



阿里Hadoop集群架构及 服务体系

梁李印（无影）

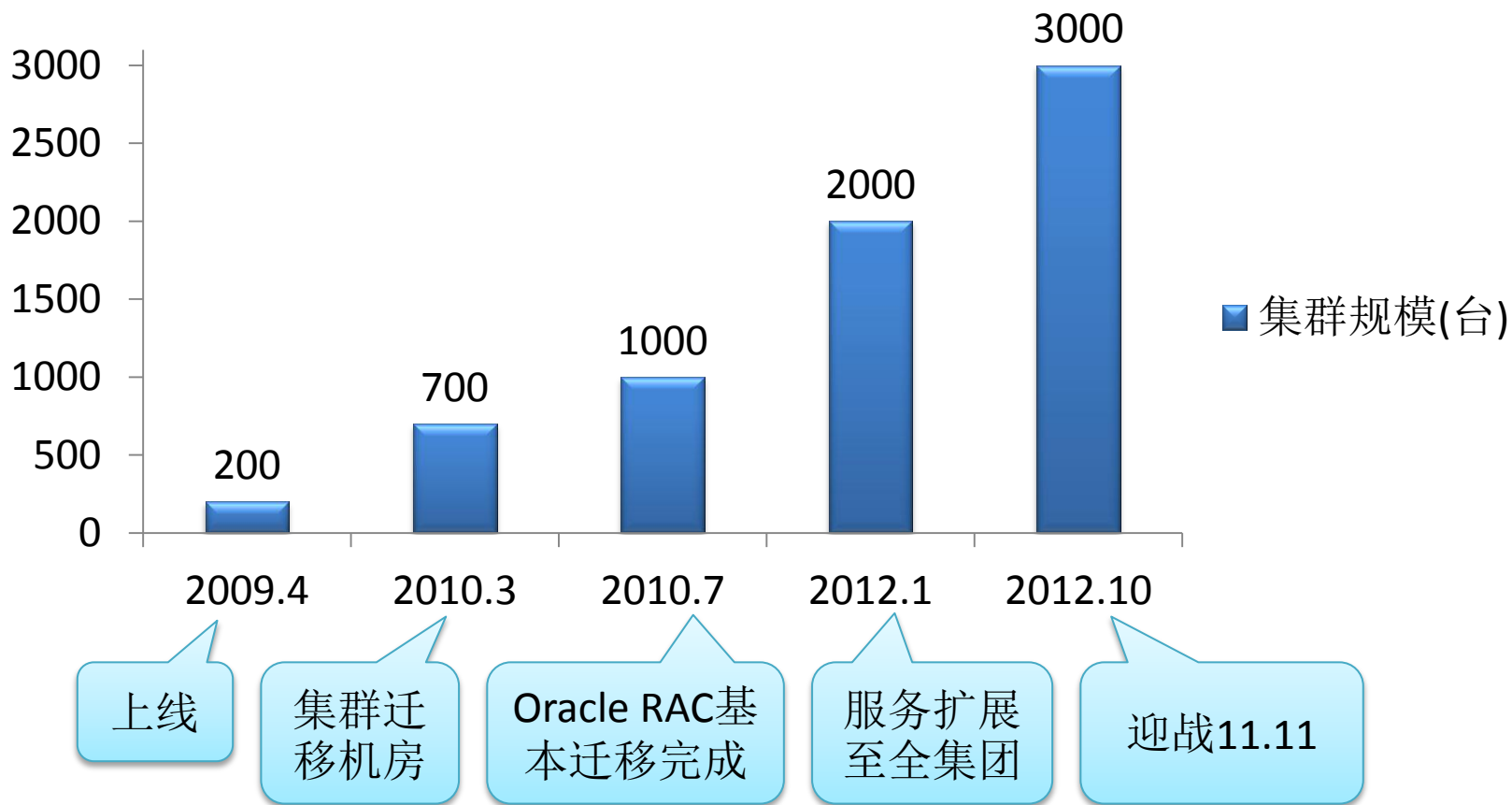
阿里巴巴集团-海量数据



大纲

1. 集群发展现状
2. 集群服务模式及挑战
3. Hadoop版本特性
4. 集群用户门户
5. 集群核心业务架构（三淘）

集群发展现状 规模





集群发展现状 容量与负载

- 集群容量

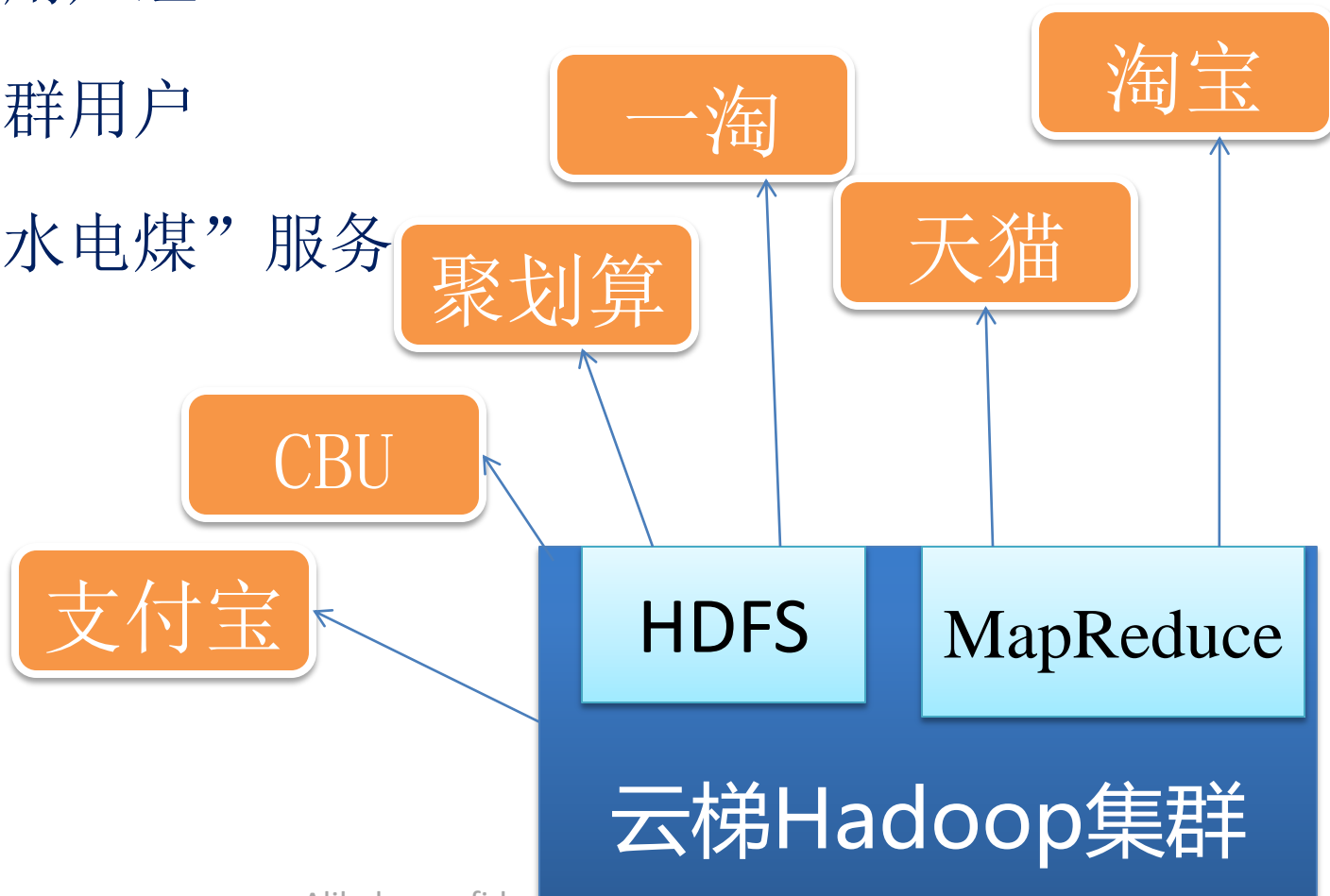
- 约3200台服务器
- 物理CPU ~30000核
- 内存 ~100TB
- 磁盘 ~36000块
- 存储容量 ~60PB

- 集群负载

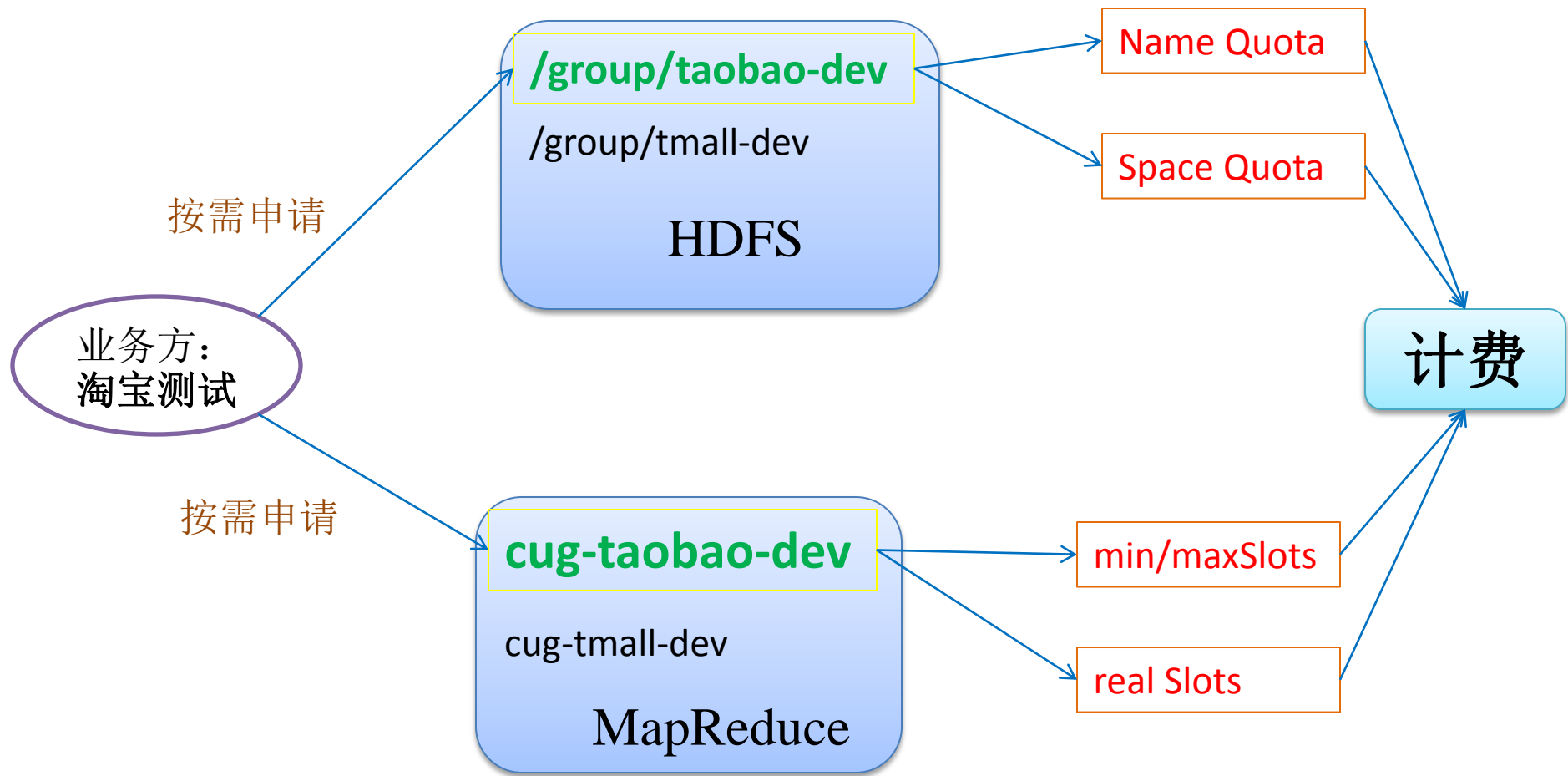
- 每天Job数 150,000+道
- 每天hive query数 6,000+
- 每天扫描数据量 ~7.5PB
- 每天扫描文件数 ~4亿
- 存储利用率 ~80%
- CPU利用率~65% 峰值80%

集群发展现状 业务

- 150+集群用户组
- 4,500+集群用户
- 阿里的“水电煤”服务

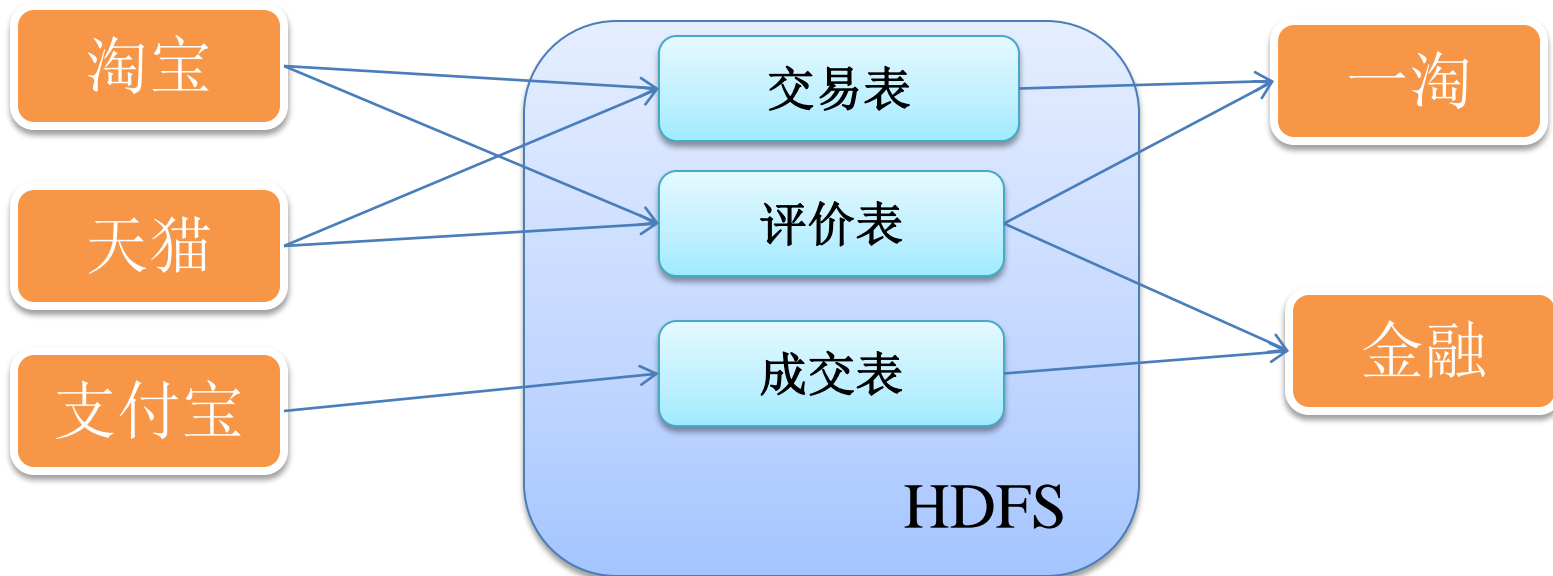


集群服务模式 分组管理



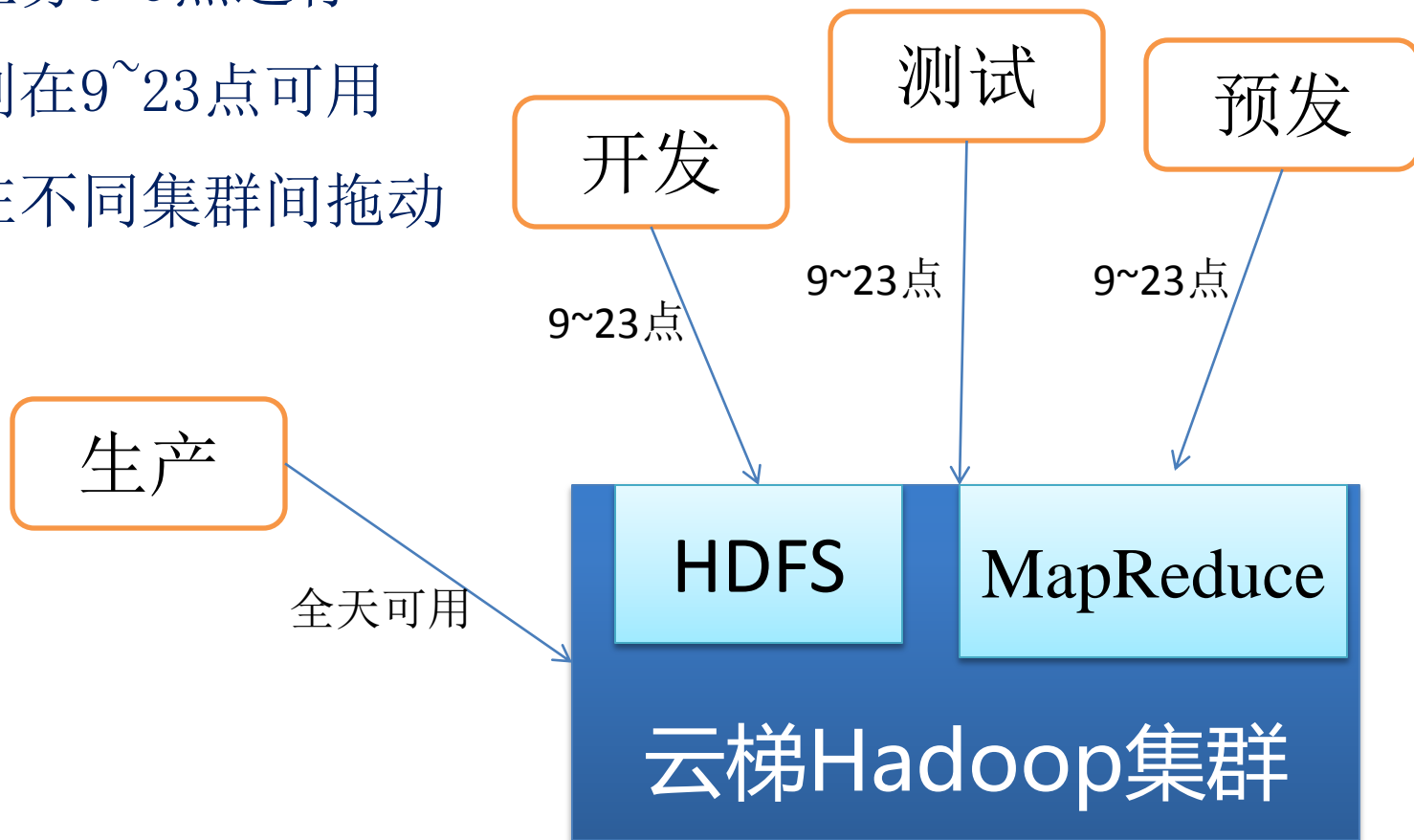
集群服务模式 数据共享

- 原始表/中间表/元数据共享
- 避免重复计算和存储
- “阿里巴巴数据交换中心”



集群服务模式 集群共享

- 生产\开发\测试\预发共享一个集群
- 重点生产业务0~9点运行
- 非生产限制在9~23点可用
- 避免数据在不同集群间拖动





集群服务模式 生态系统

Hive

基于MapReduce的SQL引擎

Streaming

可以用任意可执行程序运行MapReduce

Mahout

机器学习算法库

Pig

类似于Hive的大规模数据分析工具

HBase

准在线存储服务



集群服务模式 优势

	自建Hadoop集群	共用Hadoop集群
集群搭建	机器采购，机房布局	不用考虑
集群运维	节点宕机后需要立即介入	不用考虑
集群扩容	计算或存储资源不足，需要扩容，采购新的机器	网页申请，审批通过即可生效
代码维护	Hadoop代码维护，专业的Hadoop开发人员	不用考虑
数据复制	使用其他团队数据，需要从另外集群复制过来	集团大部分数据集中在一起
成本	11668台	2861台



集群服务挑战 稳定性

- 异常作业拖垮\拖慢集群
- 异常大作业浪费slots
- 部分Slave机器带宽被打满
- NameNode被单个作业拖慢
- 近期案例
 - 用户作业的Task中创建了大量线程
 - 用户提交的作业配置了777777700个Reduce Task
 - 用户作业的Task在slave机器上创建了1000w+临时文件



集群服务挑战 兼容性

- 1000+客户端/Gateway, 100+部门
- 客户端全量升级不现实
- 服务器端升级要保持向下兼容
- 客户端版本:
 - Hadoop 0.19.x
 - Hadoop 0.20.2
 - CDH3u3
 - 内部版本



集群服务挑战 扩展性

- NameNode压力
 - 2亿文件 + 2.8亿Block
 - RPC ops 1.8~2.4万
- JobTracker调度压力
 - 并发运行1000+ Job
 - 并发运行4万+ Map和2万+ Reduce
 - 1000+ Gateway轮询作业状态
- JVM的极限
 - 上限130G heap
 - 超过100G的JVM Heap



集群服务挑战 诊断与测试

- 上千台机器，多个 Master
- 上百个指标：系统，Java GC，Hadoop metrics...
- 集群突然变慢了？某个组新上线大规模作业？
- 大压力情况下出现bug了！
- 每个季度都有新版本发布，版本性能是否有提升？

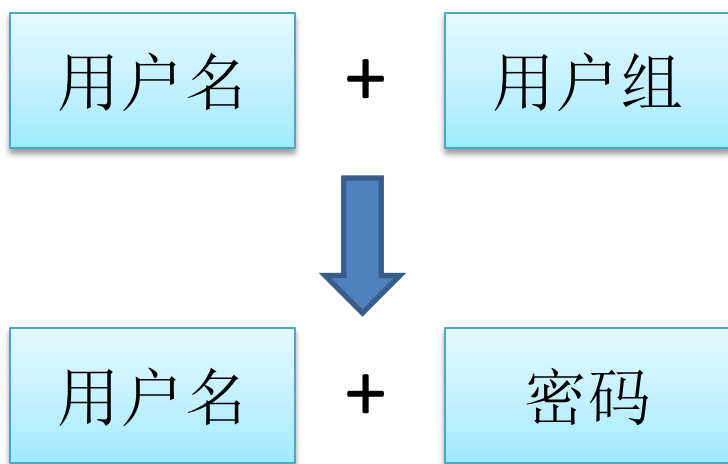


Hadoop版本特性 资源调度

- 调度器改造
 - 增加JobLevel概念，优先保证Level高的作业
 - 支持异构操作系统调度，如RTEL 4u7, 5u6
 - 服务端控制Reduce启动时机 (slow start)
- 动态调整slots配置
 - 动态增减TaskTracker上的slots个数
 - 根据集群map和reduce使用比例调整

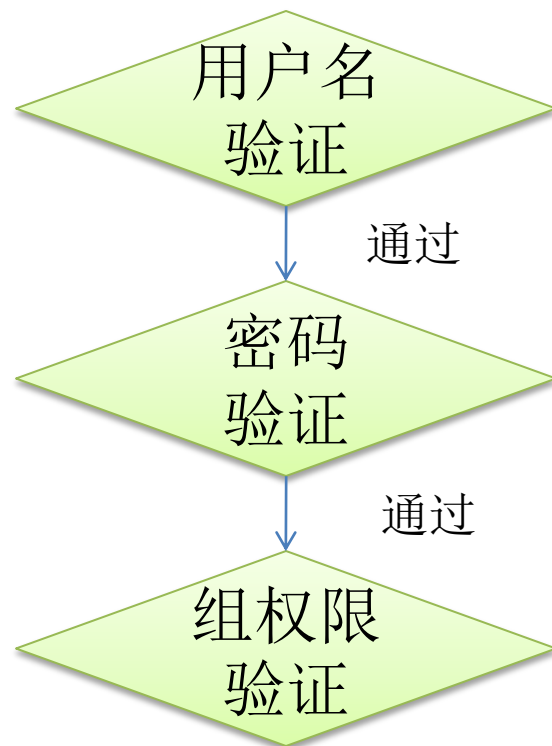
Hadoop版本特性 安全改进

客户端



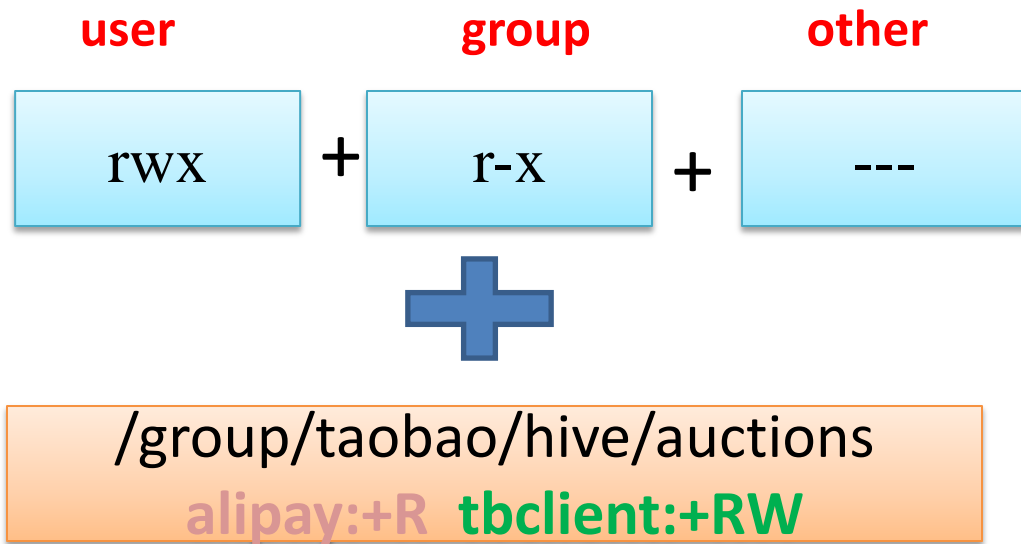
- 增加密码，加强安全性
- 将用户组信息放到服务端，方便客户端配置

服务端

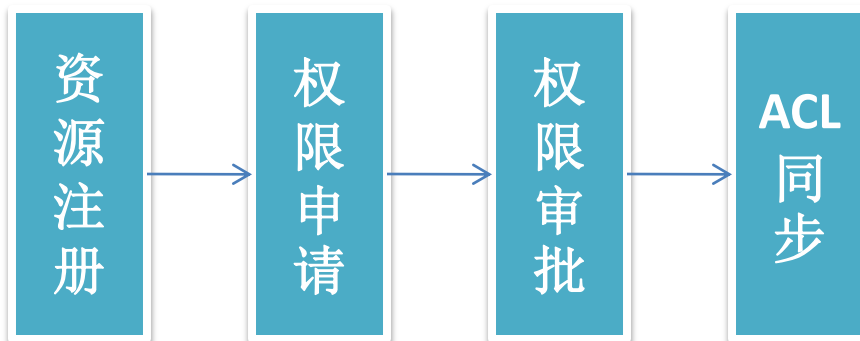


Hadoop版本特性 扩展UGO

- 满足数据的跨部门\公司访问
- 完善的外围系统



外围系统





Hadoop版本特性 稳定性

- 消除异常Job的影响
 - 内存监控
 - 单个Task内存限制
 - 计算节点内存上限控制
 - 磁盘IO监控
 - Shuffle磁盘的读取限制
 - Shuffle速度限制
 - Job的日志量限制
 - Job的Task数限制
 - Job的Counter数限制
 - Job本地文件系统数据读写量监控
 - Job创建HDFS文件数目的监控
 -

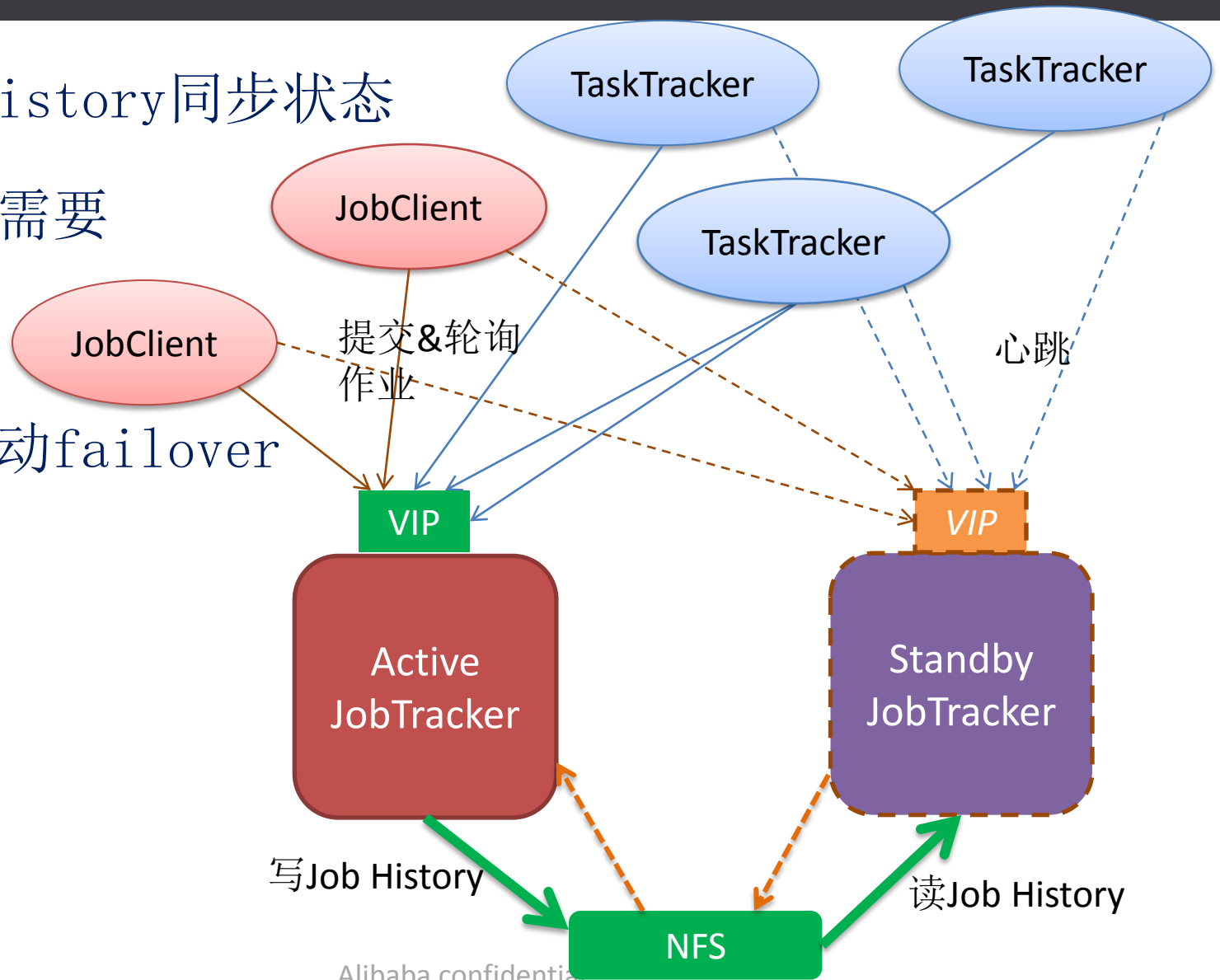


Hadoop版本特性 性能

- Master节点的单点性能压力
 - NameNode改进
 - RPC 改造, Listener 拆分成多个 Reader
 - 使用读写锁, 提高并发度
 - 写操作等待sync时释放 handler
 - JobTracker改进
 - Scheduler调度算法重写, 从 $O(n^2)$ 降低到 $O(1)$
 - 一次心跳分配多个Task
 - Job History改造成异步写
 - Out-of-bound heartbeat提高调度的效率

Hadoop版本特性 JobTracker HA

- 根据JobHistory同步状态
- 满足升级需要
- 热备
- 不支持自动failover





Hadoop版本特性 Hook机制

- pre-job hook
 - 作业提交前执行
 - 配置JobLevel
 - 自动优化Shuffle参数
 - 自动设置Reduce个数
- post-job hook
 - 作业结束后执行
 - 收集Job运行数据



集群用户门户 一站式Portal

▽用户服务

- › 申请用户
- › 申请用户组
- › 申请Gateway
- › 进度查询
- › 查询用户
- › 查询用户组
- › 查询Gateway
- › 用户手册

▽组管理员服务

- › 申请审批
- › 申请Slots
- › 申请存储

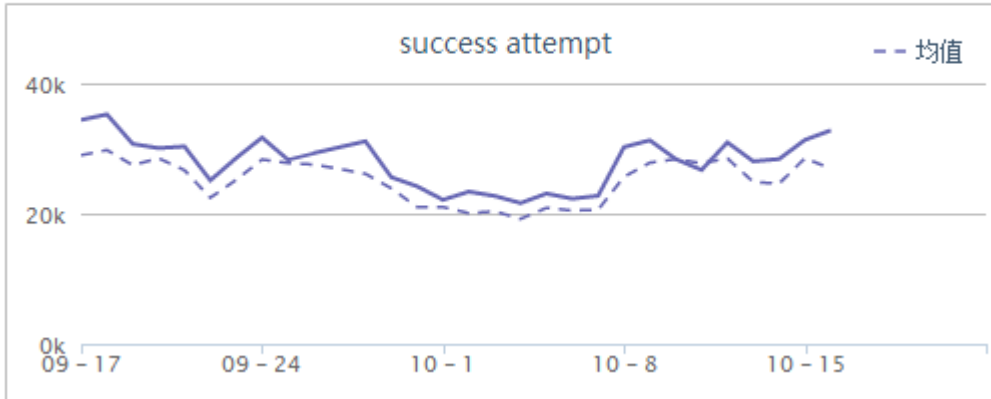
▽集群管理员服务

- › 申请审批
- › 管理用户
- › 管理用户组
- › 管理Gateway
- › Queue管理
- › Slots管理

- 集群用户一站式完成各类申请
- 组管理员负责申请计算\存储资源
- 集群管理员通过网页控制集群配置

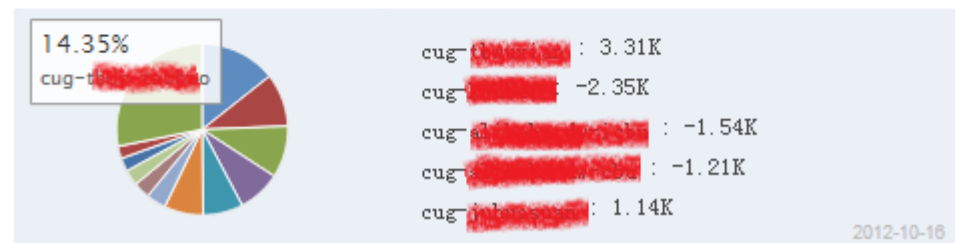
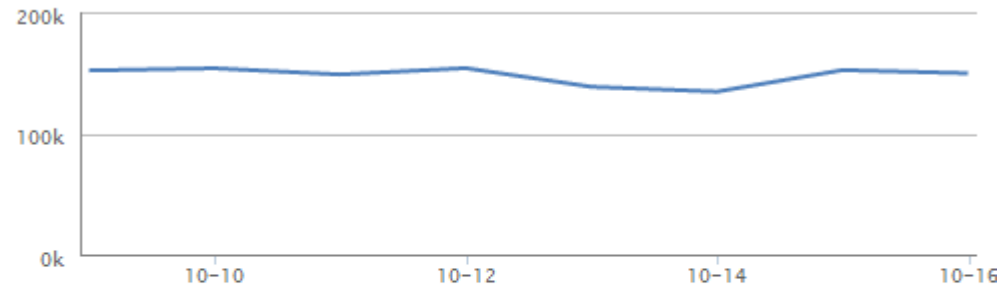
集群用户门户 云梯医生

r01b05028@ali.com 指标图

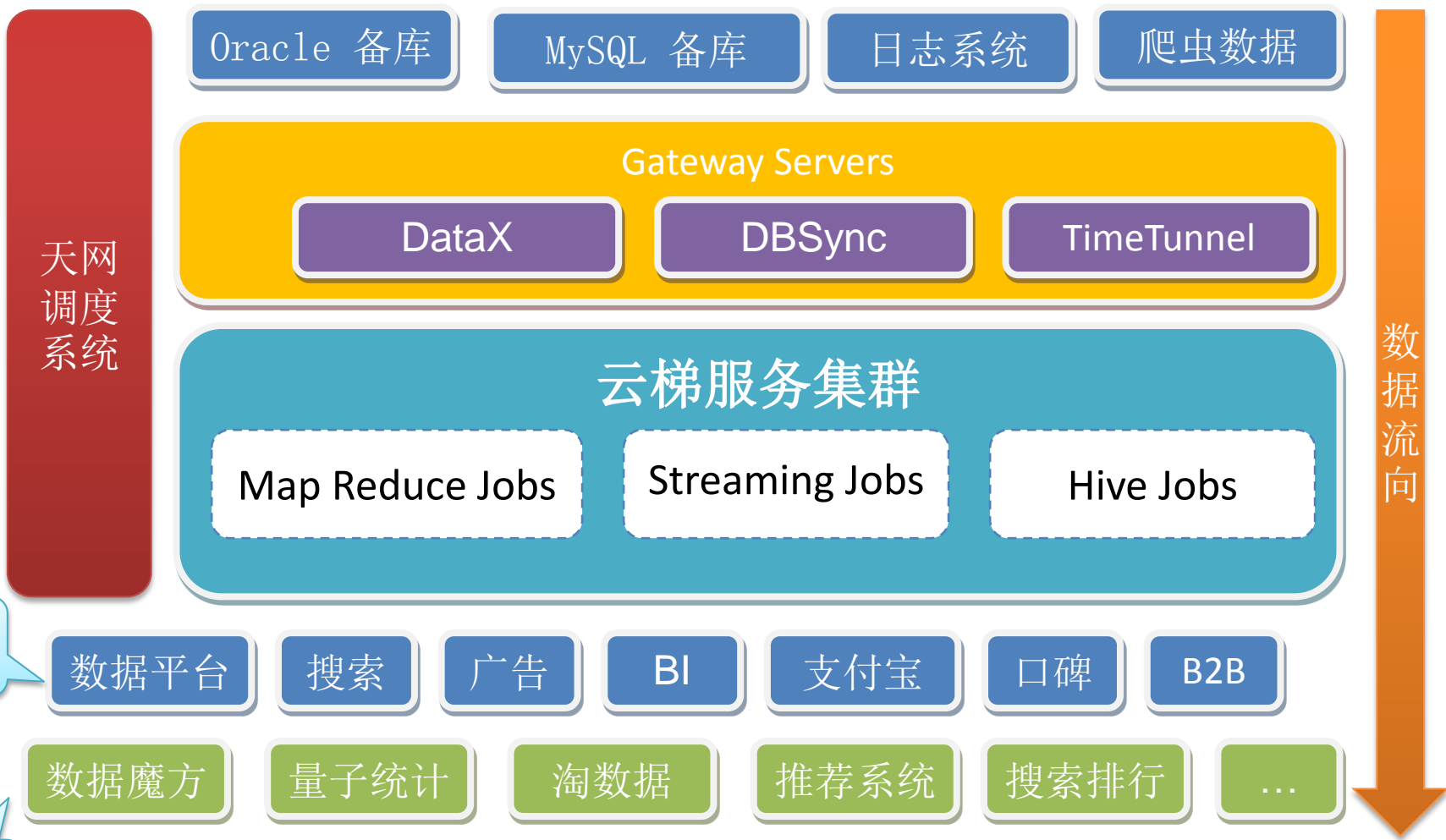


- 集群全局指标
 - 存储\计算利用率趋势
- 用户\组资源使用趋势分析
 - Slots*Sec
 - HDFS/Local r/w
- 机器\机器组视图

- 业务作业对比(前一天\前一周)
 - 数据量增长趋势
 - 不同优先级作业资源消耗
- Master节点关键指标
 - JobTracker心跳频率\时间
 - NameNode RPC各项性能指标



集群核心业务 数据平台架构



资料来源: Velocity China 2010 - 《淘宝云梯分布式计算平台整体架构》, 张清(淘宝)

集群核心业务 数据同步

- 数据流入

- 日志数据:

TimeTunnel, 分布式日志收集工具

- 数据库表:

DataX, 前台数据库<=>云梯 (双向同步)

DBsync, 增量, 大表的快速同步

- 其他数据来源

- 来自其他团队和公司的数据, 比如支付宝数据, 广告反作弊数据, 通过云梯共享

- 数据流出

- 前台业务系统, 如传统数据库、Hbase、DataX ...

- 在云梯上共享给其他团队和公司, 做进一步分析



集群核心业务 数据开发

- 数据分析
 - Hive SQL Web IDE
 - 帐号和云梯服务集成
- 知识管理
 - 元数据/数据字典/数据订阅/表字段血缘分析
- 存储管理
 - 数据生命周期管理
 - 数据保留策略：周期性删除/极限存储/压缩/HDFS
Raid

集群核心业务 Web IDE

- Web IDE 开发界面截图

The screenshot displays the Web IDE interface. On the left, there is a table management panel for the table 's_audit_log_ww'. It shows the table name, size (0.0000 M), and creation time (2012-01-21). Below this, there are tabs for '列信息' (Column Information), '分区信息' (Partition Information), and '数据预览' (Data Preview). The '分区信息' tab is active, showing two partitions: 'pt=20120221' and 'pt=20120120'.

On the right, there is a SQL editor window titled 'run-auditlog*'. It contains the following SQL code:

```
1 --alter table s_audit_log_ww add partition (pt='20120221
2 add jar /home/hive/hive/lib/hive_contrib.jar;
3 select * from s_audit_log_ww where pt = '20120221' and s:
```

Below the SQL editor is a log viewer window titled 'Log' with the filter 'selec...log%'. It displays a table of log entries:

	time1	level	class	ugi	ip	cmd	src
02-21	01:50:26,104	INFO		taobao,cug-...	/119.42.229.18	create	/group
02-21	01:50:26,158	INFO		taobao,cug-...	/119.42.229.18	rename	/group
02-21	01:51:15,044	INFO		taobao,cug-...	/119.42.229.14	listStatus	/group
02-21	02:04:04,080	INFO		taobao,cug-...	/172.24.208.63	rename	/group
02-21	02:04:46,317	INFO		taobao,cug-...	/10.249.64.13	open	/group
02-21	06:05:54,075	INFO		taobao,cug-...	/10.249.38.17	open	/group



Q&A



谢谢！