

## 提升企业灾难恢复之道： WAN 优化驱动高级复制技术

2010 年 1 月



如今的分布式和动态企业越来越依赖于 24×7 全天候访问那些日渐增多的关键业务应用和敏感数据。这些应用和数据比以往任何时候都更加分散：它们可能位于公司数据中心、远程办公室和/或用户的计算机上。此外，每一个行业领域的整体数据量都在迅速增长，而虚拟化技术的广泛采用也意味着服务器与数据的移动化程度比以往更高。而且，在艰难的经济环境下，IT 运营团队只能依靠持平甚至不断减少的预算来苦苦支撑。这些挑战使得灾难恢复 (DR) 规划比过去更为困难，但也使它比以往更加重要。

为了满足当前对应用和数据可用性的需求，成功的企业正越来越多地依靠广域网 (WAN) 作为灾难恢复的存储传输资源。这使得灾难恢复操作可以集中进行，从而减少冗余和降低管理费用，并利用领先存储厂商所提供的创新、基于磁盘的备份与复制技术。分散式的、基于磁带的灾难恢复策略成本太高，且属于劳动密集型工作。在实践中，它们不能满足恢复时间目标 (RTO) 和恢复点目标 (RPO)，而这正是企业在面对越来越严格的客户服务和法规要求时所必需的。

在本文中，我们调查了使企业范围内的灾难恢复规划变得复杂化，并使得其成本增加的业务和技术趋势，还总结了磁盘到磁盘备份与复制技术已被证实的优势。然后，我们更加深入地探讨了当与这些数据保护解决方案一起部署时，WAN 优化在释放灾难恢复的效率方面所扮演的关键角色。WAN 优化使企业能够在当前的网络容量下完成更多的工作（更频繁、更快速的备份与复制，以及更快的恢复），同时还能快速、高效地利用新的容量。我们得出的结论是：WAN 优化解决方案与高级复制技术相结合，能够为多数据中心企业的灾难恢复提供出色的灵活性、性能和成本效益。

### 传统灾难恢复的局限

过去，灾难恢复被设计为用来配合传统的 IT 领域，如物理服务器、存储阵列、网络设备、应用等，并且主要是基于资源的过度供应问题。如果可能的话，大多数的服务器和数据存储都是在本地备份到磁带，需要本地 IT 员工

来管理备份软件、计划安排、磁带库和异地归档。在发生故障时，必须协调多个复杂的流程来单独地恢复和重新配置服务器及数据集，并且通常需要在多个地点进行。因此，恢复时间往往过长并且不可预测。

## 解决方案分析

分布式、基于磁带的备份还存在地理位置上的制约：远距离运送磁带的费用可能高得让人望而却步，并且它们运送的距离越远，在发生灾难时进行恢复所需的时间就越长。这使得许多企业将恢复站点设在离主站点过近的位置，大大增加了电网故障、飓风等大面积影响事件发生时导致灾难性故障的风险。

当灾难恢复策略过度依赖于备份时，这些变通方法是必要的；理想情况下，数据应被复制到配备一组镜像服务器和存储平台的“热”或“温”远程站点。最高效的灾难恢复架构是基于数据中心到数据中心（或站点到站点）的复制，以便消除备份介质和处理过程所带来的成本和延迟。然而，对于数据中心之间的复制来说，无论同步或是异步，都只是将成本负担从备份介质转移到网络上，而这方面高昂的带宽成本仍然是为灾难恢复而广泛部署复制措施的一个主要障碍。

### 当前的行业趋势让灾难恢复变得更复杂

在 IT 运营团队努力提供更高级别可用性的同时，当前的趋势又带来了额外的挑战：

**更多的工作成为“关键任务”**：在美国大中型企业中，高达 60% 的应用被视为关键型业务，包括电子邮件、协作以及内部网应用和数据。这一演变需要灾难恢复规划覆盖到更多系统，它们位于更多的地点，依赖于更及时、敏感的数据，同时要求数据中心运营团队为关键业务应用的增长提供一级的应用支持和数据保护。

**数据量正在迅速扩大**：Taneja Group 的研究表明，超过一半的美国大型企业拥有 11 TB 甚至更多的非结构化数据，如业务文档、虚拟机映像、电子邮件、媒体文件等，并且每年还在以 60% 左右的速率增长。这还要再加上年增长率达 20% 甚至更高的交易数据，大批量的历史数据处理。随着远程办公人员配置水平的下降，IT 对这些不断增长的数据集进行追踪和保护的能力陷于危险之中。

**员工和数据的移动性**：IDC 预计，到 2011 年年底，移动员工将占到美国总劳动力的 73%。这些员工不断增加对最新的、全面的且通常是敏感的数据的访问需求，同时还要依赖于规模更小的 IT 运营团队来提供此服务。这为灾难恢复规划带来了重大挑战：即如何保持大多数的数据对于大多数移动员工的可用性及安全性水平。

**愈发严格的法规监管**：由于最近的自然灾害（卡特里娜飓风和墨西哥湾沿岸各州的其他飓风、南加州大火）以及恐怖活动（9/11 恐怖袭击后华尔街和纽约/新泽西的许多商务区被关闭），有关灾难恢复站点之间距离和冗余水平的行业和政府法规越来越严格。此外，备受关注的客户数据安全破坏事件（TJX Cos., Heartland Payment Systems）也使得公众广泛呼吁对各个行业实施更严格的法规监管控制（萨班斯法案、HIPAA、PCI DSS 和欧盟隐私法等）。

## 解决方案分析

### 新技术带来新的挑战

在这些业务挑战之外，新技术（尤其是虚拟化和云计算）为灾难恢复规划开辟了新的途径，但只有经过适当管理才能实现。

### 服务器虚拟化和工作负载移动性

服务器虚拟化将服务器和应用安装到移动工作负载中，这为工作负载保护带来了福音：任何能够运行虚拟机管理程序 (hypervisor) 的服务器均可作为恢复目标。此外，由于虚拟机 (VM) 实际上仅仅是文件集，因此它们也可以作为数据进行保护和复制。如果正确地理解和管理，这些功能便可以协助灾难恢复规划。

首先，无论是在今天还是在可预见的未来，虚拟化并非适用于所有工作负载，因此仍然需要现有的物理服务器灾难恢复，同时也需要虚拟化感知的灾难恢复。此外，服务器虚拟化提供了两个必须保持平衡的关键优势：整合带来的更高利用率以及服务器的移动性。整合极大地降低了服务器硬件的投资成本，但增加了一个物理服务器故障影响多个应用的可能性。这种更高的故障风险可以通过局域网中工作负载的移动性来加以缓解，例如使用 VMware VMotion 这样的实时迁移技术。

### 存储虚拟化和数据保护

很明显，服务器虚拟化带动了许多存储技术的进步，使得存储灵活性和性能得以显著提升，并且带来了新的数据保护技术。

存储虚拟化可动态地优化分层存储，从而在选择恢复目标时提供更好的选择和更出色的效

率。虚拟备份目标，包括虚拟磁带库，允许用户从基于磁带的备份轻松迁移到基于磁盘的备份。自动精简配置可降低虚拟机映像的磁盘空间，从而使灾难恢复操作实现高效的存储和传输。基于阵列的虚拟机快照和克隆，可将服务器保护迁移到阵列并充分利用规模优势，多种数据类型（包括虚拟机）都能够以更小的备份目标集得到保护。最后，阵列上的映像和文件级恢复选项，可以提高虚拟机以及虚拟机映像中文件的恢复速度。

借助一些高级的数据复制解决方案，上述增强功能在灾难恢复方面得以进一步扩展，这些解决方案包括 EMC 的 Symmetrix Remote Data Facility/Asynchronous (SRDF/A)、Hitachi Data Systems 的 TrueCopy 和 Universal Replicator、NetApp 的 SnapMirror、Brocade 的 FCIP Gateway，以及 Dell EqualLogic 的 Auto-Replication 等。

针对虚拟服务器，虚拟机管理程序平台厂商自己也提供了许多高级数据保护功能。例如 VMware 的 Consolidated Backup、Site Recovery Manager、Storage VMotion 和 vStorage API，以及 Citrix 的 StorageLink API。连同复制解决方案一起，这些针对虚拟化而优化的增强功能提供了出色的灵活性和可选择性，但也造成了需保护数据总量的增加，这也是一个重要的考虑因素。

### 云计算和存储弹性

新的基于设施定价的外包项目的出现，引导许多企业去寻求云计算来作为对现有灾难恢复策

## 解决方案分析

略的补充。低成本的云存储产品日趋成熟，可为企业构建专用的恢复设施提供一个有效的选择。一旦几个关键要求得到满足，云存储便可以提供一个高弹性的可信存储池。提供商的平台必须能够快速按需扩展到更大的容量，提供明确的多重租赁政策，将一家公司的数据与所有其他公司的数据隔离开，同时必须支持客户对这些数据随时进行安全访问。云存储提供商还必须提供迅速、强大的数据恢复，这要求存储基础设施具有非常高的 MTBF（平均故障间隔时间）和 MTTDL（平均数据损失时间）。

云存储在大约两年时间里已经从不被人理解的技术发展成为备受追捧的解决方案，我们期望灾难恢复成为云存储最引人注目的应用案例之一。实际上，超过 40% 的企业数据中心经理最近都表示，到 2011 年年底，他们将拥有或计划部署某种形式的云存储解决方案。当前大多数的关键业务应用都通过一定数量的异地备份来提供保护，但许多企业仍需要更先进的灾难恢复水平以满足行业法规遵从、风险缓解或业务合作伙伴的要求，他们也都在竭力证明实施此类解决方案所需的资本和运营成本是合理的。数字内容服务、视频监控和医疗影像归档是一些常见的云存储数据应用案例，它们都要求强大的数据保护和安全性，同时云存储厂商还在迅速增加更多的高级功能以用于其他的数据类型。

不论云存储如何实施，它都需要一个可靠、高效且高容量的 WAN 接口。如果能够最大限度地提高 WAN 基础设施的容量和性能，那么任何包含云存储的灾难恢复策略都会受益匪浅。

### 广域网成为灾难恢复的核心要素

这些行业趋势和技术发展综合在一起揭示了一个深层次的基础设施需求：广域网 (WAN) 比以往任何时候都更加成为灾难恢复所必不可少的存储运输资源，并且将会在释放灾难恢复操作效率方面扮演越来越关键的角色。

不过，WAN 要想成为一个核心的灾难恢复基础设施要素，它必须解决网络性能最具挑战性的问题：带宽竞争和协议延迟。

### 企业中数据保护的现状

据 Taneja Group 的研究表明，寻求优化灾难恢复的解决方案是企业最优先考虑的事项。走在前列的通常还是那些对虚拟化进行了大量投资的企业。我们的研究显示，目前超过 50% 的大中型美国企业对于虚拟服务器的备份，采用的是与物理服务器相同的流程。55% 的企业仍然使用磁带作为其主要的备份/灾难恢复介质，不过大多数企业还另外采用了至少一种磁盘到磁盘方法（图 1）。其中最常见的是二级 SAN 备份（备份到位于另一个 SAN 上的主磁盘）和二级备份（备份到位于本地站点的低成本磁盘或虚拟磁带库 (VTL)）。

解 决 方 案 分 析

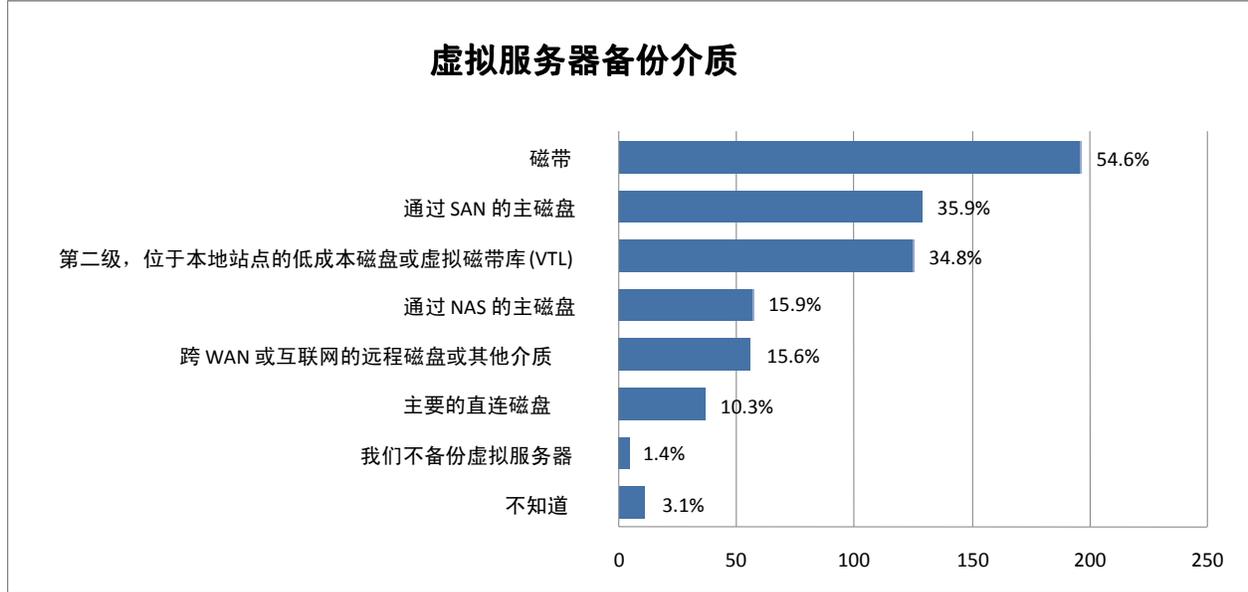


图 1: 虚拟服务器环境中使用的备份介质 (资料来源: Taneja Group, 2009 年)

然而，变革已在顺利进行：相同的研究还表明，虚拟服务器数据保护被大多数用户视为各个虚拟化阶段中一个需要改进的重大问题。

尽管大多数企业还没有在其关键灾难恢复实践中采用虚拟服务器技术，但是一个虚拟化感知的灾难恢复方法已经被广泛使用：使用 VMware VMotion 或等同的工具在服务器之间迁移虚拟机。如图 2 中所示，目前有超过

50% 的受调查企业针对灾难恢复采用手动和/或自动的虚拟机迁移方法。

我们的研究还可以证实，复制是另一个流行的灾难恢复机制，特别是在虚拟化程度最高的数据中心中越来越流行。45% 到 50% 的受访用户目前在其虚拟基础设施中采用了数据复制的方法。

## 解 决 方 案 分 析

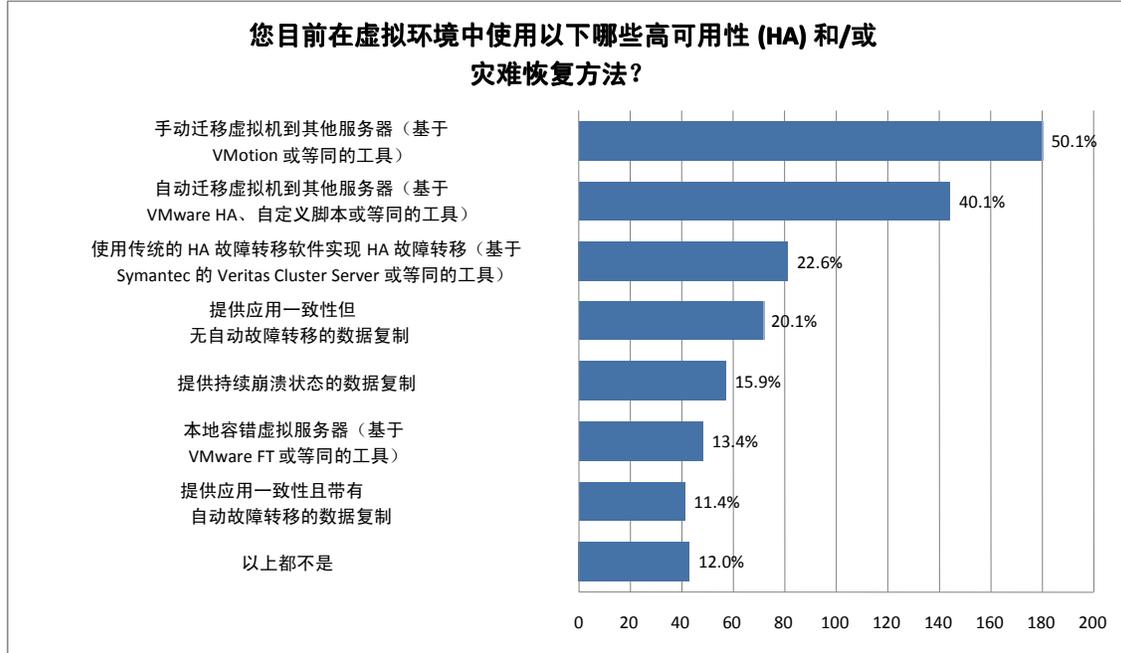


图 2: 虚拟服务器环境中部署的 HA/DR 策略 (资料来源: Taneja Group, 2009 年)

在我们看来, 以上资料表明的关键灾难恢复结论是: 深谙虚拟化之道的公司正迫切地使用虚拟化来简化和加快灾难恢复。只要存在虚拟机管理程序, 快照、克隆和按需转换 (物理到虚拟, 及虚拟到物理) 即可实现虚拟工作负载的近乎实时的复制, 并且可用性水平只受限在于在不同地点之间移动工作负载和数据的频率和速度。

显然, 网络速度和容量是制约多站点灾难恢复性能的主要因素。而能够对现有网络这两个方面进行改善的 WAN 优化解决方案将越来越受欢迎。我们对 WAN 优化将会有更高的需求持充分预期, 因为我们认识到它能够大幅提升复制水平, 进而提高数据中心的总体可用性。

## WAN 优化对灾难恢复的强大动力

企业目前应调研相应的解决方案来优化网络容量和利用率, 以确保 WAN 带宽和延迟不会成为当前或未来灾难恢复规划的制约因素。在任何灾难恢复规划工作中, 及早释放现有的 (但尚未开发的) WAN 带宽, 将有助于避免将来在其他灾难恢复解决方案组件 (例如存储或磁带接盘箱) 上花费过多。

### 谁适合 WAN 优化?

WAN 优化为灾难恢复打破了运营障碍, 解除了地域限制, 并增强了新兴的存储技术。如果以下任何一句话对您来说是对的, 那么您应当寻求 WAN 优化解决方案:

## 解决方案分析

**“我无法确保满足我的灾难恢复服务水平协议 (SLA)。”**

如果您只是复制数据中心和 LAN 上的一部分关键数据或工作负载，并且依赖于磁带备份或可移动磁盘在站点之间进行灾难恢复操作，那么您预计恢复时间和/或提供充足恢复点的能力可能会变得不确定且难以测试。WAN 优化可以帮助用户实现在 WAN 链路上提供接近于 LAN 的性能，从而支持更多的数据中心到数据中心复制，以符合更严格的 RPO/RTO 服务水平协议。此外，利用自动化的网络服务质量 (QoS) 管理，WAN 优化还允许您在不影响应用流量的情况下，更频繁地测试灾难恢复流程，并且在实际网络中进行测试。

**“很难保护在地理上分散的多个地点。”**

WAN 优化能够帮助您克服 LAN 与 WAN 环境之间的带宽不匹配问题，支持您更频繁地进行备份，并备份到更多的地点：数据中心到数据中心、分支机构到数据中心、远程员工到数据中心等。通过克服协议延迟，WAN 优化使您能够在更远的距离上传输备份和恢复数据集，而不会降低性能。此外，WAN 优化还使您能够将保存在多个地点的敏感数据持续地复制或移动到一个安全的位置，从而改善合规性审计。

**“我想要集中进行备份操作以减少冗余。”**

WAN 优化可以免除本地磁带备份和管理它们所需的人员。集中进行数据保护可以充分利用您目前已有的数据中心员工和技能，并减少对分支机构承包商和其他第三方的依赖。WAN

优化惠及数据中心的方方面面：存储团队可以在站点和阵列之间部署先进的、高带宽需求的复制技术，服务器团队可以更频繁地执行虚拟机快照并将它们备份到多个地点，而网络团队可以为在 WAN 容量方面花费的每一分钱提供更高的性能回报。

### WAN 优化能够带来什么？

WAN 优化可以立即提高您当前网络基础设施的投资回报率 (ROI)，并带来引人注目的运营优势：

- **更高的灵活性：**更多的备份和复制、更多的站点、更多的配置（1:1、1:N、N:1、P2V、V2P、移动到数据中心等）。
- **更高的成本效率：**释放额外的容量可以从您现有的资源中发掘价值，并有助于延缓额外带宽的购买需求。
- **更低的风险：**整合备份降低了复杂性、冗余和数据丢失的可能性；更少的地点带来更简单、通过率更高的合规性审计。
- **更高的可用性：**保护更多的数据和工作负载，并利用更多的复制，可提升您的整体企业可用性水平。
- **更短的 RPO：**在保护其他网络流量服务质量的同时，更频繁的复制可确保最新数据的可用性，并尽可能提供到最需要的位置。

## 解 决 方 案 分 析

- **更短的 RTO:** 克服协议低效和减少线路上的数据，可实现更快速的备份以及更频繁的复制。

### WAN 优化如何运作？

我们认为，一个全面的 WAN 优化解决方案应包括以下基本功能，以便为灾难恢复提供最大的优势：

灾难恢复目标	解决方案能力
提高 WAN 利用率和效率	通过 TCP 优化减轻网络低效问题，允许最大限度地利用现有网络的额定容量并支持更频繁地发送更多的灾难恢复数据集
克服应用协议延迟	克服固有的应用协议延迟，从而使备份/复制能够到达位于更远距离的多个数据中心和远程站点
减少线路上的数据	减少通过 WAN 发送的重复数据的数量，从而缩短备份和恢复时间，并实现接近于同步的复制
针对灾难恢复工作负载优化性能	针对突发流量和大数据集，自动地检测和调节网络性能，进而专为灾难恢复工作负载提供额外的优化
管理竞争和服务质量	协调针对共享 WAN 资源的竞争，以便在灾难恢复操作或测试运行期间，为应用工作负载保持必需的吞吐量

### 评估针对灾难恢复的 WAN 优化解决方案

如何针对您已选定的备份和恢复技术确定最佳的 WAN 优化解决方案？首先，了解所有供应商之间的关系：确认他们已对联合解决方案进行了充分测试，已按照各自的标准进行了认证和记录，并且提供了公开的基准性能。在该基准建立之后，寻找这些技术的与众不同之处：

**TCP 优化:** 该解决方案应能够最大限度地提高每个 TCP 往返周期中每个有效负载的数据量；去除冗余的 WAN 数据；提供高速 TCP 选项（HS-TCP、MX-TCP）以最有效地填充网络“管道”。领先的解决方案应能使 WAN 利用率降低 60-95%。

**数据优化:** 该解决办法应采用先进的重复数据删除算法，并优化板载处理耗能和内存的使用。重复数据删除技术应在字节层级上实现，并且其得到的数据压缩比应超过存储厂商所提供的基于块的数据差分技术。数据压缩比至少应达到 4:1 到 8:1 或更高。

**工作负载优化:** 该解决方案应提供灾难恢复特定工作负载的识别和加速功能；使用自适应算法自动检测灾难恢复特定的流量（例如请求大数据集）、拦截请求，并分隔数据以获得最高的线路效率。此种优化在基于 WAN 的备份和恢复时间方面应带来至少 5 倍，最高 50 倍的改善。

**服务质量 (QoS) 管理:** 该解决方案应能够协调针对 WAN 带宽的竞争，按应用和/或端口分配带宽，同时根据对 TCP 和 UDP 流量的

## 解决方案分析

单独延迟敏感性来独立划分数据包的优先级。领先的解决方案应能够提供创新的 QoS 调节算法，为每个应用分配所需的最小带宽，避免过度补偿。

**透明度和可用性：**在不同的层级位置（用户、部署、报告、端口和地址层级）上，该解决方案对现有的网络基础设施都应该是基本透明的；它应能够支持针对多个系统的集中管理；它应提供高级可用性、可扩展性和部署功能，例如高效地使用板载内存和固态硬盘来进行缓存及群集配置。

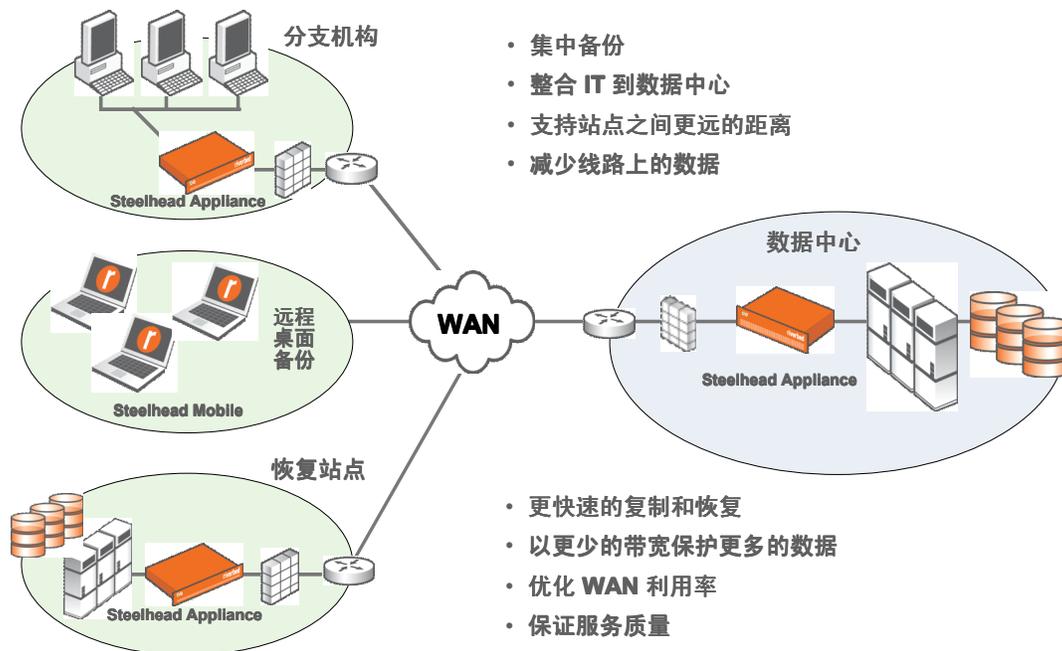


图 3: WAN 优化的灾难恢复参考架构（基于 Riverbed 的 Steelhead 系列）。设备在分支机构、远程桌面、灾难恢复站点以及主数据中心进行部署，并通过现有的 WAN 透明地工作以加速性能和最大化吞吐量。

## 解 决 方 案 分 析

数据中心正在迅速演变。因此，大多数的数据保护和灾难恢复 (DR) 计划在制定时所针对的 IT 环境，与如今动态、高度分散且越来越虚拟化的数据中心并没有多少相似之处。传统的灾难恢复规划会假设远程站点拥有专门的 IT 员工、应用和数据一般是静态的、本地数据保护最为高效，并且不同地点之间的广域网受到严格限制且成本高昂。

这些假设受到了经济压力、快速增长的数据量、广泛的虚拟化，以及更强大的存储阵列功能等方面的质疑。在较小的分支机构保留专门的 IT 员工，或者在发生灾难时依赖于冗余的、基于磁带的备份已不再具有成本效益。越来越多的移动员工和不断增多的法规监管，推动了对更高的灾难恢复效率不断增长的需求，包括多个站点之间，以及跨越广泛地理区域的工作负载和数据复制。此外，虚拟化本身也带来了新的应用移动性水平，从而支持在存储阵列层级实施创新的虚拟化感知的备份和复制策略。不过，我们复制数据和工作负载的能力正快速超越将其放在所需位置上的能力。

我们需要一种新的灾难恢复方法，这种方法可以解决企业灾难恢复可扩展性和效率上的关键制约：有限的网络容量。WAN 优化通过释放现有网络带宽，并针对备份和恢复工作负载的独特传输要求来优化 WAN，消除了这个“瓶颈”。WAN 优化已被证实的优势包括极高的数据压缩比和网络吞吐量的巨幅提升，当与来自优秀阵列厂商的存储复制解决方案相结合时，这种优势将会得到进一步放大。

这一强大的联合解决方案使得远距离、多站点复制变得可行、更简单且更加经济高效，同时在我们看来，这也为降低企业对基于磁带的灾难恢复（昂贵且容易出错）的依赖提供了明确的路径。

---

备注：TANEJA GROUP 提供的信息和产品建议是基于公开的信息和来源，也可能既包括 TANEJA GROUP 和其他方的个人观点，所有这些都是我们认为准确和可靠的。但是，由于市场情况的变化且不在我们的控制之下，因此所提供的信息和建议不带有任何形式的担保。此处提及的所有产品名称均是其各自所有者的商标。TANEJA GROUP, Inc. 对因您使用或依赖此处提供的信息和建议所造成的任何损失（包括偶然、结果性或其他方式）以及本文档中可能出现的因疏忽所致的任何错误不承担任何责任或义务。