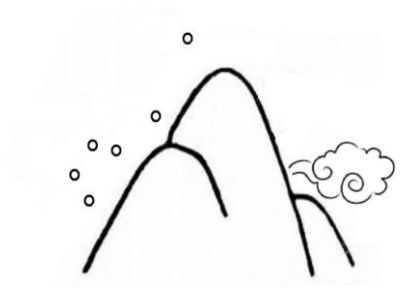


千易云预科教程

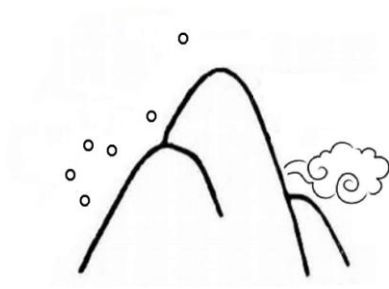
服务器篇

内容

- 服务器概述
- 服务器架构
- 服务器基准体系
- 服务器部件技术
- 主流服务器产品

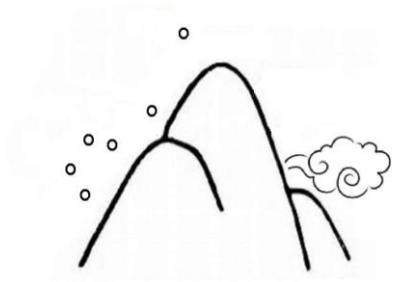


服务器概述



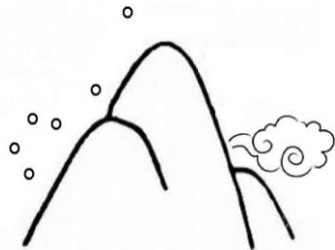
- **什么是服务器**
在网络环境下运行相应的应用软件，为网上用户提供共享信息资源和各种服务的一种高性能计算机，称为服务器，英文名称叫做Server。
- **服务器的设计思想 (RASUM)**
R:Reliability-可靠性
A:Availability-可用性
S:Scalability-可扩展性
U:Usability-易用性
M:Manageability-可管理性
- **服务器类型**
按照架构分类：X86服务器、RISC服务器、EPIC服务器
按照结构分类：塔式服务器、机架式服务器、刀片服务器
按照用途分类：数据库服务器、应用服务器、文件服务区器、网络服务器

服务器架构



- **X86架构**
复杂指令集 -- CISC(Complex Instruction Set Computer)
IA-32、EM64T、AMD64
- **RISC架构**
精简指令集 -- RISC(Reduced Instruction Set Computing)
小型机(IBM<Power>、HP<PA-RISC>、SUN<SPARC>)
专用平台、专用系统
大型应用后台密集集中处理
- **EPIC架构**
显示并行指令集 -- EPIC(Explicitly Parallel Instruction Computers)
IA-64安腾

服务器架构



- **CISC-复杂指令集**

每个微处理器的核心是运行指令的电路。指令由完成任务的多个步骤所组成，把数值传送进寄存器或进行相加运算。CISC是一种执行整套计算机指令的微处理器

- **RISC-精简指令集**

特点是所有指令的格式都是一致的，所有指令的指令周期也是相同的，并且采用流水线技术。

- **EPIC-显式并行指令集**

模式专为高效地并行处理而设计，能够同时处理多个指令或程序。并行处理可以增加每个处理器时钟周期内完成的有效工作数量，从而极大地提高应用性能。通过建立一个面向增强的并行处理能力的基础，显式并行指令计算（EPIC）支持英特尔通过提高并行指令吞吐率来扩充处理器性能。



服务器基准测试(Benchmark)体系

- **两大基准体系 -- TPC、SPEC**

TPC(Transactionprocessing Performance Council , 事务处理性能委员会)是由数十家会员公司创建的非盈利组织，总部设在美国。

SPEC(The Standard Performance Evaluation Corporation)标准性能评测机构，是国际上对系统应用性能进行标准评测的权威组织，该组织成立于1988年，是由斯坦福大学、清华大学、微软、等全球几十所知名大学、研究机构、IT企业组成的第三方测试组织。

- **四大应用中的基准测试**

在线事务处理(OLTP)：TPC-C

Web服务：SPEC Web2005

Java应用服务器：SPECjbb2005

高性能计算(HPC)：Linpark



服务器基准测试(Benchmark)体系

- **TPC-C**

TPC的成员主要是计算机软硬件厂家，而非计算机用户，其功能是制定商务应用基准程序的标准规范、性能和价格度量，并管理测试结果的发布。TPC不给出基准程序的代码，而只给出基准程序的标准规范。这便于任何厂家或其他测试者都可以根据规范，最优地构造出自己的测试系统。测试者为保证测试结果的完整性，必须提交给TPC一套完整的报告。

TPC-C使用三种性能和价格度量。其中性能由tpmC(transactions per minute，C指TPC中的C基准程序)衡量。它定义是每分钟内系统处理的新订单个数。TPC-C还经常以系统性能价格比的方式体现，其单位--\$/tmpC（即：系统的总价格/tmpC数）

- **SPEC**

SPEC体系中最广为人知的两个子项是衡量CPU、内存性能的CPU2000和WEB服务器性能的WEB2005

CPU2000;主要测试的对象是CPU、内存

WEB2005:测试原理是通过多台客户端向服务器发出http get请求，请求调用we服务器上的网页文件，大小不等。在相同的时间中，服务器回答请求越多，就表明服务器处理能力越强，系统的web性能就越好

服务器基准测试(Benchmark)体系



- **SPEC**

SPEC jbb2005:基准测试模拟一个三层架构环境来进行Java 应用服务器测试，目的是衡量应用服务器端JAVA 应用之性能

Linpack:Linpack现在在国际上已经成为最流行的用于测试高性能计算机系统浮点性能的benchmark。通过利用高性能计算机，用高斯消元法求解N元一次稠密线性代数方程组的测试，评价高性能计算机的浮点性能。

服务器CPU厂商

- 主要厂商及产品线

- Intel
 - Xeon 至强
 - Itanium 安腾
- AMD
 - Athlon MP
- IBM
 - Power

- 主要性能指标

主频、缓存、指令集、工艺、封装





服务器内存技术

- **内存技术要点**

内存由半导体器件制成，存取速度快

内存是CPU直接寻址的存储空间，是CPU与硬盘进行数据读取的桥梁

所有程序都在内存中运行，因此内存性能对应用影响非常大

内存断电后数据丢失

内存主要性能指标：主频、容量

- **内存发展史**

SDRAM、DDR、DDR2、DDR3、DDR4

- **内存技术**

服务器内存典型技术：Buffer（缓存器）、Registered（寄存器）

内存容错技术：ECC、Chipkill、内存镜像、内存热备、NVDIMM



服务器内存技术

- Buffer (缓存器) : 在计算机领域, 缓冲器指的是缓冲寄存器, 它分输入缓冲器和输出缓冲器两种。前者的作用是将外设送来的数据暂时存放, 以便处理器将它取走; 后者的作用是用来暂时存放处理器送往外设的数据。
- Registered (寄存器) : 作用是: 再次推动数据信号通过内存芯片, 使内存条上能够焊接更多的芯片。
- ECC: :ECC内存技术检测和纠正单一比特错误
- Chipkill:ECC内存技术虽然可以同时检测和纠正单一比特错误, 但如果同时检测出两个以上比特的数据错误, 则无能为力。但基于Intel处理器架构的服务器的CPU性能以几何级的倍数提高, 而硬盘驱动器的性能同期只提高了5倍, 因此为了保证正常运行, 服务器需要大量的内存来临时保存从CPU上读取的数据。这样大的数据访问量就导致单一内存芯片在每次访问时通常要提供4 (32位) 或8 (64位) 字节以上的数据。一次性读取这么多数据, 出现多位数据错误的可能性会大大提高, 而ECC又不能纠正双比特以上的错误, 这样就很可能造成全部比特数据的丢失, 系统就会很快崩溃。IBM的Chipkill技术是利用内存的子结构方法来解决这一难题的。



服务器内存技术

- 内存镜像:内存镜像是指将服务器上的内存分为两个频道。一个频道是另一个频道的镜像,用来创建内存的冗余副本。内存镜像就像硬盘存储的RAID 1。一旦其中一个频道的内存发生故障,内存控制器就会将数据转移到另一个频道上,以避免中断,并且当故障修复后,频道可以自动同步。而内存镜像的缺点也和RAID 1类似,由于内存上的内容是重复的,所以内存的成本也增加了一倍。
- 内存热备:做热备份的内存在正常情况下是不使用的,也就是说系统是看不到这部分内存容量的。每个内存通道中有一个DIMM不被使用,预留为热备内存。芯片组中设置有内存校验错误次数的阈值,即每单位时间发生错误的次数。当工作内存的故障次数达到这个“容错阈值”,系统开始进行双重写动作,一个写入主内存,一个写入热备内存,当系统检测到两个内存数据一致后,热备内存就代替主内存工作,故障内存被禁用,这样就完成了热备内存接替故障内存工作的任务,有效避免了系统由于内存故障而导致数据丢失或系统宕机。这个做热备的内存容量应大于等于所在通道的最大内存条的容量,以满足内存数据迁移的最大容量需求。
- NVDIMM: NVDIMM 是在一种集成了DRAM + 非易失性内存芯片的内存条规格,能够在完全断电的时候依然保存完整内存数据



服务器硬盘技术

- **硬盘基本参数**

容量：容量是硬盘最主要的参数，目前的硬盘容量基本都是TB级别

转速：传统机械硬盘盘片转速，单位为RPM（转/分钟）

传输率：硬盘读写数据的重要参数，分为内部传输率和外部传输率，单位一般为MB/s

缓存：硬盘缓存的目的是为了解决系统前后级读写速度不匹配的问题，以提高硬盘的读写速度。

- **硬盘类型**

按用途分：桌面级、监控级、数据中心级（企业级）

按接口分：SCSI、SATA、MSATA(Mini SATA)、SAS(Serial Attached SCSI, 由SCSI演变而来)、NVME(闪存类存储所用协议)

按介质类型分：机械硬盘、固态硬盘



硬盘技术 -- RAID

- **什么是RAID**

Redundant Array of Independent Disks -- 独立磁盘冗余阵列

RAID是将阵列中多个物理磁盘视为单一的逻辑磁盘，数据以分段方式顺序存储于磁盘阵列中

- **RAID级别**

RAID0

RAID1

RAID3

RAID5

RAID6

RAID1+0

RAID0+1

RAID50

RAID6



硬盘技术 -- RAID (续)

- RAID的实现方法

软件RAID：

硬件RAID：

- RAID技术的特点

提高系统读写性能

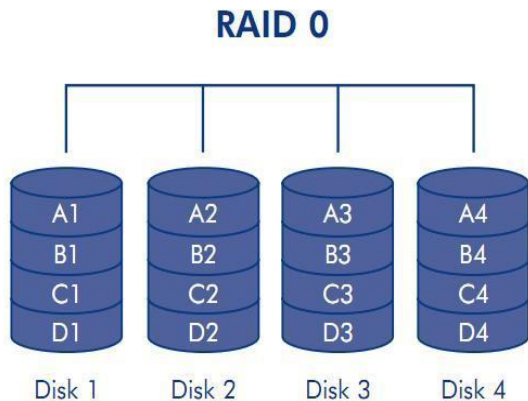
具备容错性，保护数据安全

逻辑上作为一个大容量硬盘供操作系统使用



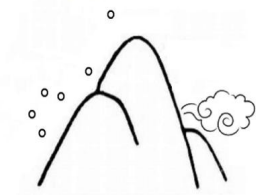
RAID技术 -- RAID0

条带化存取数据，数据块被细分在每一个磁盘，如下图：



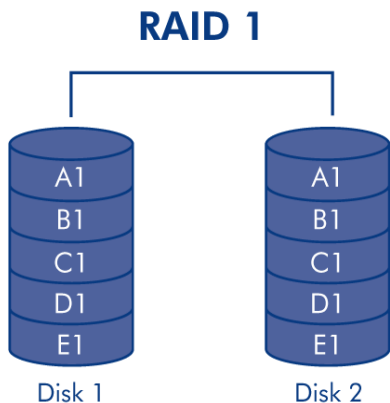
特点：

- 磁盘空间的完整利用，总容量为单个硬盘的N倍
- 读写速度最快，不允许任何一块磁盘损坏，N倍故障风险
- 高风险，用于存放安全性不高，但需要高速读写数据，经常用于图像处理和多媒体应用



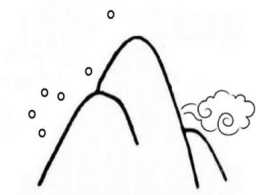
RAID技术 -- RAID1

由两块物理或逻辑硬盘组成，总容量等同于一块物理或逻辑硬盘



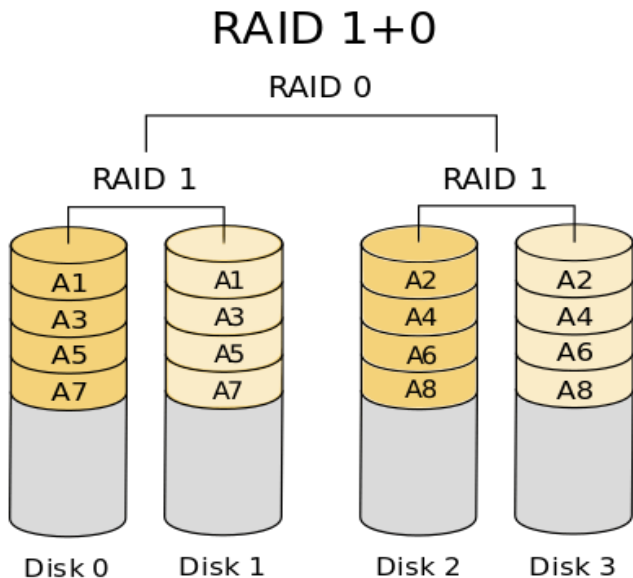
特点：

- 允许单块硬盘损坏，且硬盘故障时对性能影响小
- 两倍单盘的读性能，等同于单盘的写性能，磁盘空间利用率低
- 一般用于数据安全性要求高的环境

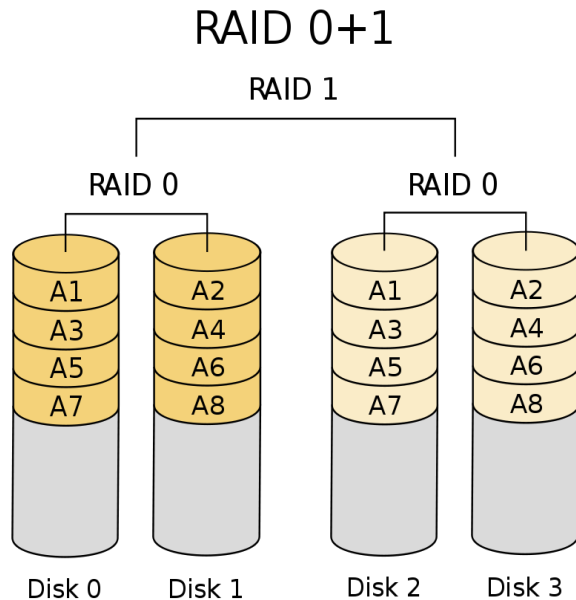


RAID技术 -- RAID10/RAID01

RAID1+0



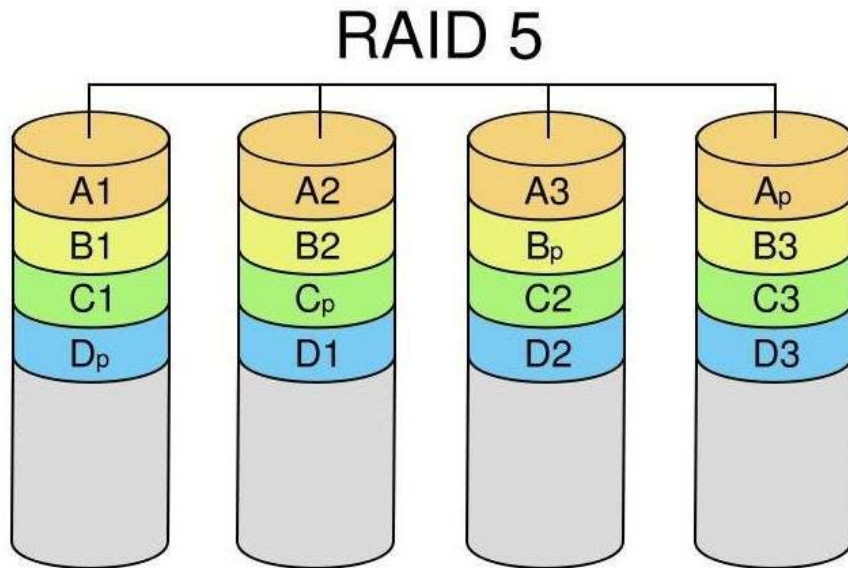
RAID0+1





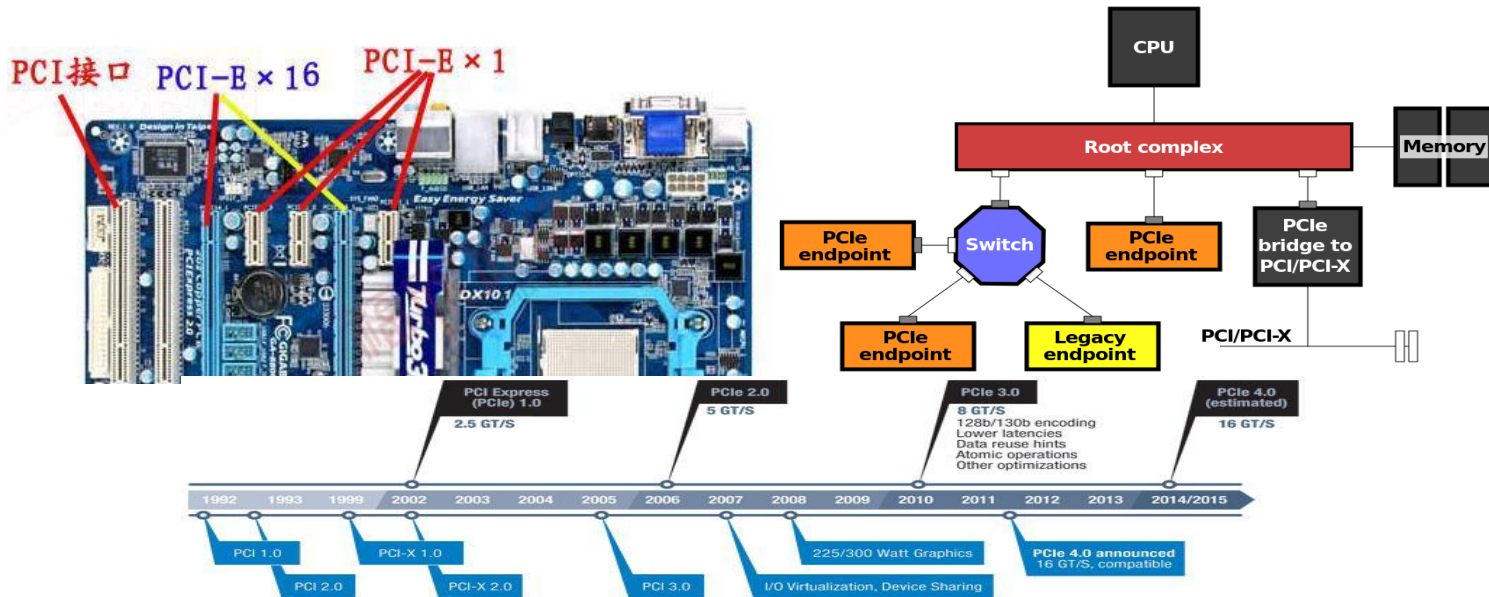
RAID技术 -- RAID5

条带化校验数据分布存储，数据块整体存在磁盘上，写操作同时生成校验数据，校验数据分布储存在不同的磁盘上



服务器PCI总线技术

- PCI总线技术要点
 - PCI插槽用途：用于连接显卡、声卡、网卡、RAID卡等各类HBA卡
 - PCI插槽规格：PCI、PCI-X、PCI-E



服务器PCI总线技术

适配器图谱

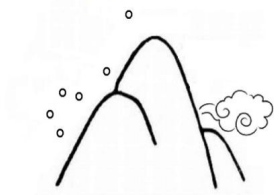




服务器PCI总线技术

各类PCI速度对比

规格	位数	频率	带宽
PCI 2.3	32 位元	33/66 MHz	133/266 MB/s
PCI-X 1.0	64 位元	66/100/133 MHz	533/800/1066 MB/s
PCI-X 2.0 (DDR)	64 位元	133 MHz	2.1 GB/s
PCI-X 2.0 (QDR)	64 位元	133 MHz	4.2 GiB/s
AGP 2X	64 位元	66 MHz	532 MB/s
AGP 4X	64 位元	66 MHz	1.0 GB/s
AGP 8X	64 位元	66 MHz	2.1 GB/s
PCI-E 1.0 1X	8 位元	1.25 GHz	250 MB/s (双工)
PCI-E 1.0 2X	8 位元	1.25 GHz	500 MB/s (双工)
PCI-E 1.0 4X	8 位元	1.25 GHz	1.0 GB/s (双工)
PCI-E 1.0 8X	8 位元	1.25 GHz	2.0 GB/s (双工)
PCI-E 1.0 16X	8 位元	1.25 GHz	4.0 GB/s (双工)



服务器 -- 网卡技术

- **网卡分类**

按介质分类：RJ-45网卡、光纤网卡

按速度分类：100Mb、1Gb、10Gb、25Gb、40Gb、100Gb

按位置分类：集成网卡(LOM)、独立网卡

端口数量分类：单口、双口、四口

- **多网卡绑定技术**

多网卡故障切换

多网卡负载均衡

- **网络加速技术**

Intel IOAT：支持Intel网卡

TOE：支持Broadcom网卡



服务器 -- 电源技术

- **服务器电源标准有两类：**

ATX标准：用于PC、工作站或低端服务器

SSI标准：目前中高端服务器常用标准，包含服务器电源规格、背板系统规格、服务器机箱系统规格和散热系统规格

- **冗余特性：**

1+1，每个模块承担50%的输出功率,当一个模块拔出时,另一个模块承担100%输出功率；

2+1，有三个模块,每个模块承担输出功率的1/3,当拔出一个模块,其余两个模块各承担50%的输出功率。

- **各种电源：**



服务器管理技术

- 标准协议

IPMI

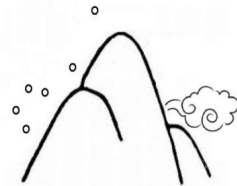
- **BMC监控产品方案**

IT基础架构的监控

应用系统的监控

事件管理/服务影响管理

资产配置管理

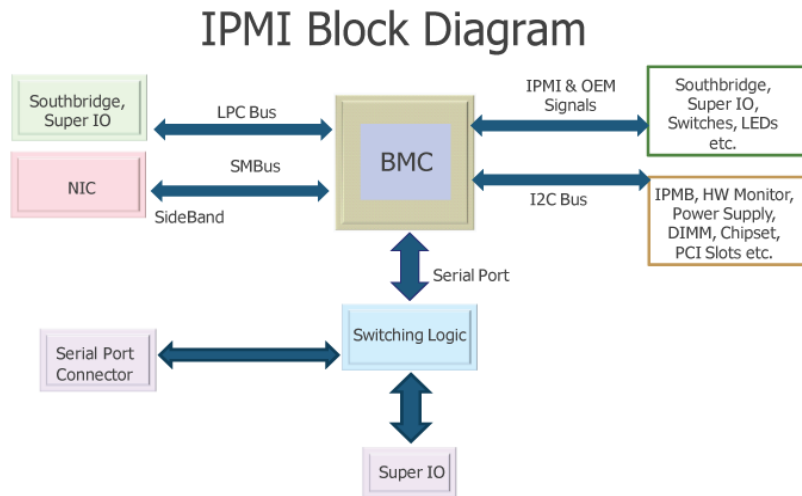


服务器管理技术 -- IPMI



智能平台管理接口(Intelligent Platform Management Interface)原本是一种Intel架构的企业系统的周边设备锁采用的一种工业标准。IPMI亦是一个开放的免费标准，用户无需支付额外的费用即可使用此标准。

IPMI能够横跨不同的操作系统、固件和硬件平台，可以只能的监视、控制和自动回报大量服务器的运作状况，以降低服务器系统成本。





服务器管理技术 -- IPMI

- **发展历史**

1998年Intel、Dell、HP及NEC共同提出IPMI规格，可以通过网络远程控制温度、电压。

2001年IPMI从1.0版改版至1.5版，新增PCI Management Bus等功能。

2004年Intel发表了IPMI 2.0的规格，能够向下兼容IPMI 1.0及1.5的规格。新增了Console Redirection，并可以通过Port、Modem以及LAN远程管理服务器，并加强了安全、VLAN和刀片服务器的支持特性。

2014年2月11日，发表了v2.0 revision 1.1，增加了IPv6的支持。

- **特性**

IPMI独立于操作系统外自行运作，并容许管理者即使在缺少操作系统或系统管理软件、或受监控的系统关机但有接电源的情况下仍能远程管理系统。IPMI也能在操作系统引导后活动，与系统管理功能一并使用时还能提供加强功能，IPMI只定义架构和接口格式成为标准，详细实现可能会有所不同。

IPMI 1.5版和之后的版本能通过直接串接、局域网、或Serial over LAN到远程提交警报，系统管理员能使用IPMI消息去查询平台状态、查看硬件的日志、或通过相同连接从远程控制台发出其他要求，这个标准也为系统定义了一个警报机制提交简单网络管理协议(SNMP) 平台事件圈套(PET, Platform Event Trap)。



服务器管理技术 -- BMC

基板管理控制器 (Baseboard Management Controller) 是一个专门的服务处理机，它利用传感器来监控一台计算机，网络服务器，或者是其他硬件驱动设备的状态，并且通过独立的连接线路和系统管理员进行通信，BMC是智能平台控制接口 (IPMI , Intelligent Platform Management Interface) 的一部分并且通常被包含在模板或者被监控的设备的主电路板里。

BMC的传感器用来测量内部物理变量，例如：温度，湿度，电源电压，风扇速度，通信参数和操作系统函数，如果这些变量中任何一个超出了指定限制的范围之外，他就会通知管理员，管理员就会利用远程控制来采取正确的措施。监控设备可以动力循环或者当必要的时候重新启动，这样，单一的管理员就可以同时远程控制无数个服务器和其他设备，节省了网络的总体成本，并且可以确保可靠性。



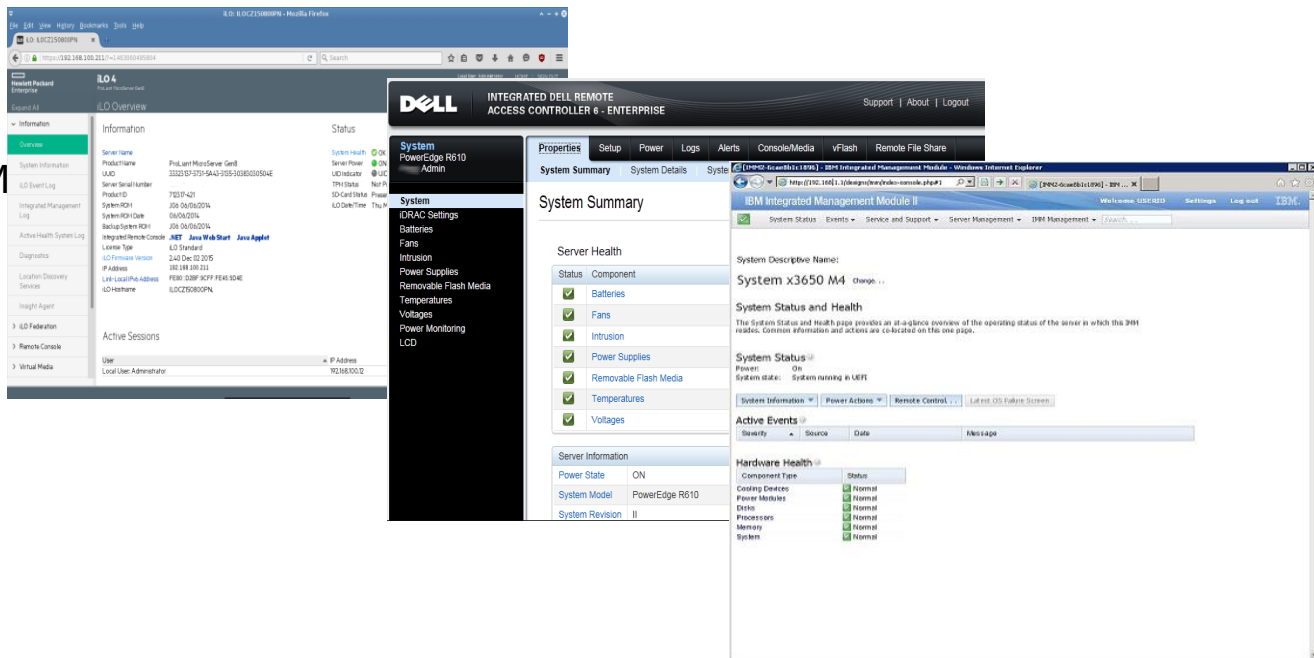
服务器管理技术 -- 主流厂商管理模块

除以上所提到的管理标准规范之外，一些主流厂商亦提供基于这些标准的自有管理模块，如以下：

HPE -- iLO

DELL -- iDrac

Lenovo/IBM -- IMM



主流服务器产品

HPE ProLiant Gen10家族 (产品主页: <http://www.hpe.com>)





主流服务器产品

HPE Integrity 家族 (产品主页: <http://www.hpe.com>)



主流服务器产品

DELL PowerEdge 家族 (产品主页: <http://www.dell.com>)



主流服务器产品

Lenovo/IBM SystemX 家族 (产品主页 : <http://b2b.lenovo.com.cn/Solution/datacenter.html>)



主流服务器产品

Lenovo ThinkServer 家族 (产品主页 : <http://b2b.lenovo.com.cn/Solution/datacenter.html>)



主流服务器产品

Lenovo ThinkSystem 家族 (产品主页 : <http://b2b.lenovo.com.cn/Solution/datacenter.html>)



主流服务器产品



IBM Power Systems家族 (产品主页 : <https://www.ibm.com/products?lnk=fdi>)





谢谢