

FusionInsight HD

解决方案介绍

www.huawei.com



目录

1. FusionInsight HD 企业版概述
2. FusionInsight HD 特性介绍
3. FusionInsight HD 成功案例

Apache Hadoop—繁荣的开源生态系统

Hadoop核心基本组件的开源代码量，接近200万行

HBase组件在快速的发展中，平均每个月一个发布版

各组件团队之间、组件内团队成员是松耦合、地理分布式运作，相关特性无有效整合

组件	开源代码量 (KLoc)
ZooKeeper	64
HDFS	438
YARN MapReduce	380
HBase	260
Hive	273
Oozie	88
汇总	1652

▶  0.95.1	Next in the 0.95	266 issues
Not for production. Release Date: 11/Jun/13 Release Notes		
▶  0.95.0	First in the 0.95	1266 issues
Not for production. Release Date: 07/Apr/13 Release Notes		
▶  0.94.8	Bug fix release	45 issues
Release Date: 31/May/13 Release Notes		
▶  0.94.7	Bug fix release	73 issues
Release Date: 25/Apr/13 Release Notes		
▶  0.94.6.1	Emergency Bug Fix	1 issues
Release Date: 04/Apr/13 Release Notes		
▶  0.94.6	Bug fix release	58 issues
Release Date: 22/Mar/13 Release Notes		
▶  0.94.5	Bug fix release	77 issues
Release Date: 15/Feb/13 Release Notes		
▶  0.94.4	Bug fix release	83 issues
Release Date: 12/Jan/13 Release Notes		
▶  0.94.3	Bug fix release	51 issues
Release Date: 03/Dec/12 Release Notes		
▶  0.94.2	Bug fix release	117 issues
Release Date: 11/Oct/12 Release Notes		



FusionInsight HD, 开源到企业级的蜕变

采纳社区精华，去除开源Bug：

- 谨慎选择稳定基线版本；
- 认真评估高版本补丁影响范围评估和回合策略
- 采用数万个测试用例，确保企业版本稳定性

3

- 易开发
- 易管理运维

2

- 统一安全管控
- RBAC授权机制
- 数据加密存储

1

- 系统无单点故障
- 地理容灾

易用

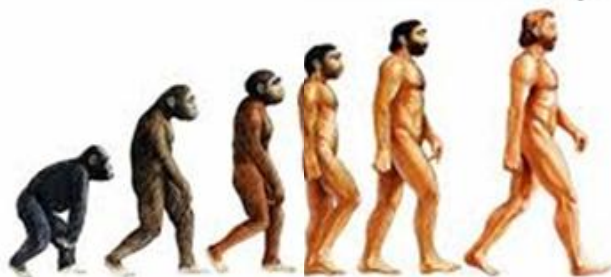
安全

可靠

初生
开源

繁荣
社区

企业
版本



安全

版本配合

配置

补丁选择

性能调优

基线选择

Hadoop

HBase

日志

FusionInsight架构：分层解耦开放



FusionInsight 大数据平台



•敏捷

- 完全开放的架构，性能线性扩展
- 强大的SQL能力，业务移植便捷
- 丰富的工具支持，开发运维高效

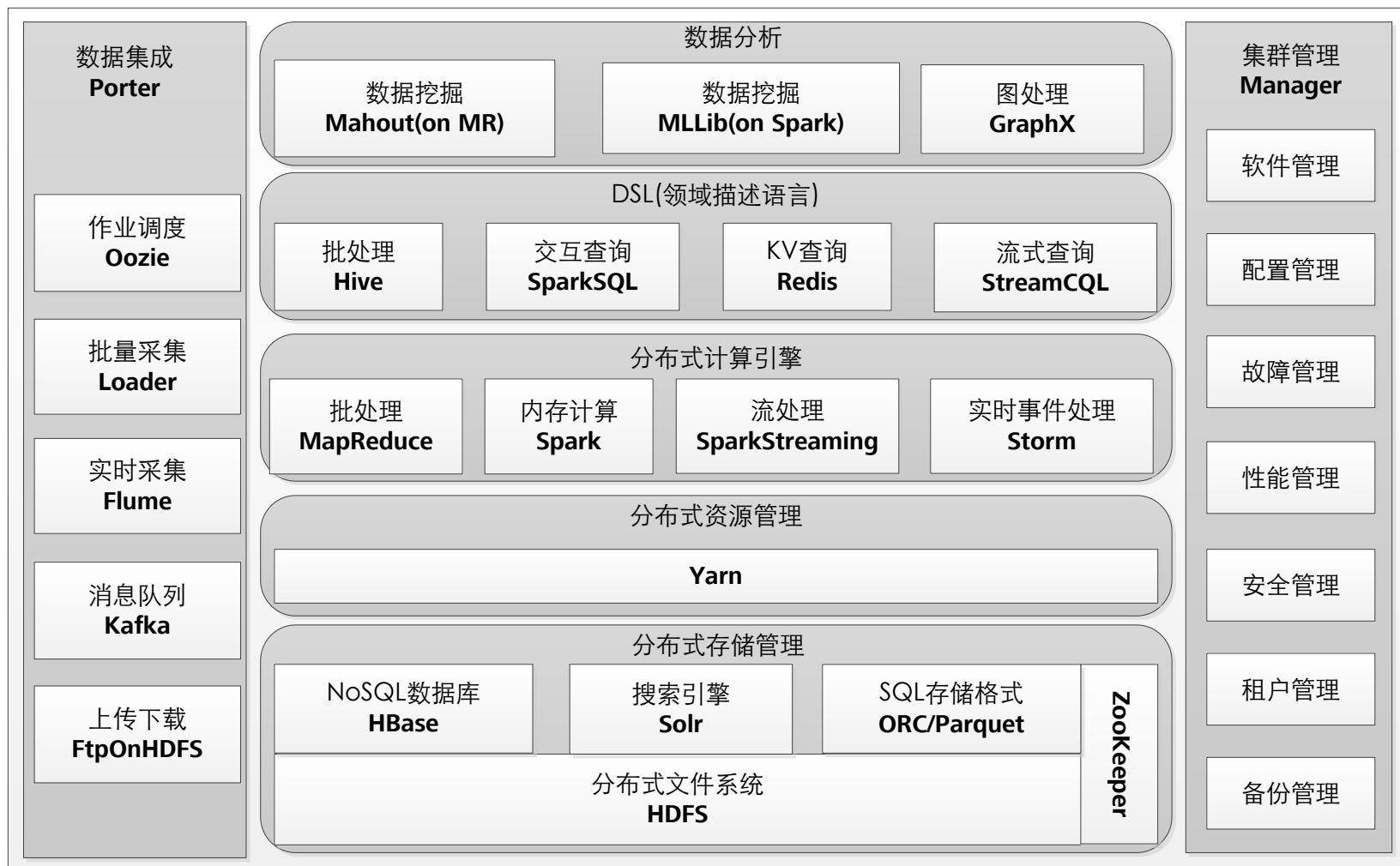
•智慧

- 全量建模，深刻洞察
- 自研算法，高效精准

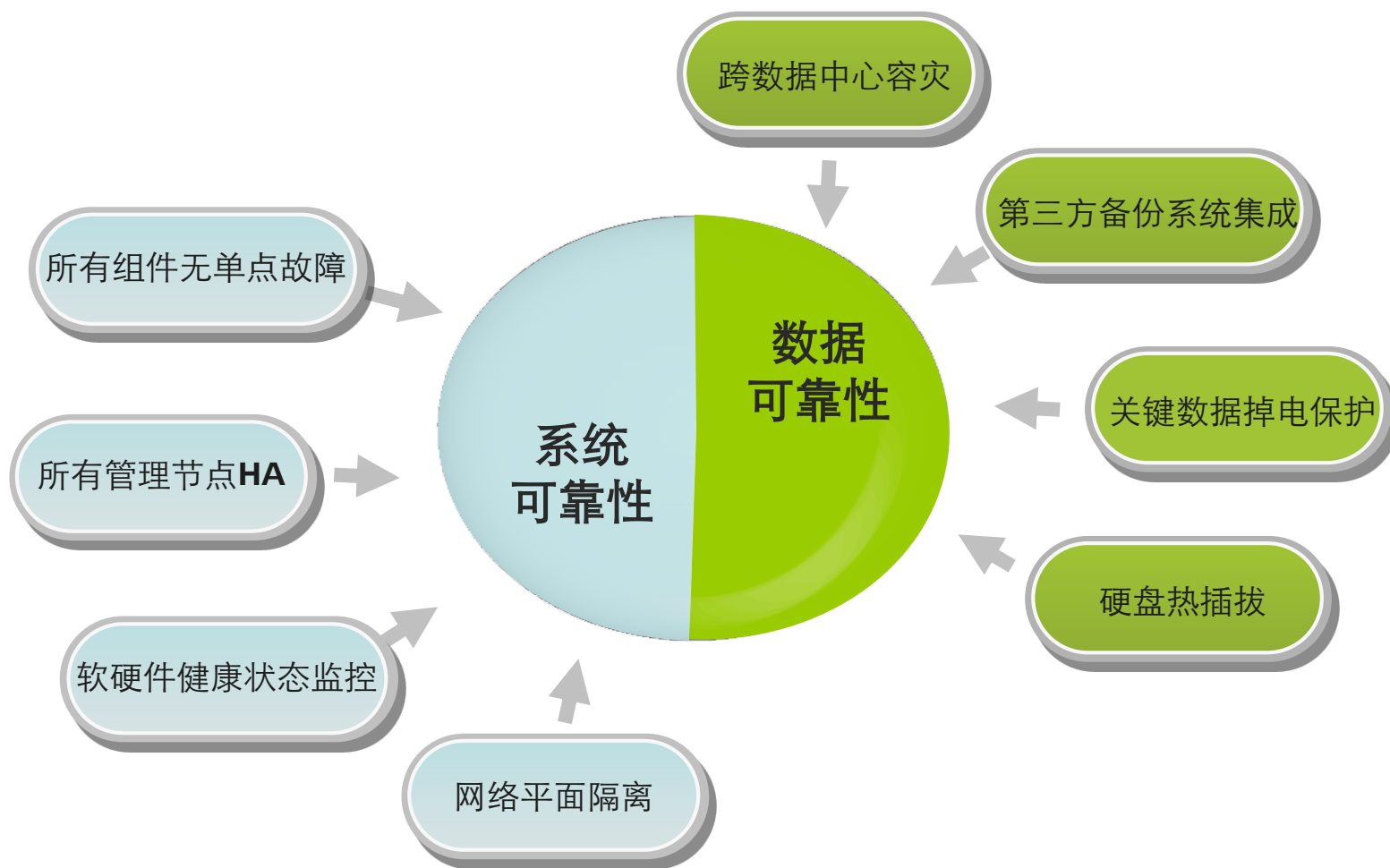
•可信

- 全组件HA、异地容灾
- 开放共赢，可信赖的合作伙伴

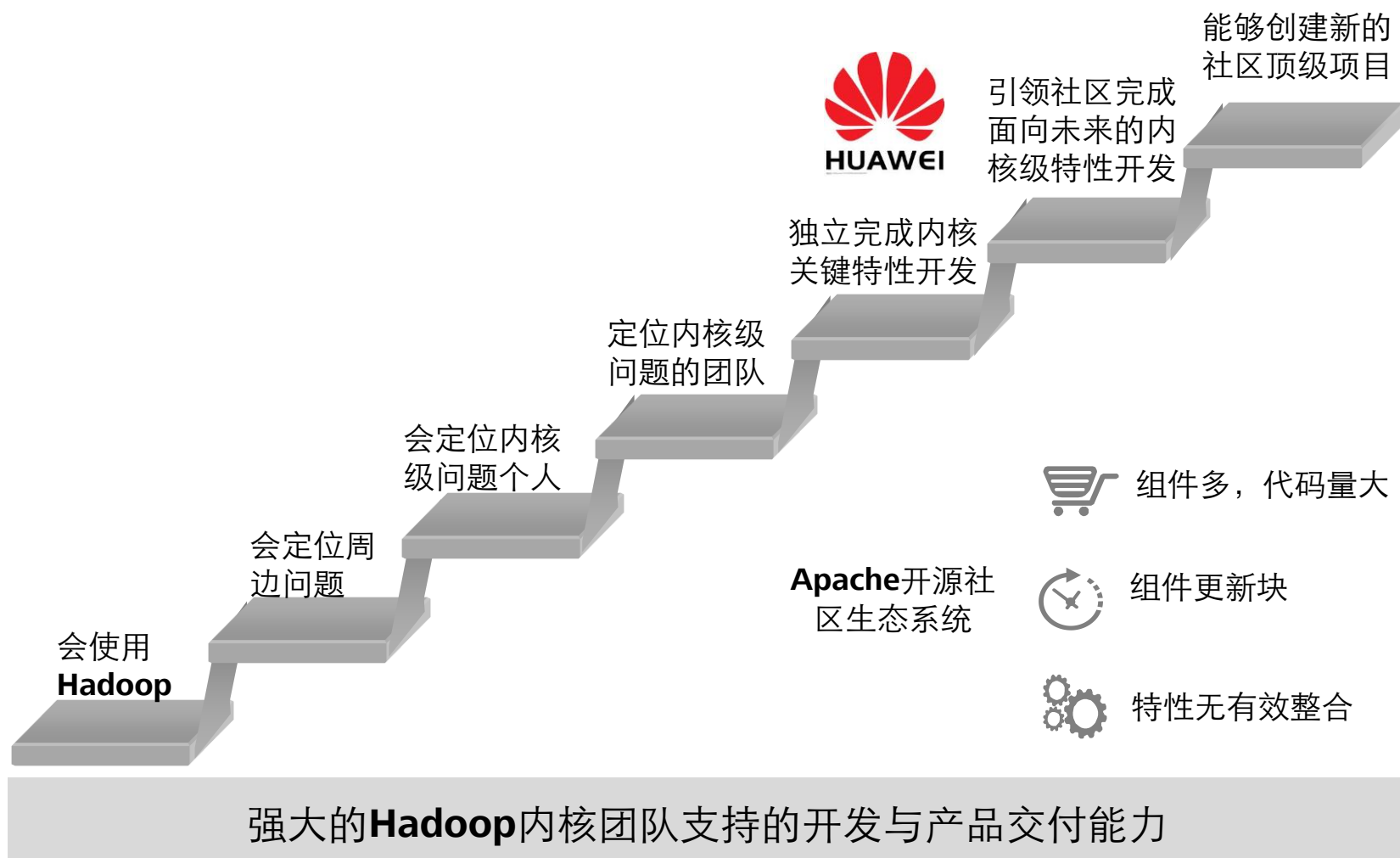
FusionInsight HD解决方案架构



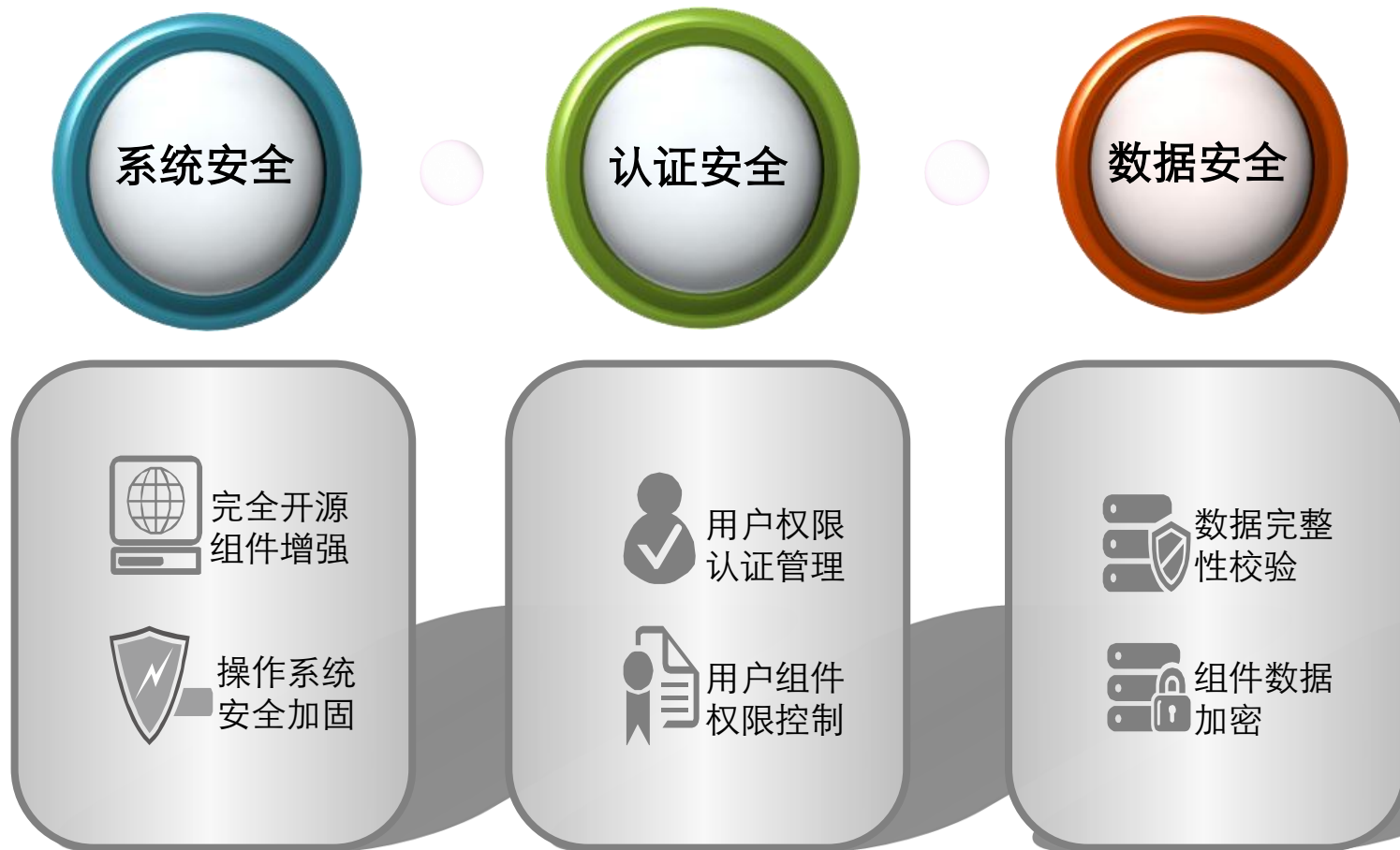
高可靠性



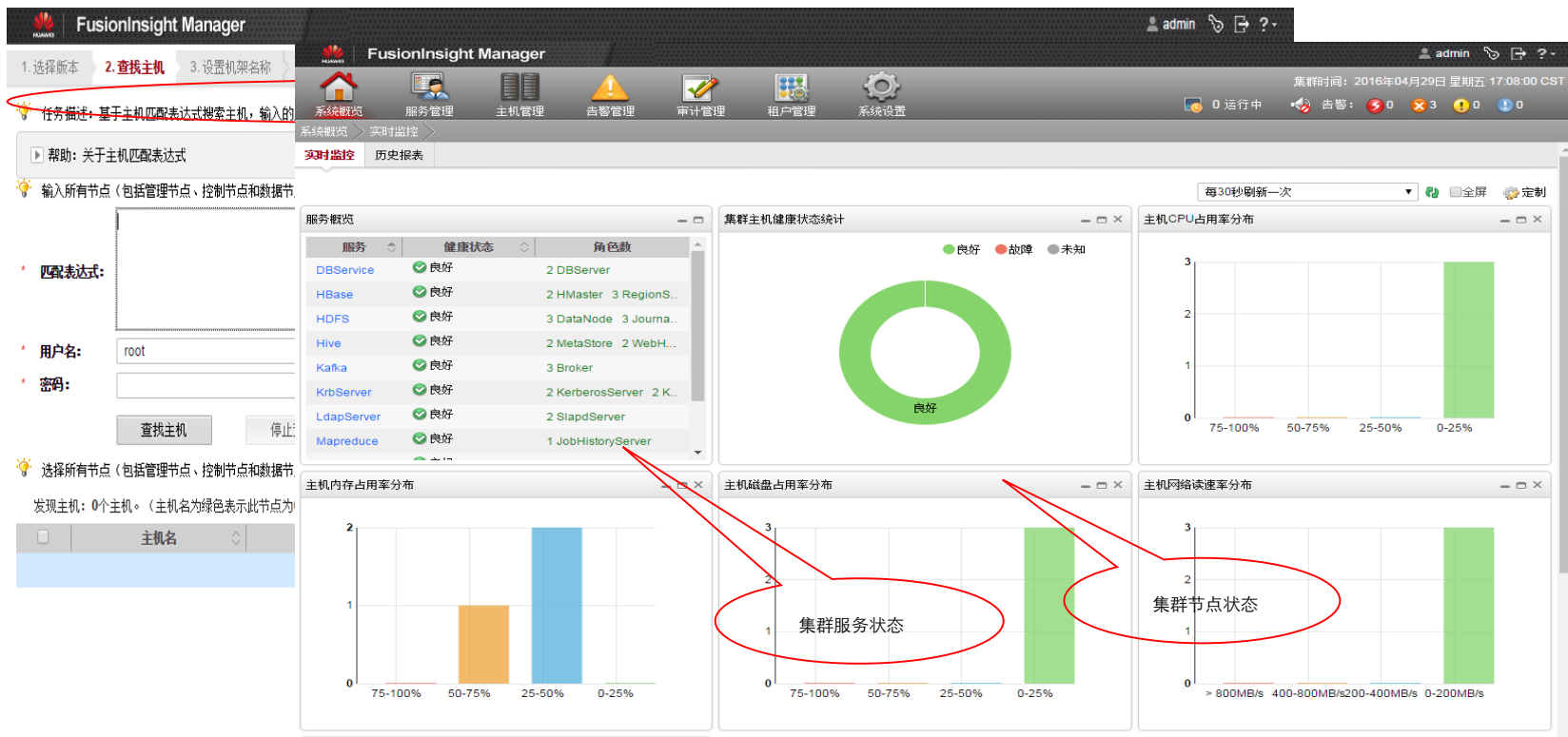
强大的组织支撑能力



安全性



可视化集群管理，易运维



- 服务组件配置项可视化
- 向导式/一键式集群安装部署和扩容，简单快速
- 全面的系统监控与告警管理，服务状态、硬件状态一目了然
- **SNMP/FTP/SYSLOG**标准接口与企业已有管理系统无缝对接

易开发

HBase原生API

```
try {
    table = new HTable(conf, TABLE);
    // 1. Generate RowKey.
    {.....}
    // 2. Create Put instance.
    Put put = new Put(rowKey);
    // 3. Convert columns into qualifiers(Need to consider
    merging cold columns).
    // 3.1. Add hot columns.
    {.....}
    // 3.2. Merge cold columns.
    {.....}
    put.add(COLUMN_FAMILY, Bytes.toBytes("QA"), hotCol);
    // 3.3. Add cold columns.
    put.add(COLUMN_FAMILY, Bytes.toBytes("QB"),
    coldCols);
    // 4. Put into HBase.
    table.put(put);
} catch (IOException e) {
    // Handle IOE. Need to re-create connections under
    some scenarios.
}
```

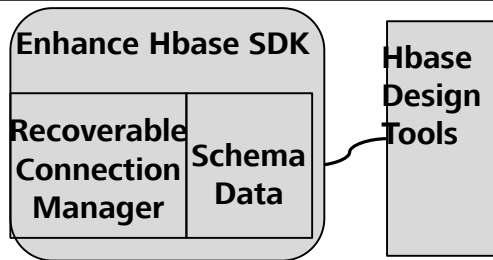
增强API

```
try {
    table = new ClusterTable(conf,
    CLUSTER_TABLE);
    // 1. Create CTRow instance.
    CTRow row = new CTRow();
    // 2. Add columns.
    {.....}
    // 3. Put into HBase.
    table.put(TABLE, row);
} catch (IOException e) {
    // Does not care connection re-creation.
}
```

HBase的表设计工具、连接池管理和增强的SDK，来简化复杂数据表的业务开发

HBase API

HBase



易开发(续)

Storm原生API

```
Def Input:
public void open(Map conf, TopologyContext context,
SpoutOutputCollector collector)
{...}
public void nextTuple()
{...}
public void declareOutputFields(OutputFieldsDeclarer declarer)
{...}
Def logic:
public void execute(Tuple tuple, BasicOutputCollector collector)
{...}
public void declareOutputFields(OutputFieldsDeclarer ofd)
{...}
Def Output:
public void execute(Tuple tuple, BasicOutputCollector collector)
{...}
public void declareOutputFields(OutputFieldsDeclarer ofd)
{...}
Def Topology:
public static void main(String[] args) throws Exception
```

StreamCQL

```
Def Input:
Create Input Stream
kafkareader...

Def logic:
Insert into filterstr select *
from kafkareader where
name="HUAWEI";

Def Output:
Create Output Stream
kafkawriter...

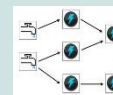
Def Topology:
Submit application test;
```

StreamCQL SDK

Storm API

基于类SQL的StormCQL
查询语言，降低流处理
业务的开发门槛

Storm

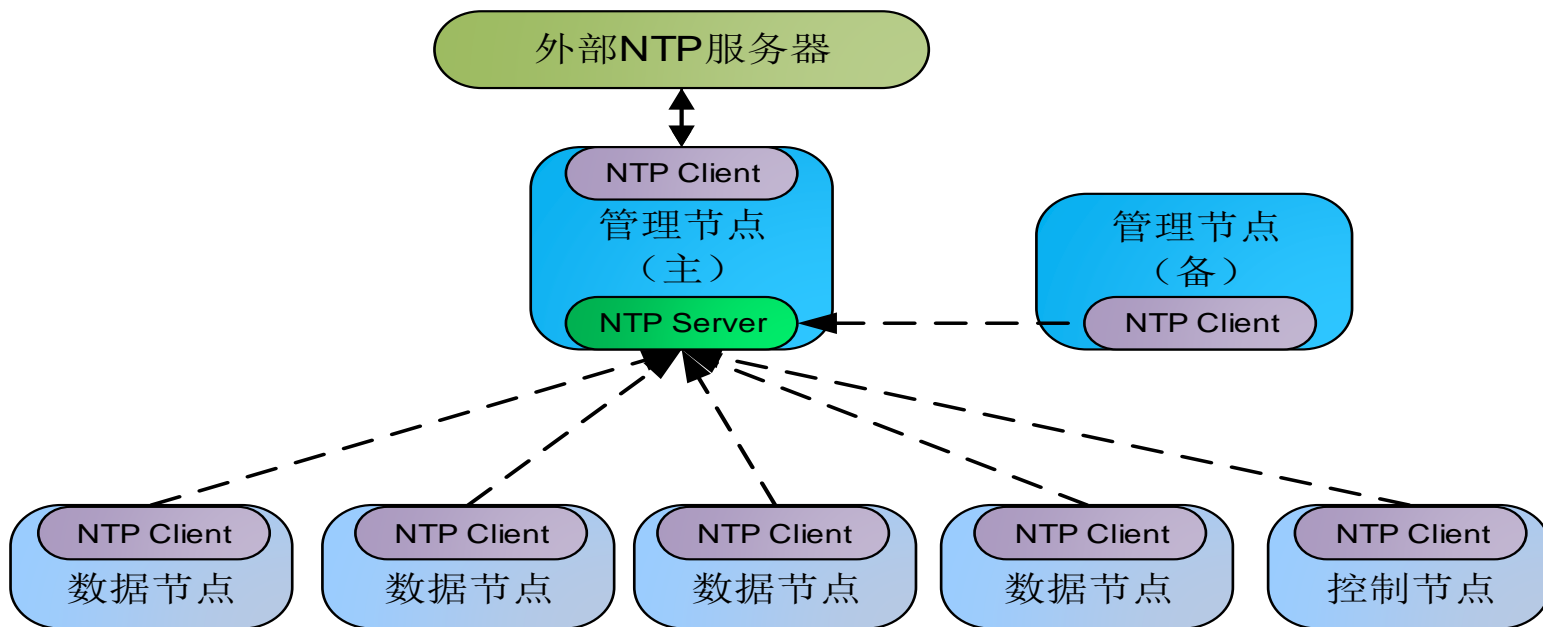




目录

1. FusionInsight HD 企业版概述
2. FusionInsight HD 特性介绍
3. FusionInsight HD 成功案例

NTP自动配置



- 用户价值

- 减少配置NTP的专业技能要求
- 加强了开源NTP在极端情况下的可靠性
- 加快了用户安装Hadoop集群的环境准备时间

自动配置主机映射关系

```
#
# hosts        This file describes the
#              mappings for the TCP/IP
#              stack used at boot time, which
#              is used by the "named" name server.
#
# Syntax:
#
# IP-Address  Full-Qualified-Hostname
#
127.0.0.1      localhost
# special IPv6 addresses
::1           localhost ipv6-localhost
fe00::0       ipv6-localnet
ff00::0       ipv6-mcastprefix
ff02::1       ipv6-allnodes
ff02::2       ipv6-allrouters
ff02::3       ipv6-allhosts
10.164.0.7    host01
10.164.0.32   host03
10.164.0.31   host02
```

- 价值

- 大大缩短了用户安装Hadoop集群的环境准备时间
- 降低用户配置出错的概率
- 减少在大规模集群下，用户扩容后再配置已稳定运行节点的风险

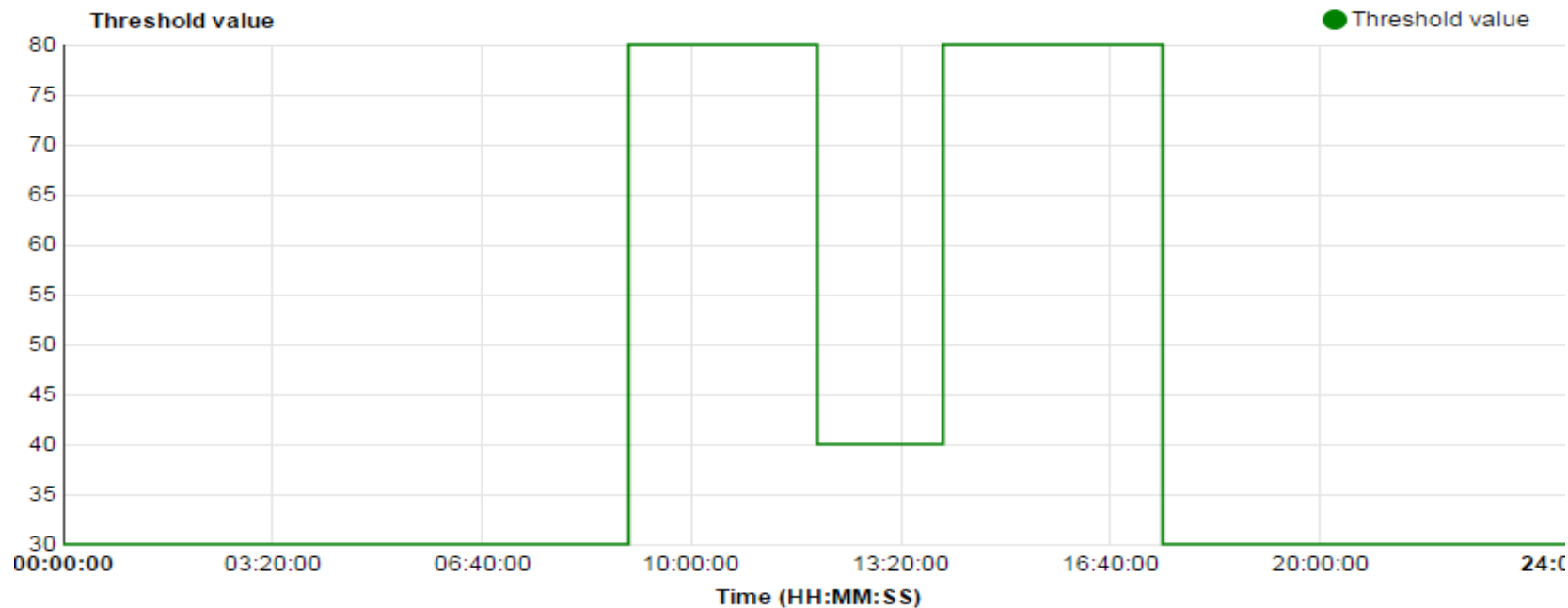
资源分布监控



- 价值

- 帮助用户快速聚焦在最关键的资源消耗上，
- 帮助用户迅速找到资源消耗最高的节点，采取适当的措施

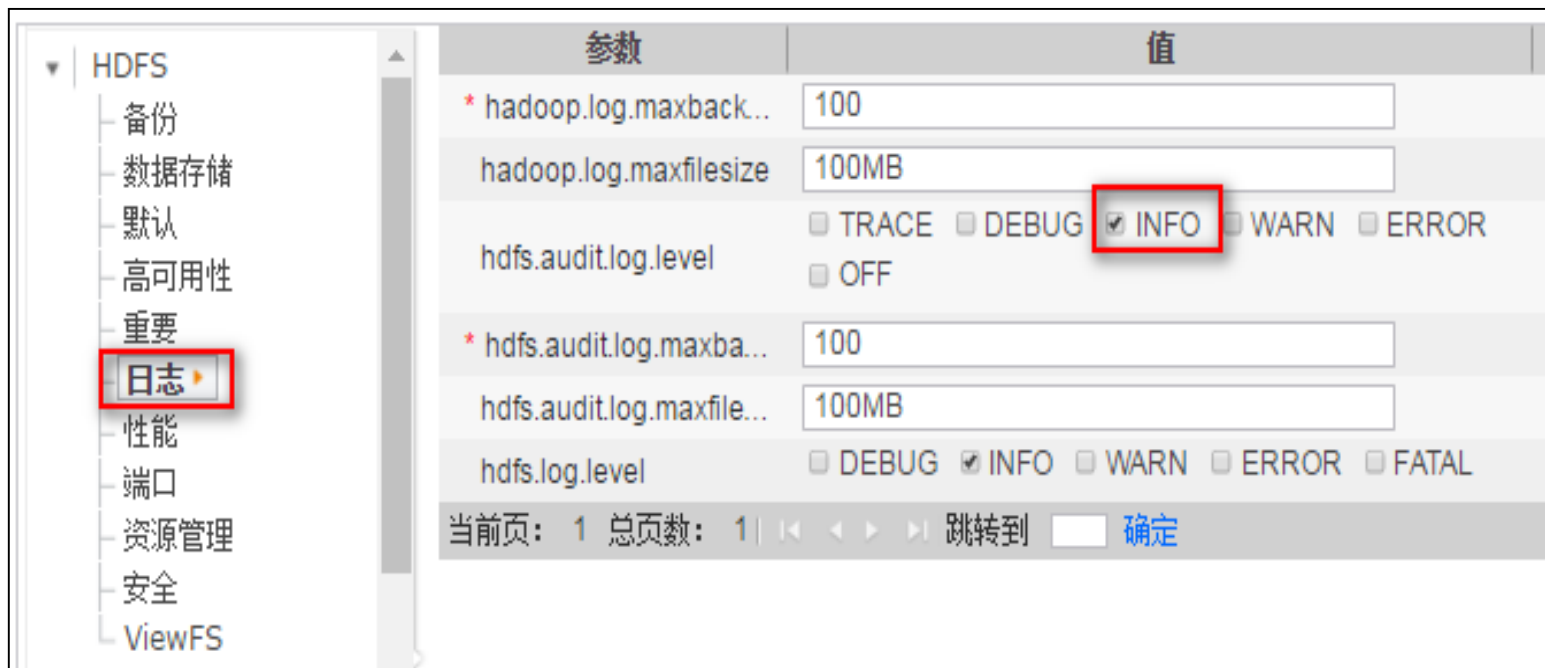
自定义监控阈值



- 价值

帮助用户按实际业务模型，精细化定制监控阈值，提前发现问题

日志级别动态调整



- **应用场景：**Hadoop集群出问题后，快速定位需要修改日志级别，将定位所需的日志打印出来又不能重启进程导致业务中断，如何处理？
- **解决方案：**提供在WebUI上动态修改日志级别的功能，用户修改
- **用户价值：**定位问题时，快速修改指定服务或者节点的日志级别，不需重启服务，不中断业务

向导式备份管理，操作简单可靠

向导式备份/恢复任务配置

创建恢复任务

* 任务名称: dddd 任务名称由数字、字母和下划线组成，长度3~128个字符

* 恢复配置: 元数据

- ☐ NameNode
☐ LdapServer
☐ OMS

业务数据

☒ HBase

路径配置

- * 路径类型: RemoteHDFS(DistCp) ✓
- * 源端NameNode IP地址: 10.64.35.146 ✓
- * 源端路径: /backup/hbase/abc_HBase_20160308163358 ✓
- * 队列名称:
- * 恢复点列表: abc_HBase_HBase_20160309140710 ✓ 刷新

数据配置

<input type="checkbox"/>	备份数据
<input type="checkbox"/>	default11

当前页: 1 总页数: 1 < > 跳转到

数据选择:

强制覆盖:

☐ 是 ☒ 否

校验

灵活的调度策略

创建备份任务

* 任务名称: aaa

备份类型: ☒ 周期备份 ☐ 手动备份

* 开始时间:

* 周期: ☒ 按小时 ☐ 按天

* 备份策略: ⚠ 元数据不支持增量备份

☒ 首次全量备份，后续增量备份

☐ 每次都全量备份

☐ 每n次进行一次全量备份

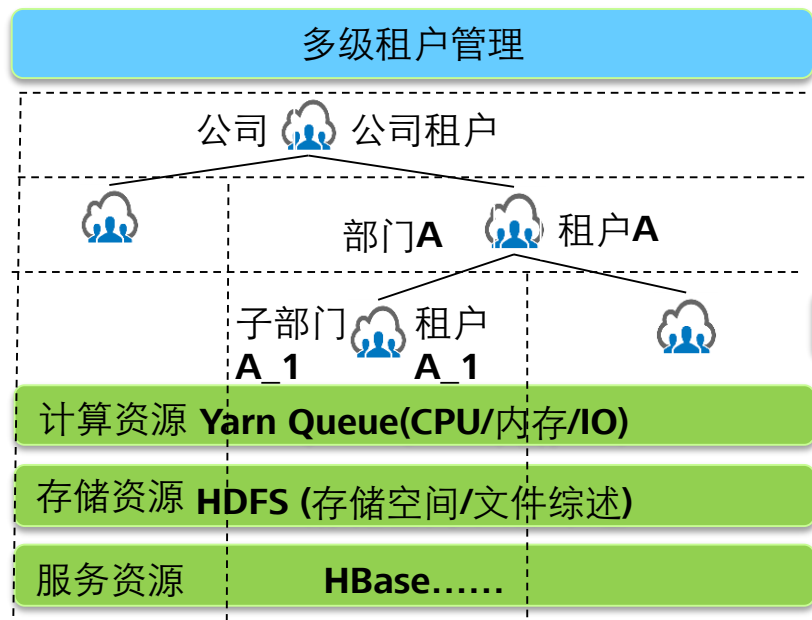
* 备份配置: 元数据

- ☐ NameNode
☐ LdapServer
☐ OMS

业务数据

- ☐ HBase
☐ HDFS

与企业组织结构匹配的多租户管理



一站式租户管理

租户管理 > 租户管理 > 添加租户

▶ 帮助: 添加租户

* 名称:

* 租户类型: ☒ 叶子租户 ☐ 非叶子租户 ☒

动态资源:

* 默认资源池容量: %

储存资源:

* 存储空间配额: MB

服务: [+ 关联服务](#)

服务名称	
HBase	独占

可视化的集中用户权限管理

可视化的集中用户权限管理，易用、灵活、精细

- 可视化的多组件统一的集中用户权限管理，易用
- 基于角色的访问控制（**RBAC**），预定义权限集（角色）可重复使用，灵活
- 多层次（数据库/表/列级）、细粒度（**Select/Delete/Update/Insert/Grant**等）授权，精细

系统设置



权限配置

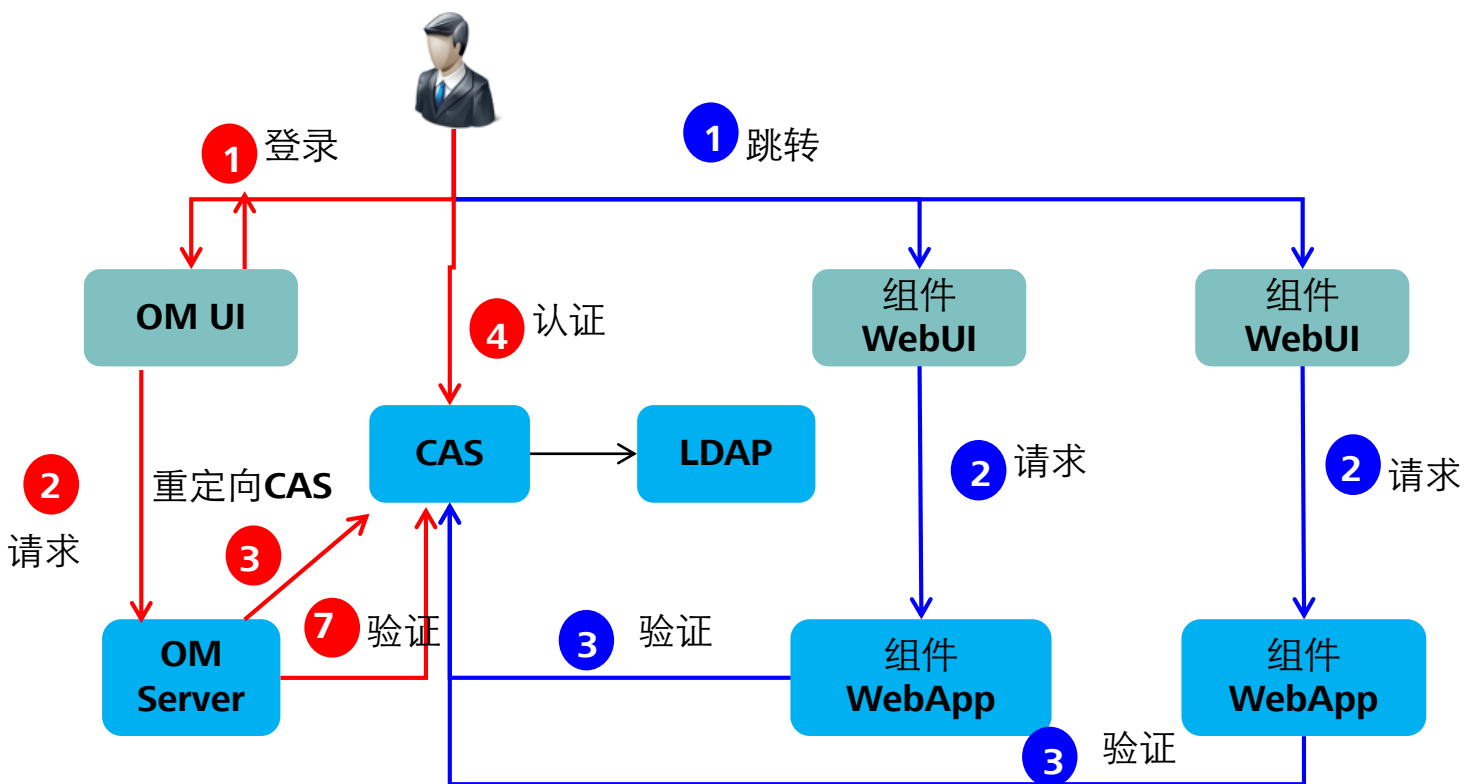
管理本地用户、用户组和角色，配置系统的密码策略。服务配置。

用户管理

用户组管理

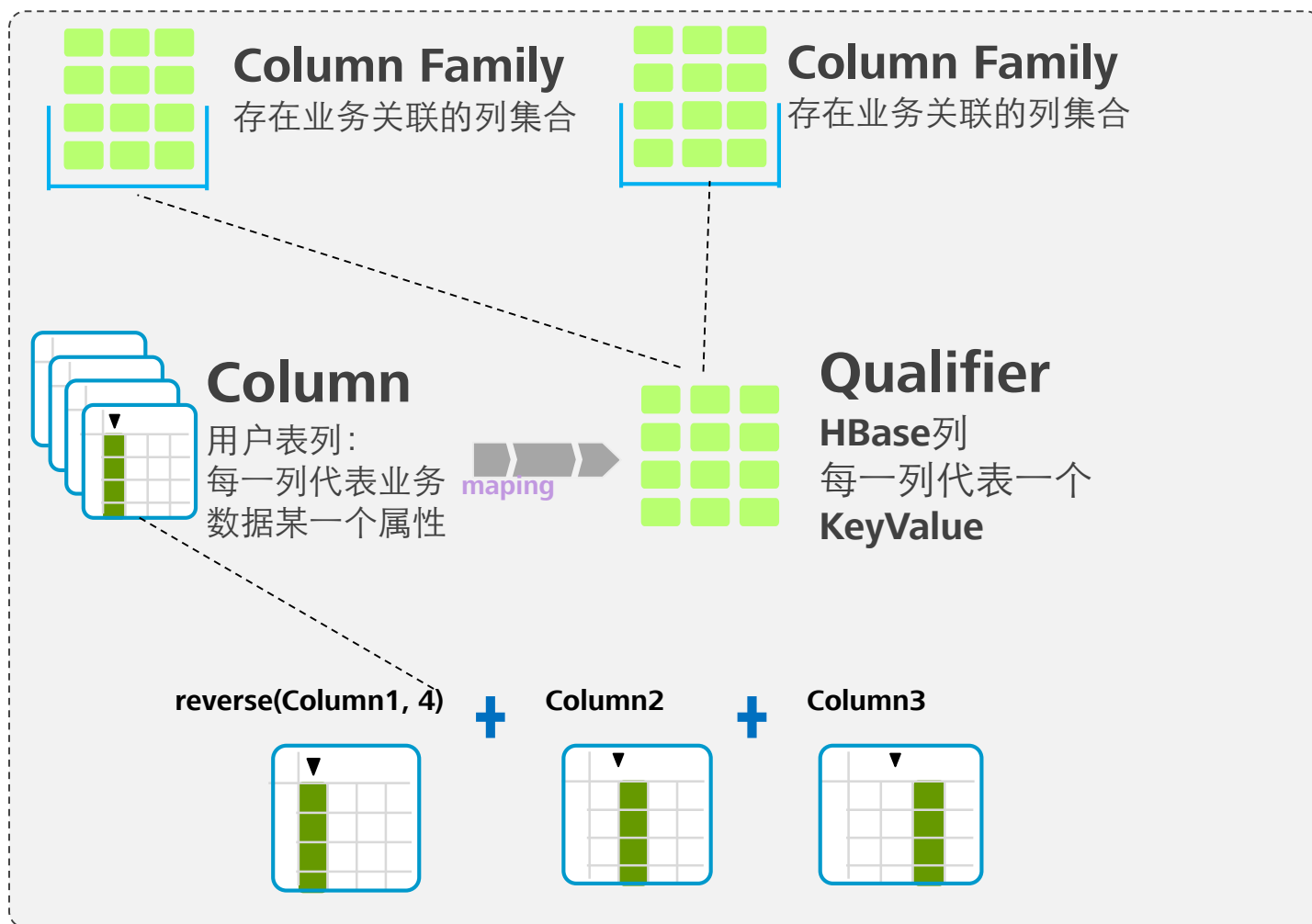
角色管理

多组件UI间单点登陆

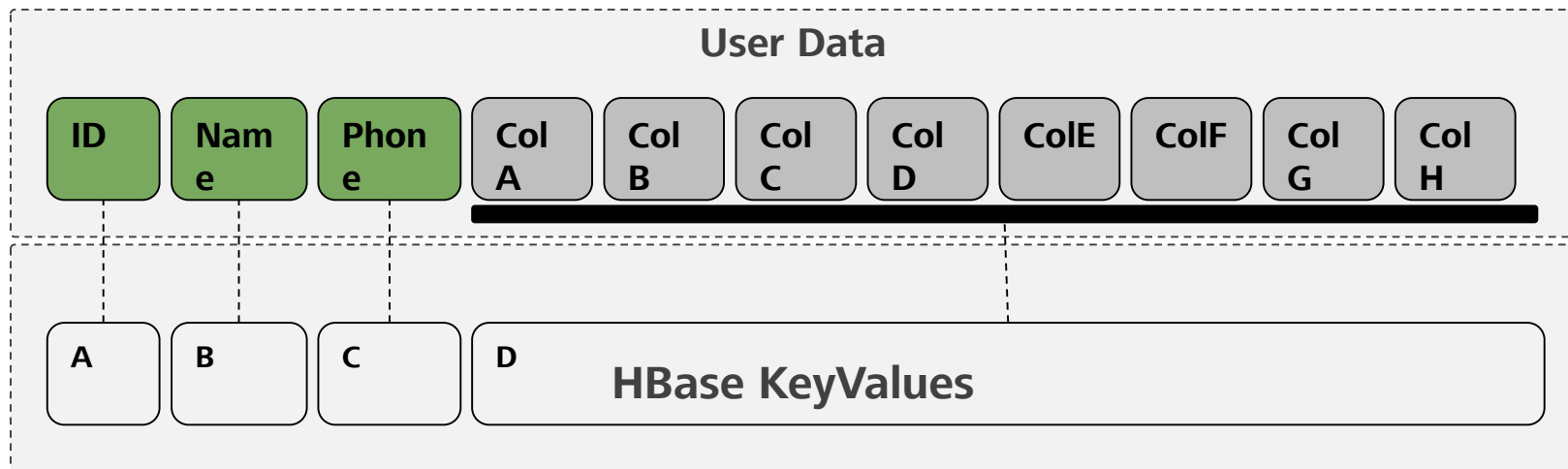


- 提供安全、可靠的统一用户身份校验机制
- 提供多组件UI间单点登陆功能，操作方便易用

HBase可视化建模



HBase应用透明的冷字段合并



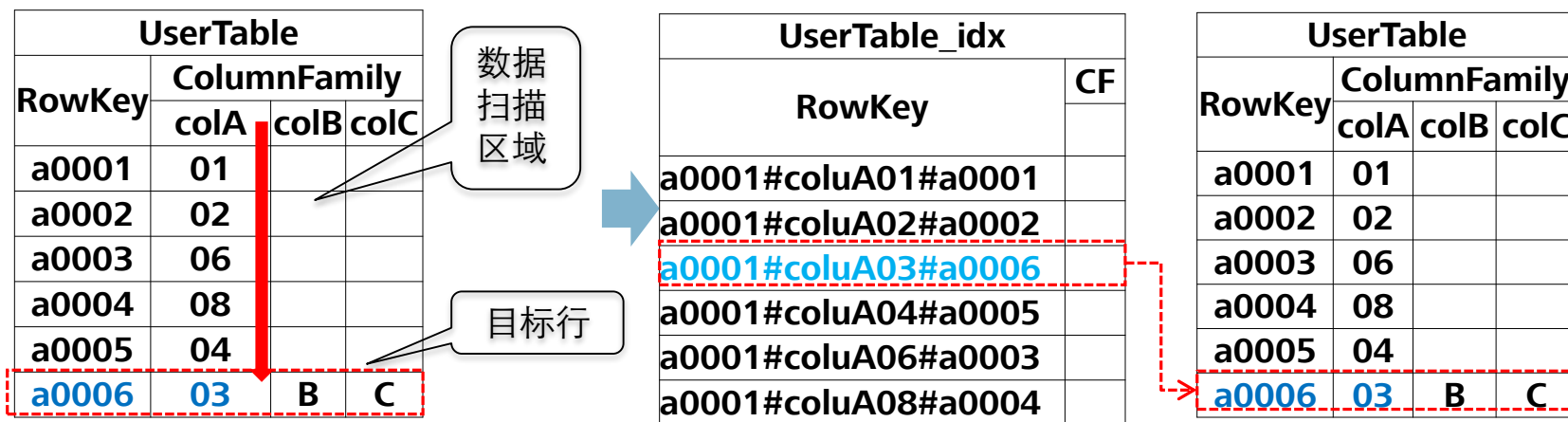
问题

- 在**HBase**列数增多导致数据膨胀率较高，查询性能低下
- 应用层自己合并冷数据列，使得开发复杂度提升，还需要维护相关元数据信息

特性功能

- 提供对应用透明的冷字段合并特性
- 支持实时写入接口及批量导入接口

HBase二级索引



无索引: “Scan + Filter”，扫描大量数据

二级索引: 两次IO定位到数据

- 索引Region与数据Region伴生，统一机制处理
- HBase原生API接口几改动，对用户友好
- 基于Coprocessor机制插件化实现，易升级
- 面向写优化，支持实时写入

Step by Step故障定位

对系统的影响

集群无法提供Yarn服务。用户无法执行新的application。已提交的application无法执行。

可能原因

- Yarn集群中没有主ResourceManager节点。
- Yarn集群中的所有NodeManager节点异常。
- ZooKeeper服务异常。

处理步骤

① 检查Yarn集群中的ResourceManager状态。

1. 在FusionInsight Manager界面，单击“Services > Yarn”。
2. 在“Yarn Summary”中，检查Yarn集群中是否存在主ResourceManager节点。
 - 是，执行[步骤 3](#)。
 - 否，执行[步骤 6](#)。

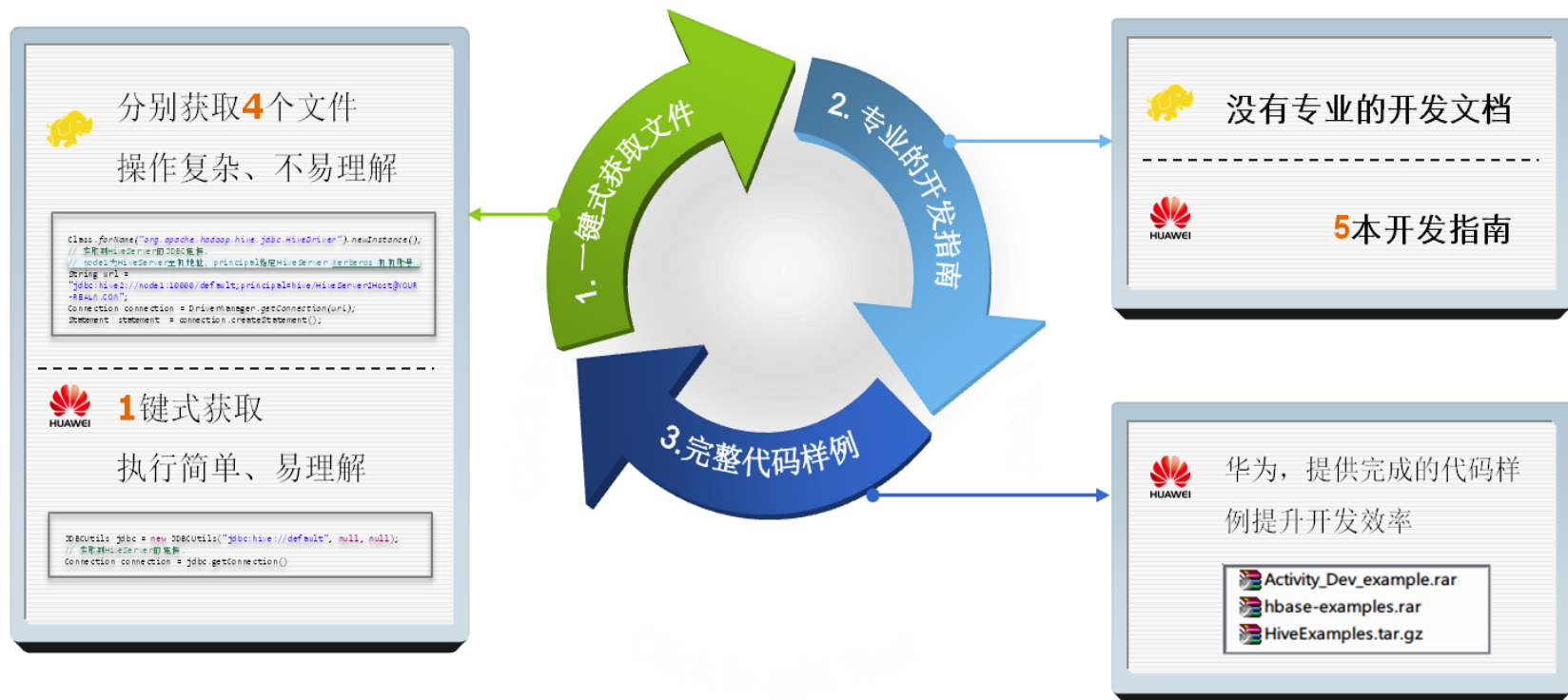
② 检查Yarn集群中的NodeManager节点状态。

3. 在FusionInsight Manager界面，单击“Services > Yarn > Instances”。
4. 查看“NodeManager Health Status”，检查是否有处于非健康状态的节点。
 - 是，执行[步骤 5](#)。
 - 否，执行[步骤 6](#)。
5. 按[ALM-18002 NodeManager心跳丢失](#)或[ALM-18003 NodeManager不健康](#)提供的步骤处理该故障。

③ 检查ZooKeeper 服务状态。

6. 在“Services”中，查看ZooKeeper服务是否处于健康状态。
 - 是，执行[步骤 8](#)。
 - 否，执行[步骤 7](#)。
7. 单击“Restart”重启ZooKeeper实例。检查该告警是否恢复？

应用开发助手



CTBase简化HBase多表业务开发

Transaction表

account_id	amount	time
A0001	\$100	12/12/2014 18:00:02
A0001	\$1020	10/12/2014 15:30:05
A0001	\$89	09/12/2014 13:00:07
A0002	\$105	11/12/2014 20:15:00

AccountInfo表

account_id	account_name	account_balance
A0001	Andy	\$100232
A0002	Lily	\$902323
A0003	Selin	\$90000



CTBase表

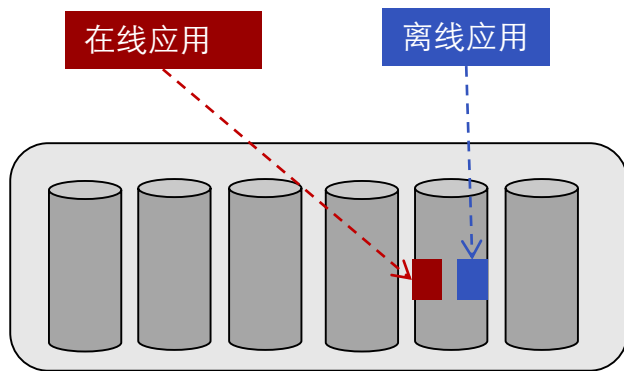
A0001	Andy	\$100232
A0001	\$100	12/12/2014 18:00:02
A0001	\$1020	10/12/2014 15:30:05
A0001	\$89	09/12/2014 13:00:07
A0002	Lily	\$902323
A0002	\$105	11/12/2014 20:15:00
A0002	\$129	11/11/2014 18:15:00
A0003	Selin	\$90000

AccountInfo
表记录

Transaction
表记录

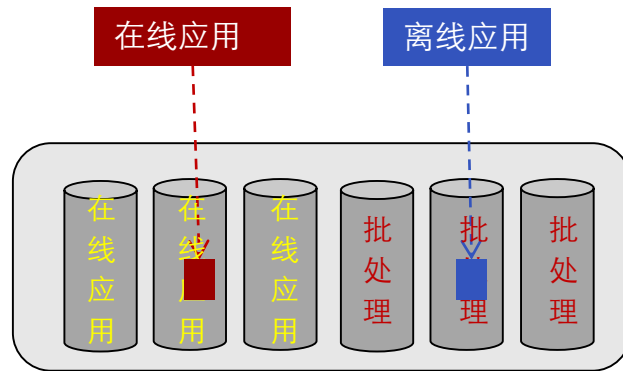
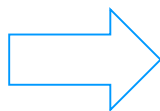
标签存储

IO冲突影响在线业务



HDFS普通存储

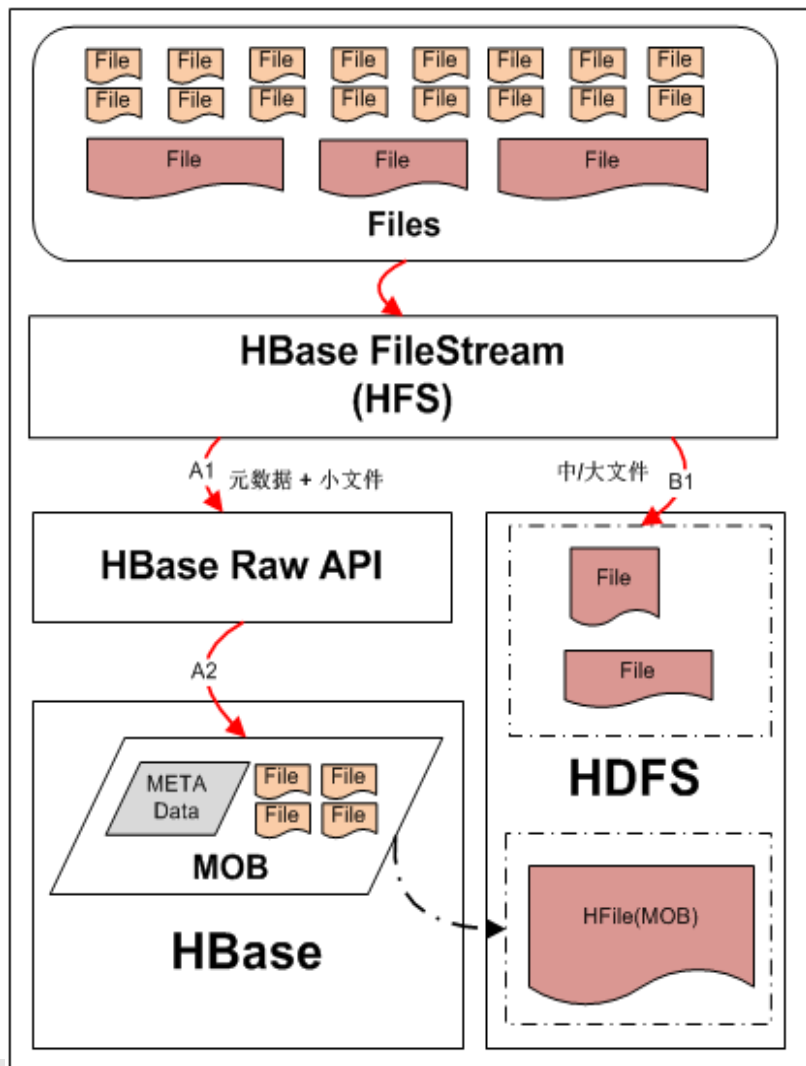
在线应用的数据只存放在有“在线应用”标签的节点上，与离线应用分开存储，避免IO竞争，提升计算的数据本地命中。



HDFS标签存储

- 方案描述：根据应用或物理特征为集群节点打上标签，如“在线应用”，应用数据存放时指定标签存储，即该应用的数据只存在含有指定标签的节点上。
- 应用场景：1.在线与离线应用共享一个集群 2. 特定业务(如在线应用)，运行在特定节点上
- 客户价值：1.不同应用IO隔离，保障应用的SLA 2.提升应用的数据命中，从而提升系统性能

HFS小文件存储与检索引擎



应用场景

- 需要存储大量小文件及关联的描述信息

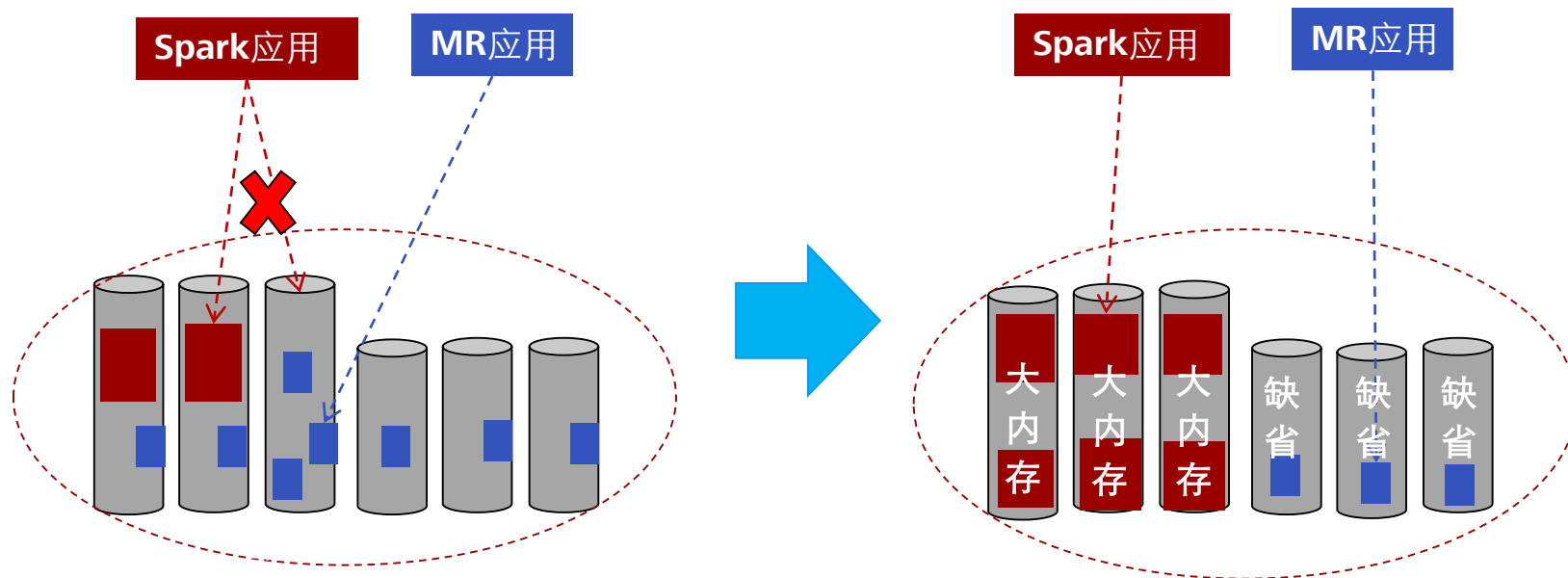
当前问题

- HDFS中存海量小文件会给NameNode带来极大的压力；HBase存海量小文件，Compaction会带来IO资源浪费。

HFS方案价值

- 既可以存储小文件，又可以存储与文件相关的元数据描述信息。
- 提供了统一且友好的访问API
- 能够基于文件的大小，灵活的选择最佳的存储方案
 - 较小的文件，直接存储在MOB中
 - 较大的文件，直接存储在HDFS中

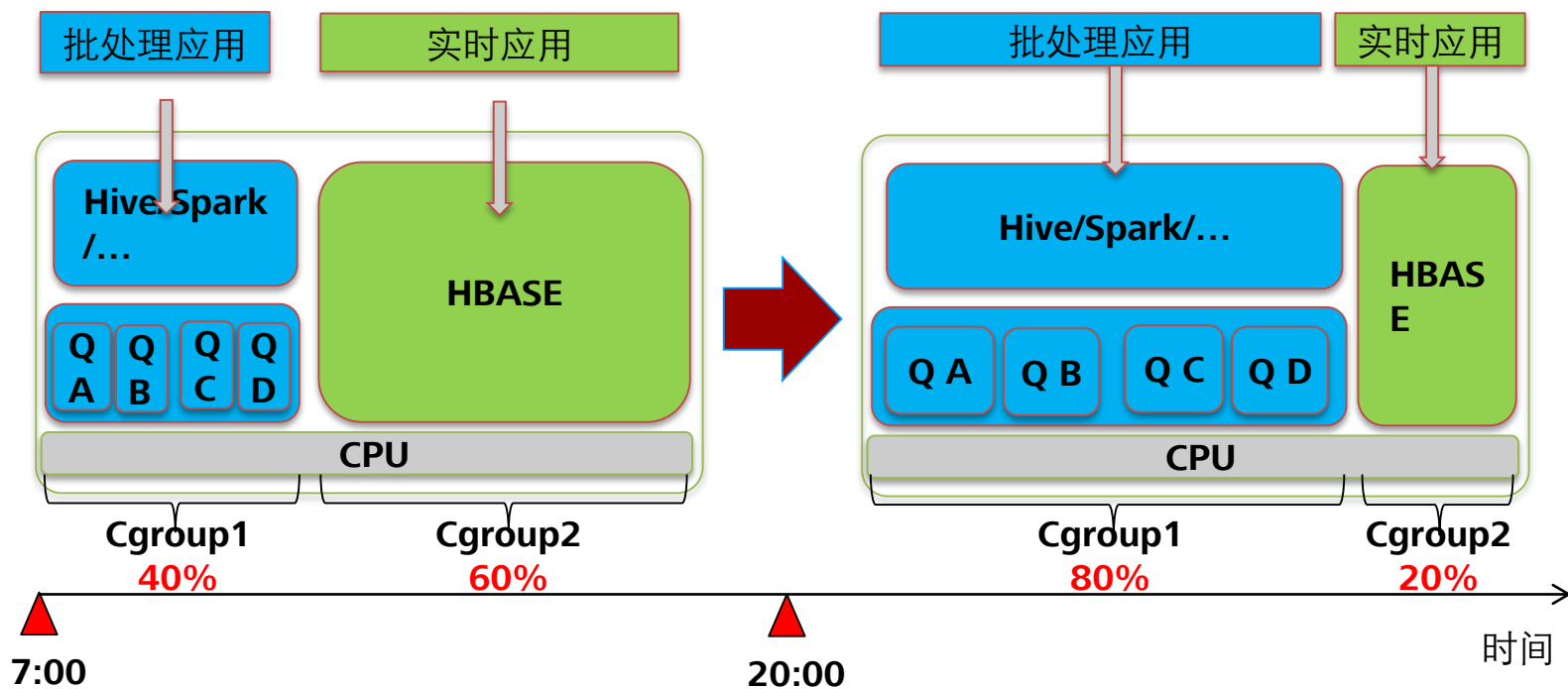
标签调度



应用感知的精细化调度，提升资源利用率

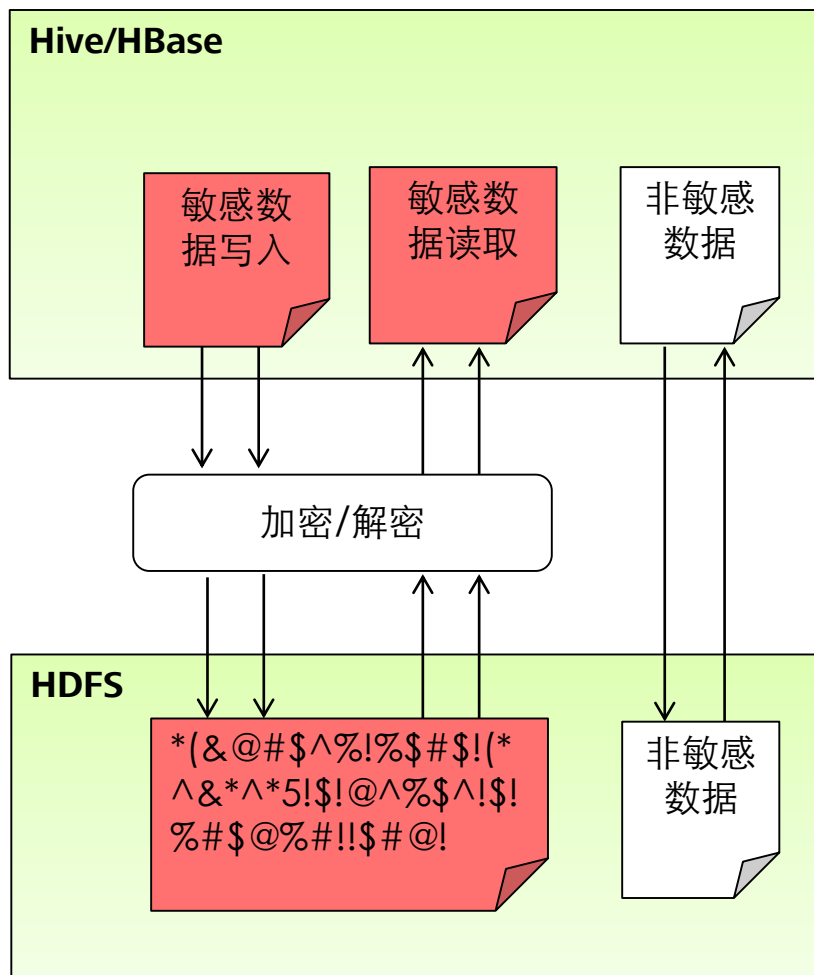
- 不同类型应用（例如在线和批处理）只运行在各自标签的节点上，实现不同应用的计算资源绝对隔离，提升业务SLA
- 对节点硬件有特殊要求的应用，只运行在含有特殊硬件的节点上（如Spark应用需要跑在大内存节点上），资源按需调度，提升资源利用率和系统性能

CPU资源配置周期调整



- **方案描述:** 不同业务在不同时间段分得不同比例的资源，如**7:00**到**20:00**之间，实时业务的高峰，可分给实时业务**60%**资源；而**20:00**到**7:00**，实时业务波谷时，可分配**80%**资源给批处理
- **应用场景:** 不同业务的波峰与波谷相互错开
- **客户价值:** 不同业务在业务波峰时都能分得资源最大化，大幅提升系统资源平均利用率

Hive/HBase细粒度加密



应用场景

- 明文存放会导致某些敏感数据存在泄密的安全隐患

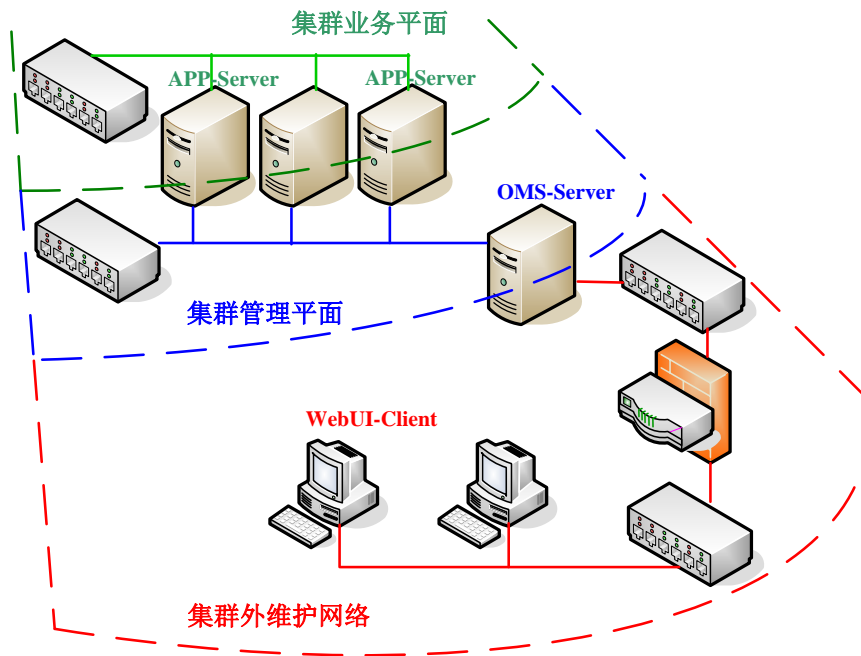
解决方案

- **Hive**支持对表、列加密
- **HBase**支持对表、列族和列加密
- 支持**AES**、国密算法**SM4**等加密算法，同时支持用户自定义的加密算法

用户价值

- 敏感数据按表/列级别细粒度加密存储
- 加密算法的多样性，系统更安全
- 业务透明，加解密过程业务无感知

管理与业务双平面组网，安全可靠



网络类别	信任度	说明
集群业务平面	高	Hadoop 集群核心部件，业务数据在其中存储、流转
集群管理平面	中	仅具备集群管理功能，不接触实际的业务数据
集群外维护网络	低	仅能访问 OMS Server 提供的 Web 服务

网络平面隔离，避免管理与业务带宽抢占，相互干扰

- 管理平面、业务平面、对外维护网络，三网相互隔离，安全可靠，**SLA**有保障
- 避免业务平面的高负载阻塞集群管理通道
- 阻止外部攻击者通过管理通道入侵实际业务数据



目录

1. FusionInsight HD 企业版概述
2. FusionInsight HD 特性介绍
3. FusionInsight HD 成功案例

案例1-某运营商详单分析系统

客户挑战

- 详单数据量年增长**30%**，每月达到**50TB**，传统数据库容量无法支撑，扩容成本高。目前只能支持在线**1**个月话费详单查询
- 传统数据库处理大数据量任务时需执行**6**小时以上，分析速度慢，而且经常由于资源抢占分析不出结果

解决方案

- 企业版**Hadoop**平台：高可靠，高安全，易管理、易开发
- 主备大数据业务集群，**200+**大数据节点
- 拥有内核级开发的工程团队和咨询、服务能力
- 线性扩容，大数据量时并发处理速度快

客户价值

- 更优的服务质量：在线查询**6~24**个月的历史话费详单
- 更佳的处理性能：最大**500**用户并发查询任务，响应时间小于**500ms**
- 提供**ad-hoc**即席查询功能
- **PB**级海量数据存储、计算能力，经分业务并发分析，由**5**天减少到**1**天

案例2-某运营商构建统一大数据平台 挖掘存量经营的价值

客户挑战

- 应用烟囱式建设，多种应用系统独立存储，数据无法共享，跨部门获取数据长达数月
- 缺乏有效的数据资产管理，不清楚到底有多少数据/模型/规则，数据安全隐患大
- 现有系统支撑数据量有限，且数据量越大，分析处理速度越慢

解决方案

- 统一的企业级大数据平台，数据分级存储，一份数据、一份存储
- 统一的数据资产管理和数据安全管控
- 标准的数据共享访问接口和能力开放接口
- 线性扩容，大数据量时并发处理速度不减

客户价值

- 平台数据存储能力达到**PB**级，多应用并发处理速度快
- 应用间数据共享，加快应用开发和部署的速度：新业务推出周期由原来的**1.5**个月减少到**1**周
- 数据资产有效管理，加速挖掘数据价值

案例3-某银行日志分析系统

客户挑战

- 应对互联网金融的竞争，需要掌握网银用户行为轨迹，精准营销，扩大销售；优化网银服务模块的质量，提升客户体验
- 安全、运维操作日志的关联分析，提高故障发生时故障点定位的准确率，提升故障响应速度

解决方案

- 企业级大数据平台：高可靠、高安全、易管理、易开发
- 分布式日志收集系统自动地将各分行的日志收集到总部大数据平台
- 统一管理：分布式日志收集系统+大数据分析平台
- 基于访问日志的用户行为统计和分析模型
- 拥有内核级开发的工程团队和咨询、服务能力

客户价值

- 统一的分布式日志收集+分析大数据平台；超过**500**节点的分布式日志收集，数据收集周期大大缩短
- 基于网银用户行为统计和分析，实时+离线的精准营销
- 安全+运维操作日志关联分析，精准定位故障点

案例4-某银行构筑大数据平台向互联网金融转型

客户挑战

- 面对互联网金融竞争，急需重构以金融大数据分析为基础的决策和服务体系，提升自身竞争力和客户满意度
- 面对金融数据量和种类不断增加，传统数据处理平台仅适合结构化数据处理，扩展性差、扩容成本高，无法满足大数据时代要求

解决方案

- 企业级大数据平台：高可靠、高安全、易管理、易开发
- 第一家支持金融等保、第一家支持**1000**公里以上异地容灾 的海量数据分析和挖掘平台，无缝衔接企业应用
- 解决方案配套能力：数据洞察平台**Miner**、数据服务平台**Farmer**
- 拥有内核级开发的工程团队和咨询、服务能力

客户价值

- 统一的金融大数据全生命周期解决方案
- 丰富的创新业务：在线明细，实时征信，精准营销 ...
- 小微贷获客预测，比传统方式提升**40**倍的转化率；或有金融资产预测误差率降低一倍；信用卡征信时间将由**2**周左右降至**2 - 5**秒

案例5-某银行构架第二数据管理平面

客户挑战

- 业务系统中，对于历史数据保存和访问的需求越来越多，但是缺少统一的数据归档和查询平台
- 数据量越来越大，特别是票据影像信息，海量小文件，存储压力大
- 原有系统对于历史数据查询流程复杂，效率低下

解决方案

- 企业级大数据平台：统一的历史数据归档、查询大数据平台。提供海量数据归档存储和高性能查询能力
- 小文件解决方案：小文件合并存储解决方案，解决海量小文件问题
- 集群部署：在线查询集群独立部署，提供高性能的数据在线查询业务体验

客户价值

- 建设统一的大数据平台，实现全行历史数据和票据影像数据的统一存储管理
- 实现建行海量的票据影像小文件存储（XXPB级）
- 大幅提升系统并发查询能力，同时支持上层柜台、监保等多个应用

案例6-某银行建设大数据平台 推动金融大数据转化为生产力

客户挑战

- 在互联网金融时代，急需建设统一的大数据平台，在此基础上开发创新的新型金融产品，推动金融大数据转化为真正的生产力
- 金融数据量和种类不断增加，传统数据平台在处理、存储、扩容能力等方面已无法满足要求

解决方案

- 企业级大数据平台：高可靠、高安全、易管理、易开发
- 提供大数据批处理、即席查询、实时查询任务的统一调度和管理
- 拥有内核级开发的工程团队和咨询、服务能力
- 多家金融企业应用，规划设计、实施经验丰富

客户价值

- 大数据平台存储扩展可达**PB**级，提供统一大数据任务管理能力
- 大数据业务：交易数据仓库、橙e网 行为日志分析、在线明细 ...
- 交易数据仓库，从只能存储**3**个月的交易数据提升到**1**年；在线历史交易明细查询，由只能查询**1**年内明细提升到**6**年内明细

思考题

1. **FusionInsight HD** 有哪些特性?
2. **Hive/HBase**细粒度加密支持哪些加密算法?
3. **HDFS**中存海量小文件会给**NameNode**带来极大的压力;
HBase存海量小文件, **Compaction**会带来**IO**资源浪费。针对这种现象时有什么技术解决?
4. 日志级别调整有哪些级别?



总结

- 本章介绍了华为大数据产品**FusionInsight HD** 企业版的相关概况，并且重点介绍了**FusionInsight HD** 特性及其使用场景。最后介绍了华为**FusionInsight HD** 在业界内取得的众多成功案例。



习题

- 判断题

1.**Hive**支持对表、列加密，**HBase**支持对表、列族和列加密。

2.用户权限管理基于角色的访问控制，提供可视化的多组件统一的集中用户权限管理。

- 选择题

FusionInsight HD 系统高可靠性体现在哪些方面（）

A.所有组件无单点故障

B.所有管理节点HA

C.软硬件健康状态监控

D.网络平面隔离

FusionInsight资料链接

http://enterprise.huawei.com/topic/FusionCloud_2013/index.html

Thank you

www.huawei.com