

# 在基于瞻博网络基础设施的数据中心接入层实施第二层局域网设计

尽管瞻博网络力图在本指南中提供准确的信息，但瞻博网络不对本指南所提供信息的准确性做出任何担保或保证。本文中提供的第三方产品描述和相关技术细节仅供参考，瞻博网络不负责为这些产品提供支持。本指南中所含信息均以“概不保证”的方式提供，瞻博网络对于其中的所有错误不做任何类型的保证，无论是明示、暗示还是法定的保证。瞻博网络及其供应商特此声明，不提供与本指南及其所含信息相关的任何明示、暗示或法定的保证，包括但不限于适销性、适用于特定用途和不侵权性，以及由交易过程、使用情况或贸易惯例而引发的事项。

## 目录

引言 .....	4
内容范围 .....	4
目标受众 .....	5
设计思路 .....	5
实施和配置指南 .....	6
通过第二层连接接入层与核心层 .....	6
第1步——通过10 Gbps上行链路连接接入层与核心层 .....	6
第2步——使用来自不同集群交换机成员的多个上行链路启用LAG .....	8
第3步——在LAG上启用dot1q中继 .....	8
第4步——在LAG上行链路上启用RTG .....	9
在核心设备上配置第二层桥接 .....	10
为了执行桥接域设置，通过MX系列以太网业务路由器上的IRB和多个物理接口连接相同的VLAN .....	10
在MX系列上配置LAG端口 .....	11
使用IRB接口配置虚拟路由器路由例程 .....	12
支持RTG所需要的配置（启用MX系列路由器之间的LAG端口上的RSTP） .....	12
在EX8200以太网交换机上配置VLAN和路由虚拟接口（RVI） .....	13
在EX8200交换机上配置LAG端口 .....	13
使用RVI接口配置虚拟路由器路由例程 .....	14
支持RTG所需要的配置（在EX8200以太网交换机之间的LAG端口上启用RSTP） .....	14
在核心层MX系列上配置第三层 .....	15
第1步——介绍IRB和VRF的配置过程 .....	15
第2步——配置MX系列之间的VRRP .....	15
在核心EX系列交换机上配置第三层 .....	16
故障切换情况概述 .....	16
接入交换机/集群交换机发生故障 .....	17
从服务器到客户端的北向流量 .....	17
从客户端到服务器的南向流量 .....	17
从接入交换机/集群交换机到MX/EX系列的活动链路发生故障 .....	18
从服务器到客户端的北向流量 .....	18
从客户端到服务器的南向流量 .....	18
活动MX系列路由器/EX8200交换设备发生故障 .....	19
从服务器到客户端的北向流量 .....	19
从客户端到服务器的南向流量 .....	19
最佳运行实践 .....	19
构建新VLAN（端到端） .....	19
针对9.2版本之后的Junos代码： .....	21
升级两个层 .....	22
总结 .....	24
关于瞻博网络 .....	24

## 图片目录

图1: EX4200位于接入层和MX960 (或EX8200) 位于核心层的数据中心.....	6
图2: 核心层MX系列路由器上的IRB和桥接域的逻辑布局 .....	11
图3: 正常状态下的通信流量 (北向——从服务器到客户端) .....	16
图4: 服务器的直连式接入交换机/集群交换机发生故障时的流量情况 .....	17
图5: 接入交换机/集群交换机与EX8200交换机或MX系列路由器之间的RTG活动链路发生故障时的流量情况.....	18
图6: 活动MX系列或EX8200交换机发生故障时的流量情况 .....	19

## 引言

本实施指南介绍了如何通过接入层上使用支持集群交换技术的瞻博网络 (Juniper Networks®) EX4200以太网交换机，在单一数据中心内实施两层局域网设计；以及如何将接入层互连到基于瞻博网络MX系列以太网业务路由器或EX8200以太网交换机的数据中心核心网络。依据这个两层设计建立的网络不仅可以完全容错，而且能够支持多种适用于数据中心的服务器技术。在传统设计中，网络超额配置 (oversubscription) 和性能由设备的特征决定；瞻博网络第二层数据中心网络采用了不同的设计原则，可通过互连接入网络层和核心网络的链路数量来控制流量超额配置。在这种设计中将不存在设备超额配置限制，因为瞻博网络系统使用的是完全无阻塞的架构。

依据《Enterprise Data Center Network Reference Architecture (企业数据中心网络参考架构)》所述，结合瞻博网络总体数据中心解决方案，我们发现将数据中心从传统的三层网络设计 (接入交换层连接至汇聚交换层，汇聚交换层又连接至核心交换和路由设备) 压缩成两层设计 (取消了汇聚层) 可以带来极大的优势。通过在接入层上使用瞻博网络集群交换技术，再结合高密度10 Gbps平台 (例如MX系列路由器和EX8200交换机)，从技术层面讲已经不再需要汇聚层。通过减少接入层上行链路的数量，可以在核心层上出色地部署服务，而且不再需要汇聚层来支持众多的物理连接。

除了考虑数据中心网络将由多少个层构成之外，网络架构师还必须考虑在何处实施开放系统互连 (OSI) 模型第三层路由，以及相对地在何处扩展OSI模型第二层广播域。在尽量靠近服务器 (接入交换层) 的网络基础设施中实施第三层路由的优势通常包括可以隔离故障、更准确地预测网络流量路径。扩展第二层广播域大小的优势通常是可以尽量将其扩展到最大，或者将第三层终端推送至汇聚层或核心层。该技术使您可以灵活地在网络中定位服务器，并且最大程度地降低配置开销。

在本文中，我们针对只有两个网络层 (接入层和核心层) 的数据中心局域网提供了具体的实施指南，其中网络基础设施的IP寻址功能由核心网络设备提供。因此，连接到接入交换机的服务器默认网关位于核心以太网业务路由器中。在此设计中，第二层广播域可以跨越多个接入元件 (单个物理交换机或作为单个逻辑交换机的集群交换机) 实施多节点服务器集群技术，但需要在这些集群中的节点之间进行第二层连接。这些技术的实例包括VMware实时迁移技术或“vmotion”、Microsoft主动式集群和其他高性能计算 (HPC) 集群或网格计算应用等等。

接入交换机经过配置，可以在上行链路层面提供完全容错功能。因此，每个接入交换元件至少有两个链路连接到核心层。每个链路的终端都位于不同的核心设备中。在此设计中，核心设备故障将不会影响到接入元件的服务器的可用性。此外，我们使用从接入交换元件到每个核心设备的连接，通过作为链路汇聚组 (LAG) 的多个物理链路，提供链路级冗余。无需多言，每个接入元件还支持多个VLAN的本地托管，并将上行链路中的所有VLAN中继到核心网络层。

## 内容范围

本文介绍了一种最佳实践方法，展示了如何才能可靠地构建数据中心LAN。本实施指南描述了一个包括多个独立功能区域的数据中心网络，这些功能区域有的可以进行物理分割，有的不行。因此，在我们的使用案例中，同一机柜中的服务器分别归属于不同的数据中心功能区域 (或应用层)。此外，本文还介绍了在实施数据中心LAN时需要考虑的第二层 (接入层) 和第三层 (核心层) 问题，并且提供了关于接入层与核心层连接性的指导。

注：以下主题不在本文的讨论范围中，我们会在数据中心系列的单独实施指南中对它们加以讨论。

- 数据中心核心网络扩展到多个地点时的设计和实施考虑因素
- 构建数据中心接入网络
- 接入网络至核心网络连接的实施方案

本文首先介绍了与接入和核心数据中心网络的物理和逻辑连接性有关的数据中心设计假设，以及连接的外部元件（例如服务器）的连接性要求。然后，我们将提供实现成功的设计所需遵循的指导和配置指导，以及与设备相关的一些故障情况，特别是详细描述了故障的原因以及这些故障对服务器和客户端之间双向通信流量产生的影响。最后，我们通过介绍设备配置以及配置/升级过程的步骤，提供了使用MX系列路由器、EX8200交换机和EX4200交换机运行数据中心网络的最佳实践。

## 目标受众

本指南主要针对以下读者：

- 评估网络设计新方法的可行性的数据中心网络架构师
- 设计和实施新数据中心网络的数据中心网络工程师和操作人员
- 研究和分析新方法以实施灵活强大的网络的技术人员

## 设计思路

以下项目符号列表描述了总体的数据中心网络环境，我们将在其中实施跨越多个接入元件的第二层连接。该列表首先描述了服务器级的连接性要求，接着介绍了到达网络核心所经过的网元。另外，该列表还介绍了根据本指南实施的网元所具备的一些功能。本节的主要目标是阐明开发这些实施指南所使用的网络环境。此实施也可能适用于其他假设条件，但在某些极端的情况下可能不适用。

- 服务器使用一个IP地址（无VLAN标签）通过单个网络接口连接至每个接入集群交换机。因此，每个服务器端口都将与一个第二层广播域（VLAN）关联。鉴于此方案的采用，本文将不再赘述服务器级的网络接口卡（NIC）组合情况。

**注：**每个服务器可以连接两个独立的集群交换机，以实现高可用性（HA）接入。

- 每个接入元件上可配置多个VLAN。这些VLAN将供不同的服务器使用，并且都是从接入元件中继到核心网络层。
- 每个集群交换机使用LAG与一个MX系列路由器连接。尽管LAG由两个链路成员组成，但并非只能有两个链路，每个LAG最多可以有8个物理链路成员。
- 每个集群交换机连接至两个MX系列路由器或EX8200交换机。这些连接（LAG如前所述）是连接至MX系列接口的相同中继器，并且与使用相同VLAN ID的路由器上的相同桥接域关联。
- 核心网络包括两台MX系列路由器或EX8200交换机，以实现冗余目标。这些冗余路由器和交换机与相同的网络和桥接域相关。
- 核心网络层包括一个与接入层上的每个VLAN相对应的桥接域。
- 核心网络连接至网络服务层（防火墙、负载均衡器、高速缓存系统和应用加速设备），并使用路由或策略来管理不同服务器网络之间的所有连接。
- 两台核心网络MX/EX8200设备使用纯第二层以太网802.1q中继或虚拟专用LAN服务标签交换路径（VPLS LSP）标签，通过所有桥接域彼此相连，使不同的网络保持分段、接入层VLAN保持互连。本文中不包括将MX系列路由器或EX8200交换机进行互连的实施细节。
- 通过在与桥接域相关的MX系列路由器的集成路由和桥接（IRB）接口上运行虚拟路由冗余协议（VRRP），或在与桥接域相关的EX8200交换机的路由虚拟接口（RVI）上运行虚拟路由冗余协议（VRRP），这两台核心网络设备经过配置后可以向服务器提供具有高可用性的默认网关IP地址。

注：下图将作为本文中不同配置代码实例的参考，以及网络接口数和设备名称的参考。注意，下图所显示的瞻博网络MX960以太网业务路由器也可换成瞻博网络EX8200以太网交换机。MX960和EX8200的具体配置将在本文中给出。

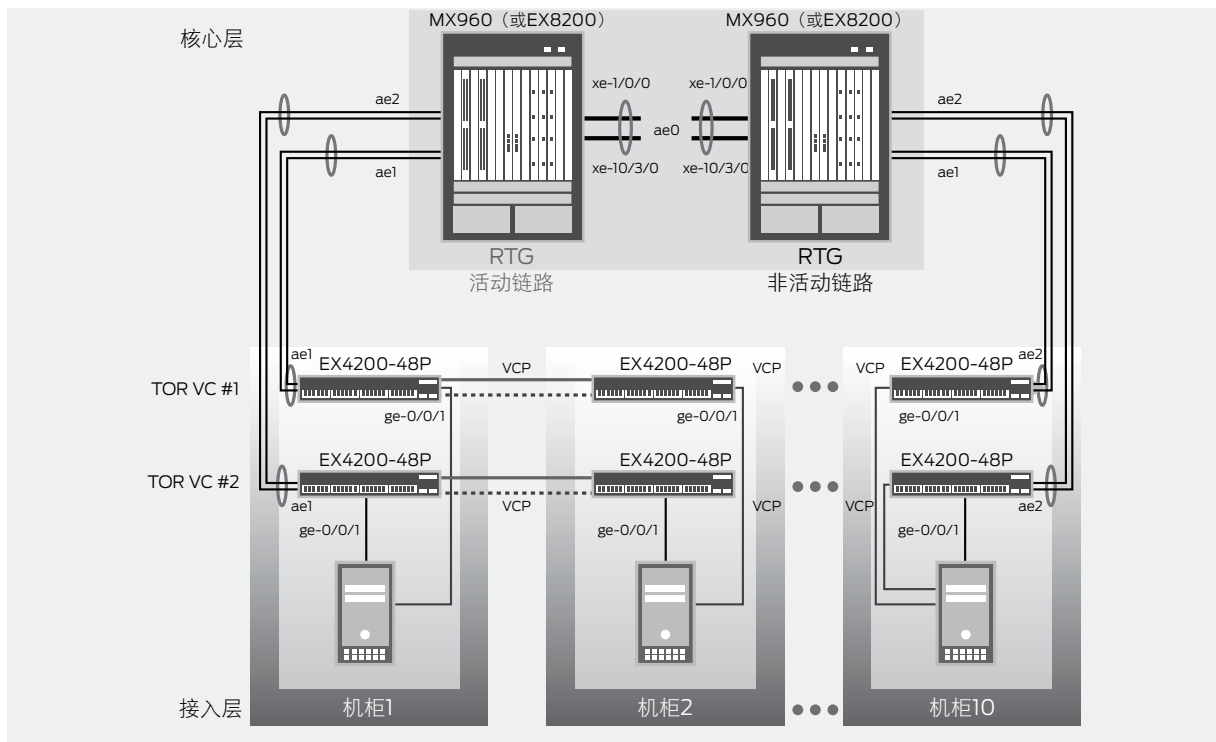


图1：EX4200位于接入层和MX960 (或EX8200) 位于核心层的数据中心

## 实施和配置指南

本节通过介绍所有相关网络组件或基础设施的设备配置，提供了解决方案实施指南。此外，还向读者展示了如何验证解决方案的运行情况。

本节中描述的主要实施和配置指南包括：

- 通过第二层连接接入层与核心层
- 通过第二层配置核心层
- 在核心层的MX系列以太网业务路由器或EX8200以太网交换机上配置第三层

### 通过第二层连接接入层与核心层

以下列表概括介绍了通过第二层连接接入层与核心层的主要步骤：

- 通过10 Gbps上行链路连接接入层与核心层
- 启用绑定了多个上行链路 (来自不同的集群交换机成员) 的LAG
- 启用LAG上的dot1q中继
- 启用LAG上行链路上的冗余中继组 (RTG)

#### 第1步——通过10 Gbps上行链路连接接入层与核心层

在本节中，我们描述了通过使用10 Gbps上行链路，将数据中心接入层的EX4200交换机与核心层的MX960路由器或EX8200交换机进行互连的最佳实践方法。如“设计思路”一节中所述，我们建议在每个机柜中配置两台EX4200集群交换机作为柜顶式交换机，其中每台集群交换机都连接至运行MX系列路由器或EX8200交换机的数据中心核心层。

我们建议您在每个EX4200集群交换机的第一台和最后一台成员交换机上建立冗余的10 Gbps上行链路电缆连接，如图1所示。瞻博网络之所以采用这种方法，是因为第一台和最后一台成员交换机不能配置为集群交换的主交换机和备用交换机。这种方法可以避免上行链路和集群交换机路由引擎同时进行故障切换。在本例中，EX4200集群交换机设计也可称为“编织环 (braided ring)”连接（连接所有其他的交换机）。通过将成员交换机之间的连接电缆长度缩短到3米以内，可以最大限度地降低集群交换机间延迟和上行链路延迟。

注：如本文“内容范围”一节所述，构建数据中心接入网络的集群交换机连接选择不在本文的讨论范围内。

在这种情况下，服务器通常通过两个NIC连接到接入层的两台EX4200集群交换机。如图1所示，每个接口卡都连接到不同的柜顶式集群交换机。有关服务器配置、代码段和相关屏幕快照的更多信息，请参阅《Implementing VMware Server Virtualization on Juniper Networks Infrastructure》（基于瞻博网络基础设施实施VMware服务器虚拟化）。

冗余布局是采用这种部署的关键原因，其中每个集群交换机都彼此备份。接入层EX4200集群交换机的默认工作方式是在任何时间都使用两个上行链路连接中的一个，从而根据LAG和RTG配置经由10 Gbps链路分发流量。

针对EX4200接入端口的瞻博网络Junos®软件实施，要求VLAN和接口一对一地进行关联，而EX4200上的中继端口只需要定义接口下的一系列VLAN成员（参见第3步）。以下列举了连接到服务器的接入端口的配置实例：

EX4200交换机上有三种不同的端口：接入端口、中继端口和路由端口。

- 接入端口是只隶属于一个VLAN的第二层端口。这是EX系列交换机（EX4200和EX8200）的默认端口模式。
- 中继端口是隶属于多个VLAN的第二层端口。第3步中对中继端口进行了详细描述。
- 路由端口是带有IP地址的第三层端口，负责在不同的IP网络之间传输流量。

EX4200交换机可以通过两种方法对VLAN成员关系进行配置——在端口下配置或在VLAN下配置（如方法1和方法2所示）。在这两个实例中，接口是一个接入端口（例如连接至接口的服务器），而且只需要是一个VLAN的一部分即可。两种方法没有优劣之分。

方法1：在端口下配置VLAN成员关系

```
{master}[edit]
root@EX4200# set vlans vlan71 vlan-id 71 interface ge-0/0/1.0 ## define vlan71 and assign
interface ##
root@EX4200# set interfaces ge-0/0/1 unit 0 family ethernet-switching port-mode access ##
Optional, to explicitly configure the port as an access port##
root@EX4200# run show vlans vlan71 ## verify interface ge-0/0/1.0 is part of vlan71##
```

方法2：在VLAN下配置VLAN成员关系

```
{master}[edit]
root@EX4200# set interfaces ge-0/0/1 unit 0 family ethernet-switching port-mode access ##
Optional, to explicitly configure the interface as an access port##
root@EX4200# set vlans vlan71 vlan-id 71 interface ge-0/0/1.0 ## define vlan71 and assign
interface ##
root@EX4200# run show vlans vlan71 ## verify interface ge-0/0/1.0 is part of vlan71##
```

技巧：瞻博网络建议在VLAN下配置接入端口的VLAN成员关系，以便于管理。接入端口的VLAN成员关系配置将会集中到一个位置。

## 第2步——使用来自不同集群交换机成员的多个上行链路启用LAG

LAG可以增加带宽并提供链路冗余。在本例中，您可以在EX4200交换机上创建最多64个汇聚接口，并将最多8个物理链路分成一组分配给每个LAG接口——称为聚合以太网 (AE) 接口。您可以在MX/EX系列平台上分别创建最多128/255个汇聚接口。

由于哈希计算是在基于第二层/第三层包头的虚拟捆绑中的成员链路上完成，因而链路汇聚还可提供负载均衡的链路利用。如IEEE 802.3ad标准中的定义，AE接口可以静态地进行创建或使用链路聚合控制协议 (LACP)。

AE虚拟捆绑接口中的物理端口不必相邻，而是可以位于集群交换机的不同成员交换机上。然而为了正确地构成虚拟捆绑，AE接口中的LAG成员端口必须是相同的物理类型、相同的速度和双工。

AE接口必须在接入层EX4200交换机和核心层MX系列路由器或EX8200交换机上进行相应的配置。LAG的Junos软件实施提供了基本的检错功能，以防止错误配置。

下面是在EX4200接入交换机上使用LACP启用LAG的配置实例：

```
{master}[edit]
root@EX4200# set chassis aggregated-devices ethernet device-count 64 ## set the total
number of AE ports ##
root@EX4200# set interfaces ae1 aggregated-ether-options minimum-links 1 ## define min. link
to keep AE port up ##
root@EX4200# set interfaces ae1 aggregated-ether-options link-speed 10g ## define bandwidth
of LAG member port ##
root@EX4200# set interfaces ae1 aggregated-ether-options lACP active ## activate LACP
protocol on LAG ports ##
root@EX4200# set interfaces xe-0/1/0 ether-options 802.3ad ae1 ## join the LAG on physical
interface ##
root@EX4200# set interfaces xe-1/1/0 ether-options 802.3ad ae1 ## join the LAG on physical
interface ##
root@EX4200# run show interface terse | match ae ## verify that AE ports are defined ##
root@EX4200# run show lACP interfaces ## verify LAG contains correct physical ports when AE
interface is up ##
```

## 第3步——在LAG上启用dot1q中继

在EX4200交换机上，分配给VLAN的端口可配置为接入端口或中继端口。为了在流量通过中继端口时确定该流量所属的VLAN，以太网VLAN上的所有帧都可以根据IEEE 802.1Q标准中定义的标签加以识别。这些帧使用802.1Q标签（有时简称为dot1q标签）进行标记和封装。dot1q中继接口通常用于交换机之间的互连。中继端口使用标签通过多路复用技术将许多VLAN中的流量汇集到单一链路（逻辑或物理）中进行传输。EX4200交换机最多支持4096个VLAN。Vlan-id号0和4095被Junos所保留。

与只支持一个VLAN的接入端口（如第1步中所述）不同，EX4200交换机上的dot1q中继端口可支持多个VLAN。中继端口的VLAN成员关系配置与第1步中的描述相同。您可以在VLAN或接口上配置VLAN成员关系。瞻博网络建议在接口下配置中继端口的VLAN成员关系。当配置EX4200交换机上的dot1q中继端口时，由于dot1q中继端口实际上允许一个接口支持4,000多个VLAN，所以这种方法可以简化中继端口VLAN成员关系的管理工作，而且需要的命令行接口 (CLI) 命令行最少。从另一方面讲，在VLAN下配置VLAN成员关系，能够将该配置扩展到多个VLAN上。

下面是在LAG接口上启用dot1q中继的配置实例：

```
{master}[edit]
root@EX4200# set interfaces ae1 unit 0 family ethernet-switching port-mode trunk ## set
interface mode to trunk ##
root@EX4200# set interfaces ae1 unit 0 family ethernet-switching vlan members [ 71-73 573 ]
## define allowed vlans ##
root@EX4200# run show vlans ## verify VLANs membership on the trunk interface ae1.0 ##
```

#### 第4步——启用LAG上行链路上的RTG

冗余中继组是EX4200交换机支持的二层链路故障切换机制，最适用于在具有双宿连接（dual home connection）的接入交换机或集群交换机上实施，其中一个链路是活动链路并负责转发流量，而另一个链路则负责阻挡流量并作为活动链路的备用链路。如果RTG先前已经启用，这一特性便消除了配置生成树协议（STP）的麻烦。

在由核心层和接入层构成的典型企业网络中，当中继端口出现故障时，冗余的中继链路可以轻松地使网络恢复正常。通过冗余链路，流量可以路由到其他中继端口，最大限度缩短网络融合时间。您可以在独立的交换机或集群交换机上配置最多16个冗余中继组。

在接入层，RTG可以替代STP。在指定端口上，RTG和STP是相互排斥的。因为EX4200交换机默认启用了快速生成树协议（RSTP）来生成无环拓扑，所以冗余组中的中继链路不能成为STP拓扑的一部分。否则交换机将会显示以下错误信息，从而导致配置变更无法生效。

```
error: XSTP : msti 0 STP and RTG cannot be enabled on the same interface ae1
error: configuration check-out failed
```

虽然用于EX4200交换机的Junos软件不允许一个接口同时归属于冗余中继组和STP拓扑，但是STP可以继续在网络的其他部分运行。在核心层，您可以启用STP以最大程度地减少人为错误，无论是否在接入层进行了RTG配置。在接入层，交换机的某些上行链路针对RTG进行了配置，其他上行链路端口（如果有）及所有下行链路端口仍然可以针对STP进行启用，以防网络中出现任何意外的第二层环路。

下面是在LAG上行链路接口禁用STP的配置实例：

```
{master}[edit]
root@EX4200# set protocols rstp interface ae1 disable ## specifically disable STP on
interface defined in RTG ##
root@EX4200# set protocols rstp interface ae2 disable ## specifically disable STP on
interface defined in RTG ##
root@EX4200# run show spanning-tree interface ## verify STP state on AE1 and AE2 interfaces
is DIS (disabled) ##
```

冗余中继组配置还要求活动和非活动的链路中继相同的VLAN成员。如果RTG组中的活动和非活动链路发生VLAN误配，交换机将会显示以下错误信息，阻止RTG配置生效：

```
error: RTG : grp DC_RTG primary ae1.0 secondary ae2.0 vlan mismatch vlan573
error: configuration check-out failed
```

要想配置冗余中继链路，首先需要创建RTG。您可以在包含两个链路——主（活动）链路和辅助链路——的接入交换机上配置RTG。如果活动链路出现故障，辅助链路将自动开始转发数据流，无需等待常规的STP聚合。您无需在接入层和核心层同时启用RTG。

数据流只在活动链路上转发。当您发出工作模式命令“show interfaces interface-name extensive”时，辅助链路上的数据流量会被丢弃并显示为丢弃的数据包。虽然辅助链路会阻止数据流量，但仍然允许第二层控制流量通过。例如，链路层发现协议（LLDP）会话可以在辅助链路上的两个EX4200交换机之间运行。

下面是在LAG上行链路接口上启用RTG的配置实例：

```
{master}[edit]
root@EX4200# set ethernet-switching-options redundant-trunk-group group DC_RTG interface
ae1 ## define RTG ##
root@EX4200# set ethernet-switching-options redundant-trunk-group group DC_RTG interface
ae2
root@EX4200# run show redundant-trunk-group ## verify RTG interfaces are correctly defined
##
```

### 在核心设备上配置第二层桥接

以下列表概括介绍了通过第二层配置核心层的主要步骤：

- 通过IRB/RVI和多个物理接口连接相同的VLAN
- 在MX系列路由器或EX8200交换机上配置LAG端口
- 使用IRB/RVI接口配置虚拟路由器路由例程
- 支持RTG (MX系列路由器或EX8200交换机之间的LAG端口上的RSTP和dot1q中继)

为了执行桥接域设置，须通过MX系列以太网业务路由器上的IRB和多个物理接口连接相同的VLAN

桥接域（仅限于MX系列路由器）是一个包括一组逻辑接口的域，这些接口共享相同的泛洪或广播特性，可以执行第二层桥接。（要使物理接口设备能够在Junos中工作，您必须在该设备上配置至少一个逻辑接口。针对每个逻辑接口，您必须指定接口支持的协议系列。）一个桥接域必须包括一组参与第二层学习（learning）和转发的逻辑接口。默认状态下，每个桥接域都维护着一个第二层转发数据库，其中包含从该桥接域的端口接收到的数据包中获知的媒体访问控制（MAC）地址。

桥接在OSI参考模型的第二层执行，路由在第三层执行。可以说，针对桥接配置的一组逻辑端口构成了桥接域。

集成的路由和桥接是针对特定的桥接域配置的路由接口，可在同一接口上同时支持第二层桥接和第三层路由。IRB使您能够将数据包路由到另一个路由接口，或路由到配置了IRB接口的另一个桥接域。IRB可执行以下功能：

- 如果目的地MAC地址是路由器的MAC地址，而且数据包以太类型是IPv4、IPv6或MPLS，则路由一个数据包
- 在第二层的桥接域中交换所有组播和广播数据包
- 如果目的地MAC地址是组播MAC地址，而且数据包以太类型是IPv4或IPv6，则路由一个数据包的副本
- 在第二层上交换所有其他的单播数据包
- 处理所支持的第二层控制数据包，例如STP和LACP
- 处理所支持的第三层控制数据包，例如OSPF和RIP

下图是核心层MX系列路由器的IRB和桥接域逻辑布局。

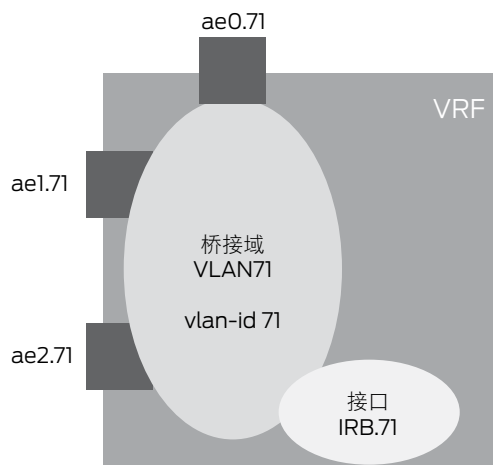


图2：核心层MX系列路由器的IRB和桥接域逻辑布局

通过配置路由例程并将例程类型指定为桥接，可创建一个桥接域。下面是针对核心层MX系列路由器多个物理接口上的相同VLAN设置桥接域的配置实例：

```
{master}[edit]
root@MX960# set bridge-domains VLAN71 domain-type bridge ## define VLAN71 as a bridge-domain
on MX Series ##
root@MX960# set bridge-domains VLAN71 vlan-id 71 ## assign vlan-id 71 to bridge-domain
VLAN71 ##
root@MX960# set bridge-domains VLAN71 interface ae0.71 ## assign interface to bridge-domain
VLAN71 ##
root@MX960# set bridge-domains VLAN71 interface ae1.71 ## assign interface to bridge-domain
VLAN71 ##
root@MX960# set bridge-domains VLAN71 interface ae2.71 ## assign interface to bridge-domain
VLAN71 ##
root@MX960# run show bridge domain interface ae0.71 ## verify VLAN71 is allowed on correct
interfaces ##
```

### 在MX系列上配置LAG端口

核心层MX系列路由器的LAG配置与EX4200接入层的配置完全相同。以下配置显示了如何在核心层MX系列路由器上配置LAG端口。

```
{master}[edit]
root@MX960# set chassis aggregated-devices ethernet device-count 128 ## define total number
of LAG ports on MX ##
root@MX960# set interfaces ae0 aggregated-ether-options minimum-links 1 ## set min link to
keep AE interface up ##
root@MX960# set interfaces ae0 aggregated-ether-options link-speed 10g ## define bandwidth
of LAG member ports ##
root@MX960# set interfaces ae0 aggregated-ether-options lacp active ## enable LACP protocol
on LAG interface ##
root@MX960# set interfaces xe-1/0/0 gigether-options 802.3ad ae0 ## join the LAG port on
physical interface ##
root@MX960# set interfaces xe-10/3/0 gigether-options 802.3ad ae0 ## join the LAG port on
physical interface ##
root@MX960# run show lacp interfaces ## verify LAG has the correct physical ports when AE
interface is up ##
```

下面是在核心层MX系列路由器之间的LAG接口上启用dot1q中继的配置实例。

针对9.2版本之前的Junos代码：

```
{master}[edit]
root@MX960# set interfaces ae0 flexible-vlan-tagging ## set flexible vlan tagging on LAG port
##
root@MX960# set interfaces ae0 encapsulation flexible-ethernet-services ## set encapsulation
on LAG port ##
root@MX960# set interfaces ae0 unit 71 encapsulation vlan-bridge ## set encapsulation for
sub-interface ##
root@MX960# set interfaces ae0 unit 71 vlan-id 71 ## allow vlan-id 71 traffic on sub-
interface ##
root@MX960# run show bridge domain interface ae0.71 ## verify correct vlans are allowed on
ae0 trunk interface ##
```

针对9.2版本之后的Junos代码：

```
{master}[edit]
root@MX960# set interface ae0 unit 0 family bridge interface-mode trunk ## set interface
mode to trunk ##
root@MX960# set interface ae0 unit 0 family bridge vlan-id-list 71-73 ## allow vlan range
on LAG port ##
root@MX960# set interface ae0 unit 0 family bridge vlan-id-list 573 ## allow a vlan on LAG
port ##
root@MX960# run show bridge domain interface ae0.0 ## verify vlans are allowed on ae0.0
trunk interface ##
```

### 使用IRB接口配置虚拟路由器路由例程

虚拟路由器路由例程与虚拟路由和转发 (VRF) 路由例程类似，为每个例程提供单独的路由表和转发表。然而，VRF路由例程要求的许多配置步骤在虚拟路由器路由例程中并无要求。特别是您不必配置路由标识、路由表策略 (vrf-export、vrf-import和route-distinguisher语句) 或电信供应商路由器之间的MPLS。

下面是使用IRB接口设置虚拟路由器路由例程的配置实例。

```
{master}[edit]
root@MX960# set routing-instances VMotionTest instance-type virtual-router ## define a
virtual router on MX ##
root@MX960# set routing-instances VMotionTest interface irb.71 ## assign interface to the
virtual router ##
root@MX960# set routing-instances VMotionTest interface irb.72 ## assign interface to the
virtual router ##
root@MX960# run show route instance VMotionTest detail ## verify virtual router is
correctly defined ##
```

### 支持RTG所需要的配置 (在MX系列路由器之间的LAG端口上启用RSTP)

传统上，预防第二层环路，同时在第二层提供设备冗余，需要采用STP。EX4200交换机和MX系列路由器支持3个标准版本的STP：802.1d STP、802.1w快速STP和802.1s多实例STP。

注：STP、RSTP和多生成树协议 (MSTP) 还可与思科的PVST+/RPVST+实现互操作。

当您在接入层EX4200交换机上启用RTG时，RSTP (MX系列上的默认生成树协议版本) 在核心层MX系列路由器之间实施，以预防人为错误导致的第二层环路。

下面是在核心层MX系列路由器之间的LAG接口上启用RSTP的配置实例。

```
{master}[edit]
root@MX960# set protocols rstp bridge-priority 4k ## set spanning-tree bridge priority to
MX ##
root@MX960# set protocols rstp interface ae0 mode point-to-point ## enable STP on interface
and set mode to p2p ##
root@MX960# run show spanning-tree bridge ## verify root bridge is the correct MX Ethernet
services router by bridge ID ##
root@MX960# run show spanning-tree interface ## verify port is in correct STP state ##
```

注：STP桥接优先级只能设置成以4096递增，范围是0到61,440。默认值为32,768。

### 在EX8200以太网交换机上配置VLAN和路由虚拟接口 (RVI)

VLAN与MX系列的桥接域概念相同；VLAN包括一组共享同一个第二层域的接口。每个VLAN将所有本地流量（单播、组播/广播）限制在各自的域内。EX8200交换机支持多达4096个VLAN，并可将其任何一个VLAN分配给接入端口或中继端口。

EX8200交换机的配置语法与EX4200交换机的相同。以下实例显示了如何在EX8200交换机上创建VLAN。

```
{master}[edit]
root@EX8200# set vlans VLAN71 vlan-id 71 ## Defines a VLAN name (VLAN71) and assigns a VLAN-
id (71) to the VLAN
```

VLAN成员关系的配置步骤与“通过第二层连接接入层与核心层”的第1步一样。瞻博网络建议在VLAN下配置接入端口的VLAN成员关系，在接口下配置中继端口的VLAN成员关系。

路由虚拟接口 (RVI) 是VLAN的逻辑第三层路由接口。与MX系列桥接域中的IRB类似，它可以加快VLAN之间的路由 (VLAN间路由)，或是VLAN到任何其他可路由网络之间的路由。EX8200交换机支持IPv4 (inet协议族) 和/或IPv6 (inet6协议族) 单播路由。以下是在EX8200交换机上配置RVI的实例。

```
{master}[edit]
root@EX8200# set interfaces vlan unit 71 family inet address 192.168.71.1/24 ## Create a
RVI interface and assigning an IP address to it##
root@EX8200# set interfaces vlan unit 72 family inet address 192.168.72.1/24 ## Create a
RVI interface and assigning an IP address to it##
root@EX8200# set vlans VLAN71 l3-interface vlan.71 ## Bind the RVI vlan.71 to the VLAN
VLAN71##
root@EX8200# set vlans VLAN72 l3-interface vlan.72 ## Bind the RVI vlan.72 to the VLAN
VLAN72##
root@EX8200# run show interfaces vlan ## Verify the RVI interface##
```

注：所有EX系列交换机的RVI配置相同。

### 在EX8200交换机上配置LAG端口

EX8200交换机上的LAG配置与EX4200类似。有关LAG的配置实例，请参见“通过第二层连接接入层与核心层”的第2步。

## 使用RVI接口配置虚拟路由器路由例程

EX8200以太网交换机的虚拟路由器例程 (VR) 与MX系列的相同。EX8200以太网交换机为每个路由器例程创建一个独立的转发表。这对于分离两个域非常有用，且无需配置防火墙过滤器。

EX8200交换机可支持多达255个VR，目前限制为IPv4物理接口 (如第三层接口、RVI、第三层子接口等)。下面是在EX8200交换机上配置VR的实例：

```

{master}[edit]
root@EX8200# set routing-instances VMotionTest instance-type virtual-router ## define a
virtual router on EX8200 ##
root@EX8200# set routing-instances VMotionTest interface vlan.71 ## assign interface to the
virtual router ##
root@EX8200# set routing-instances VMotionTest interface vlan.72 ## assign interface to the
virtual router ##
root@EX8200# run show route instance VMotionTest detail ## verify virtual router is
correctly defined ##

```

## 支持RTG所需要的配置 (在EX8200以太网交换机之间的LAG端口上启用RSTP)

RTG是交换机本地的一个特性，而不是一种协议；任何设备之间都不存在协议交换。因此任何设备——除配置RTG所在的设备以外——都没有RTG知识。RTG可以通过减少交换机上的协议数量来简化网络管理。

EX8200交换机默认支持快速生成树协议 (RSTP)。与在接入交换机上启用RTG一样，我们强烈建议您在核心交换机上启用生成树协议。未与启用了RTG的接入交换机连接的二层端口仍然容易形成二层环路，而生成树协议有助于预防该问题。当配置EX8200交换机上与接入交换机相连的端口时，有两个不同的选择。

### 方法1：在连接到接入交换机 (启用了RTG) 的端口上启用STP

鉴于EX8200的端口启用了生成树，EX4200将放弃并且不再处理来自EX8200的BPDU，因为它启用了RTG。因此除了消耗极少的带宽以外，EX8200发送的BPDU不会带来任何害处。建议在连接到接入交换机 (启用了RTG) 的EX8200交换机上启用边缘端口。这有助于缩短故障恢复聚合时间。

```

{master}[edit]
root@EX8200# set protocols rstp interface ae1.0 edge ## classify the spanning-tree port
role as edge port ##
root@EX8200# set protocols rstp interface ae2.0 edge ## classify the spanning-tree port
role as edge port ##
root@EX8200# run show spanning-tree interface ae1.0 detail ## verify the port is in the
correct STP states and role ##
root@EX8200# run show spanning-tree interface ae2.0 detail ## verify the port is in the
correct STP state and role##

```

### 方法2：在连接到接入交换机 (启用了RTG) 的端口上禁用STP

由于默认支持STP，所以需要进行的唯一配置是在端口上禁用生成树。

```

{master}[edit]
root@EX8200# set protocols rstp interface ae1.0 disable ## disable STP on an interface ##
root@EX8200# set protocols rstp interface ae2.0 disable ## disable STP on an interface ##
root@EX8200# run show spanning-tree interface ae1.0 ## verify STP is not enabled on the
interface ##
root@EX8200# run show spanning-tree interface ae2.0 ## verify STP is not enabled on the
interface##

```

## 在核心层MX系列上配置第三层

以下列表概括介绍了在核心层MX系列路由器上配置第三层的主要步骤：

- 回顾IRB和VRF的配置过程
- 配置MX系列之间的VRRP

### 第1步——回顾IRB和VRF的配置过程

核心层MX系列上的IRB在同一接口上同时支持第二层桥接和第三层IP路由。IRB使操作人员能够将本地数据包路由到另一个路由接口或另一个配置了第三层协议的桥接域。

EX4200交换机还支持称为路由VLAN接口 (RVI) 的路由接口。由于本实施指南只讨论接入层的第二层，因此RVI不在本文的范围之内，不予以介绍。与路由桥接域的IRB相对，RVI路由的是VLAN。

下面是在核心层MX系列路由器上启用第三层IRB接口的配置实例。

```
{master}[edit]
root@MX960# set bridge-domains VLAN71 domain-type bridge
root@MX960# set bridge-domains VLAN71 vlan-id 71
root@MX960# set bridge-domains VLAN71 routing-interface irb.71 ## define L3 IRB routing
interface in VLAN71 ##
root@MX960# set interfaces irb unit 71 family inet address 172.16.56.2/24 ## assign IP
address to the defined IRB interface ##
root@MX960# run show interface irb.71 terse ## verify IRB interface is correctly defined ##
```

### 第2步——配置MX系列之间的VRRP

MX系列和EX4200平台都支持VRRP。通过VRRP，构成冗余组的路由器共同负责转发数据包，就像它们都拥有与主机上配置的默认网关相对应的IP地址一样。在任何时候，冗余组中的VRRP路由器都是一部分作为主路由器，另一部分作为辅助（备用）路由器。如果主路由器发生故障，备用路由器将成为新的主路由器。这样可以确保始终提供路由器冗余，LAN上的流量路由不会依赖于单个路由器。

下面是在IRB接口上启用VRRP的配置实例：

在主MX系列设备上：

```
{master}[edit]
root@MX960# set interfaces irb unit 71 family inet address 172.16.56.2/24 vrrp-group 1
virtual-address 172.16.56.1
root@MX960# set interfaces irb unit 71 family inet address 172.16.56.2/24 vrrp-group 1
priority 190
root@MX960# set interfaces irb unit 71 family inet address 172.16.56.2/24 vrrp-group 1
preempt
root@MX960# set interfaces irb unit 71 family inet address 172.16.56.2/24 vrrp-group 1
accept-data
root@MX960# run show vrrp interface irb.71 ## verify VRRP is correctly defined ##
```

在辅助MX系列设备上：

```
{master}[edit]
root@MX960# set interfaces irb unit 71 family inet address 172.16.56.3/24 vrrp-group 1
virtual-address 172.16.56.1
root@MX960# set interfaces irb unit 71 family inet address 172.16.56.3/24 vrrp-group 1
priority 180
root@MX960# set interfaces irb unit 71 family inet address 172.16.56.3/24 vrrp-group 1
preempt
root@MX960# set interfaces irb unit 71 family inet address 172.16.56.3/24 vrrp-group 1
accept-data
root@MX960# run show vrrp interface irb.71 ## verify VRRP is correctly defined ##
```

### 在核心EX系列交换机上配置第三层

在前面关于EX8200交换机的一节中，本文介绍了VLAN、LAG、RTG、RVI和VRF。剩余唯一需要配置的特性是VRRP。VRRP向子网提供服务，并为其提供网关冗余。VRRP配置遵从与MX相同的CLI命令。请参见“在核心层MX系列上配置第三层”的第2步。

### 故障切换情况概述

本节介绍了可能出现的故障切换情况，特别描述了故障切换的原因以及故障切换对北向和南向通信流量的影响。

- 接入交换机/集群交换机发生故障
- 从接入交换机/集群交换机到MX系列路由器或EX8200交换机的活动链路发生故障
- 活动MX系列路由器或EX8200交换机设备发生故障

为了延长网络正常运行时间，我们在数据中心中使用冗余链路来互连接入层交换机与核心层，并连接双宿主服务器与接入交换机。为了在第二层上实现路径冗余和快速故障切换，支持802.3ad链路汇聚和RSTP (802.1w) 的交换机提供了比原来的STP更快的链路故障恢复速度。EX4200接入交换机还提供了RTG特性，作为比STP更快速、更易于实施的替代方案。以下几节我们将针对北向（服务器到客户端）和南向（客户端到服务器）通信流量，来讨论三种故障切换情况。

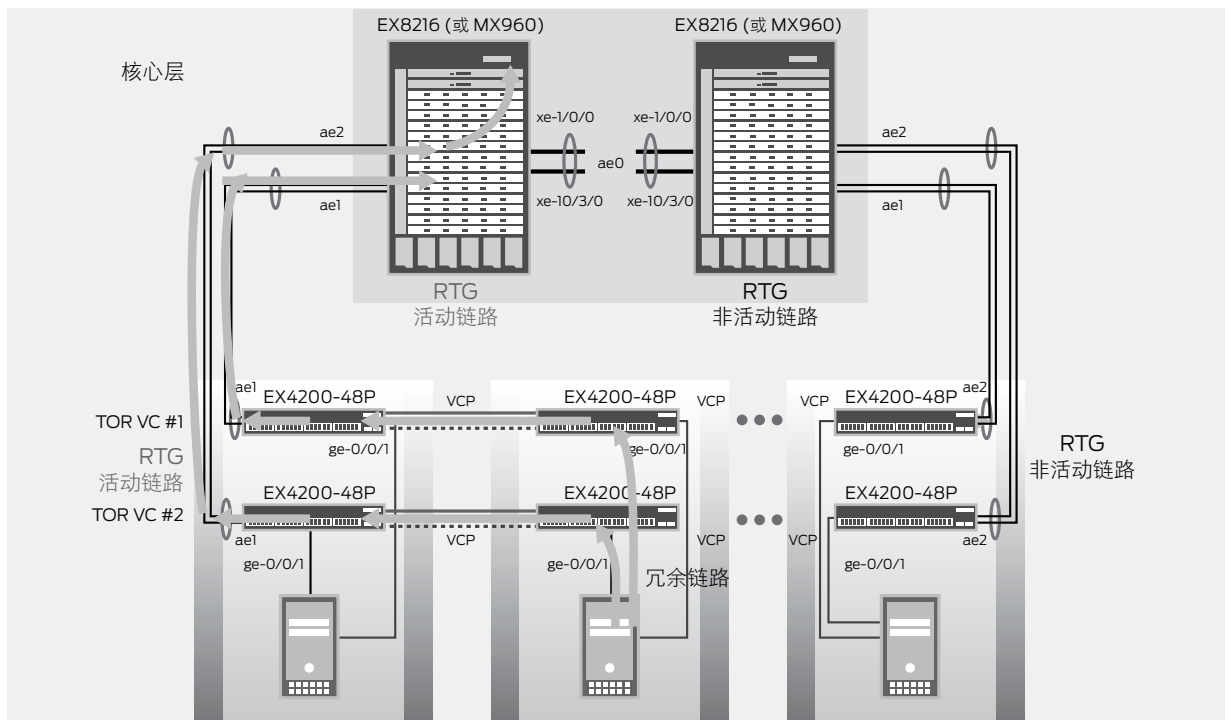


图3：正常状态下的流量（北向——从服务器到客户端）

## 接入交换机/集群交换机发生故障

### 从服务器到客户端的北向流量

按照“通过第二层连接接入层与核心层”第1步中的建议，我们使用两个冗余链路来连接服务器与接入层交换机。当两个链路都正常时，服务器的默认工作模式是在任何时候都使用一个链路，从而根据服务器端的冗余链路和端口组配置在活动链路上分发流量。

当一个直接连接的接入交换机发生故障时，从服务器到客户端的北向流量将改用连接至冗余接入集群交换机的另一个链路进行传输。服务器会在其与接入交换机之间的活动链路连接发生故障时触发这一变更。之后，流量将通过冗余集群交换机上的活动上行链路传输至核心层，然后再从数据中心网络路由到核心层MX系列路由器或EX8200交换机的客户端。如果接入集群交换机发生故障，则通信流量将按照与本例中直连式接入交换机发生故障时完全相同的路径进行传输。

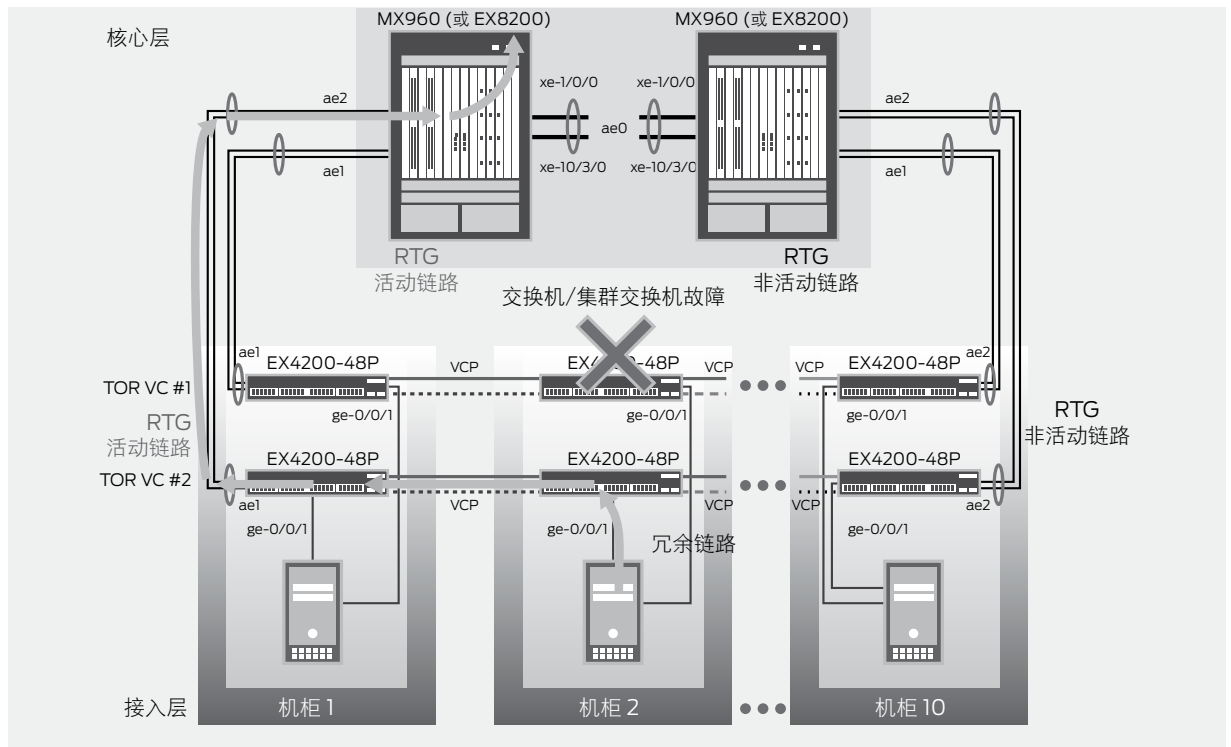


图4：服务器的直连式接入交换机/集群交换机发生故障时的流量情况

### 从客户端到服务器的南向流量

在正常运行过程中，从客户端到服务器的南向流量通过活动的RTG LAG连接传输到接入层EX4200集群交换机。当一个直接连接的接入交换机发生故障时，其他的活动交换机/集群交换机将变成活动状态，因为服务器与其他接入交换机之间的连接现在是唯一处于up/up状态的链路。从客户端到服务器的流量通过另一个活动的LAG连接传输到冗余的接入集群交换机，然后通过服务器接入链路传输至服务器。如果集群交换机发生故障，则通信流量将按照与本例中直连式接入交换机发生故障时完全相同的路径进行传输。

## 从接入交换机/集群交换机到MX/EX系列的活动链路发生故障

### 从服务器到客户端的北向流量

该故障切换情况假设将互连接入层和核心层的10 Gbps上行链路配置为LAG接口上的RTG。如果活动LAG虚拟捆绑中的物理成员链路发生故障，故障链路上的北向流量将转移到该LAG中的其他成员链路上。一个LAG接口中最多可配置8个10 Gbps上行链路。这样可以提供不间断的冗余，最大程度地减少物理链路故障导致的宕机。

在极少数情况下，活动LAG接口发生故障（由该LAG连接中的所有物理成员链路都发生故障，或链路故障超出定义的LAG阈值引起），则该接入层EX4200集群交换机与核心层MX系列路由器或EX8200交换机之间的非活动LAG连接将变成活动状态。这一变化由接入交换机/集群机上的RTG设置触发，其聚合时间比STP的更短。此时，流量将开始通过这个辅助LAG链路传输至核心层。

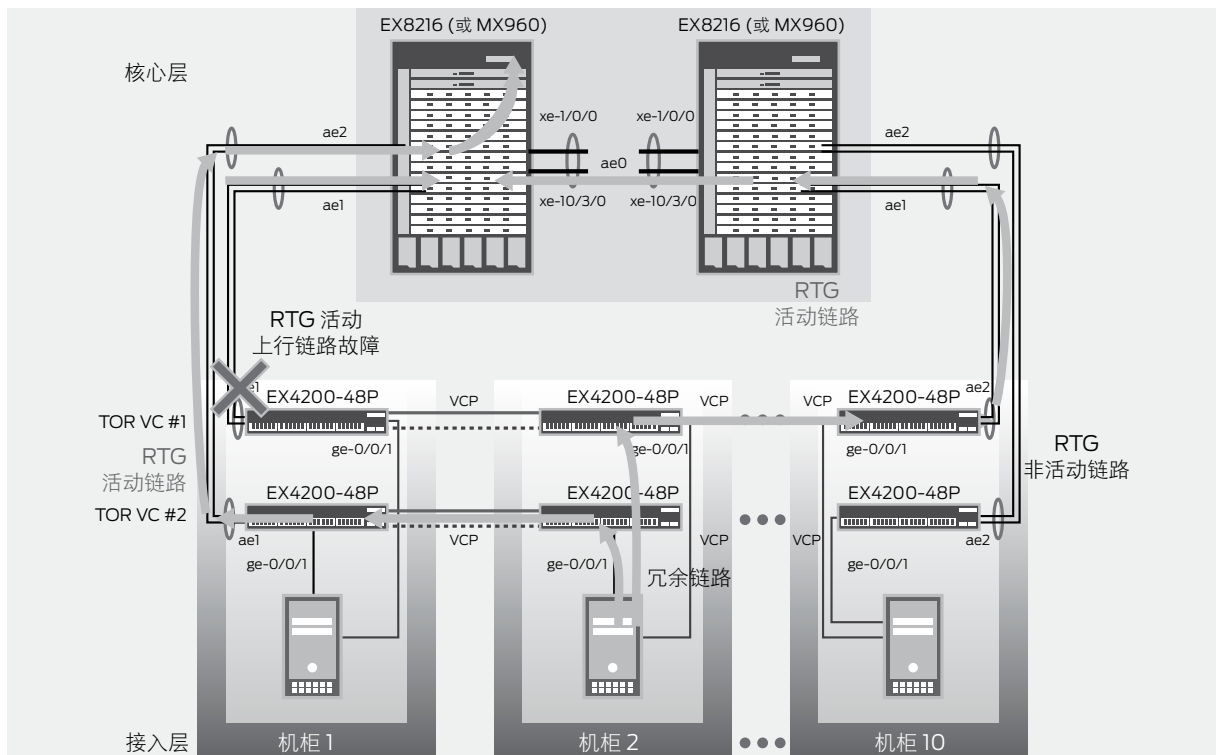


图5：接入交换机/集群交换机与EX8200交换机或MX系列路由器之间的RTG活动链路发生故障时的流量情况

### 从客户端到服务器的南向流量

在正常运行过程中，从客户端到服务器的南向流量通过两个活动的RTG LAG连接传输到接入层EX4200集群交换机。如果活动LAG接口的一个成员链路发生故障，那么从客户端到服务器的流量将通过该LAG连接中的其他成员链路传输到接入集群交换机（按照链路汇聚组中的定义），然后通过服务器接入链路传输至服务器。

当接入层EX4200集群交换机与核心层MX系列路由器或EX8200交换机上活动的AE接口发生故障时，辅助AE接口将变成活动状态（由接入交换机/集群交换机上的RTG设置进行触发），并开始将路由的客户端流量从核心层MX系列路由器或EX8200交换机传输到服务器端。

## 活动MX系列路由器/EX8200交换设备发生故障

### 从服务器到客户端的北向流量

当一个核心设备发生故障后，该MX系列路由器或EX8200交换机上的RTG LAG接口将变得不可用；其他MX/EX设备上的RTG LAG接口将变成活动状态。这一变化由接入交换机/集群交换机上的RTG设置进行触发。从服务器到客户端的北向流量通过冗余链路传输到接入层EX4200集群交换机，然后通过两个集群交换机上的活动上行链路传输至核心层EX4200交换机。最终，流量将通过数据中心网络路由到其他核心层MX系列路由器或EX8200交换机的客户端。

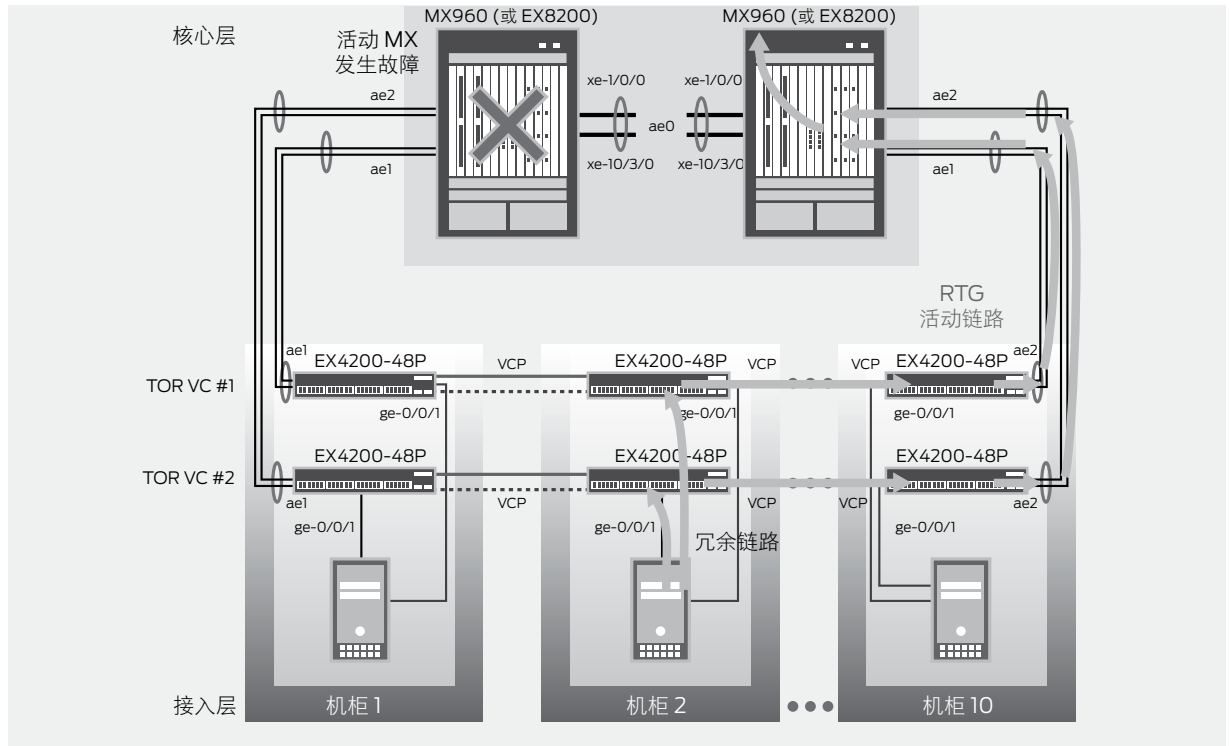


图6：活动MX系列或EX8200交换机发生故障时的流量情况

### 从客户端到服务器的南向流量

在正常运行期间，从客户端到服务器的南向流量在活动MX系列路由器或EX8200交换机上进行路由，然后经过活动的RTG LAG连接传输到接入层EX4200集群交换机。当一个核心层路由器发生故障后，来自其余MX系列路由器/EX8200交换机的连接将会变成活动状态（由接入交换机/集群交换机上的RTG设置进行触发）。从客户端到服务器的南向流量将只在这些MX系列路由器或EX8200交换机上进行路由，然后经过活动的LAG连接传输到接入集群交换机，再经过服务器接入链路到达服务器。如果集群交换机发生故障，则流量将按照与本例中直连式接入交换机发生故障时完全相同的路径进行传输。

## 最佳运行实践

本节将通过介绍设备配置和配置/升级过程的步骤，提供使用MX系列路由器和EX8200及EX4200平台运行数据中心网络的最佳实践。

### 构建新VLAN (端到端)

网络设备通过数据中心LAN相连，以共享域名系统 (DNS) 和文件服务器等公共资源，以便客户端能够通过外部网络与服务器相连接。如果没有桥接和VLAN，以太网LAN上的所有设备都将处于同一个广播域内，这些设备全部都

将检测LAN上的所有数据包。桥接可以通过创建VLAN分隔LAN上的广播域，而VLAN是将相关设备分组到不同网段的独立逻辑网络。设备在VLAN上的分组与设备在LAN上的物理位置无关。

参见图1所示的数据中心网络。为了将网络设备连接到数据中心LAN（接入层使用EX4200交换机，核心层使用MX系列路由器），您必须（至少）通过以下步骤在EX4200和MX系列设备上配置桥接和VLAN：

1. 创建新的VLAN（例如vlan71），并将该VLAN与连接服务器（位于两个EX4200柜顶式交换集群上）的接入端口进行关联。

```
{master}[edit]
root@EX4200# set vlans vlan71 vlan-id 71 interface ge-0/0/1.0
root@EX4200# set interfaces ge-0/0/1 unit 0 family ethernet-switching port-mode access
root@EX4200# set interfaces ge-0/0/1 unit 0 family ethernet-switching vlan members vlan71
```

2. 将新建的VLAN加入连接到核心层设备（位于两个EX4200柜顶式集群交换机上）的所有四个LAG上行链路中。

```
{master}[edit]
root@EX4200# set interfaces ae1 unit 0 family ethernet-switching port-mode trunk
root@EX4200# set interfaces ae1 unit 0 family ethernet-switching vlan members vlan71
root@EX4200# set interfaces ae2 unit 0 family ethernet-switching port-mode trunk
root@EX4200# set interfaces ae2 unit 0 family ethernet-switching vlan members vlan71
```

- 3a. 在核心层MX系列路由器上为新建的VLAN创建桥接域。

```
{master}[edit]
root@MX960# set routing-instances ServicesSwitch bridge-domains VLAN71 domain-type bridge
root@MX960# set routing-instances ServicesSwitch bridge-domains VLAN71 vlan-id 71
root@MX960# set routing-instances ServicesSwitch bridge-domains VLAN71 interface ae0.71
root@MX960# set routing-instances ServicesSwitch bridge-domains VLAN71 interface ae1.71
root@MX960# set routing-instances ServicesSwitch bridge-domains VLAN71 interface ae2.71
```

- 3b. 在核心层EX8200交换机上创建新的VLAN，并在中继接口上配置VLAN成员关系。

```
{master}[edit]
root@EX8200# set vlans ServicesSwitch vlan-id 71
root@EX8200# set interfaces ae1.0 family ethernet-switching vlan members 71
root@EX8200# set interfaces ae2.0 family ethernet-switching vlan members 71
root@EX8200# set interfaces ae3.0 family ethernet-switching vlan members 71
```

4. 将新建的VLAN加入到核心层MX系列路由器上的所有LAG中继中，从而连接至接入层EX4200柜顶式集群交换机和冗余核心层MX系列路由器。

针对9.2版本之前的Junos代码：

```
{master}[edit]
root@MX960# set interfaces ae0 unit 71 encapsulation vlan-bridge
root@MX960# set interfaces ae0 unit 71 vlan-id 71
root@MX960# set interfaces ae1 unit 71 encapsulation vlan-bridge
root@MX960# set interfaces ae1 unit 71 vlan-id 71
root@MX960# set interfaces ae2 unit 71 encapsulation vlan-bridge
root@MX960# set interfaces ae2 unit 71 vlan-id 71
```

针对9.2版本之后的Junos代码：

```
{master}[edit]
root@MX960# set interface ae0 unit 0 family bridge vlan-id-list 71
```

5a. 根据需要，在两个核心层MX系列路由器上为新建的VLAN创建支持VRRP的IRB接口。

在主MX系列路由器上：

```
{master}[edit]
root@MX960# set interfaces irb unit 71 family inet address 172.16.56.2/24 vrrp-group 1
virtual-address 172.16.56.1
root@MX960# set interfaces irb unit 71 family inet address 172.16.56.2/24 vrrp-group 1
priority 190
root@MX960# set interfaces irb unit 71 family inet address 172.16.56.2/24 vrrp-group 1
preempt
root@MX960# set interfaces irb unit 71 family inet address 172.16.56.2/24 vrrp-group 1
accept-data
```

在辅助MX系列路由器上：

```
{master}[edit]
root@MX960# set interfaces irb unit 71 family inet address 172.16.56.3/24 vrrp-group 1
virtual-address 172.16.56.1
root@MX960# set interfaces irb unit 71 family inet address 172.16.56.3/24 vrrp-group 1
priority 180
root@MX960# set interfaces irb unit 71 family inet address 172.16.56.3/24 vrrp-group 1
preempt
root@MX960# set interfaces irb unit 71 family inet address 172.16.56.3/24 vrrp-group 1
accept-data
```

5b. 根据需要，在两个核心层EX8200交换机上为新建的VLAN创建支持VRRP的RVI接口。

在主EX8200交换机上：

```
{master}[edit]
root@EX8200-1# set interfaces vlan unit 71 family inet address 172.16.56.2/24 vrrp-group 1
virtual-address 172.16.56.1
root@EX8200-1# set interfaces vlan unit 71 family inet address 172.16.56.2/24 vrrp-group 1
priority 190
root@EX8200-1# set interfaces vlan unit 71 family inet address 172.16.56.2/24 vrrp-group 1
preempt
root@EX8200-1# set interfaces vlan unit 71 family inet address 172.16.56.2/24 vrrp-group 1
accept-data
root@EX8200-1# set vlans ServicesSwitches 13-interface vlan.71
```

在辅助EX8200交换机上：

```
{master}[edit]
root@EX8200-2# set interfaces vlan unit 71 family inet address 172.16.56.3/24 vrrp-group 1
virtual-address 172.16.56.1
root@EX8200-2# set interfaces vlan unit 71 family inet address 172.16.56.3/24 vrrp-group 1
priority 180
root@EX8200-2# set interfaces vlan unit 71 family inet address 172.16.56.3/24 vrrp-group 1
preempt
root@EX8200-2# set interfaces vlan unit 71 family inet address 172.16.56.3/24 vrrp-group 1
accept-data
root@EX8200-2# set vlans ServicesSwitches 13-interface vlan.71
```

## 升级两个层

在升级用于数据中心MX系列路由器或EX8200交换机和EX4200平台的Junos软件之前，请确保已记录了现有系统的信息，以便在完成升级后比较相同的信息，验证所有组件是否均已安装且达到预期要求。如欲了解在升级Junos软件前记录系统信息的详细步骤，请参阅[www.juniper.net/techpubs/software/nog/nog-baseline/html/upgrade2.html](http://www.juniper.net/techpubs/software/nog/nog-baseline/html/upgrade2.html)。

您可以访问[www.juniper.net/customers/support](http://www.juniper.net/customers/support)，从“Download Software”（软件下载）菜单中下载用于MX、EX8200和EX4200平台的Junos软件。如欲下载Junos软件，您必须拥有瞻博网络的用户账号。想要了解如何获得账号，请参阅[www.juniper.net/entitlement/setupAccountInfo.do](http://www.juniper.net/entitlement/setupAccountInfo.do)。

升级交换机软件须遵循以下步骤。

1. 如上所述，下载用于相关EX4200交换机的Junos软件包。
2. 将软件包复制到交换机上。我们建议您使用FTP将文件复制到/var/tmp目录下。
3. 在运行模式下输入以下命令，开始在交换机上安装新的Junos软件包：

```
root@EX4200> request system software add source [member member_id] reboot
```

包括成员选项，只在集群交换机的一个成员上安装Junos软件包。集群交换机其他成员不受影响。

如欲在集群交换机的所有成员上安装软件，请勿包括成员选项。

使用下列一个路径取代原路径：

从交换机的本地目录安装软件包：/pathname/package-name

从远程地址下载并安装软件包：

```
ftp://hostname/pathname/package-name
http://hostname/pathname/package-name
```

例如，数据包名称为jinstall-ex-9.2R1.10-domestic-signed.tgz：

1. 按照上述步骤，在第一个EX4200柜顶式集群交换机上安装升级软件。
2. 如图1所示，首先在核心层备用MX系列上安装Junos软件包。

升级路由器软件须遵循以下步骤。

1. 使用request system software add (请求添加系统软件) 命令安装新的软件包：

```
user@host> request system software add validate /var/tmp/jinstall-9.2R1.8-domestic-signed.tgz
```

2. 使用request system reboot (请求系统重启) 命令重启路由器，开始运行新软件：

```
user@host> request system reboot
```

3. 升级EX8200交换机需要一个“维护窗口期” (maintenance window)。升级RE时，两个RE都需要重启 (这将使机箱关机)。整个升级过程可能需要5至8分钟。

```
root@EX8200> request system software add re1 <source> reboot  
root@EX8200> request system software add re0 <source> reboot
```

4. 系统重启后将加载该软件。安装过程可能需要5至10分钟。路由器随后将从刚刚安装了软件的引导设备上重启。重启后，路由器会显示登录提示。登录并输入show version (显示版本) 命令，以验证安装的软件版本。

如果网络中的MX系列路由器有两个路由引擎，则应分别在每个路由引擎上安装Junos软件，以避免中断网络运行。

在备用路由引擎上安装新的Junos软件版本，同时保持主路由引擎上当前运行的软件版本不变。

确定新软件版本在备用路由引擎上正常运行后，切换到新安装的路由引擎以激活新软件。

最后，在新的备用路由引擎上安装新的Junos软件版本。如欲了解如何针对具有冗余路由引擎的设备进行软件升级的更多信息，请参阅[www.juniper.net/techpubs/software/Junos/Junos92/swconfig-install/installing-the-software-package-on-a-router-with-redundantrouting-engines.html#jN11F4E](http://www.juniper.net/techpubs/software/Junos/Junos92/swconfig-install/installing-the-software-package-on-a-router-with-redundantrouting-engines.html#jN11F4E)。

5. 遵照以上步骤，在核心层主MX系列路由器上升级Junos软件包。

## 总结

瞻博网络为客户提供高性能的路由和交换产品，以满足当今企业网络对端到端解决方案的需求。这些产品是大型数据中心的理想选择，并支持影响数据中心的新兴技术潮流，例如节能、位置整合(location consolidation)以及服务器和应用虚拟化等。企业数据中心架构可以简化为两层（传统架构为三层），接入层为连接到线速可堆叠交换机的设备，核心层为强大的高密度线速路由器。通过紧缩数据中心内的设备和/或层，可以节省资本支出 (CAPEX) 和运营支出 (OPEX)。分层拓扑结构可将数据中心网络分割成独立的功能区，简化了网络运营并提高了可用性。分层基础设施中的每个层具有特定的作用。

本文介绍了如何在核心层上使用支持集群交换技术的瞻博网络EX4200交换机，以及在接入层与基于MX系列以太网业务路由器或EX8200以太网交换机的数据中心核心网络之间建立互连，从而在单一数据中心内实施两层LAN设计。同时，本文为读者提供了一个构建可靠数据中心局域网的折衷方法。本文中所述的实施指南定义了一个由多个功能分区构成的数据中心网络，这些分区有的可进行物理分割，有的不行。因此，我们还提供了多个用例，演示了处于同一机柜中但分属于不同数据中心功能区 (或应用层) 的服务器。

## 关于瞻博网络

瞻博网络是高性能网络领域中的领导者。瞻博网络提供高性能的网络基础架构，能够在单一网络中创建一个具有响应性的和受信赖的环境，从而加速服务和应用的部署，并推动高性能的业务进行。欲知详情，请访问[www.juniper.net/cn/zh/](http://www.juniper.net/cn/zh/)。

### 北京代表处

北京市东城区东长安街1号  
东方经贸城西三办公楼15层1508室  
邮政编码: 100738  
电话: 8610-5812-6000  
传真: 8610-8518-2626  
[www.juniper.net/cn/zh/](http://www.juniper.net/cn/zh/)

### 上海代表处

上海市淮海中路333号  
瑞安广场1102-1104室  
邮政编码: 200021  
电话: 8621-6141-5000  
传真: 8621-6141-5090

### 广州代表处

广州市天河区天河路228号  
广晟大厦28楼03-05单元  
邮政编码: 510620  
电话: 8620-8511-5900  
传真: 8620-8511-5901

如需购买瞻博网络 (Juniper Networks) 解决方案，请与当地的瞻博网络 (Juniper Networks) 销售代表或授权经销商联系。

Copyright 2009, Juniper Networks, Inc. 版权所有，保留所有权利。Juniper Networks, Juniper Networks标识, Junos, NetScreen和ScreenOS是瞻博网络 (Juniper Networks) 在美国和其他国家的注册商标。Junos是瞻博网络 (Juniper Networks) 所属商标。所有其他的商标、服务标记、注册商标或注册的服务标记均为其各自公司的财产。瞻博网络 (Juniper Networks) 不承担由本资料中的任何不准确性或引起的任何责任，瞻博网络 (Juniper Networks) 保留不做另行通知的情况下对本资料进行变更、修改、转换或以其他方式修订的权利。

文档编号: 8010014-002-SC 10/2009