

基于开放标准构建世界级的数据中心网络

IBM® 红皮书® 观点出版物，来自 IBM System Networking

作者：**CasimerDeCusatis**，博士，
IBM 杰出的工程师，系统网络战略联盟的杰出工程师与首席技术官

要点

具有可互操作网络的开放式数据中心能够带来显著效益：

- ▶ 扩展至上万台虚拟机，减少资本和运营支出达 15% 至 25%
- ▶ 通过虚拟化网络资源，在数据中心内部及数据中心之间自由移动虚拟机，优化利用率
- ▶ 将预先优化的解决方案与数据中心的其余部分无缝集成
- ▶ 自动执行网络供应服务、管理和故障诊断，以增加其灵活性并减少出错机会
- ▶ 创建更平坦的应用程序感知网络，此类网络只需一次性连线即可使用

转变数据中心网络以创造业务价值

数据中心正在经历一次重大转变，向更智慧、更动态的基础架构发展。更为灵活的 IT 架构在不断涌现，以应对当前业务应用程序和新领域（例如，云计算、多租户、自行携带设备（BYOD）、大数据与分析）的需求。

作为动态基础架构趋势的一部分，数据中心网络的角色也正在发生变化。这种变化导致企业重新评估自身当前网络，而这些网络被设计时，未曾考虑用以处理现代化的工作负载和应用程序。多年前，以太网是作为校园网络技术进行开发的，通过各种中继器、集线器和交换机使众多站点（哑终端）互联互通。传统的多层以太网网络（包含访问层、聚集层及核心层）非常适合校园网。在这些网络中，网络配置在很长时间内不会发生明显变化，用户之间的流量模式不会因延迟和预订过量而受到影响。

在此环境中，网络设计将第 2 层服务推送至访问层以下。同时，诸如生成树协议（STP）之类的协议足以提供有限的拓扑集，即使在流程中禁用交换机之间的链路。这种方法曾卓有成效，以致于数据中心内的网络也采用了相同的实践。

然而，传统的多层以太网网络并不符合现代数据中心的需求。如今，服务器已高度虚拟化，单一的物理服务器上就存在数以计百的虚拟机（VM）。由于可以随时创建、调整、移动或销毁这些虚拟机，因此网络必须变得更灵活、更动态。除了更传统的应用程序需要多租户，此概念是公共云和私有云计算的基础。

遗憾的是，网络虚拟化程度与服务器或存储器虚拟化程度并不相同。技能高超的网络工程师可能需要耗费数天甚至数周时间来供应新的功能，通常还涉及数据中心内的移动电缆。大量虚拟机在这种环境下还要求网络以经济有效且节能的方式进行扩展。

数据中心网络也存在诸多其他问题。流量模式已发生变化，在许多数据中心内，高达 75% 的数据通信发生在服务器之间（称为东西向流量）。延迟和预订过量所导致的变化会显著影响应用程序的性能。成本压力和利润的不断缩水促使将特定于应用程序的各种网络融合到公共光纤网中。最后，各组织会寻求新的途径，快速可靠地部署具有复杂功能的应用程序，例如，优质服务与资源共享。面对令人如此困惑的新应用程序集，许多组织被迫将自己有限的资源用于保证当前网络的正常运行，而非用于部署新的应用程序以产生更多收入流。



鉴于这些变化，网络行业正在经历严重的断层。旧的技术正逐渐被新的业务开展方式所取代。网络虚拟化、集成、自动化与融合将成为行业接下来的“前沿阵地”。网络虚拟化技术，例如，软件定义的网络（SDN）与 OpenFlow，结合了网络叠加，从而能够在所选的流量中实现丰富的管理、隔离和优化功能。随着网络的不断发展，注意力将转向应用程序感知网络。此次转变将对整个数据中心基础架构产生影响深远的根本性改变，需要制定路线图，在不影响日常业务运营的情况下循序渐进地实施此类改变。

正如您可能预期的那样，在征服新的领域时，必须考虑多条路径。网络行业已通过一系列新的行业标准技术和供应商专利技术对这些问题做出了回应。令人遗憾的是，并非所有供应商都使用兼容的技术，造成难以在不依附于单一供应商的情况下评估与实施适宜的网络解决方案。这个令人困惑的选项列表造成网络难以管理，且几乎无法优化。

虽然将整个网络托付给他人看似很有吸引力，但是行业分析人员已经揭示了这种战略的真相。现实中，未能实施多供应商方法意味着针对网络服务需多支付 15% - 25% 的成本。事实上，根据最近的一份 Gartner Group 报告，¹ 未重新评估长期实施的现任供应商决策的 CIO 无法履行其作为受托人的职责。（Gartner Group: 一家信息技术研究与分析的公司）

IBM 的观点是：构建一流数据中心网络最为经济有效的方法是基于多家供应商支持的开放行业标准。

为帮助引导组织通过当前网络断层与转换遗留的数据中心，IBM 建议使用端到端的数据中心网络参考架构，该架构基于开放行业标准。作为 *开放数据中心可互操作网络 (ODIN)*，它可利用现有行业标准和最佳实践，创建出具有多供应商互操作性的世界一流的数据中心网络。ODIN 并非新的标准，而是运用来自标准主体的现有标准和最佳实践，例如，IEEE、IETF、INCITS、IBTA 与 ONF。IBM 发布了一系列有关 ODIN 的五项技术概要文档，欲知详情，请参阅以下 System Networking Web 站点：

ibm.com/systems/networking/solutions/odin.html

ODIN 计划自 InterOp 2012 推出以来受到广泛认可。至今，许多行业领头羊已公开支持 ODIN，包括瞻博网络、Brocade、华为、NEC、爱德华、讯远通信、BigSwitch、美国极进网络与马里斯特学院。IBM 已建立了强大的合作伙伴系统，以帮助实现这种开放式可互操作网络的愿景。

下一代以数据为中心的光纤网

在过去几年里，数据网络已发生了天翻地覆的变化，自因特网诞生起，所涌现出的新标准和架构之规模是前所未有的。重要的是，不仅要了解这些标准能为您的业务实现什么，还要了解这些标准部署在端到端解决方案中时彼此如何交互作用。

ODIN 可解决现代数据网络所面临的诸多关键问题，包括自动化、集成和管理。与经典的以太网架构相比，ODIN 光纤网具有一些独特优势。

如第二页图 1 中所示，ODIN 可提供更平坦的网络，从而减少延迟并提高网络性能。更大的第 2 层域主要支持东西向流量。通过提高虚拟机的移动性，可以在多台服务器与存储器之间实现负载均衡。可启用任意拓扑和东西向流量模式。其范围从交换机堆叠和链路聚合技术到网状 Clos 网络，基于 IETF 控制板 IS-IS 算法（这些算法使用 TRILL 替换 STP）。

¹ 揭开单一供应商网络之谜核心研究报告，2010 年 11 月 17 日。

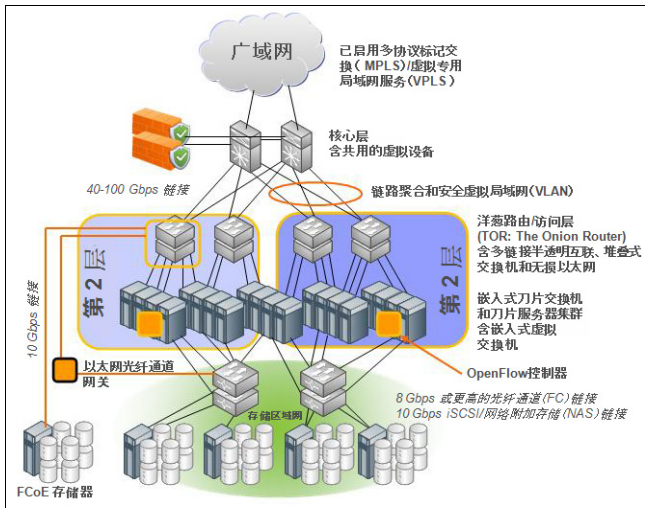


图 1 ODIN 的主要特征

此外，当使用 SDN 时，网络控制器（例如，OpenFlow）会发现交换机并创建拓扑。如果设备或电缆发生故障，离其最近的邻居会报告事件，控制器会快速重新融合时间来绕过故障。通过到 SAN（光纤通道或 iSCSI）的网桥支持现有存储器，而且这些存储器支持使用无损以太网和聚合式以太网光纤通道（FCoE）。自动启用物理断开连接的路径可实现高可用性或存储访问，从而简化管理与配置。

数据流经过优化可提高效率。同一服务器上两台虚拟机之间的流量无需离开此服务器。虚拟化的网络支持多租户，而这又反过来扩展到服务器管理程序或者相隔数百公里的多个站点。这种类型的网络使用虚拟叠加技术，称为分布式叠加虚拟以太网（DOVE）。DOVE 可用于公共和私有云服务，并且能够在合并或收购之后整合多个数据中心。

网络功能作为服务提供，包括多租户虚拟机安全和负载均衡。因此，新功能可更快地被添加至网络中，因为可复用的软件应用程序位于网络控制器、虚拟机访客系统或硬件上。动态网络意味着可以根据应用程序的需求对带宽加以管理。例如，它可将少量的可用带宽自动整合为大型的虚拟连接，以处理流量高峰。

正如您对开放标准的期待一样，ODIN 允许您选择刀片服务器或机架安装式服务器（x86、IBM POWER® 或 IBM System z®）、虚拟或物理交换机以及操作系统或管理程序。ODIN 允许您根据应用程序需求灵活地制定端到端的解决方案。ODIN 还可满足 Figure 1 中所列标准之间的交互。它支持以不中断的方式从当前网络进行迁移，并提供端到端的方法，以便从服务器管理程序扩展至城域网（MAN）或广域网（WAN）。ODIN 支持开放式应用程序开发者自动执行关键任务（例如，提供新服务），从而促进网络创新。

IBM 的观点是：数据中心网络正经历断层，或者说处在会发生颠覆性变化的时期，在此期间，网络虚拟化、自动化、集成与融合将成为接下来的前沿领域。

使用 ODIN 的智慧型数据中心基础架构

ODIN 参考设计可以解决数据中心网络转型中的三个关键问题，如图 2 中所示。

集成	简单统一的管理 网络服务敏捷性 软件定义的网络平台
自动化	应用程序感知网络 动态配置 一次性连线光纤网
优化	聚合 (FCoE, iSCSI, NAS 和 RDMA) 单一、平坦 (一个受管交换机集群) 安全、随需求而扩展的架构

图 2 智慧型数据中心基础架构的组成部分

首先，现有网络由一批独立的交换机、路由器、适配器和其他设备组成，并且每个都有其自己的元素管理器。这种配置会导致过度复杂的网络设计，无法进行有效扩展。这些网络具有多个管理域和次优点服务，由每个网络设备自行提供。ODIN 可以形成一种集成的解决方案，将服务器、网络 and 存储器与其各自的管理域结合起来组合在一起。

其次，传统的网络是静态配置，需要大量手工干预以实现配置、维护与故障诊断。手动干预使网络安全和服务质量易受到人为错误的影响。不仅如此，旧的网络从未打算支持多租户云计算应用程序。ODIN 基于行业标准 SDN，推行自动化、动态且虚拟化的网络。ODIN 光纤网在线上推出新的服务时，无需重新配置电缆。事实上，这种自动化网络是应用程序感知型的网络，也就是说，它可以在敏捷的业务环境中适应工作负载、安全威胁和其他因素的变化。

再者，一个组织具有太多的传统网络，而每个网络中均配备了太多的交换机、层和设备。这种方法浪费资本设备，占用太多的空间，使用过多的能源，并且无法实现低延迟和高效的性能。ODIN 针对更平坦的光纤网进行了优化，可处理现代化数据中心内的流量模式。随着时间的推移，多种类型的光纤网可被聚合并整合，从而为网络提供高效的野葱魁斗型方法。通过提供集成、自动化且优化的光纤网，ODIN 可提供更有弹性且更易于扩展的数据基础架构，以改善端到端的应用程序性能并减少资本与运营开支。

可选择的平台

IBM 与我们的合作伙伴为您的数据中心网络提供一系列符合 ODIN 的解决方案。您无需没日没夜地进行数据中心转型。您可以根据切身利益，从各种便捷的起点开始转型。例如，您可能希望通过服务器集成来降低资本支出和运营费用，通过 WAN 来备份重要数据，或者只是基于自身的特定需求开展概念证明实验。

IBM PureSystems 产品线是网络集成如何为企业增添价值的一个力证。通过提供出厂时经过测试和预先配置的服务器、存储器和网络，IBM 可以将部署时间从数周或数月缩短至一天或者更短。PureSystems 打造了理想的入门级云计算平台，每个计算节点可托管的虚拟机数量比原先的选择多出 54 %。

正如您对于开放式数据中心的预期，PureSystems 的核心就是灵活性。您可以选择服务器、操作系统、管理程序的类型或其他属性，以适应自身需求，而非依附于单一的解决方案。PureSystems 集群或集成系统包含虚拟网络 (IBM Distributed Virtual Switch 5000V)，刀片机箱嵌入式交换机和机架顶部交换机 (IBM RackSwitch? G8264)。利用 Flex System Manager，可以从单一界面来管理您所有的物理与虚拟资源，而 PureSystems 中的网络同样支持 ODIN。PureSystems 内部网络经过预先集成，可实现尽可能低的延迟，并最大化虚拟机的移动性。它还经过与多家 IBM 网络合作伙伴的互操作性测试，可以轻松集成您现有数据中心的资源。

又如，远程存储器备份和虚拟机移动性多年前正是由 IBM SAN Volume Controller 开创的。虚拟机移动性不仅改善了应用程序可用性，还提供了更有效的途径来使用有限的存储资源。虚拟机移动性适合用于业务连续性、避免灾难或灾难恢复，包括计划内的事件（例如，迁移数据中心或者消除由于定期维护而产生的停机时间）。

考虑到日益增长的全球劳动力，此方法还提供了负载均衡，并增强跨多个时区的用户性能（称为胰詹宦淑方法）。不仅如此，通过远距离移动工作负载，您可以优化能耗成本，因为最低成本的电力通常在夜晚供应（称为以虏宦淑）。目前，SAN Volume Controller 扩展了集群支持 ODIN 功能，包括无损以太网和行业标准存储器（光纤通道和 iSCSI）。通过使用暗光纤“波分复用 (WDM)”、FC-IP 或多协议标签交换 (MPLS) 或虚拟专用局域网服务 (VPLS) 选项，可以将距离扩展至 300 公里。同时，有多家供应商合作伙伴已通过了 IBM 的测试和资格认证。

IBM 已在多个行业中为客户设计并部署了下一代的网络解决方案，包括医疗卫生业、金融服务业、云供应商和电信业。例如，IBM 正在积极为 SDN 标准和协议做贡献。作为“开放式网络论坛 (Open Networking Forum)”标准机构的鼻祖，IBM 正致力于采用不含供应商专有扩展的开放式行业标准 SDN。

IBM 率先垂范，在市场中推出 SDN 解决方案，例如，首个商用支持 OpenFlow 交换机 - 10/40 Gbe IBM RackSwitch G8264。IBM 正与众多行业中的客户（包括华尔街的主要金融机构）合作，以实现 SDN 和 ODIN 解决方案。IBM 还在帮助全球电信公司重新定义其网络，通过使用此项技术，形成以云为中心的网络。通过与支持 ODIN 的其他供应商合作，许多领域中的业务解决方案均已获得了资格认证。

后续步骤：IBM 如何提供帮助

如果您的组织希望了解有关基于开放标准的多供应商网络如何帮助您创建灵活的应用程序感知网络的更多信息，请首先考虑您最大的困难是什么。然后为您的网络转型挑选起点。

例如，像 IBM PureSystems 之类的集成产品，提供了一种简单的途径来介绍这些概念。IBM 全球服务部门可帮助将 IBM PureFlex? 或 IBM PureApplication? 解决方案与您现有的网络进行整合。反之，您可以根据构建块设计自己的网络，例如，通过使用 IBM 5000V 虚拟管理程序交换机、刀片机箱嵌入式交换机和机架顶部交换机。IBM System Networking 能够提供最佳实践，并推荐使用这些产品的网络架构。欲知您所处行业中 IBM 解决方案的更多示例，请参阅下一部分中所列出的参考资料，或者联系当地 IBM 业务代表。

查阅参考资料，获取更多信息

欲知本文标注的概念详情，请参阅以下资源：

- ▶ 数据中心网络博客
ibm.com/connections/blogs/DCN
- ▶ Debra Curtis 和 Mark Fabbi, 揭开单一供应商网络之谜, Gartner RAS 核心研究报告, 2010 年 11 月 17 日
<http://www.dell.com/downloads/global/products/pwcnt/en/Gartner-Debunking-the-Myth-of-the-Single-Vendor-Network-20101117-published.pdf>
- ▶ IBM PureSystems, 模式的力量
ibm.com/ibm/puresystems/us/en
- ▶ IBM SVC VMware 互操作性知识库
http://kb.vmware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalId=2032346
- ▶ 针对 IBM 网络解决方案的 Juniper Networks 互操作性指南
http://kb.juniper.net/InfoCenter/index?page=answers&type=search&searchid=1343918845874&question_box=Qfabric+storage+interoperability+guide
- ▶ 网络解决方案：开放式数据中心可互操作网络 (ODIN)
ibm.com/systems/networking/solutions/odin.html
- ▶ OpenFlow 研究由 IBM 和马里斯特学院赞助：
<http://openflow.marist.edu>
- ▶ T. Bundy, M. Haley, F. Street, C. DeCusatis, The impact of data center convergence, virtualization, and cloud on DWDM optical networks both today and into the future, Proc. 太平洋电信委员会 2012 年年会, 夏威夷檀香山市 (2012 年 1 月)
- ▶ C. DeCusatis, Enterprise networks for low latency, high frequency financial trading, Proc. 企业计算社区会议, 2011 年 6 月 12 日至 14 日, 纽约州波基普西市马里斯特学院 (2011 年)
<http://ecc.marist.edu/conf2011>
- ▶ C. DeCusatis, Towards an open data center with an interoperable network: enterprise networking using open industry standards, Proc. 美国国家科学基金会企业计算会议, 马里斯特学院 (2012 年 6 月 11 日至 13 日)
- ▶ C. DeCusatis 和 B. Larson, VM mobility over distance, leveraging IBM SVC split cluster with VMWare/VSphere, Proc. IBM 存储器边缘会议, 佛罗里达州奥兰多市 (2012 年 6 月 4 日至 7 日)

声明

本信息是针对在美国提供的产品和服务编写的。

IBM 可能在其他国家或地区不提供本文中讨论的产品、服务或功能特性。有关您当前所在区域的产品和服务的信息，请向您当地的 IBM 代表咨询。任何对 IBM 产品、程序或服务的引用并非意在明示或暗示只能使用 IBM 的产品、程序或服务。只要不侵犯 IBM 的知识产权，任何同等功能的产品、程序或服务，都可以代替 IBM 产品、程序或服务。但是，评估和验证任何非 IBM 产品、程序或服务，则由用户自行负责。

IBM 公司可能已拥有或正在申请与本文档内容有关的所有专利。提供本文档并未授予用户使用这些专利的任何许可。您可以用书面方式将许可查询寄往：*IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785 U.S.A.*

本条款不适用英国或任何这样的条款与当地法律不一致的国家或地区：
INTERNATIONAL BUSINESS MACHINES CORPORATION “按现状”提供本出版物，不附有任何种类的（无论是明示的还是暗示的）保证，包括但不限于暗示的有关非侵权、适销和适用于某种特定用途的保证。某些国家或地区在某些交易中不允许免除明示或暗示的保证。因此本条款可能不适用于您。

本信息中可能包含技术方面不够准确的地方或印刷错误。此处的信息将定期更改；这些更改将编入本资料的新版本中。IBM 可以随时对本资料中描述的产品和 / 或程序进行改进和 / 或更改，而不另行通知。

本信息中对非 IBM Web 站点的任何引用都只是为了方便起见才提供的，不以任何方式充当对那些 Web 站点的保证。那些 Web 站点中的资料不是 IBM 产品资料的一部分，使用那些 Web 站点带来的风险将由您自行承担。

IBM 可以按它认为适当的任何方式使用或分发您所提供的任何信息而无需对您承担任何责任。

涉及非 IBM 产品的信息可从这些产品的供应商、其出版说明或其他可公开获得的资料中获取。IBM 没有对这些产品进行测试，也无法确认其性能的精确性、兼容性或任何其他关于非 IBM 产品的声明。有关非 IBM 产品性能的问题应当向这些产品的供应商提出。

本信息包含在日常业务操作中使用的数据和报告的示例。为了尽可能完整地说明这些示例，示例中可能会包括个人、公司、品牌和产品的名称。所有这些名字都是虚构的，若现实生活中实际业务企业使用的名字和地址与此相似，纯属巧合。

此处包含的任何性能数据都是在受控环境中测得的。因此，在其他操作环境中获得的数据可能会有明显的不同。有些测量可能是在开发级的系统上进行的，因此不保证与一般可用系统上进行的测量结果相同。此外，有些测量是通过推算而估计的，实际结果可能会有所不同。本文档的用户应当验证其特定环境的适用数据。

版权许可：

本信息包括源语言形式的样本应用程序，这些样本说明不同操作平台上的编程方法。如果是为按照在编写样本程序的操作平台上的应用程序编程接口（API）进行应用程序的开发、使用、经销或分发为目的，您可以任何形式对这些样本程序进行复制、修改、分发，而无须向 IBM 付费。这些示例并未在所有条件下作全面测试。因此，IBM 不能担保或暗示这些程序的可靠性、可维护性或功能。

本文档“REDP-4933-00”创建或更新于 2013 年 4 月 03 日。

IBM®



商标

IBM、IBM 徽标和 ibm.com 是 International Business Machines Corporation 在美国和 / 或其他国家或地区的商标或注册商标。这些术语和其他 IBM 已注册商标的术语在本信息中首次出现时都使用适当的符号（或）加以标记，以表示在本信息发布时由 IBM 在美国注册或拥有的普通法商标。这些商标也可能是在其他国家或地区的注册商标或普通法商标。在 Web 地址 ibm.com/legal/copytrade 中包含了 IBM 商标的最新列表。

以下术语是 International Business Machines Corporation 在美国和 / 或其他国家或地区的商标：

IBM®
POWER®
PureApplication™
PureFlex™
PureSystems™
RackSwitch™
Redbooks®
Redbooks® (徽标) ®
System z®

以下术语是其他公司的商标：

其他公司、产品或服务名称可能是其他公司的商标或服务标记。