火龙果•整理
uml.org.cn

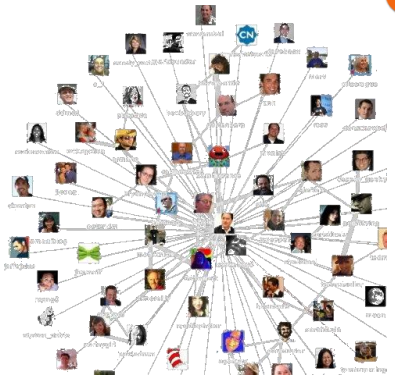
ORACLE®
甲骨文

消费大数据的智慧之路

火龙果•整理
uml.org.cn

“互联网上，没有人知道你是一条狗”。

彼得·施泰纳
1993年《纽约客》



“给我足够多的数据，我将比你更了解你自己！”。

Big Data大数据
2012年6月20日

云+大数据

火龙果•整理
uml.org.cn

云

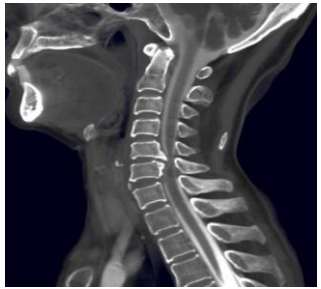
彻底改变IT

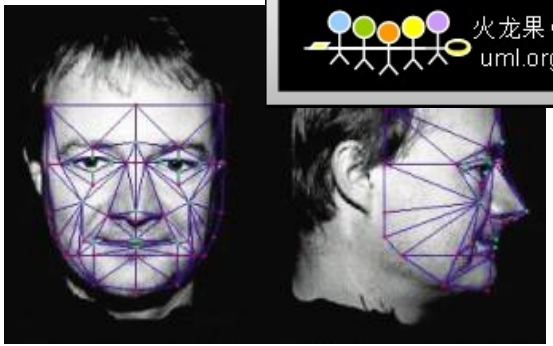
大数据

彻底改变业务

迎接大数据的挑战

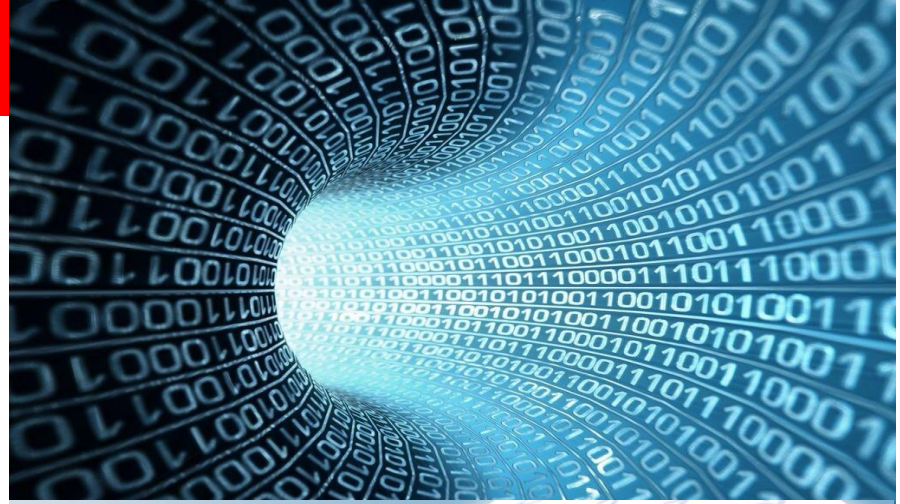
火龙果•整理
uml.org.cn





大数据意义重大

火龙果•整理
uml.org.cn

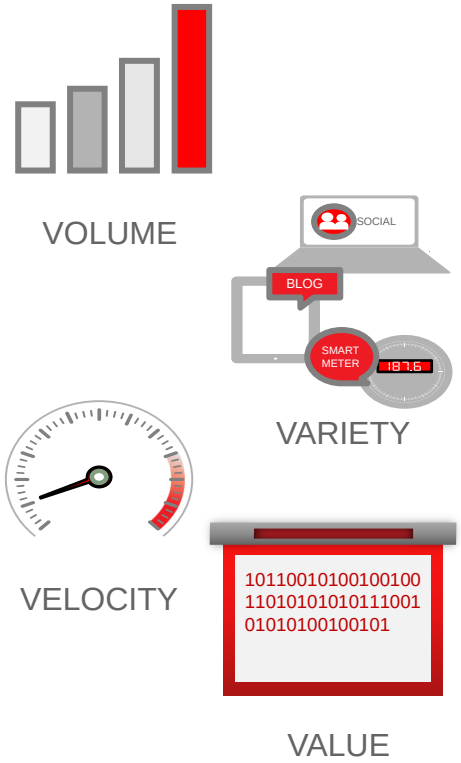


“大数据是新时代的石油。”

“大数据，改变的是业务。”

管理和处理有难度

火龙果•整理
uml.org.cn



- 巨大的数据量 **Volume**
  - 集中储存/集中计算已经无法处理巨大的数据量

3亿+用户，高峰期一天上亿条微博

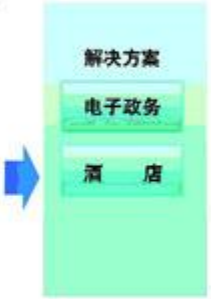
中型城市每月数十亿智能电表数据

2015年全球移动终端产生的数据量6300PB

- 多结构化数据 **Variety**
  - 文本/图片/视频/文档等
- 增长速度很快 **Velocity**
  - 海量数据的及时有效分析
  - 用户基数庞大/设备数量众多/实时海量/数据指数级别增长
- 价值密度低**Value**
  - 单条数据并无太多价值，但庞大的数据量蕴含巨大财富

熟悉的数据和系统

火龙果•整理
uml.org.cn



成熟的数据管理和处理体系

火龙果•整理
uml.org.cn

数据洞察层	商务智能	报表系统	交互式分析	实时仪表盘
数据消费层	应用系统平台	数据共享、数据服务		
数据智慧层	数据库	数据仓库	文件系统	
数据整合层	数据整合			
数据计算层	**数据处理和计算**			
数据存储层	非/半结构化数据	结构化数据库		
数据收集层	数据采集交换平台			
数据源层	数据文件数据源	RDBMS数据源	其他数据源	

新技术新方法－分布式

OLAP
+
ETL
+
Storage

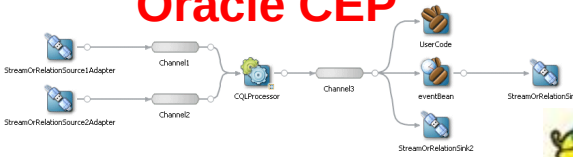
+

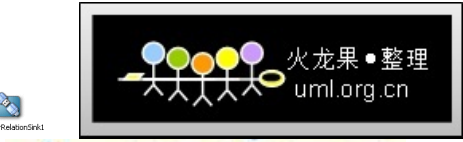
OLTP
+
Storage

狭义

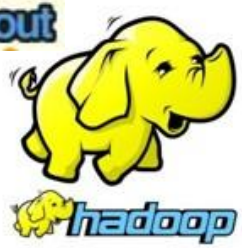
广义

Oracle CEP

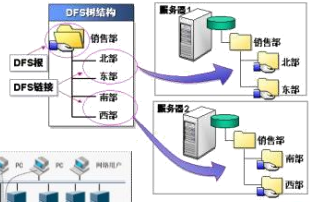




S4 distributed stream computing platform



TimesTen



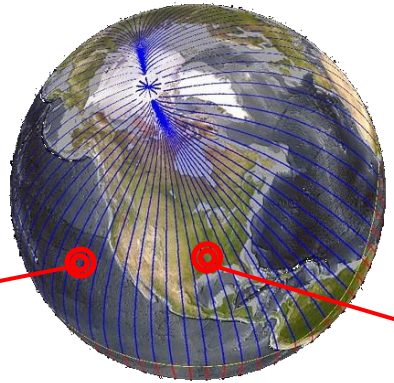
Sqoop
Big Data Connector

大数据的计算模式

火龙果•整理
uml.org.cn

持续信息，流式计算

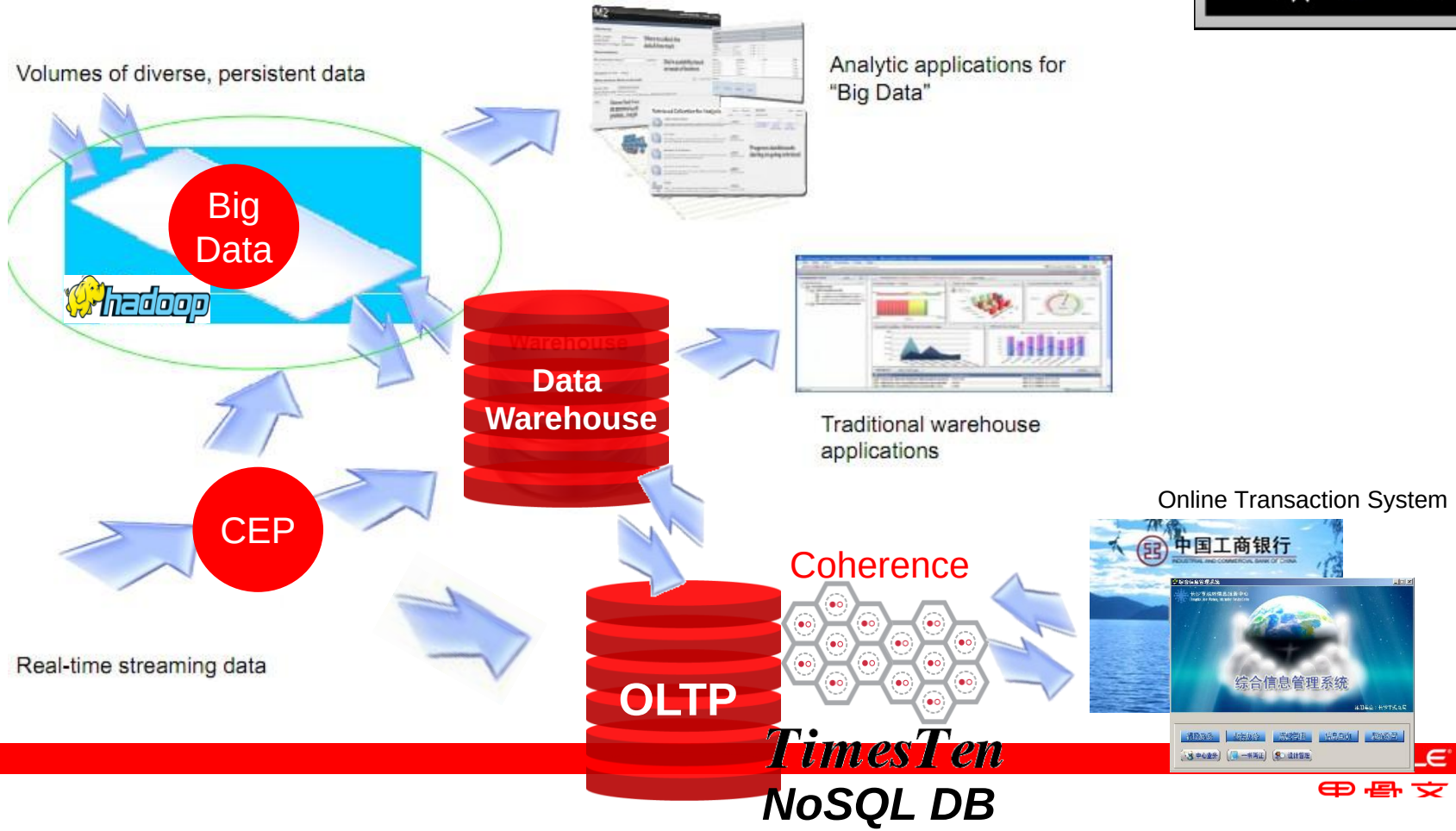
信息海洋，海量计算



大海捞针，实时计算

大数据计算模式组合

火龙果•整理
uml.org.cn



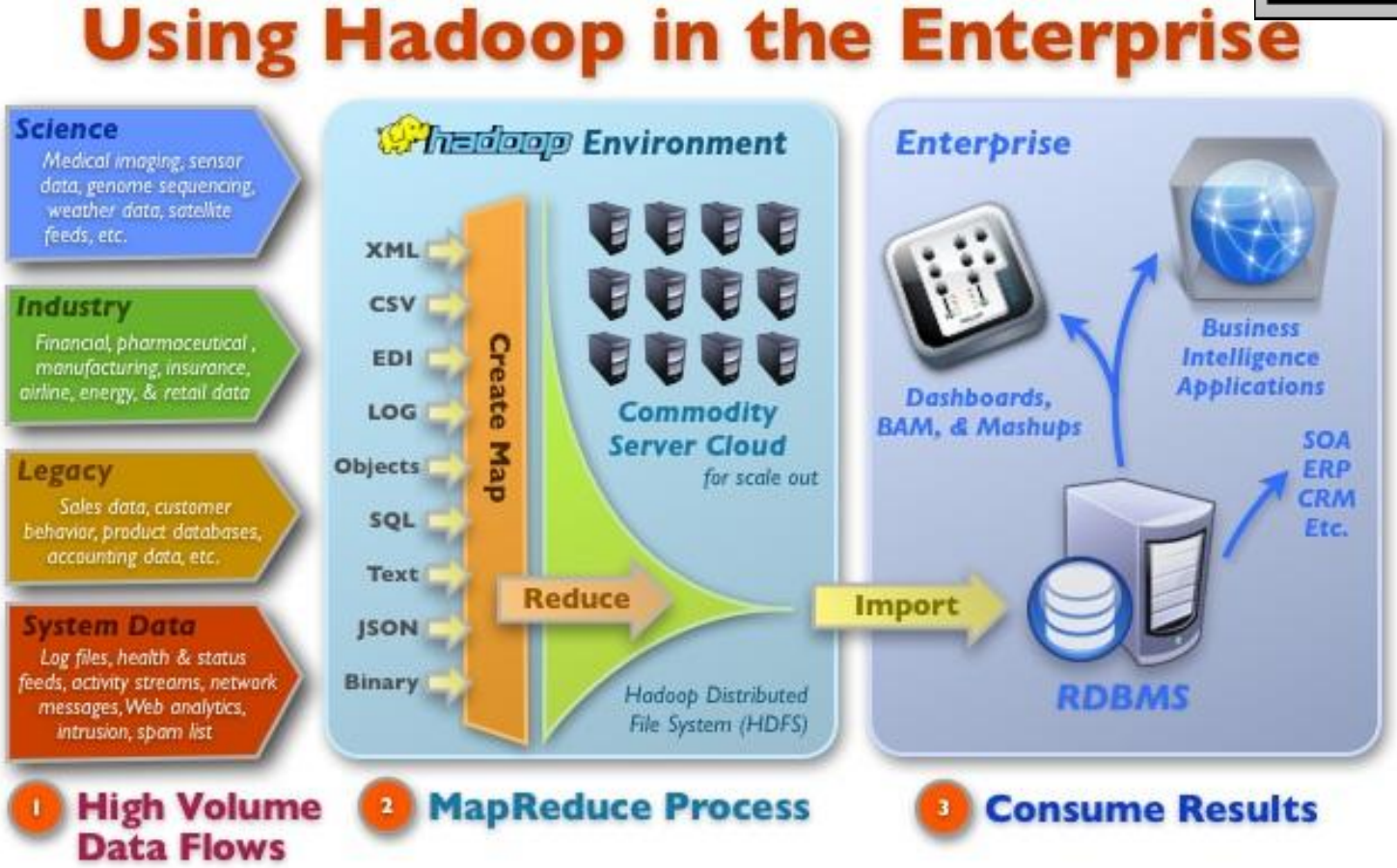
完备的数据管理和处理体系 - 多结构化

火龙果•整理
uml.org.cn

数据洞察层	商务智能	报表系统	交互式分析	数据探索	实时仪表盘
数据消费层	数据应用消费平台		数据共享、数据服务		
数据智慧层	数据库	数据仓库	灵活数据存储系统		
数据整合层	数据整合				
数据计算层	**实时计算**	**批量计算**	**流式计算**		
数据存储层	集中式文件系统	结构化数据库	分布式文件系统	灵活存储	
数据收集层	数据采集交换系统	日志收集系统	消息/事件系统		
数据源层	文件数据源	RDBMS数据源	事件数据源	其他数据源	

大数据在企业的应用形态

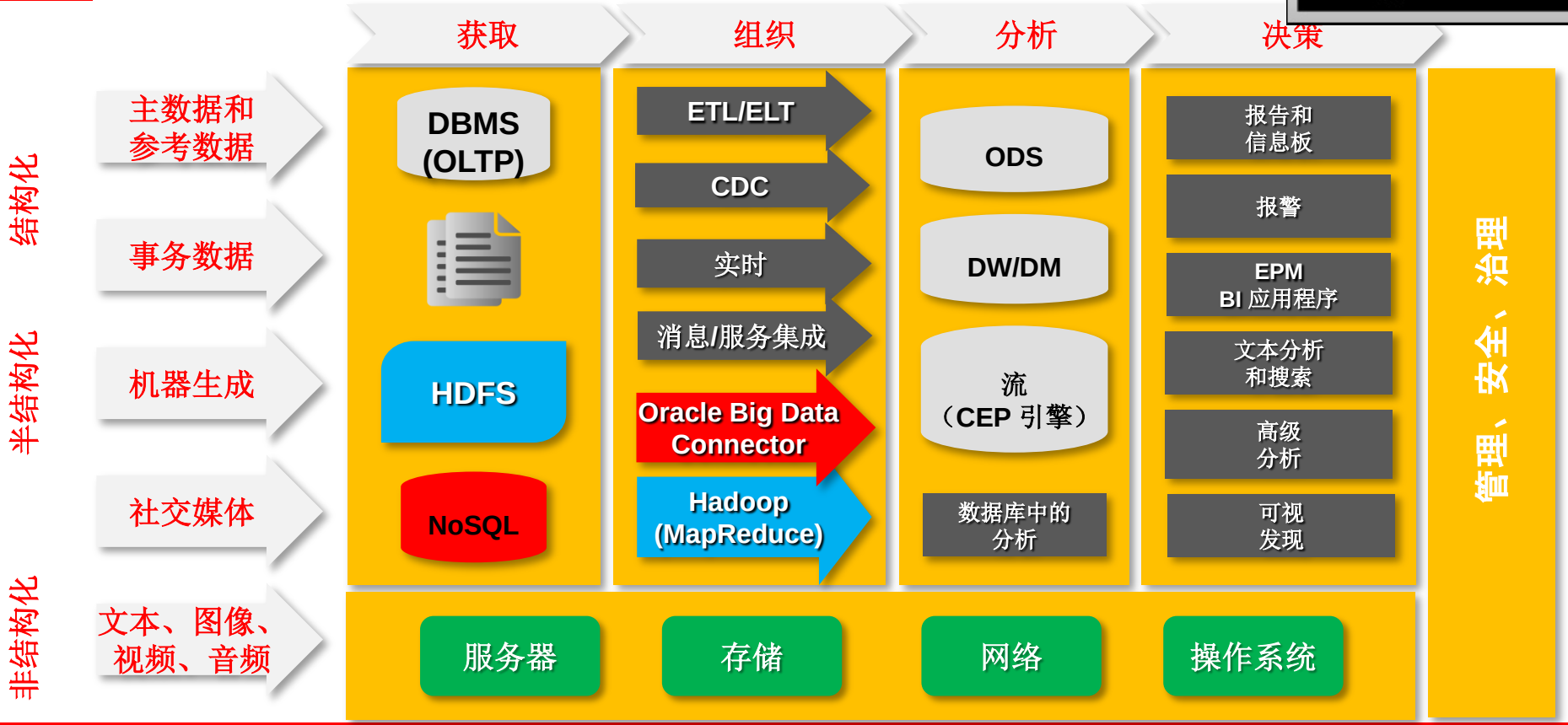
火龙果•整理
uml.org.cn



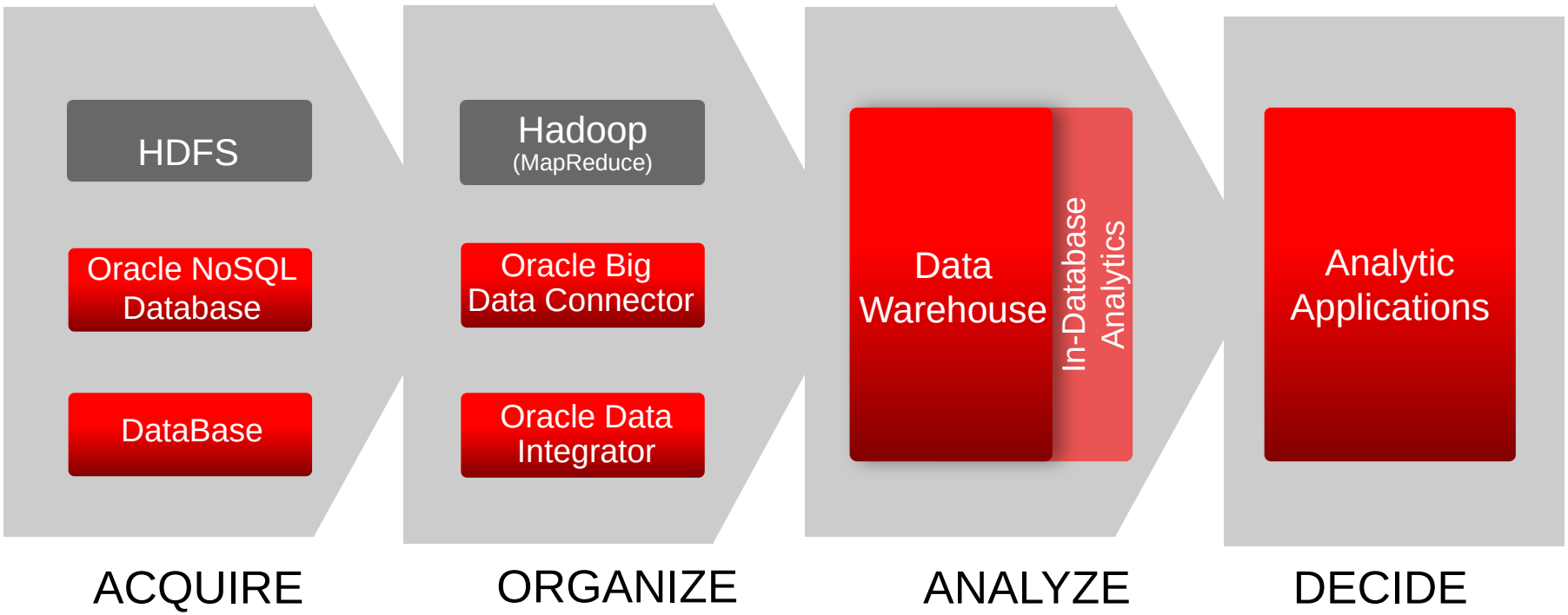
ORACLE
甲骨文

改进企业数据架构，引入大数据，数据协同

火龙果•整理
uml.org.cn



Oracle大数据解决方案



火龙果•整理
uml.org.cn

Oracle大数据解决方案

火龙果•整理
uml.org.cn

Data Variety

Schema-less

Schema

Big Data Solution
- Cloudera CDH3
  Cloudera Manager
- Big Data Connectors*
- NoSQL Database CE

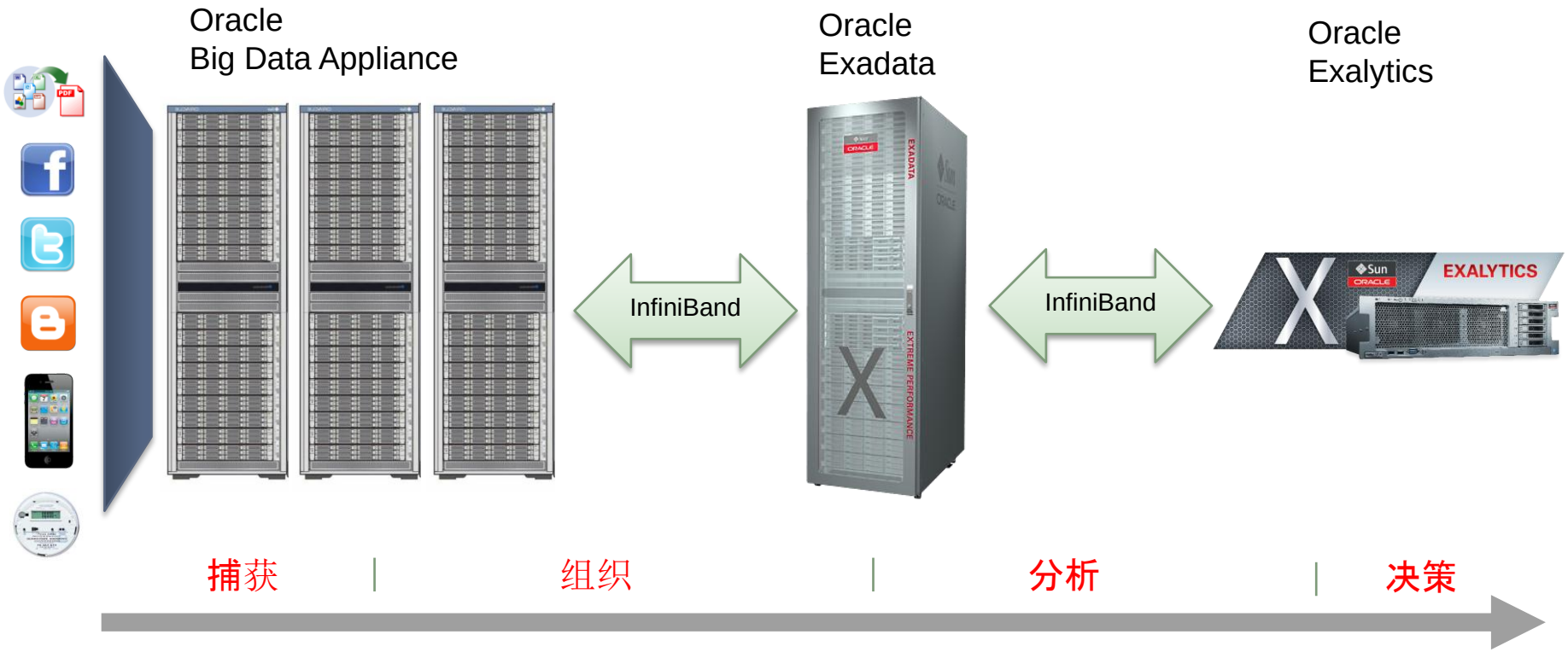
Oracle Database Solution
- OLTP & DW
- Data Mining & Oracle R
- Semantics
- Spatial

BI Solution
- Speed of Thought Analytics

Acquire *Organize* *Analyze*

软硬一体优化集成的Oracle大数据综合解决

火龙果•整理
uml.org.cn

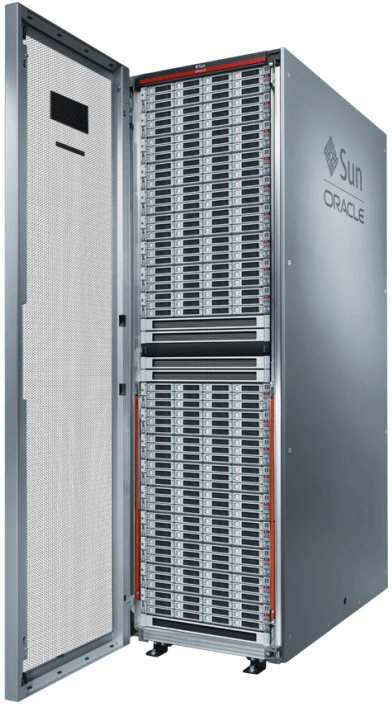


Big Data Appliance

- 硬件部分
- 18 Sun X4270 M2 服务器
- 每台2 CPUs * 6核
- 每台48 GB内存(可扩展)
- 12 * 3TB 磁盘空间

- 网络部分
- 40Gb InfiniBand
- 10Gb 以太网

Raw Storage: 648T
Core Count: 216核
Mem Count: 864G – 2592G



- 软件部分
- Oracle Linux
- Oracle JDK
- Cloudera Hadoop Distribution
- Cloudera Manager
- Open-source R distribution
- Oracle NoSQL Database Community Edition ***

- Oracle Big Data Connector ***
  - ODI Adapter for Hadoop
  - Oracle Loader for Hadoop
  - Oracle Direct Connector for HDFS
  - Oracle R Conenctor for Hadoop

*** No Support for free NoSQL CE. NoSQL EE needs license. *** Not include in BDA, sales as separated product.

最全面最强的Oracle数据管理和处理体系

火龙果•整理
uml.org.cn

数据洞察层	Exalytics/BIEE/Essbase	Endeca	Oracle BAM		Web Center
数据消费层	Exalogic	Coherence/SOA/ODI/GG			
数据智慧层	Exadata/Oracle/Mysql	Exadata/Oracle Hive	Oracle NoSQl DB HBase		BPM
数据整合层	Big Data Connector / ODI / GG / MDM	Sqoop			ECM
数据计算层	Exadata TimesTen/Coherence Oracle NoSQL DB	Exadata /ORE Hadoop MapReduce /R	CEP TimesTen Coherence		IDM
数据存储层	ZFS/Pillar	Exadata/Oracle MySQL	Lustre HDFS	Oracle NoSQL DB HBase	EM12c
数据收集层	ODI+GG+SOA	Flume	WebLogic/CEP		
数据源层	文件数据源	RDBMS数据源	事件数据源	其他数据源	

大数据平台上的行业解决方案

火龙果•整理
uml.org.cn

移动应用用户行为统计分析 | 基于日志和访问内容的用户画像 | 机顶盒用户使用习惯和精准营销 | 基于用户画像的广告精准营销 | 语义分析/搜索引擎/实时处理 | 基于互联网的舆情监控系统 | 补充非结构化数据分析的数据仓库 | 分布式大数据存储和管理系统 | 海量历史数据分析平台 | 海量指纹识别/人脸识别查询系统 ……

数据消费层	Exalogic		Coherence/SOA/ODI/GG	
数据智慧层	Exadata/Oracle/Mysql	Exadata/Oracle	Hive	Oracle NoSQl DB / HBase
数据整合层	Big Data Connector / ODI / GG / MDM		Sqoop	
数据计算层	Exadata TimesTen/Coherence Oracle NoSQL DB	Exadata /ORE	Hadoop MapReduce /R	CEP TimesTen Coherence
数据存储层	ZFS/Pillar	Exadata/Oracle MySQL	Lustre / HDFS	Oracle NoSQL DB / HBase
数据收集层	ODI+GG+SOA	Flume	WebLogic/CEP	
数据源层	文件数据源	RDBMS数据源	事件数据源	其他数据源

ORACLE
甲骨文

火龙果•整理
uml.org.cn

ORACLE®

甲骨文