

BIG-IP 应用优化为广域网带来的优势 [概述](#) [解决方案](#)

概述

支持 Web 的应用在当今的企业环境中已然非常普及。一家公司在网络上拥有数百个基于 Web 的不同应用的现象十分常见。随着部署这些应用变得越来越容易，企业的关注点正从应用转向安全和访问。基于 Web 的应用当前面临着众多的挑战，其中之一是满足数量呈指数级增长的用户的需求，这些用户通常远程访问此类应用。这不仅引入了新的安全问题，而且远程用户要求获得与通过局域网访问应用相同的性能水平。通过定制每个应用来解决性能和安全问题不仅费用高昂，而且时间和资源的利用效率也十分低下。凭借其在网络基础设施中重要的战略位置，全新的 BIG-IP 系统提供了一个快速、简单且更经济有效的解决方案来应对这些挑战。

BIG-IP 系统 9.x 版具有许多重要特性，可促进应用流量的优化并提高应用流量的速度。F5 Networks 解决方案中心进行了大量测试，来衡量 BIG-IP 系统为互联网上的应用和 web 站点带来的性能优势。这些测试证实了 BIG-IP 设备的应用加速技术的独特功能，从而将它与业界其它技术区别开来。尽管其他一些厂商声称自己的产品能够提高应用性能，但他们的测试都是在独立的实验室环境下进行的。这给用户带来了很大困难，因为他们必须设法确定这些测试结果转变为实际性能增益的情况如何。而这正是 F5 计划解决的问题。为此我们得出了一个结论，即测试用户在互联网上看到的真正性能的唯一方法，就是实际使用互联网。

为了执行这些类型的实际测试，F5 选择 Gomez 性能网络（后文将详述）来测试端到端交易。根据测试结果，BIG-IP 系统将最终用户的响应时间缩短两倍或更多倍，带宽利用率减少 75% 以上，较慢链路上的浏览器超时问题减少 80% 以上，从服务器上卸载的连接多达 98%。

解决方案

加快应用速度并缩短最终用户响应时间

BIG-IP 设备包含目标和专门的优化措施，充分利用了 F5 独特的 WAN、LAN 和数据加速技术。这些优化措施提供了无与伦比的优化和数据包丢失恢复功能，并且能够更智能地在部分优化的服务器和客户端之间进行调度。表 1 显示了一些典型应用在响应时间方面的改进。

表 1：完整应用交易性能改进示例

Application	Not Optimized	BIG-IP Optimized	Improvement
BEA WebLogic Portal 8.1	38.209 seconds	17.291 seconds	121% (2.2x)
Microsoft IIS 6.0	21.09 seconds	12.39 seconds	70% (1.7x)
Microsoft Outlook Web Access 2003	32.88 seconds	21.23 seconds	55% (1.3x)
Microsoft SharePoint Portal Services 2003	40.6 seconds	18.06 seconds	125% (2.2x)
Siebel Business Applications 7.7	88.157 seconds	33.989 seconds	126% (2.3x)

改进服务器群功能

除了提供更好的用户体验外，BIG-IP 设备还可以显著提高周围基础设施的可扩展性。

利用全新的快速缓存 (Fast Cache) 特性，BIG-IP 系统平均能够帮助后端服务器卸载 36% 的内容服务和 95% 的 TCP 连接，从而显著提高了现有基础设施的性能和可扩展性。

通过上述和其它优化技术，我们发现在网络中安装 BIG-IP 系统后，服务器能够处理的工作量一般为原来的两倍。

提高带宽效率

在未经优化时，由于协议和 WAN 效率低下，企业通常只能利用所购买的带宽的一小部分。通过流量管理操作系统 (TM/OS) 和 TCP Express 特性集，测试显示 BIG-IP 设备可显著提高站点的带宽效率。例如，通过使用 BIG-IP 系统，测试显示线缆上的数据量平均增加 224% (3.2 倍)，线缆上的数据包平均减少 50%。总体而言，BIG-IP 系统平均可将带宽的利用效率提高 322% (4 倍)。

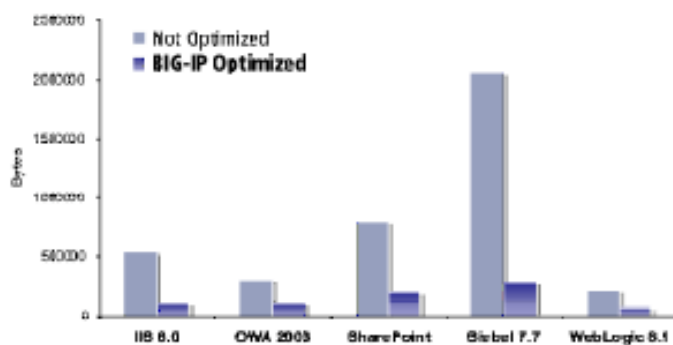


Figure 1: Sample bandwidth reductions using Intelligent Compression and TCP Optimizations

提高 WAN 连接的可靠性

利用 TCP Express 特性集，测试显示 BIG-IP 系统将客户端可能遇到的 TCP 超时和重设次数减少了多达 50%。这对于在高损耗网络上进行传输或使用低带宽连接（如拨号连接）的客户端而言尤为有益。

利用 Gomez Performance Network 在真实环境中进行测试

当 F5 计划以一种能够为客户带来价值的方式测试应用优化功能的性能时，我们必须找到值得信赖的第三方服务，使我们能够访问拥有不同网络带宽和不同操作环境的全球用户群。在对多个可用解决方案进行评估之后，我们最终选择 Gomez Performance Network 进行测试。

屡获殊荣的 Gomez Performance Network 由全球 10,000 多台计算机通过实际最终用户互联网连接构成，在我们的许多企业级客户中均得到了广泛应用。它荣获了多项大奖，其中包括《网络计算》(Network Computing) 杂志的编辑选择奖。作为测试工具，其网络使得 F5 能够通过互联网上对实际环境中的最终用户性能增益进行公平取样。

借助 GPN，F5 解决方案中心的测试结果比业界其他公司提供的市场宣传材料更加贴近现实。通过在实验室环境下谨慎选择测试的页面和动态应用内使用的数据，创建显示 10 倍、20 倍或更多倍改进的测试轻而易举。但是如果这些数字不能代表大多数用户可能遇到的情况，又有什么意义呢？我们得出的结论是，当我们的现有客户和潜在客户选择应用加速产品时，这些实验室产生的结果并没有什么帮助——人们希望看到其最终用户能够实际体验到的性能提升。

F5 认为展示应用性能提升的最佳方式就是选择用户一般在 WAN 上可以看到的应用。鉴于此，我们选择了“门户”和“协作”应用，如 BEA WebLogic、Siebel Business Applications 和 Microsoft SharePoint 等。我们还将 Microsoft Outlook Web Access 作为用户在 WAN 上使用的一个关键应用，并提供了一个基于 Microsoft Internet Information Services (IIS) 的通用网站 (www.f5.com)，以了解可为大多数静态内容带来的总体性能增益。由于企业可能运行大量应用，因此此处看到的结果将能够很好地显示出这一技术可带来的应用增益。

对于每一次测试，我们都会测量最终用户在使用应用过程中的响应时间体验。这一过程通常持续数天，并重复进行数千次测试。例如，我们用来测试 Microsoft Outlook Web Access 的脚本包括登录、信息检索、日历查看等。正如您所知，这与单页下载时间完全不同。它代表了在整个应用交易期间最终用户所体验到的响应时间，如同用户在实际环境中体验到的情况。

测试方法

Gomez 提供了一项服务，使客户能够像 Gomez 监控代理一样监控网站的性能，这些代理运行在遍布全球和整个互联网的用户计算机之中。Gomez 付费向企业和最终用户租借计算机域带宽资源。然后，Gomez 在计算机上安装一个代理，进而能够指示该计算机连接到指定的站点。每隔一段时间，Gomez 会指示一定数量的计算机请求一个特殊网站，以衡量其各个方面的性能。然后，这一信息被返回到 Gomez Performance Network，以便客户能够从最终用户的角度查看目标网站的性能。

除选择请求的网站，Gomez 还支持客户选择使用哪一用户群体（或“对等群体”），以及指定他们希望用来收集性能指标的客户端连接类型。例如，F5 选择了存在客户端的各种区域，以及它们所属的带宽类型。Gomez 提供了一个记录器，以方便对客户端使用的示例客户交易实例进行配置，简化对于在客户端上进行的客户交易实例的配置。这使得客户端能够指定在每个网站上选择哪些链路、使用什么登录证书（如果适用）、以及模拟真实用户体验所需的其它交易设置。当用户会话范例使用该记录工具进行配置后，Gomez Performance Network 上选定的代理就可以重现这一会话。

在测试用户通过互联网上体验到的响应时间的过程中，我们发现以下几点对于控制 Gomez Performance Network 最为重要：

- ☆ “最后一公里”测试——这些测试使用了真实环境中的客户端。
- ☆ 定义代表目标受众和地理位置的带宽类别。
- ☆ 关键指标“端到端”响应时间和“错误”——其它统计数据有助于减少任何问题。
- ☆ 记录用户会话实例而非进行单一页面下载——捕获典型用户会话，而不是简单的页面下载。
- ☆ 对经过数天测试得到的结果数据取平均值，以获得真正的平均数，并致力于通过修改应用或采取性能增强技术提高平均值。

☆ 正如许多最终用户一样，Gomez 代理支持所有标准浏览器行为和特性，但是在接受第三方插件方面的能力有限——应用必须能够处理此类可能无法支持插件的安全客户端。我们的测试参数相当简单，测试用户来自全球拨号上网用户、美国拨号上网用户和美国低带宽用户。对于每一应用，我们均使用了一个涉及以下步骤的用户会话实例：

- ☆ 转到主页面。
- ☆ 点击按钮，输入登录证书。
- ☆ 输入登录证书。
- ☆ 点击链路 A。
- ☆ 点击链路 B。

这只是一个高级示例，实际上，为了测试出用户通过每种应用获得的体验，这些步骤可以进行变化。

在Gomez确定了我们的范例用户群且我们记录了用户体验实例之后，我们开始利用该实例，同时衡量两个应用例程的性能指标。一个应用例程采用了 BIG-IP 系统进行加速；另一个应用流程采用了常见的服务器负载平衡器，只进行了简单的转换，而没有加速。这两个例程处于具有相同配置和硬件的同一位置。在完成所有设置之后，Gomez 将在几天内重复进行大约 2000 次测试，以获得一个有代表性的实例。

BIG-IP 优化技术的特性与优势

本文描述的测试和结果适用于 BIG-IP 应用加速器产品，以及具备相关加速与优化特性的 BIG-IP 本地流量管理设备。以下列表包含了它们可提供的特性与优势。

快速缓存：内存中 RAM Web 缓存

- ☆ 卸载服务器请求，能够独自满足有关通用网页/包含的对象/图像请求。
- ☆ 当与“智能压缩”相结合时，通过允许缓存和提供压缩的对象，能够进一步加快页面下载速度，而无需重复压缩同一个对象。
- ☆ 同时保存压缩和未压缩的内容，并根据客户端将接受的内容智能地提供正确的页面。

- ☆ 完全符合 RFC2616 要求。
- ☆ 缓存各种服务器响应代码：200, 203, 206, 300, 301 或 410。
- ☆ 能够代表服务器对条件 GET 和 HEAD 请求做出响应。
- ☆ 允许在一个系统上安装多个专用缓存库（“多存储器”），以将缓存资源导向一个共享系统上的高优先级应用。
- ☆ 支持 F5 的高级编程语言 iRules，以便加强对可缓存内容的控制和管理。

智能压缩：免代理 (Agent-less) 压缩卸载

- ☆ 缩短页面下载时间——更少的数据包，更短的往返时间。
- ☆ 减少带宽消耗——使用更少的带宽为相同数量的用户提供服务。
- ☆ 能够只针对采用拨号连接的客户端或那些属于高延迟（长距离）连接（正在申请专利）的用户进行压缩。
- ☆ 支持最近 5 年来出现的所有 Web 浏览器——无需插件或任何其它种类的软件。
- ☆ 从服务器层卸载服务器并集中管理压缩处理过程，从而提供一个成本更低、安全（通过精细的客户端定位）且易管理的解决方案。

TCP Express: 行业领先的一流 TCP 优化

- ☆ 众多 WAN 和 LAN 优化技术，可根据不同客户端和网络条件加快数据传输速度。
- ☆ 众多 TCP 互操作性改进，适用于各种商业堆栈（Windows 98、XP、2000、IBM AIX、Sun Solaris 等）。
- ☆ 集中式/无客户端 WAN 优化提供了对称和不对称优化技术，而无需客户端下载或分支设备。
- ☆ 根据开放标准 TCP 优化：
- ☆ 延迟或选择性应答 (Delayed and Selective Acknowledgments) (RFC 2018)：当处理 WAN 上丢失和重新排序的数据包时，提高性能。
- ☆ 显式拥塞通知 (Explicit Congestion Notification) (RFC 3168)：使 BIG-IP 系统能够前瞻性地同类设备发出调度路由器将超载的信号，以便它们能够采取避退 (back off) 措施，避免数据包丢失。
- ☆ 有限且快速的重新传输 (Limited and Fast Re-Transmits) (RFC 3042 和 RFC 2582)：能够高效地重新传输丢失的数据，消除由于数据包丢失导致的超时影响。
- ☆ 适应性初始拥塞窗口 (Adaptive Initial Congestion Windows) (RFC 3390)：研究表明，通过卫星链路进行的 HTTP 传输可实现 30% 的改进，28.8 bps 拨号连接可实现 10% 的改进，而跳线率没有相应增加。
- ☆ 慢启动和拥塞避免 (Slow Start with Congestion Avoidance) (RFC 2581)：
- ☆ TCP 慢启动 (TCP Slow Start) (RFC 3390)：能够提高链路间带宽利用率，以提高现有公共互联网连接和租用线路的吞吐率。
- ☆ 带宽延迟控制：改进且扩展的带宽延迟计算可以确定网络的最佳负载，从而不会使网络负载过重。
- ☆ 时间戳和窗口尺寸 (TimeStamps and Windows Scaling) (RFC 1323)：BIG-IP 允许有选择性地使用将数据添加到 TCP 片段的时间戳，来实现其它优化。

L7 速率调整：能够划分带宽利用率的优先级

☆ 能够使用 iRules 按照 L2-L7 信息划分（选择）流量——这意味着可以根据 MAC 地址、IP 信息或 L7 信息（如 HTTP cookie 等）跟踪流量。

☆ 包含允许上级-下级带宽借用关系的“速率等级”结构——例如，多个 FTP 客户可能有唯一的速率政策，但是如果资源可用，就可以允许“借用”。

☆ 能够独立对入站和出站流量进行分类。

☆ 支持“高优先级 FIFO (Priority FIFO)”和“随机平等队列 (Stochastic Fair Queue)”排队规则。

☆ 可配置的猝发、借用和极限速率。

☆ 在 iRules 内，能够动态应用到虚拟服务器。

SSL 卸载：对 SSL 进行解密，以便您的服务器无需：

☆ 增加服务器容量——通过从服务器卸载资源密集型 SSL 处理负载，服务器将有更多资源来处理其真正的特定业务对象。

☆ 缩短页面下载时间——F5 的最新高性能加密 ASIC 具备强大能力，能够以比普通服务器 CPU 高多倍的速度处理加密数据。

☆ BIG-IP 应用加速器设备可处理高达 5000 个 TPS（每秒处理交易），功能强大的 BIG-IP LTM 6800 平台可处理高达 20,000 个 TPS。为新连接和专用硬件卸载引擎内的数据加密提供了独特支持。

☆ 能够智能重写从 HTTP 到 HTTPS 的 HTTP 重导向程序，以帮助无缝将 SSL 集成到现有的 HTTP 应用中。

☆ 与 iRules 的紧密集成支持在加密强度、客户证书和其它 SSL 信息的基础上，灵活地制定政策决策。

☆ 允许对加密流量进行深入的应用流量剖析和修改。

OneConnect TCP卸载：通过多路传输 HTTP 请求减少 TCP 开销：

☆ 将多个客户端连接汇聚到较少的服务器端连接。

☆ 避免请求延迟或排队——保持开放足够的服务器端，以同时处理所有连接。

☆ 改变 HTTP 标头，以支持长时间的服务器端连接。

☆ 独立处理客户端和服务器连接——该特性能带来令人难以置信的效率，可以将数百万个连接转变为后端服务器必须处理的几百个连接。

内容缓冲——服务器缓冲：

☆ 能够根据传输速度获取服务器响应，避免了与速度较慢的客户端直接通信的负担。

☆ 卸载重新传输处理载荷和优化单个数据流，以便为最终用户带来最佳性能，同时根据实际接收速度尽快发送数据。

☆ 通过更快地从服务器读取和缓冲数据，服务器可以有更多资源来处理更多连接，从而增加服务器容量。

