



FusionInsight HD解决方案介绍

LEADING NEW ICT

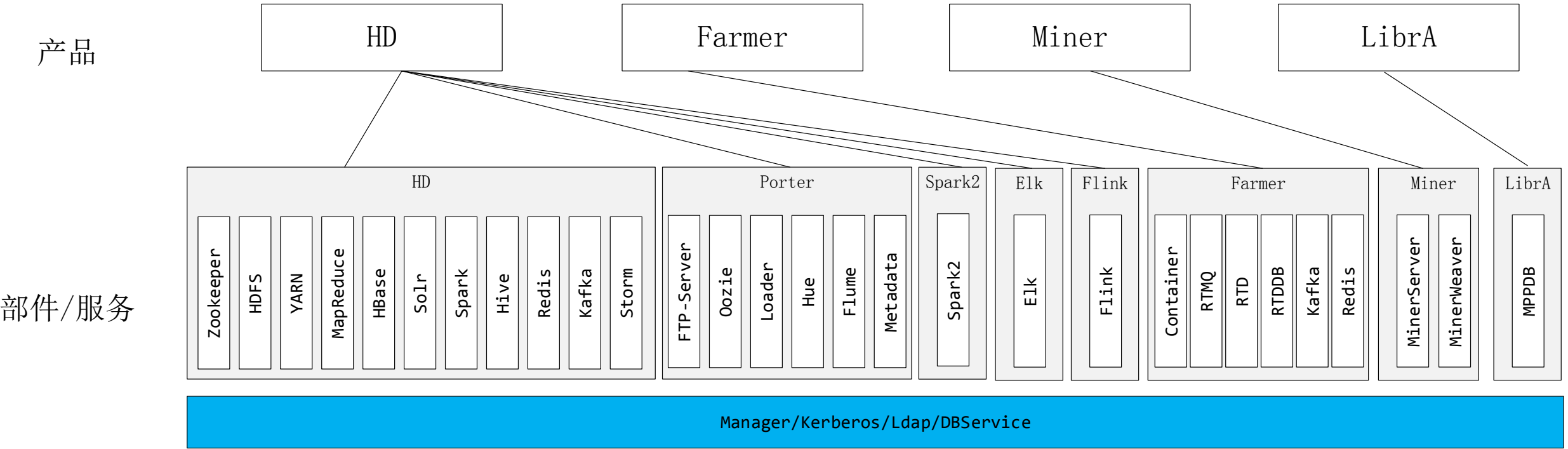
目录

1 FusionInsight 简介

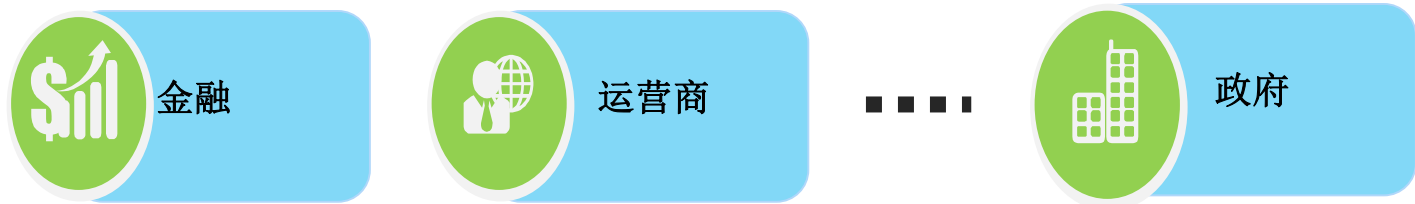
2 FusionInsight HD 介绍

3 FusionInsight HD 成功案例

产品/部件/服务视图



FusionInsight架构和策略：分层解耦开放



FusionInsight 大数据平台

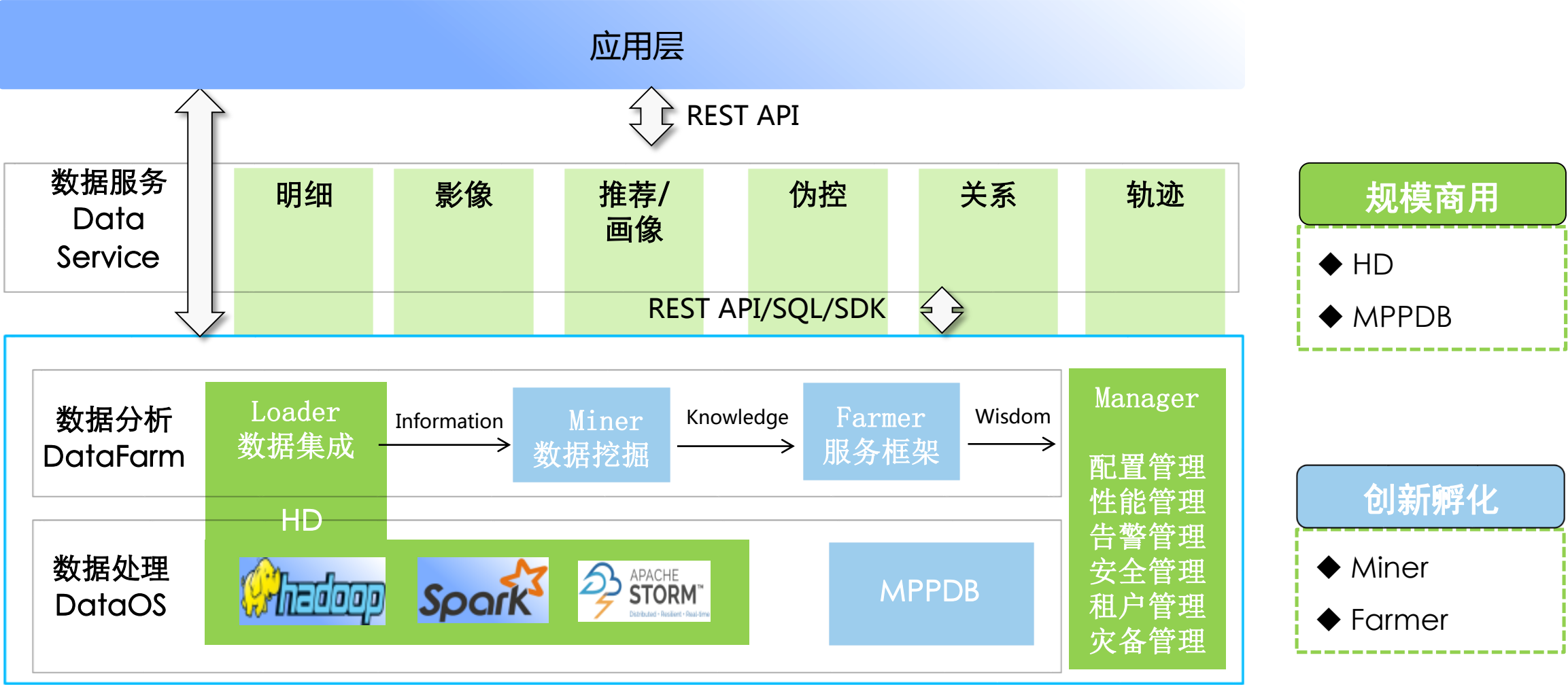


- 敏捷—行业最优化
 - 完全开放的架构，性能线性扩展
 - 强大的SQL能力，业务移植便捷
 - 丰富的工具支持，开发运维高效

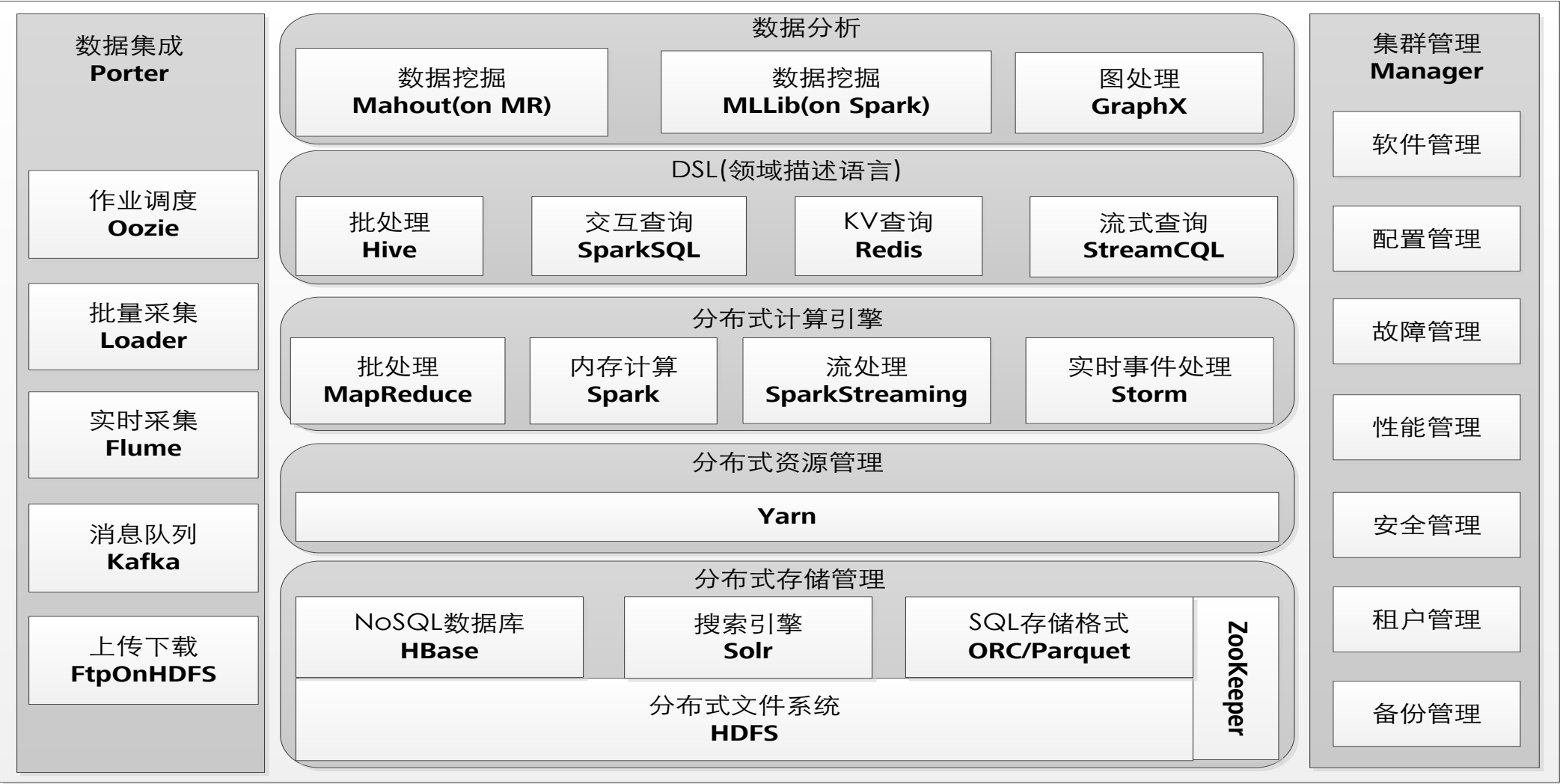
- 智慧—场景最优化
 - 全量建模，深刻洞察
 - 自研算法，高效精准

- 可信—企业级的质量和体验
 - 全组件HA、异地容灾、金融等保
 - 开放共赢，可信赖的合作伙伴

FusionInsight产品族：全栈平台产品



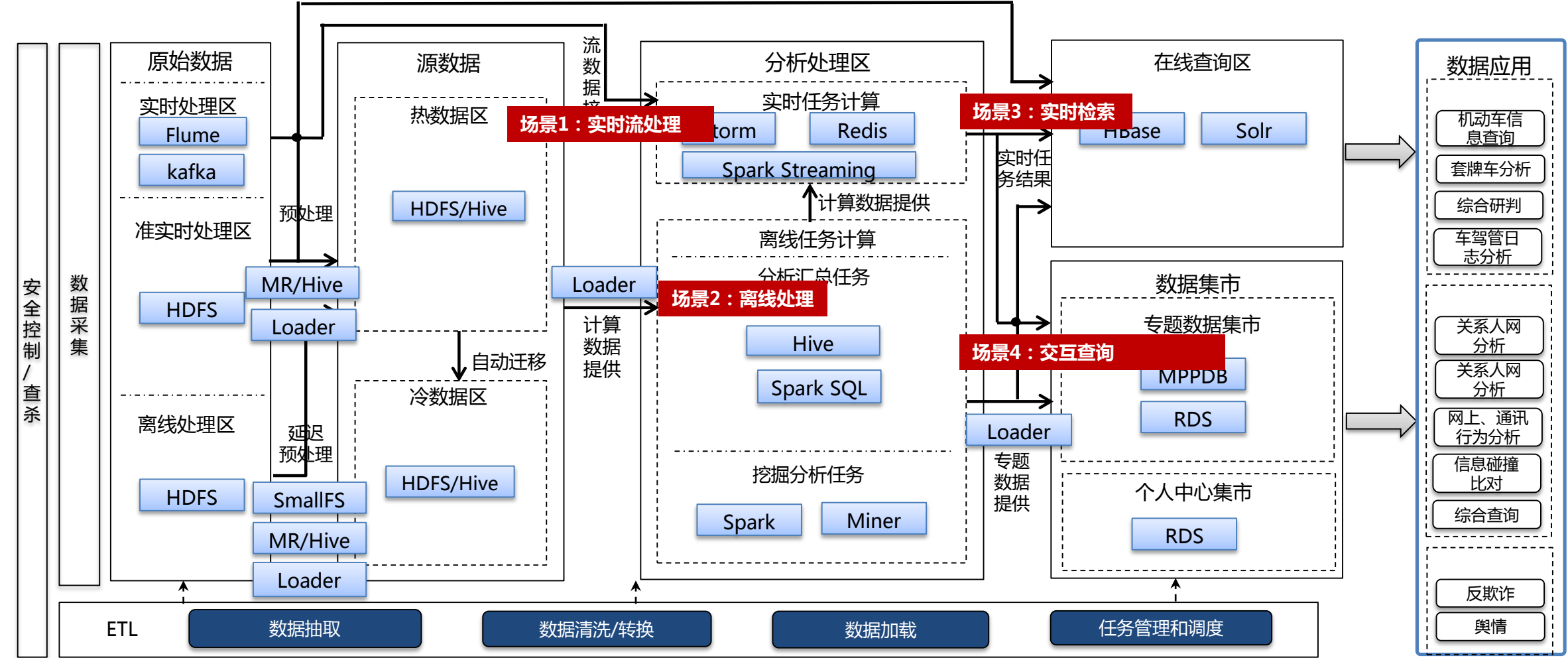
FusionInsight HD解决方案架构



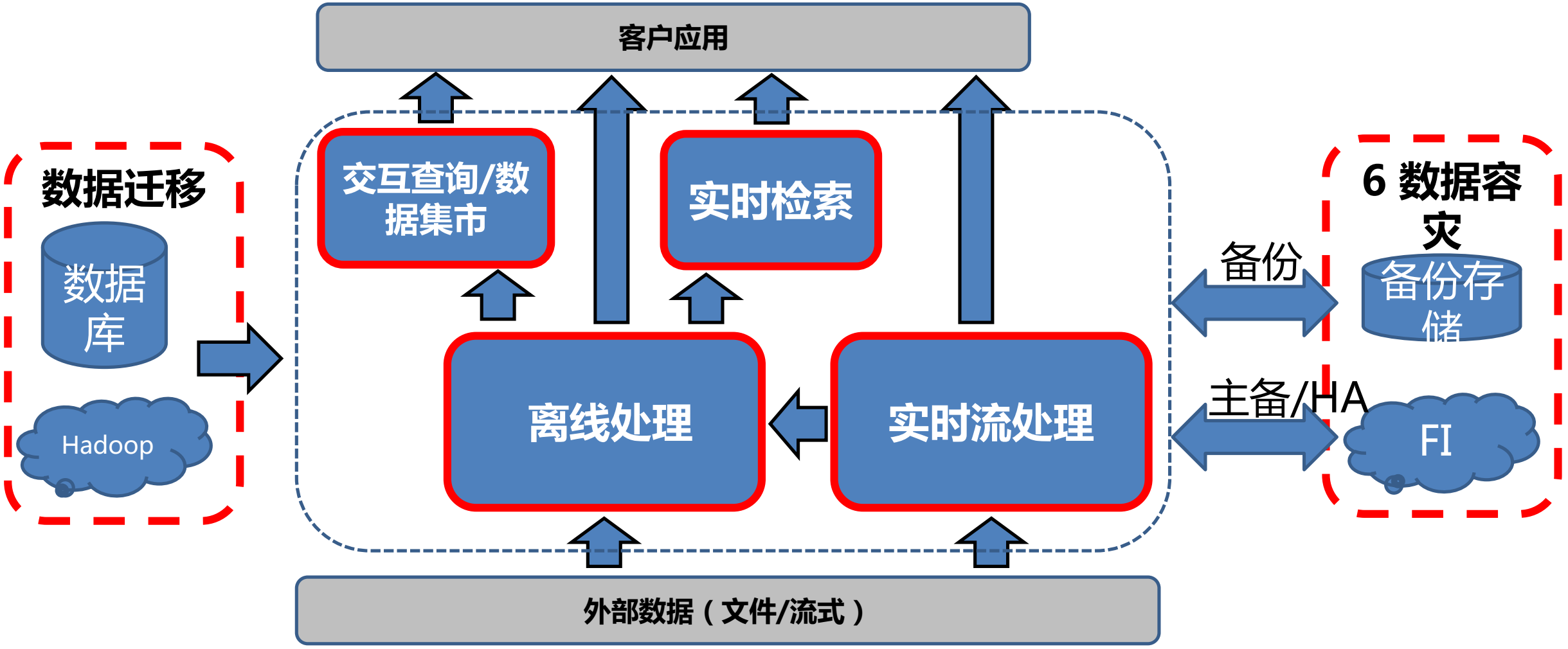
目录

- 1 FusionInsight 简介
- 2 FusionInsight HD 介绍
- 3 FusionInsight HD 成功案例

数据架构



几个解决方案的关系



离线处理场景 (场景1) 既可以独立交付, 也常与其他场景组合交付, 离线处理场景负责数据的加工处理, 为其他场景提供数据。

典型场景下的方案总览1

解决方案	主要组件	适用场景	典型特点
离线处理	HDFS Yarn Spark Hive MapReduce SparkSQL Load	对海量数据进分析和处理，形成结果数据，供下一步数据应用使用。离线处理对处理时间要求不高，但是所处理数据量较大，占用计算存储资源较多，通常通过MR或者Spark作业或者SQL作业实现	<ul style="list-style-type: none"> ● 处理时间要求不高 ● 处理数据量巨大，PB级 ● 处理数据格式多样 ● 多个作业调度复杂 ● 占用计算存储资源多 ● 支持SQL类作业和自定义作业 ● 容易产生资源抢占
实时流处理	Flume Kafka SparkStreaming Flink Storm	对实时数据源进行快速分析，迅速触发下一步动作的场景。实时数据对分析处理速度要求极高，数据处理规模巨大，对CPU和内存要求很高，但是通常数据不落地，对存储量要求不高。实时处理，通常通过Spark Streaming或者Flink任务实现。	<ul style="list-style-type: none"> ● 处理时间要求极高，毫秒级 ● 处理数据量巨大，每秒数百兆 ● 占用计算资源多 ● 容易产生计算资源抢占 ● 数据格式以各种网络协议格式为主 ● 任务相对简单 ● 数据不落地、存储量不大

典型场景化方案总览2

解决方案	主要组件	适用场景	典型特点
实时检索	ElasticSearch Solr Hbase Yarn	<p>指数数据实时写入，对海量数据基于索引主键实时查询，查询响应要求较高，查询条件相对比较简单。查询条件复杂的可以根据关键词在全域数据中通过索引搜索主键后，通过主键查询。全域数据既包含了结构化数据又包含了文本数据。</p> <p>例如：按照卡口和时间为主键快速查询过车数据，按照车辆为主键查询车辆轨迹数据</p>	<ul style="list-style-type: none"> ● 查询响应时间要求较高，几秒内 ● 高并发 ● 处理数据量巨大，PB级 ● 能够同时处理结构化和非结构化的数据 ● 全文检索功能 ● 近实时索引
交互分析	Elk	<p>对数据进行交互式的分析和查询，查询响应要求较高，能够实现人机之间交互，查询通常比较复杂。交互查询的数据通常已经被预处理过，按照适合交互查询的数据模型进行组织。交互查询数据量巨大，对CPU和内存要求很高，对于存储要求也很高。交互查询，以复杂SQL查询最为常见，多维Cube分析也比较常见</p>	<p>典型特点：</p> <ul style="list-style-type: none"> ● 查询相应时间要求较高，3秒内 ● 处理数据量巨大，百TB级 ● SQL支持度要求高，达到数据库水平 ● 支持数据更新和插入 ● 多种查询并发，容易产生资源抢占 ● 支持第三方展现和报表工具

HDFS原理简介 - 分布式文件系统

✓ HDFS主要特点：

- 存储大文件
- 将大文件分割成很多小块存储
- 流式数据读取，“write one read many”
- 本身是分布式的，具备良好的可扩展性
- 通过放开POSIX要求，极大改善数据读写性能

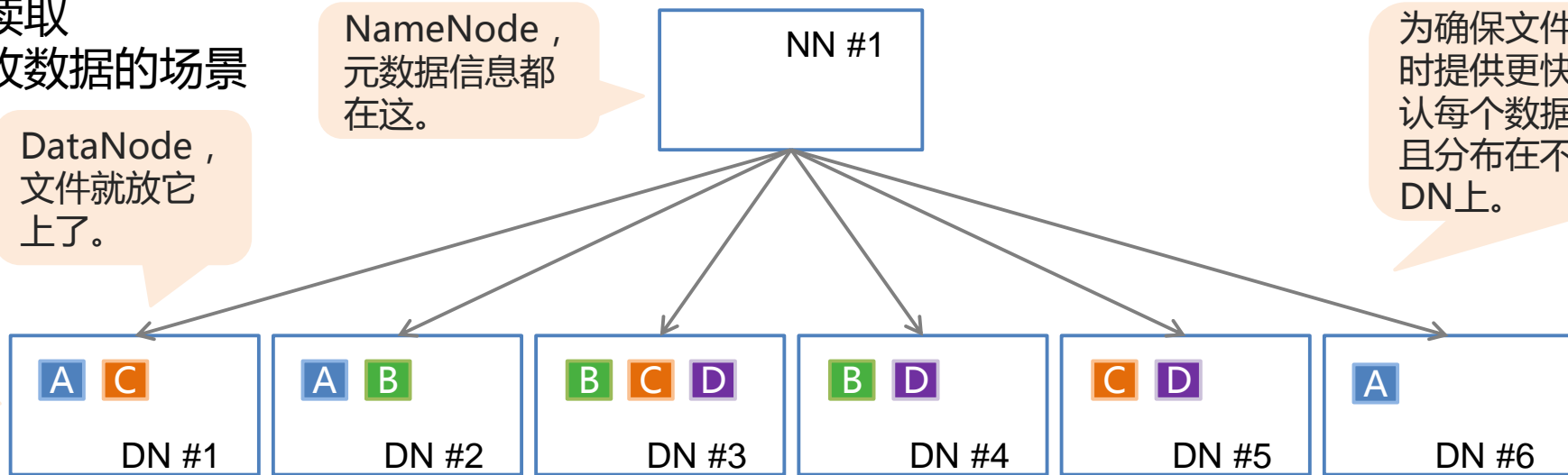
✗ HDFS**不适合**于：

- 存储大量小文件（<1MB）
- 实时数据读取
- 需经常修改数据的场景

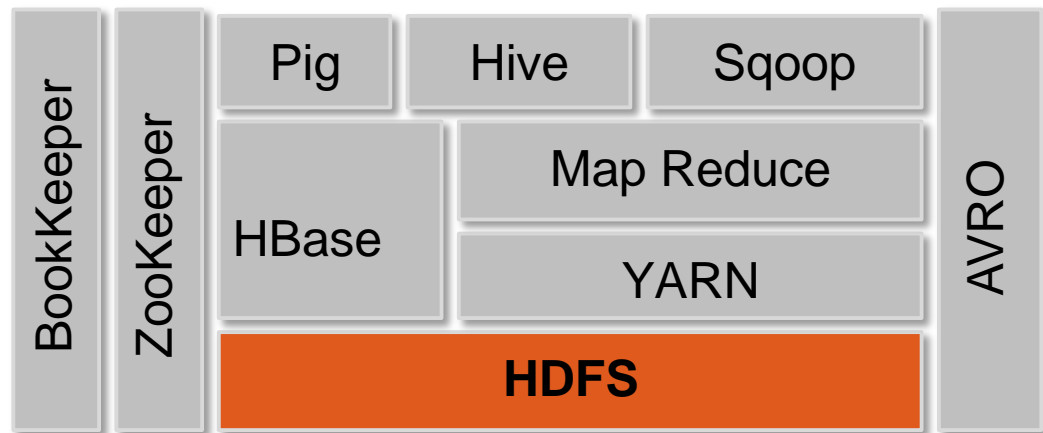
文件被切分成大小相同的块（最末尾的块可能小于块大小），并存储在不同的数据节点上。

DataNode，文件就放它上了。

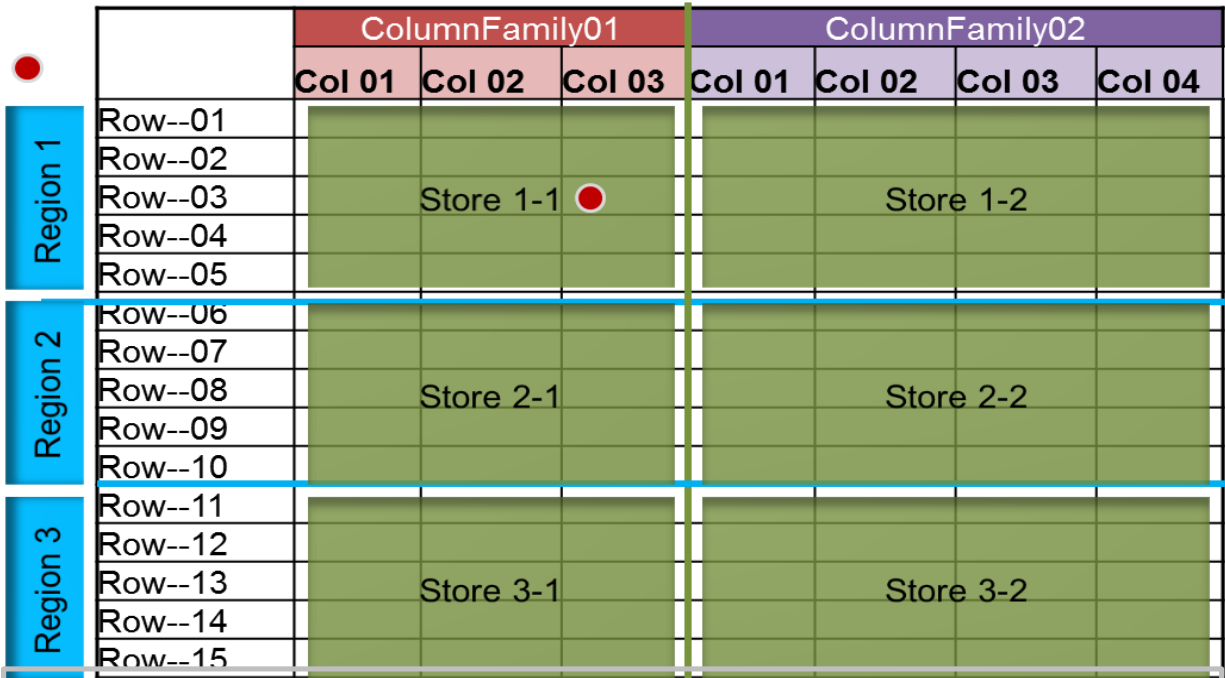
NameNode，元数据信息都在这。



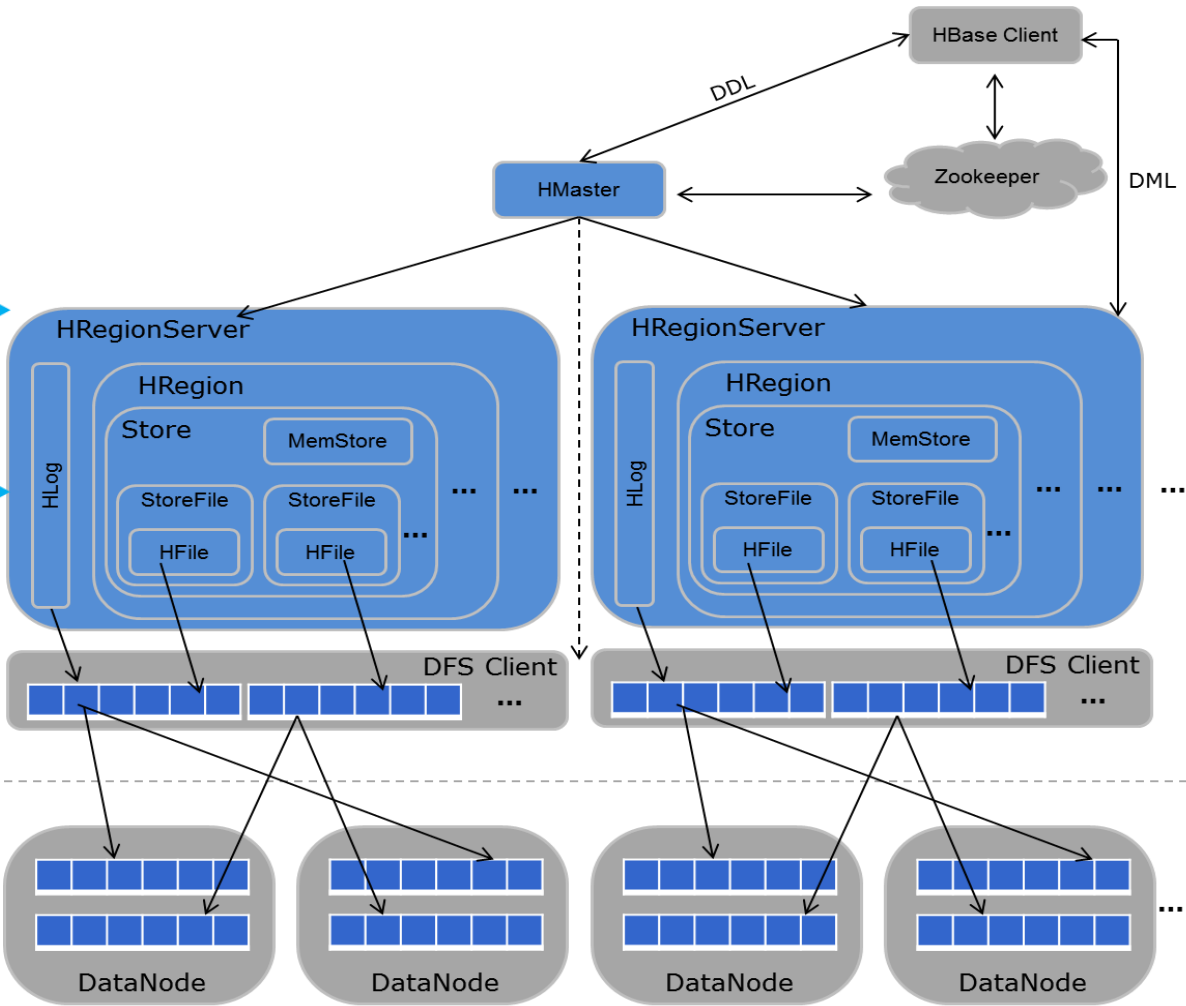
为确保文件块的容错性，同时提供更快的数据读取，默认每个数据块有3个副本，且分布在不同的数据节点DN上。



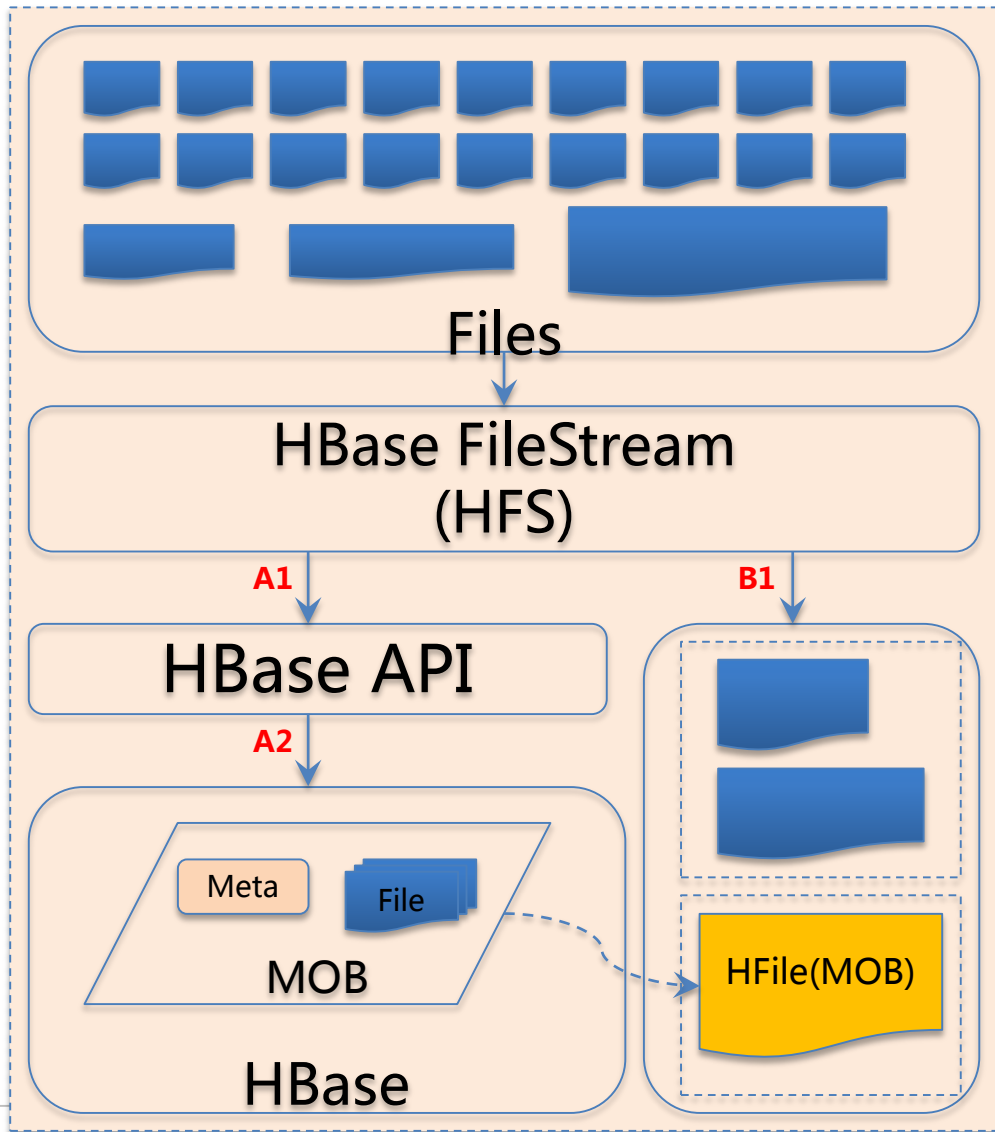
HBase原理简介 - 分布式数据库



HRegionServer内部管理了一系列HRegion对象，每个HRegion对应Table中的一个Region。HRegion由多个Store组成。每个Store对应Table中的一个Column Family的存储，即一个Store管理一个Region上的一个列族（CF）。每个Store包含一个MemStore和0到多个StoreFile。Store是HBase的存储核心，由MemStore 和 StoreFile组成。



HFS：高效的小文件存储与查询



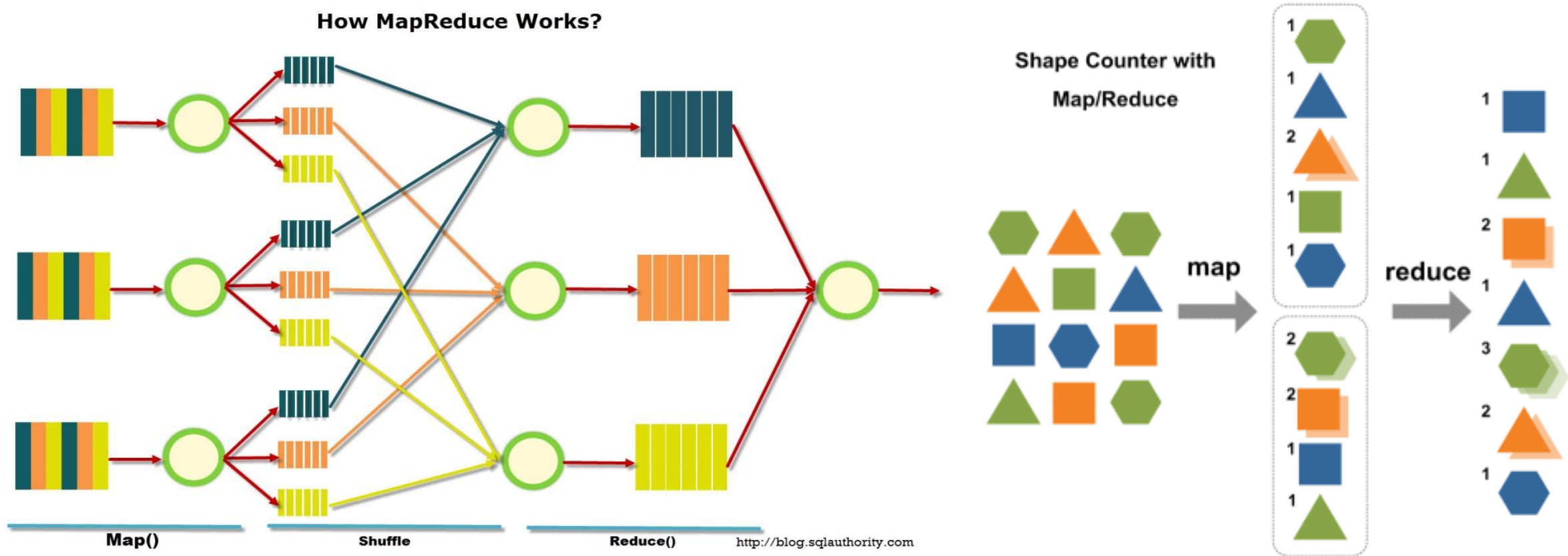
HFS方案出现的背景：

无论是HDFS，还是HBase，均不适合小文件的存储方案：如果将小文件直接存储在HDFS中，会给NameNode带来极大的压力；而如果将小文件直接存储在HBase中，也将会给HBase的Compaction带来较大的压力，浪费大量的IO资源。HFS的出现，就是为了解决在Hadoop中存储小文件的问题。

HFS方案的特点：

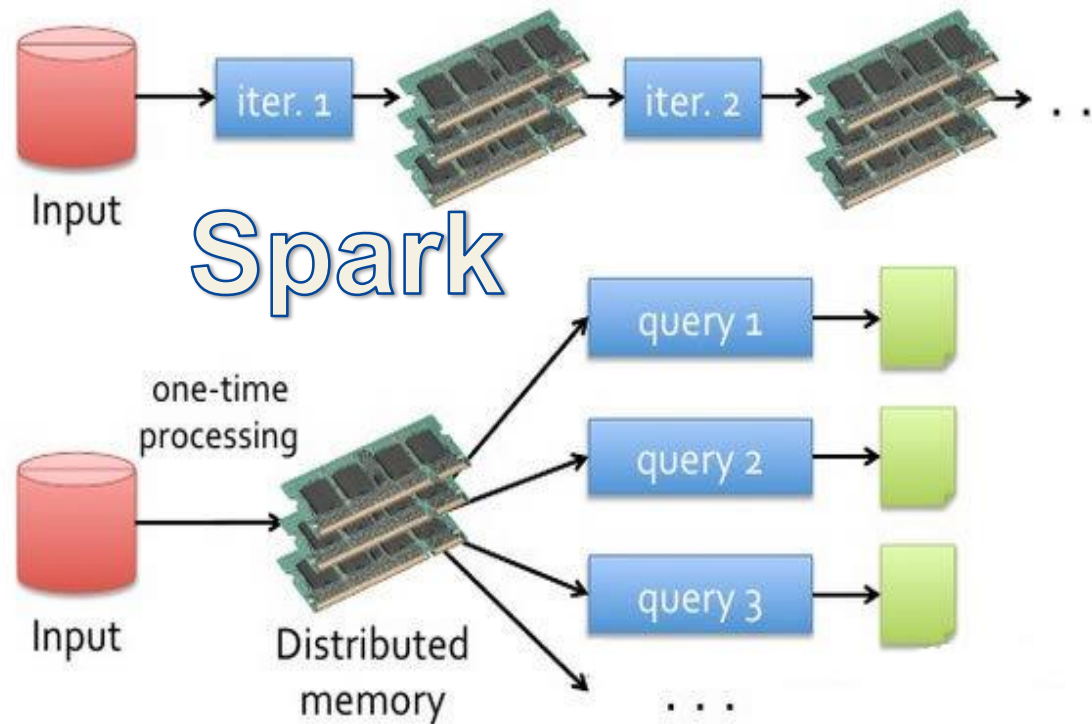
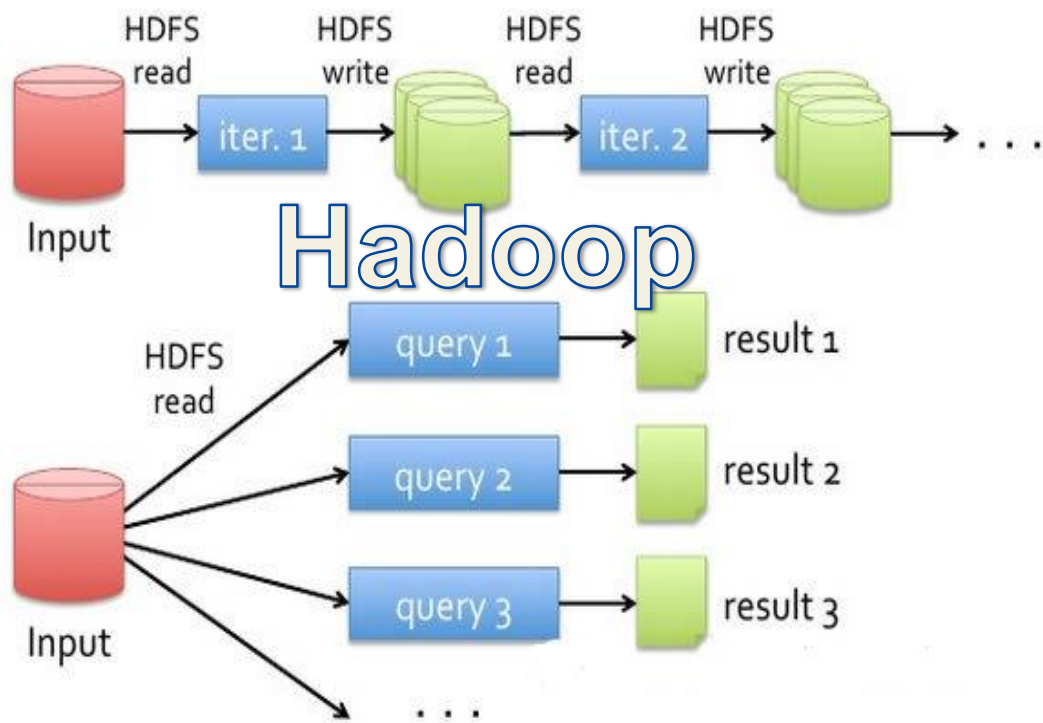
- ◆ 既可以存储小文件，又可以存储与文件相关的元数据描述信息
例如，对于图片信息，不仅仅可以存储图片本身，还可以存储针对这些图片的一些描述信息，例如，拍摄地点，拍摄时间，备注信息等等
- ◆ 提供了统一且友好的访问API:
 - 对于文件，提供了用户习惯的File Stream访问接口
 - 对于文件元数据信息，通过HBase接口进行访问
- ◆ 能够基于文件的大小，灵活的选择最佳的存储方案：
 - 较小的文件，直接使用HBase的MOB方案进行存储
 - 较大的文件，直接存储在HDFS中

MapReduce原理简介 - 分布式计算架构



Apache MapReduce是google MapReduce的开源实现。是对并行计算的封装，使用户通过一些简单的逻辑即可完成复杂的并行计算。其核心理念是将一个大的运算任务分解到集群每个节点上，充分运用集群资源，缩短运行时间。

Spark 迭代计算框架：重构M-R，优于Hadoop



- Spark是UC Berkeley AMP 实验室基于map reduce算法实现的分布式计算框架，**输出和结果保存在内存中**，不需要频繁读写HDFS，数据处理效率更高
- Spark适用于**近线或准实时、数据挖掘与机器学习**应用场景

Storm：流式数据处理框架，实时的Hadoop

Hadoop

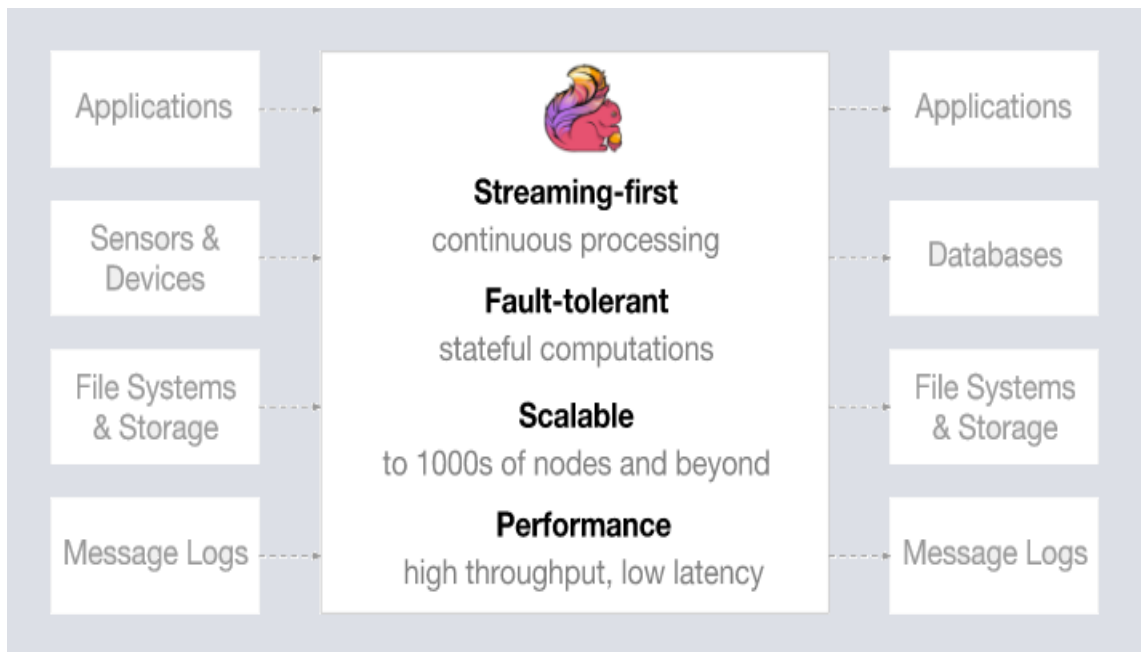


Storm



Storm 广泛应用于**实时分析，在线机器学习，持续计算、分布式远程调用**等领域。

Flink：流数据处理引擎



Flink流数据处理引擎是华为主打的流处理系统，具有**高吞吐、低时延，高可靠**等特性，在生态上具有完备的生态系统，支持和如下系统的对接：

- 消息队列（如kafka）；
- 文件存储系统（如HDFS）；
- DataBase（如cassandra）；

Flink适用于所有要求低时延的流处理场景，支持但不限于如下场景：

- 流式ETL场景：实时数据转换、抽取场景；
- 复杂事件处理场景：实时的过滤、聚合、join等场景；
- 实时检测场景：实时的检测场景；

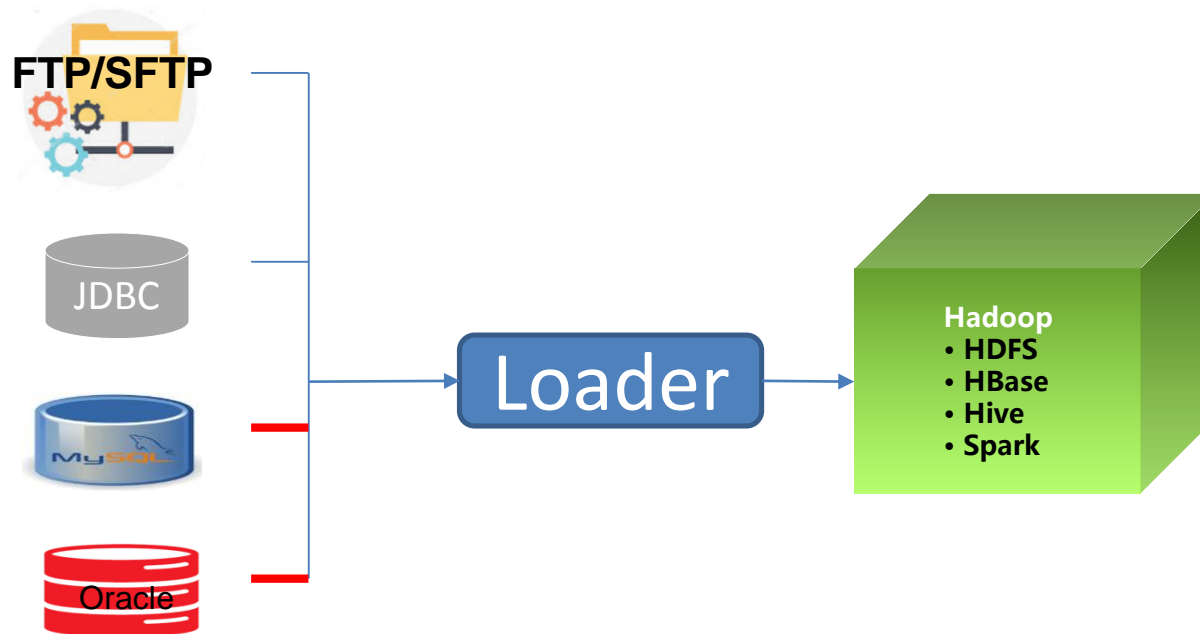
关键特性：

- **低时延**：支持ms级时延；
- **高可靠**：支持异步快照机制将用户作业状态进行备份，支持用户作业有状态的恢复，真正做到每个事件exactly once的处理；
- **高扩展**：Taskmanager支持手动水平扩展；
- **高易用**：
 - 提供Netty connector、kafka connector等connectors和外界系统进行对接；
 - 提供DataStream API、StreamSQL等API能力方便用户书写作业流程图；
 - 提供异步IO查询外部数据库（如redis）的能力；
- **高安全**：支持用kerberos等进行认证，支持用SSL加密传输；

增强能力：

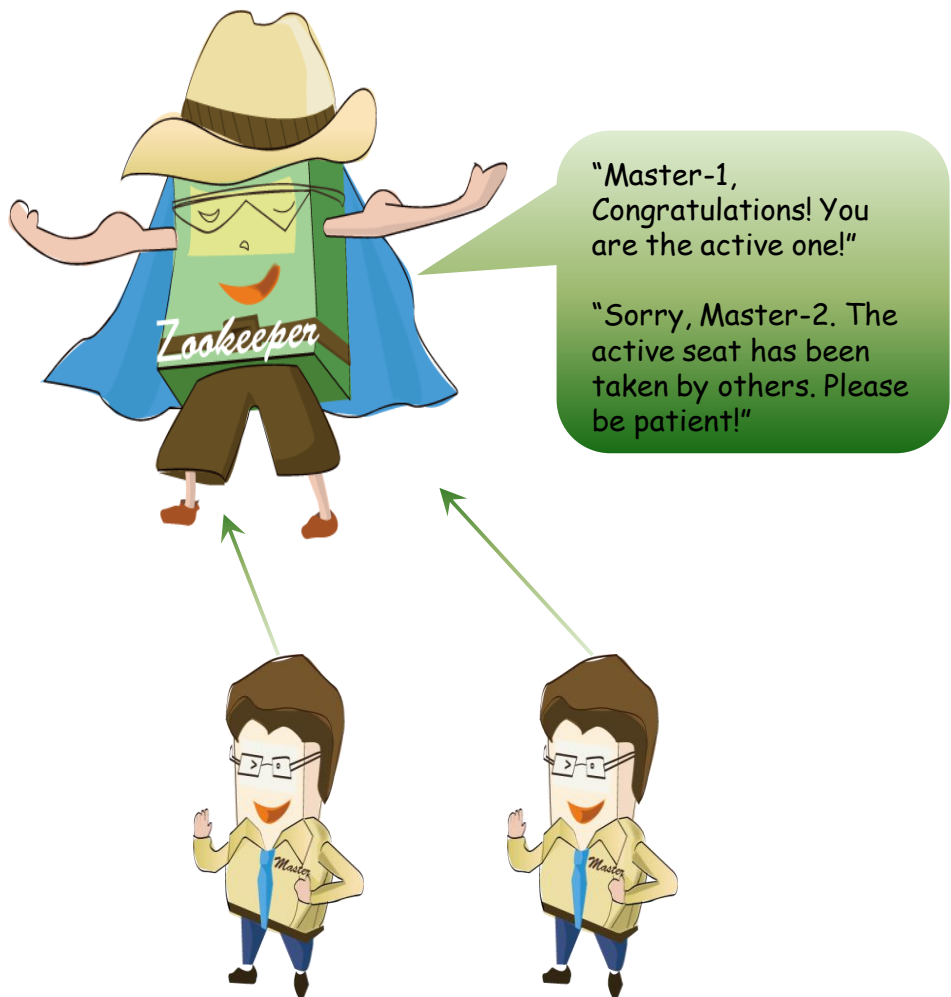
- **作业Pipeline**：多个Flink作业串联起来，完成复杂的流业务分析
- **滑动窗口优化**：重构滑动窗口组织方式，**内存节省1~10倍**
- **配置表支持**：基于DataStream API，支持导入用户配置表，流和配置表join
- **支持窗口数据共享**：对同一个数据流定义不同窗口，且窗口有重叠时，支持重叠窗口中的数据只存一份，做到数据共享，节约内存；

Loader：高性能的数据采集工具



- Loader的定位是数据加载，只能做简单的数据转换和迁移，复杂的数据加工处理需要加载到在Hadoop内部后，使用离线处理作业处理
- Loader从数据库直接加载速度慢，建议数据库数据源在接口机上落地成文件后再加载

ZooKeeper简介



- 提供分布式锁的服务。

例如，多个Master进程竞争主Master角色时，怎么样保证仅有一个Active角色存在？这就需要一个分布式的锁机制来保证。多个Master进程都尝试着去ZooKeeper中写入一个对应的节点，该节点只能被一个Master进程创建成功，创建成功的Master进程就是Active角色。

- 提供了事件侦听机制。

例如，主Master进程宕掉之后，其它的备Master如何能够快速的接管？这个过程中，备Master在侦听那个对应的ZooKeeper节点。主Master进程宕掉之后，该节点会被删除，那么，其它的备Master就可以收到相应的消息。

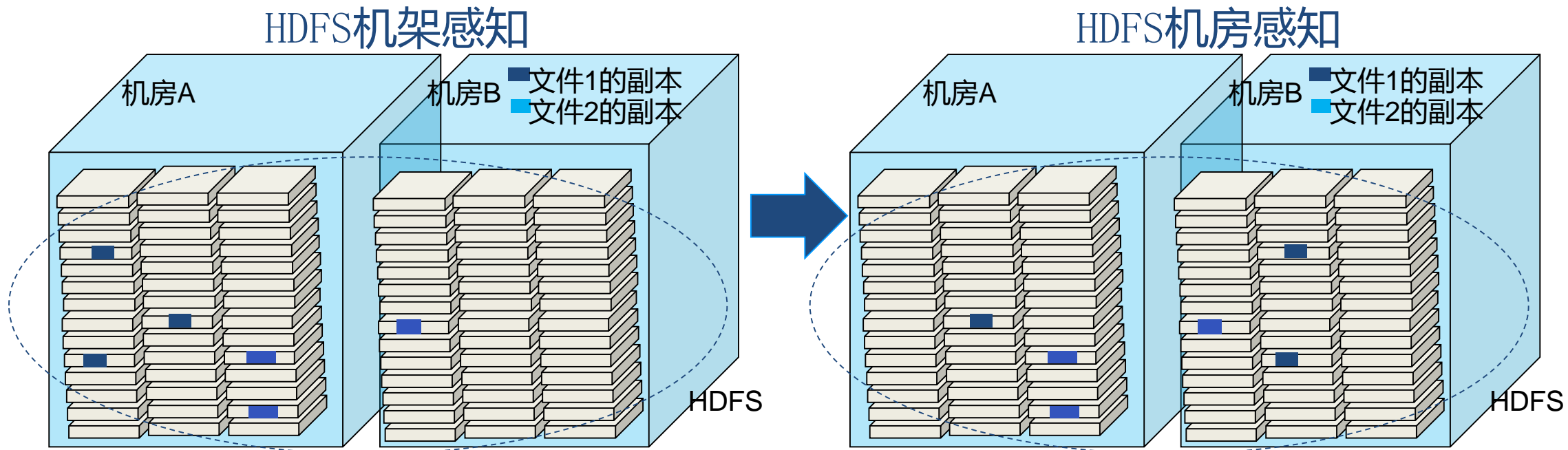
- 个别场景，可充当一个微型数据库角色。

例如，在ZooKeeper中存放了Root Region的地址（Root Region原来是存在ZooKeeper中的！），此时，可以将它理解成一个微型数据库。

HDFS机房感知，保障跨机房数据可靠性

只能保证文件的副本跨机架存放。文件1的所有副本都在机房A，当机房A整体故障（如断电），文件1有数据块丢失

保证文件副本跨机房存放，即每个文件至少有一个副本存放于另一个机房中。当机房A整体故障，HDFS上所有文件不会丢失数据块

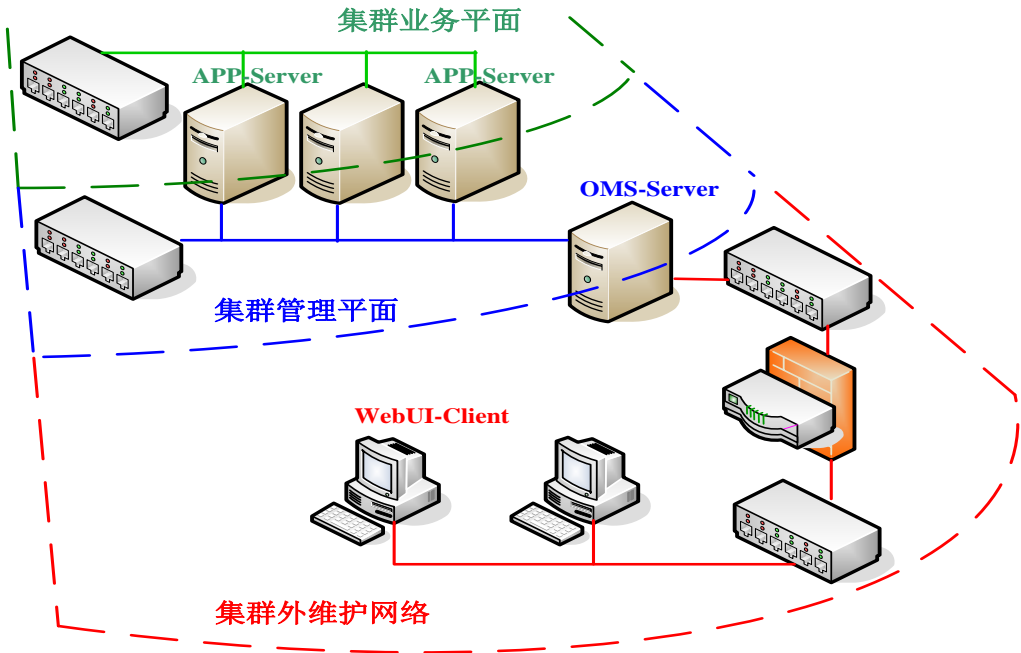


方案描述：配置服务器机房信息，HDFS第2副本选择时优先其他机房服务器，从而保证任何一个文件的多副本不会都落在同一机房，做到跨机房数据可靠性保障。

应用场景：集群服务器分布到2个及以上机房中

客户价值：实现数据跨机房容灾，即一个机房整体故障，HDFS数据不会丢

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

开放兼容

操作系统

JDK

浏览器



11.1~12.2



1.7~1.8



21+



6.4~7.3



1.7~1.8



9+



6.4~7.3

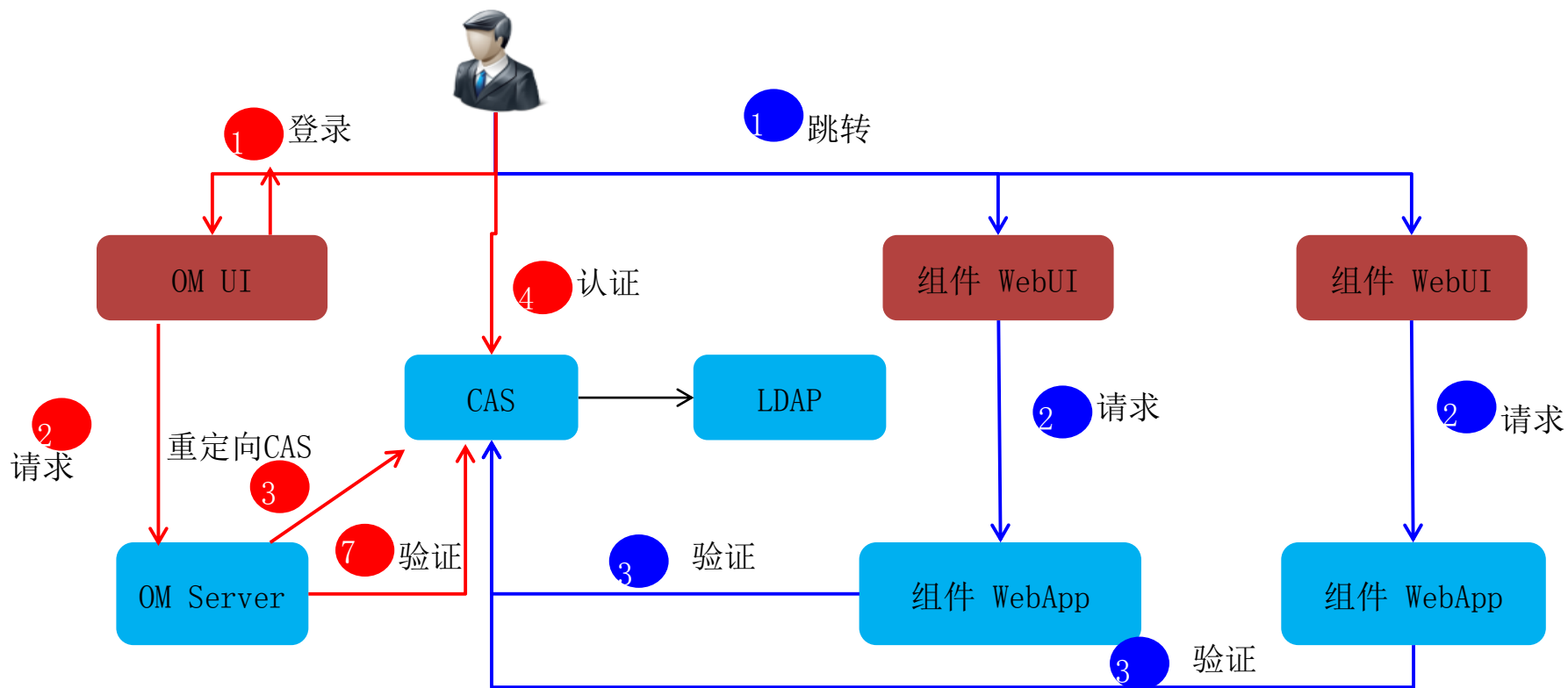


24.x/31.x/40.x



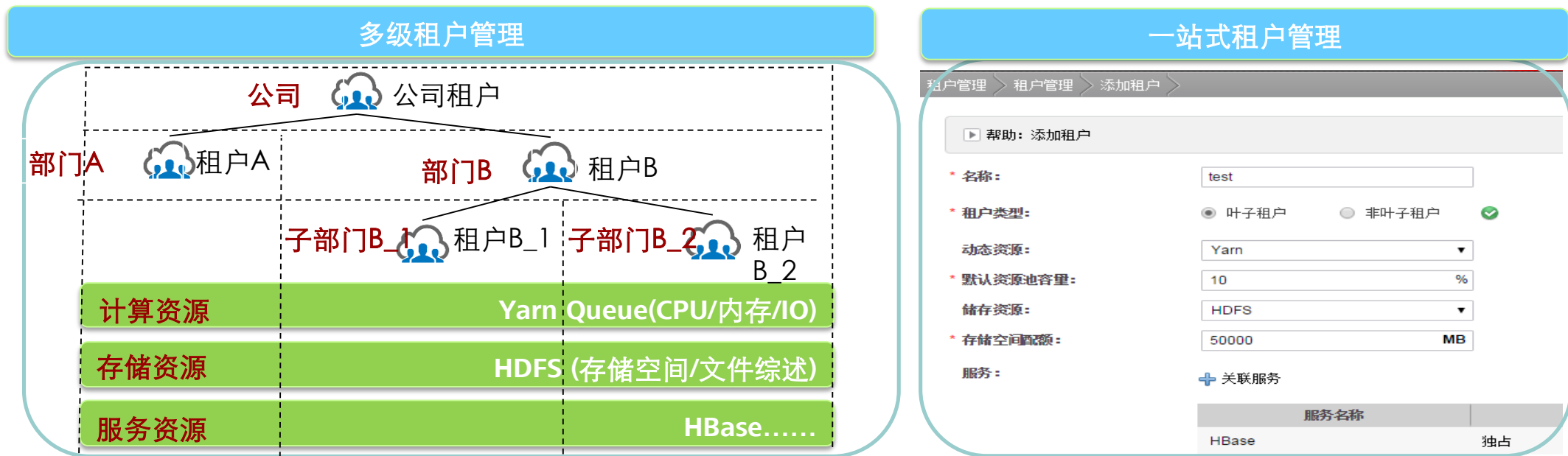
支持两款国产OS的POC能力

一次登录，访问多个组件WebUI



- 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

与企业组织结构相匹配的多租户管理



可视化的多级租户管理，与企业组织结构相匹配，简化系统资源分配与管理

- 与企业组织结构相匹配的多级的租户模型，不同部门对应不同的租户，按需动态增删租户
- 一站式管理租户资源：计算资源（CPU/内存/IO）、存储资源（HDFS）、服务资源（HBase...）
- 基于容器机制的租户资源隔离，为租户SLA保驾护航
- 租户资源使用情况实时监控

可视化的集中用户权限管理

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

系统设置 > 用户管理 > 添加用户

* 用户名: test ✓
 * 用户类型: 人机 ✓
 * 密码: ✓
 * 确认密码: ✓
 * 用户组: 选择添加的用户组

hadoop X hbase X
 spark X

* 主组: hadoop ✓
 分配角色权限: 选择绑定角色

default X Manager_operator X
 Manager_administrator X System_administrator X

系统设置 > 角色管理 > 添加角色

* 角色名称: test ✓

权限:

服务	HBase	HBase Scope	资源名称	资源类型	权限
			global	Global	✓ admin ✓ create ✓ read ✓ write ✓ execute

系统设置 > 角色管理 > 添加角色

* 角色名称: Test ✓

权限:

服务	HDFS	File System	hdfs://hadoop1	权限	
				Read Write Execute	
			apps	Folder	
			apps1	Folder	

* 角色名称: Hive ✓

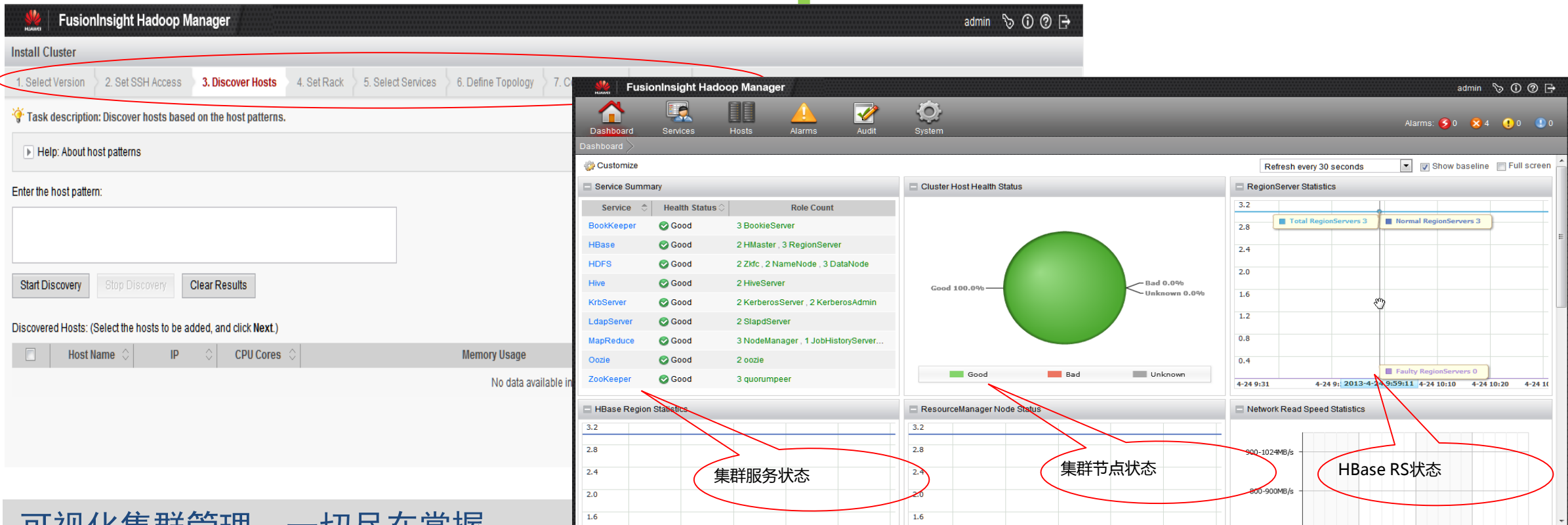
权限:

服务	Hive	Hive Read Write Priv...	default	权限	
				SELECT DELETE UPDATE INSERT Grant Of Sele... Grant Of Dele... Grant Of Upd... Grant Of Insert	
			aaa	Table	
			admin.tbl	Table	

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

可视化集群管理，运维便捷



可视化集群管理，一切尽在掌握

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

目录

1 FusionInsight 简介

2 FusionInsight HD 介绍

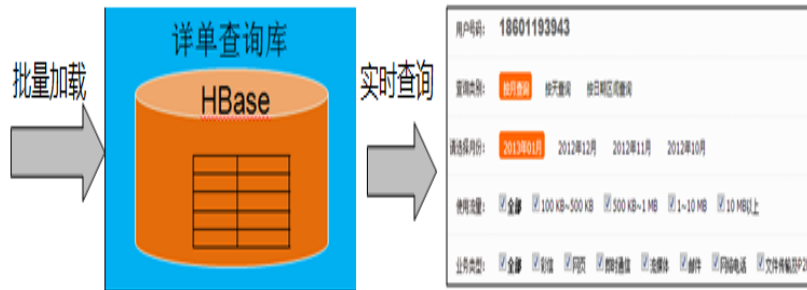
3 FusionInsight HD 成功案例

案例1-某运营商详单分析系统

详单查询场景：海量清单入库、随机点查询

价值：PB级海量数据库，低时延查询，传统DB/DWH无法处理

MSISDN	用户名	时间	流量	业务类型
1860923759	Alex	20130101010101	10KB	微信
1863240504	Ben	20130203123325	250KB	QQ
1862345234	Rudy	20130305110101	34KB	微博
1863423246	James	20130315070809	97KB	WAP访问
....



客户挑战

- 详单数据量大增长快（每月50TB，年增长30%），传统数据库无法支撑，扩容成本高。
- 在线查询只能支持1个月话费详单
- 离线分析速度慢（6小时以上）

解决方案

- 主备大数据业务集群，200+大数据节点
- 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24个月的话费详单
- 更佳的处理性能：500用户并发查询，响应时间小于500ms
- PB级数据存储与计算，经分业务由5天减少到1天

离线汇总统计场景：各维度汇总统计与并行计算

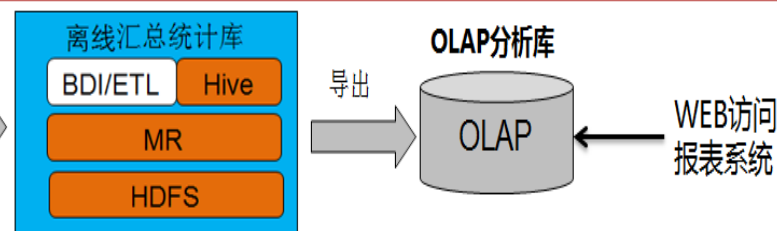
价值：PB级海量数据，分布式存储，并行计算，传统DB/DWH无法处理

5千万用户话单量

总用户数5千万	500000000
每用户每天产生话单数	200
每条话单大小(Byte)	500
存储天数	180
存储/处理总容量(TB)	818.54523

XDR:
日志/话单/信令

批量加载



原始话单

MSISDN	用户名	时间	流量	业务类型
1860923759	Alex	20130101010101	10KB	微信
1863240504	Ben	20130203123325	250KB	QQ
1862345234	Rudy	20130305110101	34KB	微博
1863423246	James	20130315070809	97KB	WAP访问
....

汇总统计后

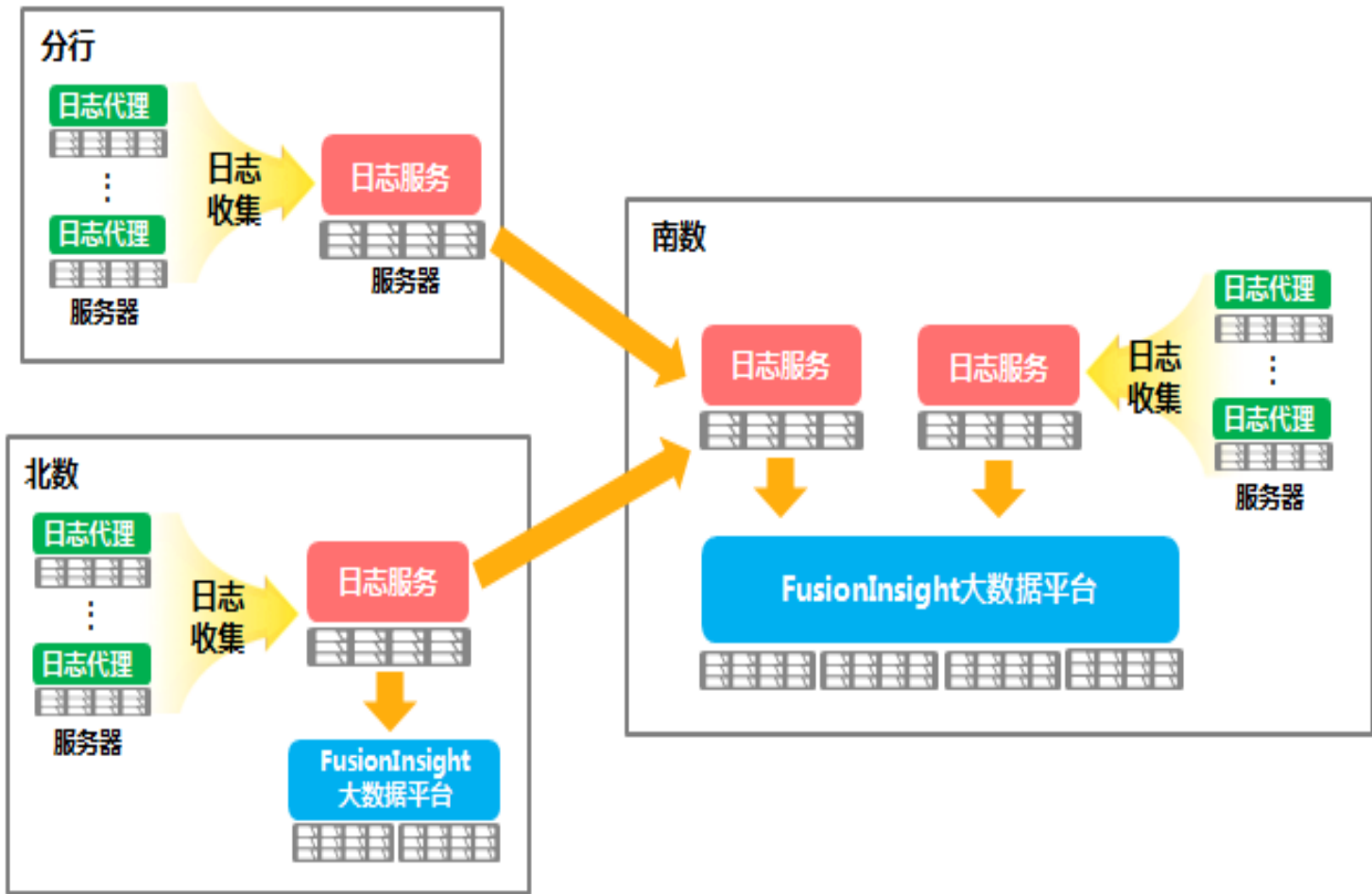
按日期和业务类型，统计数据流量

日期	数据流量	业务类型
20130101	10523 MB	微信
20121212	2355 MB	WAP访问

按用户和日期，统计数据流量

MSISDN	日期	数据流量
1860923759	20130101	35MB
1863254235	20121212	21MB

案例2-某行日志分析系统



客户挑战

- 精准营销，扩大销售
- 优化服务，提升客户体验
- 提高故障定位的准确率，提升故障响应速度

解决方案

- 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
- 基于访问日志的用户行为统计和分析模型

客户价值

- 统一日志收集+分析大数据平台；500+节点的分布式日志收集，大大缩短数据收集周期
- 基于网银用户行为统计和分析，实时+离线的精准营销
- 安全和运维日志关联分析，精准定位故障点



THANK YOU

Copyright©2016 Huawei Technologies Co., Ltd. All Rights Reserved.

The information in this document may contain predictive statements including, without limitation, statements regarding the future financial and operating results, future product portfolio, new technology, etc. There are a number of factors that could cause actual results and developments to differ materially from those expressed or implied in the predictive statements. Therefore, such information is provided for reference purpose only and constitutes neither an offer nor an acceptance. Huawei may change the information at any time without notice.