高可用性园区网络故障恢复分析

简介

无论何种规模大小的企业园区都需要高度可用及安全、智能的网络基础设施，为话音、视频、无线和关键任务数据应用等业务解决方案提供支持。要想提供这种可靠的网络基础设施，必须将由构成园区整个系统的部件故障而引发的运行中断降至最低限度。了解系统如何从部件停运（包括计划内停运和故障）中恢复，以及停运过程中可能发生的行为，在高可用安全园区网络的设计、升级和运行过程中，无疑是一个关键步骤。

本文是《设计高可用性园区网络》的附属文。

本问对这些文章中所描述的园区网络设计的故障恢复进行了分析，包括下列章节：

目标读者

本文的目标读者是负责园区网络设计的思科系统工程师和客户工程师，本文也可帮助运营人员及其他员工了解现有生产园区网络可能的收敛行为。

本文目的

本文对推荐的分级园区设计出现主要部件故障后，观察到的数据流恢复时间进行了记录和分析。旨在为园区网络构建或升级过程中的设计方案提供评估参考意见。

概述

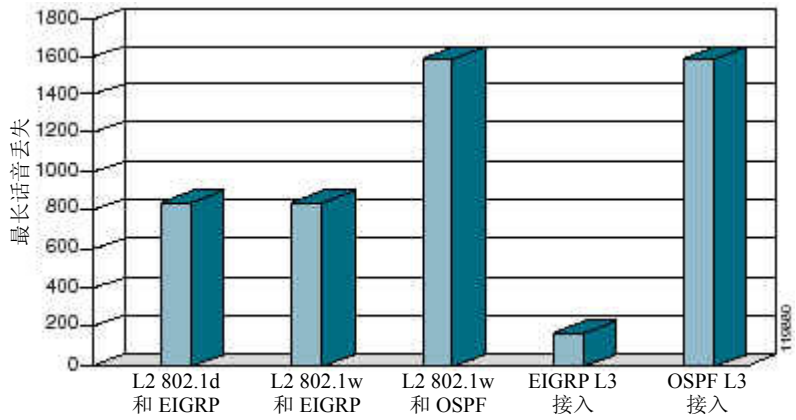
本节包括下列主题：

高可用性园区网络故障恢复分析

收敛分析综述

采用增强内部网关路由协议（EIGRP）的端到端第三层设计，在发生任意单一部件、链路或节点故障时提供了理想的恢复功能。图 1 介绍了任意单一部件故障测试过程中记录到的最长的恢复时间。

图 1
最长话音丢失时间



测试表明，在接入层运行第三层和 EIGRP 的园区网络中，在任意单一部件故障时，G.711 话音流量最长的丢失时间不足 200 毫秒。

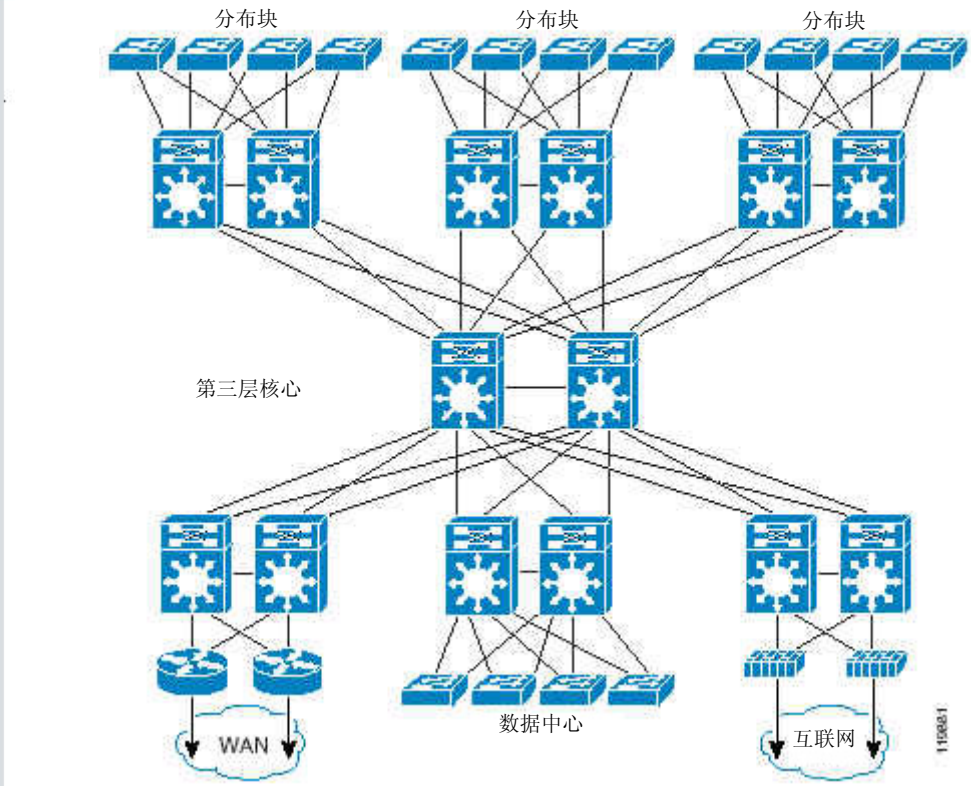
测试中观察到，采用亚秒热备份路由协议（HSRP）/网关负载均衡协议（GLBP）计时器的传统第二层接入设计的收敛，在发生任意部件故障时均不到一秒。此恢复时间在 IP 电话的可接受范围内，并将故障时终端用户对话音质量影响的感觉降至了最低限度。

注: 在上述情况中，由于交换管理引擎故障或软件问题造成的接入交换机故障，会导致连接到故障接入交换机的所有设备的话音和数据丢失时间延长。为将接入交换机发生故障问题的可能性尽量降低，思科建议每台接入交换机均使用状态化切换（SSO）或不间断转发（NSF）/SSO 等冗余交换管理引擎配置，或实施冗余堆叠。本文的分析结果中不包括对冗余交换管理引擎收敛的分析。

所测试的园区设计

选定进行测试的设计是根据《设计高可用性园区网络》中列举的层次化设计建议而确定的。所有所测试的设计都使用第三层路由核心连接其他架构模块的，如图 2 所示。

图 2 园区设计



在结构化层次设计模式中，进行了下列四种基本分布构建块的测试：

- 利用每 VLAN 生成树＋（PVST+）的第二层接入
- 运行快速 PVST+的第二层接入
- 第三层接入端到端 EIGRP
- 第三层接入端到端开放最短路径优先（OSPF）

这四种分布设计的部件故障和部件恢复测试均已完成。

除了对这四种基本分布配置测试以外，还对比较各种基本的 L2 分布块设计进行了两项额外测试。第一个测试是利用运行快速 PVST+分布块设计的 L2 接入，与将 HSRP 作为冗余缺省网关协议的 GLBP 进行了比较。第二个测试将带生成树环的快速 PVST+分布块设计的恢复与不带生成树环的设计的恢复进行了比较。

注: 有关各项设计实施的更多信息，请参考《设计高可用性园区网络》。

下列章节对观察结果进行了分析：

- 第三层核心故障分析
- 第二层分布块故障分析
- 第三层到边缘分布块的故障分析

以上每项测试都是利用穿行整个园区的网络化端到端数据流进行测试的，但是每个测试情况的分析是独立进行的。层次化设计的一项主要优势是故障域的隔离。网络核心中的节点或链路故障会

导致相同的收敛行为，对业务应用有着同样的影响，与分布块的设计无关。同样，分布块的故障也与核心隔离开来，可以进行独立检查。

注: 隔离故障事件和降低这些故障影响的能力只有在类似于《设计高可用性园区网络》中所描述的层次化设计中才能实现。

测试程序

测试网络的配置、测试流量和测试案例都是经过了精心挑选，以便尽可能真实地模拟实际的客户信息流和可用性要求。测试配置旨在证实思科的最佳实践设计在实际环境中的效能。

测试基于如下假设：

- 园区网支持 VoIP 和流视频。
- 园区网支持组播流量。
- 园区网支持无线功能。
- 园区网支持事务和批量数据应用。

测试台配置

用于评估故障恢复的测试台由服务器群块及带有相连的分布块的第三层路由核心构成的。采用的核心和分布层交换机为带有 Supervisor 720a 引擎的 Cisco Catalyst 6500 交换机。接入层由 39 个双链路连接到分布层的交换机组成。测试中使用的是以下配置：

- 核心交换机——2 个 6500，带 Sup 720（本地 IOS-12.2（17b）SXA）
- 服务器群分布层——2 个 6500，带 Sup 2/MSFC2（本地 IOS-12.1（13）E10）
- 服务器群接入交换机——2 个 6500，带 Sup 720（CatOS-8.3（1））
- 分布交换机——2 个 6500，带 Sup 720（本地 IOS-12.2（17b）SXA）
- 接入交换机
  - 1 个 2950（IOS-12.1（19）EA1a）
  - 1 个 3550（IOS-12.1（19）EA1）
  - 1 个 3750（IOS-12.1（19）EA1）
  - 1 个 4006，带 SupII+（IOS-12.1（20）EW）
  - 1 个 4507，带 SupIV（IOS-12.1（20）EW）
  - 1 个 6500，带 Sup1A（CatOS-8.3（1））
  - 1 个 6500，带 Sup 2/MSFC2（IOS-12.1（13）E10）
  - 32 个 3550（IOS-12.1（19）EA1）

每台接入交换机均配备三个配置为无环拓扑结构的 VLAN：

- 专用话音 VLAN
- 专用数据 VLAN
- 唯一的本地上行链路 VLAN

测试流量

在测试过程中，使用 180 台 Chariot 终端服务器来生成网络信息流量负载，以及收集每次故障和恢复事件的影响统计数据。

注：有关 Chariot 的具体信息，请访问 http://www.netiq.com/products/chr/default.asp

根据实际的思科客户网络情况，Chariot 端点经过配置，生成混合企业数据应用流量。

连接到 39 台接入和数据中心交换机的终端经过配置，生成下列单播流量：

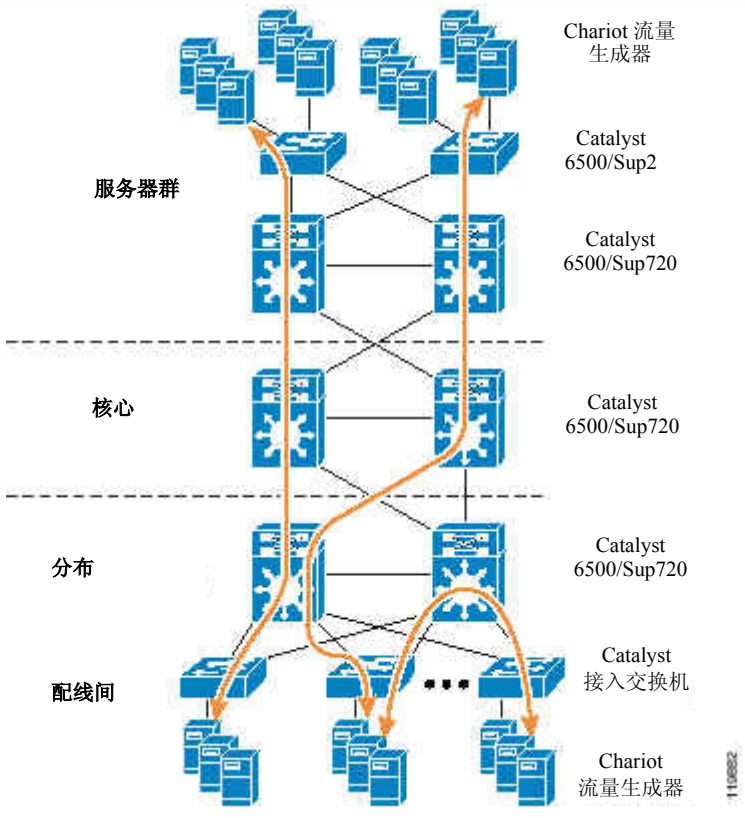
- G.711 话音呼叫——实时协议（RTP）信息流
- 94 种 TCP/UDP 数据流类型，用于模拟呼叫控制、批量数据（ftp）、关键任务数据（HTTP，tn3270）、POP3、HTTP、DNS 和 WINS。

所有流量都根据当前的“思科企业服务质量（QoS）园区设计指南”中的建议进行了标记，所生成的流量负载足以使所选的上行链路和核心基础设施发生拥塞。

流量进行了适当的定义，使得其中大部分流量利用网络核心在接入层的数据中心间传输。部分 VoIP 流被配置成利用分布交换机在接入交换机间传输，如图 3 所示。

图3

样本流量测试台



除了单播流量外，每台接入交换机还配备了 40 个组播接收器，负责接收下列混合组播流：

- 等待音乐（MoH）流，64kbps/50pps（160 字节负载，RTP=PCMU）
- IPTV 视频流，1451kbps（1460 字节负载，RTP=MPEG1）
- IPTV 音频流，93kbps（1278 字节负载，RTP=MPEG2）
- NetMeeting 视频流，64kbps（522 字节负载，RTP=H.261）
- NetMeeting 音频流，12kbps（44 字节负载，RTP=G.723）

- Real 音频流，80kbps（351 字节负载，RTP=G.729）
- Real 媒体流，300kbps（431 字节负载，RTP=H.261）
- 组播 FTP 流，4000kbps（4096 字节负载，RTP=JPEG）

所有组播 MoH 都标记为快速转发（EF），其他所有组播流量都标记为差分服务代码点（DSCP）14（AF13）。

收敛时间确定方法

为使该测试有助于了解故障事件对生产网络中应用和话音流量的影响，本文所记录的收敛结果都是根据真实的 UDP 和 TCP 测试流量的测量值得出的。每个故障情况所记录的收敛时间，由每次测试运行中对所有活动的 G.711 话音流中的最长分组丢失时间的测量来确定。

注：标准 G.711 编译码器每秒传输 50 个分组，其速度为每 20 毫秒 1 个分组。n 个连续分组丢失等于 n*20 毫秒中断。

记录的最差结果是多次执行单个测试情况时观察到的最大数值，并非平均收敛时间，而是代表了例外情况。采用最差观察值旨在为评估收敛对生产网络的影响提供保守的参数。

每项测试都至少重复了 3 次。光纤故障测试由下列三种测试情况组成：

- 链路中两条光纤故障
- 单一光纤故障，传输方向
- 单一光纤故障，接收方向

涉及节点的故障由下列三种测试情况组成：

- 电源故障
- 模拟软件（IOS/CatOS）故障
- 模拟交换管理引擎故障

另外，对于涉及接入交换机的那些测试，它们的三个子测试情况均运行了多次，针对不同的接入交换机类型（所有进行测试的接入交换机列表请见上文）。

除了最长话音丢失周期以外，所有话音呼叫的平均印象分值（MOS），以及网络抖动和延迟也都进行了记录。

同时，也收集了网络收敛对主用 TCP 信息流影响的测试数据。在所有测试案例中，丢失周期很短，不足以造成 TCP 会话丢失。网络连接丢失的确会暂时影响到这些信息流的吞吐率。值得注意的是“影响间隔时间”，即 TCP 流量并非按正常吞吐率运行的时间周期，要长于 G.711 UDP 流量的丢失周期。如同预料之中，收敛过程中的分组丢失触发了 TCP 退回算法。TCP 流量恢复到理想吞吐量的时间等于[分组丢失周期＋TCP 流量恢复所需时间]。

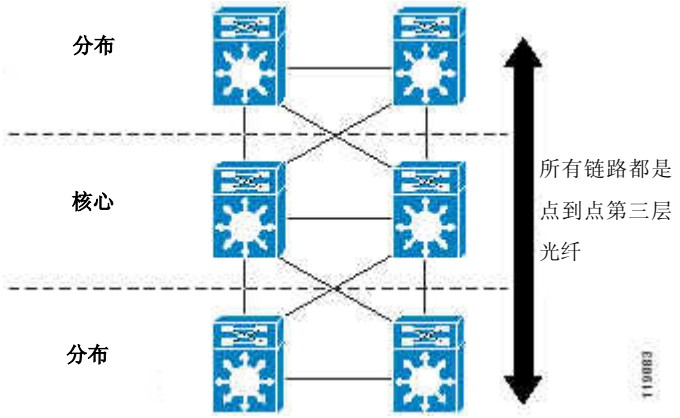
第三层核心收敛——结果与分析

本节包括下列主题：

园区核心介绍

园区核心为其他所有分级构建块间提供了冗余高速连接。建议采用一种利用点到点第三层光纤连接的全面网格化核心设计，如图 4 所示，来提供最优收敛行为。

图 4
核心拓扑结构



测试的网络核心由一对 Cisco Catalyst 6500/Supervisor 720 组成，每台分布交换机和核心交换机间有冗余点到点 10 千兆位以太网（GigE）链路。核心交换机通过点到点 10GigE 光纤相连接。虽然在纯单播环境中并非一定需要该链路来实现冗余，但在园区设计中依然属于建议采用的内容。在某些配置中，该链路被用于组播流量恢复。也有必要将核心节点配置成缺省或汇总路由信息发送至网络。

除了图 4 中所示的两个传输测试流量的分布块外，还有其他分布交换机组和后端路由器连接至核心，模拟园区到企业 WAN 的连接。没有测试流量会配置成在测试网络的这一部分转发，它只是用于将额外的路由提供给园区。测试中的园区总共拥有 3572 条路由。

Core-Switch-2#show ip route summary

IP routing table name is Default-IP-Routing-Table(0)

Route Source	Networks	Subnets	Overhead	Memory (bytes)
connected	1	10	704	1760
static	0	1	64	160
eigrp 100	1	3561	228224	569920
internal	5			5900
Total	7	3572	228992	577740

等成本路径第三层园区设计的优点

在推荐的园区设计中，每个分布和核心节点，都拥有等成本路径，用于将内容转发给所有目的地，而非本地连接的子网。这两种等成本路径是相互独立的，在发生任意单点部件故障时（链路或节点），这意味着总有路径确保提供有效的路由。在任意单点部件发生故障时，该设计中的每台交换机能够成功地恢复并绕行发生的故障下一跳。

因为每个节点拥有两条路径，能够从任意链路故障中恢复，所以下游设备从来无需因上游故障而重新计算路由，因为上游设备始终拥有一条有效的路径。网格化第三层设计的架构优势是，在发生任意单点部件故障时，所有路由收敛对于交换机都是在本地完成的，从不需要依赖路由协议来间接检测和恢复链路或节点故障。

在第三层核心设计中，从任意分布交换机流向其他任意分布交换机的流量的收敛时间，主要由分布交换机上的链路丢失检测而定。在 GigE 和 10GigE 光纤上，链路丢失检测通常是利用远程故障检测机制完成的，这种机制是由 802.3z 和 802.3ae 链路协商协议的一部分实施的。

注：有关 10GigE 和 GigE 的更多信息，请分别参见 IEEE 标准 802.3ae 和 802.3z。

当分布交换机检测到链路丢失，它会通过下列三个步骤对链路中断事件进行处理：

1. 删除路由表中与故障链路有关的内容
2. 更新软件思科快速转发（CEF）表，以显示这些受影响路由的下一个跳邻接项的丢失情况
3. 更新硬件表，以显示软件表中的有效下一跳邻接项的变化

在等成本路径核心配置中，交换机拥有两条路由和两个相关硬件 CEF 转发邻接项。在链路出现故障前，流量通过这两个转发项发送。删除了两项之一后，交换机就开始利用剩余的 CEF 项转发所有流量。网络所有流量的恢复时间，取决于检测物理链路故障，以及更新软件和相关硬件转发项所需的时间。

这里所建议的等成本路径设计的关键优势在于，网络的恢复行为既快速又具确定性。

使用等成本路径潜在的问题之一，是限制了工程师专用流量沿着某条链路传输的能力。这项设计克服了这一弱点，通过尽可能简化的配置和最快速的持续收敛时间，提供了更高的总体网络可用性。

注：虽然无法在等成本路径第三层交换园区中配置某特定信息流所用的路径，但是了解信息流确定的走向却是可行的。硬件转发算法可以一致地沿着网络中相同的路径，转发相同的信息流。这种一致的行为有助于诊断和容量规划，某种程度上抵消了冗余路径流量模式的问题。

第三层核心收敛结果——EIGRP 和 OSPF

本节包括下列几个部分：

故障分析

园区核心要检查以下三种基本部件故障：

- 核心节点故障
- 核心到分布层光纤故障
- 核心到核心互联光纤故障

表 1 总结了测试结果。

表 1 故障测试结果

故障情况	上行恢复	下行恢复	恢复机制
节点故障	200 毫秒	200 毫秒	上行——L3 等成本路径 下行——L3 等成本路径
核心到分布层链路故障	200 毫秒	200 毫秒	上行——L3 等成本路径 下行——L3 等成本路径
核心到核心链路故障	0 毫秒	0 毫秒	上行——无转发路径丢失 下行——无转发路径丢失

在推荐的园区核心设计中，所有节点都拥有冗余等成本路由。直接的效果是，单一部件故障的恢复时间不依赖于路由协议恢复来恢复流量。在单一链路发生故障时，各节点能够独立地将所有流量重新路由至剩余冗余路径。在节点故障时，当互联链路故障时，受影响的相邻交换机可以检测丢失情况，并能够独立地将所有流量重新路由至剩余冗余路径。

三种故障情况中有两种——核心节点自身故障，以及连接分布层到核心交换机的任意光纤故障——依赖于等成本路径恢复。在第三种情况中，即发生核心到核心光纤链路故障时，主用转发路径无丢失，因而在丢失事件中不会影响单播信息流动。在所有三种情况下，网络都能够成功地恢复所有单播信息流，无需等待任何路由协议拓扑结构更新和重新计算。

核心交换机间的光纤故障通常不会对单播流量造成直接的影响。鉴于核心网络采用全面网格化设计，该链路仅用于双故障情况下的单播流量。虽然纯单播环境并非一定需要此链路来实现冗余，但在园区设计中依然属于建议采用的内容。在某些配置中，该链路用于组播流量恢复。也有必要将核心节点配置成缺省或汇总路由信息发送至网络。

虽然网络因部件故障而恢复流量的能力与所使用的路由协议无关，但还是采用了路由协议收敛。EIGRP 生成拓扑结构更新，OSPF 则大量发送链路状态通知（LSA）并进行 Dijkstra 计算。为将这些事件对网络的影响降至最低限度，思科建议园区设计遵循优秀的路由协议设计准则。更多信息请参考 HA 园区和第三层接入设计指南。

导致等成本路径恢复事件的恢复时间取决于：

- 检测物理链路故障所需的时间
- 更新软件和相应硬件转发表所需的时间

为实现链路丢失的快速检测（这是实现上述记录的收敛时间所必须的），需要确保对于所有点到点链路启用 802.3z 或 802.3ae 链路协商。CatOS 和 Cisco IOS 的缺省行为是启用链路协商。禁用链路协商会增加上行和下行流量的收敛时间。

CatOS:

set port negotiation [***mod/port***] enable

show port negotiation [*mod[/port*]]

Cisco IOS:

int gig [mod/port]

[no] speed nenegotiate

恢复分析

链路和设备恢复的收敛情况与故障时的收敛相同：

- 核心节点恢复
- 核心到分布交换机光纤恢复
- 核心到核心互联光纤恢复

表 2 总结了测试结果。

表 2 恢复测试结果

故障情况	上行恢复	下行恢复	恢复机制
节点恢复	0 毫秒	0 毫秒	主用数据路径无丢失
核心到分布交换机链路恢复	0 毫秒	0 毫秒	主用数据路径无丢失
核心到核心链路恢复	0 毫秒	0 毫秒	主用数据路径无丢失

表 2 的结果表明，第三层园区链路和节点恢复通常对现有和新的数据流影响极低。第三层转发路径的激活或重新激活拥有这种固有的优势；交换机不向上游或下游相邻设备转发任何流量，直至相邻设备指示它可以转发该流量。通过在转发流量之前确保存在有效路由，交换机可以继续使用现有冗余路径，同时激活新的路径。

在新激活链路的激活过程中，交换机执行 EIGRP/OSPF 相邻点搜寻和拓扑结构交换。随着每台交换机了解这些新路由，它会在路由和 CEF 转发表中创建一个备用等成本转发项。在冗余设计中，该备用等成本转发项被添加时，无需禁用当前现有的转发项。在路由协议更新过程中，交换机可以继续使用现有硬件转发项，并且不会因为新路由的插入而丢失任何数据。

不同于上述的部件故障情况， 路由删除是独立于路由协议收敛，每台园区交换机直接依赖于路由协议，以便在激活新链路或节点时安装新的路由。网络性能与这些新路由插入的速度无关，因此，EIGRP 和 OSPF 传播和插入新路由的速度不是一个重要的跟踪数据。但是，在故障情况下，必须遵循建议的设计准则，以确保将路由收敛过程对网络总体的影响降至最低限度。

在大多数环境中，冗余的第三层园区网络设计中的链路或交换机激活不会造成影响。但是，在新路由插入的过渡阶段，无论它是更好的路径路由还是备用等成本路由，在一个超额配置的网络中，一个在全新激活路径上发送的现有信息流的分组，有可能比通过旧路径传输的分组提前到达，因而造成数据包到达时间顺序混乱。这种情况只有在原始路径发生严重拥塞，导致持续延迟时才会出现。

在高度超额配置（最差情况）负载测试中，我们发现，由于重新排序造成的话音流分组丢失占有效话音流的不到 0.003%。分组丢失水平极低，以及由于激活备用链路和话音流转发路径的动态变更，造成的相关抖动也较低，因此，对记录的测试信息流 MOS 分值不产生可度量的影响。冗余第三层园区设计中新链路或节点的激活，不对现有信息流的运行产生影响。

第三层分布和第二层接入的收敛——结果与分析

本节讨论下列主题：

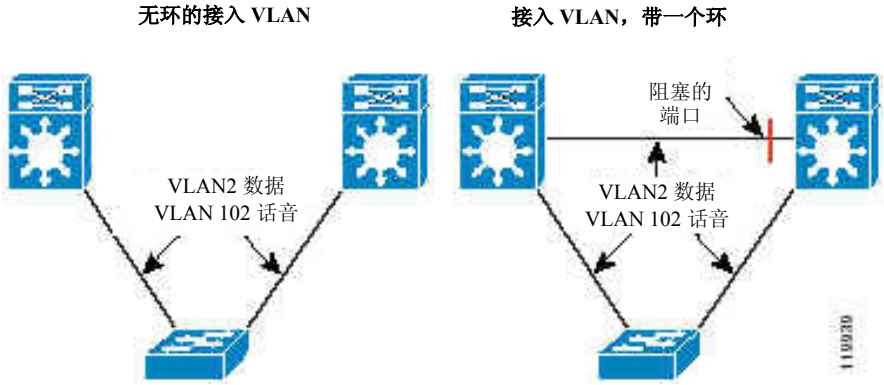
测试配置概述

构成分布层的交换机和所有相连的接入交换机通常被称为分布块。在分级设计模式中，分布块设计为连接到接入交换机的设备间传输的信息流提供了永续性，并为园区核心提供了冗余连接，以便为所有出入该园区的流量提供永续性。

在标准分布块设计中存在两种配置情况：

- 配置了第二层环的 VLAN
- 以无环拓扑结构配置的 VLAN

图 5
带和不带第二层环的标准分布构建块



这两种基本情况，由于可用的缺省网关、生成树版本和路由协议的差异，各自都可能有一些配置变化。之所以选择本节讨论的五种测试配置，是为了演示可能的各种配置间的差别（见表 3）。

表 3 五种测试配置

测试配置	缺省网关协议	生成树版本	路由协议
配置 1	HSRP	PVST+（无环）	EIGRP
配置 2	HSRP	Rapid-PVST+（无环）	EIGRP
配置 3	HSRP	Rapid-PVST+（无环）	OSPF
配置 4	GLBP	Rapid-PVST+（无环）	EIGRP
配置 5	GLBP	Rapid-PVST+（带环行拓扑结构）	EIGRP

对于五种测试配置中的每一种，都执行下列五种基本故障测试：

1. 接入交换机到主用缺省网关（HSRP/GLBP）分布交换机的上行链路光纤发生故障
2. 接入交换机到备用缺省网关（HSRP/GLBP）分布交换机的上行链路光纤发生故障
3. 主用缺省网关分布交换机发生故障
4. 备用缺省网关分布交换机发生故障
5. 交换机间分布层到分布层光纤连接发生故障

测试 1 和 2 被运行了多次，每次针对不同的接入交换机类型，由下列三种故障情况组成：

- 链路中两条光纤故障
- 单一光纤故障，传输方向
- 单一光纤故障，接收方向

在每项故障测试中，进行一台替补交换机重启或链路激活测试，以用于检查运营小组重启或更换故障部件所带来的影响。

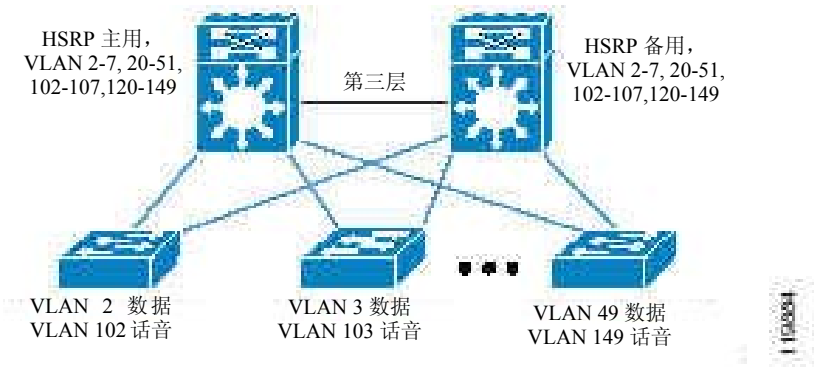
下面报告的测试结果是在多次反复测试中的最差观察值。在前四项测试中，物理拓扑结构、VLAN 和其他所有配置在整个测试中都保持一致。在第五项测试中，话音和数据 VLAN 被配置成要穿越连接 2 台分布交换机的干线。请参考下面的分布交换机和接入交换机说明和配置。

注：在下列结果与分析中，只有第一项配置收录了具体的故障和恢复结果。其他四项配置，分析部分只是介绍了配置变化对网络恢复的影响。

分布构建块说明

标准分布构建块由两台分布交换机和上行连接到这两台分布交换机的多台接入交换机组成，见图 6。

图 6
标准分布构建块，无第二层环（测试配置 1-4）



在基本物理拓扑结构的制约下，分布构建块的配置细节随着时间而有所变化。前四项测试网络配置中采用了下列设计选项：

- 每台接入交换机都配置了独特的话音和数据 VLAN

  接入和分布交换机间的上行链路是第二层干线，它的配置为了支持本地、数据和话音 VLAN。第三个独特的本地 VLAN 用于防御 VLAN 跳转攻击。更多信息，请参考《设计高可用性园区网络》和《SAFE 企业安全蓝图第二版》。

- 分布交换机间的链路是一条第三层点到点链路。

  话音和数据 VLAN 对于每台接入交换机都是独特的，在接入交换机和分布交换机间汇聚，但不在分布交换机间汇聚。分布交换机间的链路是按照第三层点到点配置的。思科最佳实践设计建议 VLAN 不要跨越多台接入交换机。在多台接入交换机间桥接的公共无线 VLAN 的使用，建议用作支持接入点（AP）间无线设备的无缝漫游的一种机制。无线 VLAN 交换模块（WLSM）的推出，提供了一种可扩展架构，可以支持快速漫游，无需用于 AP 的第二层公共 VLAN。

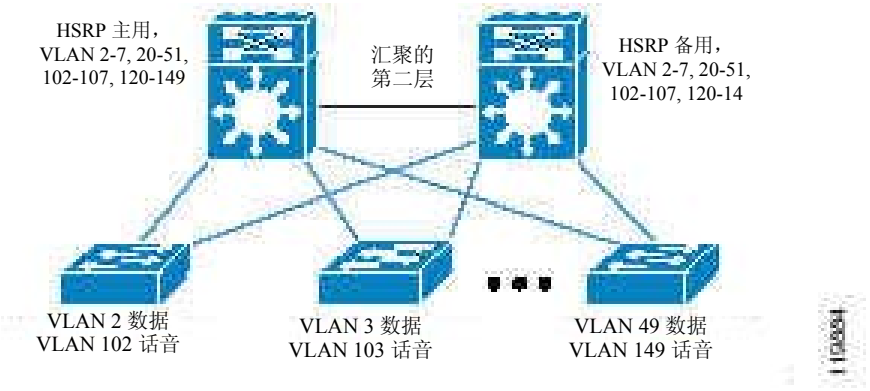
- 生成树根和 HSRP 主网关被分配给用于所有 VLAN 的分布交换机 1。

  缺省网关协议 HSRP 或 GBLP 针对每个独特的接入数据和话音 VLAN 进行了配置。在测试网络中，所有主用 HSRP 网关和每个 VLAN 相应的根网桥在分布交换机 1 上进行配置。

  当分配缺省网关位置时，通常会采用两种不同的设计方案。一种方案在两台分布交换机间更换 HSRP 网关，用于话音和数据，或是用奇偶 VLAN 来分担上行流量负载。另一种方案是将分布交换机之一作为适用于所有 VLAN 的主用网关，以便提供一致的配置和运营行为。对于那些要求负载均衡的环境，思科建议采用 GLBP，而非轮换 HSRP 组，因为它可以提供有效的上行流量负载均衡。对于那些要求更具确定性方案的环境，思科建议将所有 HSRP 组分配给单一分布交换机。

第五项测试的网络配置有所不同，见图 7。

图 7
分布构建块，带第二层环（测试配置 5）



用于第五种配置的网络不同于上述网络，差别只在于所有话音和数据 VLAN 都在两台分布交换机间的 10GigE 光纤上汇聚。这里仍旧为每台接入交换机配置了专用话音和数据 VLAN，根网桥和 HSRP 主用节点在分布交换机 1 上，针对所有 VLAN 进行了配置。

为了大幅度提高 GLBP 提供的动态缺省网关负载均衡机制的有效性，生成树被配置为阻塞连接分布层到分布层链路的端口。通过强制该链路阻塞所有 VLAN 的两条接入到分布交换机链路，网络得以在正常运行的情况下，实现上行和下行方向的共享流量加载。这一配置如下所示：

```
Distribution-Switch-2#sh run int ten 4/3

interface TenGigabitEthernet4/3
description 10GigE trunk to Distribution 1 (trunk to root bridge)
no ip address
load-interval 30
mls qos trust dscp
switchport
switchport trunk encapsulation dot1q
switchport trunk native vlan 900
switchport trunk allowed vlan 2-7,20-51,102-107,120-149
spanning-tree cost 2000 << Increase port cost on trunk to root bridge
```

注： 如图 7 所示，第二层环的使用并非最实际。虽然有多种特性，当使用正确时，可以消除使用第二层环拓扑结构的大部分风险，如环防护、单向链路检测 (UDLD)和 BPDU 防护等，但如果无扩展第二层子网的应用或业务要求，思科建议，HA 园区设计应避免任何第二层环。这一测试案例只用于提供对照分析。

欲了解这些测试中使用的设计建议的更多说明和解释，请参见《设计高可用性园区网络》。

配置 1 结果——HSRP，带 PVST+的 EIGRP

本节探讨了下列主题：

故障分析

配置 1 拥有下列特征：

- 缺省网关协议——HSRP
- 生成树版本——PVST+（每 VLAN 802.1d）
- IGP——EIGRP

表 4 总结了测试结果。

表 4 配置 1 故障测试结果

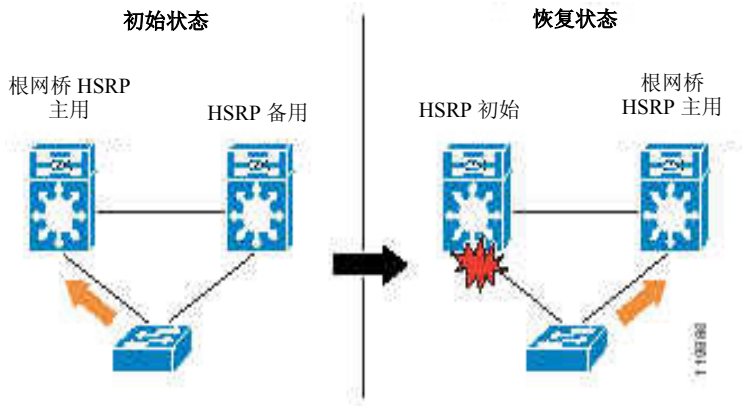
故障情况	上行恢复	下行恢复	恢复机制
主用 HSRP 的上行链路光纤故障	900 毫秒	在 700-1100 毫秒之间	上行——HSRP 下行——EIGRP
备用 HSRP 的上行链路光纤故障	0 毫秒	在 700－1100 毫秒之间	上行——无丢失 下行——EIGRP
主用 HSRP 分布交换机故障	800 毫秒	200 毫秒	上行——HSRP 下行——L3 等成本路径
备用 HSRP 分布交换机故障	0 毫秒	200 毫秒	上行——无丢失 下行——L3 等成本路径
交换机间分布光纤故障	0 毫秒	0 毫秒	主用数据路径无丢失

主用 HSRP 分布交换机上行链路光纤故障

上行信息流的恢复时间主要取决于 HSRP 计时器的配置。在正常状态下，终端站点发出的所有流量都上行传输至目的地为虚拟 HSRP MAC 地址的主用 HSRP 对等设备。当到主用 HSRP 对等设备的上行链路发生故障时，接入交换机会使所有与该链路有关的 CAM 项从其转发表中溢流，包括虚拟 HSRP MAC。

同时，由于不再从主用对等设备接收到联络信息，备用 HSRP 对等设备开始启动所配置的终止计时器。在丢失三次联络信息后，备用对等设备过渡到主用状态，传送免费 ARP，向接入交换机 CAM 表先期传播虚拟 HSRP MAC 地址的新位置，然后，开始接收和转发发送到虚拟 HSRP MAC 地址的分组（见图 8）。

图 8
主用 HSRP 分布交换机上行链路光纤故障——上行收敛



这些测试所记录的恢复时间是利用 250 毫秒联络信息和 800 毫秒终止时间间隔的 HSRP 计时器获得的。900 毫秒上行丢失是观察到的最差情况，仅发生于某一特殊的故障情况之下。两台交换机间的光纤丢失可能是一对光纤中的一条丢失，也可能是两条同时丢失。在仅是连接到主用分布交换机接收端口的光纤发生故障的情况下，还仍有少许时间供交换机传送 HSRP 联络帧，但在远程故障检测关闭该接口前，无法接收输入流量。

在主用交换机丢失传输光纤的情况下，发生了相反的效果。不发送 HSRP 联络信息，但仍能接收发送到核心的数据，使丢失周期得以缩短。虽然单一光纤故障和 HSRP 更新的传输的同步会延长最差情况下的收敛时间，但是，900 毫秒的测试结果仅是一个特殊情况，仅稍高于 780 毫秒的整体平均值。

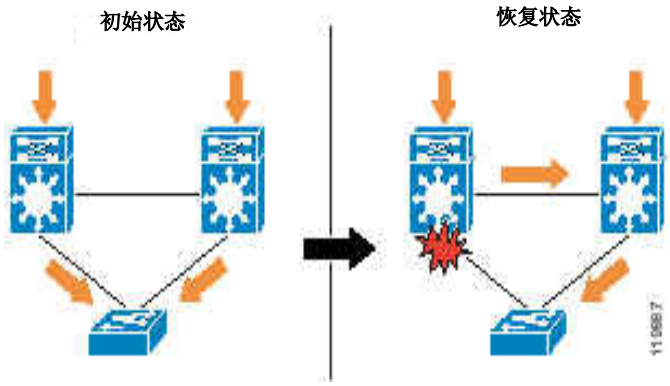
虽然主用 HSRP 对等设备的丢失还意味着生成树根网桥丢失，但这不会造成任何流量损失。生成树拓扑被配置为无环结构，无需转换任何端口的阻断状态。主用根网桥的丢失确会触发新的根网桥选择，但在 802.1d 实施中，对主用端口转发没有影响。

设计提示——虽然可以通过减少 HSRP 联络信息和终止时间间隔，以缩短话音或数据信息流上行部分的恢复时间，但是，思科建议，HSRP 终止时间应与流量下行部分的恢复时间相匹配。这些测试是采用 800 毫秒终止时间完成的，与观察到的话音 VLAN 的 EIGRP 下行恢复相对应。过多减少 HSRP 计时器，在分布交换机遭遇极高 CPU 负载时，可能导致网络不稳定。250/800 毫秒配置，在此参考测试拓扑结构中，采用 99%CPU 负载的情况下，经证实可以成功运行。

下行收敛

无环配置中，下行流量的恢复时间主要取决于路由协议收敛（见图 9）。

图 9
主用 HSRP 分布交换机上行链路光纤故障——下行收敛



检测光纤故障时，交换机处理下列事件，以恢复连接：

1. 线路协议针对受影响接口进行标记。
2. 相应的 VLAN 生成树端口被标记为禁用（停止）。
3. 由自动进程触发，与各个生成树实例有关的 VLAN 接口也被标记。
4. 与故障 VLAN 相关的 CEF 收集项被从转发表（本地连接的主机项）中清除。
5. Cisco IOS 向 EIGRP 流程通告丢失的 VLAN 接口。
6. EIGRP 清除丢失的子网路由，并向所有主用相邻设备发送对于这些路由的请求。
7. 与丢失路由相关的 CEF 项被从转发表中清除。
8. 在接到所有请求后，EIGRP 确定最佳下一跳转路由并将新路由插入路由表。
9. 与新路由匹配的 CEF 项被导入转发表。
10. 流量恢复完成。

因为分布交换机没有针对丢失网络的等成本路径或可行的替代网络，因此，EIGRP 需要启动路由收敛，以恢复信息流。

设计提示——为确保优化的收敛，思科建议，在各分布构建块中的每个分布交换机上对所有路由面向核心进行汇总。核心中的汇总路由可以防止核心节点将请求传播至网络其他区域，因此，有助于限制请求和收敛的时间。测试中还发现，利用汇总配置，EIGRP 请求生成、答复和路由插入所需的时间，对于任何丢失的相连路由来说，都不到 100 毫秒。

请求流程的快速完成还依赖于生成和接收信号的交换机处理 EIGRP 请求的能力。为确保可预测的收敛时间，还需要确保网络不会受到异常事件的攻击，如蠕虫、分布式拒绝服务（DdoS）攻击，以及可能在交换机上引发高 CPU 利用率的生成树环。

设计提示——为确保话音流量的优质收敛，思科建议，对 VLAN 号的分配进行规划，以便话音等大多数对丢失敏感的应用在各个物理接口上分配以最低的 VLAN 号，如表 5 所示。

表 5 VLAN 分配建议

VLAN 功能	VLAN 接口
有线话音 VLAN	7
无线话音 VLAN	57
有线数据 VLAN	107
无线组播 VLAN	157

并非所有汇聚在某一特定接口上的 VLAN 都同时收敛。Cisco IOS 会将发给路由流程（EIGRP/OSPF）的 VLAN 丢失通知限制在 100 毫秒/次。例如，假设每台接入交换机配置了 6 个 VLAN，一条上行链路故障时，第六个 VLAN 上的光纤流量随后会以 500 毫秒的速度收敛。

备用 HSRP 分布交换机上行链路光纤故障

上行收敛

由于所有流量都由主用交换机处理，因此，备用分布交换机故障不会影响上行流量。

下行收敛

对下行流量的影响等同于主用分布交换机上行链路故障的影响。核心交换机将继续向两台分布交换机转发流量，而下行数据路径的恢复则取决于从一台分布交换机到另一台的重新路由。

主用 HSRP 分布交换机故障

上行收敛

分布交换机全面故障后，上行流量的恢复机制就如同光纤故障案例的恢复机制完全一样。由于多个 HSRP 地址同时恢复，交换机故障的确会增加备用交换机的恢复压力。但是，在这些测试范围内，并未见到因为这一加大的处理开支而影响到恢复时间。

下行收敛

分布交换机故障后，下行流量的恢复是通过核心交换机的第三层等成本恢复实现的。如同上述第三层核心设计结果中所述，两台核心交换机拥有到使用两台分布交换机的所有接入子网的冗余路由。当两台分布交换机其中之一出现故障时，核心节点则通过另一台分布交换机转发所有下行流量，记录到的丢失周期不到 200 毫秒。

备用 HSRP 分布交换机故障

备用 HSRP 分布交换机故障对上行流量没有影响。下行流量故障同于上述主用 HSRP 交换机所述。

交换机间分布光纤故障

分布交换机间的第三层连接故障对上、下行流量都没有影响。该链路仅用于在发生上行链路故障时为分布构建块提供流量恢复功能。因为该链路的子网包含在汇总分布块地址范围内，因此，不会有 EIGRP 拓扑结构更新发送到核心。

恢复分析

配置 1 有下列特征：

- 缺省网关协议——HSRP
- 生成树版本——PVST+（每 VLAN 802.1d）
- IGP——EIGRP

表 6 总结了测试结果。

表 6 配置 1 恢复测试结果

恢复情况	上行恢复	下行恢复	恢复机制
主用 HSRP 上行链路光纤恢复	0 秒	0 秒	上行——无丢失 下行——无丢失
备用 HSRP 上行链路光纤恢复	0 秒	0 秒	上行——无丢失 下行——无丢失
主用 HSRP 分布交换机恢复	0 秒	可变（0-6 秒之间）	上行——无丢失 下行——L3 等成本路径和 ARP
备用 HSRP 分布交换机恢复	0 秒	可变（0-6 之间）	上行——无丢失 下行——L3 等成本路径和 ARP
交换机间分布光纤恢复	0 秒	0 秒	无丢失

主用 HSRP 上行链路光纤恢复

接入交换机和分布交换机间光纤连接的激活通常不会造成数据丢失。激活链路时，主分布交换机将触发根网桥的重新选择，并接管根的工作。该过程会导致生成树的逻辑收敛，但不会造成现有端口转发状态的变化，不会产生转发路径丢失。

除了生成树收敛外，一旦 HSRP 优先延迟计时器过期，主 HSRP 对等设备则会启动缺省网关接管机制。该过程在分布交换机之间是同步的，因此，不会导致分组丢失。话音和数据 VLAN 的生成树状态的转变还会引起被连接的路由插入路由表中。一旦连接的路由插入，交换机就开始向本地子网转发分组。

备用 HSRP 上行链路光纤恢复

对于主分布交换机，到备用分布交换机的上行链路光纤的激活并不影响现有话音或数据流。因此，根网桥和 HSRP 网关都无需恢复。交换机将为话音和数据 VLAN 插入连接的路由，并开始转发信息流。在上述情况中，这都不会明显地影响主用数据流。

主用 HSRP 分布交换机恢复

分布交换机的激活有可能对主用话音流的上、下行部分造成明显的影响。如果 HSRP 优先配置成主用网关，当主分布交换机、根网桥和 HSRP 高优先级激活时，可能会出现这样一段时间：该交换机接管缺省网关的工作，但还未建立到核心的 EIGRP 相邻关系。该交换机无法转发从接入子网接收到的流量，因而导致暂时路由黑洞的产生。

为避免这一问题，有许多方法。一个建议是配置一个足够大的 HSRP 优先延迟，以确保交换机上的所有线卡和接口，以及所有路由邻接项都能够激活。下面的配置示例即属此列：

```
interface Vlan20
description Voice VLAN for 3550
ip address 10.120.20.2 255.255.255.0
ip verify unicast source reachable-via any
ip helper-address 10.121.0.5
no ip redirects
ip pim query-interval 250 msec
ip pim sparse-mode
load-interval 30
standby 1 ip 10.120.20.1
standby 1 timers msec 250 msec 800
standby 1 priority 150
standby 1 preempt delay minimum 180 << Configure 3 minute delay
standby 1 authentication ese
```

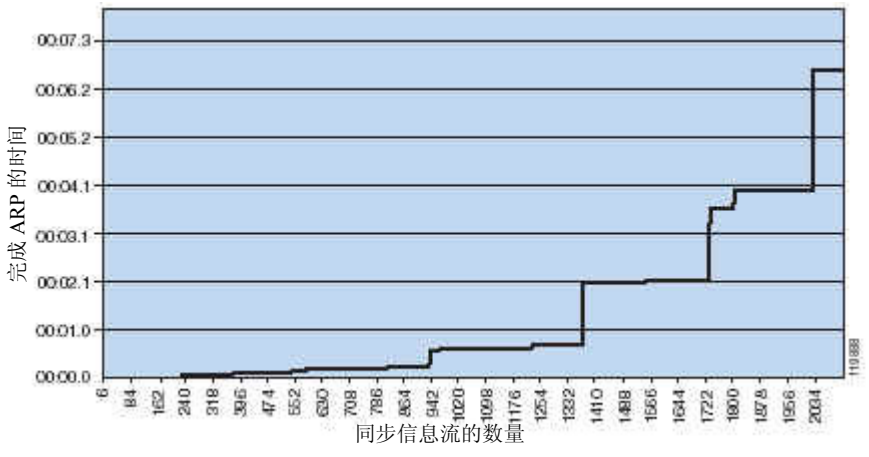
调整 HSRP 可以避免上行丢失问题；但是，高负载环境中的分布节点激活也可能造成下行流量丢失。一旦分布节点将分布块路由发送给核心交换机，它会立即开始接收所有相连子网的流量。

为转发该流量，分布交换机需要为每个信息流确定下一跳邻接项和地址解析协议（ARP）表项目。该过程的恢复时间取决于信息流数量、终端站点对 ARP 请求的响应能力，以及分布交换机的 ARP 抑制行为。

为防御 DoS 攻击，无论它是蓄意行为还是蠕虫的附带行为，所有 Cisco Catalyst 交换机都对 ARP 处理采取了速率抑制机制。虽然这些抑制功能可以在 DoS 攻击时保护交换机，但是，它们无法区分 DoS 流量造成的突发性洪泛和有效流量的突发性洪泛。在这两种情况下，交换机都会抑制 ARP 请求生成的速率。在拥有大量主用信息流的大型园区环境中，一台重启的分布交换机会遇到由内部 DoS 保护机制抑制的大量突发性 ARP 活动。

图 10 介绍了 DoS 保护对路由器性能的影响。

图 10
DoS 保护机制造成的影响



虽然这种行为会对流量造成短期的影响，但是，如果交换机在高峰活动期重启的话，DoS 保护机制提供的总体优势还是远高于损失的。

备用 HSRP 分布交换机恢复

如上所述，备用交换机的重启可能会对下行流量产生类似的潜在影响。对分布交换机计划性重启进行管理，可以消除各种潜在的影响。HSRP 和生成树状态不会因为重启而改变，因此，上行流量不受影响。

交换机间分布光纤恢复

分布交换机间链路的激活不会影响主用流量。由于分布块路由汇总到核心，新子网的激活不会导致拓扑结构更新。

配置 2 结果——HSRP，带 Rapid-PVST+的 EIGRP

本节探讨了下列主题：

故障分析

配置 2 拥有下列特征：

- 缺省网关协议——HSRP
- 生成树版本——Rapid-PVST+（每 VLAN 802.1w）
- IGP——EIGRP

测试结果总结

表 7 总结了测试结果。

表 7 配置 2 故障测试结果

故障情况	上行恢复	下行恢复	恢复机制
主用 HSRP 的上行链路光纤故障	900 毫秒	在 700-1100 毫秒之间	上行——HSRP 下行——EIGRP
备用 HSRP 的上行链路光纤故障	0 毫秒	在 700－1100 毫秒之间	上行——无丢失 下行——EIGRP
主用 HSRP 分布交换机故障	800 毫秒	200 毫秒	上行——HSRP 下行——L3 等成本路径
备用 HSRP 分布交换机故障	0 毫秒	200 毫秒	上行——无丢失 下行——L3 等成本路径
交换机间分布光纤故障	0 毫秒	0 毫秒	主用数据路径无丢失

转换到 802.1w 的影响

对于将 802.1w 作为生成树协议的分布块来说，链路或节点故障后的收敛特性与用 802.1d 配置时相同。在推荐设计中，所有第二层环已从网络中删除，该设计未利用 802.1w 相对于 802.1d 的收敛时间改进。如上所述，信息流恢复是通过 HSRP、EIGRP 或等成本路径故障转换完成的，恢复信息流时无需第二层收敛。

```
spanning-tree mode rapid-pvst <<< Enable 802.1w per VLAN spanning tree
spanning-tree loopguard default
no spanning-tree optimize bpdu transmission
spanning-tree extend system-id
spanning-tree vlan 2-7,20-51,102-149,202-207,220-249,900 priority 28672
```

恢复分析

配置 2 拥有下列特征：

- 缺省网关协议——HSRP
- 生成树版本——Rapid-PVST+（每 VLAN 802.1w）
- IGP——EIGRP

测试结果总结

表 8 总结了测试结果。

表 8 配置 2 恢复测试结果

恢复情况	上行恢复	下行恢复	恢复机制
主用 HSRP 上行链路光纤恢复	180 毫秒	180 毫秒	上行——802.1w 下行——802.1w
备用 HSRP 上行链路光纤恢复	0 毫秒	180 毫秒	上行——无丢失 下行——802.1w
主用 HSRP 分布交换机恢复	180 毫秒	可变 （在 180 毫秒-6 秒之间）	上行——802.1w 下行——802.1w，L3 等成本路径和 ARP
备用 HSRP 分布交换机恢复	180 毫秒	可变 （在 180 毫秒-6 秒之间）	上行——802.1w 下行——802.1w ，L3 等成本路径和 ARP
交换机间分布光纤恢复	0 秒	0 秒	主用数据路径无丢失

转换到 802.1w 的影响

在交换机或链路激活过程中，当用 802.1w 替代 802.1d 时，由于这两种协议在端口转换到主用转发状态的方式方面有所差异，会对话音流量产生少许可量度的影响。当主交换机（配置以适合的根网桥）重启或重新连接至接入交换机时，生成树拓扑结构就需要改变。新激活的根网桥开始以低于备用根的优先级传输网桥协议数据单元（BPDU）帧，并触发根网桥移动以及相关的第二层拓扑结构重新计算。

在 802.1d 拓扑结构中，过渡到新拓扑结构不会触发当前主用的端口转换为阻断状态，因为该拓扑结构是无环的，其结果是，流量持续传输至备用分布交换机，没有任何丢失。由于 HSRP 优先功能的使用，主用 HSRP 网关还会在备用交换机上保留 180 秒，因此，根交换机上的端口转换为转发状态并不会影响到主用流量。

在 802.1w 拓扑结构中，主交换机重新连接时，接入交换机上的上行链路端口和主交换机上的下行链路端口会进入指定的阻断状态。当接入交换机从新激活的主分布交换机接收到更好的根 BPDU 时，它首先阻断所有的非边缘指定端口（所有未被配置为 PortFast 的端口），然后开始拓扑结构转换。一旦所有非边缘端口阻断后，接入交换机就可以完成与新根网桥的协商，并安全地将相关上

行链路端口转换为转发状态，不会产生生成树环。新的根网桥变为转发状态后，接入交换机就可以在各个阻断的指定端口上完成相同的协商过程，必要的话，也可以将其转换为转发状态。

在转换到新的根端口期间，所有指定端口同步阻断，这在 802.1w 生成树拓扑结构计算中是必要步骤，但是，附带地也会导致到备用分布交换机的上行链路暂时阻断。因为备用分布交换机依然是主用 HSRP 网关，因而会造成所有上行流量的主用转发路径丢失。下行流量也会受到同样同步过程的影响。从核心经过备用分布交换机传送的流量也会在端口协商和转回到转发状态期间被阻断。

有关 802.1w 运行的更多信息，请参考《Cisco AVVID 网络基础设施：在园区网络中实施 802.1w 和 802.1s 》设计指南，网址为：
http://www.cisco.com/application/pdf/en/us/guest/tech/tk621/c1501/ccmigration_09186a0080174993.pdf

配置 3 结果——HSRP，带 Rapid-PVST+的 OSPF

本节探讨下列主题：

故障分析

配置 3 拥有下列特征：

- 缺省网关协议——HSRP
- 生成树版本——Rapid-PVST+（每 VLAN 802.1w）
- IGP——OSPF

测试结果总结

表 9 总结测试结果。

表 9 配置 3 故障测试结果

故障情况	上行恢复	下行恢复	恢复机制
主用 HSRP 的上行链路光纤故障	900 毫秒	1650 毫秒	上行——HSRP 下行——OSPF
备用 HSRP 的上行链路光纤故障	0 毫秒	1650 毫秒	上行——无丢失 下行——OSPF
主用 HSRP 分布交换机故障	800 毫秒	200 毫秒	上行——HSRP 下行——L3 等成本路径
备用 HSRP 分布交换机故障	0 毫秒	200 毫秒	上行——无丢失 下行——L3 等成本路径
交换机间分布光纤故障	0 毫秒	0 毫秒	主用数据路径无丢失

转换到 OSPF 的影响

在冗余园区设计中，只有一种情况才需要网络重新计算到任意目的地的新路由，这就是当接入交换机和分布交换机间的上行链路发生故障时。在发生所有其他故障时，转发路径的恢复依赖于缺省网关（HSRP）冗余或等成本路径重新路由。在接入交换机上行链路故障的情况下，因为没有冗余路径，网络需要启动重新路由。由于两种协议在运行方面的差异，该故障的 OSPF 收敛较 EIGRP 要差。

在接入上行链路故障时，分布交换机需要触发网络重新路由。如果网络被配置成将汇总路由传输到核心（即分布块配置为独立 OSPF 区域），那么针对该故障的重新路由就会选择连接两台分布交换机的链路。反之，如果园区整体配置成一个无路由汇总的区域，那么，一旦发现分布交换机下行路由丢失，就会在核心节点上重新路由。EIGRP 和 OSPF 的行为在这一点上是一致的：两种协议都会启动重新路由，无论是在汇总分布块中，还是在核心中（如果未配置汇总的话）。

EIGRP 完成这一重新路由所需时间主要取决于请求和应答处理的效率。在好的汇总设计中，请求过程处理的效率很高，可以有确定的收敛时间范围。

注：欲了解有关 EIGRP 收敛行为的更多信息，请参考上述故障情况 1 分析。

与 EIGRP 相比，由于需要交换 LSA 更新信息，计算最短路径优先（SPF）树，以及抑制这两种事件可能发生的速率，因此，OSPF 收敛所需时间被拉长了。

在 OSPF 配置中，实际重新路由并非由联络信息丢失和终止计时器过期引发的，而是由使接口关闭的 802.3z 或 802.3ae 远程故障检测引发的。在链路故障事件通知时，OSPF 流程会启动 LSA 传播抑制计时器。路由器并没有立即发送更新，而是等待一段时间（缺省为 0.5 秒），以便在传输前缓冲 LSA。LSA 传输后，接到新的 LSA 更新时，启动备用抑制计时器，即 SPF 计时器。

这个备用计时器过期后，执行 SPF 计算，如有新路由会被填入路由表中，创建相关的 CEF 转发项。要减少响应单一网络事件的 SPF 计算的次数，进而减少由于局部拓扑结构信息造成的不正确或不全面路由更新，必须采用 LSA 传播和 SPF 抑制计时器。

在测试过程中获得的 1600 毫秒收敛时间，为 OSPF 检测链路故障、传播 LSA、计算 SPF，以及插入新路由的全部时间，其中包括抑制计时器所花费的时间。这些测试的目的在于，网络中所有节点的 SPF 计时器时间降至 1 秒，如以下配置所示。

```
router ospf 100
 router-id 10.122.0.3
 log-adjacency-changes
 timers spf 1 1                  <<< Reduce SPF Timers to 1 second
 area 120 stub no-summary
 area 120 range 10.120.0.0 255.255.0.0
 network 10.120.0.0 0.0.255.255 area 120
 network 10.122.0.0 0.0.255.255 area 0
```

在 Cisco IOS 版本 12.2(17b)SXA 中，SPF 调整的配置语法随着亚秒 SPF 计时器的引入而改变。

```
router ospf 100
 router-id 10.122.0.3
 log-adjacency-changes
 timers throttle spf 1000 1000 1000          <<< One second SPF using 12.2(17b)SXA and later
IOS
```

```
area 120 stub no-summary
area 120 range 10.120.0.0 255.255.0.0
network 10.120.0.0 0.0.255.255 area 120
network 10.122.0.0 0.0.255.255 area 0
```

欲了解配置亚秒 SPF 抑制计时器的更多信息，请参考下面的 CCO 文档：
http://www.cisco.com/en/us/partner/products/sw/iosswrel/ps1838/products_feature_guide09186a0080134ad8.html

恢复分析

配置了有下列特征：

- 缺省网关协议——HSRP
- 生成树版本——Rapid-PVST+（每 VLAN 802.1w）
- IGP-OSPF

测试结果总结

表 10 总结了测试结果。

表 10 配置 3 恢复测试结果

恢复情况	上行恢复	下行恢复	恢复机制
主用 HSRP 上行链路光纤恢复	180 毫秒	180 毫秒	上行——802.1w 下行——802.1w
备用 HSRP 上行链路光纤恢复	0 毫秒	180 毫秒	上行——无丢失 下行——802.1w
主用 HSRP 分布交换机恢复	180 毫秒	可变 （在 180 毫秒-6 秒之间）	上行——802.1w 下行——802.1w，L3 等成本路径和 ARP
备用 HSRP 分布交换机恢复	180 毫秒	可变 （在 180 毫秒-6 秒之间）	上行——802.1w 下行——802.1w ，L3 等成本路径和 ARP
交换机间分布光纤恢复	0 秒	0 秒	主用数据路径无丢失

转换到 OSPF 的影响

在冗余园区设计中激活一条链路或重启一个节点时，网络的行为一致，与所用的路由协议（OSPF 或 EIGRP）无关。流量可能丢失的情况与上述讨论的配置测试 1 和 2 的情况相同。在根网桥拓扑结构改变期间，802.1w 同步流程会造成少量可察觉丢失，而在上述条件下，ARP DoS 保护特性也会造成相同的流量中断。

配置 4 结果——GLBP，带 Rapid-PVST+的 EIGRP

本节探讨了下列主题：

配置 4 有下列特征：

- 缺省网关协议——GLBP
- 生成树版本——Rapid-PVST+（每 VLAN 802.1w）
- IGP-EIGRP

测试结果总结

表 11 总结了测试结果。

表 11 配置 4 故障测试结果

故障情况	上行恢复	下行恢复	恢复机制
主用 GLBP 的上行链路光纤故障	900 毫秒	800 毫秒	上行——GLBP 下行——EIGRP
备用 GLBP 的上行链路光纤故障	900 毫秒	800 毫秒	上行——GLBP 下行——EIGRP
主用 GLBP 分布交换机故障	800 毫秒	200 毫秒	上行——GLBP 下行——L3 等成本路径
备用 GLBP 分布交换机故障	800 毫秒	200 毫秒	上行——GLBP 下行——L3 等成本路径
交换机间分布光纤故障	0 毫秒	0 毫秒	主用数据路径无丢失

转换到 GLBP 的影响

当采用相同的计时器配置时，HSRP 和 GLBP 的光纤和/或节点故障的最长恢复时间相同。在这两种情况中，故障转换机制取决于相邻设备丢失通知和虚拟 MAC 地址接管。在 GLBP 实施中发现有两点不同：

- 由于在单一部件故障期间，只有一半流量丢失，因此 GLBP 的平均收敛时间较短。
- 上行链路或分布交换机丢失，会使 GLBP 配置遭遇失败

动态 GLBP 负载均衡算法的性质确保了，只有一半终端站点在任意时间将每台分布交换机用作缺省网关。由于一半工作站遭遇上行流量丢失，一半工作站遭遇下行流量丢失，因此，使用 GLBP 并未使最差收敛情况得到改进。根据统计数字，受到上行或下行流量影响的站点间没有关联，而在最差情况下，每个终端站点在网络收敛过程中都会出现中断。

注：测试中选用的 GLBP 计时器是为了弥补下行收敛时间。减少 GLBP 计时器会改进上行收敛，但不会影响返回路径的流量。

恢复分析

配置 4 拥有下列特征：

- 缺省网关协议——GLBP
- 生成树版本——Rapid-PVST+（每 VLAN 802.1w）
- IGP——EIGRP

测试结果总结

表 12 总结了测试结果。

表 12 配置 4 恢复测试结果

恢复情况	上行恢复	下行恢复	恢复机制
主用 GLBP 上行链路光纤恢复	180 毫秒	180 毫秒	上行——802.1w 下行——802.1w
备用 GLBP 上行链路光纤恢复	0 毫秒	180 毫秒	上行——无丢失 下行——802.1w
主用 GLBP 分布交换机恢复	180 毫秒	可变 （在 180 毫秒-6 秒之间）	上行——802.1w 下行——802.1w，L3 等成本路径和 ARP
备用 GLBP 分布交换机恢复	180 毫秒	可变 （在 180 毫秒-6 秒之间）	上行——802.1w 下行——802.1w ，L3 等成本路径和 ARP
交换机间分布光纤恢复	0 秒	0 秒	主用数据路径无丢失

转换到 GLBP 的影响

改变缺省网关冗余协议对设备激活没有影响。网络中的丢失会造成分组丢失，这或者是由于 802.1w 同步过程，或者由于在流量极高的情况下，ARP DoS 保护机制影响了恢复。

注： 欲了解这些行为的更多信息，请参考上面的章节。

配置 5 结果——GLBP,EIGRP，带第二层环的 Rapid-PVST+

本节探讨了下列主题：

故障分析

配置 5 拥有下列特征：

- 缺省网关协议——GLBP
- 生成树版本——环形 Rapid-PVST+（每 VLAN 802.1w）
- IGP——EIGRP

测试结果总结

表 13 总结了测试结果。

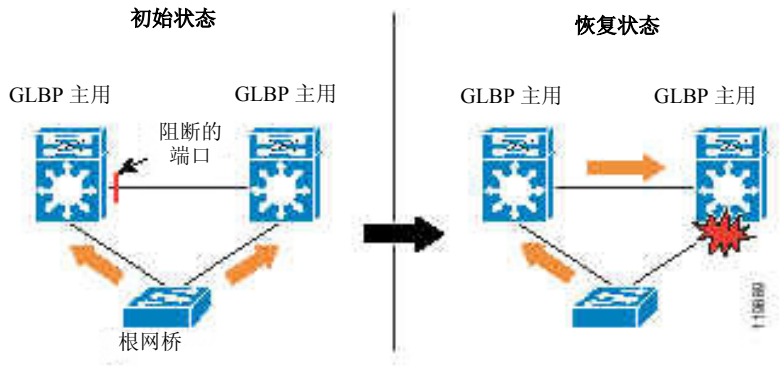
表 13 配置 5 故障测试结果

故障情况	上行恢复	下行恢复	恢复机制
主用 GLBP 的上行链路光纤故障	400 毫秒	400 毫秒	上行——802.1w 下行——802.1w
备用 GLBP 的上行链路光纤故障	400 毫秒	400 毫秒	上行——802.1w 下行——802.1w
主用 GLBP 分布交换机故障	800 毫秒	200 毫秒	上行——GLBP 下行——L3 等成本路径
备用 GLBP 分布交换机故障	800 毫秒	200 毫秒	上行——GLBP 下行——L3 等成本路径
交换机间分布光纤故障	0 毫秒	0 毫秒	主用数据路径无丢失

在接入 VLAN 中采用第二层环的影响

第二层生成树环将 VLAN 扩展到两台分布交换机间的干线，与无环设计相比，在发生单一光纤故障时缩短了上行和下行流量的恢复时间。在第二层环形配置中，对拓扑结构链路丢失的 802.1w 恢复无需 HSRP 和 EIGRP 收敛，如图 11 所示。

图 11
第二层环形配置中的 802.1w 恢复



在 802.1w 推出前，HSRP 和 EIGRP 能够以远快于 802.1d 检测故障链路和取消冗余链路的阻断的速度收敛。通过将环应用于网络，我们改变了设计的下列两个特性：

- 802.1w 能够在所配置的对等终止时间前恢复两个 GLBP 对等设备间的连接。配置以 800 毫秒终止时间和 802.1w 的 GLBP 可以在不到 400 毫秒的时间内恢复连接。
- 因为每台分布交换机在各个 VLAN 有多个端口，单一下行端口丢失不会引发自动关闭相应的交换式虚拟接口（SVI）。因此，EIGRP 不会启动路由收敛，网络在恢复受影响的流量前，要等待直至 802.1w 恢复两台交换机间的第二层连接。

采用环形第二层拓扑结构对分布交换机故障的网络收敛时间没有影响。在这种情况下，冗余第二层路径不再存在，网络要依靠 GLBP 恢复缺省网关，以及依赖核心的等成本路径恢复下行流量。

注： 目前并不建议将第二层生成树环应用于网络作为最佳实践。在利用该配置实施设计前，思科郑重提示您，生成树环存在潜在的网络停运风险，必须与所带来的流量恢复时间的少许缩短优势进行权衡。虽然有多种特性，当使用正确时，可以消除使用环形第二层拓扑结构的大部分风险（环防护、UDLD 和 BPDU 防护），但如果没有针对扩展第二层子网的应用或商业要求，思科建议，HA 园区设计应避免第二层环。建议希望提供最短收敛时间的网络工程师，考虑使用第三层接入设计，而非实施这种第二层设计。

恢复分析

配置 5 拥有下列特征：

- 缺省网关协议——GLBP
- 生成树版本——环形 Rapid-PVST+（每 VLAN 802.1w）
- IGP——EIGRP

测试结果总结

表 14 总结了测试结果。

表 14 配置 5 恢复测试结果

恢复情况	上行恢复	下行恢复	恢复机制
主用 GLBP 上行链路光纤恢复	180 毫秒	180 毫秒	上行——802.1w 下行——802.1w
备用 GLBP 上行链路光纤恢复	0 毫秒	180 毫秒	上行——无丢失 下行——802.1w
主用 GLBP 分布交换机恢复	180 毫秒	可变 （在 180 毫秒-6 秒之间）	上行——802.1w 下行——802.1w，L3 等成本路径和 ARP
备用 GLBP 分布交换机恢复	180 毫秒	可变 （在 180 毫秒-6 秒之间）	上行——802.1w 下行——802.1w ，L3 等成本路径和 ARP
交换机间分布光纤恢复	0 秒	0 秒	主用数据路径无丢失

在接入 VLAN 中采用第二层环的影响

将生成树环应用于接入 VLAN 不会改变设备激活对于主用流量的影响。网络中的丢失依然会造成分组丢失，这或者是由于 802.1w 同步过程，或者由于在流量极高的情况下，ARP DoS 保护机制影响了恢复。

注: 欲了解有关这些行为的具体信息，请参考上述章节。

第三层分布式和第三层可路由接入的收敛——结果与分析

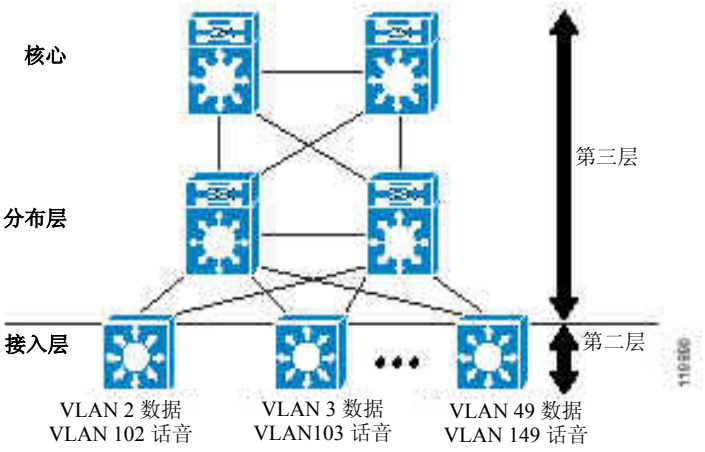
本节包括下列主题：

- 第三层路由接入概述，第 33 页
- EIGRP 结果，第 34 页
- OSPF 结果，第 38 页

第三层路由接入概述

标准第二层接入分布构建块设计的替代方案有：路由式接入或第三层到边缘部署，如图 12 所示。

图 12 第三层到接入分布构建块



在该设计中，接入交换机被配置成加入园区路由协议的完全第三层路由节点。第二/三层的边界从分布交换机迁移到园区网边缘的接入交换机。本地话音和数据 VLAN 流量被路由，而并非上行桥联到分布交换机。

该设计初看起来与标准分布构建块极为不同，但在许多方面却非常相似。在这两种设计中，每台接入交换机都配置以专用的话音和数据子网（VLAN）。路由节点间的所有链路都配置成点到点模式。它们的主要差别在于缺省网关的位置。

在传统的设计中，缺省网关驻留在分布节点中，接入交换机和分布交换机间的链路经过配置，将数据和话音 VLAN 扩展到分布式路由器。在第三层接入设计中，缺省网关驻留在接入交换机上，而上行链路则配置成采用/30 或/31 编址的专用点到点子网。

注： 欲了解第三层接入设计和推荐的最佳实践的具体信息，请参考《设计高可用性园区网络》和相关的《高可用性园区第三层接入设计指南》。

EIGRP 结果

本节探讨了下列主题：

EIGRP 故障结果

在第三层接入设计中，要考虑的故障情况较少。通过将第二/三层边界从分布交换机迁移到接入交换机，某些故障情况不会再出现。由于缺省网关功能现在指定给边缘交换机，因此，不再需要冗余缺省网关机制。HSRP/GLBP 调整和收敛时间不会影响园区设计决策。生成树实例在单一交换机中运行，不影响收敛时间。路由接入园区完全取决于园区交换机的第三层收敛行为和所部署的路由协议的收敛行为。

在接入层中运行第三层功能还可以实现所有上行和下行流量的动态负载均衡。从接入到分布层的等成本路由会在两条上行链路间分配发往核心的上行负载。各个核心节点上的等成本路径将在两台分布交换机间分配负载，从而使从分布到接入层的下行返回路径形成对称负载。这种负载均衡或对称行为意味着，在冗余第三层园区设计中，任意特定节点或链路故障只影响最多一半的网络流量。由于链路或节点故障造成的任意单一路径故障，只会影响使用该路径的流量。

注： 对称负载是为方便统计的测量方法，并不考虑网络传输的字节数，只关注流量数。在只有 G.711 话音流和站点磁盘备份这两个信息流的极端情况下，按 bps 测量的每条上行链路的负载是不等的。但是，这些流量在两条上行链路上都具有等同的转发可能性。随着流量数增加，测量到的平均负载，无论是流量数还是字节数，都会逐渐缩小差距，趋向于相同。

在第三层接入设计中，您需在评估收敛行为时考虑以下三种情况：

- 接入和分布交换机间上行链路光纤故障
- 分布交换机故障
- 分布交换机间光纤故障

因为第三层接入设计具有动态负载均衡特性，第一种情况和第二种情况表示分布节点或从接入交换机到这些节点的上行链路发生了故障。

表 15 总结了这三种故障情况的测试结果。

表 15 三种故障情况的测试结果

故障情况	上行恢复	下行恢复	恢复机制
上行链路光纤故障	150 毫秒	200 毫秒	上行——L3 等成本路径 下行——EIGRP
分布交换机故障	150 毫秒	200 毫秒	上行——L3 等成本路径 下行——EIGRP
交换机间分布光纤故障	0 毫秒	0 毫秒	主用数据路径无丢失

接入和分布交换机间上行链路光纤故障

在路由接入设计中，从接入交换机到核心的上行信息流的收敛时间取决于分布交换机链路丢失的检测。检测出链路故障后，交换机处理链路中断事件，删除所有与故障接口相关的路由和 CEF 硬件转发项。在推荐配置中，使用从两个分布交换机发出的等成本路由，随后通过其余硬件 CEF 项转发所有上行信息流。在等成本路由配置中，EIGRP 仍有到所有上行目的地的有效路由，无需软件路由重新计算。

如上面有关路由核心设计的讨论中所述，需确保所有上行链路接口上都启用 802.3z 或 802.3ae 链路协商，以便远程故障检测机制能检测出因为光纤或上行节点故障而导致的链路丢失。禁用链路协商会延长上行和下行信息流的收敛时间。

发生在接入交换机到分布交换机间光纤故障时的下行恢复时间取决于 EIGRP 重路由。在冗余第三层接入园区中，除了一种情况，您无需在其他情况下进行路由协议收敛。每个分布交换机有一条路径通往用于每个接入交换机的话音和数据子网。在此路径丢失时，分布交换机转换为 EIGRP 主用状态，用于与那一链路相关的接入网络/路由。分布节点向核心交换机和其分布对等设备请求一条替代路径。

为确保优化收敛，思科建议汇总每个分布构建块中从每个分布交换机上行到核心的所有路由。核心上的汇总路由可防止核心节点向网络的其他部分传播请求，因此可限制请求和收敛时间。请求过程能否快速完成也取决于发出和接收信息的交换机处理 EIGRP 请求的能力。

应确保网络不受可能在交换机上造成高 CPU 流量的异常事件（蠕虫、DDoS 攻击、生成树环等）的影响，对于确保可预测收敛时间也是必不可少的。

分布交换机故障

在分布节点故障时，从接入交换机到核心的上行流量的恢复依然取决于分布交换机丢失链路的检测。交换机间的光纤丢失或分布交换机故障被接入交换机视为同等事件。其行为和异议如同上所述。

分布交换机故障时的下行收敛取决于核心交换机的等成本路径故障行为。每台核心交换机拥有两条到分布块网络的等成本路由。分布交换机之一发生故障时，核心交换机会删除无效的路由和相关的硬件 CEF 输入项。EIGRP 还有一条到分布块网络的有效路由，并不会被其他路由激活，不执行路由重新计算。

分布交换机间的光纤故障

在正常运行条件下，分布交换机间的光纤路径不传输话音和应用数据流量。在接入到分布交换机链路故障时，它为应用和话音流量提供了一条备用路径（上述配置 1 所示）。因此，单一故障时，该链路丢失不会影响话音和应用数据流量。

EIGRP 恢复结果

如同上述第三层核心设计中的部件恢复检测所观察到的那样，第三层接入设计中的链路和节点恢复是一个非常稳定和可预测的过程。在路由表变更反映出新链路激活以前，不会执行第三层转发。直到新激活的链路完成了两个方向的邻接设备搜索和协商，才会更新路由表。

在所有交换机都已确定能够成功地转发话音和数据流量以前，不会改变转发表。端到端 IGP 的使用避免了与第三层静态路由设计有关的问题，该设计不是根据确认的有效网络路径，而仅仅根据激活的链路转发流量。EIGRP 接入设计避免了静态路由实施带来的黑洞问题。

链路和设备恢复的收敛情况与发生下列故障时的收敛情况（见表 16）相同：

- 接入到分布交换机上行链路光纤恢复
- 分布交换机恢复
- 交换机间光纤恢复

表 16 EIGRP 恢复案例

恢复情况	上行恢复	下行恢复	恢复机制
上行链路光纤恢复	0 秒	0 秒	主用数据路径无丢失
分布交换机恢复	0 秒	0 秒	主用数据路径无丢失
交换机间分布光纤恢复	0 秒	0 秒	主用数据路径无丢失

冗余设计中的链路或节点恢复过程中，主用转发路径无丢失。备用链路/节点激活将导致 EIGRP 邻接设备搜索/路由建立，从而在路由表中生成一条备用等成本路由或导致现有路由被一条成本更低的新路由取代。新路由建立后，每台交换机更新其反映新路由的硬件 CEF，并通过新路径转发所有现有流量和新的流量。在路由协议更新过程中，交换机继续使用现有硬件转发项，并且不会由于新路由的插入而丢失任何数据。

接入和分布交换机间的链路恢复导致备用等成本路由添加到上行网络接入交换机。分布交换机可以利用恢复的光纤了解到更好的路由，并用新的路由替代旧的路由和 CEF 项。由于在其到核心的上行链路上配置了汇总边界，因此，分布交换机不会将新路由信息传播至核心。此上行链路激活导致的影响在 EIGRP 设计中被降至了最低限度。

分布交换机的恢复会在接入和核心交换机中各插入一条备用等成本路径。核心交换机获得一条到分布块汇总地址范围的备用路由，接入交换机获得一条备用缺省路由。在这两种情况下，备用等

成本路由的插入不会产生影响。核心交换机不会将任何更新的路由信息传播到其他相连的交换机，因为它并不是获得了一条更好的路径，而只是到同一网络、成本相同的备用拓扑结构和路由项。此上行链路激活导致的影响在 EIGRP 设计中被降至了最低限度。

在新路由插入的转换过程中，无论它是更好的路径路由还是备用等成本路由，在高度超额配置的网络中，在新路径上发送的现有信息流分组可能在抵达时出现错序。这种情况只有在原始路径上的负载遭遇严重拥塞，导致持续延迟时，才会出现。

在利用高度超额配置（最差情况）负载进行测试的过程中，结果发现，主用话音流中单一分组丢失不到 0.003%。极低的分组丢失水平和由于激活备用链路造成的较低相关抖动，以及话音流转发路径的动态变更，对所记录的测试信息流 MOS 分值不产生可度量的影响。冗余第三层园区设计中的新链路或节点的激活，不对现有信息流的运行产生影响。

OSPF 结果

- OSPF 故障结果，第 38 页
- OSPF 恢复结果，第 39 页

OSPF 故障结果

在路由第三层接入设计中，IGP 选择的变化不会显著地改变网络设计，也不会改变需要评估的故障情况。对于 EIGRP 进行的三项故障分析，同样需要应用于 OSPF：

- 接入到分布交换机上行链路光纤故障
- 分布交换机故障
- 交换机间分布光纤故障

因为发生故障时路由协议行为的差异，从 EIGRP 转换到 OSPF 会影响第一种故障情况下的网络收敛行为（见表 17）。

表 17 OSPF 故障结果

故障情况	上行恢复	下行恢复	恢复机制
上行链路光纤故障	150 毫秒	1650 毫秒	上行——L3 等成本路径 下行——OSPF
分布交换机故障	150 毫秒	200 毫秒	上行——L3 等成本路径 下行——OSPF
交换机间分布光纤故障	0 毫秒	0 毫秒	主用数据路径无丢失

接入和分布交换机间上行链路光纤故障

在 OSPF 环境中，从接入交换机到核心的上行流量恢复类似于 EIGRP，即主要取决于分布交换机链路丢失检测。链路故障检测过程中，交换机将处理链路中断事件，删除与故障接口有关的所有路由和 CEF 硬件转发项。

为确保快速的上行恢复，适用于 EIGRP 网络的基本设计建议，也同样适用于 OSPF：从分布到接入交换机采用等成本路由，并确保实施链路协商。

接入到分布交换机光纤故障的下行恢复时间取决于 OSPF 重路由。在此路径丢失时，分布交换机生成一个 LSA，表示在等待一个 LSA 抑制计时器间隔后链路状态的改变。一旦交换机的流量溢出了新 LSA，它会根据 SPF 抑制计时器而等待一段已配置好的时间，等待从该区域中其他交换机处收到 LSA。此计时器过期后，执行 SPF 计算，插入新路由和相关硬件 CEF 项。

由于拓扑结构变更的重新计算对于处理器和内存的依赖性更强，并且经常涉及区域中的所有路由器，因此，OSPF 实施了旨在减少拓扑结构变更次数和频率的 LSA 和 SPF 抑制机制。

收敛和恢复下行流量所需的时间长度主要取决于 OSPF 抑制计时器的配置，还依赖于交换机处理 LSA 和完成 SPF 计算的能力。确保网络不受可能在交换机上造成高 CPU 利用率的异常事件（蠕虫、DDoS 攻击、生成树环等）的影响，对于确保可预测收敛时间也是必不可少的。

分布交换机故障

在分布节点故障时，从接入交换机到核心的上行流量的恢复依然取决于分布交换机丢失链路的检测。交换机间的光纤丢失或分布交换机故障被接入交换机视为同等事件。其行为如同上所述。

分布交换机故障时的下行收敛取决于核心交换机等成本路径故障行为。每台核心交换机拥有两条到分布块网络的等成本路由。分布交换机中的一台故障时，核心交换机会删除无效的路由和相关的硬件 CEF 输入项。由于链路丢失，核心路由器和对等分布交换机上的 OSPF 会启动 LSA 洪泛，OSPF 区域的所有交换机都必须完成 SPF 计算。在冗余园区设计中，除了三台直接连接到故障分布节点的交换机外，不会导致区域中的交换机路由变化。

分布交换机间光纤故障

在正常运行条件下，分布交换机间的光纤路径不传输话音和应用数据流量。在接入到分布交换机链路故障时，它为应用和话音流量提供了一条备用路径（上述配置 1 所示）。因此，即使导致 OSPF 区域内所有交换机出现 LSA 洪泛和 SPF 重新计算，在单一故障时，该链路的丢失也不会影响话音和应用数据流量。

OSPF 恢复结果

路由接入 OSPF 设计中的链路和节点恢复是十分容易预测的，只有一种情况可能会出现显著的意外。通常，除非为反映新链路激活而进行路由表变更，否则不会执行第三层转发。直到新激活的链路完成了两个方向的邻接设备搜索和协商，才会更新路由表。除非所有交换机都已确定能够成功地转发话音和数据流量，否则不会改变转发表。此故障的例外情况在恢复情况二——分布交换机的恢复中进行了阐述。

- 接入到分布交换机上行链路光纤恢复
- 分布交换机恢复
- 分布交换机间光纤恢复

表 18 总结了测试结果。

表 18 OSPF 恢复结果

恢复情况	上行恢复	下行恢复	恢复机制
上行链路光纤恢复	0 秒	0 秒	主用数据路径无丢失
分布交换机恢复	0-45 秒	0 秒	主用数据路径无丢失
交换机间分布光纤恢复	0 秒	0 秒	主用数据路径无丢失

接入到分布交换机上行链路光纤恢复

冗余设计中的链路或节点恢复过程中，主用转发路径无丢失。备用链路/节点激活将导致 OSPF 邻接设备搜索/路由建立，从而在路由表中生成一条备用等成本路由或导致现有路由被一条成本更低的新路由取代。新路由建立后，每台交换机更新其反映新路由的硬件 CEF，并通过新路径转发所有现有流量和新的流量。在路由协议更新过程中，交换机继续使用现有硬件转发项，并且不会由于新路由的插入而丢失任何数据。

接入和分布交换机间的链路恢复导致备用等成本路由添加到上行网络接入交换机。分布交换机可以利用恢复的光纤了解更好的路由，用新的路由替代旧的路由和 CEF 项。由于在其到核心的上行链路上配置了汇总边界，因此，分布交换机不会将新路由信息传播至核心。上行链路激活导致的影响在 OSPF 设计中被降至了最低限度。

在新路由插入的转换过程中，无论它是更好的路径路由还是备用等成本路由，在高度超额配置的网络中，在新路径上发送的现有信息流分组可能在抵达时出现错序。这种情况只有在原始路径上的负载遭遇严重拥塞，导致持续延迟时，才会出现。在利用高度超额配置（最差情况）负载进行测试的过程中，结果发现，主用话音流中单一分组丢失不到 0.003%。

极低的分组丢失水平和由于激活备用链路造成的较低相关抖动，以及话音流转发路径的动态变更，对所记录的测试信息流 MOS 分值不产生可度量的影响。冗余第三层园区设计中的新链路或节点的激活，不对现有信息流的运行产生影响。

分布交换机恢复

分布交换机恢复过程中的行为，根据分布交换机在 OSPF 区域分级中的作用而有所差异。此处的探讨是基于下列 OSPF 设计：

- 分布块作为 OSPF 区域
- 分布域作为完全剩余域运行
- 分布交换机是区域边界路由器（ABR）

分布交换机的激活为核心收敛，包括添加备用 ABR，以及被广播到区域 0 的相关汇总路由。该过程不会丢失数据。域内收敛会遇到某些与完全剩余域有关的问题。当激活分布 ABR 时，一旦建立多个邻接关系，它就会将缺省路由广播到下行相邻接入交换机中。缺省广播并非根据与核心的邻接关系而建立，而是根据与包括接入交换机在内的其他交换机的邻接关系而建立。

因此，分布交换机有可能在可向核心路由数据前，就将缺省路由广播到接入交换机。接入交换机有两个缺省路由：一个源于转发到核心的现有分布交换机，另一个源于不接入核心的新激活的交换机。分布交换机流量黑洞的时间长短取决于多种因素，包括线卡启动顺序、CPU 和其他交换机负载。

注意 欲了解这些 OSPF 设计选择的具体信息和理由，请参考高可用性园区和园区第三层接入设计指南。

分布交换机间光纤恢复

在正常运行条件下，分布交换机间的光纤路径不传输话音和应用数据流量。在发生接入到分布交换机链路故障时，它为应用和话音流量提供了一条备用路径（上述配置 1 所示）。该链路的激活会导致 OSPF 区域内所有交换机的 LSA 洪泛和 SPF 重新计算，但不会造成话音和应用数据流量的丢失。

测试所用配置

本节探讨了下列主题：

核心交换机配置

所有测试均使用下列核心交换机配置。借助下列分级设计原则，可以在分布块中调整配置，而无需调整核心配置，因此，其余的测试不必进行变更。

测试设计遵循最佳实践不在核心实施复杂策略管理的建议， 只有一种情况例外。核心交换机被配置成组播路由点。这是一种组播设计选项，作为代表性案例。

注意 欲了解组播设计选项的更多信息，请参考《组播设计指南》，网址为：http://www.cisco.com/warp/public/779/largeent/it/ese/srnd.html

本节探讨了下列主题：

核心交换机配置（EIGRP）

```
key chain eigrp
 key 100
  key-string 7 01161501
!
! Enabled spanning tree as a fail-safe practice
spanning-tree mode rapid-pvst
!
redundancy
 mode sso
 main-cpu
  auto-sync running-config
  auto-sync standard
!
```

```
! Configure necessary loopback interfaces to support Multicast MSDP and Anycast for
! RP redundancy
interface Loopback0
 description MSDP PEER INT
 ip address 10.122.10.2 255.255.255.255
!
interface Loopback1
 description ANYCAST RP ADDRESS
 ip address 10.122.100.1 255.255.255.255
!
interface Loopback2
 description Garbage-CAN RP
 ip address 2.2.2.2 255.255.255.255
!
! Configure point to point links to Distribution switches
interface TenGigabitEthernet3/1
 description 10GigE to Distribution 1
! Use of /31 addressing on point to point links optimizes use of IP address space in
! the campus
 ip address 10.122.0.27 255.255.255.254
 ip pim sparse-mode
! Reduce EIGRP hello and hold timers to 1 and 3 seconds. In a point-point L3 campus
! design the EIGRP timers are not the primary mechanism used for link and node
! failure detection. They are intended to provide a fail-safe mechanism only.
 ip hello-interval eigrp 100 1
 ip hold-time eigrp 100 3
 ip authentication mode eigrp 100 md5
 ip authentication key-chain eigrp 100 eigrp
 load-interval 30
! Reduce carrier delay to 0. Tuning carrier delay no longer has an impact on GigE and
! 10GigE interfaces but is recommended to be configured as a best practice for network
! operational consistency
 carrier-delay msec 0
! Configure trust DSCP to provide for maximum granularity of internal QoS queuing
 mls qos trust dscp
!
router eigrp 100
! Passive all interfaces not intended to form EIGRP neighbors
 passive-interface Loopback0
 passive-interface Loopback1
 passive-interface Loopback2
 network 10.0.0.0
 no auto-summary
```

```
! Explicitly configure the EIGRP router id as a best practice when using Anycast and/or
! any identical loopback address on multiple routers.
 eigrp router-id 10.122.0.1
!
! Multicast route point and MSDP configuration.
! For a detailed explanation on the specifics of the configuration below please see
! the campus chapter of the multicast design guides.
ip pim rp-address 2.2.2.2
ip pim rp-address 10.122.100.1 GOOD-IPMC override
ip pim accept-register list PERMIT-SOURCES
ip msdp peer 10.122.10.1 connect-source Loopback0
ip msdp description 10.122.10.1 ANYCAST-PEER-6k-core-left
ip msdp cache-sa-state
ip msdp originator-id Loopback0
!
ip access-list standard GOOD-IPMC
 permit 224.0.1.39
 permit 224.0.1.40
 permit 239.192.240.0 0.0.3.255
 permit 239.192.248.0 0.0.3.255
!
ip access-list extended PERMIT-SOURCES
 permit ip 10.121.0.0 0.0.255.255 239.192.240.0 0.0.3.255
 permit ip 10.121.0.0 0.0.255.255 239.192.248.0 0.0.3.255
```

核心交换机配置（OSPF）

```
!
! Enabled spanning tree as a fail-safe practice
spanning-tree mode rapid-pvst
!
redundancy
 mode sso
 main-cpu
  auto-sync running-config
  auto-sync standard
!
! Configure necessary loopback interfaces to support Multicast MSDP and Anycast for
! RP redundancy
interface Loopback0
 description MSDP PEER INT
 ip address 10.122.10.2 255.255.255.255
!
interface Loopback1
```

```
 description ANYCAST RP ADDRESS
 ip address 10.122.100.1 255.255.255.255
!
interface Loopback2
 description Garbage-CAN RP
 ip address 2.2.2.2 255.255.255.255
!
! Configure point to point links to Distribution switches
interface TenGigabitEthernet3/1
 description 10GigE to Distribution 1
! Use of /31 addressing on point to point links optimizes use of IP address space in
! the campus
 ip address 10.122.0.25 255.255.255.254
 ip pim sparse-mode
! Reduce OSPF hello and dead timers to 1 and 3 seconds. In a point-point L3 campus
! design the OSPF timers are not the primary mechanism used for link and node
! failure detection. They are intended to provide a fail-safe mechanism only.
 ip ospf hello-interval 1
 ip ospf dead-interval 3
 load-interval 30
! Reduce carrier delay to 0. Tuning carrier delay no longer has an impact on GigE and
! 10GigE interfaces but is recommended to be configured as a best practice for network
! operational consistency
 carrier-delay msec 0
! Configure trust DSCP to provide for maximum granularity of internal QoS queuing
 mls qos trust dscp
!
router ospf 100
! Explicitly configure the OSPF router id as a best practice when using Anycast and/or
! any identical loopback address on multiple routers.
 router-id 10.122.0.1
 log-adjacency-changes
! Tune the SPF throttle timers down from the defaults. Please refer to the HA Campus
! Design Guides for details on specific tuning recommendations.
 timers spf 1 1
 passive-interface Loopback0
 passive-interface Loopback1
 passive-interface Loopback2
 network 10.122.0.0 0.0.255.255 area 0.0.0.0
!
! Multicast route point and MSDP configuration.
! For a detailed explanation on the specifics of the configuration below please see
! the campus chapter of the multicast design guides.
```

```
ip pim rp-address 2.2.2.2
ip pim rp-address 10.122.100.1 GOOD-IPMC override
ip pim accept-register list PERMIT-SOURCES
ip msdp peer 10.122.10.1 connect-source Loopback0
ip msdp description 10.122.10.1 ANYCAST-PEER-6k-core-left
ip msdp cache-sa-state
ip msdp originator-id Loopback0
!
ip access-list standard GOOD-IPMC
 permit 224.0.1.39
 permit 224.0.1.40
 permit 239.192.240.0 0.0.3.255
 permit 239.192.248.0 0.0.3.255
!
ip access-list extended PERMIT-SOURCES
 permit ip 10.121.0.0 0.0.255.255 239.192.240.0 0.0.3.255
 permit ip 10.121.0.0 0.0.255.255 239.192.248.0 0.0.3.255
```

第二层接入和分布块的交换机配置

本节包括在上述第一种情况中使用的配置。测试内容如下：

- 缺省网关协议——HSRP
- 生成树版本——PVST+（每 VLAN 802.1d）
- IGP——EIGRP

其他测试情况中基本配置的所有变化都记录在上述结果中。仅包括接口和 VLAN 配置示例，以用作参考。欲了解有关推荐配置的具体信息，请参考《设计高可用性园区网络》。

本节探讨了下列主题：

分布交换机 1——根网桥和主 HSRP

```
! Use vtp transparent mode, configure all VLANs explicitly
vtp domain campus-test
vtp mode transparent
! Enable UDLD aggressive mode as a fail safe mechanism
udld aggressive

! Enable 802.1d per VLAN spanning tree enhancements.
spanning-tree mode pvst
spanning-tree loopguard default
```

```
no spanning-tree optimize bpdu transmission
spanning-tree extend system-id
spanning-tree uplinkfast
spanning-tree backbonefast
! Configure root bridge priority. Note use of extended system-id
spanning-tree vlan 2-7,20-51,102-149,202-207,220-249 priority 24576
!
! Define unique voice and data vlans for each access switch
vlan 4
 name Data_VLAN-4507-SupII+
vlan 7
 name Data_VLAN-6500-Sup2-CatOS
vlan 104
 name Voice_VLAN-4507-SupII+
vlan 107
 name Voice_VLAN-6500-Sup2
vlan 204
 name Trunk_VLAN-4507-SupII+
vlan 207
 name Trunk_VLAN-6500-Sup2-CatOS
!
! Define a local loopback address to provide a sink hole route point for
! invalid multicast groups
interface Loopback2
 description Garbage-CAN RP
 ip address 2.2.2.2 255.255.255.255
!
! Configure point to point Layer 3 interface uplinks to core switches
interface TenGigabitEthernet4/1
 description 10 GigE to Core 1
 ip address 10.122.0.26 255.255.255.254
 ip pim sparse-mode
! Reduce EIGRP hello and hold timers to 1 and 3 seconds and enable MD5
! route authentication
 ip hello-interval eigrp 100 1
 ip hold-time eigrp 100 3
 ip authentication mode eigrp 100 md5
 ip authentication key-chain eigrp 100 eigrp
! Advertise a summary route for the entire distribution block upstream to the
! core
 ip summary-address eigrp 100 10.120.0.0 255.255.0.0 5
 load-interval 30
! Reducing carrier delay to 0 as a
```

```
 carrier-delay msec 0
! Trust inbound DSCP markings
 mls qos trust dscp
!
interface TenGigabitEthernet4/2
 description 10 GigE to Core 2
 ip address 10.122.0.30 255.255.255.254
 ip pim sparse-mode
 ip hello-interval eigrp 100 1
 ip hold-time eigrp 100 3
 ip authentication mode eigrp 100 md5
 ip authentication key-chain eigrp 100 eigrp
 ip summary-address eigrp 100 10.120.0.0 255.255.0.0 5
 load-interval 30
 mls qos trust dscp
!
interface TenGigabitEthernet4/3
 description 10GigE to Distribution-2
 ip address 10.122.0.21 255.255.255.254
 ip pim sparse-mode
 ip hello-interval eigrp 100 1
 ip hold-time eigrp 100 3
 ip authentication mode eigrp 100 md5
 ip authentication key-chain eigrp 100 eigrp
 load-interval 30
 mls qos trust dscp
!
! Configure Layer 2 trunk connections to downstream access switches
interface GigabitEthernet3/3
 description to 4507_SupII+_Access
 no ip address
 load-interval 30
! Trust inbound DSCP markings
 mls qos trust dscp
 switchport
 switchport trunk encapsulation dot1q
! Configure trunk to use a dedicated native VLAN to protect against VLAN hoping
 switchport trunk native vlan 204
! Manually prune all VLANs from trunk other than dedicated voice and data
 switchport trunk allowed vlan 4,104
! Configure switchport to bypass Trunk and Etherchannel negotiation
 switchport mode trunk
 switchport nonegotiate
```

```
!
interface GigabitEthernet3/6
 description to 6500_Sup1A_Access
 no ip address
 load-interval 30
 mls qos trust dscp
 switchport
 switchport trunk encapsulation dot1q
 switchport trunk native vlan 207
 switchport trunk allowed vlan 7,107
 switchport mode trunk
 switchport nonegotiate
!
! Define the Layer 3 SVI for each voice and data VLAN
interface Vlan4
 description Data VLAN for 4507 SupII+
 ip address 10.120.4.3 255.255.255.0
! Enable loose uRPF to mitigate against spoofed source IP addressing
 ip verify unicast source reachable-via any
! Define ip-helper to forward DHCP requests
 ip helper-address 10.121.0.5
 no ip redirects
! Reduce PIM query interval to 250 msec
 ip pim query-interval 250 msec
 ip pim sparse-mode
 load-interval 30
! Define HSRP default gateway with 250/800 msec hello and hold timers
 standby 1 ip 10.120.4.1
 standby 1 timers msec 250 msec 800
! Set preempt delay large enough to allow network to stabilize before HSRP
! switches back on power on or link recovery
 standby 1 preempt delay minimum 180
! Enable HSRP authentication
 standby 1 authentication ese
!
interface Vlan7
 description Data VLAN for 6500 Sup2 CatOS
 ip address 10.120.7.3 255.255.255.0
 ip verify unicast source reachable-via any
 ip helper-address 10.121.0.5
 no ip redirects
 ip pim query-interval 250 msec
 ip pim sparse-mode
```

```
 load-interval 30
 standby 1 ip 10.120.7.1
 standby 1 timers msec 250 msec 800
 standby 1 preempt delay minimum 180
 standby 1 authentication ese
!
interface Vlan104
 description Voice VLAN for 4507 SupII+
 ip address 10.120.104.3 255.255.255.0
 ip verify unicast source reachable-via any
 ip helper-address 10.121.0.5
 no ip redirects
 ip pim query-interval 250 msec
 ip pim sparse-mode
 load-interval 30
 standby 1 ip 10.120.104.1
 standby 1 timers msec 250 msec 800
 standby 1 preempt delay minimum 180
 standby 1 authentication ese
!
interface Vlan107
 description Voice VLAN for 6500 Sup2 CatOS
 ip address 10.120.107.3 255.255.255.0
 ip verify unicast source reachable-via any
 ip helper-address 10.121.0.5
 no ip redirects
 ip pim query-interval 250 msec
 ip pim sparse-mode
 load-interval 30
 standby 1 ip 10.120.107.1
 standby 1 timers msec 250 msec 800
 standby 1 preempt delay minimum 180
 standby 1 authentication ese
!
router eigrp 100
! Passive all interfaces except the core uplinks and link top peer distribution
 passive-interface default
 no passive-interface TenGigabitEthernet4/1
 no passive-interface TenGigabitEthernet4/2
 no passive-interface TenGigabitEthernet4/3
! Specify EIGRP advertise routes for all distribution access and core uplink
! subnets
 network 10.120.0.0 0.0.255.255
```

```
 network 10.122.0.0 0.0.0.255
 no auto-summary
! Explicitly configure the EIGRP router id as a best practice when using Anycast and/or
! any identical loopback address on multiple routers.
 eigrp router-id 10.122.0.3

! Define the valid multicast RP and garbage can RP. See the Multicast Design
! Guide for details on this configuration
ip pim rp-address 10.122.100.1 GOOD-IPMC override
ip pim rp-address 2.2.2.2
ip pim spt-threshold infinity
!
!
ip access-list standard GOOD-IPMC
 permit 224.0.1.39
 permit 224.0.1.40
 permit 239.192.240.0 0.0.3.255
 permit 239.192.248.0 0.0.3.255
```

分布交换机 2——备用根网桥和备用 HSRP

```
! Use vtp transparent mode, configure all VLANs explicitly
vtp domain campus-test
vtp mode transparent
! Enable UDLD aggressive mode as a fail safe mechanism
udld aggressive
! Enable 802.1d per VLAN spanning tree enhancements.
spanning-tree mode pvst
spanning-tree loopguard default
no spanning-tree optimize bpdu transmission
spanning-tree extend system-id
spanning-tree uplinkfast
spanning-tree backbonefast
! Configure root bridge priority as secondary. Note use of extended system-id
spanning-tree vlan 2-7,20-51,102-149,202-207,220-249 priority 28672
!
vlan 4
 name Data_VLAN-4507-SupII+
vlan 7
 name Data_VLAN-6500-Sup2-CatOS
vlan 104
 name Voice_VLAN-4507-SupII+
vlan 107
 name Voice_VLAN-6500-Sup2
```

```
vlan 204
 name Trunk_VLAN-4507-SupII+
vlan 207
 name Trunk_VLAN-6500-Sup2-CatOS
!
! Define a local loopback address to provide a sink hole route point for
! invalid multicast groups
interface Loopback2
 description Garbage-CAN RP
 ip address 2.2.2.2 255.255.255.255
!
! Configure point to point Layer 3 interface uplinks to core switches
interface TenGigabitEthernet4/1
 description 10 GigE to Core 1
 ip address 10.122.0.34 255.255.255.254
 ip pim sparse-mode
! Reduce EIGRP hello and hold timers to 1 and 3 seconds and enable MD5
! route authentication
 ip hello-interval eigrp 100 1
 ip hold-time eigrp 100 3
 ip authentication mode eigrp 100 md5
 ip authentication key-chain eigrp 100 eigrp
! Advertise a summary route for the entire distribution block upstream to the
! core
 ip summary-address eigrp 100 10.120.0.0 255.255.0.0 5
 load-interval 30
! Reducing carrier delay to 0 as a
 carrier-delay msec 0
! Trust inbound DSCP markings
 mls qos trust dscp
!
interface TenGigabitEthernet4/2
 description 10 GigE to Core 2
 ip address 10.122.0.38 255.255.255.254
 ip pim sparse-mode
 ip hello-interval eigrp 100 1
 ip hold-time eigrp 100 3
 ip authentication mode eigrp 100 md5
 ip authentication key-chain eigrp 100 eigrp
 ip summary-address eigrp 100 10.120.0.0 255.255.0.0 5
 load-interval 30
 mls qos trust dscp
!
```

```
interface TenGigabitEthernet4/3
 description 10GigE to Distribution-2
 ip address 10.122.0.22 255.255.255.254
 ip pim sparse-mode
 ip hello-interval eigrp 100 1
 ip hold-time eigrp 100 3
 ip authentication mode eigrp 100 md5
 ip authentication key-chain eigrp 100 eigrp
 load-interval 30
 mls qos trust dscp
!
! Configure Layer 2 trunk connections to downstream access switches
interface GigabitEthernet3/3
 description to 4507_SupII+_Access
 no ip address
 load-interval 30
! Trust inbound DSCP markings
 mls qos trust dscp
 switchport
 switchport trunk encapsulation dot1q
! Configure trunk to use a dedicated native VLAN to protect against VLAN hoping
 switchport trunk native vlan 204
! Manually prune all VLANs from trunk other than dedicated voice and data
 switchport trunk allowed vlan 4,104
! Configure switchport to bypass Trunk and Etherchannel negotiation
 switchport mode trunk
 switchport nonegotiate
!
interface GigabitEthernet3/6
 description to 6500_Sup1A_Access
 no ip address
 load-interval 30
 mls qos trust dscp
 switchport
 switchport trunk encapsulation dot1q
 switchport trunk native vlan 207
 switchport trunk allowed vlan 7,107
 switchport mode trunk
 switchport nonegotiate
!
! Define the Layer 3 SVI for each voice and data VLAN
interface Vlan4
 description Data VLAN for 4507 SupII+
```

```
 ip address 10.120.4.2 255.255.255.0
! Enable loose uRPF to mitigate against spoofed source IP addressing
 ip verify unicast source reachable-via any
! Define ip-helper to forward DHCP requests
 ip helper-address 10.121.0.5
 no ip redirects
! Reduce PIM query interval to 250 msec
 ip pim query-interval 250 msec
 ip pim sparse-mode
 load-interval 30
! Define HSRP default gateway with 250/800 msec hello and hold timers
 standby 1 ip 10.120.4.1
 standby 1 timers msec 250 msec 800
! Raising HSRP priority forces this switch to be active HSRP gateway
 standby 1 priority 150
! Set preempt delay large enough to allow network to stabilize before HSRP
! switches back on power on or link recovery
 standby 1 preempt delay minimum 180
! Enable HSRP authentication
 standby 1 authentication ese
!
interface Vlan7
 description Data VLAN for 6500 Sup2 CatOS
 ip address 10.120.7.2 255.255.255.0
 ip verify unicast source reachable-via any
 ip helper-address 10.121.0.5
 no ip redirects
 ip pim query-interval 250 msec
 ip pim sparse-mode
 load-interval 30
 standby 1 ip 10.120.7.1
 standby 1 timers msec 250 msec 800
 standby 1 priority 150
 standby 1 preempt delay minimum 180
 standby 1 authentication ese
!
interface Vlan104
 description Voice VLAN for 4507 SupII+
 ip address 10.120.104.2 255.255.255.0
 ip verify unicast source reachable-via any
 ip helper-address 10.121.0.5
 no ip redirects
 ip pim query-interval 250 msec
```

```
 ip pim sparse-mode
 load-interval 30
 standby 1 ip 10.120.104.1
 standby 1 timers msec 250 msec 800
 standby 1 preempt delay minimum 180
 standby 1 authentication ese
!
interface Vlan107
 description Voice VLAN for 6500 Sup2 CatOS
 ip address 10.120.107.2 255.255.255.0
 ip verify unicast source reachable-via any
 ip helper-address 10.121.0.5
 no ip redirects
 ip pim query-interval 250 msec
 ip pim sparse-mode
 load-interval 30
 standby 1 ip 10.120.107.1
 standby 1 timers msec 250 msec 800
 standby 1 preempt delay minimum 180
 standby 1 authentication ese
!
router eigrp 100
! Passive all interfaces except the core uplinks and link top peer distribution
 passive-interface default
 no passive-interface TenGigabitEthernet4/1
 no passive-interface TenGigabitEthernet4/2
 no passive-interface TenGigabitEthernet4/3
! Specify EIGRP advertise routes for all distribution access and core uplink
! subnets
 network 10.120.0.0 0.0.255.255
 network 10.122.0.0 0.0.0.255
 no auto-summary
 eigrp router-id 10.122.0.4
! Define the valid multicast RP and garbage can RP. See the Multicast Design
! Guide for details on this configuration
ip pim rp-address 10.122.100.1 GOOD-IPMC override
ip pim rp-address 2.2.2.2
ip pim spt-threshold infinity
!
!
ip access-list standard GOOD-IPMC
 permit 224.0.1.39
 permit 224.0.1.40
```

```
 permit 239.192.240.0 0.0.3.255
 permit 239.192.248.0 0.0.3.255
```

IOS 接入交换机（4507/SupII+）

```
vtp mode transparent
udld aggressive
! Auto QoS defined policier
policy-map autoqos-voip-policy
   class class-default
 dbl
! Enable 802.1d per VLAN spanning tree enhancements.
spanning-tree loopguard default
spanning-tree extend system-id
spanning-tree uplinkfast
spanning-tree backbonefast
! Define dedicated voice, data and trunk VLAN for this access switch
vlan 4
 name Data
vlan 104
 name Voice
vlan 204
 name Trunk
! Define switchport trunk uplink to distribution
interface GigabitEthernet1/1
 description Uplink to Distribution Switch 1
 switchport trunk encapsulation dot1q
! Define a unique trunk vlan in order to prevent vlan hopping attacks
 switchport trunk native vlan 204
! Explicitly configure voice and data vlan on trunk
 switchport trunk allowed vlan 4,104
! Explicitly enable trunking and disable Etherchannel negotiation
 switchport mode trunk
 switchport nonegotiate
 load-interval 30
! Auto QoS defined configuration for voice enabled uplink
 qos trust cos
 service-policy output autoqos-voip-policy
 auto qos voip trust
 tx-queue 3
  bandwidth percent 33
  priority high
  shape percent 33
spanning-tree link-type point-to-point
```

```
!
interface GigabitEthernet1/2
 description Uplink to Distribution Switch 2
 switchport trunk encapsulation dot1q
 switchport trunk native vlan 204
 switchport trunk allowed vlan 4,104
 switchport mode trunk
 switchport nonegotiate
 load-interval 30
 qos trust cos
 service-policy output autoqos-voip-policy
 auto qos voip trust
 tx-queue 3
    bandwidth percent 33
    priority high
    shape percent 33
 spanning-tree link-type point-to-point
!
! Define access ports using the recommended Smartports configuration. Please
! see the following for more information on smartports
http://www.cisco.com/en/US/partner/netsol/ns439/networking_solutions_packages_list.html
!
interface FastEthernet2/1
 switchport access vlan 4
 switchport mode access
 switchport voice vlan 104
 switchport port-security
 switchport port-security aging time 2
 switchport port-security violation restrict
 switchport port-security aging type inactivity
 qos trust device cisco-phone
 qos trust cos
 service-policy output autoqos-voip-policy
 auto qos voip cisco-phone
 tx-queue 3
    priority high
    shape percent 33
 spanning-tree portfast
 spanning-tree bpduguard enable
!
! Define switch management address on data vlan
interface Vlan4
 ip address 10.120.4.4 255.255.255.0
```

CatOs 接入交换机（6500/Sup2）

```
set vtp domain campus-test
set vtp mode transparent
set vlan 7 name Data type ethernet mtu 1500 said 100007 state active
set vlan 107 name Voice type ethernet mtu 1500 said 100107 state active
set vlan 207 name Uplink type ethernet mtu 1500 said 100207 state active

#ip
set interface sc0 7 10.120.7.4/255.255.255.0 10.120.7.255

#uplinkfast groups
set spantree uplinkfast enable rate 0 all-protocols off
#module 1 : 2-port 1000BaseX Supervisor
set vlan 207 1/1-2
set udld aggressive-mode enable 1/1-2
clear trunk 1/1 1-6,8-106,108-206,208-1005,1025-4094
set trunk 1/1 nonegotiate dot1q 7,107,207
clear trunk 1/2 1-6,8-106,108-206,208-1005,1025-4094
set trunk 1/2 nonegotiate dot1q 7,107,207
set spantree guard loop 1/1-2
set port qos 1/1-2 trust trust-dscp
set port qos 1/1-2 policy-source local
set port channel 1/1-2 mode off
! Define access ports using the recommended Smartports configuration. Please
! see the following for more information on Smartports,
http://www.cisco.com/en/US/partner/netsol/ns439/networking_solutions_packages_list.html
!
set port enable 4/1
set port l2protocol-tunnel 4/1 cdp stp vtp dis
set port membership 4/1 static
set port host 4/1
set spantree bpdu-guard 4/1 enable
set vlan 7 4/1
set port auxiliaryvlan 4/1 107
set port inlinepower 4/1 auto
set cdp enable 4/1
set port qos 4/1 autoqos voip ciscoipphone
set port security 4/1 enable age 2 maximum 1 violation restrict
```

第三层接入和分布块的交换机配置

在上述第三层接入测试中采用了下列配置：

- 生成树版本——Rapid-PVST+（每 VLAN 802.1w）
- IGP——EIGRP

其他测试情况中基本配置的所有变化都记录在上述结果中。欲了解有关推荐配置的具体信息，请参考《设计高可用性园区网络》。

本节探讨了下列主题：

分布节点 EIGRP

注意 分布交换机对称配置。

```
key chain eigrp
 key 100
key-string 7 01161501
!
<Configure spanning tree as a redundant protective mechanism>
spanning-tree mode rapid-pvst
spanning-tree loopguard default
!
<Configure point to point Layer 3 links to each of the access switches>
interface GigabitEthernet3/1
 description Link to Access Switch
<configure the switch to switch link using a /30 or /31 subnet>
 ip address 10.120.0.204 255.255.255.254
 ip pim sparse-mode
<specify the use of 1 second hello and 3 second dead timers for EIGRP>
 ip hello-interval eigrp 100 1
 ip hold-time eigrp 100 3
<enable eigrp MD5 authentication>
 ip authentication mode eigrp 100 md5
 ip authentication key-chain eigrp 100 eigrp
 logging event link-status
 load-interval 30
<Set carrier delay to 0. On Catalyst 6500 this will have no effect on GigE ports however
it is necessary on 3x50 series switches and should be consistently configured for best
practices>
 carrier-delay msec 0
<Trust the dscp settings in all packets sourced from the access. We are extending the
trust boundary to the access switch>
 mls qos trust dscp
!
!
```

```
<Configure point to point L3 links to each of the core switches>
interface TenGigabitEthernet4/1
 description 10 GigE to Core 1
<configure the switch to switch link using a /30 or /31 subnet>
 ip address 10.122.0.26 255.255.255.254
 ip pim sparse-mode
<specify the use of 1 second hello and 3 second dead timers for EIGRP>
 ip hello-interval eigrp 100 1
 ip hold-time eigrp 100 3
<Configure EIGRP authentication on all links>
 ip authentication mode eigrp 100 md5
 ip authentication key-chain eigrp 100 eigrp
<Advertise a summary address for the entire distribution block upstream to the core>
 ip summary-address eigrp 100 10.120.0.0 255.255.0.0 5
 logging event link-status
 load-interval 30
 carrier-delay msec 0
<Trust all DSCP markings from the core of the network>
 mls qos trust dscp
!
!
<Configure a point to point Layer 3 link between distribution switches>
interface TenGigabitEthernet4/3
 description 10 GigE to Distribution 2
<configure the switch to switch link using a /30 or /31 subnet>
 ip address 10.122.0.21 255.255.255.254
 ip pim sparse-mode
<specify the use of 1 second hello and 3 second dead timers for EIGRP>
 ip hello-interval eigrp 100 1
 ip hold-time eigrp 100 3
<Configure EIGRP authentication on all links>
 ip authentication mode eigrp 100 md5
 ip authentication key-chain eigrp 100 eigrp
 logging event link-status
 load-interval 30
 mls qos trust dscp
!
!
router eigrp 100
<Passive all interfaces not connected to another Layer 3 switch>
 passive-interface GigabitEthernet2/1
<Specify which networks should be routed by EIGRP. Include the distribution block and the
core links>
```

```
 network 10.120.0.0 0.0.255.255
 network 10.122.0.0 0.0.0.255
<Apply a distribute list filtering all routes other than select default(s) to the access
switches>
 distribute-list Default out GigabitEthernet3/1
 distribute-list Default out GigabitEthernet3/2
 . . .
 distribute-list Default out GigabitEthernet9/14
 distribute-list Default out GigabitEthernet9/15
 no auto-summary
! Explicitly configure the EIGRP router id as a best practice when using Anycast and/or
! any identical loopback address on multiple routers.
 eigrp router-id 10.122.0.3
!
ip classless
no ip http server
ip pim rp-address 10.122.100.1 GOOD-IPMC override
ip pim rp-address 2.2.2.2
ip pim spt-threshold infinity
!
!
ip access-list standard Default
 permit 0.0.0.0
ip access-list standard GOOD-IPMC
 permit 224.0.1.39
 permit 224.0.1.40
 permit 239.192.240.0 0.0.3.255
 permit 239.192.248.0 0.0.3.255
```

接入节点 EIGRP（冗余交换管理引擎）

```
key chain eigrp
 key 100
key-string 7 01161501
!
<Configure spanning tree as a redundant protective mechanism>
spanning-tree mode rapid-pvst
spanning-tree loopguard default
!
redundancy
 mode sso
 main-cpu
  auto-sync running-config
auto-sync standard
```

```
!
<Create a local Data and Voice VLAN>
vlan 6
 name Access-Data-VLAN
!
vlan 106
 name Access-Voice-VLAN
!
interface Loopback22
 ip address 2.2.2.2 255.255.255.255
!
<Define the uplink to the Distribution switches as a point to point Layer 3 link>
interface GigabitEthernet1/1
 description Uplink to Distribution 1
 ip address 10.120.0.205 255.255.255.254
 ip pim sparse-mode
<Reduce EIGRP hello and dead timers to 1 and 3 seconds>
 ip hello-interval eigrp 100 1
 ip hold-time eigrp 100 3
<Enable EIGRP MD5 authentication>
 ip authentication mode eigrp 100 md5
 ip authentication key-chain eigrp 100 eigrp
 logging event link-status
 load-interval 30
 carrier-delay msec 0
 mls qos trust dscp
!
interface GigabitEthernet2/1
 description Uplink to Distribution 2
 ip address 10.120.0.61 255.255.255.252
 ip pim sparse-mode
 ip hello-interval eigrp 100 1
 ip hold-time eigrp 100 3
 ip authentication mode eigrp 100 md5
 ip authentication key-chain eigrp 100 eigrp
 logging event link-status
 load-interval 30
 carrier-delay msec 0
 mls qos trust dscp
!
<Define Switched Virtual Interfaces's for both access Data and Voice VLANs>
interface Vlan6
 ip address 10.120.6.1 255.255.255.0
```

```
 ip helper-address 10.121.0.5
 no ip redirects
 ip pim query-interval 250 msec
 ip pim sparse-mode
 load-interval 30
!
interface Vlan106
 ip address 10.120.106.1 255.255.255.0
 ip helper-address 10.121.0.5
 no ip redirects
 ip pim query-interval 250 msec
 ip pim sparse-mode
 load-interval 30
!
<Configure EIGRP as an EIGRP stub router, advertising connected routes upstream to the
distribution>
router eigrp 100
 network 10.120.0.0 0.0.255.255
 no auto-summary
 eigrp stub connected
 eigrp router-id 10.122.0.22
!
ip classless
no ip http server
ip pim rp-address 10.122.100.1 GOOD-IPMC override
ip pim rp-address 2.2.2.2
ip pim spt-threshold infinity
!
ip access-list standard GOOD-IPMC
 permit 224.0.1.39
 permit 224.0.1.40
 permit 239.192.240.0 0.0.3.255
 permit 239.192.248.0 0.0.3.255
```

分布节点 OSPF

注意 分布交换机对称配置。

```
key chain eigrp
 key 100
  key-string 7 01161501
!
<Configure spanning tree as a redundant protective mechanism>
spanning-tree mode rapid-pvst
```

```
spanning-tree loopguard default
!
<Configure point to point Layer 3 links to each of the access switches>
interface GigabitEthernet3/1
 description Link to Access Switch
<configure the switch to switch link using a /30 or /31 subnet>
 ip address 10.120.0.204 255.255.255.254
 ip pim sparse-mode
<specify the use of 1 second hello and 3 second dead timers>
 ip ospf hello-interval 1
 ip ospf dead-interval 3
<enable eigrp MD5 authentication>
 ip authentication mode eigrp 100 md5
 ip authentication key-chain eigrp 100 eigrp
 logging event link-status
 load-interval 30
<Set carrier delay to 0. On Catalyst 6500 this will have no effect on GigE ports however
it is necessary on 3x50 series switches and should be consistently configured for best
practices>
 carrier-delay msec 0
<Trust the dscp settings in all packets sourced from the access. We are extending the
trust boundary to the access switch>
 mls qos trust dscp
!
!
<Configure point to point L3 links to each of the core switches>
interface TenGigabitEthernet4/1
 description 10 GigE to Core 1
<configure the switch to switch link using a /30 or /31 subnet>
 ip address 10.122.0.26 255.255.255.254
 ip pim sparse-mode
<specify the use of 1 second hello and 3 second dead timers>
 ip ospf hello-interval 1
 ip ospf dead-interval 3
 logging event link-status
 load-interval 30
 carrier-delay msec 0
<Trust all DSCP markings from the core of the network>
 mls qos trust dscp
!
!
<Configure a point to point Layer 3 link between distribution switches>
interface TenGigabitEthernet4/3
 description 10 GigE to Distribution 2
<configure the switch to switch link using a /30 or /31 subnet>
 ip address 10.122.0.21 255.255.255.254
 ip pim sparse-mode
<specify the use of 1 second hello and 3 second dead timers>
```

```
 ip ospf hello-interval 1
 ip ospf dead-interval 3
 logging event link-status
 load-interval 30
 mls qos trust dscp

router ospf 100
! Explicitly configure the OSPF router id as a best practice when using Anycast and/or
! any identical loopback address on multiple routers.
 router-id 10.122.0.3
 log-adjacency-changes
 area 120 stub no-summary
 area 120 range 10.120.0.0 255.255.0.0
 timers throttle spf 1000 1000 1000
 network 10.120.0.0 0.0.255.255 area 120
 network 10.122.0.0 0.0.255.255 area 0
```

接入节点 OSPF（冗余交换管理引擎）

```
!
<Configure spanning tree as a redundant protective mechanism>
spanning-tree mode rapid-pvst
spanning-tree loopguard default
!
redundancy
 mode sso
 main-cpu
  auto-sync running-config
auto-sync standard
!
vlan 4
 name cr8-4507-1-Data-VLAN
!
vlan 104
 name cr8-4507-1-Voice-VLAN
!
<Define the uplink to the Distribution switches as a point to point Layer 3 link>
interface GigabitEthernet1/1
 description Uplink to Distribution 1
 ip address 10.120.0.205 255.255.255.254
 ip pim sparse-mode
<Reduce hello and dead timers to 1 and 3 seconds>
 ip ospf hello-interval 1
 ip ospf dead-interval 3
<Enable EIGRP MD5 authentication>
 ip authentication mode eigrp 100 md5
 ip authentication key-chain eigrp 100 eigrp
 logging event link-status
 load-interval 30
```

```
 carrier-delay msec 0
 mls qos trust dscp
!
interface GigabitEthernet2/1
 description Uplink to Distribution 2
 ip address 10.120.0.61 255.255.255.252
 ip pim sparse-mode
 ip ospf hello-interval 1
 ip ospf dead-interval 3
 ip authentication mode eigrp 100 md5
 ip authentication key-chain eigrp 100 eigrp
 logging event link-status
 load-interval 30
 carrier-delay msec 0
 mls qos trust dscp
!
<Define Switched Virtual Interfaces's for both access Data and Voice VLANs>
interface Vlan4
 ip address 10.120.4.1 255.255.255.0
 ip helper-address 10.121.0.5
 no ip redirects
 ip pim query-interval 250 msec
 ip pim sparse-mode
 load-interval 30
!
interface Vlan104
 ip address 10.120.104.1 255.255.255.0
 ip helper-address 10.121.0.5
 no ip redirects
 ip pim query-interval 250 msec
 ip pim sparse-mode
 load-interval 30
!
router ospf 100
 router-id 10.122.0.22
 log-adjacency-changes
 area 120 stub no-summary
 timers throttle spf 1000 1000 1000
 network 10.120.0.0 0.0.255.255 area 120
```

思科系统（中国）网络技术有限公司

北京
北京市东城区东长安街 1 号东方广场东方经贸城东一办公楼 19-21 层

邮政编码：100738
电话：(8610) 85155000
传真：(8610) 85181881

上海
上海市淮海中路 222 号力宝广场 32-33 层

邮政编码：200021
电话：(8621) 33104777
传真：(8621) 53966750

广州
广州市天河北路 233 号中信广场 43 楼

邮政编码：510620
电话：(8620) 85193000
传真：(8620) 38770077

成都
成都市顺城大街 308 号冠城广场 23 层

邮政编码：610017
电话：(8628) 86961000
传真：(8628) 86528999

如需了解思科公司的更多信息，请浏览 http://www.cisco.com /cn

2005 年 3 月翻译