

云数据中心

本文包括：

- 1、云数据中心代替传统数据中心
- 2、传统数据中心的组成
- 3、云数据中心的组成
- 4、云数据中心与传统数据中心的区别
- 5、云数据中心建设与演进路线
- 6、管理好云数据中心
- 7、云计算

- DC: Data Center
 - 企业IT系统的核心
 - 海量数据运算、交换、存储的中心
 - 关键信息业务应用的计算环境
 - 集中管控各种数据、应用程序、物理或虚拟化设备的环境
- DC四大焦点：可靠、灵活、绿色、资源利用率

1、云数据中心代替传统数据中心

CIA项目，IBM与Amazon竞标，AWS计算方案，云计算能力-快速水平扩展能力

2、传统数据中心的组成

- 应用软件层（Application Software Layer
 - CRM、MES
- 平台软件层（Platform Software Layer
 - 中间件
 - DBMS
- 基础设施层（IT Infrastructure Layer
 - 服务器
 - 存储设备
 - 网络设备
- 机房配套层（DC Facility Layer
 - 数据通信连接
 - 环境的控制设备
 - 监控设备
 - 独立建设的

数据中心是一个层层向上的概念

垂直扩展

- 基础设施层受到业务的强约束
- 传统的数据中心，当业务出现瓶颈时，给服务器扩内存、扩CPU、扩磁盘，还要去买更高级别的机器，称之为垂直扩展，扩展时需要暂停所有业务，有风险，且越往上扩展，成本越高

IT烟囱

- 机房配套层才可以共享，其他三层都不能共享，IT建设是割裂的，所以称之为IT烟囱、IT孤岛

3、云数据中心的组成

- 全球部署，大型数据中心+边缘数据中心互联
- 传统提供服务器、存储、网络（主要是设备），而云数据中心按需提供**IT服务**（计算、管理、数据库等等IT服务）
- 核心能力：水平扩展能力，当信息爆炸时，可以快速地通过IT服务进行平滑地水平扩展
 - 技术上的水平扩展
 - 水平扩展设备低成本
 - 资源的共享，弹性伸缩
- 云数据中心核心：云操作系统
- 多业务，共部署
- 基础设施层提供的云服务，称之为IaaS
- IaaS最为成熟，CIA当年也是选择的Amazon的IaaS
- SaaS最早提出
- PaaS发展最慢

4、云数据中心与传统数据中心的区别

- 传统数据中心
 - 面向应用软件提供设备
 - 烟囱式架构
- 云数据中心
 - 提供IT服务
 - 共享服务器架构

业务价值差异

- 传统DC
 - 资源闲置
 - 波峰波谷不均匀，所以资源利用率很低
 - 上线太慢
- 云DC
 - 实现企业级的资源共享，资源利用率很高
 - 水平扩展带来的资源弹性和杜绝浪费
 - 资源获取很快，上线很快

可扩展性的差异

- 传统DC
 - 割接时，不用停业务
 - 资源供给速率太慢
 - 成本越来越高
- 云DC
 - 资源供给速率
 - 使用效率
 - 资源扩容模式，垂直向上变成水平平滑线性扩展
 - 降低成本
 - 资源维护效率提高

云DC重点关注的问题

- 降低成本，高效建设
- 提升效率，高效管理
- 缩短上线时间，敏捷使用

5、云数据中心建设与演进路线

- 高效建设云DC的前提
 - 选好标准
 - 制定演进方式
 - 统一架构
- 云DC演进方式：
 - 全局规划，大公司如Amazon
 - 建设资源池、以点带面逐步渗透，适合稍弱的公司
- 云转型的三条路径
 - 购买云服务，企业托管云，共有云
 - 新建云DC，DC整合搬迁，适合如电信银行等企业
 - 外围IT化、核心IT化，业务云化，适合大中型企业
- 4种典型场景
 - 运营商IDC公有云
 - 园区IDC公有云
 - 政务外网IT托管私有云
 - 大企业IT托管私有云

6、管理好云数据中心

- 考虑管理模式
- 被动运维与主动运维（关键）
 - 靠工具靠手段，不是靠人
 - 事先预防，哪台DC可能会出问题，这就是主动运维
- Day And Night运行
 - Day，日出，繁忙的时候，开启所有DC
 - Night，日落，空闲的时候，关掉，节能
- DC资源健康体检可视化
 - 把运维经验写成脚本，提高自动化程度

7、云计算

- 资源抽象成为共享资源池
- 上层操作系统与硬件解耦，操作系统从资源池中分配资源
- 4个特点
 - 分区
 - 隔离
 - 封装
 - 独立
- 安全可靠
- 自动化
- 绿色
- 云计算面临的新威胁
 - 虚拟化安全威胁

- 多租户安全威胁
- 特权用户问题 云计算是下一代数据中心的核心技术，虚拟化技术是云计算的灵魂，而再云DC的环境里安全，容灾等技术是客户关注的重点技术

测试

[对] 实现云数据中心后，需防止出现虚拟化扩张造成的资源浪费，避免因过量分配资源导致资源不足而影响业务。

[错] 虚拟化使服务器工作更饱满，一般数据中心运行初期负荷很高。

下列关于云数据中心的特点，说法错误的是哪一项？C

- A 自动化管理，包括对物理服务器、虚拟服务器的管理，对相关业务的自动化流程管理、对客户服务的收费等自动化管理
- B 绿色节能，云计算数据中心在各方面符合绿色节能标准
- C 节省成本，由于云数据中心不建设在企业内部，节省相应成本
- D 高度虚拟化，这其中包括服务器、存储、网络、应用等虚拟化，使用户可以按需调用各种资源

下列关于云数据中心向用户提供服务的特性，哪一项说法是错误的？A

- A 满足客户请求，避免重复服务
- B 服务可以快速部署并能够回收
- C 需要实现自助服务
- D 需要有计费系统，按使用量和服务级别计费

与传统数据中心相比，云数据中心存在着许多优势，以下关于云数据中心的说法，哪些是正确的？ABD

- A 解除了 IT 基础设施增长过程中受到的场地、空间等制约
- B 云数据中心的基础设备更加规模化、标准化
- C 与传统物理数据中心相比，需要更多的维护人员
- D 数据中心为了节省成本、实现日益增多的业务，必须采用各种虚拟化技术

存储基础原理

本文包括：

- 1、存储的定义及其发展历程
- 2、存储与应用环境
- 3、存储介质
- 4、RAID 技术及应用
- 5、存储阵列技术
- 6、SAN技术及应用
- 7、NAS存储
- 8、备份容灾
- 9、存储高级技术

1、存储的定义及其发展历程

- 存储与数据
 - 数据是首要的，有数据才有存储，
 - 每三年数据翻倍，类型多种多样，数据爆炸年代
 - 因数据价值很高，所以需要专门的存储设备
 - 狭义存储
 - 软盘、CD、DVD、磁带

- 广义的存储
 - 存储硬件系统
 - 存储软件系统
 - 存储网络
 - 存储解决方案：

关系如下：服务器通过存储网络才能够访问存储硬件系统中的数据，存储软件系统对存储中的数据提供管理，将多种存储硬件和软件组合起来形成解决方案，可以满足业务较高的数据管理需求，比如数据整合的解决方案，容灾备份的解决方案

- 业务与存储
 - 当今的存储是广义的存储，是一种复杂的用来存储和管理企业重要信息的系统。
 - 存储对业务的进行处理和数据保护，备份容灾。
- 存储的发展历程
 - 最开始，因需求不高，服务器与存储设备是一体的
 - 后来考虑硬盘从服务器中剥离，称之为 DAS（Direct Attached Storage：直接连接存储）
 - JBOD（Just a Bunch Of Disks，磁盘簇）：数据控制管理+存储，数据控制管理仍由服务器负责，因此服务器仍有额外开销
 - 智能化：服务器只需关注数据读写即可
 - 如 VAS 智能业务存储，90年代末
 - 集中式存储（现在都用这个）：SAN 和 NAS
 - SAN：服务器存储数据，首先要在服务器用文件系统来格式化，更多用于数据库场景，来存储结构化的数据，性能要求高
 - NAS：面向用户办公的数据共享，服务器的文件系统放在存储里，面向非结构化的存储，提供网关
 - 海量存储，属于 NAS 存储形态
 - 生产中心和灾备中心
 - 统一存储：SAN + NAS

关于 DAS、NAS、SAN（from 知乎用户 LSI China: <https://www.zhihu.com/question/24335605>）

- DAS 即直接连接存储（Direct Attached Storage）
 - DAS 是指将外置存储设备通过 SCSI 或 FC 接口直接连接到应用服务器上，存储设备是整个服务器结构的一部分。
 - 在这种情况下，数据和操作系统往往都未分离。
- NAS 即网络接入存储（Network Attached Storage）
 - NAS 采用网络技术（TCP/IP、ATM、FDDI），通过网络交换机连接存储系统和服务器主机来建立存储私网。
 - 其主要特征是把存储设备、网络接口和以太网技术集成在一起，直接通过以太网网络存取数据。也就是把存储功能从通用文件服务器中分离出来。
- SAN 即存储区域网络（Storage Area Network）
 - SAN 是通过光纤交换机连接存储阵列和服务器，建立专用数据存储的存储私网。
- 三种模式比较
 - 三种模式中，DAS 模式最简单，就是直接把存储设备连接到服务，而这种模式最大的问题是：每个应用服务器都要有独立的存储设备，这样增加了数据处理的复杂度，随着服务器的增加，网络系统效率也急剧下降。
 - 为了解决上述问题，提出了 NAS 和 SAN 两种模式。
 - NAS 和 SAN 最本质的区别就是文件管理系统在哪里。如图 1 所示，SAN 结构中，文件管理系统（FS）分别在每一个应用服务器上面，而 NAS 则是每个应用服务器通过网络共享协议，使用同一个文件管理系统。即 NAS 和 SAN 存储系统的区别就是 NAS 有自己的文件管理系统。

□

2、存储与应用环境

3、存储介质

- 机械硬盘（主流）
 - 硬盘物理结构：
 - 盘片：数据存放，
 - 启停区，
 - 数据块区（磁道），磁道越多，硬盘容量越大，
 - 扇区：在磁道上可以划分扇区，一般8K，
 - 柱面：不是真实存在的

- 磁头数：盘片个数的两倍
- 主轴：驱动盘片高速运转
- 磁头驱动设备：写数据
- 接口：
- 控制电路：供电

。硬盘电子结构：

- 高速缓存芯片：影响性能
- 主控芯片

。主要参数

- 硬盘容量
- 转速：性能越好
- 缓存：性能越好
- 平均访问时间：平均寻道时间+平均等待时间（4毫秒以内）
- 数据传输率
- $IOPS = 1000 \text{ ms} / (\text{寻道时间} + \text{旋转延迟})$
 - 旋转延迟指盘片旋转将请求数据所在扇区移至读写磁头下方所需要的时间，一般用磁盘旋转一周所需时间的1/2标识
 - 比如，7200rpm的磁盘，平均寻道时间为3ms，问IOPS，
 - 则 $IOPS = 1000 / (3 + 60 * 1000 / 7200 / 2) = 140$
- 物理磁盘总的 IOPS = 物理磁盘的 IOPS × 磁盘数目

可用的 IOPS = (物理磁盘总的 IOPS × 写百分比 ÷ RAID 写惩罚) + (物理磁盘总的 IOPS × 读百分比)

假设组成 RAID-5 的物理磁盘总共可以提供 500 IOPS，使用该存储的应用程序读写比例是 50%/50%，那么对于前端主机而言，实际可用的 IOPS 是：

$$(500 \times 50\% \div 4) + (500 \times 50\%) = 312.5 \text{ IOPS}$$

• SSD硬盘

- 。有很多芯片构成
- 。分为控制单元与存储单元，其中存储单元包括两种：
 - FLASH式（大部分用这种） 应用在可移动中
 - 内存式 应用在系统中
- 。各种接口

- ATA接口
 - 高级技术附加装置
- SCSI接口
 - 优点：适用性广、高性能、支持热插拔
 - 缺点：价格昂贵，安装复杂
- SATA接口
 - 速率至少是ATA接口的两倍
 - 支持点对点
 - 支持热插播
 - 应用于可移动
- SAS接口
 - 串行连接SCSI接口
 - 支持全双工
 - 支持点对点
 - 支持双端口
 - 价格更高
 - 速率慢
- FC接口
 - 光线通道接口
 - 长距离传输
 - 传输带宽高

- 全双工
- 热插拔
- 价格昂贵，成本高
- 组建复杂
- 1、SAS； 2、SATA； 3、NL-SAS； 4、SCSI； 5、IDE
- 速率从高到低：54231

4、RAID 技术及应用

- RAID: Redundant Arrays of Independent Disks 独立冗余磁盘阵列，简称磁盘阵列，条带化
- RAID实现方式两种：
 - 硬件RAID，安装RAID卡
 - 软件RAID，通过CPU处理
- RAID组的四种工作状态
 - RAID组创建
 - RAID组正常工作：创建成功或重建成功（降级后重建）后
 - RAID组降级：成员盘掉线或故障
 - RAID组失效：故障盘数超过冗余盘数时，失效

结论：当有降级时，及时更换

- RAID级别
 - RAID0
 - 条带化重组
 - 至少需要两块硬盘
 - 多块硬盘同时工作
 - 没有校验，没有镜像，若有任何一个盘出问题，则需要对整个RAID0重新格式化
 - 正是因为没有校验没有镜像读写性能最高，磁盘利用率最高
 - RAID1
 - 条带化重组
 - 至少需要两块硬盘
 - 提供镜像（复制），允许一份数据丢失
 - 安全高与RAID0，性能低于RAID0
 - 数据库用RAID1
 - RAID3
 - 奇偶校验
 - 校验信息存放于专用硬盘，至少需要3块硬盘，但有可能负载太大
 - **写性能最低！**
 - 冗余备份，允许一块盘出故障
 - RAID5
 - 校验信息分布式存放，至少需要3块硬盘，安全优于RAID3
 - 数据读取比RAID3快
 - 允许一块盘出故障
 - RAID6
 - 两次校验，分布式存放，至少需要4块硬盘
 - P+Q方式，两次校验恢复允许两块盘同时出故障
 - 数据写入一般，读取比较快
 - DP方式，两次校验恢复允许两块盘同时出故障
 - 出故障概率比较大
 - 数据写入一般，读取比较快
 - RAID10（目前比较常见）
 - 先RAID1，再RAID0，需要4块硬盘
 - 读写比较快
 - 结合RAID0，RAID1的优点
 - 数据库可以用RAID10

- RAID50
 - 先RAID5（至少需要3个），在RAID0（至少需要2组），至少需要6块硬盘

5、存储阵列技术

- 控制框
 - 管理网口
 - 维护网口
 - 电源模块（1+1冗余
 - 风扇模块（1+1冗余
 - SAS级联端口
 - 串口
- 硬盘框：主要用于容纳各种硬盘，为应用服务器提供重组存储资源

6、SAN技术及应用

- 存储结构
 - NAS：网络连接存储
 - DAS：直接连接存储
 - SAN：存储区域网络storage area network
- 组网结构
 - 直接组网：一台应用服务器-光线卡-存储设备
 - 单交换组网：多台服务器-光线交换机-存储设备
 - 双交换组网：多台服务器-多台光纤交换机-存储设备
- SAN组成部分
 - 主机总线设配卡
 - 光纤线缆、IP线缆
 - 光纤交换机、以太网交换机
 - 存储阵列设备
- IP上的架构
 - 点对点
 - 双交换
 - 单交换
- FC上的架构
 - 光纤上的通道
 - FC协议栈，五层，FC0~FC4，一般三层：FC0~FC2
 - FC三种拓扑结构
 - 点对点：只能连接2个设备，直接连接，缺点：所有存储只能为同一个应用服务
 - 仲裁环（单交换）：光纤集线器，环路工作，一台有问题就会导致环路出问题
 - 交换式（双交换）：FC交换机冗余设计
- FC SAN 与IP SAN对比
 - 传输距离：IP更高
 - 速率：FC速率更高
 - 效率：FC更高
 - 成本：FC更高
 - 安全：FC更难被攻击
 - FC适用于企业内部
 - IP适用于容灾备份

7、NAS存储

- NAS：Network AttachedStorage 网络附属存储

- NAS拥有自己的文件系统和操作系统，在内部完成
- NAS拥有众多协议应对不同的操作系统和应用场景
 - CIFS: NAS对Windows共享文件时，采用CIFS（common internet file system）通用internet文件系统，CIFS可以看成是FTP和HTTP的一个实现，传输协议TCP/IP
 - NFS: NAS对Linux、Unix共享文件时，NFS（network file system），传输协议TCP或UDP，网络安全性比CIFS低
- NAS组成
 - NAS引擎：提供文件系统以及承载文件系统、
 - 网络接口：对外提供服务，体现在两个协议上，NFS、CIFS
 - 存储：RAID、SCSI、SAS、FC，可以包容SAN
- NAS文件系统IO与性能影响
 - 客户端与主机
 - 网络
 - NAS本身内部（包括NAS引擎与存储阵列）
- NAS与SAN比较
 - 速度：SAN快
 - 成本：SAN高（SAN主要用FC架构）
 - 客户端资源占用：SAN占用高（SAN没有自己的文件系统，而NAS有自己的文件系统）
 - 协议：
 - SAN: FC,SCSI,ISCSI
 - NAS: CIFS,NFS,HTTP,FTP
 - 共享：
 - SAN需客户端安装专用共享文件系统
 - NAS自己的文件系统就可以提供文件共享
 - 应用场景：
 - SAN: 高速信息存储（DB，企业数据）
 - NAS: 偏重文件共享（文件共享，视频等）

8、备份容灾

- 备份：将文件系统或数据库系统中的数据复制，一旦发生灾难或错误操作时，可以放面及时地恢复系统的有效数据保证业务正常运行
 - 备份服务器
 - 备份软件
 - 备份PC
- 备份结构-LAN BASED
 - 优点
 - 备份系统和应用系统分开
 - 缺点
 - 同时进行，影响业务
- 备份结构-LAN FREE
 - 不占用LAN资源，由专用的高速存储网络
 - 优点
 - 不受影响，业务正常运行
 - 影响
 - 下层的存储网络需要搭建
- 备份方式
 - D2D备份
 - 磁盘disk to disk，成本高，速度快
 - D2T备份
 - 磁盘-物理磁带，成本低，速度慢
 - D2V备份

- 磁盘-虚拟磁带库，成本适中，速度适中
- D2D2T备份（目前使用广泛）
 - 磁盘-虚拟磁带库-物理磁带库，兼顾可靠性、可管理性、性能
 - 每段时期要备份，每段时期要归档，适用于大企业
- 备份要素
 - 数据类型
 - 备份介质
 - 数据类型
 - 数据保留时间
 - 备份周期
 - 备份窗口：什么时候开始备份
- 备份类型
 - 完全备份
 - 差异性备份：所要求空间小，每次备份时与某次完全备份时的差值
 - 增量备份：磁盘利用率最高，每次备份时只备份与上次相比增加的部分，数据恢复要求很高
- 容灾级别
 - 业务级 级别最高，成本最高，相当于在异地创建新的数据中心
 - 应用级 级别中等
 - 数据级 级别最低
- 容灾系统建设-系统指标
 - RPO恢复点目标：灾难发生后，系统和数据必须恢复到的时间点要求，值越小越好
 - RTO恢复时间目标：灾难发生后到完成容灾操作的所需时间，值越小越好

9、存储高级技术

- 硬盘坏道检测技术与修复及时
 - 每30天为周期对硬盘进行周期性扫描
 - 读写失败自动分析
 - 磁盘介质自动扫描
- 磁盘健康分析技术
 - 上传磁盘的信息至分析中心，如华为的ISM软件
- 磁盘预拷贝技术
 - 降低重构（重建？）的概率
 - 当发现某个硬盘疑似故障时，将该盘上的迁移数据到热备盘，迁移完成后，用新盘替换掉故障盘，数据会被拷贝到新盘
- 重复数据删除技术
 - 减少数据存放的空间，节省空间
 - 窗口：备份的时间
 - 衡量维度
 - 位置
 - 时刻
 - 在线重删
 - 后处理重删
 - 粒度
 - 文件集重删，视频音频可以用这种
 - 块级重删，哈希算法，比较哈希值，数据库可以用这种
 - 字节级重删
 - 范围
 - 本地重删，比较本地硬盘
 - 全局重删，比较整个磁盘组或阵列