

# 宏杉科技数字图书馆存储架构设计

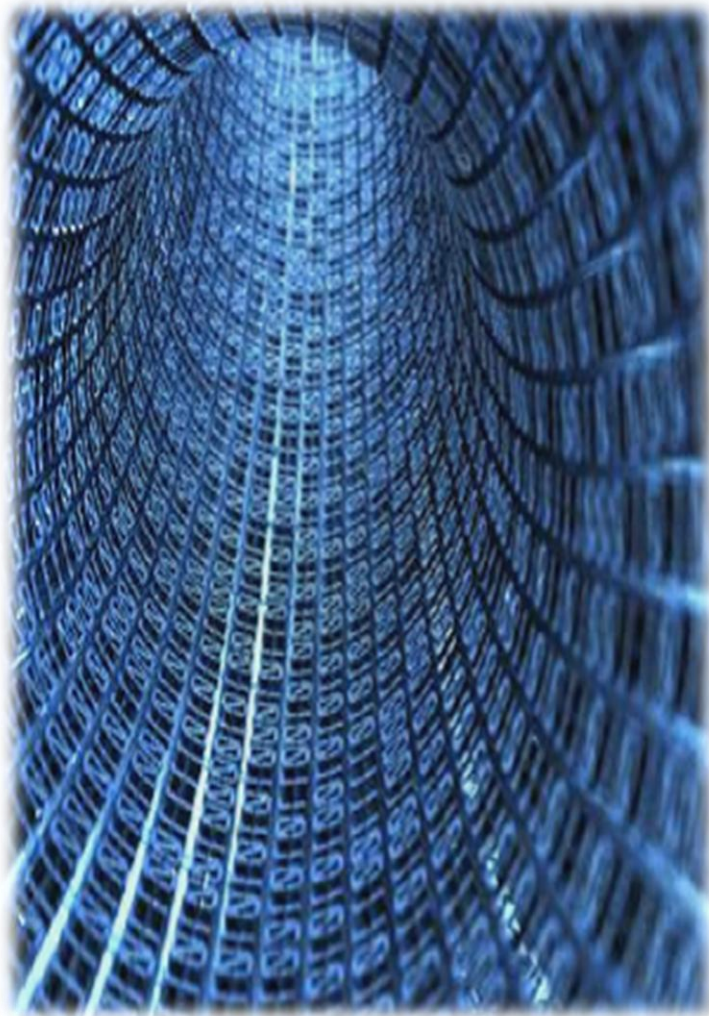


# 信息存储的本质：时空结合的信息传递

**信息存储：跨越时间进行信息传递的过程。**

- **跨越时间的传递**：我们称之为记忆、存储。例：读《离骚》，信息跨越时间传递
- **跨越空间的传递**：我们称之为通讯、传输。例：打越洋电话，信息跨越空间传递

**通讯是传播知识，存储是积累知识**



数字图书馆是**没有时空限制**的知识网络和知识中心，是一套完整的知识定位系统。

传统图书馆



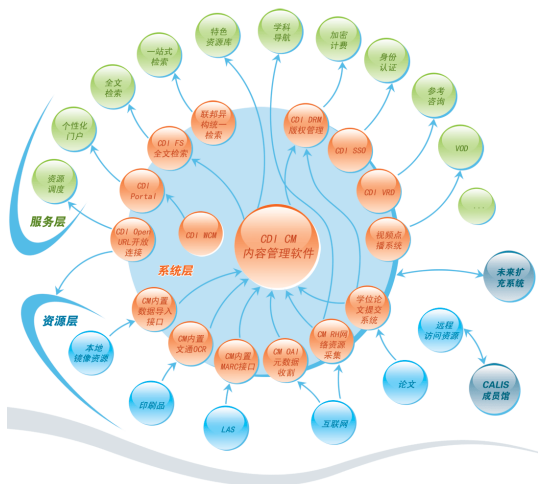
数字图书馆



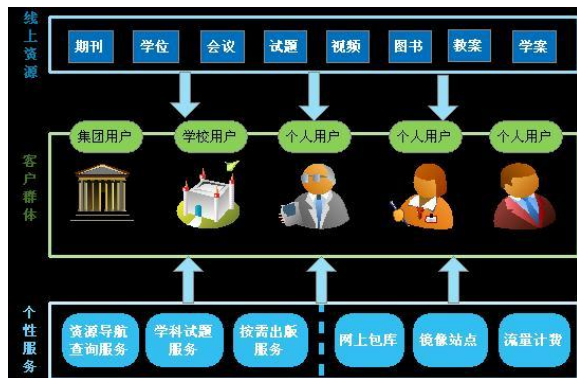
- 馆舍
- 读物：书、刊、报、资料...
- 管理人员
- 目录卡片：书名、作者、出版社、内容简介...

- 数字资源永久保存体系
- 知识库
- 智能检索
- 网上文献传递，网上导航
- 个性化服务...

# 数字图书馆体系结构

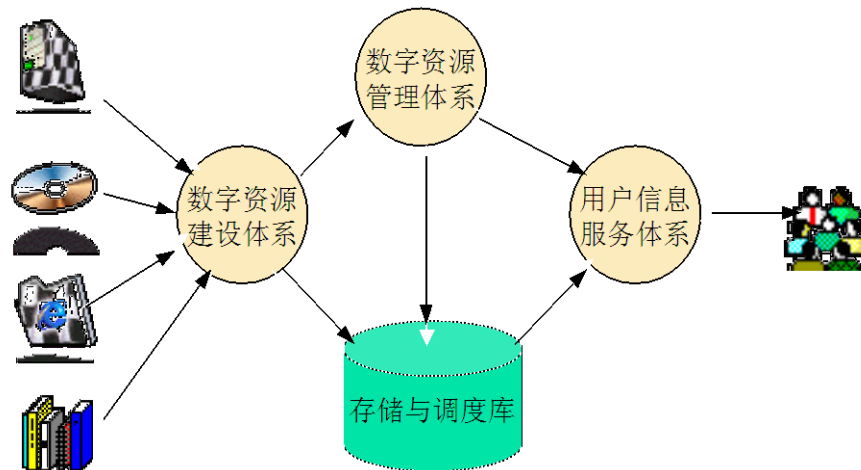


数字图书馆服务架构举例



数字图书馆服务架构举例

存储作为知识信息的载体，  
是不可或缺的关键



数字图书馆基本体系架构

杭州宏杉科技有限公司 ( MacroSAN Technologies Co.,Ltd. ) 成立于2010年5月，总部位于杭州，是**具有自主知识产权，从事全系列存储产品的研发、生产、销售、服务的高科技企业。**

杭州宏杉科技致力于专业存储设备及相关软件的开发，在杭州和深圳设立研发中心。研发团队拥有十年存储设备研发经验和技術积累，在存储软硬件核心技术领域拥有数十项发明专利。

杭州宏杉科技2012年销售额过亿，服务客户超过1000个。

**愿景：立足中国放眼世界，做专业存储领域领导品牌**



## 应用 存储

### NAS网关



NS2000G

### 集群存储



MC系列

### 大数据存储



MP系列

## 存储 阵列

### 单控系列



MS1000

### 双控系列



MS2000



MS3000



MS5000

### 多控系列



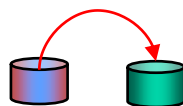
MS7000

## 软件 系列

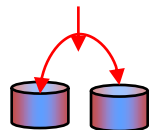
### 功能套件



快照套件



复制套件



镜像套件

### 平台软件



ODSP开放数据存  
储平台



IDSMS智能数据服  
务管理套件

# MS7000：宏杉科技新一代高端企业存储阵列



MS7020/7040

业界最高硬件规格的  
新一代高端存储阵列

## 高规格

- 2-16个引擎
- 6TB缓存扩展
- 4096Gb总线带宽
- 384个主机接口
- >400万IOPS

## 多引擎

- 16个引擎扩展
- 分布式共享缓存
- 多缓存镜像
- 交换式后端网络
- 负载均衡与冗余

## 虚拟化

- 控制器集群(虚拟控制器)
- 智能缓存池(虚拟缓存)
- HotCache技术(二级缓存)
- ICMT技术(虚拟Cell)
- CRAID技术(虚拟RAID)

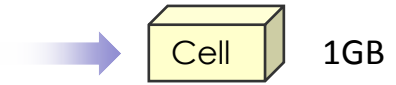
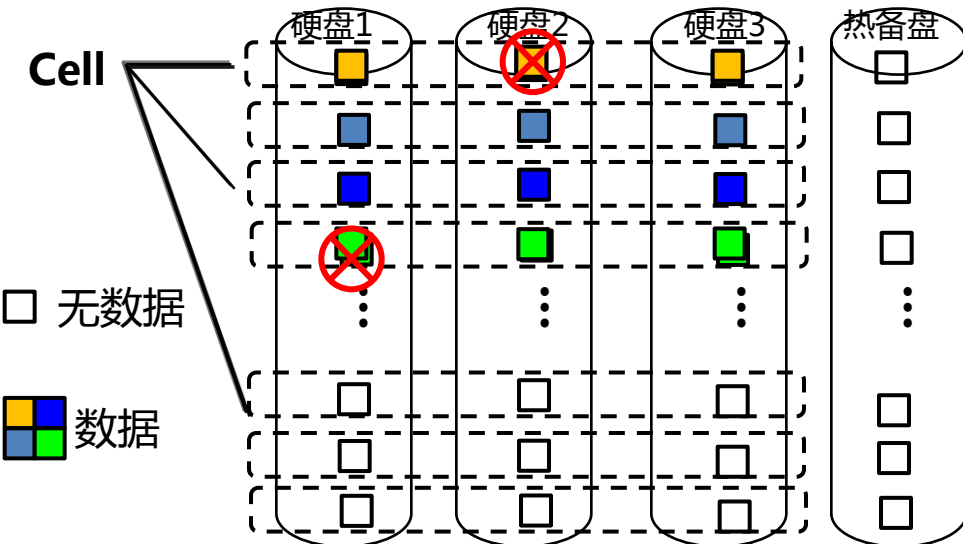
# CRAID : 革命性的创新技术



基于Cell的RAID技术，能够将RAID重建时间从数十小时降为数分钟

RAID类型	允许同时介质故障的硬盘数	允许同时整盘物理损坏的硬盘数
RAID10	极端情况2块，最多1半	极端情况2块，最多1半
RAID5	1块	1块
RAID6	2块	2块
<b>CRAID</b>	<b>多块，最多所有硬盘</b>	<b>1-3块</b>

■CRAID技术（基于Cell的RAID）：降低RAID失效的概率(年故障率 < 0.5%)

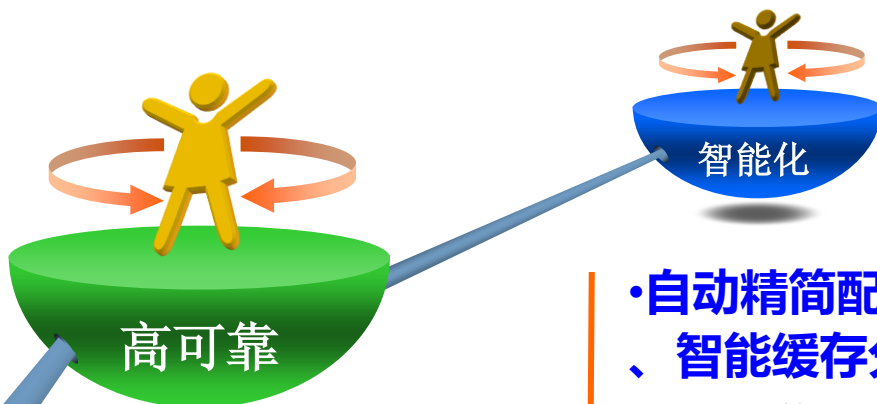


RAID组的最小维护单位：  
由硬盘变成了更小更灵活的Cell



•CRAID、IDDC、多路径、双活控制器、双机双阵列

- ✓RAID重建效率提高数十倍
- ✓减少80%硬盘故障宕机时间
- ✓业务不中断的故障处理



•自动精简配置、分层存储、智能缓存分区

- ✓容量弹性分配，按需调配资源
- ✓简化管理，提升性能



•存储专用处理器、大容量缓存、百Gb无阻塞带宽、SSD高速硬盘

- ✓性能比传统产品提升3-5倍
- ✓单台设备具备百万级IO能力，带宽大于5GB

# 大数据时代对图书馆业务的挑战

90%的数字内容属于图片、音像等非结构化内容。

挖掘数据价值



数据集成

挖掘数据价值举例：

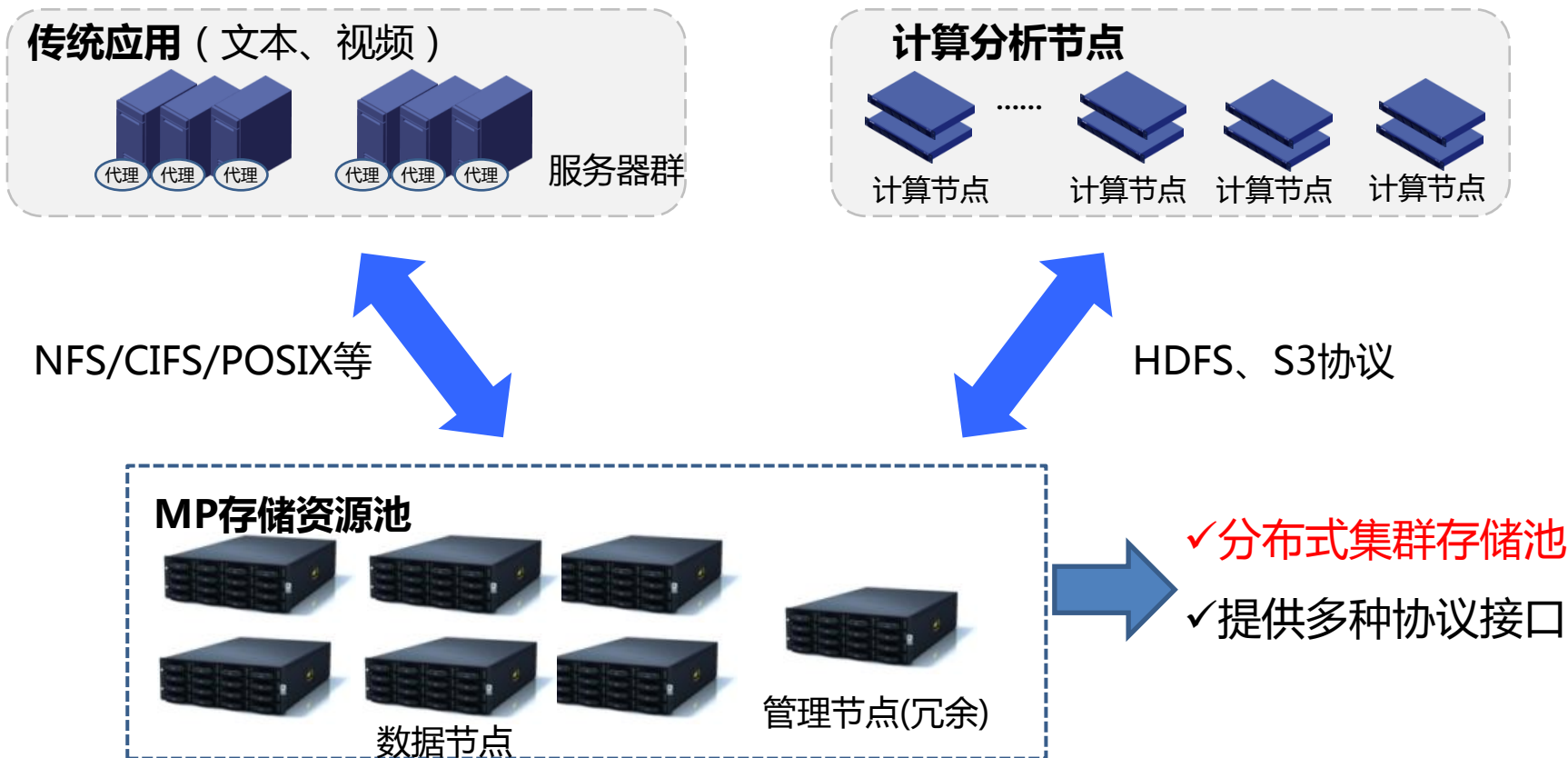
- ✓资源及学术搜索
- ✓资源及服务推荐
- ✓知识服务社区行为智能分析
- ✓用户知识需求预测

2010年全球数据总量达600EB ( 1EB=10<sup>18</sup>

Byte ) , 全球信息总量每两年就可以翻一番

# 宏杉科技大数据统一存储架构

传统应用进行数据集成，大数据分析节点用于数据挖掘



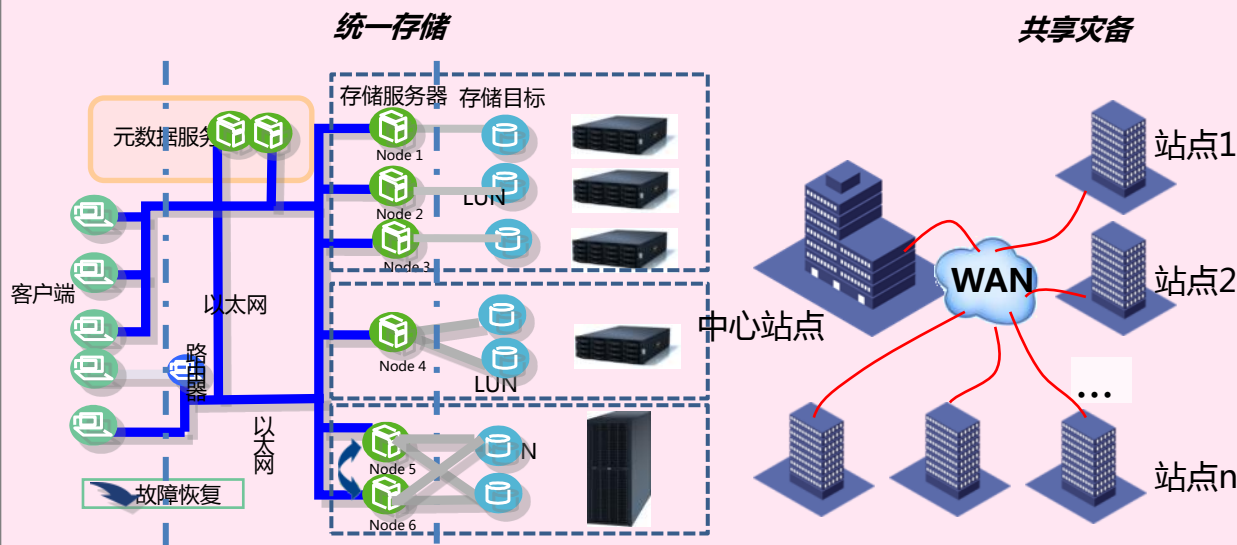
为传统应用、大数据分析提供统一存储空间

# 宏杉科技数字图书馆存储架构



- 容量管理
- 报警管理
- 统计报表
- 业务管理
- 性能统计
- 设备拓扑
- 业务拓扑
- 全局资源
- 接口开放
- 统一配置

智能数据管理平台IDSM



多种解决方案  
架构灵活



接口卡



MS5000/MS3000/MS2000



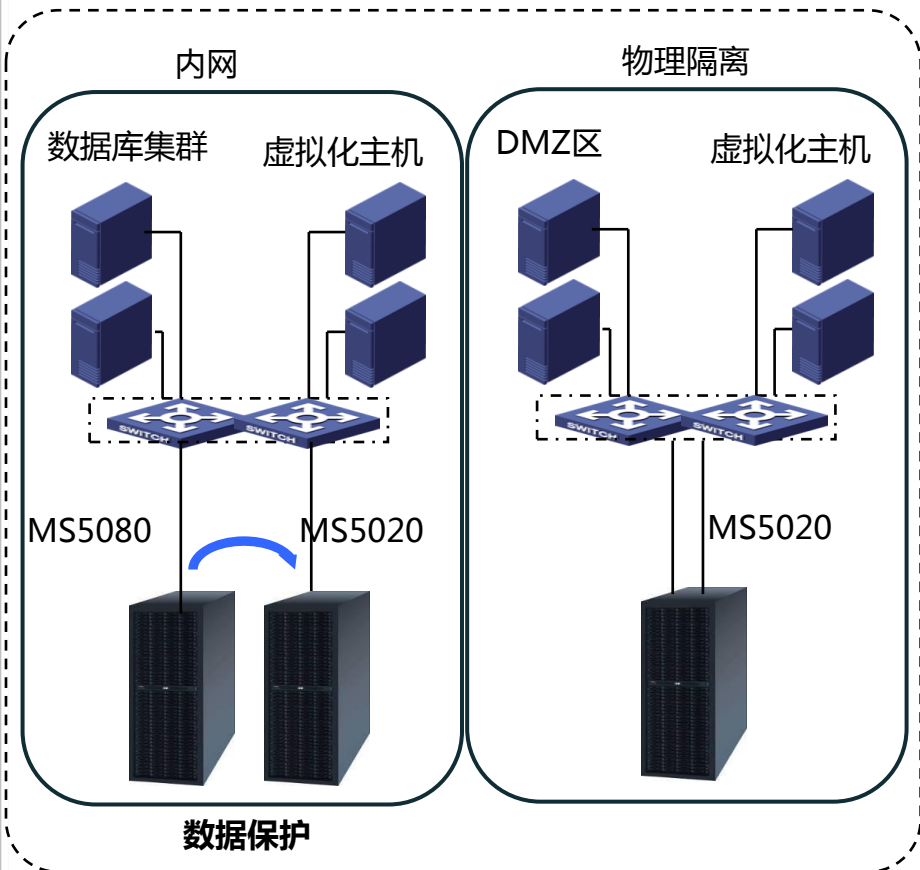
集群NAS

复制和镜像

快照

ICMT

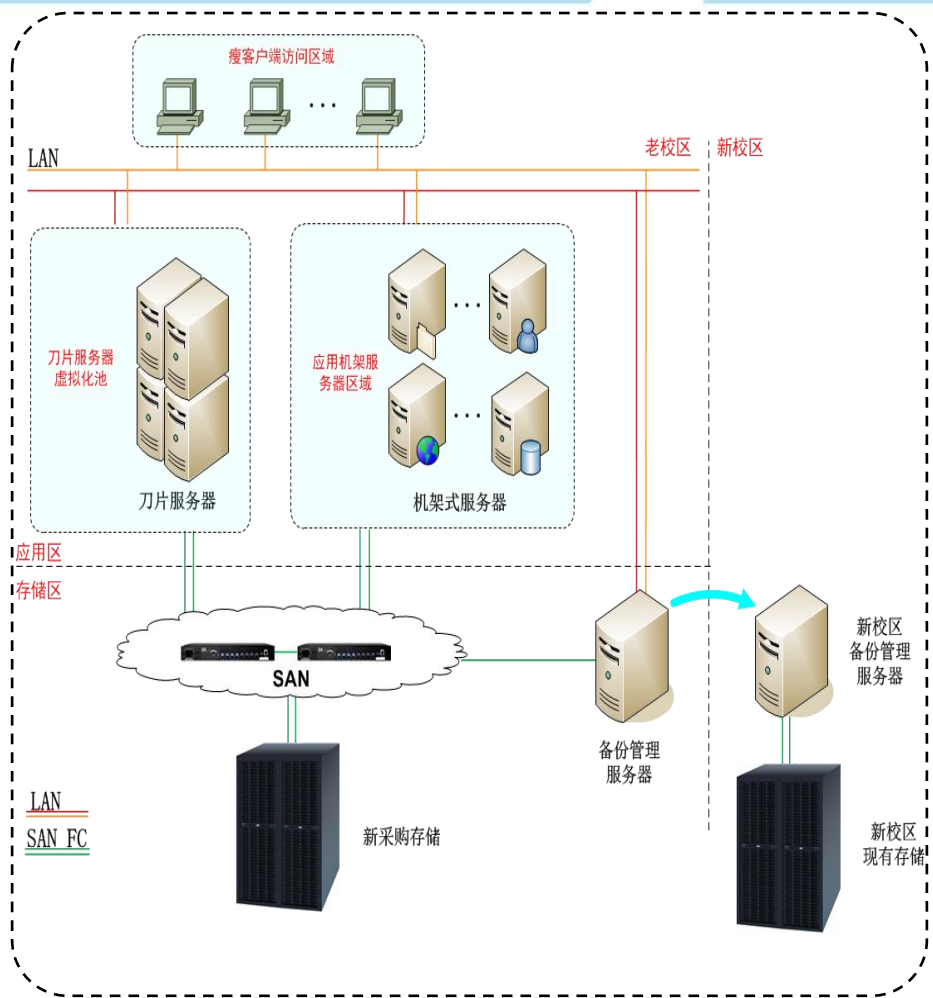
丰富的数据中心  
存储产品



- **主要应用**
  - 应用与教育部教育服务与监管体系基础设施建设项目（金教数据中心），为22个信息系统、14个主题数据库提供存储支持
- **面临的挑战**
  - 优异的存储性能，支撑大量业务正常运行
  - 完善的数据保护手段，保证数据安全
  - 自主可控，开放管理接口，实现统一网管
- **产品和方案**
  - 1台MS5080，2台MS5020搭建集中存储平台，应用于不同的区域和系统
  - 配置快照复制软件提供连续数据保护解决方案
- **用户价值**
  - 成熟的产品和方案，保障系统快速上线
  - 丰富的数据管理功能，提高运维水平

宏杉科技存储系统，分别部署在清华机房和国家教育部，为22个信息系统、14个主题数据库提供集中存储支持，配置快照、复制软件提供连续数据保护。





## 主要应用

- 基于V基于VWware vSphere、CtriX桌面虚拟化的IT虚拟化基础架构平台

## 面临的挑战

- 大规模服务器虚拟化、大规模桌面虚拟化
- 业务连续性和数据安全性保障

## 产品和方案

- 宏杉基于Vmware vSphere5的存储解决方案，MS5040存储系统及配套管理软件

## 用户价值

- 弹性化的资源平台，能够快速满足任何突发系统需求
- 快速恢复故障应用系统，确保应用业务的连续性

共采用宏杉科技2台存储，一主一备搭建基于VWware vSphere、CtriX桌面虚拟化的IT虚拟化存储平台，并配套自动分层软件满足系统性能及可靠性要求，两台阵列之间通过复制软件实现数据的容灾备份。



# 部分教育科研存储应用案例

- 北京大学医学部
- 北京邮电大学
- 中国政法大学
- 北京市教委
- 北京航空航天大学
- 中国人民公安大学
- 北京师范大学
- 中国地质大学
- 山东大学
- 四川外国语学院
- 塔里木大学
- 教育部小金教工程
- 中国科学院
- 南京市图书馆
- 南京博物院
- 山东省农科院
- 青海师范大学
- 南京沿江教师进修学校
- 北京信息科技大学
- 东莞理工学院
- 中国科学院光电研究院
- 河北经贸大学
- 陕西科技大学
- 空军工程大学
- 第四军医大学
- 中国科学院自动化研究所
- 湖北省图书馆
- 浙江省图书馆
- 江宁电教馆
- 中国科学院云南天文台
- 西南石油大学
- 西安工业大学
- 西安交通大学
- 云南省司法警官学院
- 承德石油学院
- 浙江工商大学
- 广东外语艺术学校
- 咸阳师范学院
- 中国人民解放军边防学院



**自主创新，扬帆起航**