

DDR/FDR IB网络最强Linpack效率 92% (400节点)



南京大学
HS22 + Infiniband
效率： 92%

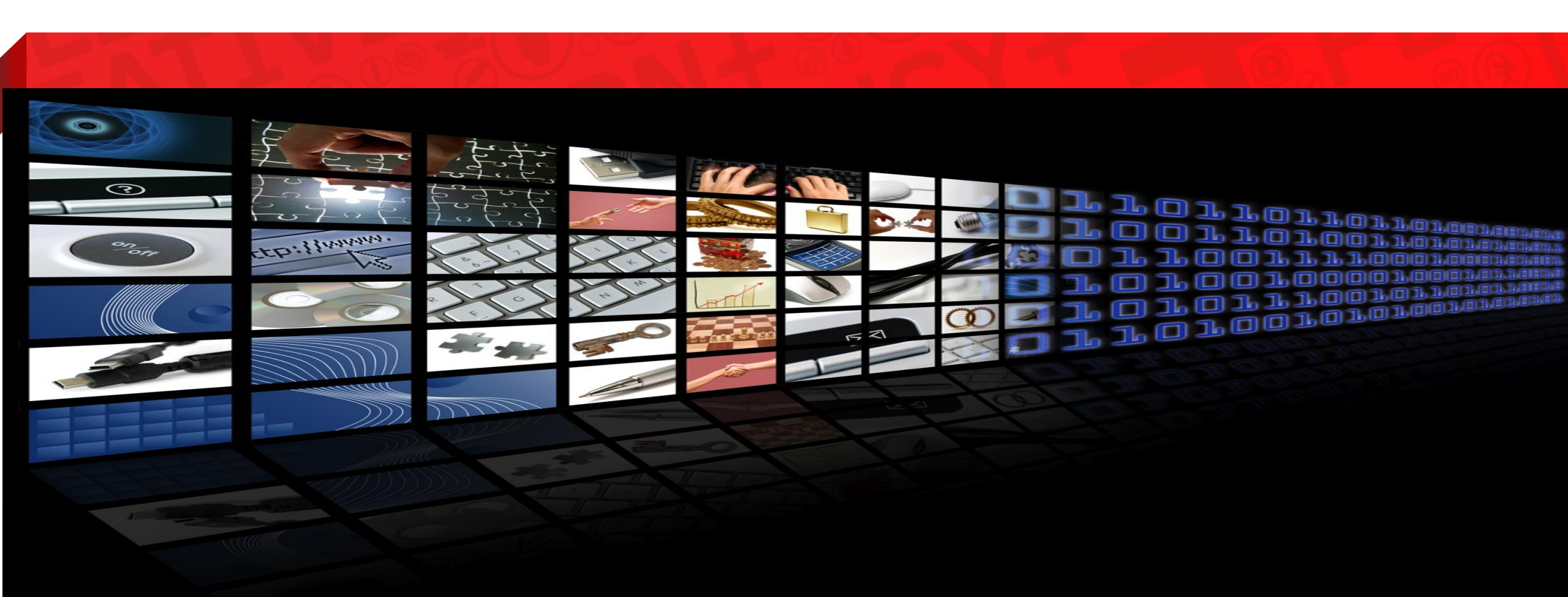
- 2016/4/15日完成了高校第一的北京大学高性能计算中心的验收
- 624个计算节点，447.6万亿次
- 4PB容量，70GB/s带宽
- 56Gb/s IB 网络
- Linpack效率81% 创Intel E5-2600V3 记录

排名	厂商	型号	安装地点	年份	应用领域	处理器核	Linpack (Gflops)	来源	Peak (Gflops)	效率
1	国防科大	天河一号/3072x2 Intel Quad Core Xeon E5540 2.53GHz/E5450 3.0GHz+ 2560 ATI Radeon 4870X2@575MHz/Infiniband	国家超算天津中心	2009	科学计算/工业	24576	563100.00	C	1206210.00	0.467
2	曙光	曙光5000A/1920x4 AMD Quad Core Barcelona 1.9GHz/DDR Infiniband/WCCS+Linux	上海超级计算中心	2008	科学计算/工业	30720	180600.00	C	233472.00	0.774
3	联想	深腾7000/1240x2 Intel Xeon Quad-Core E5450 3.0GHz/140x4 Intel Xeon Quad-Core X7350 2.93GHz Infiniband 4xDDR	中国科学院超级计算中心	2008	科学计算	12160	106500.00	C	145293.00	0.733
4	IBM	BladeCenter HS22 Cluster/Intel Xeon QC GT 2.53 GHz/Giga-E	网络公司	2009	工业/游戏	7168	38790.00	T	72540.00	0.535
5	IBM	BladeCenter HS22 Cluster/Intel Xeon QC GT 2.53 GHz/Giga-E	网络公司	2009	工业/游戏	7168	38790.00	T	72540.00	0.535
6	IBM	BladeCenter HS22 Cluster/Intel Xeon QC GT 2.53 GHz/Giga-E	网络公司	2009	工业/游戏	7168	38790.00	T	72540.00	0.535
7	IBM	IBM BladeCenter HS22 Cluster/400x2 Intel Xeon 5550 2.66GHz/Infiniband	南京大学	2009	科学计算/教育	3200	31310.00	C	34048.00	0.920
8	曙光	曙光5000/420x2 Intel Quad Core 3.0GHz/DDR Infiniband	计算物理国家重点实验室	2009	科学计算	3360	31048.80	C	40320.00	0.770
9	IBM	BladeCenter HS22 Cluster/Intel Xeon QC GT 2.53 GHz/Giga-E	网络公司	2009	工业/游戏	5376	31030.00	T	54410.00	0.570
10	IBM	BladeCenter HS22 Cluster/Intel Xeon QC GT 2.53 GHz/Giga-E	网络公司	2009	工业/游戏	5376	31030.00	T	54410.00	0.570

联想具有最丰富最完善的高性能计算产品

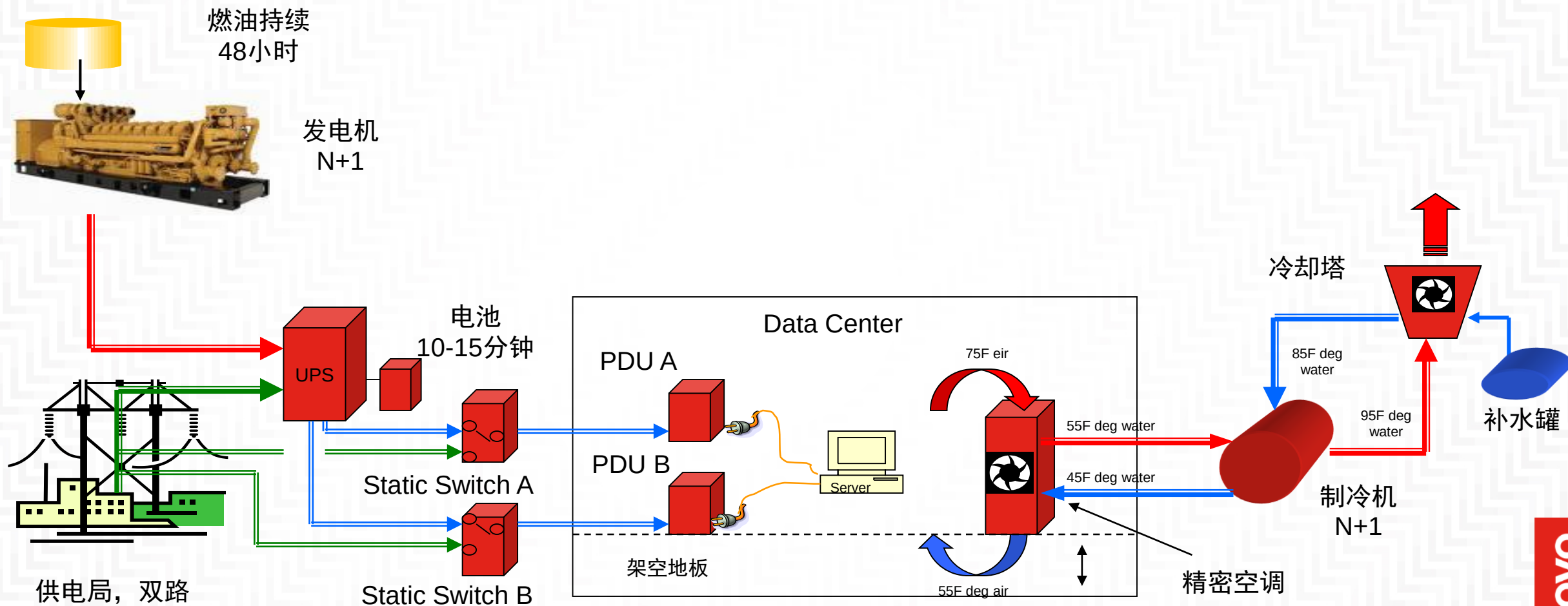


- 经过了工厂集成和兼容性测试的完整的系统（包含节点，存储，网络和集群管理）
- 完全满足您应用的要求
- 预装常见开源HPC应用
- 作为一个整体系统为您提供服务
- 7x24小时在线运维服务



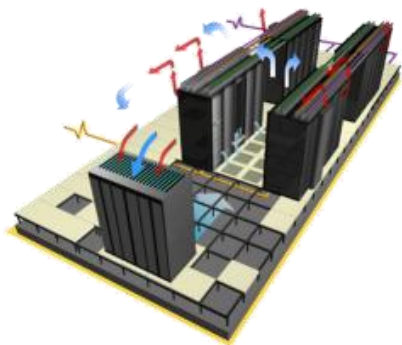
联想水冷解决方案

+ 能源, 计算和制冷



+ 数据中心制冷方式的选择 (PUE = 数据中心总设备能耗/IT设备能耗)

风冷



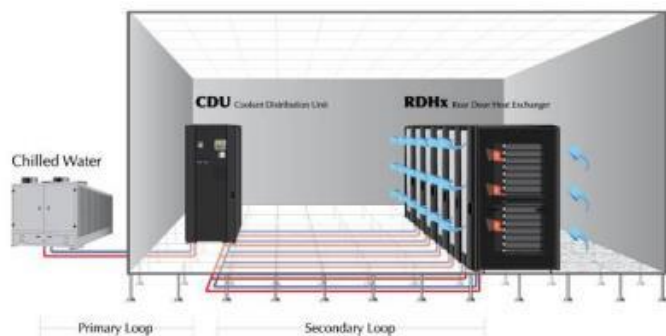
- 依靠服务器内部风扇产生的气流
- 适用任何的机房
- 具有最大的灵活性
- 广泛的支持各种设备
- Supports Native Expansion nodes (Storage NeX, PCI NeX)

PUE = ~1.5

广泛的选择

国内大部分数据中心在2到3

行间空调



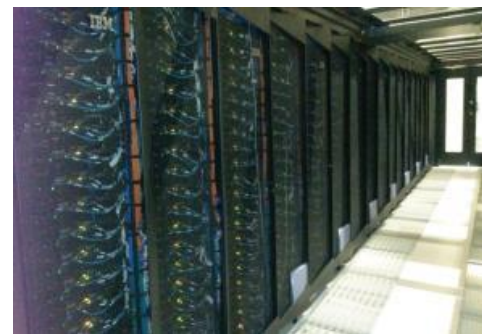
- 风冷, 通过机柜水冷和空间封闭增强散热效果
- 制冷机运行更经济 (18C water)
- 机柜摆放密度更高

PUE = ~1.2

兼顾灵活性和能效

实际环境中在1.4到2

直接水冷



- 没有内部风扇
- 更高的每瓦特性能
- Free cooling (45C water)
- 能源循环利用成为可能
- 高密度
- 高电价地区和新机房的理想选择
- 支持最高功耗的处理器

PUE <= 1.1

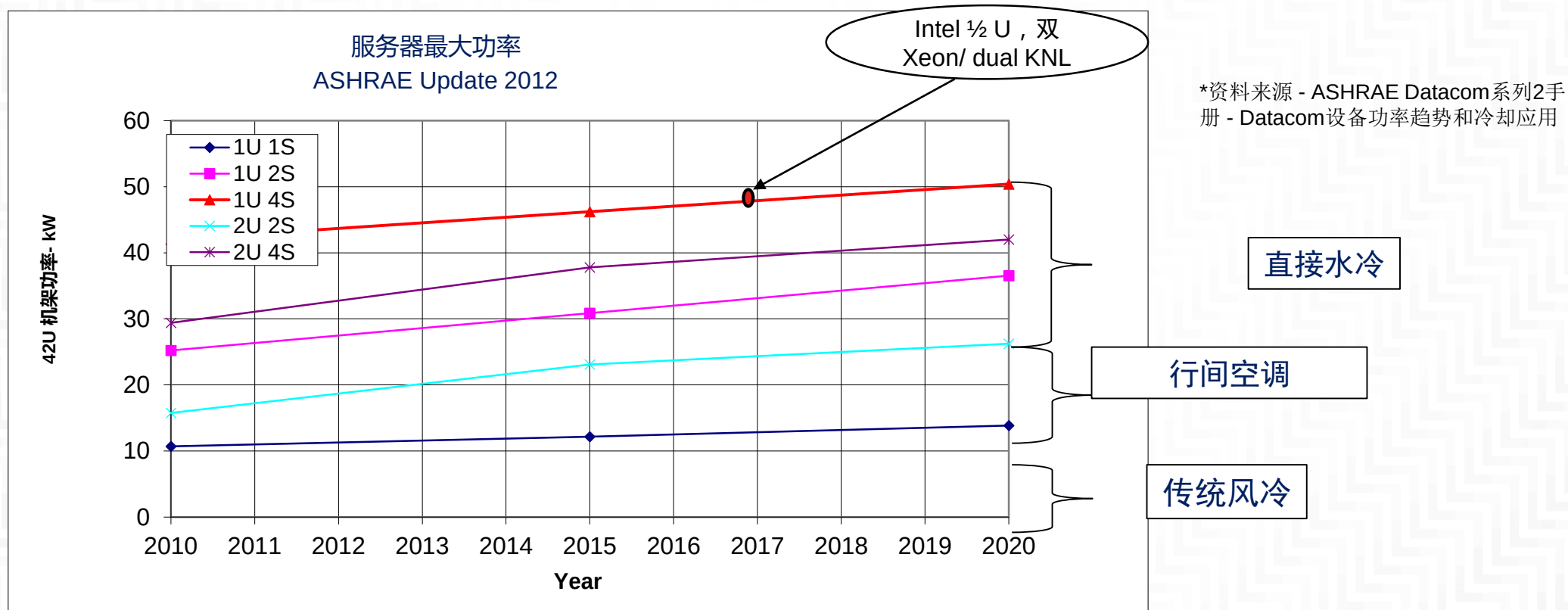
最高性能和能耗的选择

实际全年在1.4以下

或者是这三种方式的混合制冷

+ 数据中心能耗和制冷技术趋势

- ASHRAE预计，到2020年，最大机架功率约为**50kW**
- 功率跟踪（低于最大值）：HPC约低15%；企业应用比目标值约低50%
- 传统的空气冷却达到了每个机架约**12 kW**的冷却极限，这限制了节点和机架的密度
- 高密度机架需要非传统的冷却系统





System-X 直接水冷技术发展

- 利用液体冷却技术对风冷服务器进行优化
- 支持45° C的进水口温度
- 最大限度地回收热量
- 对流和空气辐射热损失约为10-15%



2013: NextScale nx360M5 WCT
45°C-50°C; 约85%的热回收
100% w/ RDHx @ 25°C

2011: iDataplex dx340M4
45°C; 约85%的热回收



2009: x3550 M3 - 1U
35°C; 100%的热回收



2009: Blade HS-22
45°C; 90%的热回收

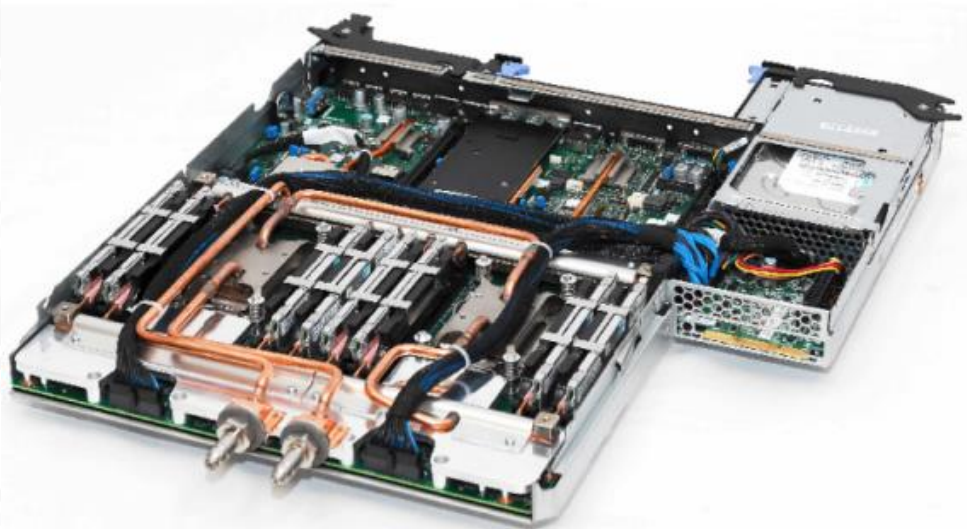


System-x DWC演化

CPU	60%
内存	18%
VR-CPU	6%
PCH/BMC/FPGA	4%
PSU	6%
IO: IB-EDR或100GbE	3%
VR内存	2%
PSU风扇	< 1%

🔴 业内第一家大规模生产和安装的X86水冷产品iDataplex dx360M4 (2010-2013)

- 具有84台dx360M4服务器的iDataplex机架
- dx360 M4节点、2个CPU（130W，115W）、16xDIMMS（4GB/8GB）、1HDD/2SSD、网卡
- 85%的热回收，水温18° C-45° C，0.5 lpm/节点



dx360M4服务器



iDataplex机架

+ 联想NextScale nx360M5 WCT 直接水冷 (2013-2016)

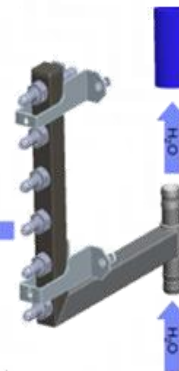
- NextScale机箱（6U/12节点），每个托盘2个节点
- nx360M5 WCT 2xCPU（最高165W）、16个DIMM（8GB/16GB/32GB）、1HDD/2SSD、1个ML2或PCIe网卡
- 82%的热回收，水温18° C-45° C，0.5 lpm/节点



托盘中的2个nx-360M5 WCT节点



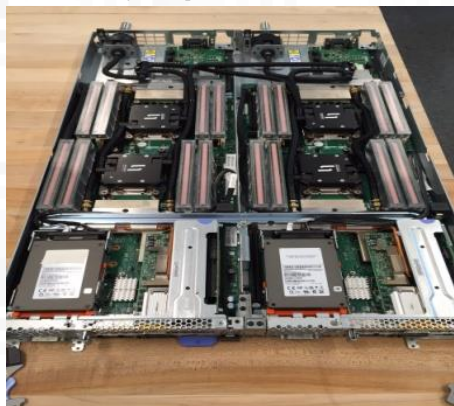
NextScale机箱



可扩展歧管



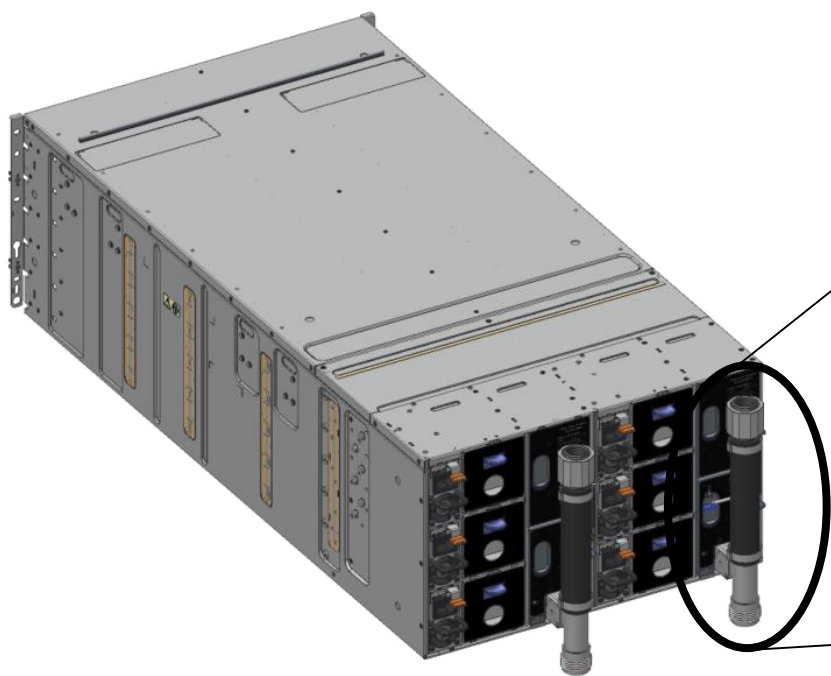
机架配置



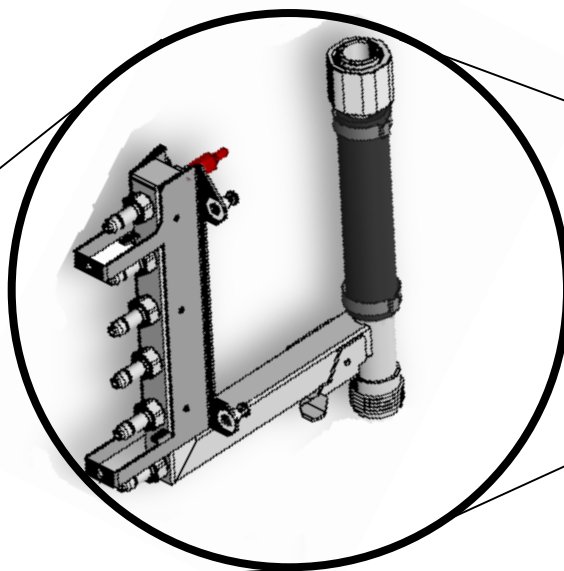
nx360M5（具有2个SSD）

+ 联想nx360 M5 WCT 主要水冷部件

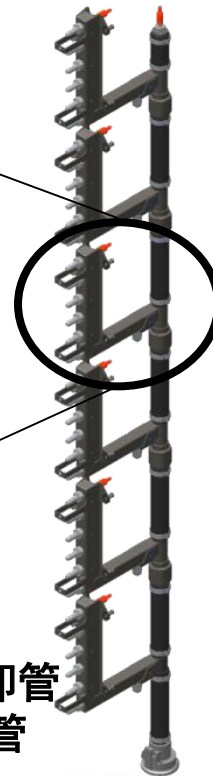
- **歧管**通过进水口与出水口之间的快速连接直接为机箱内的每个计算节点供水和排水
- 由于采用了模块化设计，因此可通过每个机箱**冷却管**的子装配构件支持多种配置
- 6种型号：1、2、3、4、5或6个机箱**冷却管**



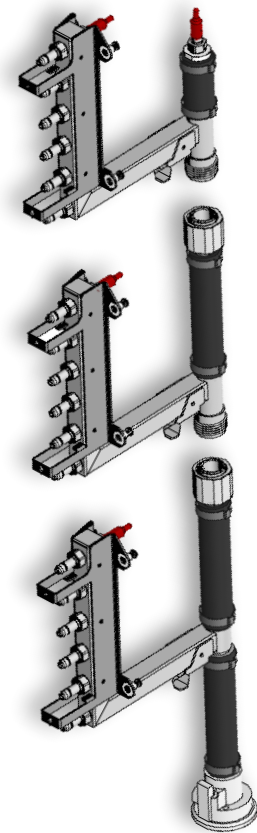
n1200 WCT机箱



单个歧管冷却管（每个机箱1个）



6冷却管
歧管



+nx360 M5 水冷节点

WCT Compute Node



- WCT计算节点
- 每个全宽1U托盘2个计算节点 – Grantley
- 通过冷却管进行水循环，以实现组件级冷却
- Intel E5-2600 v4双路处理器
- 16个DIMM插槽（DDR4）
- Infiniband支持：
 - FDR: ConnectX-3（ML2）
 - FDR: Connect-IB（PCIe）
 - EDR: PCIe
- Omnipath
- 板载GbE NIC

2016 Lenovo Confidential. All rights reserved.



风冷前面板*



1GbE/共享NIC

存储/PCI

LED
（电源、位置、日
志、错误）

双端口ML2
（IB/以太网）

用于添加
Infiniband的PCI
插槽

WCT前面板*



节点1

节点2



WCT后视图

出水口

出水口

*因配置而异配置包括MI2和PCI适配器



N1200 WCT机箱

- 6U机箱，6个托盘
- 每个托盘2个节点（每个6U机箱12个节点）
- 最多6颗900W/1300W/1500W电源，N+N或N+1配置
- 无风扇（PSU除外）
- 风扇和电源控制器
- 滴速传感器、错误LED以及用于检测漏水的Web链路

N1200风冷机箱

- 6U机箱，12个托架
- 最多6颗900W/1300W/1500W电源，N+N或N+1配置
- 最多10个热插拔风扇
- 风扇和电源控制器
- 混合搭配计算、存储或GPU节点



风扇和电源控
制器

3颗电源



风扇和电源控
制器

5个80毫米
风扇



联想水冷客户案例--SuperMUC System



第1阶段 (IBM System x iDataPlex) :

- 最佳性能为3.2 PFlops
- 18个计算节点岛中具有9216个IBM iDataPlex dx360M4节点
- 每个计算节点具有2个Intel Xeon E5-2680处理器以及32 GB内存
- 147,456个计算内核
- Infiniband FDR10网络 (胖树拓扑)



第2阶段 (Lenovo NeXtScale WCT) :

- 最佳性能为3.6 PFlops
- 6个计算节点岛中具有3072个Lenovo NeXtScale nx360M5 WCT节点
- 每个计算节点具有2个Intel Xeon E5-2697v3处理器以及64 GB内存
- 86,016个计算内核
- Infiniband FDR14网络 (胖树拓扑)

共用GPFS文件系统，分别具有10 PB和5 PB的可用存储空间

共用编程环境

直接温水冷却系统技术

“

THANK YOU

DAKUJEM DANK BEDANKT MERCI TAKK 谢谢
ありがとう СПАСИБО GRACIAS DZIĘKUJĘ DANKE
OBRIGADO БЛАГОДАРЯ GRAZIE תודה GRACIAS



Lenovo™

The image shows the Lenovo logo, which consists of the word "Lenovo" in a white, bold, sans-serif font, followed by a trademark symbol (TM). The logo is centered within a solid black rectangular background.

Lenovo™