

2013年11月15日星期五

HUAWEI ENTERPRISE ICT SOLUTIONS **A BETTER WAY**

云引擎 承未来

华为 CloudEngine系列交换机产品介绍

enterprise.huawei.com
HUAWEI TECHNOLOGIES CO., LTD.



目录

1

云计算时代数据中心网络面临的挑战

2

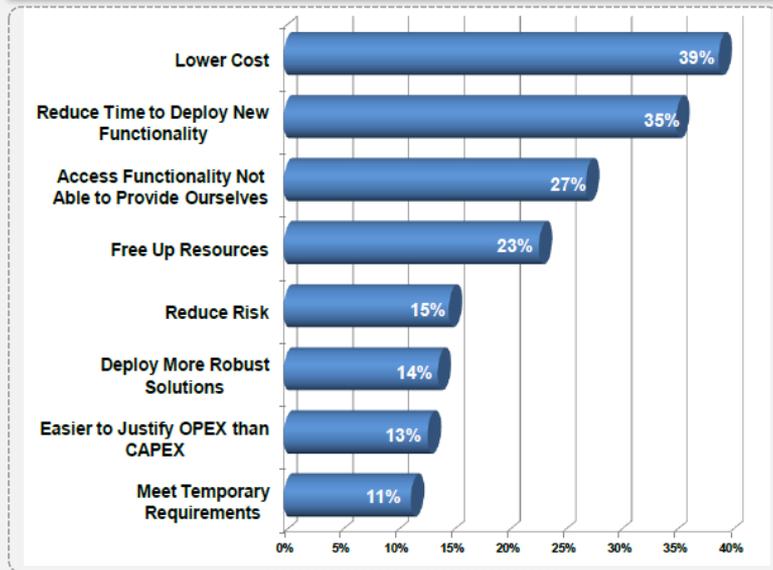
CloudEngine 交换机亮点介绍

3

成功故事

云计算时代数据中心网络面临的挑战

为什么需要云计算？



Data from Webtorials "The 2011 Cloud Networking Report"

数据中心网络面临的挑战

大量高速服务器接入
流量模型的改变

虚拟机大范围的迁移
高效的资源利用

高密网络，出现故障影响范围大
业务多样化，对网络质量要求高

服务器高速接入带来的挑战

业务发展趋势

数据大集中
高速业务

- 排名全球2000名以内的企业均在往**数据中心整合**的方向发展
- 2011年，电子商务及视频行业市场规模增长均超过**60%**
- 到2015年，全球Internet总流量将达到当前的四倍，高达**966EB**

服务器发展趋势

高密度10GE、
40GE服务器接入

Long term, we expect **10G** penetration for DC Ethernet switches to reach **~74% by 2016** ; we forecast **40G** share to grow to 4% by **2016** ; We expect the first **100G** ports to ship by **2013**

-来自Infonetics 2012年对数据中心网络的预测报告

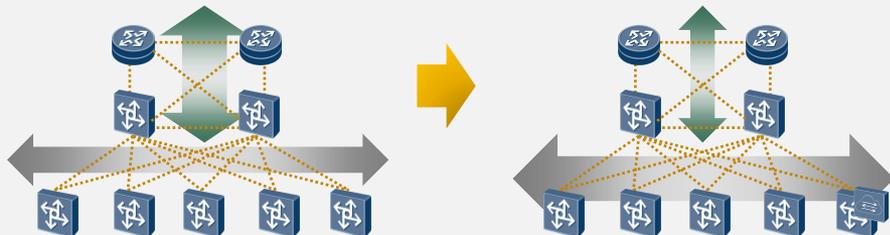
网络发展趋势

10GE、40GE TOR接入
40GE、100GE网络互联

- 服务器**高密度10GE/40GE接入**，网络侧**40GE/100GE互联**是云计算网络未来10年的演进方向
- 网络平台要求支撑服务器GE、10GE、40GE、100GE至少**10年的生命周期演进**

流量模型演进带来的挑战

East-to-West Traffic

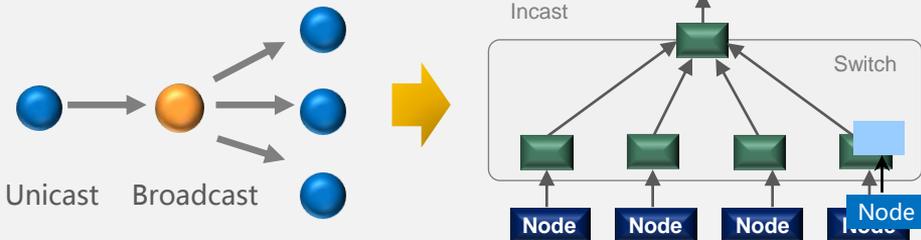


Gartner predicted that east to west traffic will account for 80% of the total traffic in a data center by 2014

服务器间通信，需要无阻塞的网络架构

- 并行计算、3D渲染、搜索等业务，服务器集群需要协同运算，产生大量的**横向交互数据**
- 虚拟机的自由部署和动态迁移，VM间需要**实时同步**大量的数据

Incast Model



Incast模式，需要可动态部署的大容量缓存

- 服务器协同计算，集群中的多个节点同时向主节点发送TCP数据
- 并发的**多对一**的通信产生瞬时突发流量，这种通信模式导致网络拥塞

数据中心扩张带来的挑战

数据中心扩张带来的问题

问题1



1

服务器、网络采购投资飞速增加

问题2

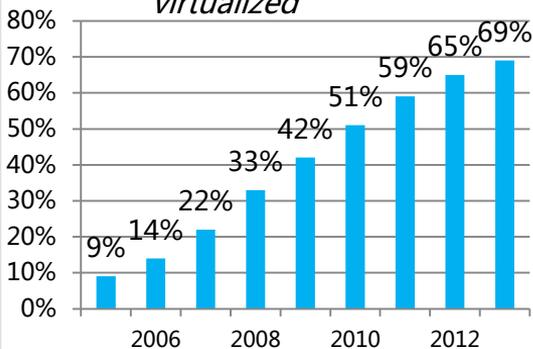


2

现有硬件设备资源利用率低下

物理设备虚拟化成为必然选择

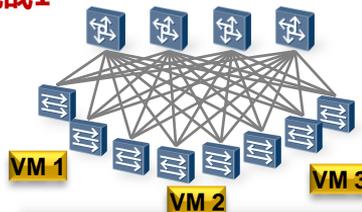
WW installed workloads virtualized



- 到2013年，**69%**的计算都将在虚拟机上进行
- 服务器租用正从物理服务器向**虚拟机租用**，甚至是应用程序租用转变

网络面临的挑战

挑战1



1

大二层
虚拟机任意部署，自由迁移

挑战2



IT管理员

网络管理员

2

快速部署
网络迅速识别和响应虚拟机迁移

云计算时代需要高品质的数据中心网络



终端客户/租户：需要提供**7*24**小时无中断业务



ISP/ICP：网络故障意味着收入损失和客户的流失



金融/政府等行业客户：需要提供低延时，高可靠，高安全的数据中心，保证业务的高效运转



目录

1 云计算时代数据中心网络面临的挑战

2 CloudEngine 交换机亮点介绍

- 弹性云网络
- 虚拟云网络
- 品质云网络

3 成功故事

Cloud Fabric 数据中心网络解决方案全景图

云应用



公有云

私有云

混合云

网络方案



网络设备

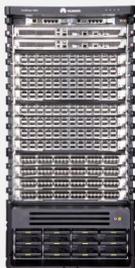


CloudEngine高性能云级交换机

核心交换机



CE12816



CE12812



CE12808



CE12804

10GE TOR交换机



CE6850-48T4Q-EI



CE6850-48S4Q-EI

GE TOR交换机



CE5850-48T4S2Q-EI



CE5850-48T4S2Q-HI(2014Q1)



CE5810-24T4S-EI



CE5810-48T4S-EI

弹性云网络

- 高速线卡:
8*100GE, 24*40GE/96*10GE
- **64T**超大容量, 2T槽位带宽(可扩展到4T)
- **3大类交换网**, 多样化TOR接入, 网络灵活扩展

覆盖10年4代服务器生命周期

虚拟云网络

- 网络虚拟化：**1虚8 VS, 多虚1**
CSS有效降低投资成本
- **TRILL**:为虚拟机迁移构建大二层网络
- **nCenter** 协助批量虚拟机快速迁移

降低网络投资, 云业务敏捷部署

品质云网络

- **五大无阻塞架构**特质
- 整机五大硬件系统全热备
- **DCB**保障存储业务无丢包
- **2us**超低时延, 高价值业务转发无延时

可靠网络, 高价值业务承载

目录

1 云计算时代数据中心网络面临的挑战

2 CloudEngine 交换机亮点介绍

- 弹性云网络
- 虚拟云网络
- 品质云网络

3 成功故事

弹性云网络：10年稳定的数据中心网络平台

10年稳定网络平台，支撑云业务**长期**发展

- 4款核心交换机
- 12种交换网板（A/B/C三类）
- 10倍容量平滑扩展



19年的稳定软件VRP平台



1T线卡能力

2T槽位带宽

64T交换容量

18K+ 10GE服务器接入能力

10年稳定的硬件平台—40GE上行TOR交换机

40GE上行10GE TOR 交换机

•40GE 上行, 10GE电口TOR
48*10GE Base-T, 4*40GE QSFP+



CE6850-48T4Q-EI

•40GE 上行, 10GE光口TOR
48*10GE SFP/SFP+, 4*40 QSFP+



CE6850-48S4Q-EI

GE TOR:40GE端口无阻塞堆叠

•GE TOR: 40GE上行
48*10/100/1000Base-T, 4*10GE
SFP/SPF+, 2*40GE QSFP+



CE5850-48T4S2Q-EI

CE6850: 1 x 40 GE = 4 x 10 GE

- 单个40GE QSFP+可分为4个10GE使用, 提高网络灵活性

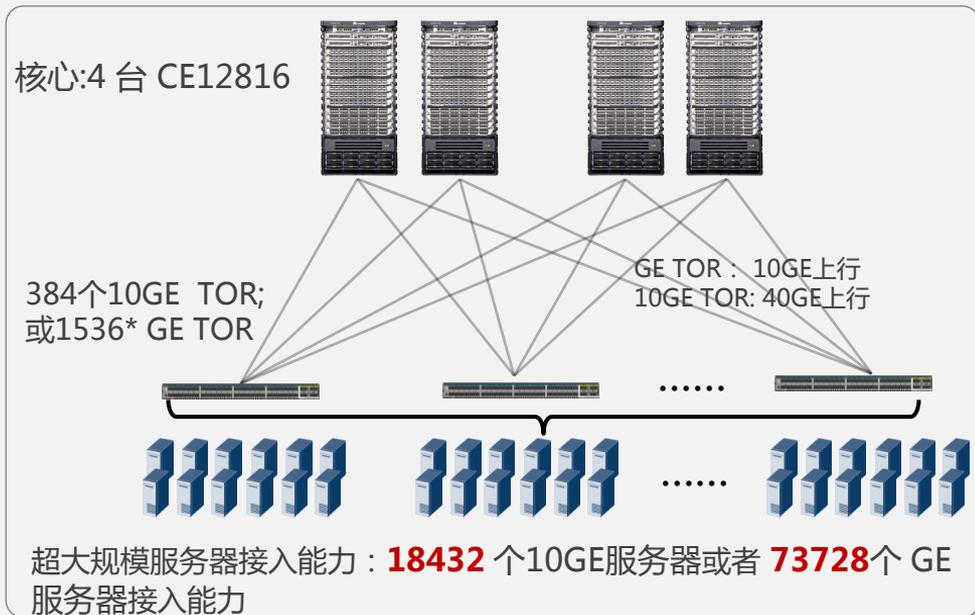


1分4高速线缆



1分4多模光纤

适应未来快速增长的无阻塞网络



40GE 线速互联

- CE12800单板支持24*40GE
- 10GE TOR CE6800支持4*40GE上行
- GE TOR CE5850支持2*40GE上行

超大规格服务器接入能力

- 高密度40G端口接入
- **18432** 个10GE服务器, 或**73728** 个 GE 服务器接入能力

满足云计算场景东西向流量

- 10GE/40GE/100GE高速端口大缓存应对突发流量

久经验证的稳定软件平台

应用于华为**5**大产品族



19年技术积累



服务于全球**100+**个国家和地区，
7,000,000+台设备应用



2000+研发工程师，持续创新；
众多标准组织积极参与者



国际权威机构发布CE12800测试报告



Key findings and conclusions:

- Huawei CloudEngine 12812 proved full line rate throughput at 2 Tbps per line card without loss using traffic pairs on the same line card as well as traversing cards
- CE12800 is a high performance switch supporting up to 1152*10GbE or 288*40GbE port connections, up to 47.52 Tbps of line rate traffic forwarding rate
- Ultra low latency of 2-5 usec and high reliability in high volume traffic tests and regard to high reliability
- Highly reliable Core Data Center Switch with hot swappable components including line cards, power supplies, fans and redundant processing
- CE12800 is highly energy efficient, up to 50% better than the industry average

2 Tbps每槽位带宽

高密10GE/40GE接入

2~5us低时延

高可靠

低于业界50%的功耗

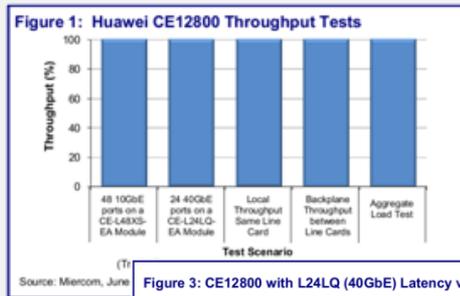


Figure 3: CE12800 with L24LQ (40GbE) Latency vs. Frame Size

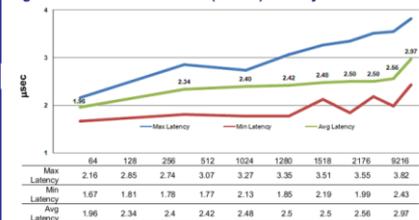
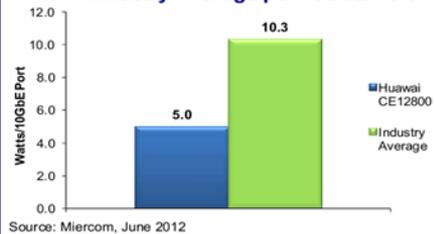


Figure 4: CE12800 Power Efficiency vs Industry Average per 10GbE Port



目录

1 云计算时代数据中心网络面临的挑战

2 CloudEngine 交换机亮点介绍

- 弹性云网络
- 虚拟云网络
- 品质云网络

3 成功故事

虚拟云网络：VS + CSS 降低投资成本

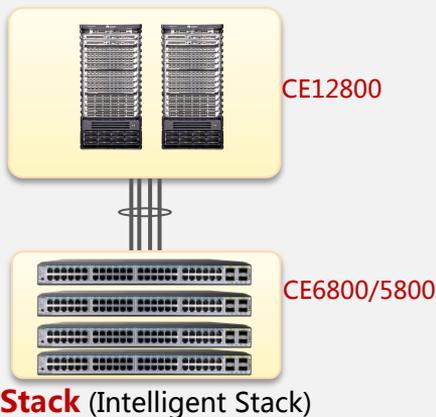
独一无二的 VS+CSS 整网虚拟化

VS (Virtual System)



- VS 支持一台设备虚拟成**8**台虚拟交换机
- CSS多虚一简化网络管理
- 支持 **VS + CSS 组合**应用场景，用户定义更灵活

CSS (Cluster Switch System)



价值

降低CAPEX

- VS: 减少物理设备，提高资源利用率
- CSS: 100%带宽利用率

降低OPEX

- VS: 减少需要维护的物理节点
- CSS: 简化网络管理

安全可靠

- VS: 网络区域隔离；网络故障隔离
- CSS: 无环高可靠的网络

VS协助用户降低网络投资

降低CAPEX

- 1虚8, 交换机物理复用, 节省设备成本
- 按需灵活分配资源, 提高设备利用率



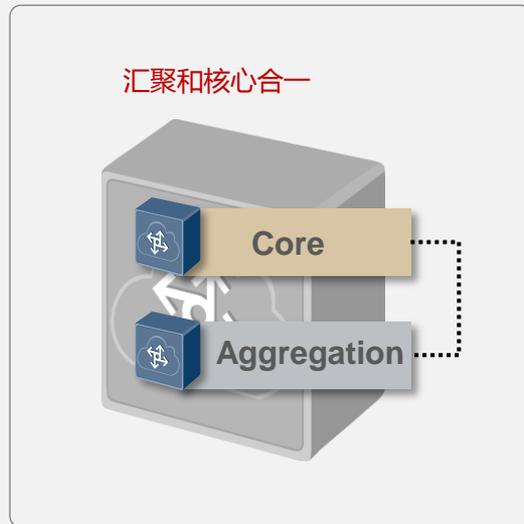
降低OPEX

- 物理设备共用, 降低功耗
- 减少占用的物理空间
- 降低维护的网络物理节点



安全隔离

- 网络划分与隔离, 提高网络安全性
- 各个VS故障隔离



CSS & iStack灵活部署方案



多虚一规模

- CE12800支持多框集群
- 业务口集群，集群带宽可达1.28Tbps
- 接入交换机最多可支持16框堆叠

长距集群

- 集群线缆可采用专用的集群线缆或者普通的光纤
- 采用业务口光纤集群，单跳最远可拉至80Km (10GE SFP+光模块)
- 可应用于不同的场景，灵活部署

本地优先转发

- 集群设备的多条链路负载分担
- 本地选路转发
- 有效提高带宽利用率

虚拟云网络：虚拟机大范围快速迁移

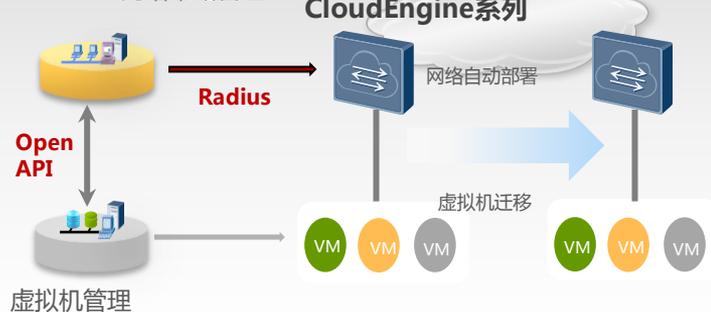
超大二层网络



- 提供高达**512**个节点的大二层网络,虚拟机自由部署
- **10GE/GE TOR**及核心交换机均支持TRILL
- 遵循**IETF标准**协议,支持**4K**租户

网络策略迁移

nCenter: 网络策略管理



- **nCenter**: 实现**自动**网络策略部署
- 高速 Radius接口:策略下发速度快,虚拟机快速上线
- 开放API接口:可**兼容**业界主流的虚拟机管理平台

目录

1 云计算时代数据中心网络面临的挑战

2 CloudEngine 交换机亮点介绍

- 弹性云网络
- 虚拟云网络
- 品质云网络

3 成功故事

品质云网络：网络无阻塞，业务无中断

无间断交换核心

全热备

系统全热备

交换网

主控板

监控板

电源

风扇

全业务不间断倒换

NSR

NSB

NSF

NSP

- 整机五大硬件系统全热备，三大总线全冗余
- 全生命周期可靠性保障

无阻塞交换架构

五特质



正交网板设计，提升系统带宽，可演进

动态Clos架构，负载更均衡

信元交换，提高交换网利用率

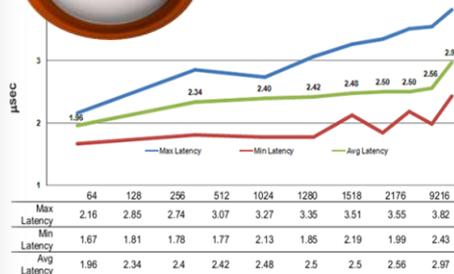
VoQ(Virtual Output Queue) 机制，
交换网有序调度

12GB大缓存，突发流量无阻塞

无损交换网络

零丢包

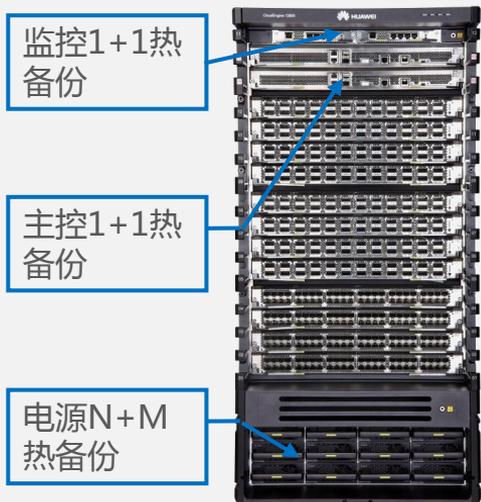
Fig. LQ (40GbE) Latency vs. Frame Size



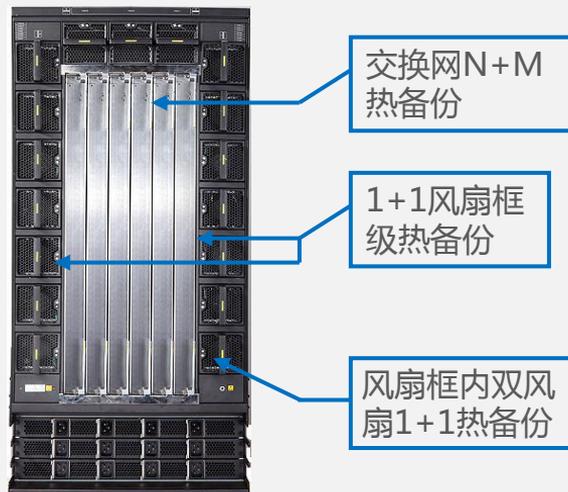
Source: Microm, June 2012

- DCB数据中心桥：存储业务转发无丢包
- 2us超低时延：业务转发效率为业界2倍

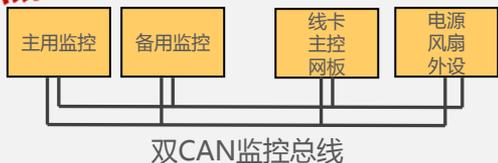
硬件高可靠性: 所有部件冗余热备



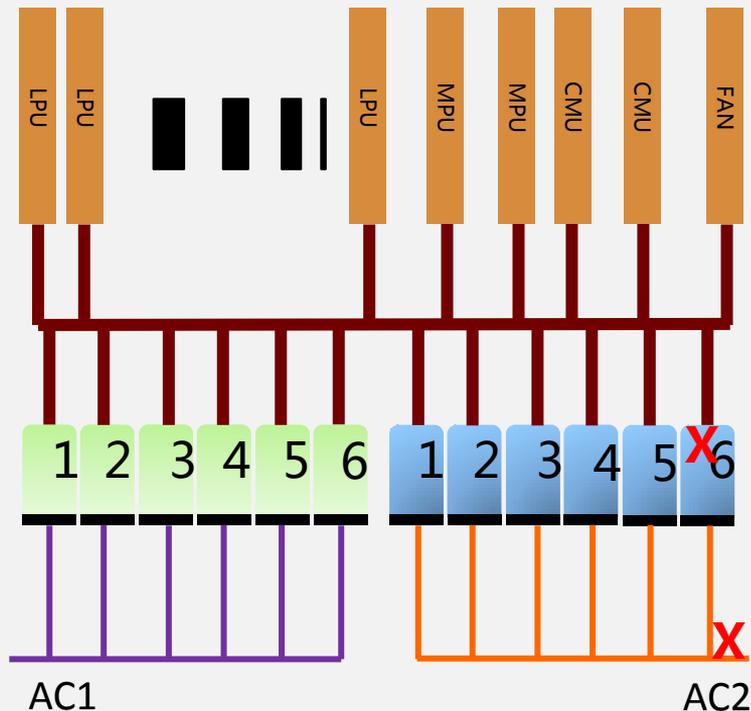
系统级热备份



系统内热备份



高可靠性设计——双路供电



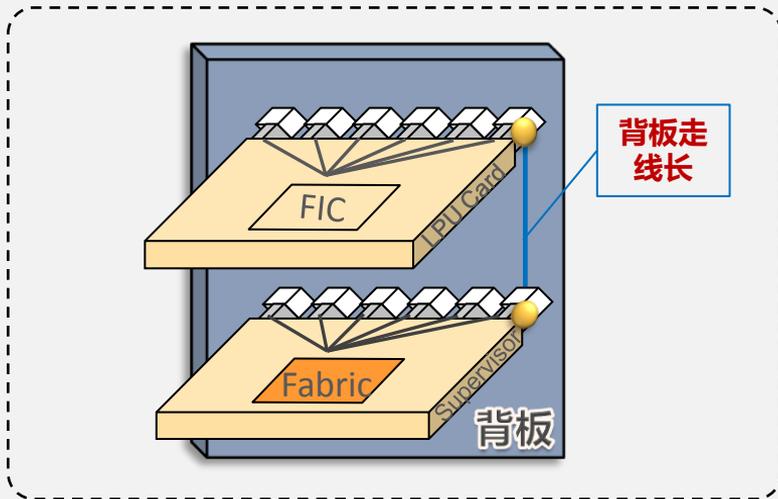
高效可靠的供电架构

- 双路电源输入，N+M备份(CE12812/12808/12804支持的最大电源模块数量分别为12块/8块/4块)
- 动态休眠：可跟据负载休眠部分模块，提高其余电源负载，进而提高电源转换效率

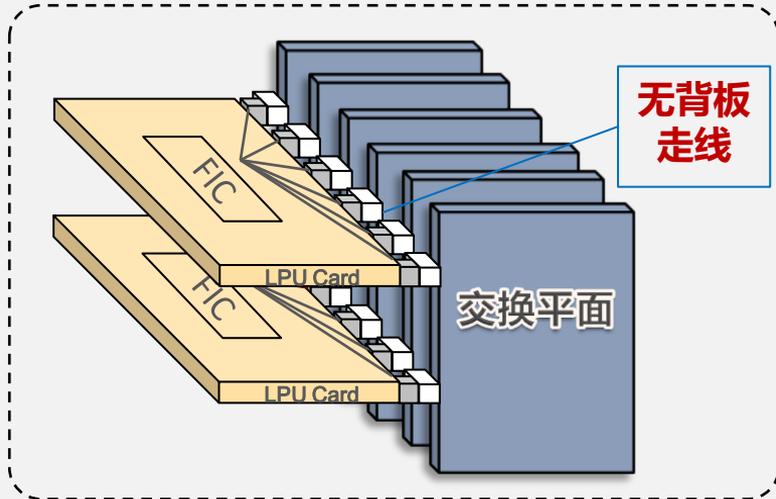
高效率新一代电源

- 电源效率 $\geq 96\%$ ，降低转换损耗
- 宽电压输入范围：最大90~290V AC，-48V/-60V DC 有效缓解电网波动
- 自主均流：避免不同模块压差过大，部分模块损耗大
- 智能管理：可远程管理、可休眠、实时功耗检测，降低功耗

卓越的正交网板架构设计



传统架构



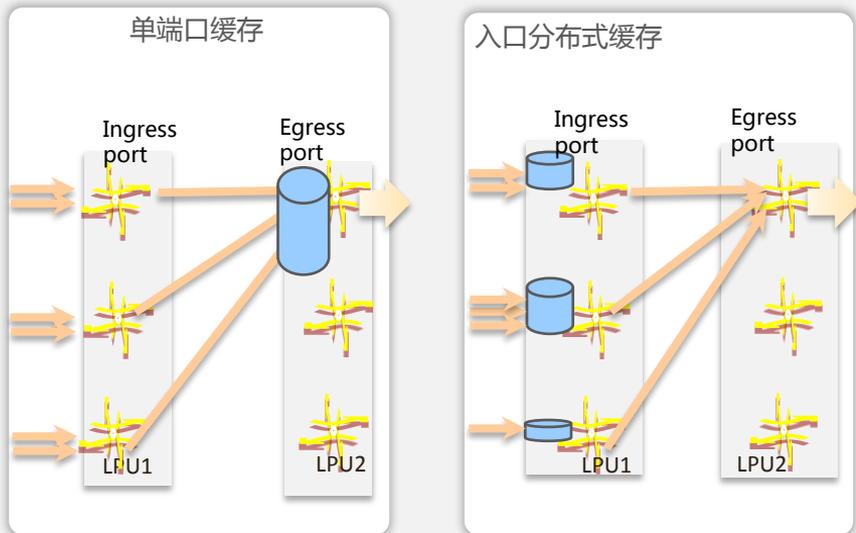
CE12800全正交架构

线卡和交换网全正交架构设计

- 核心交换机中，线卡到交换网的走线是影响槽位带宽的最重要因素之一，背板链路的长度和速率越大，则损耗越大
- CE12800采用正交架构设计，突破技术难点，把背板走线长度降为0，省去背板走线，极大提升系统的带宽和演进能力

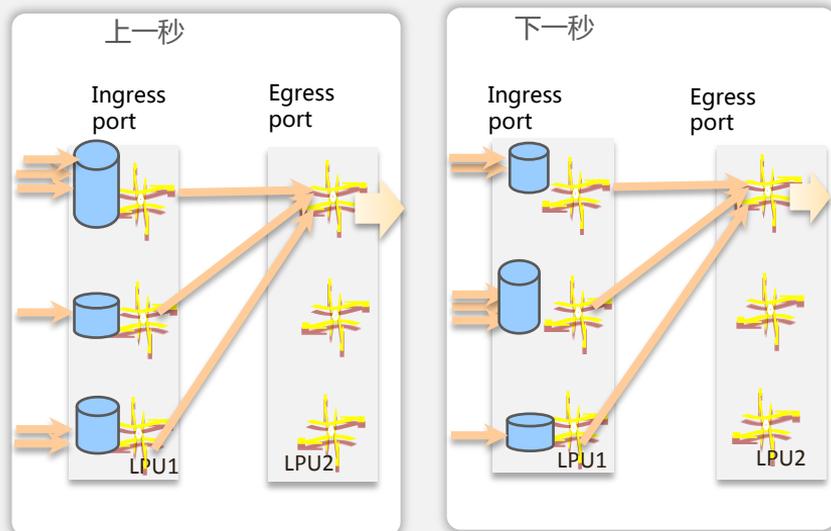
动态大缓存协助构建无阻塞网络

分布式缓存更好的吸收突发流量



入口分布式缓存，有效降低单端口缓存压力

动态缓存提高缓存的利用率



缓存可按需分配，动态调整。

→ 数据流 缓存

目录

1 云计算时代数据中心网络面临的挑战

2 CloudEngine 交换机亮点介绍

3 成功故事

中国农业银行一级数据中心

背景

- 中国农业银行在英国《银行家》杂志全球银行1,000强税前利润排名中，位列第7，品牌价值99.29亿美元，位列第18；
- 农行即将在全行范围内上线新一代的核心业务系统。作为核心业务系统上线的基础，其网络架构需要匹配核心业务系统部署的要求，并且满足未来5年的业务发展；
- 此次农行全网一级分行和骨干网改造，将为境内总行本部、36个一级分行、19,573个基层营业机构、23,461个分支机构提供强有力的服务保障；

方案

- 共计144台CE12800，每个一级中心部署4台CE12800，2台用于生产网核心，2台作为各一级分行核心；
- 每个一级中心与华为高端路由器无缝对接，与全国两大中心高速互联；
- CE12800上通过VRF虚拟化技术，彻底隔离来自不同分支及不同安全域的业务流量。

价值

- 超大容量：每单板2T流量吞吐能力，高密线速万兆接口，满足未来10年发展需求；
- 深度虚拟，安全隔离：匹配农行2013年“控制风险、降低成本、商业可持续”诉求。



华为助力浙江省财政厅新大楼网络建设

背景

- 浙江省财政厅新大楼项目，水平需接入信息点一千二百余个，整体分为两个独立的网络区域，物理分开隔离，其中一个为财政业务专网（内网），一个是互联网（外网）；
- 通过多功能、高效率与高舒适性的办公网络智能化建设，达到和实现便捷、高效、安全、健康的办公环境的需要；

方案

- 华为选用高可靠性网络产品，合理设计网络架构，制订可靠的网络备份策略，保证网络具有故障自愈的能力；
- 核心使用先进的CE12808交换机，采用集群技术，保障网络的稳定性。内网汇聚交换机使用S5700系列千兆三层交换机通过万兆线路连接核心，采用链路聚合，加大网络带宽，保证网络的高效和稳定。

价值

- 提供高可靠冗余核心，无环路的拓扑结构，各种业务稳定运行；
- 无线网络和有线网络的无缝融合，在网络互联、认证管理、安全防御等方面与有线网络进行良好的兼容和互补。



广东广电高清互动项目

背景

- 广东广电集团对广电省干数据网进行升级改造，把省干数据网建设成一张互动电视承载网和一张宽带IP承载网。计划新增22台交换机，部署在11个地市，替换11个地市原有的设备。
- 客户需求：地市级数据网容量需要扩容，以满足各地市双平面建设需要。

方案

- 每个地市级数据网采用CE12808双核心CSS结构组网，消除网络单点故障；
- 互动电视承载网和城域网独立部署的地市分别上行至互动高清业务出口和省干数据网络出口；
- 互动电视承载网和城域网共享的地市，互动高清业务通过省干汇聚交换机上行至互动高清业务出口；
- 原有核心设备路由器作为出口路由器与其他地市互联；

价值

- CE12808核心交换机双核心提供了广电网络的高带宽和超高可靠性，2us的超低延时及分布式缓存，提供VOD业务良好的业务体验。





中科院国家蛋白质科学中心园区网络

背景

- 中科院国家蛋白质科学中心位于上海市张江高科技园区，新建的园区网络覆盖5栋楼宇，约1000个信息点；
- 网络建设要求：光纤覆盖学院所有楼宇，园区网络万兆到楼层，千兆到桌面；无线全覆盖，为移动终端提供便利的网络接入；网络出口高安全，阻止各类攻击；

方案

- 核心采用两台CE12808，CE6850作为10GE汇聚交换机，接入采用S5700
- 防火墙串接在核心交换机和出口交换机之间；VPN旁挂在CE12800之下；
- 无线控制器旁挂在CE6850之下，S5700为室内和室外AP供电；

价值

- CE12800满足了园区网络中大带宽业务应用需求；
- CE6850高密万兆，1U设备即可满足千兆接入的汇聚。





HUAWEI ENTERPRISE ICT SOLUTIONS **A BETTER WAY**

Copyright©2012 Huawei Technologies Co., Ltd. All Rights Reserved.

The information in this document may contain predictive statements including, without limitation, statements regarding the future financial and operating results, future product portfolio, new technology, etc. There are a number of factors that could cause actual results and developments to differ materially from those expressed or implied in the predictive statements. Therefore, such information is provided for reference purpose only and constitutes neither an offer nor an acceptance. Huawei may change the information at any time without notice.

附录：CloudEngine系列产品介绍

CloudEngine系列是华为公司面向下一代数据中心和高端园区推出的

“云”级高性能交换机，支持丰富的数据中心和园区业务特性。

旗舰级核心交换机



CE12816

CE12812

CE12808

CE12804

高性能TOR交换机



CE6850-48T4Q-EI

CE6850-48S4Q-EI

CE5850-48T4S2Q-HI(2014Q1)

CE5850-48T4S2Q-EI

CE5810-24T4S-EI

CE5810-48T4S-EI

- 数据中心核心交换机——CloudEngine 12800系列
- 数据中心TOR交换机——CloudEngine 6800/5800系列
- 成熟稳定的软件平台——VRP8 软件系统平台

CE12800系列产品规格

规格项	CE12804	CE12808	CE12812	CE12816
交换容量	16T	32T	48T	64T
线卡槽位	4	8	12	16
槽位带宽	2T（单向），未来可演进到4T（单向）			
板卡类型	24*40GE/96*10GE、48/24*10GE、48*GE，8*100GE均为全线速板卡			
高度	11U	17U	22U	36U
系统电源	交流：110/220VAC电源模块，2700W功率； 直流：-48V/-60VDC电源模块，2200W功率；			
整机端口数	192*10GE 96*40GE/384*10GE 32*100GE	384*10GE 192*40GE/768*10E 64*100GE	576*10GE 288*40GE/1152*10GE 96*100GE	768*10GE/GE 384*40GE/1536*10GE 128*100GE
缓存容量	平均100ms/port（10GE/40GE/100GE均具有超大缓存）			

CE12800面向下一代数据中心和高端园区核心，具备业界最高的交换容量和超高密度线速板卡，满足未来网络长期发展需求。

线卡：高密40GE/10GE/GE



24*40GE光口板



48*10GE光口板



48*GE 电口板



24*10GE光口板



48*GE 光口板

名称/特性	24*40GE	24*10GE	48*10GE	48*GE 光口板	48*GE 电口板
接口密度	24端口40GE/96端口10GE	24端口GE/10GE	48端口GE/10GE	48端口GE SFP	48*10/100/1000Base-T
接口类型	QSFP+	SFP/SFP+	SFP/SFP+	SFP	RJ45

单板特点：功耗小、缓存大、时延低

Flex Port: 适应不同场景混合组网

10GE Flex Port



SFP-T:
RJ45, GE电口



距离: 100m

SFP: GE光
口互联



距离: 500m,
10Km, 40km,
80km, 100km

10GBase-CR:
铜缆, 10m内
低成本互联



距离: 1m, 3m,
5m, 7m, 10m

SFP+:
10GE光口互
联



距离: 220m,
300m, 10km,
40km, 80km

40GE Flex Port



QSFP+: 150m,
10Km互联



距离: 150m,
10km

40GBase-CR4:
5m内互联



距离: 1m, 3m,
5m

铜缆、光模块1分4:与
10GE接口对接

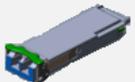
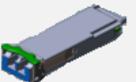
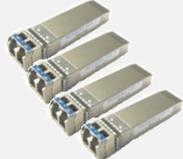


距离: 150m,
10km



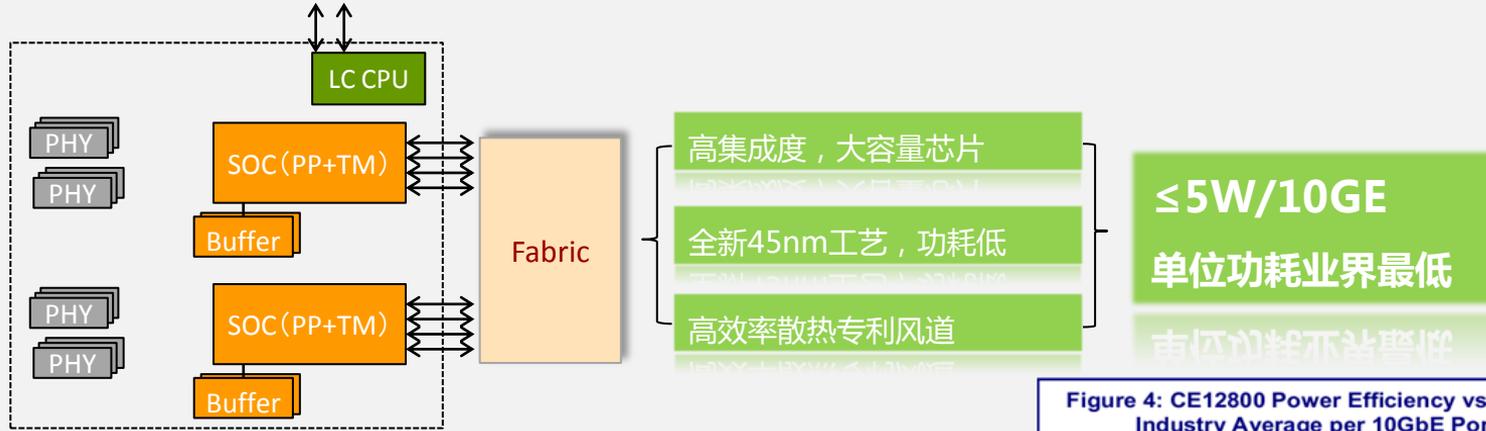
距离: 1m,
3m, 5m

Flex Port: 40GE灵活互联方案

光纤互联				
端口类型	光模块	光纤	光模块	端口类型
40GE 端口	40GBase-SR4(MPO) 	多模光纤: OM3:100米; OM4:150m 	40GBase-SR4(MPO) 	40GE 端口
40GE 端口	40GBase-LR4(LC) 	单模光纤: 10Km 	40GBase-LR4(LC) 	40GE 端口
40GE 端口	40GBase-SR4(MPO) 	MPO-4*DLC 多模光纤: OM3:100米 	4*10GBase-SR 	4个 10GE 端口

高速线缆互联		
端口类型	高速线缆	端口类型
40GE 端口	1米/3米/5米40GBase-CR4 高速线缆 	40GE 端口
40GE 端口	1米/3米/5米1分四高速线缆 	4个 10GE 端口

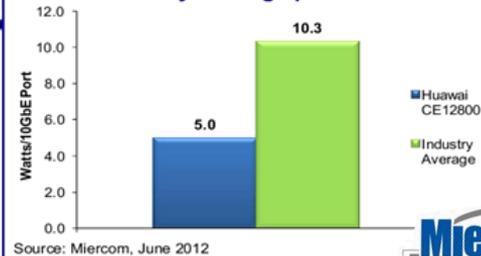
Green Port : 单位10GE功耗最低



Key findings and conclusions:

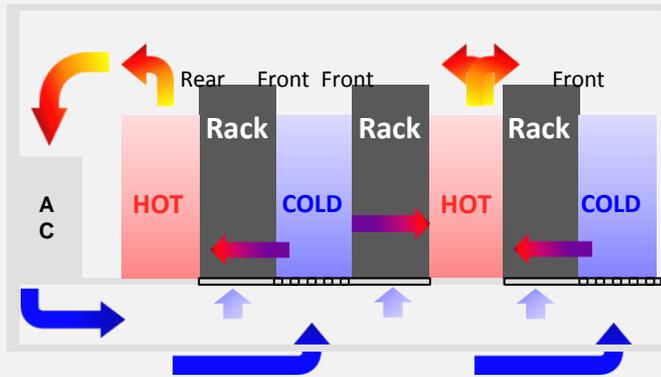
- Highly reliable Core Data Center Switch with hot swappable components including line cards, power supplies, cooling fans and redundant processing modules
- CE12800 is highly energy efficient with power consumption 50% better than the industry average for core switches

Figure 4: CE12800 Power Efficiency vs Industry Average per 10GbE Port



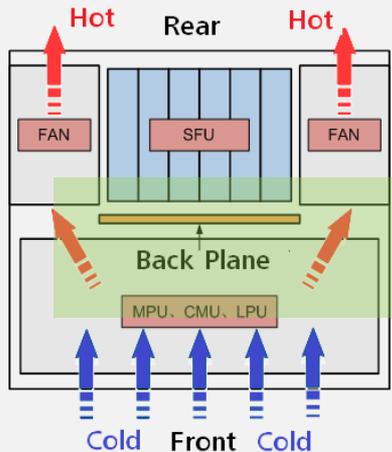
专利散热风道:严格的前后风道

数据中心机房风道图



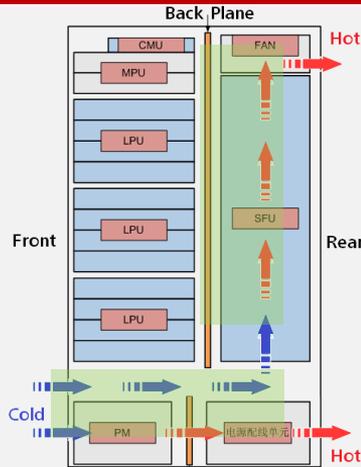
典型数据中心冷热风道

散热风道1：MPU/CMU/LPU



CE12800俯视图

散热风道2：SFU



CE12800侧面图

- 严格的前后风道，利于机柜布局和机房规划，降低机房空调能耗
- 线卡风道无迂回，提升散热效率
- 风扇智能调速，分区散热;对旋风扇设计，强劲高效散热

高性能万兆TOR 交换机



CE6850-48T4Q-EI



CE6850-48S4Q-EI

	CE6850-48S4Q-EI	CE6850-48T4Q-EI
下行端口	48*10G SFP+	48*10GBase-T
上行端口	4*40G QSFP+	4*40G QSFP+
散热风道	前后 / 后前 风道	
堆叠	支持16台堆叠	
设备尺寸(宽*深*高mm)	442*600*43.6	

高性能，高密度

- 最大容量：1.28Tbps；
- 高密接入：高达64个10GE端口

40GE上行

- 40GE上行，搭建无阻塞40GE全互连网络，1+1链路备份；
- 无收敛堆叠，不占用10GE端口

丰富的DC特性

- 严格前后风道
- 512个节点的TRILL大二层
- 协同nCenter，实现虚拟机批量迁移

高性能千兆TOR 交换机



CE5850-48T4S2Q-EI



CE5850-48T4S2Q-HI (2014Q1)



CE5810-48T4S-EI



CE5810-24T4S-EI

	CE5850-48T4S2Q-EI/HI	CE5810-48T4S-EI	CE5810-24T4S-EI
下行端口	48*10/100/1000BASE-T	48*10/100/1000BASE-T	24*10/100/1000BASE-T
上行端口	4*10G SFP+, 2*40G QSFP+	4*10G SFP+	4*10G SFP+
散热风道	前后 / 后前 风道	前后 / 后前 风道	前后 / 后前 风道
尺寸(宽*深*高mm)	442*420*43.6	442*420*43.6	442*420*43.6

CE5850 : 40GE上行

- 业界首款40GE上行的GE接入交换机
- 40GE上行优势
 - 搭建无阻塞40GE全互连网络, 1+1链路备份;
 - 可采用40GE端口无收敛堆叠

高可靠

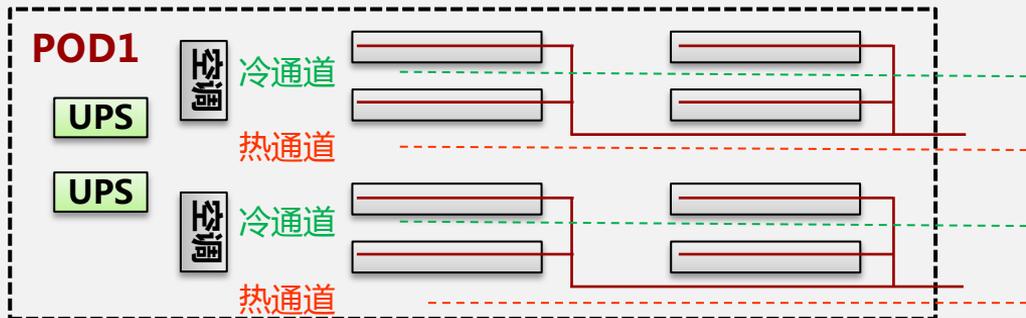
- 双电源热插拔, 1+1冗余
- 双风扇热插拔, 1+1冗余, CE5850每个风扇框含两个风扇

丰富的DC特性

- 严格前后风道
- CE5850支持512个节点大二层网络TRILL
- 协同nCenter, 实现虚拟机批量迁移

前后/后前风道，满足DC冷热分离风道需求

- 通过选配风扇模块，灵活支持前后/后前风道
 - 风扇和电源模块必须配套；



CE5800/CE6800设备采用**严格的前后风道设计**，无需机架改造就支持冷热风道隔离，提升散热的效率，降低散热功耗



HUAWEI ENTERPRISE ICT SOLUTIONS **A BETTER WAY**

Copyright©2012 Huawei Technologies Co., Ltd. All Rights Reserved.

The information in this document may contain predictive statements including, without limitation, statements regarding the future financial and operating results, future product portfolio, new technology, etc. There are a number of factors that could cause actual results and developments to differ materially from those expressed or implied in the predictive statements. Therefore, such information is provided for reference purpose only and constitutes neither an offer nor an acceptance. Huawei may change the information at any time without notice.