

基于 Riverbed 的 ITER 广域网优化方案^①

杜庆庆^{1,2}, 谭海波¹, 李晓风¹

¹(中国科学院 合肥物质科学研究院, 合肥 230031)

²(中国科学院研究生院, 北京 100049)

摘 要: 为了满足参与国际热核聚变实验堆(ITER)各采购包对 CN DA 和 ITER 之间的网络性能不断增长的需求, 给出一个基于 Riverbed 的 ITER 网链路优化加速解决方案. Riverbed 实现了三层优化系统 RIOS, RIOS 在数据层、传输层和应用层实现了数据优化和协议优化, 此外 RIOS 支持丰富的 QoS 标志和加速设备负载均衡和冗余的高可用性集群技术, 同时 RIOS 实现了管理优化, 使得网络部署更加容易和快捷.

关键词: ITER 网链路优化; RIOS; 数据优化; 协议优化

Optimization of ITER Wide Area Network (WAN) Based on Riverbed

DU Qing-Qing^{1,2}, TAN Hai-Bo¹, LI Xiao-Feng¹

¹(Information Center, Hefei Institutes of Physical Science, Chinese Academy of Sciences, Hefei 230031, China)

²(Graduate University, Chinese Academy of Sciences, Beijing 100049, China)

Abstract: To meet the growing demand for network performance between CN DA and ITER for each procurement package of participating in the International Thermonuclear Experimental Reactor (ITER), an optimization of WAN based on Riverbed is given. Riverbed has a three-layer optimizing system RIOS, which implements data streamlining, transport streamlining and application streamlining to achieve data and protocol optimization. In Addition, RIOS supports diverse Quality of Service (QoS) and high-availability-clustering functionality which makes system load-balancing and redundant, meanwhile, RIOS has realized the management optimization which makes the network deployment easier and faster.

Key words: WAN optimization; RIOS; data optimization; protocol optimization

国际热核聚变反应实验堆(以下简称 ITER)是目前当今世界最大的国际大科学工程项目之一, ITER 计划由欧盟, 中国, 美国, 韩国, 俄罗斯, 印度以及日本等七方共同参与 ITER Tokamak 装置的设计, 加工, 制造, 并实现稳态运行. 为了实现中方能够与 ITER IO 和其他各参与方之间远程协同设计, 需要通过一个专用高速数据网络实现异地同步的共享环境, 完成项目的设计, 研制, 项目合同, 计划进度管理, 质量保障管理, 文献管理, 以及数据共享.

为了满足上述应用需求, 专用数据网络的带宽必须达到 100Mbps, 并且与 ITER IO 之间的互访将通过 Orinet 实现, 误码率小于 10e-7, 网络实用性达到 99.9%.

通过广域网加速技术和 IPSEC 等加密保护技术, 以合肥为中心, 成都, 北京为分站点建设一个高速接入, 安全, 存储以及数据共享等功能的网络架构接入 ITER 专用高速数据网络中.

1 研究背景

随着 Internet 的迅猛发展, 高带宽网络的出现, 以及很多对数据敏感的应用越来越多, 如音视频应用等, 使得拥塞变得越来越严重, 特别是高带宽和高时延网络, 这种网络通常称之为长肥网络(Long Fat Network, LFN), 也称之为高带宽时延乘积网络 (High-Bandwidth-Delay-Product Network, BDP). 带宽时延乘积(BDP)通常

① 基金项目: ITER 国际高速专用数据网的设计、测试与构建资助项目(2008GB111000)

收稿时间: 2013-01-26; 收到修改稿时间: 2013-03-11

表示网络通道的容量,也就是能够在网络中缓冲的数据量,显然带宽增大一倍或者时延增大一倍都会使得通道的容量加倍。当这个乘积变得越来越大时,传统 TCP 协议的端到端应答机制以及其慢启动和拥塞控制机制,使得端到端延迟增大,导致传输效率及应用程序效率低下和带宽不足的情况,传统 TCP 协议^[3]的局限性暴露无遗。

根据 ITER 国际组(简称 ITER IO)IT 方面的管理要求,需要在各成员国内部署 Satellite 服务器组,从而实现 ITER IO 与各成员国设计人员之间的远程协同设计。ITER 专网与 ITER IO 之间的网络面临两大问题:1)地域造成的超长传输距离;2)跨越多个国际运营网络所带来的传输带宽不确定性。为降低以上问题对网络传输质量的影响,需在网内接入加速设备,提高网络的传输速率。

2 优化系统RIOS

2.1 RIOS 优化、加速方案

Riverbed 针对广域网带宽不足、高延迟环境中传输协议效率低下和高延迟环境中应用程序协议效率低下^[1]等问题,研发了 RIOS 优化系统,该系统从数据层、传输层和应用层实现广域网链路优化,并提供管理优化功能实现简单的、友好管理。RIOS 在数据层面上实现了基于数据优化的缓存机制和数据压缩机制,并结合使用连接池和扩展 TCP 窗口实现传输层的基于 TCP 协议的优化,以及在应用层上实现了 HTTP、HTTPS 和 Lotus Notes 等应用层协议的加速。RIOS 结合三个层级的优化,并支持服务质量(QoS)、服务等级 (CoS)、MPLS 标签或其他使用 DSCP 字段的任何分类机制使用 DSCP 位标记^[1],最大限度的提高 WAN 带宽利用率,保证广域网数据服务(WDS)实时性和可靠性,同时在数据安全方面实现了 SSL 和 IPSEC 加速功能,在保证链路优化的同时增强了数据安全性能。RIOS 还提供了服务平台(RSP),为特定的软件提供数据和应用程序交互的专有平台,实现服务扩展和优化。

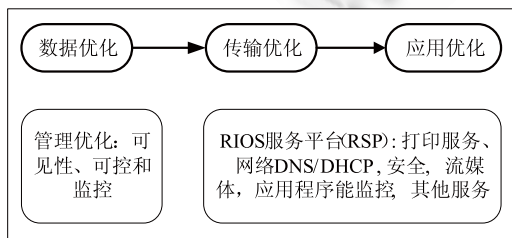


图1 RIOS 优化系统

2.2 RIOS TCP 代理体系结构

RIOS 用作透明的 TCP 代理,其体系结构如下图 2 所示,客户端或服务器端在与 Riverbed 设备监视 TCP 连接期间,其通过三个背对背的 TCP 连接从逻辑上实现单个端到端的 TCP 连接。TCP 代理连接按 1:1 的比例建立,无需任何封装、流量混合或隧道配置。在客户端或服务端看到的两个“外部”连接与未优化的单个连接相同,而“内部”连接对于客户端或服务器是不可见的,这样 RIOS 就可以对通过广域网的传输流量进行各种性能改善,而不会影响也无需重新配置客户端、服务器端或路由器。

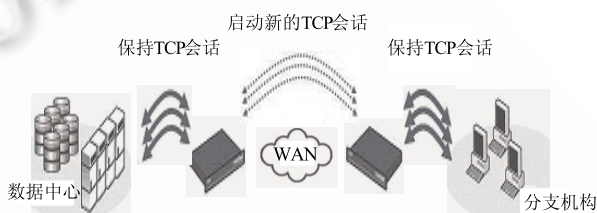


图2 RIOS TCP 代理体系结构

3 RIOS优化技术

3.1 数据优化

RIOS 对所有基于 TCP 的应用程序和应用层协议数据进行优化处理,RIOS 获取 TCP 数据流并分析 TCP 数据,以此建立数据的引用,数据和数据的引用会同时发送到远端设备上,并且通信的两端都会保存数据和数据引用,构成数据缓存。当发生数据改变时,比较已引用数据和磁盘数据,如果磁盘数据与引用数据不同,数据的改变部分以原始数据和数据引用的方式被发送到远端,而没有改变的数据,数据引用替代磁盘数据被程序发送到远端。

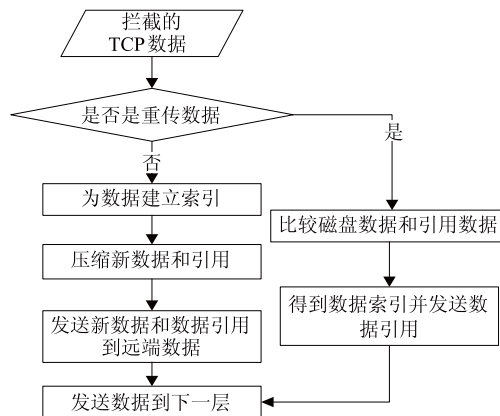


图3 RIOS 数据压缩流程

RIOS 采用分层结构模式的数据引用, 一个引用可以表示数兆字节的数据, 如此对于需要重传的数据, 传输数据引用的代价将远远低于传输数据本身的代价, 因此能够增大网络的有效负载, 此外 RIOS 采用无损压缩算法 Lempel-Ziv (LZ)^[2]对数据进行压缩, 压缩率峰值大于等于 100:1, 由于 RIOS 数据优化作用在字节序列层, 所以 RIOS 只传输字节序列发生变化的数据, 因此 RIOS 的数据优化和传统的 Cache 技术有本质的区别, 能够最大限度的减少 WAN 上传输的重复数据, 例如当文件只改变文件名时, RIOS 不需要重新构建数据引用, 也不需要重新传输所有的数据, 因此能够有效利用本地缓存, 从而减少网络传输重复数据, 并且 RIOS 还支持所有 TCP 和 UDP 的 QoS 标记, 以此实现更加灵活的数据包传输控制。

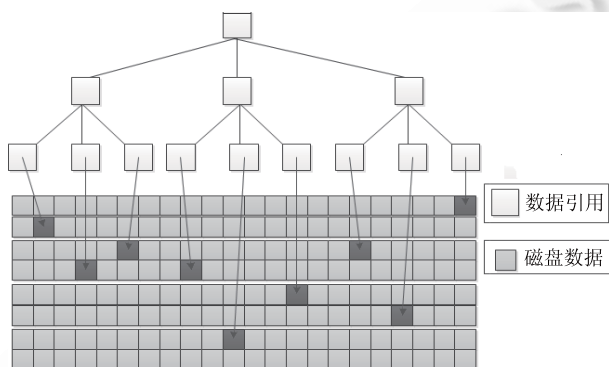


图 4 数据引用和磁盘数据

3.2 传输优化

RIOS 针对不同的网络环境提供了 MX-TCP 和 HS-TCP 技术^[12], 在高速链路的网络中使用 HS-TCP, 能够保持较高的传输窗口^[11]; 而在高延时、高丢包率的网络中使用 MX-TCP, 这两种方案都能有效的解决传统 TCP 不能够充分利用网络带宽资源的问题。同时 RIOS 维护一组打开的 TCP 连接, 构成 TCP 连接池 (Connection pool), 当用户请求 TCP 连接时, 设备使用一打开的连接, 以减少 TCP 建立连接时的资源消耗。TCP 连接池的连接都是短暂 TCP 连接, 并且不是通过回收已使用的 TCP 连接, 而是在数据传输前提前打开的连接^[1]。

同时, RIOS 传输优化通过自动侦测来自动调整扩展 TCP 协议窗口、智能数据包重新封装机制、链接管理和其他的协议优化技术相结合来减少广域网传输信息的往返次数^[1], 同时保持传输的可靠性和弹性。RIOS 扩展 TCP 窗口, 改善 TCP 的吞吐量, 以提供带宽的利用率, 并借助软件虚拟扩展(VWE)技术, 进一步增大

TCP 的有效负荷。RIOS 通过数据优化后, 传输层上传数据是由磁盘数据和数据引用构成有效负载, 而没有经过 RIOS 优化的 TCP 传输数据的有效负载是磁盘数据, 因此数据优化进一步增加了 TCP 的有效负载。

3.3 应用程序优化

RIOS 应用程序优化模块可消除应用程序协议之间交互的往返次数, 并支持多种应用层协议优化, 包括 CIFS、NFS、MAPI、HTTP(S)、MS-SQL、Oracle 11i 等, 例如 RIOS 正对 HTTP, 提供了一个并行算法用于追踪网页请求对象和并行预取。RIOS 能够动态实现数据简化和压缩, 能够有效缩短了备份或复制操作所需的时间, 减少处理数据灾难时执行数据备份和复制时间, 保障和优化广域网数据服务(WAN Data Service, WDS)实时性、可用性和安全性。

3.4 网络部署

Riverbed 广域网加速设备在实际网络中有三种部署模式, 一是物理的 in-of-path 模式, 二是逻辑的 in-of-path 模式, 三是 out-of-path 模式, 在 in-of-path 模式中, Riverbed 设备作为数据链路(data-path)中的设备之一, 服务器和客户端是可以见的, 即通过 IP 地址可以发现服务器和客户端, 然而物理 in-of-path 模式和逻辑 in-of-path 模式在实现上有本质的不同, 物理 in-of-path 直接在数据链路上转发数据, 而逻辑 in-of-path 则需要启用 4 层交换、路由策略、WCCP(Web Cache Communication Protocol)等包转发机将数据转发到加速设备中, 对于 out-of-path 模式, Riverbed 设备则作为一个代理服务器(proxy server), 服务器和客户端是不可见的, 即客户端和服务器的通信必须借助于作为中继代理服务器的 Riverbed 加速设备。三种部署模式有不同的网络应用环境, 当链路带宽小于 Riverbed 加速设备的处理能力时, 可以使用物理 in-of-path 模式, 反之则使用逻辑 in-of-path, 而当要求服务器和客户端不可见或者不能够使用 in-of-path 时, 采用 out-of-path 模式, 并且要求客户端工作在 in-path 模式。

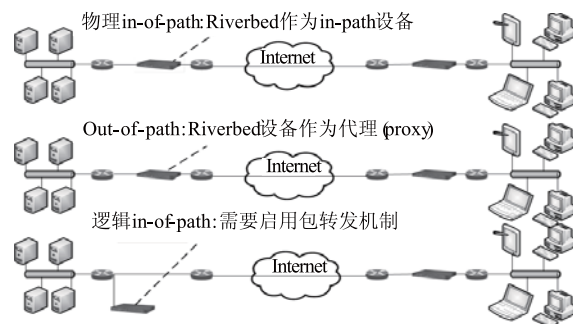


图 5 Riverbed 设备部署网络拓扑

4 测试结果及分析

本次测试以 ITER 专网和 ITER IO 之间的国际链路为测试背景,从 ITER 专网的实际应用环境出发,测试网内加入 Riverbed 设备后一个月内 LAN 和 WAN 的吞吐量,比较之间的差异.以下图 6 和图 7 分别是一个月内 LAN 和 WAN 的吞吐量示意图,图 8 为带宽优化情况.

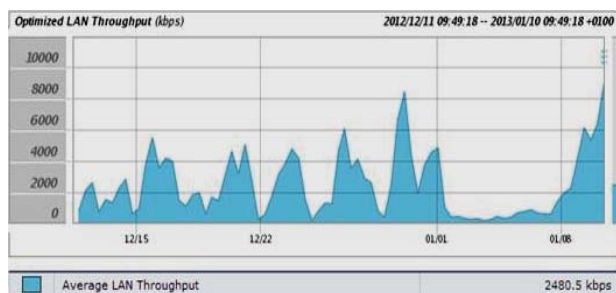


图6 LAN 吞吐量

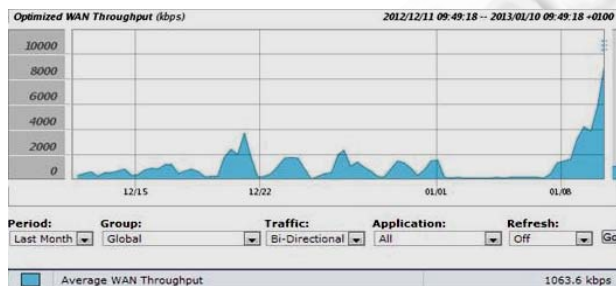


图7 WAN 吞吐量

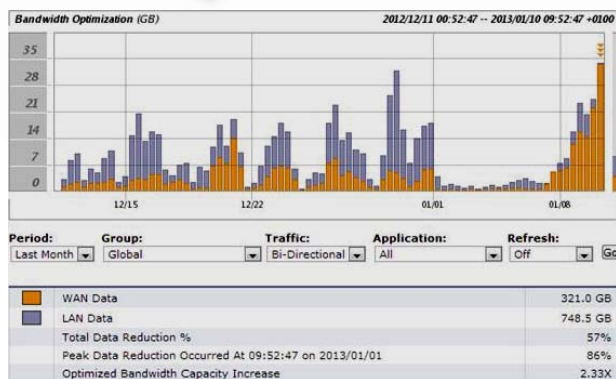


图8 带宽优化

Riverbed 加速设备使用 MX-TCP 和 HS-TCP 代替传统的 TCP,能够充分利用广域网带宽,增强超长距离的数据传输,同时通过本地的缓存和其索引机制能够有效减少重复数据的重传,由上图 6 和图 7 可看出,一个月内,LAN 端的平均吞吐量在 2.48M,但与此对应的发生在 WAN 端的平均吞吐量则为 1.06M,大大减少了 WAN 上数据量的传输.由上图 8 中从带宽的优化看,总的数据传输优化了 57%,减少的最大数据传输量达

到 86%,优化后带宽容量增加了 2.33x.由此可见,基于 Riverbed 的广域网链路优化加速方案通过数据优化、传输优化、应用优化等技术减少了 WAN 链路中数据量的传输,增强了 ITER 国际链路中超长距离的数据传输、显著提高了链路带宽的利用率,有效的承载了 ITER 大科学工程的数据通信任务.

5 结语

本文介绍和分析了基于 Riverbed 的广域网优化技术和部署方案, Riverbed 的优化核心是 Riverbed 优化系统 RIOS, RIOS 采用三层优化技术相结合来实现广域网链路优化,并支持丰富的 QoS 标志,同时在数据安全上提供了 SSL 和 IPSec 加速,实现了广域网数据服务的安全性和实时性.同时 RIOS 提供多种部署方案,适应不同网络的需求,以及加速设备的高可用性.集群实现设备的负载均衡和冗余,此外 Riverbed 提供了管理优化模块,实现加速设备的简单部署和管理.

参考文献

- 1 Riverbed Technology Riverbed.优化系统(RiOS)6.5,2009.
- 2 Zehe C. The Lempel Ziv Algorithm January16, 2003.
- 3 高文宇,李绍华.TCP 拥塞控制研究综述.信息技术,2009,(3).
- 4 任勇毛,唐海娜,李俊,钱华林.高速网络 TCP 改进协议 NS2 仿真性能比较.计算机工程,2009,(2).
- 5 孙珍,李世银,王秀娟,徐京.基于自相似流量预测的 TCP 拥塞控制算法.计算机工程设计,2010,(12).
- 6 翟冰,杨阳,孟晓波.高速网络中 TCP 拥塞控制研究综述.计算机与数字工程,2008,(8).
- 7 马宗鄂,谭海波,李晓风.基于 XTP 的 ITER 广域网链路优化方案研究.计算机工程与计,2012,(1).
- 8 冯开明.可控核聚变和国际热核试验堆(ITER 计划书).中国核电,2009,(3).
- 9 Gartner RAS Core Research Note,AndyRolfe,JoeSkorupa,Severine Real Magic Quadrant for WAN Optimization Controllers,2009.
- 10 贾亚军.一种新的网络传输协议-XTP.计算机系统应用,1997,16(5).
- 11 任勇毛,唐海娜,李俊,钱华林.高速长距离网络传输协议软件学报,2010,21(7):1576-1588.
- 12 黄小猛,林闯,任丰源.高速传输协议研究进展.计算机学报,2008,(1).