

第30届中国数据库学术会议



曙光公司简介

曙光公司
刘冠川

»» 与中国同步前行
Synchronization with China forward

不断促进中国自主科研实力提升
不断推动民族信息产业实力提升

»» 1979
改革开放

玻璃房子
的启示

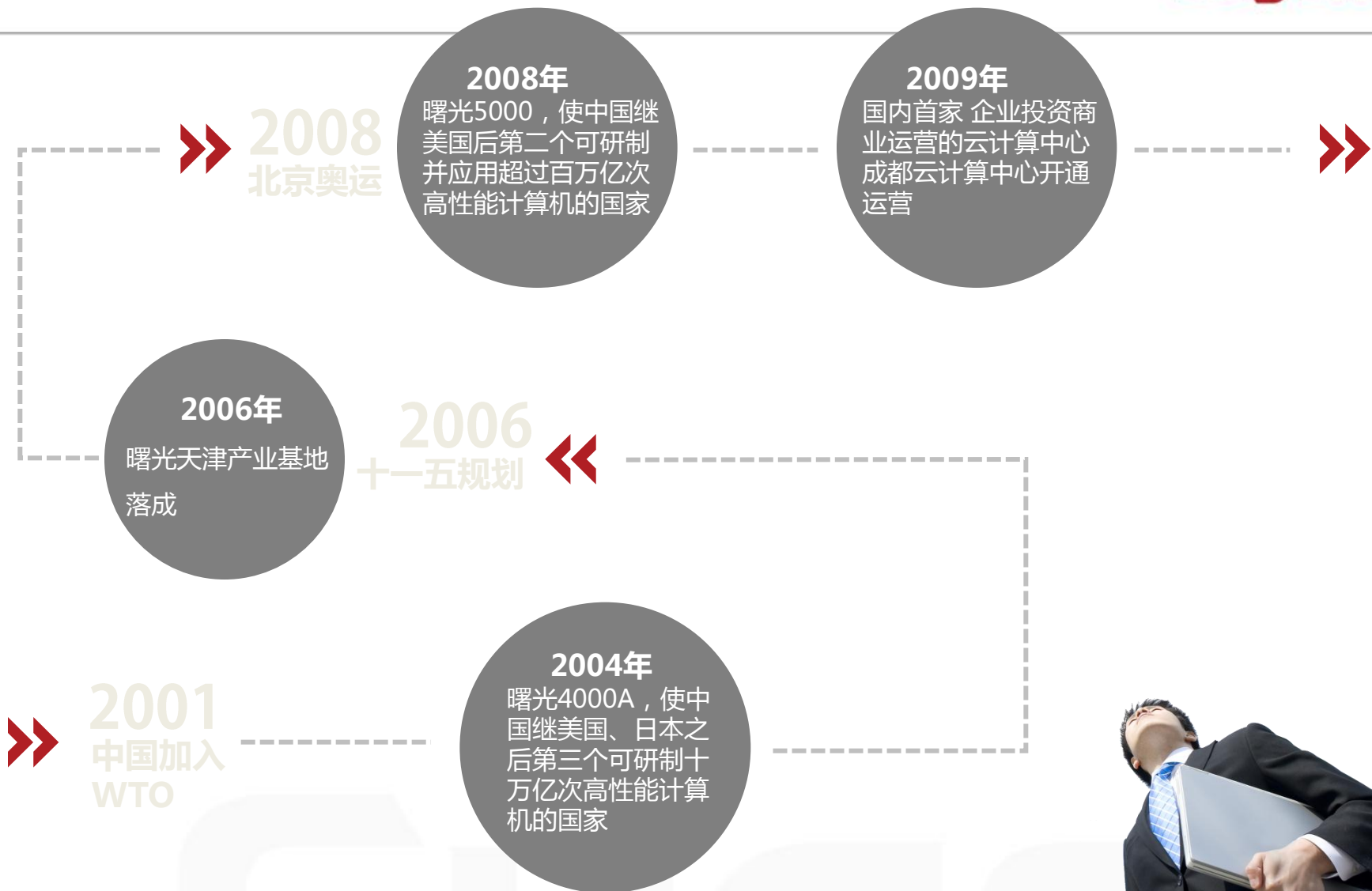
1994
中国科技
十大进展

1993年
中国第一台SMP
架构曙光一号计
算机问世

1996年
曙光公司成立

1997年
曙光1000A落户辽
河油田，我国高性
能计算机首次独立
进入市场





2010
上海世博

2010年

曙光星云，国内首
台全球第三台实测
性能超千万亿次高
性能计算机，位列
世界TOP500第二位

2011年

曙光换标，迈向
国际化

2011年

曙光星云在深圳
部署完成

2011年

曙光投资并运营
的中国物联网无
锡城市云计算中
心建成

2011
十二五规划

十二五期间，计划
在全国建成并投运
使用15个云计算中
心，云服务将成为
曙光支柱性业务模
块之一

2012
党的十八大

2012年

北京研发基地投入
使用

2012年

曙光连续四年蝉
联中国TOP100数
量份额第一





我们的业务布局

The layout of our business

曙光公司自1996年成立以来，始终倡导着“自主创新 服务中国”的品牌理念，目前已拥有1000+员工。

- +在北京，拥有1个研发中心
- +在天津，拥有1个制造物流中心
- +在盘锦，拥有1个生产基地
- +在北京、包头、南京、无锡、成都拥有全资的云计算公司
- +在香港拥有一个全资子公司
- +全国设有**29个办事处、分公司**，服务范围覆盖中国主要市和区域

我们的产品
Our products



服务器

+

存储

+

HPC

+

云计算

+

大数据

+

安全

在同行业产品中领先，拥有国内最全的服务器产品线





国家认定
企业技术中心

国家发展改革委 科技部
财政部 海关总署 国家税务总局

国家高性能工程技术研究中心

National High Performance Computer
Engineering Technology Research Center

中国电子工业标准化技术协会
高性能计算机标准工作委员会

High Performance Computer Standardization Committee
of Chinese Electronics Standardization Association

曙光信息产业(北京)有限公司
中关村科技园区海淀园
博士后工作站分站

中华人民共和国人力资源和社会保障部



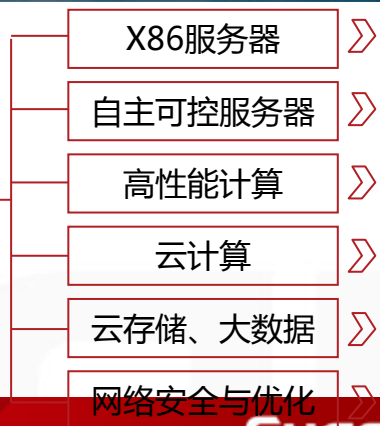
我们的研发

Our research and development

与众不同的研发资源和技术积累，支持国家重大项目的研发

- + 与中科院计算所、国家智能计算机研究开发中心、国家高性能计算机工程中心紧密合作
- + 研发投入经费占销售收入**10%**以上
- + 主持**26**项重大研发及产业化项目

研发布局

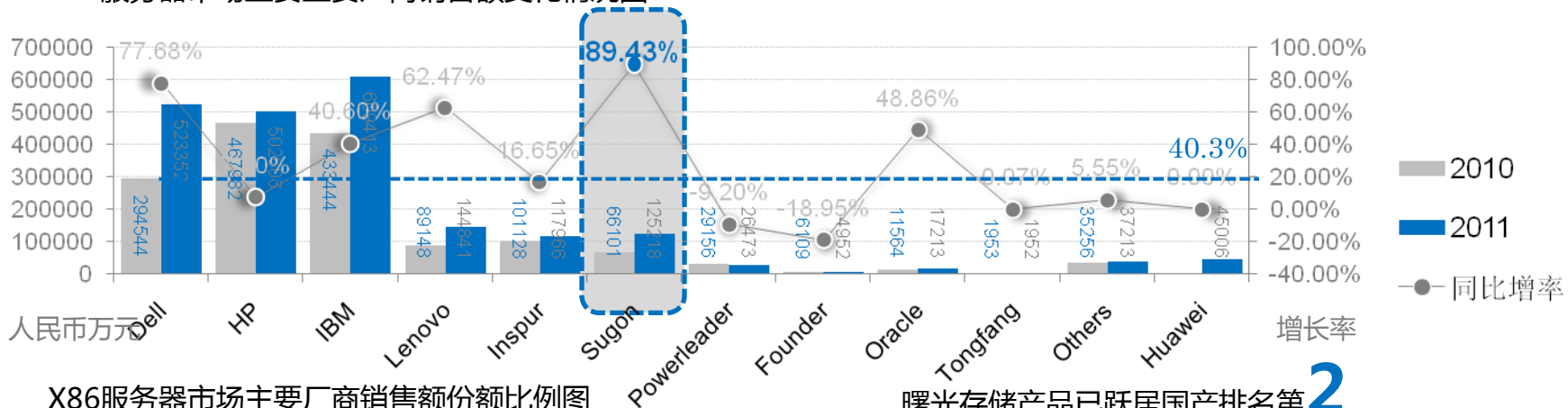




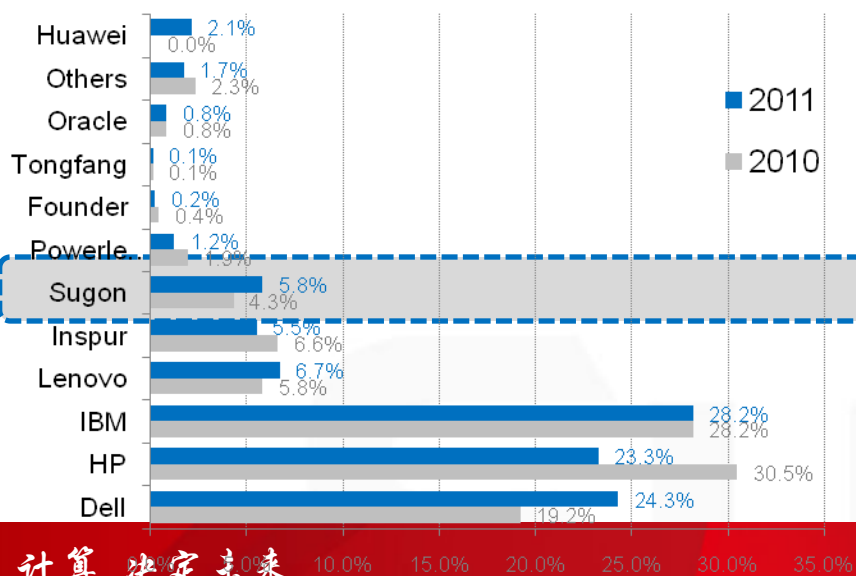
2011年业绩 2011 results

营收近15亿元

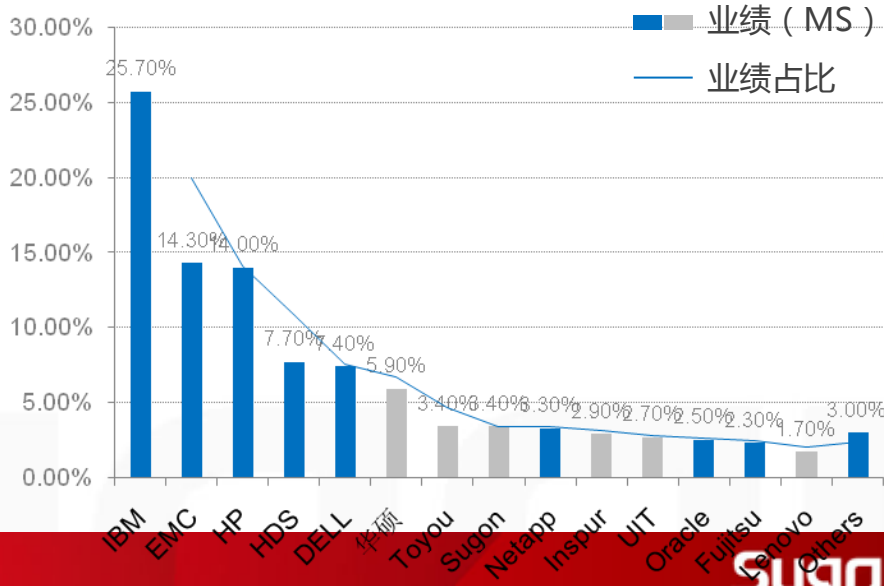
X86服务器市场主要厂商销售额变化情况图



X86服务器市场主要厂商销售额份额比例图



曙光存储产品已跃居国产排名第2

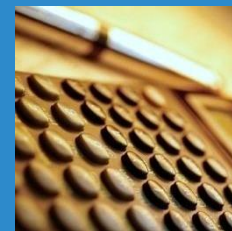


曙光公司教育行业发展

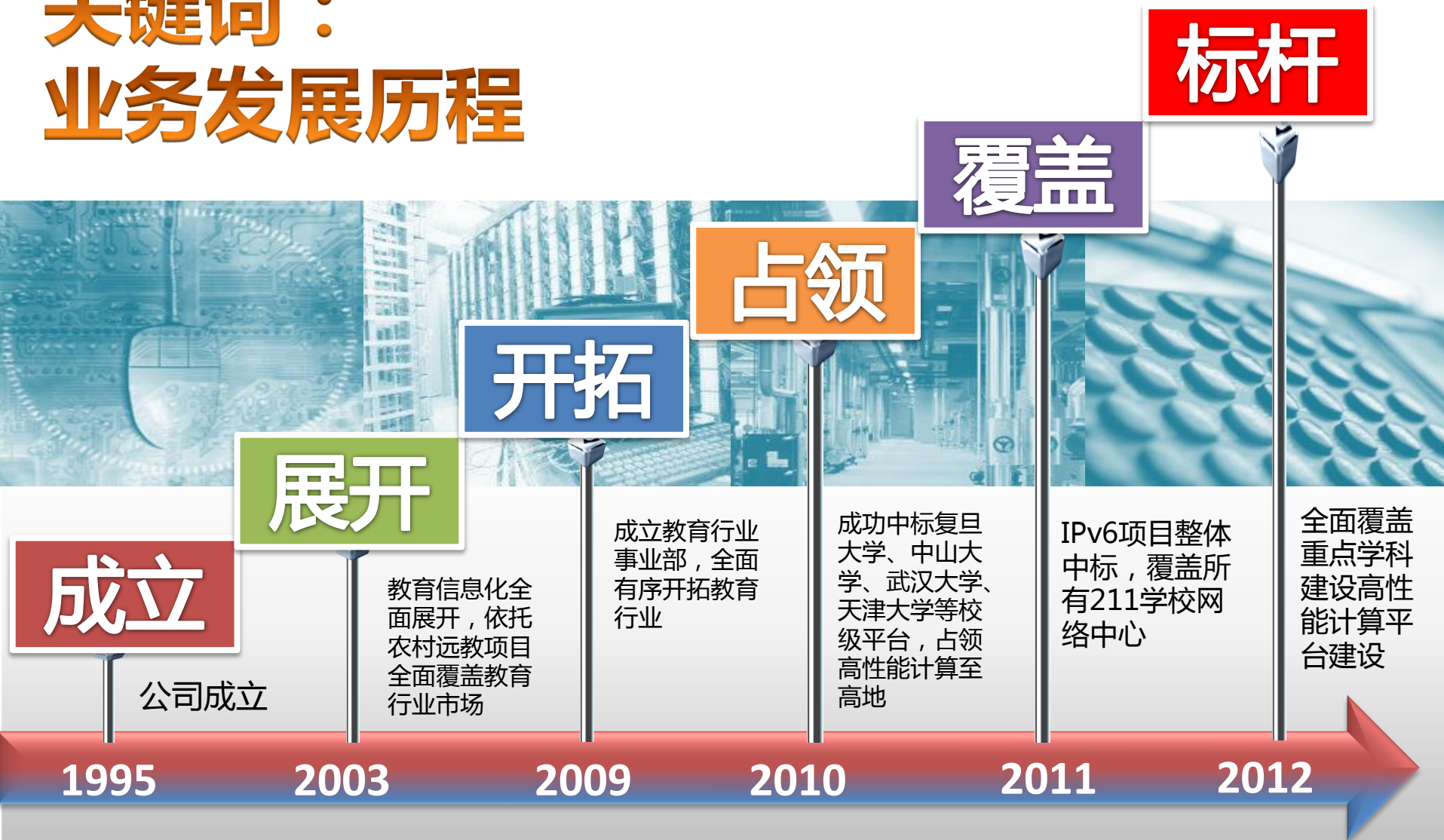
C o r p o r a t e

P r o f I l e

2 0 1 2



关键词： 业务发展历程



教育行业的成功范例



同济大学云计算平台



复旦大学高性能计算中心

中山大学

高性能计算GPU、Parastor200存储



ChinaGrid

重点学科研究平台



教育科研网建设

教育科研基础设施 IPv6技术升级和应用示范

- 利用已经CNGI主干网和CNGI高校驻地网等互联基础设施，全面推进校园网IPv6技术升级改造，
- 100余所学校的校园网升级到IPv6下一代互联网，用户达到100万人以上



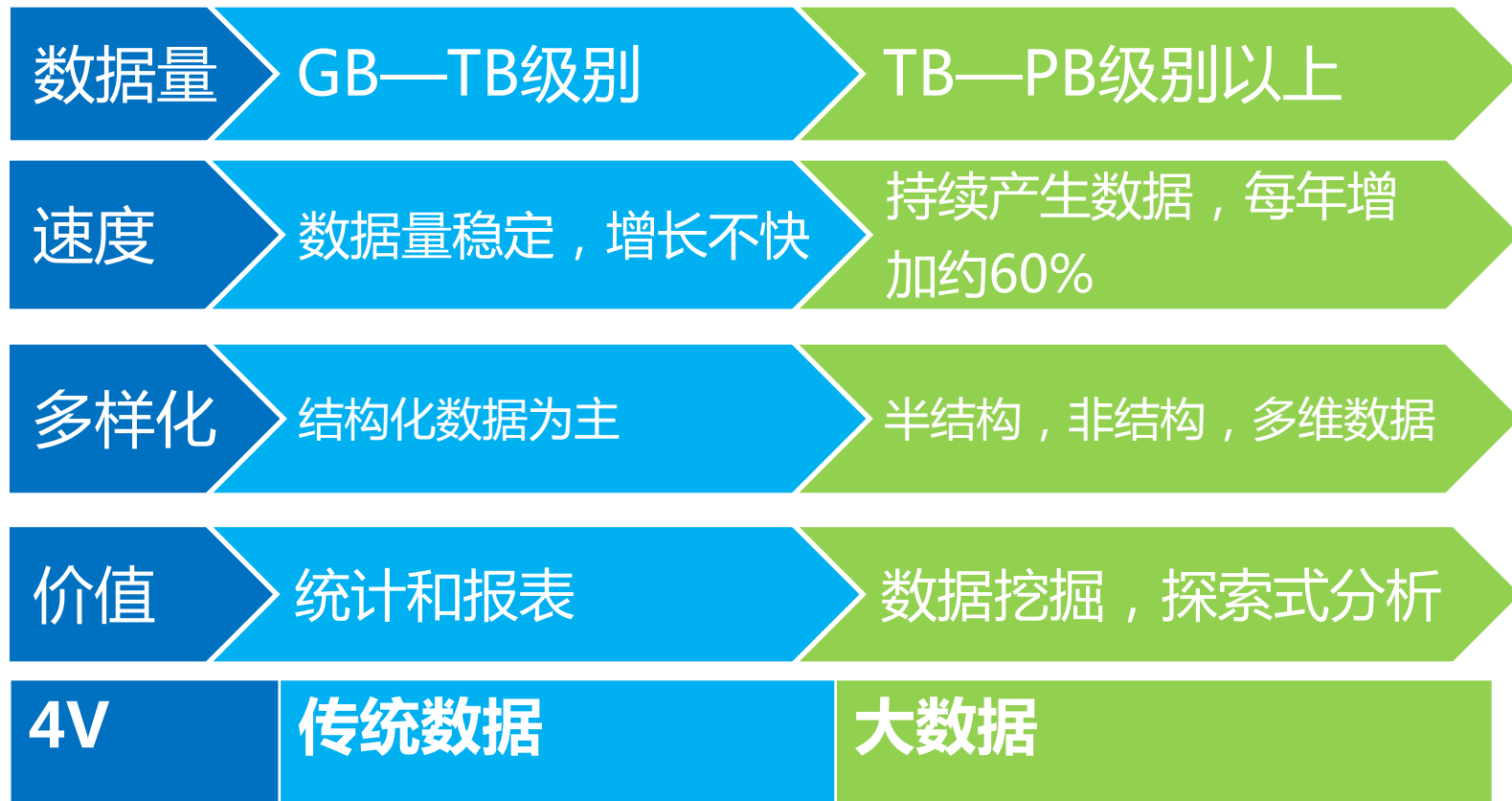
- 整合了全国几十所高校的大量网格资源，建立了资源共享、配置灵活、跨学科、跨地域的高效网格环境，部署了生物信息学、图像处理、计算流体力学、海量信息处理、大学课程在线等特色网格应用



从清华北大到偏远农村中小学，曙光为各类学校和教育机构提供提供定制的解决方案，让学生为将来在全球劳动力市场中竞争做好充分的准备

- + 为高校的顶级科研部门提供高性能计算解决方案
- + 助力校园网IPv6改造，使100余所学校、100万用户享受到IPv6下一代互联网
- + 助力远程教育工程，缓解不发达地区农村中小学教育资源短缺和师资不足，促进师资水平和教学质量提高
- + 100+所高等学校，1000+所中小学校

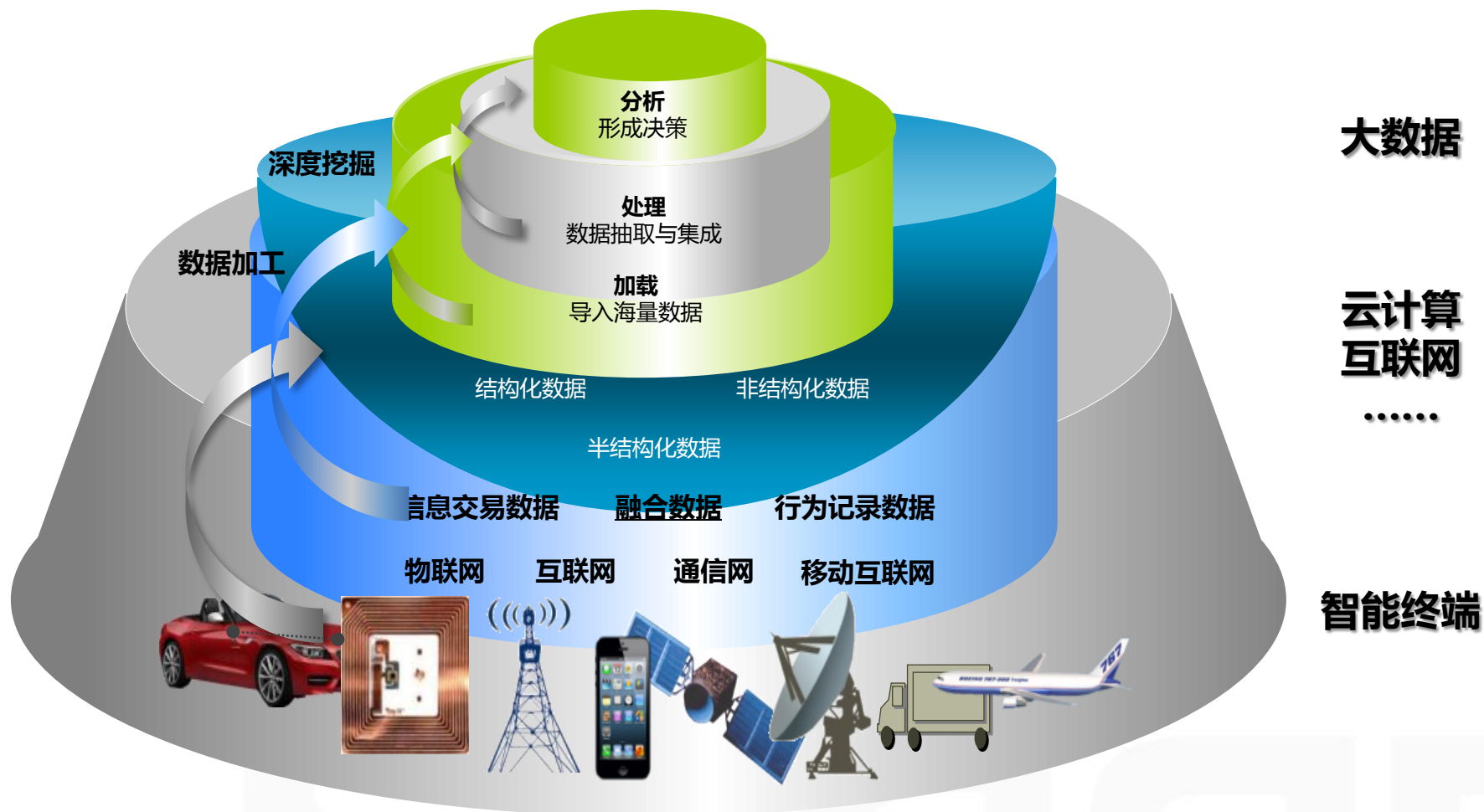
xData通用大数据处理平台

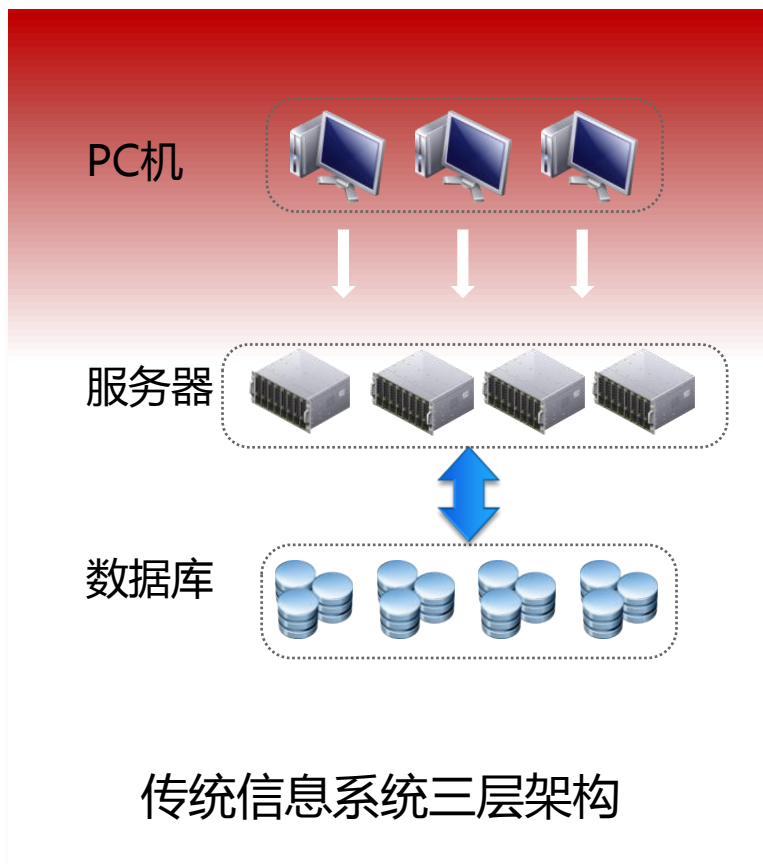


大数据是指数据集的大小超过了现有典型的数据库软件和工具处理能力，与此同时，及时捕捉，存储，聚合，管理这些大数据以及对数据深度分析的新技术和新能力，正在快速增长，正像预测芯片增长的摩尔定律一样。

---McKinsey Global Institute

大数据问题如何产生







- 对不同数据类型，提供统一的数据处理模式
 - 类JDBC的接口
 - 类MR执行机制，支持复杂处理

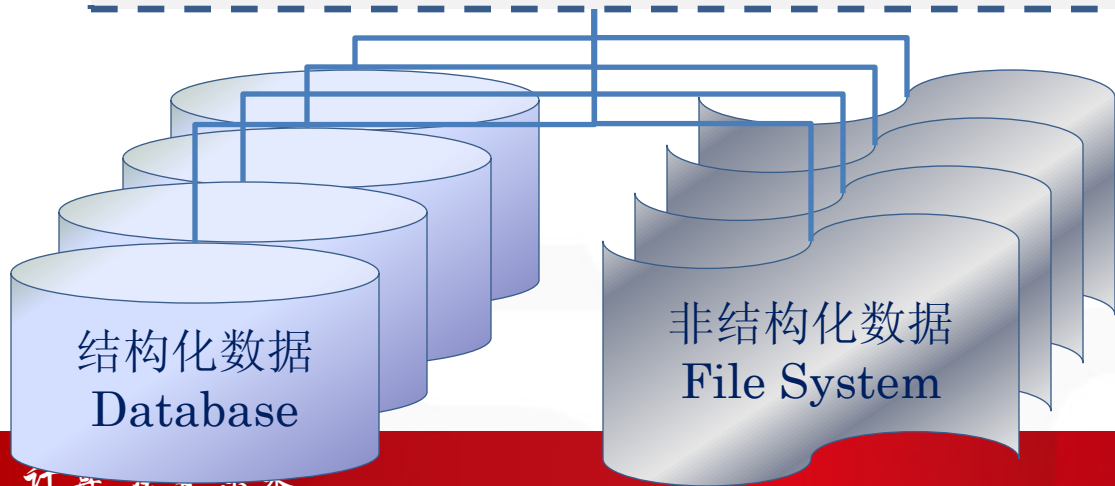
通用的大数据处理平台

数据节点采用SN结构+类MR的并行查询执行引擎

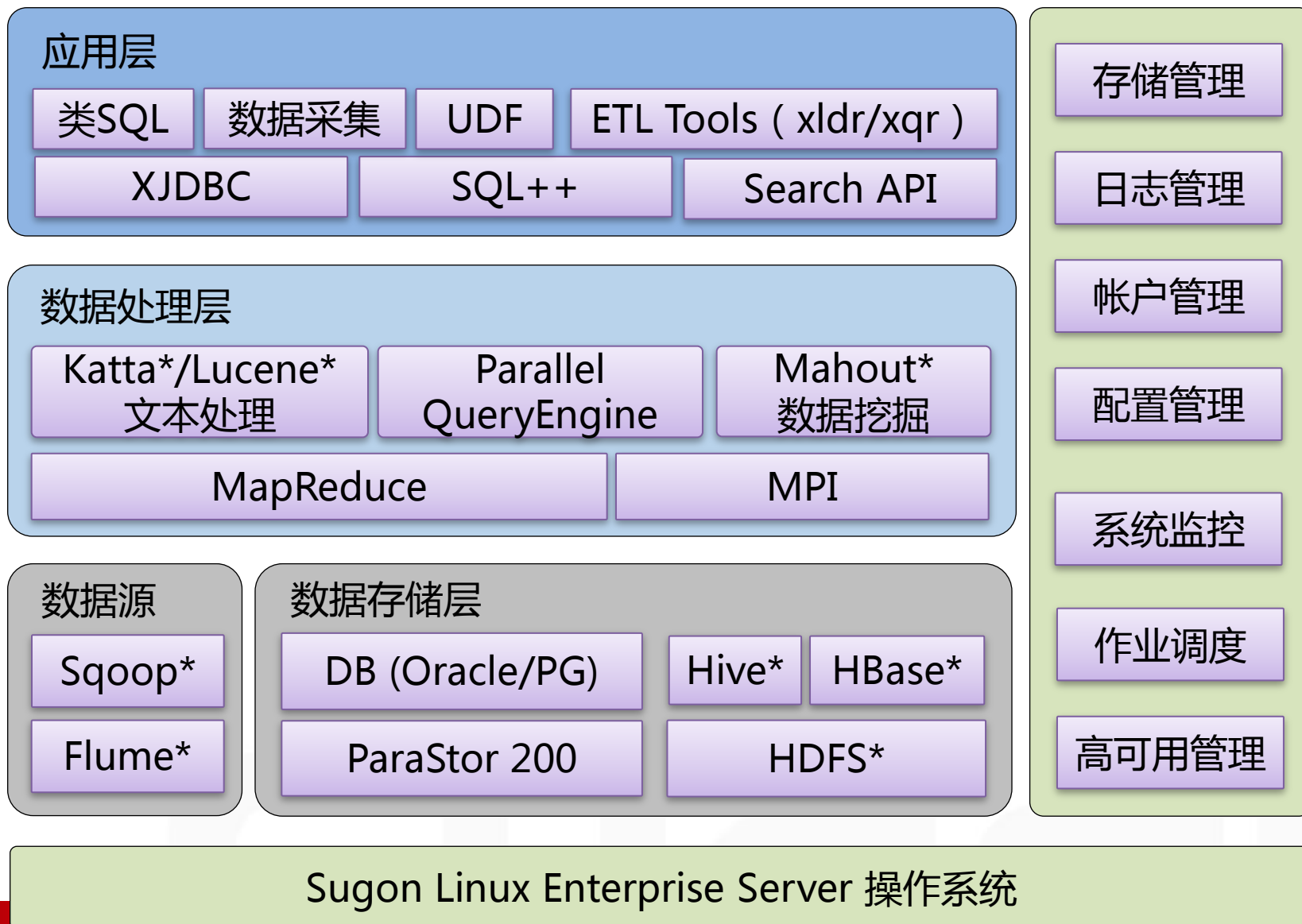
向上：对用户提供统一的数据处理接口

向下：管理不同类型的数据，提供通用的执行框架

里
訓
和



- 结构化数据：采用标准SQL语句
- 文本数据：采用lucene文本进行检索
- 音视频数据：数据库+文件系统
- 其他：支持自定义数据处理任务



XDATA大数据一体机

XData 曙光通用海量数据处理平台欢迎 admin

Sugon

XData

曙光通用海量
数据处理平台

用户名: admin 密码: *****

监控中心

- Hadoop集群
- Hadoop监控
- HDFS
- MapReduce
- Zookeeper
- HBase
- 资源监控
- 服务器监控
- 热图
- 告警
- 实时告警
- 历史告警
- 报表
- 自定义报表
- 预定义报表

Hadoop监控

组件信息

名称	状态	角色配置
Zookeeper	运行中	3个Zookeeper
HDFS	运行中	1个SecondaryNameNode, 1一个NameNode, 3个DataNode
MapReduce	运行中	1个JobTracker, 3个TaskTracker
Hbase	未运行	3个HMaster, 3个RegionServer
Hive	未配置	未配置

HDFS存储信息

Namenode

总容量: 100TB 已用容量: 10TB

10%

Datanode

总容量: 100TB 已用容量: 10TB

10%

MapReduce信息

Slots

总数: 100 已用: 10

10%

Jobs

Tasks

Map Tasks

Reduce Tasks

集群性能

平均负载

CPU使用情况

内存使用情况

环境检查

节点名	网络	操作系统	SSH无密...	Oracle用...
1 dbnode01	-	-	-	-
2 dbnode02	-	-	-	-
3 dbnode03	-	-	-	-
4 dbnode04	-	-	-	-
5 dbnode05	-	-	-	-
6 dbnode06	-	-	-	-
7 dbnode07	-	-	-	-
8 dbnode08	-	-	-	-
9 config	-	-	-	-
10 load01	-	-	-	-

显示第 1 条到 10 条记录, 共 14 条

部署说明

配置填写

环境检查

开始部署

- 支持多种数据分布策略

- 事实数据分片存储

- hash, range, 和round-robin分布
- 支持系统规模动态伸缩

- 配置数据全复制

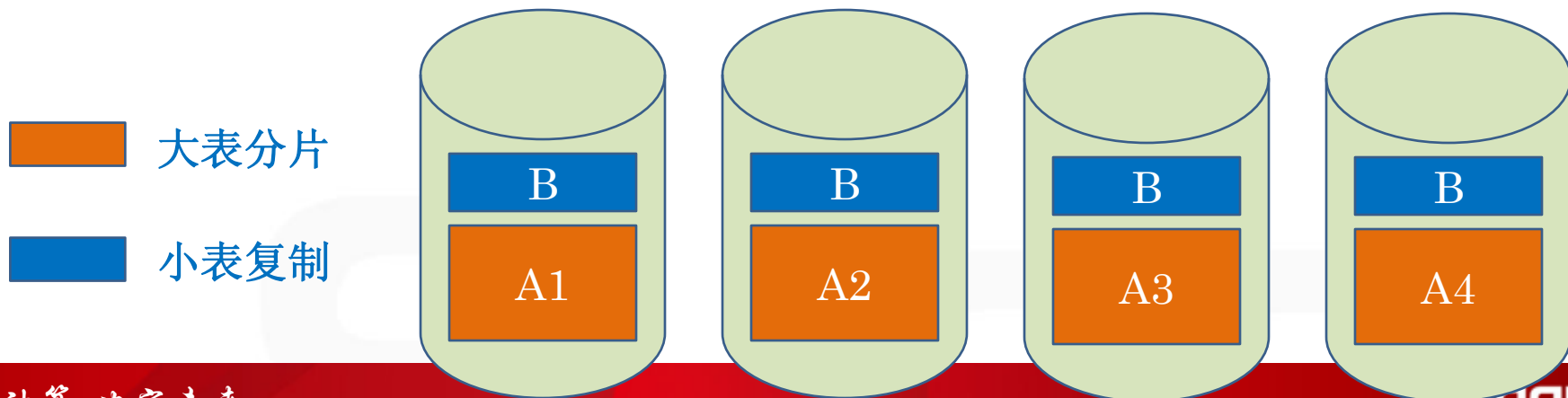
- 大量DML操作
- 一致性控制

- 提供数据双写的支持

- 支持本地双副本
- 数据写入异地容灾机制
- 异地数据批量迁移的支持

- 节点故障和恢复

- 数据不丢失
- 不影响当前执行的任务



数据分区的存储划分

● 存储划分

- 四个级别：记录，分区，节点，节点组
- 可扩展的虚分区技术，分区数是节点数的4~8倍（建议为数据节点的核数），支持数据节点规模的动态伸缩

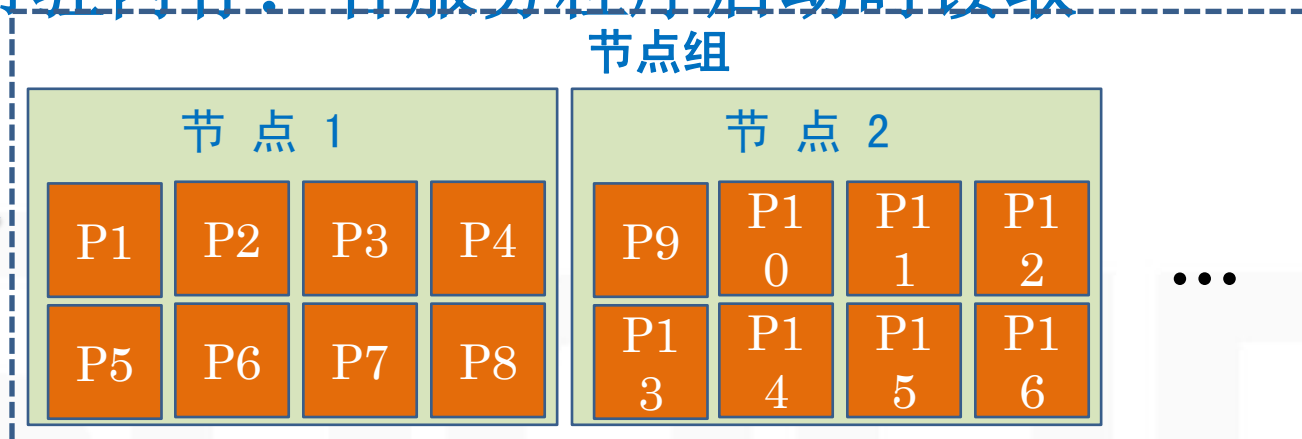
● 分区信息表示

- 一个节点组包含多个节点<nodegroup1, node_list>
- 一个节点包含多个分区<node1, part_list>

● 元数据信息常驻内存：各服务程序启动时读取

存储划分：

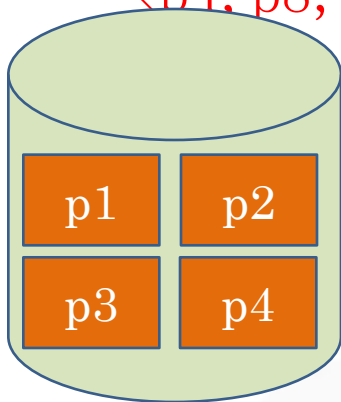
- ✓记录
- ✓分区
- ✓节点
- ✓节点组



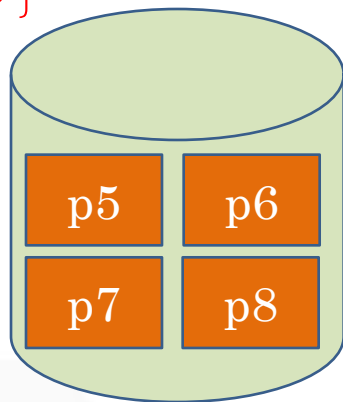
● 分区重新映射

- 最小化数据迁移，分区整体迁移
- 迁移完成，修改分区节点映射表

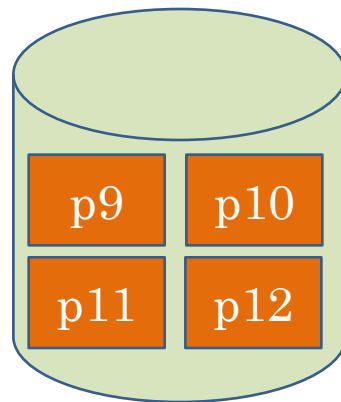
$\{\langle p1, p2, p3, p4 \rangle, \langle p5, p6, p7, p8 \rangle, \langle p9, p10, p11, p12 \rangle\}$
 $\Rightarrow \{\langle p1, p2, p3 \rangle, \langle p5, p6, p7 \rangle, \langle p9, p10, p11 \rangle,$
 $\langle p4, p8, p12 \rangle\}$



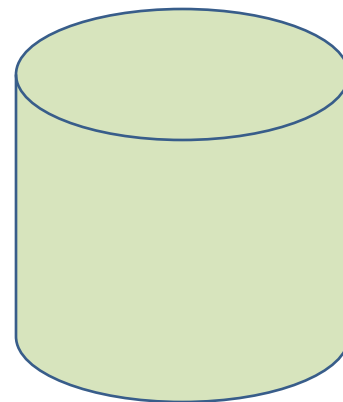
节点1



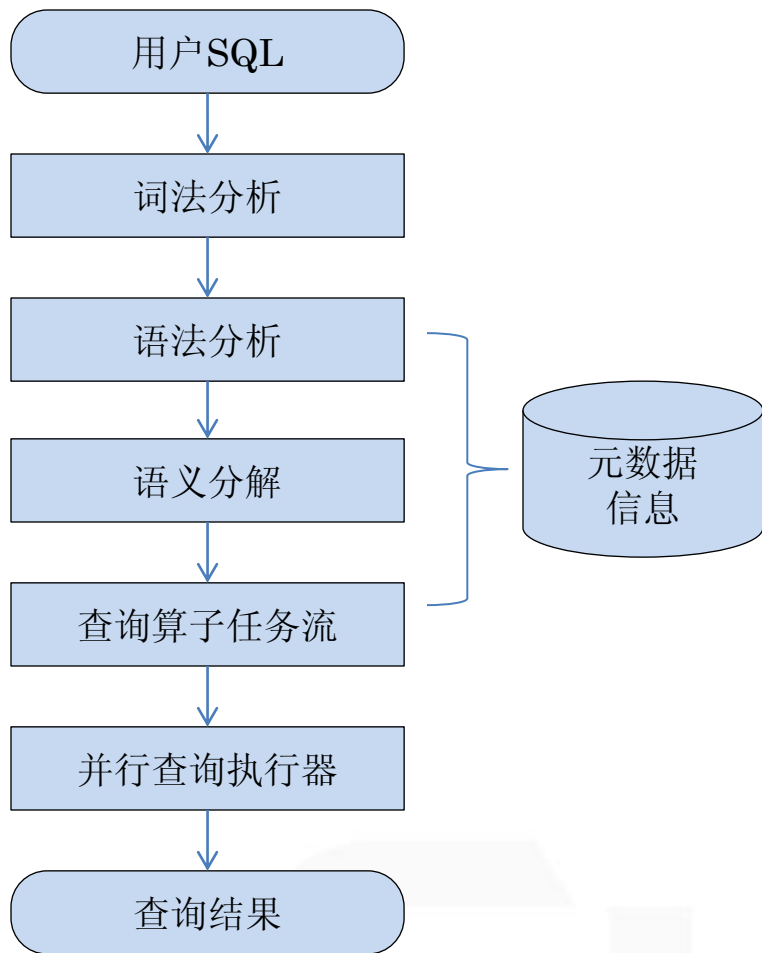
节点2



节点3



新增节点



●词法分析

- 采用LEX分解SQL语句

●语法分析

- 将查询语句按照关键字分成拆分成标准结构
- 数据分布分析

●语义分解

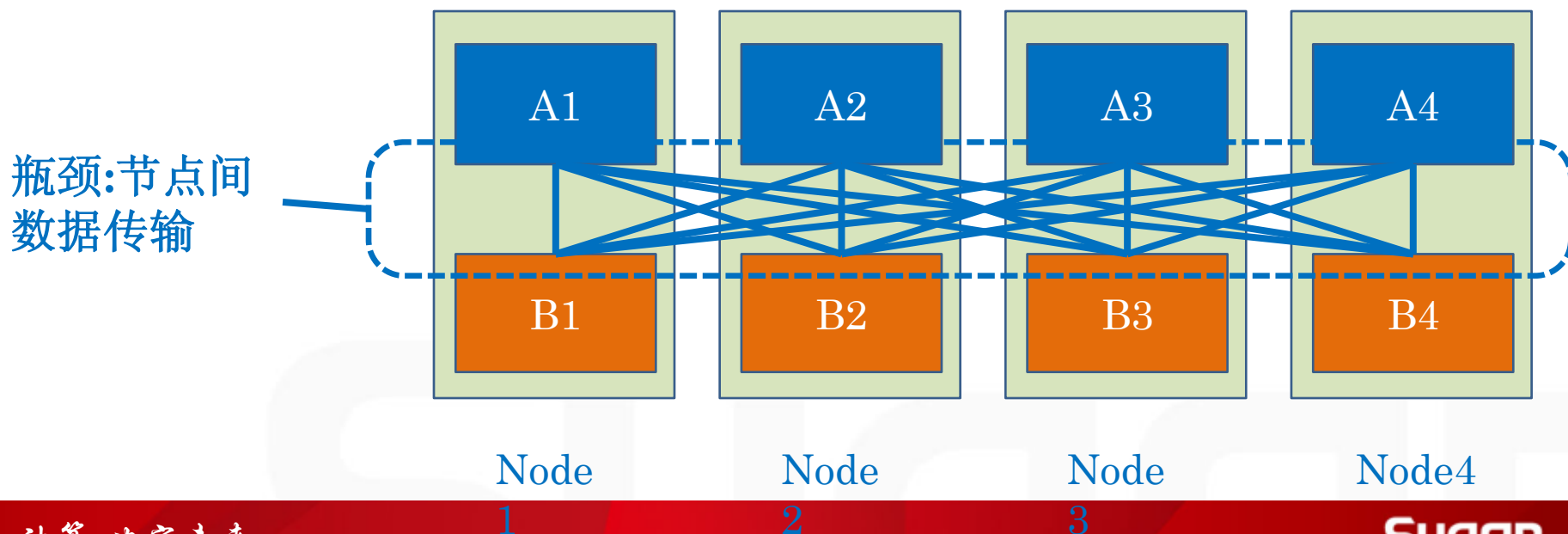
- 按语义拆分成并行查询算子流
- 算子任务的执行节点
- 算子任务的数据流向

●查询算子任务流

- 查询执行计划，表达了查询处理的逻辑顺序
- 作为并行查询执行器的输入

- 大表关联查询是海量数据处理的难点，也是提高核心竞争力的关键

- 大量数据交互，对磁盘读写和网络传输性能要求高
- 应用广泛：大量关联分析类查询和数据挖掘类应用



用户自定义查询任务流

- 支持用户自定义查询任务流，适合广泛的应用需求

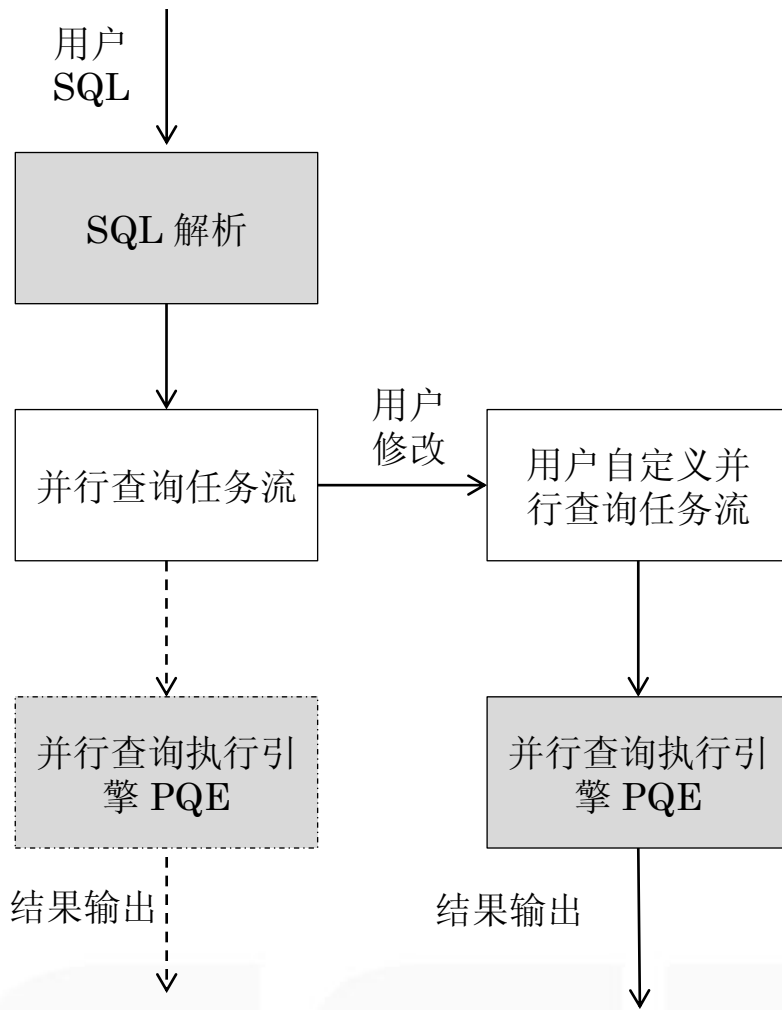
- 自定义查询类任务

- SQL语句查询任务节点
- 支持任何复杂的关联和嵌套查询处理
- 直接支持自定义任务节点

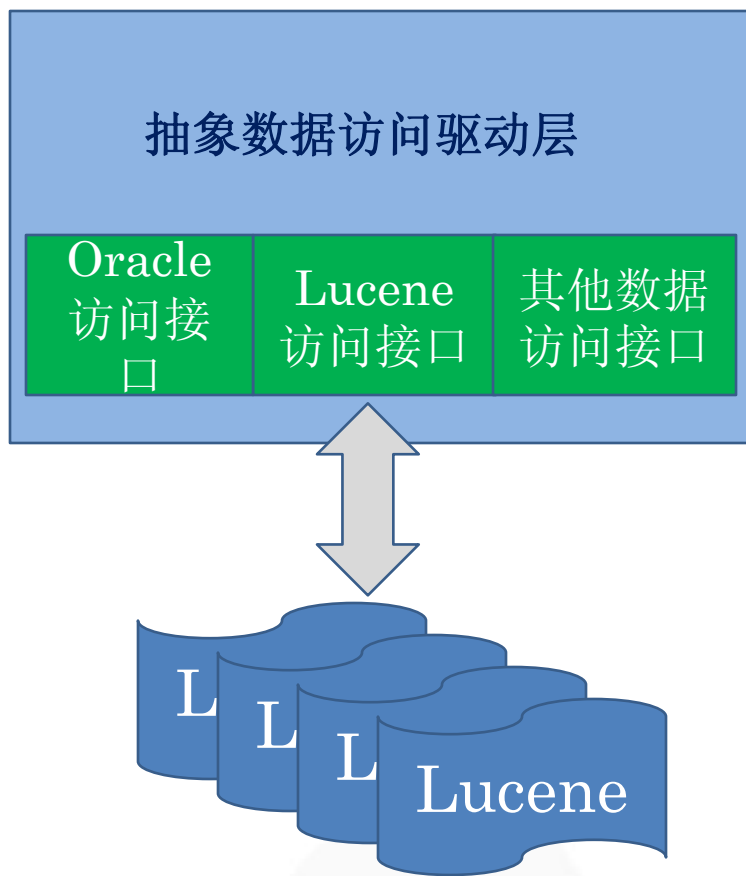
- 自定义挖掘类任务

- 自定义数据集操作函数
- 支持嵌入各种数据挖掘算法
- 需要操作函数或者挖掘算法的库函数支持

- 可利用开源算法



修改查询任务，满足定制处理



- 抽象数据访问驱动层，支持加载不同类型数据的访问

- 文本检索

- 采用文本索引开源软件lucene
 - lucene文本索引访问接口
- 关联查询和分析(先后顺序)
 - 关键字检索+数据库查询
如查找含有特殊内容的数据库记录
 - 数据库查询+关键字检索
如查找某些数据库记录的详细文本内容

- 音视频数据检索

- 通用的框架设计，可以装载不同数据类型的访问驱动

- 实现对SQL/Mapreduce的统一处理

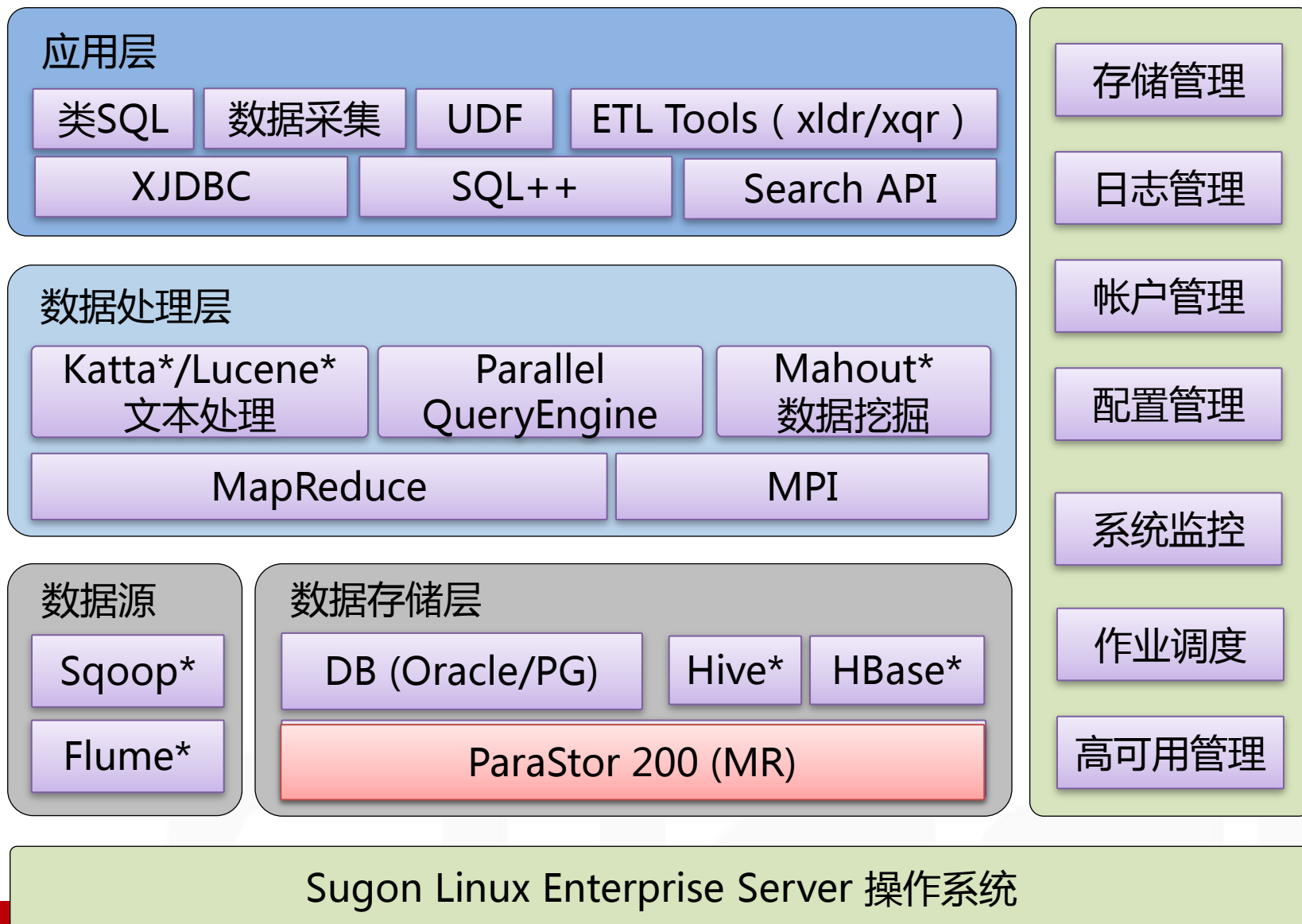
- 支持MR任务预定义，简化的用户请求
- 采用类SQL的方式，实现SQL/MR任务的级联和关联等复杂数据处理

HDFS是大数据
据唯一选择

No



xData + Parastor完美结合



更稳定

更高易用性

企业级软硬件优化

曙光一体化大数据方案

完备集群监控管理功能

更成熟的全活索引集群

ParaStor200 集群存储系统

简单：4类基本组件

- 应用控制器：提供访问接口
- 索引控制器：数据管理
- 数据控制器：数据存储
- 管理控制器：集群管理和监控

高性能：

文件在集群内均匀分布，并行读写
支持千兆、万兆和IB网络环境

高可用：

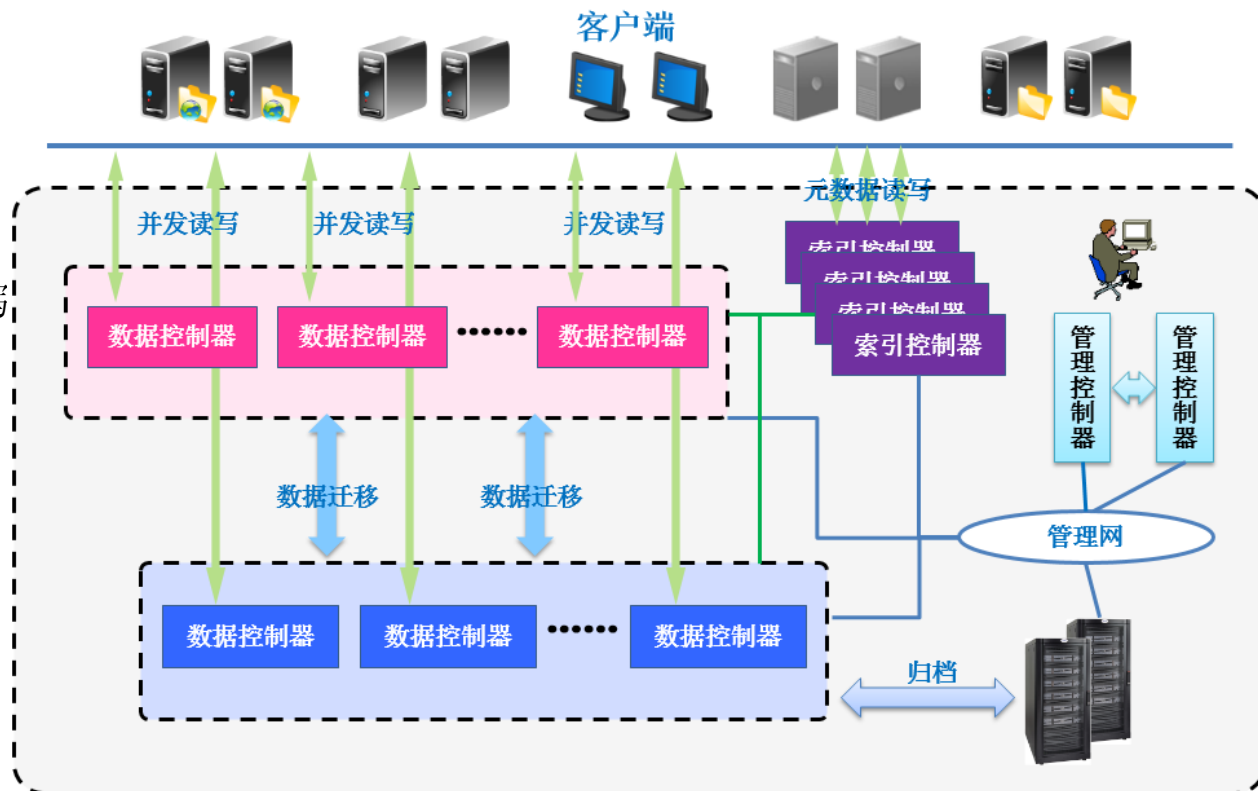
节点、磁盘、网络多方位冗余
系统内无单一故障点

高可扩展：

在线线性扩展
扩展后存储容量自动均衡

易用性：

全局统一命名空间
策略化的分级存储
图形化的管理和监控界面



系统特点及优势

中科曙光
Sugon



单系统容量
16PB!
目前国内最大!

国家超算中心深圳中心
国家超算中心深圳中心

“八仙过海”——丰富的访问接口

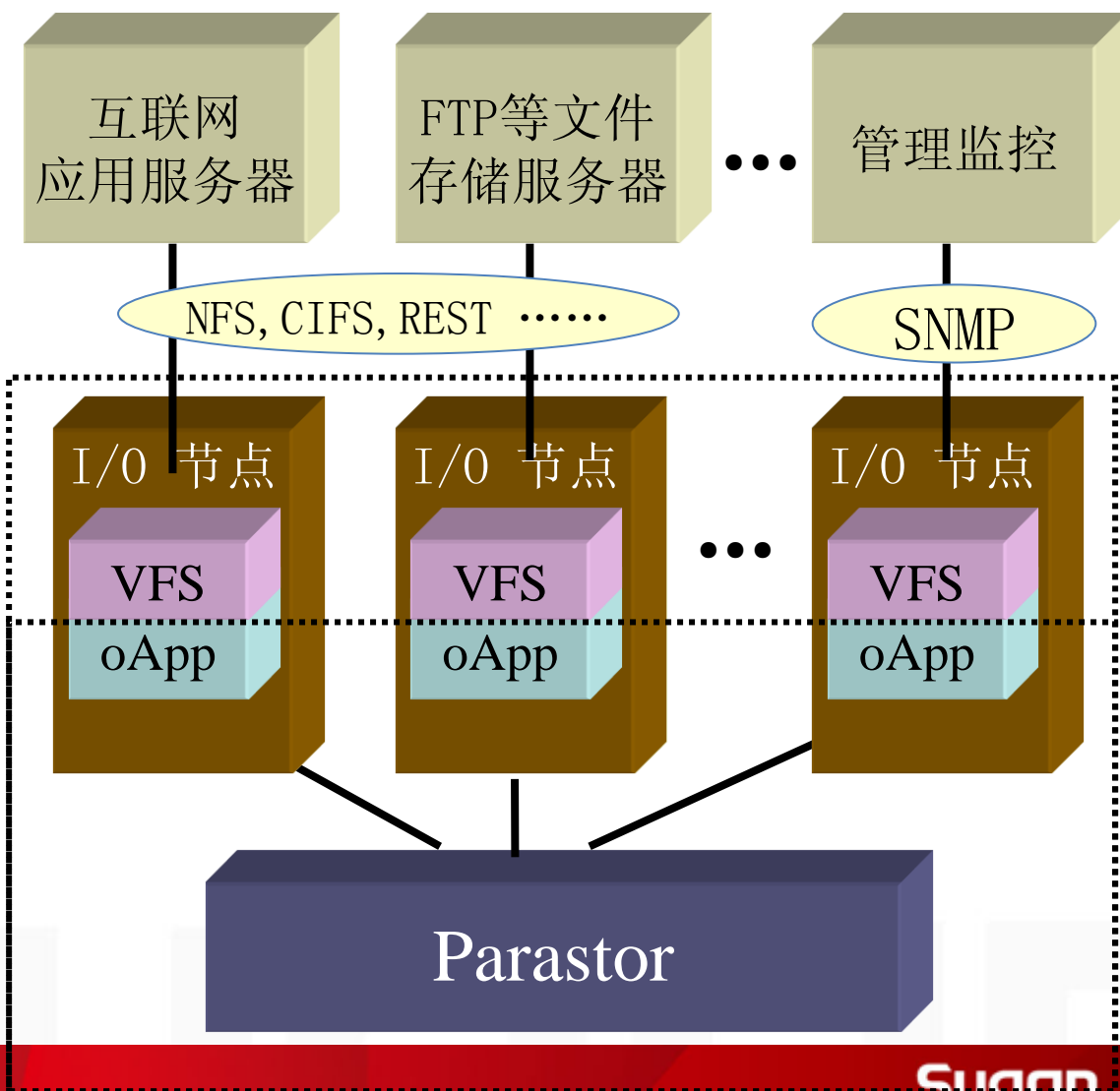
丰富的接口

访问接口

- 专有Linux、Windows内核态访问接口
- 标准NFS, CIFS接口
- POSIX API
- MapReduce编程接口
- REST编程接口
- SOAP编程接口
- SNMP接口

网络接口

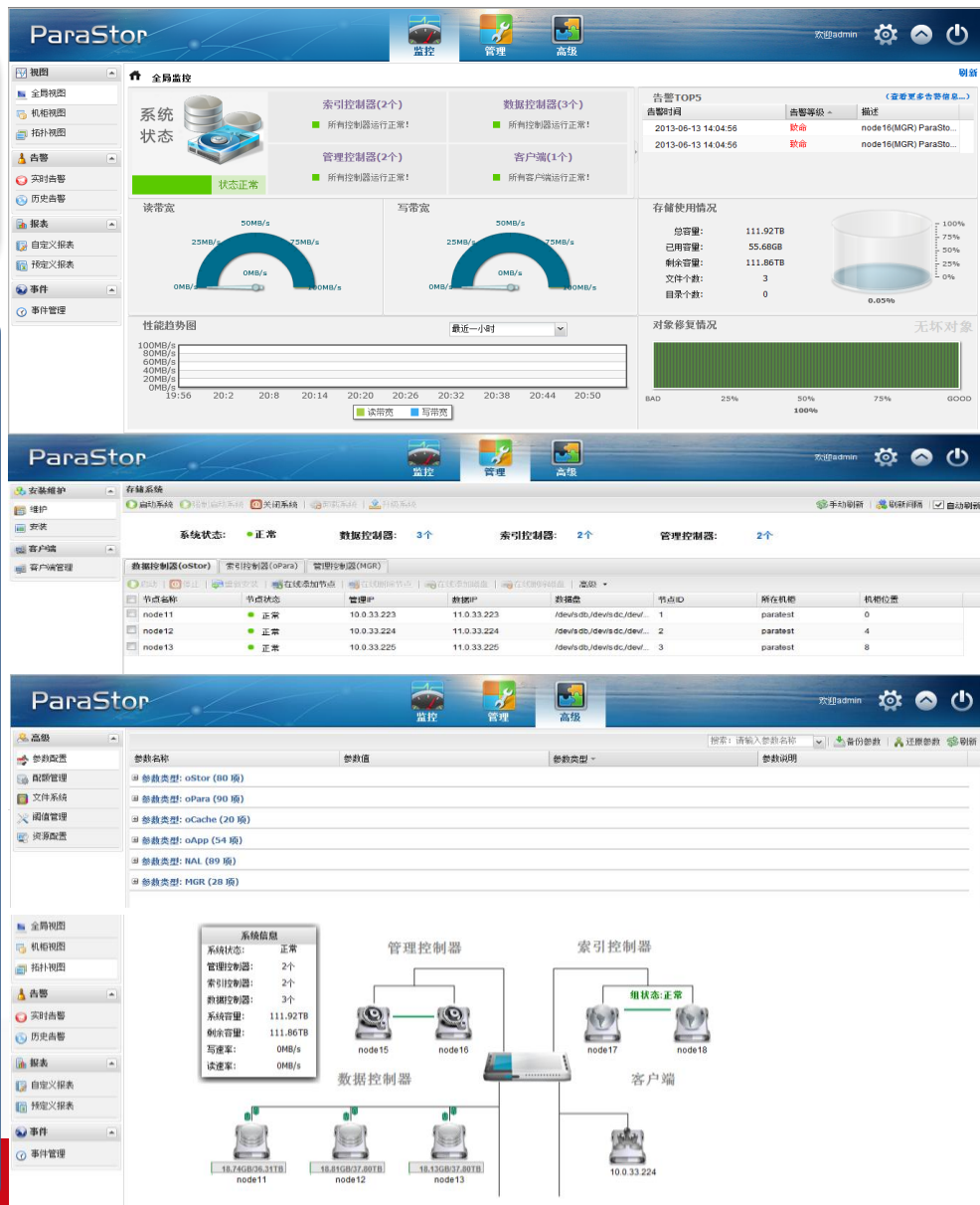
- 20Gb/40Gb/56Gb IB
- 10Gb/1Gb以太网
- 支持负载均衡和冗余



ParaStor并行存储管理系统是曙光公司专为ParaStor系列产品开发的统一监控管理平台。直观易懂的图形界面方便用户管理和监控系统的软硬件资源。

管理功能

- **监控系统**：网络、节点磁盘、内存、RAID卡状态监控，节点服务状态监控，系统故障告警，运维报表，管理事件记录。
- **系统管理**：系统的启停、卸载、升级，异常情况下进行强制启动。客户端授权、挂载及状态管理，管理控制器、索引控制器和数据控制器的增删、启停以及更换。
- **高级管理**：配额管理，文件系统创建、删除、配置，在线参数配置，阈值管理，资源配置。



ParaStor



监控



管理



高级

欢迎admin



- 视图
 - 全局视图
 - 机柜视图
 - 拓扑视图
- 告警
 - 实时告警
 - 历史告警
- 报表
 - 自定义报表
 - 预定义报表
- 事件
 - 事件管理

全局监控

系统状态



状态正常

索引控制器(2个)

所有控制器运行正常!

数据控制器(3个)

所有控制器运行正常!

管理控制器(2个)

所有控制器运行正常!

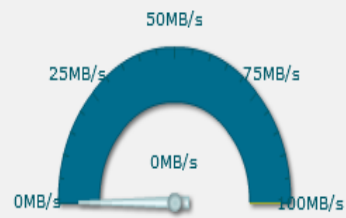
客户端(1个)

所有客户端运行正常!

读带宽



写带宽



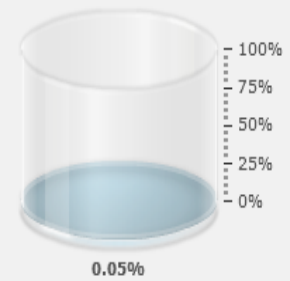
告警TOP5

(查看更多告警信息...)

告警时间	告警等级	描述
2013-06-13 14:04:56	致命	node16(MGR) ParaSto...
2013-06-13 14:04:56	致命	node16(MGR) ParaSto...

存储使用情况

总容量: 111.92TB
 已用容量: 55.68GB
 剩余容量: 111.86TB
 文件个数: 3
 目录个数: 0



性能趋势图

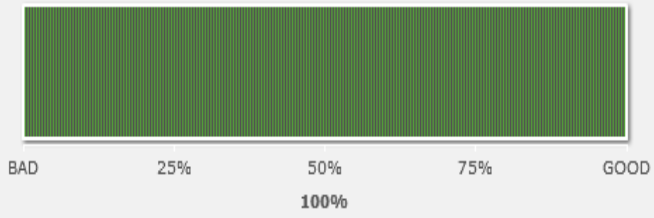
最近一小时

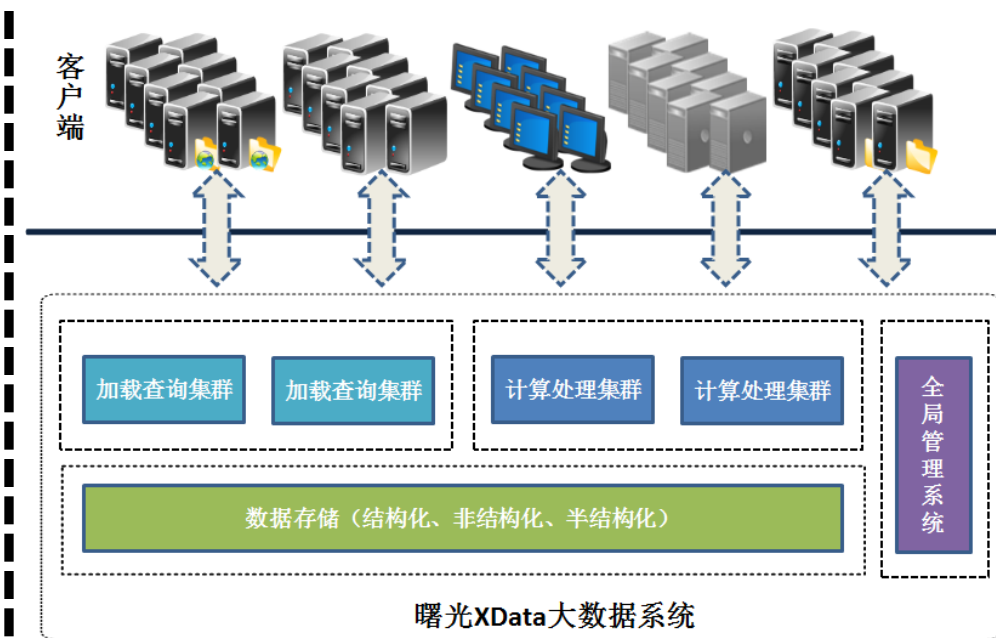


读带宽 写带宽

对象修复情况

无坏对象





XData大数据处理机是曙光公司立足于当前大数据处理的时代背景下，自主研发的通用海量数据处理平台，可广泛的应用在通信数据统计，互联网/移动互联网的日志和用户行为分析，物联网/传感器网络的数据监控和追踪分析，以及金融交易数据的离线统计和挖掘等众多领域。

谢谢