



ARISTA 白皮书

软件驱动的云网络

Arista Networks，高速、高度可编程数据中心交换领域的领导者，对与软件定义的网络 (SDN) 技术（包括控制器、虚拟机监控程序、云编排中间件和定制的基于流量的转发代理）集成提出了一些指导原则。这些指导原则采用 Arista 的行之有效、可扩展、基于标准的控制和数据平面交换技术。

新兴的 SDN 技术通过在更广泛集成的云基础架构生态系统内自动化网络策略和服务开通来完善数据中心交换机。Arista 将 SDN 技术与 Arista 可扩展操作系统 (Arista EOS[®]) 的组合 Arista EOS+ 平台定义为软件驱动的云网络 (SDCN)。

由此带来的 Arista SDCN 方法的好处包括：网络应用程序和开放式、可扩展标准与众多云服务开通和编排工具集成；多租户虚拟机的无缝移动性和可见性用于实现开放式工作负载移动性；实时网络遥测数据用于实现与云运营工具的同类最佳关联。

ARISTA

云技术变迁

自 20 世纪 80 年代诞生以来，以太网网络已获得显著进化，许多发展变化造就了今天各种各样可用的交换机类别（参见图 1）。数据中心交换已形成一个独特的类别，高密度 10Gbps、40Gbps 以及现在的 100Gbps 端口到端口线速交换是重要的以太网网络产品领域之一。除了这些值得关注的速度级数以外，数据中心交换还提供了次微秒交换机延迟时间、具有广泛的等价多路径路由架构的更富弹性的架构、网络虚拟化的集成（以支持简化的服务开通）和数据中心基础设施上的网络应用程序的集成（使 IT 运营与网络行为相对应）。

这些最先进的交换功能充分利用了 30 年来不断进步的硬件和软件技术演变，因而成功实施 Arista SDCN 需要实现从封闭的、供应商特定的专有单一设备中心操作系统到开放式、可扩展、外部可编程网络操作系统的根本性转变。这种开放式、可扩展性要求由云数据中心的指导原则促成，

其中，资源将被作为一个由计算、网络和存储构成的集成系统进行动态管理。

云控制器，作为底层基础设施的编排中间件和资源管理器，推动实现工作负载放置和移动性的服务开通决策。这包括工作负载驻留在网络边缘的什么位置。每当工作负载被移动时，必须更新网络以确保所需资源的适当服务开通或虚拟工作负载的可见性。交换机必须与这些控制器实时交互，因为基于机架级和数据中心私有及混合优化技术，这些工作负载越来越高度可移动。Arista 称此为开放式工作负载移动性。

构建于旧的设计原理的封闭网络操作系统，最多能够提供一次性实施，难以支持越来越多的不同 SDN 控制器形态规格。然而，Arista 处于独特的领导地位——业界屡获奖项的模块化 Arista EOS 能够与多个系统同时进行交互，处理外部控制器更新和管理高度分布式交换机转发状态，这两者都是实时进行的。Arista 的方法能够做到两全其美，为外部控制器提供服务控制，同时通过叶子/骨干交换架构实现最苛刻的运营级云数据中心的扩展。

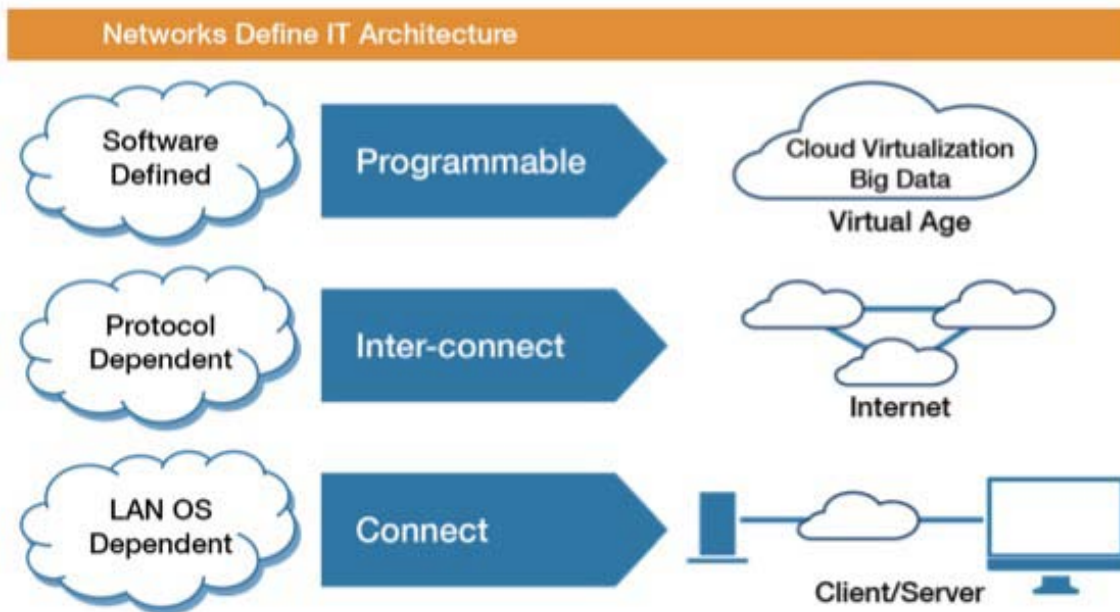


图 1：软件驱动的云网络演变

表 1: Arista SDCN 的多种方法

多种方法	示例	好处
虚拟机	Web 应用程序框架	根据需要扩展/收缩
SDN 控制器	OpenFlow、OpenStack、vCloud Suite、vSphere	编排、服务抽象和服务开通
网络虚拟化	可扩展、多租户虚拟网络	实现开放式工作负载移动性
服务器虚拟机监控程序	x86 裸机服务器抽象	弹性计算、资源优化、不中断的服务器升级
存储	网络、直接连接、SSD、Hadoop、大数据	集中式 VMDK 实现应用程序移动性、软件修补程序
启用云的网络	Arista EOS	开放式和可编程实现自定义流量、VM 移动性、自动化租户载入

平均而言，对于任何新的或更新的应用程序，使用传统的网络操作系统需要花费两到四个星期来配置一个完全集成的数据中心基础设施并将其发布投入生产。这两到四个星期的时间周期多半是基于管理员相互之间以特别的方式对变更控制问题进行的协调。

云数据中心与这种模式完全相反，它具有高度虚拟化和弹性的工作负载、当天实现应用程序需求，以及由服务目录面向 Web 的前端驱动的快速开通服务需求。管理员不再需要手动协调服务开通事件、手动更新配置数据库，和在托管实时生产中环境之前进行全面的系统测试。高度虚拟化云基础架构推动对实时配置、交换机拓扑数据的需求，以及跨物理和虚拟资源端到端跟踪虚拟机 (VM) 的能力、根据服务质量 (QoS) 和安全策略更改或重新映射租户与 VM 的能力。网络管理员无法瞬间执行这些功能，也不能孤立地执行这些功能。与外部控制器、云编排和服务开通中间件及服务集成，需要花费几周时间进行配置、测试和发布投入生产。服务级别协议 (SLA) 管理工具已成为一项核心云基础架构要求（参见表 1）。

考虑以下数据中心或云管理员情况。基础设施管理员非常了解服务器、交换机和互连的物理特性。在许多情况下，每个服务器的 MAC 地址、其物理位置（包括层次、行和机架信息）、分配的 IP 地址、到交换机的物理和逻辑连接，以及配置文件，被导入到资产跟踪和配置数据库应用程序中。此数据库信息对于查明问题和以有效方式执行中断/修复任务非常重要。在非虚拟化、非云环境中，此数据是静态的，可以由管理员轻松维护。在云中，服务器是虚拟化的，这些 VM 的放置经常变化，其中有基于这些位置变化的底层基础设施平台（如服务器、交换机、防火墙和负载均衡器）。

在大规模虚拟化环境中，运营商将无需担心 MAC 学习、老化、地址解析协议 (ARP) 更新，和上载任何 VM 位置变化到配置数据库中。从一个设备到另一个设备的路径在集中式拓扑数据库内应该是已知的，而且会实时更新。当外部集成到云控制器时，该路径会使网络配置、定制转发和故障排除变得更加容易。整个行业内大多数数据中心交换机都不允许任何转发路径的外部可编程性。它们是由供应商控制的封闭黑盒，具有预先设定的转发路径软件和硬件算法。这是外部控制器体现价值的一个明显例子。

同样地，也有其他用例-在将测试访问点 (TAP) 聚合到集中式收集点的流量工程中，加入特定标头以将 2 层流量重叠到 3 层网络上、根据内容进行流量分类、监视链路聚合组 (LAG) 或等价多路径 (ECMP) 组上的拥塞和散列效率，等等。由外部控制器管理的可编程交换机可以解决许多这类情况。

SDN 重叠、承载或两者兼有？

企业包含传统的和较新的基于云的应用程序的复杂组合。随着网络架构进化到支持云和虚拟网络的下一阶段，需要将这两重世界融合起来，因此考虑软件定义的网络中的两个主要要素十分重要。

重叠 SDN

SDN 重叠使用网络虚拟化从‘逻辑’或‘虚拟’拓扑中分离、抽象和解耦网络的物理拓扑（通过使用封装隧道）。这种逻辑网络拓扑通常称为‘重叠网络’。在这些体系结构中，一些网络特性和功能被移到重叠中，以控制数据、特定流量或转发路径。这可能包括：

1. 软件重叠，以便从交换机的控制平面转移管理功能到服务器
2. 特定网络服务，如负载平衡、访问控制、用作防火墙以及可见性

SDN 重叠的控制器利用现有物理网络来交付功能，如服务开通和可见性。

承载 SDN

控制器不解决存在于 L2/L3/L4 上的更广泛的复杂高性能网络问题集。为实现该目标，我们需要一个不折不扣的线速 SDN 数据平面作为物理‘承载网络’。为充分利用 SDN 重叠的功能，同时提供最大限度的透明性和性能，底层物理网络必须以无缝方式通过编程来线性扩展接口，使新网络虚拟化功能基本不存在争用并具有最少的端到端延迟时间。采用开放标准和开放式 API，使得与控制器无关成为可能，并且能够与其他网络基础设施元件互操作，如 Wi-Fi、负载平衡器和防火墙。

在每个云的核心，客户需求可扩展性、弹性和关键业务一整年每天 24 小时正常运行。数以百计的交换机可以轻松安置在同一数据中心内，并且需要对拓扑内任何位置的改变事件瞬间作出反应，不丢失数据包或造成拥塞状况。为实现这些需求，网络平台已经进化，包含许多分布式和嵌入式数据平面控制器功能。链路聚合控制协议 (LACP)、开放式最短路径优先 (OSPF)、ECMP，以及边界网关协议 (BGP) 是基于标准的分布式控制器功能（通常称为流量工程协议）的主要实例。由于大多数改变事件通常发生在本地，分布式方法可允许受影响的网络节点独立运作，在一瞬间作出反应并解析变化，几乎不会出现数据包丢失。这种分布式方法提供实现每一天连续不断的正常运行时间所需的高弹性行为。因此，当今网络很少成为出现应用程序中断状况的根本原因。

两全其美：软件驱动的云网络

网络对每个正在构建云的 IT 企业至关重要，无论云规模大小。因此，降低弹性而实现流量优化可能性不大。非常适合大多数公司的做法是让网络层根据标准协议执行其智能转发，并使用 Arista SDCN 依靠在应用程序层更紧密的集成来提升网络行为，利用开放式 API 实现每一层（无论是控制平面、管理平面或数据平面）的承载网络的可编程性。

比较常见的 Arista SDCN 用例包括以下内容：

- 多租户配置、移动性和 VM 管理的网络虚拟化
- 服务器和监视/统计工具之间的定制流量（或可定制的数据分流）
- 到位于互联网边缘的负载均衡器和防火墙的服务路由
- 大数据/Hadoop 搜索位置和实时诊断

Arista SDCN 可大大增强和自动化与这些用例相关联的操作。与外部控制器的集成提供用于映射、连接，和追踪高度可移动 VM 的定制智能，而网络设备中的分布式协议提供跨整个大型分布式拓扑结构的最佳路径数据转发和网络弹性智能。

迁移到私有、公共或混合云是革命性技术，但其采纳正在逐步进化。Arista 认可广泛的数据、管理和控制平面功能，赋予客户灵活性，如图 2 所示。

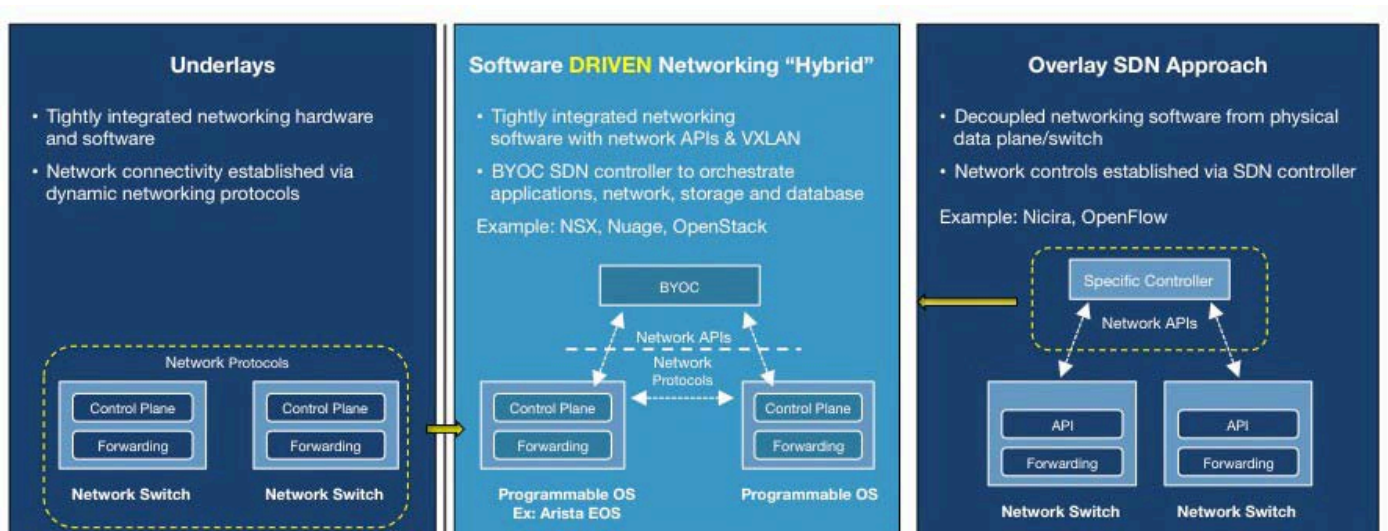


图 2：SDN 方法

ARISTA SDCN 的四大支柱

Arista Networks 认为，从 10Gb 到 40Gb 再到 100Gb 扩展的以太网—甚至万亿位—具有良好定义的 2 层和 3 层标准与协议，是适合大多数正在构建云的公司最佳方法。今天这种扩展允许 1 万个或更多物理与虚拟服务器和存储节点的大型云网络，在未来将扩展到 10 万个或更多节点，无需重塑互联网或引入专有 API。

在 VMworld 2012 中，Arista 展示了其高度分布式 2 层和 3 层叶子/骨干架构与 VMware 的虚拟可扩展 LAN (VXLAN) 集中控制的覆盖传输技术的集成。

这种集成提供无与伦比的多租户可扩展性，在同一 2 层广播域内支持最多 1600 万个逻辑分区的 VM。VXLAN 体现了 Arista SDCN 的多个设计原则，它是由 VMware、Arista 和其他几家公司的 IETF 提交的成果。

认识到构建这种高度可扩展的密集云只是该模型的一部分是很重要的。应用程序移动性、存储便携性、自助服务开通和自动化，以及动态资源优化带来了不同于许多传统数据中心（包括 20 世纪 90 年代末设计的基于客户端/服务器架构的数据中心）的新管理和运营挑战。

Arista 已经明确了这些云挑战，并且正在采用软件驱动的云网络的四大支柱有条不紊地解决它们（参见表 3）：

支柱 1：通用云网络

经由 2 层多机箱链路聚合组 (MLAG) 或 3 层等价多路径 (ECMP) 跨多个机箱扩展云网络是可扩展云网络的一种基于标准的方法。这种方法可以确保以非阻塞模式有效使用所有可用带宽，同时在任何单个机箱或端口出现中断状况时提供故障转移和弹性。MLAG 和 ECMP 以切合实际的方式涵盖了所有重要的多路径部署情况，无需引入任何专有发明。这些技术目前可扩展到 10 万个或更多物理计算和存储节点，并且可扩展到超过 100 万个虚拟机。

Arista 高密度主机聚合平台引入的新 Arista Spline® 架构，使大密度直接连接的主机能够连接到单层或双层网络。

随着新一代多核服务器 CPU 以及密集 VM 和存储的出现，这种具有无超额订阅容量、上行、下行和对等端口的不折不扣的叶子/骨干或样条拓扑至关重要。这些技术通常与服务器链路冗余集成（包括物理和逻辑集成）。行业标准为 LACP。Arista 具有完全互操作性，包括与 VMware 的 vSphere 5.1 版本的配置自动化。这种互操作性和配置自动化可以确保在虚拟网络接口卡 (vNIC) 级别链路正确配置为负载分担和冗余。

支柱 2：单一映像 2/3/4 层控制平面

一些网络供应商正试图用三十年的网络控制平面架构（非模块化、非数据库中心、专有）来应对 SDN。对于这些供应商，SDN 集成需要几年、昂贵的投入才能实现。客户将接收专有实施，在控制器级别具有供应商锁定，及其许多非标准分布式转发协议。Arista 已经亲眼目睹了这些问题。客户已经要求与 Arista 交换机以及与其他供应商的交换机的 2 层和 3 层控制互操作性。Arista 不得不对许多这些非标准协议。总之，其他供应商的交换机很难实施作为 SDN 架构的一部分，而且他们具有用于配置和管理的专有工具。这不是有发展前景的解决方案。

取代这些被吹捧的专有“构造”方法，基于标准的 2 层和 3 层 IETF 控制平面规格加上 OpenFlow 选项可谓是一个很有前途的开放方法，跨 Arista 交换机系列提供单一映像控制平面。未来几年，OpenFlow 实施将根据具体用例和控制器能够载入交换机的导向来进行。操作创新的实例包括：Arista 零接触服务开通 (ZTP) 功能，用于自动化网络和服务器服务开通，以及 Arista 延迟分析器 (LANZ) 产品，用于检测应用程序引起的拥塞。

支柱 3：全网络虚拟化

通过从应用程序解耦“物理基础设施”，获得更大移动性和资源池，从而使全网络虚拟化扩展了完全优化及分摊计算和存储资源的能力。因此，这对在整个网络中部署精心定义的细分和安全，以在网络上的任何位置无缝管理任何应用程序都是有意义的。这种解耦推动了云运营商的规模经济。全网络虚拟化是一个理想的用例，由外部控制器从网络中抽象出 VM 要求并定义移动性和优化策略，具有比目前可用的更大的网络灵活性。这种虚拟化需要一种隧道方法来提供跨 3 层域的移动性以及 API 支持，其中外部控制器可以定义转发路径。Arista 一直引领这一努力方向，提供了多种主要的虚拟机监控程序产品。这种集成的关键是 Arista CloudVision，它增强了多供应商开放式 API 支持，包括 eAPI（使用 JSON 的 RESTful api）、OVSDB、Openstack (Neutron ML2) 和开放式管理接口 OMI。

支柱 4：全网络云自动化

正在部署下一代数据中心的客户将面临数以百计（或许数以千计）的网络设备的管理和服务开通挑战，而且这种挑战的变化步伐越来越快。在当今的网络中，需要的不仅仅是少数 CLI 初始化功能，以便根据管理和自动化网络的需求进行扩展。应用程序移动性、存储便携性、自助服务开通和自动化，以及动态资源优化需要网络可编程性，而且，对于许多客户来说，需要可轻松根据其独特需求进行定制的预构建解决方案。Arista EOS+ 平台和 Arista EOS CloudVision 提供构建于开放式网络操作系统 Arista EOS 的开放式可编程网络。通过实现动态工作负载和 workflows 集成、可扩展性和可见性，对网络进行编程来获得业务敏捷性。—这是真正的 SDN 和软件驱动的云网络的核心，通过 Arista CloudVision 交付。

CloudVision

Arista 的 CloudVision 平台提供一整套服务，简化了虚拟化数据中心中的监视、管理和控制器集成。CloudVision 提供完全标准化的 API，支持商用或开源控制器。CloudVision 的开放式 API 可以确保系统架构师不被局限于潜在的专有虚拟化系统架构。除了 OVSDB 服务以外，CloudVision 还提供 RESTful，基于 JSON 的命令行 API，它允许管理员设计定制的网络管理和 service 开通工具。CloudVision 的开放式 API 基础设施降低了编排工具的开发成本。CloudVision 同时与多个控制器进行通信以适应异构数据中心，启用公共基础设施以经济高效地服务于云数据中心工作负载。Cloudvision 通过启用单点网络状态来交付此公共基础设施，没有专有锁定，可实现开放式合作伙伴集成，包括 Dell ASM、HP OneView、F5、Palo Alto Networks Panorama、Microsoft（System Center 和网络控制器）、Openstack Neutron ML2（RedHat、Rackspace、Mirantis、SUSE、Vmware）、Nuage Networks Virtualized Service Platform 和 Infinera Transport SDN。

CloudVision 提供一个网络服务抽象层，从物理数据中心基础设施中解耦控制器。这种抽象可从基础设施操作系统中隔绝控制器依赖，消除交换机操作系统和控制器版本依赖，并降低控制器认证和网络维护的相关费用。这包括通过智能系统回滚和升级以及集成的 workflow 遥测实现工作负载移动性、网络配置管理和控制的连网应用程序。CloudVision 的网络抽象极大地提高了控制器扩展，最多可达竞争产品的 10 倍，以单个接触点来控制云数据中心中的 Arista 交换机。

CloudVision 通过利用 Arista 的开放式和可编程的可扩展操作系统 EOS，带来软件驱动的云解决方案的好处，从而降低运营成本。CloudVision 通过利用 Arista 的全网络状态同步到可扩展操作系统 EOS 中，带来软件驱动的云解决方案的好处，降低运营成本。它为客户提供了云状工作负载自动化和 workflow 可见性的“交钥匙”解决方案。

CloudVision 构建在此网络抽象层上，还提供了全网络自动化的“交钥匙”方法。从经由零接触服务开通 (ZTP) 功能实现的初始设备服务开通，到进一步的全网络配置、软件版本和脚本的服务开通，CloudVision 在运营生命周期内都将有助于减少更改控制的时间和提高准确性。由于 CloudVision 构建在包括实时和历史状态的全网络数据库上，运营商实际上可以回滚其整个网络（包括配置和软件版本）到以前已知的工作状态。这是一个特别有用的工具，可在维护窗口期间发生这种需求时还原网络服务。这些功能专门构建用来帮助云状自动化成为主流，它们被预构建到 CloudVision 套件中，因此客户可以利用功能优势，而无需自定义集成和脚本编制本身。

Arista CloudVision

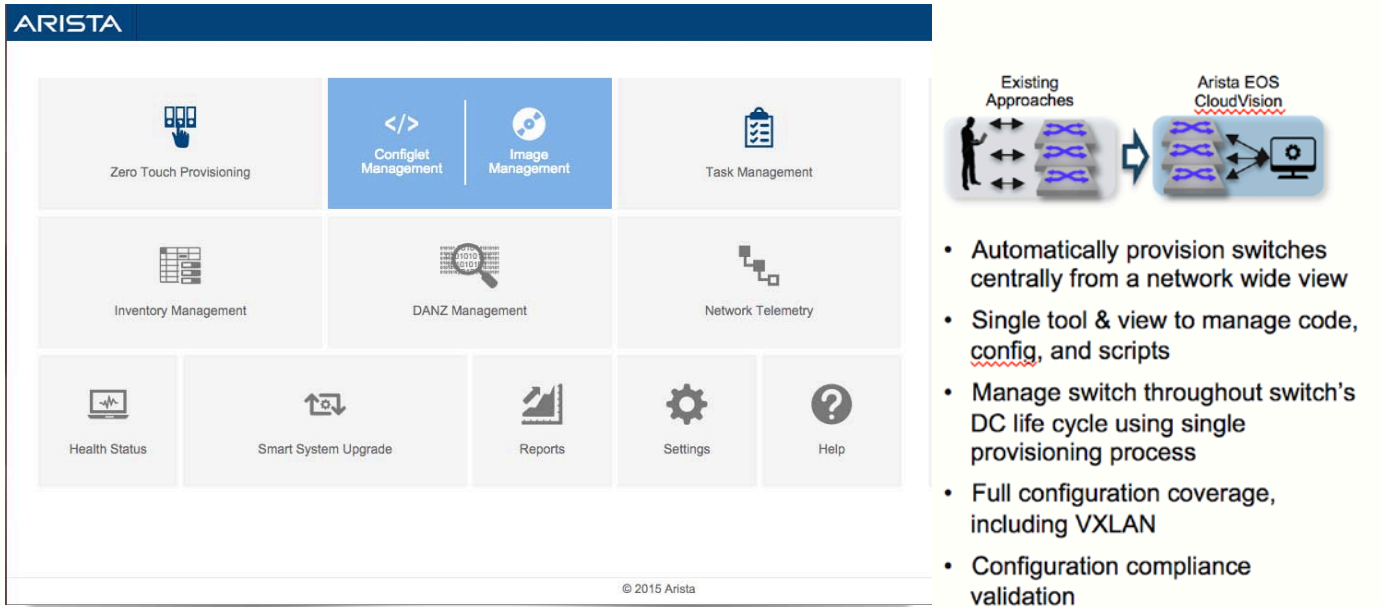


图 3: Arista CloudVision

Arista EOS，可扩展操作系统

Arista EOS 本质上是一个开放式、可编程的网络操作系统。它构建在标准 Linux 发行版上，从而具有与广泛的、基于 Linux 的 DevOps 工具生态系统集成的能力。Arista EOS 拥有丰富的 API 集，使用标准的和众所周知的管理协议，如 OpenFlow、可扩展通讯和表示协议 (XMPP)、系统网络管理协议 (SNMP)，并且本身能够支持常见脚本编制语言（如 Python）。而且，Arista EOS 提供单点管理，可与各种云堆栈架构轻松集成。不需要任何专有构造技术，并且不存在需要将每个交换机功能转变为复杂的分布式系统的问题。

Arista 的 EOS 软件架构特别适合于网络的扩展自动化或编程。Arista

EOS 部署一个内存驻留数据库 Arista sysDB 来管理系统状态—从路由表到 ACLS 再到计数器。数据库发布-订阅事件模型从本质上解决了在关注的进程和应用程序间协调与沟通状态改变的复杂问题。即使添加更多功能，Arista EOS 仍旧能够保持其可靠性和弹性。

Arista EOS 的开发提供了专为今天的云数据中心需求而设计的第一个网络操作系统。

传统网络操作系统

Arista 操作系统

封闭的/专有	开放式、基于标准的 linux 发行版
多个映像和软件系列	跨整个产品组合的单一映像
臃肿的内核	内存内系统数据库，用于存储状态
IPC 消息传递	可扩展的、弹性发布-订阅通信
中断驱动的	事件驱动的
零散的策略	开发人员友好
意图锁定	旨在集成

Arista EOS 可扩展性

Arista EOS 在所有层上提供可扩展性，如图 4 所示。

1. 通过 API（如 EOS API (eAPI) 和 SNMP）提供管理平面扩展能力。Arista API 利用简单、编写良好且广泛使用的编程模型，如 Java-Script Object Notation (JSON)、eXtensible Markup Language (XML)、Python、XMPP 等，与 EOS 管理平面互动，允许对 Dell ASM、HP OneView、EMC Smarts、VMware vCenter/vRealize、IBM Tivoli 和 Splunk 等管理系统直接进行编程访问。
2. 通过高级事件管理 (AEM) 提供控制平面扩展能力。AEM 是一个完整的事件处理程序子系统，支持对控制平面进行实时和事件驱动的编程。Arista EOS 与 SysDB 互动，支持网络交换机通过一组预定义触发信号对系统的任何状态变化作出响应。
3. 通过在线编程提供数据平面扩展能力。希望在网络上调整应用程序性能的客户，可以定制流量流向，即利用业界标准 OpenFlow 或无控制器的 Arista DirectFlow 结构体过滤和重定向流量。
4. 利用 Arista vEOS 和 VM Tracer 提供虚拟机扩展能力。Arista vEOS 控制平面允许 EOS 软件在任何虚拟化环境内像 VM 一样运行。这就为客户提供了虚拟机灵活性以支持实验室认证或开发工作。
5. 针对第三方开发的应用程序层扩展能力。Arista EOS 应用程序门户向通过 SDK 工具套件、脚本和 API 进行的第三方开发开放了 Arista EOS，使得新型应用程序和便捷集成同类最佳解决方案成为可能。
6. 访问 Linux 操作系统的所有工具，包括 shell 级访问。Arista EOS 可通过不经修改的 Linux 应用程序和越来越多的开源管理工具来扩展，以便满足网络工程和运营的需求。此外，Arista EOS 提供对全套 Linux 工具的直接访问，例如通过我们的完全二进制 Fedora 兼容能力执行 cpdump。

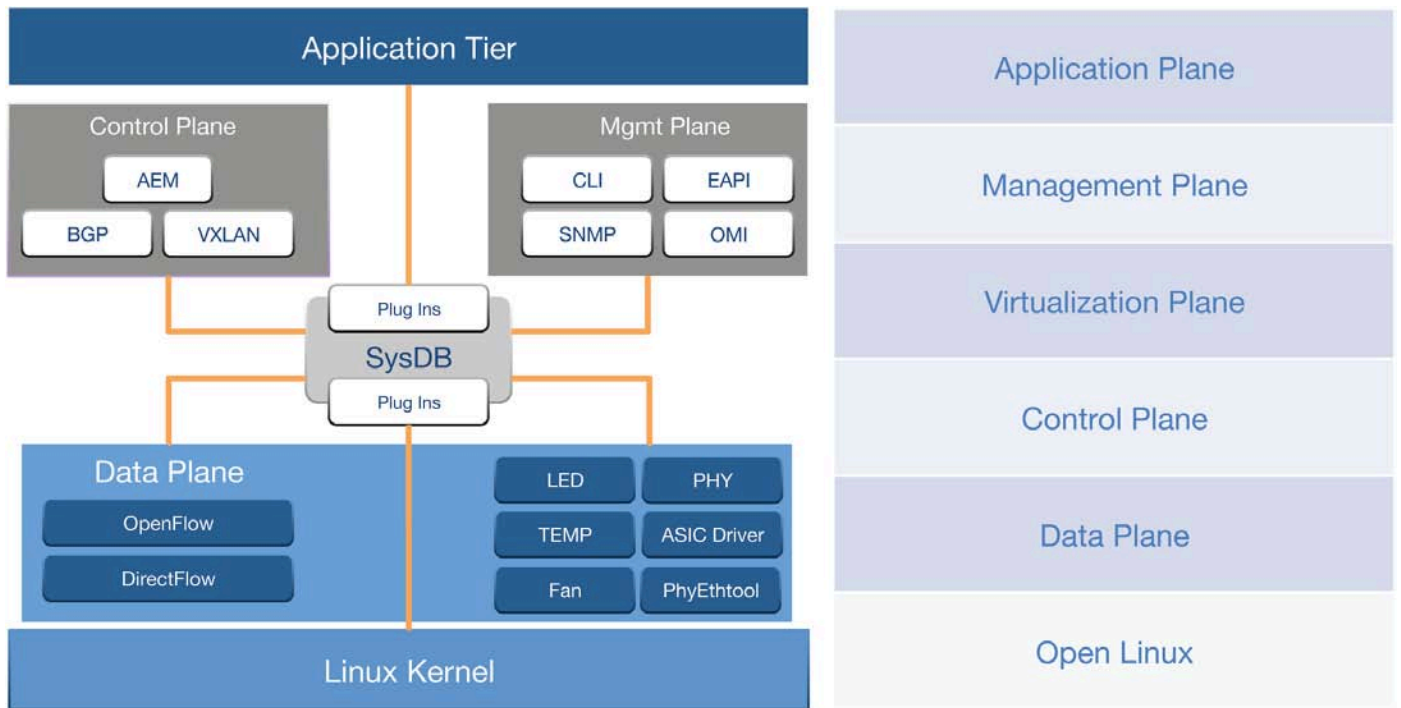


图 3: 所有层上的可扩展性

Arista EOS+ 平台

Arista EOS+ 平台通过构建在最佳网络操作系统上的丰富开发环境以及预构建的应用程序简化了网络编程。EOS+ 支持紧密集成到云堆栈、DevOps、网络控制器和应用程序/服务工作流程中。Arista 咨询服务可用于定制应用程序、集成堆栈，或按照完全自定义要求进行开发。

Arista EOS+ 包括

- **EOS SDK 和 eAPI:** 这些可编程接口通过提供到 Arista EOS SysDB 的直接接口，允许客户开发直接与交换机操作系统集成的应用程序。
- **vEOS:** 允许快速自动化部署、多拓扑设置、测试和认证的虚拟 EOS 实例。所有 EOS 功能，包括 STP、LACP、BGP、OSPF、JSON、OVSDB、Openstack 和完全 Linux 访问都是可用的。
- **EOS 应用程序:** Arista 技术协助中心 (TAC) 工程师为“交钥匙”端到端软件解决方案提供支持。EOS 应用程序实例包括，用于服务开通的 ZTPServer，以及用于 Splunk® Enterprise 的网络遥测解决方案。
- **EOS 咨询:** 专注于“企业”和“服务提供商”客户、提供自定义开发和集成的专业服务实践。EOS 咨询支持快速原型设计和开发，提供云应用程序工作流程的定制、DevOps 和第三方服务集成。

EOS 功能和应用程序

利用 EOS 和 EOS+，Arista 已开发核心网络功能和应用程序，其设计超过和超越网络操作系统的传统角色，源于其集成了三个核心组件：

1. 专用于配合重要 IT 工作流程和操作任务的 EOS 功能集合
2. 集成主要的同类最佳合作伙伴，使整个生态系统融合在一起
3. 网络应用程序的可扩展性，以便它可以适应任何网络操作环境，并扩展传统 IT 工作流程

通过使用 EOS, Arista 可实现开放式工作负载移动性。Arista EOS 连接到最广泛的网络控制器, 并将该集成与 VM 感知、自动服务开通和网络虚拟化连接在一起; 因此 Arista EOS 能够提供与当今编排和虚拟化平台最紧密且最开放的集成。总之, 网络运营商获得了在网络中的任何位置部署任何工作负载的能力, 凭借软件配置和可扩展的 API 结构, 所有服务开通在很短时间内即可完成。

Arista 智能系统升级 (SSU) 是一系列正在申请专利的技术, 使网络运营商能够将网络运作中最具挑战性的时期之一, 升级和变更管理操作, 与网络的运作行为无缝配合。采用 SSU 的网络, 能够从容地退出拓扑结构, 从直接连接的主机移出工作负载, 并使服务器负载均衡器虚拟 IP (VIP) 老化退役, 在此期间不会发现任何中断。许多网络运营商为实现最大系统运行时间所经历的多步骤、数小时过程成为了默认操作方法。SSU 已经证明与 F5 负载均衡器、VMware vSphere、OpenStack, 等等具有互操作性。智能系统升级已经预集成到 CloudVision 解决方案中, 因此客户拥有一套“交钥匙”解决方案来自动化全网络升级与回滚。

网络遥测完全是关于数据: 生成、收集和分发关于问题可能发生在什么地方作出消息灵通的网络决策所必需的数据, 从而保证数据可用、可轻松访问, 和建立了索引, 因此这些热点或有问题的区域被迅速修复, 并且故障排除简单快捷。网络遥测集成了 Splunk Enterprise 以及其他几个日志管理和循环/索引工具。

表 2: Arista SDCN 四大网络支柱

云网络需求	Arista EOS 支柱
1.高弹性、链路优化的、可扩展拓扑结构	IEEE 和 IETF 标准协议 MLAG 和 ECMP 拓扑协议
2.云适应、控制平面	所有平台使用一个二进制映像 零接触服务开通用于快速交换机部署 为 OpenFlow 和 OpenStack 提供行业支持
3.网络虚拟化	基于硬件的 VXLAN、NVGRE VM Tracer 用于故障排除 与虚拟机监控程序控制器集成 开放式工作负载部署和编排
4.开放式网络可编程性	所有层上的可扩展性 到 Arista EOS 的知名接口包括 Arista CloudVision 标准 Linux 工具 XMPP、XML、RESTful API、eAPI、EOS SDK EOS 的工具、解决方案和服务 vEOS、EOS 应用程序、EOS 咨询服务

ARISTA SDCN 用例

网络虚拟化

网络虚拟化至关重要，因为网络必须随 VM 和租户分区数量以及与其移动性、邻接、资源和安全策略相关联的相似规则扩展。此外，IP 移动性，比如无论 VM 被放置在 2 层或 3 层网络上，是否在同一数据中心中或被移动到不同数据中心，VM 保持相同 IP 地址，具有十分重要的意义。此外，从特设方法到一个基于预留的方法，分配带宽的能力正逐渐成为服务产品真正的区别因素（参见图 3）。

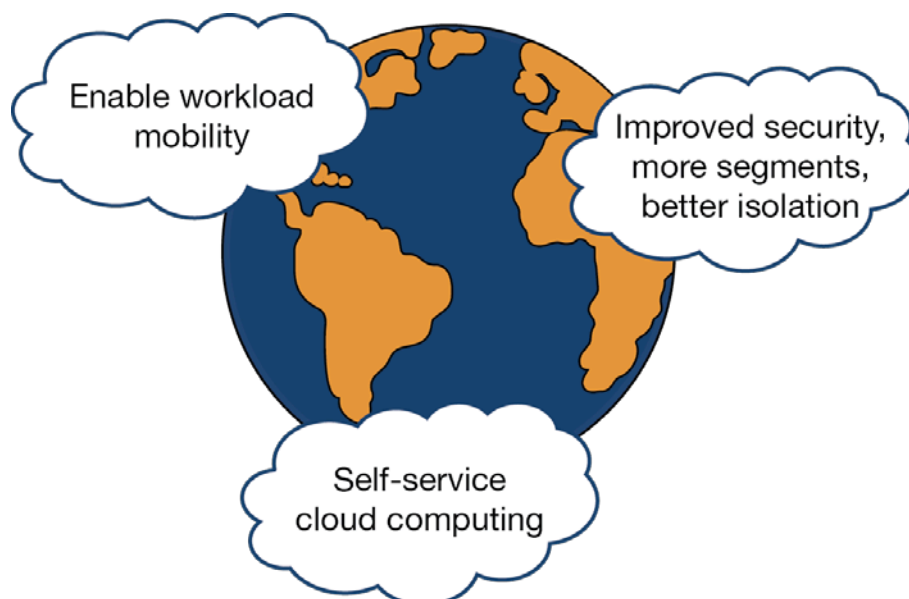


图 3：网络虚拟化及用例

虚拟化网络面临多重挑战。第一，每个叶子/骨干数据中心交换核心必须支持租户池远高于目前的 4K VLAN 限制，因为这是用于网络虚拟化的 VXLAN 和 NVGRE 协议的要求。第二，这些交换核心（或有界限的 2 层和 3 层交换域）必须提供大型交换表以便扩展到 1 万台物理服务器和 10 万个 VM。

第三，必须易于对交换核心的拓扑结构、位置、资源和服务感知实时数据库进行集中编程。第四，交换核心必须支持能够在外部控制器的三态内容寻址存储器 (TCAM) 中对定制流量进行编程。最后，必须有基于角色的安全配置模型，其中仅部分服务可用于外部控制器，而网络合规性由网络管理员管理和严格维护（且不可用于外部控制器）。

使用重叠隧道方法提供超过 4K VLAN 限制的租户池扩展，并且在物理上和逻辑上支持大型主机表，非常依赖硬件。交换机必须在交换芯片内支持这些功能。这是 Arista 云功能交换产品的核心支柱之一——用于处理大型交换表具有超低延迟效率的高度可扩展、分布式协议。对交换机进行实时编程，从集中控制器辐射到拓扑内数以百计的交换机，需要一种消息传递总线方法以及一个实时数据库。这是另一个 Arista SDCN 核心支柱——Arista EOS 以开放式可编程接口技术引领整个行业，包括能够将共驻在 Arista EOS 内的应用程序作为 VM 运行。此外，CloudVision 提供到外部控制器的接口，因为对转发表（或 TCAM）进行编程需要为 OpenFlow 和其他控制器形式因素提供支持。再则，作为一个 SDCN 核心支柱，Arista 已经证明能够使用外部控制器对交换机表内的主机和流量条目进行编程（参见图 4）。

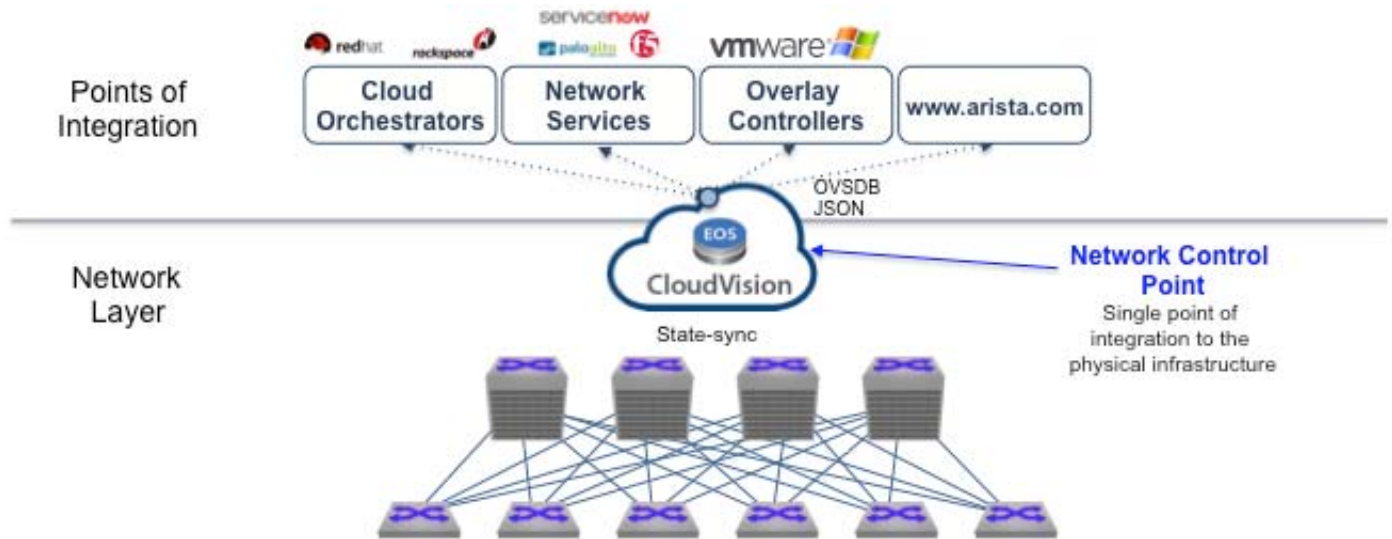


图 4：Arista 控制器集成

Arista 提供业界领先的转发平面隧道技术（如 VXLAN）并集成了网络虚拟化控制器。Arista EOS 是实时通信事件变化、通知以及通过外部控制器更新的主要促成因素之一。Arista 从跟踪与故障排除角度提供其屡获殊荣的 VM Tracer 应用程序。Arista VM Tracer 支持标准 VLAN 多租户虚拟交换机分段，并且已扩展为还可根据 VXLAN 身份跟踪和追踪 VM。

可定制的数据分流

收集和存档应用程序流量的需求在许多垂直市场内已成为一项基本的合规性要求。金融交易、医疗保健患者互动、数据库请求，通话录音和呼叫中心响应都成为被审核和记录的事件。此外，云运营管理器必须根据客户 SLA、带宽订阅率以及容量管理从云基础架构中收集流量数据。

对于大多数这些垂直市场合规性和 SLA 管理要求，网络是指导、收集、过滤、分析，和报告的理想场所。然而，考虑到通信量、每个云租户的应用程序和相关联 VM 数量，以及高速数据传输速率，则变得难以捕获、收集和存档流经网络的每一个数据包。这是一个典型的数据过载问题。

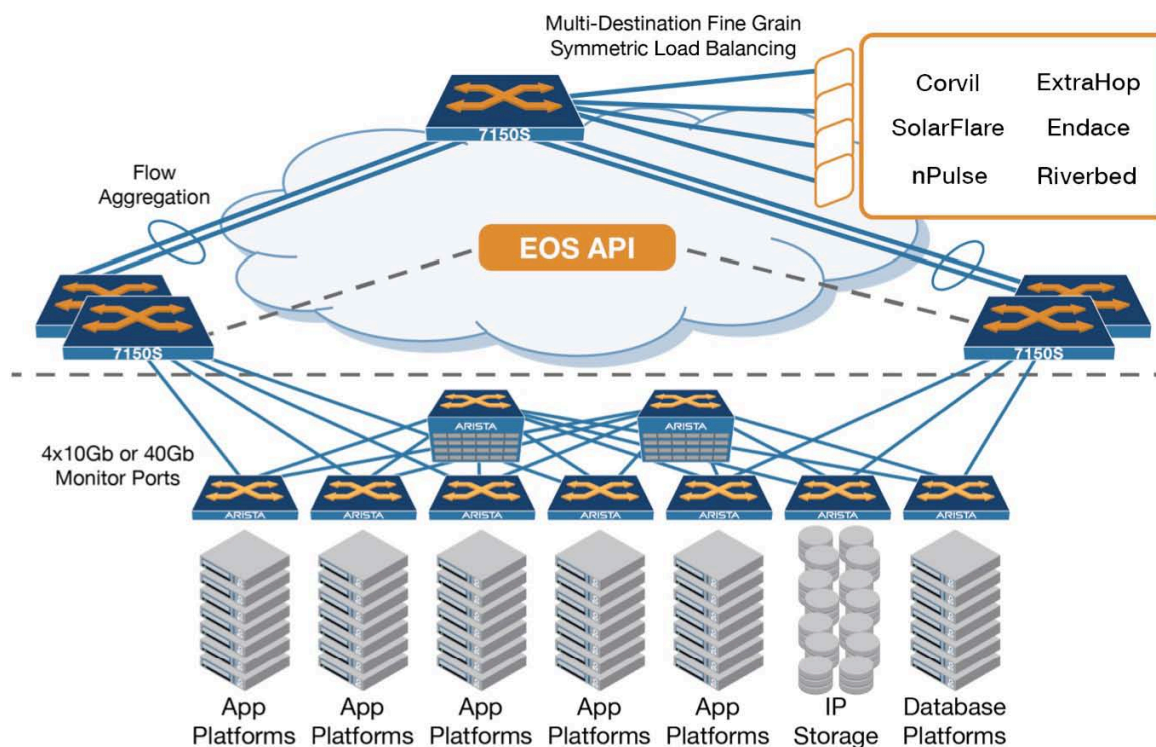


图 5: TAP 和 SPAN 聚合

减少这一问题的范围的一种方法是提供定制数据 TAP（参见图 5）——具体而言，就是对生成流量的端点与捕获、存储、分析和报告数据的收集器设备之间的数据流进行编程。这是外部控制器的一个理想用例。该控制器提供了应用程序端点间的中介层。它根据各个垂直市场合规性规则，确定需要捕获流量的端点、一天中需要进行收集的时间，以及专门设计用于收集、过滤和报告的收集设备。理想情况下，该控制器集成了基于信息技术基础设施库 (ITIL) 的服务目录租户载入接口，并且，基于一组收集选项可以捕获这些合规性要求作为一组可操作的配置事件（在每个 VM 激活启动和停止基础上）。

每次启动、移动或停止 VM 时，控制器将端点信息传达给交换机基础设施。之后在 VM 激活事件驱动时，交换机转发表被特别定制用于将流量跨非生产的通信端口重定向到特定行业的收集器（通常称为工具）。当 VM 启动并且运行 VM 的物理计算机位置被确定时，将会设置定制数据流和分流。当 VM 迁移到另一位置或停止使用时，它们将被删除和重新编程。

与外部控制器集成的定制数据分流是用于监视、报告和警示 VM 流量更具扩展性、有效性的行业标准方法。这对于在大型多租户云基础架构中扩展到 10 万个或更多 VM 的客户来说更是如此。这个用例运用了几个 Arista SDCN 核心支柱，包括在启动、移动或停止 VM 时对网络监视流量进行编程的需要；基于多租户标头和数据包信息以线速率镜像、转发和重定向流量的能力；以及实时检测拥塞状况并将警示发回控制器以进行实时修复的能力。今天，Arista EOS 提供了这些功能。

云服务

云主机以基于网络优势的应用服务极大地推动了技术转变，包括防火墙、负载均衡器、文件压缩和文件缓存设备。这些转变有两种形式。第一，许多这类服务变为虚拟化，在虚拟机监控程序内运行，该虚拟机监控程序与他们所服务的 VM 共驻和相邻（相对于中心）。第二，这些服务继续设在专用设备的 WAN 边缘，但是他们需要根据 VM 移动性变化具有动态的邻近意识。在第一个方案中，服务与 VM 一起移动。在第二个方案中，需要根据 VM 的新位置瞬间更新一个或多个边缘设备上的服务。从控制器到网络的数据包流量角度来看，第二个方案最令人信服，因为存在拓扑结构依存关系。

网络的控制平面保存拓扑位置信息，当 VM 在拓扑内移动时，它会从拓扑结构上最先获知。虽然应用服务管理平台也可以基于与外部虚拟化平台的集成确定 VM 的新位置，但是拓扑结构内的映射以及什么位置最合适提供该服务不是立即获知的。这可能导致应用程序中断、客户端可访问性问题，甚至应用程序性能问题，持续的时间已不可接受。

作为外部控制器和应用程序服务管理平台间的中介，Arista 已开发 JSON (JavaScript Object Notation) 远程过程调用 (RPC) API 来提供瞬时位置信息给云应用边缘服务。Arista 基于 Arista EOS 内的实时交互模型及其在与多个外部系统（包括控制器和其他管理平台）通信时和更新转发表并行工作的能力来提供此功能。Arista 已通过与 F5 Networks、Palo Alto Networks 等的密切合作开发了此基于 Arista SDCN 控制器的应用程序服务 API。

APACHE HADOOP 大数据

Apache Hadoop 通常被部署在专用机架中且未在虚拟化云基础架构中集成，许多客户构建出多个大数据计算机架并将其作为一项服务提供给他们业务分析团队。他们使这种技术可作为一个实用工具，而不是由单个业务团队拥有这些计算机架。业务团队利用一种分时方法，以便他们能够在专用时间段内加载其数据集和运行分析，之后根据队列中另一团队的指示将其从群集中删除。

作业完成的时间是 SLA 的关键要求，因为每个团队只能在给定时间段内依据结构化和非结构化的数据搜索和分析来发掘可操作的业务数据。结构化和非结构化搜索越快可以完成越好。网络在其中起着至关重要的作用，因为它提供拓扑位置数据，这有助于使每个搜索的定位最靠近数据存储的位置。关键技术要素是 MapReduce 以及将网络拓扑数据加入这些搜索算法的能力。此外，处理和报告微爆状况以确定瓶颈有利于搜索位置的决策。

Apache Hadoop 大数据需要几个云网络支柱。分布式拥塞、微爆，和负载均衡控制在交换机控制平面和转发平面内确定，对于确保没有数据包丢失和实现完成结果的最佳时间至关重要。提供拓扑数据的实时外部接口以及节点映射感知，培养了 Hadoop 开源开发人员并促进了商业应用程序（例如，Cloudera）与 MapReduce 技术的集成。提供基于拥塞和超额订阅的事件触发器，有助于在其实时发生时将搜索重定向到其他具有更多网络容量的机架。所有这些都是 Arista EOS 的组件。

SDN 控制器

对于云控制器的需求十分明确且在不断增长。VM 移动性、多租户流量隔离、实时网络路径追踪、防火墙规则更新，和定制数据捕获等用例，推动了对更大可编程性的需求。网络平台的转发和控制平面的外部控制器，提供 VM 服务需求与托管 VM 的基础架构间的可编程中介层。控制器将这些服务需求转化为可操作的控制和转发逻辑到计算、网络、存储和应用程序服务平台。这些基础设施平台（包括网络交换机），基于来自控制器的输入执行操作。

由于用例、载入技术以及用户社区（私有、公共和混合）的日益多样化，因而没有通用的形态规范或公认的一套标准来界定控制器如何介导和交互。控制器市场还处于其起步阶段，开源产品、客户开发的产品和基础设施系统产品具有专用的嵌入式控制器（参见表 3）。这需要一个开放式、高度可编程的方法来集成各种控制器形态规格以及用于用例的实现。Arista CloudVision® 提供了用于与所有网络虚拟化控制器紧密集成的通用框架。

表 3: Arista CloudVision 集成

控制器	详细信息
OpenFlow	Open Flow 1.3 集成包括 Open Daylight、NEC。
OpenStack	OpenStack Neutron ML2 插件；合作伙伴包括：VMware、Red Hat、Rackspace、Mirantis、SUSE
VMware	与 vSphere 5.0/6.0、vCloud、NSX-MH、NSX-V 集成的原生 VMware
Microsoft	与 Microsoft System Center、Hyper-V、网络控制器和 Azure 集成云服务
Nuage Networks	与 Nuage Networks Virtualized Services Platform 集成
Infinera	与 Infinera Transport SDN 集成
网络遥测、监视、 管理和安全- F5、 Riverbed、Palo Alto Networks、 VMWare Vrealize Operations、A10、 Fortinet、HP OneView、Dell ASM、 EMC SMARTS、SPLUNK、 ServiceNow、VM Turbo	正在开发针对主要合作伙伴的本地 API 调用；实现工作负载移动性和可见性、网络管理和服务插入，包括安全性的网络自动化。

Arista 将工作重点放在最佳配合这些用例的控制器供应商以及 Arista 交换机最优化适合的市场上。这种集成的基础集中在 Arista EOS 以及在跨整个拓扑更新控制和转发平面时与外部控制器实时交互的能力上。这种集成需要高度可扩展、基于事务处理的实时数据库和现代消息传递网络操作系统架构。这是 Arista EOS 的核心技术组件。

Arista 正在与许多不同的控制器形态规格和行业领导者集成。这种集成包括与 OpenFlow 开源发行版（特别是版本 1.3）的基于 CloudVision 代理的集成，以及几个也是基于代理 (OpenFlow) 的独特用例，包括与 Open Daylight 和 NEC 控制器的集成。此外，Arista 一直是 OpenStack Neutron 项目中的积极贡献者，已开发了双堆栈 ML2 驱动程序用于统一物理和虚拟网络设备配置。OpenStack 对许多想要为自己的定制品牌服务提供开源服务目录、服务开通、运营管理架构以及 Arista CloudVision（支持 Redhat、Rackspace、Mirantis、VMware、SUSE openstack 发行版）的服务提供商来说极富吸引力。最终，Arista 已开发出一种方式采用无控制器操作来扩展 OpenFlow 的功能，这种无控制器操作使用 Arista DirectFlow 实现对特定流量交换操作的直接 CLI 和 eAPI 控制。此接口提供机对机通信从而对防火墙、负载平衡，和其他应用程序层服务优化器之间的服务路径进行动态编程。

总结：网络即是应用程序

Arista SDCN 方法体现了软件定义的网络的许多理论和设计原则；然而，该公司采取了基于云计算特定的可扩展性、虚拟化、移动性和自动化需求的更偏外科的角度。以太网交换已取得良好进展，有许多分布式转发功能为世界上很多最大的数据中心提供可扩展性和弹性。

显然，云技术和云自动化与优化的操作优势催生了对外部控制器的新需求，用于通过单点管理对服务进行抽象，或用于为高度定制的应用程序定义独特的转发路径。Arista 完全采纳了这些原则。Arista 已定义了基于高度模块化、弹性、开放式、以状态为中心的网络操作系统 Arista EOS 的四大支柱，开发人员和最终用户可以在其中添加自己的脚本和管理工具。Arista 继续基于此操作系统展开构建工作，它是 SDCN 的关键构建模块。Arista 的独特产品包括提供工作负载移动性、监视和可见性，以及实时网络遥测的应用程序，用于与云运营和管理工具集成。

术语表

命令行接口 (CLI): CLI 是配置、检查、存档, 和获得交换机状态的实际标准。CLI 是一个提示驱动、字符驱动的基本编程语言, 需要非常了解底层交换机操作系统的技术和语法。CLI 通常基于每设备使用, 提供用于更改和获得逐个特性交换机信息的快速、直接接口。

在技术上先进掌握 CLI 使用的系统管理员, 对交换机的功能有着深刻理解。

简单网络管理协议 (SNMP): SNMP 编写于 20 世纪 80 年代末期, 它是用于管理交换机和路由器的、与 CLI 相比更高层次、更抽象的接口。SNMP 是许多基于 GUI 的管理应用程序的实际接口。SNMP 在交换机设备上需要一个代理 (SNMP 代理)。代理可以支持只读和读写操作。SNMP 代理公开管理数据, 特别是管理信息库 (MIB) 内包含的信息。MIB 打包一系列低层次信息并将该信息发送给已注册和被授权接收 MIB 数据的集中式管理工作站。

可扩展通讯和表示协议(XMPP): XMPP 是用于即时通讯和表现技术的一项 IETF 认可的标准。XMPP 成为用于从交换机到集中式控制点 (控制器) 传达状态信息的一项正式协议。XMPP 采用客户端/服务器架构。交换机传达到一个或多个中央控制器, 但它们不作为相互间的对等点进行通信。不存在一个权威 (服务器) 控制器, 从而提供非常适合于云应用程序的各种实施方式。XMPP 提供一个多交换机消息总线方法用于从控制器发送 CLI 命令到任何参与的交换机或交换机组。

OpenFlow 协议: OpenFlow 协议提供一种在交换机和一个或多个集中控制器之间进行通信的方法。此协议与其他协议一样, 都是基于 TCP/IP, 具有安全和加密定义。该协议使用一个熟知的 TCP 端口 (6633) 用于传达到控制器。交换机和控制器通过交换由一个位点特异性私钥签署的证书来互相验证。该协议通过一个良好定义的标头字段和标记交换交换机和流量信息。有关更多信息, 请参阅《OpenFlow 交换机规格》。

OpenStack: OpenStack 处于更广泛的程序级别。它不仅仅定义通信接口和一套用于与集中控制器进行通信的标准。OpenStack 拥有超过 135 家公司对其进行积极贡献, 包括服务器、存储、网络、数据库、虚拟化和应用程序公司的代表。OpenStack 目标是, 使任何公共或私营企业能够在标准硬件上提供云计算服务。Rackspace Hosting 和 NASA 于 2010 年正式推出了 OpenStack。OpenStack 是免费的、模块化、开源软件, 用于开发公共和私有云计算架构、控制器、自动化技术、业务流程, 以及云应用程序。

虚拟化 API: 虚拟机监控程序和虚拟机监控程序管理工具中有一些可用的 API, 适用于与以太网交换机和集中控制器进行通信。这些 API 和工具定义 SLA 的关联性规则、资源池、租户组和业务规则。此外, 这些工具可按照业务策略和服务级别自动化低级别服务器、网络和存储配置。每次添加或更改新的 VM, 当它在云内运行时, 这种自动化将可减少管理点和降低运营成本。

ARISTA

圣克拉拉市—公司总部
5453 Great America Parkway
Santa Clara, CA 95054
电话: 408-547-5500
www.arista.com

爱尔兰—国际总部
4130 Atlantic Avenue
Westpark Business Campus
Shannon
Co. Clare, Ireland

新加坡—亚太地区行政办事处
9 Temasek Boulevard
#29-01, Suntec Tower Two
Singapore 038989

版权所有 © 2015 Arista Networks, Inc. 保留所有权利。CloudVision 和 EOS 是注册商标, Arista Networks 是 Arista Networks, Inc. 的商标。其他所有公司名称是其各自持有者的商标。本文档中的信息如有更改, 恕不另行通知。有些功能可能尚不可用。Arista Networks, Inc. 对于本文档中可能出现的任何错误不承担任何责任。 MM/YY