

# Kapazitäts- und Performance-Management: Whitepaper zu Best Practices

## Inhalt

[Einleitung](#)

[Kapazitäts- und Performance-Management - Überblick](#)

[CPU](#)

[Rückwandplatine oder E/A](#)

[Arbeitsspeicher](#)

[Schnittstellen- und Rohrgrößen](#)

[Warteschleifen, Latenz und Jitter](#)

[Geschwindigkeit und Entfernung](#)

[Anwendungseigenschaften](#)

[Best Practices für Kapazitäts- und Performance-Management](#)

[Servicelevel-Management](#)

[Netzwerk- und Anwendungs-Was-wäre-wenn-Analyse](#)

[Baselining und Trending](#)

[Ausnahmeverwaltung](#)

[QoS-Management](#)

[Erfassen und Berichten von Kapazitätsinformationen](#)

[Ermitteln Sie Ihre Anforderungen](#)

[Definieren eines Prozesses](#)

[Kapazitätsbereiche definieren](#)

[Kapazitätsvariablen definieren](#)

[Daten interpretieren](#)

[Zugehörige Informationen](#)

## Einleitung

Hohe Netzwerkverfügbarkeit ist eine geschäftskritische Anforderung in großen Unternehmens- und Service-Provider-Netzwerken. Netzwerkmanager sehen sich mit wachsenden Herausforderungen bei der Bereitstellung höherer Verfügbarkeit konfrontiert, darunter ungeplante Ausfallzeiten, Fachkräftemangel, unzureichende Tools, komplexe Technologien, Geschäftskonsolidierung und konkurrierende Märkte. Das Kapazitäts- und Performance-Management unterstützt Netzwerkmanager dabei, neue Geschäftsziele zu erreichen und eine konsistente Netzwerkverfügbarkeit und -leistung sicherzustellen.

In diesem Dokument werden folgende Themen behandelt:

- Allgemeine Kapazitäts- und Leistungsprobleme, einschließlich Risiken und potenzielle Kapazitätsprobleme innerhalb von Netzwerken.
- Best Practices für Kapazitäts- und Performance-Management, darunter "Was-wäre-wenn"-Analysen, Baselining, Trends, Ausnahmemanagement und QoS-Management.

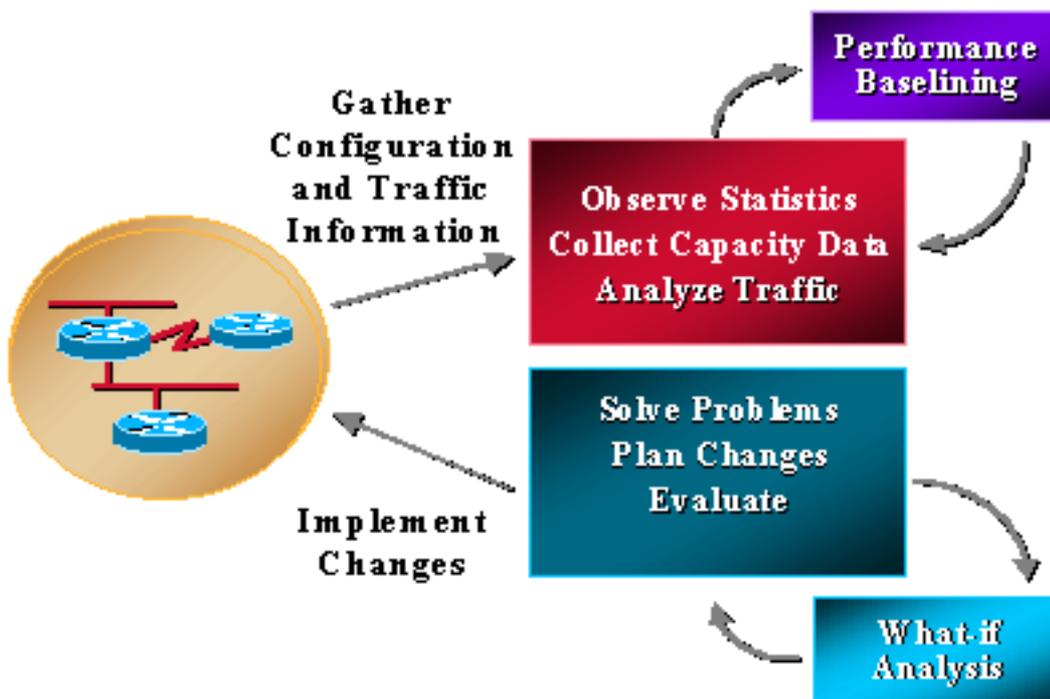
- Entwicklung einer Strategie für die Kapazitätsplanung, die die bei der Kapazitätsplanung verwendeten gängigen Techniken, Tools, MIB-Variablen und Schwellenwerte umfasst.

## Kapazitäts- und Performance-Management - Überblick

Bei der Kapazitätsplanung werden die erforderlichen Netzwerkressourcen ermittelt, um Performance- oder Verfügbarkeitsbeeinträchtigungen für geschäftskritische Anwendungen zu vermeiden. Performance-Management bezeichnet die Verwaltung der Reaktionszeit, Konsistenz und Qualität von Netzwerkservices für einzelne und allgemeine Services.

**Hinweis:** Leistungsprobleme hängen normalerweise mit der Kapazität zusammen. Anwendungen sind langsamer, da Bandbreite und Daten in Warteschlangen verbleiben müssen, bevor sie über das Netzwerk übertragen werden können. Bei Sprachanwendungen beeinflussen Probleme wie Verzögerungen und Jitter direkt die Qualität des Sprachanrufs.

Die meisten Unternehmen erfassen bereits einige kapazitätsbezogene Informationen und arbeiten konsequent daran, Probleme zu lösen, Änderungen zu planen und neue Kapazitäts- und Leistungsfunktionen zu implementieren. Unternehmen führen jedoch nicht routinemäßig Trendanalysen und Was-wäre-wenn-Analysen durch. Was-wäre-wenn-Analyse ist der Prozess der Bestimmung der Auswirkungen einer Netzwerkänderung. Unter Trending versteht man den Prozess der Durchführung einer Reihe von Problemen im Zusammenhang mit der Netzwerkkapazität und -leistung sowie die Überprüfung der Grundlagen für Netzwerktrends, um die zukünftigen Upgrade-Anforderungen zu ermitteln. Das Kapazitäts- und Leistungsmanagement sollte auch das Ausnahmemanagement umfassen, bei dem Probleme erkannt und behoben werden, bevor sich Benutzer einwählen, sowie das QoS-Management, bei dem Netzwerkadministratoren individuelle Probleme mit der Serviceleistung planen, verwalten und identifizieren. Die folgende Grafik zeigt die Prozesse des Kapazitäts- und Performance-Managements.



Das Kapazitäts- und Performance-Management unterliegt ebenfalls gewissen Einschränkungen, die typischerweise mit CPU und Arbeitsspeicher zusammenhängen. Folgende Bereiche könnten Anlass zur Sorge geben:

- CPU
- Rückwandplatine oder E/A
- Speicher und Puffer
- Schnittstellen- und Rohrgrößen
- Warteschleifen, Latenz und Jitter
- Geschwindigkeit und Entfernung
- Anwendungseigenschaften

Einige Verweise auf Kapazitätsplanung und Performance-Management erwähnen auch die so genannte "Datenebene" und die "Kontrollebene". Bei der Datenebene handelt es sich schlicht um Kapazitäts- und Leistungsprobleme, wenn die Daten das Netzwerk durchlaufen, während die Kontrollebene Ressourcen umfasst, die für die Aufrechterhaltung des ordnungsgemäßen Funktionierens der Datenebene erforderlich sind. Die Funktionalität der Kontrollebene umfasst Service-Overhead wie Routing, Spanning Tree, Schnittstellen-Keepalives und SNMP-Management des Geräts. Für diese Anforderungen der Kontrollebene werden CPU, Arbeitsspeicher, Pufferung, Warteschlangenverwaltung und Bandbreite genau wie für den Datenverkehr im Netzwerk benötigt. Viele der Anforderungen an die Kontrollebene sind auch für die Gesamtfunktionalität des Systems wichtig. Wenn sie nicht über die benötigten Ressourcen verfügen, fällt das Netzwerk aus.

## **CPU**

Die CPU wird in der Regel sowohl von der Kontroll- als auch von der Datenebene eines Netzwerkgeräts verwendet. Beim Kapazitäts- und Leistungsmanagement müssen Sie sicherstellen, dass das Gerät und das Netzwerk jederzeit über eine für den Betrieb ausreichende CPU verfügen. Eine unzureichende CPU kann ein Netzwerk häufig lahmlegen, da sich unzureichende Ressourcen auf einem Gerät auf das gesamte Netzwerk auswirken können. Unzureichende CPU kann auch die Latenz erhöhen, da die Daten warten müssen, um verarbeitet zu werden, wenn es kein Hardware-Switching ohne die Haupt-CPU gibt.

## **Rückwandplatine oder E/A**

Backplane oder I/O bezieht sich auf den gesamten Datenverkehr, den ein Gerät verarbeiten kann. Diese Angabe bezieht sich in der Regel auf die Busgröße oder die Backplane-Funktion. Eine unzureichende Backplane führt normalerweise zu Paketverlusten, die zu erneuten Übertragungen und zusätzlichem Datenverkehr führen können.

## **Arbeitsspeicher**

Der Arbeitsspeicher ist eine weitere Ressource, für die Anforderungen auf Datenebene und Kontrollebene gelten. Für Informationen wie Routing-Tabellen, ARP-Tabellen und andere Datenstrukturen ist Speicher erforderlich. Wenn nicht genügend Arbeitsspeicher zur Verfügung steht, können einige Vorgänge auf dem Gerät fehlschlagen. Der Vorgang kann je nach Situation die Prozesse auf der Kontroll- oder Datenebene beeinflussen. Wenn die Prozesse auf der Steuerungsebene fehlschlagen, kann sich das gesamte Netzwerk verschlechtern. Dies kann beispielsweise der Fall sein, wenn für die Routing-Konvergenz zusätzlicher Speicher erforderlich ist.

## **Schnittstellen- und Rohrgrößen**

Schnittstellen- und Leitungsgrößen beziehen sich auf die Datenmenge, die gleichzeitig über eine

Verbindung gesendet werden kann. Dies wird oft fälschlicherweise als die Geschwindigkeit einer Verbindung bezeichnet, aber die Daten werden nicht mit unterschiedlichen Geschwindigkeiten von einem Gerät zum anderen übertragen. Die Geschwindigkeit des Siliziums und die Hardwarefunktionen helfen dabei, die verfügbare Bandbreite auf Basis der Medien zu ermitteln. Darüber hinaus können Software-Mechanismen Daten drosseln, um bestimmte Bandbreitenzuweisungen für einen Service zu erfüllen. Dies wird in der Regel in Service-Provider-Netzwerken für Frame-Relay oder ATM angezeigt, die inhärent Geschwindigkeiten von 1,54 kbps bis 155 mbs und höher aufweisen. Wenn Bandbreitenbeschränkungen bestehen, werden Daten in eine Übertragungswarteschlange gestellt. Eine Übertragungswarteschlange kann über verschiedene Softwaremechanismen verfügen, um Daten innerhalb der Warteschlange zu priorisieren. Wenn sich jedoch Daten in der Warteschlange befinden, muss sie auf vorhandene Daten warten, bevor sie die Daten über die Schnittstelle weiterleiten kann.

## **Warteschleifen, Latenz und Jitter**

Warteschlangen, Latenz und Jitter wirken sich ebenfalls auf die Leistung aus. Sie können die Übertragungswarteschlange so einstellen, dass sie die Leistung auf verschiedene Weise beeinflusst. Wenn die Warteschlange beispielsweise groß ist, warten die Daten länger. Bei kleinen Warteschlangen werden Daten verworfen. Dies wird als Taildrop bezeichnet und ist für TCP-Anwendungen akzeptabel, da die Daten erneut übertragen werden. Die Sprach- und Videoübertragung funktioniert jedoch nicht gut, wenn Warteschlangen verloren gehen oder sogar eine erhebliche Warteschlangenlatenz auftritt, was ein besonderes Augenmerk auf die Bandbreite oder die Leitungsgröße erfordert. Eine Warteschlangenverzögerung kann auch bei Eingangswarteschlangen auftreten, wenn das Gerät nicht über ausreichende Ressourcen verfügt, um das Paket sofort weiterzuleiten. Dies kann auf CPU, Speicher oder Puffer zurückzuführen sein.

Latenz beschreibt die normale Verarbeitungszeit vom Zeitpunkt des Empfangs bis zur Weiterleitung des Pakets. Normale moderne Daten-Switches und Router haben unter normalen Bedingungen und ohne Ressourcenbeschränkungen eine extrem niedrige Latenz (< 1 ms). Moderne Geräte mit digitalen Signalprozessoren zur Umwandlung und Komprimierung analoger Sprachpakete können sogar bis zu 20 ms länger dauern.

Jitter beschreibt die Lücke zwischen Paketen für Streaming-Anwendungen, einschließlich Sprach- und Videoanwendungen. Wenn Pakete zu unterschiedlichen Zeiten mit unterschiedlichen Inter-Packet-Gap-Timings eintreffen, tritt starker Jitter auf, und die Sprachqualität verschlechtert sich. Jitter ist vor allem ein Faktor für Warteschlangenverzögerung.

## **Geschwindigkeit und Entfernung**

Geschwindigkeit und Entfernung sind ebenfalls ein Faktor für die Netzwerkleistung. Datennetze verfügen über eine konsistente Datenweiterleitungsgeschwindigkeit, die auf der Lichtgeschwindigkeit basiert. Das sind ungefähr 100 Meilen pro Millisekunde. Wenn eine Organisation international eine Client-Server-Anwendung ausführt, kann sie mit einer entsprechenden Verzögerung bei der Paketweiterleitung rechnen. Wenn Anwendungen nicht für die Netzwerkleistung optimiert sind, können Geschwindigkeit und Entfernung ein entscheidender Faktor für die Anwendungsleistung sein.

## **Anwendungseigenschaften**

Die Anwendungseigenschaften sind der letzte Bereich, der sich auf Kapazität und Leistung auswirkt. Probleme wie die geringe Fenstergröße, Anwendungs-Keepalives und die Menge der

über das Netzwerk gesendeten Daten im Vergleich zu den erforderlichen Daten können die Leistung einer Anwendung in vielen Umgebungen, insbesondere WANs, beeinträchtigen.

## Best Practices für Kapazitäts- und Performance-Management

In diesem Abschnitt werden die fünf wichtigsten Best Practices für das Kapazitäts- und Leistungsmanagement ausführlich erläutert:

- [Servicelevel-Management](#)
- [Was-wäre-wenn-Analyse von Netzwerk und Anwendungen](#)
- [Baselining und Trendanalyse](#)
- [Ausnahmeverwaltung](#)
- [QoS-Management](#)

### Servicelevel-Management

Das Service Level Management definiert und reguliert andere erforderliche Kapazitäts- und Performance-Management-Prozesse. Netzwerkmanager wissen, dass sie eine Kapazitätsplanung benötigen. Budgetäre und personelle Einschränkungen verhindern jedoch eine Komplettlösung. Das Service Level Management ist eine bewährte Methodik, die bei Ressourcenproblemen hilft, indem ein Leistungsumfang definiert und eine bidirektionale Verantwortlichkeit für einen mit diesem Leistungsumfang verbundenen Service geschaffen wird. Dies lässt sich auf zwei Arten erreichen:

- Erstellen Sie einen Service Level Agreement zwischen Benutzern und der Netzwerkorganisation für einen Service, der Kapazitäts- und Leistungsmanagement umfasst. Der Service umfasst Berichte und Empfehlungen zur Aufrechterhaltung der Servicequalität. Die Benutzer müssen jedoch bereit sein, den Service und alle erforderlichen Upgrades zu finanzieren.
- Die Netzwerkorganisation definiert ihren Service für das Kapazitäts- und Leistungsmanagement und versucht dann, diesen Service und die Upgrades von Fall zu Fall zu finanzieren.

In jedem Fall sollte die Netzwerkorganisation zunächst einen Service für Kapazitätsplanung und Leistungsmanagement definieren, der die Aspekte des Services berücksichtigt, den sie derzeit bereitstellen können, und der für die Zukunft geplant ist. Ein vollständiger Service würde eine Was-wäre-wenn-Analyse für Netzwerk- und Anwendungsänderungen, Baselining und Trendanalysen für definierte Leistungsvariablen, Ausnahmeverwaltung für definierte Kapazitäts- und Leistungsvariablen und QoS-Management umfassen.

### Netzwerk- und Anwendungs-Was-wäre-wenn-Analyse

Durchführung einer "Was-wäre-wenn"-Analyse von Netzwerk und Anwendungen, um das Ergebnis einer geplanten Änderung zu bestimmen. Ohne "Was-wäre-wenn"-Analyse gehen Unternehmen erhebliche Risiken ein, um den Erfolg und die allgemeine Netzwerkverfügbarkeit zu verändern. In vielen Fällen führten Netzwerkänderungen zu einem Zusammenbruch des Netzwerks, der viele Stunden Ausfallzeit verursachte. Darüber hinaus schlägt eine verblüffende Anzahl von Anwendungseinführungen fehl und verursacht Auswirkungen auf andere Benutzer und Anwendungen. Diese Ausfälle halten in vielen Netzwerkorganisationen an, können jedoch mit wenigen Tools und zusätzlichen Planungsschritten vollständig vermieden werden.

Normalerweise benötigen Sie einige neue Prozesse, um eine Was-wäre-wenn-Analyse durchzuführen. Der erste Schritt besteht darin, die Risikostufen für alle Änderungen zu ermitteln und eine gründlichere Was-wäre-wenn-Analyse für Änderungen mit höherem Risiko zu verlangen. Die Risikostufe kann ein Pflichtfeld für alle eingereichten Änderungen sein. Änderungen höherer Risikostufen würden dann eine definierte Was-wäre-wenn-Analyse der Änderung erfordern. Eine Was-wäre-wenn-Analyse des Netzwerks bestimmt die Auswirkungen von Netzwerkänderungen auf die Netzwerknutzung und Probleme mit Ressourcen auf der Steuerungsebene. Eine "Was-wäre-wenn"-Analyse der Anwendungen würde den Erfolg des Projekts, die Bandbreitenanforderungen und etwaige Probleme mit den Netzwerkressourcen bestimmen. Die folgenden Tabellen enthalten Beispiele für die Zuweisung von Risikostufen und entsprechende Testanforderungen:

Risikostufe	Definition	Planungsempfehlungen ändern
1	<ul style="list-style-type: none"> <li>• Hohe potenzielle Auswirkungen für eine große Anzahl von Benutzern (über 500) oder geschäftskritische Services aufgrund der Einführung neuer Produkte, Software, Topologie oder Funktionen</li> <li>• Die Änderung bringt eine erwartete Netzwerkausfallzeit mit sich.</li> </ul>	<ul style="list-style-type: none"> <li>• Validieren Sie das Labor der neuen Lösung. Die Laborvalidierung umfasst dokumentierte Tests und Validierungen von Lösungen sowie Was-wäre-wenn-Analysen, die die Auswirkungen auf die vorhandene Infrastruktur aufzeigen. Wir empfehlen Pilotprojekte. Für neue Lösungen muss ein Dokument zur Betriebsunterstützung ausgefüllt werden.</li> <li>• Cisco NSA Designprüfung durchführen.</li> <li>• Erstellen Sie einen Backup-</li> </ul>

		<p>Plan.</p> <ul style="list-style-type: none"> <li>• Erstellung eines Implementierungsplans</li> <li>• Änderungsprozess erstellen.</li> </ul>
2	<ul style="list-style-type: none"> <li>• Große potenzielle Auswirkungen auf eine große Anzahl von Benutzern (über 500) oder geschäftskritische Services aufgrund einer starken Zunahme des Datenverkehrs oder der Anzahl von Benutzern, Backbone-Änderungen oder Routing-Änderungen</li> <li>• Veränderungen können einige Ausfallzeiten erfordern.</li> </ul>	<ul style="list-style-type: none"> <li>• Durchführung einer Was-wäre-wenn-Analyse, um die Auswirkungen auf die vorhandene Umgebung zu ermitteln (sollte in der Laborumgebung durchgeführt werden).</li> <li>• Testen und Überprüfen von Routingänderungen auf Funktionalität</li> <li>• Erstellen Sie einen Backup-Plan.</li> <li>• Entwurfsüberprüfung für größere Änderungen am Routing oder am Backbone durchführen.</li> <li>• Erstellung eines Implementierungsplans</li> <li>• Änderungsprozess erstellen.</li> </ul>
3	<ul style="list-style-type: none"> <li>• Mittlere potenzielle Auswirkungen auf eine</li> </ul>	<ul style="list-style-type: none"> <li>• Durchführen einer</li> </ul>

	<p>geringere Anzahl von Benutzern oder Geschäftsservices aufgrund von nicht standardmäßigen Änderungen.</p> <ul style="list-style-type: none"> <li>• Umfasst neue Produkte, Software, Topologie, Hinzufügen von Funktionen oder neuen Benutzern, erhöhten Datenverkehr oder eine vom Standard abweichende Topologie.</li> <li>• Veränderungen können einige Ausfallzeiten erfordern.</li> </ul>	<p>technischen Analyse der neuen Lösung (erfordert möglicherweise eine Laborvalidierung)</p> <ul style="list-style-type: none"> <li>• Erstellung eines Implementierungsplans</li> <li>• Änderungsprozess erstellen.</li> </ul>
4	<ul style="list-style-type: none"> <li>• Geringere Auswirkungen auf Services oder Benutzer</li> <li>• Umfasst das Hinzufügen neuer Netzwerkmodule mit Standardvorlage, z. B. für die Erstellung von Switches oder Server-Hubs auf Routern.</li> <li>• Ermöglicht die Einrichtung neuer WAN-Standorte oder zusätzliche bewährte Access Services.</li> <li>• Alle Änderungen der Risikostufe 3 wurden in der Produktionsumgebung technisch nachgewiesen.</li> <li>• Veränderungen können einige Ausfallzeiten erfordern.</li> </ul>	<ul style="list-style-type: none"> <li>• Erstellung eines Implementierungsplans</li> <li>• Änderungsprozess erstellen.</li> </ul>
5	<ul style="list-style-type: none"> <li>• Keine Auswirkungen auf Benutzer oder Services.</li> <li>• Dies umfasst das Hinzufügen einzelner Benutzer zum Netzwerk sowie Änderungen an der Standardkonfiguration, z. B. Kennwort, Banner, SNMP oder andere</li> </ul>	<ul style="list-style-type: none"> <li>• Änderungsvorgang optional.</li> </ul>

	Standardkonfigurationsparameter. • Keine Ausfallzeiten.	
--	------------------------------------------------------------	--

Sobald Sie festgelegt haben, wo Sie die Was-wäre-wenn-Analyse benötigen, können Sie den Service definieren.

Sie können eine Netzwerk-Was-wäre-wenn-Analyse mit Modellierungstools oder mit einem Labor durchführen, das die Produktionsumgebung nachahmt. Modellierungstools sind davon abhängig, wie gut die Anwendung die Probleme mit den Gerätesressourcen versteht, und da die meisten Netzwerkänderungen neue Geräte sind, versteht die Anwendung die Auswirkungen der Änderung möglicherweise nicht. Die beste Methode besteht darin, eine Darstellung des Produktionsnetzwerks in einem Labor zu erstellen und die gewünschte Software, Funktion, Hardware oder Konfiguration unter Last mithilfe von Traffic-Generatoren zu testen. Das Austreten von Routen (oder anderen Steuerungsinformationen) aus dem Produktionsnetzwerk in das Labor verbessert ebenfalls die Laborumgebung. Testen zusätzlicher Ressourcenanforderungen mit verschiedenen Datenverkehrstypen, einschließlich SNMP, Broadcast, Multicast, verschlüsseltem oder komprimiertem Datenverkehr. Mit all diesen unterschiedlichen Methoden können Sie die Ressourcenanforderungen von Geräten in Situationen analysieren, in denen möglicherweise Stress auftritt, wie z. B. bei der Routenkonvergenz, beim Flapping von Verbindungen und beim Neustarten von Geräten. Zu den Problemen bei der Ressourcenauslastung gehören Ressourcenbereiche mit normaler Kapazität wie CPU, Arbeitsspeicher, Backplane-Auslastung, Puffer und Warteschlangen.

Neue Anwendungen sollten außerdem eine Was-wäre-wenn-Analyse durchführen, um den Anwendungserfolg und die Bandbreitenanforderungen zu bestimmen. Normalerweise führen Sie diese Analyse in einer Laborumgebung mithilfe eines Protokoll-Analysators und eines WAN-Verzögerungssimulators durch, um die Auswirkungen von Entfernungen zu verstehen. Sie benötigen nur einen PC, einen Hub, ein WAN-Verzögerungsgerät und einen Lab-Router, der mit dem Produktionsnetzwerk verbunden ist. Sie können die Bandbreite im Labor simulieren, indem Sie den Datenverkehr durch generisches Traffic-Shaping oder durch Ratenbegrenzung auf dem Testrouter drosseln. Der Netzwerkadministrator kann gemeinsam mit der Anwendungsgruppe die Bandbreitenanforderungen, Fensterprobleme und potenzielle Leistungsprobleme der Anwendung in LAN- und WAN-Umgebungen ermitteln.

Führen Sie vor der Bereitstellung einer Geschäftsanwendung eine Was-wäre-wenn-Analyse der Anwendung durch. Andernfalls gibt die Anwendungsgruppe dem Netzwerk die Schuld für die schlechte Leistung. Wenn Sie eine Anwendungs-Was-wäre-wenn-Analyse für neue Bereitstellungen über den Change-Management-Prozess anfordern können, können Sie nicht erfolgreiche Bereitstellungen verhindern und plötzliche Erhöhungen der Bandbreitennutzung für Client-Server- und Batch-Anforderungen besser verstehen.

## Baselining und Trending

Baselining und Trendanalysen ermöglichen Netzwerkadministratoren die Planung und Durchführung von Netzwerk-Upgrades, bevor ein Kapazitätsproblem zu Netzwerkausfällen oder Leistungsproblemen führt. Vergleichen Sie die Ressourcennutzung in aufeinander folgenden Zeiträumen oder destillieren Sie Informationen im Zeitverlauf in einer Datenbank, und ermöglichen Sie Planern, die Parameter der Ressourcennutzung für die letzte Stunde, den letzten Tag, die letzte Woche, den letzten Monat und das letzte Jahr anzuzeigen. In jedem Fall muss die Information wöchentlich, zweimal wöchentlich oder monatlich überprüft werden. Das Problem mit Baselining und Trends ist, dass in großen Netzwerken eine überwältigende Menge an

Informationen überprüft werden muss.

Sie können dieses Problem auf verschiedene Weise beheben:

- Bauen Sie ausreichend Kapazität auf, und schalten Sie in die LAN-Umgebung um, sodass die Kapazität kein Problem darstellt.
- Teilen Sie die Trendinformationen in Gruppen auf, und konzentrieren Sie sich auf hochverfügbare oder kritische Bereiche des Netzwerks, z. B. kritische WAN-Standorte oder Rechenzentrums-LANs.
- Berichtsmechanismen können Bereiche hervorheben, die einen bestimmten Schwellenwert überschreiten, um besondere Aufmerksamkeit zu erhalten. Wenn Sie zuerst kritische Verfügbarkeitsbereiche implementieren, können Sie die Menge der zu überprüfenden Informationen erheblich reduzieren.

Bei allen vorherigen Methoden müssen Sie die Informationen regelmäßig überprüfen. Baselineing und Trendanalysen sind eine proaktive Maßnahme. Wenn das Unternehmen nur über Ressourcen für reaktiven Support verfügt, werden die Mitarbeiter die Berichte nicht lesen.

Viele Netzwerkmanagement-Lösungen bieten Informationen und Diagramme zu Kapazitätsressourcenvariablen. Leider verwenden die meisten Menschen diese Tools nur zur reaktiven Unterstützung eines bestehenden Problems. Dies verfehlt den Zweck des Baselineing und Trends. Zwei Tools, die Informationen zum Kapazitätstrend bei Cisco Netzwerken liefern, sind Concord Network Health und INS EnterprisePRO. In vielen Fällen führen Netzwerkanbieter einfache Skriptsprachen aus, um Kapazitätswerte zu erfassen. Nachstehend finden Sie einige Beispielberichte, die über Skript zur Verbindungsauslastung, CPU-Auslastung und Ping-Leistung gesammelt wurden. Andere Ressourcenvariablen, die für den Trend wichtig sein können, sind Speicher, Warteschlangentiefe, Broadcast-Volumen, Puffer, Frame-Relay-Überlastungsbenachrichtigung und Backplane-Nutzung. Informationen zur Verbindungsauslastung und CPU-Auslastung finden Sie in der folgenden Tabelle:

### Link-Nutzung

Ressource	Adresse	Segment	Durchschnittliche Auslastung (%)	Spitzenauslastung (%)
JTKR01S2	10.2.6.1	128 Kbit/s	66.3	97.6
JKR01S0	10.2.6.2	128 Kbit/s	66.3	97.8
FMCR18S4/4	10.2.5.1	384 Kbit/s	51.3	109.7
PACR01S3/1	10.2.5.2	384 Kbit/s	51.1	98.4

### CPU-Auslastung

Ressource	Polling-Adresse	Durchschnittliche Auslastung (%)	Spitzenauslastung (%)
-----------	-----------------	----------------------------------	-----------------------

FSTR01	10.28.142.1	60.4	80
NERT06	10.170.2.1	47	86
NORR01	10.73.200.1	47	99
RTCR01	10.49.136.1	42	98

## Link-Nutzung

Ressource	Adresse	AvResT (mS) 09-09-98	AvResT (mS) 09-09-98	AvResT (mS) 09-09-98	AvResT (mS) 10-01-98
ADR 01	10.190.56.1	469.1	852.4	461.1	873.2
ABNR01	10.190.52.1	486.1	869.2	489.5	880.2
APRIL 01	10.190.54.1	490.7	883.4	485.2	892.5
ASAR01	10.196.170.1	619.6	912.3	613.5	902.2
ASR 01	10.196.178.1	667.7	976.4	655.5	948.6
ASYR01S					503.4
ZWRT01	10.177.32.1	460.1		444.7	
BEJR01	10.195.18.1	1023.7	1064.6	1184	1021.9

## Ausnahmeverwaltung

Das Ausnahmemanagement ist eine nützliche Methode zur Identifizierung und Behebung von Kapazitäts- und Performance-Problemen. Ziel ist es, eine Benachrichtigung über Verstöße gegen Kapazitäts- und Leistungsschwellenwerte zu erhalten, um das Problem sofort untersuchen und beheben zu können. Beispielsweise kann ein Netzwerkadministrator einen Alarm für einen Router mit hoher CPU-Auslastung erhalten. Der Netzwerkadministrator kann sich beim Router anmelden, um festzustellen, warum die CPU so hoch ist. Anschließend kann sie eine Korrekturkonfiguration durchführen, die die CPU reduziert, oder eine Zugriffsliste erstellen, die den Verkehr verhindert, der das Problem verursacht, insbesondere wenn der Verkehr nicht geschäftskritisch zu sein scheint.

Sie können die Ausnahmeverwaltung für kritischere Probleme einfach mithilfe der RMON-Konfigurationsbefehle eines Routers oder mithilfe erweiterter Tools wie dem Service Level Manager von Netsys in Verbindung mit SNMP-, RMON- oder Netflow-Daten konfigurieren. Die meisten Netzwerkmanagement-Tools können Grenzwerte und Alarme für Verstöße festlegen. Der wichtige Aspekt des Ausnahmemanagementprozesses besteht darin, das Problem nahezu in Echtzeit zu melden. Andernfalls kann das Problem verschwinden, bevor jemand bemerkt, dass die Benachrichtigung empfangen wurde. Dies ist innerhalb eines NOC möglich, wenn die Organisation über eine konsistente Überwachung verfügt. Andernfalls wird eine Pager-

Benachrichtigung empfohlen.

Im folgenden Konfigurationsbeispiel wird eine Benachrichtigung über einen steigenden oder fallenden Schwellenwert für die Router-CPU an eine Protokolldatei gesendet, die konsistent überprüft werden kann. Sie können ähnliche RMON-Befehle für Verstöße gegen kritische Schwellenwerte bei der Verbindungsnutzung oder andere SNMP-Schwellenwerte einrichten.

```
rmon event 1 trap CPUtrap description
"CPU Util >75%"rmon event 2 trap CPUtrap description
"CPU Util <75%"rmon event 3 trap CPUtrap description
"CPU Util >90%"rmon event 4 trap CPUtrap description
"CPU Util <90%"rmon alarm 75 lsystem.56.0 10 absolute rising-threshold
75 1 falling-threshold 75 2rmon alarm 90 lsystem.56.0 10 absolute rising-threshold
90 3 falling-threshold 90 4
```

## QoS-Management

Zur Verwaltung der Quality of Service (QoS) werden innerhalb des Netzwerks bestimmte Datenverkehrsklassen erstellt und überwacht. Datenverkehr ermöglicht eine konsistentere Leistung für bestimmte Anwendungsgruppen (definiert in Datenverkehrsklassen). Traffic Shaping-Parameter bieten eine hohe Flexibilität bei der Priorisierung und Datenverkehrsformung für bestimmte Datenverkehrsklassen. Zu diesen Funktionen gehören Funktionen wie Committed Access Rate (CAR), Weighted Random Early Detection (WRED) und Class-Based Fair Weighted Queuing. Verkehrsklassen werden in der Regel auf Basis von Performance-SLAs für geschäftskritische Anwendungen und spezifische Anwendungsanforderungen wie Sprache erstellt. Nicht kritischer oder nicht geschäftskritischer Datenverkehr wird so gesteuert, dass Anwendungen und Services mit höherer Priorität nicht beeinträchtigt werden.

Das Erstellen von Datenverkehrsklassen erfordert grundlegende Kenntnisse der Netzwerkauslastung, der spezifischen Anwendungsanforderungen und der Prioritäten für Geschäftsanwendungen. Zu den Anwendungsanforderungen gehören Kenntnisse der Paketgröße, Timeout-Probleme, Jitter-Anforderungen, Burst-Anforderungen, Batch-Anforderungen und allgemeine Leistungsprobleme. Mit diesem Wissen können Netzwerkadministratoren Traffic-Shaping-Pläne und Konfigurationen erstellen, die eine konsistentere Anwendungsleistung in einer Vielzahl von LAN-/WAN-Topologien bieten.

Ein Unternehmen verfügt beispielsweise über eine 10-Megabit-ATM-Verbindung zwischen zwei Hauptstandorten. Der Link ist manchmal bei großen Dateiübertragungen überlastet, was zu Leistungseinbußen bei der Verarbeitung von Online-Transaktionen und schlechter oder unbrauchbarer Sprachqualität führt.

Die Organisation hat vier verschiedene Verkehrsklassen eingerichtet. Der Sprachdatenverkehr erhielt die höchste Priorität und konnte diese beibehalten, selbst wenn er die geschätzte Datenverkehrsmenge überschritt. Die Klasse der kritischen Anwendungen erhielt die nächsthöhere Priorität, es war jedoch nicht zulässig, über die gesamte Verbindungsgröße abzüglich der geschätzten Anforderungen an die Sprachbandbreite zu springen. Wenn es platzt, wird es fallen gelassen. Der Datenverkehr zur Dateiübertragung wurde einfach mit einer niedrigeren Priorität versehen, und der gesamte übrige Datenverkehr lag irgendwo in der Mitte.

Die Organisation muss nun für diesen Link eine QoS-Verwaltung durchführen, um den Umfang des von den einzelnen Klassen abgewickelten Datenverkehrs zu bestimmen und die Leistung innerhalb der einzelnen Klassen zu messen. Wenn die Organisation dies nicht tut, kann es bei einigen Klassen zu Engpässen kommen, oder die Leistungs-SLAs werden innerhalb einer

bestimmten Klasse möglicherweise nicht erfüllt.

Die Verwaltung von QoS-Konfigurationen gestaltet sich aufgrund des Fehlens von Tools nach wie vor schwierig. Eine Möglichkeit besteht darin, mit dem Internet Performance Manager (IPM) von Cisco unterschiedlichen Datenverkehr über die Verbindung zu senden, die in die einzelnen Verkehrsklassen fällt. Anschließend können Sie die Leistung für jede Klasse überwachen. IPM bietet dann Trendanalysen, Echtzeitanalysen und Hop-by-Hop-Analysen, um Problembereiche zu identifizieren. Andere verlassen sich möglicherweise weiterhin auf eine manuellere Methode, z. B. die Untersuchung der Warteschlangen und verlorenen Pakete innerhalb jeder Datenverkehrsklasse anhand von Schnittstellenstatistiken. In einigen Unternehmen können diese Daten über SNMP erfasst oder in einer Datenbank für Baselines und Trends analysiert werden. Es gibt auch einige Tools auf dem Markt, die bestimmte Datenverkehrstypen über das Netzwerk senden, um die Leistung für einen bestimmten Dienst oder eine bestimmte Anwendung zu bestimmen.

## Erfassen und Berichten von Kapazitätsinformationen

Die Erhebung und Meldung von Kapazitätsinformationen sollte mit den drei empfohlenen Bereichen des Kapazitätsmanagements verknüpft werden:

- Was-wäre-wenn-Analyse, bei der es um Netzwerkänderungen und deren Auswirkungen auf die Umgebung geht
- Baselining und Trendanalyse
- Ausnahmeverwaltung

In jedem dieser Bereiche sollten Sie einen Plan zur Informationssammlung entwickeln. Im Fall von "Was-wäre-wenn"-Analysen für Netzwerke und Anwendungen benötigen Sie Tools, mit denen Sie die Netzwerkumgebung nachahmen und die Auswirkungen der Änderung im Hinblick auf potenzielle Ressourcenprobleme innerhalb der Geräte- oder Datenebene verstehen können. Im Falle von Baselining und Trends benötigen Sie Snapshots für Geräte und Links, die die aktuelle Ressourcennutzung anzeigen. Anschließend prüfen Sie die Daten im Laufe der Zeit, um sich über die potenziellen Upgrade-Anforderungen zu informieren. So können Netzwerkadministratoren Upgrades richtig planen, bevor Kapazitäts- oder Leistungsprobleme auftreten. Wenn Probleme auftreten, benötigen Sie eine Ausnahmeverwaltung, um die Netzwerkadministratoren zu benachrichtigen, damit sie das Netzwerk abstimmen oder das Problem beheben können.

Dieser Prozess kann in folgende Schritte unterteilt werden:

1. Bestimmen Sie Ihre Anforderungen.
2. Definieren Sie einen Prozess.
3. Definieren Sie Kapazitätsbereiche.
4. Definieren Sie die Kapazitätsvariablen.
5. Daten interpretieren

## Ermitteln Sie Ihre Anforderungen

Die Entwicklung eines Kapazitäts- und Performance-Managementplans erfordert ein Verständnis der benötigten Informationen und des Zwecks dieser Informationen. Teilen Sie den Plan in drei erforderliche Bereiche auf: jeweils eine für "Was-wäre-wenn"-Analysen, Baselining/Trends und Ausnahmemanagement. Informieren Sie sich in jedem dieser Bereiche über die verfügbaren Ressourcen und Tools sowie deren Bedarf. Viele Unternehmen scheitern bei der Bereitstellung

von Tools, weil sie die Technologie und Funktionen der Tools berücksichtigen, aber nicht die Mitarbeiter und das Fachwissen, die für die Verwaltung der Tools erforderlich sind. Berücksichtigen Sie bei der Planung nicht nur die erforderlichen Personen und Fachkenntnisse, sondern auch die erforderlichen Prozessverbesserungen. Zu diesen Mitarbeitern gehören u. a. Systemadministratoren für die Verwaltung der Netzwerkmanagementstationen, Datenbankadministratoren für die Unterstützung der Datenbankadministration, geschulte Administratoren für die Verwendung und Überwachung der Tools und übergeordnete Netzwerkadministratoren für die Festlegung von Richtlinien, Grenzwerten und Informationserfassungsanforderungen.

## **Definieren eines Prozesses**

Sie benötigen außerdem einen Prozess, der sicherstellt, dass das Tool erfolgreich und konsistent eingesetzt wird. Möglicherweise müssen Sie die Prozesse optimieren, um festzulegen, was Netzwerkadministratoren tun sollten, wenn Schwellenwertverletzungen auftreten, oder um festzulegen, welcher Prozess für Baselineing, Trending und Netzwerkupgrades zu verfolgen ist. Sobald Sie die Anforderungen und Ressourcen für eine erfolgreiche Kapazitätsplanung ermittelt haben, können Sie die Methodik berücksichtigen. Viele Unternehmen entscheiden sich dafür, diese Art von Funktionen an ein Netzwerkdienstleistungsunternehmen wie INS auszulagern oder ihr Know-how intern aufzubauen, da sie den Service als Kernkompetenz betrachten.

## **Kapazitätsbereiche definieren**

Der Plan für die Kapazitätsplanung sollte auch eine Definition der Kapazitätsbereiche enthalten. Die folgenden Netzwerkbereiche können eine gemeinsame Strategie für die Kapazitätsplanung nutzen: z. B. LAN, WAN-Außenstellen, WAN-Standorte und Einwahlzugang. Das Definieren verschiedener Bereiche ist aus mehreren Gründen hilfreich:

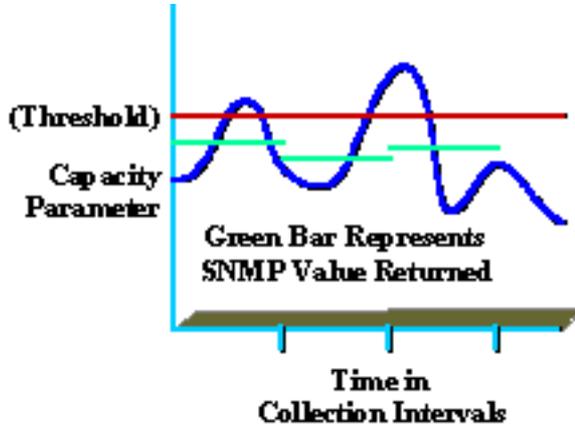
- Unterschiedliche Bereiche können unterschiedliche Schwellenwerte aufweisen. Die LAN-Bandbreite ist beispielsweise wesentlich billiger als die WAN-Bandbreite. Daher sollten die Auslastungsschwellenwerte niedriger sein.
- In verschiedenen Bereichen müssen möglicherweise verschiedene MIB-Variablen überwacht werden. FECN- und BECN-Zähler in Frame Relay sind beispielsweise für das Verständnis von Kapazitätsproblemen bei Frame-Relays entscheidend.
- Die Aktualisierung einiger Bereiche des Netzwerks ist möglicherweise schwieriger oder zeitaufwendiger. So können beispielsweise internationale Schaltungen wesentlich längere Vorlaufzeiten aufweisen und eine entsprechend höhere Planung benötigen.

## **Kapazitätsvariablen definieren**

Der nächste wichtige Bereich ist die Definition der zu überwachenden Variablen und der Schwellenwerte, die eine Aktion erfordern. Die Definition der Kapazitätsvariablen hängt maßgeblich von den im Netzwerk verwendeten Geräten und Medien ab. Generell sind Parameter wie CPU, Arbeitsspeicher und Verbindungsauslastung nützlich. Andere Bereiche können jedoch für bestimmte Technologien oder Anforderungen wichtig sein. Hierzu zählen Warteschlangentiefen, Leistung, Frame-Relay-Überlastungsbenachrichtigung, Backplane-Nutzung, Puffer-Nutzung, NetFlow-Statistiken, Broadcast-Volumen und RMON-Daten. Denken Sie dabei an Ihre langfristigen Pläne, doch zunächst sollten Sie sich auf einige wenige Kernbereiche beschränken, um Ihren Erfolg sicherzustellen.

## Daten interpretieren

Ein Verständnis der erfassten Daten ist auch für die Bereitstellung eines qualitativ hochwertigen Service von entscheidender Bedeutung. Viele Unternehmen verstehen beispielsweise die Spitzenauslastung und die durchschnittliche Auslastung nicht vollständig. Das folgende Diagramm zeigt einen Spitzenwert für Kapazitätsparameter basierend auf einem SNMP-Erfassungsintervall von 5 Minuten (grün dargestellt).



Auch wenn der gemeldete Wert unter dem Schwellenwert lag (rot dargestellt), können innerhalb des Erfassungsintervalls weiterhin Spitzen auftreten, die über dem Schwellenwert liegen (blau dargestellt). Dies ist wichtig, da es während des Erfassungsintervalls möglicherweise zu Spitzenwerten kommt, die sich auf die Leistung oder Kapazität des Netzwerks auswirken. Achten Sie darauf, ein sinnvolles Auflistungsintervall auszuwählen, das nützlich ist und keinen übermäßigen Aufwand verursacht.

Ein weiteres Beispiel ist die durchschnittliche Auslastung. Wenn Mitarbeiter nur von acht bis fünf im Büro sind, die durchschnittliche Auslastung jedoch 7x24 beträgt, können die Informationen irreführend sein.

## Zugehörige Informationen

- [Technischer Support – Cisco Systems](#)

## Informationen zu dieser Übersetzung

Cisco hat dieses Dokument maschinell übersetzen und von einem menschlichen Übersetzer editieren und korrigieren lassen, um unseren Benutzern auf der ganzen Welt Support-Inhalte in ihrer eigenen Sprache zu bieten. Bitte beachten Sie, dass selbst die beste maschinelle Übersetzung nicht so genau ist wie eine von einem professionellen Übersetzer angefertigte. Cisco Systems, Inc. übernimmt keine Haftung für die Richtigkeit dieser Übersetzungen und empfiehlt, immer das englische Originaldokument (siehe bereitgestellter Link) heranzuziehen.