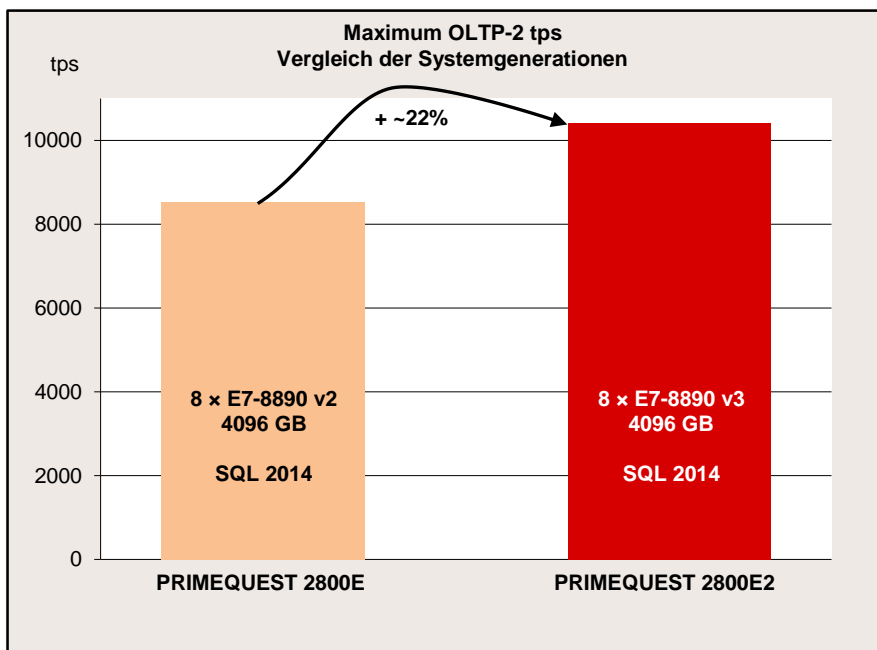
Anhand der Anzahl Cores lassen sich die Prozessoren in verschiedene Leistungsgruppen unterteilen:

Den Einstieg stellt Xeon E7-8893 v3 als Prozessor mit vier Kernen und Hyper-Threading dar.

Die Gruppen der 10-, und 16-Core-Prozessoren bieten in dieser Prozessorserie eine OLTP-2-Performance im mittleren Bereich. Durch die unterschiedlichen technischen Eigenschaften der Prozessoren in diesen Gruppen (vgl. „Technische Daten“) kann je nach Einsatzszenario die passende CPU ausgewählt werden.

Am oberen Ende der Leistungsskala liegt die Gruppe der Prozessoren mit 18 Kernen. Durch die gestaffelten CPU-Taktfrequenzen wird eine OLTP-2-Leistung von 9657.13 tps (8 x Xeon E7-8870 v3) bis 10419.03 tps (8 x Xeon E7-8890 v3) erreicht.

Vergleicht man die maximal erreichbaren OLTP-2 Werte der aktuellen Systemgeneration mit den Werten, die auf den Vorgängersystemen erreicht wurden, so ergibt sich eine Steigerung von ca. 22%.



vServCon

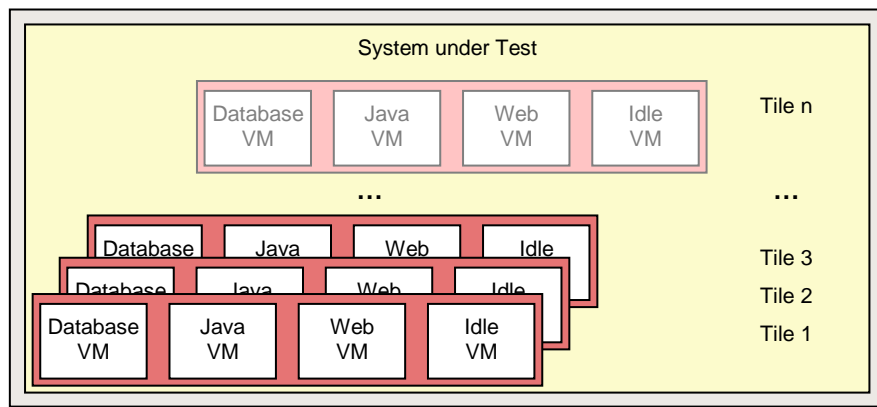
Benchmark-Beschreibung

vServCon ist ein bei Fujitsu verwendeter Benchmark zum Vergleich von Serverkonfigurationen mit Hypervisor in Bezug auf ihre Eignung für Server-Konsolidierung. Hiermit ist sowohl der Vergleich von Systemen, Prozessoren und I/O-Technologien möglich, wie auch der Vergleich von Hypervisor-en, Virtualisierungsformen und zusätzlichen Treibern für virtuelle Maschinen.

Bei vServCon handelt es sich nicht um einen neuen Benchmark im eigentlichen Sinn. Es ist vielmehr ein Framework, das bereits etablierte Benchmarks, ggf. auch in modifizierter Form, als Workloads zusammenfasst, um die Last einer konsolidierten und virtualisierten Serverumgebung nachzubilden. Es kommen drei bewährte Benchmarks zum Einsatz, die die Anwendungsszenarien Datenbank, Applikationsserver und Web-Server abdecken.

Anwendungsszenario	Benchmark	Anzahl logischer CPU-Cores	Memory
Database	Sysbench (angepasst)	2	1.5 GB
Java-Applikationsserver	SPECjbb (angepasst, mit 50% - 60% Last)	2	2 GB
Webserver	WebBench	1	1.5 GB

Jedes der drei Anwendungsszenarien wird jeweils einer dedizierten virtuellen Maschine (VM) zugeordnet. Hinzu kommt eine vierte, so genannte Idle-VM. Diese vier VMs bilden eine „Tile“ (engl. Kachel). Durch die Leistungsfähigkeit der zugrunde liegenden Server-Hardware ist es meist notwendig, dass im Rahmen einer Messung mehrere identische Tiles parallel gestartet werden müssen um eine maximale Gesamt-Performance zu erreichen.



Jedes der drei vServCon-Anwendungsszenarien ergibt für die jeweilige VM ein spezifisches Ergebnis in Form von applikationsspezifischen Transaktionsraten. Um hieraus eine normalisierte Bewertungszahl zu bilden, werden die einzelnen Ergebnisse für eine Tile in Relation zu den jeweiligen Ergebnissen eines Referenzsystems gesetzt. Die daraus resultierenden relativen Performance-Werte werden geeignet gewichtet und über alle VMs und Tiles aufsummiert. Das Ergebnis ist eine Bewertungszahl, „Score“ genannt, für diese Tile-Anzahl.

Diese Prozedur wird – in der Regel beginnend mit eins – für steigende Tile-Anzahlen durchgeführt, bis keine signifikante Steigerung dieses vServCon-Scores mehr eintritt. Der finale vServCon-Score ist dann das Maximum über die vServCon-Scores aller Tile-Anzahlen. Diese Bewertungszahl spiegelt somit den maximalen Gesamtdurchsatz wider, den man durch den Betrieb des in vServCon definierten Mixes aus vielen Anwendungs-VMs bis zur möglichst vollständigen Ausnutzung der CPU-Ressourcen erzielen kann. Dabei ist die Messumgebung für vServCon so ausgelegt, dass nur die CPU der begrenzende Faktor ist und keine Limitierungen durch andere Ressourcen eintreten.

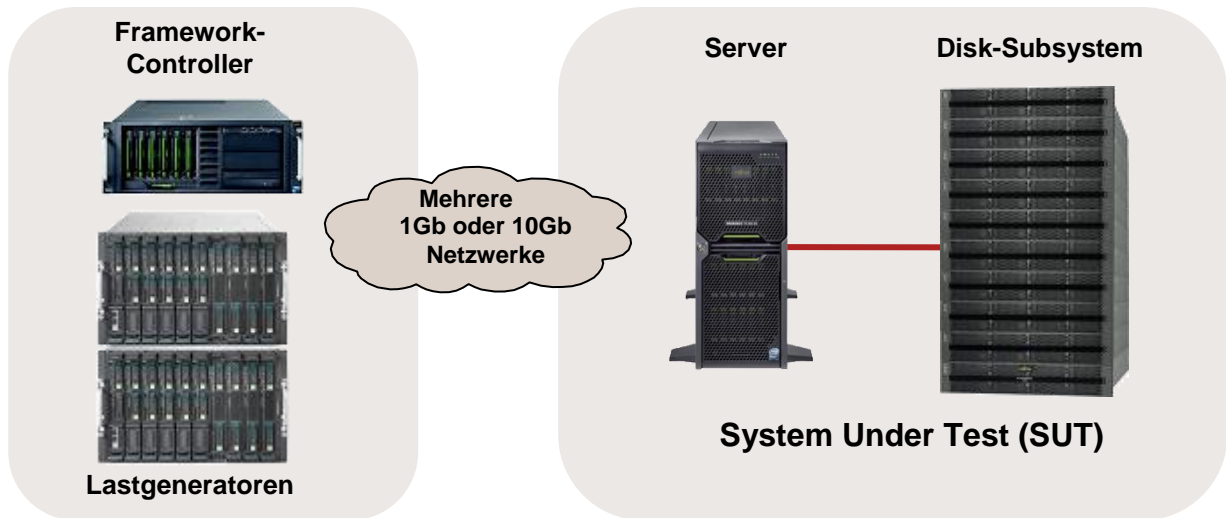
Der Verlauf der vServCon-Scores über die Tile-Anzahlen liefert nützliche Informationen über das Skalierungsverhalten des „System unter Test“.

Ferner werden bei vServCon die Gesamt-CPU-Auslastung des Hosts (VMs und alle übrigen CPU-Aktivitäten) und soweit möglich die elektrische Leistungsaufnahme dokumentiert.

Eine ausführliche Beschreibung von vServCon ist zu finden im Übersichtsdokument: [Benchmark-Überblick vServCon](#).

Benchmark-Umgebung

Der Messaufbau wird symbolisch durch folgende Grafik veranschaulicht:



System Under Test (SUT)	
Hardware	
Prozessor	Intel® Xeon® Processor E7 v3 Family
Speicher	2 TB: 64 x 32GB (2x16GB) 2Rx4 DDR4-2133 R ECC
Netzwerk-Interface	1 x dual port 1GbE adapter 1 x dual port 10GbE server adapter
Disk-Subsystem	1 x dual-channel FC-Controller Emulex LPe16002 LINUX/LIO basiertes Flash Storage System
Software	
Betriebssystem	VMware ESX 6.0.0 Build 2724185

Lastgenerator (inkl. Framework-Controller)	
Hardware (Shared)	
Gehäuse	PRIMERGY BX900
Hardware	
Modell	18 x PRIMERGY BX920 S1 Server-Blades
Prozessor	2 x Xeon X5570
Speicher	12 GB
Netzwerk-Interface	3 x 1 Gbit/s LAN
Software	
Betriebssystem	Microsoft Windows Server 2008 R2 Enterprise mit Hyper-V

Lastgenerator-VM (pro Tile 3 Lastgenerator-VMs auf verschiedenen Server-Blades)	
Hardware	
Prozessor	1 x logische CPU
Speicher	512 MB
Netzwerk-Interface	2 x 1 Gbit/s LAN
Software	
Betriebssystem	Microsoft Windows Server 2003 R2 Enterprise Edition

Einige Komponenten sind möglicherweise nicht in allen Ländern / Vertriebsregionen verfügbar.

Benchmark-Ergebnisse

Das hier behandelte PRIMEQUEST Acht-Sockel-System basiert auf Prozessoren der Intel® Xeon® Processor E7 v3 Family. Die Eigenschaften der Prozessoren sind im Kapitel „Technische Daten“ zusammengestellt.

Die verfügbaren Prozessoren dieser Systeme mit ihren Ergebnissen zeigt folgende Tabelle.

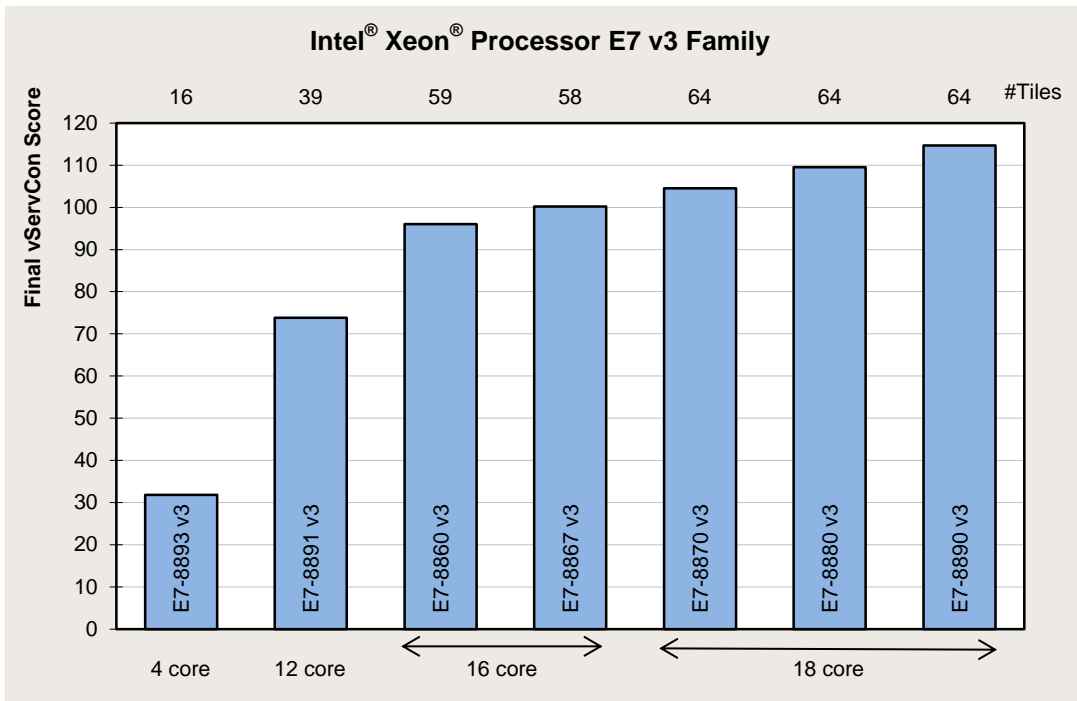
Prozessor		Score	#Tiles	
Intel® Xeon® Processor E7 v3 Family	4 Cores Hyper-Threading, Turbo-Modus	E7-8893 v3	31.8	16
	12 Cores Hyper-Threading, Turbo-Modus	E7-8891 v3	73.8	39
	16 Cores Hyper-Threading, Turbo-Modus	E7-8860 v3	96.0	59
		E7-8867 v3	100.2	58
	18 Cores Hyper-Threading, Turbo-Modus	E7-8870 v3	104.5	64
		E7-8880 v3	109.5	64
		E7-8890 v3	114.7	64

Diese PRIMEQUEST Acht-Sockel-Systeme sind durch weitere Fortschritte in der Prozessortechnologie gut für die Virtualisierung von Anwendungen geeignet. Verglichen mit einem System basierend auf der vorherigen Prozessorgeneration ist eine etwa 40.5% höhere Virtualisierungs-Performance (gemessen in vServCon-Score in der jeweils größten Konfiguration) erreichbar.

Die relativ großen Performance-Unterschiede zwischen den Prozessoren sind durch ihre Eigenschaften zu erklären. Die Werte skalieren aufgrund der Anzahl der Cores, der Größe des L3-Caches und der CPU-Taktfrequenz sowie durch die bei den meisten Prozessortypen vorhandenen Features Hyper-Threading und Turbo-Modus. Darüber hinaus bestimmt auch die Datenübertragungsrate zwischen den Prozessoren („QPI Speed“) die Performance. Grundsätzlich hat auch die Speicherzugriffsgeschwindigkeit Auswirkungen auf die Leistung. Als Richtschnur für die Auswahl von Arbeitsspeicher gilt im Virtualisierungsumfeld, dass eine ausreichende Menge wichtiger ist als die Geschwindigkeit der Speicherzugriffe.

Näheres zur Thematik „Speicher-Performance“ ist im White Paper [Speicher-Performance Xeon E7 v3 \(Haswell-EX\) basierter Systeme](#) zu finden.

Die erste Grafik vergleicht die mit den hier betrachteten Prozessortypen erreichbaren Werte der Virtualisierungs-Performance.

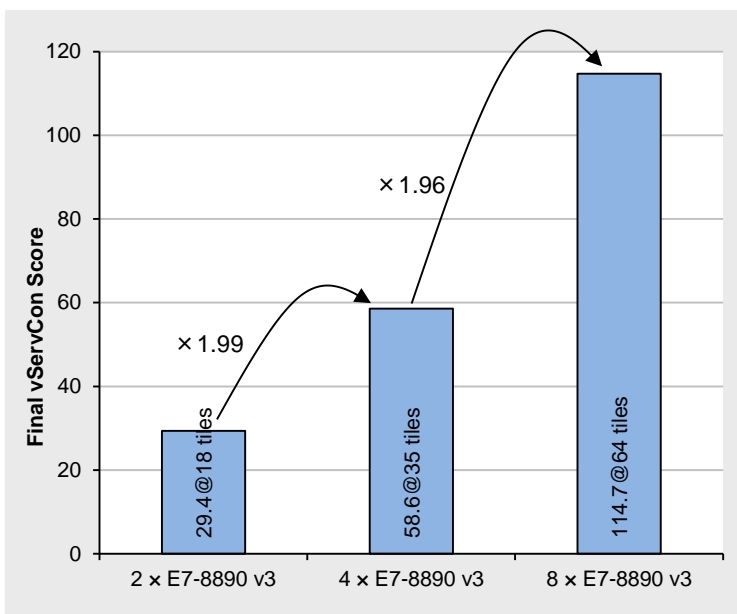


Den Einstieg stellt der Xeon E7-8893 v3 als Prozessoren mit nur vier Kernen dar.

Einen Leistungssprung erreicht der Prozessor mit zwölf Kernen (Xeon E7-8891 v3).

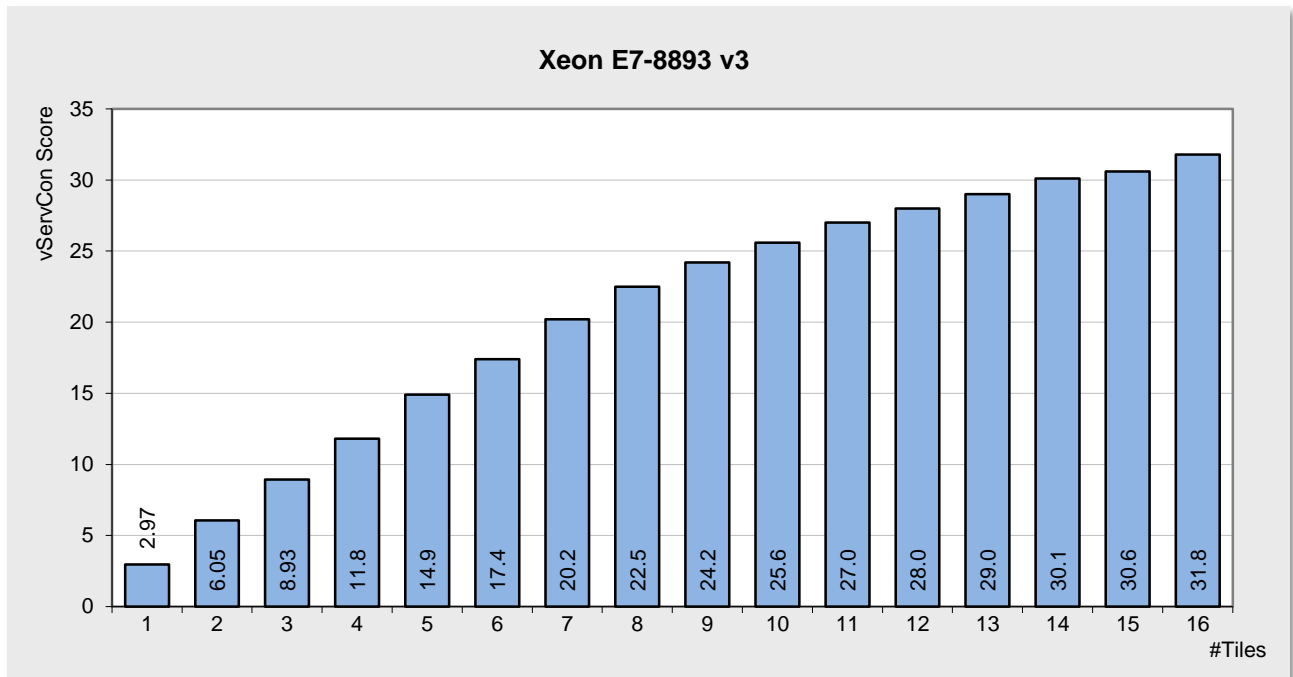
Am oberen Ende der Leistungsskala findet man die Gruppen der Prozessoren mit 18 Kernen, die eine höhere Leistung erreichen als die 16-Core Prozessoren.

Innerhalb einer Gruppe von Prozessoren mit gleicher Core-Anzahl sieht man eine Skalierung über die CPU-Taktfrequenz.



Bisher wurde die Virtualisierungs-Performance eines voll ausgebauten Systems betrachtet. Bei einem Server mit acht Sockeln stellt sich jedoch auch die Frage, wie gut die Performance von zwei auf vier bzw. acht Prozessoren skaliert. Je besser die Skalierung, desto geringer ist der Overhead, der durch die gemeinsame Nutzung der Ressourcen innerhalb eines Servers üblicherweise entsteht. Der Skalierungsfaktor hängt auch von der Anwendung ab. Dient der Server als Virtualisierungsplattform für die Server-Konsolidierung, skaliert das System mit dem Faktor 1.99 bzw. 1.96. Beim Betrieb mit vier bzw. acht Prozessoren erreicht das System also fast die doppelte Leistung wie mit zwei bzw. vier Prozessoren, wie die nebenstehende Grafik am Beispiel der Prozessorvariante Xeon E7-8890 v3 verdeutlicht.

Die nächste Grafik veranschaulicht die Virtualisierungs-Performance für wachsende VM-Anzahlen am Beispiel des Prozessors Xeon E7-8893 v3 (4-Core).



Neben der erhöhten Anzahl physikalischer Cores ist das Hyper-Threading, das alle Xeon E7 Prozessoren unterstützen, ein weiterer Grund für die hohe Anzahl betreibbarer VMs. Hierdurch wird bekanntermaßen ein physikalischer Prozessorkern in zwei logische Cores unterteilt und damit die für den Hypervisor verfügbare Anzahl Cores verdoppelt. Dieses standardmäßig eingestellte Feature steigert daher im Allgemeinen die Virtualisierungs-Performance eines Systems.

Der in der vorangegangenen Grafik dargestellte Verlauf der Skalierungskurve über die Tile-Anzahl ist spezifisch für Systeme mit Hyper-Threading. Bei den Prozessoren Xeon E7-8893 v3 stehen 32 physikalische und damit 64 logische Cores zur Verfügung, und pro Tile werden etwa vier davon verwendet (siehe [Benchmark-Beschreibung](#)). Das bedeutet, dass bis etwa acht Tiles eine parallele Nutzung gleicher physikalischer Cores durch mehrere VMs vermieden wird. Daher skaliert die Performance in diesem Bereich nahezu ideal. Darüber verläuft der Performance-Zuwachs bis zur CPU-Sättigung flacher.

Das vorige Bild hat den Aspekt der summierten Performance über alle Anwendungs-VMs eines Hosts betrachtet. Genauso interessant ist aber auch die Performance aus Sicht einer einzelnen Anwendungs-VM. Diese Information lässt sich ebenfalls aus dem vorigen Bild entnehmen. Im oben dargestellten Fall des Xeon E7-8893 v3 beispielsweise wird bei 48 Anwendungs-VMs (16 Tiles, die Idle-VMs nicht mitgezählt) das Gesamtoptimum erreicht; der Niedriglastfall wird durch drei Anwendungs-VMs (eine Tile, die Idle-VM nicht mitgezählt) repräsentiert. Man rufe sich in Erinnerung: der vServCon-Score für eine Tile ist ein Durchschnittswert über die drei Anwendungsszenarien in vServCon. Diese durchschnittliche Performance einer einzelnen Tile sinkt beim Übergang vom Niedriglastfall zum Gesamtoptimum des vServCon-Scores von 2.97 auf $31.8/16=2.05$, also auf 45%. Dabei können die einzelnen Typen von Anwendungs-VMs im Hochlastfall durchaus unterschiedlich reagieren. Hierdurch wird deutlich, dass man bezüglich der VM-Anzahlen auf einem Virtualisierungs-Host im konkreten Fall die Performance-Anforderungen einer einzelnen Anwendung gegen die Gesamtanforderungen abwägen muss.

VMmark V2

Benchmark-Beschreibung

VMmark V2 ist ein von VMware entwickelter Benchmark zum Vergleich von Serverkonfigurationen mit Hypervisor-Lösungen von VMware in Bezug auf ihre Eignung für Server-Konsolidierung. Neben der Software zur Lastgenerierung besteht der Benchmark aus einem definierten Lastprofil und aus einem verbindlichen Regelwerk. Die Benchmark-Ergebnisse können bei VMware eingereicht werden und werden nach einem erfolgreich durchlaufenen Review-Prozess auf deren Internet-Seite veröffentlicht. Nachdem der bewährte Benchmark „VMmark V1“ im Oktober 2010 eingestellt wurde, gibt es den Nachfolger „VMmark V2“, der ein Cluster aus mindestens zwei Servern voraussetzt und Datacenter-Funktionen wie Cloning und Deployment von virtuellen Maschinen (VMs), Load Balancing sowie die Verschiebung von VMs durch vMotion und auch Storage vMotion mit abdeckt.

Neben dem „Performance Only“ Ergebnis kann ab Version 2.5 von VMmark auch wahlweise die elektrische Leistungsaufnahme mit gemessen werden und als „Performance with Server Power“ (Leistungsaufnahme nur der Server-Systeme) und/oder „Performance with Server and Storage Power“ Resultat (Leistungsaufnahme der Server-Systeme und aller Storage-Komponenten) veröffentlicht werden.

Bei VMmark V2 handelt es sich nicht um einen neuen Benchmark im eigentlichen Sinn. Es ist vielmehr ein Framework, das bereits etablierte Benchmarks als Workloads zusammenfasst, um die Last einer konsolidierten und virtualisierten Serverumgebung nachzubilden. Drei bewährte Benchmarks, die die Anwendungsszenarien Mail-Server, Web 2.0 und E-Commerce abdecken, wurden in VMmark V2 integriert.

Anwendungsszenario	Last-Tool	# VMs
Mail-Server	LoadGen	1
Web 2.0	Olio client	2
E-Commerce	DVD Store 2 client	4
Standby-Server	(IdleVMTest)	1

Die drei Anwendungsszenarien werden jeweils insgesamt sieben dedizierten VMs zugeordnet. Hinzu kommt eine achte VM, der so genannte Standby-Server. Diese acht VMs bilden eine „Tile“ (englisch für „Kachel“).

Durch die Leistungsfähigkeit der zugrunde liegenden Server-Hardware ist es meist notwendig, dass im Rahmen einer Messung mehrere identische Tiles parallel gestartet werden müssen um eine maximale Gesamt-Performance zu erreichen.

Neu bei VMmark V2 ist eine Infrastruktur-Komponente, die einmal je zwei Hosts vorhanden ist. Diese misst Fähigkeiten der Datacenter-Konsolidierung durch VM Cloning und Deployment, vMotion und Storage vMotion. Zusätzlich wird die Load Balancing Fähigkeit des Datacenters eingesetzt (DRS, Distributed Resource Scheduler).

Das Ergebnis von VMmark V2 im Testtyp „Performance Only“ ist eine Zahl, „Score“ genannt, die Aufschluss über die Leistungsfähigkeit der vermessenen Virtualisierungslösung gibt. Der Score spiegelt für eine Server-Konfiguration mit Hypervisor den maximalen summarischen Konsolidierungs-Nutzen über alle Hosts und VMs wider und dient als Vergleichskriterium von verschiedenen Hardwareplattformen.

Dieser Score wird aus den Einzelergebnissen der VMs und einem Infrastruktur-Ergebnis ermittelt. Jede der fünf VMmark V2 Anwendungs- bzw. Front End-VMs ergibt für jede VM ein spezifisches Ergebnis in Form von anwendungsspezifischen Transaktionsraten. Um hieraus eine normalisierte Bewertungszahl zu bilden, werden die einzelnen Ergebnisse für eine Tile in Relation zu den jeweiligen Ergebnissen eines Referenzsystems gesetzt. Die daraus resultierenden dimensionslosen Performance-Werte werden dann für diese Tile geometrisch gemittelt und als letztes über alle Tiles aufsummiert. Dieser Wert geht mit einer Gewichtung von 80% in den Gesamt-Score ein. Der Infrastruktur-Workload ist im Benchmark nur einmal pro zwei Hosts vorhanden; er bestimmt das Resultat zu 20%. Für die Infrastruktur-Workload-Komponenten werden jeweils die Anzahl der Transaktionen pro Stunde und die durchschnittliche Dauer in Sekunden für den Score ermittelt.

Neben dem eigentlichen Score wird bei jedem VMmark V2 Ergebnis die Anzahl Tiles mit angegeben. Das ausgewiesene Resultat hat dann die Form „Score@Number of Tiles“, beispielsweise „4.20@5 Tiles“.

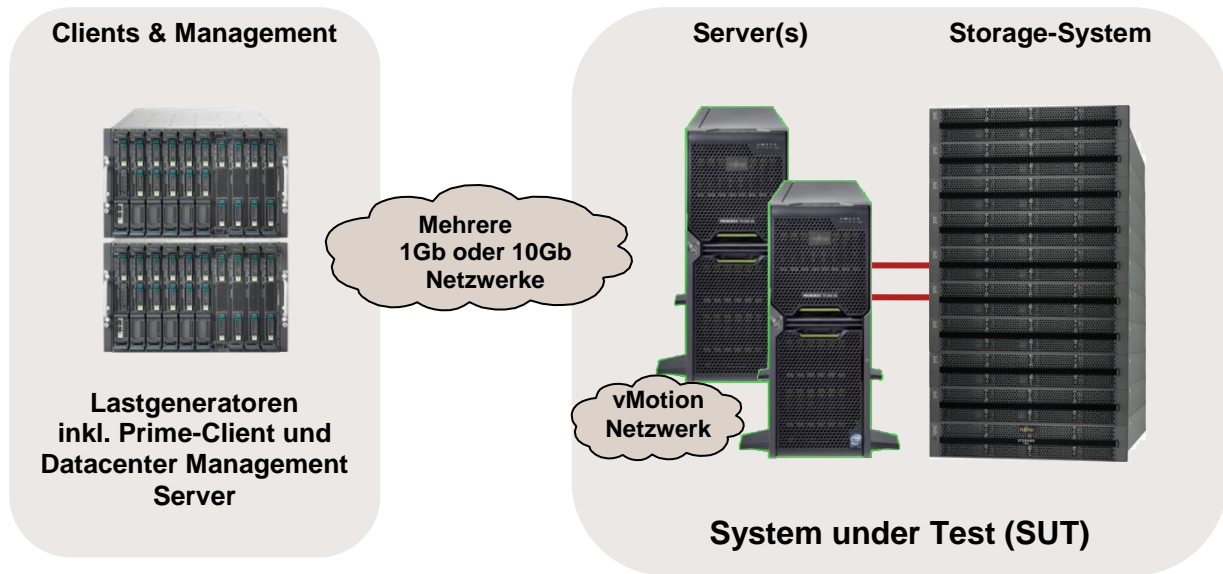
Bei den zwei Testtypen „Performance with Server Power“ und „Performance with Server and Storage Power“ wird ein sogenannter „Server PPKW Score“ bzw. „Server and Storage PPKW Score“ ermittelt, dies ist der Performance Score dividiert durch die durchschnittliche Leistungsaufnahme in Kilowatt (PPKW = Performance per Kilowatt (KW)).

Die Ergebnisse der drei Testtypen dürfen untereinander nicht verglichen werden.

Eine ausführliche Beschreibung von VMmark V2 ist im Übersichtsdokument [Benchmark Overview VMmark V2](#) zu finden.

Benchmark-Umgebung

Der Messaufbau wird symbolisch durch folgende Grafik veranschaulicht:



System Under Test (SUT)	
Hardware	
Anzahl Server	2
Modell	PRIMEQUEST 2800B2
Prozessor	8 x Xeon E7-8890 v3
Speicher	2048 GB: 64 x 32GB (2x16GB) 2Rx4 DDR4-2133 R ECC
Netzwerk-Interface	2 x Dual port Emulex OCe14102 10GbE Adapter
Disk-Subsystem	Dual port PFC EP LPe16002 4 x PRIMERGY RX300 S8 konfiguriert als Fibre Channel Target, Details siehe Disclosures
Software	
BIOS	Version 1.22
BIOS-Einstellungen	Siehe Details
Betriebssystem	VMware ESXi 6.0.0b Build 2809209
Betriebssystem-einstellungen	ESXi-Einstellungen: siehe Details

Datacenter Management Server (DMS)	
Hardware (Shared)	
Gehäuse	PRIMERGY BX600
Netzwerk-Switch	1 x PRIMERGY BX600 GbE Switch Blade 30/12
Hardware	
Modell	1 x Server-Blade PRIMERGY BX620 S5
Prozessor	2 x Xeon X5570
Speicher	24 GB
Netzwerk-Interface	6 x 1 Gbit/s LAN
Software	
Betriebssystem	VMware ESXi 5.1.0 Build 799733

Datacenter Management Server (DMS)-VM	
Hardware	
Prozessor	4 x logische CPU
Speicher	10 GB
Netzwerk-Interface	2 x 1 Gbit/s LAN
Software	
Betriebssystem	Microsoft Windows Server 2008 R2 Enterprise x64 Edition

Prime-Client	
Hardware (Shared)	
Gehäuse	PRIMERGY BX600
Netzwerk-Switch	1 x PRIMERGY BX600 GbE Switch Blade 30/12
Hardware	
Modell	1 x Server-Blade PRIMERGY BX620 S5
Prozessor	2 x Xeon X5570
Speicher	12 GB
Netzwerk-Interface	6 x 1 Gbit/s LAN
Software	
Betriebssystem	Microsoft Windows Server 2008 Enterprise x64 Edition SP2

Lastgenerator	
Hardware	
Modell	3 x PRIMERGY RX600 S6 1 x PRIMERGY RX500 S7
Prozessor	4 x Xeon E7-4870 (PRIMERGY RX600 S6) 4 x Xeon E5-4650 (PRIMERGY RX500 S7)
Speicher	PRIMERGY RX600 S6: 512 GB PRIMERGY RX500 S7: 256 GB
Netzwerk-Interface	PRIMERGY RX600 S6: 2 x 10 Gbit/s LAN PRIMERGY RX500 S7: 2 x 10 Gbit/s LAN
Software	
Betriebssystem	VMware ESX 4.1.0 U2 Build 502767

Lastgenerator-VMs (pro Tile 1 Lastgenerator-VM)	
Hardware	
Prozessor	4 x logische CPU
Speicher	4 GB
Netzwerk-Interface	1 x 10 Gbit/s LAN
Software	
Betriebssystem	Microsoft Windows Server 2008 Enterprise x64 Edition SP2

Details	
Siehe Disclosure	http://www.vmware.com/a/assets/vmmark/pdf/2015-08-04-Fujitsu-PRIMEQUEST2800B2.pdf

Einige Komponenten sind möglicherweise nicht in allen Ländern / Vertriebsregionen verfügbar.

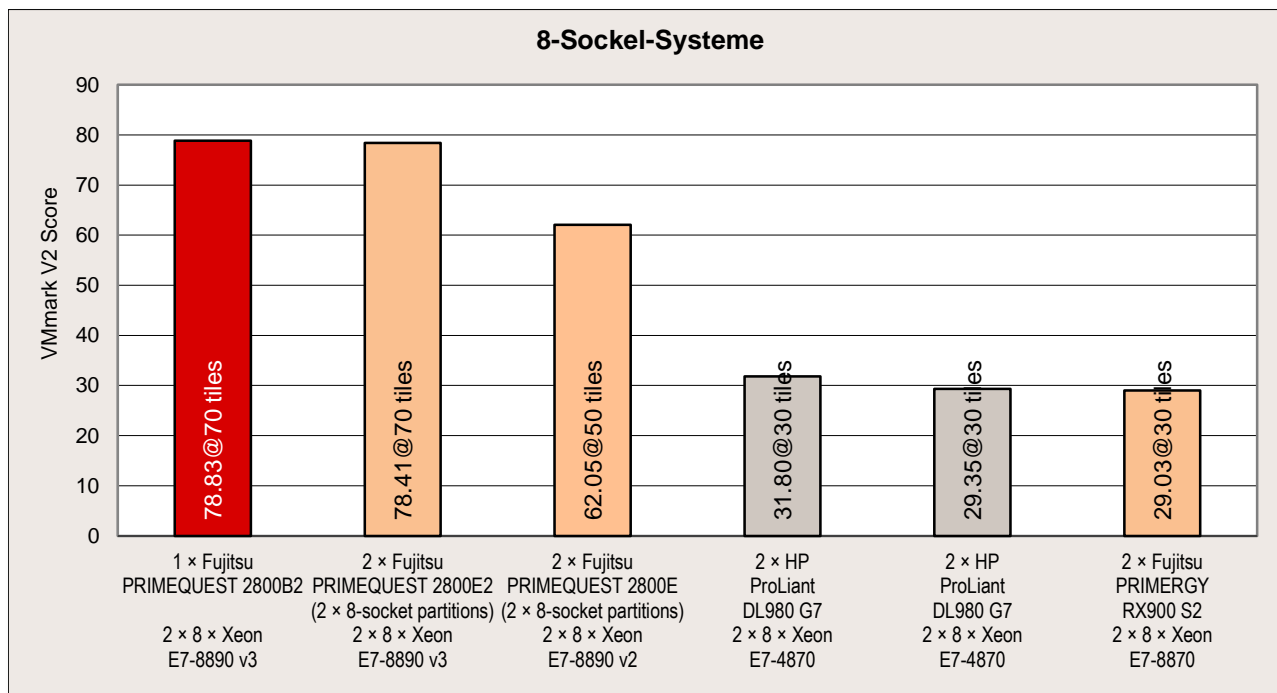
Benchmark-Ergebnisse



Am 4. August 2015 erzielte Fujitsu mit zwei PRIMEQUEST 2800B2 Systemen mit Xeon E7-8890 v3 Prozessoren und VMware ESXi 6.0.0 einen VMmark V2-Score von „78.83@70 tiles“ in einer Systemkonfiguration mit insgesamt 2 x 144 Prozessorkernen beim Einsatz von zwei identischen Servern/Partitionen im „System under Test“ (SUT). Mit diesem Ergebnis ist die PRIMEQUEST 2800B2 in der offiziellen VMmark V2-Rangliste der leistungsstärkste 8-Sockel-Server weltweit und auch der leistungsstärkste Server weltweit in einer „matched pair“ Konfiguration aus zwei identischen Hosts (zum Zeitpunkt der Veröffentlichung des Benchmark-Ergebnisses).

Alle genannten Vergleichswerte zu Wettbewerbsprodukten geben den Stand vom 4. August 2015 wieder. Die aktuellen VMmark V2-Ergebnisse sowie die ausführlichen Resultate und Konfigurationsdaten sind zu finden unter <http://www.vmware.com/a/vmmark/>.

Die Grafik zeigt das Ergebnis der PRIMEQUEST 2800B2 im Vergleich zu allen 8-Sockel-Systemen.



Die Differenz des Scores (in %) des Fujitsu Systems zu den anderen 8-Sockel-Systemen zeigt die nebenstehende Tabelle.

8-Sockel-Systeme	VMmark V2 Score	Differenz
Fujitsu PRIMEQUEST 2800B2	78.83@70 tiles	
Fujitsu PRIMEQUEST 2800E2	78.41@70 tiles	0.54%
Fujitsu PRIMEQUEST 2800E	62.05@50 tiles	27.04%
HP ProLiant DL980 G7	31.80@30 tiles	147.89%
HP ProLiant DL980 G7	29.35@30 tiles	168.59%
Fujitsu PRIMERGY RX900 S2	29.03@30 tiles	171.55%

Wesentliche Voraussetzungen zur Erreichung aller Ergebnisse der PRIMEQUEST 2800B2 waren die verwendeten Prozessoren, die bei gut eingestelltem Hypervisor ihre Prozessor-Features inklusive Hyper-Threading optimal nutzen konnten. All dies wirkt sich speziell bei der Virtualisierung positiv aus.

Alle VMs, deren Anwendungsdaten, das Host-Betriebssystem sowie weitere erforderliche Daten befanden sich auf einem leistungsfähigen Fibre-Channel Disk-Subsystem. Die Einrichtung des Disk-Subsystems berücksichtigt möglichst die spezifischen Anforderungen des Benchmarks. Der Einsatz von Flash-Technologie in Form von SAS-SSDs und PCIe-SSDs in dem leistungsfähigen Fibre-Channel Disk-Subsystem brachte weitere Vorteile bei den Antwortzeiten des eingesetzten Speichermediums.

Der Netzwerkanschluss an die Lastgeneratoren wurde über 10Gb LAN Ports realisiert. Die Infrastruktur-Workload-Verbindung zwischen den Hosts erfolgte per 1Gb LAN Ports.

Alle verwendeten Komponenten wurden dabei optimal aufeinander abgestimmt.

STREAM

Benchmark-Beschreibung

STREAM ist ein seit vielen Jahren eingesetzter synthetischer Benchmark zur Ermittlung des Speicherdurchsatzes, der von John McCalpin während seiner Professur an der Universität in Delaware entwickelt wurde. Heute wird STREAM an der Universität von Virginia betreut. Dort kann der Quellcode wahlweise in Fortran oder C heruntergeladen werden. Besonders im HPC-Umfeld spielt STREAM nach wie vor eine wichtige Rolle. So ist er z.B. Bestandteil der HPC Challenge Benchmark-Suite.

Der Benchmark ist so konzipiert, dass er sowohl auf PCs als auch auf Serversystemen eingesetzt werden kann. Die Maßeinheit des Benchmarks ist GB/s, also die Anzahl Gigabytes, die pro Sekunde gelesen und geschrieben werden kann.

STREAM misst den Speicherdurchsatz bei sequentiellen Zugriffen. Diese können generell effizienter durchgeführt werden als Zugriffe, die zufällig auf den Speicher verteilt sind, da bei sequentiellem Zugriff die Prozessor-Caches genutzt werden.

Vor der Ausführung wird der Quellcode der zu vermessenden Umgebung angepasst. So muss die Größe des Datenbereiches mindestens 12mal höher sein als die Summe aller „last-level“ Prozessor-Caches, damit diese einen möglichst geringen Einfluss auf das Ergebnis nehmen. Mit Hilfe der OpenMP-Programmibibliothek können ausgewählte Programmteile während der Laufzeit des Benchmarks parallel ausgeführt werden, wodurch eine optimale Lastverteilung auf die verfügbaren Prozessorkerne erreicht wird.

Bei der Ausführung wird der definierte Datenbereich, bestehend aus 8-Byte-Elementen, nacheinander auf vier Arten kopiert, wobei teilweise zusätzlich arithmetische Berechnungen durchgeführt werden.

Art	Ausführung	Bytes je Schritt	Gleitkommarechnung je Schritt
COPY	$a(i) = b(i)$	16	0
SCALE	$a(i) = q \times b(i)$	16	1
SUM	$a(i) = b(i) + c(i)$	24	1
TRIAD	$a(i) = b(i) + q \times c(i)$	24	2

Bei jeder Berechnungs-Art wird der Durchsatz in GB/s ausgegeben. Die Unterschiede der verschiedenen Werte sind auf modernen Systemen in der Regel nur gering. Zum Vergleich wird im Allgemeinen nur noch der ermittelte TRIAD-Wert verwendet.

Die Messergebnisse hängen in erster Linie von der Taktfrequenz der Speichermodule ab, die Prozessoren beeinflussen die arithmetischen Berechnungen.

In diesem Kapitel sind Durchsätze durchgängig zur Basis 10 angegeben ($1 \text{ GB/s} = 10^9 \text{ Byte/s}$).

Benchmark-Umgebung

System Under Test (SUT)	
Hardware	
Modell	PRIMEQUEST 2800B2
Prozessor	8 Prozessoren der Intel® Xeon® Processor E7-8800 v3 Product Family
Speicher	64 x 32GB (2x16GB) 2Rx4 DDR4-2133 R ECC
Software	
BIOS-Einstellungen	EnergyPerformance = Performance
Betriebssystem	Red Hat Enterprise Linux Server release 6.6
Betriebssystem-einstellungen	Transparent Huge Pages inactivated
Compiler	Intel C++ Composer XE 2015 for Linux
Benchmark	STREAM version 5.10

Einige Komponenten sind möglicherweise nicht in allen Ländern / Vertriebsregionen verfügbar.

Benchmark-Ergebnisse

Prozessor	Speicherfrequenz [MHz]	Max. Speicherbandbreite [GB/s]	Cores	Prozessorfrequenz [Ghz]	Anzahl Prozessoren	TRIAD [GB/s]
Xeon E7-8893 v3	1600	102	4	3.20	8	397
Xeon E7-8891 v3	1600	102	10	2.80	8	436
Xeon E7-8860 v3	1600	102	16	2.20	8	427
Xeon E7-8867 v3	1600	102	16	2.50	8	442
Xeon E7-8870 v3	1600	102	18	2.10	8	443
Xeon E7-8880 v3	1600	102	18	2.30	8	442
Xeon E7-8890 v3	1600	102	18	2.50	8	442


Literatur


PRIMEQUEST Server

<http://www.fujitsu.com/de/primequest>

PRIMEQUEST 2800B2

Dieses White Paper:

 <http://docs.ts.fujitsu.com/dl.aspx?id=e23ba4d1-936d-4e2d-98bd-d63ed173ec4a>

 <http://docs.ts.fujitsu.com/dl.aspx?id=135a7dc2-78c8-48ae-badc-96110c7bb12a>

Datenblatt

<http://docs.ts.fujitsu.com/dl.aspx?id=8d7f4737-c19e-4074-ad96-8a5d5d856bef>

PRIMEQUEST Performance

<http://www.fujitsu.com/de/x86-server-benchmarks>

Performance von Server-Komponenten

<http://www.fujitsu.com/de/products/computing/servers/mission-critical/benchmarks/x86-components.html>

Speicher-Performance Xeon E7 v3 (Haswell-EX) basierter Systeme

<http://docs.ts.fujitsu.com/dl.aspx?id=8bd39bba-41e6-4158-83b3-d8d630dfc41>

RAID-Controller-Performance

<http://docs.ts.fujitsu.com/dl.aspx?id=71fac54b-7ec3-4b3f-b13d-f80fbb42d583>

Disk-I/O: Performance von Speichermedien und RAID-Controllern

Grundlagen Disk-I/O-Performance

<http://docs.ts.fujitsu.com/dl.aspx?id=3d4fbad8-2a7e-465f-b9ee-d43b711f636d>

Informationen über Iometer

<http://www.iometer.org>

OLTP-2

Benchmark-Überblick OLTP-2

<http://docs.ts.fujitsu.com/dl.aspx?id=743d7d46-56e8-41d2-9d50-9ab29ccf4d18>

SPECcpu2006

<http://www.spec.org/osg/cpu2006>

Benchmark Überblick SPECcpu2006

<http://docs.ts.fujitsu.com/dl.aspx?id=04351fd2-8a69-42a3-ba1c-4342dcc89b89>

STREAM

<http://www.cs.virginia.edu/stream/>

VMmark V2

Benchmark-Überblick VMmark V2

<http://docs.ts.fujitsu.com/dl.aspx?id=ea18bef0-c1ff-46ae-81b3-c47811f866de>

VMmark V2

<http://www.vmmark.com>

vServCon

Benchmark-Überblick vServCon

<http://docs.ts.fujitsu.com/dl.aspx?id=214ee9dc-9239-4985-86e4-f0f9ac78ea25>

Kontakt

FUJITSU

Website: <http://www.fujitsu.com/de/>

PRIMEQUEST Product Marketing

<mailto:Primergy-PM@ts.fujitsu.com>

PRIMERGY Performance und Benchmarks

<mailto:primergy.benchmark@ts.fujitsu.com>