

万兆位以太网：打造面向巨型数据的存储网络

前言

巨型数据(Big Data)要求使用巨型网络。

如今数据的爆炸式增长已成为不争的事实，而且这种增长势头愈发强劲。为适应和支持数据的巨幅增长，部署功能更加完善和强大的网络已迫在眉睫。随着数据的增多和数据用途的多样化发展，用户需要在数据中心部署基于角色的存储解决方案。由于这些因素的影响，加之数据中心逐渐迁移到高度虚拟化的环境，企业购买和管理服务器、存储系统和网络资产的方式随之发生改变，巨型数据也逐渐成为日常运营流程的一部分。未来，只有云计算才能为巨型数据提供理想的运作环境。

巨型数据(Big Data)由一些大型数据集组成，这些数据集极为庞大，现有的数据库管理工具已完全不能适应需求，这些数据集的捕获、存储、搜索、共享、分析和可视化都构成了巨大的困难和挑战。但是由于更大的数据集能够帮助分析人员更精确地掌握商业趋势并解决相关问题，这种增长趋势将会延续下去。目前，业界将数据的规模限制在太字节（万亿字节-Terabytes）、艾字节（百亿亿字节-Exabytes）和泽字节（十万亿亿字节-Zetabytes），但实现此目标的时间表并不确定。

巨型数据在许多领域的用处极大，包括气象学、基因组学、生物研究、互联网搜索、金融和商业信息研究与分析。Arista是巨型数据管理领域的领先企业，在以上这些市场具有强大的竞争优势。目前我们在使用巨型数据时面临的主要挑战是如何有效使用关系数据库和桌面统计数据/可视化数据包，这些数据要求在几十、几百、甚至几千个服务器上并行运行软件。网络和云存储系统的功能恰恰迎合了这种需求，Arista提供的网络产品具有高密度、无阻塞特性和高可扩展性，而且体积小巧，因此针对以上问题提供了理想的解决方案。

用户、应用或机器产生的数据几乎无处不在，而且数量急速增长，这种趋势覆盖了所有的市场或行业。因此，所有的IT机构都不得不正视眼前的挑战，寻求合适的解决方案来存储、管理并充分利用这些数据，同时最大限度节约成本。一场角逐云计算的竞赛由此展开，其中计算框架需要能够近实时地处理庞大的数据，而且必须将成本大幅度降低。

推动存储网络发展的主要应用

推动存储网络升级的驱动因素不一而足，最显著的推动力来自于云计算、Hadoop集群、存储虚拟化、高性能计算和多媒体/视频应用。此外，由于以下趋势的出现，企业对于面向巨型数据的高性能、低延迟和高可用性网络的需求也日益强烈，这些趋势包括数据中心越来越多地采用基于Web的应用架构，企业纷纷采用虚拟化工具来整合服务器，以及核心关键任务应用开始使用HPC。

推动巨型数据迅速增长的机构主要是互联网服务供应商、应用服务供应商和存储服务供应商。这些供应商的网络核心通常需要万兆以太网(10GbE)带宽，以支持各种应用以及远程复制和磁盘到磁盘备份。此外，其他一些在短时间内消耗或产生大量数据的应用也将受益于10GbE网络，包括视频编辑和多媒体应用。

Hadoop集群

在过去十年，数据分析已成为业务决策流程的关键组成部分。目前，在数据库内存储的数据报表的传统方式正面临挑战，而面对数量庞大的非结构化数据以及将数据迅速投入使用的要求，过去的的数据收集与采集方法也明显不再适用。这种分析方法的主要局限性在于无法获得必要的计算和存储资源来及时得出分析结果。专为Hadoop应用（而非标准企业应用）而设计的网络能够极大地提升Hadoop集群的性能。Arista Networks能够满足Hadoop集群网络的要求，而且已经在许多Hadoop环境中实现成功部署。随着成本的降低和企业逐渐采用新的数据关联

与分析方法，巨型数据分析法将逐渐普及开来。随着企业能够以低成本方式管理和分析巨型数据，他们将是最大的获益者。

Hadoop是一种功能强大的分布式计算框架，能够处理各种各样的数据集，包括从几千兆字节(Gigabytes)到几千兆兆(Petabytes)字节的结构化和非结构化数据。Hadoop在业界迅速受到广泛欢迎，现在各种科学计算和商业分析都纷纷采用Hadoop作为平台。虽然基于商用Linux的服务器具有出色的可用性，允许构建超大型集群，但网络却构成了瓶颈，经常造成拥塞，导致无法高效使用集群。

Hadoop由两种关键服务组成：基于Hadoop分布式文件系统 (HDFS)的可靠数据存储和基于MapReduce的高性能并行数据处理。MapReduce是Hadoop平台的处理引擎，赋予平台出色的速度和灵活性。利用MapReduce，开发人员能够创建强大的应用程序，可在包含处理器或独立计算机的大型分布式集群中并行处理大量非结构化数据。Hadoop使企业能够使用针对其信息和问题的定制分析方法来研究和运用这些复杂的数据。巨型数据带来的挑战不仅仅在于庞大的数据量，处理性能以及处理速度也是备受关注的焦点问题。

Arista提供高性能的1/10 GbE无阻塞、超低延迟解决方案，能够从几个机架扩展到最大型的Hadoop环境。此外，多机箱LAG (MLAG)在每个机架提供真正的Active/Active上行链路连通性，允许在一个扁平的第二层网络中使用网络的全部上行带宽。Arista的可扩展操作系统能与Ganglia和Nagios轻松集成。最后，Arista的网络解决方案在服务器互连带宽价格方面实现了真正的零增长。这些优势使Arista成为任何Hadoop环境的理想网络解决方案。

云计算

云计算能够帮助企业以一种合理和简单易用的方式存储、管理、共享和分析其巨型数据。在按需式分析解决方案供应商的支持下，当今的基础设施即服务 (IaaS) 提供商已大幅度降低了巨型数据的分析成本。由于与位置无关的计算要求使用共享的服务器，按需为系统和设备提供资源、软件和数据，云计算将极大地推动高性能存储网络的应用。虚拟化技术与公用计算的广泛普及带动了这种自然的发展演变，使客户无需掌握专业知识或控制为其提供支持的技术基础设施。但是，为帮助客户获得最佳的性能和服务，需要配置强大的高性能网络来为其提供支持，比如，Arista Networks的整个交换解决方案产品线就专门用于打造这种高性能网络。

存储虚拟化

虚拟化是云计算必须具备的特性，因为虚拟节点所需要的性能仅是目前的CPU的一小部分。向虚拟存储的急速转型和融合也带来某些问题。首先，这种融合意味着存储和网络容量的过量供应(over-provisioning)。此外，光纤通道(FC)存储和以太网数据网络设施的数量过多，iSCSI等新存储网络技术的普及速度也超乎想象。以太网iSCSI是一种适合采用虚拟化的理想环境，因为它允许管理员在物理机器间移动虚拟服务器，同时无需在存储网络中重新配置分区和逻辑单元号(LUN)掩码。随着虚拟服务器部署范围的扩大，除非网络能够发挥重要作用，成功推动这种转型的实现，否则这些问题将使我们无法真正受益于新的存储模式。Arista的万兆以太网解决方案能够减轻这些问题的影响，确保我们获得存储虚拟化的所有优势。

高性能计算 (HPC)

高性能计算即利用超级计算机或强大处理器的集群来解决一些计算密集型问题。为了让数量庞大的处理器能够协同运行，集群必须配置互连来支持高带宽和超低延迟通信。Arista Networks的交换机产品具有当今市场上最低的延迟，能够用于HPC集群互连，帮助气象学、基因组学、石油和天然气、生物研究行业以及其他任何需要处理大量数据的企业处理数量庞大的数据。

多媒体/视频

Arista Networks能够提供完全冗余、基于云的内容原创服务，这些服务支持所有的视频和高清晰照片内容。Arista Networks提供的以太网网络能够将流服务器彼此互连，这些服务器要求使用高性能1/10GbE交换，并提供大型的数据包缓冲。这些网络非常适合支持高带宽的视频流和存储互连，同时能够保持高水平的系统可用性，将设备

与系统灵活地集成到多媒体/视频基础设施的设计中。此外，机构通常需要对这些数据进行汇聚和分析，从中获取商业情报或用于后期制作。由于可用性和灵活性至关重要，Arista可扩展操作系统[EOS]及其服务中软件升级(In-Service Software Upgrades)和状态故障修复(Stateful Fault Repair)功能对于满足服务水平协议和总体运行时间要求将发挥巨大作用。

巨型数据存储网络选项

随着性价比的提升和新型光技术对部署的驱动，万兆以太网的部署呈现急速上升趋势。此外，新应用的增长也将继续推动带宽要求的提高。虽然万兆以太网是最佳的网络接口，但它其实是在交换解决方案的总体架构中发挥作用。成功的万兆以太网部署还必须支持一些领先的智能交换服务，例如高性能、低延迟、高可用性，同时为新型应用提供必要的可管理性服务。

当今市场上存在各种各样的存储解决方案，它们采用的方法和技术各有不同。基于iSCSI的块存储和NAS网络存储的云存储领域的主流技术是采用万兆以太网。Arista Networks交换机具备无阻塞吞吐率、无与伦比的高密度、低延迟和最低的总体拥有成本的特点，是云存储应用的理想平台。

根据接入方法的不同，存储系统可分为存储区域网络(SAN) 和网络连接存储(NAS)解决方案。在SAN环境中，存储设备虽然位于远端，但看起来如同与客户端直接相连，而且对存储设备的访问是基于块的访问。与之相反，在NAS环境中，客户端利用一个基于网络的文件系统从远端访问文件。

互连技术 协议	映射
光纤通道(FC) 协议(FCP)	光纤通道
以太网 TCP/IP	iSCSI
以太网	FCoE
Infiniband	iSER

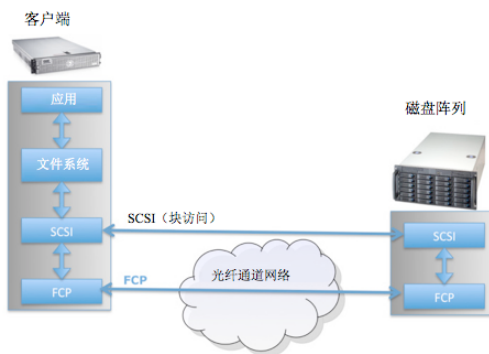
存储区域网络

在存储区域网络(SAN)架构中，服务器通过一个专用互连访问远程磁盘块。多数SAN利用SCSI协议在服务器和磁盘之间建立通信。用户可以使用各种互连技术，其中每种技术都需要一种特定的SCSI映射协议，如左图所示。

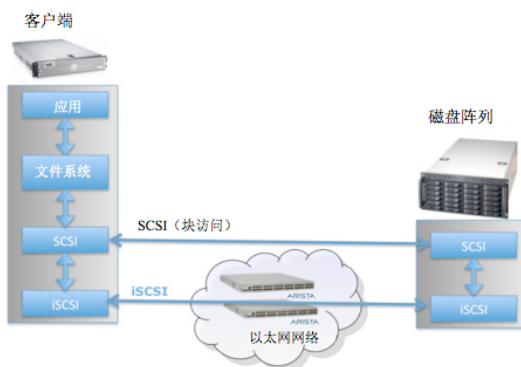
SAN是一个专用网络，允许服务器快速可靠地访问外部或独立的存储资源。在企业网络环境中可根据存储数据需要不断增加，SAN为这些数据提供了理想的存储解决方案。通过实施SAN，用户可以将存储流量从日常网络运行中卸载下来，并在存储组件和服务器之间建立直接连接。SAN能够将存储接口互连在许多网络中，而且可认长距离进行。这些互连还将SAN接口连接到了SAN交换平台(Fabric)。

Switched SCSI、FCS和Switched SSA共同组成了最普通的SAN交换平台。通过网关，SAN也可以扩展到广域网。交换机允许建立可供大量集中式、中央管理的整合存储库，实现在各类应用中共享。用于建立SAN的网络技术必须具备出色的可扩展性、性能和可靠性，以便将传统存储环境的突发性和速度特性与网络的连通性完美结合。

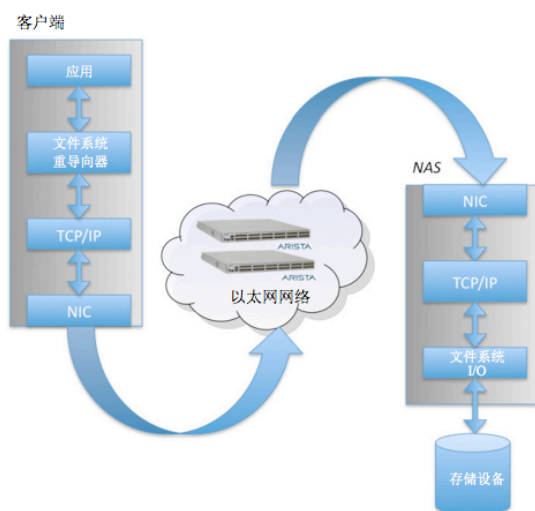
光纤通道（FC）存储区域网络（SAN）



iSCSI存储区域网络（SAN）



网络附加存储（NAS）



光纤通道协议(FCP)

如今，多数SAN在一个专用光纤通道（FC）上利用FCP映射到SCSI。部署光纤通道的企业需要部署多个网络，包括采用以太网技术的局域网（以太网是全球85%的网络的基本组件，也是目前应用最广泛的网络协议）和专用FC网络。

iSCSI

由于数据中心需要降低存储成本，iSCSI开始受到普遍关注。iSCSI基于1/10Gigabit以太网传输技术，从而减轻了需要建立独立的传统光纤通道SAN的复杂性。是许多中小企业理想存贮解决方案，iSCSI采用了TCP/IP协议，因此成为支持专用和公共云通信的首选通信技术。

iSCSI的性能优势非常显著。新型的多核心处理器和堆叠软件与存贮阵列结合，目前能够生成数百万的iSCSI IOPS。配备1/10 Gbps以太网控制器的新型iSCSI阵列包括Dell Equal Logic 6100阵列、EMC VNX存储阵列和Netapp FAS 6000系统，这些产品与Arista Networks 7000交换机相结合，能够提供无阻塞的存储访问。

网络连接存储（NAS）

目前网络连接存储(NAS)作为一种存储资源整合方法日益受到青睐。NAS设备是一个服务器，专门用于为网络上的其他设备提供基于文件的数据存储服务。NAS是一个远程文件系统I/O，文件请求将在网络上重新定向。

NAS以三大优势闻名，这些优势共同降低了总体TCO：

- o 存储整合
- o 部署简单
- o 管理容易

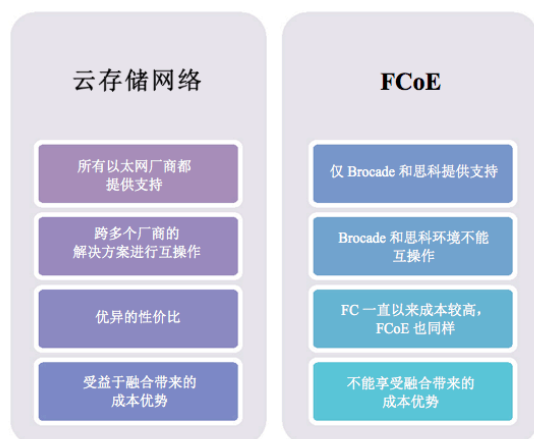
NAS系统经过不断改进，已经能够通过标准的以太网网络（存储分层）实现高性能和高可用性，这些优势以前仅在SAN中提供。这些特性与TCO方面的优势使

NAS在企业中的部署率越来越高。

以太网光纤通道(FCoE)

FCoE协议是FCP在以太网上的一种封装。FCoE使习惯于使用光纤通道的企业客户能够直接在他们的局域网以太网上运行光纤通道协议，从而能够在同一个网络基础设施上整合其局域网和存储网络。FCoE的目标客户是那些希望运行一个高端FC SAN，而同时对于LAN和SAN融合感兴趣的企业组织。目前FCoE尚处在早期发展阶段，需要对标准以太网进行特殊改造（如优先级流控制），这些改造可通过Arista交换机实现。

存储需求推动了万兆以太网的发展。Arista认为存储是10GbE云计算网络发展的自然推动力。7000系列专为数据中心解决方案而优化，提供了重要的存储特性，如灵活性、高吞吐率、用于处理无损1/10GBE流量的大缓冲区、PFC/ DCBX支持和可预测的低延迟。Arista为iSCSI和NAS存储设计了一种令人惊叹的1/10千兆以太网架构，同时为FCoE流量提供了透明的传输支持。



要想在现实环境中成功部署和运用FCoE，需要满足几个要求并排除几个问题。首先，存储流量必须与局域网上的其他流量分开。其次，绝对不能丢弃任何存储数据包，因为光纤通道协议在发生数据包错误之后的恢复速度极慢。最后，不同厂商的HBA和网关产品之间必须能够互操作。任何两个FCOE部署环境都不可能完全相同，也无法保证互操作性。但是Arista交换机支持FCOE标准，并能在10GEe网络上传输FCoE流量。

综上所述，与FC和FCoE不同，由于iSCSI和NAS符合业界标准，多厂商支持而且具有最好的性价比，因此已成为“云存储”的最佳选择。左侧的表列出了iSCSI和NAS的优势特性以及它们成为最佳选择的原因。

Arista Networks的优势和独有特点

Arista Networks致力于为当今的数据中心和云计算环境构建可扩展的高性能和超低延迟网络，并在该领域居于领先地位。专门定制的硬件和Arista可扩展操作系统提供了一个面向所有平台的统一二进制系统映像，同时支持最长的系统运行时间、状态化错误修复、服务过程中升级，而且完全支持对Linux shell的访问。即使您的工作负载有最严苛的要求，Arista 10GbE交换机也能为您提供完美的网络解决方案。Arista产品支持VMware虚拟化和几百种Linux应用，这些应用集成在专为满足当今数据中心最严苛的电源与制冷要求而设计的硬件平台中，因此Arista的10GbE平台在能效和性能方面堪称业界之冠。

存储以数据为核心。从主机移动到目标设备的数据越多，主机处理应用的效率将越高。Arista交换机被公认为是市场上性能最高、最受好评的10GbE交换机。存储设备执行写操作与存储目标发出确认之间的延迟越短，主机处理数据和应用的效率也将更高。Arista 10GbE交换机是延迟最低的产品之一，在24端口交换机中，延迟仅为500ns；在1RU 48端口10GbE交换机中延迟低至1.2usec，而在384端口的Arista 7500系列交换机中，延迟只有4.5usec。

如果存储设备和主机的连接速度不同，则应该重点考虑使用缓冲区。Arista的交换机系列提供极大的缓冲区，比其他厂商的产品大100多倍。这样的大型缓冲区能够吸收突发的读/写流量，不会丢弃帧，因此减少了重发次数，提高了效率和应用性能。Arista致力于建立一个开放的存储市场—该市场将同时支持iSCSI、NAS、SAN、FC和

FCoE。因此，Arista与EMC、Dell/EqualLogic、NetApp和Panasas等著名存储系统供应商以及QLogic和Emulex等一流NIC厂商建立了合作伙伴关系，确保客户能够成功部署这些关键任务应用。



Arista不但提供简单易用的产品和解决方案，还能自动开通(provision)您的网络，免除了这项工作通常涉及的人工干预。Arista帮助客户实现了零接触开通[ZTP] – 云数据中心网络自动化部署。利用基于标准的协议(如DHCP、T/FTP、HTTP)，网络的开通可快速完成。借助高级脚本功能，管理员能够根据各种参数定制启动配置，满足最复杂的数据中心部署的需求。在其他Arista特性的帮助下，如VMTracer自适应VLAN配置，数据中心经理可以完全将网络组件和虚拟服务器的启动和运行自动化。Arista独特的“无需插手”开通方法实现了云数据中心的自动启用。

应用

虚拟数据中心的完全自动化部署：借助ZTP部署的Arista交换机将获得一组VMTracer配置参数。在VMware vSphere服务器启动和实例化虚拟机时，VMTracer将为VM自动配置VLANS。在此过程中无需人工配置，交换机将自动启动，准备投入使用。

利用Puppet/Cobbler模板在HPC集群中配置交换机：通过对集群服务器配置方法进行扩展，HPC经理可以利用模板在HPC集群中快速部署其他的服务器机架。除服务器模板之外，HPC经理还可以增加一个用于机架顶端(ToR)交换机的模板。现在，服务器机架安装到集群之后，网络和服务器平台都已配置完毕并集成到了现有集群中。

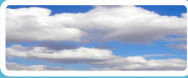
利用脚本将Linux扩展程序安装到EOS：某个服务供应商编写了一个shell脚本，可在平台开启时通过ZTP下载到一个交换机。shell脚本将下载一个OpenVPN RPM，然后对其进行安装和配置。该脚本的最后一个功能是将日志和配置信息发送给NOC。交换机重启之后，HPC经理就可以从远程NOC安全地访问交换机，进行开通。

功能	以太网	FCoE	Infiniband
端口速度	1Gbps / 10Gbps 40Gbps / 100 Gbps (未来)	10Gbps	8/16/32Gbps
第3层	是	否	否
交换机延迟	~1usec	3.5usec - 10's usecs	~ 500ns
与基于以太网的应用的互操作	无缝	目前不能互操作	要求使用网关
可管理性	与传统局域网使用的工具相同	与传统局域网使用的工具相同	有限的运营工具
IEEE 802.1Qbb/az/au	支持	支持	N/A
PFC, 增强的传输选择, QCN	是 (L2和L3)	是 (仅支持L2)	Fat Tree

总结

顺应多协议SAN与NAS的发展，Arista Networks通过一些具体措施扩展了FC和10GbE存储网络的价值，更好地满足了快速演变的存储数据中心环境的需求。Arista Networks投入巨资在全球大力拓展了存储基础设施商业合作伙伴网络。这些合作伙伴致力于为拥有高级数据中心和存储需求的大型企业提供服务。Arista Networks与合作伙伴共同提供的基于网络的存储服务系列为新环境带来了全新的高级功能，对于10 GbE部署项目来说，Arista Networks产品无疑是最佳的选择。

面向巨型数据的云存储：

	iSCSI与NAS，与FC或FCoE相反
	目前10GbE的多数机会来源于iSCSI和NAS
	推动数据中心部署 10GbE 的主要因素
	Arista与众多云存储厂商建立了合作伙伴关系，包括BlueArc、Coraid、Dell EqualLogic、EMC、Isilon (已并入EMC)、NetApp和Panasas

基于Arista的云服务器/存储设计

设计优势：

- ✓ 出色的可扩展性
- ✓ 无损传输
- ✓ 高可用性
- ✓ 配置简单
- ✓ 虚拟化
- ✓ 投资回报时间大大缩短

10 GbE使企业能够扩展应用功能，快速解决复杂的财务、云计算和集群应用问题，同时及时满足客户需求，迅速应对市场变化。高带宽、低延迟和高能效是Arista Networks交换机产品的主要竞争优势，能够帮助企业从容应对挑战。Arista的7000系列万兆以太网交换机在满足云存储需求方面具有无可比拟的独特优势，在每个端口上提供无阻塞的10 Gbps以太网吞吐率，延迟低于一微秒，而且无论是固定配置还是模块化配置，每RU的端口密度都高居市场首位。Arista交换机提供灵活性极高的可扩展操作系统(EOS)、自恢复及实时修补功能、而且支持动态缓冲，能有效应对速度不匹配问题或避免数据包丢弃或丢失。



综上所述，Arista的产品系列专门面向存储数据中心和大企业网络，可帮助存储数据中心部署高性能的计算集群，而在大企业网络中，则可帮助企业将各自独立的SAN解决方案整合为一个统一的低成本高速10 GbE IP存储网络，同时从根本上改变了巨型数据的联网方式。转型和升级的时机已经成熟，Arista将始终为您提供性能最高、最优秀的解决方案，有效满足您的巨型数据存储和管理需求。

总部

5470 Great America
Parkway
Santa Clara, California
USA 95054
408-547-5500

支持

support@aristanetworks.com
408-547-5502
866-476-0000

www.aristanetworks.com

销售

sales@aristanetworks.com
408-547-5501
866-497-0000