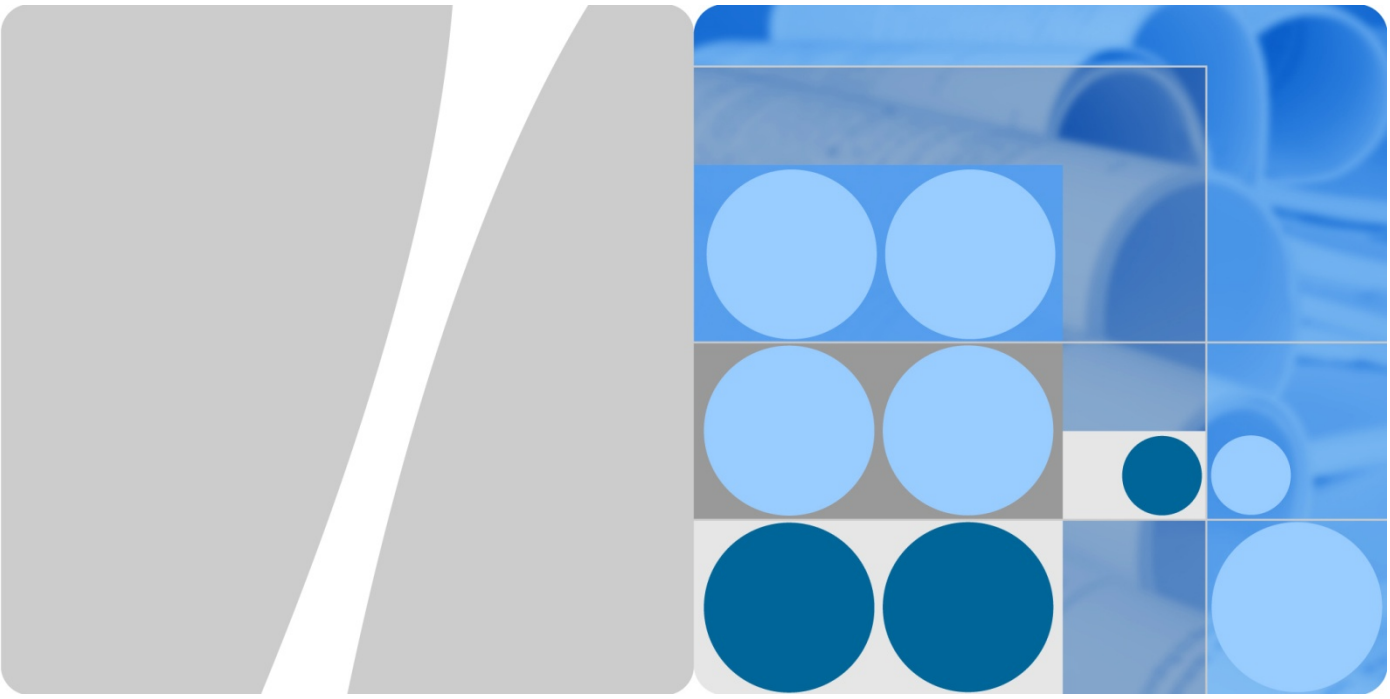


数据中心解决方案
V100R001C01
技术建议书

文档版本 01

发布日期 2011-07-22

华为技术有限公司

商标声明


和其他华为商标均为华为技术有限公司的商标。

本文档提及的其他所有商标或注册商标，由各自的所有人拥有。

注意

您购买的产品、服务或特性等应受华为公司商业合同和条款的约束，本文档中描述的全部或部分产品、服务或特性可能不在您的购买或使用范围之内。除非合同另有约定，华为公司对本文档内容不做任何明示或默示的声明或保证。

由于产品版本升级或其他原因，本文档内容会不定期进行更新。除非另有约定，本文档仅作为使用指导，本文档中的所有陈述、信息和建议不构成任何明示或暗示的担保。

华为技术有限公司

地址： 深圳市龙岗区坂田华为总部办公楼 邮编：518129

网址： http://www.huawei.com

客户服务邮箱： support@huawei.com

客户服务电话： 0755-28560000 4008302118

客户服务传真： 0755-28560111

目 录

1 数据中心网络概述

1.1 数据中心网络简介

当今社会的竞争是信息化的竞争，企业信息化程度越高，竞争力就越大。随着网络技术、通信技术的发展，数据中心已经成为企业信息化的核心，数据中心的建设好坏直接影响企业的效率和发展。

数据中心是为企业的关键业务系统提供承载的最重要的 IT 基础设施。是企业核心数据管理中心。企业的数据中心需要集中处理因企业业务需求而产生的接入控制、安全过滤、服务应用、信息计算、存储备份等环节。

数据中心主要组件包括: 土建（房间场地）、供电系统、网络设备（数据网、计算网、存储网）、服务器（包含相关操作系统和应用软件）、存储器、安全系统以及相应的运维系统等。

对于企业来讲，随着“数据大集中”的进展，越来越多的业务和数据都集中到若干个数据中心。数据中心承载网络的高性能和高可靠性成为企业业务开展的核心问题。

华为数据中心网络方案主要侧重于对建设数据中心的网络部分提出整体建议，保证数据中心的高性能、安全、可靠，从而使数据中心能承载更多高品质的业务。

1.2 数据中心网络总体需求

企业数据中心从物理上看是聚集了大量服务器的一个地方，它不但是企业网络的逻辑中心，更是企业提供业务的源头。企业数据中心应具备十分丰富的带宽资源、安全可靠的机房设施、高水平的网络管理、十分完备的增值服务。数据中心的实质是创造尽可能多的基于带宽的增值价值。因此，下面几点需求对数据中心特别重要：

a. 可靠性

   高可靠性是数据中心运营成功的关键，企业对外提供的业务如电子商务、视讯类业务，如果因为数据中心网络故障而导致用户体验差，这势必影响企业业务的扩展，严重的甚至导致企业盈利空间的下降。所以可靠性是企业部署数据中心的重要原则。

   可靠性设计包括：链路冗余、关键设备冗余和重要业务模块冗余。

b. 可扩展性

数据中心方案设计中，每个层次的设计所采用的设备本身都应具有极高的端口密度，为数据中心的扩展奠定基础。

在 Internet 互联层、Intranet 互联层、核心/汇聚层的设备都采用模块化设计，可根据数据中心网络的发展进行灵活扩展。

功能的可扩展性是数据中心提供增值业务的基础。实现负载均衡、动态内容复制、VLAN 等功能，为数据中心增值业务的扩展提供基础。

c. 可管理性

网络的可管理性是数据中心运营管理成功的基础。数据中心应提供多种优化的可管理信息。数据中心应具备完整的 QoS 功能、完整的 SLA 管理体系、多厂家网络设备管理能力、相对独立的后台管理平台，方便数据中心及其用户的网络管理。

d. 安全性

安全性是数据中心的用户特别是电子商务用户最为关注的问题，也是数据中心建设中的关键，它包括物理空间的安全控制及网络的安全控制。数据中心应有完整的安全策略控制体系以实现数据中心安全控制。

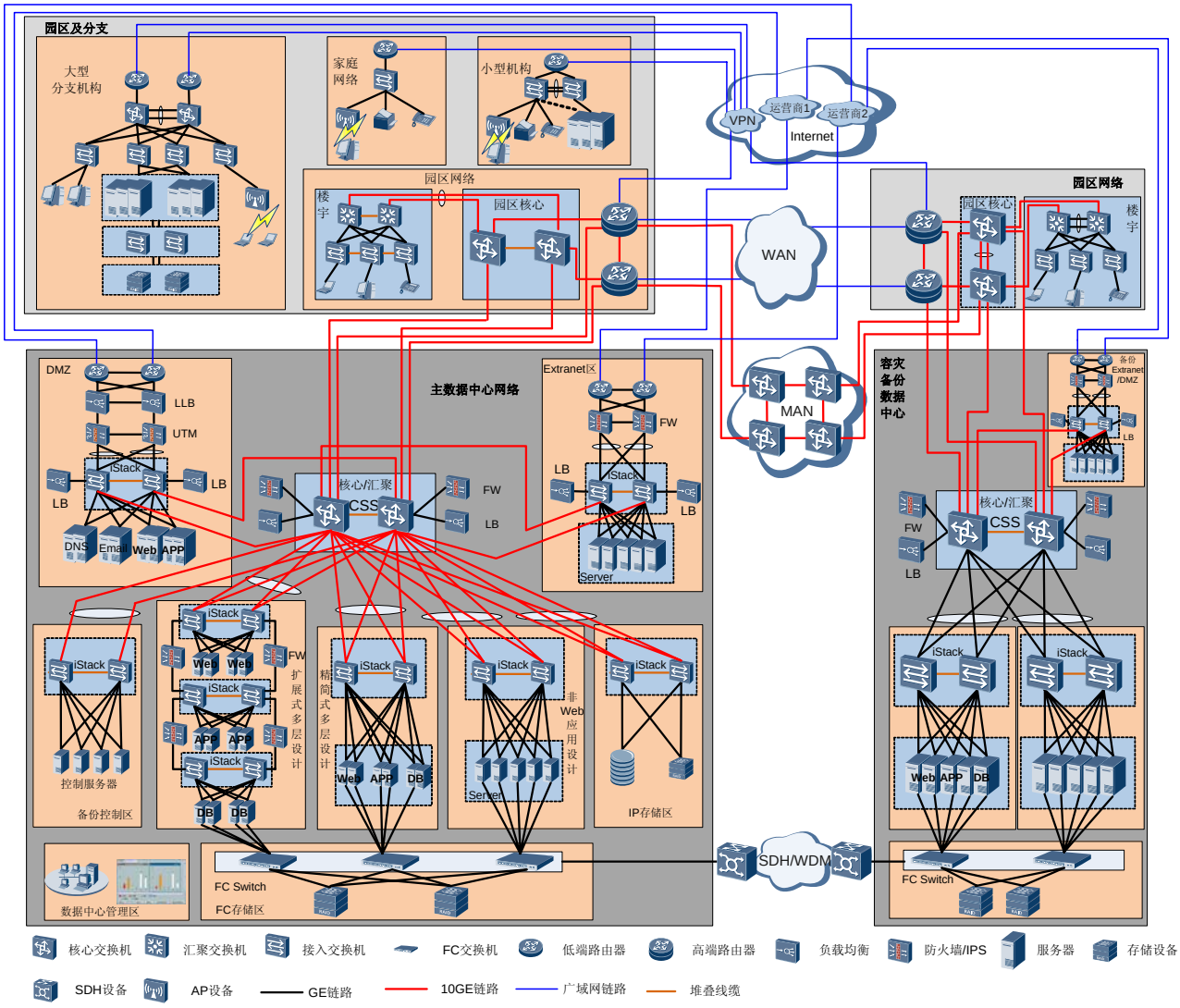
1.3 数据中心网络解决方案

总体方案描述

华为全业务数据中心解决方案，具有以下特点：

- 模块化、层次化结构
- 前台业务网络与后台管理网络相分离，保证了业务网络的高性能和高安全性
- 业务承载网络按功能划分不同的业务区，为不同业务区用户提供差异化服务

图1-1 数据中心网络整体解决方案组网图



如图 1-1 所示，为了增强网络安全性、可扩展性和可维护性，华为数据中心解决方案可分为前端业务网络、后端管理网络和存储网络。

- 前端业务网络主要由网络连接模块、服务器接入模块构成；
- 后端管理网络由后端管理模块构成；
- 存储网络由存储系统和 SAN（Storage Area Network）网络组成。

该技术建议书重点考虑前端业务网络和后端管理网络。

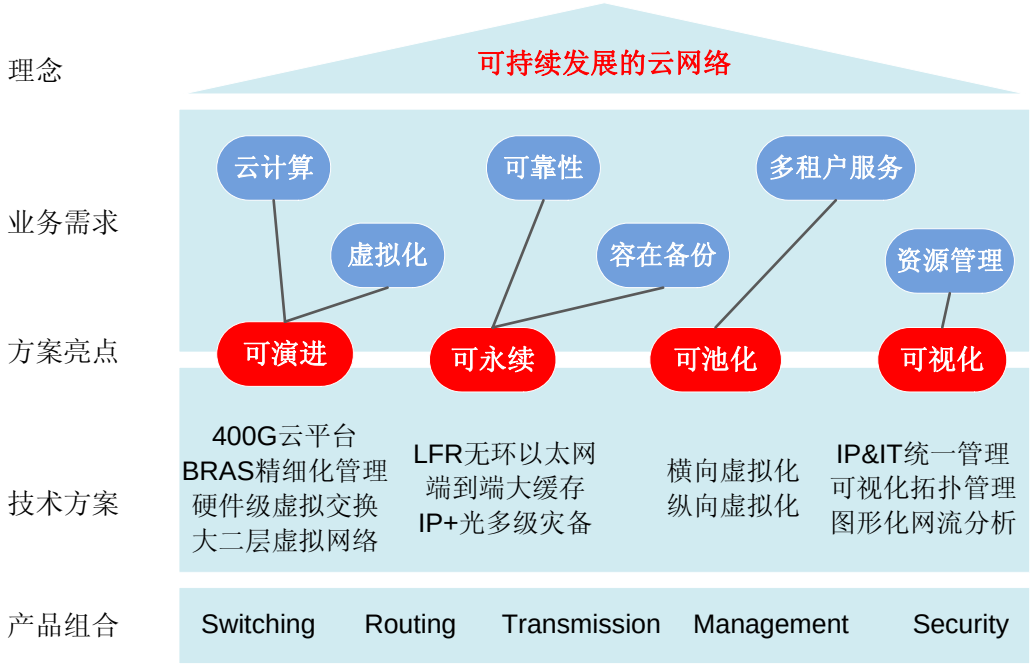
网络连接模块通过路由器、交换机、防火墙、负载均衡器、统一威胁管理 UTM（Unified Threat Management，含防火墙、IDS/IPS、防病毒、URL 过滤、SSL VPN 等）提供高性能、高密度、高可用性、高安全的网络基础架构。

服务器接入模块根据客户提供业务的不同，根据业务类别和业务特征又划分成多个不同的业务区块，根据需要每个业务区在物理上或逻辑上做到隔离。

数据中心网络解决方案亮点

华为秉承“Cloud Network”的核心理念，提出“可持续发展”的数据中心网络解决方案，提供“可演进”、“可永续”、“可池化”和“可视化”四大亮点，以迎接“云计算”时代面临的强劲挑战。

图1-2 数据中心网络解决方案亮点



- 可演进-----面向 “云计算”和“虚拟化”数据中心

  400G 云网络平台：华为数据中心核心交换机基于 10Tbit 级无阻塞 CLOS 架构，可平滑升级到业界最佳的 400G 平台，支持高密度 40*10GE 业务板和 100GE 端口，充分满足“云计算”超宽带数据中心的容量需求。

  虚拟化新技术演进：硬件交换机上实现虚拟交换和策略感知（802.1Qbg VEPA），管理界面清晰，流量可管可控，并大幅提高虚拟机交换性能；硬件支持基于 ISIS 的透明路由桥协议（802.1AQ/TRILL），可软件升级提供网络扩展能力，为云计算 DC 提供超大范围的虚拟机无缝迁移方案。

  桌面云精细化管理：针对桌面云数据中心，创新性地引入运营级 BRAS（Broadband Remote Access Server）部署经验，支持海量桌面云虚拟机的准入管理；同时基于虚拟机用户和业务，实现精细化的带宽控制和 SLA（Service-Level Agreement）保证（H-QoS）。

- 可永续----- LFR（Loop Free Reliable）无环以太网打造无间断数据中心

  端到端高可靠架构：端到端 200 毫秒的可靠性架构保证了数据中心的持续运营，实现业务永续。L2 交换（如汇聚到接入层）采用业界领先的 LFR 无环以太网技术，打造快速收敛的无环网络；L3 路由（如核心层以上）采用毫秒级的 BFD（Bidirectional Forwarding Detection）+FRR（Fast Reroute）快速重路由技术；同时配合设备级的 ISSU（In-Service Software Upgrade）和关键组件冗余等技术，打造无间断的数据中心。

LFR 无环以太网：数据中心交换机采用“CSS（Cluster Switch System）集群 + LAG（Link Aggregation Group） + iStack 堆叠”技术组合，构建业界最佳的 LFR 无环以太网：a）最可靠的交换网“硬”集群、b）最快的 200 毫秒级收敛、c）最大的 256G 集群带宽。

扁平化无丢包网络：数据中心高端交换机支持 10GE/GE 接口 200ms 缓存，通过 S12700（核心）+S9300（EOR（End of Rack）接入）扁平化组网，实现整网端到端大缓存部署，满足分布式计算等业务的低时延和流量突发无丢包需求。

“IP+光”多级灾备：集成光传输和路由器，形成全方位的数据级和业务级快速容灾备份能力；OTN 设备提供全面的 14 种专业级存储接口（FC/FICON/ESCON 等），支持主备中心间的实时硬件备份；NE40E 路由器提供多数据中心间的灵活互联和 IP SAN 备份。

- 可池化-----池化网络实现数据中心资源按需调度

  资源按需调度：横向角度，多个交换机可通过 CSS 集群或 iStack 堆叠技术虚拟为一台设备，完全共享 100%的网络总带宽，避免传统 STP 方式 50%的带宽浪费；纵向角度，MPLS VPN/ MCE 等多实例技术确保不同业务可以灵活调度数据中心网络资源池，实现按需分配。

  简化网络架构：CSS 和堆叠技术将多个交换机虚拟为 1 台逻辑设备，网管角度可看作单一网元，从而简化了逻辑网络架构，降低用户管理和配置的复杂度。

  业务有效隔离：MPLS VPN 和 MCE 等虚拟化多实例技术充分保证数据中心业务的有效隔离和安全，并可通过灵活设置 VPN 访问策略，控制不同部门和数据中心服务器间的互访行为。

- 可视化-----智能可视网管实现 IP&IT 统一管理

  IP&IT 统一管理：eSight 智能网管集成数据中心多系统管理能力（网络设备、服务器及企业应用系统），降低建设成本，提高运维效率；同时提供开放的第三方平台，和业内主流 IT 厂商（IBM/HP/Oracle 等）进行深度集成和广泛合作。

  OSS 合作伙伴：http://www.huawei.com/partners/integrated_with_oss.do

  可视化拓扑管理：eSight 智能网管提供数据中心网络拓扑和业务视图，使得业务部署和网络配置更加直观便捷。

  图形化网流分析：数据中心交换机/路由器通过随板或集成的 Netstream 功能，可随时洞察数据中心的业务分布，为用户提供图形化的网流分析报告，帮助客户轻松制定业务规划。

2 业务需求

2.1 概述

企业建设数据中心，一方面通过将企业的各种业务系统集中部署到数据中心，完成业务系统的整合，进一步支撑业务分析、决策支持，最大化地提升信息的生产力。

另一方面，通过 WEB 提供信息门户，建立和客户沟通的渠道，提供企业宣传、产品推广、客户服务，进一步支撑电子商务，完成基于互联网的业务运营。

数据中心还能够提供高性能计算的业务，如：3D 渲染、药物研究、基因分析、WEB 搜索业务。

在一个企业的数据中心中，可能同时有这三种典型业务，这些业务可能相对独立，也可能融合在大的业务系统中，在数据中心网络规划时需要根据具体情况综合分析。

2.2 数据业务

2.2.1 数据业务简介

数据业务是数据中心中最基本的业务。对于企业来说，文件存储、邮件系统、ERP（Enterprise Resource Planning）都是典型的数据业务。其基本业务模型是 C/S 模式的。

图2-1 C/S 业务模式

C/S 模式由两部分构成：

- 前端是客户机，通常是 PC，一般就在企业的园区网或企业分支中。

- 后端是服务器，部署在数据中心，由独立的存储设备为服务器提供扩展存储。例如：数据库应用的服务器就是数据库服务器，数据库的数据存放在专用的存储设备中（上图中省略了）。

2.2.2 数据业务网络需求

数据业务的典型流程是：

a. 客户端发出请求

b. 服务器和存储共同完成数据处理

c. 服务器向客户端发送数据

对网络的需求主要反映在：

- 流量需求

  数据业务的流量主要是客户端和服务器之间的数据请求和应答，其特征是流量极不均衡，一般会在业务系统的特殊日期或时间段（如结帐日）出现数据流量的高峰期。网络流量的规划需要按照峰值，并为未来业务的发展留有余量。

  另外，网络设备的流量规划还需要参考客户端数量、多个业务的并发情况进行规划。由于没有服务器间的流量，网络各层设备间的带宽收敛比的设计主要依据业务的并发情况进行设计。

  例如：各业务系统的结帐日就是业务数据流量的高峰期，如果生产、销售、考勤的结帐日规定在不同日期，则网络带宽的峰值可以取三个业务系统中数据流量峰值最大的一个。如果几个业务的结帐日在同一天，则网络带宽的峰值可以取三个业务系统数据流量峰值的总和。

  数据业务一般对时延没有特别要求，以满足使用者的业务体验为目的，一般数据库应用的响应时间需要在 2 秒内，数据中心网络转发的时延在整个响应时间中只占很小一部分，其他如广域网时延需要约 300 毫秒，数据处理时长可能达到数十毫秒，而数据中心中网络转发的时延是小于 1 毫秒的。

  也有一些特殊行业要求时延特别低，如证券交易要求网络转发时延在 5 微秒内。

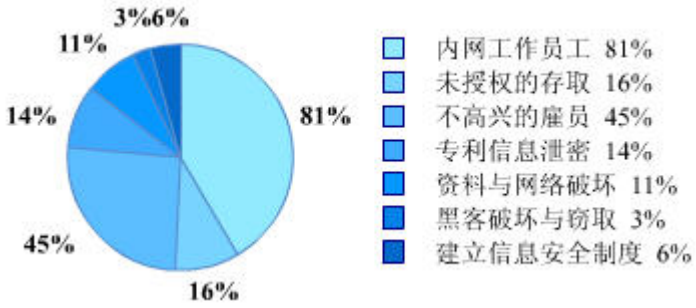
- 安全性需求

  数据中心是业务和数据的中心，在 IT 系统中是安全性要求最高的。

  如图 2-2 所示，据美国 CSI/FBI 的调查结果显示，企业和政府机构因重要信息被窃所造成的损失，已远远超过病毒感染和黑客攻击所造成的损失，80%以上的安全威胁来自内部。

  同时，据中国公安部最新统计，70％的泄密犯罪来自于内部；电脑应用单位 80％未设立相应的安全管理系统、技术措施和制度。

图2-2 企业中影响信息安全的因素



企业的财务等关键业务一般都是以数据业务的形式承载的，对安全性的要求是非常高的。除了物理安全的手段，网络也需要对业务安全有必要的支撑，包括：不同业务的隔离，网络攻击流量、病毒传播的识别处理等等。

业务的隔离要求各种终端只能访问相应的业务的服务器。

- 可靠性要求

  数据业务的可靠性主要是对数据可靠性的要求，对网络的可靠性根据业务是对内还是对外要求不同。

  对于企业内部使用的业务系统，网络可靠性的要求不高，数据中心内部的故障在20～30 分钟内恢复，数据中心整体故障需要在 4～8 小时内从备用数据中心恢复业务。

  对于对外提供服务的业务系统，网络可靠性的要求较高，数据中心内部的故障应能够自动切换修复，或者在 10 分钟内手工恢复，数据中心整体故障需要在 2 小时内从容灾中心恢复。

  按照国际标准 share78，将容灾分为 7 个层次，分别为：

  - 层次 0：没有异地数据（No off-site Data）
  - 层次 1：PTAM 卡车运送访问方式（Pickup Truck Access Method）
  - 层次 2：PTAM 卡车运送访问方式+热备份中心（PTAM + Hot 中心）
  - 层次 3：电子链接（Electronic Vaulting）
  - 层次 4：活动状态的备份中心（Active Secondary 中心）
  - 层次 5：双重在线存储（Two-Site Two-Phase Commit）
  - 层次 6：零数据丢失（Zero Data Loss），应用系统自动切换

  另外，衡量容灾备份性能有两个主要指标：

  - RPO（Recovery Point Objective）：数据恢复点目标，指的是业务系统所能容忍的数据丢失量。
  - RTO（Recovery Time Objective）：恢复时间目标，指的是所能容忍的业务停止服务的最长时间，也就是从灾难发生到业务系统恢复服务功能所需要的最短时间周期。

  RPO 针对的是数据丢失，而 RTO 针对的是服务丢失，二者没有必然的关联性。RTO 和 RPO 的确定必须在进行风险分析和业务影响分析后根据不同的业务需求确定。

  RTO 越短，RPO 越新，业务损失就越小，但相应的系统开发、建设成本就越高，需要综合权衡。

- 云计算需求

  数据业务的多个业务系统，一般不会同时并发业务，因此，在一台物理服务器上部署多个虚拟服务器，分别承担不同的业务系统，可以充分利用服务器资源，也是企业内部云计算的最容易开展起来的应用方式。在物理服务器上部署多个虚拟服务器时，需要考虑业务的带宽需求，避免同一个物理服务器上一个业务挤占了其他业务的带宽。网络的带宽规划上需要考虑虚拟服务器的业务需求。

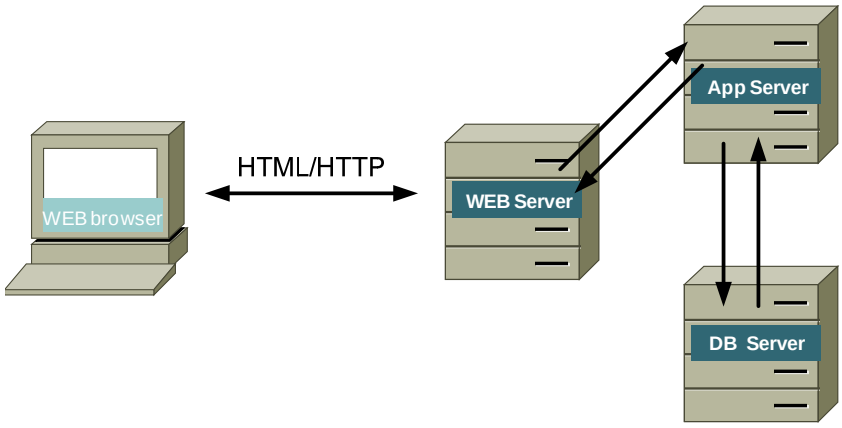
综上所述，数据业务对网络的主要需求是带宽保证（包括云计算场景）和安全。

2.3 WEB 业务

2.3.1 WEB 业务简介

随着互联网的兴起和发展，WEB 业务在企业的业务系统中所占的比重越来越大。这一方面是方便客户通过互联网访问企业信息，进行电子商务交易。另一方面，C/S 方式的一些缺点（客户端软件维护工作量大等），用 WEB 方式能够很好的解决，因此，越来越多的企业内部业务也开始通过 WEB 提供了。

图2-3 WEB 业务模式



WEB 业务模式增加了 WEB 服务器、应用服务器两层，变成了三层结构。其业务处理分工可以总结为：界面（WEB 服务器负责完成）、业务处理（应用服务器负责完成）和数据库（数据库服务器和存储系统负责完成）。

三层结构增加了业务系统部署的灵活性，业务系统的发展变化可以在 WEB 服务器上实现，也可以在应用服务器或数据库中实现，客户端不需要改变（只需要刷新一下 WEB 页面）。

2.3.2 WEB 业务网络需求

WEB 业务和数据业务相比较，在数据中心中增加了服务器部署，WEB 服务器和应用服务器之间、应用服务器和数据库服务器之间又数据流量。

对网络的需求主要反映在：

- 流量需求

  WEB 业务的流量既有客户端和服务器之间的数据请求和应答，也有服务器之间的流量。除了和数据业务一样有流量不均衡的特征外，还增加了服务器间的流量。

  在流量规划是需要首先看部署模型。WEB 业务一般有两种部署模型：分层部署模型和扁平部署模型，分别如图 2-4 和图 2-5 所示。

  - 对于大型数据中心，WEB、应用、数据库三类服务器数量都比较多，足够分别部署在不同分区，可以采用分层部署模型。
  - 对于中小型数据中心，服务器总数不多，则建议采用扁平部署模型。

图2-4 分层部署模型

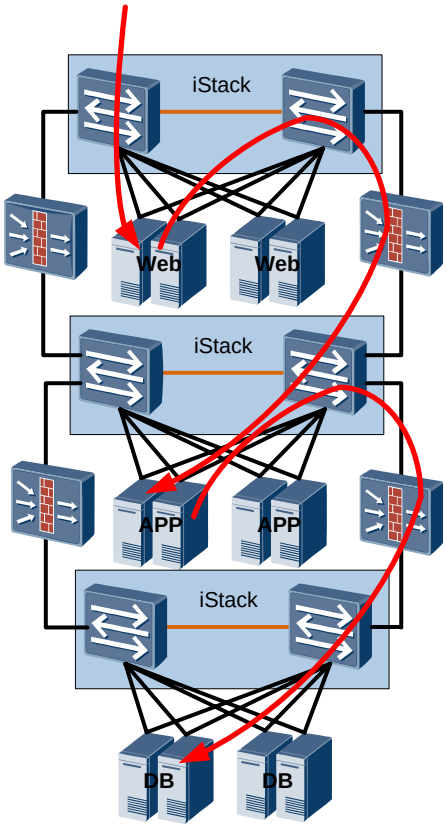
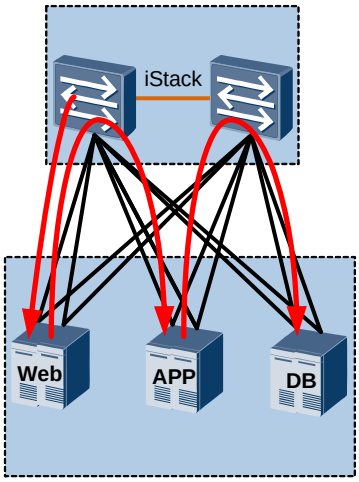


图2-5 扁平部署模型



分层部署模型在规划流量时只需要按各部分的流量进行规划。扁平部署模型则是将服务器间流量叠加到同一台设备上，需要按照总和进行规划。客户端和数据中心的（纵向）流量相比数据中心内部的（横向）流量要小得多。

由于 WEB 业务经过的服务器和网络设备相比数据业务有较大幅度的增加，因此，对网络时延的要求有一些提高。但由于 WEB 业务在和客户端的交互设计上有所变化，先由 WEB 服务器对客户端做出响应，再由 APP 服务器、DB 服务器处理后逐步完成 WEB 页面内容。因此，对时延的要求主要集中在 WEB 服务器对客户端的响应上。

- 安全性需求

  WEB 业务通过 WEB 服务器和 APP 服务器将客户端和数据库服务器隔离开，从业务模型上提高了数据库服务器及最终数据的安全性。但 WEB 服务器、APP 服务器和数据库服务器之间有逐跳的网络通道，也给攻击者提供了逐跳攻击的途径。

  WEB 业务特别是面向互联网用户的业务，带来更大的安全威胁：攻击来源更加组织化、产业化，可能来自互联网的如何一个角落；业务系统组成更加复杂，操作系统、WEB 服务器、通用 APP 引擎、数据库都可能有漏洞，任何一个漏洞都可能导致相关的几个系统逐个被攻破；内部用户在访问 Internet 过程中被入侵，成为攻击的桥梁。

- 可靠性要求

  WEB 业务的三层结构，将业务处理流程分解为多段，对客户端的业务提供需要三层服务器共同完成。交互的增加加大了对网络可靠性的要求。总体的故障恢复时间的要求虽然并没有提高，但按照串联系统的可用性，必须提高网络的可靠性，才能保证数据中心的可用性指标不降低。

  以交换机到服务器的链路故障概率为 1 小时/1 千小时。由于 WEB 业务涉及交换机和 WEB 服务器、交换机和 APP 服务器、交换机和 DB 服务器 3 条链路。则，故障概率就是 1－（1－1 小时/1 千小时）^{3}≈3 小时/1 千小时。要想仍然保持整个业务的故障概率为 1 小时/1 千小时，就必须将交换机到服务器的链路故障概率降低为 20 分钟/1 千小时。

综上所述，WEB 业务对网络的主要需求是带宽保证和安全需求。

2.4 计算业务

2.4.1 计算业务简介

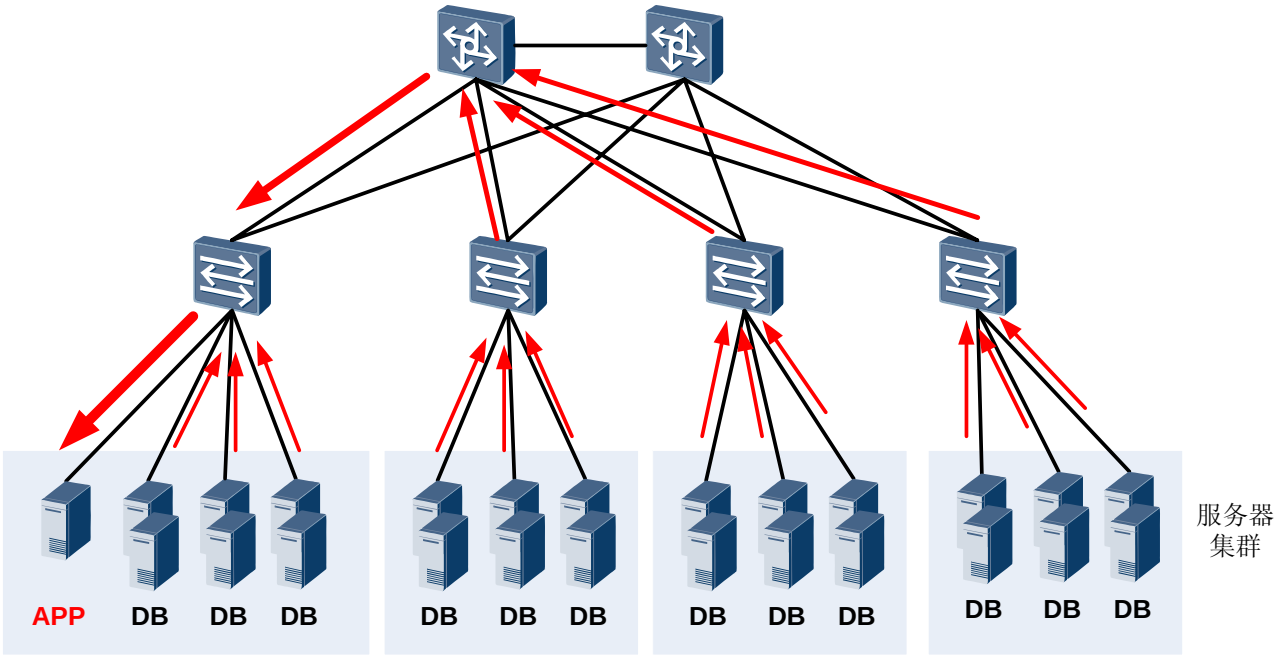
计算业务是需要高性能计算的业务，典型的有 3D 渲染、药物研究、基因分析、WEB 搜索业务等。

计算业务的典型模式是以大量的普通服务器协同工作组成集群，共同完成一项计算任务。

2.4.2 计算业务网络需求

计算业务的流量主要是服务器之间的流量，参考 WEB 业务举例如图 2-6 所示。

图2-6 计算业务流量示意图（服务器集群）



APP 服务器将计算业务分发给大量 DB 服务器处理，DB 服务器将处理结果返回给 APP 服务器。对网络的需求主要是两点：

- 网络对瞬时流量的缓存能力

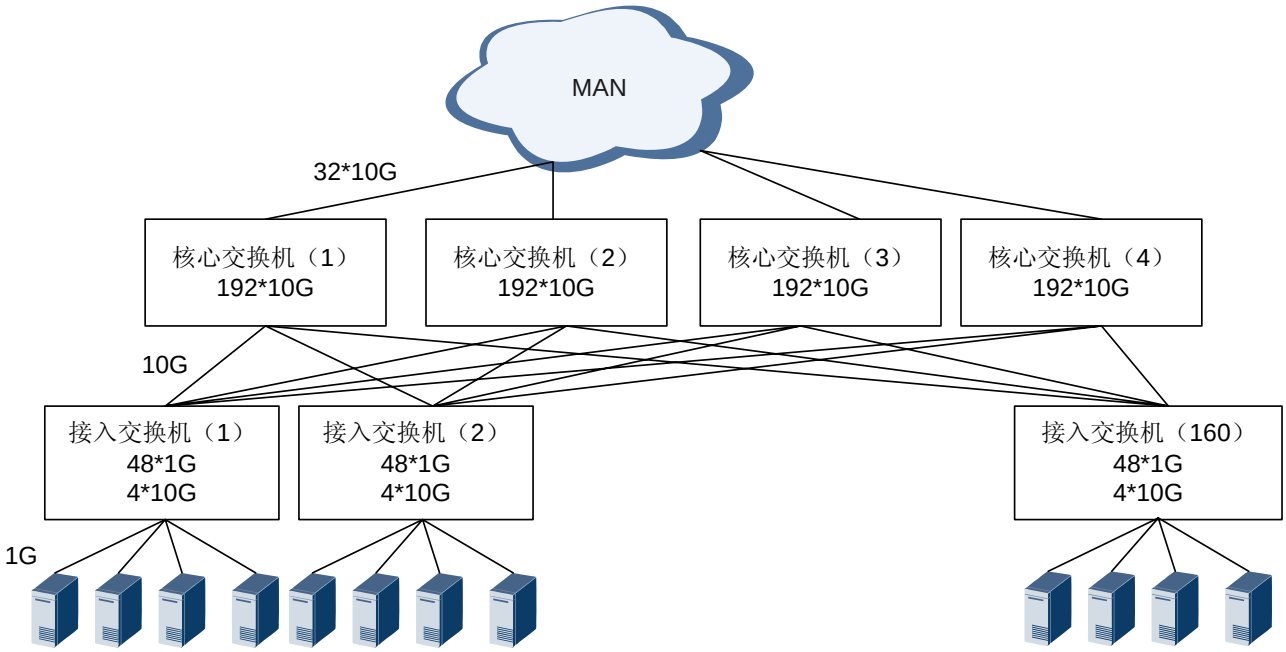
  除非 APP 服务器在业务分发时有很好的调度机制，否则，DB 服务器返回的处理结果会集中在一个很小的时间区间内同时到达 APP 服务器。数据流量瞬间会超过 APP 服务器的网络接口带宽，如果网络不能缓存流量，则必然造成丢包，导致 APP 无法完成完整的业务处理。这势必增加 APP 和 DB 间的交互，导致业务总的处理时间延长。因此需要网络进行流量缓存，保证不丢包。

- 网络无阻塞

  和上面的集群模型不同，另一种集群模型是服务器之间业务是全互联的。业务系统需要使用点对点的通信模式，任何两个服务器节点之间都有可能建立连接。

由于任意两个服务器之间都有业务交互，这就要求在网络带宽规划时，需要做到转发性能的位置无关性，也就是通常所说的无阻塞。

图2-7 互联网业务服务器集群



3 数据中心网络方案

3.1 数据中心网络架构

数据中心网络设计原则

数据中心网络的设计需要遵循以下设计原则：

- 模块化

  应考虑到业务的调整及发展，网络结构和系统结构要模块化、易于扩展。

- 高可靠

  网络设计中采用冗余网络设计，实现关键设备、链路冗余；关键设备选用高可靠性产品，可实现单板、模块热拔插、控制模块设计冗余、电源冗余；减少网络层级，简化网络结构，从组网架构上提高可靠性。

- 安全隔离

  数据中心网络应具备有效的安全控制。按业务、按权限进行分区逻辑隔离，对特别重要的业务采取物理隔离。

  以服务器为中心的业务、IP 存储备份、管理网络等多个网络进行逻辑隔离，管理网络采取物理隔离。

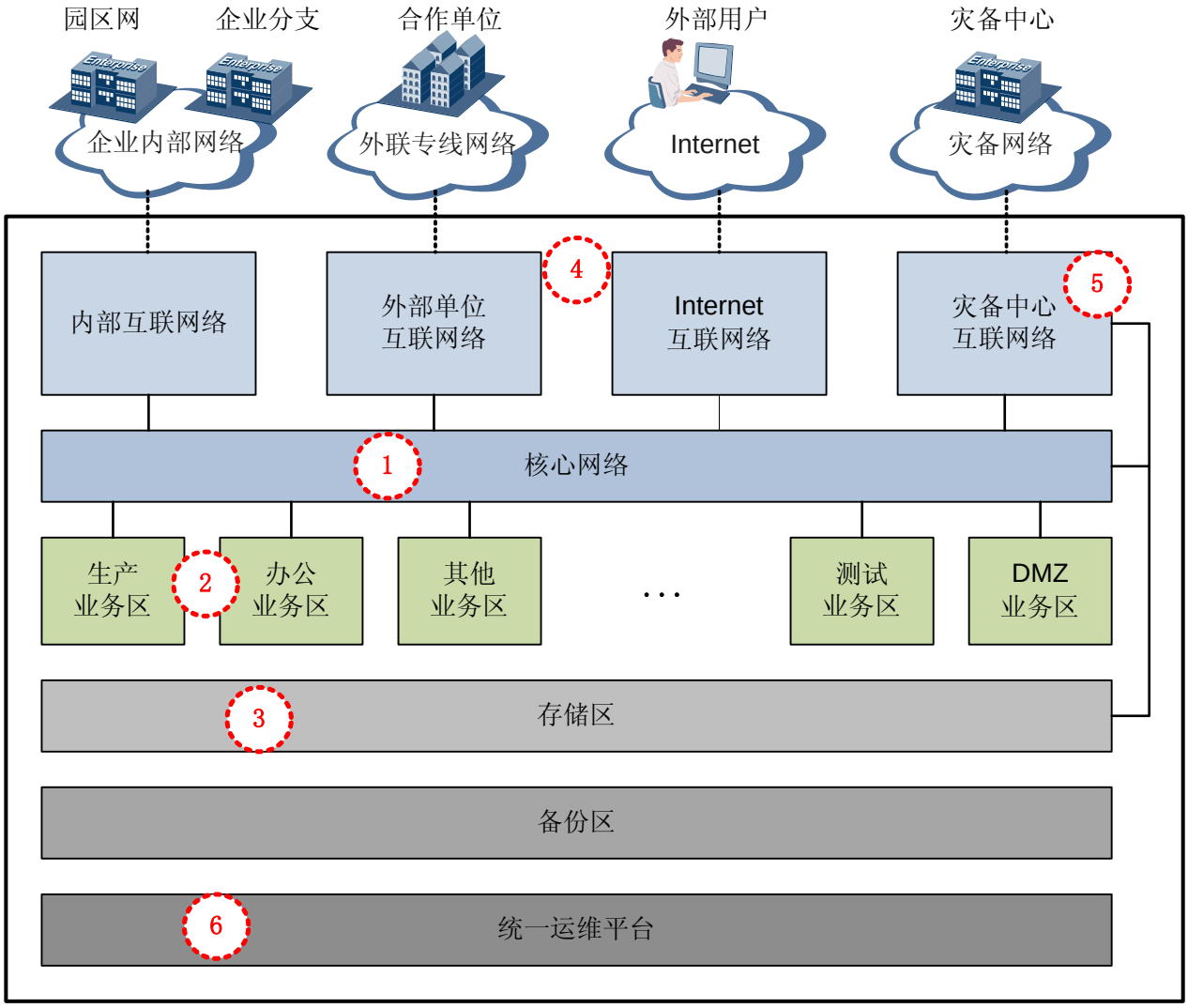
- 可管理性和可维护性

  网络应当具有良好的可管理性。为了便于维护，应尽可能选取集成度高、模块可通用的产品。

数据中心逻辑架构

数据中心的逻辑架构如下图 3-1 所示，包括六大部分。

图3-1 数据中心的逻辑架构



- 核心网络区

  是数据中心网络的核心，连接内部各个服务器区域、企业的内部网络、合作单位的网络、灾备中心和外部用户接入的网络等。

- 服务器区域

  部署服务器和应用系统的区域。出于安全和扩展性的考虑，可以根据应用的类型分为：生产业务区、办公业务区、测试业务区、DMZ 区等。

- 存储区域

  包括 FC SAN 和 IP SAN 的存储设备和网络。

- 互联网络区域

  是把企业内部用户和外部用户接入到数据中心的区域。出于安全和扩展性的考虑，根据互联的用户类型分为：内部互联网络，合作单位互联网络，Internet 互联网络。

  - 内部互联网络通过园区网、广域网和企业总部、分支机构的网络互联。
  - 合作单位互联网络通过城域专线、广域专线和合作单位的网络互联。

- Internet 互联网络实现互联网公众用户的接入、出差员工通过互联网安全接入、没有通广域网的办公点通过互联网安全接入。

- 灾备中心互联网络区域

  是实现灾备数据中心互联的区域，主要是以传输设备实现与同城灾备中心的互联，以广域网专线实现与异地灾备中心的互联。

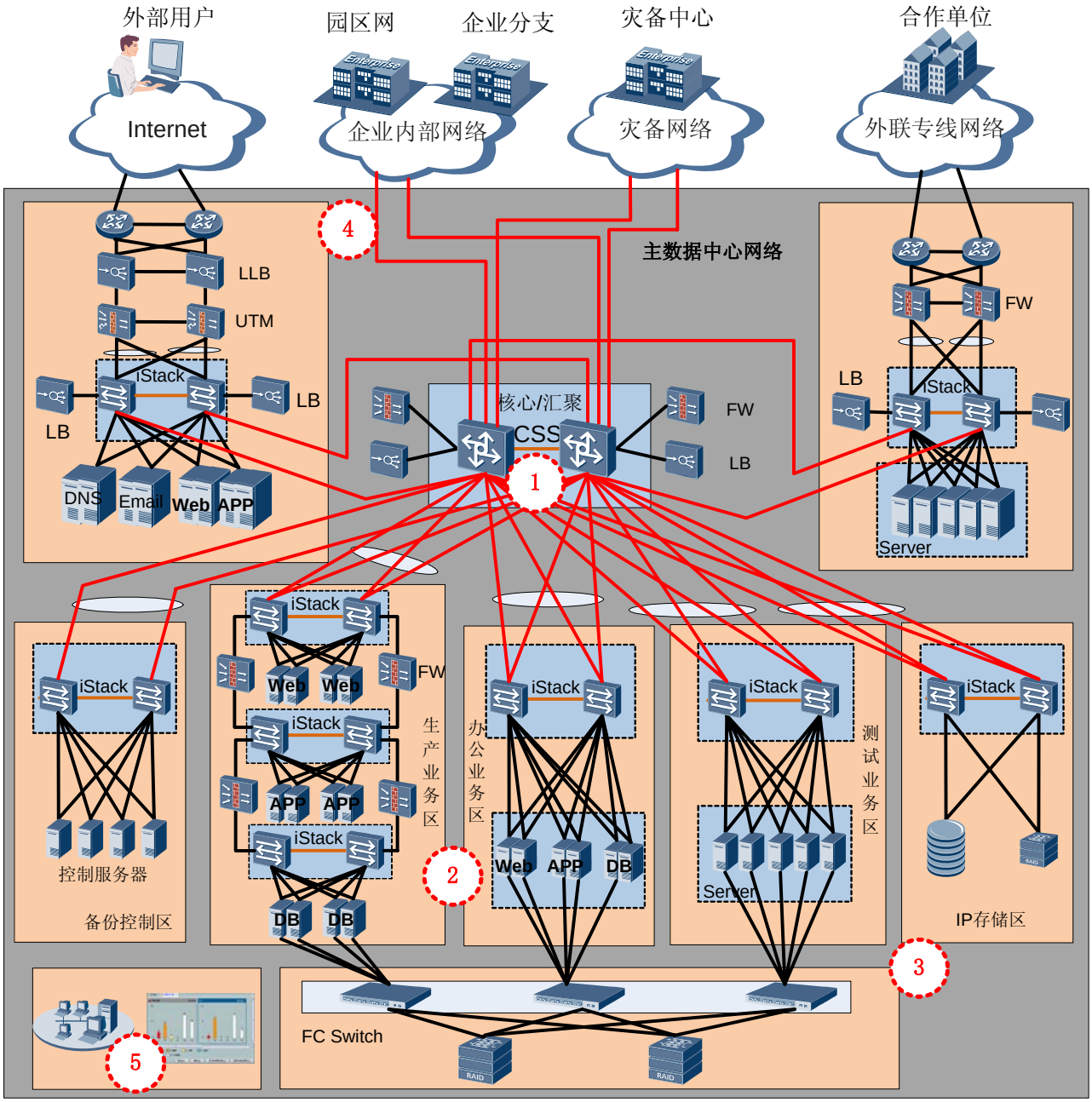
- 运维管理区

  对网络、服务器、应用系统、存储管理的区域。包括故障管理，配置管理，性能管理，安全管理等。

数据中心物理架构

对应逻辑架构，数据中心的物理架构如图 3-2 所示。

图3-2 数据中心物理架构



这个模块化数据中心的架构，具有如下特点：

- 整体架构可扩展
  - 分为 5 大区域（核心区、服务器区、存储区、互联区和管理区），各个区域独立扩展
  - 以核心节点为“根”的星型拓扑
- 核心区域：流量的枢纽
  - 采用大容量，高性能的核心交换机

  - 采用高密度的万兆接口
- 服务器区域/管理区域
  - 多个业务区独立扩展
  - 以服务器为中心的数据、管理、存储网络独立扩展
- 互联区域
  - 分为四个独立的互联区域，各区域独立扩展
  - 灾备互联网络，保证业务平滑向其他数据中心扩展

3.2 核心区网络规划

核心区是整个数据中心网络的枢纽，主要连接着服务器区和互联区。承担了内部数据流量和对外数据流量，在逻辑上成为可靠性、安全设计的中心。

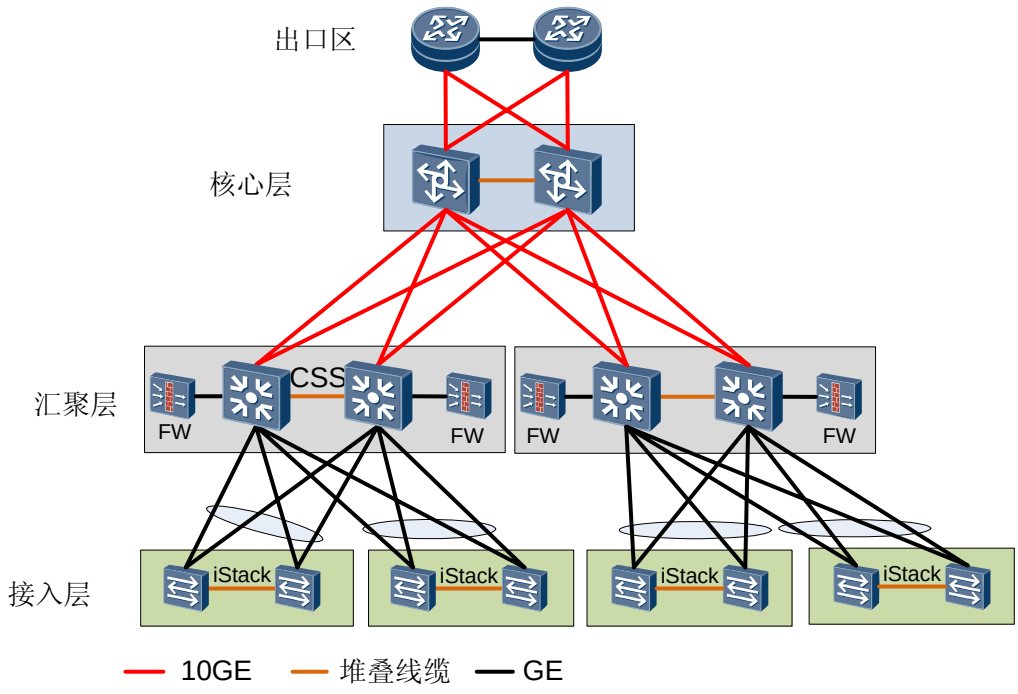
3.2.1 物理组网规划

核心区和服务器区的关系有两种物理组网方式，一种是核心/汇聚/接入三层方式，一种是将核心/汇聚整合的方式又称扁平化方式。

三层方式

三层方式组网图如图 3-3 所示。这种方式有核心层、汇聚层两层设备，每个汇聚区各自部署防火墙等安全设备。

图3-3 核心/汇聚/接入三层方式组网图

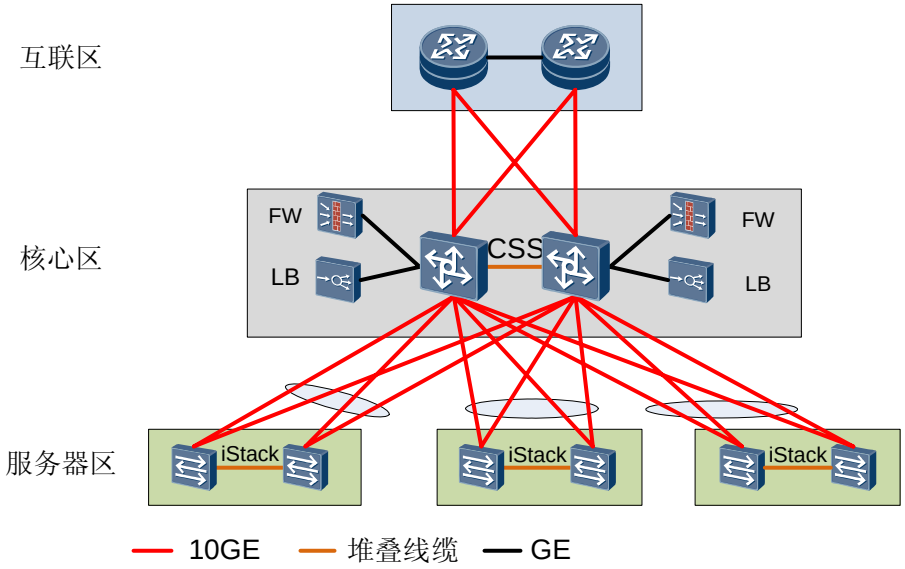


扁平化方式

扁平化方式的物理组网如图 3-4 所示。这种方式将核心/汇聚用一个超大容量交换机（物理上一般是两台设备）代替，在核心区集中部署超大规格的防火墙等安全设备。

扁平化方式降低了网络复杂度，简化了网络拓扑，提高了转发效率，是推荐的组网方式。

图3-4 扁平化方式核心区组网图



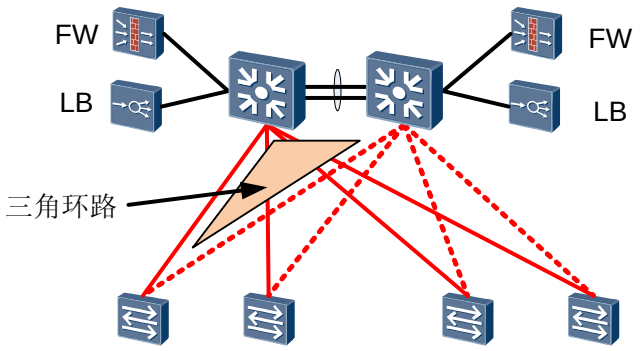
3.2.2 可靠性规划

从上述组网图可以看到，网络可靠性由双设备、链路冗余来保证。

对于双设备、链路冗余的网络，如果接入层进三层，这在接入层和核心层之间采用三层路由的方式，通过等价路径再辅助部署 BFD 快速检测故障，就能够保证链路故障、设备故障的快速切换，同时也能够充分利用冗余链路。

更多的组网是在核心层进三层，这样就需要解决接入层和核心层之间二层流量的环路问题。传统的方案是 STP＋VRRP 的方案，如图 3-5 所示。

图3-5 二层环路 STP 方案



如上图所示，虚线是被 STP 协议阻断的链路。这个方案的主要优点是采用的是标准的协议，多个厂家的设备可以混合组网。

这个方案的缺点也比较明显：

- 收敛时间

  传统的 STP 技术收敛速度慢， 在故障发生时，故障恢复收敛时间> 10 秒；虽然采用 RSTP 进行优化，但收敛时间大于秒级，秒级别的业务中断，会造成用户的体验降低。

- 链路利用率低

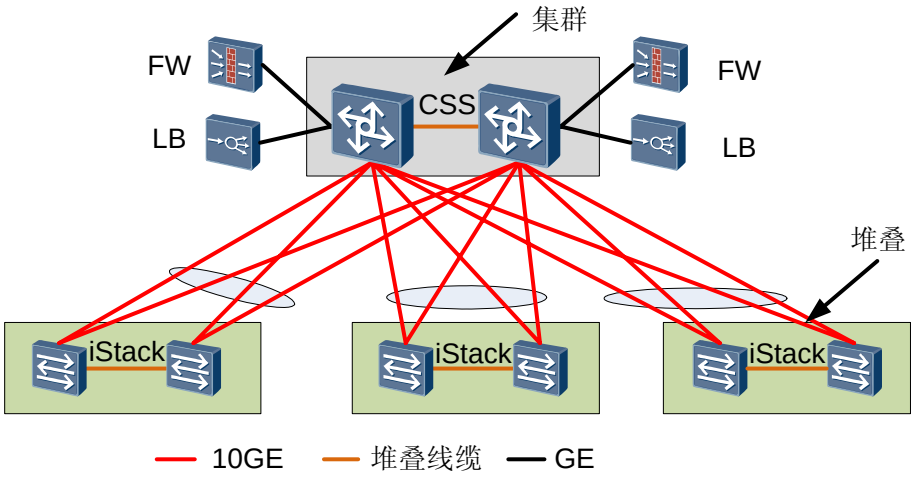
  如果同一机架内的服务器属于同一 VLAN，则有一个上行链路的带宽无法利用。带宽利用率只有 50%；虽然 MSTP 基于 VLAN 进行优化，但不能从根本上解决问题。

- 配置维护复杂，网络故障率高

  每个接入交换机和汇聚交换机都需要运行 STP 协议，随着接入交换机的增加，交换机需要处理的 STP 也越来越复杂，会导致出现可靠性问题。

集群+堆叠的无环网络方案可以解决上面这些缺陷。

图3-6 无环网络方案



核心/汇聚采用两台框式交换机集群。接入层采用盒式交换机，盒式交换机每两台堆叠。接入层交换机和核心/汇聚层交换机间的链路进行链路捆绑。。

这个方案有四大优势：

- 减化管理和配置

  首先，集群和堆叠技术将需要管理的设备节点减少一半以上。

  其次，组网变得简洁不需要配置复杂的协议，包括 STP/SmartLink/VRRP 等。

- 快速的故障收敛

  链路故障收敛时间可以控制在<10ms，大大降低了网络链路/节点的故障对业务的影响。

- 带宽利用率高

  采用链路 Trunk 的方式，带宽利用率可以达到 100%。

- 扩容方便、保护投资

  随着业务的增加，当用户进行网络升级时，只需要增加新设备既可，不需要更改网络配置的情况下，平滑扩容，很好的保护了投资。

该方案极大提高了可靠性，以单链路故障率为 1 小时/1 千小时为例，增加到两条链路，就可以将故障率降低到 3.6 秒/1 千小时，可靠性从 3 个 9 提高到 6 个 9。

可靠性的另一个重要方面是设备可靠性，核心区设备一般为框式设备，在可靠性方面的要求包括：

- 主控单元的备份；
- 支持电源模块的备份；
- 需要提供模块化的风扇设计，支持单风扇失效；
- 支持所有模块的热插拔；
- 支持 CPU 防攻击；
- 需要提供完善的各种告警功能。

3.2.3 安全规划

核心区部署防火墙，主要解决几个安全问题：

- 服务器区不同分区间互访的控制，不同业务间的安全隔离
- 园区网和服务器区之间互访的控制，客户端和服务器安全分组隔离
- 分支机构和服务器区之间互访的控制，客户端和服务器安全分组隔离

3.3 服务器区网络规划

3.3.1 物理组网规划

接入层交换机部署在服务器机架内或者独立的网络机柜中，部署在服务器机架内的一般称为 TOR（Top Of Rack），部署在列头柜中的一般称为 EOR（End Of Row），一般提供二层交换功能。

TOR 的部署模式一般适合高密度的机架服务器的接入；EOR 模式一般适合低密度的服务器，如小型机的接入。两者的区别如所示。

表3-1 EOR 接入模式和 TOR 接入模式的区别

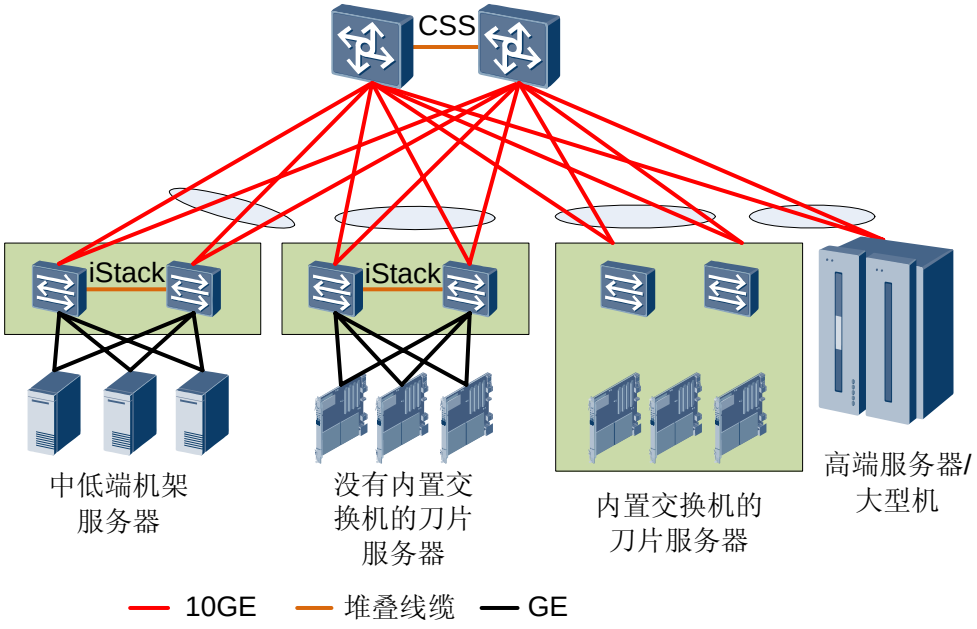
项目	EOR 接入模式	TOR 接入模式
接入交换机部署位置	专门的网络机柜	与服务器同机柜
交换机类型	框式交换机 盒式交换机	盒式交换机

项目	EOR 接入模式	TOR 接入模式
布线	服务器到 EOR 交换机的布线距离长、布线数量大	服务器到 TOR 布线在本机柜内、距离短、 到核心交换机布线数量少
管理	EOR 服务器接入密度高、数量少、管理成本低	TOR 数量多、管理成本高
适合场景	机柜内服务器密度低	机柜内服务器密度高

服务器的接入方式分为下面四种情况:

- 中低端机架服务器，数量众多，通过接入层交换机接入；
- 高端服务器/大型机，数量较少且重要性高，直接接在核心/汇聚层交换机上，保证带宽；
- 没有内置交换机的刀片服务器，通过接入层交换机接入；
- 内置交换机的刀片服务器，直接接在核心/汇聚层交换机上，减少交换网络的层级，提升网络性能；

图3-7 不同服务器的接入方式



3.3.2 服务器多通道分离规划

从服务器的发展趋势谈起。服务器的 CPU 技术多核化的趋势，从单核到目前的 128 核，使得 CPU 处理能力大大提高。但服务器的 IO 的性能发展比较缓慢，远远赶不上 CPU 的发展。这就造成了 IO 的瓶颈。所以，服务器必须采用多通道的方式，采用物理隔离的多个网络接口，才能充分发挥 CPU 的高性能。

图3-8 服务器多通道分离示意图

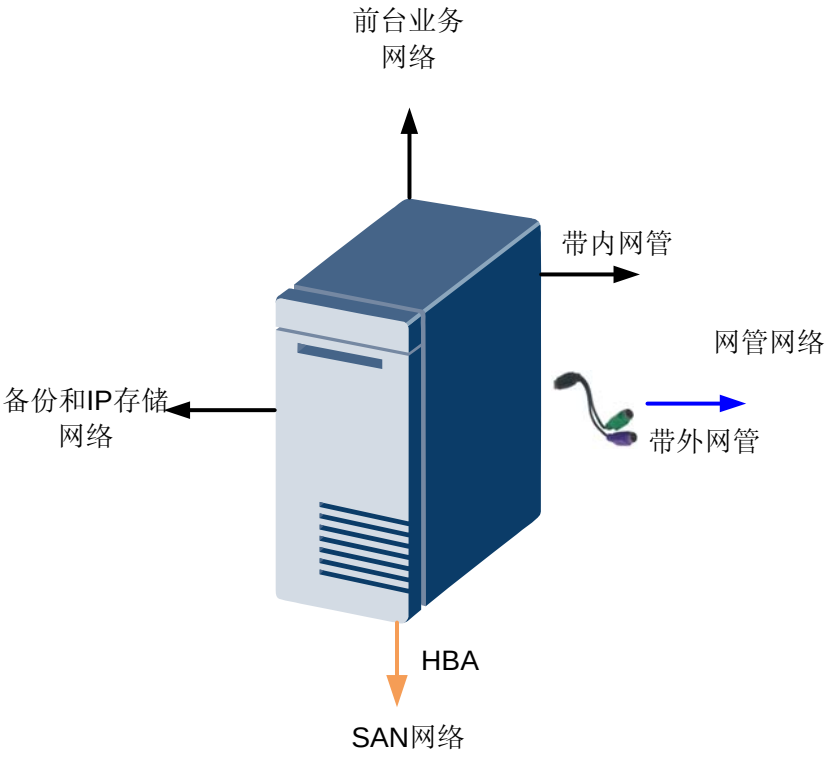


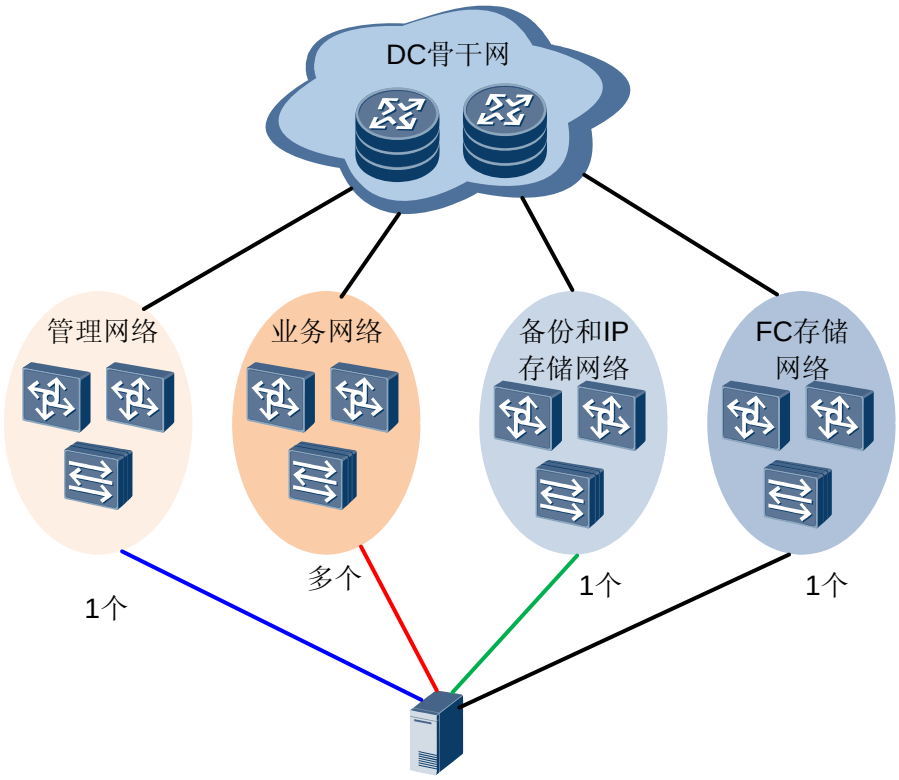
图 3-8 为服务器多通道分离示意图。服务器包括四类接口，分别接入到业务网络 、网管/KVM 网络、SAN 网络、备份和 IP 存储网络。

服务器的多通道有两个优势：

- 提高 IO 能力
- 不同类型的业务流量安全隔离

服务器多通道，相应的网络逻辑结构如图 3-9 所示。

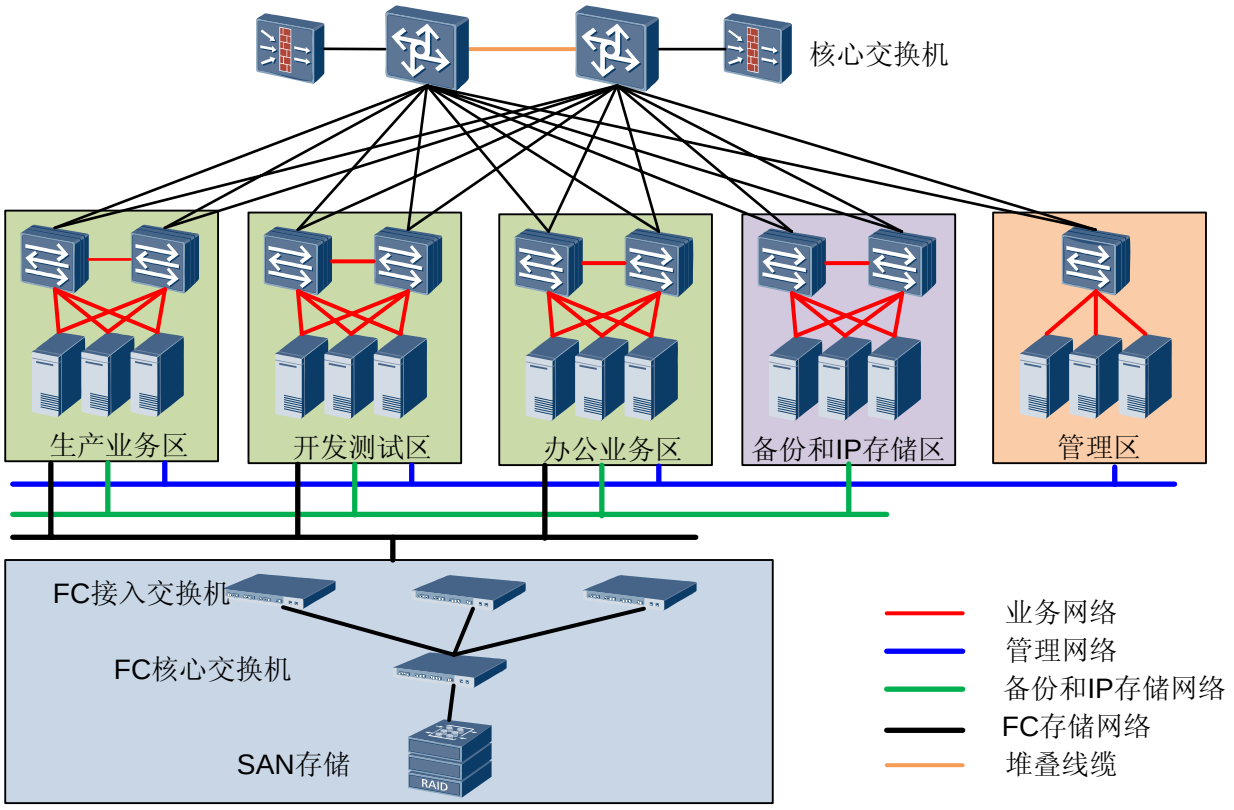
图3-9 四网分离网络



整个服务器区的网络架构是四网分离的架构，也就是说，网络分为：业务网络、管理网络、存储网和备份网络，四张网络物理隔离。服务器通过不同的网卡分别接入不同的网络。

网络物理拓扑图如图 3-10 所示。

图3-10 网络物理拓扑图



3.3.3 可靠性规划

总体可靠性规划

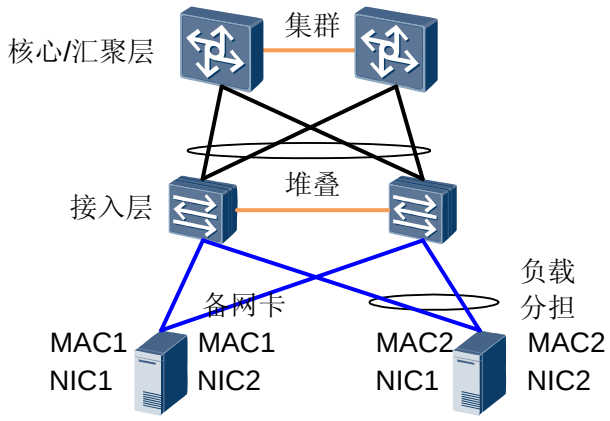
服务器区的可靠性包括网络可靠性、设备可靠性和服务器可靠性：

- 网络可靠性通过集群＋堆叠的无环网络提供，具体见核心区可靠性规划。
- 设备可靠性采用接入交换机堆叠。
- 服务器可靠性是通过服务器双网卡来支持。

  服务器网络驱动程序将多个网卡捆绑成一个虚拟的网卡，对外提供一个唯一的 IP 地址。需要服务器支持网卡聚合特性（NIC Teaming）：当一个网卡失效，另一个网卡接管它的 MAC 地址。两个网卡采用主备或者负载分担的方式。

双网卡主备方式

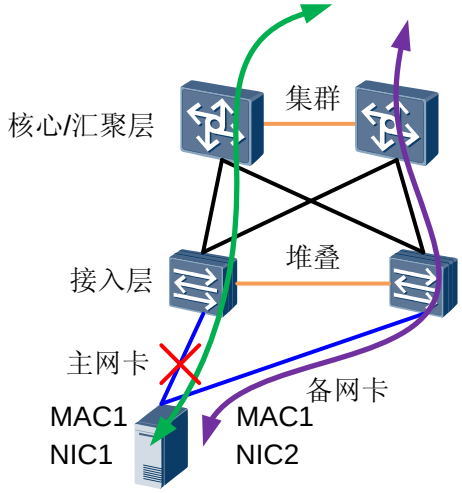
图3-11 服务器可靠性组网图



对于主备方式的双网卡，两个网卡的 MAC 相同（如上图，都是 MAC1）。服务器在发现主网卡故障后，切换到备网卡。并通过备网卡发出免费 ARP。网络设备必须正确处理这个免费 ARP 报文，才能将发给服务器的流量切换到新的转发路径上。

如图 3-12 所示，主网卡故障后，转发路径需要从绿色曲线切换到紫色曲线。

图3-12 主备方式故障切换



接入层交换机在处理免费 ARP 报文时，需要将 MAC1 的出接口刷新到连接备网卡的链路上，因此要求接入层交换机配置时将对应服务器主备网卡的两个端口配置在同一个 VLAN，不配置成链路捆绑（否则不会刷新 MAC1 的出接口）。

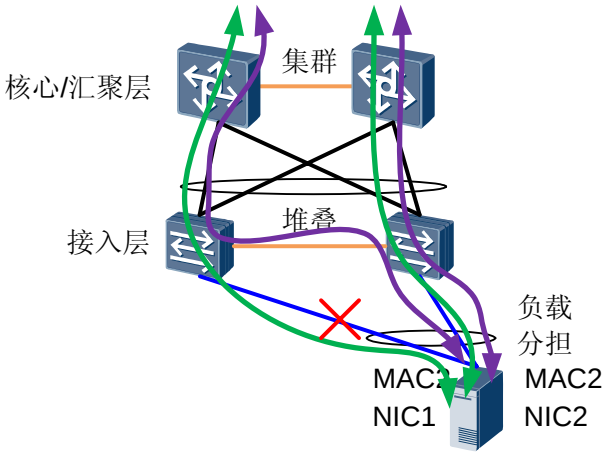
核心/汇聚层交换机在处理免费 ARP 报文时，由于核心/汇聚层交换机和接入层交换机之间的是多条链路捆绑成的 Trunk 链路，因此，核心/汇聚层交换机不会感知到变化。

双网卡负载分担方式

对于负载分担方式的双网卡，两个网卡的 MAC 相同（如图 3-13，都是 MAC2），而且两个网卡都可以发送和接收流量。接入层交换机必须配置成堆叠，并将对应服务器主备网卡的两个端口配置成链路捆绑。才能屏蔽 MAC 地址在两个交换机端口间不断“跳跃”的处理。

如图 3-13 所示，没有故障时转发路径时绿色曲线，两个网卡都有流量。左边网卡故障后，转发路径需要从绿色曲线切换到紫色曲线。

图3-13 负载分担方式故障切换



由于核心/汇聚层交换机和接入层交换机之间的是多条链路捆绑成的 Trunk 链路，因此，核心/汇聚层交换机感知不到接入层的变化，仍然会将流量发给左边的接入层交换机。这个流量通过接入层交换机之间的堆叠链路转发给右边的接入层交换机，由右边的接入层交换机转发给服务器。

3.3.4 安全规划

服务器区用 VLAN 隔离不同业务，同一业务的 WEB、APP 和 DB 也建议用 VLAN 隔离开。

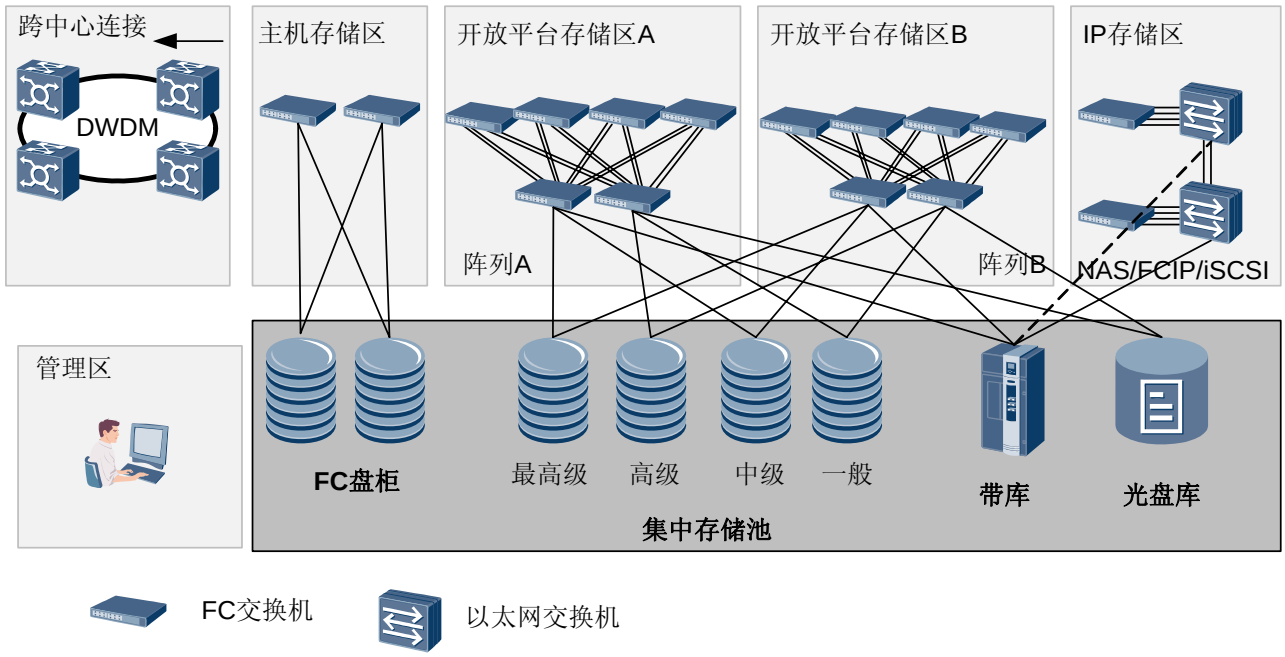
3.4 存储区网络规划

3.4.1 物理组网规划

存储网络包括 IP 存储网和 SAN 存储网。

IP 存储网络用于承载服务器访问的 NAS 存储的业务流量。NAS 主要承载 2 种业务类型：第一种为特定应用服务器与 NAS 存储之间的数据交互的流量，第二种为虚拟化业务产生的大量网络流量。

图3-14 存储区网络架构



存储区域基本规划

- 小型机存储区：单独分区
- 开放应用平台区
  - 采用双陈列、每个陈列双核心结构，保障高可用性
  - 边缘设备与核心设备可采用多条链路捆绑连接，实现低超载比
  - 集中存储池按服务等级分类
- IP 存储区：单独分区

  IP 存储区可部署压缩加速设备，对异地灾备进行 FCIP 流量的压缩。通过 IP/MPLS 网络备份同步。由于虚拟化的需求，极大的增进了服务器与存储的交换数据。
- 管理区：对存储网络及存储资源进行管理、调配
- 同城灾备：通过 DWDM 连接同城数据中心或灾备中心

主备数据中心之间要实现数据的实时或者准实时交互，建议：

- 对于 IP 存储部分，如服务器之间的交互数据，可以通过目前运营商的 MPLS VPN 或者 VPLS 虚拟专线，实现主备两个数据中心之间的互通；
- 对于 SAN 存储部分，则建议采用裸光纤甚至 DWDM，提供主备数据中心之间高速、低时延的互联互通，以满足存储数据的准实时备份需求。

由于虚拟化的需求，极大的增进了服务器与存储的交换数据，对于采用 NAS 方式的 IP 存储网络，至少要保证接入交换机采用 10G 的链路接入。

3.4.2 可靠性规划

IP 存储网络的可靠性规划和服务器区业务网络的可靠性规划相同。可靠性的设计依然采用集群＋堆叠的无环网络方案，具体请参见 3.3 服务器区网络规划。

3.4.3 安全规划

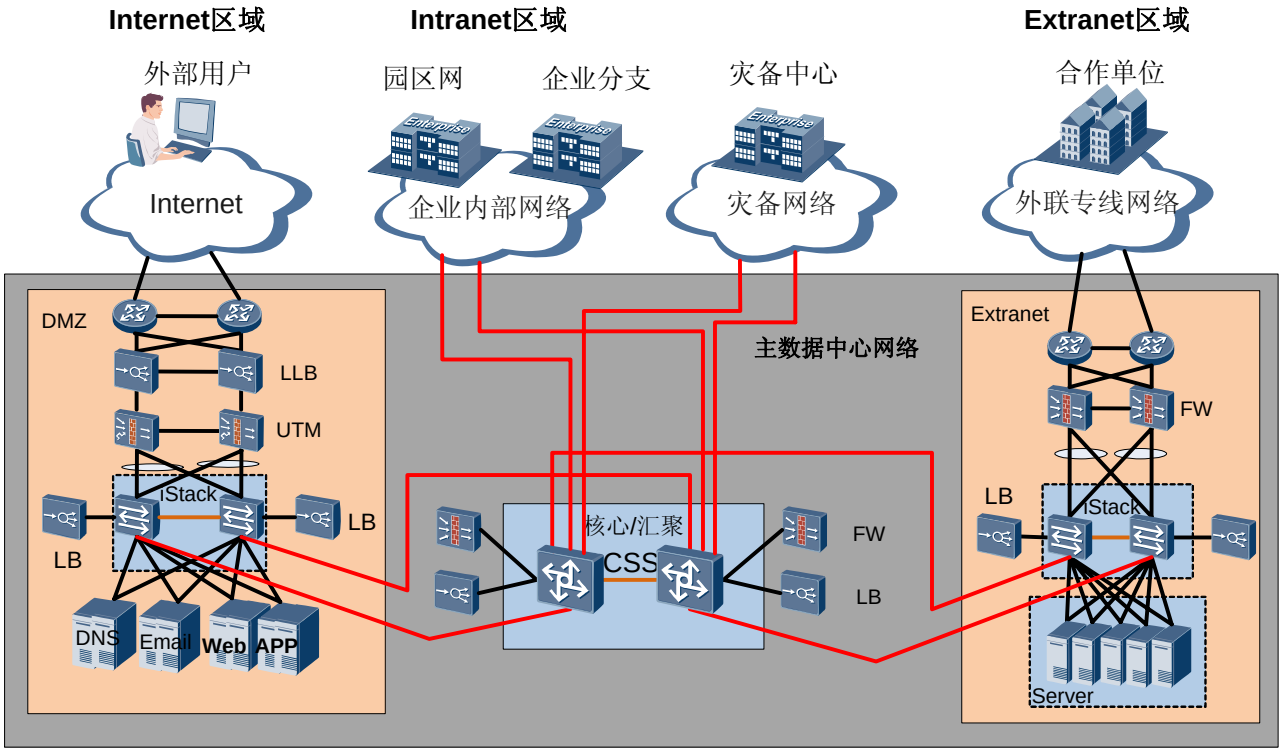
SAN 存储使用专用技术实现隔离。

IP 存储网络通过使用 VLAN、VPN 等技术对于用户及不同的 IP 存储区域进行划分，严格控制存储的访问权限。

3.5 互联区网络规划

3.5.1 物理组网规划概述

图3-15 互联区域网络



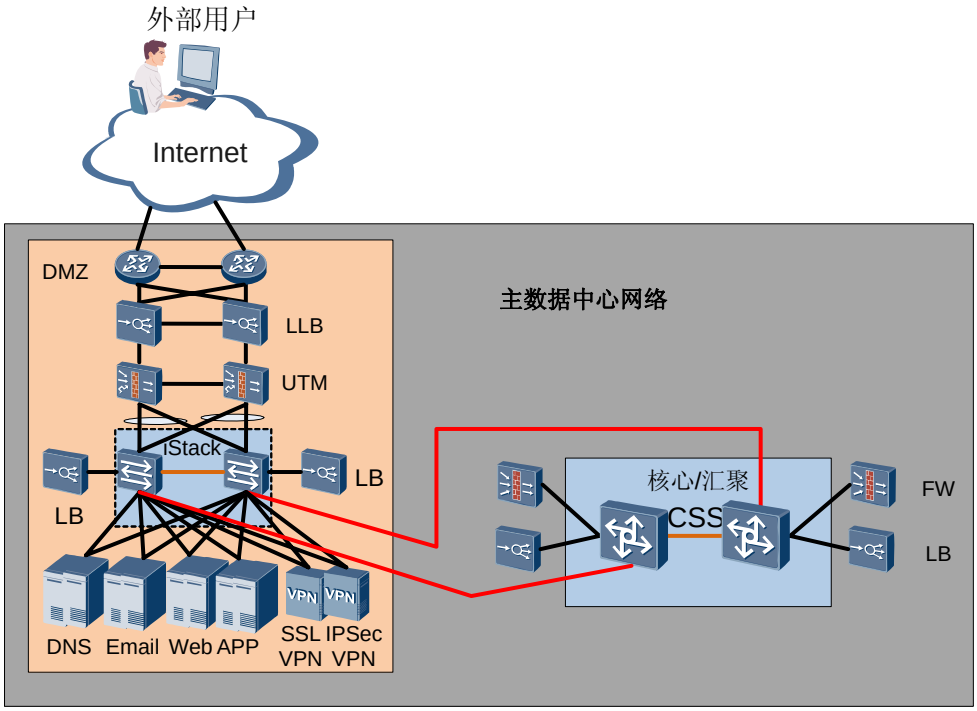
根据接入类型及服务类型划分多个不同的互联接入区域：

- Intranet 互联
  企业内部用户通过广域或局域网访问数据中心
- Internet 互联
  企业外部用户通过 Internet 访问数据中心
- Extranet 互联
  合作单位用户通过广域或局域网访问数据中心

说明
对应使用 Internet 网络的 VPN 用户，可在 Internet 区域内单独规划区域。

3.5.2 Internet 互联

图3-16 Internet 互联区域的网络架构



如图 3-16 所示，Internet 互联区包括路由器、链路负载均衡、UTM 等设备。其中 UTM 至少要包括防火墙和 IPS 两项功能。

- 链路负载均衡设备用于在租用了两个运营商出口时，对来自不同运营商的请求从相应出口回应。在只有一个运营商出口时不需要部署。
- 入侵检测系统 IPS（ Intrusion Detection System）对掺杂在应用数据流中的恶意代码、攻击行为、DDOS 攻击等进行侦测，并实时进行响应。
- 防火墙在网络层面，过滤非法流量、抵御外部的攻击，保护内部资源。

防火墙和 IPS 本身都是重要的网络设备，而且其位置一般都是作为网络的出口。其位置和功能决定了防火墙和 IPS 设备应该具有非常高的可靠性。

为了保证 Internet 互联区域的可靠性，所有设备均需要成对部署，即：两台路由器，两台链路负载均衡设备，两台 UTM（至少含防火墙和 IPS）。

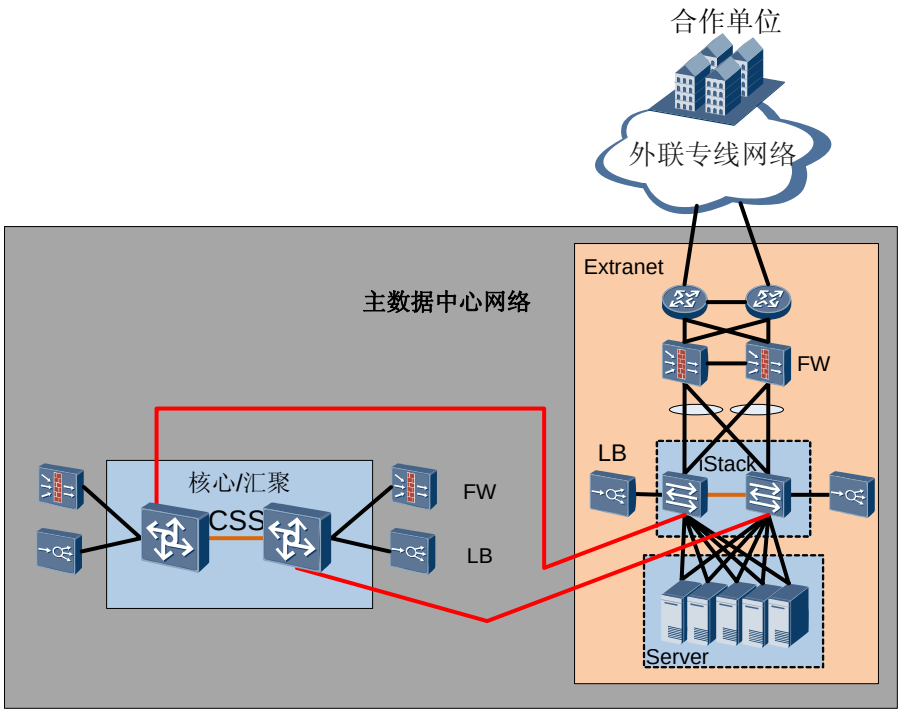
VPN 接入区根据需要提供 IPsec VPN 和 SSL VPN 两种接入功能，解决移动用户的安全接入问题。

- IP Sec VPN 主要适用 Site-to-Site 方式接入
- SSL VPN 主要适用于 Client-to-Site 方式接入

可以部署独立的 IPsec VPN 网关和 SSL VPN 网关，也可以采用 UTM 设备统一接入。

3.5.3 Extranet 互联

图3-17 Extranet 互联区域的网络架构



如图 3-17 所示，是 Extranet 区组网图。

由于 Extranet 区域属于企业外部用户接入的区域，从网络信任关系上讲，安全等级与 DMZ 相同，都属于非可信网络，不能直接与内部数据中心连接。访问权限应限制在本区域内部及 DMZ 区域，内网访问应严格控制。

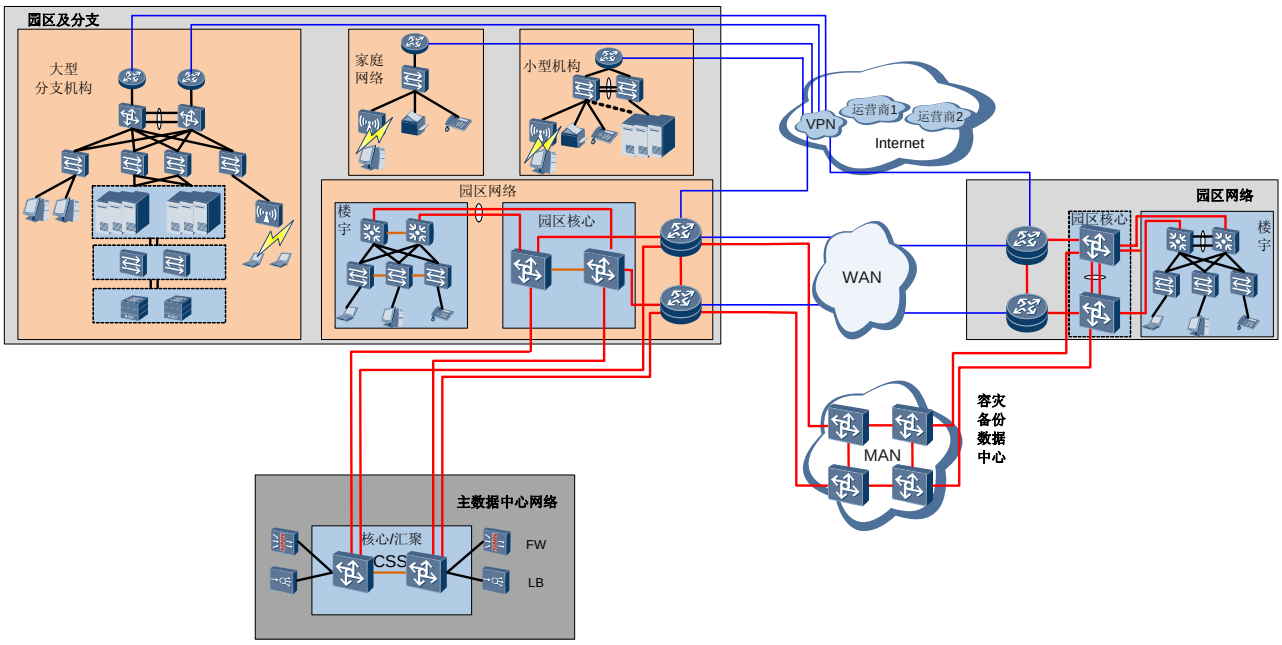
业务隔离：同 DMZ 区域，主要提供对外服务，因此对该区域网络应与内网隔离，必要的业务可以通过严格控制访问 DMZ 区域。

防火墙：在网络层面，通过 NAT 技术隐藏内网拓扑，保护内部资源控制访问权限。

3.5.4 Intranet 互联

企业内部用户通过广域或局域网访问数据中心。

图3-18 Intranet 互联区域的网络架构



网络方面主要考虑线路双归，路由及设备的冗余备份。

多分支机构网络互联应考虑多出口线路备份，并在出口考虑路由备份、负载分担，同时对于广域网链路考虑还应考虑 QoS 保证线路及不同业务的服务质量。

要部署独立的互联接入设备并部署 2 台进行热备，提供设备的可靠性。

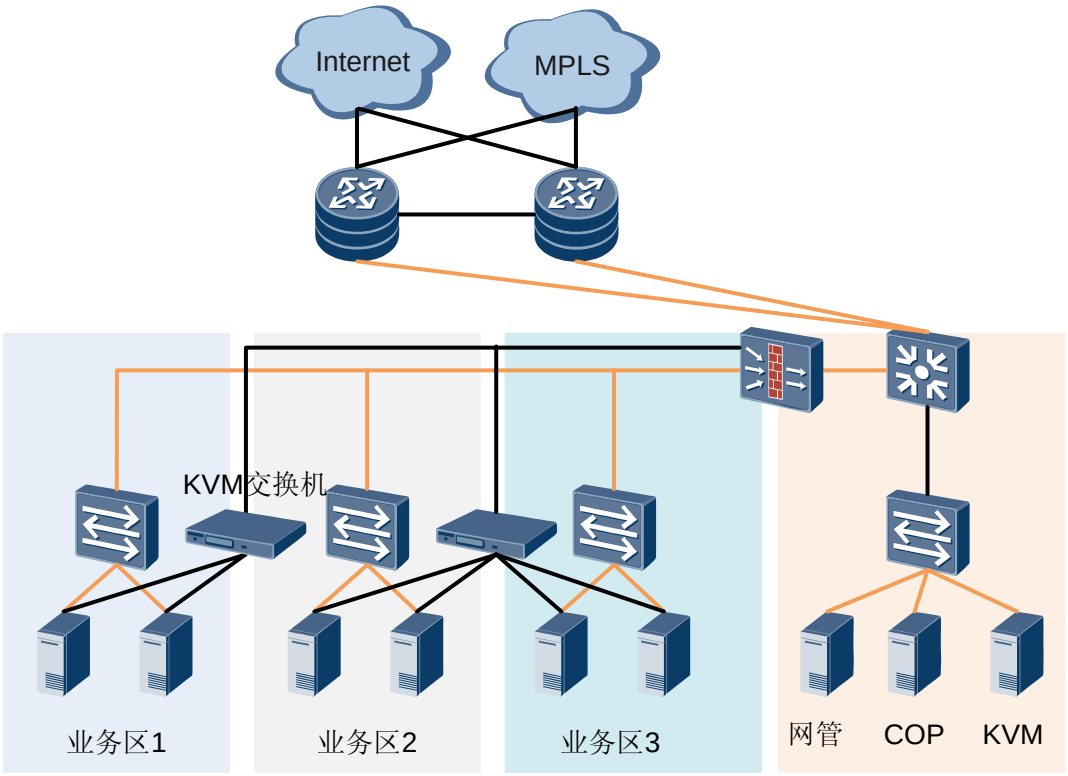
内部网络属于比较安全的区域，是绿色区域，风险比较低。主要的安全风险来自内部网络自身的用户，如用户未经授权的存取。在接入设备中，根据实际需求控制不同分支之间的数据互通的访问控制。

3.6 管理区网络规划

3.6.1 物理组网规划

管理网络总体要求：带外管理，分权访问，安全审计。

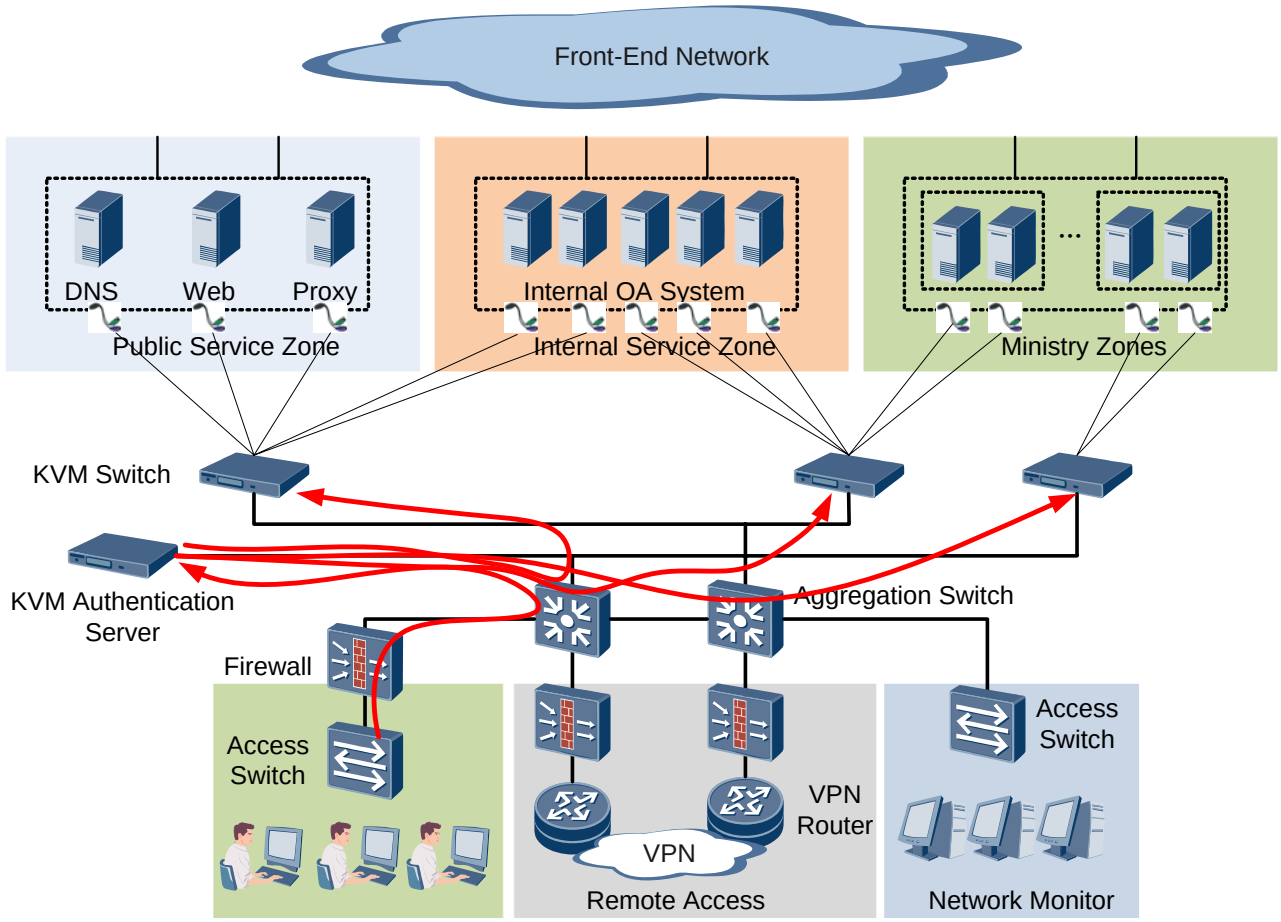
图3-19 管理网络架构



管理网络网络架构如图 3-19 所示。管理网络通过管理网口及 KVM 交换机连接所有设备，进行网络管理、数据收集和实时监控等。

非管理员用户不能访问管理网络，管理网络通过 VPN 或防火墙等隔离手段与内部数据中心连接。管理网络根据管理员的分工不同，授予其不同的访问权限，从而限制其访问不同的设备。

图3-20 KVM 管理网络示意图



网络管理：实现了对交换机、路由器、防火墙等设备的全方位管理，提供了丰富的拓扑、配置、资产、故障、性能、事件、流量、报表等网络管理功能。

流量管理：提供网络流量监测、流量门限、协议分析、Web 上网行为审计等功能。结合 NetFlow 网络流量分析器实现更为细化、便捷的全网流量分析功能。

应用管理：实现了对多种系统及上层应用监控管理功能，包括服务器、数据库、邮件服务器、WEB 服务器、应用服务器、操作系统、网站监控等。

3.7 VLAN 规划

3.7.1 VLAN 概述

VLAN 是将 LAN 内的设备逻辑地而不是物理地划分为一个个网段，从而实现在一个 LAN 内隔离广播域的技术。既隔离了广播域，减少了广播风暴，又增强了信息的安全性。当网络规模越来越庞大时，局部网络出现的故障会影响到整个网络，VLAN 的出现可以将网络故障限制在 VLAN 范围内，增强了网络的健壮性。

3.7.2 VLAN 规划原则

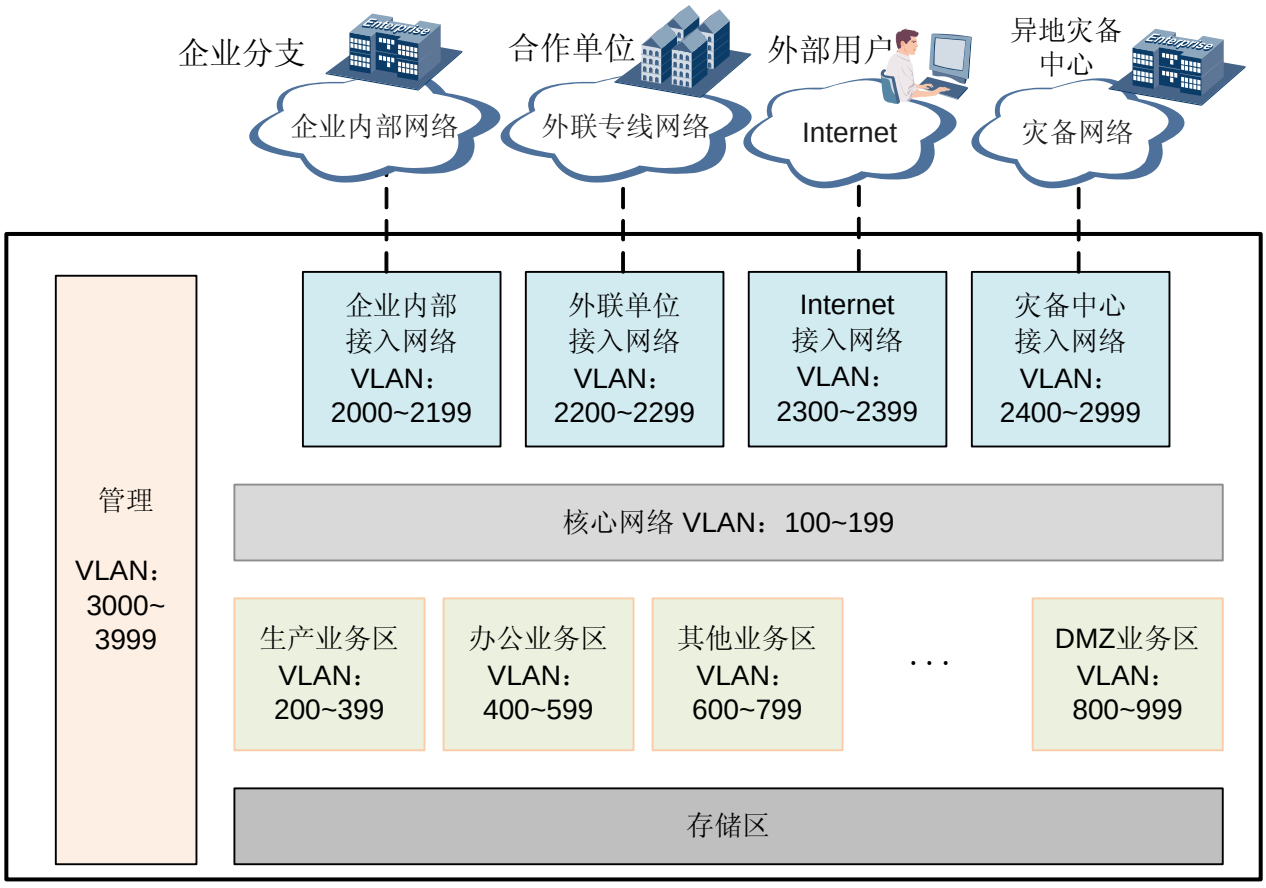
- 区分业务 VLAN、管理 VLAN 和互联 VLAN
- 按照业务区域划分不同的 VLAN
- 同一业务区域按照具体的业务类型（如 WEB、APP、DB）划分不同的 VLAN
- VLAN 需连续分配以保证 VLAN 资源合理利用
- 预留一定数目 VLAN 方便后续扩展

3.7.3 VLAN 规划建议

如图 3-21 所示，按照区域划分 VLAN 范围：

- 核心区：100～199
- 服务器区：200～999，预留 1000～1999
- 接入网络：2000～2999
- 管理网络：3000～3999

图3-21 VLAN 规划示意图



3.8 IP 规划

数据中心中在 Internet 互联区有少量设备使用公网 IP，数据中心内部使用的则是私网 IP。由于私网 IP 地址空间很大，如 10.0.0.0 是 1 个 A 类，因此在数据中心内部 IP 地址规划时主要以易管理为主要目标。

3.8.1 IP 地址规划

- 唯一性

  一个 IP 网络中不能有两个主机采用相同的 IP 地址。即使使用了支持地址重叠的 MPLS/VPN 技术，也尽量不要规划为相同的地址。

- 连续性

  连续地址在层次结构网络中易于进行路径叠合，大大缩减路由表，提高路由算法的效率。

- 扩展性

  地址分配在每一层次上都要留有余量，在网络规模扩展时能保证地址叠合所需的连续性。

- 实意性

  “望址生意”，好的 IP 地址规划使每个地址具有实际含义，看到一个地址就可以大至判断出该地址所属的设备。

3.8.2 DNS 规划

DNS 服务器的角色

DNS 系统中的 DNS 服务器分为如下的角色：

- Master 服务器：主服务器

  作为 DNS 的管理服务器，可以增加、删除、修改域名。修改的信息可以同步到 Slave 服务器。一般部署 1 台。

- Slave 服务器：从服务器

  从 Master 服务器获取域名信息，采用多台服务器形成集群的方式，统一对外提供 DNS 服务，一般采用基于硬件的负载均衡器提供服务器集群的功能。一般部署 2 台从服务器。

- Cache 服务器：缓存服务器

  缓存内部用户的 DNS 请求结果，加快后续的访问。一般部署在 Slave 服务器上。

DNS 服务器的 IP 地址

- Master 服务器：采用企业内网地址。
- Slave 服务器：分配企业私网地址。并在负载均衡器上分配一个虚拟的企业内网地址。

Internet 域名地址有两种方案，一种是在防火墙上做 NAT 映射，把 Slave 服务器的虚拟地址映射为一个公网 IP 地址，用于外部 Internet 用户的访问；另一种是在链路负载均衡设备上通过智能 DNS 为外部 Internet 用户提供服务。

用 Slave 服务器为 Internet 用户提供 DNS 服务

图3-22 数据中心 DNS 部署

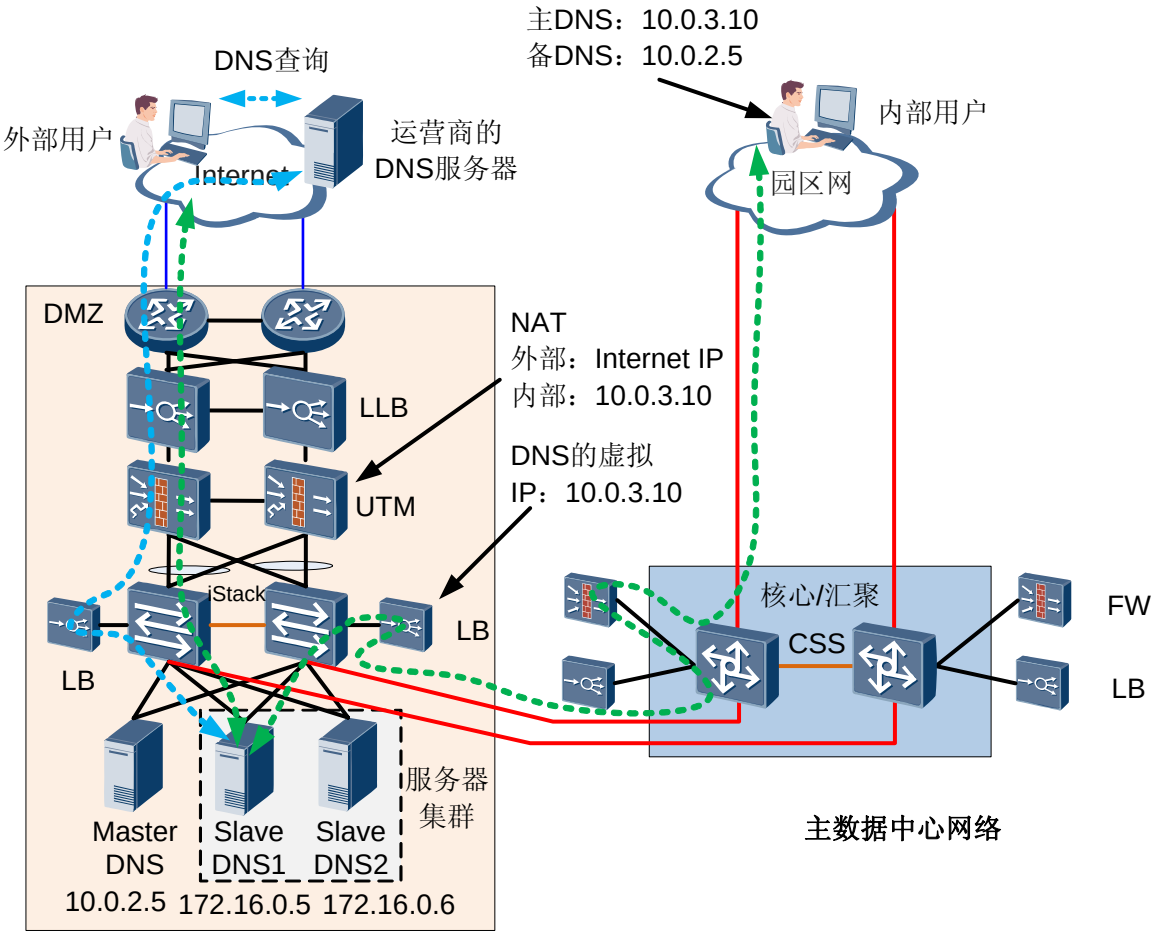


图 3-22 中绿色是用 Slave 服务器为 Internet 用户提供 DNS 服务的情况。

数据中心的 DNS 可靠性设计：众多内部用户发送 DNS 请求，被均匀分担到 Slave DNS1 和 DNS2。当 Slave DNS1 服务器故障后，所有的 DNS 请求被分发给 Slave DNS2。最终 DNS 服务器必须与外部 DNS 通讯。

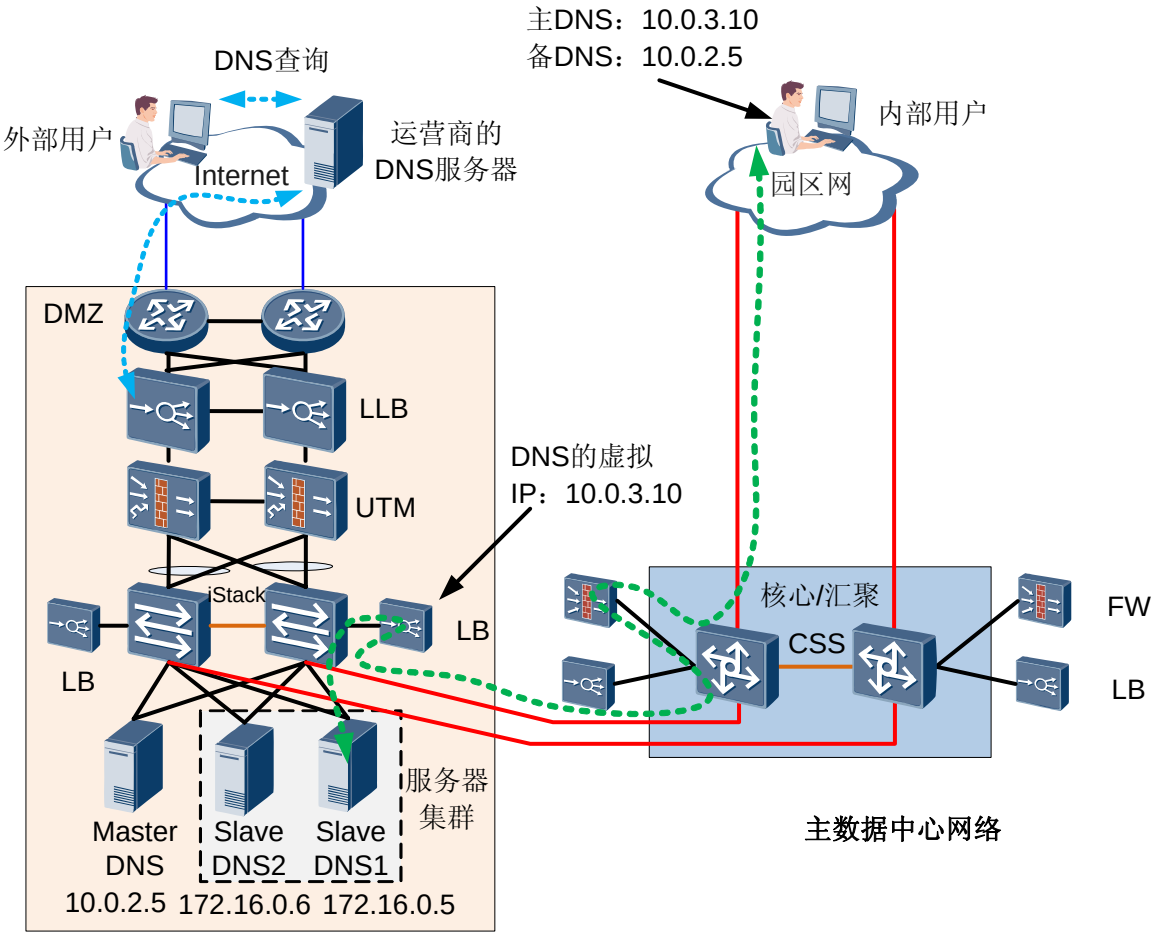
主 DNS 服务器，建议放置在 DMZ 区域，并在同区内部建立备份 DNS 服务器。如只对内提供服务的 DNS 服务器，可以作为二级的 DNS 服务器，放入其他非 DMZ 区域。

当所有的 Slave DNS 都故障后，用户发送的 DNS 请求无响应。用户就切换到备 DNS，由 Master DNS 处理所有的请求。

通过智能 DNS 为外部 Internet 用户提供服务

在链路负载均衡设备上通过智能 DNS 为外部 Internet 用户提供服务的方案如下图：

图3-23 采用智能 DNS 为 Internet 用户服务



外部 Internet 用户向运营商的 DNS 服务器发起 DNS 请求查询华为公司的域名，比如 www.huawei.com，运营商的 DNS 服务器发现是 huawei.com，就请求华为公司数据中心内部的 DNS 服务器来进行域名解析。如上图中蓝色线条所示。

链路负载均衡器中的智能 DNS 收到外部请求，完成 DNS 解析。

智能 DNS 通过判断访问用户的来源，把域名分别解析成不同的 IP 地址。如访问者是网通用户，DNS 策略解析服务器会把域名对应的网通 IP 地址解析给这个访问者。如果用户是电信用户，DNS 策略解析服务器会把域名对应的电信 IP 地址解析给这个访问者。

同时，智能 DNS 还通过对运营商侧链路质量的检测，当某个运营商的链路中断后，就完全返回另一个运营商的 IP 地址，实现业务连续。

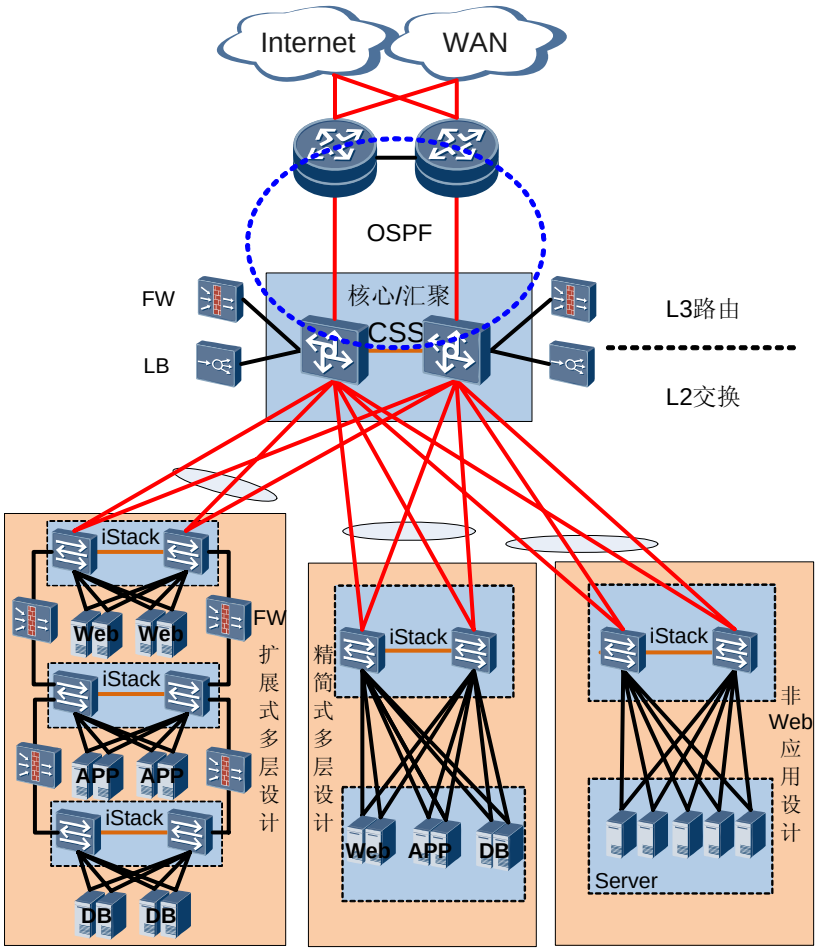
3.9 路由规划

3.9.1 路由概述

如图 3-24 所示，建议采用在汇聚/核心交换机上作为路由和交换的分界点：

- 在汇聚/核心交换机以下，使用 L2 交换设计
- 在汇聚/核心交换机之上，采用 L3 路由设计

图3-24 路由交换分界点设计



这种设计方法有如下的优点：

- 路由配置简单

  只需要在 2 台汇聚/核心交换机上，配置路由。 在大量的接入交换机上，只是做二层交换，配置简单。便于采用接入交换机的“自动配置”功能。减少配置维护工作量。

- 扩展性好

  在同一个汇聚/核心交换机下的服务器扩容方便。

一个业务系统新增的服务器，可以部署在任意的机架中，服务器的 IP 地址与原来业务系统的 IP 地址连续，可以统一规划。

并且，随着业务的变化，服务器位置调整后，不需要更改服务器和网络的配置，即插即用。

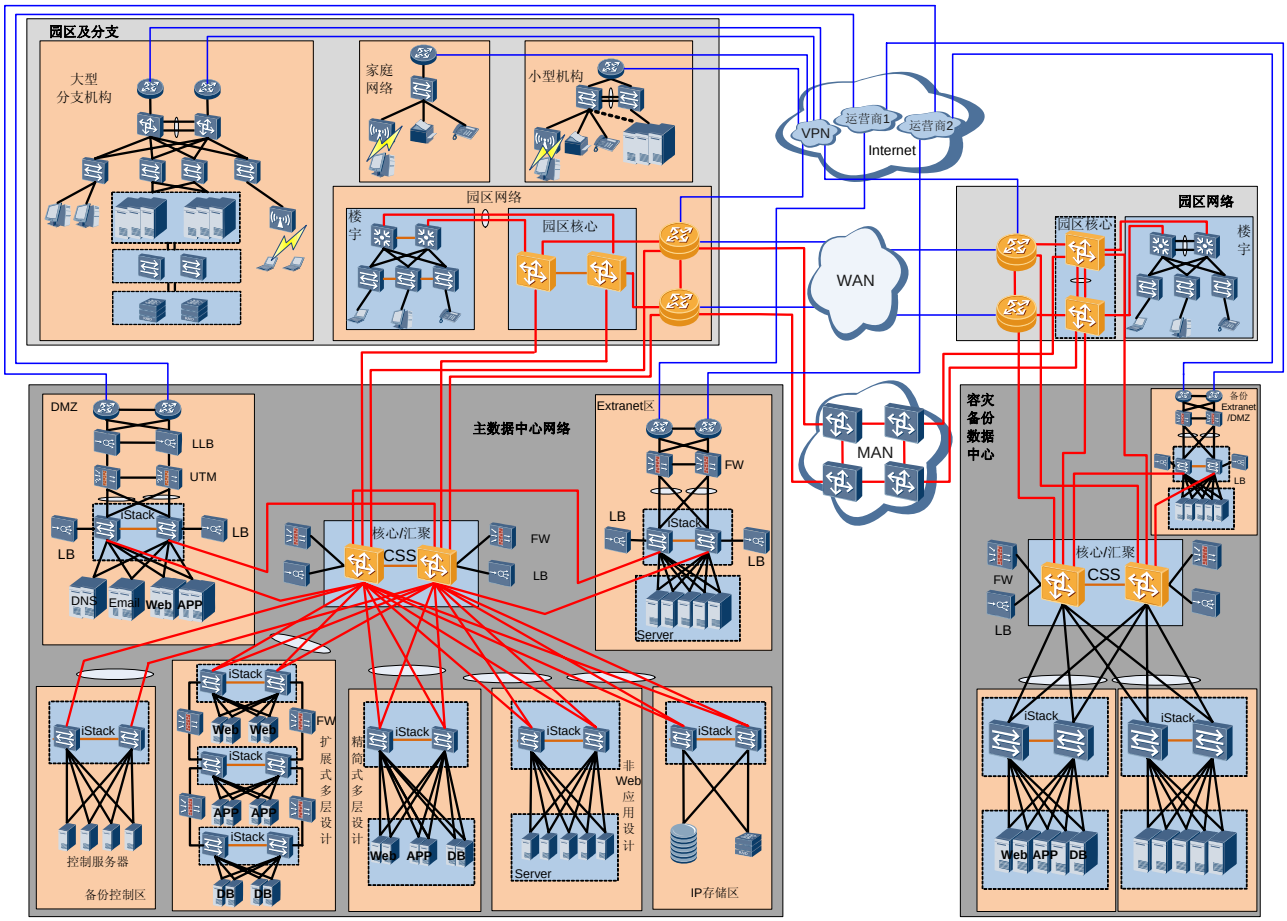
另外，随着下一代服务器的“虚拟化”，也需要大的 L2 网络，便于进行“在线迁移”。

3.9.2 IGP 设计

在数据中心内部，考虑到网络的稳定性和路由的快速收敛，方便以后的维护和管理，建议采用 OSPF 动态路由协议。

如图 3-25 所示，黄色图标所示的数据中心的两台核心交换机、园区网骨干组成 OSPF 的骨干区域（Area 0）。

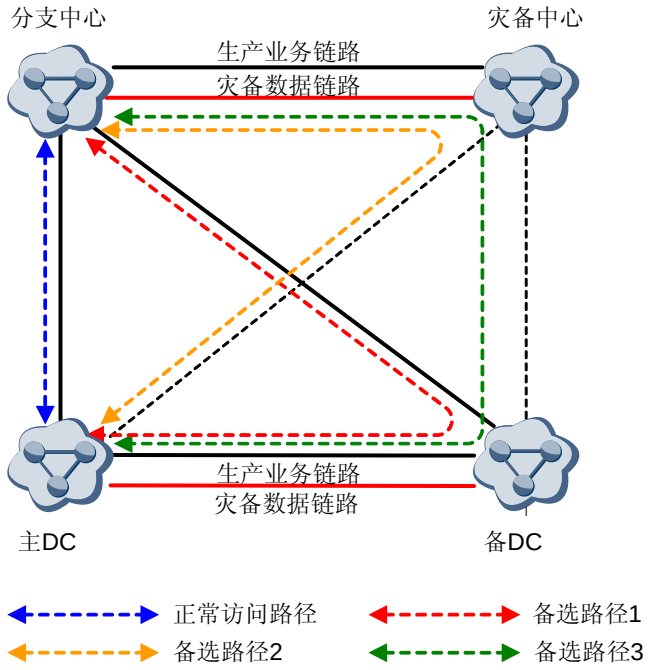
图3-25 数据中心路由规划



3.9.3 BGP 设计

数据中心的内部网络接入路由器与企业的分支数据中心、灾备中心之间建立 EBGP，进行路由通告。主数据中心、备份数据中心、分支数据中心、灾备中心形成的网络拓扑如图 3-26 所示。

图3-26 数据中心间主备路径规划



从主 DC 到分支中心，如图 3-26 所示有 4 条路径，规划的优先级如下：

- 主 DC 直接到分支中心，优先级最高
- 主 DC 经备份 DC 到分支中心，优先级第二
- 主 DC 经灾备中心到分支中心，优先级第三
- 主 DC 经备份 DC，再经灾备中心到分支中心，优先级最低

通过 EBGP 的 AS-Path 和 MED，可以实现 4 条路径的优先级控制。

3.10 VPN 及业务区隔离规划

3.10.1 VPN 概述

VPN 主要实现业务隔离，访问控制，安全隔离的作用。

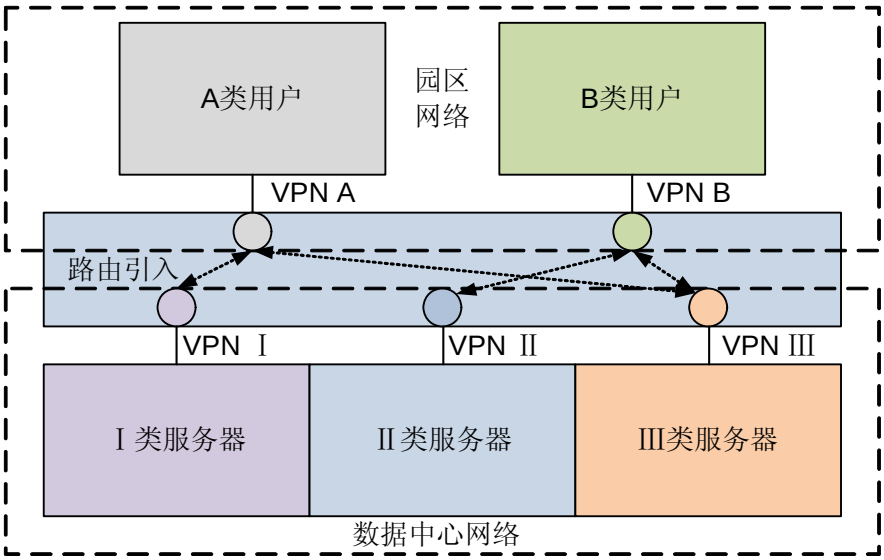
BGP/MPLS IP VPN 是一种 L3VPN（Layer 3 Virtual Private Network）。它使用 BGP（Border Gateway Protocol）在服务提供商骨干网上发布 VPN 路由，使用 MPLS（Multiprotocol Label Switch）在服务提供商骨干网上转发 VPN 报文。IP VPN 中的 IP 则是指 VPN 承载的是 IP 报文。

3.10.2 内部业务 VPN 业务隔离规划

如图 3-27 所示，首先把划分的用户类和服务器类都设置到不同的 VPN 中。这样，默认情况下，A 类用户、B 类用户、I 类服务器、II 类服务器、III 类服务器之间的路由都是隔离的，之间不能访问。

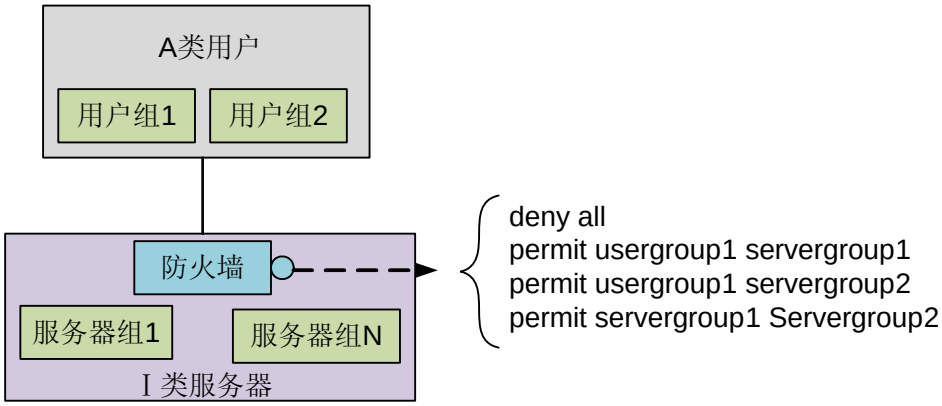
根据制定的“用户和服务器控制策略”，按需在用户 VPN 和服务器 VPN 之间互相引入路由。如 A 类用户能访问 I 类和 III 类服务器，就在 VPN A、VPN I 和 VPN III 之间相互引入路由。

图3-27 数据中心服务器基于路由的隔离方案



如图 3-28 所示，对于服务器组的精细的权限控制，采用防火墙的方式进行控制。根据制定的“用户组和服务器组的权限表”，在防火墙上配置具体的安全策略。防火墙按照“默认关闭，设置打开”的策略，保证只有设置了安全策略，用户才能访问。

图3-28 数据中心服务器基于防火墙的隔离方案



3.11 QoS 规划

3.11.1 QoS 概述

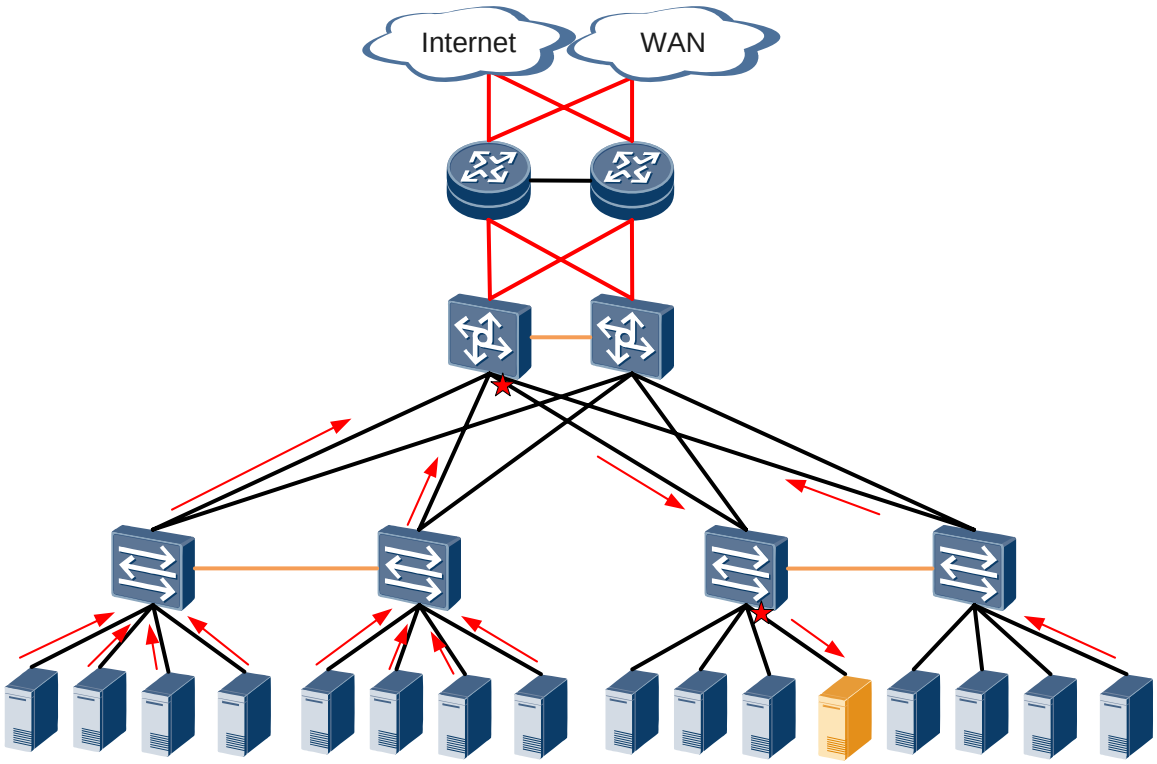
总的来说，数据中心规划建设是是要保证峰值业务的需求的，不需要再做 QoS 的处理。某几个特殊应用情况需要考虑 QoS。包括：协同计算、多租户。

其中多租户的情况主要是对多个租户的带宽管理，不属于版本初始范围，后续再补充完善。

3.11.2 协同计算的 QoS 规划

在协同计算（如搜索引擎计算、石油勘探计算、气象计算等复杂科学计算需要将计算任务分发到多台服务器协作进行）业务场景下，会存在多个服务器在瞬间同时向一个服务器发送计算结果数据，从而产生突发流量，这会引发在某个网络节点上的某个出端口上的数据出现拥塞而丢包。

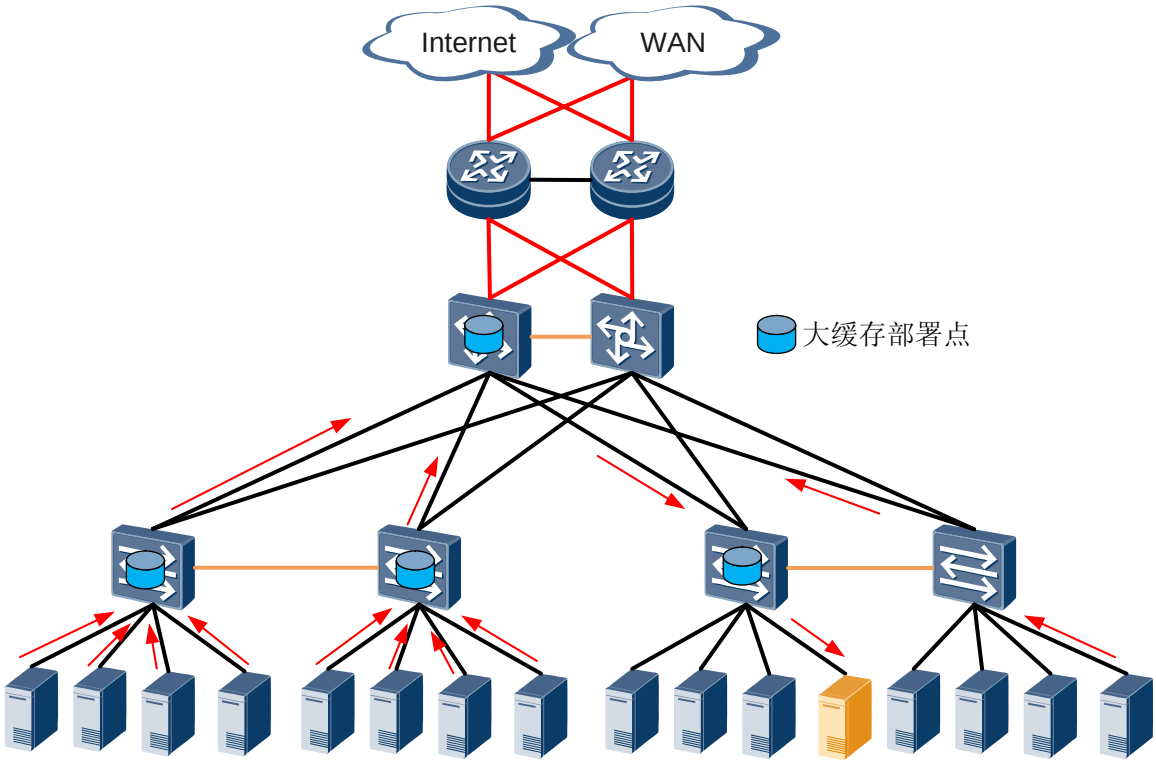
图3-29 多个端口向一个端口发送流量时的拥塞



如图 3-29 所示，蓝色的服务器向黄色的服务器反馈信息，在打星的位置发生拥塞，网络节点转发队列不够就会丢包。

解决这个问题的一个方法，在 EOR 交换机和核心交换机上，选择使用带大缓存的线卡，以对突发的数据进行缓存，避免报文丢失。

图3-30 EOR 和核心交换机提供大缓存避免拥塞



4 桌面云网络方案

4.1 桌面云业务概述

桌面云是云计算的一种应用形态。关于云计算的定义很多，大家广泛认可的云计算定义是“云计算（cloudcomputing），是一种互联网上的资源利用新方式，可为大众用户依托互联网上异构、自治的服务进行按需即取的计算，云计算的资源是动态易扩展而且虚拟化的，通过互联网提供”。

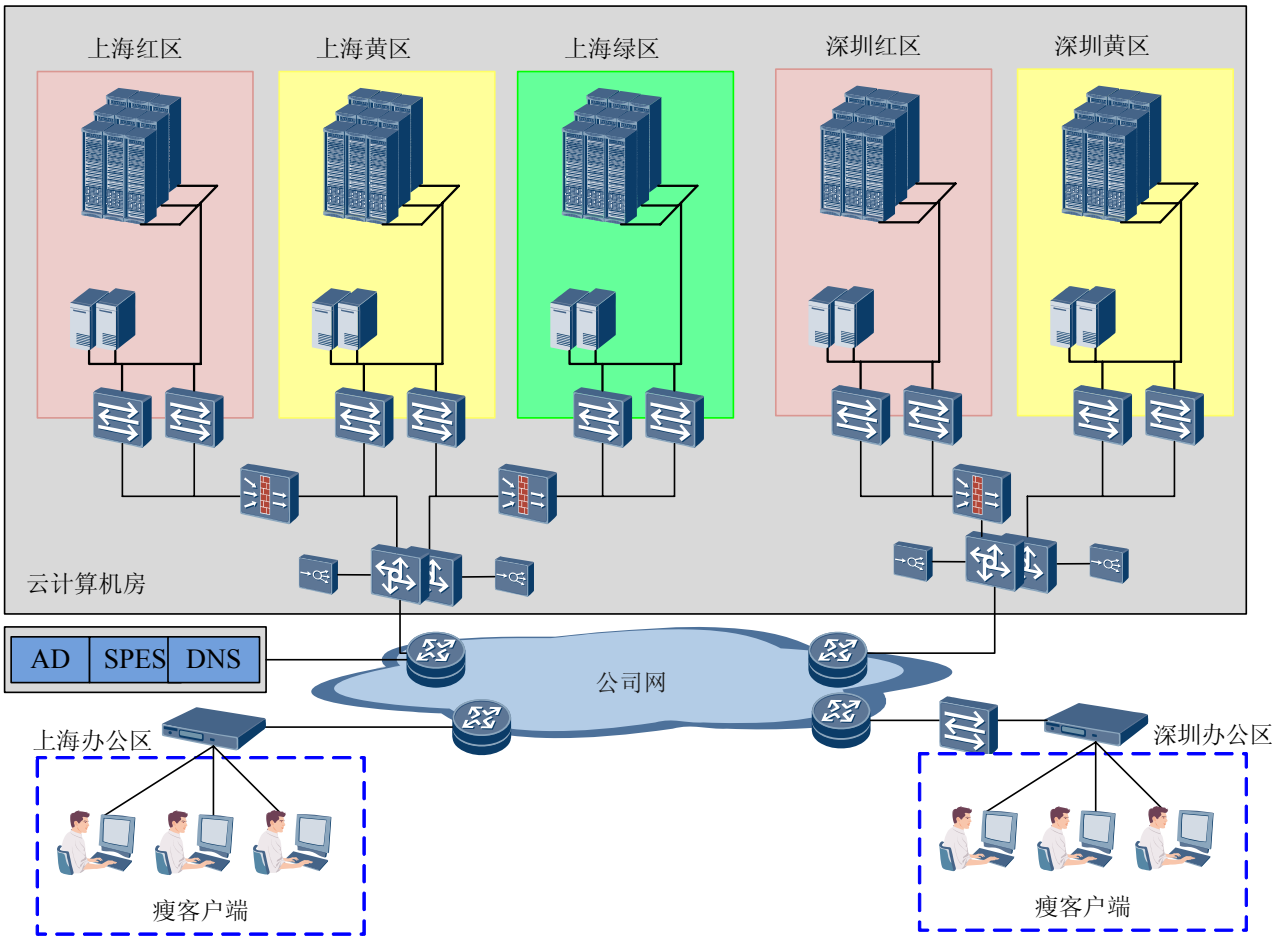
桌面云是合乎上述云计算定义的一种，它具备云计算的三大特征：对用户呈现为桌面服务、资源可弹性管理、通过网络提供，是一种云化的服务。也就是说我们只需要一个瘦客户端设备，或者其他任何可以连接网络的设备，通过专用程序或者浏览器，就可以访问驻留在服务器端的个人桌面以及各种应用，且用户体验和我们使用传统的个人电脑是等同的。

桌面云是将个人计算机桌面环境通过云计算模式从物理机器分离出来，成为一种可以对外提供桌面服务；个人桌面环境所需的计算、存储资源集中于中央服务器上，以取代客户端的本地计算、存储资源；中央服务器的计算、存储资源同时也是共享的、可伸缩的，使得不同个人桌面环境资源按需分配、交付，达到提升资源利用率，降低整体拥有成本的目的。

图4-1 桌面云服务示意

桌面云通过企业内部网络或 Internet，将办公、前台等个人工作区的瘦客户端连接到数据中心。在数据中心中采用桌面虚拟化技术，将一台物理服务器虚拟出几十台虚拟桌面。

图4-2 桌面云数据中心和瘦客户端组网图

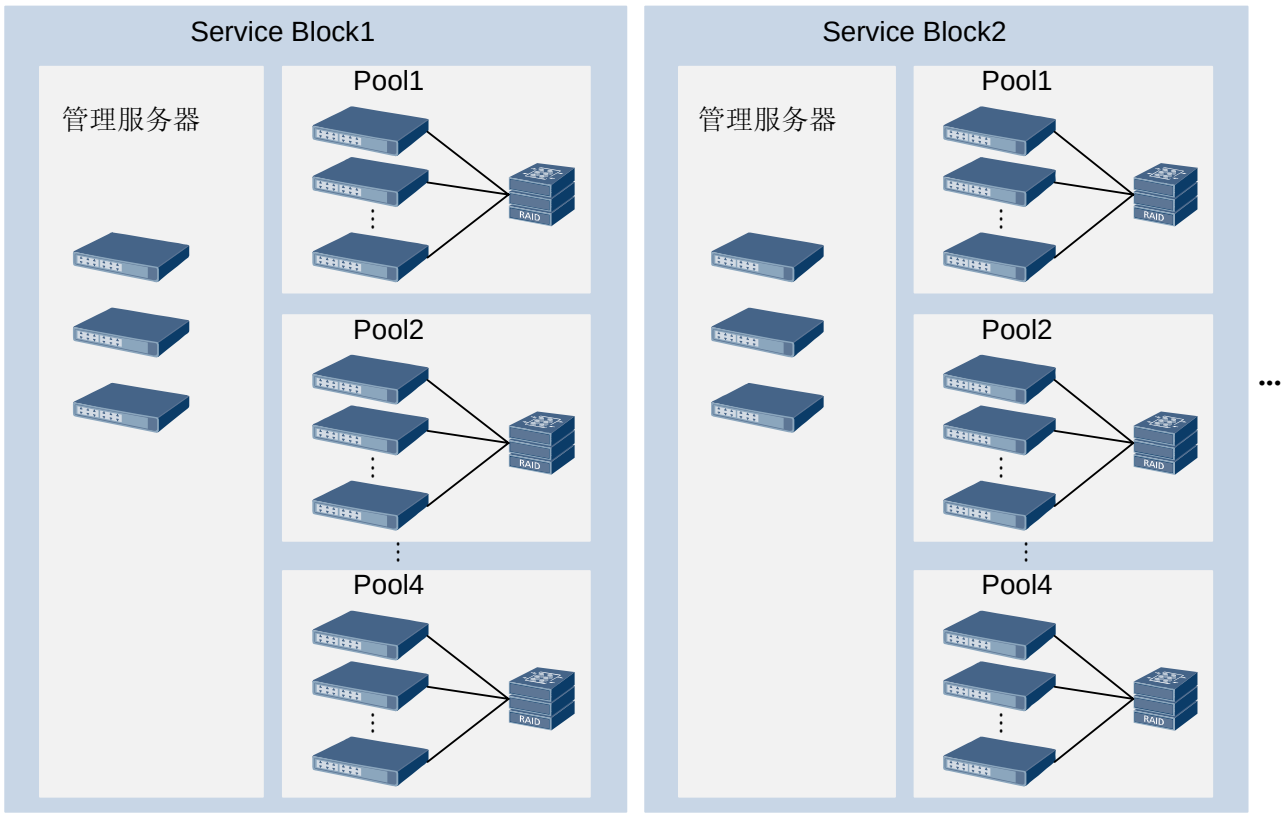


桌面云数据中心一般按照业务、安全的需要分成多个区，分区的设置还受虚拟化管理的规模限制。

这里，简单解释一下桌面云解决方案涉及的一些基本概念：

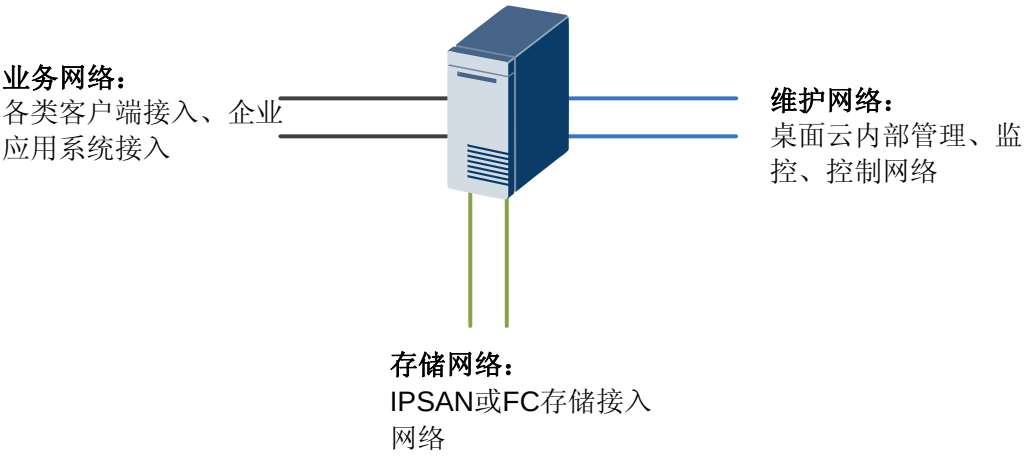
- 管理服务器：除桌面虚拟化服务器之外的其它组件，如 WI、DDC、AD、License Server 等
- Service Block：业务块，是解决方案中的扩展单元，每个 Service Block 支持 2000 个并发桌面用户。它包含 1-4 台管理服务器、1-4 个 Pool。
- Pool：由 1-20 台服务器、1 套存储（1 个控制器框、若干个扩展框）组成。每个 Pool 支持 400-500 并发桌面用户。该 Pool 为管理的逻辑。

图4-3 桌面云逻辑概念示意



如图 4-3 所示，桌面云服务器一般采用刀片服务器，管理服务器、桌面虚拟化服务器都是在刀片服务器中虚拟化。

图4-4 服务器网络平面



刀片服务器采用缺省配置 3 个网络平面：业务网络、维护网络和存储网络，每个网络平面为 1+1 冗余。3 个网络平面相互独立，互不干扰，保障了网络的稳定性与可用性。

4.2 桌面云网络架构

典型桌面云组网如下所示。

图4-5 桌面云数据中心和瘦客户端组网图

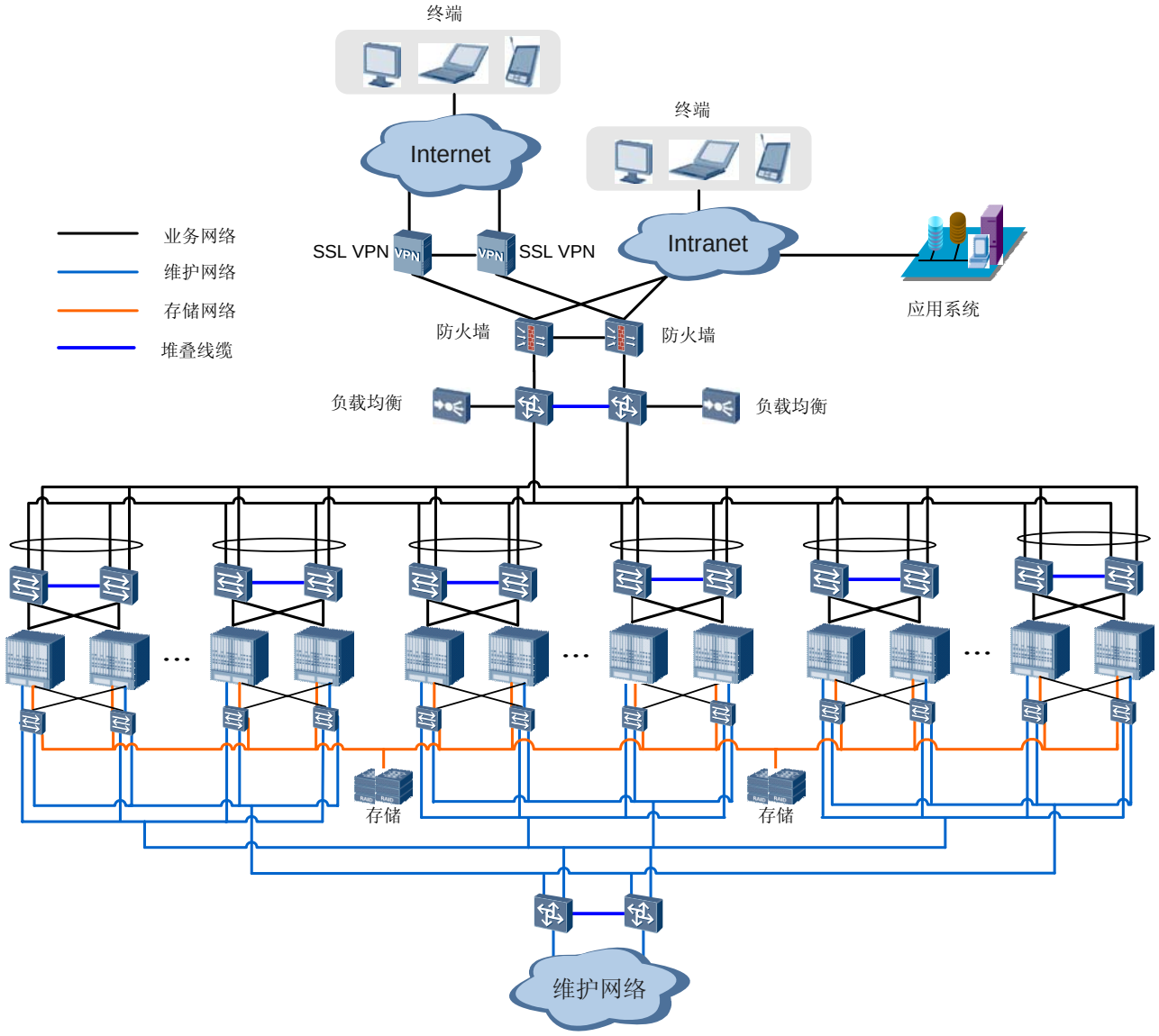


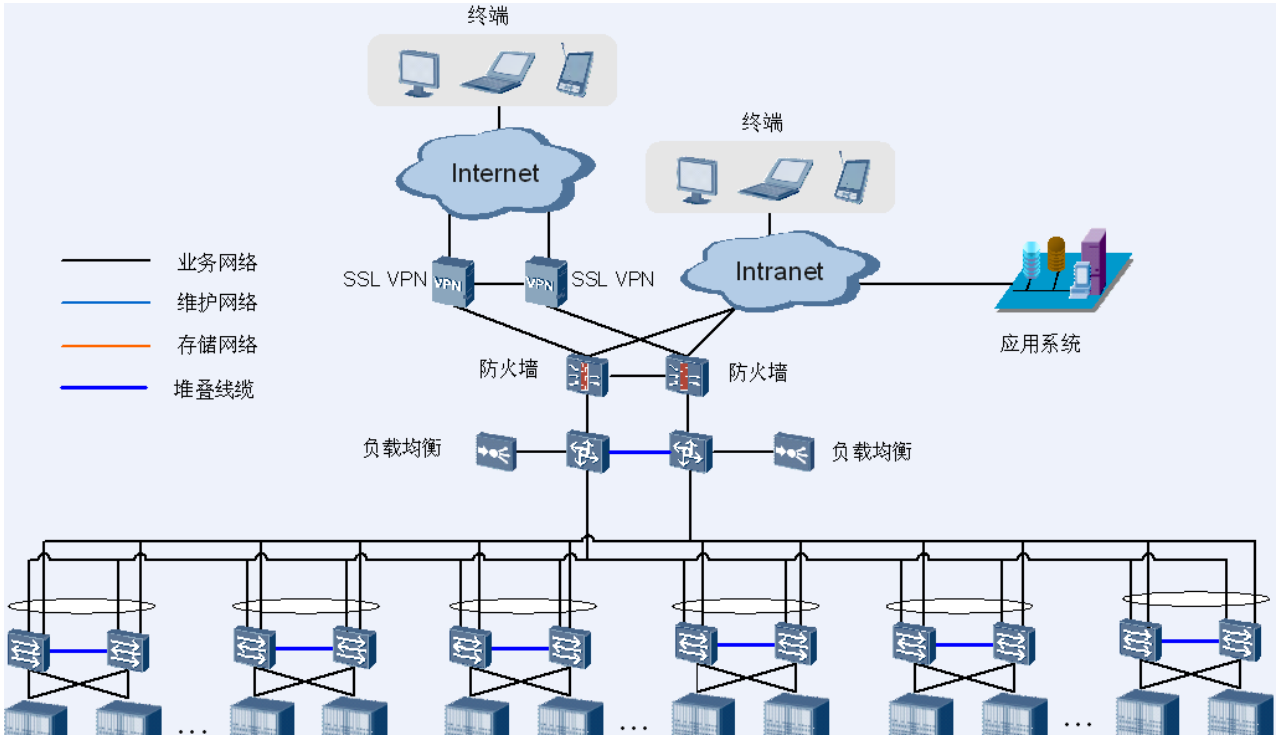
图 4-5 是 3 个 Service Block 组网情况，每个 Service Block 支持 2000 并发桌面用户。桌面云网络设计为业务、维护、存储网络分离设计。其中业务网络、维护网络均为两层结构：接入、核心/汇聚。防火墙、负载均衡全网共享。

桌面云的数据可靠性通过存储系统的可靠性来保证，一般不部署备份系统，也就不需要部署备份网络。

桌面云访问的应用系统，服务器不是部署在云桌面数据中心，而是部署在原来的普通数据中心中。当然，云桌面数据中心也可以作为一个大的分区部署在普通数据中心中。

4.3 业务网络规划

图4-6 桌面云数据中心业务网络组网图



桌面云的业务网络类似于园区网，和园区网相比，桌面的部署密度集中在数据中心中，而且每个服务器支持 20 多个虚拟桌面，因此需要将园区网的盒式交换机集中部署在桌面云数据中心中。

以桌面云推荐的 E6000 机框为例，配置的服务器刀片 10 个，每个服务器刀片支持 20～23 个并发桌面。即每个 E6000 机框支持 200～230 个并发桌面。

配置 6 块 NX910 电口直通模块，每块 NX910 出 10 个 GE 电口。也就是每 200～230 个并发桌面出 60 个 GE 口。参考下文中的带宽统计，上行采用 2 个 GE 就可以满足业务需求。

因此，接入采用 S5700，核心/汇聚也采用 S5700。接入到核心/汇聚间为 GE。

负载均衡和防火墙的性能按分区数×GE 来规划。

4.3.1 业务网络带宽

对于普通办公场景，用户的各种使用行为对流量的占用情况如表 4-1 所示。

表4-1 用户的各种使用行为占用的流量

用户行为	接收 kbps	发送 kbps
打开文件夹	40.40	28.90

用户行为	接收 kbps	发送 kbps
Word 编辑、浏览	346.51	13.51
PPT 浏览	535.37	15.64
网页浏览	201.71	22.20
音乐播放	1658.90	51.91
普通视频播放	1293.66	36.28
高清视频播放	9824.12	331.21
VoIP 语音通话	182	180

根据我们统计，桌面访问流量为 150kbps/用户左右。桌面访问流量与业务应用及用户行为相关。总体来讲，200kbps/用户可以满足大部分网页应用和企业内部应用的需要。

那么，业务网络所需要承载的总流量为：350Kbps*并发桌面数量。

例如：2500 个并发桌面，则需带宽为：350Kbps*2500＝875Mbps。那么，业务网络配置 2 个 1GE 端口互联即可满足。

同样，SSL VPN 网关、负载均衡器均只需配置 1GE 吞吐流量即可满足要求。

4.4 安全规划

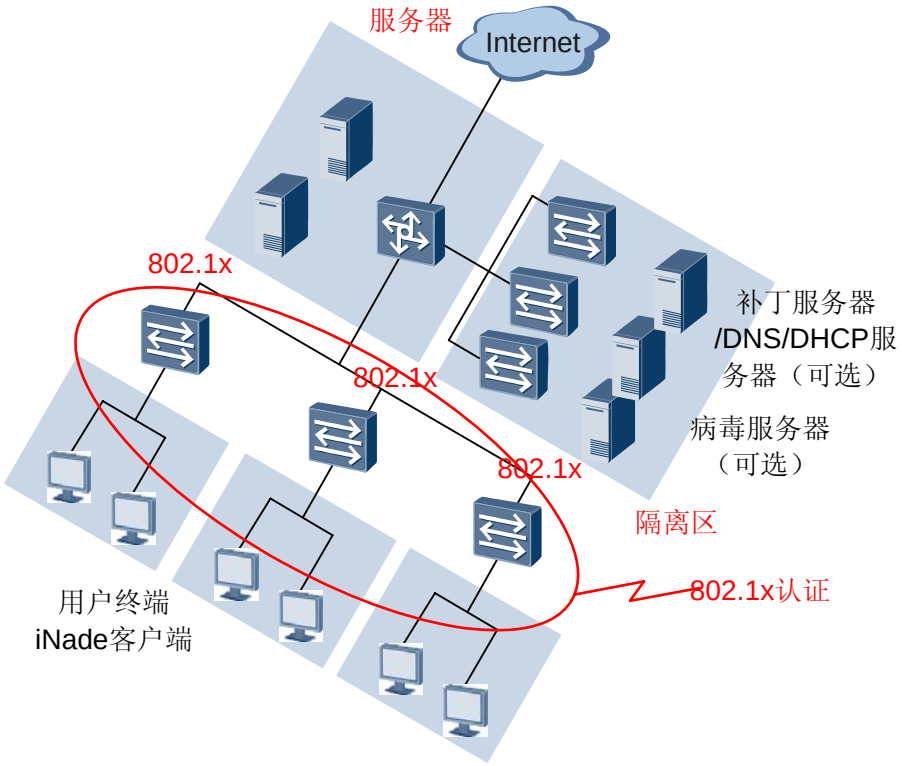
桌面云数据中心中虚拟机实际上是桌面系统，因此安全控制主要是常用于园区网的 NAC。和园区网不同的时，桌面云数据中心中不需要进行移动存储设备和外设接口的管理控制，需要实施的是：

- 网络行为管理：
  - 控制网络流量
  - ARP 防护
  - 控制各类 IM、炒股、P2P 软件和网游
  - 控制 Web 访问、IP 访问
- 终端行为管理
  - 文件操作监控
  - 强制运行信息安全控制软件
  - 阻止不合规软件安装
  - 强制关闭危险服务（如 DHCP 等），或启用必要服务

网络设备上需要部署策略强制执行点，有两种方案：

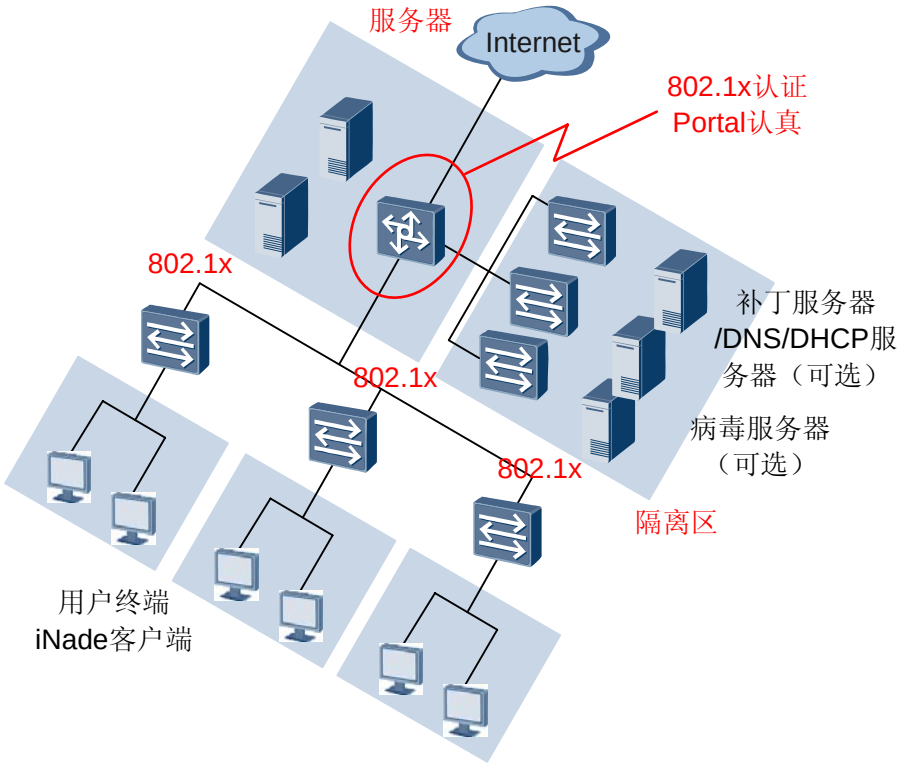
- 一种方案是在接入层交换机上部署 802.1x，能有效隔离不符合安全策略的用户可以防止来自网络内部的安全威胁。

图4-7 桌面云接入层部署 802.1x



另一种方案是部署在汇聚层交换机上（基于 MAC 地址的 802.1X），接入层交换机上启用端口隔离、PVLAN 等安全功能，来防止同一接入交换机下的虚拟机间的相互影响。

图4-8 桌面云汇聚层部署 802.1x



5 多数据中心规划建议

5.1 多中心网络架构

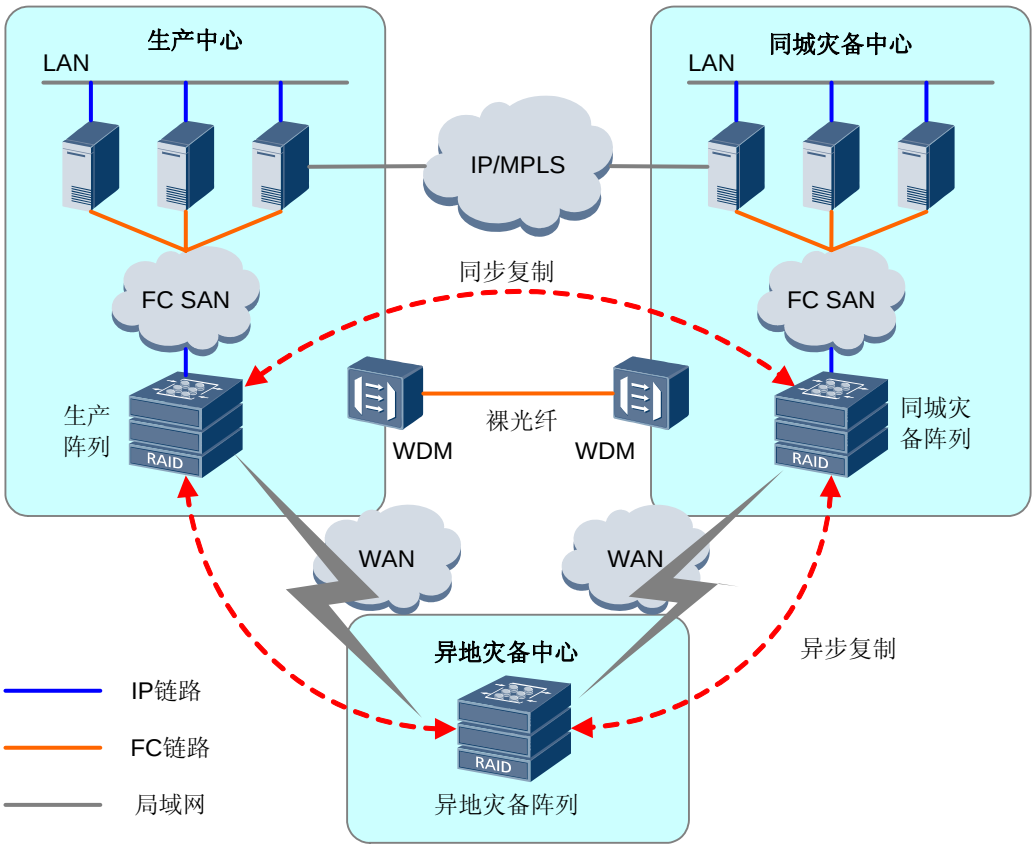
数据中心承载着企业的核心业务和存储海量的业务数据。从可靠性来讲，数据中心需要提供 7×24 小时不间断服务。

- 企业数据中心一般选择在同城 50km 范围内再建立备份数据中心，作为主数据中心的备份，通过专线或者上传输设备，对业务数据进行实时复制。
- 在备份的同时，也可以将部分业务中心转移到备份数据中心上，到达主备数据中心双活状态。

另外考虑到不可抗拒的自然灾害（如地震）对地区城市的毁灭性破坏，建议企业在有能力的情况下，在另外一个距离大于 400km 的城市建立灾备中心，用于主备双中心的备份，定时同步生产中心和同城灾备中心的数据。当发生灾难时，尽量保证重要数据得以保持不被破坏，异地灾备中心可以用备份数据进行业务的恢复。

多数据中心的网络架构如图 5-1 所示。

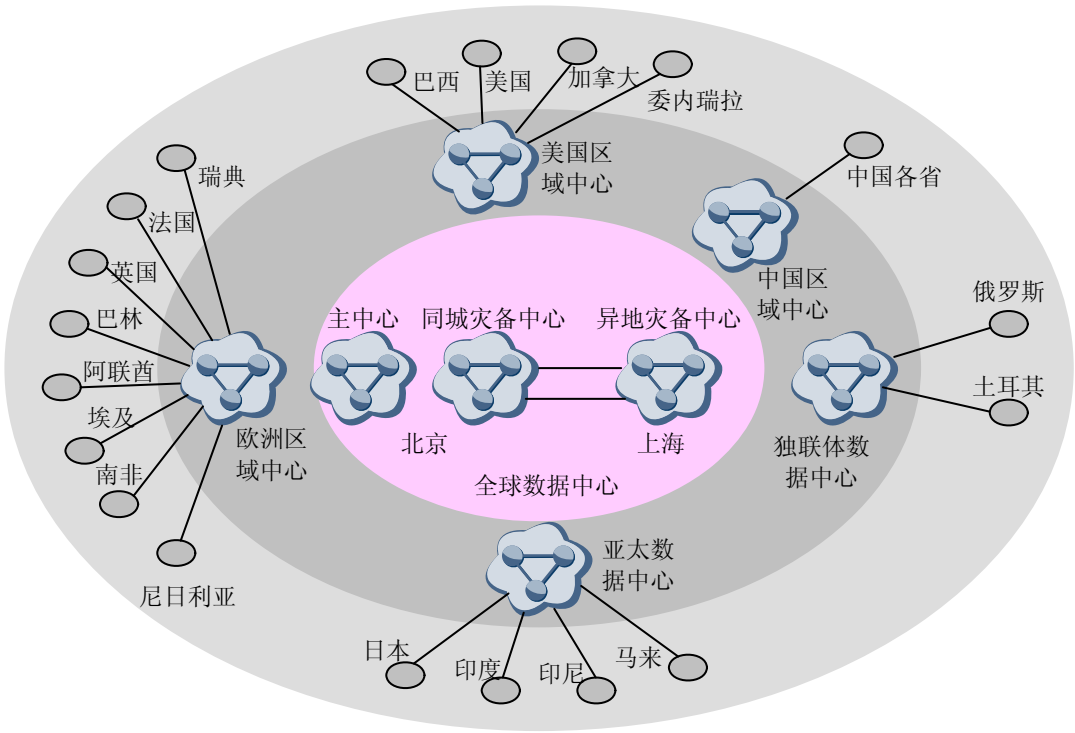
图5-1 两地三中心网络架构



随着企业业务的全球化发展，“两地三中心”的数据中心架构已经不能满足其发展需求，数据中心架构将向“分级多中心”发展。在每个区域中心建立分级的数据中心，可以减轻全球数据中心的负载，节省宝贵的广域网带宽，提高区域业务的响应时间，区域中心故障不会影响到其他区域的业务。

多级多中心网络结构如图 5-2 所示。

图5-2 多级多中心网络结构



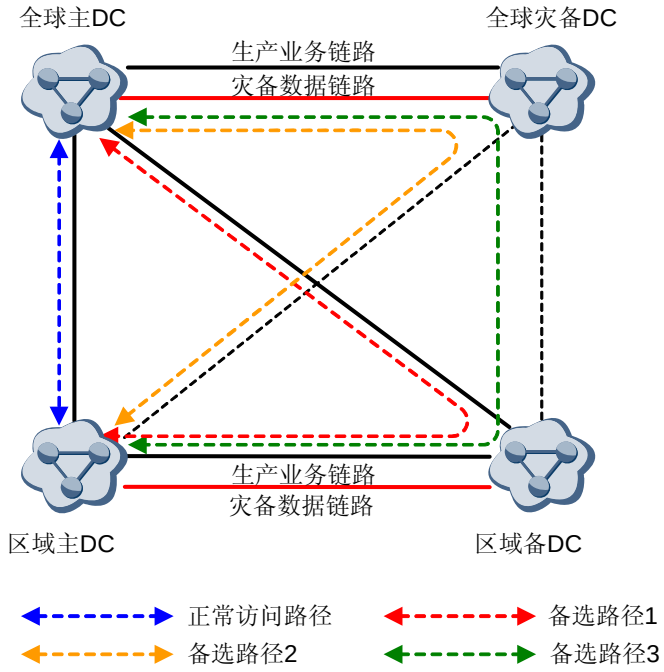
5.2 网络可靠性设计

5.2.1 区域中心和全球中心间的可靠性

全球数据中心主中心跟灾备中心双链路互联，生产业务链路和灾备数据链路独立，安全隔离，带宽保证。区域数据中心同样采用双链路互联，并且主业务链路接入全球数据中心主中心，备业务员链路接入全球数据中心灾备中心。

全球数据中心跟地区数据中心形成的网络拓扑如图 5-3 所示。

图5-3 数据中心间主备路径规划



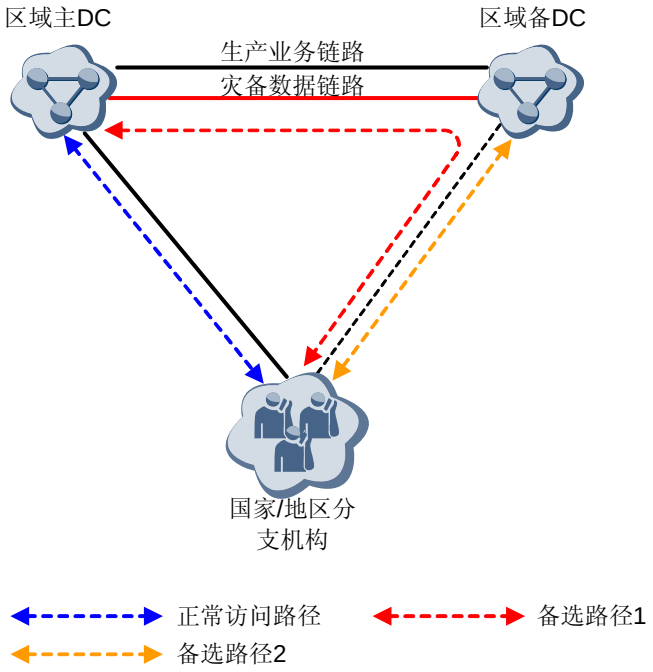
四个数据中心作为不同的 AS 区域，之间采用 EBGP 发布路由，从区域主 DC 到全球主 DC，如图 5-3 所示有 4 条路径，规划的优先级如下：

- （正常访问路径）正常无链路故障情况下，区域主 DC 直接连接全球主 DC，优先级最高。
- （备选路径 1）当区域主 DC 出口设备故障或出口链路故障情况下，区域主 DC 经区域备 DC 到北京中心，优先级第二。
- （备选路径 2）当全球主 DC 接入设备故障情况下，区域主 DC 经全球灾备 DC 到全球主 DC，优先级第三。
- （备选路径 3）当上述故障同时出现情况下，区域主 DC 经区域备 DC，再经全球灾备 DC 到全球主 DC，优先级最低。
- 上述各数据中心间的链路故障切换通过控制 EBGP 的 AS-Path 和 MED，可以实现 4 条路径的优先级控制。

5.2.2 国家/地区到区域中心的可靠性

国家/地区分支机构选择不同的运营商主备链路接入到区域中心的主备 DC 上，网络拓扑如图 5-4 所示。

图5-4 国家/地区接入区域数据中心



国家/地区主接入链路连接到区域主 DC，备接入链路连接到区域备 DC。采用 EBGP，区域主 DC、区域备 DC、国家/地区分别配置不同的 AS。

- （正常访问路径）正常情况下，国家/地区通过主接入链路直接到区域主 DC。
- （备选路径 1）当主接入链路故障时，国家/地区通过备接入链路经区域备 DC 到区域主 DC。
- （备选路径 2）当区域主中心整体故障时，通过智能 DNS 机制，进行应用层次的切换，流量切换到备选路径 2 上。

5.3 路由规划

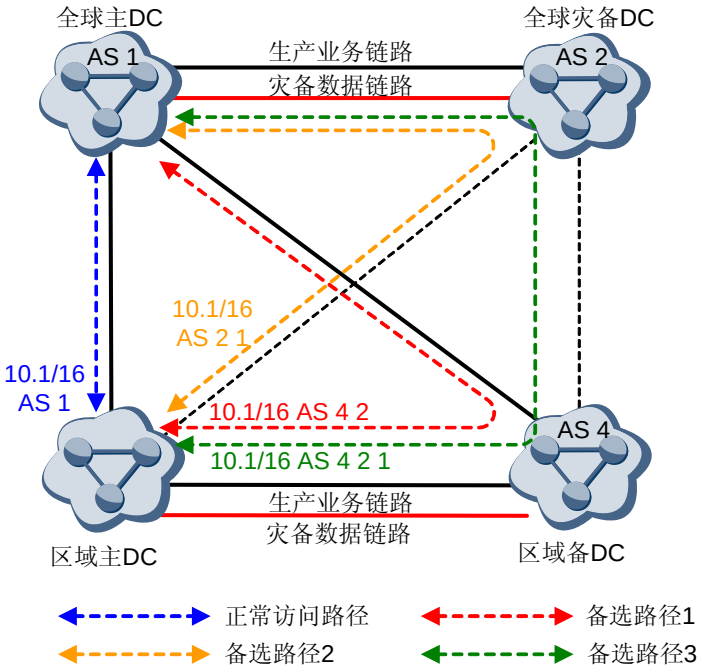
5.3.1 路由概述

一般在数据中心内部部署 IGP 路由协议即可，可以使用 OSPF、ISIS。考虑到网络的稳定性和路由的快速收敛，方便以后的维护和管理，建议采用 OSPF 动态路由协议。在多数据中心之间使用 BGP 传递路由，BGP 在路由控制与策略功能方案具有很强的能力，适合大型网络互联。

5.3.2 BGP 设计

区域数据中心跟全球数据中心互联时，每个数据中心作为一个独立 AS，AS 之间通过 EBGP 传递各自数据中心的路由。在 EBGP 的设计中，通过 AS-Path 和 MED 路由属性，对路由进行控制选择，提高链路的可靠性保护。如图 5-5 所示。

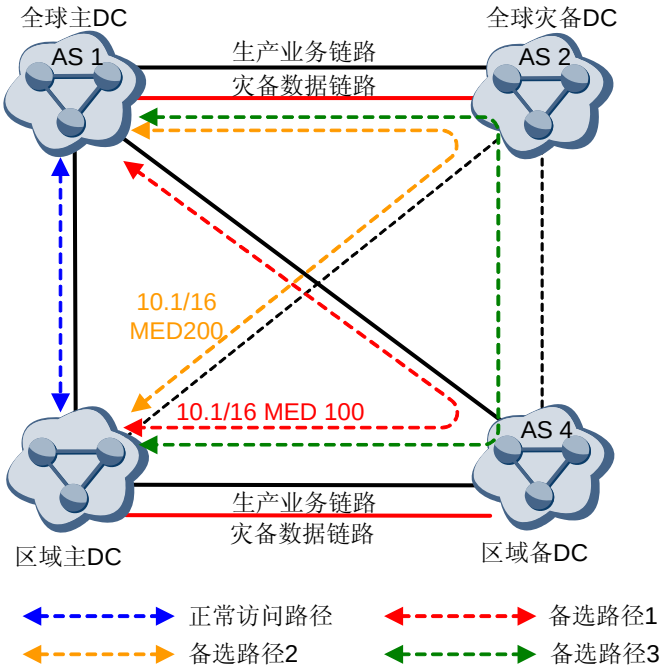
图5-5 BGP AS-Path 选择路径



EBGP 优先选择 AS-Path 长度短的路由。图 5-5 中，AS 3 分别从 AS 1、AS 2、AS4 收到了 10.1/16 的路由信息。这几条路由的 AS-Path 分别是 AS 1、AS 2 1、AS 4 1、AS 4 2 1。

- AS 1 直接发布过来的路由 AS-Path 最短选择 AS 1 直接发布的路由，优先级最高。
- 从 AS 4 2 1 发布的路由 AS-Patch 最长，备选路径 3 优先级最低。
- 从 AS 2 1、AS 4 1 的路由 AS-Patch 相同，需要进一步采用 BGP MED 属性区分优先级。如图 5-6 所示。路由 10.1/16 从 AS 4 传递过来的路由属性 MED 为 100，比从 AS2 传递过来的路由属性 MED 小，所以备选路径 1 的优先级比备选路径 2 的优先级高。

图5-6 BGP MED 选择路径



BGP 具有强大的路由控制选择能力，通过 AS-Path 和 MED 属性的控制，能够很好解决多数据中心的路由选择问题，同时也解决了多数据中心之间的链路可靠性问题。

5.4 容灾规划

5.4.1 容灾概述

灾备中心，也称为灾难恢复中心或容灾中心，就是指企业除了拥有一套完整的计算机网络系统(称为生产中心)之外，另外建立一套计算机网络系统。这套系统能在突发性灾难发生，造成生产中心停止工作时，迅速并及时地接管原来运行在生产中心的所有或部分业务，达到减少或避免灾难事件发生时所造成的损失，为企业用户提供完善、优质服务的目的。

按照国际标准 share78，将容灾分为 7 个层次，分别为：

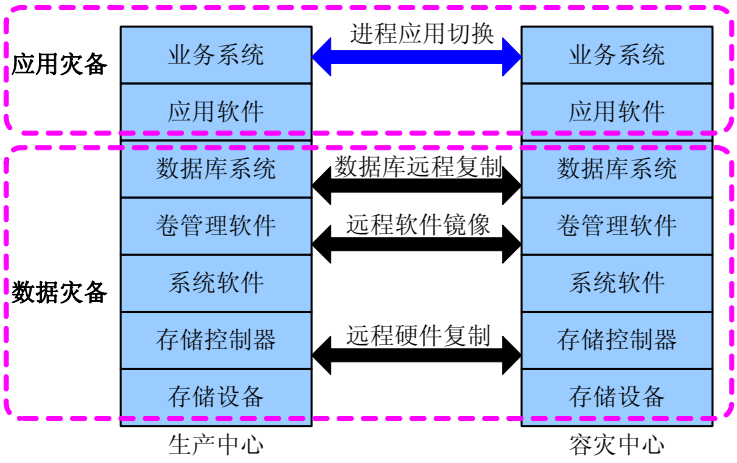
- 层次 0：没有异地数据(No off-site Data)
- 层次 1：PTAM 卡车运送访问方式 (Pickup Truck Access Method)
- 层次 2：PTAM 卡车运送访问方式+热备份中心 (PTAM + Hot 中心)
- 层次 3：电子链接 (Electronic Vaulting)
- 层次 4：活动状态的备份中心 (Active Secondary 中心)
- 层次 5：双重在线存储（Two-Site Two-Phase Commit）
- 层次 6：零数据丢失(Zero Data Loss)
- 层次 7：零数据丢失，自动系统故障切换

按照每个层级数据和业务的特点，华为将容灾简分为了 3 个等级：

- 层次 0～2 为备份级容灾。
- 层次 3～5 为数据级容灾。数据级容灾是指建立一个异地的数据系统，该系统是对本地系统关键应用数据实时复制。当出现灾难时，可由异地系统迅速接替本地系统而保证业务的连续性。
- 层次 6、7 为应用级容灾。应用容灾比数据容灾层次更高，即在异地建立一套完整的、与本地数据系统相当的备份应用系统（可以同本地应用系统互为备份，也可与本地应用系统共同工作）。在灾难出现后，远程应用系统迅速接管或承担本地应用系统的业务运行。

数据级容灾和应用级容灾的业务框架如图 5-7 所示。

图5-7 容灾业务分类



衡量容灾备份的有两个技术指标：

- RPO（Recovery Point Objective）：即数据恢复点目标，主要指的是业务系统所能容忍的数据丢失量。
- RTO（Recovery Time Objective）：即恢复时间目标，主要指的是所能容忍的业务停止服务的最长时间，也就是从灾难发生到业务系统恢复服务功能所需要的最短时间周期。

RPO 针对的是数据丢失，而 RTO 针对的是服务丢失，二者没有必然的关联性。RTO 和 RPO 的确定必须在进行风险分析和业务影响分析后根据不同的业务需求确定。对于不同企业的同一种业务，RTO 和 RPO 的需求也会有所不同。

5.4.2 容灾技术介绍

通常说来，对于远程灾难恢复方案建议用户建立两个数据中心，主中心和备份中心。正常情况下，应用运行在主数据中心的计算机系统上，数据也存放在主中心的存储系统中。当主数据中心由于断电，火灾甚至地震等灾难无法工作时，则立即采取一系列相关措施，将网络、电话线路切换至备份中心，并且利用备份中心计算机系统重新启动应用系统。

而这里最关键的问题就是切换过程时间最短，同时尽可能保持主数据中心和备份中心数据的连续性和完整性。

传统的磁带备份方式一般采取定点备份，而当系统崩溃时。距最近一次备份时间之间的数据将全部丢失，无法恢复，而且磁带备份恢复时间比较长。由于速度慢，缺乏实时性，无法满足用户大数据量数据恢复及数据库连续性，实时性的要求。

而现在主流的灾难恢复方案主要是采用实时的数据备份的方式。它的主要原理是通过通信线路，实时地将主中心更新数据拷贝至备份中心存储系统中，保证主、备中心数据的实时一致性。当主中心无法工作时，备份中心可以立即接管业务，并且确保数据的最大完整性。

分层数据复制技术

根据信息系统中的不同层次，可采用不同的 IT 技术进行数据同步或者复制。通常将其分为六个层次：

- 磁盘存储层（Disk Array）
- SAN 存储网络层（SAN Network）
- 操作系统逻辑卷层（Volume Manager）
- 文件系统层（File System）
- 数据库层（Database）
- 应用系统层（Application）

a. 基于存储镜像复制技术

基于存储镜像复制技术的灾备方案的核心是利用存储阵列自身的盘阵对盘阵的数据块复制技术实现对生产数据的远程拷贝，从而实现生产数据的灾难保护。在主数据中心发生灾难时，可以利用灾备中心的数据在灾备中心建立运营支撑环境，为业务继续运营提供 IT 支持。同时，也可以利用灾备中心的数据恢复主数据中心的业务系统，从而能够让业务运营快速回复到灾难发生前的正常运营状态。

盘阵之间的镜像复制技术的主要特点是不占用主机 CPU、内存、I/O 资源，并且对主机操作系统无关，对应用系统影响比较小。这也是目前最成熟，应用最广泛的灾备技术。但是其缺点是生产中心和备份中心需要采用同厂商同型号的存储设备。

现在主流的存储厂商都支持磁盘阵列级的镜像复制技术，例如 EMC DMX 系列的 SRDF，EMC CX 系列的 MirrorView，IBM DS8000 的 MetroMirror、GlobalMirror，IBM DS4000 的 ERM，HP XP 系列的 ContinuousAccess，HDS USP 系列的 TrueCopy 等。

b. 基于 SAN 网络复制技术

基于 SAN 网络复制技术，是近年来比较新的一种技术，此技术实质是在 SAN 网络中增加一个虚拟存储管理设备，根据厂商的不同可以直路部署或旁路部署。

基于 SAN 网络的复制技术支持异构存储设备，并且对于主机端来说是透明的，在数据中心拥有多个厂商的磁盘阵列时，比较适合，但是缺点是对后端存储 I/O 速度有影响，成熟度还有待提高。

支持此技术的厂商有 IBM SVC、EMC invista、Falcon Ipstor 等。

c. 基于操作系统卷复制技术

基于操作系统卷复制技术工作在主机的卷管理器这一层，通过磁盘卷的镜像或复制，实现数据的容灾。这种方式也不需要在两边采用同样的存储设备，具有一定的灵活性，但复制功能会占用一些主机的 CPU 资源，对主机的性能有比较大的影响。因此，这种方法的可扩充性较差，实际运行的性能不是很好。基于主机的方法也有

可能影响到系统的稳定性和安全性，因为有可能导致不经意间越权访问到受保护的数据。

常见的卷复制软件有 Symantec Veritas Volume Relicator 等。

d. 基于文件系统复制技术

基于文件系统的灾备复制是指通过复制数据文件的方式，从生产中心向容灾中心进行数据容灾。基于文件的数据复制备份需要基于文件的存储系统，这可能表现为多种形式：文件服务器、网络附加存储（NAS）、NAS 设备，或者使用文件虚拟化形成的组合体。

使用基于文件的存储复制用于数据备份保护越来越流行，这主要有两个原因：

- 易于部署，使用标准协议，支持原生复制功能，还可与众多驱动器技术配合使用。
- 解决了企业使用基于块的存储系统时面临的一些主要难题，像存储资源的利用、跨介质服务器资源共享、及时为介质服务器配置存储容量等。

e. 基于数据库逻辑复制技术

基于数据库的复制技术是一种逻辑复制技术，支持异构存储、甚至是异构操作系统平台，它的工作原理为通过分析生产数据库的重做日志，生成通用或私有的 SQL 语句，然后传输到备份数据库上进行 Apply 应用。

这种数据复制优点是可以与底层存储无关，跨平台，速度较快，但缺点就是占用主机资源，并且对某些特殊数据类型支持不好，有些 DDL 操作语句也不支持，并且如果业务系统中有随机产生的数据时，数据一致性无法得到保证。

常见的数据库逻辑复制技术有 Oracle DataGuard，Oracle Stream，Quest shareplex for Oracle，DSG RealSync for Oracle，IBM DB2 HA/DR。

f. 基于应用系统技术

基于应用系统的技术，应用系统必须支持交易的分发，利用交易中间件软件，将在线交易同时在生产中心和灾备中心执行；或者通过交易中间件软件将任何主中心的数据改变发送到备份中心，从而保证生产中心和灾备中心的数据一致性。

这种方式的优点是对网路带宽的要求较低，缺点是需要修改应用，在现有应用的情况下，比较难实现。

数据备份模式

数据备份可以建立在本地，也可以备份到远程。根据保护要求级别，数据备份可以分为同步模式或异步模式。

- 同步模式

  是指在向磁盘进行下一次写操作之前，本地和远程卷都必须进行上次写操作的更新。这提供了最高级别的保护，但可以会因为在相隔两地的阵列之间传送数据的延迟导致应用性能的降低。

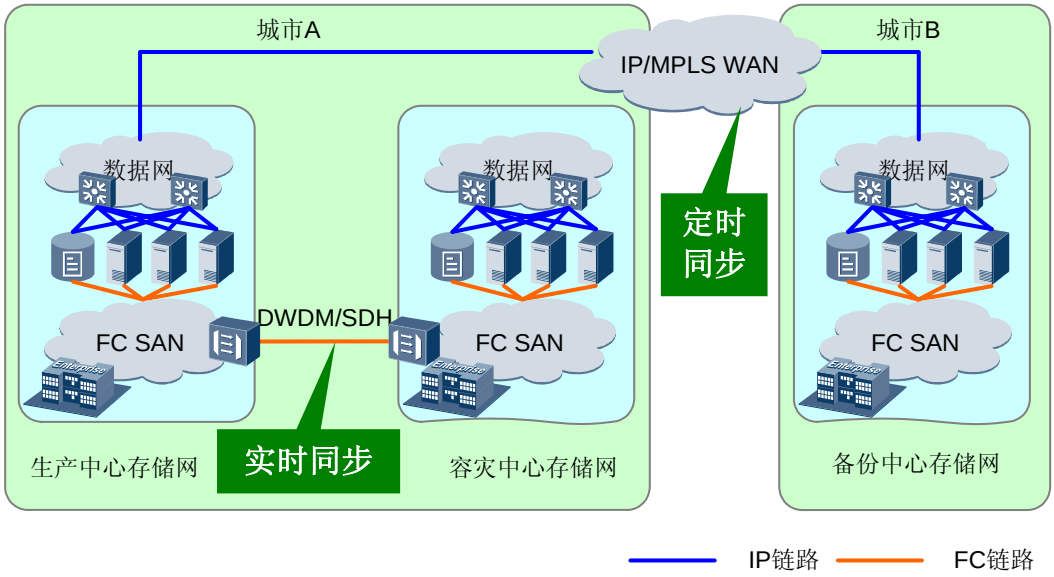
- 异步模式

  是指本地卷可以继续进行写操作，即使远程卷还没有被更新；远程可以在延迟一段时间慢慢更新。这种方法提供了比较高的应用性能，但如果灾难发生，就会有丢失一些在远程卷上还未更新的数据。

5.4.3 容灾网络规划

容灾网络设计考虑同城实时容灾和异地备份容灾两种方式。一般同城容灾采用实时同步，异地备份容灾采用定时同步。如图 5-8 所示。

图5-8 同城/异地灾备网络规划



同城容灾

在城域容灾方案中，根据灾备地点和目前生产中心之间的物理距离，建议在城域的模式下，对核心业务数据采用同步/异步保护模式。

对于 FC SAN 存储网络，可以使用 WDM/SDH 技术用作远程备份网络，采用同步数据复制方式，可以选择基于存储镜像硬件复制技术。

如果站点距离在 100 公里之内，而且链路仍然采用光纤链路的话，考虑光纤信号的时延问题，可以对部分核心业务数据采用同步数据模式，其他数据采用异步模式。

如果采用基于 IP 数据链路，可以利用基于 IP 的 SAN 的互连协议，如 FCIP、iFCP、Infiniband、iSCSI 等。并且最好采用异步方式。

异地容灾

在异地容灾方案中，由于考虑到异地之间的距离比较长，备份流量通过租用线路和 ATM 网络。如果用户资金比较充足，建议采用租用点到点线路。并且可以采用广域网加速设备降低租用的广域网专线带宽，以最小代价提高网络性能，为远程数据备份提高速度和服务效率。

异地容灾出于数据容灾的带宽和时延性能要求，基本采用异步数据复制方式，当备份中心存储的数据量过大时，可利用快照技术将其备份到磁带库或光盘库中。

异地容灾远端数据会比本地生产端数据落后一定时间，这个时间随采用的技术，带宽、距离、数据流特点的不同而不同。一般而言，通过软件方式的数据复制技术较容易实现

完整的数据包的排队和断点重发机制，在灾难情况下可以保证灾难时间点的数据一致性。

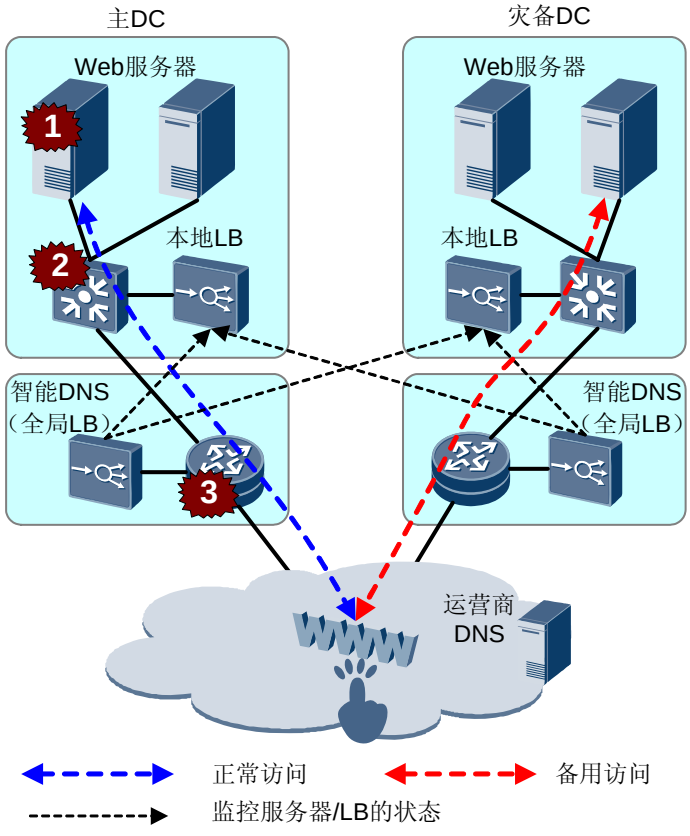
另外与同步传输方式相比，异步传输方式对带宽和距离的要求低很多，它只要求在某个时间段内能将数据全部复制到异地即可，同时异步传输方式也不会明显影响应用系统的性能。其缺点是在本地生产数据发生灾难时，异地系统上的数据可能会短暂损失（如果广域网速率较低，交易未完整发送的话），但不影响一致性（类似本地数据库主机的异常关机）。

5.4.4 容灾业务规划

在数据实时同步的基础上，可以实现业务的实时自动切换和双活负载均衡。如图 5-9 所示。智能 DNS（全局均衡负载器）需要监控服务器或者本地均衡负载器的状态，并基于服务器/LB 的状态给予 DNS 解析结果。

故障 1 是一个 Web 服务器故障，这种情况是通过本地的负载均衡器直接切换到本地的另外一台服务器。故障 2 和故障 3 是整个业务区域故障，这时就需要通过全局负载均衡器，自动切换到备业务区域。

图5-9 基于主备智能 DNS/GSLB 实现业务系统的自动切换和双活负载均衡



由于 DNS 对数据中心的业务影响巨大，因此，需要重点考虑 DNS 服务器的容灾。在多数据中心的情况下，建议 DNS 的部署时，把 Slave 服务器部署在主数据中心，把 Master 服务器部署在备数据中心。这样，在整个主中心故障时，也可以保证 DNS 服务的联系性。

5.5 业务分布规划

5.5.1 业务分布概述

企业的发展壮大，促使企业从单数据中心逐渐走向两地三中心、多中心的模式，数据中心承载的业务也将根据需要调整分布在主数据中心或者区域数据中心上。

用户的体验和业务本身的特点造就不同业务对网络的带宽、时延需求各不相同，因此对应数据中心的部署模式也不相同。如 OA 业务中的 Notes、Email 业务应用，对网络延迟敏感，对网络带宽要求较大，采用分布式部署有利于减少区域数据中心跟主数据中心的专线带宽，缓解主数据中心带宽压力。

5.5.2 业务分布规划

从业务的层面看，数据中心采用“集中+分布”的架构，是一种理想的方式，可以满足不同业务员对网络的需求，提高用户对业务的满意度。如下表格分析了数据中心部分应用业务的特点及其部署建议。

表5-1 基于业务的分布式和集中式部署方式

应用	架构	特点	部署方式
Notes Email OA 系统	C/S	交互式操作：延迟敏感 大数据操作：带宽和延迟敏感	分布式部署： 全球主数据中心 区域数据中心
Web	B/S	交互式操作：延迟敏感 大数据操作：带宽和延迟敏感	集中+分布部署： DB/APP 服务器集中部署； HTTP 服务器分布式部署
ERP	B/S	延迟/误码敏感	集中部署： 全球主数据中心
视频 VOIP	/	带宽和抖动	集中+分布部署： GK 集中部署，MCU 分布式部署
交互类生产业务	/	交互式操作，低带宽	集中式数据中心

集中式和分布式部署分别适用的业务类别和特点如表 5-2 所示。

表5-2 集中/分布部署方式所适用的业务

部署方式	适用业务
分布部署	• 分布在地区部的业务 • 仅限于地区部内部的业务 • 业务量大、流量大、交互频繁的业务
集中部署	• 业务量小、流量小的业务（如早期拓展期业务） • 极其重要，需要总行监管的业务

分布式业务可以通过全局负载均衡技术，实现为企业“就近”服务的要求，同时实现多个物理位置的数据中心之间互相备份，实现多数据中心的负载分担。

- 正常情况下，各个数据中心就近提供服务。
- 某个数据中心服务器故障或者网络故障时，在用户不感知的情况下，通过异地数据中心提供服务。
- 通过全局负载均衡，实现智能 DNS 功能，分担企业业务，实现多个物理位置的服务器之间负载均衡。

通过全局负载均衡 GSLB，还可以实现重定向功能：

a. 用户发送 HTTP/RTSP 请求。
b. GSLB 之间交互信息，选择最合适给该用户提供业务的数据中心。
c. 最靠近的 GSLB 通过 HTTP/RTSP 302 重定向消息，给用户返回被选中的 IDC 的虚拟 IP 地址。
d. 用户的 HTTP/RTSP 自动重定向到被选中的 IDC 的虚拟 IP 地址。

6 数据中心网络维护建议

6.1 网络管理

eSight 是华为面向企业网管理推出的新一代面向企业园区和分支网络管理系统，实现对企业资源、业务、用户的统一管理以及智能联动。

eSight 支持对 IT&IP、以及第三方设备的统一管理，同时对网络流量、接入认证角色等进行智能分析，自动调整网络控制策略，全方位保证企业网络安全。同时 eSight 提供灵活的开放平台，为企业量身打造自己的智能管理系统提供基础。

6.1.1 网络日常维护场景

日常维护概述

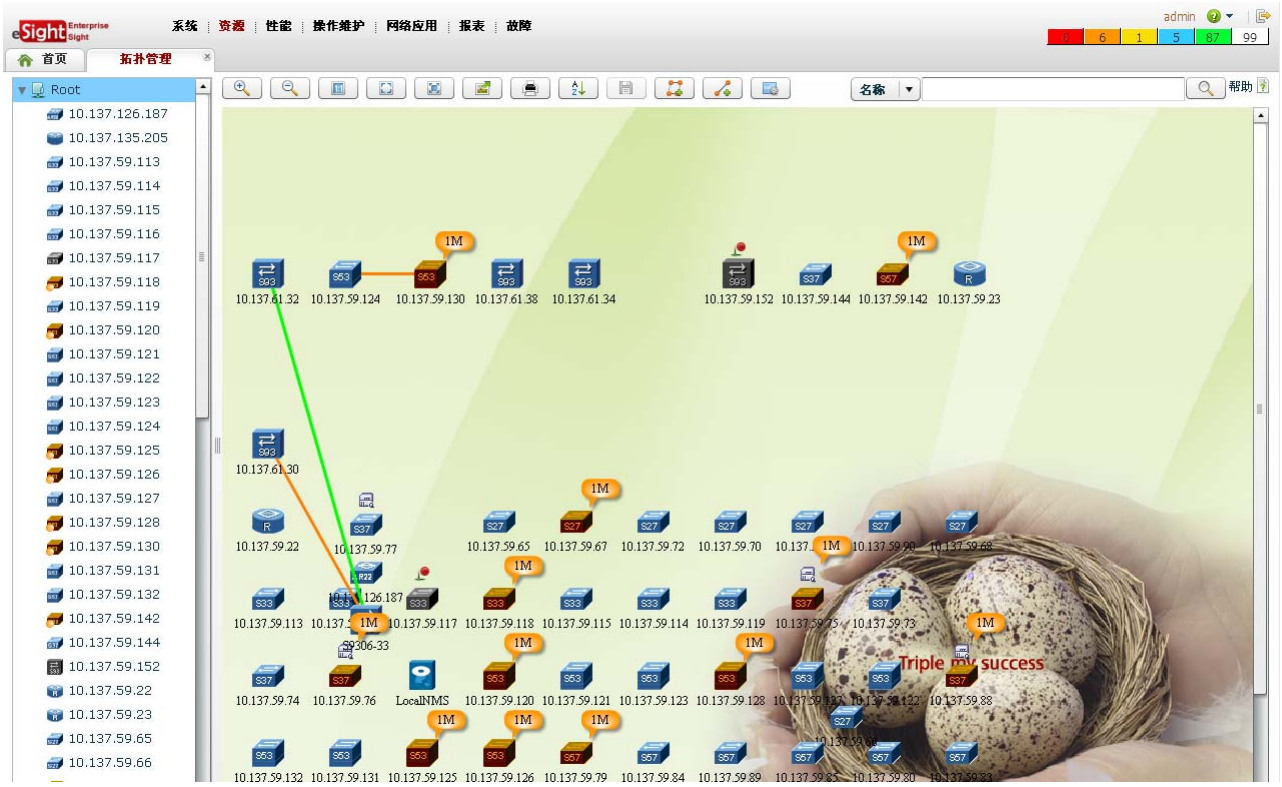
对于网络运维人员而言，日常维护工作不仅繁杂，而且工作量大，涉及的工作内容包括监控拓扑对象、监控网元、配置网元、监控业务、诊断故障、监控性能、查看资源、报表生成等。

华为公司推荐 eSight 网络管理系统，可以准确、快捷的提供运维人员所需要的信息，大大减轻运维人员的工作量。通过 eSight 网络管理系统丰富的管理功能和灵活多样的维护手段，可以轻松实现网络日常维护。

拓扑管理

eSight 以左树右图的方式组织整个视图，其中左导航树以树型直观的体现出网络结构的层次关系。右视图在背景图上将指定网络层次的对象显示在不同的坐标上，可直观了解对象部署。如图 6-1 所示。

图6-1 监控 TOPO 对象



eSight 的拓扑图提供以下功能：

- 支持对拓扑上子网、网元、链路、虚拟网元等的增删改查。
- 支持移动拓扑上的元素。
- 支持显示告警状态及 Tips 信息。
- 支持排列、浏览属性、放大缩小、打印等常用基本操作能力。
- 支持在拓扑图中提供其他功能的快捷操作入口，如进入网元管理器，查看设备相关告警等功能。

eSight 的拓扑告警提供以下功能：

- 支持通过拓扑节点的颜色监控设备的轮询状态（正常、未知、离线等）。
- 支持屏蔽显示低级别告警，当网元或子网同时产生多条告警时，系统只显示最高级别告警。

网元监控

网元管理器首页提供设备基本信息、TOPN 告警以及接口流量、带宽利用率、CPU、内存等性能图表。各图表用户可进行定制是否显示。

图6-2 网元管理

eSight 针对各种不同类型的设备，支持丰富的网元监控和管理功能，如表 6-1 所示。

表6-1 eSight 网元监控功能

设备类型	支持的功能
华为路由器、交换机	• 提供完整的性能采集、告警监控能力。 • 提供设备基本信息管理功能。 • 支持通过适用仿真图片的设备面板查看设备状态，并支持单板、端口状态的联动显示。 • 支持查看设备的接口数据、IP 地址数据。 • 支持单网元的配置管理功能。 • 提供设备配置文件的查看、备份、恢复、比较的功能。 • 提供设备、机框、单板、子卡、端口的资源管理功能。
华为防火墙	• 提供基于标准实现的性能采集、告警监控能力。 • 提供设备基本信息管理功能。 • 支持通过仿真图片的设备面板查看设备状态，并支持单板、端口状态的联动显示。 • 支持查看设备的接口数据、IP 地址数据。 • 支持调用设备的 WEB 网管提供单网元的配置管理功能。 • 提供对设备配置文件的查看、备份、恢复、比较的功能。

设备类型	支持的功能
预集成的主流CISCO、H3C设备	• 提供基于标准实现的性能采集、告警监控能力。 • 提供设备基本信息管理功能。 • 支持通过使用仿真图片的设备面板查看设备状态，并支持单板、端口状态的联动显示。 • 支持查看设备的接口数据、IP地址数据。 • 支持调用设备的WEB网管提供单网元的配置管理功能。 • 提供应设备配置文件的查看、备份、恢复、比较的功能。 • 提供设备、机框、单板、子卡、端口的资源管理功能。
未预集成的第三方设备	• 提供基于标准实现的性能采集、告警监控能力。 • 提供设备基本信息管理功能。 • 支持通过基本图片查看设备面板，基于设备定制提供单板、端口状态的联动显示。 • 基于设备定制功能，用户可以通过输入定制数据实现并支持设备图标展示、设备自身的性能采集、告警上报、配置文件备份。
服务器、打印机	• 提供基于标准实现的性能采集能力。 • 提供设备基本信息管理功能，例如设备基本属性。 • 支持通过使用仿真图片的设备面板查看设备状态，并支持单板、端口状态的联动显示。 • 支持查看设备的接口数据、IP地址数据。 • 支持调用设备的WEB网管提供单网元的配置管理功能。 • 提供服务器、打印机的设备存量管理功能。

配置网元

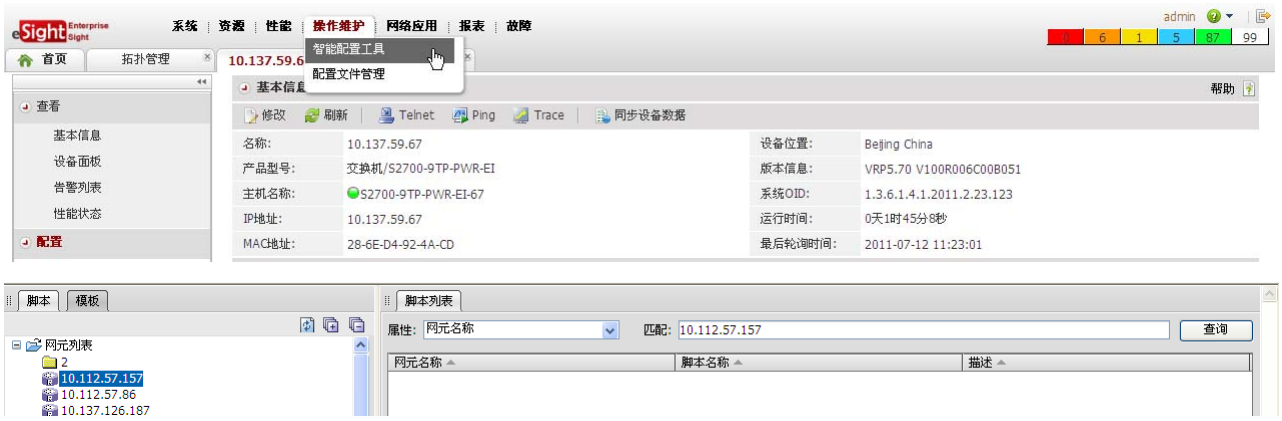
eSight网络管理系统可以通过三种方式完成单点网元配置工作。

1. 使用简单配置框架实现接口、路由等配置。如图 6-3 所示。

图6-3 网元单点配置（简单配置框架方式）

2. 使用智能配置工具进行设备单点配置。如图 6-4 所示。

图6-4 网元单点配置（智能配置工具方式）

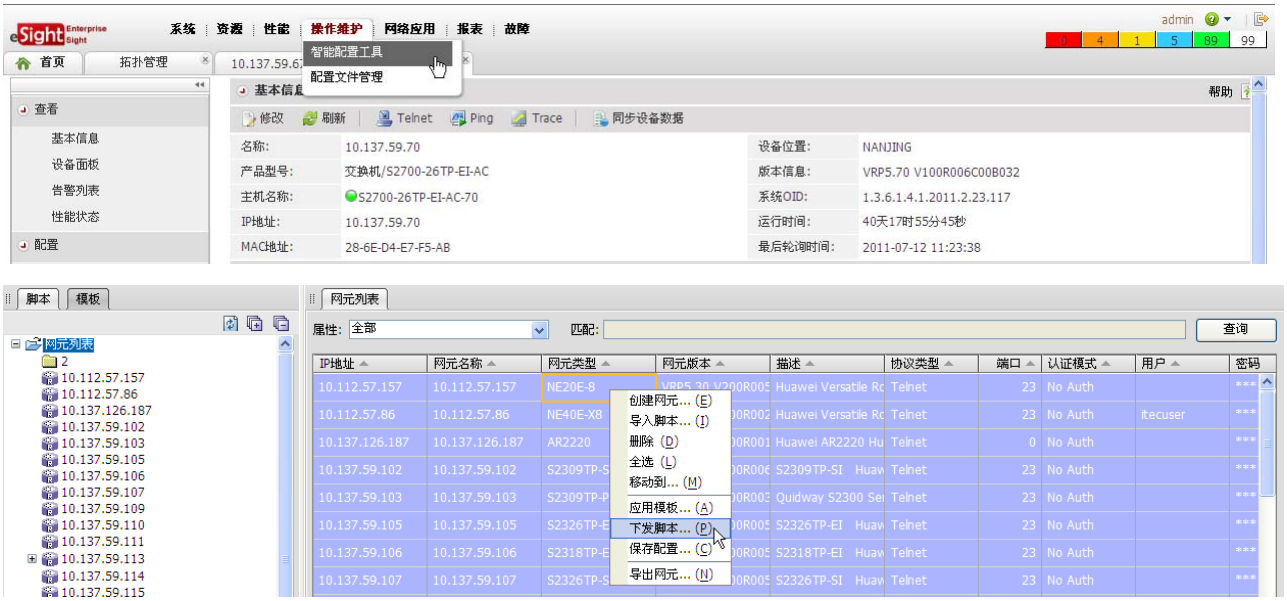


3. 交换机、AR 和安全设备通过 Web 网管进行单点配置。如图 6-5 所示。

图6-5 网元单点配置（Web 网管方式）

在开局、网络维护等多个场景，用户有对集中部署的设备的业务进行批量操作的需求，用户通过智能配置工具能够对多台设备的业务进行批量配置，提高用户的运维效率。如图 6-6 所示。

图6-6 网元批量配置

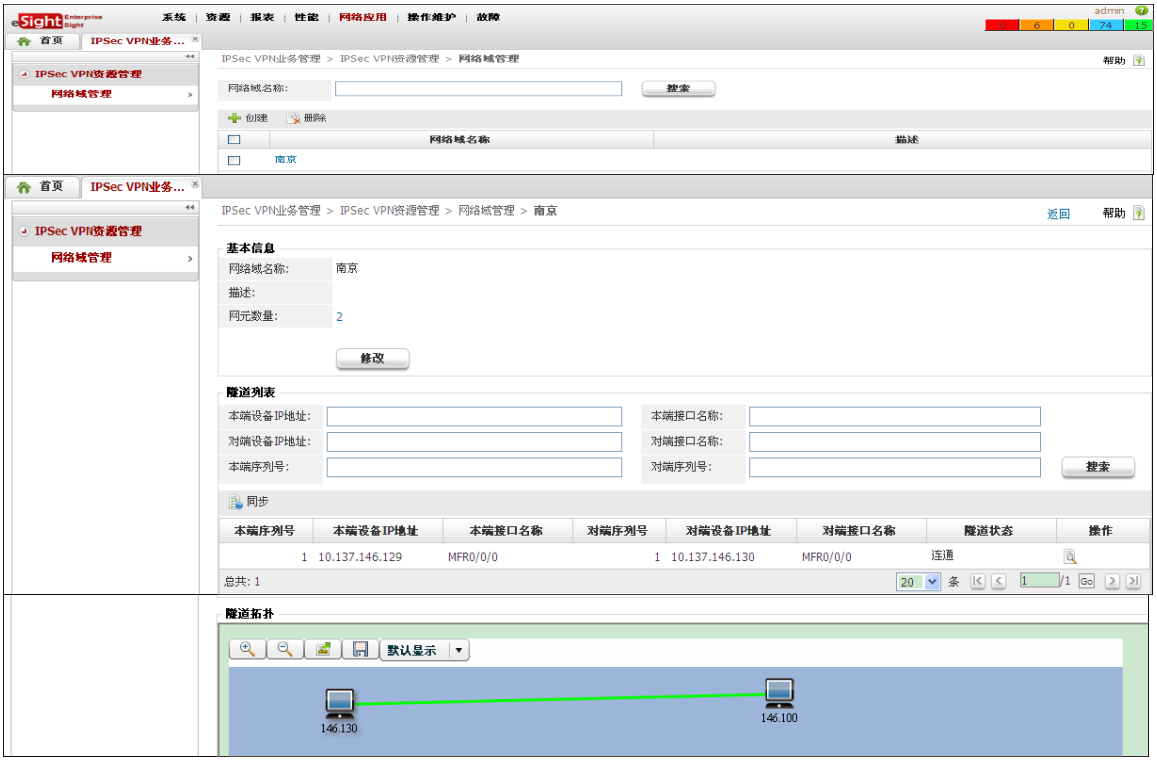


----结束

监控业务

eSight 网络管理系统能够对业务进行实时监控，根据业务类型进行流量、信息统计，极大的方便网络运维人员实时监控业务状况。如图 6-7 所示。

图6-7 监控业务



监控性能

eSight 可以对网络的关键性能指标进行监控，并对采集到的性能数据进行统计。通过可视化的操作界面，方便用户对网络性能进行管理。

通过监视模板管理性能监视指标，并设定告警的阈值。通过性能监视模板，用户可以方便的将性能采集规则应用到多个对象中。性能监视模板包括以下内容：

- 性能指标组

  将多种性能指标集成到一个性能指标组中，可以支持分场景定制指标组，包含场景相关的所有性能指标，便于根据业务场景建立对应的监视任务。

- 性能指标

  定义具体的性能采集的指标。

- 采集周期

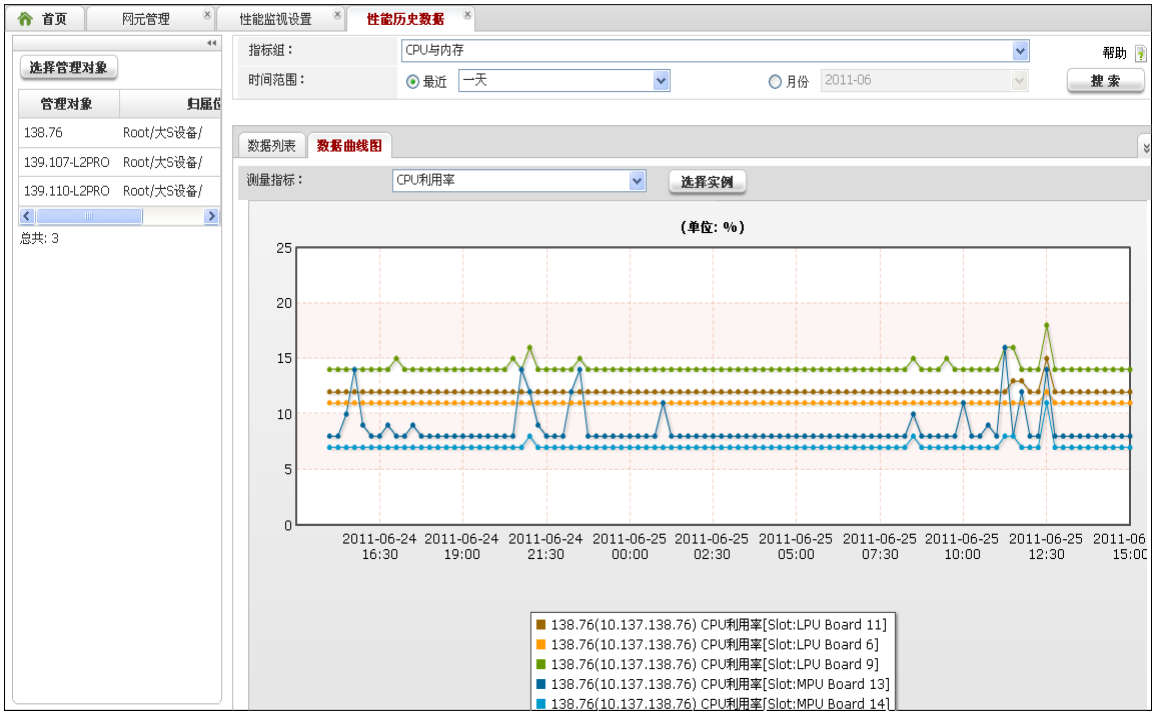
  提供多种采集周期供采集性能指标时选择。

- 性能阈值

  通过设置性能门限值，可以在网络的性能数据低于门限值时及时预警，避免网络性能的持续恶化。

通过性能监视的设置，实现网络性能数据的采集。支持周期性性能指标采集，可以了解网络在指定时间范围内的性能状况，并为预测网络的性能变化提供数据依据。如图 6-8 所示。

图6-8 性能监控



通过性能监视设置获取网络性能数据后，可以通过性能监视视图以图形化的方式进行指标值查看。用户可以了解网络在指定时间范围内的性能状况，为预测网络的性能变化提供数据依据。

资源查看、报表管理

eSight 提供丰富的资源查看和预定义报表，同时提供强大易用的报表设计功能，用户可根据行业特点和自身运维要求进行客户报表定制。如图 6-9 和图 6-10 所示。

图6-9 查看物理资源

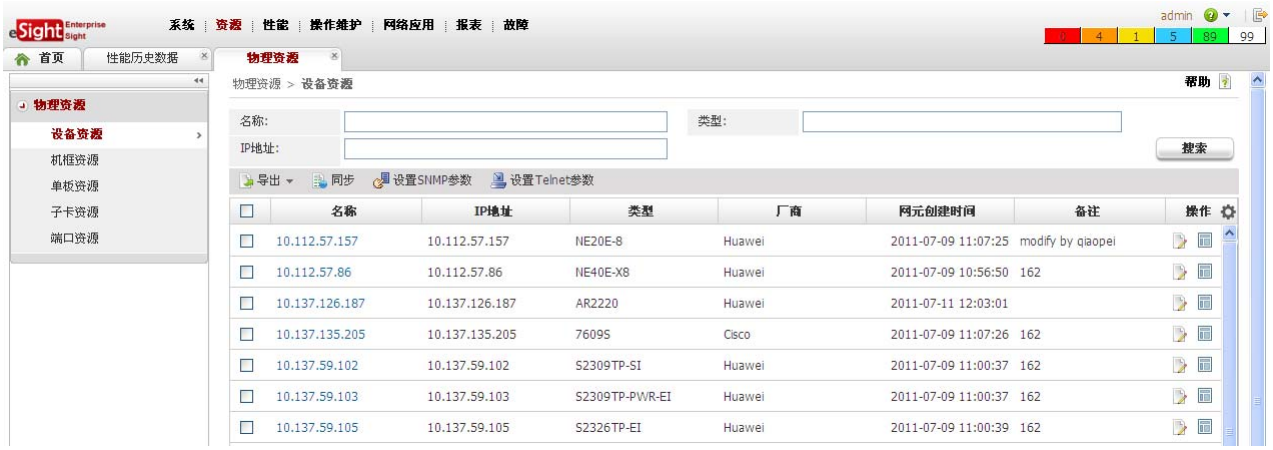
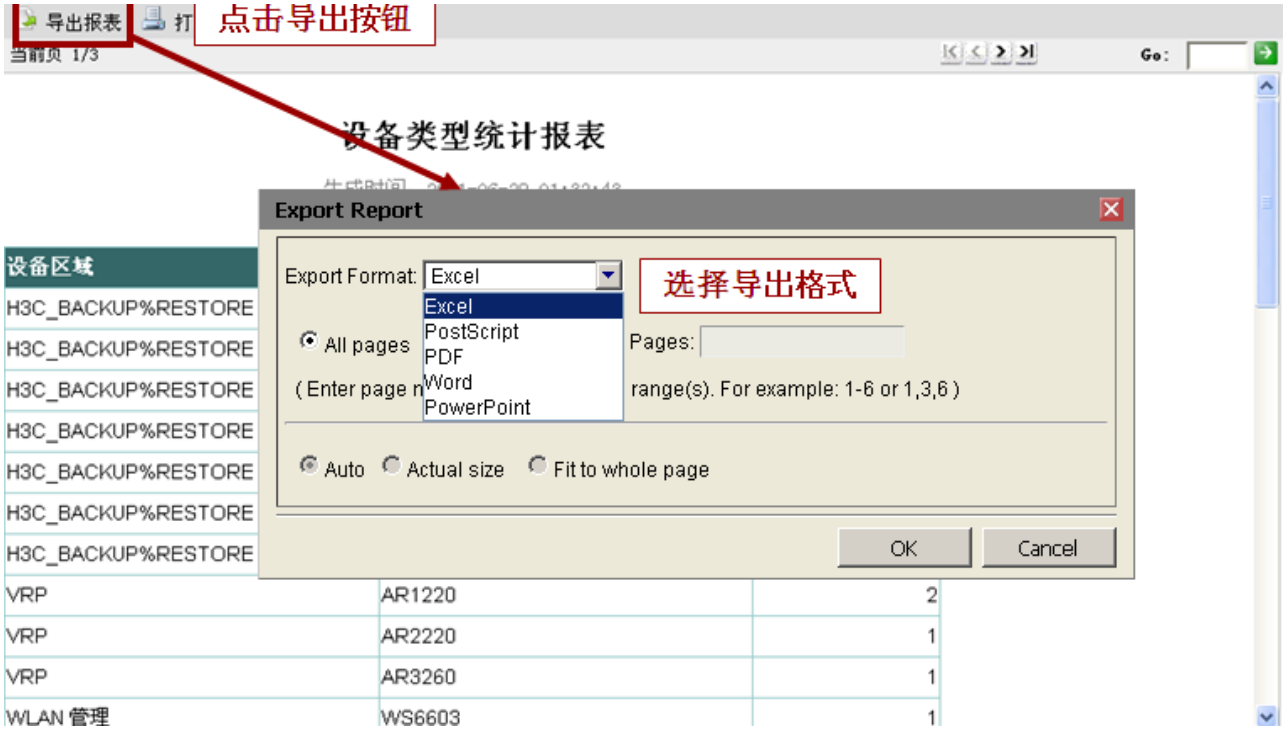


图6-10 查看导出报表



阶段维护

eSight 提供配置文件管理和备份功能，可以快速的进行文件备份和设备登陆管理。同时还提供系统巡检工具，能够定时的对设备进行自检，减轻网络维护人员的工作量。如图6-11、图 6-12 所示。

图6-11 配置文件管理

图6-12 配置文件备份

6.1.2 第三方设备定制场景

第三方设备定制概述

数据中心网络设备来自不同厂商，无法统一采用预集成的方式管理第三方设备，需要提供定制的能力，如果使用各自的网管系统进行管理，不仅增加了运维成本，而且极大的增加了网络维护人员的工作量。

华为公司 eSight 网管系统提供了对第三方设备管理能力的定制功能，包括对设备厂商信息、设备型号信息、告警参数、性能指标、设备面板、设备配置文件管理的定制功能，方便用户实际网络设备进行定制化的管理。满足对第三方设备的管理需求。

厂商信息定制

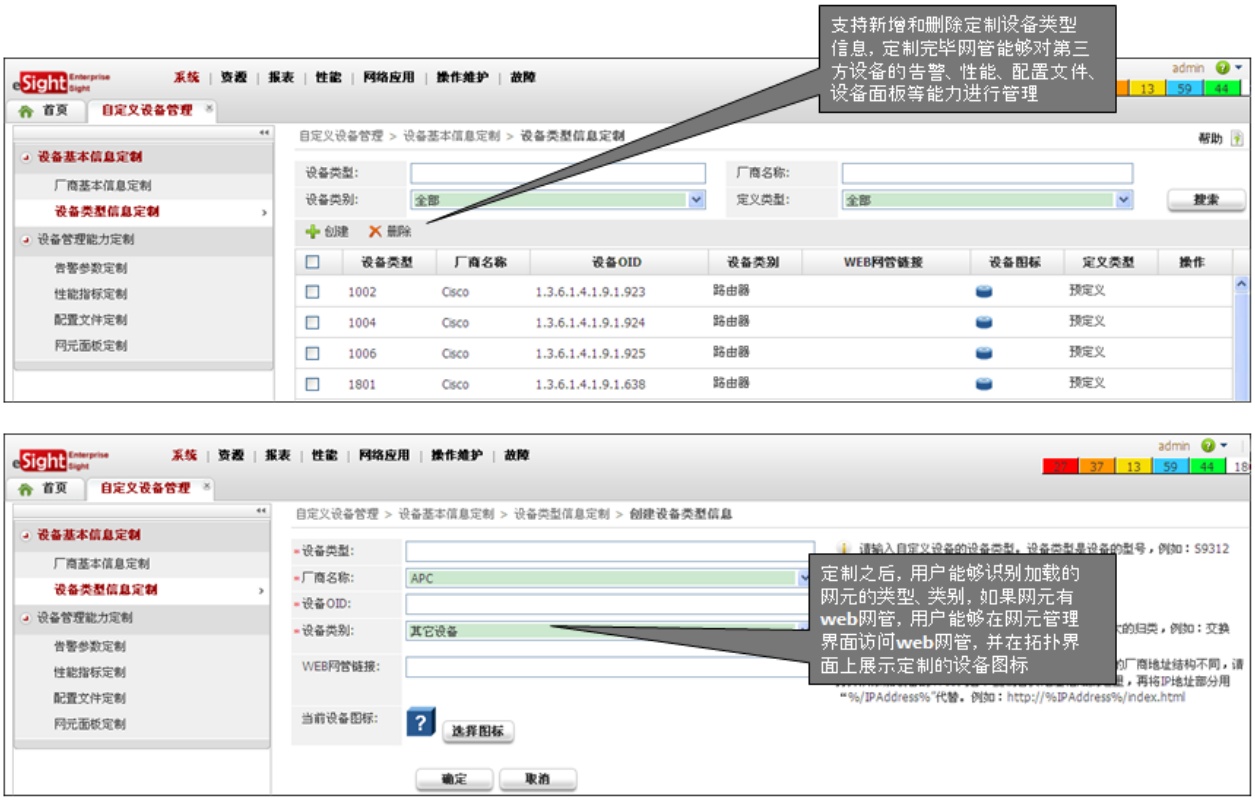
eSight 网管系统可以定制厂商的名称、联系人等信息，用于后续的设备类型定制。如图6-13 所示。

图6-13 厂商信息定制

设备类型定制

eSight 网管系统可以定制设备类型的描述、设备图标、Web 网管链接信息，定制的设备图标能在拓扑上显示。如图 6-14 所示。

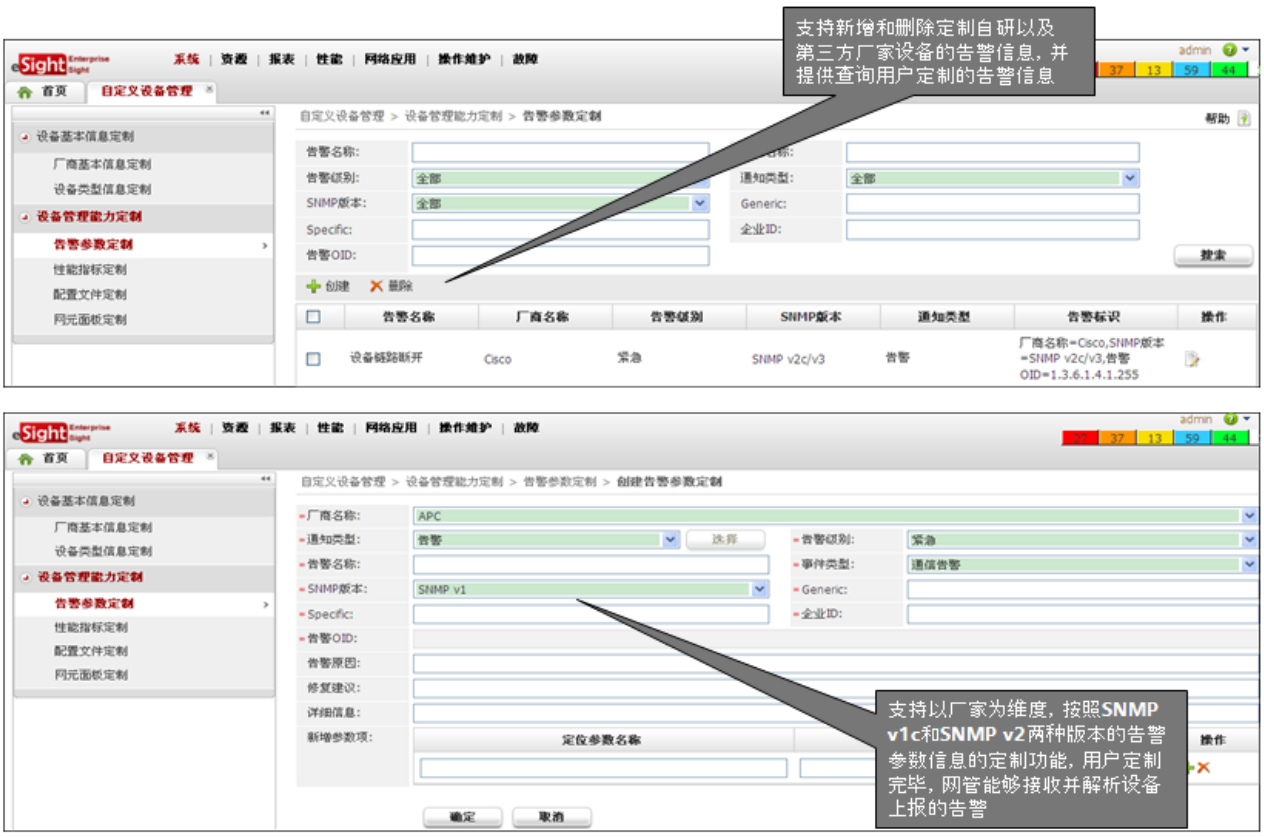
图6-14 设备类型定制



告警定制

eSight 网管系统可以对上报告警格式进行定制，定制后的告警能支持告警报文解析，并在在告警管理界面上进行显示。如图 6-15 所示。

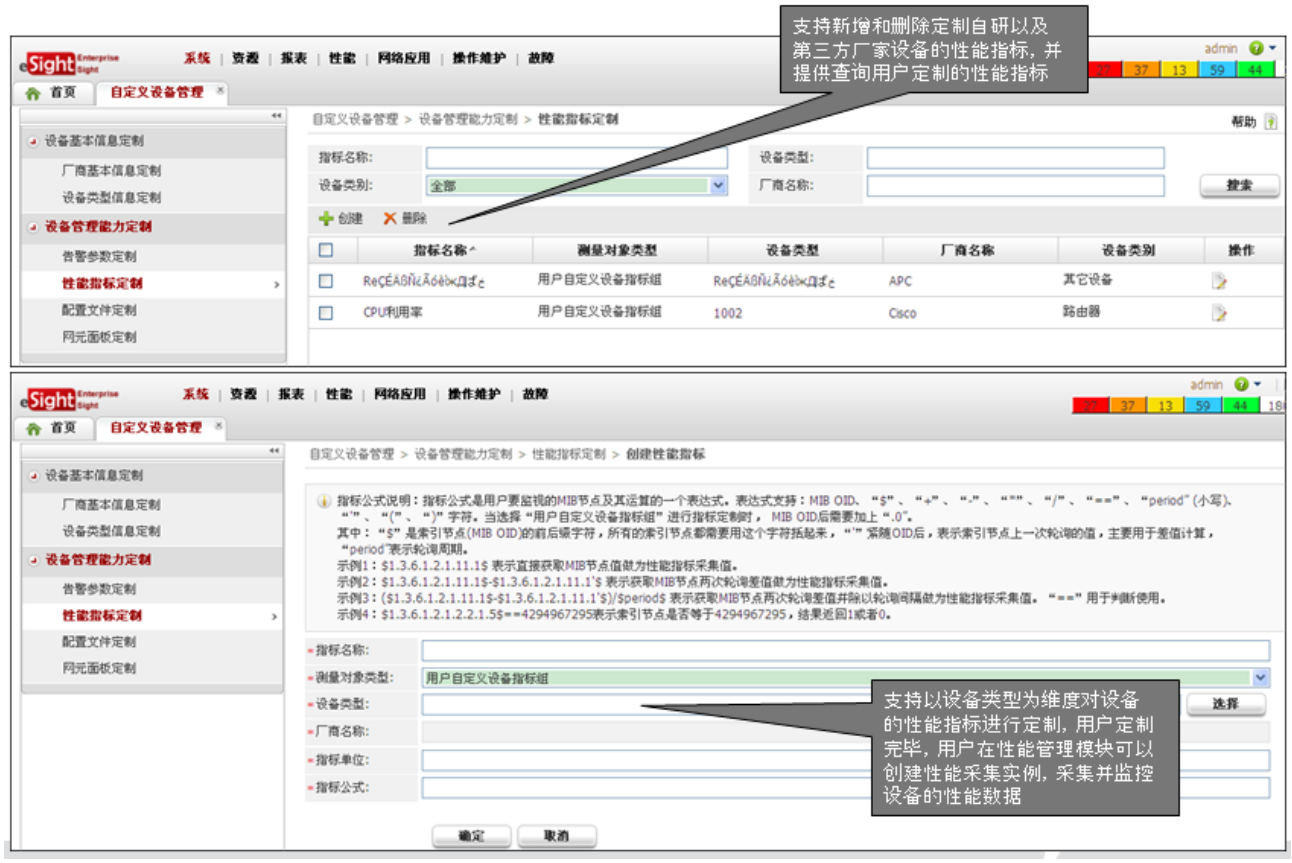
图6-15 告警定制



性能指标定制

eSight 网管系统可以对设备上支持的采集指标进行定制，定制后的性能指标能通过性能任务进行采集，在性能界面中进行数据浏览。如图 6-16 所示。

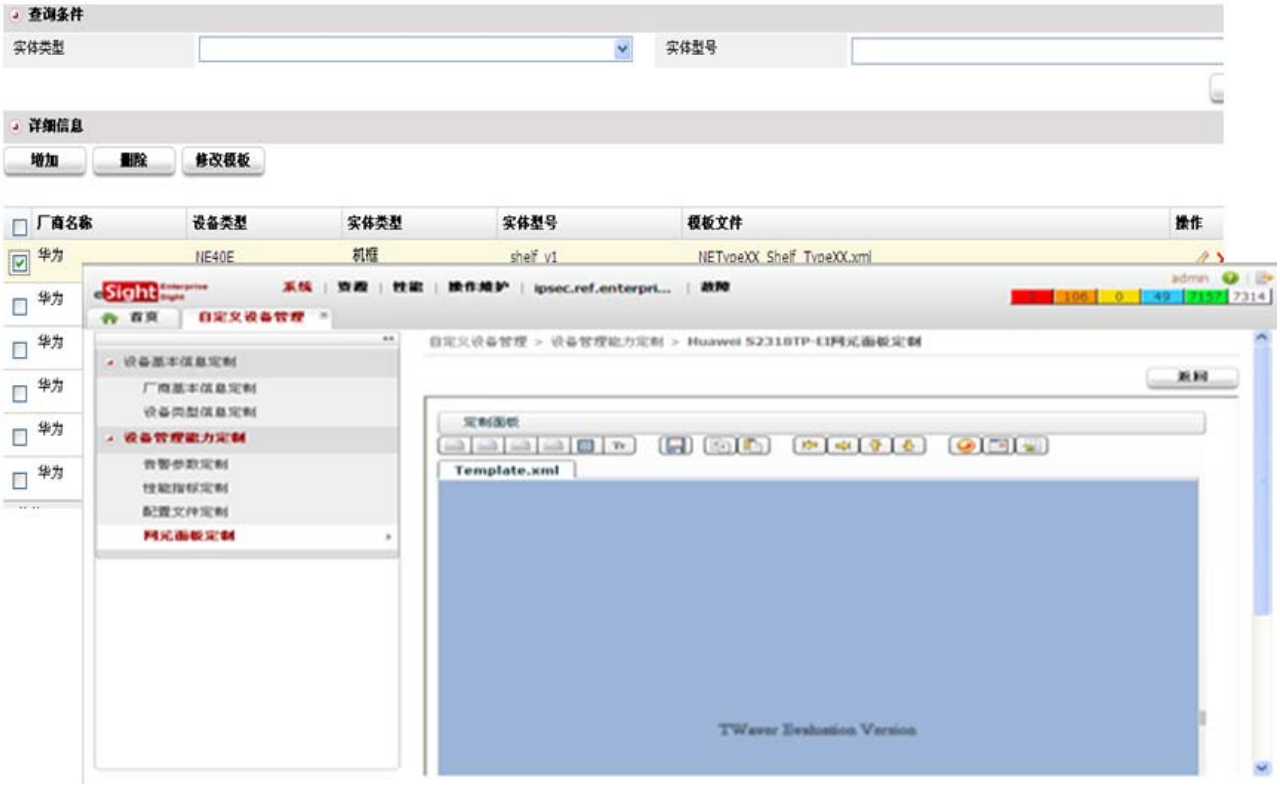
图6-16 性能指标定制



设备面板定制

eSight 网管系统可以对设备框、单板、子卡、端口进行仿真图定制，定制后的面板将显示新的仿真图。如图 6-17 所示。

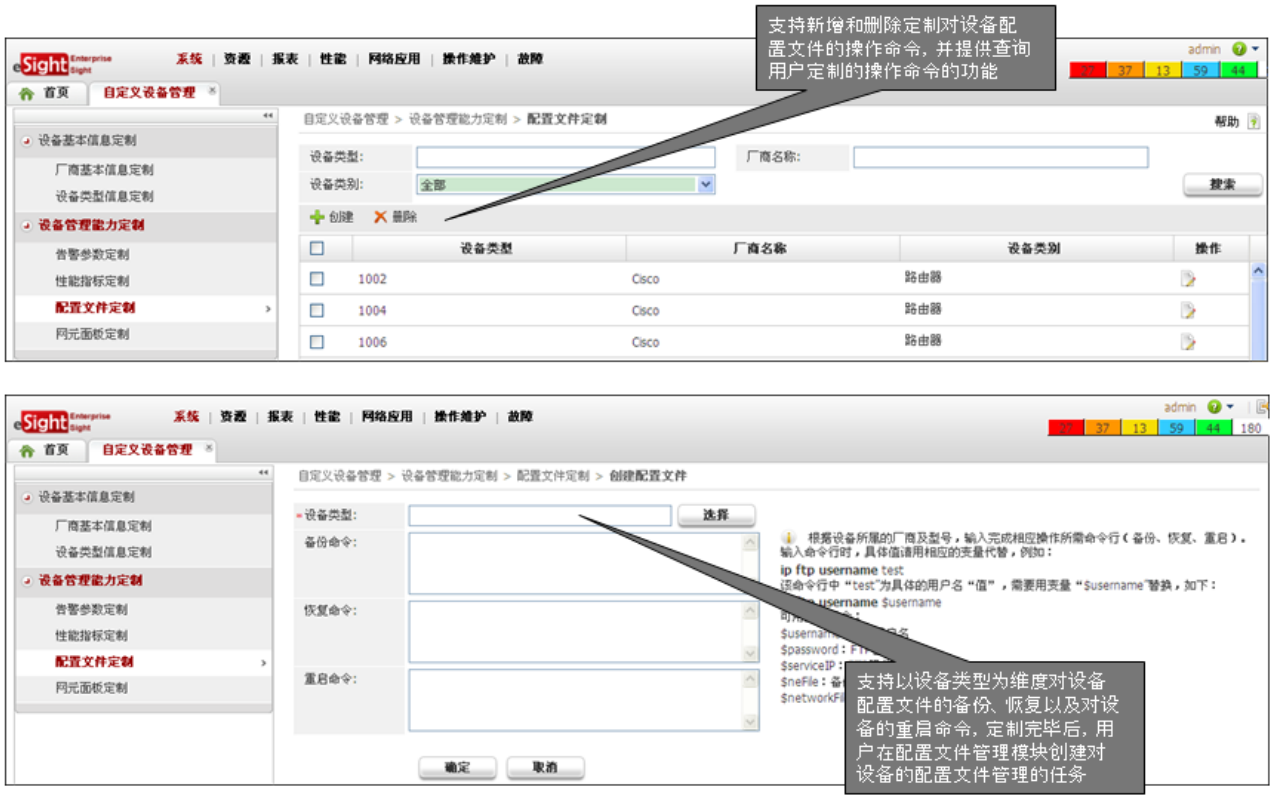
图6-17 设备面板定制



设备配置文件定制

eSight 网管系统可以针对第三方设备定制关于配置文件备份的备份、恢复、重启命令，支撑配置文件自动备份。如图 6-18 所示。

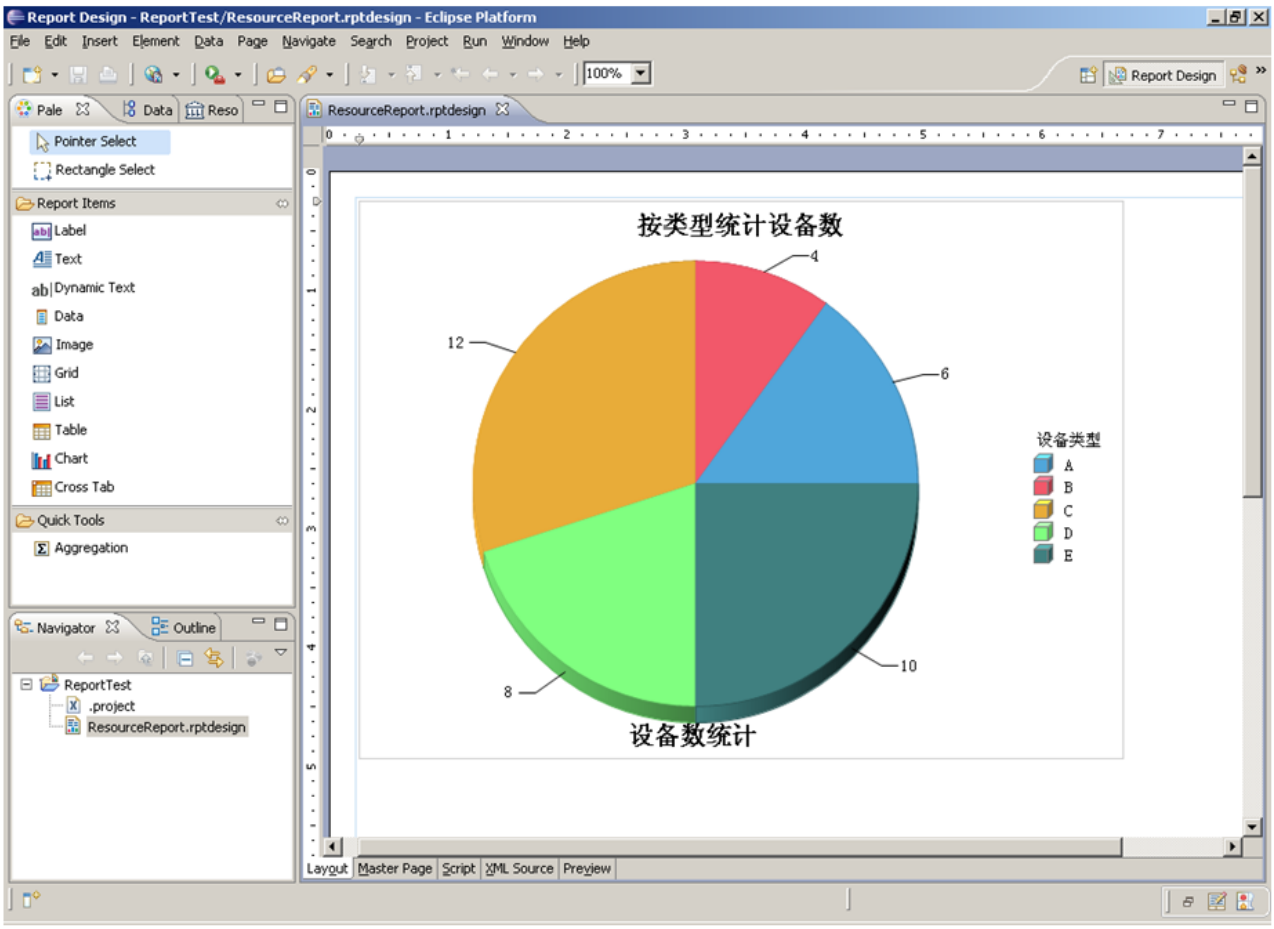
图6-18 配置文件备份和恢复定制



报表定制

eSight 提供强大的自定义报表能力。提供所见即所得的报表设计环境，可以修改现有的报表设计文件，生成新的设计文件。如图 6-19 所示。

图6-19 报表定制



6.1.3 软件升级和补丁加载场景

软件升级和补丁加载概述

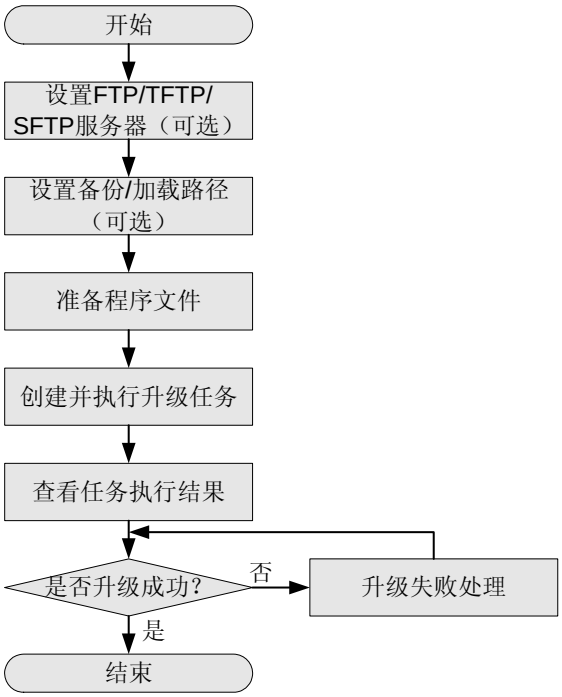
数据中心网络设备众多，如果使用一台台的方式去升级和加载补丁，不仅耗时耗力，而且容易出现人为原因造成的升级失败，需要考虑通过远程集中式进行统一升级和加载补丁。

eSight 网管系统提供了远程集中式软件升级和补丁加载机制，极大的减轻网络维护人员工作量，避免了人为原因造成的升级失败和补丁加载失败。

软件升级

eSight 网管系统提供远程集中式的软件升级功能，按照操作向导，轻松完成设备升级，并且对升级失败有相应的处理，避免升级失败后的设备状态异常。如图 6-20 所示。

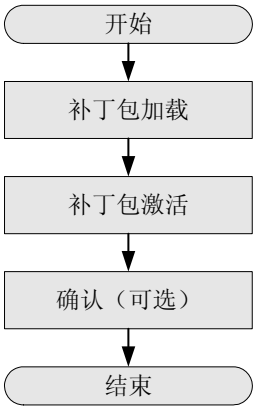
图6-20 软件升级流程



补丁加载

eSight 网管系统提供远程集中式的补丁加载功能，按照操作向导，轻松完成补丁加载，并且具有补丁回滚功能，可以将网元恢复到补丁升级前的状态。如图 6-21 所示。

图6-21 补丁升级流程



6.2 故障处理

数据中心网络系统是由网络设备、连接设备间链路和一些相关服务器组成。因此出现网络系统故障的原因也基本上从链路、网络设备状态、是否受病毒攻击、服务器状态等方面来查找。这些组件的任何一个出现故障，都会导致上层应用无法正常工作。

6.2.1 网络设备故障处理

网络设备发生故障可以分为几种：

- 设备宕机：设备上的电源或者其他指示灯都不亮，没有任何工作时的声响。
- 设备 CPU 使用率高：监控软件或者登陆设备时，发现设备的 CPU 利用率很高，同时相关应用响应较慢。
- 有错误消息：查看日志服务器或者登陆设备时，发现设备有错误消息。
- 有报警信息：设备状态指示灯报警，显示为红色等。

设备宕机

如果发现一旦发现设备宕机，首先检查电源连接线和机房电源。如果电源连接线和电源均正常，立即拨打设备提供商和服务提供商的服务号码，请求支持。如果发现设备硬件存在问题，可要求设备提供商和服务提供商在最短时间内做备件更换服务。

设备 CPU 利用率高

立即报告服务提供商，要求提供技术支持。待技术支持工程师远程处理或到场后，协助工程师找出设备 CPU 利用率高的原因。一般情况下，可以判断为设备受到病毒的攻击。

有错误消息

将错误消息发送给服务提供商，并跟踪进度。经过服务提供商分析后，给出错误消息的原因，如果设备有隐形的故障，可以预先做好相应的准备工作或者更换设备。

报警信息

报告服务提供商和设备提供商，要求对设备进行报警故障排除或者更换硬件。

6.2.2 服务器故障处理

跟网络系统相关的服务器主要有 DHCP 服务器、ACS 服务器、外网代理服务等。常见的故障现象包括：

- 不能正确获取 IP 地址。
- 不能正常登陆网络设备。
- 不能通过代理服务器上网。

不能正确获取 IP 地址

可以按照如下步骤进行故障处理：

1. 首先查看 DHCP 服务器的连通性，可以用 ping 的办法确定。
2. 如果 DHCP 服务器连通性正常，则可以登陆服务器，查看该服务器的 DHCP 服务是否正常；如果服务正常，可以查看是否网络当中有病毒，导致 DHCP 的请求消息超时。
3. DHCP 服务器在市局有备份服务器，在当前服务器不可用的情况下，可以替换当前的服务器。
4. 在 DHCP 服务器恢复正常工作前，我们也可以采用手动静态配置 IP 地址的方法来临时解决电脑访问网络的问题。

----结束

不能正常登陆网路设备

1. 首先查看该网络设备是否具有连通性。
2. 如果该设备可以 ping 通，可以尝试登陆服务器，看 ACS 的服务器的服务是否正常。
3. 如果服务不正常，可以考虑使用网络设备上的 Console 端口登陆设备，临时去掉 AAA 认证服务相关配置，启用网络设备的内置本地认证数据库进行临时登录认证。

----结束

不能通过代理服务器上网

1. 首先查看网络是否连通，是否可以访问其他的应用；然后测试到代理服务器的连通性。
2. 如果代理服务器连通性正常，则可以登陆服务器，查看该服务器的代理服务和相关系统服务是否正常。如果发现服务不正常的，可以尝试重启服务或者服务器来解决。
3. 如果重启代理服务器服务或系统后问题仍无法解决的，则需进一步检查代理服务器硬件是否存在故障，例如网卡等关键硬件。
4. 如果代理服务器硬件或者系统存在问题的，我们可以临时使用备用代理服务器来满足代理上网的需求。
5. 如果前述问题均不存在，一切正常，则可以测试到 Internet 的访问是否正常。我们应该对提供 Internet 服务的 DNS 和 ISP 网关进行 Ping 测试。如果 DNS 或 ISP 网关存在连通性问题，则及时联系 ISP 商排查解决。
6. 对于 ISP 提供的当前线路出现问题的，我们可以使用备份线路进行 Internet 访问。

----结束

6.3 网络扩容

网络扩容概述

随着数据中心业务、规模的不断增加，现有网络容量已经不能满足数据中心的长期发展，对数据中心网络的扩容迫在眉睫，在不影响现有业务的情况下实现平滑扩容，是数据中心网络扩容的基本要求。

数据中心网络扩容包含如下三种场景：

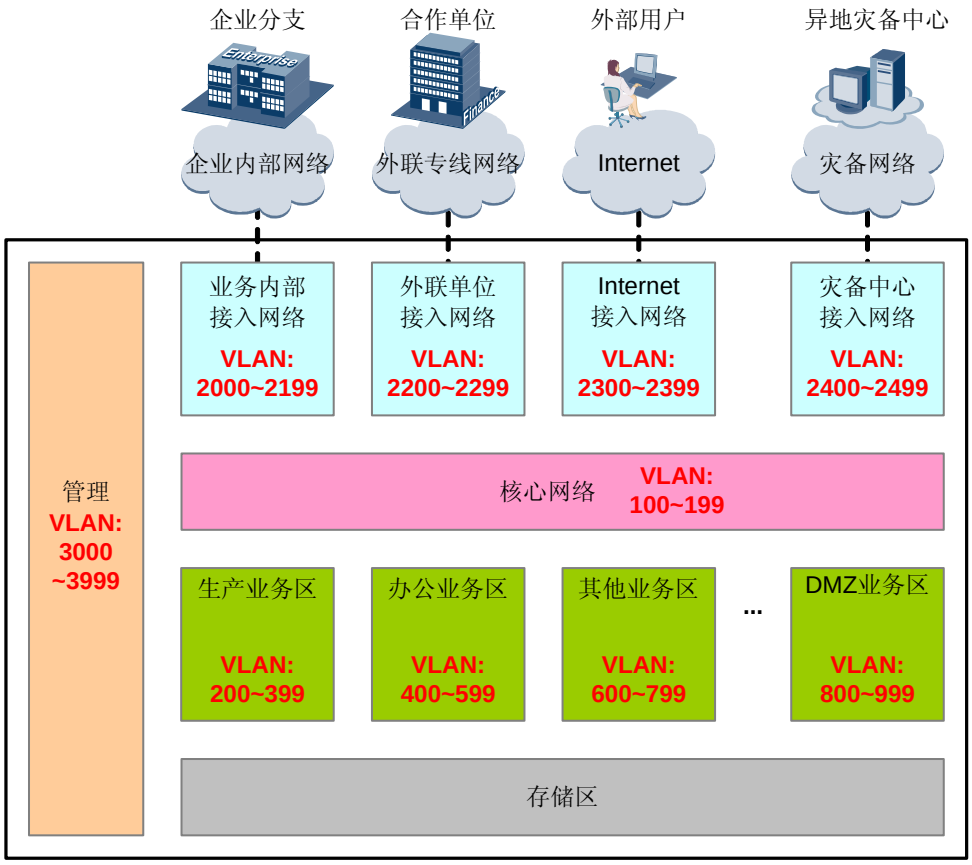
- 服务器扩容
- 网络设备扩容
- 链路带宽扩容

针对不同的扩容场景需要使用不同的扩容策略，实现业务的平滑过渡。

服务器扩容

服务器扩容包含在原区域扩容服务器和在新区域新建服务器两种情况，针对这两种情况，所采取的扩容策略不尽相同。

图6-22 数据中心内部架构



- 原区域扩容服务器

  随着生产业务的不断发展，当前生产业务区的服务器资源已经不能满足业务发展需要，需要进行生产业务区服务器的扩容，实现平滑扩容需要使用该区域初期规划好的 VLAN，保持 VLAN 的连续性，并且 IP 地址使用该区域初期规划好的地址段，这样做可以保证上游路由和防火墙策略不需要进行修正，便于维护的同时也减轻了扩容工作量。

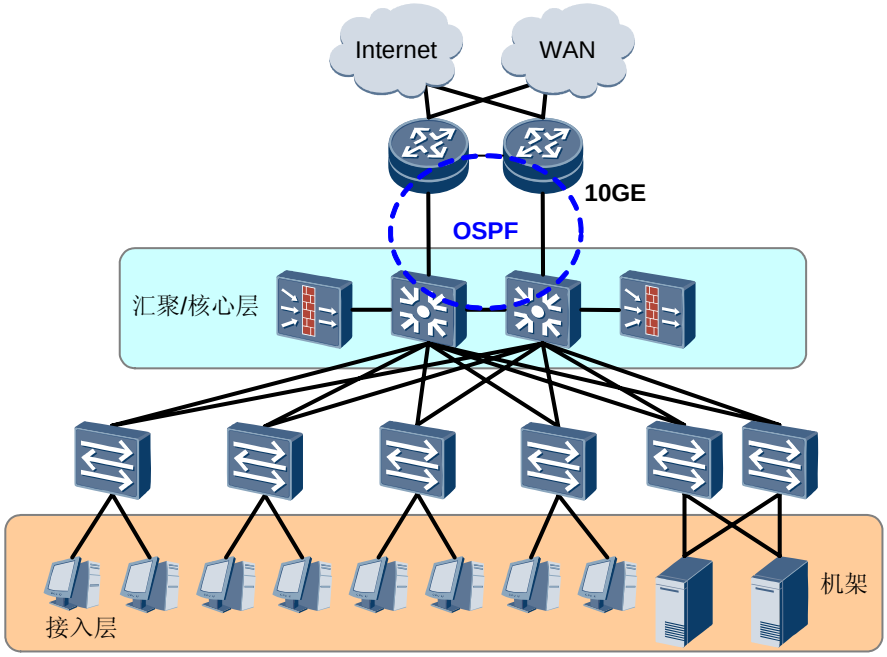
- 新区域新建服务器

假设 DMZ 区是新建区域，那么就需要为该区域重新规划 VLAN 资源和 IP 地址资源，重新进行路由和防火墙策略规划，这样做可以确保新区域的建设不会影响到现有业务，实现现有业务的平滑扩容，划分新的区域也便于今后运维管理。

网络设备扩容

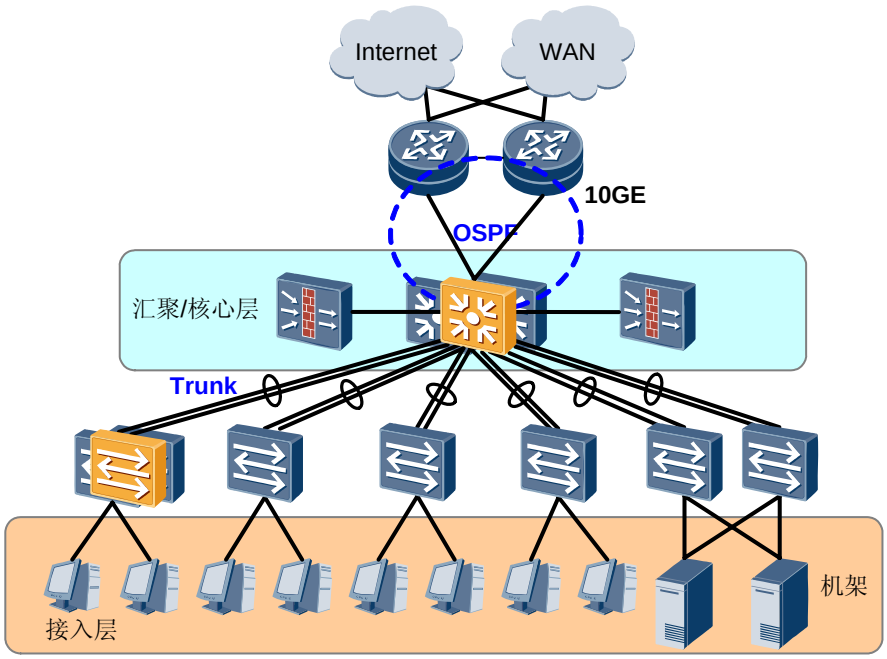
图 6-23 所示为通用的数据中心网络架构。可以看到在接入与汇聚层存在众多环网，一旦由于业务增长需要增加服务器资源，就需要增加接入层交换机并连接到汇聚/核心层，增加网络复杂度的前提下还要使用破环技术，不可避免的会对现网的业务产生影响。

图6-23 通用数据中心网络架构



华为公司推荐在数据中心网络架构采用堆叠和集群技术，首先消除破环协议，其次简化网络规模，并且利于网络设备扩容。如图 6-24 所示。

图6-24 堆叠/集群化的数据中心网络架构



使用堆叠和集群技术后，网络结构由环形简化为树形。首先利于网络运维管理，其次网络设备扩容时，只需要在原有的堆叠环境下新增设备，对网络结构不产生影响，也不需要添加物理链路到汇聚/核心层，实现数据中心业务的平滑扩容。

链路带宽扩容

随着数据中心业务的扩展，链路带宽也会成为数据中心业务的瓶颈，除了使用更换高性能、高带宽单板外（比如 GE 换成 10GE 单板，10GE 换成 40GE 单板），还可以通过链路捆绑技术进行链路带宽扩容，在不影响现网业务的情况下实现链路带宽的平滑扩容。

6.4 灾难应急

灾难应急概述

当数据中心出现意外灾害时（例如地震、火灾等），如何及时的进行应急处理，及时恢复数据中心业务，将数据中心业务损失降低到最小，这些都需要在网络设计时进行充分考虑。

灾难应急建议

业界主推的“两地三中心”方案已经充分考虑到意外灾害的发生，主中心、备份中心、容灾中心之间如何进行数据存储，某中心发生灾害时如何进行业务切换等请参见“5 多数据中心规划建议”。

7 产品建议

数据中心解决方案涉及的部件和产品如下。

- 核心交换机：S9300 系列核心交换机
- 接入交换机：S6700 系列接入交换机
- 接入交换机：S5700 系列接入交换机
- Mini OTN 系统：OSN 1800
- 光传送平台：OSN 6800 智能光传送平台

7.1 S9300 系列核心交换机

7.1.1 产品概述

Quidway® S9300 系列运营级园区核心交换机是由华为公司自主开发的新一代高性能核心路由交换机产品，提供大容量、高密度、模块化的二到四层线速转发性能，具有强大组播功能，完善的 QoS 保障、有效的安全管理机制和电信级的高可靠设计，满足高端用户对多业务、高可靠、大容量、模块化的需求，降低运营商的建网成本和维护成本。

S9300 可广泛应用于构建各种类型园区网核心层和汇聚层交换机，对于接入交换机性能和接口密度要求高的某些大型园区网，也可使用 S9300 系列交换机作为汇聚交换机使用。

7.1.2 产品型号

S9300 系列核心交换机的产品型号如下。

表7-1 S9300 系列产品型号

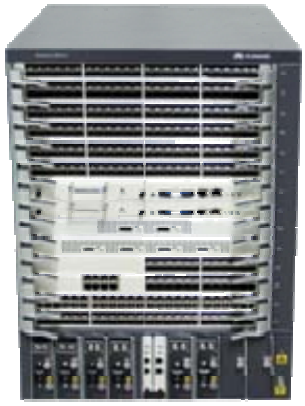
产品型号	描述
S9303	支持 3 块 LPU 交换网容量 720Gbit/s 背板容量 3Tbit/s 转发能力 540Mpps

产品型号	描述
S9306	支持 6 块 LPU 交换网容量 2Tbit/s 背板容量 6Tbit/s 转发能力 1080Mpps
S9312	支持 6 块 LPU 交换网容量 2Tbit/s 背板容量 6Tbit/s 转发能力 1080Mpps

图7-1 S9303 外观图

图7-2 S9306 外观图

图7-3 9312 外观图



7.1.3 产品特点

先进的体系结构、高性能、配置灵活

S9300 系列交换机采用先进的全分布式体系结构设计，采用业界最新的硬件转发引擎技术，所有端口支持的业务能够线速转发，业务包括 IPv4/MPLS/二层转发等。支持 ACL 线速转发。

S9300 系列交换机实现组播线速转发，硬件完成两级复制：交换网板复制到接口板和转发引擎复制到接口。

S9300 支持 2Tbps 交换容量，支持多种高密度板卡，满足核心、汇聚层设备大容量、高端口密度的要求，可以满足用户日益增长的带宽需求，能够极大的保护和节约用户投资。

完善的安全机制

支持完善的 AAA（Authentication，Authorization and Accounting）机制，根据策略对接入用户进行认证、授权和计费。支持 802.1X、Portal、Guest VLAN，支持用户动态接入认证，与其他主流厂商的 NAC 互通。

支持路由协议加密、MAC 地址过滤、动态 ARP 检测、ACL 等等一系列安全特性，可以为服务提供商以及网络的最终用户提供数据保护。基于硬件的包过滤和采样，可实现高性能和高扩展性。

提供 2 级 CPU 保护机制，支持 1K CPU 硬件保护队列，可实现数据和控制的分离处理，防止拒绝服务攻击、非法接入以及控制平面过载等安全威胁，提供业界领先的一体化安全解决方案。

全面的可靠性

9300 具备超越 5 个 9 的运营级高可靠性，主控、电源、风扇等关键部件采用冗余设计，所有模块均支持热插拔。基于分布式的硬件转发架构，路由平面和数据交换平面严格分离，保证业务流永续畅通。

独立的故障检测定位硬件，提供 3.3ms 高精度硬件级以太 OAM 功能，实现快速故障检测与定位，与其他倒换技术联动可有效保证毫秒级网络保护。

能够在冗余控制引擎间实现无缝切换，设备优雅重启无中断转发。支持 ISSU 业务无缝升级，减少关键业务和服务中断。

支持 IEEE 802.3ad 链路汇聚、IEEE 802.1s/w 和虚拟路由器冗余协议（VRRP），同时支持丰富的毫秒级倒换技术如 RRPP、SmartLink、IP FRR、TE FRR、VPN FRR 等，实现运营级级高可靠性。

7.1.4 主要指标/规格

表7-2 S9300 系列产品主要指标/规格

指标/规格	S9303	S9306	S9312
背板容量	1.2Tbps	2.4Tbps	4.8Tbps
业务槽位	3	6	12
GE 端口密度	144	288	576
10G 端口密度	36	72	144
VLAN	支持 Access、Trunk、Hybrid 方式 支持 default VLAN 支持 VLAN 交换 支持 QinQ、增强型灵活 QinQ		
MAC 地址功能	支持 MAC 地址自动学习和老化 支持静态、动态、黑洞 MAC 表项 支持源 MAC 地址过滤 支持基于端口和 VLAN 的 MAC 地址学习限制		
STP	支持 STP，RSTP 和 MSTP 支持 BPDU 保护、Root 保护、环路保护 支持 BDPU Tunnel		
IP 路由	支持 RIP、OSPF、ISIS、BGP 等 IPv4 动态路由协议 支持 RIPng、OSPFv3、ISISv6、BGPv4 等 IPv6 动态路由协议		
组播	支持 IGMP Snooping 功能 支持用户快速离开机制 支持组播流量控制 支持组播查询器 支持组播协议报文抑制功能 支持组播 ACL		

指标/规格	S9303	S9306	S9312
MPLS	支持 MPLS 基本功能 支持 MPLS OAM 支持 MPLS TE 支持 MPLS VPN/VLL/VPLS		
时钟特性	支持同步以太时钟 支持 1588V2		
QoS	支持基于 Layer2 协议头、Layer3 协议、Layer4 协议、802.1p 优先级等的组合流分类 支持 ACL、CAR、Remark、Schedule 等动作 支持 PQ、WRR、DRR、PQ+WRR、PQ+DRR 等队列调度方式 支持 WRED、尾丢弃等拥塞避免机制 支持流量整形		
配置与维护	支持 Console、Telnet、SSH 等终端服务 支持 SNMPv1/v2/v3 等网络管理协议 支持通过 FTP、TFTP 方式上载、下载文件 支持 BootROM 升级和远程在线升级 支持热补丁 支持用户操作日志		
安全和管理	802.1x 认证，Portal 认证 支持 RADIUS 和 HWTACACS 用户登录认证 命令行分级保护，未授权用户无法侵入 支持防范 DoS 攻击、TCP 的 SYN Flood 攻击、UDP Flood 攻击、广播风暴攻击、大流量攻击 支持 CPU 通道的保护 支持 ICMP 实现 ping 和 traceroute 功能 支持 RMON		
机箱尺寸 mm （宽×深×高）	442×476×175	442×476×442	442×476×664
机箱重量（空配）	<15Kg	<30Kg	<45Kg
工作电压	DC：–38.4V～–72V AC：90V～264V		
典型功耗	180W	<350W	<650W
整机供电能力 （不含 PoE）	800W	1600W	1600W

7.2 S6700 系列接入交换机

7.2.1 产品概述

Quidway® S6700 系列交换机是华为公司自主开发的下一代全万兆盒式交换机，可用于数据中心万兆服务器接入、城域网汇聚及园区网的核心。

S6700 系列是业内最高性能的交换机之一，同时提供最多 24/48 个全线速万兆接口，使万兆服务器高密度接入和园区网高密度万兆汇聚成为可能。同时，S6700 支持丰富的业务特性、完善的安全控制策略、丰富的 QoS 等特性以满足数据中心扩展性、可靠性、可管理性、安全性等诸多挑战。

7.2.2 产品型号

S6700 目前有两款主机形态。

- S6748-EI：支持 48 个 GE SFP/10GE SFP+端口，双电源槽位，含 USB 接口。
- S6724-EI：支持 24 个 GE SFP/10GE SFP+端口，双电源槽位，含 USB 接口。

图7-4 S6748-EI 外观图

图7-5 S6724-EI 外观图

7.2.3 产品特点

大容量高密万兆灵活接入

随着用户端带宽不断提高，服务器万兆网卡的应用越来越广泛，尤其是数据中心服务器交换机需要提供更高的转发性能和万兆端口扩展能力。S6700 系列盒式交换机拥有业内同档次设备最高的万兆端口密度、最大交换容量，单台设备可以最多支持 48 个全线速转发的万兆端口。

端口支持 GE 和 10GE 灵活接入，自动识别安装光模块类型，亦可通过配置光电转换模块接入千兆光/电口服务器。从而最大程度保护用户投资和确保使用的灵活性。

针对于数据中心数据流量大和无阻塞传输的要求，S6700 交换机可以提供强大的缓存能力，并且支持先进的缓存调度机制可以保证设备缓存能力有效利用的最大化。

完善的安全控制策略

S6700 提供多种安全保护功能。支持 DoS（Denial of Service）类防攻击、网络的防攻击、用户的防攻击等功能。其中 DoS 类防攻击主要包括 SYN Flood、Land、Smurf、ICMP Flood。网络的防攻击主要是指 STP 的 bpdu/root 攻击。用户的防攻击涉及 DHCP 仿冒攻击、中间人攻击、IP/MAC Spoofing 攻击、DHCP request flood、改变 CHADDR 值的 DoS 攻击等等。

S6700 支持通过建立和维护 DHCP Snooping 绑定表，侦听接入用户的 MAC/IP 地址、租用期、VLAN-ID、接口等信息，解决 DHCP 用户的 IP 和端口跟踪定位问题；对不符合绑定表项的非法报文（ARP 欺骗报文、擅自修改 IP 地址等）直接丢弃，有效防止黑客或攻击者通过 ARP 报文实施园区网常见的“中间人”攻击。利用 DHCP Snooping 的信任端口特性还可以保证 DHCP Server 的合法性。

S6700 支持 ARP 表项严格学习功能，可以防止因 ARP 欺骗攻击将交换机 ARP 表项占满，导致正常用户无法上网。同时，支持 IP Source Check 特性，防止包括 MAC 欺骗、IP 欺骗、MAC/IP 欺骗在内的非法地址仿冒带来的 DoS 攻击；S6700 支持 URPF 功能，保证端口接收到数据包时，会反向查找路径验证数据包的真实性，从而有效地杜绝了网络中日益泛滥的源地址欺骗。

S6700 支持集中式 MAC 地址认证和 802.1X 认证，支持用户账号、IP、MAC、VLAN、端口、客户端是否安装病毒防范等用户标识元素的，同时实现用户策略（VLAN、QoS、ACL）的动态下发。

S6700 支持基于端口的源 MAC 地址学习限制功能，有效防止用户源 MAC 欺骗冲击设备 MAC 表项，导致正常用户无法学到 MAC 表而泛洪的问题等。

完备的可靠性保护机制

S6700 支持双电源冗余供电，也可以交、直流同时输入。用户可灵活选择单电源工作模式或者双电源工作模式，提高了设备可靠性；内置两个风扇提升了稳定性，设备 MTBF 时间业界领先。

S6700 对传统的 STP/RSTP/MSTP 生成树进行了增强，支持 MSTP 多进程，大大提高接入环子环实例数目。还支持 SmartLink 和 RRPP 等增强型以太网技术，可以实现毫秒级链路保护倒换，保证高可靠性的网络质量。此外，针对 SmartLink 和 RRPP 均提供多实例功能，可实现链路负载分担，进一步提高了链路带宽利用率。

S6700 支持增强 E-Trunk 功能。此功能应用于 CE 双归接入 VPLS、VLL、PWE3 网络时，CE 与 PE 间的链路保护以及对 PE 设备节点故障的保护。利用 E-trunk 技术可以实现跨设备的链路聚合，从而将链路可靠性提高到了设备级。

S6700 支持智能以太保护 SEP（Smart Ethernet Protection），SEP 是一种专用于以太网链路层的环网协议。适用于半环组网场景，部署时可独立于上层汇聚设备，并提供毫秒级的快速业务倒换性能。保证业务的不中断。在华为设备上已经利用 SEP 协议实现了以太网链路管理。SEP 协议简单可靠、倒换性能高、维护方便、拓扑灵活，可以大大方便用户进行网络的管理和规划。

S6700 支持 VRRP 虚拟路由冗余协议，构建 VRRP 备份组，保持通讯的连续性和可靠性，有效保障网络稳定。支持在设备上配置多条等价路由的方式实现上行路由的冗余备份，当主上行路由发生故障时自动切换到下一个备份路由上去，实现上行路由的多级备份。

丰富的 QoS 控制

S6700 系列交换机支持多种 QoS 控制手段，可以基于五元组、IP 优先级、TOS、DSCP、IP 协议类型、ICMP 类型、TCP 源端口、VLAN、以太网帧协议类型、CoS 等信息，实现复杂流流分类功能，支持双向 ACL。S6700 支持基于流的双速三色限速功能，每端口支持 8 个优先级队列，支持 WRR、DRR、SP、WRR＋SP、DRR+SP 多种队列调度算法和 WRED 拥塞避免机制，有效地保证了话音、视频和数据等网络业务不同的质量要求。

良好的可扩展性

S6700 支持长距智能堆叠 iStack 功能，普通端口可以通过命令行配置为堆叠端口，使端口应用更加灵活。通过光纤进行堆叠还可大幅增加堆叠的距离，突破了传统堆叠距离的限制。智能堆叠和单一设备相比，在扩展性、可靠性、性能等方面均具有强大的优势。

当客户需要扩容或者有单个设备故障需要替换时可实现“新增设备”的热插拔，从而减少了业务中断对客户的影响；相对于框式交换机来说，智能堆叠在性能和端口密度方面突破了硬件架构的限制。而对管理来说，堆叠后的数台设备在逻辑上可认为是一台，减轻了网络管理和配置的工作量。

贴心的可管理性

S6700 采用“一次进站”方案，支持自动配置、即插即用、USB 开局、自动批量远程升级功能，便于部署升级和业务发放，简化后续的管理和维护性能。从而大大降低了维护成本。

S6700 支持 SNMP V1/V2/V3，CLI 命令行、Web 网管、TELNET、HGMP 集群管理等多样化的管理和维护方式，设备管理更加灵活。支持 NTP、SSHv2.0、TACACS+、RMON、多日志主机、基于端口的流量统计，支持 NQA 网络质量分析，有利于进一步作好网络规划和改造。

S6700 支持 GVRP。GVRP 是一种 VLAN 的动态配置技术，在复杂的组网环境中应用 GVRP，能够简化 VLAN 配置管理，减少因为配置不一致而导致的网络互通问题。从而达到减少网络管理员的手工配置量及保证 VLAN 配置正确的目的。

S6700 支持 MUX VLAN 功能。MUX VLAN 提供了一种在 VLAN 的端口间进行二层流量隔离的机制。采用两层 VLAN 隔离技术，只有上层 VLAN 全局可见，下层 VLAN 相互隔离。MUX VLAN 通常用于企业内部网，客户端口可以同服务器端口通讯，但客户端口之间不能通讯。用来防止连接到某些接口或接口组的网络设备之间的相互通信，但却允许与默认网关进行通信。从而对企业内部的资源共享及安全通信起到了至关重要的作用。

S6700 支持 BFD 链路快速检测功能，能为 OSPF、ISIS、VRRP、PIM 等协议提供毫秒级检测机制，提高了网络可靠性。S6700 遵循 IEEE 802.3ah 和 802.1ag 提供点到点以太网故障管理功能，可以用于检测用户链路上的故障。以太网 OAM 能够有效提高以太网的网络管理维护能力，保障网络的稳定运行。

丰富的 IPv6 特性

S6700 提供双协议栈，可平滑升级。硬件支持 IPv4/IPv6 双栈和 IPv6 over IPv4 隧道（包括手工 Tunnel、6to4 Tunnel、ISATAP Tunnel），三层线速转发。既可以用于纯 IPv4 或 IPv6 网络，也可以用于 IPv4 到 IPv6 共存的网络，组网方式灵活，充分满足当前网络从 IPv4 向 IPv6 过渡的需求。

S6700 系列交换机支持丰富的 IPv6 路由协议，包括 RIPng、OSPFv3。支持 ND，管理邻居节点的交互。支持 PMTU，可以找到从源端到目的端的路径上一个合适的 MTU 值，以便有效地利用网络资源并得到最佳的吞吐量。

7.2.4 主要指标/规格

表7-3 S6700 系列产品主要指标/规格

项目	S6724-EI	S6748-EI
端口描述	24 个 GE SFP/10GE SFP+端口	48 个 GE SFP/10GE SFP+端口
转发性能（PPS）	358M	715M
端口交换容量（bps）	480G	960G
MAC 地址表	128K 地址容量 支持 MAC 地址自动学习和老化 支持静态、动态、黑洞 MAC 表项 支持源 MAC 地址过滤	
VLAN 特性	支持 4K 个 VLAN 支持 Guest VLAN、Voice VLAN 支持基于 MAC/协议/IP 子网/策略/端口的 VLAN 支持 1:1 和 N:1 VLAN 交换功能 支持基本、灵活 QinQ 功能	
IPv4 路由	静态路由、RIP V1/2、ECMP、支持 URPF OSPF、IS-IS、BGP 支持 VRRP 支持策略路由 支持路由策略	
IPv6 路由	支持静态路由 支持 RIPng 支持手动隧道 支持 6to4 隧道 支持 ISTAP 隧道	

项目	S6724-EI	S6748-EI
IPv6 特性	支持 ND（Neighbor Discovery） 支持 PMTU 支持 IPv6 Ping、IPv6 Tracert、IPv6 Telnet 支持 6to4、ISATAP、手动配置 Tunnel 支持基于源 IPv6 地址、目的 IPv6 地址、四层端口、协议类型等 ACL 支持 MLD v1/v2 snooping	
组播	支持二层静态组播 MAC 支持 MAC 模式转发 支持 IGMP Snooping 和快速离开机制 支持组播 VLAN 支持 MLD Snooping 支持 IGMP Proxy 支持可控组播 基于端口的组播流量统计 支持 IGMP v1/v2/v3 支持 PIM-SM、PIM-DM、PIM-SSM 支持 MSDP	
QoS/ACL	支持对端口接收和发送报文的速率进行限制 支持报文重定向 支持基于端口的流量监管，支持双速三色 CAR 功能 每端口支持 8 个队列 支持 WRR、DRR、SP、WRR＋SP、DRR+SP 队列调度算法 支持 WRED 支持报文的 802.1p 和 DSCP 优先级重新标记 支持 L2（Layer 2）~L4（Layer 4）包过滤功能，提供基于源 MAC 地址、目的 MAC 地址、源 IP 地址、目的 IP 地址、端口、协议、VLAN 的非法帧过滤功能 支持基于队列限速和端口整形功能	

项目	S6724-EI	S6748-EI
可靠性	支持 STP/RSTP/MSTP 协议 支持 BPDU 保护、根保护和环回保护 支持 RRPP 环型拓扑和 RRPP 多实例 支持 SmartLink 树型拓朴和 SmartLink 多实例，提供主备链路的毫秒级保护 支持智能以太保护协议(SEP) 支持 BFD for OSPF/ISIS/VRRP/PIM 协议 支持增强 Trunk（E-trunk）	
安全特性	用户分级管理和口令保护 支持防止 DOS、ARP 攻击功能、ICMP 防攻击 支持 IP、MAC、端口、VLAN 的组合绑定 支持端口隔离、端口安全、Sticky MAC 支持黑洞 MAC 地址 支持 MAC 地址学习数目限制 支持 IEEE 802.1X 认证，支持单端口最大用户数限制 支持 AAA 认证，支持 Radius、TACACS+、NAC 等多种方式 支持 SSH V2.0 支持 HTTPS 支持 CPU 保护功能 支持黑名单和白名单	
管理和维护	支持堆叠（业务口实现） 支持 MAC 地址强制转发（MFF） 支持虚拟电缆检测(VCT) 支持以太网 OAM（802.3ah 和 802.1ag） 支持本地端口镜像和远程端口镜像（RSPAN），支持观察端口正常转发报文 支持 Telnet 远程配置、维护 支持 SNMPv1/v2/v3 支持 RMON 支持网管系统、支持 WEB 网管特性 支持集群管理 HGMP 支持系统日志、分级告警 支持 GVRP 协议 支持 MUX VLAN 功能 支持 802.3az（能效以太网 EEE）	

项目	S6724-EI	S6748-EI
环境要求	温度范围：长期工作温度：0 OC～45 OC；短期工作温度：-5 OC～50 OC；相对湿度：10%～90%（无凝露）	
输入电压	AC： 额定电压范围：100-240V AC；50/60Hz 最大电压范围：90-264V AC；50/60Hz	
外形尺寸 mm （宽×深×高）	442×420×43.6	
功耗	165w	237w

7.3 S5700 系列接入交换机

7.3.1 产品概述

Quidway® S5700 系列全千兆交换机（以下简称 S5700），是华为公司为满足大带宽接入和以太网多业务汇聚而推出的新一代全千兆高性能以太网交换机，可为客户提供强大的以太网功能。S5700 基于新一代高性能硬件和华为公司统一的 VRP®（Versatile Routing Platform）软件，具备大容量、高密度千兆端口，可提供万兆上行，充分满足客户对高密度千兆和万兆上行设备的需求。S5700 可满足运营商园区网汇聚、企业网汇聚、IDC 千兆接入以及企业千兆到桌面等多种场合的需求。

S5700 系列以太网交换机为盒式设备，机箱高度为 1U，提供精简版（LI）、标准版（SI）和增强版（EI）、高级版（HI）四种产品版本。精简版提供完备的二层功能；标准版支持二层和基本的三层功能；增强型支持完善的路由协议和丰富的业务特性；高级版除了提供上述增强版本的功能外，支持 MPLS、硬件 OAM 等高级功能。

7.3.2 产品外观

S5700 系列交换机包括如下型号。

表7-4 S5700 系列交换机型号

设备型号	设备外观	描述
S5706TP-LI		• 4 个 10/100/1000Base-T • 2 个 1000M Combo 口 • 交流供电

设备型号	设备外观	描述
S5724TP-SI		• 20 个 10/100/1000Base-T • 4 个 100/1000Base-X 千兆 Combo 口 • 分交流供电和直流供电两种机型 • 支持 RPS 12V 冗余电源 • 支持 USB 口
S5724TP-PWR-SI		• 20 个 10/100/1000Base-T • 4 个 100/1000Base-X 千兆 Combo 口 • 可插拔双电源、交流供电 • 支持 PoE • 支持 USB 口
S5748TP-SI		• 44 个 10/100/1000Base-T • 4 个 100/1000Base-X 千兆 Combo 口 • 分交流供电和直流供电两种机型 • 支持 RPS 12V 冗余电源 • 支持 USB 口
S5748TP-PWR-SI		• 44 个 10/100/1000Base-T • 4 个 100/1000Base-X 千兆 Combo 口 • 交流供电 • 支持 PoE • 支持 USB 口
S5728C-SI		• 24 个 10/100/1000Base-T • 4 个 100/1000 Base-X 千兆 Combo 口 • 上行支持 2×10GE XFP、4×1000Base-X SFP、2×10GE SFP+或4×10GE SFP+插卡 • 双电源，可插拔 • 支持 USB 口
S5728C-PWR-SI		• 24 个 10/100/1000Base-T • 4 个 100/1000 Base-X 千兆 Combo 口 • 上行支持 2×10GE XFP、4×1000Base-X SFP、2×10GE SFP+或、4×10GE SFP+插卡 • 双电源，可插拔，交流供电 • 支持 PoE • 支持 USB 口

设备型号	设备外观	描述
S5752C-SI		• 48 个 10/100/1000Base-T • 上行支持 2×10GE XFP、4×1000Base-X SFP、2×10GE SFP+或4×10GE SFP+插卡 • 双电源，可插拔 • 支持 USB 口
S5752C-PWR-SI		• 48 个 10/100/1000Base-T • 上行支持 2×10GE XFP、4×1000Base-X SFP、2×10GE SFP+或4×10GE SFP+插卡 • 双电源，可插拔，交流供电 • 支持 PoE • 支持 USB 口
S5728C-EI		• 24 个 10/100/1000Base-T • 上行支持 2×10GE XFP、4×1000Base-X SFP、2×10GE SFP+或4×10GE SFP+插卡 • 双电源，可插拔
S5728C-PWR-EI		• 24 个 10/100/1000Base-T • 上行支持 2×10GE XFP、4×1000Base-X SFP 或 2×10GE SFP＋插卡 • 双电源，可插拔，交流供电 • 支持 PoE
S5728C-EI-24S		• 24 个 100/1000Base-X • 4 个 10/100/1000Base-T 千兆 Combo 口，上行支持 2×10GE XFP、4×1000Base-X SFP、2×10GE SFP+或4×10GE SFP+插卡 • 双电源，可插拔
S5752C-EI		• 48 个 10/100/1000Base-T • 上行支持 2×10GE XFP、4×1000Base-X SFP、2×10GE SFP+或4×10GE SFP+插卡 • 双电源，可插拔

设备型号	设备外观	描述
S5752C-PWR-EI		• 48 个 10/100/1000Base-T • 上行支持 2×10GE XFP、4×1000Base-X SFP 或 2×10GE SFP+插卡 • 双电源，可插拔，交流供电 • 支持 PoE
S5728C-HI		• 24 个 10/100/1000Base-T • 上行 4×1000Base-X SFP、2×10GE SFP+、4×10GE SFP+插卡 • 双电源，可插拔
S5728C-HI-24S		• 24 个 100/1000Base-X • 上行 4×1000Base-X SFP、2×10GE SFP+或 4×10GE SFP+插卡 • 双电源，可插拔

7.3.3 产品特点

良好的可扩展性

S5700 支持智能堆叠 iStack 功能，完全即插即用，插好堆叠线缆即可自动组建堆叠虚拟框式架构。

智能堆叠和单一设备相比，在扩展性、可靠性、性能等方面均具有强大的优势。当客户需要扩容或者有单个设备故障需要替换时可实现“新增设备”的热插拔，从而减少了业务中断对客户的影响；相对于框式交换机来说，智能堆叠在性能和端口密度方面突破了硬件架构的限制。而对管理来说，堆叠后的数台设备在逻辑上可认为是一台，减轻了网络管理和配置的工作量。

强大的多业务支持能力

S5700 支持增强型灵活 QinQ 功能，确保灵活的外层 VLAN 标签功能，同时不占用 ACL 资源,充分满足多业务承载的要求。

S5700 支持 IGMP v1/v2/v3 snooping/Filter/Fast Leave/Proxy 等完备的组播协议。同时，S5700 支持线速的跨 VLAN 组播复制功能，捆绑端口的组播负载分担，支持可控组播，可以充分满足 IPTV 等组播业务的需求，确保高质量的视频服务感受。

S5700 支持 MCE 功能，实现了不同 VPN 用户在同一台设备的隔离，有效解决用户数据安全问题，同时降低用户投资成本。

S57HI 系列支持基本的 MPLS 和 VLL 功能，可作为高质量企业专线接入设备，也可助力运营商打造高品质 MPLS 到边缘的网络，是业界为数不多的高性价比盒式 MPLS 交换机。

S5700 有多款设备支持 PoE 功能，遵循 IEEE802.3af 及 802.3at (PoE+)标准。可通过以太网对所连接的标准 PD 设备（如 IP Phone、WLAN AP、Bluetooth AP 等终端）供电，单端口供电能力可高达 30w。有效地简化了终端设备的电源布线和管理成本；同时还可通过配置实现按时、按需进行供电。

完备的高可靠保护机制

S5700 支持双电源冗余供电，也可以交、直流同时输入。用户可灵活选择单电源工作模式或者双电源工作模式，提高了设备可靠性；内置三个风扇提升了稳定性，设备 MTBF 时间业界领先。

S5700 对传统的 STP/RSTP/MSTP 生成树进行了增强，支持 MSTP 多进程，大大提高接入环子环实例数目。还支持 SmartLink 和 RRPP 等增强型以太网技术，可以实现毫秒级链路保护倒换，保证高可靠性的网络质量。此外，针对 SmartLink 和 RRPP 均提供多实例功能，可实现链路负载分担，进一步提高了链路带宽利用率。

S5700 支持增强 Trunk（E-Trunk）功能。此功能应用于 CE 双归接入 VPLS、VLL、PWE3 网络时，CE 与 PE 间的链路保护以及对 PE 设备节点故障的保护。利用 E-trunk 技术可以实现跨设备的链路聚合，从而将链路可靠性提高到了设备级。

S5700 支持智能以太保护 SEP（Smart Ethernet Protection），SEP 是一种专用于以太网链路层的环网协议。适用于半环组网场景，部署时可独立于上层汇聚设备，并提供毫秒级的快速业务倒换性能。保证业务的不中断。在华为设备上已经利用 SEP 协议实现了以太网链路管理。SEP 协议简单可靠、倒换性能高、维护方便、拓扑灵活，可以大大方便用户进行网络的管理和规划。

S5700 支持 VRRP 虚拟路由冗余协议，构建 VRRP 备份组，保持通讯的连续性和可靠性，有效保障网络稳定。支持在设备上配置多条等价路由的方式实现上行路由的冗余备份，当主上行路由发生故障时自动切换到下一个备份路由上去，实现上行路由的多级备份。

多样的安全机制和 QoS 策略

S5700 提供多种安全保护功能。支持 DOS（Denial of Service）类防攻击、网络的防攻击、用户的防攻击等功能。其中 DOS 类防攻击主要包括 SYN Flood、Land、Smurf、ICMP Flood。网络的防攻击主要是指 STP 的 bpdu/root 攻击。用户的防攻击涉及 DHCP 仿冒攻击、中间人攻击、IP/MAC Spoofing 攻击、DHCP request flood、改变 CHADDR 值的 DoS 攻击等等。

S5700 支持通过建立和维护 DHCP Snooping 绑定表，侦听接入用户的 MAC/IP 地址、租用期、VLAN-ID、接口等信息，解决 DHCP 用户的 IP 和端口跟踪定位问题；对不符合绑定表项的非法报文（ARP 欺骗报文、擅自修改 IP 地址等）直接丢弃，有效防止黑客或攻击者通过 ARP 报文实施园区网常见的“中间人”攻击。利用 DHCP Snooping 的信任端口特性还可以保证 DHCP Server 的合法性。

S5700 支持 ARP 表项严格学习功能，可以防止因 ARP 欺骗攻击将交换机 ARP 表项占满，导致正常用户无法上网。同时，支持 IP Source Check 特性，防止包括 MAC 欺骗、IP 欺骗、MAC/IP 欺骗在内的非法地址仿冒带来的 DoS 攻击；S5700 支持 URPF 功能，保证端口接收到数据包时，会反向查找路径验证数据包的真实性，从而有效地杜绝了网络中日益泛滥的源地址欺骗。

S5700 支持集中式 MAC 地址认证和 802.1X 认证，支持用户账号、IP、MAC、VLAN、端口、客户端是否安装病毒防范等用户标识元素的，同时实现用户策略（VLAN、QoS、ACL）的动态下发。

S5700 支持基于端口的源 MAC 地址学习限制功能，有效防止用户源 MAC 欺骗冲击设备 MAC 表项，导致正常用户无法学到 MAC 表而泛洪的问题等。

S5700 系列交换机支持多种 QoS 控制手段，可以基于五元组、IP 优先级、TOS、DSCP、IP 协议类型、ICMP 类型、TCP 源端口、VLAN、以太网帧协议类型、CoS 等信息，实现复杂流流分类功能，支持双向 ACL。S5700 支持基于流的双速三色限速功能，每端口支持 8 个优先级队列，支持 WRR、DRR、SP、WRR+SP、DRR+SP 多种队列调度算法，有效地保证了话音、视频和数据等网络业务不同的质量要求。

易部署简单维护

S5700 采用“一次进站”方案，支持自动配置、即插即用、USB 开局、自动批量远程升级功能，便于部署升级和业务发放，简化后续的管理和维护性能。从而大大降低了维护成本。S5700 支持 SNMP V1/V2/V3，CLI 命令行、Web 网管、TELNET、HGMP 集群管理等多样化的管理和维护方式，设备管理更加灵活。支持 NTP、SSHv2.0、TACACS+、RMON、多日志主机、基于端口的流量统计，支持 NQA 网络质量分析，有利于进一步作好网络规划和改造。

S5700 支持 GVRP。GVRP 是一种 VLAN 的动态配置技术，在复杂的组网环境中应用 GVRP，能够简化 VLAN 配置管理，减少因为配置不一致而导致的网络互通问题。从而达到减少网络管理员的手工配置量及保证 VLAN 配置正确的目的。

S5700 支持 MUX VLAN 功能。MUX VLAN 提供了一种在 VLAN 的端口间进行二层流量隔离的机制。采用两层 VLAN 隔离技术，只有上层 VLAN 全局可见，下层 VLAN 相互隔离。MUX VLAN 通常用于企业内部网，客户端口可以同服务器端口通讯，但客户端口之间不能通讯。用来防止连接到某些接口或接口组的网络设备之间的相互通信，但却允许与默认网关进行通信。从而对企业内部的资源共享及安全通信起到了至关重要的作用。

S5700 支持 BFD 链路快速检测功能，能为 OSPF、ISIS、VRRP、PIM 等协议提供毫秒级检测机制，提高了网络可靠性。S5700 遵循 IEEE 802.3ah 和 802.1ag 提供点到点以太网故障管理功能，可以用于检测用户链路上的故障。以太网 OAM 能够有效提高以太网的网络管理维护能力，保障网络的稳定运行。

S57HI 和 S5706 还提供硬件级 3.3ms 高精度以太 OAM 功能和 Y.1731 性能检测，实现快速故障检测与定位，OAM 功能与其他倒换技术联动可有效保证毫秒级网络保护。

丰富的 IPv6 特性

S5700 提供双协议栈，可平滑升级。硬件支持 IPv4/IPv6 双栈和 IPv6 over IPv4 隧道（包括手工 Tunnel、6to4 Tunnel、ISATAP Tunnel），三层线速转发。既可以用于纯 IPv4 或 IPv6 网络，也可以用于 IPv4 到 IPv6 共存的网络，组网方式灵活，充分满足当前网络从 IPv4 向 IPv6 过渡的需求。

S5700 系列交换机支持丰富的 IPv6 路由协议，包括 RIPng、OSPFv3。支持 ND，管理邻居节点的交互。支持 PMTU 发现（Path MTU Discovery）机制，可以找到从源端到目的端的路径上一个合适的 MTU 值，以便有效地利用网络资源并得到最佳的吞吐量。

7.3.4 产品规格

表7-5 S5700 系列产品主要指标/规格

项目	S5706TP-LI 系列	S5700-SI 系列	S5700-EI 系列	S5700HI 系列
扩展插槽	S5706 无扩展插槽 S57TP 系列提供一个堆叠扩展插槽 S57C 系列提供两个扩展插槽，分别支持上行插卡和堆叠卡 S57HI 系列提供一个扩展插槽，支持上行插卡			
转发性能（PPS）	S5706：9M S5724TP-SI/S5724TP-PWR-SI：36M S5748TP-SI/S5748TP-PWR-SI：72M S5728C-SI/S5728C-PWR-SI/S5728C-EI/S5728C-PWR-EI/ S5728C-EI-24S/S57HI：96M S5752C-SI/S5752C-PWR-SI/ S5752C-EI/S5752C-PWR-EI：132M			
端口交换容量（bps）	S5706：12G S5724TP-SI/S5724TP-PWR-SI：48G S5748TP-SI/S5748TP-PWR-SI：96G S5728C-SI/S5728C-PWR-SI/S5728C-EI/S5728C-PWR-EI/ S5728C-EI-24S/S57HI：128G S5752C-SI/S5752C-PWR-SI/ S5752C-EI/S5752C-PWR-EI：176G			
背板交换容量	256G			
MAC 地址表	LI/SI 系列：16K；EI/HI 系列：32K 支持 MAC 地址自动学习和老化 支持静态、动态、黑洞 MAC 表项 支持源 MAC 地址过滤			
VLAN 特性	支持 4K 个 VLAN 支持 Guest VLAN、Voice VLAN 支持基于 MAC/协议/IP 子网/策略/端口的 VLAN 支持 1:1 和 N:1 VLAN 交换功能 支持基本、灵活 QinQ 功能			
MPLS 特性	不支持	不支持	不支持	支持基本 MPLS 支持 MPLS VLL

项目	S5706TP-LI 系列	S5700-SI 系列	S5700-EI 系列	S5700HI 系列
IPv4 路由	静态路由	静态路由、RIP V1/2、ECMP、支持 URPF	OSPF、IS-IS、BGP 支持 VRRP 支持策略路由 支持路由策略 其他同 SI	同 EI
IPv6 路由	静态路由	支持 RIPng 支持手动隧道 支持 6to4 隧道 支持 ISTAP 隧道	支持 OSPFv3 其他同 SI	同 EI
IPv6 特性	支持 ND（Neighbor Discovery） 支持 PMTU 支持 IPv6 Ping、IPv6 Tracert、IPv6 Telnet 支持 6to4 、ISATAP 、手动配置 Tunnel 支持基于源 IPv6 地址、目的 IPv6 地址、四层端口、协议类型等 ACL 支持 MLD v1/v2 snooping			
组播	支持二层静态组播 MAC 支持 MAC 模式转发	支持 IGMP Snooping 和快速离开机制 支持组播 VLAN 支持 MLD Snooping 支持 IGMP Proxy 支持可控组播 基于端口的组播流量统计	支持 IGMP v1/v2/v3 支持 PIM-SM、PIM-DM 、PIM-SSM 支持 MSDP 其他同 SI	同 EI

项目	S5706TP-LI 系列	S5700-SI 系列	S5700-EI 系列	S5700HI 系列
QoS/ACL	支持对端口接收和发送报文的速率进行限制 支持报文重定向 支持基于端口的流量监管，支持双速三色 CAR 功能 每端口支持 8 个队列 支持 WRR、DRR、SP、WRR＋SP、DRR+SP 队列调度算法 支持 WRED（S5706 和 S57HI 支持） 支持报文的 802.1p 和 DSCP 优先级重新标记 支持 L2（Layer 2）~L4（Layer 4）包过滤功能，提供基于源 MAC 地址、目的 MAC 地址、源 IP 地址、目的 IP 地址、端口、协议、VLAN 的非法帧过滤功能 支持基于队列限速和端口整形功能			
可靠性	支持 STP/RSTP/MSTP 协议 支持 BPDU 保护、根保护和环回保护 支持 RRPP 环型拓扑和 RRPP 多实例 支持 SmartLink 树型拓朴和 SmartLink 多实例，提供主备链路的毫秒级保护 支持智能以太保护协议(SEP) EI/HI 系列支持 BFD for OSPF/ISIS/VRRP/PIM 协议 支持增强 Trunk（E-trunk）			
安全特性	用户分级管理和口令保护 支持防止 DoS、ARP 攻击功能、ICMP 防攻击 支持 IP、MAC、端口、VLAN 的组合绑定 支持端口隔离、端口安全、Sticky MAC 支持黑洞 MAC 地址 支持 MAC 地址学习数目限制 支持 IEEE 802.1X 认证，支持单端口最大用户数限制 支持 AAA 认证，支持 Radius、TACACS+、NAC 等多种方式 支持 SSH V2.0 支持 CPU 保护功能 支持黑名单和白名单			
OAM	硬件实现 EFM OAM CFM OAM Y．1731 性能检测：支持硬件级时延和抖动检测	软件实现	软件实现	硬件实现 EFM OAM CFM OAM Y．1731 性能检测：支持硬件级时延和抖动检测

项目	S5706TP-LI 系列	S5700-SI 系列	S5700-EI 系列	S5700HI 系列
管理和维护	支持智能堆叠（S57HI/S5706 系列除外） 支持 MAC 地址强制转发（MFF） 支持虚拟电缆检测(Virtual Cable Test) 支持以太网 OAM（802.3ah 和 802.1ag） 支持本地端口镜像和远程端口镜像（RSPAN），支持观察端口正常转发报文 支持 Telnet 远程配置、维护 支持 SNMPv1/v2/v3 支持 RMON 支持网管系统、支持 WEB 网管特性 支持集群管理 HGMP 支持系统日志、分级告警 支持断电告警 Dying gasp 功能（仅 S5706 支持） 支持 GVRP 协议 支持 MUX VLAN 功能 支持 HTTPS 支持 802.3az 能效以太网 EEE（仅 S57HI 和 S5706 支持支持）			
环境要求	温度范围：长期工作温度：0 OC～50 OC；短期工作温度：- 5OC～55 OC；相对湿度：10%～90%（无凝露）			
输入电压	AC： 额定电压范围：100-240V AC；50/60Hz 最大电压范围：90-264V AC； 50/60Hz DC： 额定电压范围：-48- -60V DC 最大电压范围：-36- -72V DC 注：PoE 机型无 DC 电源			
外形尺寸 mm（宽×深×高）	S5706：250×180x43.6 S5724TP-SI/S5724TP-PWR-SI/S57HI：442×220×43.6 其他：442×420×43.6			

项目	S5706TP-LI 系列	S5700-SI 系列	S5700-EI 系列	S5700HI 系列
功耗	S5706：<40w	S5724TP-SI：<40w S5724TP-PWR-SI：<455w S5748TP-SI：<64w S5748TP-PWR-SI：<907w S5728C-SI：<56w S5728C-PWR-SI：<891w S5752C-SI：<78w S5752C-PWR-SI：<917w	S5728C-EI：<60w S5728C-PWR-EI：<472w S5728C-EI-24S：<63w S5752C-EI：<88w S5752C-PWR-EI：<930w	S57HI：<93w

7.4 OSN 1800

7.4.1 产品概述

Mini OTN 系统 OptiX OSN 1800，率先结合了 OTN 及 WDM 特性，协助能源、教育、金融、政府、大企业等行业将接入传送统一到一张网络上，解决了接入传送中面临的各种问题。同时，创新性地将 ITU-T G.709 OTN 标准适配范围扩展至 10M～10G。

在降低网络建设、运维成本方面，OSN 1800 另辟蹊径，从比例最大的机房费用入手。创新性的技术 PON over OTN，解决了 FTTx 局点多、租金高、维护工作大的难题，形成“小节点、大局所、集中运营”的理想局面。

OSN 1800 II

OSN 1800 I

7.4.2 产品特点

各种业务统一传送，简洁组网

- 业务统一承载，简化组网。通过“All over WDM/OTN”方式，从低速率业务（如E1）到大带宽业务（如任意协议的10G业务）全部封装到OTN帧格式中进行统一传送，可以广泛传送DSL、FTTx、专线等业务；
- 长距传送，减少网络节点。业界首次在接入WDM上引入G.709 OTN标准。支持标准OTN FEC，可以实现120km（33dB）传送距离，远超过传统CWDM传送80km的限制；
- 业界最强大的业务汇聚能力和集成度，减少设备数量。OSN 1800支持单卡2*GE+2*FE、2*GE+2*STM-1、4*GE、8*GE、4*Any、8*Any、8*EPON、4*GPON等高密度汇聚能力。所有板卡只占用一个槽位，5G线路速率以下的板卡均内置“双发优收”保护。

减少维护投入，降低维护成本

- 网络简洁化，减少节点费用。WDM/OTN的大容量统一传送和长距离传送能力使得网络简洁化成为可能：组网上，业务设备与网管和OSS系统集中配置在中心节点；传送等功能简洁的设备，分散部署在无人的机房，实现广覆盖。这种“头脑集中、四肢简单”的网络结构有助于大幅提升维护效率，减少维护工作量；
- OTN GCC带内开销，可靠管理。支持传统光监控信道（OSC）和G.709 OTN 带内管理（ESC）两种模式，确保在不增加任何网管系统投资的情况下，即可实现所有SDH、WDM/OTN设备统一管理，统一维护。此时网管信息不通过IP网络，严格保证安全性；
- 支持无风扇设计，无须维护。由于消除了风扇本身故障，进一步提高可靠性；

绿色环保，节电降耗

相比核心层设备，接入层设备数量庞大，节能尤为重要。

- 统一组网，整网降耗。OSN 1800的多业务统一传送特性把接入传送网络简化到一张网，避免建设多张传送网络，大大减低整网功耗。
- 网络瘦身，精简节点。创新的“All over WDM/OTN传送拉远”技术实现网络瘦身，业务设备可以集中放置在大局点，通过拉远方式直接连接到用户端，节省中间节点数量30%~90%，减少节点建设费用和管理费用；
- 台灯式低功耗设计。1U设备在单站点2*GE的典型配置下，功耗约25W，还不到一盏台灯的耗电量。

平滑升级，保护投资

- OSN 1800支持“单波白光、18波CWDM、40波DWDM”的升级扩容，以及CWDM、DWDM混合组网，能够满足网络容量逐步发展要求，保护客户的投资；
- 光模块全部可插拔，即插即用，可重用降低了备件投入。

7.4.3 产品特性

7.4.3.1 保护类型

产品提供了完善的网络级保护机制。

表7-6 OptiX OSN 1800 系列支持的业务保护机制及其应用场合（WDM 保护）

保护类型		应用场合
光线路保护	光线路保护	• 保护对象为整个光纤线路。 • 运用 OLP 单板的双发选收功能，在相邻站点间利用分离路由对线路光纤提供保护。
光通道保护	板内 1＋1 保护	• 保护对象为具有双发选收功能的单块 OTU 单板，也可运用 OLP 单板的双发选收功能保护只具有一组波分侧收发光口的单块 OTU 单板。 • 利用分离路由对同一业务采用两个不同波长进行业务保护。
	客户侧 1+1 保护	• 保护对象为具有汇聚功能的 OTU 单板。 • 可通过 SCS 单板实现。 • 可对单个客户侧业务进行保护。
	子架间波长保护	• 通过 OLP 单板实现。 • 保护对象为具有汇聚功能的 OTU 单板，对单个客户侧业务进行 1＋1 保护，工作和保护 OTU 单板可放置在不同机盒上。 • 保护对象为不具有汇聚功能的 OTU 单板，对单个通道进行 1＋1 通道保护，工作和保护 OTU 单板可放置在不同机盒上。
SNCP 保护	SW SNCP 保护	保护对象为具有汇聚、交叉功能的 OTU 单板，可对单个客户侧业务配置交叉及保护。

表7-7 OptiX OSN 1800 系列支持的业务保护机制及其应用场合（SDH 保护）

保护类型	应用场合
子网连接保护 SNCP	运用电层交叉的双发选收功能，通过业务源的多发和业务宿的选收实现对业务的保护。

表7-8 OptiX OSN 1800 系列支持的业务保护机制及其应用场合（以太网保护）

保护类型	应用场合
以太环网保护（ERPS）	以太环网保护基于传统的以太网机制，利用环网自动保护倒换 R-APS 协议，实现以太环网的快速保护倒换。

保护类型	应用场合
链路聚合组（LAG）	指多条连接到同一设备的链路捆绑在一起，便于增加带宽和改善链路的可靠性。
多生成树协议（MSTP）	对于存在环路的以太网用户网络，MSTP 可以基于以太网报文的 VLAN ID 生成相应的树型拓扑，防止广播风暴的产生，同时基于用户报文的 VLAN ID 实现负载分担。

保护倒换概述

OptiX OSN 1800 系列提供保护机制状态监控功能，支持保护倒换状态，保护机制可靠性和资源可用性的检测，保证网络的可靠性。

各类型保护倒换的执行共有五种形式，根据倒换优先级，由高到低依次排列为：清除倒换→锁定倒换→强制倒换→自动倒换→人工倒换。其中自动倒换是由系统内部根据倒换条件自动触发，锁定、强制及人工倒换需要通过网管命令下发，作为系统测试及维护的方法。通过网管可下发清除倒换命令清除以上三种人为下发的外部倒换命令。

7.4.3.2 WDM 传输技术特性

产品提供了 WDM 传输技术方面的多种特性，如 FEC，监控信道和自动激光器关断等。

OTN 技术

产品采用 OTN 传送网络架构。

OptiX OSN 1800 系列全面支持 OTN 技术，关键的技术包括：

- 客户业务映射：对于 G.709 建议明确定义了映射过程的客户业务，OptiX OSN 1800 系列采用了完全符合建议要求的映射处理方法，这些业务包括 SDH 业务和以太网业务。对于速率比 ODU1 容器小的业务，把 ODU1 等分为 16 个时隙来承载，不同速率业务分配不同的时隙数，例如用 1 个时隙承载 STM-1，用 4 个时隙承载 STM-4 业务，用 6 个时隙承载 FC100 业务。采用划分子时隙的方法提高了 ODU 通道的使用效率。
- 通道映射：OptiX OSN 1800 系列支持 Any 业务，OPU1，ODU0，ODU1，OTU1 上下行两个方向的逐层映射。
- OAM：OptiX OSN 1800 系列全面支持 G.709 定义的各种管理开销，主要包括：通过 GCC 字节实现 ESC 管理，可任意选择 GCC0、GCC1 和 GCC2 字节做为管理信息传送通道；支持 SM、PM 性能监视和上报；支持 FEC 和纠错结果上报。

通过采用 OTN 相关技术，使 OptiX OSN 1800 系列具备了以下几大方面的技术优势：

- 通过 OPUk 容器，实现了真正透明的任意客户业务适配和传送，而不更改客户侧业务的任何净荷和开销信息，并提供有效的管理和业务质量监视，也能够比较容易的兼容未来可能出现的各种新业务。
- 采用了异步映射、异步复用机制，不再需要全网络同步，消除了由于同步带来的限制，简化了系统设计。

- 通过 ODU0 通道的映射和复用，使得子速率业务可以在不同 OCh 通道和客户侧端口间进行灵活的调度，同时兼顾了波长带宽高利用率和端到端的灵活调度两方面的要求。
- 通过 OTN 提供的标准 FEC 功能，实现了最大 6.2dB（BER＝10E-15）编码增益，降低了光通道的 OSNR（Optical Signal-to-noise Ratio）容限，延长电中继距离，减少系统站点个数，同时提高光功率预算增益。

链路状态穿通（LPT）功能

通过在波分侧信号帧格式中添加支持 LPT 协议的开销字节，产品可对网络接入点或者服务网络的运行状态进行监视。

正常情况下，上游站点的光波长转换类单板将表示波分侧传输线路正常的 LPT 协议信息传送至下游站点的光波长转换类单板。当上游波分侧传输线路发生状态变化（如故障产生或者故障消除）时，上游站点的光波长转换类单板向下游站点的光波长转换类单板发送网络状态变化的 LPT 报文，告知下游站点传输线路状态出现变化，下游站点便可通过启用（或放弃）备用的传输线路资源，保证传输线路中业务的畅通。

LPT 主要实现两个功能：监视服务网络的运行状态和监视接入点业务的运行状态。

7.4.3.3 光功率管理特性

OptiX OSN 1800 系列产品支持的光功率管理特性为通道增益锁定（AGC）。

AGC 简介

AGC（Automatic Gain Control），即通道增益锁定。无论光纤内传输多少波长，AGC 都可以实现单通道的增益锁定。单个或多个波长发生增波或掉波、光信号波动都不会影响其他通道的信号增益。

WDM 系统的 EDFA 光放大器工作模式为增益锁定。在此模式下，放大器的输出功率随输入功率的变化而变化，增益保持不变。即在波长数发生变化时，增益锁定调整的时间在 1ms 以内，可以确保系统中其他通道的光功率不受影响，从而避免上下波变化时突发误码。

WDM 系统的 EDFA 光放大器工作在增益锁定模式下，光放大器中嵌入了前向和后向反馈控制环路，提供了对于输入功率的快速响应。当输入功率变化在 1dB 以内时，使能后向反馈环路以提供精确的功率控制。反之，使能前向反馈环路，提供对于输入功率的快速响应。

借助于增益锁定功能，系统支持单波传送，也可满足在波数的增加、减少时不影响现有业务。借助于光放大器的内嵌抑制机制，某个跨短的业务突然变化或光放大器劣化，不会对其它跨段的业务造成影响。

极端情况，当系统中仅有 1 波正常，其他掉波时，AGC 功能可以确保该波业务不受影响。

16 波系统，每通道的发送光功率是+5dBm。当其中的 15 波在瞬间全部掉波时，剩下的 1 波业务不受影响。

增益锁定模式比功率锁定模式在功率效率上更有效。因为在功率锁定模式下，泵浦光功率总是按满波情况输出，与实际的工作波长数不相关。

当某些波长发生变化时，AGC 功能可以保证剩余通道的光功率不受影响，有利于避免主通道由于加波或掉波引起的误码突增。

OptiX WDM 产品的所有光放大板都工作在 AGC 模式。AGC 功能自动启动，不需要在 U2000 上配置。

7.4.3.4 物理层时钟

OptiX OSN 1800 支持的物理层时钟包括 SDH 时钟同步、同步以太网时钟。

SDH 时钟同步

SDH 时钟同步为传统时钟同步技术，保证接入 SDH 业务的传送质量。

支持以下方式提取 SDH 时钟：

- 从网元的外时钟口接收的 2M 定时信号
- 从 TSP 单板接收到的光信号中提取的定时信号

支持 1 路 120 欧姆外部时钟源输入和输出。

支持跟踪、保持和自由振荡三种工作模式，线路时钟和 2Mbit/s 时钟可以处理和传递 SSM(Synchronization Status Message)。

SDH 时钟特点：

- 实现简单，可靠性高。
- 使用 SSM（synchronization status information）信息来表示时钟质量等级，通过 SDH 开销来传递 SSM 信息。

同步以太网时钟

同步以太网时钟是一种物理层频率同步技术，类似于 SDH 时钟。

OptiX OSN 1800 的 LEM18 单板支持同步以太网时钟，可从 GE 端口、10GE 端口和 OTU2 端口提取时钟频率及时钟质量等级信息。

同步以太时钟特点：

- 不支持外时钟源。
- 实现简单，端口可直接提取物理层时钟，并且时钟质量满足时钟源要求。
- 使用 SSM（synchronization status information）信息来表示时钟质量等级，通过专用的以太网报文或 OTN 开销来传递 SSM 信息。
- 若需实现同步以太网时钟，要求同步信息所经过的每个网络节点都支持同步以太技术。

7.4.3.5 数据特性

OptiX OSN 1800 支持以太网特性。

业务

OptiX OSN 1800 支持的以太网业务类型如表 7-9 所示。

表7-9 OptiX OSN 1800 支持的以太网业务类型

业务类型	定义	特点
EPL(Ethernet Private Line)业务	以太网专线业务	✓ 点到点透传，物理链路独立，不共享传输带宽 ✓ 为银行、证券等用户提供严格安全、高 QoS 业务
EVPL(Ethernet Virtual Private Line) 业务	以太网虚拟专线业务	✓ 点到多点业务汇聚，共享传输带宽，使用 VLAN ID 等标签隔离 ✓ 为企业等用户提供高 QoS 业务
EPLAN(Ethernet Private LAN)业务	以太网专网业务	✓ 多点到多点通信，不同用户不共享带宽 ✓ 具有严格的带宽保障和用户隔离，为企业用户提供 LAN 互联业务
EVPLAN(Ethernet Virtual Private LAN)业务	以太网虚拟专网业务	✓ 多点到多点通信，多个用户共享带宽 ✓ 使用 VLAN/QinQ 机制来区分不同用户的数据，为企业用户提供 LAN 互联业务

QoS

服务质量 QoS（Quality of Service）：通信网络在各种情况下都能保证可预期的带宽、延迟、延迟抖动、丢包率等方面的服务水平，使应用的请求和响应满足可预知的服务级别。

在传统的 IP 网络中，所有的报文都被无区别的等同对待，每个路由器对所有的报文均采用先入先出 FIFO（First in First out）的策略进行处理，它尽最大的努力（Best-Effort）将报文送到目的地，但对报文传送的可靠性、传送延迟等性能不提供任何保证。

为了支持具有不同服务需求的语音、视频以及数据等业务，要求网络能够区分出不同的通信，进而为之提供相应的服务。

使用优先级队列来支持 QoS 的报文，在报文发送时，将其中一个队列设为严格优先级队列 SP（Strict-Priority Queue），首先保证该优先级队列中的报文得到调度，可以满足关键业务报文得到优先处理。其余队列采用加权轮询 WRR（Weighted Round Robin）调度算法，保证每个队列中的报文都得到一定的服务时间。

以太网业务处理单板提供了 QoS 功能，帮助用户针对不同客户灵活提供分级服务质量业务，包括提供专用带宽、减少报文丢失率、降低报文传送时延及时延抖动等。

保护

OptiX OSN 1800 对以太网业务实现了保护。具体支持的保护参见表 7-8。

管理和维护

OptiX OSN 1800 提供了在设备和网络两个层面上的管理和维护，可通过 U2000 完成，以下仅介绍设备方面的重点管理和维护方案。

a. ETH-OAM

ETH-OAM 完善了以太网二层维护手段，为业务连通性验证、开局业务调测和网络故障定位提供了强大的维护功能。

OptiX OSN 1800 的以太网业务处理单板可实现 ETH-OAM，遵循标准 IEEE 802.1ag 和 ITU-T Y.1731。提供了以太网 OAM 解决方案，能实现故障自动发现和故障定位功能。

IEEE 802.1ag 和 ITU-T Y.1731 的 ETH-OAM 主要的实现方式有：

- LT 链路追踪测试：用于故障点问题定位
- LB 环回测试：用于双向连通性检测
- CC 连通性测试：用于单向连通性检测

b. RMON

RMON（Remote Monitoring）即远程监控，并且该功能可以在不同的网段间的传送网络监视数据。

RMON 自定义了一系列的统计形式和功能，用以在各个符合 RMON 标准的控制站点和检测站点进行数据交换，实现以太网端口的管理。RMON 提供灵活自由的检测模式和控制机制以适用于各种类型网络的需要。同时，RMON 还提供了全网错误诊断，规划和性能事件信息接收等功能。

c. 流量控制

以太网流量控制使用 Pause 帧来控制对端设备的发送速率，该实现基于 IEEE 802.3x 标准。

如 OptiX OSN 1800 设备绑定 50Mbps 带宽，交换机向 OptiX OSN 1800 设备以 100Mbps 速率发送报文，如果不配置流控，报文将因为带宽不足被丢弃。当配置流控后，OptiX OSN 1800 设备检测到带宽不足，向交换机反馈 Pause 帧，交换机收到 pause 帧后，降低发送速率到 50Mbps，可保证报文的有效传送。

7.4.3.6 调测和配置特性

OptiX OSN 1800 系列提供下列特性，简化了调测以及配置的操作。

PRBS 误码检测功能

系统部分 OTU 单板支持 PRBS（Pseudo Random Bit Sequence）误码检测功能。

通过网管可设置启停 OTU 单板客户侧端口的 PRBS 误码测试，从而实现设备开局时不在设备上挂接额外仪表便完成传输链路的误码测试。

该功能通过 PRBS 信号产生器与 PRBS 信号监视器配合实现。支持此功能的 OTU 单板的 PRBS 信号产生器发送 PRBS 信号，PRBS 信号监视器监视发送和从对端站环回的 PRBS 码，将发送信号和环回信号进行比较，由此判断设备或传输线路是否正常。

业务自动配置

业务自动配置方便快捷，避免了复杂的配置过程。

OptiX OSN 1800 系列网元初次上电时，单板默认配置初始业务，硬件安装人员在现场只需要将光功率调测正常即可。设备正常运行后，用户可根据网元使用场景，通过网管远程下发命令，配置此网元单板的业务类型。

此外，为方便用户配置业务，产品采用“业务套餐”向导设计，通过一键式形式选择业务组合，从而完成业务配置。

OptiX OSN 1800 系列的 LQM/LQM2/LWX2 单板可支持一键式业务套餐场景如下：

- GE 透传场景
- GE/STM-1 混传场景

端到端业务配置

系统提供方便用户使用的 OTN 端到端业务配置管理功能，可以简化用户配置过程，缩短网络部署时间，实现网络自动化管理。

OTN 端到端业务配置支持业务路径跨层创建，直接创建 Client 业务路径。用户无需关心中间 OTN 层次的业务调度，不用逐层创建 ODU0、ODU1 业务服务层路径。创建 Client 业务路径后，多层业务路径就会同时自动生成，简化了用户业务配置操作。

OTN 端到端业务配置包括创建、查询、删除、修改端到端业务和优化端到端业务。

7.4.3.7 易安装特性

OptiX OSN 1800 系列提供下列特性，简化了安装的操作。

可插拔模块

产品支持 SFP（Small Form-Factor Pluggable）、XFP（10 Gbit/s Small Form-Factor Pluggable）、SFP+（Small Form-factor Pluggables Plus）和 TXFP（Tunable 10 Gbit/s Small Form-Factor Pluggable）四种可插拔模块。

OSN OptiX 1800 系列部分光波长转换类单板的客户侧和波分侧可采用可插拔模块，若需要调整业务接入类型或需要更换故障模块，则只需直接更换模块，而不需要更换单板。

TXFP 模块通过波长的调节大大降低备件成本，减少库存。

易识别波长信息标签

光波长转换类单板的可插拔光模块采用标签上的波长编号标识，方便现场调测时识别单板上下波长。光波长转换类单板还通过颜色区分，方便标识单板光接收机采用的是 APD 模块还是 PIN 模块。

纤缆接口安全设计

产品所有线缆均提供外形差别明显的端口且使用缆线标签，可清晰辨识安装位置。

OptiX OSN 1800 系列线缆端口结构上考虑了防误插设计，插反插错都无法正常完成安装。

OptiX OSN 1800 系列对外接口电源电路提供了过流保护，如果出现误插情况，即便设备上电也不会损坏设备，造成严重后果。

7.4.3.8 主从子架特性

OptiX OSN 1800 系列支持主从子架。当需要多个机盒组成一个网元时，为了能进行统一的管理，需要采用主从子架级联。在主从子架模式下，多个机盒在网管上显示为一个网元。OptiX OSN 1800 系列最多可支持 1 个主子架管理 6 个从子架。

OptiX OSN 1800 主从子架级联方式可采用环型组网或链型组网的方式，推荐使用环型组网。

7.4.3.9 运行管理特性

产品支持 DCN 通信，激光器自动关断功能（ALS），智能光纤功能。

DCN 通信技术

WDM 设备支持采用 OSC 和 ESC 技术来承载 ECC（Embedded Control Channel），实现 DCN。

ECC 为嵌入控制通道，用于实现网元之间 OAM（Operation, administer and Maintenance）通信功能。ECC 建立在数据通信通路 DCC（Data Communications Channel）上，对于 SDH 设备来说，使用段开销字节中的 D1-D12 来承载，一般使用再生段开销 D1-D3。对于 WDM 设备来说，使用 OSC 或 ESC 监控信道来承载。

华为公司 ECC 通信实现有多种方式：

- HWECC 协议（华为私有协议）
  用于华为设备单独组网情况，或者不需要和其他厂商设备 OAM 互通的情况。
- IP over DCC 协议（标准化协议）
  用于华为和其他厂商设备 OAM 信息互通情况。
- OSI over DCC 协议（标准化协议）
  用于华为和其他厂商设备 OAM 信息互通情况。

华为设备对以上几种协议栈都可以支持，默认支持 HWECC，根据具体组网情况，可以选择 IP over DCC 或 OSI over DCC。

激光器自动关断功能

OptiX OSN 1800 系列的 OTU 单板提供 WDM ALS（Automatic Laser Shutdown）功能。

WDM ALS 功能使 OTU 单板具有根据光信号输入情况自动关闭和开启激光器的功能。

WDM ALS 功能适用于 OTU 单板客户侧和波分侧的发送光口，可以通过网络管理系统设置为“使能”或“禁止”。

ALS 功能实现方式如下：

- 当对端 OTU 单板的某一路客户侧接收光口无光信号输入时，本端 OTU 单板相对应的一路客户侧发送光口将自动关闭激光器，
- 当 OTU 单板的波分侧接收光口无光信号输入时，ALS 功能设置为“使能”的所有客户侧发送光口将自动关闭激光器，

智能光纤

OptiX OSN 1800 系列的 OTU 单板提供智能光纤功能。智能光纤功能使 OTU 单板具有根据上游客户侧或线路侧输入故障在下游客户侧光口自动下插维护码流的功能，使故障信息能够传递到下游客户侧。

智能光纤功能实现方式如下：

- 智能光纤功能“使能”
  - 当对端 OTU 单板的某一路客户侧接收光口无光信号输入时，若本端 OTU 单板相对应的一路客户侧发送光口智能光纤功能设置为“使能”，将下插维护码流。
  - 当 OTU 单板的波分侧接收光口无光信号输入或存在 ODU 层故障或存在 OTU 层故障时，智能光纤功能设置为“使能”的所有客户侧发送光口将下插维护码流。
- 智能光纤功能为“禁止”
  - 当对端 OTU 单板的某一路客户侧接收光口无光信号输入时，若本端 OTU 单板相对应的一路客户侧发送光口智能光纤功能设置为“禁止”，将下插 K28.5 码流。
  - 当 OTU 单板的波分侧接收光口无光信号输入或存在 ODU 层故障或存在 OTU 层故障时，智能光纤功能设置为“禁止”的所有客户侧发送光口将下插 K28.5 码流。

7.4.3.10 升级与维护

产品支持软件包加载，热补丁，备份和恢复网元配置数据。

软件包加载

包加载软件升级是通过一次加载操作完成网元上所有主机和单板软件的加载，替换原来的旧软件。这种加载方式避免了逐个单板加载过程中的重复性操作，提高了升级的效率。

软件包加载包括两种模式：包加载模式和包扩散模式。

- 使用包加载模式下加载软件包功能时，可以一次性加载网元各单板的所有软件，避免了逐块单板加载时的重复操作。
- 使用包扩散模式下加载软件包功能时，分为板级激活和网元级激活，采用板级激活，使用合适的激活分组，可以大大提高加载效率。

为保证升级正常完成，OptiX OSN 1800 系列可在升级之前进行网元健康性检查，包括网元告警检查、网元软件检查等。

软件包加载主要有以下特点：

- 操作时只针对网元，通过统一的操作界面进行加载。
- 自动管理整个网元，新插入单板的软件如果和网元中软件包中的软件版本不一致，则会自动更新，升级效率大大提高。
- 软件包加载是增量加载，只加载需要的文件。
- 软件包加载支持回滚功能，当系统的软件异常等原因导致加载失败时，网元软件通过回滚功能恢复到加载前的状态。

软件包加载主要应用于以下场景：

- 网元软件升级

- 更换设备软件版本

热补丁

产品支持热补丁技术。

对于一些要求长时间不间断工作的设备，当发现软件有缺陷或新需求时，需要在不中断业务的情况下，用新代码来替换正在运行的旧代码，解决这些缺陷或者实现新需求，而这段新代码，就称为热补丁。

热补丁技术主要有以下特点：

- 可以在不影响业务的情况下在线解决大部分软件问题。
- 有效减少发布的软件版本数目，避免频繁的软件版本升级。
- 补丁操作不影响业务，可以远程操作，具有回退功能，可以有效的降低升级成本避免升级风险。
- 可以作为一种有效的定位问题手段，提高解决问题的速度。

备份和恢复网元配置数据

产品网元配置数据可在本地网元 SCC 单板 Flash 中实现网元数据备份和恢复。

为保证已配置网络数据的安全，OptiX OSN 1800 系列的网管软件支持网元配置数据（包括网元的单板配置、时钟配置、保护关系等）远程备份到网管数据库。

设备运行中，若网元主控板丢失数据或设备掉电，恢复上电后用户可通过网管软件远程操作网元，实现网管向网元的配置数据恢复操作。

此外，若网元 15 分钟内反复复位 5 次（系统默认该现象为网元数据库损坏，导致网元异常脱管），该网元即默认进入安全模式。该模式下用户可通过网管软件远程接入设备，采用软件包加载方式更新系统软件，并完成网元数据恢复，避免现场操作，从而降低维护成本。

7.4.4 技术指标

机盒指标

提供机盒的尺寸、重量、功耗、电流、电压等指标。OptiX OSN 1800 I 机盒

表7-10 OptiX OSN 1800 I 直流机盒技术指标

项目	指标
外形尺寸	44mm（高）×442mm（宽）×220mm（深）
重量（空机盒）	4.5kg
最大功耗	150W
额定电流	3A
标准工作电压	-48V～-60V DC

表7-11 OptiX OSN 1800 I 交流机盒技术指标

项目	指标
外形尺寸	44mm（高）×442mm（宽）×220mm（深）
重量（空机盒）	4.5kg
典型功耗	100W
额定电流	1A
标准工作电压	100V～240V AC

- OptiX OSN 1800 II 机盒

表7-12 OptiX OSN 1800 II 直流机盒技术指标

项目	指标
外形尺寸	88mm（高）×442mm（宽）×220mm（深）
重量（空机盒）	7kg
最大功耗	300W
额定电流	6A
标准工作电压	-48V～-60V DC

表7-13 OptiX OSN 1800 II 交流机盒技术指标

项目	指标
外形尺寸	88mm（高）×442mm（宽）×220mm（深）
重量（空机盒）	7kg
典型功耗	200W
额定电流	2.5A
标准工作电压	100V～240V AC

- OptiX OSN 1800 OADM 插框

表7-14 OptiX OSN 1800 OADM 插框技术指标

项目	指标
外形尺寸	44mm（高）×442mm（宽）×220mm（深）
重量（空机盒）	4.5kg
最大功耗	<3.6W
额定电流	0.3A
标准工作电压	12V DC

系统技术指标

本节列出系统参考点 MPI-S 或 S'及 MPI-R 或 R'点的光接口特性和主光通道性能指标。

对于 2.5Gbit/s 速率和 10Gbit/s 速率 16 波系统，支持最大 1×36dB 单跨距离传输。

表7-15 OptiX OSN 1800 系列 DWDM 系统主光通道指标规范（G.652 光纤）（带光放）

项目	单位	性能指标	
线路跨段	-	7×22 dB	6×22 dB
通道数	-	16	16
比特速率	Gbit/s	2.5	10
MPI-S 和 S'点的光接口			
每通路发送光功率	dBm	≥1 dBm	≥1 dBm
最大总发送功率	dBm	17	17

项目	单位	性能指标	
MPI－S 点的最大通路功率差	dB	8	8
光通道（MPI-S-MPI–R）			
光通道代价	dB	≤2	≤2
线路色散容限	-	11200 ps/nm	9600 ps/nm
最大反射系数	dB	-27	-27
MPI－R 和 R’点的光接口			
每通路接收灵敏度	dBm	-30 dBm (2.5 Gbit/s APD)	-22 dBm (10 Gbit/s APD)
		-21 dBm (2.5 Gbit/s PIN)	-16 dBm (10 Gbit/s PIN)
MPI-R 点每通路最小光信噪比	dB	15	20
MPI-R 点的最大通路功率差	dB	10	10

表7-16 OptiX OSN 1800 系列 CWDM 系统主光通道指标规范（G.652 光纤）

项目	单位	性能指标		
线路跨段	-	1x27 dB	1x21 dB	1x16 dB
通道数	-	8	8	2
比特速率	Gbit/s	2.5	5	10
MPI-S 和 S'点的光接口				
每通路发送光功率	dBm	≥2 dBm	≥1 dBm	≥1 dBm
最大总发送功率	dBm	14	14	6
MPI－S 点的最大通路功率差	dB	5	5	5
光通道（MPI-S-MPI–R）				
光通道代价	dB	≤2	≤2	≤2
线路色散容限	-	2000 ps/nm	1400 ps/nm	1200 ps/nm
最大反射系数	dB	-27	-23	-27

项目	单位	性能指标		
MPI－R 和 R’点的光接口				
每通路接收灵敏度	dBm	-30 dBm (2.5 Gbit/s APD)	-28 dBm (5 Gbit/s APD)	-24 dBm (10 Gbit/s APD)
		-21 dBm (2.5 Gbit/s PIN)		
MPI-R 点的最大通路功率差	dB	5	5	5