

Pivotal.

Pivotal Greenplum和HDB简介

陆公瑜 (blu@pivotal.io)

目录

- Pivotal公司简介
- Pivotal Greenplum简介
- Pivotal HDB（基于Apache HAWQ）简介

Pivotal公司简介

Pivotal

大数据套件



内存计算：数据加载、查询、实时处理

可伸缩的存储和计算：
HDFS/Object

Lab服务



自动化：应用供应和生命周期管理

CF

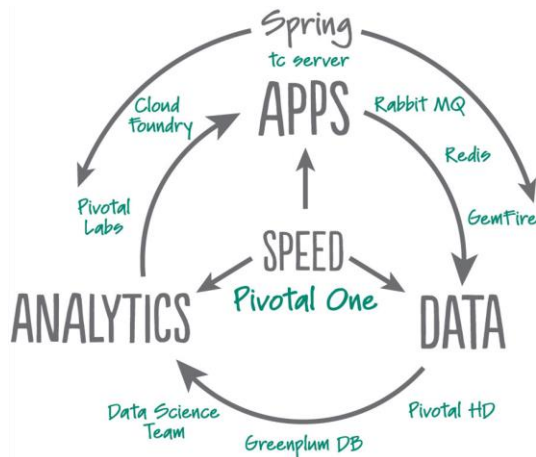
服务注册和服务目录

IaaS云抽象 (可迁移)

Pivotal Cloud Foundry



员工数	公司市场估值
2000+	\$2.8B+



Pivotal Labs现代软件开发



敏捷开发



产品设计



产品管理



创新工场



发现与架构



敏捷沉浸



质量管理



部署



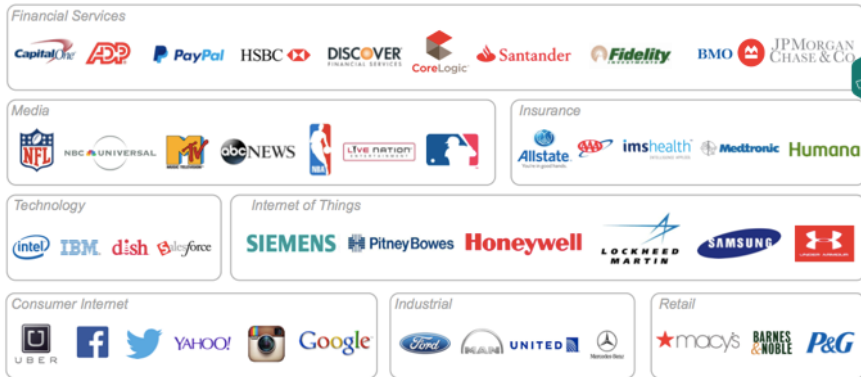
数据科学

Pivotal中国研发中心

中国研发中心主要集中在以下业务

- Pivotal Big Data Suite(BDS):
 - Pivotal Greenplum Database(GPDB)
 - Pivotal HDB(基于Apache HAWQ)
- Pivotal Cloud Foundry
- Pivotal Labs极限编程咨询
- Pivotal全球客户支持服务和关键业务支持服务

我们全球的用户



全球范围不同行业的用户采用了
Pivotal的产品和服务

Pivotal Greenplum简介

Pivotal Greenplum数据库

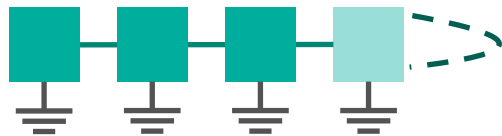


- 在全球34个国家运营
- 超过1000个企业安装实例
- 大规模并行处理（MPP）无共享（Share Nothing）架构
- 对硬件无要求
- 多种使用方式：EMC一体机，软件，云
- 完整的ACID支持，为大数据处理和分析优化
- 兼容SQL标准
- 基于PostgreSQL
- 超过10年的研发投入
- 120+代码贡献者自从2015年10月开源
- 2016年发布了10个版本，还有新组件的发布

Greenplum架构

In-Memory Data Grid

High Speed
Ingestion



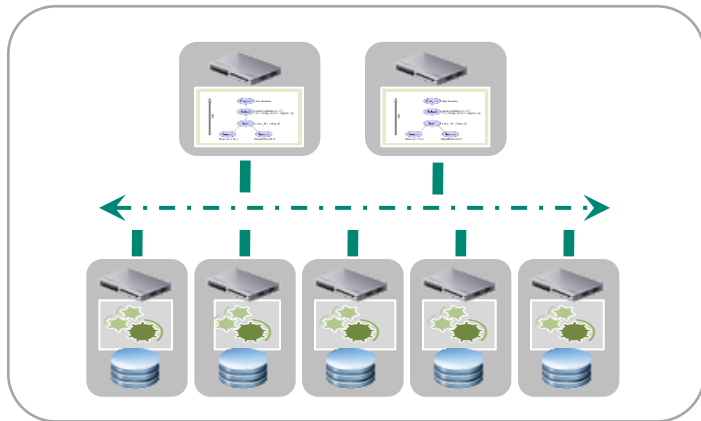
Analytical
Data to cache



Pivotal
GemFire

Parallel Configurable Data Load

Massively Parallel Data Warehouse



In-DB Predictive Analytics

Predefined Libraries



GPTeXt



Programmatic



Pivotal
Greenplum

Parallel Data Load and External Tables

Hadoop Data Lakes



cloudera

MAPR

Public Cloud Data Lakes



Data Temperature

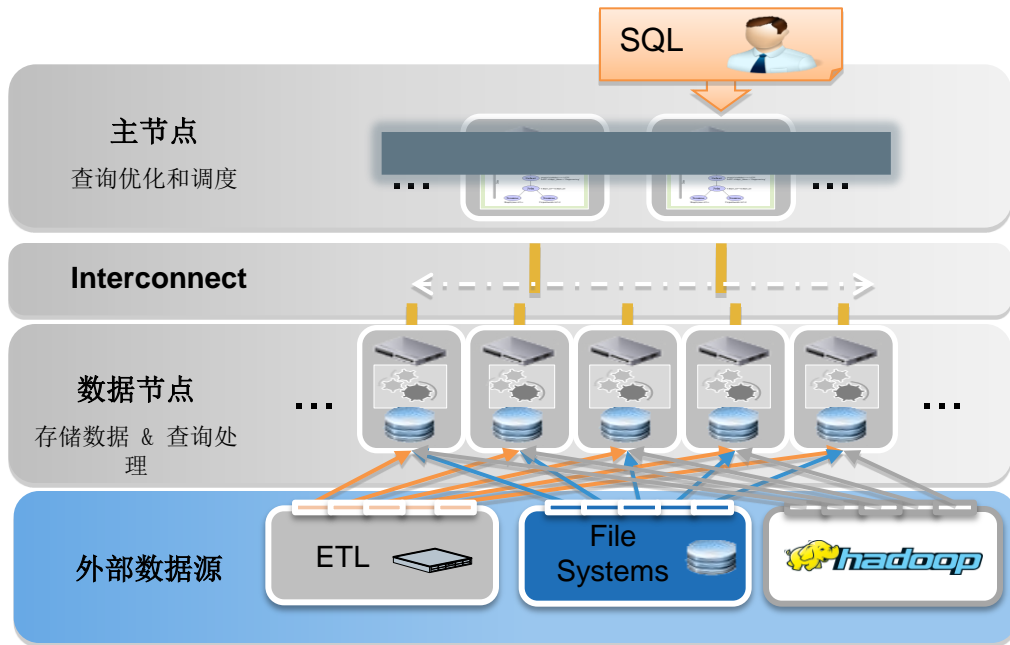
Cold

Warm

Hot

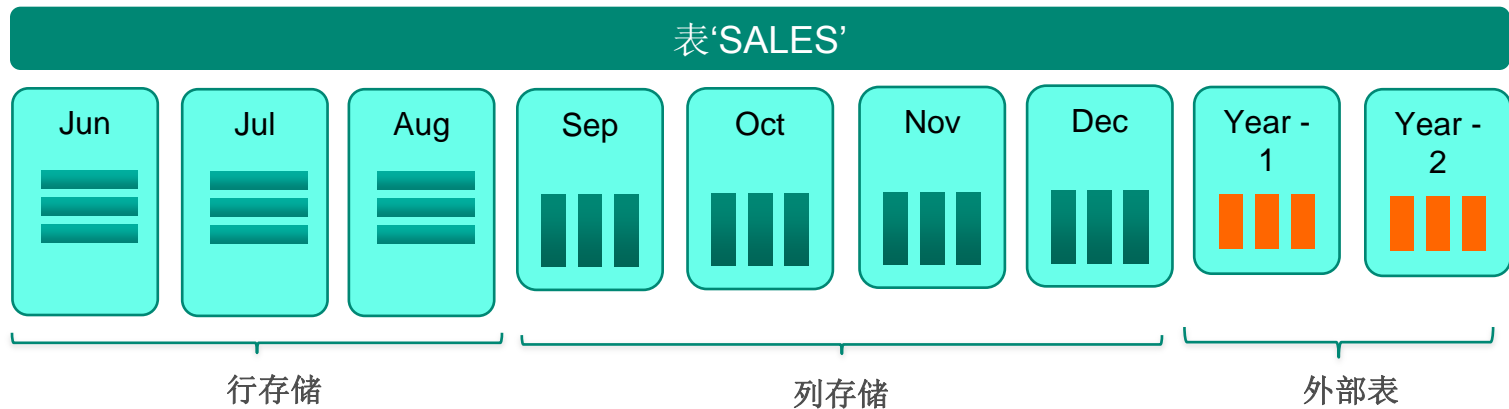
大规模并行数据加载

- 高速数据导入和导出
 - 主节点不是瓶颈
 - 10+TB/小时/Rack
 - 线性扩展
- 低延迟
 - 加载后立刻可用
 - 不需要中间存储
 - 不需要额外数据处理
- 导入/导出 到&从：
 - 文件系统
 - 任意ETL产品
 - Hadoop发行版
 - Amazon S3



多态存储

用户自定义数据存储格式



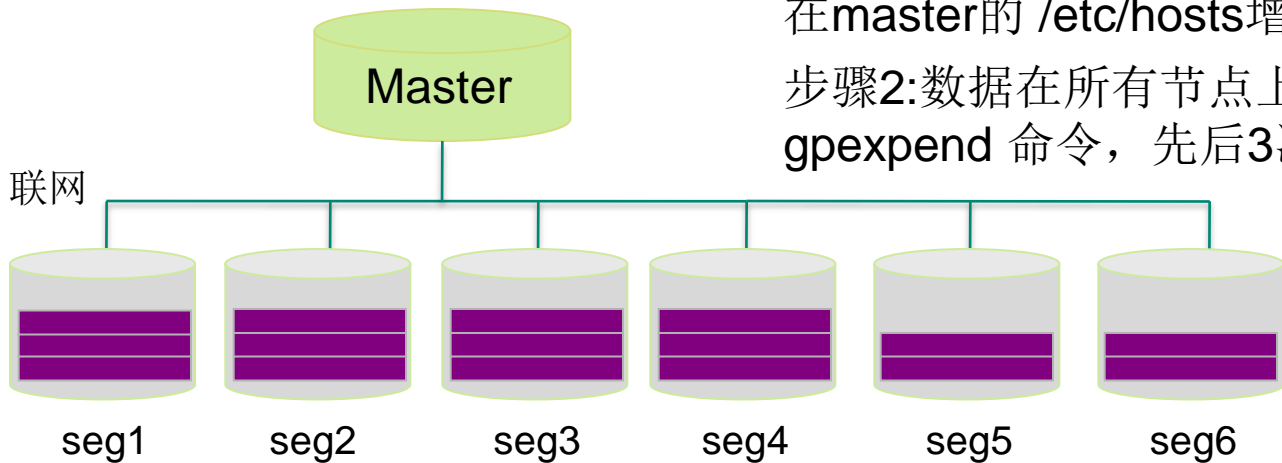
- 访问多列时速度快
- 支持高效更新和删除
- AO主要为插入而优化

- 列存储更适合压缩
- 查询列子集时速度快
- 不同列可以使用不同压缩方式: gzip (1-9), quicklz, delta, RLE

- 历史数据和不常访问的数据存储在HDFS, Amazon S3或者其他外部系统中
- 无缝查询所有数据
- Text, CSV, XML, Binary, Avro, Parquet

动态在线系统扩容

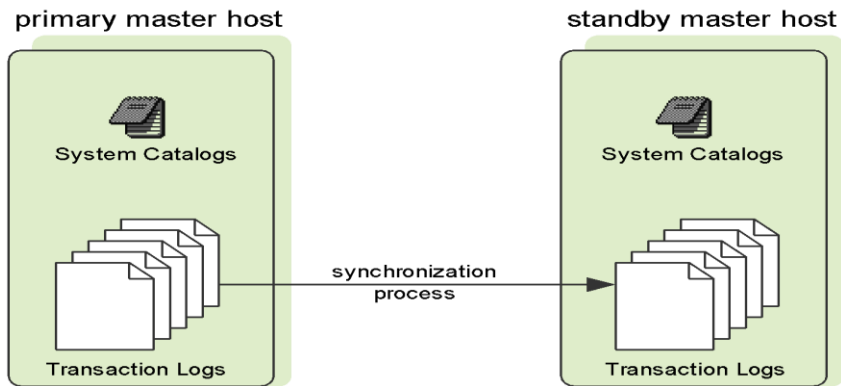
- 数据自动在所有节点上重新分布
- 容量和性能在扩展后线性增长



步骤1:新节点初始化加入MPP集群(
在master的 /etc/hosts增加IP)

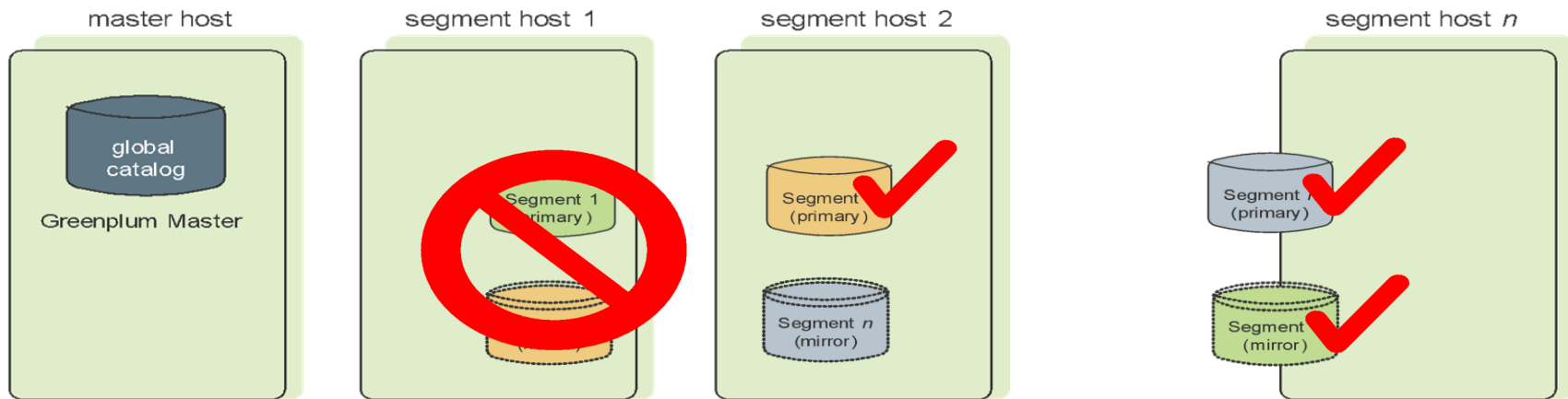
步骤2:数据在所有节点上重分布(运行
gpexpand 命令, 先后3次)

Master节点冗余镜像



- 当主用Master出故障时，热备份Master节点承担它工作
- 热备份Master节点通过复制进程，保持与Master的交易日志同步

Segment节点冗余镜像



Madlib: 运行在GPDB上的机器学习算法库

Apache MADlib (incubating): Big Data Machine Learning in SQL for Data Scientists

Open source,
commercially friendly
Apache license

Supports PostgreSQL,
Greenplum Database™,
and Apache HAWQ
(incubating)

Powerful analytics for
big data

- 运行在GP集群中
- 利用GP的并行运行框架实现并行计算
- 并行数据访问，提高速度

• Data Modeling

- Supervised Learning
 - 朴素贝叶斯分类器
 - 线性及逻辑斯回归
 - Support Vector Machines
 - 决策树
- Unsupervised Learning
 - k-Means Clustering
 - SVD Matrix Factorisation

• Descriptive Statistics

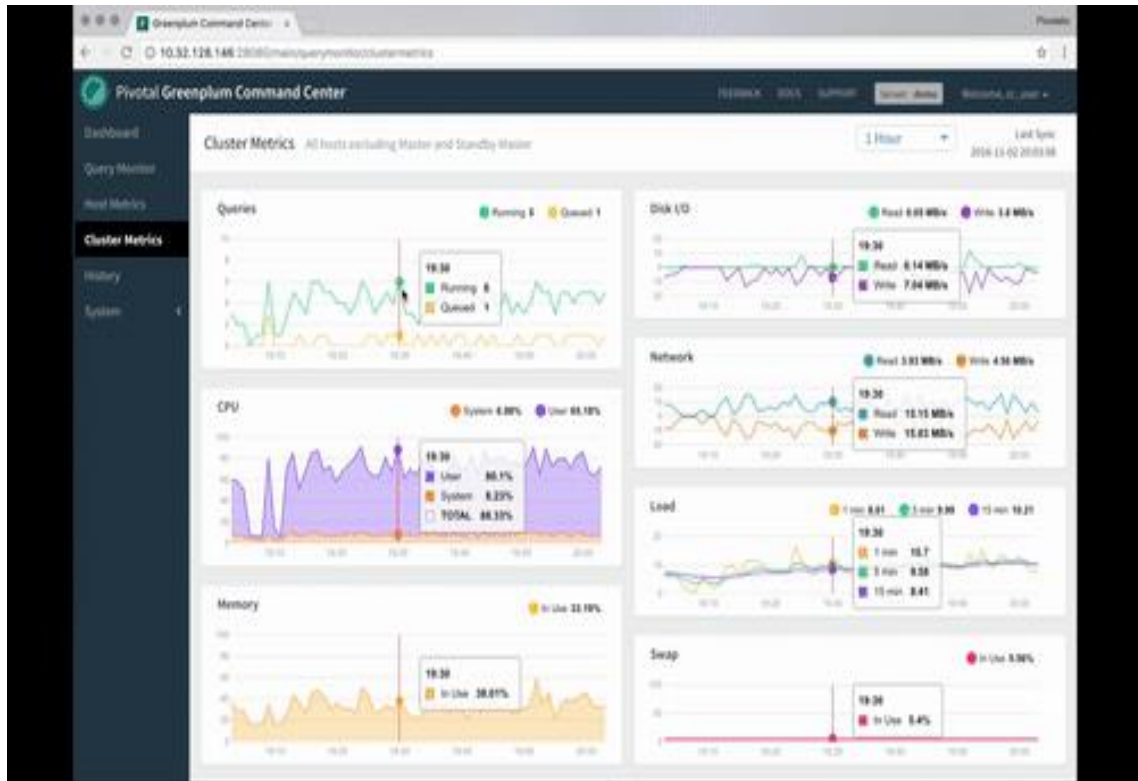
- Sketch-based Estimators (CountMin, Flajolet-Martin, MFV)
- Profile
- Quantile

• Support Modules

- Sparse Vectors

Greenplum管理控制台

- 管理
 - 启动、停止数据库
 - 恢复、rebalance 数据节点
- 交互式系统Metrics
 - 实时
 - 历史 (可以根据时间段配置)
- 系统健康状态监控
 - 硬件状态
 - 软件状态
- Query监控
 - 搜索、调整优先级、取消
 - 查看查询计划
- 工作量管理
 - 配置资源队列
 - 调整优先级



GPText



用于SQL和文本检索的集成数据仓库

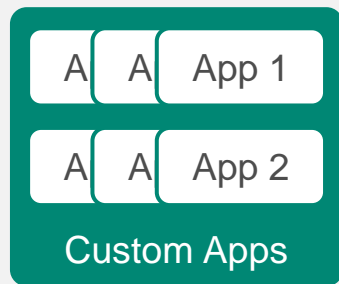
- 集成了Apache Solr and GPDB的优势
- 5年商业化生产经验
- Madlib集成用于文本数据的机器学习
- PL/Python和PL/Java集成用于自然语言处理

适用场景

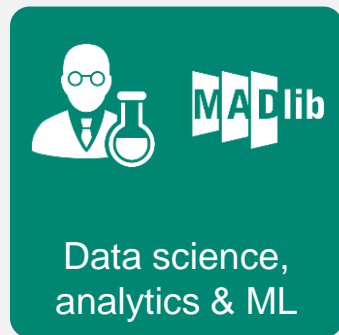
- 通信合规性和监控
- 用户情绪分析
- 文档搜索和查询
- 社交媒体处理，等等



Greenplum/Gemfire连接器



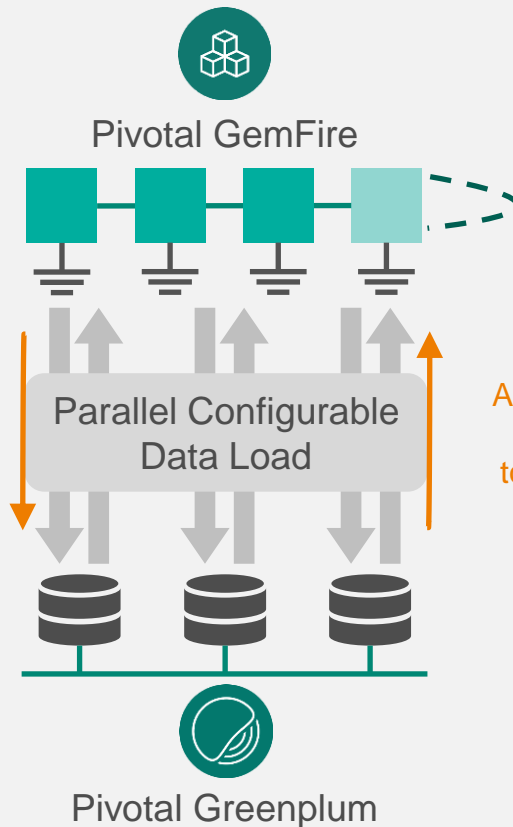
Transactional
Native API
Rest / HTTP



Analytical
ANSI SQL

Transactional
data
Write behind

Analytical
Data
to cache



Hot

Data

Temperature

Warm

Greenplum对云的支持

Greenplum对云的支持非常完善

- Amazon AWS
- 微软Azure
- Google云
- 阿里云
- Vmware

Greenplum开源和未开源组件

- 开源: Greenplum数据库, ORCA查询优化器
- 未开源
 - GPCC (管理控制台)
 - GPText
 - G2C (Greenplum/Gemfire连接器)
 - WLM (工作负载管理器)
 - PL/Containers
 - 第三方组件, 比如一些压缩工具

DCA一体机Greenplum Segment节点配置

Component	V1 Segment Servers	V2 Segment Servers	V3 Segment Servers
Processor	Dell R510, 2 X5670	2 E5-2660	2 E5-2680v3
Memory	48GB (6 x 8GB DIMMS)	64GB (DDR2) 256GB	256GB of DDR4 @ 2133MHz (8x32GB DIMMs),
Size	2U	2U	2U
Sockets	Dual Socket	Dual Socket	Dual Socket
Cores	12	16	24
Network	2 x 1GB/s network interfaces on mother board, dual-port 10gb CNA	Single and Dual 10G Nlc Quad GE ports on-board	Single 10G Dual Port NIC for internal connection
Controller	H700 Internal RAID card	2 RAID Controllers (Mezzanine + PCIe based) and 36 ports Expander board (PCI based). 3 x 8 ports backplane	1 x 12Gb x8 SAS controller (LSI based, Intel branded) w/ 12Gb SAS expander
Disks	12x 3.5" 600GB SAS	(24) 2.5" SAS drives (300GB or 900GB)	(24) 1.8Tb 10k RPM SAS
Servers in Full Rack	16	16	16

Greenplum中国研发团队

Greenplum中国研发团队成员全部来自于北大、北航、中科院等一流院校数据库团队负责GPDB内核研发，全部为开源greenplum.org社区贡献者GPDB相关的2个重要产品GPText和Command Center全部由中国团队负责

GPDB中文blog: <http://gpdb.rocks>

微信群: Greenplum官方纯技术讨论群, Greenplum官方纯技术讨论二群

QQ: 99194625



Greenplum官方纯技术讨论二群



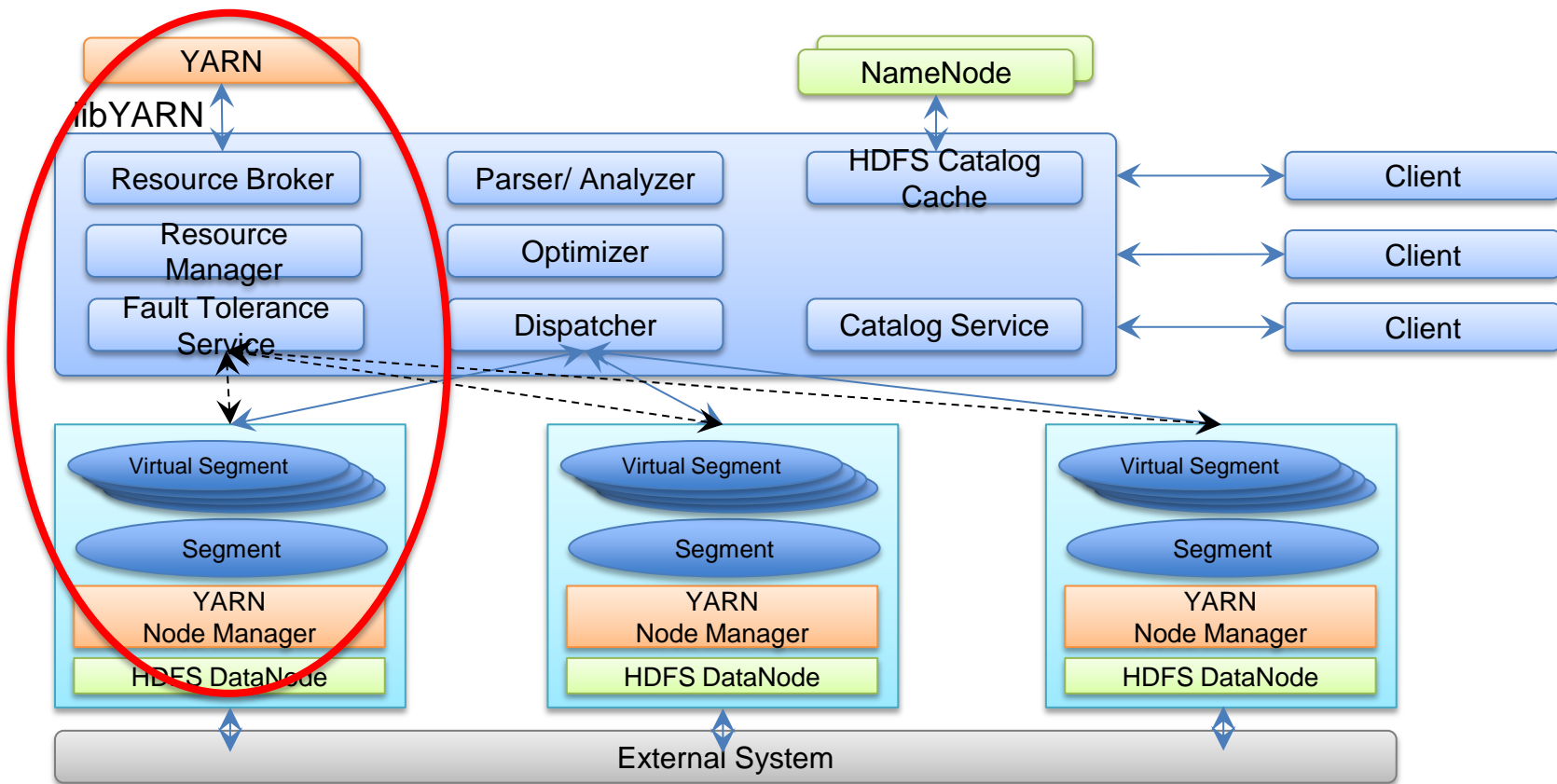
Pivotal HDB(Apache HAWQ)简介

Pivotal HDB (基于Apache HAWQ)产品定位

- 针对分析和机器学习的原生Hadoop SQL并行数据库
- Pivotal HDB是一个完整兼容ANSI-SQL标准的高性能原生Hadoop MPP分析型数据库。提供对Hadoop上PB级数据的高性能交互式查询能力。并且提供对主要BI工具的描述性分析支持，以及支持预测型分析的机器学习库。



Pivotal HDB架构



Pivotal HDB 2.X新特性(1)



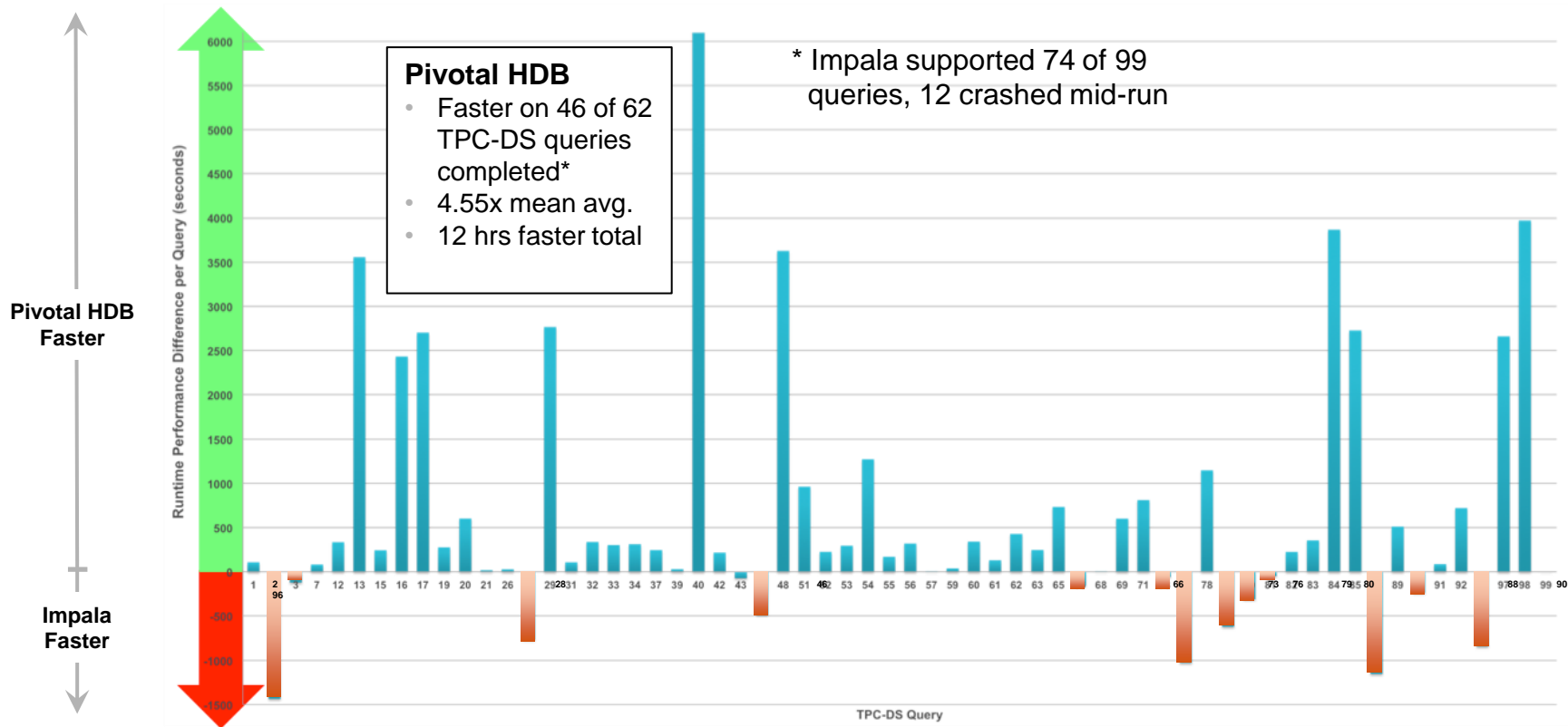
- 资源管理与YARN集成
 - 包含了一个新的资源管理器，支持多级资源管理: YARN级别/查询级别/操作符级别
- 弹性查询执行
 - 基于查询的代价, 可以动态分配若干个virtual segment来执行查询
- 动态集群伸缩
 - 解耦计算和存储, 解耦主节点和从节点, 从而实现秒级扩容
- 块级别存储
 - 优化了AO以及Parquet表存储格式, 使得一个大文件可以被多个Scanner进程以块粒度访问, 从而可以达到优化的查询并行度

Pivotal HDB 2.X新特性(2)



- 表数据文件单目录存储
 - 把一个表的所有数据文件存放在HDFS的一个目录中。这样方便了与外部系统的数据交换
- 新调度器
 - 为了支持弹性查询执行功能, 包含了一个新的调度器。新的调度器可以动态组合不同节点上的查询执行进程来动态调度和管理查询的执行
- 新的容错服务
 - 重新设计了容错服务, 新的容错基于心跳和动态探测。现在可以动态检测到新加的节点并且可以简单的把失效节点从集群中移除
- HDFS元数据缓存
 - 提升了计算到数据调度中数据局部性匹配方法的效率

Pivotal HDB与Impala的性能对比



SQL-on-Hadoop引擎比较

私有软件 & 闭源
& 非线性可扩展

原生Hadoop & 开放
& 线性可扩展

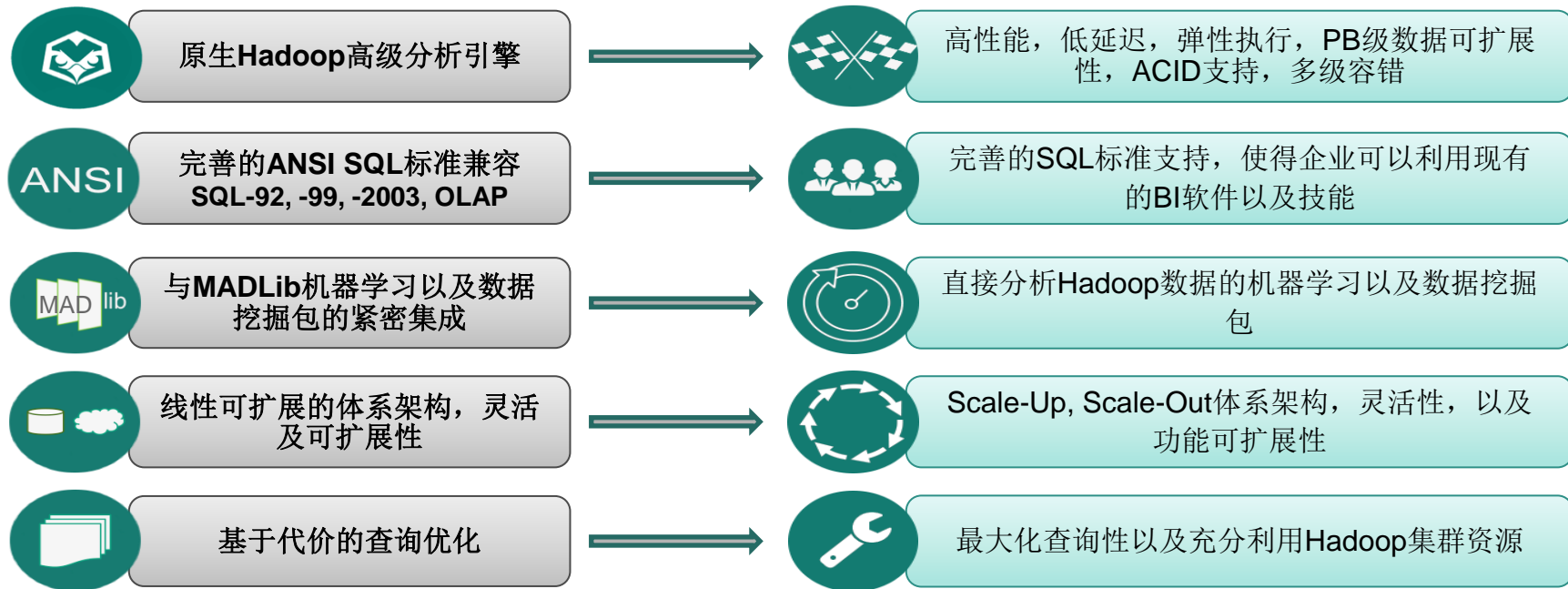
高性能及
SQL兼容性



受限的性能
及SQL兼容性



Pivotal HDB: 优点与好处

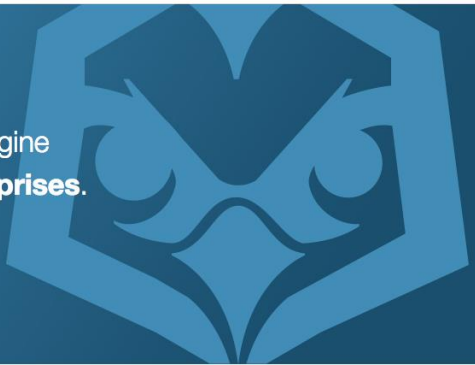


Apache HAWQ开源社区

Apache Hadoop Native SQL.

Advanced, MPP, elastic query engine
and **analytic database** for **enterprises**.

Now **incubating** with Apache.



- Pivotal准备把HAWQ以Apache顶级项目毕业并且持续大力发展开源社区



Pivotal HDB中国研发团队

初创团队

团队成员全部为Apache Committer

承担核心研发工作

成果在国际数据库顶级会议发表

拥有数十项美国专利

微信群：HAWQ官方纯技术讨论群



HAWQ官方纯技术讨论群



Thanks/Q&A

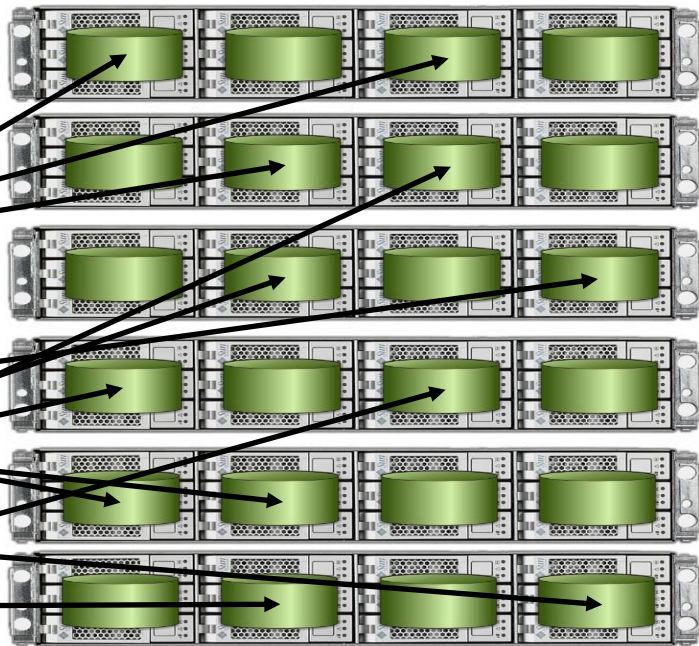
Pivotal®

补充页

数据分布：并行化的根基

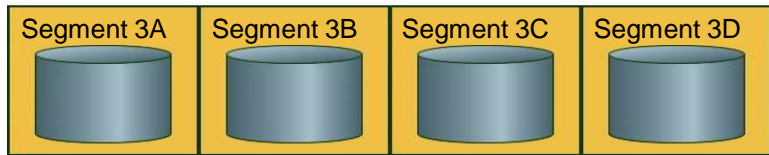
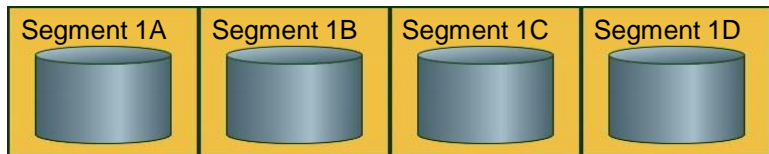
最重要的策略和目标是均匀分布数据到各个数据节点。

Order		
Order #	Order Date	Customer ID
43	Oct 20 2005	12
64	Oct 20 2005	111
45	Oct 20 2005	42
46	Oct 20 2005	64
77	Oct 20 2005	32
48	Oct 20 2005	12
50	Oct 20 2005	34
56	Oct 20 2005	213
63	Oct 20 2005	15
44	Oct 20 2005	102
53	Oct 20 2005	82
55	Oct 20 2005	55



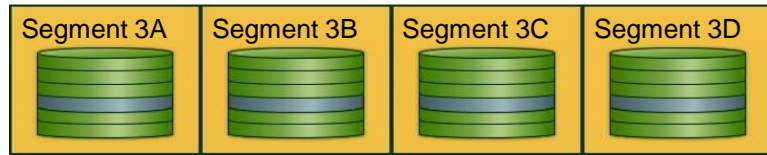
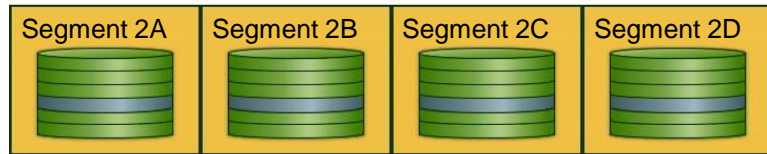
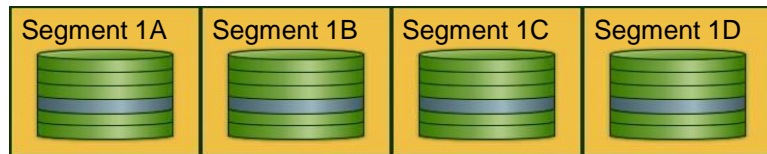
分布和分区

```
SELECT COUNT(*)  
FROM orders  
WHERE order_date >= 'Oct 20 2007'  
AND order_date < 'Oct 27 2007'
```



orders表数据均匀分布于各个节点

&



仅仅扫描orders表相关的分区

Greenplum平台概况

客户端访问和工具	客户端访问 ODBC, JDBC, OLEDB, etc.	第三方工具 BI 工具, ETL 工具 数据挖掘等	管理工具 GP Command Center GP Workload Manager
产品特性	加载 & 数据联邦 高速数据加载 任意系统数据访问	存储 & 数据访问 混合存储引擎（行存&列存） 多种压缩 多级分区表 索引（B树，位图，GiST）	语言支持 标准 SQL 支持 SQL 2003 OLAP 扩展 扩展编程语言 (Python, R, Java, Perl, C)
服务	多级容错机制	在线系统扩展	任务管理
核心MPP架构	无共享大规模并行处理 先进的查询优化器 多态存储系统 并行数据流引擎		

Pivotal HDB (基于Apache HAWQ)



- 具有MPP（大规模并行处理系统）的性能，比其他Hadoop里面的SQL引擎快数倍
- 完善的标准支持：ANSI SQL标准，标准JDBC/ODBC支持，比其它Hadoop SQL引擎都要完善
- 支持ACID事务特性
- 动态数据流引擎：基于UDP的高速互联网络
- 支持多种分区方法及多级分区
- 支持多种压缩方法：snappy, gzip, quicklz, RLE等
- 支持访问任何HDFS及其他系统的数据
- 原生的机器学习数据挖掘库MADLib支持：易于使用及高性能
- 支持多种第三方工具：比如Tableau, SAS, 较新的Apache Zeppelin等
- 支持在本地、虚拟化环境或者在云端部署