北京大学
PEKING UNIVERSITY

PART 02 ▸

数据仓库架构和技术介绍

1.2 数据仓库简介

北京大学 PEKING UNIVERSITY

- 数据库/数据仓库
- 并行数据库/分布式数据库
- 数据划分/并行处理
- 吞吐量/响应时间
- 存储与计算分离/共享存储

- OLAP/DSS
- MPP
- SQL on Hadoop
- Cloud Native

2 数据仓库架构和技术介绍

北京大学 PEKING UNIVERSITY

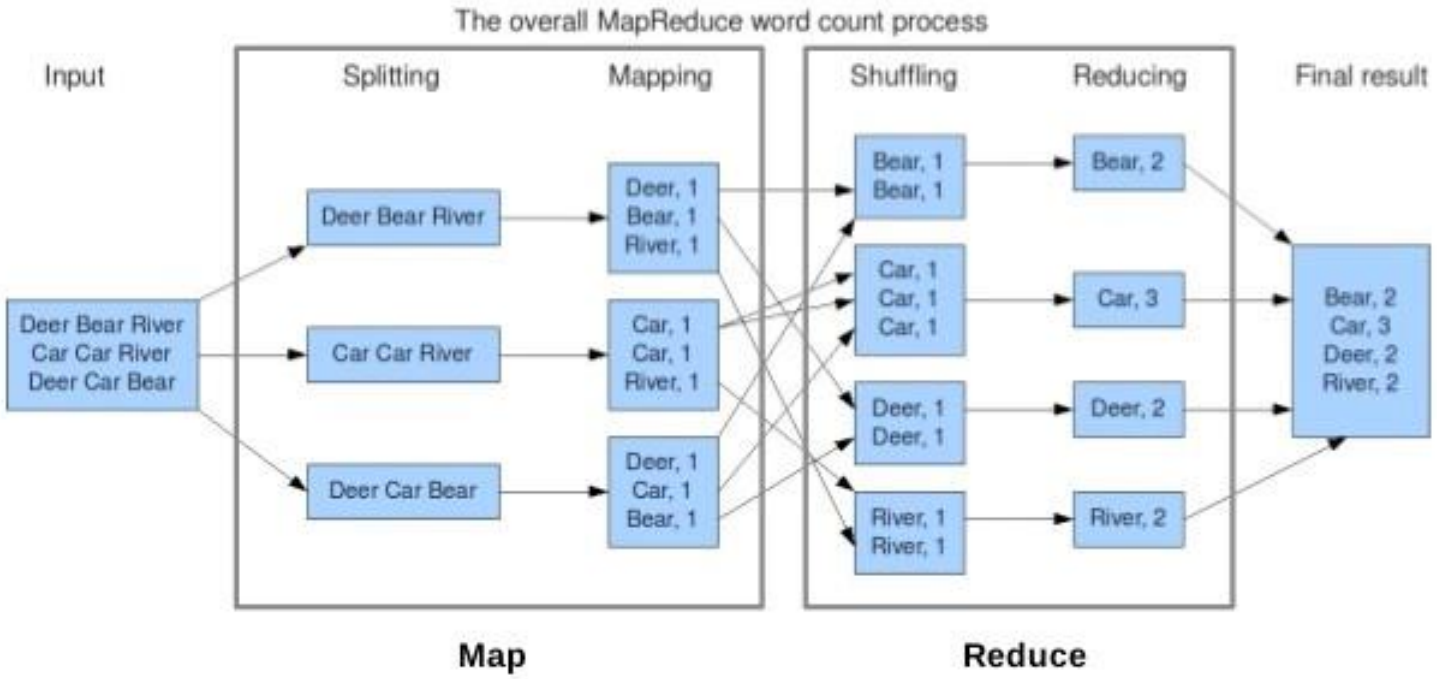
- 主从结构和P2P结构
- 单点失效

- 数据划分与偏斜
- 算子内并行，算子间并行，语句并行

- 从MapReduce说起
- Hadoop生态系统
- MPP
- SQL on Hadoop

2.1 从MapReduce说起

2.1 从MapReduce说起



2.1 从MapReduce说起

```
select word, count(*) from t
                groub by word;

for word in t.readlines():
    counter[word]+=1
for word, count in counter:
    print(word, counter[word])
```

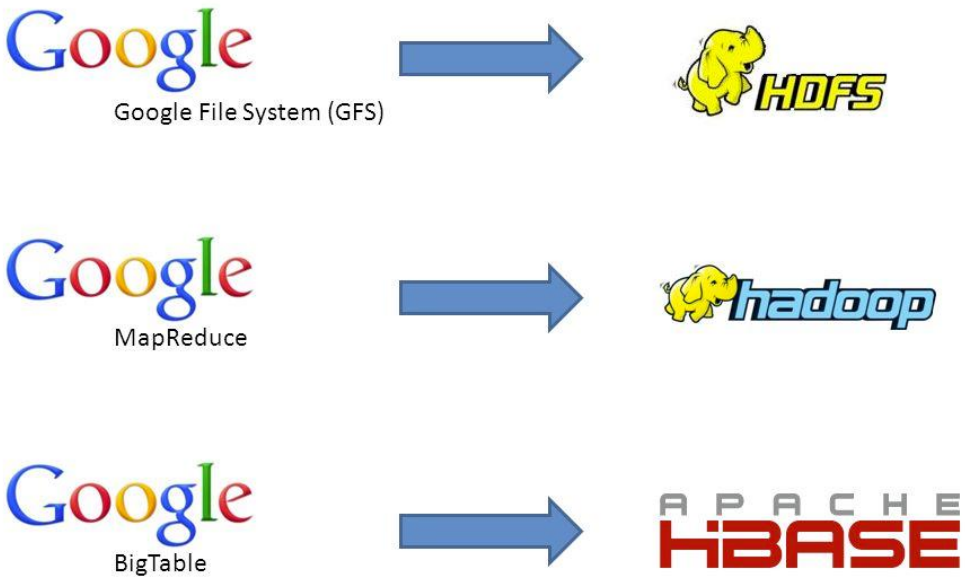
2.1 从MapReduce说起

北京大学 PEKING UNIVERSITY

- why MapReduce?
- 数据倾斜?

2.1 从MapReduce说起

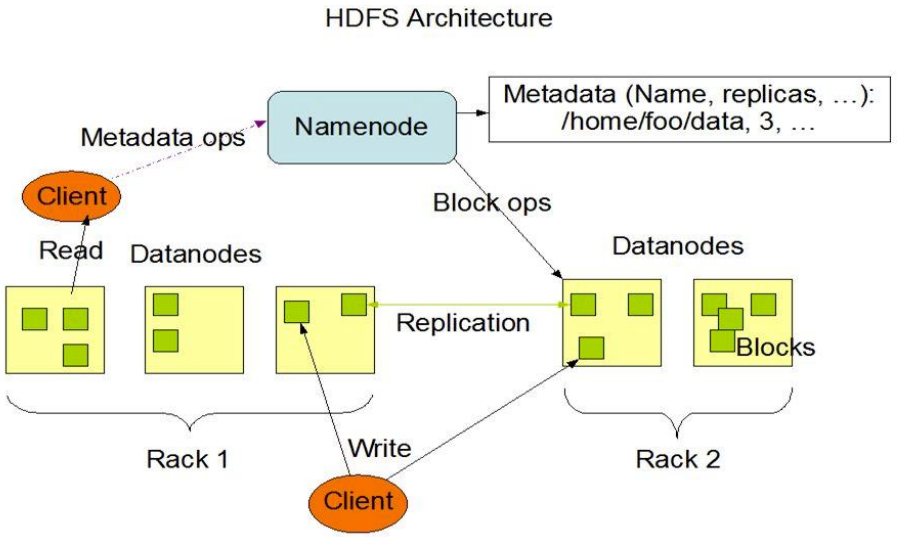
Google *Re*-Engineering



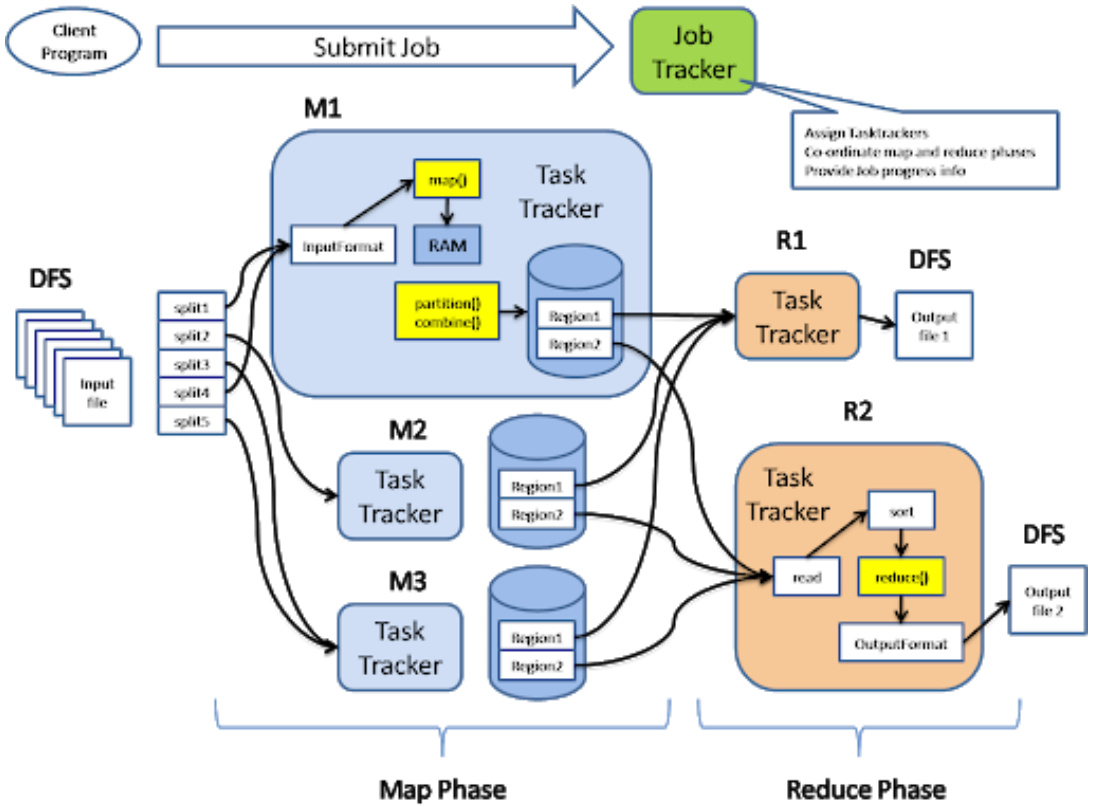
2.2 Hadoop生态系统

2.2 Hadoop生态系统

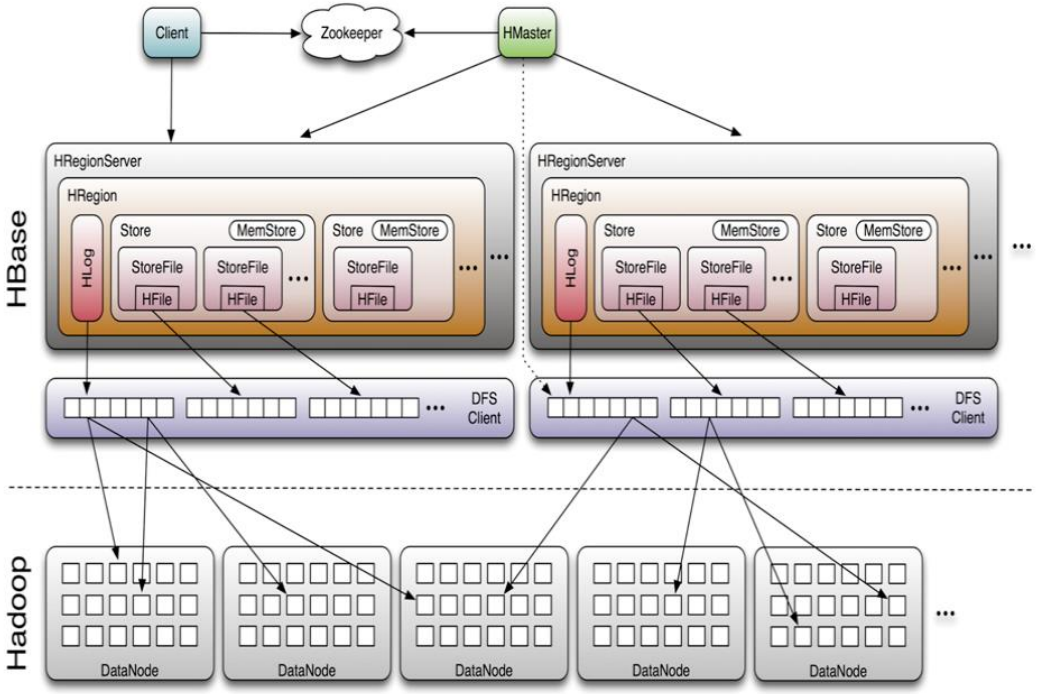
HADOOP ARCHITECTURE



2.2 Hadoop生态系统



2.2 Hadoop生态系统



2.2 Hadoop生态系统

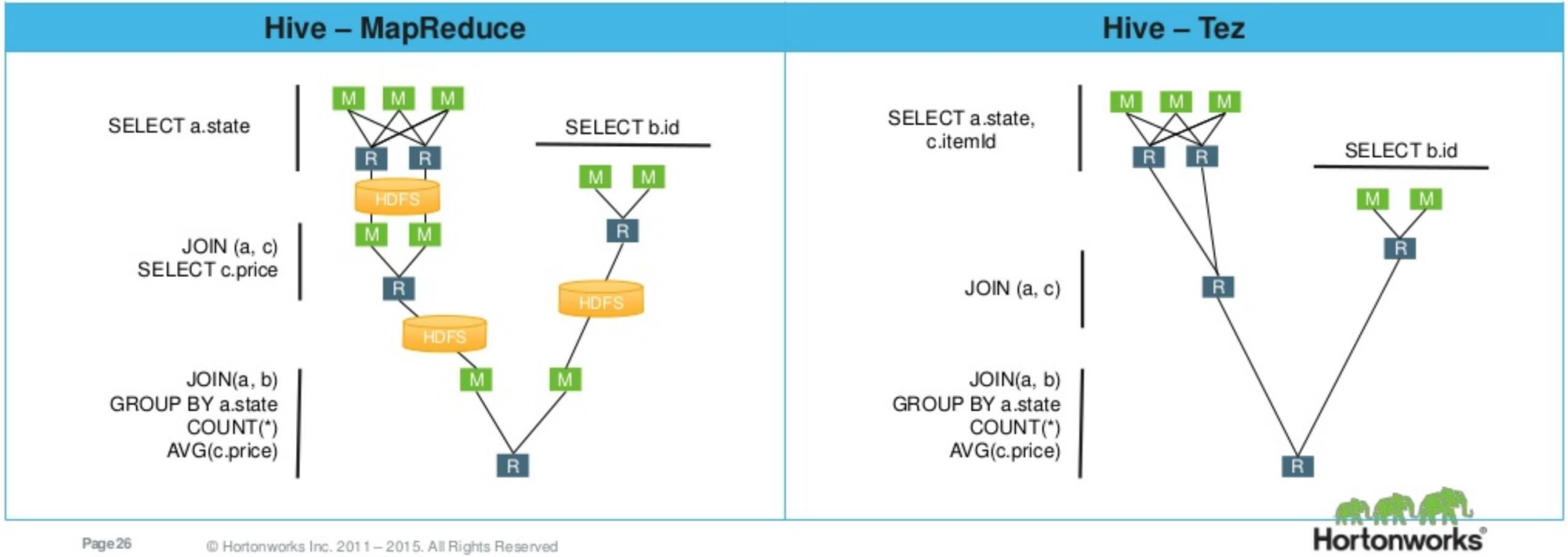
北京大学 PEKING UNIVERSITY

- 回忆一下Snowflake提到的 always on

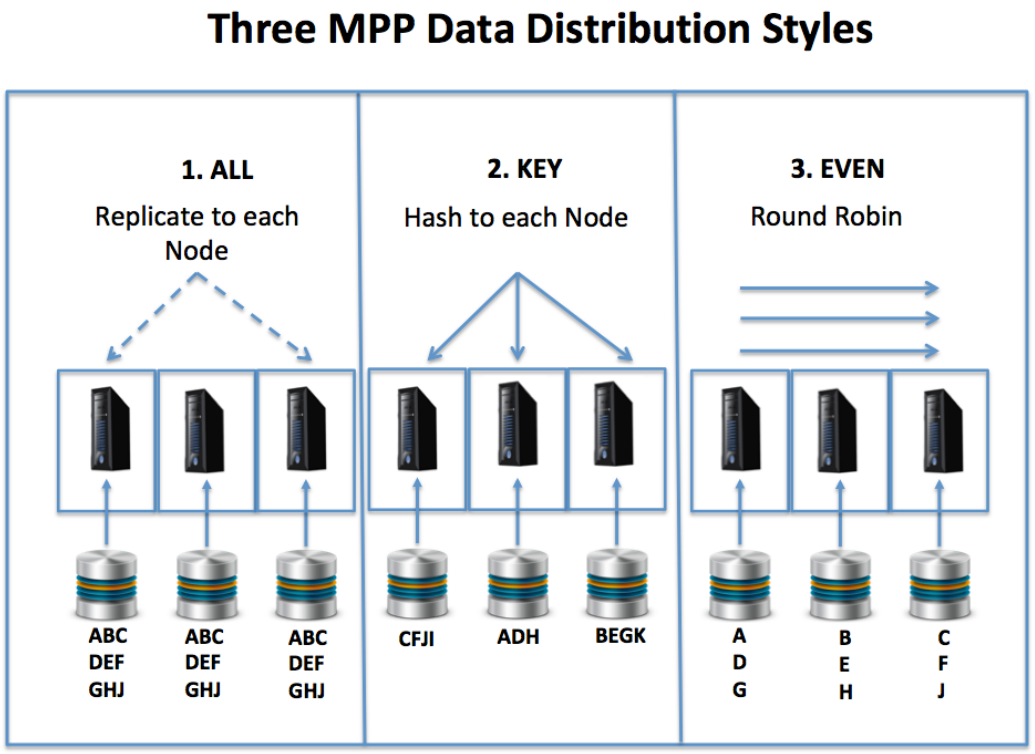
3.1 开源数据仓库实例

```
SELECT a.state, COUNT(*), AVG(c.price)
FROM a
JOIN b ON (a.id = b.id)
JOIN c ON (a.itemId = c.itemId)
GROUP BY a.state
```

Tez avoids unneeded writes to HDFS



2.3 MPP数据库


Three MPP Data Distribution Styles
1. ALL
Replicate to each Node
ABC DEF GHJ
ABC DEF GHJ
ABC DEF GHJ
2. KEY
Hash to each Node
CFJI
ADH
BEGK
3. EVEN
Round Robin
A D G
B E H
C F J

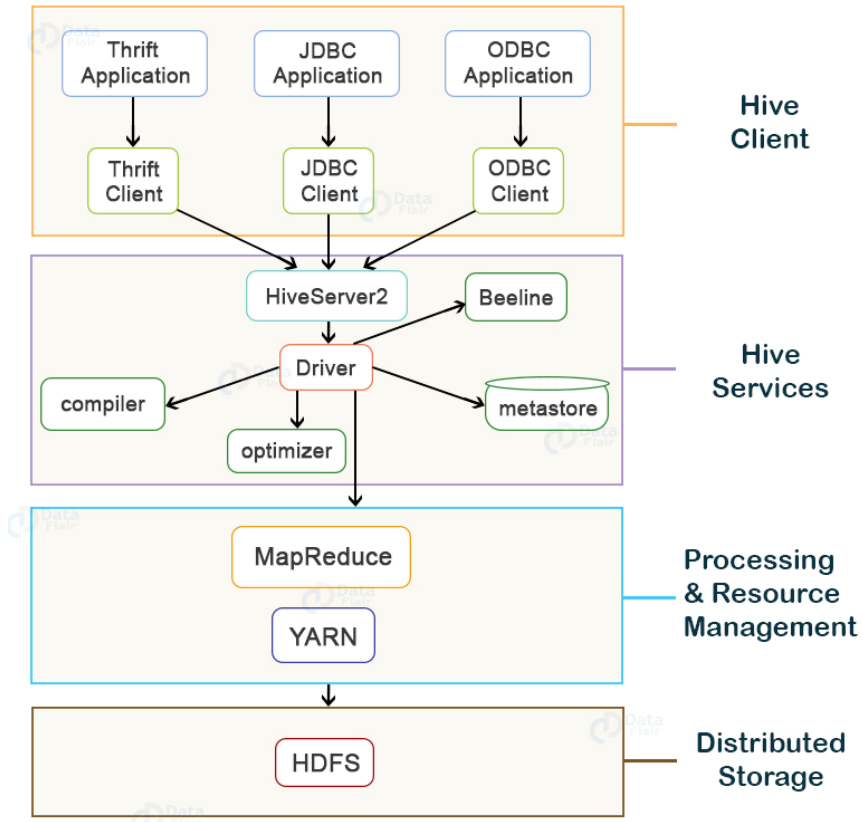
2.4 SQL on Hadoop

系统和技术实践

- Greenplum
- Hive
- Impala

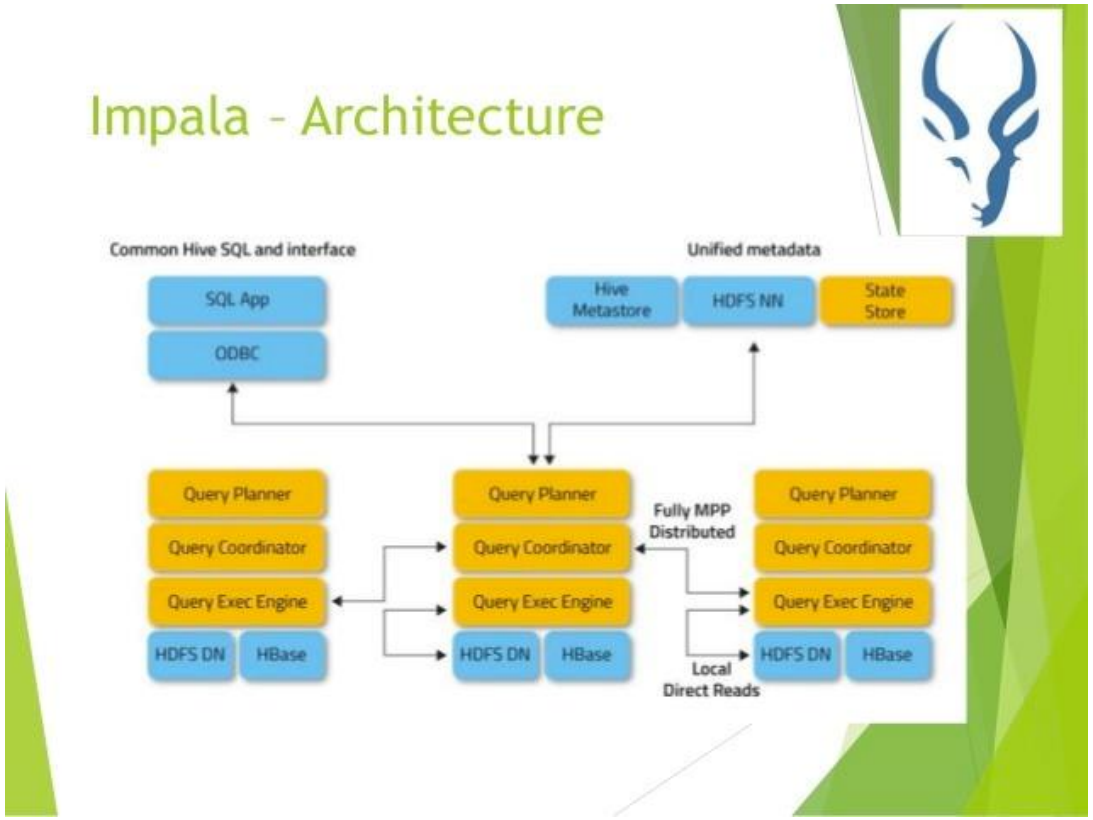
- IO划分技术
  - hash
  - random
  - range
- 互联网络
  - RPC
  - udpifc

Hive

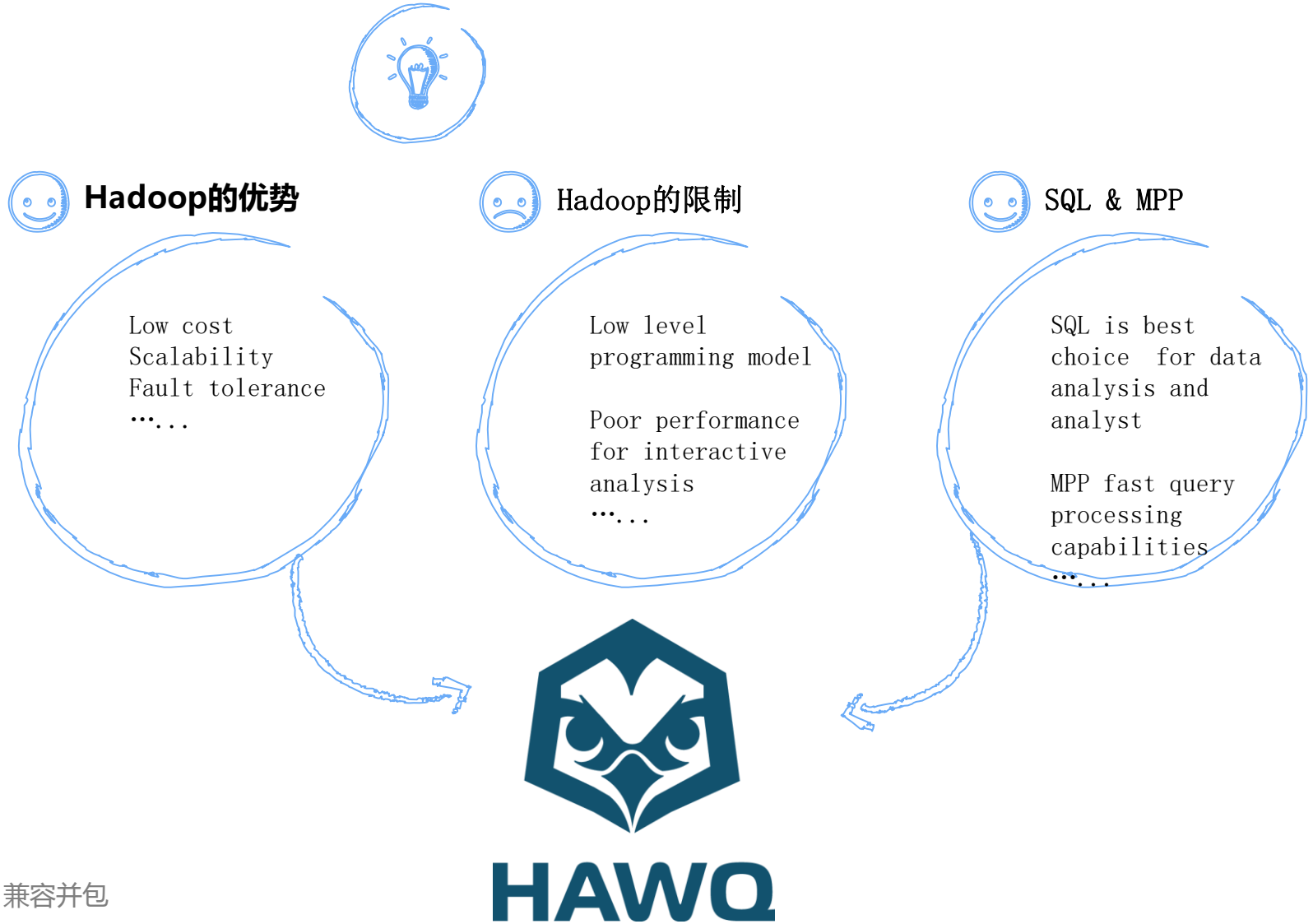


Hive Architecture & Its Components

Impala


Impala - Architecture
Common Hive SQL and interface
SQL App
ODBC
Unified metadata
Hive Metastore
HDFS NN
State Store
Query Planner
Query Coordinator
Query Exec Engine
HDFS DN
HBase
Query Planner
Query Coordinator
Query Exec Engine
HDFS DN
HBase
Fully MPP Distributed
Query Planner
Query Coordinator
Query Exec Engine
HDFS DN
HBase
Local Direct Reads

2.4 HAWQ项目背景


Hadoop的优势
Low cost
Scalability
Fault tolerance
……
Hadoop的限制
Low level programming model
Poor performance for interactive analysis
……
SQL & MPP
SQL is best choice for data analysis and analyst
MPP fast query processing capabilities
……
HAWQ

2.5 个体重复系统发育

2.3 MPP数据库

北京大学 PEKING UNIVERSITY

- MapReduce: A major step backwards
  D DeWitt, M Stonebraker - The Database Column, 2008

- MapReduce: A Flexible Data Processing Tool
  J Dean, S Ghemawat - Communications of the ACM, 2010

2.5 分布式 vs 并行