Greenplum 集群安装配置及最佳实践

目录

1 总体介绍

1.1 硬件平衡

1、性能
2、容量
3、成本

1.2 高可用

1、节点
2、网络
3、磁盘

1.3 部署方案

1、Master 与 Standby Master 分级部署
2、Primary Segment 与 Mirror Segment 分机部署
3、Segment Mirroring 部署方案
3.1、Group Mirroring
3.2、Spread Mirroring
3.3、Group + Spread Mirroring

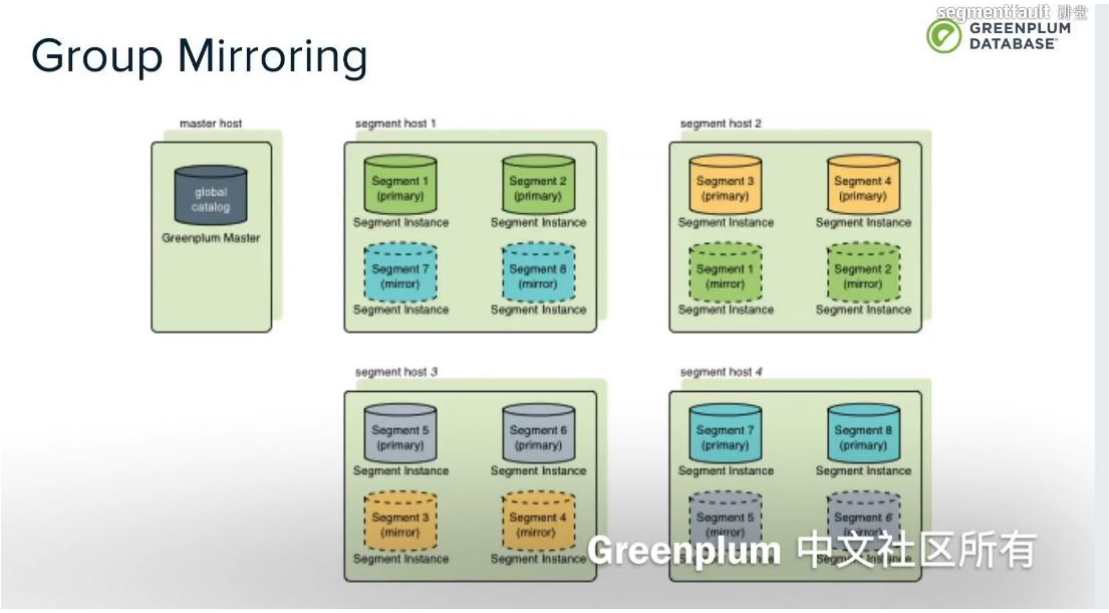
4、Pivotal Supported Greenplum 必须部署 Mirroring Segment
5、铜一主机 Segment 个数

5.1 CPU/Core 数据
5.2 查询并发数
5.3 查询复杂度
5.3 单机 Primary Segment 总数不能过多

1.3.1 Group Mirroring 部署方案

按照以下 4 台机器 Group Mirroring 的部署方案总结

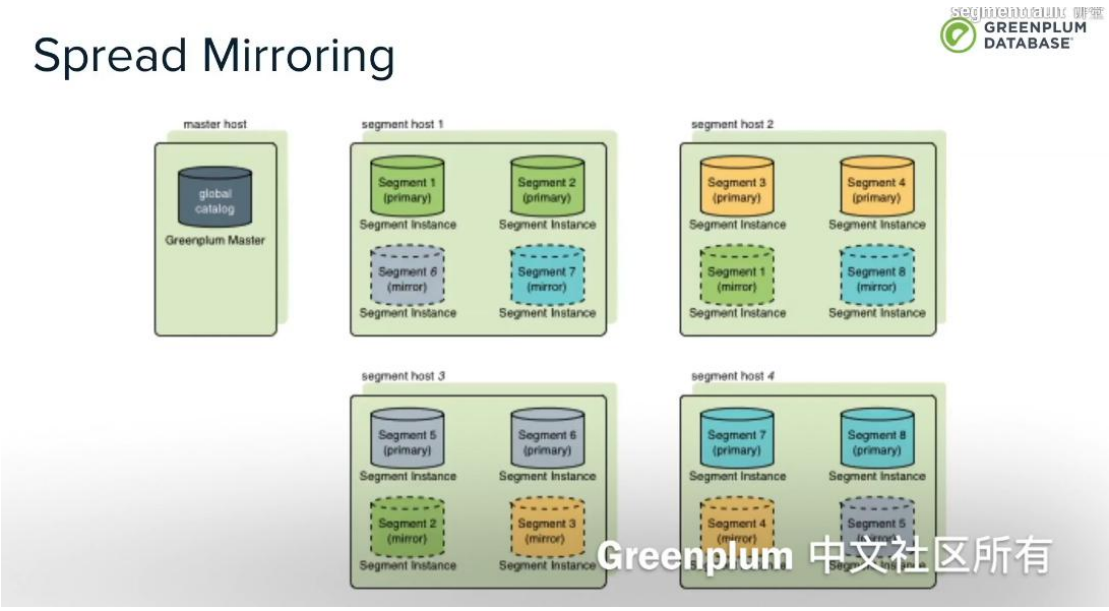
缺点: 一台机器 down 掉后，会把流量全部放在下一个节点，下一个节点的流量会变成 2 倍的流量
优点: down 掉一台机器后，集群能正常的提供服务,如果再 down 掉第二台集群就不可用



1.3.2 Spread Mirroring 部署方案

按照以下 4 台机器 Spread Mirroring 的部署方案总结

缺点: 一台机器 down 掉后，会把流量全部放在下两个节点
优点: down 掉一台机器后，集群能正常的提供服务,如果再 down 掉第二台集群就不可用



1.3.3 Group + Spread Mirroring 部署方案

如果集群比较大建议使用 Group + Spread Mirroring 部署方案，如果集群由 down 流量会分流道其他的机器上,集群不可用的几率比较小。

2 硬件选型

2.1 Master 节点

2.1.1 网卡

1、2 块万兆网卡内部互联
2、1-2 块千兆网卡带外管理及接入客户网络

2.1.2 内存

DDR4 64GB 以上，建议 256G

2.1.3 磁盘

1、6 块 600G/900G　10k RPM SAS 盘

2、采用 RAID5 或 RAID10
3、单独预留 hotspare 盘
4、1 块 RAID 卡，cache 1GB 以上，带有掉电保护功能

2.1.4 CPU

1、2 路 8 核及以上
2、主频 2.5G HZ 以上

2.2 Segment 节点

2.2.1 网卡

1、2 块万兆网卡内部互联
2、1-2 块千兆网卡带外管理及接入客户网络

2.2.2 内存

DDR4 64GB 以上，建议 256G

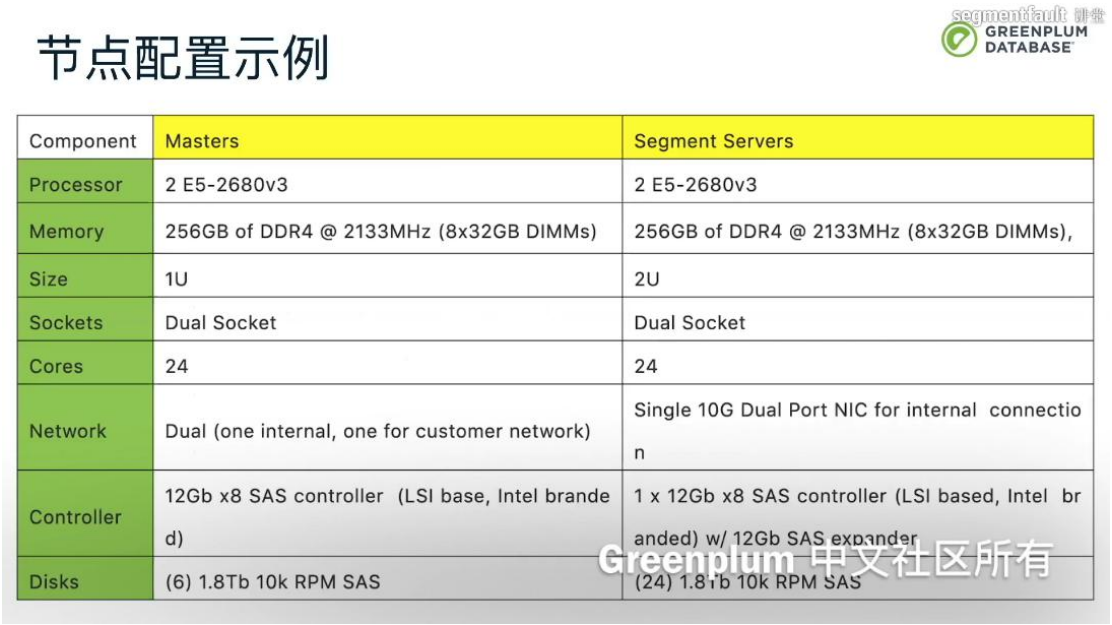
2.2.3 磁盘

1、24 块 600G/900G 10k RPM SAS 盘
2、采用 RAID5 或 RAID10
3、单独预留 hotspare 盘
4、1-2 块 RAID 卡，cache 1GB 以上，带有掉电保护功能

2.2.4 CPU

1、2 路 8 核及以上
2、主频 2.5G HZ 以上

2.3 节点配置实例



2.4 硬件配置经验总结

1、磁盘故障时 Greenplum 集群最常见的故障

1.1 分析性查询: SAS 盘 > SATA 盘

1.2 高并发小 IO 查询: 优先 SSD 或 NVMe

2、RAID 级别

2.1 RAID-5 VS RAID-10

3、RAID 卡一定带 Cache 功能，能提高磁盘的读写性能。

4、硬件监控

5、预留灾备机

RAID-5 VS RAID-10 的区别:

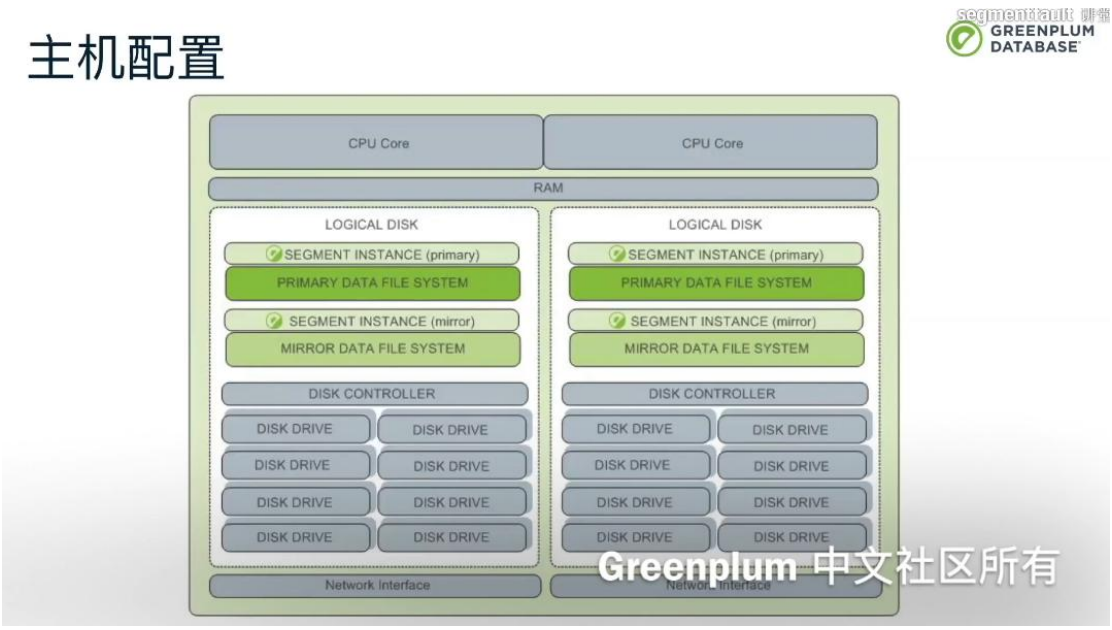
RAID-5:在容量上会比较大，储存会达到 90%的利用率，在读写的性能上比 RAID10 会好一点,RAID5 的可靠性会差很多。

RAID-10:冗余级别更高，储存会达到 50%的利用率，当有的盘出现坏的情况下，磁盘的可靠性有保证。

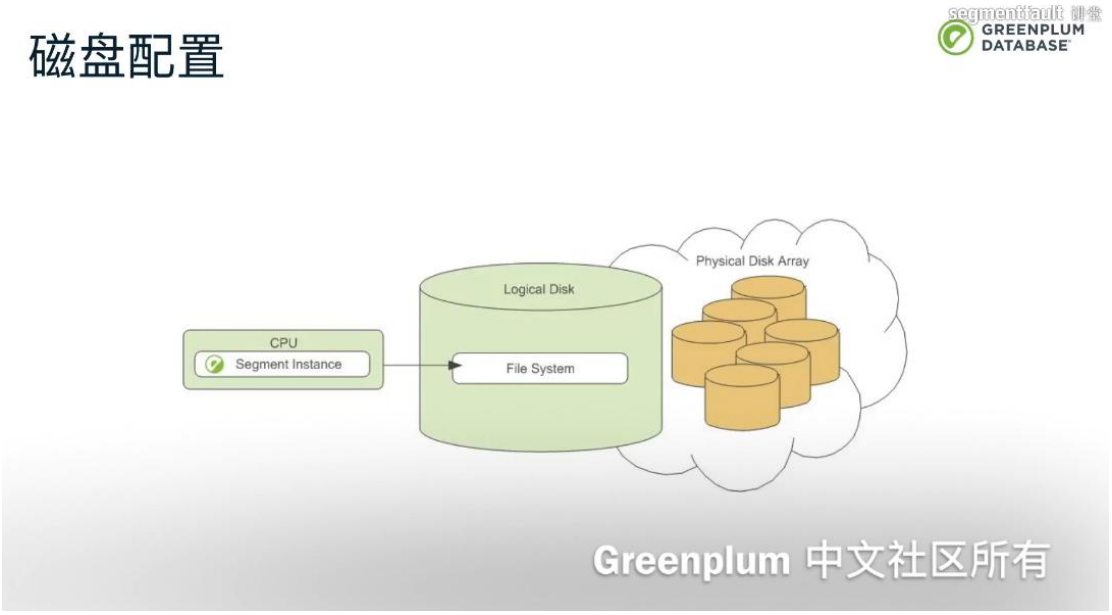
3 硬件配置

1 主机配置实例

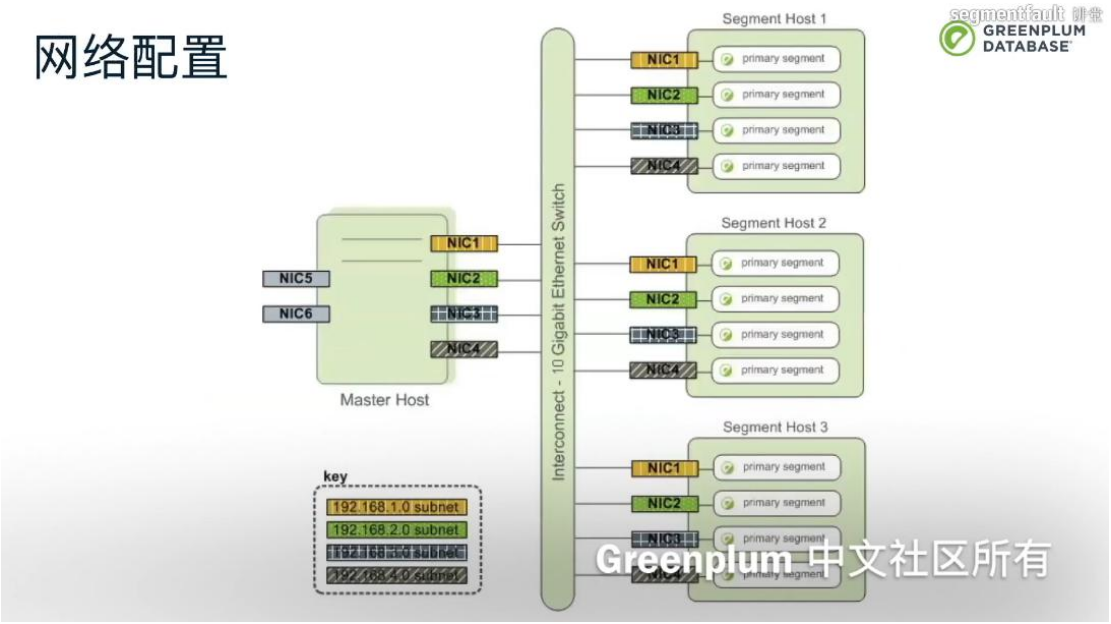
一下的配置两个 CPU 主机的例子,两个 segment 两个 primary 两个 mirror 以及两个网卡的例子



2 磁盘的配置

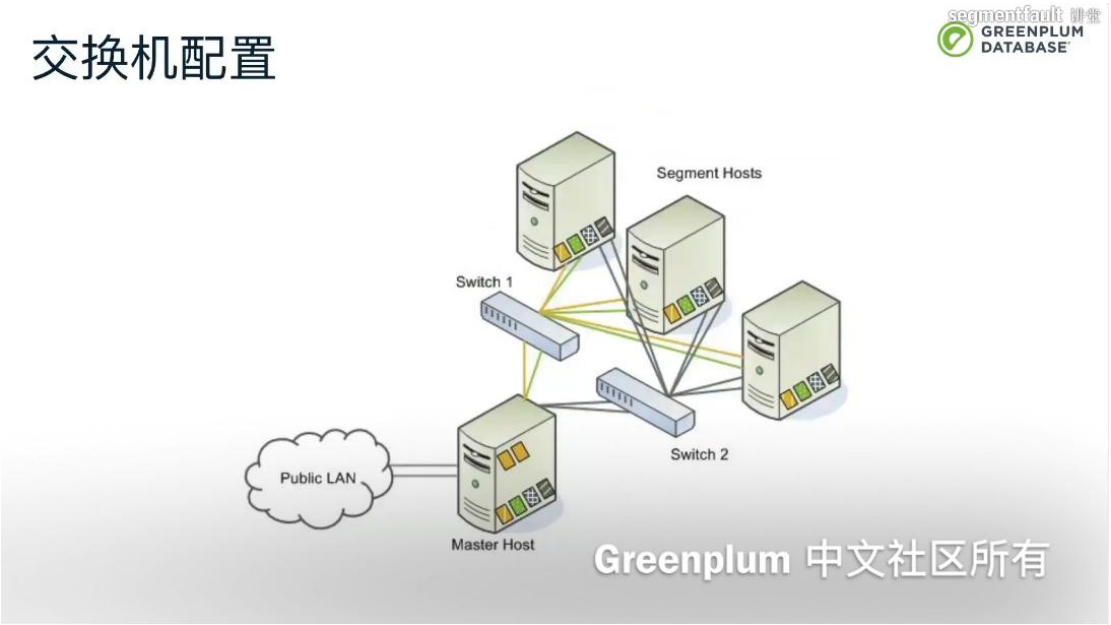


3 网络配置



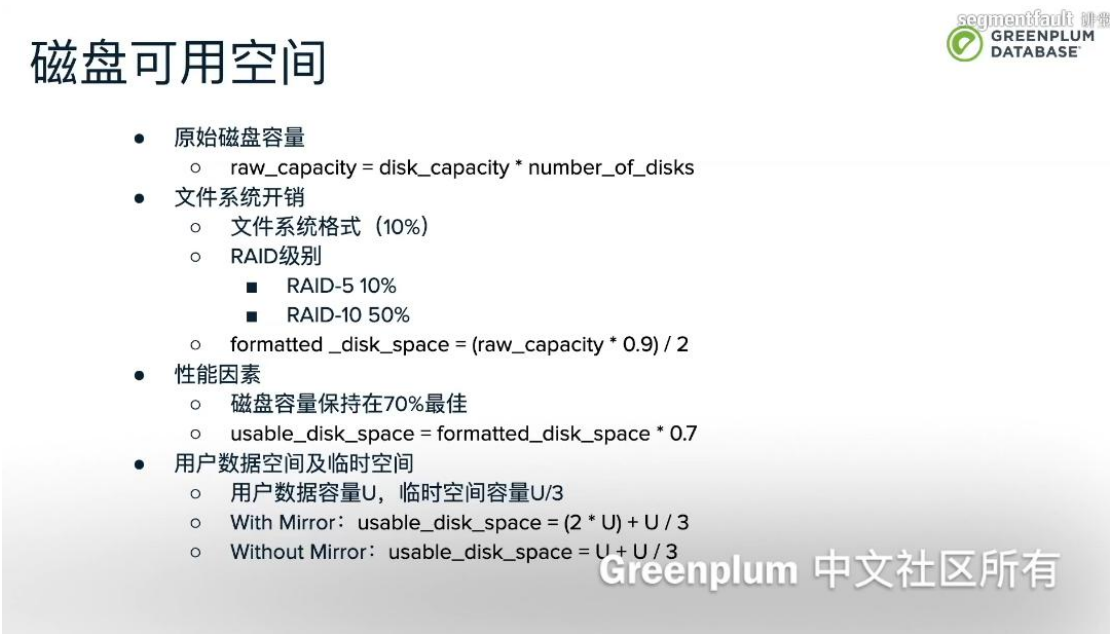
4 交换机配置

以下配置是两个交换机，配置的流量是均衡的。

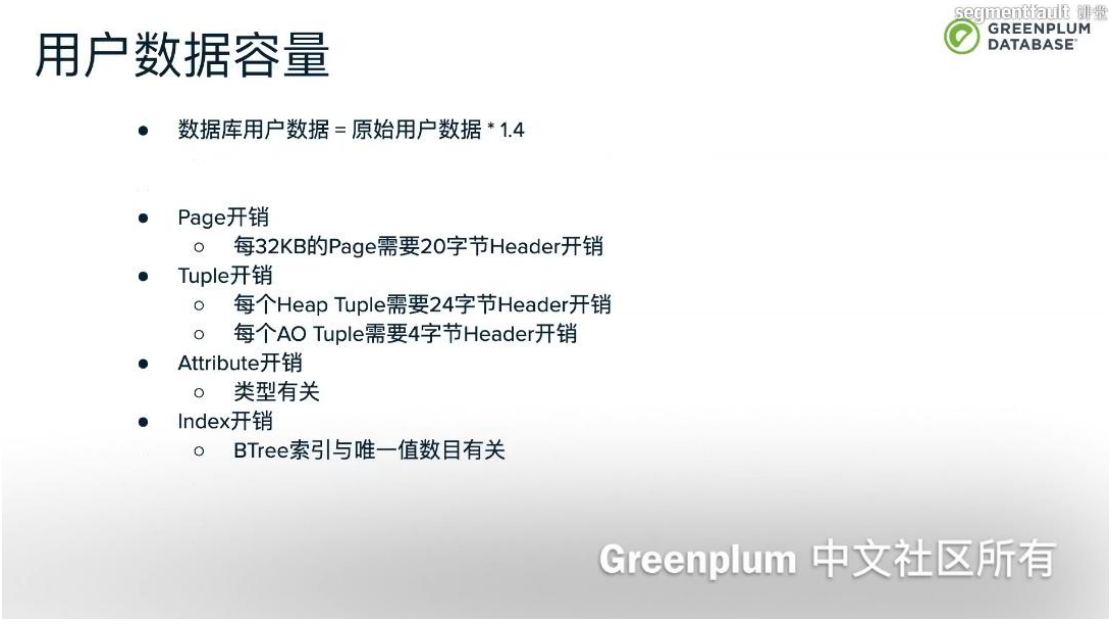


4 储存规划

4.1 磁盘可用空间



4.2 用户数据容量

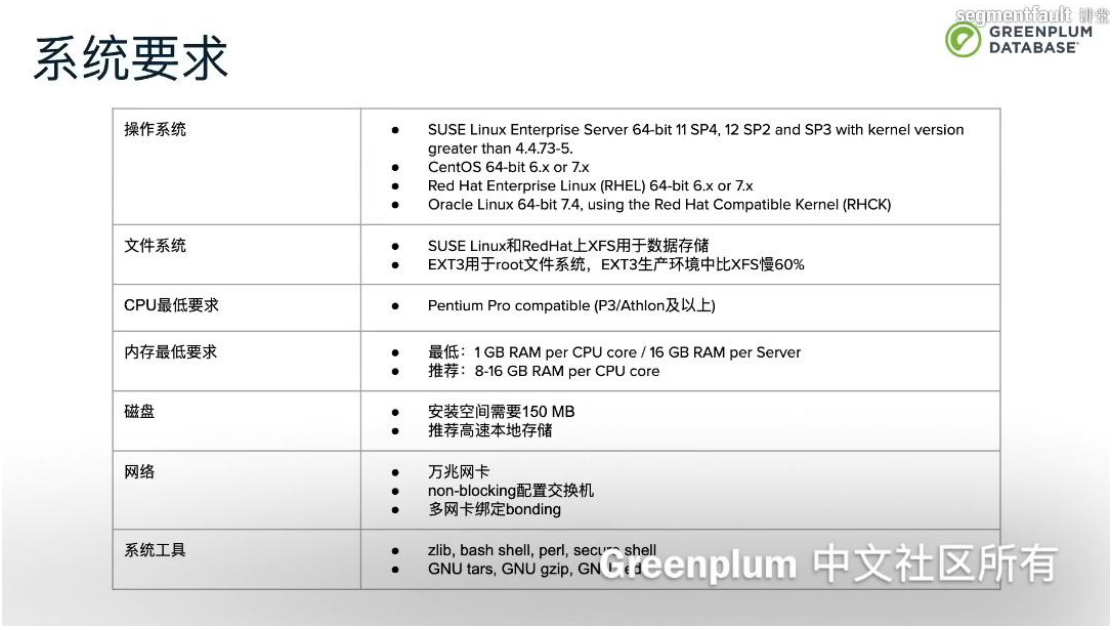


4.3 系统数据容量

5 集群的配置

5.1 最低系统要求

以下是系统的最低的配置



5.2 系统设置

5.3 操作系统参数设置

5.4 I/O 设置

I/O设置

- XFS挂载选项

```
rw,nodev,noatime,nobarrier,inode64
```

- 设置blockdev预读尺寸

```
# /sbin/blockdev --setra 16384 /dev/sdb
```

- 设置磁盘I/O调度器为deadline

```
# echo deadline > /sys/block/sbd/queue/scheduler
```


I/O设置

- 禁用Transparent Huge Page (THP)
  - RHEL 6.X配置/boot/grub/grub.conf

```
kernel /vmlinuz-2.6.18-274.3.1.el5 ro root=LABEL=/
elevator=deadline crashkernel=128M@16M  quiet console=tty1
console=ttyS1,115200 panic=30 transparent_hugepage=never
initrd /initrd-2.6.18-274.3.1.el5.img
```

  - RHEL 7.X运行grub2命令

```
# grubby --update-kernel=ALL --args="transparent_hugepage=never"
```

- /etc/systemd/logind.conf中禁用RemoveIPC

```
RemoveIPC=no
# service systemd-logind restart
```

5.5 其他设置

6 集群安装

6.1 安装方式

6.2 数据库目录

数据库目录

bin	postgres程序，数据库管理工具等
sbin	内部工具
lib	库文件，Extension等
include	头文件
share	内部共享文件
ext	内置程序，如Python
docs	命令行工具帮助文件，集群配置文件示例
greenplum_path.sh	环境变量文件

7 集群验证

7.1 硬件性能验证

硬件性能验证

- gpcheckperf
  - 网络性能

```
$ gpcheckperf -f hostfile_gpchecknet_ic1 -r N -d /tmp > subnet1.out
```

  - 磁盘IO
  - 内存带宽

```
$ gpcheckperf -f hostfile_gpcheckperf -r ds -D -d /data/primary -d /data/mirror
```

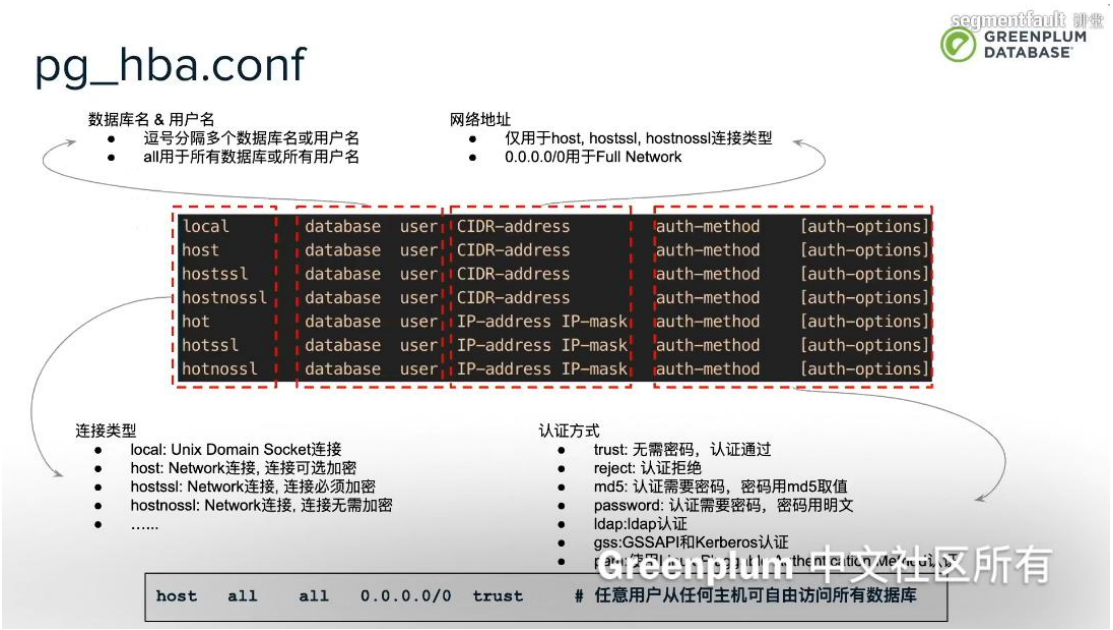
7.2 集群初始化

7.3 配置文件

7.4 数据目录

base	数据库数据文件
global	全局数据表文件，如pg_database
pg_log	数据库查询日志文件
pg_xlog	数据库WAL日志文件
pg_clog	事务提交信息日志文件
pg_tblspc	数据库所有tablespace的符号链接
postgresql.conf	数据库主配置文件
postmaster.opts	数据库postmaster进程启动命令
pg_hba.conf	客户端连接认证配置

7.5 pg_hba.conf 配置文件



pg_hba.conf

- 处理客户端连接认证
- 采用First Matching策略
- 默认客户端连接策略为Deny
- 策略更改后需要重启集群或者gpstop -u重新加载

7.6 安装完后的环境变量

环境变量

变量名	变量值示例	变量含义
GPHOME	/usr/local/greenplum-db	Greenplum安装主目录
MASTER_DATA_DIRECTORY	/data/master/gpseg-1	Master节点数据目录
PGHOST	mdw	Master节点主机地址
PGPORT	5432	Master数据库主进程端口号
PGDATABASE	postgres	默认数据库
PGUSER	gpadmin	默认数据库用户名
PGPASSWORD	******	用户对应密码
PGOPTIONS	'-c gp_session_role=utility'	数据库连接选项

8 可选组件安装

8.1 安装外部支持的语言

8.2 使用 gppkg 命令安装

8.2 故障诊断


常见故障
segmentfault 排雷
GREENPLUM DATABASE
集群初始化失败
检查$HOME/gpAdminLogs
系统参数设置不正确或没有重新加载生效
检查数据目录不存在，不为空，权限不正确
远程客户端无法连接数据库
检查pg_hba.conf设置
检查postgresql.conf设置
查询性能问题
使用gpcheckperf逐个排查网卡，磁盘，内存等硬件问题
Greenplum 中文社区所有