

容量和性能管理：最佳实践白皮书

文档ID20769

已更新：2005年10月04日

 [下载 pdf文档](#)

 [打印](#)

[Feedback](#)

相关产品

- [高可用性](#)
- [QoS 拥塞管理 \(排队 \)](#)

Contents

[Introduction](#)

[容量和性能管理概述](#)

[CPU](#)

[背板或I/O](#)

[内存](#)

[接口和管道大小](#)

[排队，潜伏期和抖动](#)

[速度和距离](#)

[应用特性](#)

[容量和性能管理最佳实践](#)

[服务级别管理](#)

[网络 and 应用程序假设分析](#)

[基线和趋向](#)

[异常管理](#)

[QoS管理](#)

[收集和报告容量信息](#)

[确定您的需要](#)

[定义一个进程](#)

[定义容量区域](#)

[定义容量变量](#)

[编译数据](#)

[Related Information](#)

[相关的思科支持社区讨论](#)

[Introduction](#)

高网络可用性是在大型企业和服务提供商网络内的一个目标关键需求。对提供的网络管理器表面增长的挑战高性能，包括不定期的停工、缺乏专业技术，不足的工具、复杂技术、商业合并和竞争的市场。容量和性能管理帮助网络管理器得到新的全球业务目标和一致网络可用性和性能。

本文检查以下题目：

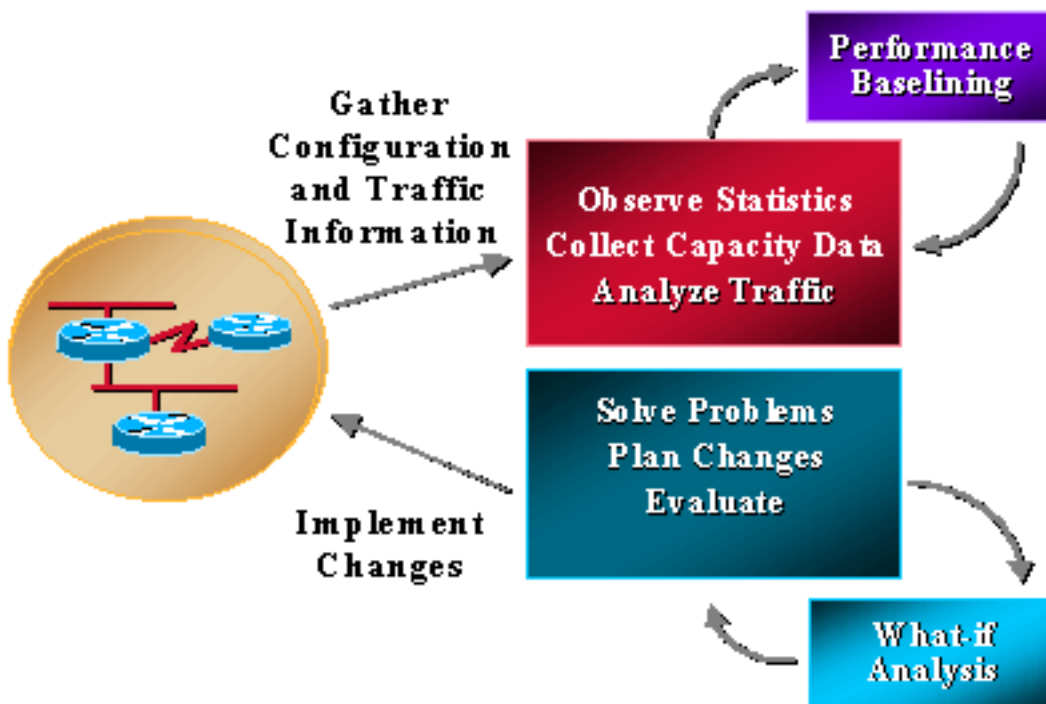
- 一般容量和性能问题，包括风险和可能性容量问题在网络内。
- 容量和性能管理最佳实践，包括若分析，基线，趋向，异常管理和QoS管理。
- 如何开发一个容量计划策略，包括普通的技术、工具、用于容量设计的MIB变量和阈值。

容量和性能管理概述

容量设计是确定网络需的资源的过程防止性能或可用性影响在商业关键应用。性能管理是管理网络服务响应时间实践、一致性和质量单个和整体服务的。

Note: 性能问题通常与容量有关。应用程序是更慢的，因为带宽和数据在队列必须等待在传输通过网络前。在语音应用，问题类似延迟和抖动直接地影响语音呼叫的质量。

多数组织一致已经收集若干涉及容量的信息和工作解决问题，计划更改，并且实现新增效益和性能功能。然而，组织若不定期地执行趋向和分析。假设分析是确定网络更改的影响的进程。趋向是实行的进程包括网络容量基准，并且性能问题和查看网络的基准趋向了解将来升级需求。容量和性能管理应该也包括问题识别的异常管理，并且解决在用户前请呼叫和网络管理员计划，管理，并且识别各自的服务性能问题的QoS管理。以下图说明容量和性能管理进程。



容量和性能管理也有其限制，典型地与CPU和内存有关。下列是关心的潜在的区域：

- CPU
- 背板或I/O
- 内存和缓冲区
- 接口和管道大小
- 排队，潜伏期和抖动
- 速度和距离

- 应用特性

在容量设计和性能管理的一些参考也提及称为“数据层面”和“控制层面的”某事。当控制层面意味着需的资源维护数据层面的适当的功能时，数据层面是容量和性能问题涉及与穿程网络的数据。控制层面功能包括服务开销例如路由、生成树、接口keepalive和设备的SNMP管理。这些控制层面需求使用CPU、内存，缓冲，排队和带宽正如穿程网络的数据流。许多控制层面需求对系统的整体功能也是重要的。如果他们没有资源他们需要，网络发生故障。

CPU

CPU乘控制层面和数据层面典型地使用在所有网络设备。在容量和性能管理，您必须保证设备和网络有一直作用满足的CPU。因为在一个设备的不适于的资源可能影响整个网络，不足的CPU能经常崩溃网络。不足的CPU能也增加潜伏期，因为数据必须等待被处理没有主CPU时，当没有硬件交换。

背板或I/O

背板或I/O是指设备能处理的数据流总量，通常被描述根据BUS大小或背板功能。不足的背板通常导致丢弃的数据包，可能导致重发和另外的数据流。

内存

内存是有数据层面和控制层面需求的另一种资源。内存对于信息是必需的例如路由表、ARP表和其他数据结构。当设备用尽内存时，在设备的一些操作可以发生故障。操作能根据情况影响控制层面进程或数据层面进程。如果控制层面进程发生故障，整个网络能降低。例如，当额外内存对于路由聚合时，是必需的这能发生。

接口和管道大小

接口和管道大小是指在所有一连接可以同时被发送的相当数量数据。这经常不正确地指连接的速度，但是数据确实不传播以从一个设备的不同的速度到另一个。硅加速，并且硬件功能帮助确定根据媒体的可用的带宽。另外，软件机制能“节流”数据依照服务的特定带宽分配。固有地有高1.54kpbs的速度功能对155mbps的和的您在服务提供商网络典型地看到此为帧中继或ATM。当有带宽限制时，数据在传输队列排队。传输队列可能有优先安排不同的软件机制在队列内的数据;然而，当有在队列时的数据，它必须等待现有数据，在能转发数据接口前。

排队，潜伏期和抖动

排队，潜伏期和抖动也影响性能。您能调整传输队列影响性能用不同的方式。例如，如果队列大，然后数据长期等待。当队列是小的时，数据下降。因为数据将被重新传输，这称为尾部丢弃并且对于TCP应用程序是可接受。然而，语音和视频不很好实行与要求对带宽或管道大小的队列丢弃甚至重大的队列潜伏期特别注意。如果设备没有充足的资源立即转发信息包，队列延迟能也发生在输入队列。这可以归结于CPU、内存或者缓冲区。

被接受的潜伏期描述从时间的正常处理时间，直到转发信息包的时间。正常现代数据交换和路由器有极低的潜伏期(< 1ms)通常情况下，不用资源约束。有数字式信号处理器转换的和压缩模拟语音信息包的现代设备可能花费很多时间，甚至至20ms。

抖动描述流应用程序的包间间隙，包括语音和视频。如果信息包到达与不同的包间间隙定时的不同的时刻，则抖动高，并且语音质量降低。抖动是排队延迟主要要素。

速度和距离

速度和距离也是在网络性能的一个要素。数据网有根据光速的一致数据转发速度。这是大约每毫秒100英里。如果组织国际上地运行客户端服务器应用程序，则他们能期待对应的数据包转发延迟。当应用程序没有为网络性能时，优化速度和距离可以是在应用程序性能的一个极大的要素。

应用特性

应用特性是影响容量和性能的最后区域。问题例如小的窗口大小、应用程序Keepalive和在网络被发送的相当数量数据与需要什么在许多环境里能影响一个应用程序的性能，特别是WAN。

容量和性能管理最佳实践

此部分详细讨论五主要容量和性能管理最佳实践：

- [服务级别管理](#)
- [网络 and 应用程序假设分析](#)
- [基线和趋向](#)
- [异常管理](#)
- [QoS管理](#)

服务级别管理

服务级别管理定义了并且调控其他必需的容量和性能管理进程。网络管理器了解他们需要容量设计，但是他们面对预算和给防止一完整的解决方案的约束雇用职员。服务级别管理是与资源问题的帮助通过定义服务的一个宣传品和创建的双向责任制附加对该宣传品的一证明的方法。您能完成此用两种方式：

- 创建在用户和网络组织之间的服务级别协议包括容量和性能管理的服务的。服务将包括报告和推荐维护服务质量。然而，用户必须准备着资助服务和所有必需的升级。
- 网络组织定义了他们的容量和性能管理服务然后尝试该服务的资助并且根据具体情况升级。

无论如何，网络组织应该通过定义包括的容量设计和性能管理管理业务开始服务的什么方面他们能当前提供，并且什么在将来计划。一项完全服务将包括网络更改的假设分析和应用程序更改，基线和趋向被定义的性能变量，异常管理的被定义的容量和性能变量和QoS管理的。

网络 and 应用程序假设分析

执行网络 and 应用程序假设分析确定一个计划内更改的结果。没有假设分析，组织感受重大的风险对更改成功和整体网络可用性。在许多情况下，网络更改导致导致许多小时生产停工期的堵塞瘫痪。另外，惊人数量的应用程序引入失效并且导致影响其他用户和应用程序。这些故障在许多网络组织继续，他们是完全可预防的用一些个工具和一些另外的规划步骤。

您通常需要一些更新过程执行质量假设分析。第一步将识别所有更改的风险级别和更高的风险更改的要求更加详细的假设分析。风险级别可以是所有更改提交的一个要求的字段。更高的风险级别更改然后将要求更改的被定义的假设分析。网络假设分析确定网络更改影响关于网络利用率和网络控制面板资源问题的。应用程序假设分析将确定项目应用成功、带宽需求和所有网络资源问题。下面的表是风险级别分配和对应测试需求示例：

风定义	更改计划推荐
-----	--------

<p>风险级别</p>	
<p>1</p> <ul style="list-style-type: none"> • 对用户(500+)或企业重要服务的大量的巨大的潜在影响由于新产品、软件、拓扑或者功能简介。 • 更改介入预计网络中断时间。 	<ul style="list-style-type: none"> • 验证新的解决方案实验室。实验室验证包括描述的显示影响的解决方案测试和验证和假设分析现有基础设施。我们推荐解决方案试验。新的解决方案要求操作支持文件的完成。 • 执行Cisco NSA设计评论。 • 创建复原计划。 • 创建实施计划。 • 创建更改过程。
<p>2</p> <ul style="list-style-type: none"> • 对用户(500+)的大量的巨大的潜在影响或企业重要服务由于数据流一个大增量或用户、骨干网变动或者路由更改。 • 更改可能需要相当的停工期。 	<ul style="list-style-type: none"> • 执行假设分析确定影响到现有的环境(在实验室环境里应该执行)。 • 功能的测试和复核路由更改。 • 创建复原计划。 • 执行专业路由或骨干网变动的设计评论。 • 创建实施计划。 • 创建更改过程。
<p>3</p> <ul style="list-style-type: none"> • 对用户或商业服务的更小的编号的中等潜在影响由于任何非标准更改。 • 包括新产品、软件、拓扑、增加功能或者新用户、增加的数据流或者非标准拓扑。 • 更改可能需要相当的停工期。 	<ul style="list-style-type: none"> • 执行对新的解决方案的工程分析(可以要求实验室验证)。 • 创建实施计划。 • 创建更改过程。
<p>4</p> <ul style="list-style-type: none"> • 降低潜在服务或用户影响。 • 添加新的标准模板网络模块，例如建立或服务器交换机/集线器的Includes在路由器。 • 带动新的广域网站点或另外的证 	<ul style="list-style-type: none"> • 创建实施计划。 • 创建更改过程。

	<p>明的访问服务的 Includes。</p> <ul style="list-style-type: none"> • 所有风险级别3变化技术上证明在生产环境上。 • 更改可能需要相当的停工期。 	
5	<ul style="list-style-type: none"> • 没有用户或服务影响。 • 添加个人用户的Includes到网络和标准配置配置更改例如密码、横幅、SNMP，或者其他标准配置参数。 • 没有停工期。 	<ul style="list-style-type: none"> • 可选的更改过程。

一旦定义了您需要假设分析的地方，您能定义服务。

您可执行网络假设分析用建模工具或与仿造生产环境的实验室。建模工具由应用程序多么恰当限制了设备资源问题，并且，因为多数网络更改是新的设备，应用程序可能不了解更改的效果。佳方法是建立生产网络的某个表示法在实验室和测试所需的软件、功能、硬件或者配置在负荷下通过使用数据流生成器。泄漏路由(或其他控制信息)从生产网络到实验室里也提高实验室环境。测试用不同的话务类型的其他资源需求，包括SNMP，播放，组播，加密或者压缩的数据流。在潜在的重点情况期间例如路由集中、链路飘荡和设备重新启动，使用所有这些不同的方法学，请分析设备资源需求。资源利用问题包括正常容量资源区域例如CPU、内存、底板利用率，缓冲区和排队。

新应用应该也执行假设分析确定应用成功和带宽需求。您在实验室环境通常执行此分析使用协议分析程序和广域网延迟模拟程序了解距离效果。您只需要PC、集线器，广域网延迟设备和实验路由器被连接到生产网络。您能通过节流在测试路由器的数据流使用通用流量整形或速率限制模拟带宽在实验室里。网络管理员能与应用组一道工作了解带宽需求、多窗口问题和潜在的性能问题应用程序的在LAN和广域网环境里。

在实施任何商业应用前执行一个应用程序假设分析。如果不执行此，应用组由于低性能责备网络。如果能为新的配置莫名其妙地需要应用程序假设分析通过变更管理流程，您可帮助防止不成功的配置和更好了解在带宽消耗的突然的增量客户端服务器和批处理需求的。

基线和趋向

在容量问题引起网络中断时间或性能问题前，基线和趋向允许网络管理员计划和完整网络升级。比较资源利用率在连续的时间或随着时间的推移蒸馏信息下来在数据库并且允许计划程序查看资源利用率参数小时、日、星期、月和去年。无论如何，某人必须查看关于一个每周，双周或者月度基本类型的信息。关于基线和趋向的问题是在大型网络要求占优势的量信息查看。

您能解决此问题是几个方式：

- 修造大量容量和交换到LAN环境，因此容量里是不是问题。
- 划分趋向信息成组并且集中网络的高可用性或重要区域，例如重要广域网站点或数据中心LAN。
- 报告机制能突出显示在特别注意的一特定的阈值上下跌的区域。如果实现重要可用性首先区域，您可以极大减少对于复核是必需的信息量。

使用所有早先方法，您仍然需要查看信息定期。基线和趋向是一个前期努力，并且，如果组织只有响应式支持服务的资源，单个不会阅读报告。

许多网络管理解决方案在容量资源变量提供信息和图形。不幸地，多数人只使用响应式支持服务的这些工具对一个现有问题;这阻挠目的对于基线和趋向。是有效的在提供容量趋向信息为Cisco网络的两个工具是和谐的网络状况产品和INS EnterprisePRO产品。在许多情况下，网络组织运行简单脚本语言收集容量信息。下面通过链路利用率、CPU利用率和ping性能的脚本收集的一些示例报告。可能是重要趋向的其他资源变量包括内存、队列深度、广播数量、缓冲区、帧中继拥塞通知和底板利用率。参考这些表关于链路利用率和CPU利用率的信息：

链路利用率

资源	地址	分段	平均利用率 (%)	峰值使用率 (%)
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

资源	轮询地址	平均利用率 (%)	峰值使用率 (%)
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

链路利用率

资源	地址	AvResT (毫秒) 09-09-98	AvResT (毫秒) 09-09-98	AvResT (毫秒) 09-09-98	AvResT (毫秒) 10-01-98
AADR01	10.190.56.1	469.1	852.4	461.1	873.2

ABNR0 1	10.190.5 2.1	486.1	869.2	489.5	880.2
APRR0 1	10.190.5 4.1	490.7	883.4	485.2	892.5
ASAR0 1	10.196.1 70.1	619.6	912.3	613.5	902.2
ASRR0 1	10.196.1 78.1	667.7	976.4	655.5	948.6
ASYR0 1S					503.4
AZWR T01	10.177.3 2.1	460.1		444.7	
BEJR0 1	10.195.1 8.1	1023. 7	1064. 6	1184	1021.9

异常管理

异常管理是识别的和解决的容量和性能问题的一重要的方法。想法是接收容量和性能门限值超越的通知为了立即调查和解决问题。例如，网络管理员也许接受高CPU的一个警报在路由器。网络管理员能登录到路由器确定CPU为什么很高。她可然后执行减少CPU的某种补救配置或建立防止引起问题的数据流的访问列表，特别是如果数据流不看来商业危急。

您能相当只是设定关键问题的异常管理使用RMON配置on命令路由器或使用更加先进的工具例如NetSys服务水平管理器与SNMP、RMON或者NetFlow数据一道。多数网络管理管理工具有设置的功能阈值和警报在侵害。异常管理进程的重要方面是提供接近实时问题的通知。否则，问题可能在任何人前消失注意通知接收了。如果组织有一致监控，这可以在NOC内执行。否则，我们推荐分页程序通知。

以下配置示例为路由器CPU提供上升的和下降阈值通知给可能一致查看的日志文件。您能设置重要链接利用率门限值超越或其他SNMP阈值的相似的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

服务质量管理介入创建和监控在网络内的特定的流量组。数据流为特定应用组提供恒定的性能(被定义在话务类别内)。流量整形参数为特定流量等级提供在优先级和流量整形的极大的灵活性。这些功能包括功能例如承诺接入速率(CAR)，加权随机早期检测(WRED)和组基于市场被加权的排队。话务类别根据更多企业重要应用和特定应用程序需求的性能SLAS通常被创建例如语音。不重要或非事务数据流也将被控制用方式不能影响更加高优先级的应用程序和服务。

创建话务类别要求对网络利用率、特定应用程序需求和商业应用优先级的基准了解。应用程序需求包括信息包大小、超时问题、抖动需求、突发传输需求、批处理需求和整体性能问题知识。使用此知识，网络管理员能创建提供在各种各样的LAN/WAN拓扑间的更加一致的应用程序性能的流量整形计划和配置。

例如，一个组织有在两个主要站点之间的10兆比特ATM连接。链路有时变得拥塞从大文件传输，导致联机事物处理和恶劣或不可用的语音质量的性能降低。

组织设置四个不同的话务类别。产生了语音最高优先级并且准许维护该优先级，即使在估计的数据流容量费率破裂了。产生了重要应用组其次高优先级，但是未允许在总链路的大小破裂估计的语音带宽需求。当它破裂，将下降。产生了文件传输流量一更加低优先级，并且其他数据流在中部适合某处。

组织必须当前执行在此链路的QoS管理确定多少数据流每个组采取和测量在每个组内的性能。如果组织不能执行此，缺乏可能为一些组发生或性能SLAS不可以在一个特定类内满足。

管理QOS配置仍然是一项困难任务由于缺乏工具。一个方法将使用思科的互联网Performance Manager (IPM)发送在落入其中每一个话务类别的链路间的另外数据流。您可能然后监控每组的性能和IPM提供趋向，实时分析和逐跳分析精确定位问题区域。其他可能仍然依靠一个手工方法类似调查在根据接口统计数据的每个话务类别内的排队和丢弃的数据包。在一些组织，此数据可能通过SNMP收集或被解析到基准和趋向的一个数据库。发送在间网络的特定的流量类型确定一个特定服务或应用程序的性能的一些工具在市场上也存在。

收集和报告容量信息

应该与三个建议使用的领域的容量管理连接收集和报告容量信息：

- 假设分析，在网络更改附近集中，并且更改如何影响环境
- 基线和趋向
- 异常管理

在这些区域中的每一个内，请开发一个信息收集计划。一旦网络或应用程序分析，您若需要工具仿造网络环境和了解更改相对潜在的资源问题的影响在设备控制层面或数据层面内的。一旦基线和趋向，您需要设备的显示当前资源利用率的快照和链路。您随着时间的推移然后查看数据了解潜在的升级需求。在容量或性能问题出现前，这允许网络管理员适当地计划升级。当问题出现时，您需要异常管理警告网络管理员，因此他们能调整网络或解决问题。

此进程可以被划分成以下步骤：

1. 确定您的需要。
2. 定义一个进程。
3. 定义容量区域。
4. 定义容量变量。
5. 编译数据。

确定您的需要

开发容量和性能管理计划要求了解您需要和该信息的目的信息。分裂计划成三个需要的区域：一每假设分析，基线/趋向和异常管理的。在这些区域中的每一个内，请发现什么资源和工具是可用的，并且什么是需要的。许多与工具配置的组织失效，因为他们考察工具的技术和功能，但是不考虑必要的人民和专业技术管理工具。包括必需的的人民和专业技术在您的计划，以及进程改进。这些人可能包括系统管理员管理网络管理位置、数据库管理员帮助与数据库管理，被培训的管理员使用和监控工具的和高水平网络管理员确定策略、阈值和信息收集需求。

定义进程

您也需要进程保证顺利和经常地使用工具。您可以要求进程改进定义跟随的什么网络管理员应该执行，当门限值超越发生时或什么进程为基线，趋向和升级网络。一旦确定需求和资源的成功的容量设计，您能考虑方法。因为他们认为服务核心能力，许多组织选择外包此种功能到一个网络服务组织例如INS或构件专业技术内部。

定义容量区域

容量设计的计划应该也包括容量区域的定义。这些是能共享一个普通的容量计划策略网络的区域：例如，公司LAN，广域网现场办公室、重要广域网站点和拨号接入。定义不同的区域由于几个原因是有用的：

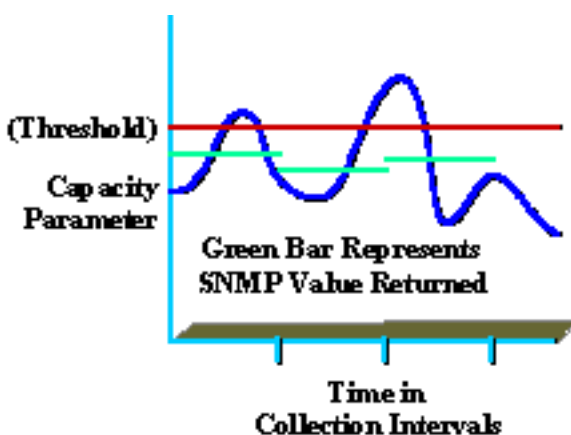
- 不同的区域可能有不同的阈值。例如，LAN带宽比WAN带宽便宜，因此利用率阈值应该是更低的。
- 不同的区域可能要求监控不同的MIB变量。例如，在帧中继的FECN和BECN计数器是重要在了解帧中继容量问题。
- 它可能更加困难或费时升级网络的一些区域。例如，国际电路能有更久的订货到交货时间和需要对应高水平计划。

定义容量变量

下个重要区域定义了要求动作的变量监控和门限值。定义容量变量极大取决于在网络内和媒体使用的设备。在一般参数例如CPU，内存和链路利用率有价值。然而，其他区域可能是重要对特定技术或需求。这些可能包括队列深度、性能、帧中继拥塞通知、底板利用率、缓冲区利用率、Netflow统计数据、广播数量和RMON数据。记住您的长期计划，但是从仅一些个关键领域开始帮助保证成功。

编译数据

了解收集的数据也是关键对提供一项优质服务。例如，许多组织不充分地了解高峰和平均利用率级别。以下图表显示根据5分钟SNMP收集间隔的一容量参数高峰，(显示为绿色)。



即使报告的值是较少比在门限值的阈值(显示为红色)峰顶能在收集间隔内仍然发生(用蓝色显示)上。这是重大的在收集间隔期间，因为，组织可能体验影响网络的性能或容量的峰值。小心选择是有用的，并且不导致额外的开销的有意义的收集间隔。

另一个示例是平均利用率。如果员工仅是在从八的办公室到五，但是平均利用率是7X24，信息可能是令人误解的。

[Related Information](#)

- [Technical Support - Cisco Systems](#)

本文档是否是有用？[有](#) [没有](#)

感谢您的反馈。

[打开支持案例](#)（需要[思科服务合同](#)。）

相关的思科支持社区讨论

[思科支持社区](#)是提出和解答问题、分享建议以及与同行协作的论坛。

有关本文档中所用的规则信息，请参阅 [Cisco Technical Tips Conventions](#)。

已更新：2005年10月04日

文档ID20769