

Energieeffiziente IT und Infrastruktur für Rechenzentren und Serverräume

PrimeEnergyIT
EFFICIENT DATA CENTERS

Impressum:

Für den Inhalt verantwortlich: PrimeEnergyIT Projektkonsortium, November 2011

Projektkoordination: Dr. Bernd Schäppi, Austrian Energy Agency, Wien

Nachdruck nur auszugsweise und mit genauer Quellenangabe gestattet. Gedruckt auf chlorfrei gebleichtem Papier.

Die ausschliessliche Verantwortung für den Inhalt dieser Publikation liegt bei den Autoren. Sie gibt nicht notwendigerweise die Meinung der Europäischen Union wieder. Weder EACI noch die Europäische Kommission sind verantwortlich für irgendeinen Gebrauch der in dieser Publikation enthaltenen Information.

Effiziente Technologie für Energie- und Kosteneinsparungen in Rechenzentren und Serverräumen

Der Energieverbrauch von Rechenzentren und Serverräumen ist in den letzten zehn Jahren signifikant gestiegen. Immer leistungsfähigere Ausstattung und komplexere IT-Services treiben den Stromverbrauch in die Höhe. Da die Infrastruktur- und Energiekosten für Rechenzentren zu einem zentralen Faktor im Facility- und IT-Management geworden sind, wurde mittlerweile eine Reihe von Technologien entwickelt, um die Energieeffizienz zu erhöhen.

Gesamt gesehen ist das Energiesparpotential in Rechenzentren und Serverräumen hoch und beträgt in manchen Fällen sogar mehr als 50%, je nach spezifischer IT und Infrastruktur. In der Vergangenheit lag der Fokus von Energiesparmaßnahmen überwiegend auf effizienten Lösungen für die Stromversorgung und Kühlung. Mittlerweile werden zunehmend auch Maßnahmen zur Erhöhung der IT-Hardware-Effizienz forciert. Aktuelle Studien zeigen, dass bisherige Effizienzmaßnahmen – im Vergleich zu einem Business-as-usual-Szenario – bereits zu einer merklichen Reduktion des Energieverbrauchs führen [1]. Trotzdem ist das verbleibende Einsparpotential groß und neue Technologien ermöglichen noch effektivere Effizienzmaßnahmen.

Die vorliegende Broschüre bietet einen Überblick über aktuelle Energieeffizienztechnologien für IT und Infrastruktur in Rechenzentren und Serverräumen, mit Fokus auf der IT-Technologie, d.h. Server, Datenspeicherung und Netzwerkausstattung. Die Effizienzmaßnahmen inkludieren effektives Systemdesign, Powermanagement auf Hardware-Ebene und auf Ebene des gesamten Rechenzentrums sowie Konsolidierung und Virtualisierung.

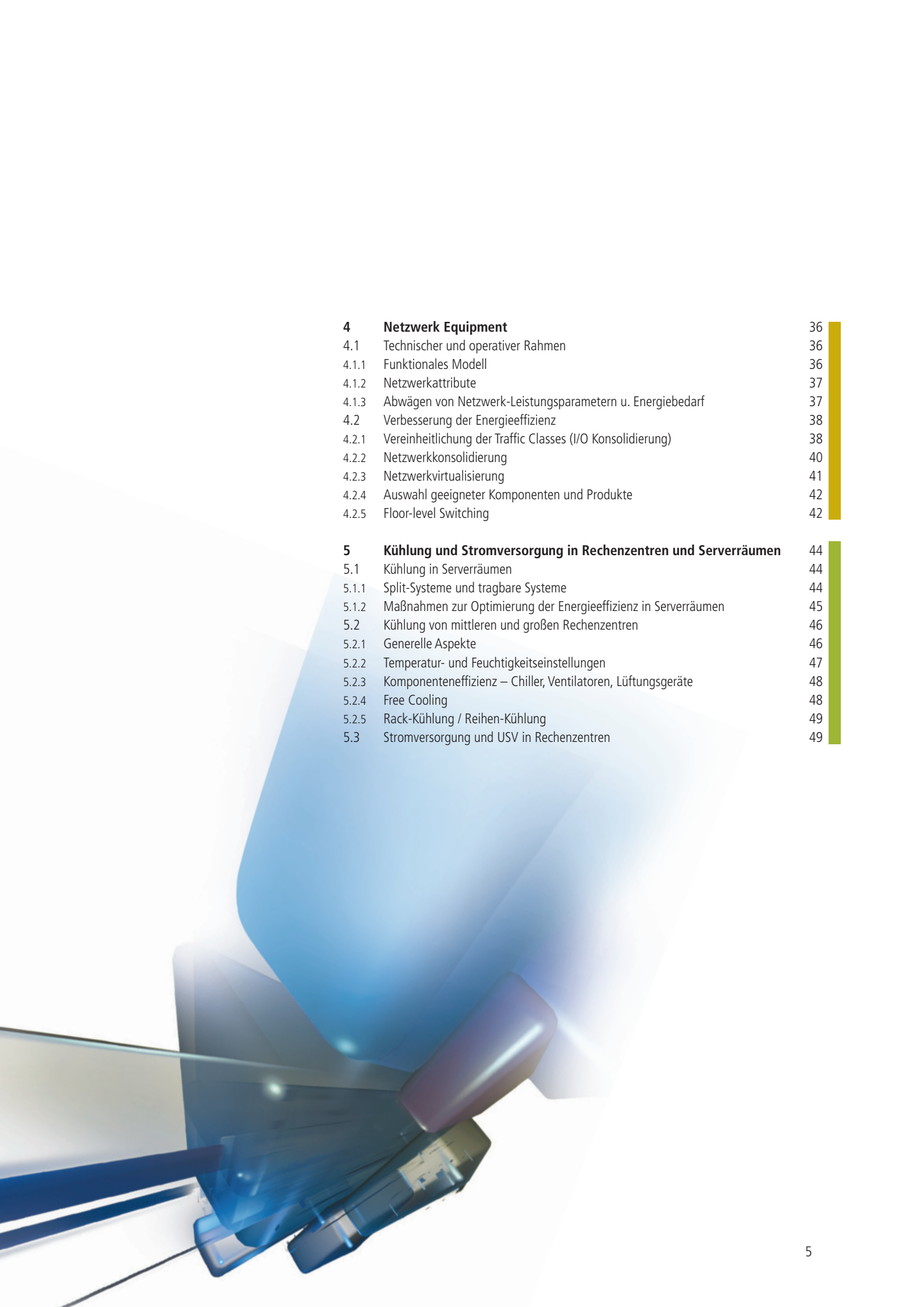
Die Broschüre bietet Basisinformationen für IT- und Infrastrukturmanager und soll damit Maßnahmen zur Verbesserung der Energie- und Kosteneffizienz in Rechenzentren und Serverräumen unterstützen. Best-Practice-Empfehlungen zeigen bewährte Maßnahmen auf, die man im Rechenzentrumsbetrieb und in der Beschaffung berücksichtigen sollte. Zu jedem Technologiebereich werden weiterführende Literaturhinweise angeboten.

Die vorliegende Broschüre entstand im Rahmen des internationalen Projekts PrimeEnergyIT mit finanzieller Unterstützung des EU-Programms „Intelligent Energy Europe“ und des Bundesministeriums für Wirtschaft, Familie und Jugend.

1) Koomey, J. (2011): Growth in Data center electricity use 2005 to 2010, Jonathan Koomey, Analytics Press, Oakland, CA, August 1, 2011

Inhalt

1	Monitoring des Energieverbrauchs in Serverräumen und Rechenzentren	6
1.1	Monitoringkonzepte	6
1.2	Messgeräte	9
2	Server Equipment	10
2.1	Energieeffizienz und Powermanagement auf Server- und Komponentenebene	10
2.1.1	CPU-Effizienz	12
2.1.2	Effizienz der Stromversorgung	13
2.2	Powermanagement auf Rack- bis Rechenzentrumsebene	14
2.2.1	Kapazitätsplanung und Energiemanagement	14
2.2.2	Power Capping	16
2.3	Powermanagement-Optionen für Blade Server	16
2.3.1	Blade Chassis und Bladedkomponenten	17
2.3.2	Bladesysteme – Energieversorgung und Kühlung	19
2.4	Servervirtualisierung	21
2.4.1	Energieeinsparungspotential von Virtualisierung	22
2.4.2	Anforderungen und Werkzeuge für die Planung von Virtualisierungsprojekten	23
2.4.3	Powermanagement in virtualisierten Umgebungen – Migration virtueller Server	24
2.4.4	Kühlung und Infrastruktur für virtualisierte Systeme	25
3	Equipment für Datenspeicherung	28
3.1	Speichergeräte und -medien	28
3.1.1	Bandsysteme	28
3.1.2	Festplattenlaufwerke (HDDs)	29
3.1.3	Solid State Drives (SSDs)	31
3.1.4	Hybrid Hard Drives (HHDs)	31
3.2	Datenspeicher-Gesamtlösungen	32
3.2.1	Laufwerke mit großer Kapazität und kleiner Baugröße	32
3.2.2	Massive Arrays of Idles Disks (MAIDs)	32
3.2.3	Effiziente RAID Level	32
3.2.4	Horizontales Speicher -Tiering, Speichervirtualisierung und Thin-Provisioning	33
3.2.5	Konsolidierung auf Speicher- und Strukturebene	34
3.2.6	Datendeduplizierung	34



4	Netzwerk Equipment	36
4.1	Technischer und operativer Rahmen	36
4.1.1	Funktionales Modell	36
4.1.2	Netzwerkattribute	37
4.1.3	Abwägen von Netzwerk-Leistungsparametern u. Energiebedarf	37
4.2	Verbesserung der Energieeffizienz	38
4.2.1	Vereinheitlichung der Traffic Classes (I/O Konsolidierung)	38
4.2.2	Netzwerkconsolidierung	40
4.2.3	Netzwerkvirtualisierung	41
4.2.4	Auswahl geeigneter Komponenten und Produkte	42
4.2.5	Floor-level Switching	42
5	Kühlung und Stromversorgung in Rechenzentren und Serverräumen	44
5.1	Kühlung in Serverräumen	44
5.1.1	Split-Systeme und tragbare Systeme	44
5.1.2	Maßnahmen zur Optimierung der Energieeffizienz in Serverräumen	45
5.2	Kühlung von mittleren und großen Rechenzentren	46
5.2.1	Generelle Aspekte	46
5.2.2	Temperatur- und Feuchtigkeitseinstellungen	47
5.2.3	Komponenteneffizienz – Chiller, Ventilatoren, Lüftungsgeräte	48
5.2.4	Free Cooling	48
5.2.5	Rack-Kühlung / Reihen-Kühlung	49
5.3	Stromversorgung und USV in Rechenzentren	49

1 Monitoring des Energieverbrauchs in Serverräumen und Rechenzentren

Carlos Patrao, University of Coimbra

1.1 Monitoringkonzepte

Ein entsprechendes Monitoring des Energieverbrauchs in Serverräumen und Rechenzentren ist wesentlich, um Energiesparpotentiale zu identifizieren und die Effektivität von Effizienzmaßnahmen evaluieren zu können. Monitoringkonzepte sollten sorgfältig entwickelt werden, um sicherzustellen, dass die richtigen Daten erfasst werden und so entsprechend effektive Maßnahmen gesetzt werden können. Folgende Aspekte sind zu berücksichtigen [1]:

- Erforderliche Genauigkeit/Auflösung der Daten
- Aufschlüsselung der Datenerfassung: Erfassung aller gewünschten Geräte
- Nutzerfreundlichkeit u. leichte Datenintegration
- Skalierbarkeit für Massenanwendung und Anwendung an mehreren Standorten
- Anpassungsfähigkeit an neue Messerfordernisse
- Optionen für Datenanalyse und Integration in Kontroll-/Steuerungssysteme
- Fähigkeit zur Problemerkennung und Problem- benachrichtigung
- Investitionskosten und Amortisationszeit

Folgende unterschiedlich aufwändige Monitoringmethoden können angewendet werden:

Minimales Monitoring – Die Durchführung periodischer punktueller Messungen mit mobilen Messinstrumenten ist in erster Linie eine Methode für sehr kleine Systeme. Es sind keine Investitionen für fix installierte Ausrüstung und Infrastruktur erforderlich.

Advanced Monitoring – Die Daten werden durch eine permanent installierte Messeinrichtung in Echtzeit erfasst. Es sind nur begrenzte Modifikationen an der Infrastruktur erforderlich.

State-of-the-Art Monitoring – Die Daten werden mittels automatisierter, permanenter Aufzeichnungssysteme in Echtzeit erfasst, unterstützt von Online-Software mit umfassenden Analyse-möglichkeiten. Dazu sind Modifikationen an der Infrastruktur und Unterstützung durch technische Experten nötig.

Das Monitoringsystem muss über die notwendige Anzahl von Monitoringpunkten verfügen, um die erforderlichen Informationen für eine umfassende Energieverbrauchsanalyse zu liefern. Bei größeren Anlagen sollte die Auswahl der "Informationsknoten" bei den – unter dem Gesichtspunkt des Energieverbrauchs – repräsentativsten Teilsystemen beginnen (siehe auch Abb. 1.1).

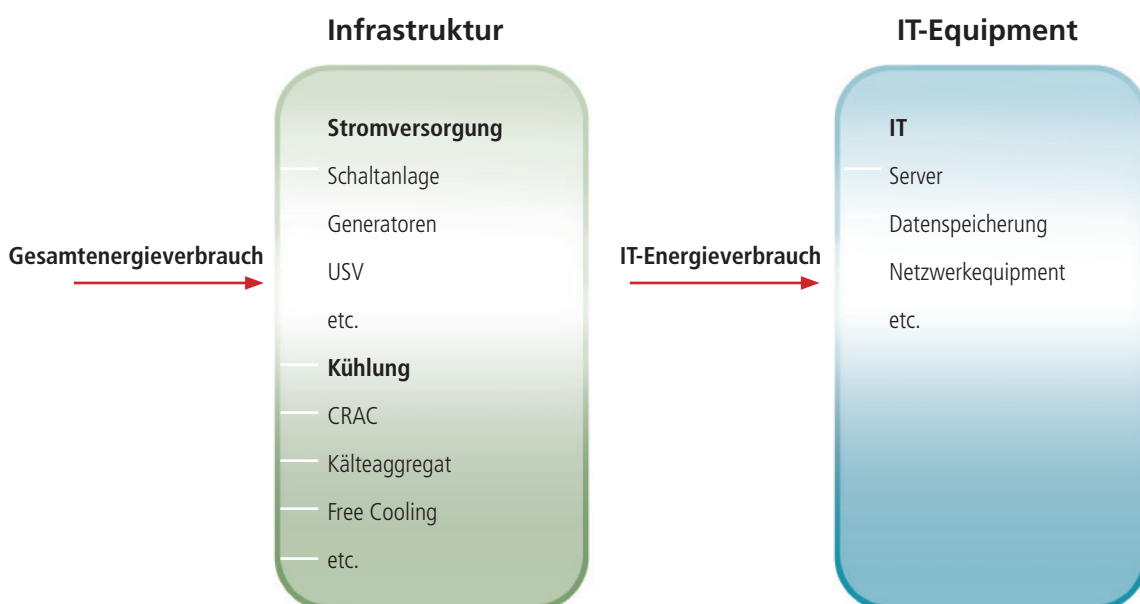


Abb. 1.1 Schematische Darstellung der zentralen Subsysteme [Quelle: ASHRAE [2]].

Die Datenerfassung, -verarbeitung und -auswertung erfolgt üblicherweise mithilfe von Software-Tools. Zu diesem Zweck hat beispielsweise das "Save Energy Now Program" (US Department of Energy) ein Softwarepaket namens "DC Pro" entwickelt. Die Software enthält einen Bewertungsprozess, Tools für Benchmarking und Performance-Überwachung wie auch Maßnahmenempfehlungen. Sie ist kostenlos erhältlich.

(<http://www1.eere.energy.gov/industrydatacenters/software.html>)

WEITERE NÜTZLICHE SOFTWARE-PAKETE:

- Power Nutzungs Tool
<http://estimator.thegreengrid.org/puee>
- PUE Reporting Tool
<http://www.thegreengrid.org/en/Global/Content/Tools/PUEReporting>
- PUE Metrik und Statistik Spreadsheet
http://www.thegreengrid.org/library-and-tools.aspx?category=MetricsAndMeasurements&range=Entire%20Archive&type=Tool&lang=en&paging=All#TB_inline?&inlined=sign_in
- PUE/DCiE Tool
<http://www.42u.com/measurement/pue-dcie.htm>

BEST PRACTICE EMPFEHLUNGEN

Für ein effektives Monitoringkonzept ist eine entsprechende Definition der Monitoringziele wesentlich.

Typische Ziele sind beispielsweise:






- Feststellung des Gesamtenergieverbrauchs von IT und Infrastruktur
- Analyse von Energieverbrauchstrends im zeitlichen Verlauf
- Erfassung des aktuellen Strombedarfs von spezifischem IT- und Infrastrukturequipment
- Abrechnung
- Berechnung von Energieeffizienzindizes und -messgrößen

Das Software-/Hardwarekonzept für das Energiemonitoring sollte die folgenden Möglichkeiten bieten (Quelle ASHRAE):

- Zuverlässige Datenerfassung und -speicherung in der erforderlichen Häufigkeit und Genauigkeit
- Normalisierung von Daten aus unterschiedlichen Geräten, Schnittstellen und Protokollen.
- Datenspeicherung für lange Messperioden
- Analyse und Visualisierung von Daten in Form von Tabellen und Grafiken
- Erweiterung der Architektur mit dem Ausbau des Rechenzentrums

Kernpunkte bei der Auswahl der Geräte für das Monitoringsystem sind u. a. Messbereich, Auflösung und Genauigkeit der Instrumente.

Tab. 1.1 Beispiele für Energiemessgeräte

Bezeichnung	Beispiel	Beschreibung	Monitoring Methode
Tragbares Messgerät	 <p>Quelle: Chauvin Arnoux</p>	Tragbare Messgeräte umfassen eine breite Produktpalette, angefangen von handlichen Einphasen-Mehrfachmessgeräten bis zu ausgeklügelten Dreiphasen-Stromanalysegeräten mit Aufzeichnungs- und Auslesemöglichkeiten.	Minimum und fortgeschrittenes Monitoring.
Schalttafelmessgerät	 <p>Quelle: Chauvin Arnoux</p>	Anzeigergeräte sind üblicherweise in den Schaltanlagen zur Messung von USV Systemen, Generatoren oder anderen Geräten fest installiert. Diese Messgeräte haben ein Display, das die aktuellen Messdaten und den aggregierten Energieverbrauch anzeigt. Man kann sie installieren, um den Gesamtverbrauch und den Verbrauch von Subsystemen aufzunehmen und zu verarbeiten.	Diese Messgeräte können für Best-Practice und für State-of-the-Art-Monitoring verwendet werden.
Hauptzähler	 <p>Quelle: Itron</p>	Gebührenzähler werden zumeist von Energieversorgern, Vermietern u. a. zur Kundenabrechnung verwendet. Kommen nur selten in Monitoringsystemen von Rechenzentren zum Einsatz, können Daten über den Gesamtenergieverbrauch einer Anlage zur Verfügung stellen. Eine digitale Kommunikationsschnittstelle macht die Auslesung und externe Speicherung von Daten für spätere Analysen möglich (z.B. alle 15 Minuten).	Bei allen Methoden verwendbar.
Intelligente Stromverteilungsleisten	 <p>Quelle: Raritan</p>	Intelligente Stromverteilungsleisten (PDUs) stellen Stromverbrauchsdaten zur Verfügung und ermöglichen es, Entscheidungen über Lastausgleich und die Restrukturierung von IT-Umgebungen zu treffen und so die Gesamtbetriebskosten zu verringern. PDUs können ein Monitoring für Strom, Spannung, Leistung, Leistungsfaktor und Energieverbrauch (kWh) mit ISO/IEC +/- 1% Abrechnungsgenauigkeit unterstützen. Nutzer können auf diese Steckdosenleisten über sichere Web, SNMP-, oder Telnet-Schnittstellen zugreifen und sie konfigurieren.	Bei allen Methoden verwendbar.
Serverintegrierte Messung		Durch Service-Prozessor unterstützte Messung.	Minimum und Advanced Monitoring.
Messumformer	 <p>Quelle: Chauvin Arnoux</p>	Der Messumformer ist üblicherweise ein Gerät ohne Display, das permanent in eine Schaltanlage eingebaut ist. Solche Geräte werden in Monitoringsystemen oft verwendet, um Leistungsmessungen von verschiedenen Punkten des Rechenzentrums zu erhalten.	Bei allen Methoden verwendbar.

1.2 Messgeräte

Für die Messung der Basisparameter wie Energieverbrauch, Temperatur, Luftdurchsatz und Luftfeuchtigkeit steht heute eine Vielzahl an Messgeräten zur Verfügung. Einige Beispiele für Energiemessgeräte sind in der Tabelle 1.1 dargestellt. Literaturhinweise finden sich im folgenden Abschnitt.

Literaturhinweise

ASHRAE (2010): Real-Time Energy Consumption Measurements in Data Centres, ASHRAE – American Society of Heating, Refrigerating and Air-Conditioning Engineers, 2010. ISBN: 978-1-933742-73-1

Stanley, J. and Koomey, J. (2009): The Science of Measurement: Improving Data Centre Performance with Continuous Monitoring and Measurement of Site Infrastructure, Stanley John and Koomey Jonathan, October 2009
www.analyticspress.com/scienceofmeasurement.html

Ton, M. et al (2008): DC Power for Improved Data Centre Efficiency, Ton, My, Fortenbery, Brian and Tschudi, William, Ecos Consulting, EPRI, Lawrence Berkeley National Laboratory, March 2008
http://hightech.lbl.gov/documents/data_centres/dcdemofinalreport.pdf

The Green Grid (2008): Green Grid Data Centre Power Efficiency Metrics. White Paper 6, The Green Grid, White Paper 6. December 30, 2008
<http://www.thegreengrid.org/Global/Content/white-papers/The-Green-Grid-Data-Centre-Power-Efficiency-Metrics-PUE-and-DCIE>

Rasmussen N. (2009): Determining Total Cost of Ownership for Data Centre and Network Room Infrastructure, Neil Rasmussen, APC by Schneider Electric, White paper #6 – Revision 4
http://www.apcmedia.com/salestools/CMRP-5T9PQG_R4_EN.pdf

Rasmussen N. (2010): Avoiding Costs From Oversizing Data Centre and Network Room Infrastructure, Neil Rasmussen, APC by Schneider Electric, 2010. White paper #37 – Revision 6
http://www.apcmedia.com/salestools/SADE-5TNNEP_R6_EN.pdf

Schneider Electric (2011): E-learning website (Energy University) that provides the latest information and training on Energy Efficiency concepts and best practice
www.myenergyuniversity.com

Webinar: „The Data Centre in Real Time: Monitoring Tools Overview & Demonstration“
<http://www.42u.com/webinars/Real-Time-Measurement-Webinar/playback.htm>

Zitierte Literatur

- [1] **Stanley, J. and Koomey, J. (2009):** The Science of Measurement: Improving Data Centre Performance with Continuous Monitoring and Measurement of Site Infrastructure. October 2009.
[2] **ASHRAE (2010):** Real-Time Energy Consumption Measurements in Data Centres: ASHRAE- American Society of Heating, Refrigerating and Air-Conditioning Engineers, 2010. ISBN: 978-1-933742-73-1.

2 Server Equipment

Bernd Schäppi, Thomas Bogner, Hellmut Teschner, Austrian Energy Agency

Server-Equipment verbraucht ca. 30–40% des gesamten Stroms in Rechenzentren und Serverräumen. Daher ist es einer der vorrangigen Bereiche, um effektive Energiesparmaßnahmen zu setzen. Das Energieeffizienzpotenzial ist häufig hoch und je nach Art des IT-Systems und der angewendeten Maßnahmen können 20–60% Energieeinsparung oder sogar mehr erzielt werden. Die primären Optionen zur Verbesserung der Energieeffizienz reichen von der Auswahl energieeffizienter Hardware und einem energieeffizienten Systemsdesign über Powermanagement auf allen Systemebenen bis zu Hardwarekonsolidierung und Virtualisierung.

Das folgende Kapitel bietet Informationen über Stromspartechnologien und entsprechende Effizienzoptionen von der Komponenten- bis zur Systemebene. Es werden Optimierungsoptionen für alle Ebenen, vom Server übers Rack bis hin zum gesamten Rechenzentrum oder Serverraum vorgestellt. Spezielle Abschnitte sind den Themen Blade-Server-Technologie und Servervirtualisierung gewidmet. Jeder Abschnitt bietet darüber hinaus Best-Practice-Empfehlungen.

2.1 Energieeffizienz und Powermanagement auf Server- und Komponentenebene

Die Energieeffizienz von Servern wurde in den letzten Jahren deutlich verbessert, vor allem dank der Entwicklung von effektivem Powermanagement für Hardwarekomponenten. Die bisherigen Standards für die Bewertung und Deklaration der Energieeffizienz von Servern basieren auf Energy Star und dem Power Benchmark von SPEC (SPEC: Standard Performance Evaluation Corporation). Die aktuellen Energy-Star-Anforderungen für Enterprise Server [1] definieren Energieeffizienzkriterien für Rack- und Tower-Server mit bis zu 4 Prozessorsockeln. In der aktuellen Version der Anforderungen sind bislang nur Kriterien für den maximalen Stromverbrauch der Server im Leerlaufmodus, für die Effizienz der Netzteile und für Powermanagementoptionen festgelegt (Tab. 2.1 und 2.4). Die Kriterien für den Leerlaufmodus sind in erster Linie als Effizienzindikator beim Betrieb mit geringen durchschnittlichen Rechenlasten interessant. Geringe Rechenlasten auf Servern sind immer noch üblich. Hardware-Konsolidierung und damit bessere Auslastung der Server sollte daher ein generelles Ziel sein.

Energieeffizienz von Servern bei höheren Rechenlasten bzw. für konsolidierte Systeme wird bislang mit dem SPECpower-Benchmark bewertet (s. Information unten). Ein umfassendes Server Efficiency Rating Tool (SERT), das alle wesentlichen Server-Hardwarekomponenten adressiert, ist derzeit bei SPEC [2] in Entwicklung und wird ab 2012 erhältlich sein. Das SERT Tool bewertet die Servereffizienz auf Basis von partiellen Benchmarks für CPU, Memory, Disk und System (Tab. 2.2). Das Tool wird IT-Manager auch dabei unterstützen, energieeffiziente Hardware für spezifische Applikationen auszuwählen.

Tab. 2.1 Energy Star Kriterien für Leerlauf-Modus

Kategorie	Anzahl Prozessoren	Managed Server	Leistungsaufnahme Leerlaufbetrieb (W)
A	1	nein	55
B	1	ja	65
C	2	nein	100
D	2	ja	150

Tab. 2.2 Konzept des SERT-Tools

	Server	
System-Benchmark	CPU	Benchmark
	Speicher	Benchmark
	Disk	Benchmark
	IO	

SPECpower_ssj2008 [2] war der erste Standard-Benchmark, der zur Energieeffizienzbewertung von Servern eingeführt wurde. Er adressiert vor allem die CPU-Effizienz und bietet eine gute Bewertungsmethode hinsichtlich CPU-intensiver Workloads. Der Benchmark wird von Herstellern derzeit nur für ausgewählte Hardware veröffentlicht. Die typische SPECpower Benchmark-Grafik bietet Information über die durchschnittliche Rechenleistung pro Watt über die gesamte Bandbreite der Rechenlast, d.h. für insgesamt 10 Lastniveaus. So kann die Energieeffizienz der Server auf verschiedenen Lastlevels vom Leerlaufbetrieb bis 100% Last verglichen werden. Für die Hardware-Beschaffung sollte die komplette SPECpower Information (inklusive detaillierter Konfigurationsinformationen) berücksichtigt werden. Darüber hinaus ist zu bedenken, dass Produkte oft in niedrigen Konfigurationen getestet werden.

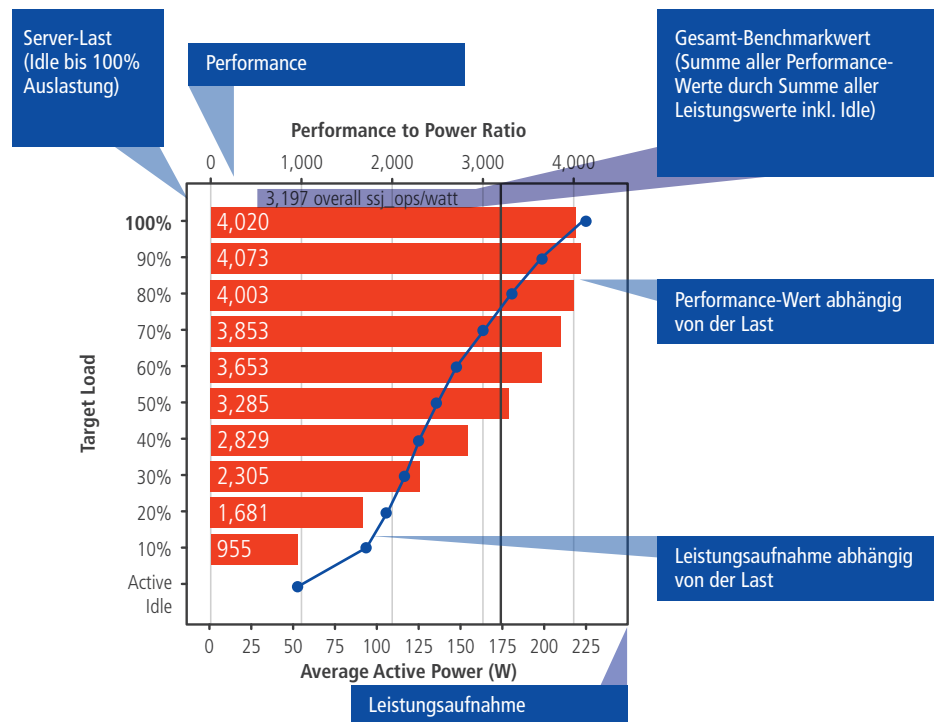


Abb. 2.1 SPECpower Diagramm

Best Practice Empfehlungen Energieeffizienzkriterien und Benchmarks für die Auswahl der Hardware

- Berücksichtigen Sie bei der Hardware-Beschaffung die Energieeffizienzkriterien von Energy Star. Für Server, die mit geringer Rechenlast betrieben werden, können bereits die aktuellen Kriterien für den Leerlauf-Modus als brauchbare Effizienzindikatoren herangezogen werden. Die Anforderungen für Netzteile sind generell anwendbar.
- Verlangen Sie vom Hersteller Benchmark-Daten für SPECpower_ssj2008 bzw. sobald verfügbar auch für SPEC-SERT. Hinsichtlich der Interpretation von SPECpower sind folgende Punkte zu beachten:
 - Es handelt sich um einen CPU-zentrierten Benchmark, d.h. repräsentativ für CPU-intensive Workloads.
 - Server wurden möglicherweise in eher niedriger Konfiguration getestet (daher Konfiguration überprüfen).
 - Berücksichtigen Sie nicht nur den SPEC-Gesamtbenchmark, sondern auch die Detaildaten.

2.1.1 CPU-Effizienz

CPUs sind die Serverkomponenten mit dem höchsten Energieverbrauch. Energieeffiziente CPU-Modelle mit effektivem Powermanagement können daher wesentlich zur Gesamteffizienz beitragen. Der Energieverbrauch von CPUs hängt von der spezifischen Spannung und der Taktfrequenz ab. Powermanagement auf CPU- oder Kern-Level basiert daher auf Dynamic Voltage und Frequency Scaling (DVFS) oder einem Abschalten der Kerne. Der Energieverbrauch von CPUs wird oftmals auf der Basis der Thermal Design Power (TDP) verglichen, einem Indikator für die maximale thermische Verlustleistung als Basis für die Auslegung des Server-Kühlsystems. Die TDP bietet allerdings nur begrenzte Information, da die Effizienz auch vom Powermanagement abhängt. Hersteller bieten auch spezielle Low-Power CPU Versionen an, die signifikante Energieeinsparungen erlauben. Die Energieeffizienz von CPUs hängt stark von der effektiven Implementierung des Powerma-

agements ab. Herkömmliche Betriebssysteme unterstützen Powermanagement auf der Basis der Advanced Configuration and Power Interface (ACPI) Spezifikationen für die Prozessorleistung, den Stromverbrauch (P-States) und das Wärme-management (C-States). Die neuen Steuerungskomponenten unter ACPI Vs 3 bieten High-Level Powermanagement, das eine feinkörnigere Leistungsanpassung je nach Bedarf ermöglicht. In vielen Servermodellen können vordefinierte Power-Profile eingestellt werden, z.B.:

- "High performance": geeignet für Server, die mit hoher Auslastung laufen und ohne Rücksicht auf Stromkosten maximale Leistung erbringen müssen.
- "Power saver mode" / "Minimum power usage": geeignet für Server, die mit geringer Auslastung laufen und mehr Leistungskapazität haben, als gebraucht wird. Mit diesem Modus sind stufenweise Stromeinsparungen möglich.
- "Balanced power and performance"

Abb. 2.2 zeigt die positiven Effekte von modernem CPU Powermanagement anhand von Benchmark-Ergebnissen (SPECpower) für die Server-Produktfamilie HP ProLiant DL 380. Die Leistungsaufnahme im Leerlaufbetrieb und bei niedrigen Lasten wurde bei dem spezifischen Servermodell von Produkt-Generation G5 zu G7 deutlich reduziert. Beim Server DL 380 G5 war der Leerlaufverbrauch (keine Last) nur um 33% (170 Watt) niedriger als die Leistungsaufnahme unter Volllast (253 Watt). Beim Modell G7 beträgt der Unterschied bereits 75%. Das zeigt, dass die heutige Servertechnologie dank intelligentem Powermanagement im Bereich niedriger Rechenlasten und im Leerlaufbetrieb viel energieeffizienter geworden ist. Gleichzeitig wurde die Rechenleistung des spezifischen Servermodells um mehr als den Faktor 3 erhöht.

Für die Hardwarekonfiguration bei der Beschaffung ist es wesentlich, die konkreten Performance-

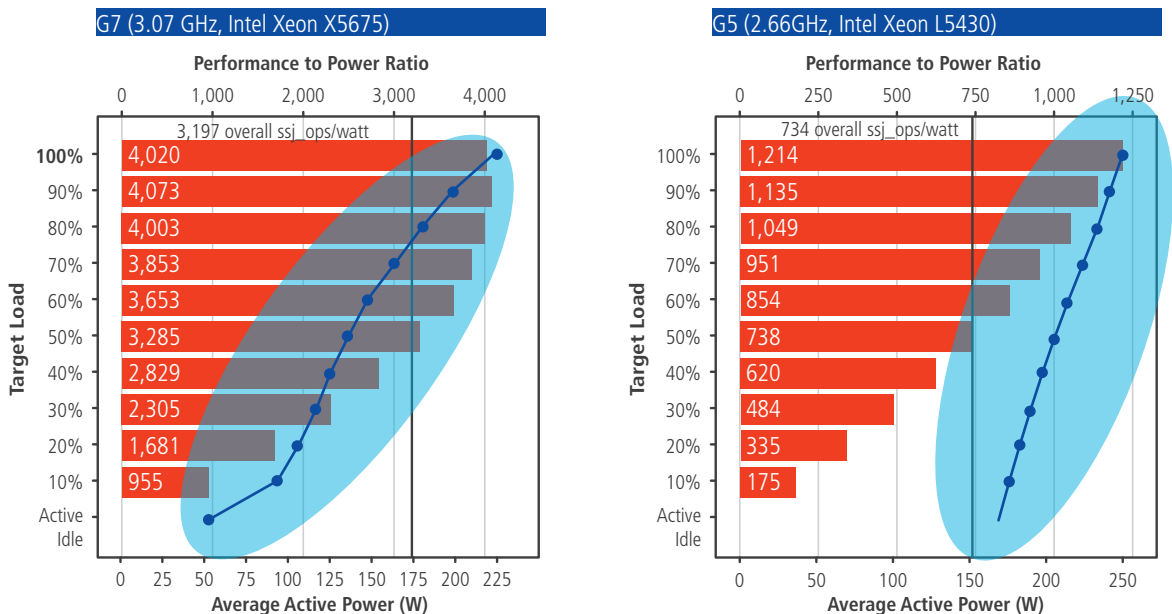


Abb. 2.2 SPECpower-Benchmark für verschiedene Server Generationen (G5, G7 Server from HP) [SPEC (2010, www.spec.org)]

anforderungen der vorgesehenen Applikationen an die Hardware-Komponenten zu berücksichtigen. Unterschiedliche Applikationen stellen unterschiedliche Anforderungen an die Hardware-Performance. Die Berücksichtigung des konkreten Bedarfes erlaubt es daher, eine Überdimensionierung bzw. überhöhte Performance-Kapazitäten zu vermeiden. Tab. 2.3 bietet eine grobe Übersicht über Hardware-Performanceanforderungen für unterschiedliche Workloads.

Tab. 2.3 Performance Anforderungen von Server-Anwendungen [5]

Kategorie	CPU	RAM	Hard Disks	IO
File/print server	0	+	++	+
Mail server	+	+	++	0
Virtualization server	++	+++	++	++
Web server	+	+	0	+
Database server	++	++	+++	+
Application server	++	++	0	+
Terminal server	++	++	+	+

2.1.2 Effizienz der Stromversorgung

Das Energy Star Programm für Server [1] definiert Anforderungen an die Netzteil-effizienz und dabei Kriterien für die Lastniveaus 10%, 20%, 50% und 100%. Das 80 PLUS-Programm [3] bietet ebenfalls Energieeffizienz-Anforderungen für Netzteile, allerdings ohne Berücksichtigung des 10%-Lastniveaus. Für die Beschaffung empfiehlt sich die

Wahl von Netzteilen, die dem 80 PLUS Gold Level entsprechen, d.h. eine Effizienz von 88% bei 20% Last und 92% bei 50% Last.

Standard-Rackserver, die im Allgemeinen unter geringer Last laufen, sind oft mit überdimensionierten Netzteilen ausgestattet. Das ist ineffizient. Die richtige Dimensionierung der Netzteile ist daher wesentlich. Unterstützung dafür bieten beispielsweise Power-Konfigurationstools der Her-

steller und Tools für Power Capping.

Einige Hersteller bieten spezielle Hardware-Features an, die es erlauben, unnötige Verluste bei redundanten Netzteilen zu vermeiden. Derartige Hardware verfügt über einen Betriebsmodus, bei dem nur ein Netzteil genutzt wird, solange die Last einen bestimmten Schwellenwert nicht überschreitet. Das zweite Netzteil befindet sich im Standby-Modus und bewahrt gleichzeitig die Redundanz im Falle eines Strom- oder Netzausfalls.

Tab. 2.4 Effizienzanforderungen für Netzteile: Energy Star und 80 PLUS Initiative [1, 3]

	Netzteil- Typ	Nennleistung	10% Last	20% Last	50% Last	100% Last
Energy Star Vs1	Multi-output (AC-DC & DC-DC)	Alle Level	N/A	82%	85%	82%
	Single-output (AC-DC & DC-DC)	≤ 500 W	70%	82%	89%	85%
		>500–1.000 W	75%	85%	89%	85%
		> 1.000 W	80%	88%	92%	88%
Energy Star Vs2 Draft	Multi-output (AC-DC & DC-DC)	Alle Level	N/A	85%	88%	85%
	Single-output (AV-DC & DC-DC)	Alle Level	80%	88%	92%	88%
80 PLUS	Bronze	Alle Level	N/A	81%	85%	81%
	Silber	Alle Level	N/A	85%	89%	85%
	Gold	Alle Level	N/A	88%	92%	88%
	Platin	Alle Level	N/A	90%	94%	91%

2.2 Powermanagement auf Rack- bis Rechenzentrums-ebene

Powermanagement ist auch auf Systemebene wesentlich, um die Gesamtenergieeffizienz zu optimieren. Wie bereits oben dargestellt, wird eine große Zahl von Servern nach wie vor unter geringer Last betrieben und es besteht ein großes Potenzial für Energieeinsparungen durch Hardware-Konsolidierung oder Powermanagement auf Systemebene. Ähnlich wie auf der Komponentenebene passt das Powermanagement auf höheren Ebenen die Leistungsaufnahme an den aktuellen Bedarf an und schaltet Ressourcen ab oder drosselt sie, wenn sie nicht gebraucht werden. Tab. 2.5 zeigt die verschiedenen Methoden von Powermanagement auf unterschiedlichen Ebenen [7]. Einige Optionen werden in den folgenden Abschnitten und in späteren Kapiteln vorgestellt.

2.2.1 Kapazitätsplanung und Energiemanagement

Servermanagement-Software stellt wesentliche Tools für den sicheren Serverbetrieb, aber auch für ein gesamtheitliches Powermanagement zur Verfügung. Servermanagement-Tools können einen wesentlichen Beitrag leisten, um den Energieverbrauch zu reduzieren. Sie unterstützen die Anwendung von Energiekonzepten im Serversystem und Features wie Provisioning, Monitoring und Konfigurationsmanagement. Wesentliche Features sind im Allgemeinen:

- Provisioning
- Monitoring
- Deployment
- Konfigurationsmanagement
- Update
- Powermanagement
- Workloadmanagement

Alle größeren Hardware-Hersteller bieten mächtige Servermanagement-Tools an. IBM (Systems

Director) und HP (Systems Insight Manager mit Insight Dynamics) bieten umfassende Managementlösungen an, die auch Systeme von Drittanbietern integrieren können. Fujitsu (Server View Site) bietet Produkte mit Basisfunktionen, die sich in vorhandene Managementkonsolen anderer Anbieter integrieren lassen. DELL verwendet die Altiris Total Management Suite. Sun und Acer bieten Konsolen für ihre eigenen Umgebungen an.

Energiemanagement Suites

(z.B. IBM Energy Manager)

Neben anderen Features unterstützt diese Art von Tools das Monitoring und die Erfassung von Stromverbrauchsdaten, weiters das Powermanagement inkl. Festlegung von Stromsparoptionen und Power Caps ebenso wie die Automatisierung von anderen Funktionen zur Energieüberwachung und -einsparung. Zu letzteren zählen die Konfiguration von Messzählern wie PDUs und Sensoren, die Festlegung von Schwellenwerten, die Erstellung und Festlegung von Energiekonzepten und die Berechnung der Energiekosten. Weitere Information zu den Energiemanagement Suites, s.u.

Tab. 2.5. Power Management Optionen von der IT-Komponente zum gesamten Rechenzentrum [7]

Komponenten Level	System Level	Rack Level	RZ Level
<ul style="list-style-type: none"> • CPU (Core C-states, P-States, T-states, Thermal throttle) • Andere Komponenten (D-states, L-states) 	<ul style="list-style-type: none"> • S-States • Platform-based Power Management • Workload Scheduler • Regelung der Lüfter 	<ul style="list-style-type: none"> • System oder Node Management • Application/Load Balancing • Chassis Management 	<ul style="list-style-type: none"> • Application/Load Balancing • Facility und Equipment Monitoring • Datenduplizierung • Multi-Rack Management, dynamische Konsolidierung

Kapazitätsplanungstools

(z.B. HP Capacity Planner)

Neben anderen Features unterstützen Kapazitätsplaner IT-Manager hinsichtlich der optimalen Auslastung der IT-Ressourcen, der Energieverbrauchsreduktion und der Verbesserung der Performance. Sie ermöglichen die Erfassung von Nutzungsdaten für CPU-Kerne, Speicher, Netzwerk, Festplatten und Netzteile. Darüber hinaus unterstützen sie die Workload-Planung oder Systemänderungen und eine Bewertung der Auswirkungen auf die Ressourcennutzung. Sie evaluieren auch Trends zur Abschätzung des zukünftigen Ressourcenbedarfs. Weitere Informationen zu Kapazitätsplanungstools, s. u..

Das HP Tool bietet beispielsweise auf Basis von Auslastungsaufzeichnungen eine gute Grundlage für die Bewertung des Ressourcenbedarfs für zusammengeführte Anwendungen und damit eine gute Entscheidungsbasis für Konsolidierungsmaßnahmen. Abb. 2.3 zeigt als Beispiel den Vergleich der Auslastung von zwei Systemen. Es wird ersichtlich, dass die jeweiligen Leistungspeaks zeitlich unterschiedlich liegen und die durchschnittliche Last im Falle einer Hardware-Konsolidierung nur geringfügig anwachsen würde.

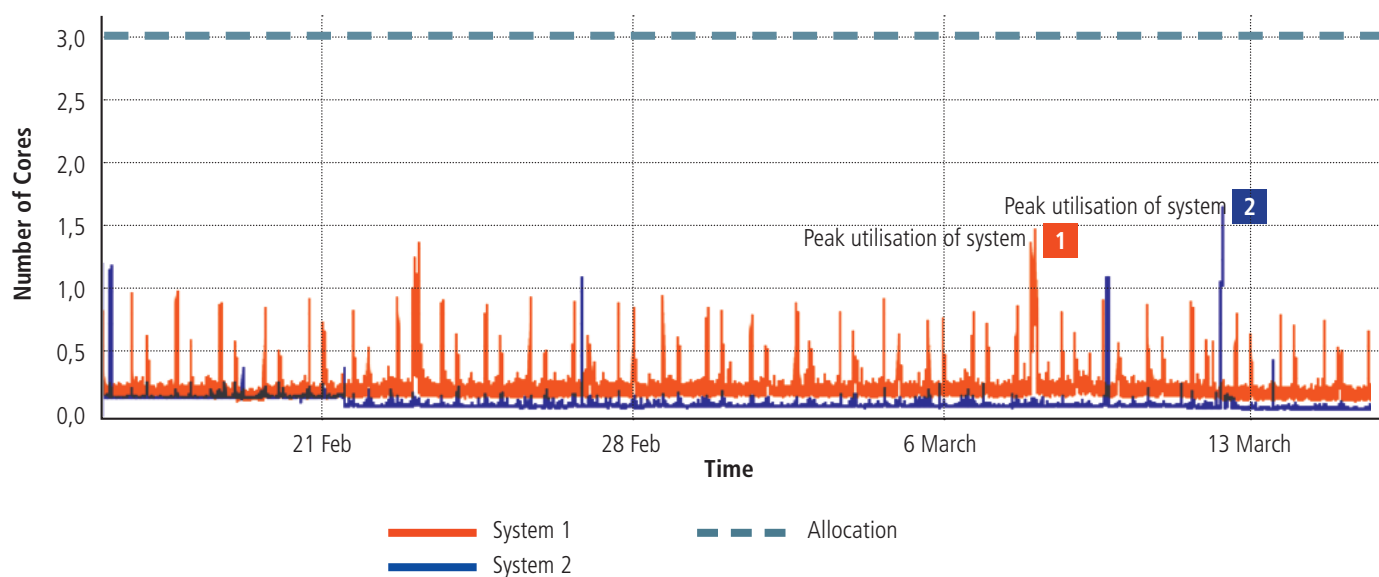


Abb. 2.3 Vergleich der CPU Nutzung für "System 1" und "System 2" im Hinblick auf eine Konsolidierung (HP Capacity Planner)

2.2.2 Power Capping

Die aktive Zuteilung von Power Budgets an Server ist auch unter der Bezeichnung Power Capping bekannt. IT-Manager können entsprechend den tatsächlichen Erfordernissen Power Caps für Server spezifizieren. Dynamic Power Capping reduziert die Spitzenlast für das System und kann damit entsprechend auch die Optimierung der Dimensionierung der Stromversorgung unterstützen.

Die konkreten Einsparungen in der Praxis hängen vom Level der Caps ab. Die Caps sollten so festgelegt werden, dass Lastspitzen gekappt werden, jedoch die Rechenleistung nicht merklich beeinträchtigt wird.

Ein optimales Capping erfordert die Bewertung der Workload- und Energieverbrauchsmuster. Bei relativ gleichförmigen Workloads können die Caps für eine durchschnittliche Serverauslastung festgelegt werden, ohne eine spürbare Auswirkung auf die Performance. Als Faustregel gilt, dass Caps nicht niedriger als in der Mitte zwischen minimaler und maximaler Leistungsaufnahme des Servers angesetzt werden sollten. Einige Management-Tools bieten auch die Möglichkeit, Caps zeitlich festzulegen, d.h. unterschiedliche Caps für verschiedene Tagesabschnitte, je nach Auslastungsmuster, Energiekosten etc.

BEST PRACTICE EMPFEHLUNGEN

Energieeffiziente Rechenzentrumsplanung und -management

- Verwenden Sie Servermanagement-Tools für die Kapazitätsplanung, für Workload- und Lastmonitoring und für spezifisches Powermanagement.
- Verwenden Sie Anwendungs- und Load Balancing, um die Nutzung der Hardware-Ressourcen zu optimieren.
- Verwenden Sie Power Capping, um den Energiebedarf des gesamten Systems auf dem gewünschten Level zu halten.
- Optimieren Sie den Grad der Ausfallssicherheit. Legen Sie das Niveau der Ausfallssicherheit dem Bedarf entsprechend fest.
- Beseitigen Sie nicht mehr genutzte Dienste und die entsprechende Hardware.

2.3 Powermanagement-Optionen für Blade Server

Blade Server werden sowohl in Rechenzentren als auch in Serverräumen eingesetzt. Der Markt für Blade Server hat sich in den vergangenen Jahren am schnellsten entwickelt und es ist dementsprechend wesentlich, dass Blade-Technologie so energieeffizient eingesetzt wird, wie möglich.

Blade Chassis (siehe Abbildung 2.4.) können 7, 14 oder mehr Server aufnehmen und beinhalten ein oder mehrere Managementmodule sowie verschiedenste Schnittstellen. Die Chassis unterstützen Server-, Datenspeicher- und Netzwerkmodule und können anwendungsspezifisch konfiguriert werden. Blade-Technologie erlaubt im Vergleich zu Standard-Servern eine Reduktion von Hardwarekomponenten wie Netzteilen, Netzwerkschnittstellen und Verkabelung, die im Chassis von den Servern gemeinsam genutzt werden.

Zentrale Vorteile von Blades sind:

- Hohe Rechendichte und reduzierter Platzbedarf
- Reduzierter Wartungsaufwand und Aufwand für Erweiterungen dank Hot-plug Ersatz von Modulen und integrierten Management Features
- Etwas höhere Energieeffizienz als bei Rack-Servern, sofern das Powermanagement und die Kühlung optimiert sind.

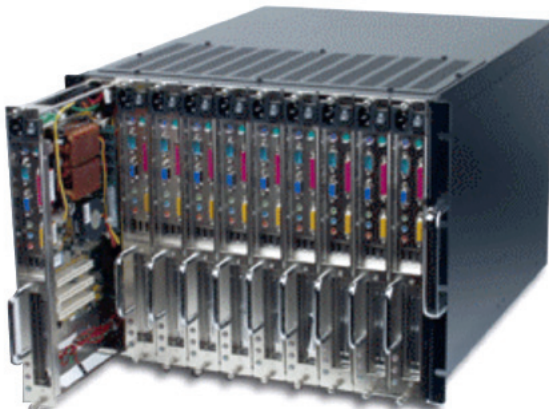


Abb. 2.4 Blade Chassis

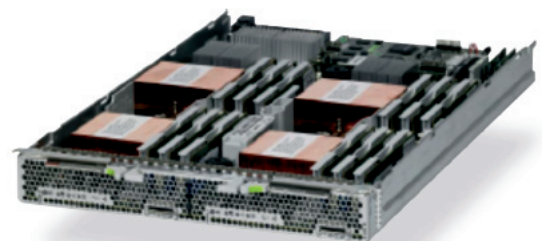


Abb. 2.5 Dual-node Blade Server

Wenn hohe Bladedichten implementiert werden sollen, steigen allerdings die Anforderungen an die Infrastruktur und Kühlung deutlich. Eine entsprechend hohe Rechendichte kann die Energie-dichte auf 10–25 kW/Rack erhöhen. In solchen Fällen ist die Standardkühlung in Rechenzentren und Serverräumen oft nicht mehr ausreichend und spezielle Kühlungskonzepte sind erforderlich. Die Energieeffizienz eines Bladesystems hängt somit auch sehr stark vom Systemdesign insgesamt ab. Dual-Node- und Multi-Node-Server basieren auf einem ähnlichen Grundkonzept wie Bladeserver. Im Multi-Node-Konzept wird jedoch eine fixe Anzahl von Servereinheiten (typischerweise 2 oder 4) in einem Serverchassis verbaut. Ähnlich wie bei Blades nutzen die Server vorhandene Netz-teile und Ventilatoren gemeinsam. Es bestehen allerdings nur geringe Erweiterungsoptionen. Multi-Node-Technologie bietet somit einen Ansatz zur Implementierung höherer Rechendichten zu vergleichsweise niedrigen Kosten, geeignet insbesondere auch für die Anforderungen in kleinen und mittleren Unternehmen. Parallel dazu werden auch High-performance Multi-Node-Server beispielsweise für Bladesysteme angeboten, die zwei Nodes in einem Blade kombinieren. Die Hauptvorteile von Dual- und Multi-node-Systemen sind:

- Niedrigere Kosten und geringerer Platzbedarf als bei Standard-Rackservern.
- Etwas niedrigerer Energieverbrauch aufgrund gemeinsam genutzter Komponenten.

BEST PRACTICE EMPFEHLUNGEN

Auswahl der Blade-Technologie basierend auf klaren Entscheidungskriterien

- Definition und Bewertung der Gründe für die Anwendung von Blade-Technologie in Rechenzentren (z.B. Raumbedarf etc).
- Abschätzung der Vorteile im Vergleich zur Standard-Rackservertechnologie
- Prüfung, ob Servervirtualisierung eine Alternative darstellen kann.
- Evaluierung der erwarteten Total Cost of Ownership (TCO) und der Energieeffizienz (basierend auf Anbieterinformationen).

2.3.1 Blade Chassis und Bladekomponenten

Größere Netzteile sind häufig effizienter als kleine, und die geringere Zahl größerer Netzteile in Blade-Systemen kann daher die Energieeffizienz im Vergleich zu Standardservern verbessern. Die Effizienz hängt insgesamt jedoch auch vom Verhältnis Leistungsbedarf zur installierten Netzteilleistung ab. Abbildung 2.6 zeigt die Effizienzkurve eines mit Effizienzklasse "Platin" gelabelten Netzteils [3] mit 2.990 W Nennleistung für ein Blade Chassis. Es zeigt sich ein Effizienzlevel von 92% bis 95% über das gesamte Lastspektrum hinweg. Effiziente

Netzteile für Blades erreichen somit ein Effizienzniveau über 90% für den Lastbereich zwischen 20% und 100%.

Für neue Produktgenerationen von Blade- und Multinode-Servern werden von einigen Herstellern mehrere Netzteile mit unterschiedlicher Nennleistung angeboten, die eine richtige Dimensionierung der Leistung entsprechend dem Bedarf der spezifischen Hardwarekonfiguration ermöglichen. Die Auswahl der Netzteile wird durch Online-Konfigurationstools der Hersteller unterstützt.

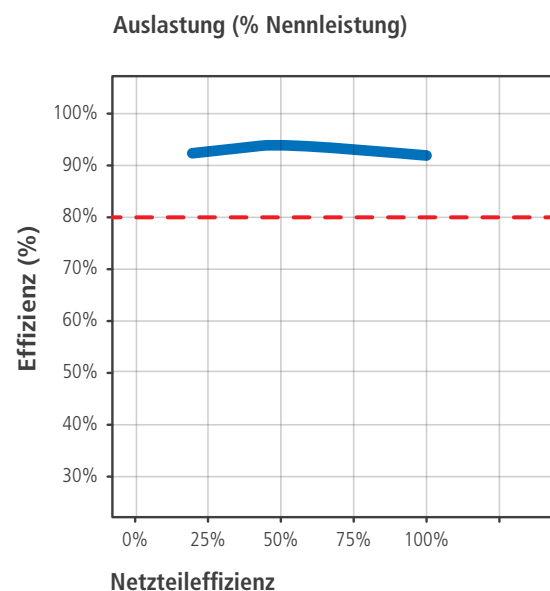
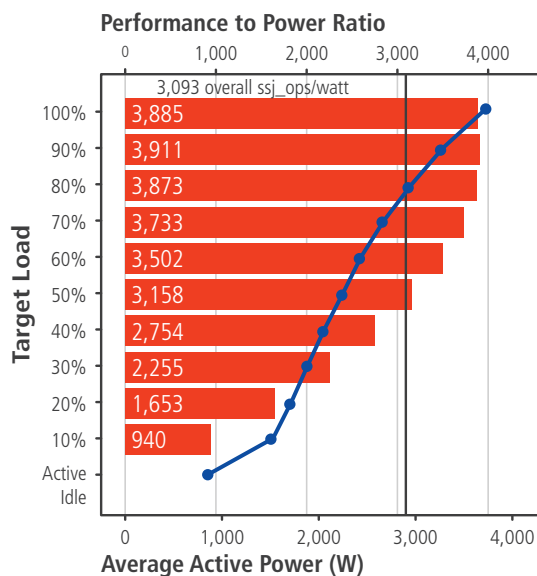


Abb. 2.6 Blade-Netzteil-effizienz [3]

Dell M610 Blade Server



R610 1U Rack Server

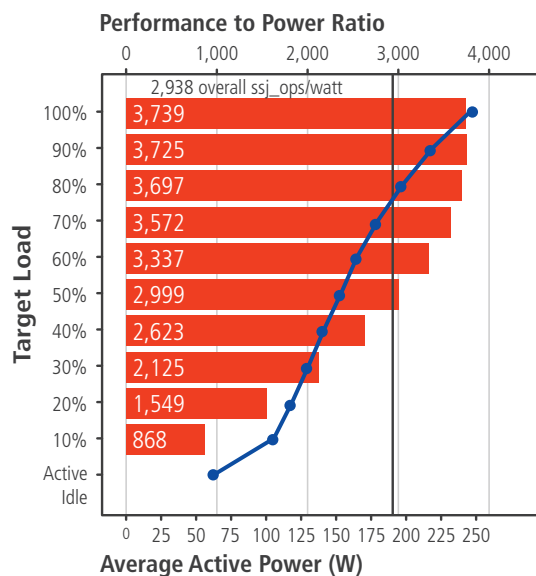


Abb. 2.7 SPECpower_ssj2008 für einen Dell M610 Blade Server und R610 1U Rack Server. Das Blade-System besteht aus 16 Blades mit identischer CPU-Konfiguration wie der Rack-Server (2 x Intel Xeon 5670, 2.93GHz). SPEC (2010, www.spec.org)

Weniger und effizientere Netzteile, effizientere Ventilatoren und umfangreichere Powermanagement-Optionen im Blade Chassis erlauben zumindest theoretisch eine höhere Energieeffizienz als bei Standardservern. Die tatsächliche Effizienz im praktischen Betrieb hängt jedoch von der entsprechenden Konfiguration und der Verwendung der Powermanagement-Optionen ab. Chassis mit nur wenigen Blades sind aufgrund der Überkapazität in den Bereichen Kühlung, Netzteile und Netzwerkkapazität nicht effizient.

Ein grober Vergleich der Energieeffizienz von Bladeservern und Standard-Rackservern kann anhand eines vollkonfigurierten Chassis durchgeführt werden. Eine solche grobe Gegenüberstellung ist anhand von publizierten Daten von Dell in Abbildung 2.7 dargestellt. Dell hat 2010 SPECpower-Ergebnisse (SPECpower_ssj2008) für Bladeserver und vergleichbare Rackserver publiziert (www.spec.org). Die SPECpower Daten zeigen eine maximale Performance von 3.885 ops/watt bei 100% Last für

das Bladesystem und 3.739 ops/watt für die Racklösung. Die Performance pro Watt oder Energieeffizienz ist daher bei voller Last im Bladesystem nur 4% besser als bei der Racklösung. Die Differenz erhöht sich auf 8% bei niedriger Rechenlast (10% Last) und auf 11% im Idle-Betrieb.

Obwohl dieser einfache Vergleich nicht überinterpretiert werden sollte (zumal SPECpower nur einen Teil der Servereffizienz bewertet), weisen die Daten darauf hin, dass auch voll konfigurierte und testoptimierte Bladesysteme nur eine leicht höhere Energieeffizienz aufweisen als Standard-Server. Der größere Effizienzunterschied bei niedrigen Rechenlasten unterstreicht das bessere Powermanagement bei Bladesystemen im unteren Lastbereich.

Bladelösungen bieten damit beispielsweise im Vergleich mit Virtualisierung nur beschränktes Potenzial zur Verbesserung der Energieeffizienz. Ähnlich wie bei Rackservern bietet sich jedoch auch bei Blades die Möglichkeit zur Virtualisie-

rung, die wiederum sehr große Effizienzverbesserungen ermöglicht.

Herausforderungen, die sich im Zusammenhang mit hohen Wärmedichten auf Rack- und Reihenebene ergeben, werden in Abschnitt 2.3.2 behandelt.

Moderne Blade-Chassis verfügen über Hardware- und Softwarekomponenten, die in Kombination mit Remote Access Controllern in den Server Blades ein Inventar des Leistungsbedarfes und entsprechendes Powermanagement für die individuellen Blades unterstützen. Managementkarten ermöglichen ein Hardware- und Leistungsbedarfsinventar für die Blades. Der Remote Access Controller übermittelt die Information zur Chassis-Management-Karte, die wiederum die Verfügbarkeit der Leistung auf Systemebene bestätigt.

Über die Managementkarte können Powermanagementregimes auf Systemebene gesetzt und der aktuelle Energieverbrauch jedes Servermoduls

überwacht werden, um zu vermeiden, dass der Gesamtverbrauch im Chassis definierte Levels überschreitet.

Die Basisfunktionen des Powermanagements im automatischen Modus sind für den Administrator normalerweise nicht sichtbar. Prioritäten für einzelne Servermodule können jedoch auch manuell definiert werden, beispielsweise durch die Festlegung, dass die Blades mit niedrigster Priorität jeweils als erste in einen Low-Power-Modus versetzt werden.

Dynamisches Power Capping kann in Blade-Chassis noch effizienter eingesetzt werden als bei Standard-Rackservern, da das dynamische Power Cap für mehrere Server im Chassis gleichzeitig spezifiziert werden kann. Power Caps können über den Onboard Administrator und Service Prozessor dynamisch angepasst werden. Blades mit niedrigerer Rechenlast erhalten niedrigere Leistungscaps. Da die Intensität und Dynamik der Rechenlast für verschiedene Blades unterschiedlich ist, treten Peaks im Leistungsbedarf zu unterschiedlichen Zeiten auf. Das Gesamt-Power-Cap für ein Chassis kann somit niedriger gesetzt werden als die Summe der Caps der einzelnen Blades. HP hat Energieeinsparungen und reduzierte TCO für ein Bladesystem berechnet, bei dem das Design der Energieversorgung mit Power Capping unterstützt worden ist. Der Leistungsbedarf und die Kosten für die Energiebereitstellung konnten um ca. 20% reduziert werden [HP2011].

BEST PRACTICE EMPFEHLUNGEN

Berücksichtigung von Energieeffizienzkriterien in der Hardwarebeschaffung

- Definition der Workloads und erwarteten Workload-Levels für die Bladesysteme
- Kosten- und Effizienzvergleich von Bladesystemen verschiedener Anbieter.
- Berücksichtigung folgender Produktinformationen von Anbietern:
 - Total Cost of Ownership (TCO)
 - Energieeffizienz (z.B. basierend auf SPECpower_ssj2008, SPEC-SERT sobald verfügbar)
 - Energieeffiziente Hardware-Komponenten, (z.B. Effizienz und Dimensionierung von Netzteilen).
 - Management Tools für Power-Management und Optimierung des Systemdesigns.

2.3.2 Bladesystem – Energieversorgung und Kühlung

Das Design von effizienten Bladesystemen wird insbesondere bei großen Systemen mit hoher Leistungsdichte gerne unterschätzt. Die wesentlichen Herausforderungen sind hier:

- Ausreichende Kühlkapazität und ein entsprechendes Kühlungsdesign für hohe Wärmedichten
- Entsprechende Kapazität der Stromversorgung (PDU-Kapazität, Verkabelung etc.)

Traditionelle Kühlungssysteme erlauben oft nur 2-3kW/Rack, d.h. 10 mal weniger als ein voll ausgebautes Bladerack. Standard-Kühlungskonzepte in Rechenzentren und Serverräumen sind daher für große, dichte Bladesysteme nicht geeignet und müssen entsprechend modifiziert werden.

BEST PRACTICE EMPFEHLUNGEN

Verwendung von Management-Tools zur Optimierung der Energieeffizienz

- Verwendung von Management-Tools sowie intelligentem Netzwerk- und Energieversorgungsequipment für das Monitoring des Energieverbrauches und der Rechenlast.
- Analyse von Optionen zum Management der Rechenlast und des Energieverbrauches innerhalb und zwischen Blade Chassis und Racks.
- Verwendung von Power Capping und Power Balancing in Blade Chassis.
- Abschätzung des tatsächlichen Kühlungs- und Energieversorgungsbedarfes mittels Kalkulationstools der Anbieter.

Tabelle 2.6 zeigt die typischen Optionen für die Implementierung von unterschiedlichen Bladedichten abhängig von den Anforderungen und Einschränkungen betreffend Infrastruktur und Kühlkapazität. Je nach Bladedichte bieten sich folgende Konzepte an [Rasmussen 2010]:

- Die Verteilung der Wärmelast auf verschiedene Racks: Die einzelnen Blade-Chassis werden auf verschiedene Racks verteilt, um die Wärmelast zu verteilen. Für diesen Ansatz muss der Anteil der Blade-Chassis im Gesamtsystem niedrig sein.
- Spezielle Zuordnung vorhandener Kühlkapazität: Überschüssige Kühlkapazität wird den Blades speziell zugeordnet. Für dieses Konzept muss der Anteil der Blade-Chassis im System ebenfalls niedrig sein.
- Installation von zusätzlicher Kühlung: Für die Blade-Racks wird zusätzliche Kühlkapazität zur Verfügung gestellt. Die Leistungsdichte kann bei diesem Konzept bis zu 10kW pro Rack

betragen. Der Ansatz erlaubt eine gute Flächennutzung und hohe Effizienz.

- Design eines eigenen High-Density-Bereiches: Eine bestimmte Fläche im Rechenzentrum wird speziell dem High-Density-Blade-System gewidmet. Die Effizienz und Flächennutzung ist hoch, die Leistungsdichte kann bis zu 25kW pro Rack betragen.
- Gestaltung eines High Density Centres: Nutzung von voll ausgebauten Blade-Racks im gesamten Rechenzentrum. Dieser Ansatz ist eher ungewöhnlich und führt meist zu hohen Kosten bei schlechter Nutzung der vorhandenen Infrastruktur.

Bei bereits in Betrieb befindlichen Rechenzentren sind die Einschränkungen für die Nutzung von Bladetechnologie bereits durch die spezifische Infrastruktur festgelegt. Beispielsweise erlaubt ein Standard-Doppelbodensystem häufig nur Leistungsdichten bis zu 5kW pro Rack. Eine saubere Spezifikation der Leistungs- und Wärmedichte

ist eine wichtige Voraussetzung für ein energie-, raum- und kosteneffizientes Systemdesign.

Ein weiterer wichtiger Punkt ist die Vermeidung von Überkapazitäten im Bereich Infrastruktur und Kühlung. In der Planung sollten räumliche und zeitliche Variabilitäten berücksichtigt werden, z.B. unterschiedliche Leistungsdichten in Blade- und Standard-Racks und Veränderungen bzw. eine Zunahme der Wärmedichte im Laufe der Zeit. Die Leistungsdichte muss pro Rack oder pro Rackreihe spezifiziert werden. Für größere Systeme bietet sich eher eine reihenbasierte Planung an, da die Stromversorgung und Kühlung normalerweise pro Rackreihe designed wird. Die Spezifikation der Leistungsdichte sollte idealerweise für die Dauer des Betriebs eines Racks oder einer Reihe beibehalten werden. Die Implementierung von neuen Servertypen mit anderer Leistungsdichte sollte daher nach Möglichkeit in einem neuen Rack oder einer neuen Reihe vorgenommen werden. Alternativ dazu gibt es jedoch auch Ansätze, die

Tab. 2.6 Konfiguration von Bladesystemen auf Rackebene und Anforderungen an die Kühlung [nach Rasmussen 2010]

No Chassis/Rack	Verteilung der Last über verschiedene Racks	Zuordnung von Kühlkapazität	Zusätzliche Kühlung	High Density Bereich	High Density Centre
1	Meistens möglich	Meistens möglich	Meistens möglich	Nicht kosteneffizient	Nicht kosteneffizient
2	Selten möglich	Meistens möglich	Meistens möglich	Nicht kosteneffizient	Nicht kosteneffizient
3	Nicht möglich	Meistens möglich	Meistens möglich, abhängig von spezifischer Lösung	Maximum für optimierte Raised floor Systeme	Nicht kosteneffizient
4	Nicht möglich	Selten möglich	abhängig von spezifischer Lösung	Hot Air Scavenging	Hot Air Scavenging, Raum Design
5	Nicht möglich	Nicht möglich	Nicht möglich	Hot Air Scavenging	Hot Air Scavenging, Raum Design
6	Nicht möglich	Nicht möglich	Nicht möglich	sehr hohe Kosten	sehr hohe Kosten

eine Variation von Leistungsdichten in Racks und Reihen erlauben:

- Einsatz von zusätzlichen Hot Plug USV Modulen
- Einsatz von Hot Swap PDUs
- Einsatz von zusätzlicher Kühlkapazität auf Rackebene (direkte Rackkühlung)

Für die Definition von Wärmedichten in Reihen ist es empfehlenswert, ein maximales Verhältnis von Maximal- zu Durchschnittsleistung von 2 festzulegen. In Racks, in denen die doppelte Durchschnittsleistung überschritten wird, sollten die IT-Lasten innerhalb der Reihe oder zwischen Reihen umverteilt werden. Insgesamt macht es trivialerweise Sinn, Racks mit höherer Wärmedichte in der Reihe zu verteilen. Leistungs- und Kühlungsmanagementsysteme können eingesetzt werden, um Regeln für die optimale Nutzung der vorhandenen Kapazitäten zu entwickeln.

Ein weiterer wesentlicher Punkt ist die geeignete Vorgehensweise bei der Erweiterung der IT-Systeme. Es ist klarerweise nicht empfehlenswert, von Anfang an maximale Kühlkapazitäten zur Verfügung zu stellen, da dies über lange Zeit eine Überversorgung und damit überhöhte Kosten und Energieverschwendung bedeuten würde. Daher ist zu empfehlen, dass zunächst nur die gebäudetechnischen Einrichtungen (Verkabelung, Verrohrung etc.) voll implementiert werden, Energieversorgung und Kühlung jedoch modular entsprechend dem Bedarf ausgebaut werden. Dieser Ansatz erlaubt einen bedarfsspezifischen Ausbau im Laufe der Zeit.

2.4 Servervirtualisierung

Servervirtualisierung bietet ein hohes Potential für Energieeinsparungen. Diese Technologie erlaubt die Konsolidierung von Lasten auf weniger physischer Hardware und reduziert dadurch den Energieverbrauch und den Kühlaufwand stark. Virtualisierung bietet große Vorteile beim tatsächlichen Design von IT-Systemen in Serverräumen und Rechenzentren, beispielsweise:

- Reduktion von Anforderungen an Hardware und Platzbedarf durch virtuelle Maschinen (VMs), die geschützt auf gemeinsam genutzter Hardware laufen und die Auslastung der Server von häufig nur 5–15% auf 60–80% erhöhen können.
- Optimierung von Testumgebungen und Entwicklung durch schnellen Roll-out von vor-konfigurierten Systemen, wobei standardisierte Entwicklungskonfigurationen verwendet werden können.
- Reduktion von Kosten und Komplexität des Betriebs von geschäftskritischen Anwendungen durch Zusammenfassen von ganzen Systemen in einzelne Files, die leicht repliziert und auf die Zielsysteme rückgespeichert werden können.

Etablierte Virtualisierungsplattformen, wie VMware, Microsoft Hyper-V und Citrix XEN bieten viele Eigenschaften wie Hochverfügbarkeit, Ausfallsicherheit, verteilte Ressourcenzuteilung, Lastverteilung, automatisierte Backup-Funktionen, verteiltes Leistungsmanagement, Server-, Storage- und Network Vmotion etc.

Es gibt die folgenden grundsätzlichen Optionen für Servervirtualisierung:

- Physische Partitionierung
- Betriebssystembasierte Virtualisierung
- Applikationsbasierte Virtualisierung: z.B. Microsoft Terminalserver, Citrix XenApp
- Hypervisor-basierte Virtualisierung:
 - VMware ESX
 - Citrix /Open-Source: XENServer 5
 - Microsoft Hyper-V

Ausgehend von der aktuellen Marktsituation, die von wenigen Hypervisor-basierten Produkten bestimmt wird, konzentrieren sich die folgenden Kapitel auf VMware ESX, Microsoft Hyper-V und Citrix XEN Server.

Diese drei marktführenden Virtualisierungsplattformen unterstützen die meisten verbreiteten Gastbetriebssysteme. Sie enthalten Managementkonsolen für die Verwaltung kleinerer Serverumgebungen, sowie für die Administration von Rechenzentren.

VMware war 2001 das erste Produkt am Markt. Seine Architektur beruht auf für Virtualisierung vorbereiteten Betriebssystemen und Prozessoren wie Intel VT und AMD-V. VMware ESX/VSphere4 bietet umfangreiche Administrationswerkzeuge, wie VMotion (Verschiebung von virtuellen Maschinen zwischen Servern), Storage VMotion, Desktop- und Netzwerkvirtualisierung, Storage Overprovisioning, Virtual Security Technologie und bietet eine komplette Virtualisierungsplattform in den Bereichen Desktop bis zu Cloud Computing.

Microsoft Hyper-V Server kommt mit dem Windows Hypervisor, dem Windows Treibermodell und Virtualisierungskomponenten. Es weist einen kleinen Footprint und minimalen Overhead auf und fügt sich in existierende IT-Umgebungen nahtlos ein, indem es bereits existierende Technologien für Patchverwaltung, Roll-out von VMs, Management, Support Tools und Prozesse nützt.

Einige der Haupteigenschaften von Microsoft Hyper-V Server 2008 R2 sind Live Migration, Unterstützung von gemeinsam in Clustern genutzten Volumes und erweiterter Support für Prozessor und Speicher von Hostsystemen.

Live Migration ist in das Hostbetriebssystem Windows Server 2008 R2 integriert und kann laufende virtuelle Maschinen ohne Unterbrechung zwischen Servern verschieben.

Abhängig von den Benutzeranforderungen kann Citrix XENServer einen kostengünstigen Weg für die Implementierung von Virtualisierung darstellen, da Basisfunktionalitäten wie der Hypervisor, die verteilte Managementarchitektur, sowie zugehörige Tools für Management und Migration von virtuellen Maschinen gratis angeboten werden. Erweiterte Management- und Automatisierungstools, wie Virtual Provisioning Services, Distributed Virtual Switching, XENMotion Live Migration, erweiterte Memory Snapshots, Tools für Leistungsmessung und Dynamic Workload Balancing machen XENServer mit den beiden anderen Produkten vergleichbar. Diese Produkte sind allerdings Bestandteil der kostenpflichtigen kommerziellen Editionen.

BMC Software, Eucalyptus Systems, HP, IBM, Intel, Red Hat, Inc. and SUSE haben die Gründung der Open Virtualization Alliance bekanntgegeben. Dieses Konsortium befasst sich mit der Adaptierung von Open Source Virtualisierungstechnologien, wie Kernel-based Virtual Machine (KVM). Es erweitert die existierenden Open Source Communities durch Förderung der Entwicklung des KVM Hypervisors und zugehöriger Management-Funktionalitäten, die eine sich ständig weiterentwickelnde Technologie für die Virtualisierung von Linux und Windows Applikationen darstellen.

Das Konsortium möchte die Verbreitung von Drittanbieterlösungen rund um KVM fördern und wird technische Unterstützung und Best Practise Beispiele zur Verfügung stellen.

2.4.1 Energieeinsparungspotential von Virtualisierung

Virtualisierung ist eine der wirksamsten Technologien, um den Energieverbrauch in Rechenzentren und Serverräumen zu reduzieren. Die Konsolidierung von Serverhardware durch Zusammenfassung der Lasten auf einer kleineren Anzahl physischer Server ermöglicht oft Energieeinsparungen von 40 bis 80% und mehr, abhängig von der jeweiligen Einsatzart. Die derzeit verfügbare Technologie ermöglicht Konsolidierungsfaktoren von zumindest 10–20, je nach der Art der jeweiligen Systeme und Anforderungen.

Abb. 2.8 zeigt ein Beispiel einer Serverkonsolidierung durch Virtualisierung im Deutschen Umweltministerium. Die getroffenen Maßnahmen

ermöglichten Energieeinsparungen von ca. 68%. Dies wurde durch eine Reduktion der Hardware auf 2 physische Server unter VMWare ESX 4 [4] erreicht.

In einem weiteren Beispiel für Virtualisierung [5] wurden Einsparpotenziale von ca. 90% ermittelt, wenn alle Maßnahmen auf Hard- und Softwareebene berücksichtigt werden.

Dies zeigt, dass Konsolidierung durch Virtualisierung eine der Hauptmöglichkeiten darstellt, um die Energieeffizienz in Rechenzentren signifikant zu erhöhen.

Ähnlich wie bei anderen Effizienzmaßnahmen, kann das volle Einsparungspotenzial allerdings nur dann ausgeschöpft werden, wenn gleichzeitig die gesamte Infrastruktur (Stromversorgung und Kühlung) mit betrachtet und optimiert wird.

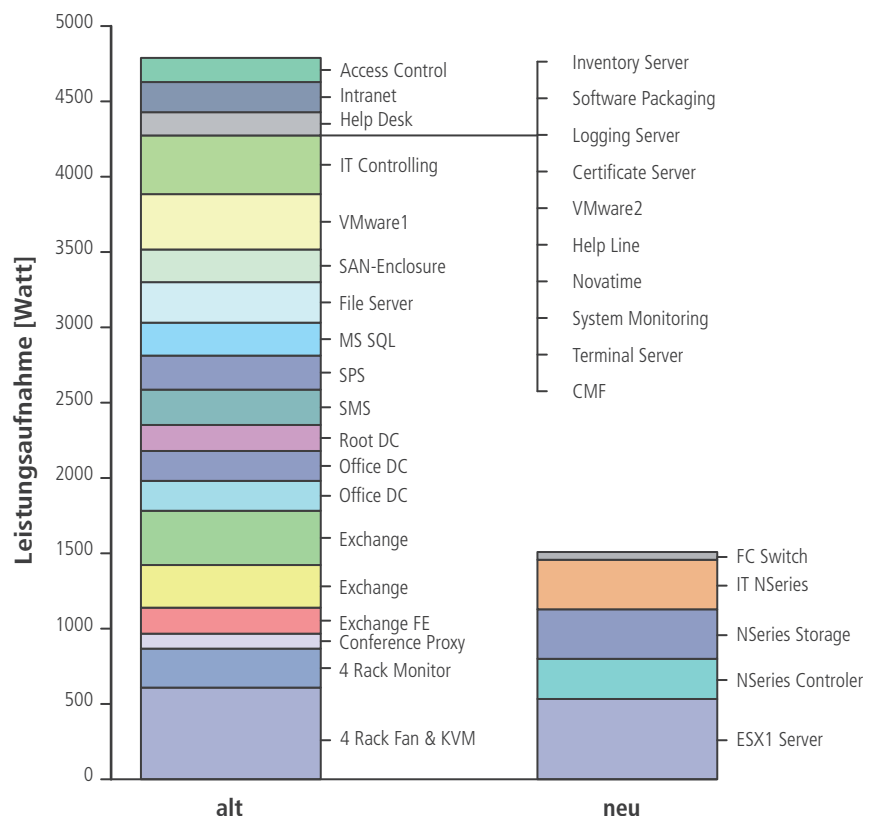


Abb 2.8 Reduktion des Energieverbrauchs durch Virtualisierung in einem Fallbeispiel [4]

2.4.2 Anforderungen und Werkzeuge für die Planung von Virtualisierungsprojekten

Die Einführung von Virtualisierung in Rechenzentren sollte auf einer Virtualisierungsstrategie basieren, die die Identifizierung und Evaluierung von geeigneten Serverkandidaten ermöglicht. Für eine solche Evaluierung werden Daten in den Bereichen Leistung, Systemauslastung, Lebensdauererwartung, Geschäftsdaten und Art der Applikation gesammelt. Wenn die Kandidaten für Virtualisierung identifiziert sind, werden die Art der Applikation und die Maschinenauslastung analysiert. Weiters wird eine Leistungsevaluierung durchgeführt, die unter anderem die folgenden Erfordernisse als Basis für die Hardwareauswahl berücksichtigt:

- CPU Leistung
- Erforderlicher Speicher
- Disk I/O Intensität
- Netzwerkerfordernisse
- Betriebssystemkonfiguration

Verschiedene Applikationen können typischerweise auf einem einzelnen physischen Server konsolidiert werden, der gegen Hardwareausfälle und Spannungsunterbrechungen abgesichert ist und die Möglichkeit von Load Balancing bietet. Um diese Ziele zu erreichen, können Server mit redundanten Netzteilen, gespiegelten Festplatten und redundanten Netzwerkkarten ausgestattet werden. Ein fehlertolerantes Storage Area Network (SAN) kann als zentrale Speicherlösung genutzt werden. Load Balancing kann durch Verschieben von virtuellen Maschinen zwischen physischen Servern erreicht werden.

Abhängig von der Art der Lasten können häufig Konsolidierungsfaktoren von 1:10 bis zu 1:20 erreicht werden.

Im Bereich der Speichererfordernisse bieten viele Virtualisierungsumgebungen das Feature „Memory Overprovisioning“. Durch Nutzung die-

ses Features kann die Summe des für alle virtuellen Maschinen angelegten Speichers den physischen Speicher um den Faktor 2 bis 3 überschreiten.

Virtualisierung wird selten allein zum Zweck der Energieeinsparung eingesetzt. Obwohl meist hohe Energieeinsparungen erzielt werden können, erfordern erfolgreiche Virtualisierungsprojekte üblicherweise eine ausführliche Planung, die auch Berechnungen von Lebenszykluskosten sowie Rentabilitätsberechnungen einschließt.

Die Lebenszykluskosten für den Roll-out von neuen virtuellen Servern werden durch Summierung aller relevanten Kostenfaktoren berechnet. Um zeitbezogene Kosten zu ermitteln, können Kurz- und Langzeitrentabilitätsrechnungen durchgeführt werden.

Der Schlüssel zu erfolgreichen Rentabilitätsberechnungen ist das Verständnis der Virtualisierungskosten. Die augenfälligen Kosten für Virtualisierungsprojekte bestehen aus Kosten für Hardware, Lizenzkosten und Arbeitskosten.

Die Einführung von Virtualisierung erfordert oft den Kauf von neuen, stärkeren Servern, Erweiterung des Speichers, Netzwerkupgrades und erhöhte Sicherheitsmaßnahmen. Das erforderliche Training der Mitarbeiter stellt oft einen zusätzlichen Kostenfaktor dar. Alle diese Aspekte müssen in die Rentabilitätsberechnung mit einfließen.

Für die Planung von Virtualisierung sowie für Rentabilitätsberechnungen und Lebenszykluskostenberechnungen gibt es auf dem Markt verschiedene Softwarelösungen. Das von Microsoft angebotene Assessment and Planning (MAP) Toolkit unterstützt beispielsweise die Planung der Migration, sowie Rentabilitäts- und Lebenszykluskostenberechnungen. Er stellt ein Werkzeug für Inventarisierung, Beurteilung und Berichtswesen dar, welches IT Umgebungen ohne die Hilfe von Software Agenten auf Eignung für Migration und Virtualisierung analysiert. Es werden in sogenann-

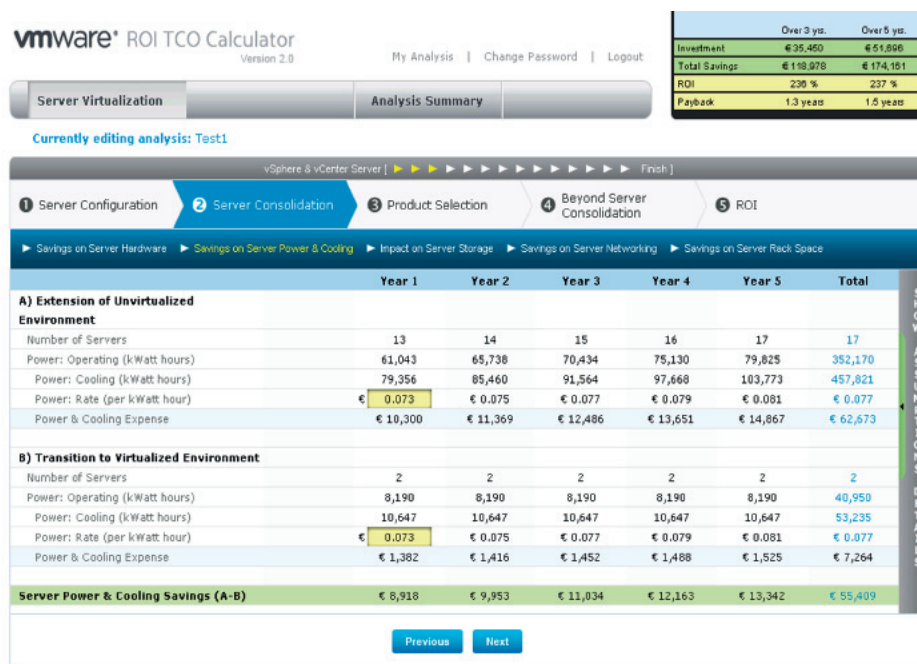


Abb. 2.9 Beispiel ROI/TCO Calculator von VMWare [6]

ten *inventory and readiness assessment reports* Empfehlungen für die Migration und Virtualisierung von Windows Vista und Windows Server 2008 Betriebssystemen gegeben. Im Speziellen werden Empfehlungen generiert, wie physische Server in Microsoft Hyper-V Umgebungen konsolidiert werden können. Zusätzlich unterstützt das Microsoft Integrated Virtualization ROI Tool die Berechnung von möglichen Energiekosteneinsparungen noch vor der tatsächlichen Migration. Es können bestehende Produktions- und Entwicklungsserver, sowie Desktop-Umgebungen auf ihre Eignung untersucht werden, indem potentielle Einsparungen, Service-Level-Erweiterungen, Investitionen und Rentabilitätswerte quantifiziert werden.

Die von VMWare als Online Tool angebotene Methode zur Rentabilitätsberechnung und Berechnung von Lebenszykluskosten, erlaubt es, Kosten, erforderliche Investitionen und Geschäftsvorteile zu vergleichen. Sie basiert auf Standard-Finanzinstrumenten, sowie von VMWare gesammelten Feld- und Kundendaten und Benutzermetriken. Auf der Grundlage von benutzerspezifischen Daten werden zentrale Finanzkennzahlen wie Einsparungen, Investitionen und Amortisationszeiträume berechnet.

2.4.3 Powermanagement in virtualisierten Umgebungen – Migration virtueller Server

Moderne Virtualisierungsumgebungen unterstützen die Migration von virtuellen Maschinen, wie auch das vorübergehende Herunterfahren von physischen Servern zum Zweck der Energieeinsparung. Beispielsweise bietet VMWare Vsphere 4 das Feature Distributed Power Management (DPM). DPM verfolgt die Auslastung der laufenden virtuellen Maschinen in einem Cluster. Falls überschüssige Kapazität vorhanden ist, wird die Migration von virtuellen Maschinen zwischen Hosts ausgelöst und werden nicht mehr benötigte Hosts in den Standby-Modus gebracht. Bei wieder zunehmender Auslastung werden diese Hosts aus dem Standby-Modus wieder hochgefahren.

DPM kann sowohl in manuellem, als auch in vollautomatischem Modus betrieben werden. Im automatischen Modus werden virtuelle Maschinen automatisch migriert und Hosts in den Standby-Modus gebracht und wieder hochgefahren. Die Einstellungen für die Auslösung dieser Vorgänge können pro Host definiert werden und der ganze Prozess kann zeitbasierend ein- und ausgeschaltet werden.

Ziel von DPM ist es, die Auslastung der ESX Hosts in einem Cluster innerhalb eines bestimmten Zielbereichs zu halten. Damit dieser Vorgang reibungslos funktioniert, müssen die folgenden Voraussetzungen erfüllt sein:

- Genaue Erfassung der Nachfrage nach Ressourcen durch die Anwendungen. Ein Überschätzen der Nachfrage reduziert die Energieeinsparungen. Ein Unterschätzen kann zu schlechter Performance und Verletzung von Service Level Agreements führen.
- Zu häufiges Abschalten und Hochfahren von Servern ist zu vermeiden, selbst wenn die Last stark variiert.
- Schnelle Reaktion auf plötzliches Anwachsen der Last, um einen Einbruch der Leistung zu vermeiden.
- Auswahl der Hosts, die abgeschaltet und gestartet werden sollen. Das Herunterfahren eines größeren Hosts kann zum Überschreiten der Zielauslastung eines oder mehrerer kleinerer Hosts führen.
- Genau geregelte Verteilung der virtuellen Maschinen, nachdem Hosts abgeschaltet oder gestartet wurden, durch nahtlos wirksames DRS.

EMPFEHLUNGEN FÜR BEST PRACTICE

Bemessung und Auswahl von Virtualisierungslösungen:

- Entwickeln Sie eine Virtualisierungsstrategie und bewerten Sie die Server, um gute Kandidaten für die Virtualisierung auszusuchen.
- Bewerten Sie die Anforderungen hinsichtlich CPU Leistung, Speicher-, Diskintensität, Netzwerkerfordernisse und Betriebssystemkonfiguration.
- Bedenken Sie das entsprechende Virtualisierungsverhältnis und die Durchmischung der Lasten (1:6 bis 1:20 abhängig von der Lastcharakteristik).
- Überprüfen Sie Produkte verschiedener Anbieter in Bezug auf die für Ihren speziellen Zweck erforderlichen Eigenschaften; Vergleichen Sie Lizenzpolitik, Power Management und Preis. Die verschiedenen Hauptprodukte auf dem Markt haben verschiedene Vorteile, abhängig von den spezifischen Anforderungserfordernissen.
- Machen Sie TCO und ROI Berechnungen, um die Vorteile geringerer Kosten für die Stromversorgung und Kühlung herausfinden zu können. Die Modelle der Hersteller sollten in Bezug auf die Bedürfnisse der jeweiligen Organisation angepasst werden.
- Überlegen Sie die Anwendung von Powermanagementoptionen, die das Migrieren von virtuellen Maschinen und das vorübergehende Herunterfahren von Servern erlauben.
- Überlegen Sie, die Anforderungen für Kühlung und Stromversorgung zu ändern (reduzierte und dynamisch änderbar Nachfrage nach Energie und Kühlung) und prüfen Sie Optionen für Designänderungen beim Kühlsystem.

Normalerweise wird DPM basierend auf einer typischen Lastverteilung während eines Arbeitstages oder einer Arbeitswoche verwendet. Beispielsweise werden Services wie Email, Fax, Intranet und Datenbankabfragen während der typischen Arbeitsstunden von 9h bis 17h intensiv genutzt. Zu anderen Zeiten kann die Auslastung stark absinken, was zu schlecht ausgelasteten Hosts führt. Hauptsächlich werden während dieser Zeiten Backup- und Archivierungsaufgaben abgearbeitet, sowie Anfragen aus Übersee beantwortet. Durch Konsolidierung von virtuellen Maschinen und Herunterfahren von Hosts während dieser Zeiten kann der Energieverbrauch reduziert werden.

Die folgenden Kriterien sind zu beachten, um die DPM-Aktivität zu justieren:

- Erhöhen des Zielwertes für das Verhältnis zwischen nachgefragter und vorhandener Leistung: Um mehr Energie einzusparen und die Auslastung der Hosts zu erhöhen (durch Konsolidierung von mehr virtuellen Maschinen auf weniger Zielhosts), kann der Default-Wert von normalerweise 63% auf 70% erhöht werden.

- alle Hosts können mit VMWare DPM vor Beginn der Arbeitszeiten gestartet und nach Ende der Zeit mit höchster Leistungsnachfrage selektiv heruntergefahren werden. Dies stellt einen mehr proaktiven Ansatz dar, um Leistungseinbrüche durch Wartezeiten beim Hochfahren von Hosts durch VMWare DPM als Reaktion auf plötzliche Lastspitzen zu vermeiden.

Die Auslastung jedes ESX Hosts wird als Verhältnis von Nachfrage zu Kapazität für jede Ressource (CPU und Speicher) berechnet, wobei Nachfrage als Summe des Ressourcenbedarfs aller laufenden virtuellen Maschinen und Kapazität als Summe aller Ressourcen, die auf dem jeweiligen Host verfügbar sind, definiert werden. Somit wird das Powermanagement von Hosts unter Berücksichtigung von CPU- und Speicherauslastung des Zielhosts im Verhältnis zu einem vordefinierten Auslastungsbereich ausgeführt. Für jeden Host, der für eine Abschaltungsempfehlung herangezogen wird, vergleicht DPM die erwarteten Kosten und allfälligen Risiken mit einer konservativen Schätzung der möglichen Energieeinsparungen.

2.4.4 Kühlung und Infrastruktur für virtualisierte Systeme

Obwohl Virtualisierung den Gesamtenergieverbrauch signifikant reduziert, kann es speziell in größeren Systemen zu einer erhöhten Energiedichte in den Racks kommen.

Zusätzlich führt Powermanagement durch Migration von virtuellen Maschinen zu variablen lokalen Veränderungen von Leistungs- und Wärmedichten und zu einer lokal erhöhten Nachfrage nach Energie und Kühlung. Daher müssen entsprechende Kühlkonzepte verwendet werden, um den variablen Anforderungen gerecht zu werden und Bereiche mit Überhitzung entsprechend zu vermeiden.

Falls die Energieversorgungs- und Kühlkapazität nicht an den niedrigeren Bedarf angepasst wird, kann sich der PUE-Wert nach Einführung von Virtualisierung entsprechend verschlechtern. Eine zu niedrige Kühllast in Relation zum Kühlsystem kann negative Effekte hervorrufen. Die Anpassung und richtige Dimensionierung von Energieversorgung und Kühlung ist daher mit dafür entscheidend, dass die Energiesparpotenziale auch weitgehend ausgeschöpft werden können. Die Reduktion von Verlusten durch die folgenden Maßnahmen ist somit wichtig:

- Anpassung von Energieversorgung und Kühlkapazität entsprechend der Last

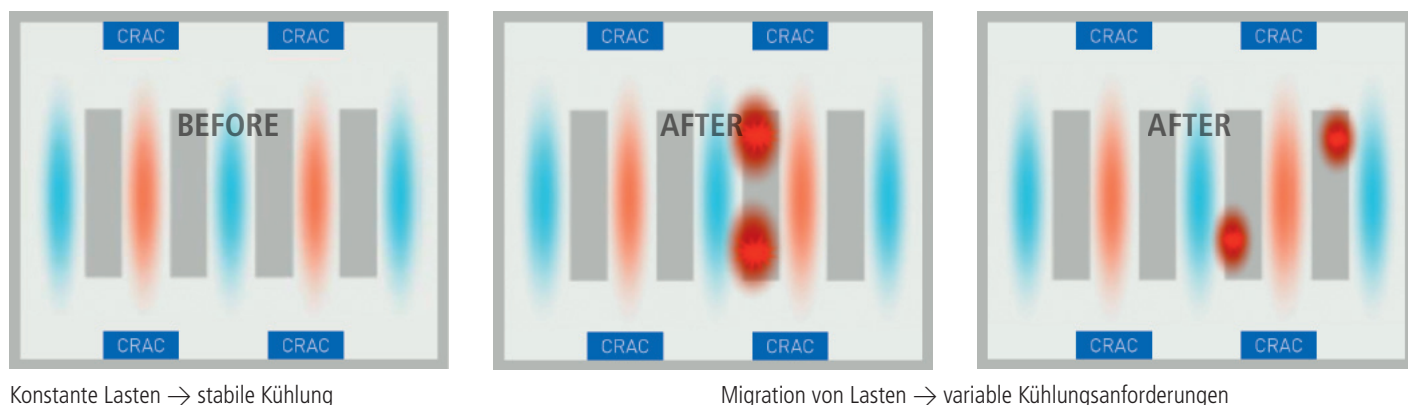


Abb. 2.10 Wärmeverteilung vor und nach Migration von virtuellen Servern [5]

- VFD-Ventilatoren und Inverter Pumpen, die über den Kühlbedarf geregelt werden.
- Nutzung von Equipment mit hoher Energieeffizienz
- Kühlungsarchitektur mit Nutzung kürzerer Luftwege (z.B. reihenbasiert)
- Kapazitätsmanagementsystem, um die Kapazität der Nachfrage anpassen zu können.
- Abdeckungen zur Reduktion der internen Luftvermischung

In einer konventionellen Umgebung mit Zwischenböden kann die raumbasierte Kühlung zur entsprechenden Kühlung von Hotspots durch Neuordnung der Kühlauslässe angepasst werden. Die dynamische Migration von virtuellen Maschinen erfordert entsprechend anpassungsfähige Kühllösungen.

Eine Möglichkeit dazu bietet die Positionierung von Kühleinheiten innerhalb der Schrankreihen. Diese müssen mit entsprechender Sensorik ausgestattet werden, um auf Temperaturunterschiede reagieren zu können. Geringe Abstände zwischen Kühleinheiten und Servern erlauben kurze Luftwege.

Die dynamische Veränderung von Lasten in virtualisierten Umgebungen ist ein wesentlicher Grund für die Umstellung auf Schrankreihen- bzw. Schrank-basierte Kühlung.

Genauere Informationen zum aktuellen Energie- und Kühlungsbedarf sind für eine schnelle Reaktion auf zeitliche Änderungen der Lastprofile entscheidend.

Ein entsprechendes Kapazitätsmanagement ermöglicht die Echtzeiterfassung und Analyse von Energie-, Kühlungs- und Platzkapazitäten, sowie deren effiziente Nutzung im gesamten Rechenzentrum. Weiters können Bereiche mit hoher oder gefährlich geringer Kapazität lokalisiert werden. Kapazitätsmanagementsysteme müssen insbesondere die folgenden Situationen beherrschen:

- Änderung der lokalen Lastverteilung: Virtualisierung kann Hotspots erzeugen, beispielsweise bei Migrationen von VMS
- Dynamische Systemänderungen: Die Aufrechterhaltung der Systemstabilität kann zur Herausforderung werden, wenn Änderungen an der Systemkonfiguration von verschiedenen Seiten ohne zentrale Koordination erfolgen.
- Virtualisierung macht Abhängigkeiten und Sekundäreffekte in Bezug auf Energie-, Kühlungs- und Platzanforderungen komplexer.
- Vermeidung von Überdimensionierung: Nach Virtualisierung reduziert sich der Energie- und Kühlungsbedarf zunächst und steigt mit zunehmender Zahl von neuen virtuellen Maschinen wieder an. Eine bedarfsgerechte Kapazität kann durch Verwendung von skalierbaren Energieversorgungs- und Kühlsystemen erreicht werden.

BEST PRACTICE EMPFEHLUNGEN

Energieeffiziente Verwaltung virtualisierter Systeme:

- Definieren Sie strikte Vorgaben für die Implementierung und Verwaltung virtualisierter Server. Vermeiden Sie unkontrolliertes Serverwachstum
- Verwenden Sie Werkzeuge für die Migration virtueller Maschinen, um Hardware in Zeiten niedriger Last herunterzufahren. Verwenden Sie anfangs die automatischen Powermanagement-Einstellungen und mit zunehmender Erfahrung angepasste Einstellungen, entsprechend den typischen Betriebserfordernissen.
- Passen Sie die Kühlkapazität an den Bedarf an und verwenden Sie Equipment, welches dynamische Kühlung erlaubt. Bedenken Sie lokale Dynamik im Kühlungsbedarf.
- Passen Sie IT-Prozesse und Workflows in Bezug auf den Roll-Out von virtuellen Maschinen, Backup- und Wiederherstellungsprozessen, Patch Administration und Verfügbarkeitsüberlegungen an.

Literaturhinweise

HP (2011): HPPowercappingandHPDynamicpower capping for ProLiant servers. Hewlett Packard Development company.

SPEC (2011): Server Efficiency Rating Tool (SERT) TM Design Document. 3rd draft. Standard Performance Evaluation Cooperation

Rasmussen, N. (2010): Strategies for deploying blade servers in existing data centres. White paper 125. APC Schneider Electric

80 PLUS (2011): 80 PLUS power supplies.

www.plugloadsolutions.com

Schäppi B. et al (2009) Energy and cost savings by energy efficient servers. IEE E-Server best practice cases. Brochure 2009

IBM (2011) Server Management suite, Module Active Energy Manager

www-03.ibm.com/systems/software/director/aeem/

HP (2011) Server management suite «Systems Insight Manager» www.hp.com

VMware DPM: Information Guide: VMware Distributed Power Management Concepts and Use.

www.vmware.com

VMware TCO: VMware ROI TCO Calculator, Overview and Analysis.

<http://roitco.vmware.com/vmw/>

Zitierte Literatur

[1] **EPA (2010):** Energy Star ENERGY STAR® Program Requirements for Computer Servers (vers 1.1)

[2] **SPEC (2010):** SPEC power and performance. Benchmark methodology 2.0. Standard Performance Evaluation Cooperation

[3] **80 PLUS (2011):** 80 PLUS power supplies.

www.plugloadsolutions.com

[4] **Schäppi B. et al (2009):** Energy and cost savings by energy efficient servers. IEE E-Server best practice cases. Brochure 2009

[5] **BITKOM (2010):** Bitkom/Beschaffungsamt des Bundesministeriums des Innern, Leitfaden Produktneutrale Leistungsbeschreibung x86-Server, 2010

[5] **Comtec Power:** Overcoming the Challenges of Server Virtualization. www.comtec.com

[6] **VMware TCO:** VMware ROI TCO Calculator, Overview and Analysis.

<http://roitco.vmware.com/vmw/>

[7] **The Green Grid (2010):** White paper Nr. 33 "A roadmap for the adoption of power-related features in servers", Pflueger, J., et al., The Green Grid, 2010

3 Equipment für Datenspeicherung

Marcos Dias de Asuncao, Laurent Lefevre, INRIA

Information ist ein Kernstück jeglicher Geschäftstätigkeit, aber die für heutige Geschäfte erforderlichen Informationen zu speichern und längerfristig verfügbar zu machen, wird zunehmend zur Herausforderung. Man rechnet damit, dass der Bedarf an Kapazität für Datenspeicherung in Organisationen zwischen 2010 und 2020 um den Faktor 44 anwachsen wird [1]. Dennoch sind Strategien für eine Verbesserung der Datenspeichereffizienz bisher eher ein Randthema. Der kontinuierliche Preisverfall pro MB Speicher führte zu einem Szenario, wo es einfacher und billiger ist, Speicherkapazität zu erweitern, als die Speichereffizienz zu verbessern. Seit jedoch der Strombedarf und die Kühlung für Hardware zur Datenspeicherung immer mehr zum Thema werden, hat sich das Bewusstsein für effiziente Lösungen erhöht. Studien zeigen, dass große Unternehmen derzeit Schwierigkeiten haben, ausreichend Strom- und Kühlkapazität zur Verfügung zu stellen, während mittlere Unternehmen damit kämpfen, ausreichend Platz für ihre Speichersysteme zu finden [2]. Da die Datenspeicherung mittlerweile einen signifikanten Teil des Energieverbrauchs von Rechenzentren verursacht, ist es wesentlich, dass Speichersysteme möglichst effizient gestaltet werden.

In diesem Kapitel werden einige Technologien zur Unterstützung der Energieeffizienz von Datenspeicherlösungen behandelt. Technologien und Maßnahmen für energieeffiziente Datenspeicherung bieten sich sowohl auf der Ebene der einzelnen Speicher-Medien und Devices, als auch auf der Ebene der Gesamtlösungen. Im folgenden werden praxisorientierte Optionen auf beiden Ebenen vorgestellt.

3.1 Speichergeräte und Speichermedien

3.1.1 Bandsysteme

Bänder werden oft als eines der kosteneffizientesten Medien für eine langfristige Datenspeicherung genannt. Untersuchungen [3][4] zeigen, dass:

- für langfristige Speicherszenarien wie Backup und Archivierung in mittelgroßen Rechenzentren Festplattenlaufwerke durchschnittlich 23 Mal so teuer kommen wie Bandlösungen und 290 Mal mehr Kosten für Energie und Kühlung verursachen als Bänder.
- Datenkonsolidierung mittels Bandarchivsystemen die Betriebskosten für Speicherzentren beträchtlich senken kann. Tape Libraries mit großer Speicherkapazität können durch die Konsolidierung von Backup-Operationen ganze Dateninseln ersetzen und so die Infrastrukturkosten senken und die Energieeffizienz erhöhen.

Mit einer Lebensdauer von 30 Jahren und großen Speicherkapazitäten stellen Bänder daher eine attraktive Lösung für Rechenzentren mit hohem langfristigem Backup- und Archivierungsbedarf dar. Für eine Umgebung mit mehreren technologischen Ebenen der Datenspeicherung bilden Bandsysteme nach wie vor die energieeffizienteste Lösung, wenn man die langfristige Archivierungsdauer und die niedrige Abfragefrequenz in Betracht zieht. Es gibt auch Plattenarchiv-Lösungen, die versuchen, den hohen Energieverbrauch von Plattenlaufwerken durch Techniken wie Reduktion der Plattengeschwindigkeit und Plattenabschaltung zu minimieren (siehe nächste Kapitel).

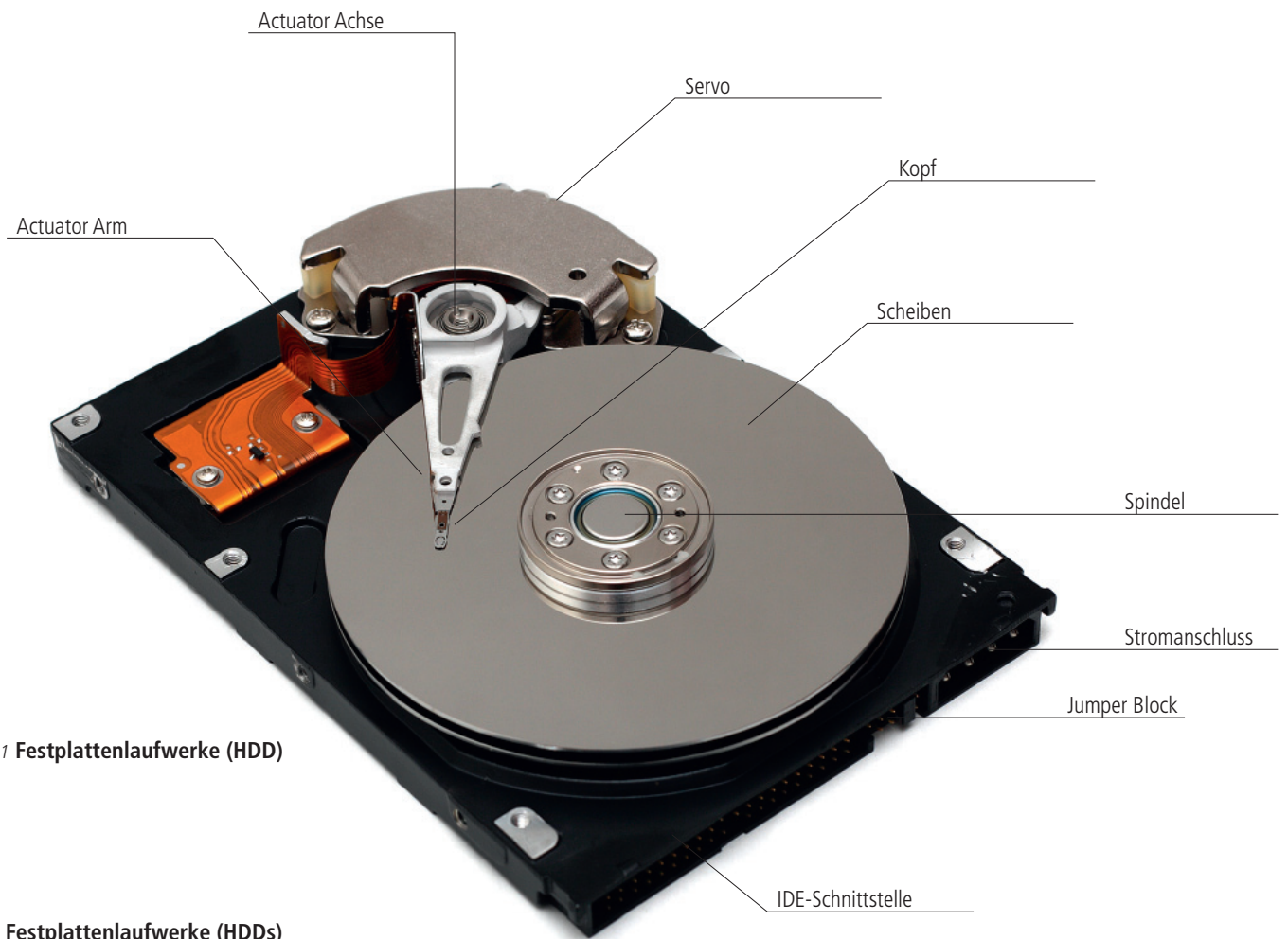


Abb. 3.1 Festplattenlaufwerke (HDD)

3.1.2 Festplattenlaufwerke (HDDs)

HDDs waren lange Zeit das bevorzugte Medium für permanente Datenspeicherung, mit raschen Schreib- und Zugriffszeiten. Bewegliche Teile wie Motor und Plattenarme verursachen bei dieser Technologie den meisten Energieverbrauch (s. Abb. 3.1). Um den Datendurchsatz von HDDs zu beschleunigen, erhöhen Hersteller die Rotationsgeschwindigkeit, damit zugleich aber auch den Energieverbrauch. Platten mit einer Umdrehungsgeschwindigkeit von 15K RPMs sind für aktuelle HDDs mit hoher Durchsatzgeschwindigkeit üblich.

Zur Erhöhung der Energieeffizienz von HDDs werden verschiedene Technologien eingesetzt,

wie beispielsweise die Speicherung von Daten in bestimmten Plattenbereichen, um den mechanischen Aufwand beim Zugriff auf die Daten zu verringern, Steuerung/Kontrolle der Rotationsgeschwindigkeit der Platten und Reduktion des Stromverbrauchs während des Leerlauf bzw. Idle-Betriebs. Eine geläufige Technik, auch als Disk-Spin-Down bezeichnet, besteht darin, Platten nach einer gewissen Zeit der Inaktivität zu verlangsamen und die Leseköpfe in der sicheren Zone zu parken. Daneben gibt es auch Laufwerke, welche die Platten je nach Lese-/Schreiblast unterschiedlich schnell oder langsam drehen, anstatt sie anzuhalten.

Manche HDDs verfügen über ein mehrstufiges Powermanagement für Leerlauf- und Standby-Betriebszustände und passen den Betriebsmodus an den Aktivitätslevel an (z.B. Deaktivieren des Servosystems, Parken der Köpfe, Verlangsamen und Anhalten der Platten). Seagate's PowerChoice Technologie [5] ist ein Beispiel, wo schrittweise Komponenten deaktiviert werden, wenn die Platte ein bestimmtes Niveau der Inaktivität erreicht. Die Reaktivierung aus den verschiedenen Standby-Betriebsmodi ist generell kürzer als wenn die Platte gänzlich angehalten wird.

Tabelle 3.1. zeigt, dass der Standby-Verbrauch um rund 50% geringer ist als der Leerlauf-Verbrauch.

Tab. 3.1 PowerChoice Technologie für ein Constellation 2.5-Zoll Laufwerk

Betriebszustand	Leistungsaufn. (W)	Einsparung* (%)	Reaktivierung (sec.)	Default Timer
Leerlauf	2,82	0	0	n/a
Leerlauf_A	2,82	0	0	1 sec.
Leerlauf_B	2,18	23	0,5	10 min.
Leerlauf_C	1,82	35	1	30 min.
Standby_Z	1,29	54	8	60 min.

* Seagate Constellation SAS 2.5-inch hard drive.

Solche Technologien können auf RAID Systemen und Massive Arrays of Idle Disks (MAIDs) substanzielle Energieeinsparungen bewirken.

Da das Verlangsamen der Platten die Performance beeinträchtigen kann, wenden Hersteller zusätzliche Techniken an, wie größere Caches und Lese-/Schreibbefehl-Queuing. Um von Techniken wie Spin-Down und variablen Plattendrehzahlen verstärkt zu profitieren, gibt es inzwischen Programme auf Betriebssystem- und Anwendungsebene, die die Dauer der inaktiven Perioden der Platten verlängern. Einige dieser Ansätze beruhen darauf, die Datenzugriffe zu verschieben, indem man den Anwendungscode oder Datenlayouts modifiziert. Es gibt aber auch weniger invasive Techniken, die Compiler-Anpassungen erlauben, die die Datenzugriffsanfragen bei der Kompilierung neu reihen, ohne dass der Anwendungs Quellcode modifiziert werden muss.

Die genannten Techniken sind durchwegs geeignet, den Stromverbrauch zu reduzieren. Es wurde jedoch teilweise auch argumentiert, dass zu häufige Ein-Aus-Zyklen die Lebensdauer der HDDs reduzieren können.

Da Motoren und Plattenarme bei HDDs den Großteil des Energieverbrauchs verursachen, besteht eine mögliche Herangehensweise, die Laufwerke energieeffizienter zu machen, darin, kleinere Baugrößen einzusetzen. Da 2,5-Zoll HDDs ungefähr nur ein Viertel der Größe von größeren Festplattenlaufwerken haben (3,5-Zoll, s. Abb. 3.2), könnte ein Chassis, das ausreichend Volumen für 16 3,5-Zoll Laufwerke bietet, so umgebaut werden, dass es bis zu 48 2,5-Zoll Disks aufnehmen kann. Hochleistungs-Laufwerke in 2,5-Zoll Gehäusen weisen einen geringeren Energieverbrauch auf, da die Motoren und Plattenarme kleiner sind und daher weniger Abwärme entwickeln. Hersteller geben an, dass die IOPS/W (Input/Output-Operationen pro Sekunde pro Watt) für Tier-1 2,5-Zoll Plattenlaufwerke 2,5 Mal besser sein können



Abb. 3.2 2.5-Zoll HDD auf 3.5-Zoll HDD (aus Wikipedia)

als für vergleichbare Tier-1 3,5-Zoll Laufwerke [6]. Darüber hinaus wird weniger Strom für die Kühlung verbraucht, da weniger Abwärme entsteht. Tabelle 3.2 zeigt den ungefähren Stromverbrauch von zwei leistungsstarken Plattenmodellen von Seagate. Die kleinere Baugröße verbraucht im aktiven Betriebsmodus ungefähr 46% weniger Strom als das 3,5-Zollmodell, im Leerlaufmodus sogar bis zu 53% weniger. Rechnet man die

Kosten, um nur 24 Laufwerke im Jahr mit Energie zu versorgen, basierend auf dem Energieverbrauch im Standardbetriebsmodus und einem Preis von 0,11 Euro pro kWh, so beträgt der Unterschied zwischen den 3,5-Zoll-Laufwerken und 2,5-Zoll-HDDs ungefähr 140 Euro/Jahr. In Rechenzentren mit Hunderten oder Tausenden Platten, können sich die Einsparungen auf Tausende oder Zehntausende Euro belaufen.

Tab. 3.2 Energieverbrauch von Seagate High Performance Festplatten

Spezifikationen	Cheetah 15K.7 300GB*	Savvio 15K.2 146GB*	Einsparung
Baugröße	3,5"	2,5"	–
Kapazität	300GB	146GB	–
Schnittstelle	SAS 6Gb/s	SAS 6Gb/s	–
Geschwindigkeit (RPM)	15K	15K	–
el. Leistung Idle (W)	8,74	4,1	53% weniger
el. Leistung Aktiv (W)	12,92	6,95	46,2% weniger

* Datenquelle: Produktdatenblätter der Hersteller.

3.1.3 Solid State Drives (SSDs)

SSDs sind neben anderen Komponenten mit Flash-Speicher-Paketen und einem für verschiedene Tasks zuständigen Controller ausgestattet. SSDs bauen auf NAND-basiertem Flash Speicher Einheiten auf, der eine von zwei Typen von Speicherzellen verwendet, entsprechend der Bit-Anzahl, die eine Zelle speichern kann. Ein Single-Level Cell (SLC) Flash-Speicher speichert 1 Bit/Zelle, Multi-Level Cell (MLC) Speicher hingegen 2 oder 4 Bit/Zelle. Die meisten leistbaren SSDs verfügen über MLC, während die High-End-Produkte oftmals auf SLC basieren.

SSDs sind energieeffizienter und zuverlässiger, da sie keinerlei mechanische Teile wie Motoren und Antriebe enthalten. Sie entwickeln zudem weniger Abwärme, können in kleinere Gehäuse gepackt werden und verringern so den Platz- und Kühlbedarf.

Tabelle 3.3 zeigt einen einfachen Vergleich zwischen einer Seagate's Pulsar enterprise SSD und einer Hochleistungs- SAS 15k-RPM HDD. Die SSD verbraucht im aktiven Betriebsmodus ungefähr 87% weniger Strom als die 15k-RPM HDD und ca. 82% weniger im Leerlauf-Modus. In der Praxis hängen die Energieeinsparungen jedoch davon ab, wie die Speicherlösungen die SSDs und HDDs nutzen sowie im Weiteren von der Charakteristik der Anwendungen, die auf die Festplatten zugreifen.

Tab. 3.3 Vergleich Seagate Pulsar SSD und Savvio 15K HDDs

Spezifikationen	Savvio 15K.2 73GB*	Pulsar SSD 50GB*	Unterschied
Baugröße	3,5"	2,5"	–
Kapazität	73GB	50GB	–
Schnittstelle	SAS 3Gb/s SAS 6Gb/s	SATA 3Gb/s	–
Geschwindigkeit (RPM)	15K	–	–
NAND Flash Typ	–	SLC	53% weniger
Leistung Leerlauf (W)	3,7	0,65	82,4% weniger
Leistung Aktiv (W)	6,18	0,8	87% weniger

* Datenquelle: Produktdatenblätter der Hersteller.

3.1.4 Hybrid Hard Drives (HHDs)

HHDs sind HDDs, die über große Puffer aus permanentem Flash Speicher verfügen, die das Daten-Schreiben oder -Lesen auf den Platten reduzieren [7]. Der Puffer ermöglicht es, Platten längere Zeit im Ruhezustand halten. Der zusätzliche Flash-Speicher kann damit den Stromverbrauch für Motoren und mechanische Arme reduzieren. Diese Laufwerke haben also im Vergleich zu HDDs geringeren Strombedarf. Es werden derzeit aber nur wenige Produkte für die Anwendung im professionellen Bereich angeboten.

BEST PRACTICE EMPFEHLUNGEN

Erwägen Sie die Vorteile verschiedener Speichertechnologien bei der Anschaffung und im Systemdesign

- Bänder weisen die höchste Energieeffizienz für langfristige Speicherung auf.
- Aktuelle HDDs können mit verschiedenen Geschwindigkeiten betrieben werden und daher bei geringeren Geschwindigkeiten Energie sparen.
- Die vielfältigen Energiesparmodi in HDDs ermöglichen beträchtliche Energieeinsparungen, wenn man sie beispielsweise in Speicherlösungen wie Disk Arrays und MAIDS verwendet.
- Obwohl teurer, sind SSDs viel effizienter als HDDs.
- Ziehen Sie SSDs für High-Performance-Datenspeicherung in Betracht.

3.2 Datenspeicher-Gesamtlösungen

Dieser Abschnitt zeigt technische Möglichkeiten, die eingesetzt und kombiniert werden können, um die Energieeffizienz von Gesamtlösungen (Disk Arrays, SAN etc.) zu verbessern.

3.2.1 Laufwerke mit großer Kapazität und kleiner Baugröße

Für Anwendungen, die keine hohe Performance erfordern, ist es üblicherweise energieeffizienter, Laufwerke mit größerer Kapazität zu verwenden. Typische SATA-Plattenlaufwerke verbrauchen bis zu 50% weniger Strom pro Terabyte Speicherung als Fibre Channel Drives [8]. Wie bereits dargelegt,

können Speicherlösungen mit Laufwerken kleiner Baugröße auch den Raumbedarf in Rechenzentren reduzieren und den energetischen Fußabdruck verkleinern. Dell hat unter Verwendung des SPC-1C Benchmark (Storage Performance Council (SPC) zwei Disk Arrays verglichen, eines mit 3,5 Zoll HDDs und ein anderes mit 2,5 Zoll HDDs [9]. Wie die Ergebnisse zeigten, erbrachten die 2,5 Zoll-Laufwerke nicht nur 93% höhere Leistung als die 3,5 Zoll-Arrays, sondern verbrauchten auch 40% weniger Energie.

3.2.2 Massive Arrays of Idle Disks (MAIDs)

MAID ist eine Technologie, die eine Kombination aus Cache-Speicher und Idle Disks verwendet, um Zugriffe abzuarbeiten und Disks nur nach Bedarf

aktiviert. Die Reduktion der Festplattengeschwindigkeit oder Abschaltung der Platten bei wenig Zugriffen spart entsprechend Strom (s. Abb. 3.3) Wieviel Strom mit MAID-Funktionalitäten gespart werden kann, hängt von der Anwendung ab, die auf die Platte zugreift, und davon, wie oft auf die Laufwerke zugegriffen wird. Spin-down oder Spin-up der Platten wirken sich hinsichtlich Energieverbrauch und Plattenleistung aus. Ursprünglich erlaubte die MAID-Technologie nur An- und Abschaltung der Platten, was entsprechend zu Leistungseinbußen führen konnte, wenn Daten von einem gerade abgedrehten Laufwerk benötigt wurden. MAID-Technologien der zweiten Generation ermöglichen jedoch bereits Intelligentes Power Management (IPM) mit unterschiedlichen Stromsparmodi und Leistungslevels. MAID 2.0, wie es oft genannt wird, kann den Stromverbrauch verschiedenen QoS-Bedürfnissen zuweisen. Der Benutzer kann den Trade-off zwischen Reaktionszeit und Stromsparen individuell konfigurieren. Die Stromsparmodi nutzen unter anderem die verschiedenen HDD-Betriebszustände, die schon oben beschrieben wurden.

Andere Stromspartechniken für Disk Arrays sind beispielsweise die so genannte Popular Data Concentration (PDC) [10] und weitere File Allokations-Mechanismen [11]. Dahinter steht das Konzept, häufig verwendete Daten durch die Speicherung oder Migration in eine Untergruppe von Platten zu konsolidieren und die übrigen Platten in einen Energie sparenden Modus zu versetzen.

3.2.3 Effiziente RAID Level

Unterschiedliche RAID Levels bieten unterschiedliche Speichereffizienz. Wenn es um Datenschutz geht, erfordern manche RAID Levels, wie z.B. RAID6 einen bedeutenden Overhead für das Datenprocessing. Hochleistungs-RAID6-Lösungen können jedoch beispielsweise die gleiche Leistung wie RAID 5 bieten, bei bis zu 48% Reduktion der Anforderungen an die Plattenkapazität im Vergleich zu RAID 10.

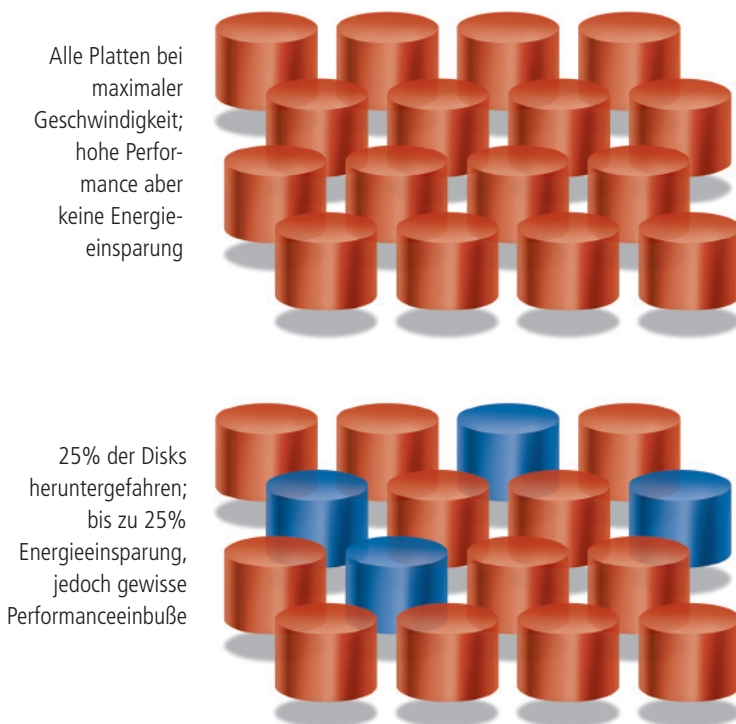


Abb. 3.3 Schematische Darstellung von MAIDs

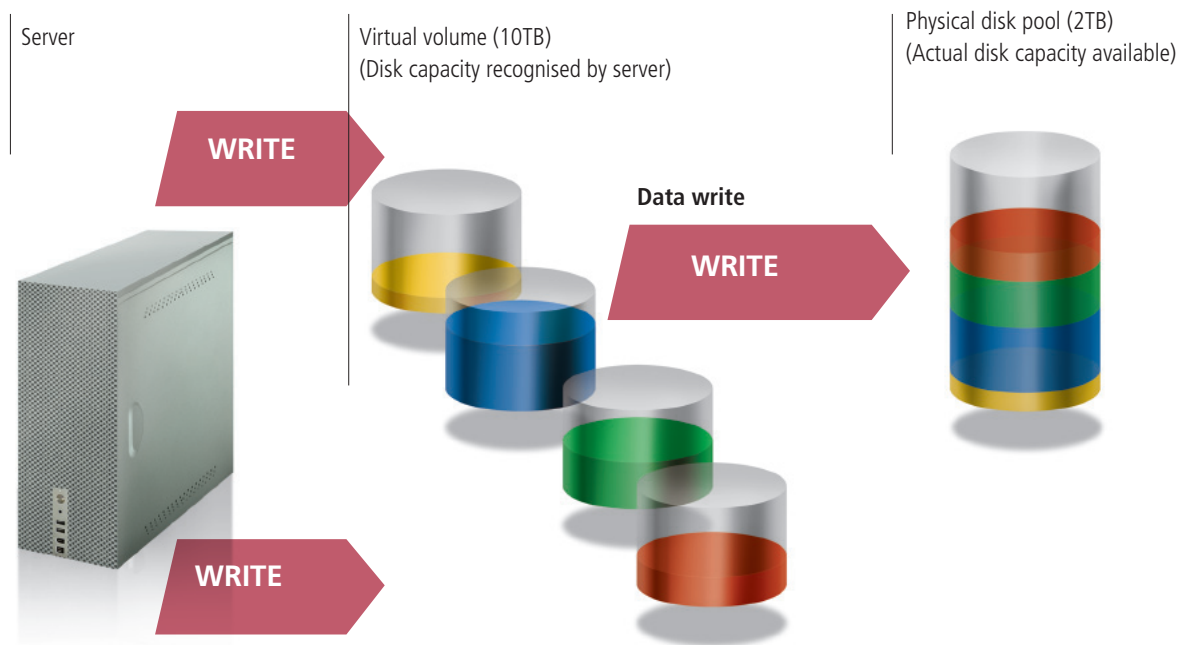


Abb. 3.4 Thin Provisioning (von Fujitsu ETERNUS solutions)

3.2.4 Horizontales Speicher-Tiering, Speichervirtualisierung und Thin-Provisioning

Für eine effiziente Nutzung der Speicherinfrastruktur ist es wichtig, Strategien für ein umfassendes Datenmanagement zu entwickeln, die mehrere Ebenen bzw. Speichertechnologien verwenden, die je nach Datenzugriffsrate, erforderlicher Geschwindigkeit und Dauer der Datenaufbewahrung zum Zug kommen. Hersteller von Datenspeicherlösungen bieten auch Softwaresysteme an, die eine automatische Abstufung ermöglichen, in dem die Daten auf Basis eines kontinuierlichen Performance-Monitorings automatisch auf die entsprechende Ebene migriert werden. Beispiele dafür sind EMC2's Fully Automated Storage Tiering (FAST), IBM's System Storage Easy Tier, Compellent's Data Progression und SGI's Data Migration Facility (DMF).

Durch die Kombination von Servervirtualisierung mit Speichervirtualisierung ist es möglich, Plattenpools und virtuelle Plattenspeicher zu erzeugen, deren Kapazität nach Wunsch und Bedarf der An-

wendungen erweitert werden kann. Die typische Effizienz von traditionellen Speicher-Arrays liegt zwischen 30 und 40%. Untersuchungen zufolge [12] kann durch Speichervirtualisierung die Effizienz bis zu 70% oder noch mehr erhöht werden, wodurch sich der Speicherbedarf verringert und Energieeinsparungen erzielt werden.

Die Virtualisierung von Speicherebenen, auch als Hierarchisches Speicher Management (HSM) bekannt, ermöglicht es, Daten automatisch zwischen verschiedenen Arten von Datenspeichern hin und her zu bewegen, ohne dass der Anwender etwas davon merkt. Softwaresysteme für automatisches Tiering werden verwendet, um solche Datenmigrationen durchzuführen. Diese Technologie kann die Kosten und den Stromverbrauch reduzieren, da sie es ermöglicht, dass nur oft benötigte Daten auf hochperformanten Datenspeichern gelagert werden, während weniger oft gebrauchte Daten auf energieeffizienteren Medien/Lösungen gespeichert werden, die beispielsweise Techniken wie MAID und Datenduplizierung verwenden.

Thin Provisioning ergänzt die Speichervirtualisierung und zielt primär darauf ab, die Speicher-

nutzung zu maximieren und vordefinierte, aber nicht verwendete Kapazität zu eliminieren. Beim Thin Provisioning wird Speicherplatz zur Verfügung gestellt, wenn die Daten geschrieben werden. Es wird keine Reservekapazität entsprechend dem maximalen Speicherbedarf der Anwendungen definiert. Die Reservekapazität wird auf Null gesetzt. Die Speicherkapazität wird online erweitert bzw. Kapazität wird dynamisch hinzugefügt, um Änderungen ohne Unterbrechung zu bewerkstelligen (s. Abb. 3.4). Thin Provisioning kann Energieeinsparungen unterstützen, da der Bedarf nach überdimensionierter Speicherkapazität für Anwendungen reduziert wird.

3.2.5 Konsolidierung auf Speicher- und Strukturebene

Speicherkonsolidierung ist kein neues Thema, da Storage Area Networks (SANs) schon seit mehreren Jahren ein gewisses Niveau an Speicherkonsolidierung und verbesserter Effizienz geboten haben, indem sie Platten-Arrays über mehrere Server auf einem lokalen privaten Netzwerk teilen, um so Dateninseln zu vermeiden. Der Wechsel von direkt verbundenen Speichern (DAS) zu Speichernetzwerken bietet eine Reihe von Vorteilen und kann die Energieeffizienz erhöhen. Konsolidierung der Datenspeicherung kann sowohl zu substantiellen Einsparungen beim Platzbedarf als auch beim Energieverbrauch führen. Einige Hersteller argumentieren, dass durch ein Multi Protokoll Netzwerk Equipment die Netzwerkstruktur auf weniger Ressourcen konsolidiert werden kann, wodurch auch der Platzbedarf, der Stromverbrauch und der Kühlbedarf reduziert werden.

3.2.6 Daten-Deduplizierung

Speicherinfrastrukturen speichern in vielen Fällen mehrfache Kopien derselben Daten. Datenduplikation erfolgt in Rechenzentren oft auf mehreren Ebenen, einige davon sind notwendig, um die Zuverlässigkeit und den Datendurchsatz zu erhöhen. Es gibt jedoch auch "Datenabfall", der minimiert werden kann, sodass wieder Speicherkapazität zur Verwendung frei wird. Aktuelle SAN-Lösungen verwenden Daten-Deduplizierungstechniken, um Duplikate zu reduzieren. Diese Techniken arbeiten vor allem auf Ebene von Datenblöcken und Files.

Abgesehen von der Ebene, auf der die Daten-deduplizierung stattfindet, unterscheiden sich die Deduplikationsmethoden auch hinsichtlich des Zeitpunktes der Deduplizierung, die vor oder nach der Speicherung der Daten auf der Platte erfolgen kann. Beide Methoden haben Vor- und Nachteile. Obwohl sie zu reduziertem Speicherbedarf führt, erfordert die Deduplizierung nach der Datenspeicherung eine Cache-Speicherung für die Entfernung der Duplikate. Bei Backup-Anwendungen führt Deduplikation nach der Speicherung der Daten üblicherweise zu kürzeren Backup-Perioden und geringerem Leistungsabfall. Datendeduplikationsmethoden unterscheiden sich aber auch noch darin, wo Deduplizierung durchgeführt wird, d.h. auf der Ausgangsseite (der Quelle), der Zielseite (Server), oder durch eine Deduplikationsanwendung, die mit dem Server verbunden ist.

Daten-Deduplikationslösungen ermöglichen es Organisationen, Speicherkapazität wiederzuverwenden und Speichermedienbedarf zu reduzieren und sie eignen sich damit auch als Methode zur Energieverbrauchsreduktion. Die Speichereinsparungen durch Daten-Deduplikationslösungen unterscheiden sich abhängig von ihrer Granularität. Lösungen, die Hashing und Deduplikation auf Datei-Ebene durchführen, sind tendenziell weniger effizient, sie erfordern jedoch einen geringeren Overhead. Bei Methoden auf Block-Ebene ist die Effizienz im Allgemeinen indirekt proportional zur Blockgröße.

Obwohl Daten-Deduplizierung eine vielversprechende Technologie darstellt, um „Datenabfälle“ zu reduzieren und den Energieverbrauch zu verringern, können nicht alle Anwendungen davon profitieren. So kann eine Deduplizierung vor der Datenspeicherung zu einem gravierenden Leistungsabfall führen, was beispielsweise für Datenbankanwendungen inakzeptabel wäre. Anwendungen, die große Datenvolumina für lange Zeitperioden lagern, profitieren entsprechend stärker. Je mehr Daten vorhanden sind und je länger diese aufbewahrt werden müssen, desto bessere Resultate lassen sich mit Datendeduplikation erzielen.

Literaturempfehlungen

McClure T. (2009): Driving Storage Efficiency in SAN Environments, Enterprise Strategy Group - White Paper, November 2009.

Craig B. and McCaffrey T. (2009): Optimizing Nearline Storage in a 2.5-inch Environment Using Seagate Constellation Drives, Dell Power Solutions, Jun. 2009.

SNIA (2010): Storage Power Efficiency Measurement Specification: Working Draft Version 0.2.10, SNIA Green Storage Initiative, August 2010. Storage Tiering with EMC Celerra FAST, EMC2 www.snia.org/sites/default/files/Storage_Power_Efficiency_Measurement_Spec_v0.2.10_DRAFT.pdf

Clark T. and Yoder A. (2008): Best Practices for Energy Efficient Storage Operations Version 1.0, SNIA Green Storage Initiative, October 2008.

Freeman L. (2009): Reducing Data Centre Power Consumption Through Efficient Storage. White Paper. NetApp, July 2009.

Zitierte Literatur

- [1] **IDC (2010):** The Digital Universe Decade – Are you ready? IDC, May, 2010.
- [2] **McClure T. (2009):** Driving Storage Efficiency in SAN Environments, Enterprise Strategy Group - White Paper, November 2009.
- [3] **Reine D. and Kahn M. (2008):** Disk and Tape Square Off Again – Tape Remains King of the Hill with LTO-4. Clipper Notes, February 2008.
- [4] **ORACLE (2010):** Consolidate Storage Infrastructure and Create a Greener Datacentre. Oracle White Paper, April 2010.
- [5] **Seagate (2011):** PowerChoice Technology Provides Unprecedented Hard Drive Power Savings and Flexibility - Technology Paper, Seagate, 2011.
- [6] **Seagate (2010):** Seagate Savvio 15K.2 Data Sheet, Seagate, 2010.
- [7] **Bisson T., Brandt S., Long D. (2006):** NVCache: Increasing the Effectiveness of Disk Spin-Down Algorithms with Caching, 14th IEEE International Symposium on Modeling, Analysis, and Simulation, pp. 422-432, 2006.
- [8] **Freeman L. (2009):** Reducing Data Centre Power Consumption Through Efficient Storage. White Paper. NetApp, July 2009.
- [9] **Craig B. and McCaffrey T. (2009):** Optimizing Nearline Storage in a 2.5-inch Environment Using Seagate Constellation Drives, Dell Power Solutions, Jun. 2009.
- [10] **Pinheiro E. and Bianchini R. (2004):** Energy Conservation Techniques for Disk Array-Based Servers. 18th Annual International Conference on Supercomputing (ICS 2004), pp. 68-78. Malo, France, 2004.
- [11] **Otoo E. D., Rotem D. and Tsao S.C. (2009):** Analysis of Trade-Off between Power Saving and Response Time in Disk Storage Systems, IEEE International Symposium on Parallel Distributed Processing (IPDPS 2009), pp. 1-8, May 2009.
- [12] **Blade Network (2009):** Storage Consolidation for Data Centre Efficiency, BLADE Network Technologies White Paper, Jun. 2009.



4 Netzwerk Equipment

Alexander Schlösser, TU Berlin, Lutz Stobbe, Fraunhofer IZM

Nach aktuellen Informationen beträgt der Energieverbrauch, welcher Switchen, Routern und weiteren Netzwerkgeräten zugeordnet werden kann, ca. 8 bis 12 Prozent des Gesamtenergieverbrauchs der IT-Ausstattung in Rechenzentren. Aufgrund dieses relativ geringen Anteils am Gesamtenergieverbrauch in Rechenzentren stand das Netzwerkequipment bislang nicht im Fokus gezielter Energieeffizienzmaßnahmen. Derzeit ändert sich jedoch die Situation gerade in mittleren und großen Rechenzentren.

Es gibt mehrere Gründe, warum der Strombedarf von Netzwerkgeräten und die energiebezogenen Effekte bei der Implementierung entsprechender Netzwerkarchitekturen aktuell eine wachsende Bedeutung bei der Planung, dem Design und dem Betrieb von Rechenzentren erfahren.

Mit steigenden Anforderungen an die Dienstgüte (Quality of Service) in Verbindung mit zeitkritischen Applikationen rücken Netzwerkgeräte und das Netzwerk an sich wieder stärker in den Fokus für Designverbesserungen. Der Stromverbrauch variiert in Bezug auf die ausgewählte Technologie und Architektur inklusive der Verkabelung, Stromversorgung und Kühlung.

4.1 Technischer und operativer Rahmen

4.1.1 Funktionales Modell

Abbildung 4.1 zeigt ein vereinfachtes funktionales Modell der energierelevanten Aspekte für Netzwerke und Netzwerkequipment in Rechenzentren. Das funktionale Modell veranschaulicht die sich überschneidenden Aspekte der Stromversorgung und Kühlungsinfrastruktur. Des Weiteren können mit Hilfe des Modells die Wechselwirkungen zwischen Netzwerk, Server und Speichersystem an Hand horizontaler Themen erklärt werden. Das umfasst die Auswahl einer geeigneten Netzwerkarchitektur, -topologie, der physikalischen Infrastruktur, entsprechender Hardwarekomponenten und Kabel sowie der Softwarekonfiguration und der Virtualisierungskapazität.

Die Energieeffizienz der Netzwerkinfrastruktur und der Netzwerkgeräte wird auch durch die verwendeten Applikationen, Service Level Agreements, Bandbreite und Latenzanforderungen beeinflusst, welche durch den Betreiber des Rechenzentrums definiert werden. Diese auf die Performance be-

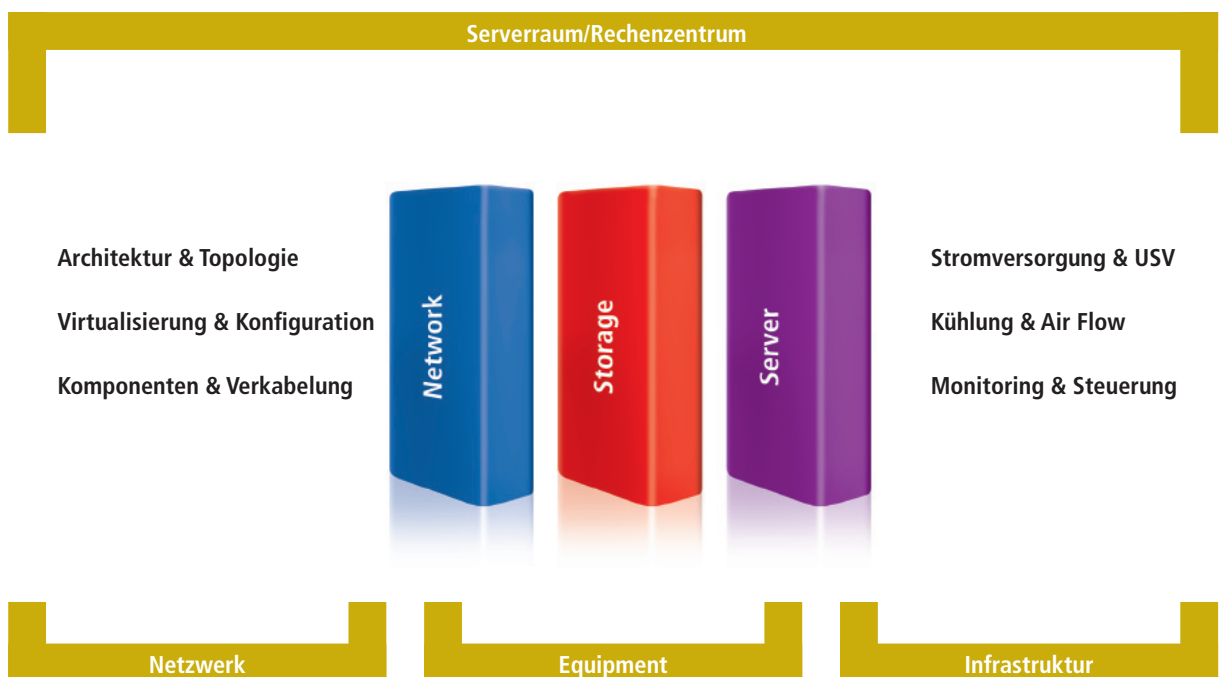


Abb. 4.1: Funktionales Modell für Netzwerke in Rechenzentren

zogenen Aspekte müssen bereits im Planungsprozess berücksichtigt werden, um die Energieeffizienz des gesamten Netzwerkes zu verbessern.

4.1.2 Netzwerkeigenschaften

Die Verbesserung der Energieeffizienz unter Berücksichtigung der Netzwerkinfrastruktur erfordert einen strukturierten Ansatz. Die Planung eines neuen Netzwerkes sollte eine strategische oder Langzeit-Ausrichtung enthalten, da die Netzwerkinfrastruktur typischer Weise länger im Einsatz ist als entsprechende Server- oder Speichertechnologien. Es wird angenommen, dass die Basisnetzwerkinfrastruktur heutzutage mehr als 8 Jahre im Einsatz ist, bevor sie durch eine neue ersetzt wird. Der Austausch der grundlegenden Netzwerkarchitektur und -topologie mit den dazugehörigen Komponenten ist mit erheblichen Investitionen verbunden. Ungeachtet dessen schafft die Aufrüstung des Netzwerkes nicht nur einen Performancegewinn, sondern ermöglicht zugleich in vielen Fällen auch eine Optimierung der Energieeffizienz. Einer geplanten Effizienzsteigerung sollte deshalb immer eine strategische Analyse vorausgehen, welche Energieeffizienz Aspekte mit berücksichtigt.

Der Rechenzentrumsbetreiber muss im Vorfeld seine Netzwerkeigenschaften und Performance-

anforderungen festlegen. Die IT-Welt befindet sich gerade in einem enormen Wandel hin zu einer zentralisierten Bereitstellung von Applikationen, was im Resultat zu wachsenden und neuen Datenvolumina und -mustern führt. In anderen Worten, die Applikationen werden nicht mehr dezentral auf der Anwenderseite bereitgestellt. Damit kann die am Arbeitsplatz bisher notwendige hohe Rechenleistung und Software drastisch gesenkt werden. Mit dem Einsatz von Software-as-a-Service und Cloud Computing werden Applikationen und der zugehörige Datenverkehr zukünftig in Rechenzentren und RZ-Clouds verlagert. Eine notwendige Voraussetzung für diesen Trend sind eine entsprechend hohe Breitbandanbindung und geringe Latenzzeiten.

Dieser allgemeine Trend führt nicht nur zu erhöhtem Datenverkehr zwischen Clients und Servern, sondern auch zu höheren Datenströmen zwischen verschiedenen Servern sowie zwischen Servern und Speichersystemen. Enterasys zeigt in diesem Zusammenhang auf, dass sich Netzwerkarchitektur und -konfiguration verändern müssen, um den wachsenden Server-zu-Server- und Speicher-zu-Server-Datenverkehr zu unterstützen. Um die Performance zu steigern, lässt sich aktuell ein technischer Trend hin zu einer aggregierten (bottom-up) und virtualisierten Vernetzung (top-down) ableiten.

Die Netzwerkarchitektur wird zukünftig aus weniger Ebenen bestehen, indem Zugangs- und Aggregationsebene sowie das Kernnetzwerk mit dem Aggregationsnetzwerk zusammengelegt werden (siehe Abbildung 3). Dieser Trend hat das Potential, den Energieverbrauch mittels einer einheitlichen und flachen Netzwerkarchitektur zu senken. Dies ist natürlich ein Balanceakt zwischen Performance- und Sicherheitsanforderungen sowie Effizienzverbesserungen. Mit nur wenigen verfügbaren Informationen und Daten kann derzeit keine herausragende Lösung am Markt identifiziert werden. Nach [1] umfassen die Designziele für RZ-Netzwerke:

- Hohe Bandbreite, geringe Latenzzeit (Auswahl der Netzwerktechnologie)
- Skalierbarkeit und Agilität (Netzwerkarchitektur und -management)
- Flexibilität zur Unterstützung div. Services (in Bezug auf Konsolidierung)
- Sicherheit (zunehmende Bedeutung, beeinflusst den Overhead)
- Hohe Verfügbarkeit und Redundanz (QoS-Anforderungen)
- Steuerung und Transparenz (wird durch Virtualisierungslösungen unterstützt)
- Kostenoptimierung (CAPEX und OPEX reduzieren)

EMPFEHLUNGEN FÜR NETZWERK-DESIGN

Da sich Netzwerkprodukte durch eine große Produktvielfalt mit zahlreichen Kombinationsmöglichkeiten auszeichnen, wird RZ-Betreibern und IT-Administratoren empfohlen, eine Prioritätenliste für entsprechende Netzwerkeigenschaften aufzustellen, wie:

- Netzwerk-Services
- Latenzzeitanforderungen
- Quality-of-Service
- Virtualisierungsunterstützung
- Weitere Performance- oder Interoperabilitätsaspekte

Am besten geeignet ist ein System-Ansatz. Dieser berücksichtigt nicht nur die Wechselwirkungen zwischen der Netzwerkinfrastruktur und Performance, sondern auch die Wechselwirkung mit dem IT-Equipment und der Kühl- und Stromversorgungsinfrastruktur.

4.1.3 Abwägen von Netzwerk-Leistungsparametern und Energiebedarf

Bandbreite, hohe Geschwindigkeit, geringe Latenzzeit und verlustfreier Datentransport sind die wichtigen Performancekriterien für RZ-Netzwerke. Kundenzufriedenheit bzw. die sogenannten Service-Level-Agreements (SLA) sind hier weitere bedeutende Performanceanforderungen. QoS wird oft innerhalb der Service-Level-Agreements festgelegt und umfasst Kenngrößen wie minimaler Datendurchsatz, maximale Reaktionszeit oder Latenzzeit. Eine schnelle, konvergierte und intelligente Netzwerkarchitektur, die den Netzwerkverkehr dynamisch entsprechend den vereinbarten SLAs behandelt, ist nicht nur für die zukünftige Wettbewerbsfähigkeit entscheidend, sondern könnte auch eine Grundlage für einen systematischen Energieeffizienzansatz bilden.

Allerdings kann die Implementierung der QoS den Netzwerkverkehr erhöhen und in diesem Zusammenhang auch eine Zunahme des Energieverbrauchs im Rechenzentrum verursachen. Individuelle Netzwerktechnologien und -geräte haben diesbezüglich unterschiedliche Vor- und Nachteile. Ein genereller Trend ist die Etablierung der 10 Gigabit-Ethernet-Technologie (10 GbE) für RZ-Netzwerke. Das Ethernet verbindet nicht nur die Server in einem LAN, sondern wird neuerdings auch verstärkt in Speichernetzwerken (SAN) eingesetzt.

Geringe Latenzzeit und verlustfreie Netze sind grundlegende Voraussetzungen für den Speicherverkehr. Nach [2] haben heutige 10 GbE-Switches eine Latenzzeit von 400 bis 700 ns. Für 2014 wird angenommen, dass 100 GbE-Switches die Latenzzeit auf 100 ns verringern werden. Daraus geht hervor, dass mit steigender Bandbreite sich die Latenzzeit der Netzwerkgeräte verbessert. Aus der Perspektive des Energieverbrauchs ist es daher notwendig, eine Balance zwischen einer verbesserten Latenz (Netzwerktechnologie) und dem potentiell höheren Stromverbrauch hoher Bandbreiten (Komponenten) zu finden. Die

Auswahl geeigneter Komponenten und I/O-Konsolidierung sind zwei Hauptaspekte, die in diesem Kontext angewendet werden.

Ebenso ist es notwendig, die verlustfreie Vernetzung (Verfügbarkeit) gegenüber der Bandbreitenperformance und der resultierenden Energieeffizienz zu untersuchen. Eine verlustfreie Vernetzung bedeutet im Allgemeinen komplexere Protokolle und zusätzliche Latenz in Verbindung mit höherer Rechenleistung und geringerer Bandbreiteneffizienz. Verlustfreie Vernetzung ist in jedem Fall eine Voraussetzung für Speichernetzsysteme. In der Vergangenheit war das nicht verlustfreie Ethernet-Protokoll ein Hinderungsgrund für dessen Anwendung in Speichernetzen.

Fibre Channel (FC) und Infiniband (IB) sind die verbreitetsten Netzwerktechnologien für Speichernetze. Inzwischen gibt es aber auch eine Vielzahl von Speichernetzwerk-Lösungen, die auf dem Ethernet-Standard basieren. Dazu zählen unter anderem Converged Enhanced Ethernet (CEE), Fibre Channel over Ethernet (FCoE), Internet Small Computer System Interface (iSCSI) over Ethernet, ATA over Ethernet (AoE) und Network-attached Storage (NAS). Diese Optionen ermöglichen es, die unterschiedlichen RZ-Netzwerke zu vereinheitlichen (keine zusätzlichen Adapter notwendig). Gleichzeitig verursachen sie einen zusätzlichen Overhead, der wiederum die Bandbreiteneffizienz absenkt. Potentielle Zielkonflikte hinsichtlich der Energieeffizienz sind bislang nicht bekannt.

Zusammenfassend ist festzustellen, dass der Betreiber die energetischen Einflüsse für eine gesteigerte Performance, Skalierbarkeit und neue konsolidierte Netzwerk-Lösungen berücksichtigen sollte. Es ist wahrscheinlich, dass neue Netzwerk-Lösungen zu einer verbesserten Energieeffizienz führen, wobei eine angemessene Dimensionierung grundlegend ist. Es wird empfohlen, dass Betreiber, die neue Netzwerkprodukte oder komplette Netzwerk-Lösungen beziehen, sich nach dem Gesamteinfluss hinsichtlich des Energieverbrauchs erkundigen.

4.2 Verbesserung der Energieeffizienz

4.2.1 Vereinheitlichung der Traffic Classes (I/O Konsolidierung)

Rechenzentrumsnetzwerke müssen verschiedene Arten von Datenverkehr innerhalb verschiedener Anwendungsgebiete übertragen. Dies hat in der Vergangenheit zur Entwicklung von speziellen Protokollen und Netzwerkarchitekturen geführt. Als Ergebnis existieren komplexe Netzwerke, die ihre Ressourcen nicht teilen. Daher ergeben sich als grundlegende Verbesserungsziele die Reduzierung von physikalischen Komponenten und die gemeinsame Nutzung von Netzwerkkapazität durch verschiedene funktionelle Einheiten. Der generelle Trend hin zu weniger bzw. einfacheren Ebenen und I/O Converged Networking basierend auf Ethernet ist ebenfalls durch den Energieeffizienzgedanken bedingt. Das übergeordnete Thema ist Netzwerkkonsolidierung, was Server- und Speichernetzwerke sowie Netzverteilungsarchitektur adressiert. Convergent Network Adapter (CAN) verbinden ehemals getrennte Schnittstellen:

- Host Bus Adapter (HBA) zur Unterstützung von SAN-Übertragung
- Network Interface Controller (NIC) zur Unterstützung von LAN-Übertragung
- Host Channel Adapter (HCA) zur Unterstützung von IPC-Übertragung

I/O-Konsolidierung ist die Möglichkeit eines Switches oder eines Host Adapters, die gleiche physikalische Infrastruktur zu nutzen, um verschiedene Arten von Datenströmen zu transportieren, wobei jeder der Datenströme über eigene Charakteristiken und Handling Requirements verfügt. Hinsichtlich des Netzwerks bedeutet dies, dass nur ein Netzwerk anstelle von dreien (wie in Abbildung 2) installiert und betrieben werden muss. Hinsichtlich der Hosts und Speicherarrays resultiert dies in weniger Convergent Network Adapters (CAN) gegenüber Ethernet NICs, FC HBAs und IB HCAs. Ein typischer Fibre Channel HBA verbraucht ca. 12,5 W. [3] Bezüglich Netzwerkredundanz müssen verschiedene Optionen abgewogen werden, um zuverlässige Netzwerke zu designen.

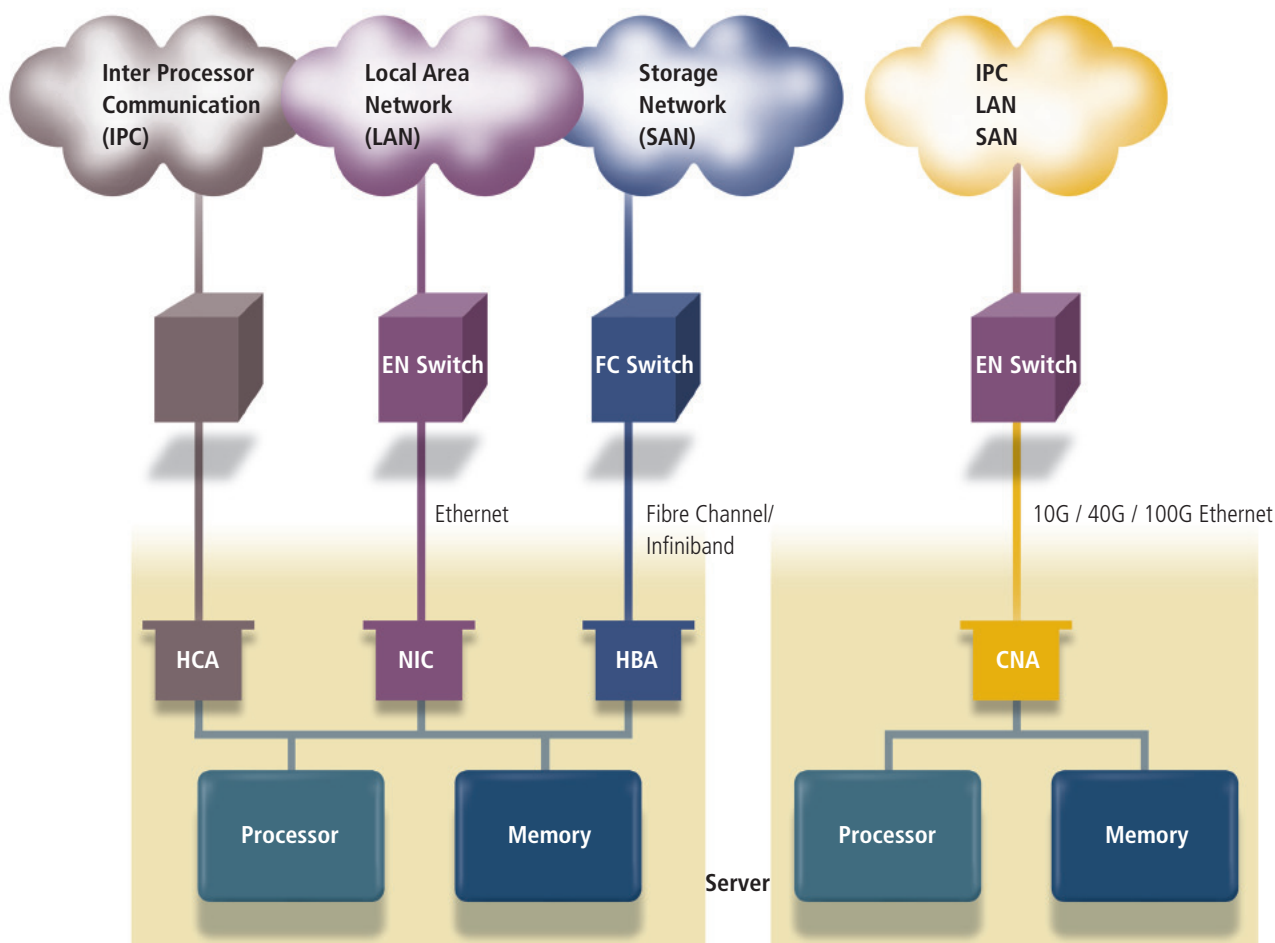


Abb. 4.2 I/O Konsolidierung und Netzwerkkonvergenz in Rechenzentrumsnetzwerken

VORTEILE VON KONVERGIERTEN NETZWERKEN

I/O-Konsolidierung ermöglicht die Konsolidierung von verschiedenen Netzwerktypen (LAN, SAN) auf einem höheren Level als Vorbereitung zur Systemvirtualisierung. Darüber hinaus wird die Zahl der physikalischen Infrastrukturelemente wie Switches, Ports, Verbindungen und Kabel innerhalb verschiedener Netzwerke signifikant reduziert. Konvergierte Netzwerke resultieren in:

- bis zu 80% weniger Adapter und Kabel
- bis zu 25% Reduktion der Switches, Adapter und Rackfläche
- bis zu 42% Reduktion der Strom- und Kühlungskosten [4]

4.2.2 Netzwerkkonsolidierung

Der Ansatz zur Optimierung des Energieverbrauchs von Rechenzentrumsnetzwerken zielt im Wesentlichen auf neue Netzwerkarchitekturen und die Zusammenführung von ehemals getrennten Netzwerken (in eine gemeinsame Technologie). Eine typische Architektur besteht aus verschiedenen Routing- und Switching-Komponenten (auf verschiedenen Ebenen) mit spezialisierteren und teureren Komponenten am oberen Ende der Netzwerkhierarchie. Das Ziel sollte es sein, die Netzwerkinfrastruktur so zu konsolidieren, dass eine flache Netzwerkarchitektur basierend auf einer funktionellen Netzwerkstruktur entsteht.

Folgende Maßnahmen sind möglich:

- Aggregation von Switches: mehrere physikalische Switches, die als ein logisches Element agieren.
- Reduzierung der Ebenen/TIERS: Nutzung eines aggregierten Switches, um mehrere Switch-Ebenen zu ersetzen, unter Berücksichtigung der Netzwerkdienste und Sicherheitsanforderungen
- Ausbildung einer vereinheitlichten Netzwerkstruktur: dies verbindet zwei Ansätze und ermöglicht einen einfachen Betrieb und hohe Performance, wieder unter Berücksichtigung der Netzwerkdienste und Sicherheitsanforderungen.

Die Zusammenführung von Server- (LAN) und Speicher- (SAN) Netzwerken ist ein Trend mit Energiesparpotenzial. Das Beibehalten von zwei getrennten Netzwerken würde die Betriebskosten und den Stromverbrauch durch die größere Anzahl von Adaptern, Kabel und Switchanschlüssen, die zum direkten Anschluss jedes Servers mit den unterstützenden LANs und SANs notwendig wären, erhöhen. Um die Netzwerkstruktur des Rechenzentrums zu vereinfachen bzw. abzufachen, werden momentan Converged Networking Technologies wie iSCSI, Fibre Channel over Ethernet (FCoE) und Data Center Bridging (DCB) implementiert.

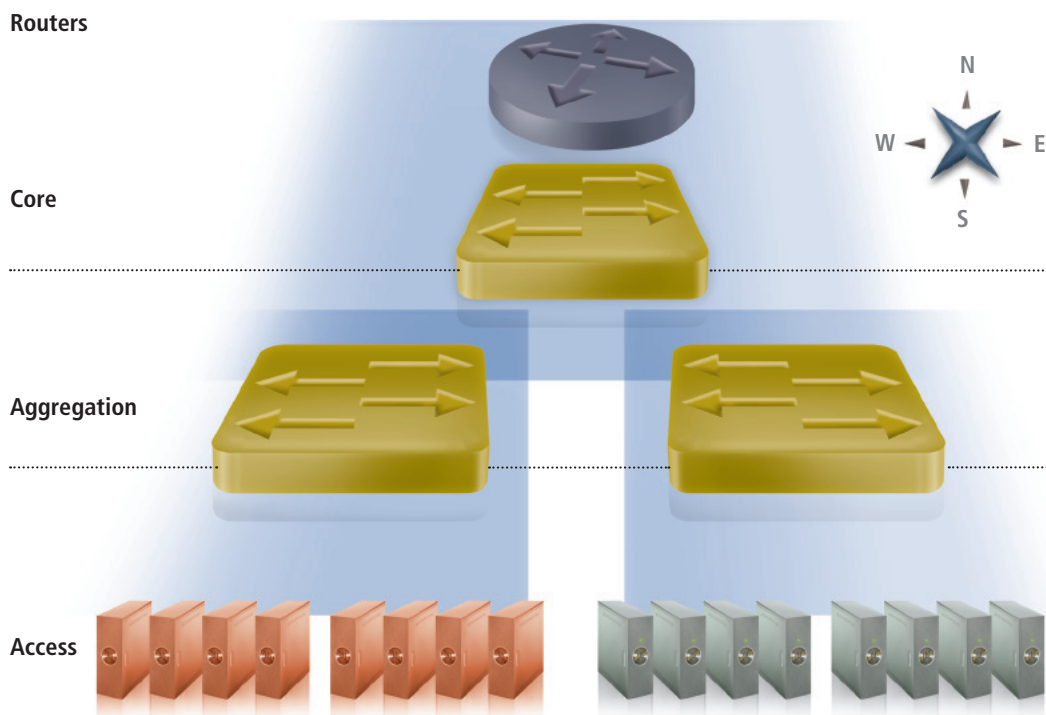


Abb. 4.3 Netzwerkkonsolidierung

4.2.3 Netzwerkvirtualisierung

Virtualisierung ist inzwischen eine etablierte Technologie, um physische Server durch verschiedene virtuelle Maschinen zu konsolidieren. Die Netzwerkvirtualisierung folgt hierbei dem gleichen Prinzip und beschreibt verschiedene Hardware- und Softwareansätze, um die Netzwerkressourcen als logische Einheiten unabhängig von ihrer physischen Topologie zu verwalten. Daraus folgt ein verringertes Netzwerkverkehr, vereinfachte Sicherheit und verbesserte Netzwerkkontrolle. Schlüsselemente für hohe Netzwerkeffizienz sind network level awareness und eine gute Sichtbarkeit der virtuellen Maschinen. Die Möglichkeit, das Netzwerk und die Ressourcen der Ports individueller VM frei zu konfigurieren sowie die dynamische Rückverfolgung VMs, während sich diese über das gesamte RZ verteilen, sind entscheidend für die effiziente Verwaltung virtualisierter Umgebungen. Die Energieeffizienz wird überwiegend durch die Konsolidierung der Router, der physikalischen I/O-Ports und der zusätzlichen Hardware für spezifische Netzwerkdienste erreicht. Die erweiterte Systemvirtualisierung der Netzwerke beinhaltet:

- Virtuelle Router (Software mit Routingfunktionalität, multiple Systeme auf einer realen Maschine)
- Virtuelle Links (logische Verbindungen der virtuellen Router)
- Virtuelle Netzwerke (virtuelle Router verbunden über virtuelle Links)

Die wachsende Servervirtualisierung verursacht eine zusätzliche Komplexität und Overhead für das Netzwerk. Überholte Netzwerkschichten können virtuelle Maschinen (VM) nicht bedienen, weshalb die daraus resultierenden fehlerhaften Netzwerkkonfigurationen ein Risiko für Serviceausfälle und Sicherheitslücken darstellen. Die Vernetzung ist ein Schlüsselbereich, der ebenso virtualisiert werden muss wie die Server- oder Speichersysteme, um ein einheitliches Level von Agilität, Bandbreite und Performance zu erreichen.

Netzwerkvirtualisierung ist eine Strategie, um das Netzwerk zu vereinfachen und mehrere Geräte zu konsolidieren. Die Virtualisierung der Firewallmodule oder Intrusion Prevention Systeme (IPS) erfolgt über Software-Images, die unterschiedli-

chen Anwendungen über eine einzige Netzwerkhardware zur Verfügung gestellt werden. Durch die effektive Nutzung von Software über die gleiche Hardware wird der Bedarf an separaten Geräten deutlich reduziert.

Die Reduzierung des Strombedarfs wird hierbei über die Konsolidierung mehrerer Geräte in ein physisches Gerät erreicht. Dabei kann auf die Zuweisung einer dedizierten Hardware für jeden möglichen Fall verzichtet werden. Durch die Vermeidung zusätzlicher physikalischer Geräte entfallen auch zusätzliche Netzteile sowie Kühlung und Rackfläche, welche sonst notwendig gewesen wären.

Vorteile der Netzwerkvirtualisierung:

- Flexiblere Verwaltungsschnittstellen
- Geringere Anschaffungskosten durch die Nutzung von Software
- Verbesserte Applikationsperformance durch vereinfachte Serviceerweiterung und -zuordnung
- Potenziell verringerter Strombedarf durch Gerätekonsolidierung

Die erfolgreiche Implementierung der Netzwerkvirtualisierung ist abhängig von Aspekten wie Kapitaleinsatz, der Definition klarer Ziele oder der Kompatibilität mit vorhandener Hardware. Deshalb benötigen Virtualisierungsprojekte eine gut ausbalancierte Kosten-Nutzen-Analyse, ein übergreifendes Projektmanagement und eine konsequente Berücksichtigung möglicher Sicherheitslücken.

EMPFEHLUNGEN FÜR KLEINERE BIS MITTLERE RECHENZENTREN

Für kleine bis mittlere Rechenzentren hängt die Wahl zwischen FCoE und iSCSI hauptsächlich von Anwendungsanforderungen und der Verfügbarkeit von in Fibre Channel ausgebildetem Personal ab.

- Wenn die entscheidenden Anwendungen High-Performance-Datenbanken sind, sind FCoE und iSCSI mögliche Lösungen, um den Service-Level zu verbessern und den Energieverbrauch zu reduzieren.
- Zentralisierte Speicher und Disaster Recovery (Datenwiederherstellung nach Ausfall) benötigen gemeinsamen SAN --> iSCSI ist zu bevorzugen.
- Bei Fibre Channel-dominierten Netzwerken wird der Einsatz von FCoE empfohlen [5].

4.2.4 Auswahl geeigneter Komponenten und Produkte

Der Stromverbrauch der Netzwerkgeräte wird im Allgemeinen durch die Auswahl und Konfiguration der im System verwendeten Komponenten beeinflusst. Ein Haupttreiber ist der verwendete Netzwerkstandard (z.B. 10 GbE). Das Chipdesign und die Tiefe der Systemintegration haben hierbei den größten Hebel für das Produktdesign. Dieser Trend wird durch die Performanceverbesserungen in der Halbleitertechnologie getrieben und folgt dabei nach wie vor Moores Gesetz. Das beinhaltet auch die thermische Performance des Chips und der Aufbau- und Verbindungstechnologie. Die Zuverlässigkeit erlangt in diesem Kontext immer größere Bedeutung. Ein weiterer wichtiger Faktor ist die Systemkonfiguration des Netzwerkes, wobei die Anschlussarten und die Anzahl der Ports innerhalb der Geräte entscheidend sind. Schließlich wird der Energiebedarf der Netzwerkgeräte auch durch die Effizienz der integrierten Netzteile und der verfügbaren Powermanagementfunktionen bestimmt.

Powermanagement

Das Ausmaß des Energiebedarfs der Netzwerkgeräte hat einen direkten Bezug zur aktiven Nutzungsphase und den entsprechenden Idle-Phasen. Der unterschiedliche Strombedarf zwischen der aktiven Phase (100% Auslastung) und der Idle-Phase (bestehende Verbindung) beträgt typischer Weise weniger als 10%. Sobald eine Verbindung deaktiviert wird, sinkt der Stromverbrauch um 50% gegenüber der aktiven Verbindung.

Es wird angenommen, dass in kleineren Einrichtungen (z.B. Serverräumen und kleinen RZ) Idle-Phasen verstärkt während der Nacht auftreten. Ein hierfür notwendiges erweitertes Powermanagement beinhaltet eine Art „networked standby“, für welches es bislang noch keinen einheitlichen Standard gibt. Der Begriff „networked standby“ wurde innerhalb der Vorbereitungsstudie für die europäische EuP/ErP Rahmenrichtlinie ENER Los 26 eingeführt. Diese Studie argumentiert auf der

BEST PRACTICE EMPFEHLUNGEN

Prüfen Sie Ihre Beschaffungskriterien hinsichtlich energieeffizienter Netzwerk-Hardware und insbesondere hinsichtlich effizienter Netzteile

- Wähle Sie Equipment mit Power Management Funktionen und vergleichen Sie den Stromverbrauch verschiedener Produkte hinsichtlich der Idle und Standby Modi
- Vergleichen Sie die Kosten und Energieeffizienz von Netzwerksystemen verschiedener Hersteller
- Fordern Sie von ihrem Lieferanten Produktinformationen hinsichtlich:
 - Gesamtenergieeffizienz (z.B. auf Basis von ECR, TEER sowie verfügbar)
 - Effizienz und Modularität der Netzteile
 - Effizienz und Skalierbarkeit der Lüftereinheiten (drehzahlgeregelte Lüfter.)

Grundlage, dass die „resume-time-to-application“ ein Schlüsselkriterium für Einführung von „networked standby“ ist. Das Powermanagement des Netzwerk-Equipments hat einen direkten Bezug zu den Server- und Speichersystemen, welche es miteinander verbindet.

Im Fall, dass ein Cluster aus Server- und Speichermodulen in einen Standby- (Sleep-) Mode heruntergefahren wird, sollte es möglich sein, genauso Teile des Access-Switches in diesen Energiesparmodus zu fahren. Der kritische Faktor ist hierbei wieder die Latenzzeit und Zuverlässigkeit beim Reaktivieren des Systems. Mit der Einführung des IEEE 802.3az „Energy Efficient Ethernet“ Standards und des ECMA-393 Standards „proxZzy™ for sleeping hosts“ sind bereits zwei Energieeffizienzansätze im Bereich des Powermanagements für Netzwerk-Equipment verfügbar.

Power Supply Unit

Zuverlässigkeit und Effizienz der Netzteile sind bedeutende Aspekte für den Gesamtenergieverbrauch. Die Wandlungseffizienz großer Netzteile (>500W Ausgangsleistung) hat sich innerhalb der letzten Jahre auf Werte von über 85% bis über

90% erhöht. Angesichts der Tatsache, dass Core-Switches und -Router eine Leistungsaufnahme bis zu einigen KW haben, führen selbst kleine Verbesserungen der Wandlungseffizienz (selbst bei nur 1%) zu deutlichen Energieeinsparungen. Es ist zu beachten, dass die Produktspezifikationen der Hersteller nicht zwingend Informationen über die Effizienz der Netzteile enthalten.

4.2.5 Floor-level Switching

Es existieren zwei grundlegende Arten der Switch-Verteilungen für die Boden- bzw. Applikations-ebene: End-of-Row und Top-of-Rack. End-of-Row (EoR) Switching beschreibt den konventionellen Ansatz, wobei ein großer Switch ein oder mehrere Racks vernetzt. Aus der Perspektive der Energieeffizienz müssen in Bezug auf EoR zwei Punkte berücksichtigt werden:

- Vorteil: zentralisiertes Switching mit Vorteilen bei der Skalierbarkeit und der Energieeffizienz gegenüber suboptimalen ToR-Lösungen
- Nachteil: erheblicher Verkabelungsaufwand mit potentiell Mehraufwand für die Luftstromführung der Kühlung

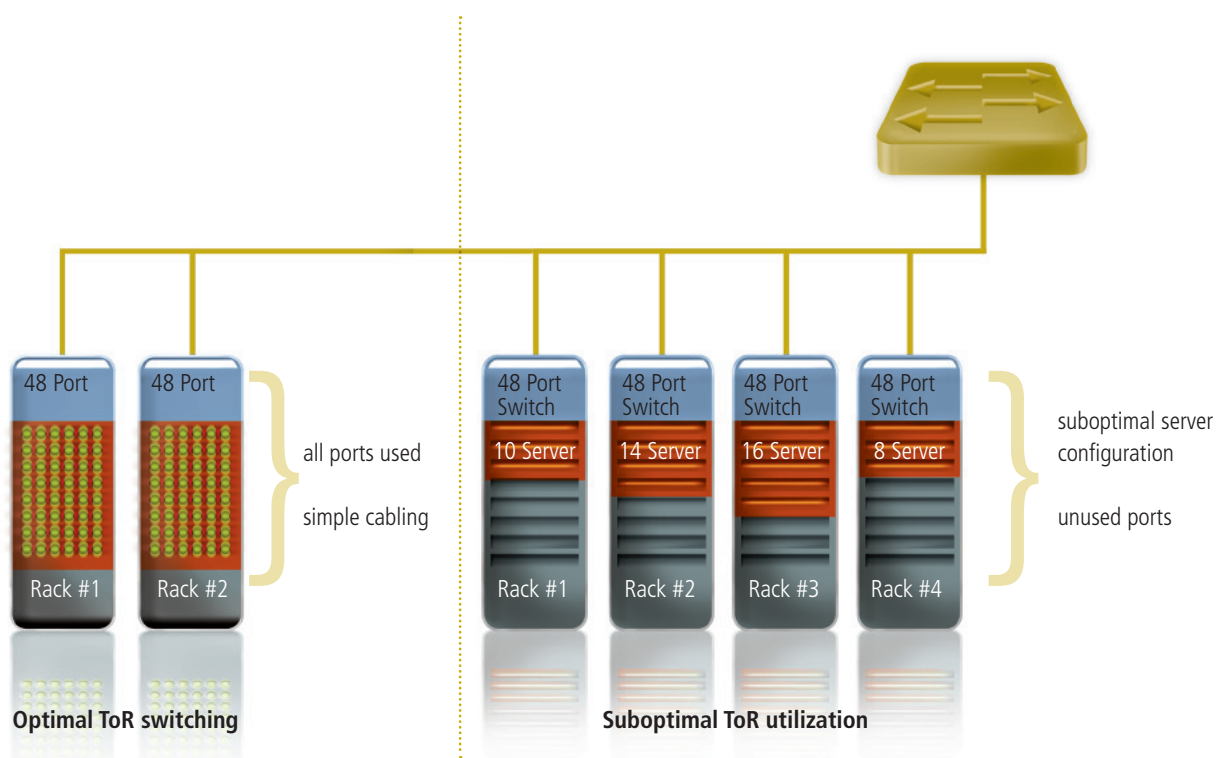


Abb. 4.4: Floor Level Top-of-Rack Switching

Top-of-Rack (ToR) beschreibt ein System, bei dem die Switches in jedem Rack integriert sind. Dieses Konzept gewährleistet geringe Latenzzeiten und einen hohen Datendurchsatz. Die Vor- und Nachteile der ToR-Lösung hinsichtlich Energieeffizienz sind:

- Vorteil: dezentrales Switching für Serverumgebungen mit einer hohen Gerätedichte (I/O Konsolidierung), welcher den Verkabelungsaufwand reduziert. Die kurze Verkabelungslänge zwischen Server und Switch verbessert den Datendurchsatz und reduziert somit den Energieverbrauch für die Übertragung.
- Nachteil: Wenn die ToR-Lösung in Umgebungen mit geringer Serverdichte (geringe Serveranzahl im Rack) eingesetzt wird, ist das System schnell überdimensioniert. Die Energieeffizienz ist geringer durch die suboptimale Ausnutzung der verfügbaren Ports.

Abschließend haben ToR-Systeme Effizienzvorteile, sofern sie in entsprechend dimensionierten Umgebungen verwendet werden. Abbildung 4.4 zeigt ein solches ToR-Konzept und dessen angemessene Auslastung.

Literaturempfehlungen

Hintemann R. (2008): Energy Efficiency in the Data centre, A Guide to the Planning, Modernization and Operation of Data centres, BITKOM, Berlin, online available:

http://www.bitkom.org/de/publikationen/38337_53432.aspx

EC JRC ISPRA (2011): Best Practices for the EU Code of Conduct on Data centres

European Commission (2011), EC Joint Research Centre, Ispra, online available:

http://re.jrc.ec.europa.eu/energyefficiency/html/standby_initiative_data_centres.htm

Juniper (2010): Government Data centre Network Reference Architecture, Using a High-Performance Network Backbone to Meet the Requirements of the Modern Government Data centre

Juniper (2010), Juniper Networks, Inc., Sunnyvale, available online:

<http://www.buynetscreen.com/us/en/local/pdf/reference-architectures/8030004-en.pdf>

Literaturzitate

[1] **Enterasys (2011):** Data centre Networking – Connectivity and Topology Design Guide; Inc Enterasys Networks, Andover.

[2] **Lippis (2011):** Open Industry Network Performance & Power Test Industry Network Performance & Power Test for Private and Public Data centre Clouds Ethernet Fabrics Evaluating 10 GbE Switches; Lippis Enterprises, Inc, Santa Clara.

[3] **Cisco (2008):** Converging SAN and LAN Infrastructure with Fibre Channel over Ethernet for Efficient, Cost-Effective Data centres; Intel, Santa Clara.

[4] **Emulex (2008):** Sheraton Case Study. Virtual Fabric for IBM BladeCentre Increases Server Bandwidth, Reduces Footprint and Enables Virtualization for High-performance Casino Applications; Emulex, Costa Mesa 2010.

[5] **Blade.org (2008):** Blade Platforms and Network Convergence; Blade.org, White Paper 2008.

5 Kühlung und Stromversorgung in Rechenzentren und Serverräumen

Andrea Roscetti, Politecnico di Milano, Thibault Faninger, Bio Intelligence Service

Kühlung kann bis zu 50% des gesamten Energieverbrauchs in Serverräumen und Rechenzentren ausmachen. Konzepte für energieeffizientes Kühlen sind daher für kleine wie auch größere IT-Systeme wichtig.

Im folgenden Kapitel wird eine Reihe genereller Möglichkeiten zur Reduktion des Energieverbrauches für Kühlung in Rechenzentren und Serverräumen im Überblick dargestellt.



Abb. 5.1: Split-System



externer Teil (Quelle: Daikin)

5.1 Kühlung in Serverräumen

Serverkammern oder kleine Serverräume sind üblicherweise mit Komfortkühlungssystemen¹ ausgestattet. Kleine Serverräume verfügen typischerweise über 1–5 Serverracks, mit einem Gesamtleistungsbedarf von maximal 20 kW.

5.1.1 Split-Systeme und tragbare Systeme

Split-Kühlsysteme werden üblicherweise in kleinen Serverräumen verwendet. Die Kühlleistung dieser Technologie reicht von 1 bis 100 kW. Im Allgemeinen haben Split-Kühlsysteme² folgende Vorteile:

- Geringe Investitionskosten
- Einfaches Design und einfache Installation
- Geringer Platzbedarf für die Installation (üblicherweise an der Wand)
- Lassen sich in fast allen Situationen/Umgebungen installieren.
- Einfache und rasche Wartung und Austausch der Systeme

Andererseits müssen folgende Nachteile berücksichtigt werden:

- Geringe Gesamteffizienz, vor allem bei kleinen, älteren oder überdimensionierten Systemen
- Geringe Feuchtigkeitskontrolle
- Begrenzte Länge und Höhe der Verrohrung zwischen den externen und internen Gerätekomponenten

Tragbare Systeme können unter anderem angebracht werden, um Hot-Spots zu vermeiden. Die Technologie bietet folgende Vorteile:

- Geringe Investitionskosten
- Einfache Installation
- Geringer Platzbedarf
- Einfache und rasche Wartung und Austausch des Systems

1) Komfortkühlungssystemen/HVAC: Heating, ventilation and air conditioning

2) Split-Kühlsysteme/DX: direct expansion

Folgende Nachteile sind zu berücksichtigen:

- Geringe Gesamteffizienz: Mobile Systeme der Klasse A sind weniger effizient als Split-Systeme der Klasse D.
- Geringe Feuchtigkeits- und Temperaturkontrolle der Kühlung
- Installation nur möglich, wenn die Luft nach draußen abgeleitet werden kann.

5.1.2 Maßnahmen zur Optimierung der Energieeffizienz in Serverräumen

Bei kleineren Serverräumen ist es gängige Praxis, die Kühlung überzudimensionieren. Um eine Überdimensionierung der Kühlung bei gut isolierten Serverräumen zu vermeiden, gilt die Faustregel, dass die Kühlleistung 120% der gesamten installierten IT-Leistung nicht übersteigen sollte.

Beim Kauf neuer Geräte bis zu 12 kW Kühlleistung kann das EU-Energielabel als Orientierungshilfe für die Auswahl energieeffizienter Ausstattung herangezogen werden. Dabei ist ein hoher EER³ bzw. SEER⁴ und Effizienzklasse A, oder darüber, die richtige Wahl. Der SEER und der entsprechend dem Label geschätzte Energieverbrauch (in kWh/a) sind die wichtigsten Vergleichskriterien. Tabelle 5.1 zeigt die Effizienz der derzeit besten Technologie am Markt.

Das neue Label wird mit einer Übergangsperiode bis 1. Jan 2013 implementiert. Bis dahin kann das Label von Herstellern bereits verwendet werden, ist aber noch nicht verpflichtend. In der Übergangsphase kann auch das alte Label für Klimaanlage noch zur Anwendung gelangen (2002/31/EC).



Abb. 5.2: **Energielabel für Klimageräte** (Quelle: Entwurfsdokumente zur neuen Labelling-Verordnung für Klimageräte)

Tab. 5.1: **State-of-the-art Effizienzwerte für kleine Kühlsysteme <12 kW** (Quelle: Entwurfsdokumente zur neuen Ecodesign-Verordnung für Klimageräte)

Benchmarks für Klimageräte		
Splitgeräte	Zweischlauchgeräte	Einschlauchgeräte
SEER	SEER	SEER
8,50	3.00	3.15

3) Energy Efficiency Ratio: Kälteleistung zur Kühlleistung für einen definierten Betriebspunkt

4) Seasonal EER: Jahresbezogene Energieeffizienz an einem bestimmten Standort

BEST PRACTICE EMPFEHLUNGEN

Bestehende Serverräume

- Verringern Sie Sonneneinstrahlung, Wärmeübertragung und Lüftungsverluste zu anderen Räumen bzw. von oder nach draußen.
- Kontrollieren Sie die Umweltbedingungen: die Temperatur der Zuluft zur IT sollte mindestens 18 bis 27 °C betragen, empfohlen wird jedoch 24 bis 27 °C.
- Überprüfen Sie die Isolierung der Verrohrung/Schächte (Kalt- und Warmluft/Wasser/Flüssigkeit)
- Überlegen Sie den Ersatz veralteter oder weniger effizienter Komponenten des Kühlsystems
- Kontrollieren Sie das Layout des installierten Kühlsystems (z.B. Distanz zwischen Kühlsystem und Lasten).
- Drehen Sie das Licht ab und entfernen Sie andere mechanische oder elektrische Lasten und Wärmequellen, sofern möglich.

Neue Serverräume

- Verwenden Sie ev. Präzisionskühlsysteme (um Wärme effektiv von der IT zu entfernen und eine zu hohe Entfeuchtung zu vermeiden).
- Definieren und überprüfen Sie die Raum- und IT Charakteristika, beachten Sie dabei Platzbeschränkungen und den Abstand zwischen den Wärmelasten und den externen Kühlaggregaten.
- Verwenden Sie keine mobilen Kühlgeräte oder Split-Geräte mit niedrigem EER (Achtung: mobile Systeme der Klasse A sind weniger effizient als Split-Systeme der Klasse D)
 - Entscheiden Sie sich für die höhere Energielabel-Klasse
 - Maximieren Sie die Kühleffizienz (SEER), s. BAT Tabelle.
- Prüfen Sie den Einsatz von Free Cooling.

5.2 Kühlung in mittleren und großen Rechenzentren

5.2.1 Generelle Aspekte

Die IT-Hardware wird normalerweise in Reihen angeordnet und zwar so, dass jeweils Kalt- und Warmgänge entstehen. Die Luftansaugung erfolgt vom kalten Gang her. Kühlluft wird dem kalten Gang zugeführt, strömt durch das Equipment und wird im heißen Gang abgeleitet.

Ein entsprechend effizientes Luftstrom-Design ist wesentlich. Typische Maßnahmen dafür sind beispielsweise die Warmgang- oder Kaltgangeinhausung (siehe [2]). Weitere zentrale Aspekte sind die Lufttemperatur der Zuluft und die Strömungs-

geschwindigkeit (siehe folgende Abschnitte).

Wenn unterschiedliches Equipment mit unterschiedlichen Betriebsbedingungen oder Luftströmungsrichtungen im selben Raum installiert werden muss, sollte jeweils ein eigener Bereich geschaffen werden. Im Weiteren sollte man vorzugsweise getrennte Steuerungen für die Kühlung einsetzen, um Ineffizienzen zu vermeiden, die aus niedrigeren Temperatureinstellungen oder unzureichender Steuerung des Luftstroms resultieren. Details s. Referenz [1].

5.2.2 Temperatur- und Feuchtigkeiteinstellungen

Rechenzentren sollten unter den gegebenen klimatischen Bedingungen auf höchstmögliche Effizienz ausgelegt und betrieben werden. Nach ASHRAE liegt der "empfohlene" Bereich für die Ansaugtemperatur bei den Servern (nicht Raumtemp.) zwischen 18 °C und maximal 27 °C. Dieser "empfohlene" Bereich entspricht jenen Bedingungen, für die ein höchst zuverlässiger Betrieb des IT Equipments bei gleichzeitiger Effizienz gewährleistet ist. Die relative Feuchtigkeit sollte bei maximal 60 % und der Taupunkt zwischen 5,5 und 15 °C liegen. ASHRAE definiert überdies auch einen "erlaubten" Bereich. Um Economizer bestmöglich zur Energieeinsparung auszunutzen, können die Ansaugtemperaturen damit für beschränkte Perioden auch über den oben spezifizierten Bereich auf bis zu 32 °C angehoben werden (siehe auch Free Cooling).

ASHRAE betont, dass der Betrieb außerhalb des empfohlenen Bereiches für kurze Perioden ohne Einschränkung der Betriebssicherheit möglich ist.

Bei höheren Temperaturen jenseits 30°C kann der steigende Energieverbrauch der eingebauten Ventilatoren bei Servern und anderem IT-Equipment jedoch die Effizienzgewinne im externen Kühlsystem übertreffen. So kann sich die Leistungsaufnahme der Server um ca. 4 bis 8% (7 - 20%) erhöhen, wenn die Ansaugtemperatur beispielsweise auf 30°C (35°C) angehoben wird (bezogen auf einen Betrieb bei 15 °C) (s. Lit.). Zu niedrige Temperatureinstellungen verschwenden Energie durch Überkühlung.

Höhere Temperatureinstellungen erfordern eine optimierte Lüftung, um die Vermeidung von lokalen Hotspots sicherzustellen. Bei großen Wärmedichten (z.B. 25 kW pro Rack), ist eine traditionelle Raumkühlung nicht mehr ausreichend (siehe [3], [4], [5]). In diesem Fall kann eine spezielle in Racks oder Reihen angeordnete Kühlung erforderlich sein.

BEST PRACTICE EMPFEHLUNGEN

Kühlsystem-Management:

- Regulieren Sie die Umgebungsbedingungen auf geeignetem Niveau (Sollwerte, Zeitplan, Position und Anzahl der Sensoren).
- Ersetzen Sie veraltete oder weniger effiziente Komponenten des Kühlsystems durch effizientere (vergleichen Sie die Effizienzklasse vorhandener Systeme).
- Überprüfen Sie die Isolierung von Rohrleitungen und Schächten.
- Platzieren Sie CRAC am Ende des Warmganges (die Kühlaggregate sind senkrecht/im rechten Winkel zu den heißen Gängen zu montieren).
- Separieren Sie Equipment mit unterschiedlichen Luftstrom- oder Lufttemperaturerfordernissen.
- Luftströme:
 - Platzieren Sie die Luftzufuhr (durchbrochene Bodenfliesen oder Luftverteiler) nur in den Kaltgängen in der Nähe zum IT Equipment.
 - Installieren Sie Luftstrombarrieren in Form einer Kaltgang- oder Warmgang-Einhausung, um die Vermischung von heißer Abluft mit kühler Zuluft zu verhindern.
- Verkabelung:
 - Verwenden Sie Overhead-Cable-Trays.
 - Kontrollieren Sie die Anordnung und die Abdichtung von offenen Kabeln und Bodenfliesen.

Kriterien für die Auswahl neuer energieeffizienter Kühlsysteme:

- Vergleichen Sie die Effizienz der Kühlaggregate (s. Referenzen Kühlerfordernisse).
- Vergleichen Sie die unterschiedlichen Luftstromdesignmöglichkeiten (Kalt/Warmgänge, Doppelboden/ Doppeldecken-Konzepte).
- Überprüfen Sie den Einsatz von:
 - Rack-basierter Kühlung (insbesondere für High-density Systeme)
 - Free Cooling (direkt/indirekt)
 - Freie Wasserkühlung (Free Watercooling)
 - Installation von Luft-Wasser-Wärmetauschern (Flüssigkühlung; direkt/indirekt)
 - Abwärmerückgewinnung/nutzung
- Richten Sie ein modulares Kühlsystem ein (angepasst an das IT Designkonzept).
- Verwenden Sie Simulationssoftware (Computational Fluid Dynamics (CFD) für die Optimierung des Kühlprozesses.

5.2.3 Komponenteneffizienz – Kälteaggregat, Ventilatoren, Lüftungsgeräte

Luftkühlung und Luft-Wasser-Wärmetauscher unterscheiden sich im Hinblick auf ihre EER (Energy Efficiency Ratio⁶), die für Wassersysteme typischerweise ca. 3,5 und für Luftsysteme ca. 2,5 beträgt. Der europäische Verband Eurovent stellt Daten zur Verfügung, die einen Vergleich der charakteristischen Effizienz zahlreicher Kühl- und Ventilations-systeme oder -komponenten ermöglichen (www.eurovent-certification.com). Wassergekühlte Kühlaggregate sind gegenüber luftgekühlten Aggregaten und DX dank ihrer höheren thermodynamischen Effizienz zu bevorzugen. Es sollte die Möglichkeit zur Verringerung der Kondensationstemperatur oder zur Erhöhung der Verdampfungstemperatur überprüft werden. Die Verringerung des Delta-T zwischen diesen Temperaturen bedeutet, dass

weniger Arbeit für den Kühlzyklus erforderlich ist und daher die Energieeffizienz erhöht wird. Die Temperaturen sind von den erforderlichen internen Lufttemperaturen abhängig (s. Temperatur- und Feuchtigkeitseinstellungen).

Die Effizienz der Ventilatoren hängt primär von der Motoreffizienz ab. Der Einsatz von Ventilatoren mit konstanter Drehzahl verbraucht viel Strom und erschwert das effiziente Management der Temperatur im Doppelboden. Ventilatoren mit variabler Geschwindigkeit sind insbesondere im Falle von hoher Redundanz im Kühlsystem und variabler IT Last sehr effektiv. Ventilatoren können durch die Rücklufttemperatur oder den Kaltluftkammerdruck gesteuert werden.

5.2.4 Free Cooling

Free Cooling ist eine Technik, die für die Kühlung im Vergleich zu den im RZ geforderten Bedingungen kühlere Außenluft oder Wasser verwendet. Je niedriger die durchschnittliche Außentemperatur über das Jahr ist, desto günstiger sind die Möglichkeiten für Free Cooling. Die klimatischen Bedingungen definieren damit die wirtschaftliche Effizienz und die Amortisationsdauer entsprechender Maßnahmen. Reines Free Cooling ist dann möglich, wenn die Differenz zwischen der Rücklauf-temperatur des Wassers und der Umgebungstemperatur größer als 11 K ist. Folgedessen sind die Energieeinsparungen umso größer, je höher die Zulufttemperatur bei den Servern eingestellt wird. Wenn die Konzeption des Kühlsystems eine höhere Serverraumtemperatur erlaubt, kann Free Cooling über längere Zeiträume im Jahr verwendet werden. Der Einsatz von Free Cooling erfordert vorab eine Machbarkeits- und Wirtschaftlichkeitsüberprüfung. Zur Abschätzung von zu erwartenden Einsparungen können auch entsprechende Evaluierungstools eingesetzt werden (siehe auch The Green Grid).

Spezifische Hinweise zu Free Cooling finden Sie unter "Literaturempfehlungen".

6) Energy Efficiency Ratio: Verhältnis zwischen abgegebener Kühlleistung und elektrischer Leistungsaufnahme an einem bestimmten Betriebspunkt (Innen- und Aussen-Temperatur- und Feuchtigkeitsbedingungen)

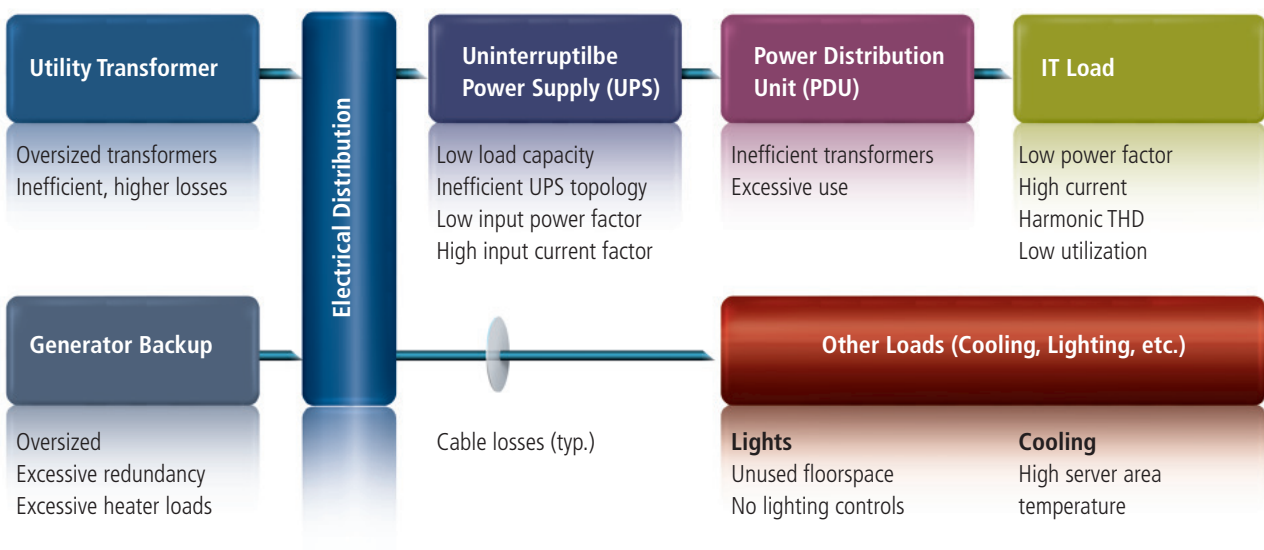


Fig. 5.3: Elektrische Infrastrukturkomponenten und Ineffizienzen im Rechenzentrum (ASHRAE: Save Energy Now Presentation Series, 2009).

5.2.5 Rack-Kühlung / Reihen-Kühlung

Wenn die Wärmedichte pro Rack sehr hohe Werte erreicht (z.B. 15-25kW pro Rack bei dichten Blade-Systemen), reicht traditionelle Kühlung nicht mehr aus und es sind spezielle Konzepte erforderlich (siehe Literatur).

5.3 Stromversorgung und USV in Rechenzentren

Abb. 5.3 zeigt die typische Stromumwandlung und -verteilung in Rechenzentren. Im weiteren sind häufige Ursachen für Ineffizienz für alle Komponenten dargestellt.

Systeme zur unterbrechungsfreien Stromversorgung (USV) bieten oftmals großes Potenzial für Energieeinsparungen.

USVs bieten neben ihrer primären Funktion, nämlich kurzfristig Strom zur Verfügung zu stellen, wenn die normale Stromzufuhr ausfällt, auch verschiedene Funktionen zur Korrektur von Frequenzen, Überspannungen etc. Es sind je nach gewünschter Anwendung drei Haupttopologien für USVs erhältlich:

- USV für Bereitschaftsbetrieb (passiver Standby): auch als spannungs- und frequenzabhängig bezeichnet, dient ausschließlich dazu, Stromunterbrechungen zu überbrücken (Stromausfall, Spannungseinbrüche, Überspannungen). Bei normaler Stromversorgung hat die USV keine Interaktion mit der Stromversorgung. Wenn die Stromzufuhr außerhalb der eingestellten USV-Lasttoleranz liegt, wird der Last über einen Bypass der Stromversorgung Strom zur Verfügung gestellt. Diese Topologie ist bei Anwendungen mit geringem Stromverbrauch eher üblich.
- USV für linearen Bereitschaftsbetrieb (line interactive): auch als spannungsunabhängig bezeichnet, bietet zusätzlich Schutz für die Last, indem sie die Frequenz innerhalb optimaler Grenzen regelt. Insbesondere bietet sie Schutz vor kontinuierlicher Unterspannung oder Überspannung. Diese Topologie ist üblicherweise nicht über 5.000 VA in Verwendung [7].

- USV im Online-Betrieb mit Doppelwandlung (double-conversion): auch als spannungs- und frequenzunabhängig bezeichnet, kann die Last vor den negativen Effekten von Spannungs- oder Frequenzänderungen schützen, indem sie den Versorgerstrom reguliert, bevor dieser die Last erreicht. Diese Topologie ist selten bei Lasten unterhalb von 750 VA.

Jede Topologie hat ihre Vor- und Nachteile. Im Bereich 750 VA–5.000 VA tendieren line-interactive USVs zu längeren Betriebsdauern und höherer Zuverlässigkeit mit geringeren Gesamtbetriebskosten, während double conversion online USVs weniger Platz brauchen und die Frequenz regulieren können.

Es gibt unterschiedliche Energiespeichermechanismen für USV-Lösungen:

- Batterien: die elektrische Energie durch die Umwandlung chemischer Energie speichern;
- Schwungräder: die kurzfristige Energiespeicherung in Form einer sich drehenden massiven Platte bieten.

Es stehen zwei Möglichkeiten zur Verfügung, um den Strom an die Last zu übermitteln:

- Statische USV (keine beweglichen Teile im Strompfad, ausgenommen die Ventilatoren für die Kühlung): wandelt Wechselstrom in Gleichstrom und anschließend wieder in Wechselstrom für die Netzteile in den Servern.
- Schwungrad: überträgt Strom via Motor/Generator und wird zur Überbrückung kurzfristiger Stromausfälle, Spannungseinbrüche etc. eingesetzt.

Tab. 5.2: Charakteristische Effizienzen von USV-Topologien

USV Topologie	Effizienz bei 25% Last	Effizienz bei 50% Last	Effizienz bei 75% Last	Effizienz bei 100% Last
Double-conversion	81–93%	85–94%	86–95%	86–95%
Line-interactive	n.a.	97–98%	98%	98%

Tab. 5.3: Minimalanforderungen zur Durchschnittseffizienz der USV gemäß EnergyStar Kriterien (P ist Power in Watt (W), ln ist der natürliche Logarithmus)

Minimale Durchschnittseffizienz (EffAVG_MIN),				
USV Klasse	Ausgangsleistung	Kriterium nach ENERGY STAR Test Methode		
		VFD	VI	VFI
Rechenzentrum	P > 10 kW	0,97	0,96	$0,0058 \times \ln(P) + 0.86$

Die USV-Energieverluste entstehen aufgrund von Ineffizienzen bei der Stromumwandlung (in Ladegerät und Umwandler) sowie Batterieladeverlusten oder Energieverlusten in Intertialsystemen (Schwungräder). Die elektrischen Verluste (und die Wärmeerzeugung) sind bei double-conversion USVs bedeutender (Verluste bei Rectifier, Inverter, Filter, Vernetzung) als bei line-interactive und Standby-USVs. Bei manchen Anwendungen können Gleichstrom-USVs (Gleichrichter) und kombinierte Gleichstrom-Wechselstrom-USVs eingesetzt werden und Verluste beim Wechselrichter und beim Gleichrichter vermeiden.

Die meisten USV-Hersteller geben die USV-Effizienz bei 100% Last an. Unter Teillastbedingungen fällt die Effizienz jedoch deutlich (die meisten USV arbeiten bei 80% und im Falle von Redundanz kann die Last sogar auf 50% und weniger fallen). Bei Lasten von 50% oder weniger laufen sowohl moderne als auch ältere USV-Systeme weniger effizient, mit signifikanten Effizienzverlusten bei Lasten unter 20%. Best-Practice würde bedeuten,

BEST PRACTICE EMPFEHLUNGEN

Für neue Installationen

- Überprüfen Sie Ihren Bedarf und dimensionieren Sie die USV Systeme richtig. Schätzen Sie ab, ob multiple oder modulare USV möglich sind bzw. skalierbare und ausbaufähige Lösungen. Berücksichtigen Sie Backup Zeit der Batterie, Kosten, Größe, Anzahl der Ausgänge etc.
- Analysieren Sie die USV Technologie und Effizienz. Berücksichtigen Sie die Teillasteffizienz der USV
- Wählen sie die richtige Topologie für die Stromversorgungssysteme
- Wählen sie USV Systeme in Übereinstimmung mit dem EU Code of Conduct for UPS oder Energy Star

Für die Optimierung

- Analysieren Sie die USV Technologie und Effizienz
- Bewerten Sie Optionen und Vorteile eines Ersatzes mit neuem Equipment
- Bewerten Sie die Kosten und Vorteile von Redundanz

die USV-Nennleistung so nahe wie möglich an die IT-Lasten des Rechenzentrums anzugleichen. Für die effiziente Dimensionierung der USV-Kapazität sind skalierbare USV-Lösungen erhältlich. Die Minimalanforderungen an die Effizienz von USVs sind im EU Code of Conduct for UPS (Edition 2011) und in den Anforderungen des Energy Star Programms spezifiziert. Neue Energy Star Energieeffizienzanforderungen für Gleichstrom- und Wechselstrom USV (Entwurfs-Version 2011) werden derzeit ausgearbeitet (Tab. 5.3).

In die neuen Energy-Star-Anforderungen werden gegebenenfalls auch Anforderungen für multi-mode USVs aufgenommen. Dieser USV-Typ arbeitet mit mehr als einem Set von Input-Dependency Charakteristika (z.B. als VFI oder VFD). Multi-mode USV kann in effizienteren, weniger geschützten Betriebsmodi betrieben werden und bei Bedarf in weniger effiziente, geschütztere Modi wechseln. Auf diese Weise sind signifikante Energieeinsparungen möglich.

Literaturempfehlungen

ASHRAE (2011): Thermal Guidelines for Data Processing Environments – Expanded Data centre Classes and Usage Guidance – ASHRAE, 2011, online available at: <http://tc99.ashraetcs.org/documents/ASHRAE%20Whitepaper%20-%202011%20Thermal%20Guidelines%20for%20Data%20Processing%20Environments.pdf>

EU Code of conduct for data centres (2009): Full list of identified best practice options for data centre operators as referenced in the EU Code of Conduct:

<http://re.jrc.ec.europa.eu/energyefficiency/pdf/CoC/Best%20Practices%20v3.0.1.pdf>

The Green Grid (2011): Evaluation tool for free cooling.

http://cooling.thegreengrid.org/europe/WEB_APP/calc_index_EU.html

ENERGY STAR (2011): UPS efficiency

http://www.energystar.gov/index.cfm?c=new_specs_uninterruptible_power_supplies

The Green Grid (2011): Evaluation tool for power supply systems

<http://estimator.thegreengrid.org/pcee>

High Performance Buildings: Data centres Uninterruptible Power Supplies (UPS)

http://hightech.lbl.gov/documents/UPS/Final_UPS_Report.pdf

EU CODE of CONDUCT (2011): EU code of conduct on Energy Efficiency and Quality of AC Uninterruptible Power Systems (UPS):

http://re.jrc.ec.europa.eu/energyefficiency/html/standby_initiative.htm

Zitierte Literatur

- [1] **ASHRAE:** Save Energy Now Presentation Series, 2009.
- [2] **Niemann, J. et al. (2010).** Hot-Aisle vs. Cold-Aisle Containment for Data centres; APC by Schneider Electric White Paper 135, Revision 1.
- [3] **Rasmussen, N. (2010).** An improved architecture for High-efficiency High-density data centres; APC by Schneider Electric White Paper 126, Revision 1.
- [4] **Blough, B. (2011).** Qualitative analysis of cooling architectures for data centres; The Green Grid White Paper #30.
- [5] **Bouley, D. and Brey, T. (2009).** Fundamentals of data centre power and cooling efficiency zones; The Green Grid White Paper #21.
- [6] **Rasmussen, N. (2011).** Calculating Total Cooling Requirements for Data centres; APC by Schneider Electric White Paper 25, Revision 3.
- [7] **ENERGY STAR Uninterruptible Power Supply Specification Framework (2010).** Available at: www.energystar.gov/ia/partners/prod_development/new_specs/downloads/uninterruptible_power_supplies/UPS_Framework_Document.pdf
- [8] **Ton, M. and Fortenbury B. (2008).** High Performance Buildings: Data centres - Uninterruptible Power Supplies. Available at http://hightech.lbl.gov/documents/UPS/Final_UPS_Report.pdf
- [9] **Samstad, J. and Hoff M.;** Technical Comparison of On-line vs. Line-interactive UPS designs; APC White Paper 79. Available at <http://www.apcdistributors.com/white-papers/Power/WP-79%20Technical%20Comparison%20of%20On-line%20vs.%20Line-interactive%20UPS%20designs.pdf>

PrimeEnergyIT

EFFICIENT DATA CENTERS



Partner



Unterstützt durch



Kontakt: Austrian Energy Agency | Dr. Bernd Schäppi | Mariahilfer Straße 136 | 1150 Wien, Austria
T +43 1 586 15 24 | bernd.schaepfi@energyagency.at | www.efficient-datacenters.eu