

容量和效能管理：最佳實踐白皮書

目錄

[簡介](#)

[容量和效能管理概述](#)

[CPU](#)

[底板或I/O](#)

[記憶體](#)

[介面和管道尺寸](#)

[佇列、延遲和抖動](#)

[速度和距離](#)

[應用特性](#)

[容量和效能管理最佳實踐](#)

[服務級別管理](#)

[網路和應用假設分析](#)

[基線化和趨勢](#)

[例外管理](#)

[QoS管理](#)

[收集和報告容量資訊](#)

[確定您的需求](#)

[定義流程](#)

[定義容量區域](#)

[定義能力變數](#)

[解釋資料](#)

[相關資訊](#)

簡介

在大型企業和服務提供商網路中，高網路可用性是一項關鍵任務要求。網路經理在提供更高可用性方面面臨著日益嚴峻的挑戰，包括計畫外停機時間、缺少專業技能、工具不足、技術複雜、業務整合以及市場競爭。容量和效能管理可幫助網路經理實現新的全球業務目標以及一致的網路可用性和效能。

本檔案將研究以下主題：

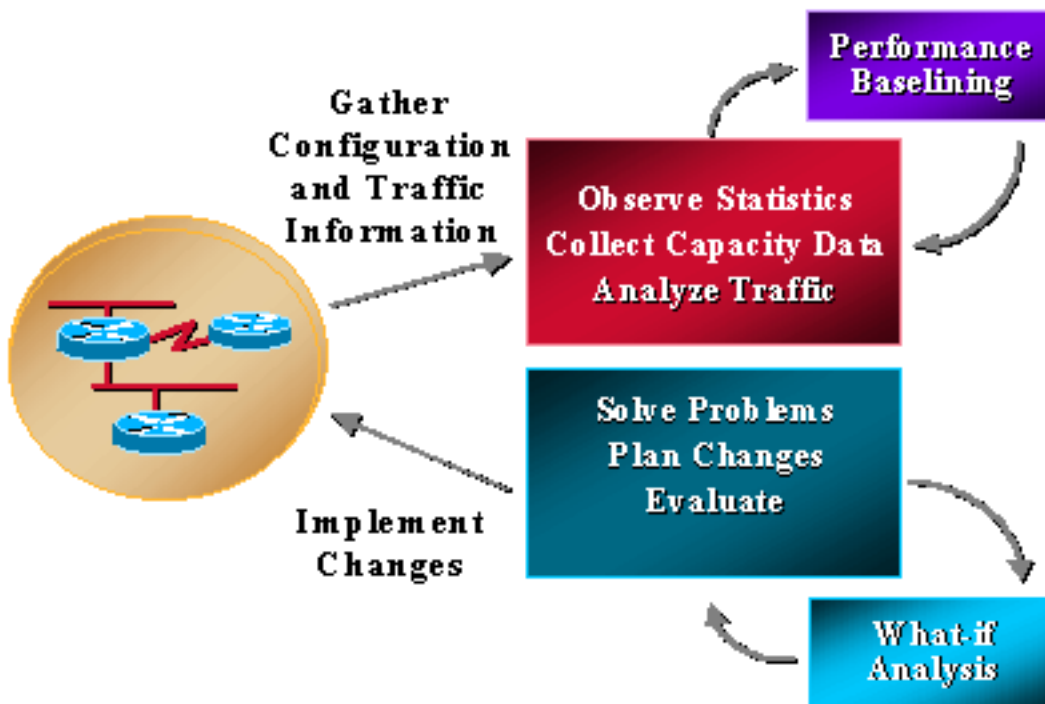
- 一般容量和效能問題，包括網路中的風險和潛在的容量問題。
- 容量和效能管理最佳實踐，包括假設分析、基準測試、趨勢分析、異常管理和QoS管理。
- 如何制定容量規劃策略，包括容量規劃中使用的常用技術、工具、MIB變數和閾值。

容量和效能管理概述

容量規劃是確定網路資源所需的過程，以防止對關鍵業務應用程式的效能或可用性產生影響。效能管理是對單個服務和整體服務的網路服務響應時間、一致性和品質進行管理的實踐。

注意：效能問題通常與容量有關。應用速度較慢，因為頻寬和資料在通過網路傳輸之前必須在隊列中等待。在語音應用中，延遲和抖動等問題直接影響語音呼叫的品質。

大多陣列組織已經收集了一些與容量相關的資訊，並始終如一地努力解決問題、規劃更改並實施新的容量和效能功能。但是，組織不會定期執行趨勢分析和假設分析。假設分析是確定網路更改影響的過程。趨勢分析是一個過程，該過程由網路容量和效能問題的基線組成，並檢查網路趨勢的基線以瞭解將來的升級要求。容量和效能管理還應該包括異常管理（即在使用者撥入之前識別並解決問題），以及QoS管理（即網路管理員規劃、管理和識別各個服務效能問題）。下圖說明了容量和效能管理流程。



容量和效能管理也有其限制，通常與CPU和內存有關。以下是潛在的關注領域：

- CPU
- 底板或I/O
- 記憶體和緩衝區
- 介面和管道尺寸
- 隊列、延遲和抖動
- 速度和距離
- 應用特性

一些對容量規劃和效能管理的引用也提到了一種稱為「資料平面」和「控制平面」的東西。資料平面只是通過網路傳輸的資料涉及的容量和效能問題，而控制平面意味著維持資料平面正常功能所需的資源。控制平面功能包括服務開銷，例如路由、生成樹、介面保持活動和對裝置的SNMP管理。這些控制平面要求使用CPU、記憶體、緩衝、隊列和頻寬，就像流經網路的流量一樣。許多控制平面要求對於系統的整體功能也是至關重要的。如果他們沒有所需的資源，網路就會發生故障。

CPU

CPU通常由任何網路裝置上的控制平面和資料平面使用。在容量和效能管理中，必須確保裝置和網路具有足夠的CPU在任何時候都能正常工作。CPU不足通常會造成網路崩潰，因為一台裝置上的資源不足可能會影響整個網路。CPU不足也會增加延遲，因為當沒有主CPU的硬體交換時，資料必須等待處理。

底板或I/O

背板或I/O是指裝置可以處理的流量總量，通常以匯流排大小或背板功能來描述。背板不足通常會導致丟包，從而導致重新傳輸和其他流量。

記憶體

記憶體是另一個具有資料平面和控制平面要求的資源。路由表、ARP表和其他資料結構等資訊需要記憶體。當裝置記憶體不足時，裝置上的某些操作可能會失敗。此操作可能會影響控制平面進程或資料平面進程，具體取決於實際情況。如果控制平面進程失敗，整個網路可能會降級。例如，當路由收斂需要額外記憶體時，可能會發生這種情況。

介面和管道尺寸

介面和管道大小是指可以在任何一個連線上同時傳送的資料量。這通常被錯誤地稱為連線速度，但資料實際上不會以不同的速度從一個裝置傳輸到另一個裝置。矽速度和硬體功能有助於根據介質確定可用頻寬。此外，軟體機制可以「限制」資料，使其符合特定頻寬分配服務。在幀中繼或ATM的服務提供商網路中，您通常可以看到這一點，這些網路本身具有速度為1.54kpbs到155mbs及更高的功能。當存在頻寬限制時，資料會在傳輸隊列中排隊。傳送隊列可以具有不同的軟體機制以優先化隊列中的資料；但是，當隊列中有資料時，它必須等待現有資料，然後才能將資料從介面轉發出去。

佇列、延遲和抖動

佇列、延遲和抖動也會影響效能。可以調整傳輸隊列以通過不同方式影響效能。例如，如果隊列很大，則資料等待時間更長。當隊列較小時，資料將被丟棄。這稱為taildrop，對於TCP應用程式是可接受的，因為資料將被重新傳輸。但是，由於隊列丟棄甚至隊列延遲過長，語音和影片的表現不佳，需要特別注意頻寬或管道大小。如果裝置沒有足夠的資源立即轉發資料包，輸入隊列也可能發生隊列延遲。這可能是由於CPU、記憶體或緩衝區。

延遲是指從收到資料包到轉發資料包的正常處理時間。正常現代資料交換機和路由器在正常情況下具有極低的延遲(< 1ms)，且沒有資源限制。採用數位訊號處理器轉換和壓縮模擬語音資料包的現代裝置可能需要更長的時間，甚至高達20毫秒。

抖動描述了流應用（包括語音和影片）的資料包間間隙。如果資料包到達的時間不同，且資料包間的時間不同，則抖動會很高，語音品質會下降。抖動主要是排隊延遲的一個因素。

速度和距離

速度和距離也是影響網路效能的一個因素。根據光速，資料網路具有一致的資料轉發速度。這大約是每毫秒100哩。如果組織在國際上運行客戶端 — 伺服器應用程式，則它們可能會出現相應的資料包轉發延遲。當應用未針對網路效能進行最佳化時，速度和距離可能會成為影響應用效能的巨大因素。

應用特性

應用特性是影響容量和效能的最後一個方面。諸如小視窗大小、應用程式keepalive以及通過網路傳送的資料量與所需資料量之類的問題，會影響應用程式在許多環境（尤其是WAN）中的效能。

容量和效能管理最佳實踐

本節詳細討論了五大容量和效能管理最佳實踐：

- [服務級別管理](#)
- [網路和應用假設分析](#)
- [基線化和趨勢化](#)
- [例外管理](#)
- [QoS管理](#)

服務級別管理

服務級別管理定義並管理其他所需的容量和效能管理流程。網路經理瞭解他們需要容量規劃，但是他們面臨著預算和人員限制，無法實現完整的解決方案。服務級別管理是一種行之有效的方法體系，通過定義交付項並為與該交付項相關的服務建立雙向責任制，可以幫助解決資源問題。您可以通過兩種方式完成此操作：

- 為包括容量和效能管理的服務建立使用者和網路組織之間的服務級別協定。該服務將包括保持服務品質的報告和建議。但是，使用者必須準備好為服務和任何所需的升級提供資金。
- 網路組織定義其容量和效能管理服務，然後逐一嘗試為該服務和升級提供資金。

在任何情況下，網路組織都應當從定義容量規劃和效能管理服務開始，該服務包括他們目前可以提供的服務的哪些方面以及未來的規劃內容。完整的服務包括網路更改和應用更改的假設分析、已定義效能變數的基線化和趨勢分析、已定義容量和效能變數的異常管理以及QoS管理。

網路和應用假設分析

執行網路和應用程式假設分析以確定計畫更改的結果。如果不進行假設分析，組織就會冒極大的風險，來改變成功情況和整體網路可用性。在許多情況下，網路更改會導致擁塞的崩潰，從而導致大量生產中斷。此外，數量驚人的應用程式引入會失敗，並會對其他使用者和應用程式造成影響。這些故障在許多網路組織內仍然存在，但只需一些工具和一些額外的規劃步驟，它們就可以完全避免。

通常需要一些新流程來執行品質假設分析。第一步是確定所有更改的風險級別，並要求對較高風險更改進行更深入的假設分析。風險級別可以是所有提交變更的必填欄位。然後，更高的風險級別變更要求定義變更的假設分析。網路假設分析確定網路更改對網路利用率和網路控制平面資源問題的影響。應用假設分析將確定專案應用的成功、頻寬要求以及任何網路資源問題。下表列出了風險級別分配和相應的測試要求的示例：

風險水準	定義	變更計畫建議
1	由於新產品、軟體、拓撲或功能的引入，對大量使用者（500多名）或關鍵業務服務產生了巨大的潛在影響。	驗證新解決方案的實驗室。實驗室驗證包括已記錄的解決方案測試和驗證，以及顯示對現有基礎架構的影響的假設分析。我們推薦解決方案試用。新的解決方案需要完成操作支援文檔。

	<ul style="list-style-type: none"> • 更改涉及預期的網路停機時間。 	<ul style="list-style-type: none"> • 執行思科NSA設計稽核。 • 建立回退計畫。 • 建立實施計畫。 • 建立變更進程。
2	<ul style="list-style-type: none"> • 由於流量或使用者的增加、主幹更改或路由更改，對大量使用者（500以上）或業務關鍵型服務有較大的潛在影響。 • 更改可能需要一些停機時間。 	<ul style="list-style-type: none"> • 執行假設分析以確定對現有環境的影響（應在實驗室環境中完成）。 • 測試並檢查路由更改的功能。 • 建立回退計畫。 • 對主要路由或主幹更改進行設計稽核。 • 建立實施計畫。 • 建立變更進程。
3	<ul style="list-style-type: none"> • 由於任何非標準更改，對較少的使用者或業務服務的潛在影響不大。 • 包括新產品、軟體、拓撲、功能新增新使用者、流量增加或非標準拓撲。 • 更改可能需要一些停機時間。 	<ul style="list-style-type: none"> • 執行新解決方案的工程分析（可能需要實驗室驗證）。 • 建立實施計畫。 • 建立變更進程。
4	<ul style="list-style-type: none"> • 降低潛在的服務或使用者影響。 • 包括新增新的標準模板網路模組，例如在路由器上構建或伺服器交換機/集線器。 • 包括部署新的WAN站點或其他成熟的接入服務。 • 所有第3級風險更改已在生產環境中經過技術驗證。 • 更改可能需要一些停機時間。 	<ul style="list-style-type: none"> • 建立實施計畫。 • 建立變更進程。
5	<ul style="list-style-type: none"> • 不影響使用者或服務。 • 包括將單個使用者新增到網路並 	<ul style="list-style-type: none"> • 更改進程可選。

<p>更改標準配置 ，如密碼、橫幅 、SNMP或其他 標準配置引數。 • 沒有停機時間。</p>	
--	--

一旦定義了需要模擬分析的位置，就可以定義服務。

您可以使用建模工具或模擬生產環境的實驗室執行網路假設分析。建模工具受限於應用程式對裝置資源問題的理解程度，而且由於大多數網路更改都是新裝置，因此應用程式可能無法理解更改的影響。最佳方法是在實驗室中構建一些生產網路的表示形式，並使用流量發生器測試負載下所需的軟體、功能、硬體或配置。從生產網路向實驗室洩露路由（或其他控制資訊）也會增強實驗室環境。測試不同流量型別（包括SNMP、廣播、組播、加密或壓縮流量）的其他資源要求。使用所有這些不同的方法，分析潛在壓力情況（如路由收斂、鏈路抖動和裝置重新啟動）下的裝置資源需求。資源利用率問題包括CPU、記憶體、背板利用率、緩衝區和隊列等正常容量資源區域。

新應用還應執行模擬分析，以確定應用成功和頻寬要求。通常在實驗室環境中使用協定分析器和WAN延遲模擬器執行此分析，以瞭解距離的影響。您只需要將PC、集線器、WAN延遲裝置和實驗室路由器連線到生產網路。您可以在實驗中使用測試路由器上的通用流量整形或速率限制來限制流量，從而模擬實驗中的頻寬。網路管理員可以與應用組合作，瞭解在LAN和WAN環境中應用的頻寬要求、視窗問題和潛在效能問題。

在部署任何業務應用程式之前執行應用程式假設分析。如果不這樣做，應用程式組會將效能不佳歸咎於網路。如果您能夠通過變更管理流程以某種方式要求對新部署進行應用程式假設分析，則您可以幫助防止不成功的部署，並更好地瞭解客戶端——伺服器和批處理要求的頻寬消耗的突然增加。

基線化和趨勢

基線化和趨勢分析允許網路管理員在容量問題導致網路停機時間或效能問題之前規劃和完成網路升級。比較連續時段內的資源利用率或提取資料庫中一段時間內的資訊，使規劃者能夠檢視上一小時、天、周、月和年的資源利用率引數。無論哪種情況，都必須有人每週、每兩週或每月複查一次資訊。基線化和趨勢分析的問題在於，在大型網路中，需要檢視大量資訊。

可以通過多種方法解決此問題：

- 構建足夠的容量並切換到LAN環境，因此容量不是問題。
- 將趨勢資訊劃分為不同的組，並集中關注網路的高可用性或關鍵區域，如關鍵WAN站點或資料中心LAN。
- 報告機制可以突出超過某個門檻的領域值得特別注意。如果您先實施關鍵可用性領域，則可以顯著減少稽核所需的資訊量。

使用前面的所有方法，您仍然需要定期檢視資訊。基線化和趨勢分析是一項積極的工作，如果組織僅擁有用於被動支援的資源，個人將不會閱讀報告。

許多網路管理解決方案提供有關容量資源變數的資訊和圖形。不幸的是，大部分人只使用這些工具被動支援現有問題；這破壞了基線化和趨勢化的目的。兩種有效提供思科網路容量趨勢資訊的工具是Concord Network Health產品和INS EnterprisePRO產品。在許多情況下，網路組織運行簡單的指令碼語言來收集容量資訊。以下是一些通過指令碼收集的鏈路利用率、CPU利用率和ping效能報告的示例。其他對趨勢很重要的資源變數包括記憶體、隊列深度、廣播量、緩衝區、幀中繼擁塞通知和背板利用率。有關鏈路利用率和CPU利用率的資訊，請參閱下表：

鏈路利用率

資源	地址	網段	平均利用率 (%)	峰值利用率 (%)
JTKR01S2	10.2.6.1	128 Kbps	66.3	97.6
JYKR01S0	10.2.6.2	128 Kbps	66.3	97.8
FMCR18S4/4	10.2.5.1	384 Kbps	51.3	109.7
PACR01S3/1	10.2.5.2	384 Kbps	51.1	98.4

CPU利用率

資源	輪詢地址	平均利用率(%)	峰值利用率(%)
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

鏈路利用率

資源	地址	AvResT(mS)09-09-98	AvResT(mS)09-09-98	AvResT(mS)09-09-98	AvResT(mS)10-01-98
AADR.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
2001年4月	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
ASR01	10.196.178.1	667.7	976.4	655.5	948.6
ASYSR01S					503.4
AZWR	10.177.32.	460.1		444.7	

T01	1				
BEJ R01	10.19 5.18. 1	1023.7	1064.6	1184	1021.9

例外管理

例外管理是識別和解決容量和效能問題的重要方法。其思想是接收容量和效能閾值違規通知，以便立即調查和解決該問題。例如，網路管理員可能收到路由器上CPU使用率較高的警報。網路管理員可以登入到路由器以確定CPU為什麼這麼高。然後，她可以執行一些補救配置，以減少CPU或建立訪問清單，防止導致問題的流量，尤其是當流量看起來不是業務關鍵型時。

只需在路由器上使用RMON配置命令，或者使用更高級的工具（如Netsys服務級別管理器和SNMP、RMON或Netflow資料），即可為更重要的問題配置例外管理。大多數網路管理工具都能夠設定違規閾值和警報。例外管理流程的重要方面是提供近即時問題通知。否則，問題會在有人注意到收到通知之前消失。如果組織具有一致的監控功能，則可在NOC內完成此操作。否則，建議使用尋呼機通知。

以下配置示例為路由器CPU提供一條對日誌檔案進行上升和下降的閾值通知，該日誌檔案可能會持續檢視。您可以為嚴重鏈路利用率閾值違規或其他SNMP閾值設定類似的RMON命令。

```
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管理

服務品質管理包括建立和監控網路內的特定流量類。流量為特定應用組（在流量類中定義）提供更一致的效能。流量整形引數為特定流量類別的優先順序劃分和流量整形提供了極大的靈活性。這些功能包括諸如承諾接入速率(CAR)、加權隨機早期檢測(WRED)和基於類的公平加權排隊等功能。流量類別通常基於效能SLA建立，用於更多關鍵業務應用和特定應用需求（如語音）。對非關鍵或非業務流量的控制方式也不會影響優先順序更高的應用程式和服務。

建立流量類別需要具備對網路利用率、特定應用要求以及業務應用優先順序的基本瞭解。應用要求包括瞭解資料包大小、超時問題、抖動要求、突發要求、批處理要求和總體效能問題。藉助此知識，網路管理員可以建立流量整形計畫和配置，從而在多種區域網/廣域網拓撲中提供更一致的應用效能。

例如，一個組織在兩個主要站點之間有一個10兆位ATM連線。鏈路有時因大型檔案傳輸而擁塞，導致線上交易處理的效能下降，語音品質差或不可用。

該組織設定了四種不同的流量類別。語音被賦予最高的優先順序，並且允許保持該優先順序，即使語音突發超過估計的流量速率。關鍵應用類具有下一個最高優先順序，但不允許在總鏈路大小減去估計的語音頻寬要求後突發流量。當它發生突發時，將被丟棄。檔案傳輸流量僅具有較低的優先順序，所有其他流量則位於中間的某個位置。

現在，組織必須對此連結執行QoS管理，以確定每個類所承受的流量大小，並測量每個類內的效能。如果組織無法做到這一點，則某些類可能會發生飢餓，或者某個特定類可能無法滿足效能SLA。

由於缺乏工具，管理QOS配置仍是一項困難的任務。一種方法是使用思科的網際網路效能管理員 (IPM)透過屬於每個流量類別的連結傳送不同的流量。然後，您可以監控每個類別的效能，IPM提供趨勢分析、即時分析和逐跳分析，以查明問題區域。其他方法可能仍依賴於更為手動的方法，如根據介面統計資訊調查每個流量類中的排隊和丟棄的資料包。在某些組織中，可以通過SNMP收集此資料，或者將這些資料解析為資料庫，以獲取基線和趨勢。市場上也存在一些工具，它們通過網路傳送特定的流量型別，以確定特定服務或應用的效能。

收集和報告容量資訊

收集和報告能力資訊應連結到建議的三個能力管理領域：

- 假設分析，以網路更改為中心，以及更改對環境的影響
- 基線化和趨勢化
- 例外管理

在上述每個領域內，制定資訊收集計畫。在網路或應用程式假設分析的情況下，您需要一些工具來模擬網路環境，並理解更改對裝置控制平面或資料平面內潛在資源問題的影響。在基線化和趨勢分析的情況下，您需要為顯示當前資源利用率的裝置和連結建立快照。然後檢視隨時間變化的資料，瞭解潛在的升級要求。這樣，網路管理員就可以在出現容量或效能問題之前正確規劃升級。出現問題時，您需要進行異常管理以警告網路管理員，以便他們調整網路或解決問題。

此過程可分為以下步驟：

1. 確定您的需求。
2. 定義流程。
3. 定義容量區域。
4. 定義能力變數。
5. 解釋資料。

確定您的需求

制定能力和績效管理計畫需要瞭解您需要的資訊以及資訊的用途。將計畫分為三個必需區域：分別用於假設分析、基線化/趨勢分析和例外管理。在上述每個領域中，瞭解有哪些可用的資源和工具以及需要什麼。許多組織在使用工具部署時都會失敗，因為它們會考慮工具的技術和功能，但不考慮管理工具所需的人員和專業技能。在計畫中納入所需人員和專業技能，並改進流程。這些人可能包括管理網路管理站的系統管理員、幫助進行資料庫管理的資料庫管理員、使用和監視工具的受過培訓的管理員，以及確定策略、閾值和資訊收集要求的高層網路管理員。

定義流程

您還需要一個流程來確保成功且一致地使用工具。可能需要改進流程，以定義發生閾值違規時網路管理員應該做什麼，或者執行什麼流程進行網路基線、趨勢分析和升級。確定成功容量規劃的需求和資源後，可以考慮方法。許多組織選擇將此型別的功能外包給網路服務組織（如INS），或建立內部專業知識，因為他們認為服務是一項核心能力。

定義容量區域

能力規劃計畫還應包括能力領域的定義。以下是可以共用通用容量規劃策略的網路區域：例如，企業LAN、WAN現場辦公室、關鍵WAN站點和撥號接入。定義不同區域有幫助，原因如下：

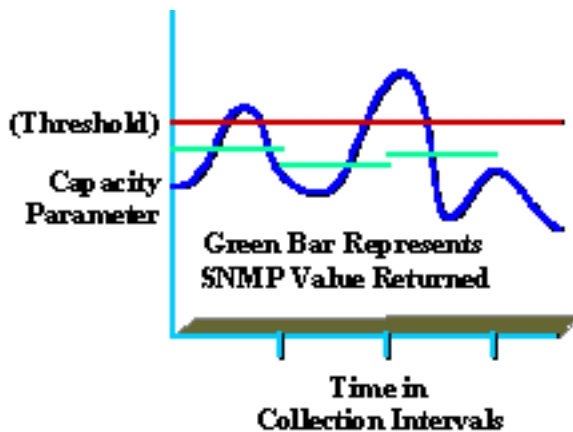
- 不同區域可能具有不同的閾值。例如，LAN頻寬比WAN頻寬便宜得多，因此利用率閾值應較低。
- 不同區域可能需要監控不同的MIB變數。例如，幀中繼中的FECN和BECN計數器對於瞭解幀中繼容量問題至關重要。
- 升級網路的某些區域可能更困難或更耗時。例如，國際電路可以有更長的提前期，並且需要相應的更高水準的規劃。

定義能力變數

下一個重要領域是定義要監控的變數和需要操作的閾值。定義容量變數在很大程度上取決於網路中使用的裝置和介質。通常，CPU、記憶體和鏈路利用率等引數是有價值的。但是，對於特定技術或要求而言，其他領域可能很重要。這些可能包括隊列深度、效能、幀中繼擁塞通知、底板利用率、緩衝區利用率、netflow統計資訊、廣播量和RMON資料。請牢記您的長期計畫，但首先從幾個關鍵領域幫助確保成功。

解釋資料

瞭解收集的資料也是提供高品質服務的關鍵。例如，許多組織並不完全了解峰值和平均利用率級別。下圖顯示了基於5分鐘SNMP收集間隔的容量引數峰值（以綠色顯示）。



雖然報告的值小於閾值（顯示為紅色），但是仍然可以在高於閾值（顯示為藍色）的收集間隔內出現峰值。這一點非常重要，因為在收集間隔期間，組織可能遇到了影響網路效能或容量的峰值。請注意選擇有意義的收集間隔，該間隔既有用，又不會造成過多開銷。

另一個示例是平均利用率。如果員工在辦公室的時間從8到5不等，但平均利用率是7X24，則資訊可能會產生誤導。

相關資訊

- [技術支援 - Cisco Systems](#)

關於此翻譯

思科已使用電腦和人工技術翻譯本文件，讓全世界的使用者能夠以自己的語言理解支援內容。請注意，即使是最佳機器翻譯，也不如專業譯者翻譯的內容準確。Cisco Systems, Inc. 對這些翻譯的準確度概不負責，並建議一律查看原始英文文件（提供連結）。