DTCC 2014中国数据库技术大会
DATABASE TECHNOLOGY CONFERENCE CHINA 2014
大数据技术探索和价值发现

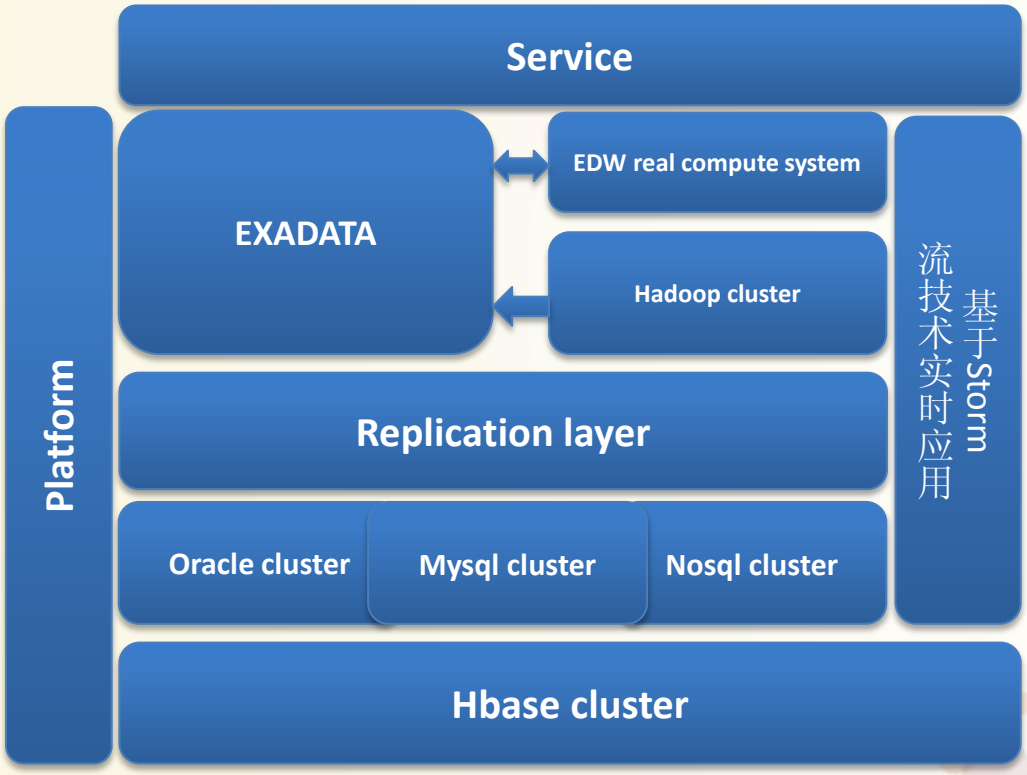
Exadata在1号店数据平台的应用实践

1号店 李勇

overview

- **as an Hardware with Intelligence**
  - offloading
  - smart scan
  - storage index
- **why we used for BI OLAP**
  - above all
  - ehcc

1号店数据平台整体架构



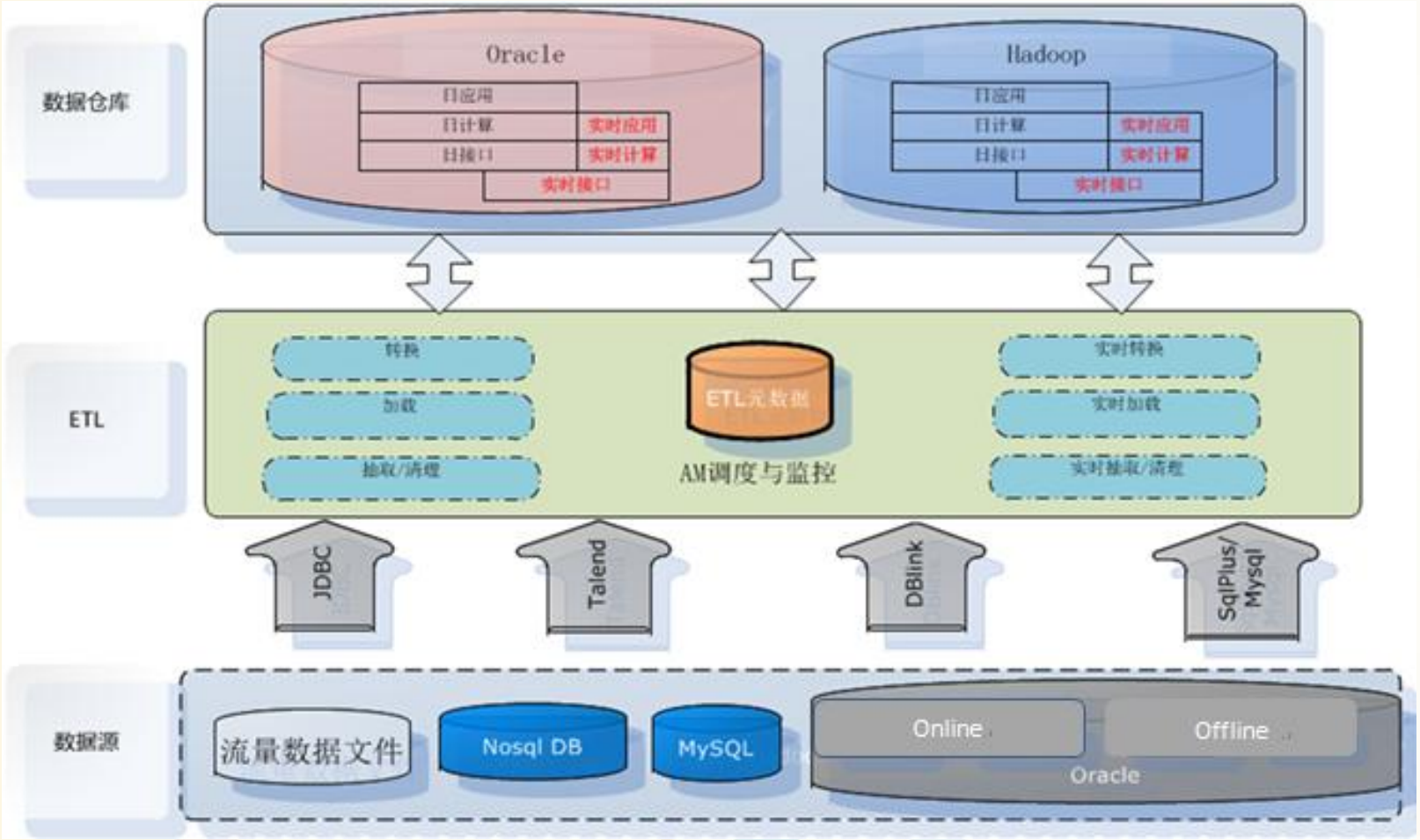
- *Maximum Availability Architecture*
- *High performance*
- *Scalable*
- *Exadata OLAP*
- *Distributed Database Systems*


Thousand ETL jobs
reports
数据分析挖掘
backend报表查询
自助取数平台
CMS
PIS
SCM
个性精准化&广告业务
We need?

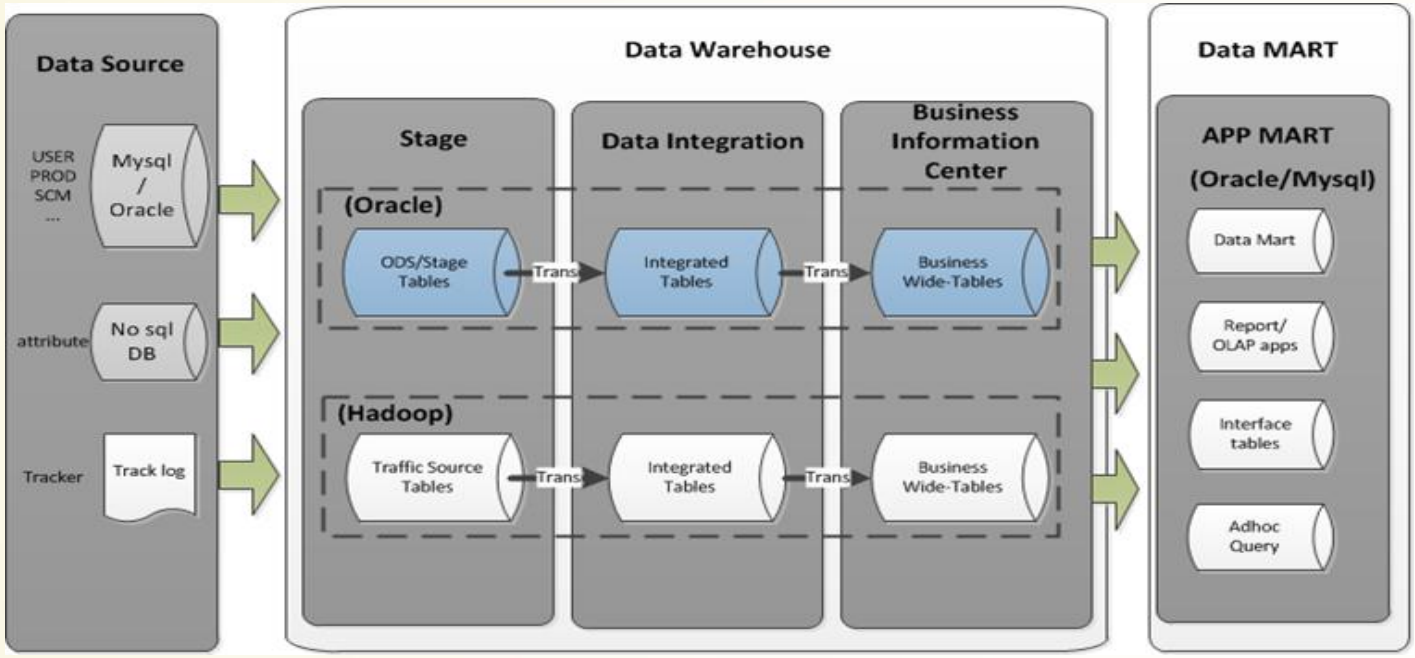
How to do?

- 订单交易处理
- 财务、SCM报表
- 流量数据挖掘
- 日志数据分析

structured and unstructured Integration



1号店DW系统框架



Exadata Hadoop Mysql 构成:

- Exadata：非流量数据的数据镜像层、数据整合层，数据中间层，应用层
- Hadoop：流量数据的数据镜像层、数据整合层
- Mysql：依据应用做数据集市

1号店Exadata

migrate

Quarter Rack
X2-2

Half Rack
X4-2

- X2 active data guard
- X4 one database many services
- Why migrate

Use Exadata Better——smart scan

- The Bottleneck on Many (Most) Large Databases is between the Disk and the DB Server(s)!
- How to Speed Up?
  - Make the Pipe Bigger/Faster
  - Reduce the Volume
- 触发smart scan的前提(三条都满足):
  - Full scan: Full Table Scan or Fast Full Index Scan
  - Direct Path Reads: 并发读取数据到program global area(PGA)
  - 数据保存在exadata storage上: smart scan offloading是storage server独有
- Need Index or not?
  - When indexes are present, Exadata databases tend to over use them
  - Make them invisible first, and then remove them

Use Exadata Better——EHCC

- The hybrid columnar compression (HCC) Requirements:
  - Direct path loads
  - Impdp with DIRECT option or DIRECT path inserts
  - Alter table move
- Data is stored in compression unit
  - not real columnar compression
  - advantage vs. disadvantage
- If you update rows in HCC table, The entire compression unit is locked. The rows that are updated may be moved to a lower compression level such as No compression/ OLTP compression
- How to use better

Use Exadata Better——ESFC

- 一种减少数据库访问IO cost的方式
  - 不依赖smart scan，一般还会被忽略
  - 最有效的单block离散读访问方式，降低小IO读请求响应时间
  - Write-through缓存，写操作绕过缓存直接写磁盘
- The OLTP acceleration(flash disk)
  - 以SSD的方式提供给ASM用于数据库存储
  - 高并发oltp中，redo log在ESFC上的表现
- 一个巨大的cache
  - cellsrv 同时向磁盘和闪存发送异步IO请求
  - ALTER TABLE xxx STORAGE (CELL_FLASH_CACHE KEEP)
- ESFC到底能提升多少性能

Use Exadata Better——IORM

- DBRM and IORM

DBRM and IORM design to allocate and limit resources within and across databases

- 通过iDB发送IO请求给存储节点上的cellsrv

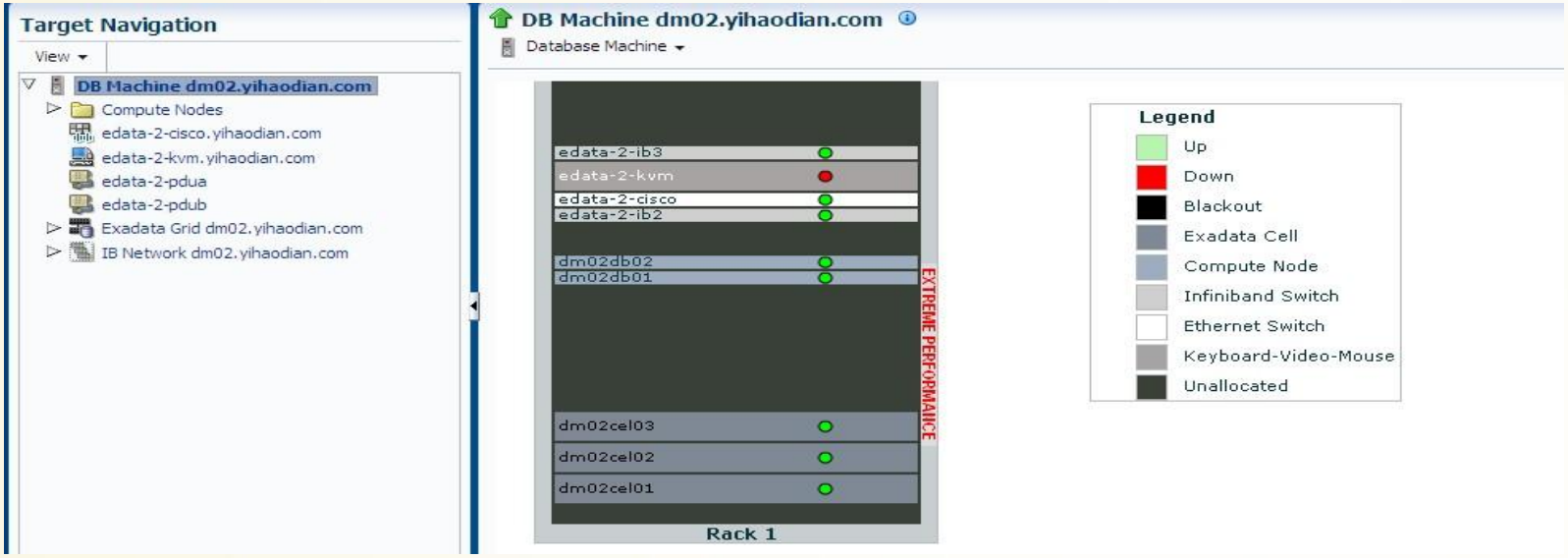
- Interdatabase IORM
- Category IORM
- Intrdatabase IORM

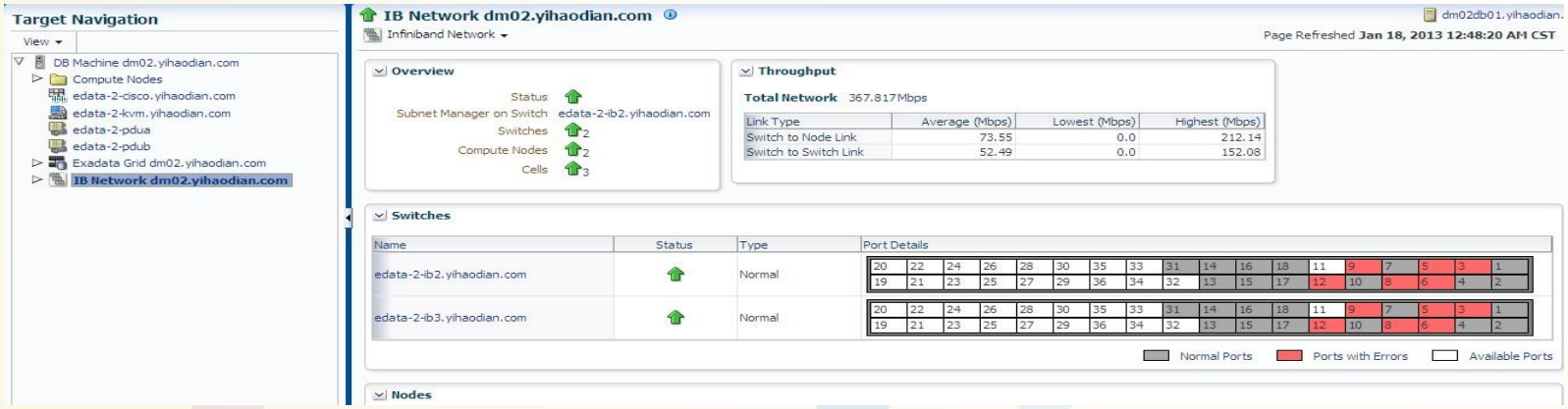
Exadata migration

- **Migration Methods - Physical Migration**
  - Physical Standby (ARCHIVELOG and LOGGING)
  - Transportable Database (Oracle 11.2 on any little endian)
  - Transportable Tablespaces (Oracle 10.1 or later on any platform)

- **Migration Methods - Logical Migration**
  - Logical Standby
  - Golden Gate
  - Data Pump

Exadata monitor





Exadata 所遇到的问题

- 高并发大语句的性能问题
- some bugs
  - ORA-00600: internal error code, arguments: [kjblpkeydrmqscchk:pkey], [341954], [287], [5698091], [0], [6], [0], [], [], [], [], []
  - LMS1 process terminating the instance caused by Bug 14409183 : SESSION HANGS ON GC BUFFER BUSY ACQUIRE
  - ORA 600 [ktspScanInit-l1]
  - ORA 600 [kolaslFree: kolasl templob ref count]
- upgrade
  - Multiple Patches
  - Infiniband (once per year)
  - DB Nodes / Storage Server (quarterly)
  - Bundle Patch (BP) DB Software (quarterly)
  - Additional components (Ethernet switch, KVM, PDU)

Q&A

THANKS