中科曙光
Sugon

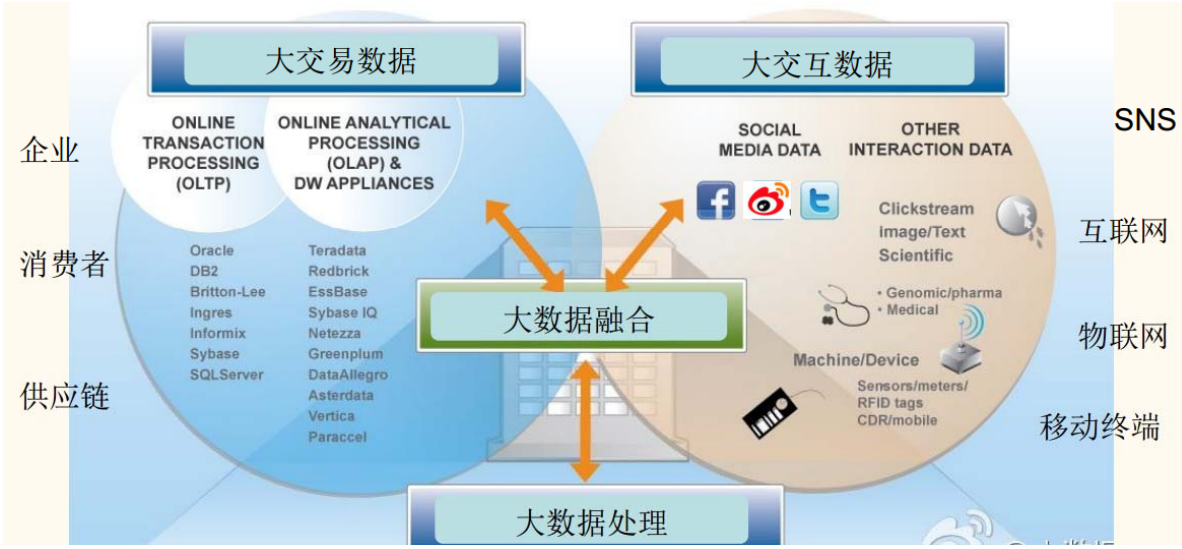
大数据系统安全技术

中科曙光

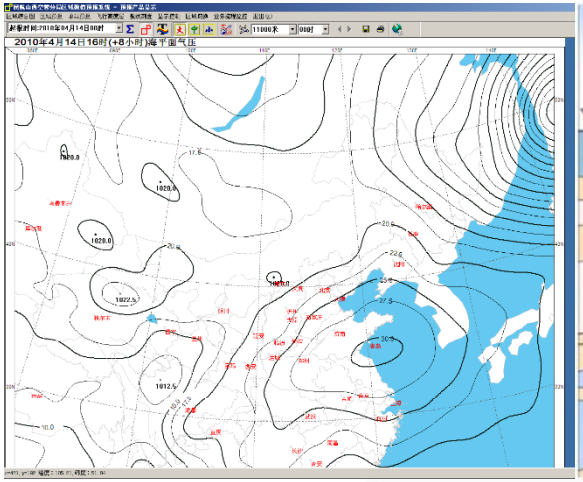
2014-05

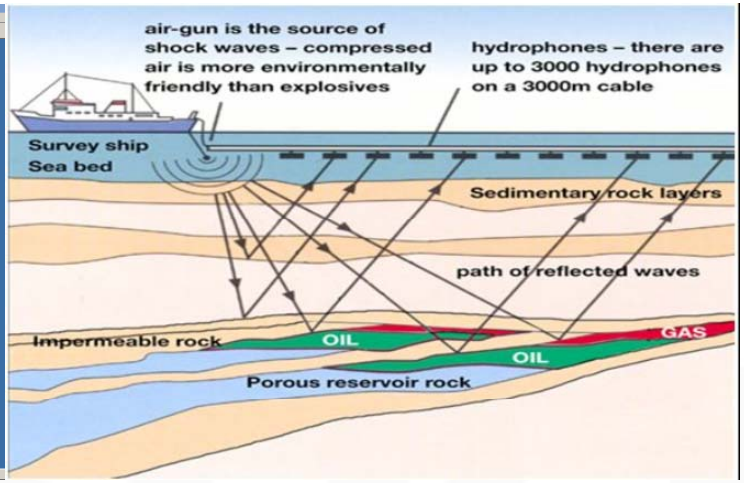
目录

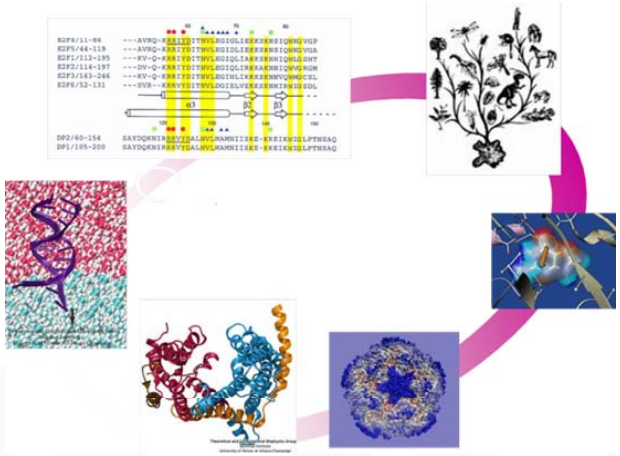
大数据无处不在


大交易数据
大交互数据
企业
消费者
供应链
ONLINE TRANSACTION PROCESSING (OLTP)
ONLINE ANALYTICAL PROCESSING (OLAP) & DW APPLIANCES
SOCIAL MEDIA DATA
OTHER INTERACTION DATA
Oracle
DB2
Britton-Lee
Ingres
Informix
Sybase
SQLServer
Teradata
Redbrick
EssBase
Sybase IQ
Netezza
Greenplum
DataAllegro
Asterdata
Vertica
Paraccel
Clickstream
image/Text
Scientific
• Genomic/pharma
• Medical
Machine/Device
Sensors/meters/
RFID tags
CDR/mobile
大数据融合
大数据处理
SNS
互联网
物联网
移动终端

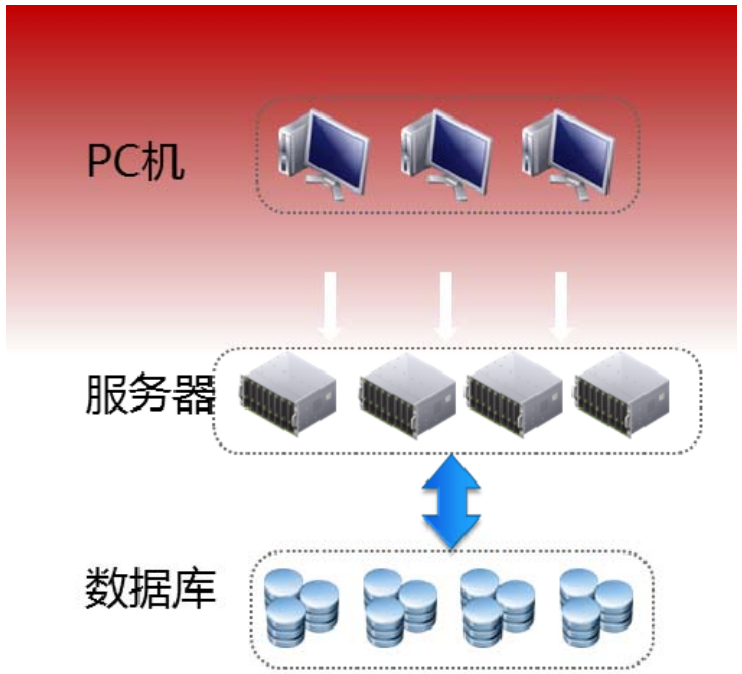

国金证券
SINOLINK SECURITIES
大数据时代的三大发展趋势和投资方向
智慧的商业
数据成为资产
行业垂直整合
泛互联网化
大数据
Hadoop
大融合数据
大交易数据
大行为数据
Oracle
DB2
Teradata
移动互联
社会化网络
车联网
云计算
www.gjzq.com.cn
@社会标TMT
DoMarketing.Org


2010年4月14日16时(+8小时)海平面气压


air-gun is the source of shock waves – compressed air is more environmentally friendly than explosives
hydrophones – there are up to 3000 hydrophones on a 3000m cable
Survey ship
Sea bed
Sedimentary rock layers
path of reflected waves
Impermeable rock
OIL
GAS
OIL
Porous reservoir rock



信息系统架构演变



传统信息系统三层架构

未来信息系统三层架构

大数据系统安全重要性

中科曙光 Sugon

数据爆发式增长，信息成为战略资产

- 大数据市场年增迅速，近5年平均增速50%以上
- 大数据技术影响到国家治理、企业决策和人民生活

大数据给信息安全提出了新的挑战

- 数据泄密影响重大：Saleforce，Google gtalk，CSDN，天涯等相继被曝用户数据泄漏
- 制约大数据业务的融合和应用发展

安全威胁大大提高，攻击者背景更加复杂

- 安全威胁的目标性、隐蔽性、破坏性都大大增加，攻击者的动机、目的、方法变得更加复杂
- 针对云计算和大数据应用的攻击成为新的攻击方向

大数据系统安全技术体系

中科曙光 Sugon

保证系统管理和运维安全

- 安全策略管理
- 系统安全审计
- 用户和权限管理
- 配置基线检查
- 漏洞和补丁管理

保证访问控制的安全

- 访问权限认证
- 流量和访问质量控制
- 用户访问行为监控
- 访问敏感信息告警、阻断和追踪

保证数据内容的安全

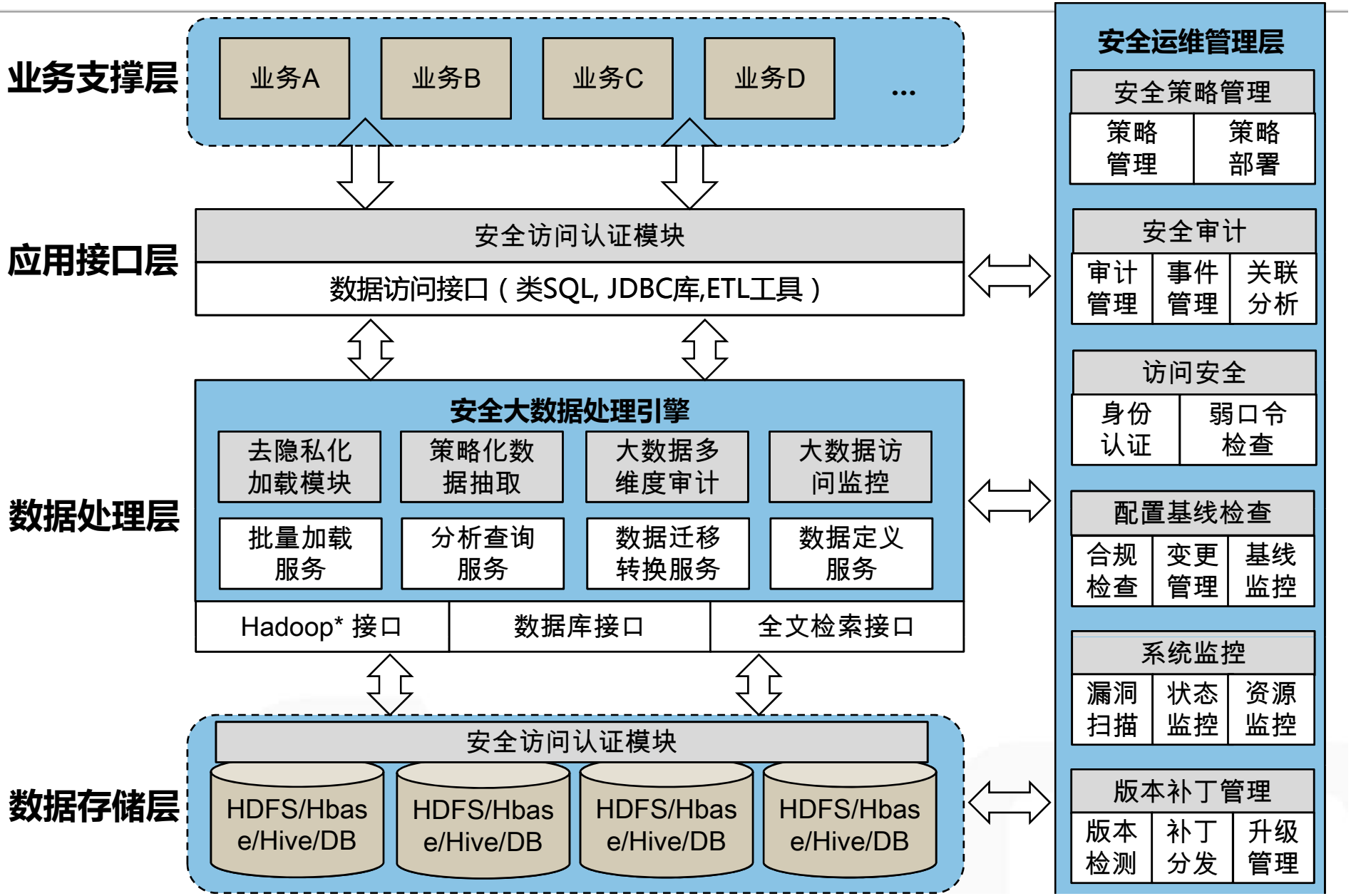
- 大数据去隐私化技术：数据加密，限制发布，数据失真
- 多维度审计技术：用户、数据对象、字段、敏感内容等审计

保证数据存储的安全

- 存储介质加密访问技术
- 文件加密存储技术
- 分布式存储分片加密和解密技术
- 数据备份和容灾

大数据安全处理平台系统架构

中科曙光 Sugon



大数据系统安全关键技术

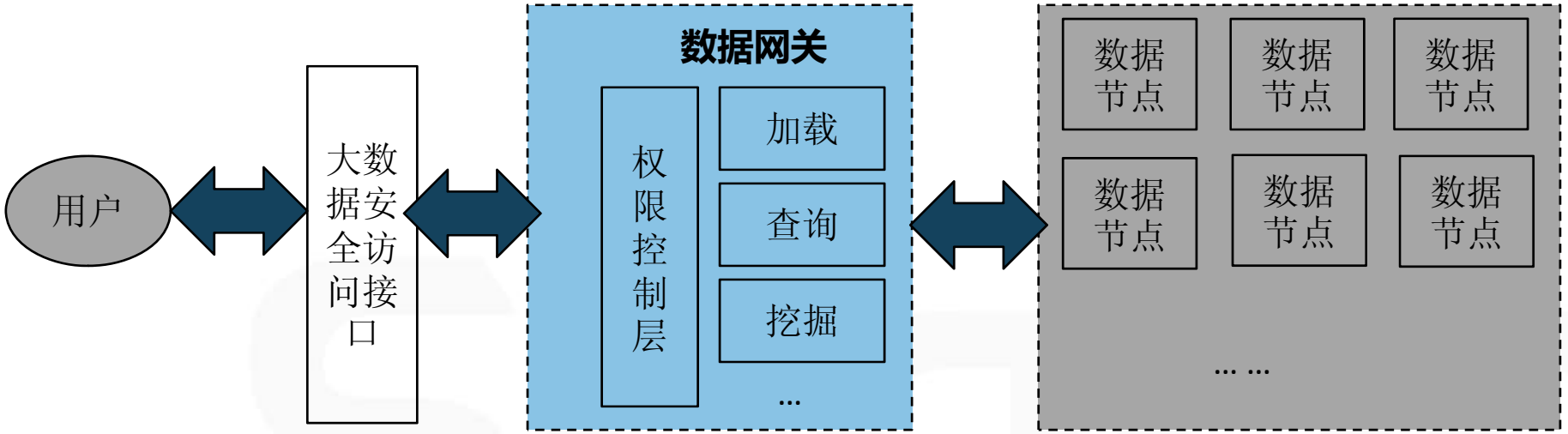
- 各类型数据存储和处理技术
- 大数据系统统一策略管理
- 配置基线检查和监控技术
- 大数据并行去隐私化技术
- 策略化抽取和集成技术
- 多维度大数据审计技术
- 访问监控和报警技术
- 访问行为追踪技术

目录

大数据系统访问安全

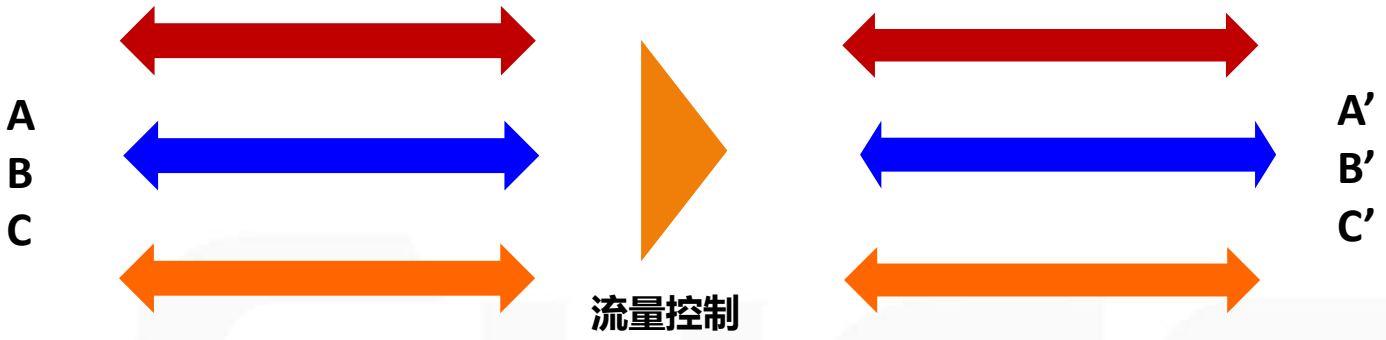
数据访问权限控制

- 分权分域
  - 针对不同的数据对象、用户、角色分配访问权限
  - 面向系统管理员，各类数据分析人员，审计人员等
- 数据网关
  - 聚合数据访问，支持内外网分离，多网络负载均衡
  - 数据访问方式，清洗、转换、加载、查询、挖掘等



数据访问流量控制

- 流量控制技术（基于DPI的协议识别技术）
  - 基于TCP窗口整形的流控技术
  - 基于队列的流控技术
  - 基于干扰的流控技术
- 防止互联网广播风暴，或者病毒/木马造成网络瘫痪



数据访问传输控制

软件加密传输

- 传输之间进行数据加密，如：S/MIME加密邮件传输

安全网络协议

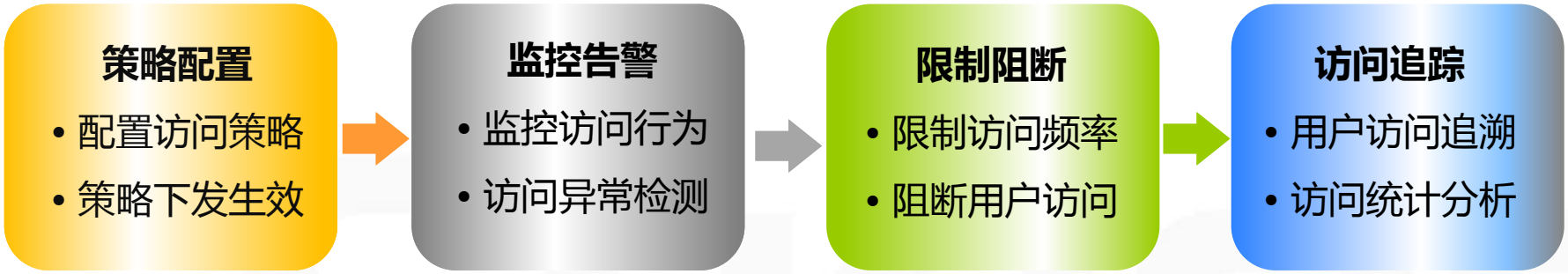
- 建立安全信息通道，SSL，安全套接字层等

系统安全认证

- 基于口令的安全认证
- 基于密钥的安全认证

敏感信息访问控制

- 访问敏感信息监控和告警
  - 访问内容和访问行为的监控
  - 访问权限、频率、敏感字段、敏感操作、异常操作等可进行告警
- 针对异常访问的操作限制
  - 防止非法访问和非法操作等
  - 告警访问、限制访问、阻断访问



目录

大数据系统内容安全

大数据系统包含大量的敏感信息，须加强对数据内容的保护

- 即使获得数据，也不能造成安全威胁
- 防止内部人员对数据的泄露

大数据内容保护在数据读写时，对内容进行相应的处理

- 组件的部署方式，读取和写入时进行处理
- 对数据内容进行审计、监控、告警、阻断和追踪

大数据系统提供对内容保护的算法优化

- 提供新计算模型下的算法优化技术

大数据去隐私化

基于失真的隐私保护技术

- 随机化：随机扰动，随机化应答
- 阻塞、凝聚、交换等技术
- 支持度和置信度

基于加密的隐私保护技术

- 安全多方计算：SMC问题，分布式计算协议
- 分布式匿名化：k-TTP模型
- 分布式关联规则挖掘和分布式聚类

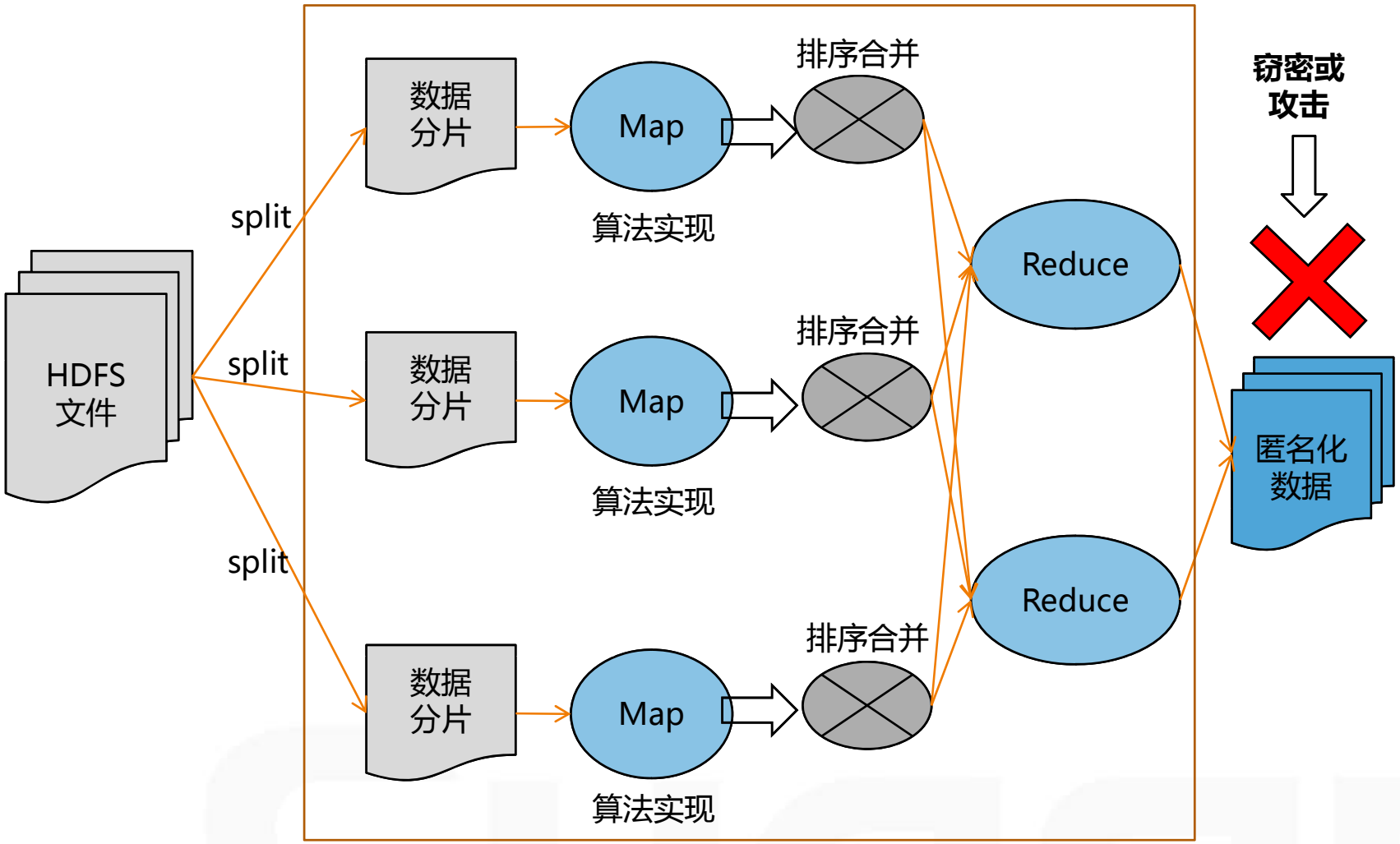
基于限制发布的隐私保护技术

- 两种基本操作：抑制、泛化
- K-匿名、l-diversity、t-近邻

大数据多维度审计


对象审计
敏感字段
敏感内容
访问行为
字段属性和约束
访问和操作权限
访问记录和监控
访问频率统计
用户访问轨迹
用户行为趋势
对象属性和约束
访问和操作权限
访问阻断和告警
敏感记录访问
敏感信息过滤
敏感信息统计

大数据内容保护优化算法



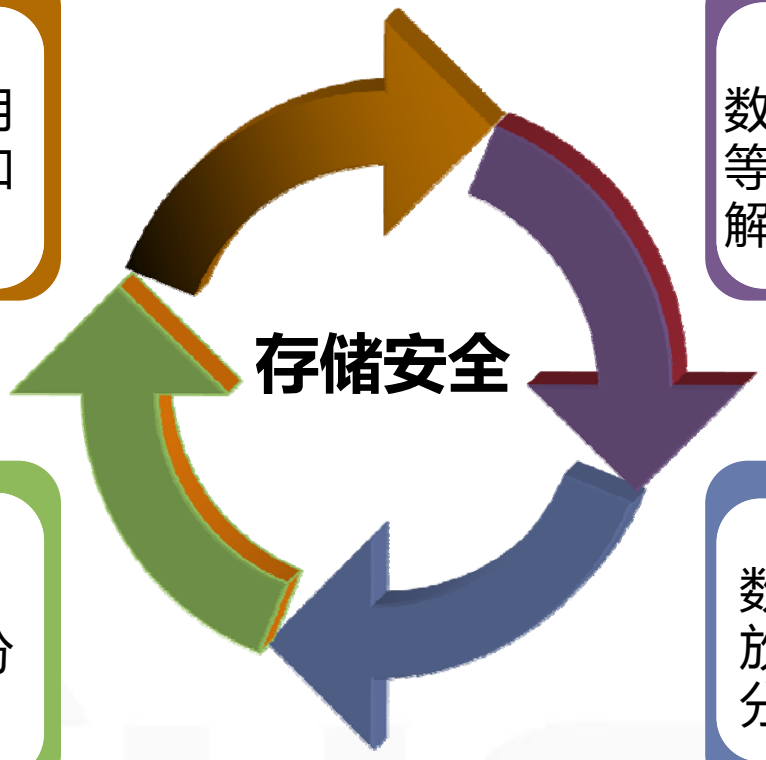
目录

大数据系统存储安全

中科曙光 Sugon

存储认证
第三方认证系统对用户和权限进行设置和认证

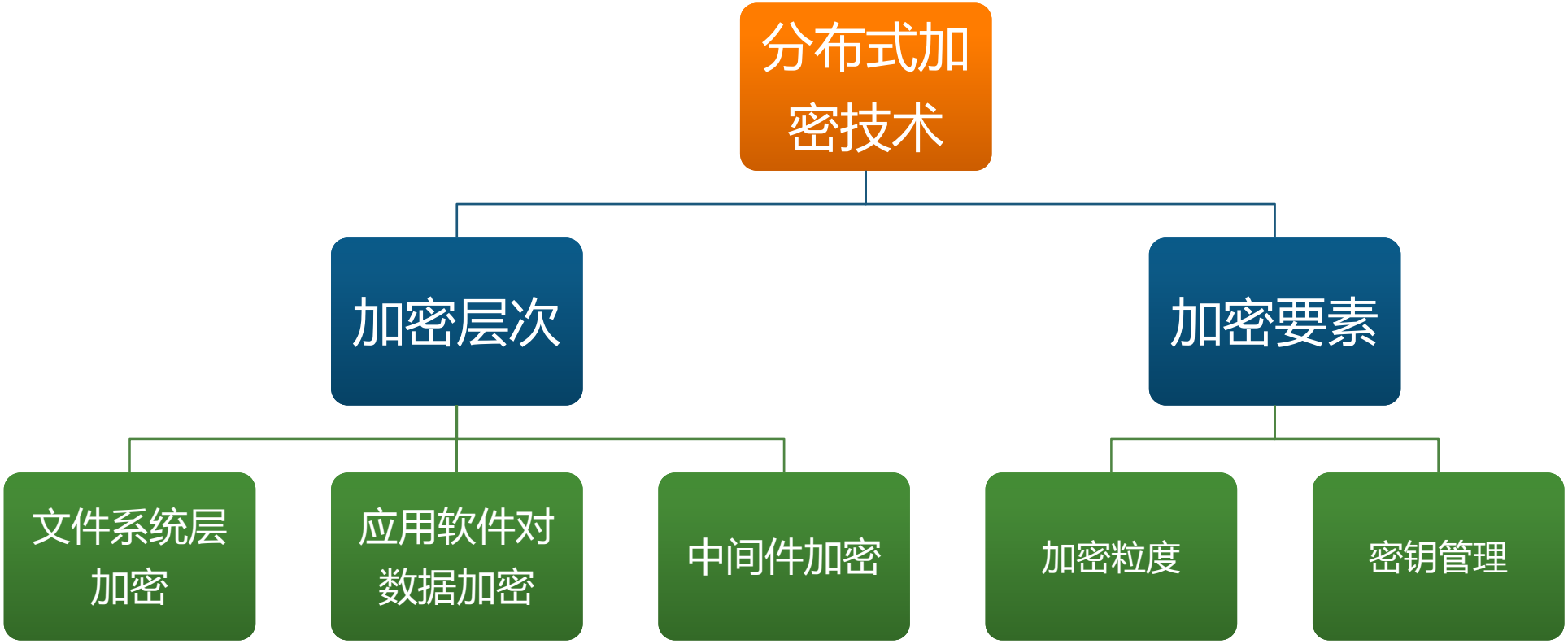
数据加密
数据条、对象、文件等不同粒度的加密和解密技术



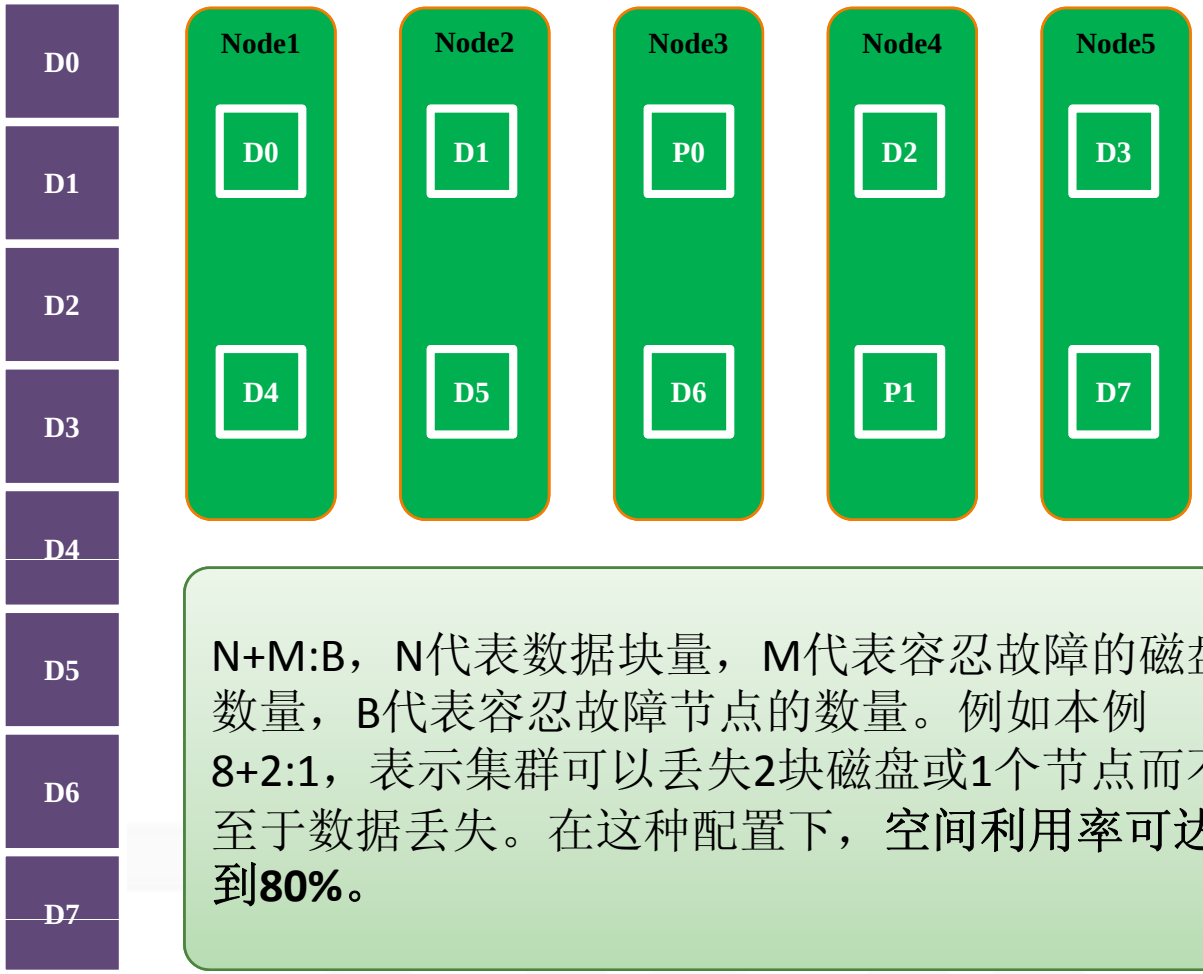
备份容灾
数据导入导出，备份恢复，容灾方案等

副本和分片
数据多副本级副本摆放和访问策略，数据分片和冗余存储

数据分布式加密



数据分布式存储



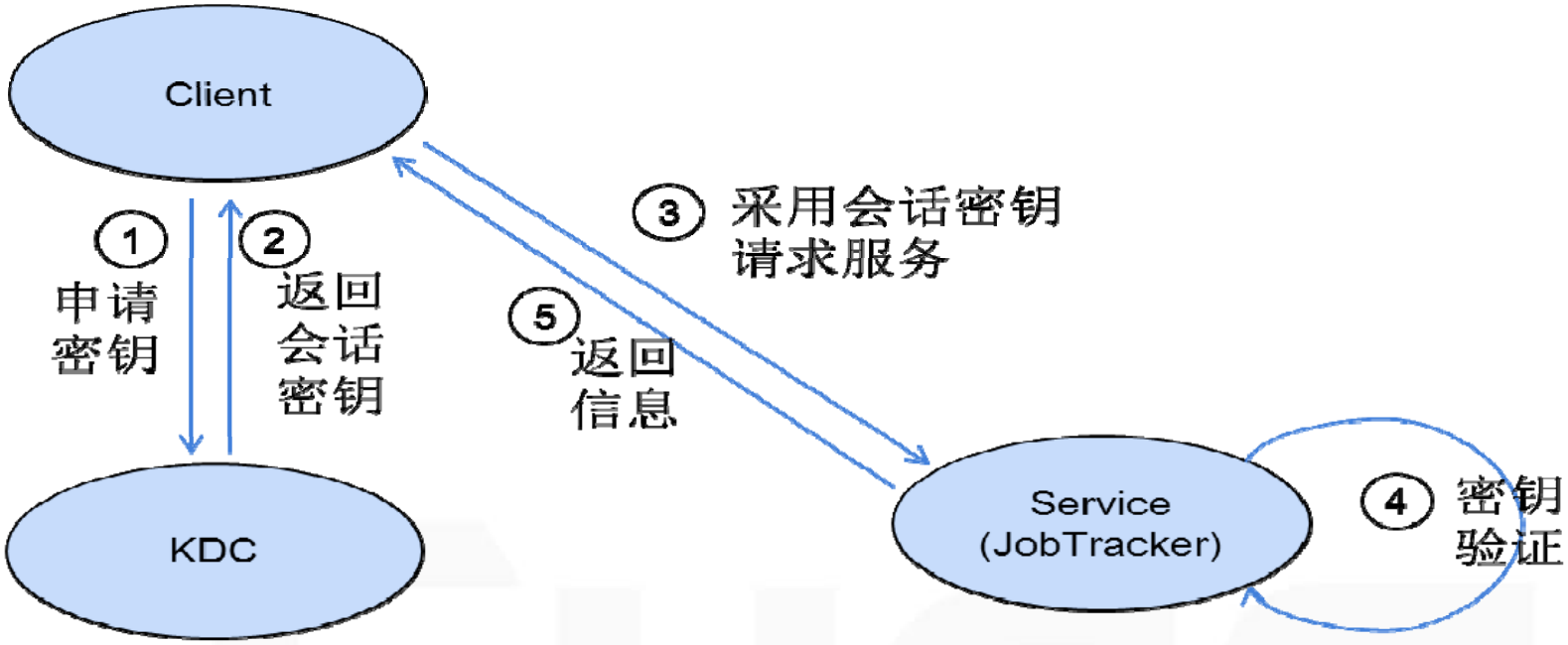
N+M:B，N代表数据块量，M代表容忍故障的磁盘数量，B代表容忍故障节点的数量。例如本例8+2:1，表示集群可以丢失2块磁盘或1个节点而不至于数据丢失。在这种配置下，**空间利用率可达到80%。**

数据备份容灾


数据备份
数据快照
完全备份
增量备份
条件导出
副本和压缩
数据恢复
数据导入
数据迁移
批量导入
索引维护
数据容灾
数据同步
一致性检查
远程镜像
异地容灾
故障切换

存储认证

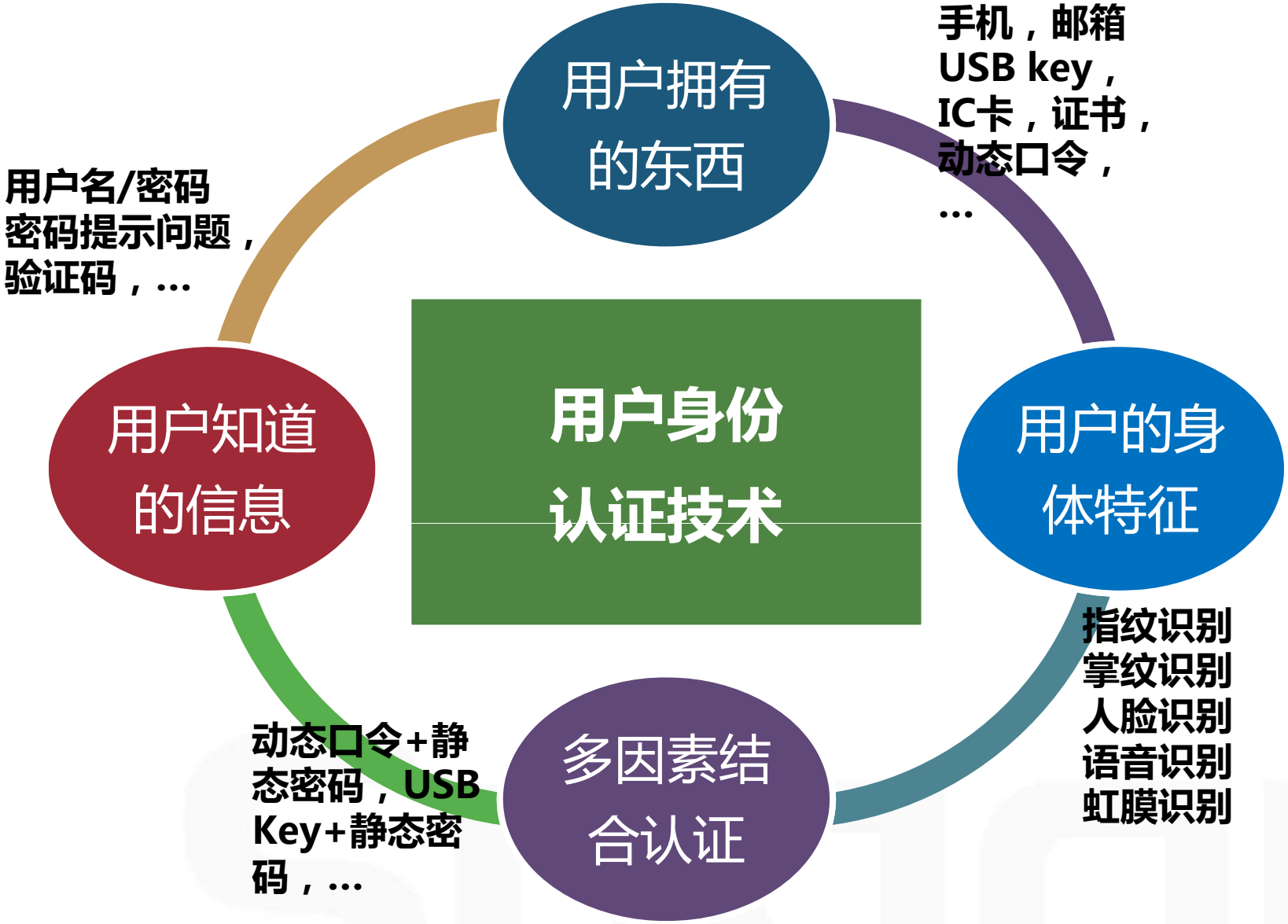
- 大数据系统依靠外围可靠的认证系统。
- 使用对称钥匙操作，比SSL的公共密钥快。
- 操作简单，如废除一个用户只需要从KDC数据库中删除即可。



目录

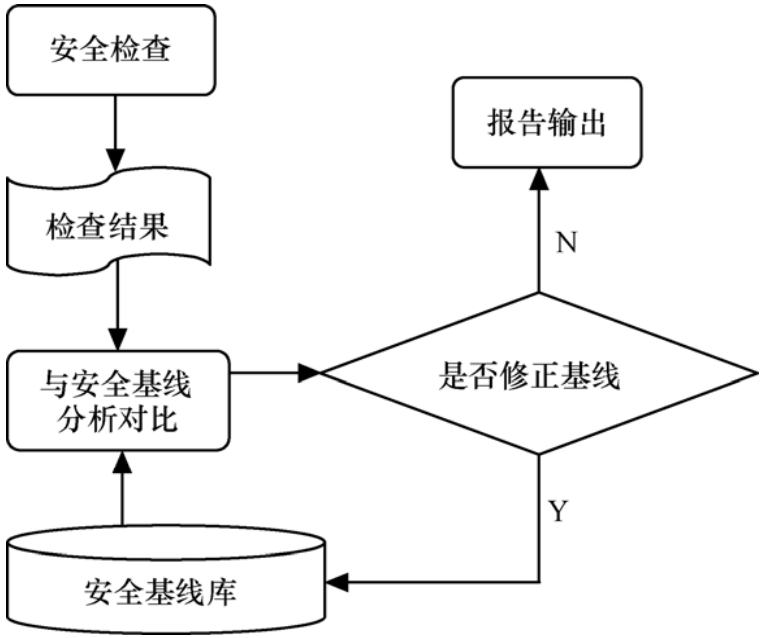
大数据系统运维安全

身份认证


用户拥有的东西
手机，邮箱
USB key，
IC卡，证书，
动态口令，
…
用户知道的信息
用户名/密码
密码提示问题，
验证码，…
用户身份认证技术
用户的身体特征
指纹识别
掌纹识别
人脸识别
语音识别
虹膜识别
多因素结合认证
动态口令+静态密码，USB Key+静态密码，…

配置基线

- 快速检索系统安全配置的自动解决方案
- 确保关键的可执行文件，配置文件的内容、权限、属性等不被恶意修改
- 配置基线的完整性/合法性检查、变更管理和监控



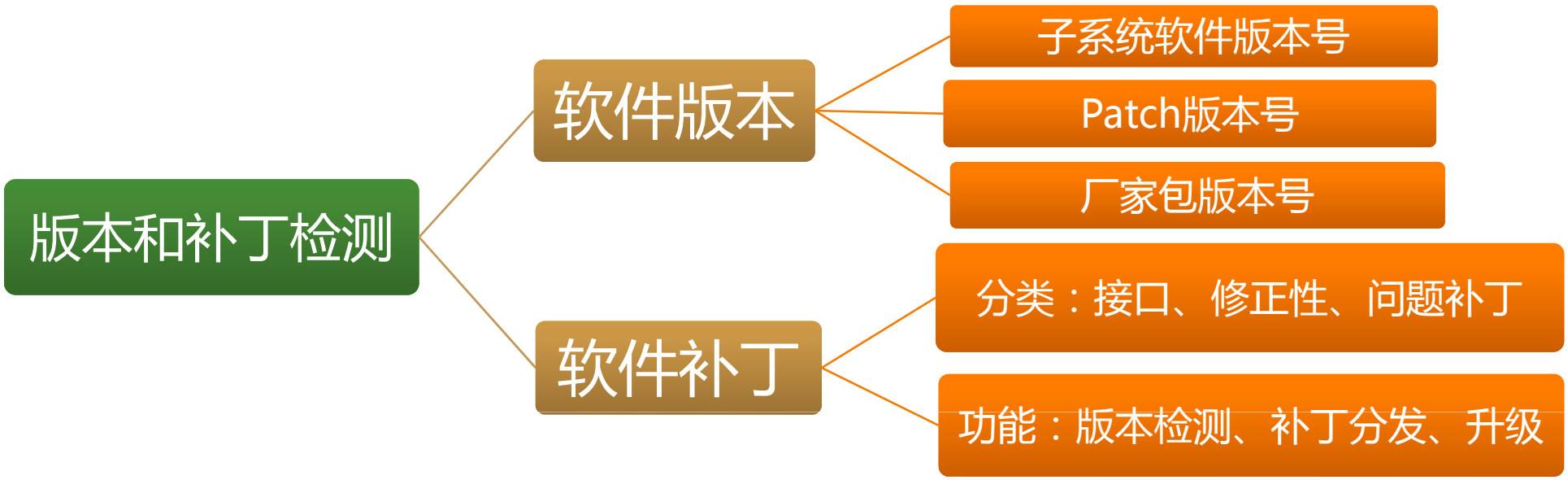
配置完整性检验
- 可执行文件
- 系统配置文件

配置变更管理
- 配置文件基线
- 配置基线变更

基线合规性检查
- 系统基线构建
- 内容权限检查

配置基线监控
- 自身安全脆弱性
- 定时检测和告警

版本和补丁管理



大数据系统版本和补丁管理

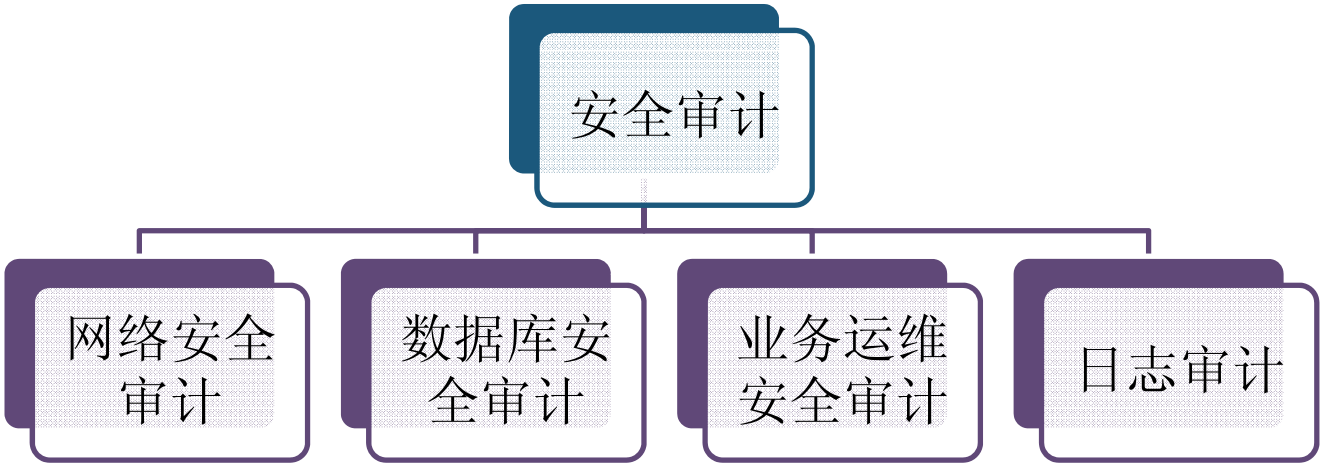
涉及多个组件（子系统）：HDFS、Hbase、Hive、MapReduce、DB、Zookeeper等

软件开发升级主要问题：兼容性

数据格式兼容：备份和迁移访问和计算兼容

数据是应用的关键，涉及到数据组建的升级需要谨慎，做好数据备份和方法验证

安全审计



漏洞扫描

漏洞分类

- 应用软件漏洞：www,FTP,SMPT等
- 操作系统漏洞：windows中RPC,NETBOIS漏洞等

扫描方法

- 特征匹配：基于规则的模式特征匹配
- 插件技术：调用插件进行检测，包括错误配置、简单口令、网络协议漏洞等

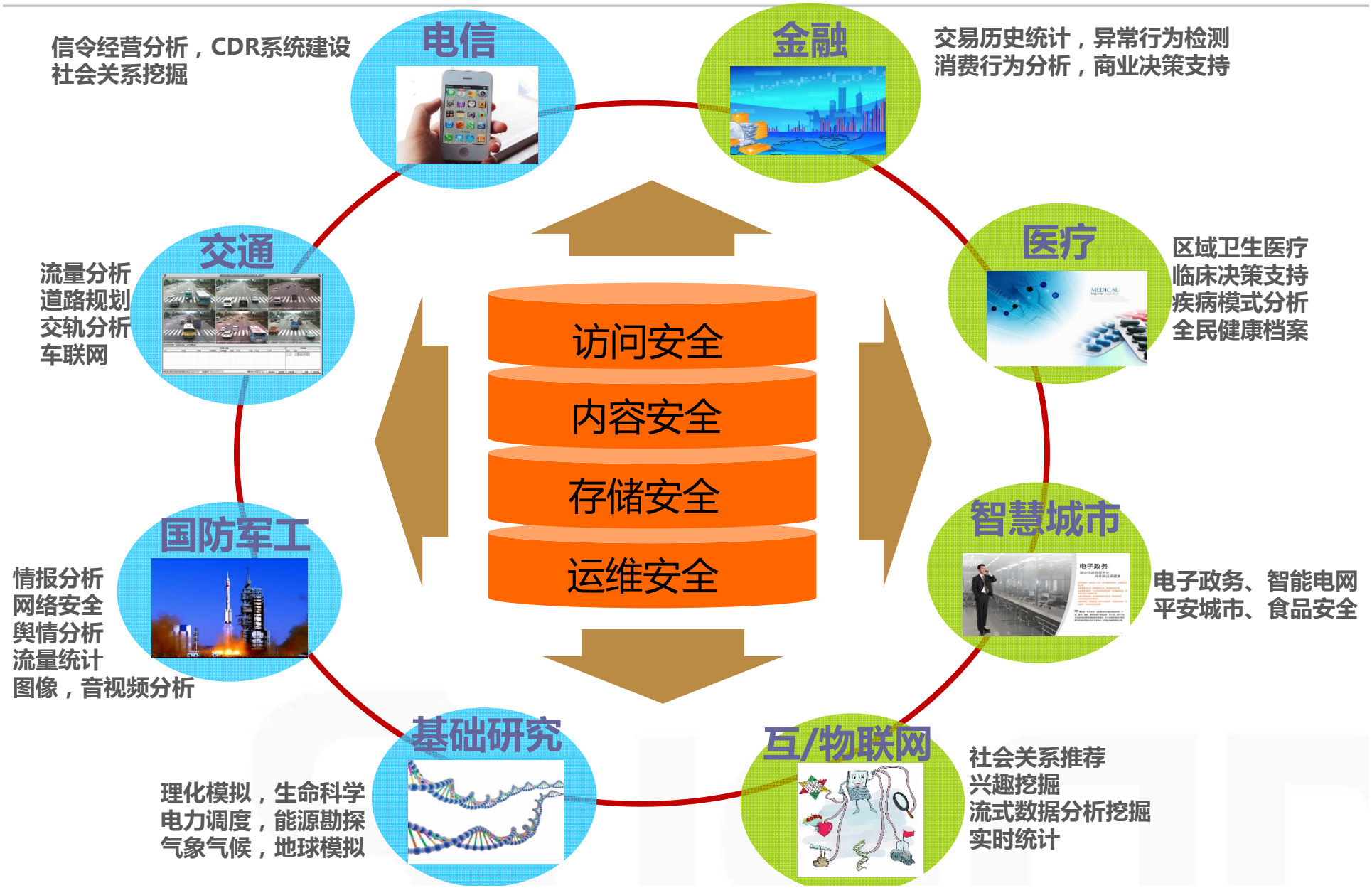
系统监控


集中存储
数据汇聚、信息规整、分布存储，批量入库
聚合分析
分类统计，聚集计算，异常检测等
系统告警
告警级别，告警方式，告警响应机制
运维优化
管理策略调整，故障处理，参数优化
信息采集
CPU、内存、网络、操作系统、中间件、数据库、应用业务
系统监控

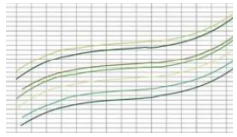
目录

典型应用场景

中科曙光 Sugon



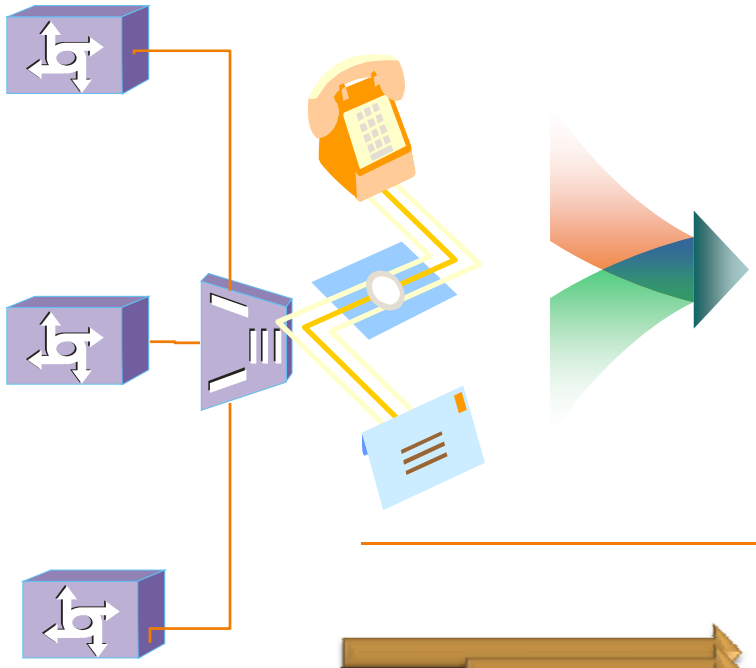
公共安全-某大型网络安全监控系统



流量事件

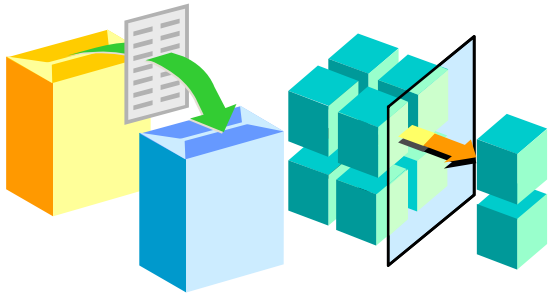
特征事件

分布事件



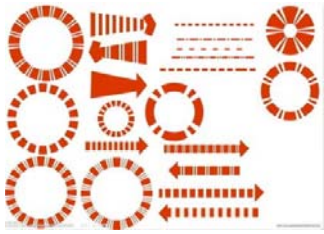
SN-MPP架构结构化大数据平台

曙光XData



采集 存储 分析

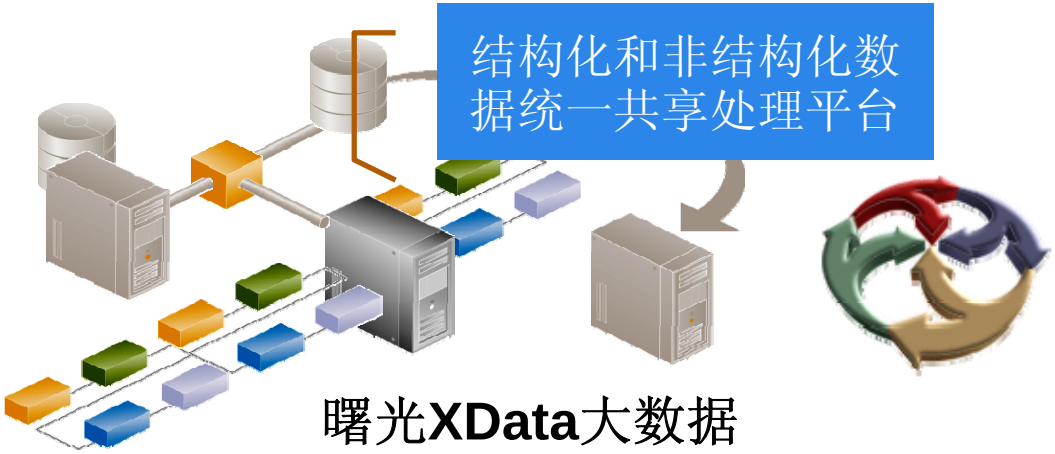
采集清洗、入库分析



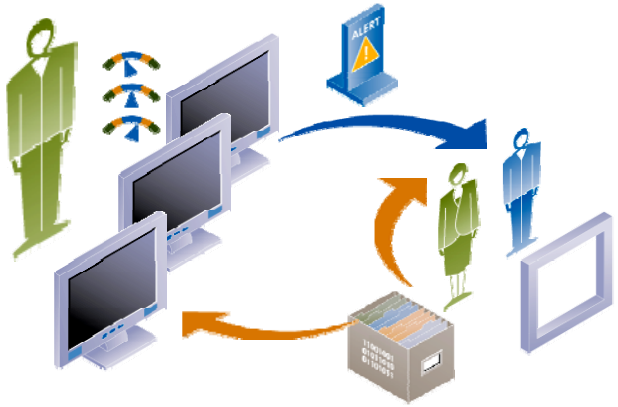
流量统计，事件分析，分布统计

定位追踪、趋势分析、预警

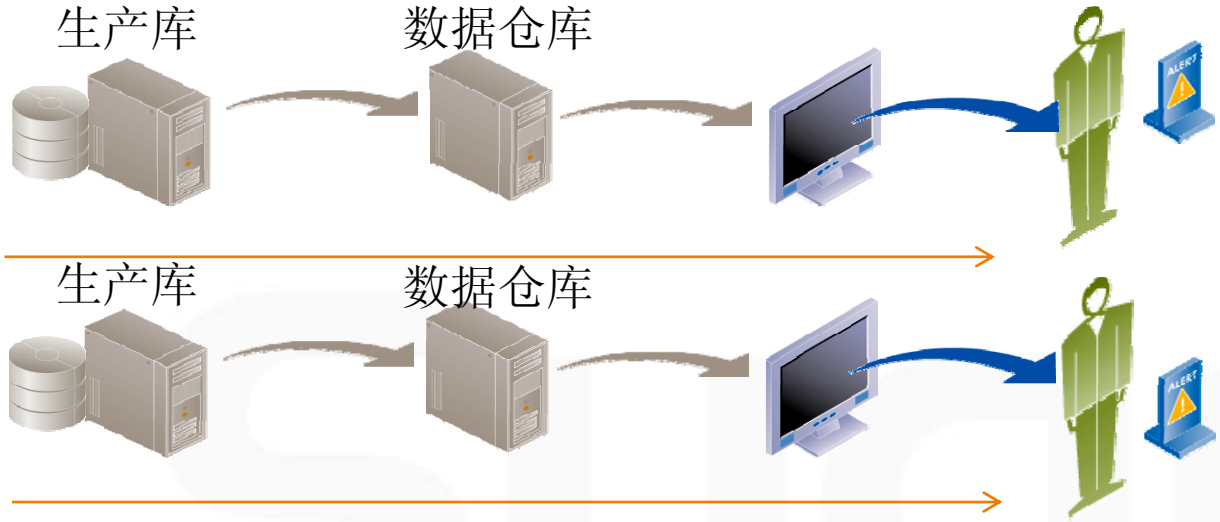
金融行业-银联离线交易数据分析平台



曙光XData大数据处理平台



基于大数据处理的应用及其价值展现



传统垂直建设方式

THANK

通讯地址：北京市海淀区东北旺西路8号中关村软件园36号

邮政编码：100094　联系电话：010-56308000　微博：http://weibo.com/zksugon

EMAIL:MARKET@SUGON.COM　网站(web)：Http://www.sugon.com