

# VMware Greenplum 6 server 新特性与新版本 展望

Wen Wang  
Vmware GP Product Manager

12/09/2022

# Disclaimer

- This presentation contains statements which are intended to outline the general direction of certain of VMware's offerings.
- It is intended for information purposes only and may not be incorporated into any contract.
- Any information regarding the pre-release of VMware offerings, future updates or other planned modifications is subject to ongoing evaluation by VMware and is subject to change.
- All software releases are on an if and when available basis and are subject to change.
- This information is provided without warranty or any kind, express or implied, and is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions regarding VMware offerings.
- Any purchasing decisions should only be based on features currently available.
- The development, release, and timing of any features or functionality described for VMware's offerings in this presentation remain at the sole discretion of VMware.

# 目录

## Greenplum 6 Core Server new features

- Greenplum to Greenplum Data Federation (GP2GP)
- Resource management 资源管理改进
- SCRAM-SHA-256身份验证方法的支持
- AO 表上的vacuum 索引功能增强

## Greenplum 7 Roadmap

- PG12 support
- Resource management 资源管理持续改进
- AO/CO enhancement: Primary Key/ Unique Indexes for AO/CO tables
- Altering the storage type of the table is supported

# Greenplum Database Version 6

The world's best and only OSS Multi-Platform Data Warehouse



Greenplum is the Massively Parallel Processing Workhouse for big database analytics

Running Billions of SQL Daily Globally  
on Exabytes Data

*Greenplum Version 6 has achieved massive success in customer adoption running billions of diverse complex SQL daily with minimal support intervention*

Powered by:

1. PostgreSQL base code line
2. ORCA Optimizer
3. Write Ahead Log High Availability
4. PXF, Spark and Streaming
5. Managed by Command Center

Any Data | Any Size | Anywhere

# Greenplum to Greenplum Data Federation (GP2GP)

## 1. GP2GP:

1. 帮助客户轻松快速地跨不同的Greenplum集群进行查询
2. 实现高速数据联合和数据共享
3. 扩大了 Greenplum cluster的集群规模
4. 为客户节省了资金

## 2. 特性:

- 正确性:
  - 保证全局事务和MVCC
- 兼容性:
  - 通过libpq协议控制和传输数据
- 效率:
  - 将数据 Pushdown , segment 之间直接相互传输

## 3. 组成:

1. Parallel retrieve Cursor
2. Greenplum Foreign Data Wrapper (FDW)



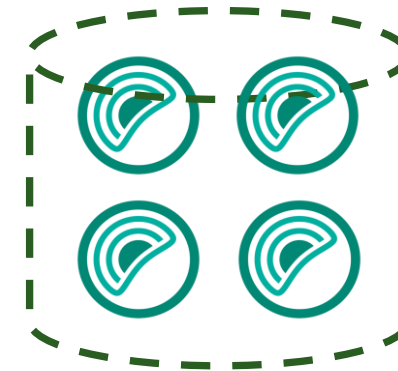
Scale



Efficiency



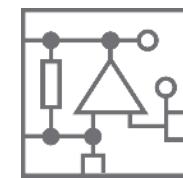
Parallel Cursor



FDW



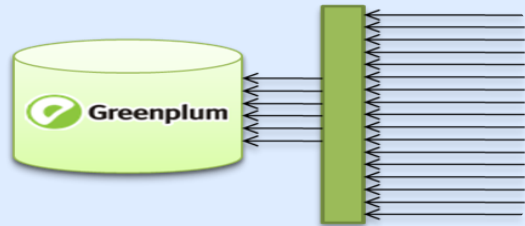
Transaction Safe



Heterogenous Cluster Configuration

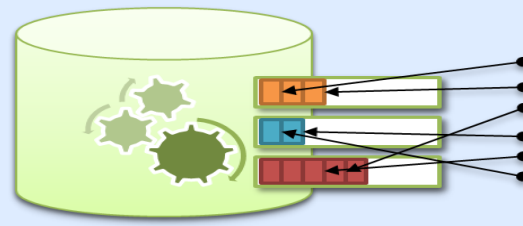
# Resource management 资源管理改进

## Resource Group



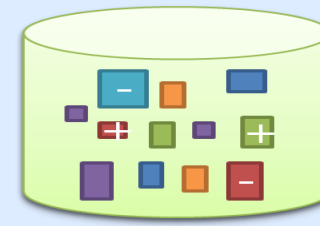
### Connection Management

- Control over how many users can be connected.
- Provides pooling (to allow large numbers) and caps (to restrict numbers if desired)
- Intelligently frees and reacquires temporarily idle session resources
- PGBouncer or Heimdahl



### User-Based Resource Groups

- Each user is assigned to a resource group that performs 'admission control' of queries into the database
- Allows DBAs to control the total number or total cost of queries allowed in at any point in time
- Memory allocations per group



### Dynamic Query Prioritization

- CPU Levels per group
- Allows DBAs to control query priorities in real-time, or determine default priorities by resource group

# Resource management 资源管理改进

## 管理范围

- CPU, Memory, Concurrency

## 新特性

- 精细化配置, 资源得到更充分的利用
  - 根据Coordinator和primary segments的内存配置, 分别计算每个query的最大内存分配额
- 高负载下改善性能, Hyper-Threading Best Practices
  - 更好的单个查询性能、
  - 资源组(RG)之间的隔离
  - 以及较低的运行与运行差异
- Debug信息细化
  - 根据Vmem、系统、查询和资源组内存分配的当前状态输出不同的错误消息

# SCRAM-SHA-256 认证支持

## GP6新安全密码功能

- 背景：
  - Md5: 不再能满足高安全用户的需求
  - Sha-256认证方式: 网络发送密码格式为明文, 必须与SSL一起工作
- SCRAM-SHA-256
  - 密码哈希算法和客户端身份验证方法的支持
  - 最安全
  - Backport from PG10



# AO 表上的vacuum 索引功能增强

## 收益:

- Vacuum index on AO 实现方法改进
- 提高运行效率, 改善性能

## 分析:

- AO 表特点, 分片文件
- Vacuum on AO
- Vacuum on AO index

GP 7 more features in coming



**GREENPLUM  
DATABASE®**

**Version 7.0**

# Greenplum and PostgreSQL

The only OSS PostgreSQL based Data Warehouse for Enterprise



Postgres 8.2 Circa  
2006 (-4)



Postgres 8.3  
Circa 2008 (-9)



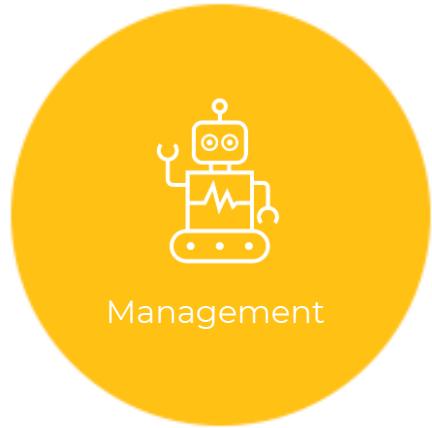
Postgres 9.4  
Circa 2014 (-5)



Postgres 12  
Circa 2019 (-4)

Greenplum v8 based on Postgres 16  
2 year trailing cadence

# PG12 support



REINDEX CONCURRENTLY without blocking writes

Progress reporting statistics for CREATE INDEX, REINDEX, CLUSTER, VACUUM FULL, and pg\_checksums



MCV statistics for query tuning



Row level security control

Scram SHA-256 Authentication

Performance

JIT



Partition table enhancement



Support for the SQL/JSON path language



Windows functions support enhancement



Upsert



Brin index



Vacuum enhancement

Automatic inlining of CTEs

Optimizations to space utilization and read/write performance for B-tree indexes

# Resource management 资源管理持续改进

- CGroup V2 support
- Memory 控制 更易用, 更灵活, 高负载下更稳定
- CPU 控制改进 更稳定, 资源更充分调度
- 增强短query 的并发支持量
- 重构和抽象接口和流程, 进程资源分配

## AO/CO enhancement: unique indexes on AO/CO tables

- 压缩表上使用主键，支持唯一的索引
- Upsert for AO
- 适用于更多的数据模型，业务更简洁

## Altering the storage type of the table is supported

- 支持修改表的存储类型,
  - altering from heap to ao/co, ao to heap/co, co to heap/ao.
- 增强数据兼容性
- 操作简单, 易用性更强
- 高效运行
- 灵活方便
- 对新客户有很大吸引力





# Thank You