System Preparation and Verification

Pivotal® Greenplum Database

Pivotal

Agenda

- **Greenplum Software and Hardware Requirements**
- System Preparation Steps

Pivotal

Greenplum Hardware Considerations

The Greenplum software-only solution:

- Gives customers a choice of hardware platforms
- Is dependent on hardware solutions to optimize performance
- Performs well on certified reference architecture

The following statements are true for general hardware configuration:

- Segment servers should have identical hardware specifications
- Master server requires fast CPU and lots of RAM

Pivotal

Greenplum Database 4.3.x.x System Requirements

Operating System	• SuSE Linux Enterprise Server 64-bit 11 SP1, 11 SP2, 11 SP4 • CentOS 64-bit 5.0 or higher • RedHat Enterprise Linux 64-bit 5.0 or higher • Oracle Unbreakable Linux 64-bit 5.5
File System	XFS required for data storage on SUSE Linux and Red Hat (ext3 supported for root file system)
Minimum CPU	Pentium PRO compatible (P3/Athlon and above)
Memory	16 GB RAM per server (minimum) 64+ GB RAM per server (recommended)
Disk Requirements	• 150 MB per host for Greenplum installation • Approximately 300 MB per segment instance for meta data • Appropriate free space for data with disks at no more than 70% capacity • High-speed, local storage
Network Requirements	• Gigabit Ethernet within the array (10-gigabit network for support) • Dedicated, non-blocking switch

Pivotal

Estimating Storage

Total raw disk capacity required must take into account the following:

Type	Amount of Storage to Allocate
RAID parity/mirroring	Depends on RAID type chosen; i.e., 50% of storage for RAID 10
Greenplum Database segment mirrors	50% of storage
File system overhead	About 10% of storage
Used capacity	Less than 70% recommended
Raw data size and database storage overhead	Raw data may be 1.4 times larger on disk; depends on table types, indexes, compressions, etc.
System metadata	About 20 MB per segment and master instance
Write ahead log (WAL)	About 1088 MB per segment and master instance
Greenplum Database log files	About 10 MB per segment and master instance

Pivotal

Calculating Usable Disk Capacity

1. **Calculate raw capacity across all segments**

```
raw_capacity = disk_size * number_of_disks
```

2. **Calculate formatted disk space with appropriate overhead and RAID taken into account**

```
formatted_disk_space = (raw_capacity * .9) / 2
```

3. **Calculate usable disk space (less than 70 %)**

```
usable_disk_space = formatted_disk_space * 0.7
```

4. **Compare against user data (U) and work area (1/3U) with and without mirrors**

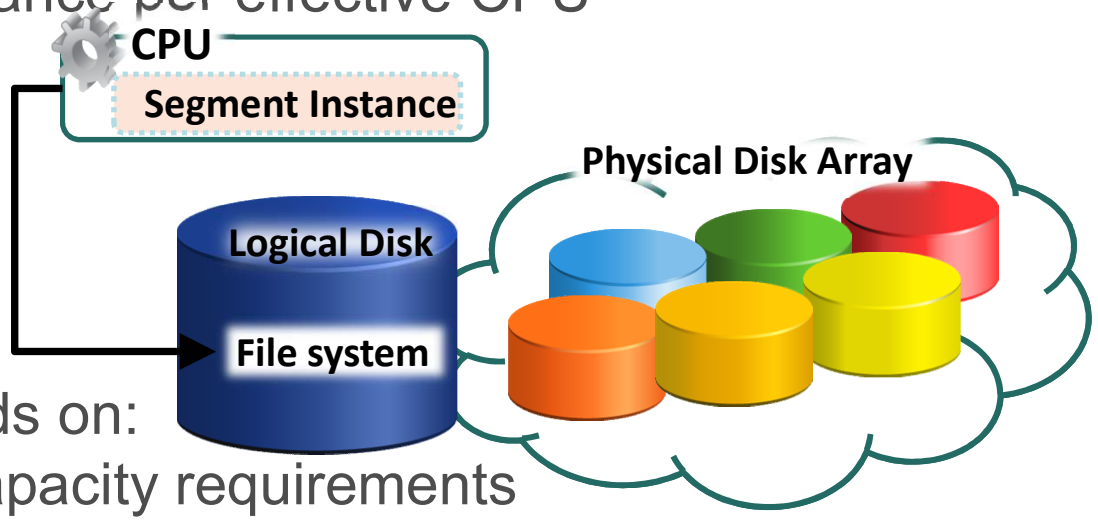
Without mirror: `U + U/3` With mirror: `2U + U/3`

Pivotal

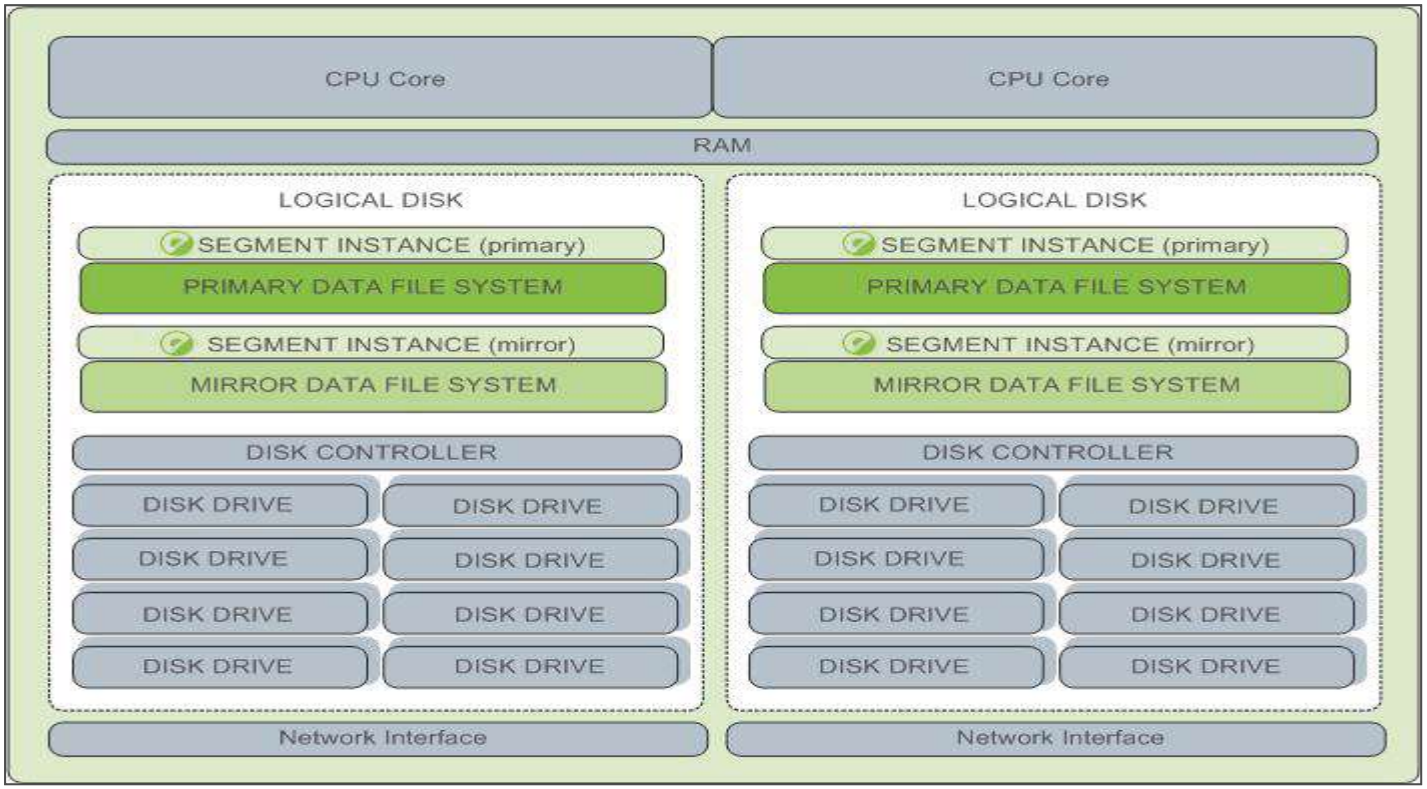
Segment Host – Disk Layout

An optimal disk layout includes:

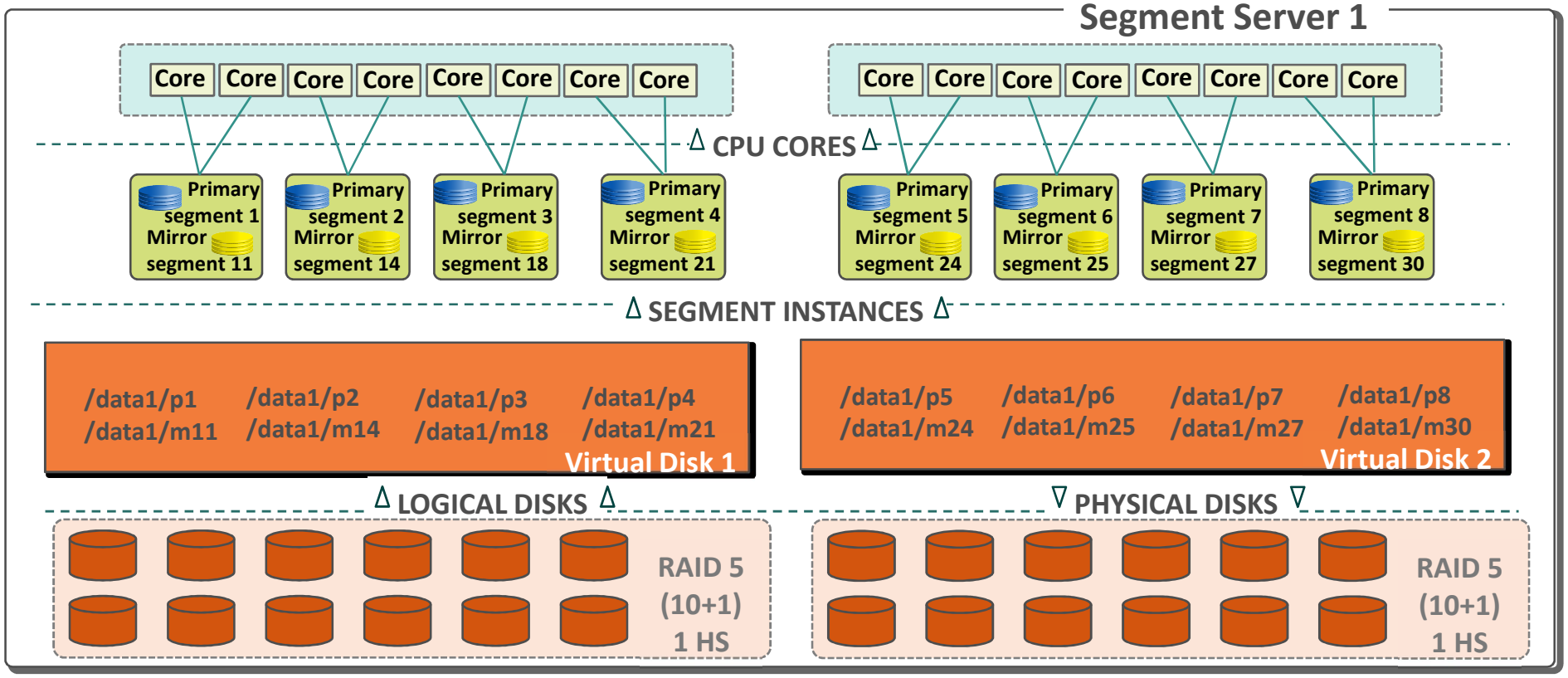
- One primary segment instance per effective CPU
- Primary segment mapped to a file system within a logical disk drive
- Logical drive uses groups of physical disks (RAID)
- RAID level chosen depends on:
  - Performance versus capacity requirements (RAID-10 or RAID-5)
  - Data protection and disk fault tolerance requirements



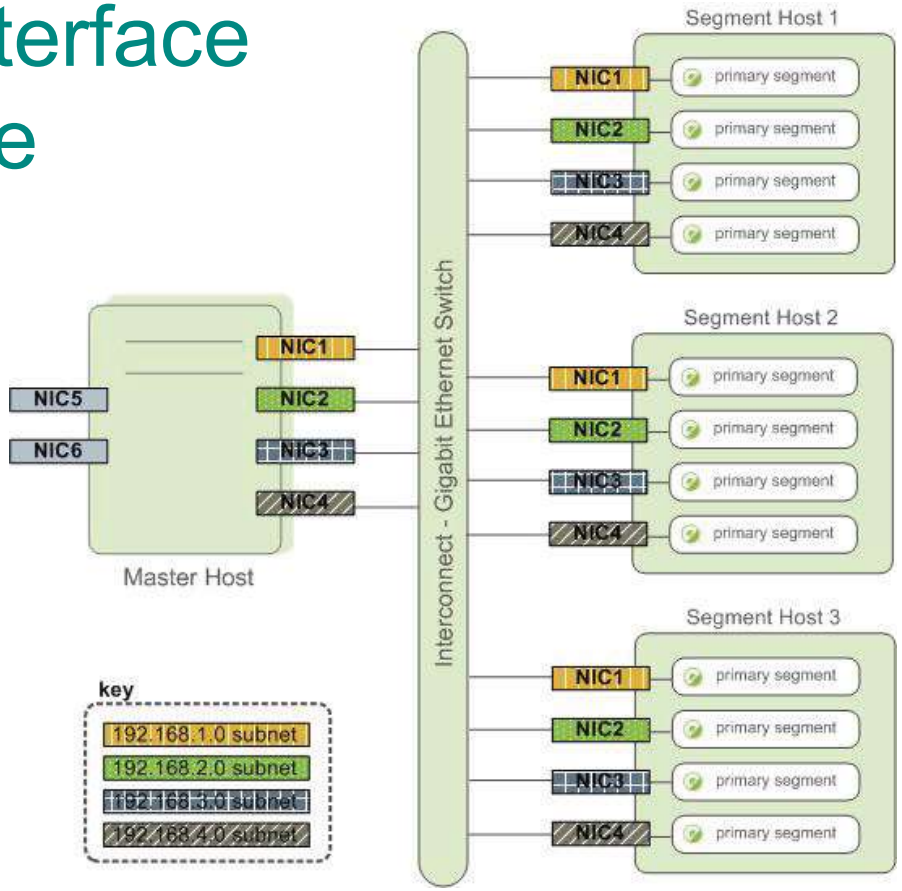
Segment Host Configuration



Anatomy of a Segment Server



Network Interface Architecture



Agenda

- Greenplum Software and Hardware Requirements
- **System Preparation Steps**

Configuring the System for Greenplum

To prepare the Greenplum environment, you must:

1. Verify the system meets the base system requirements
2. Tune the kernel for your operating system
3. Install the Greenplum binaries on the master and segments and create the Greenplum administrative user
4. Perform hardware verification tests

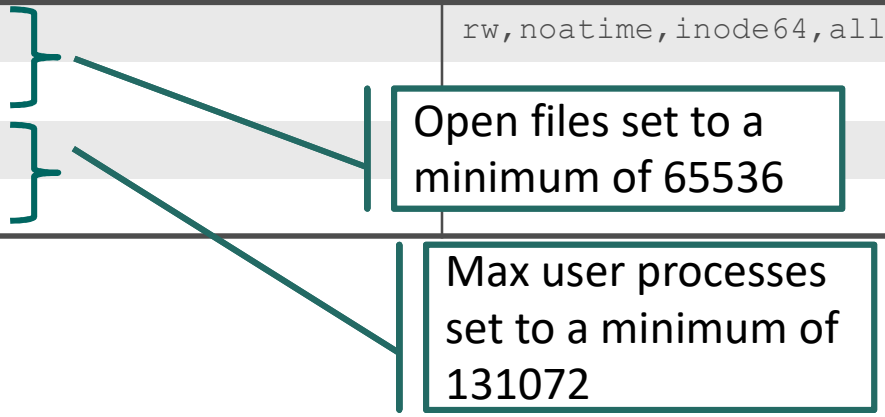
Linux Operating System Kernel Tuning

Shared Memory Kernel Parameters	Networking Kernel Parameters
kernel.shmmax = 500000000	net.ipv4.tcp_syncookies = 1
kernel.shmmni = 4096	net.ipv4.ip_forward = 0
kernel.shmall = 4000000000	net.ipv4.conf.default.accept_source_route = 0
kernel.sem = 250 512000 100 2048	net.ipv4.tcp_tw_recycle=1
kernel.sysrq = 1	net.ipv4.tcp_max_syn_backlog=4096
kernel.core_uses_pid = 1	net.ipv4.conf.all.arp_filter = 1
kernel.msgmnb = 65536	net.ipv4.ip_local_port_range = 1025 32000
kernel.msgmax = 65536	net.core.netdev_max_backlog=10000
kernel.msgmni = 2048	net.core.rmem_max = 2097152
vm.overcommit_memory=2	net.core.wmem_max = 2097152

Pivotal

Linux Operating System Kernel Tuning (Cont)

User Limits (Defined in `/etc/security/limits.conf`)	XFS Mount Options
`* soft nofile 65536`	`rw,noatime,inode64,allocsize=16m`
`* hard nofile 65536`	
`* soft nproc 131072`	
`* hard nproc 131072`	



Block Device Options	Value
I/O Scheduler	`deadline`
Block device read-ahead value	`16385`

Pivotal

Disk Device and OS Settings

1. Update mount options for XFS file systems

Set mount option to: `rw, noatime, inode64, allocsize=16m`

2. Set the I/O scheduler to `deadline` for all devices

Use the following command on each device:

`echo deadline > /sys/block/`*`device name`*`/queue/scheduler`

3. Change the read-ahead value for each block device to 16385

Use the following command on each block device:

`/sbin/blockdev --setra 16385 /dev/`*`block_device_name`*

4. Disable Transparent Huge Pages (THP). (ReadHat Linux 6.0 and higher)
   THP degrades Greenplum Database performance. Applicable on HUGE memory systems

Use the following command on to disable THP:

```
echo never > /sys/kernel/mm/redhat_transparent_hugepage/enabled
```

Pivotal

Greenplum Database Installation Overview

1. Install the Greenplum Database binaries on the master server

   a. Download the Greenplum Database binary
   b. Unzip and execute the installation program as `root`

2. Install the Greenplum Database binaries on the standby and segment servers

   a. Access the master server as root and source `/usr/local/greenplum-db/greenplum_path.sh`
   b. Verify that the /etc/hosts file on all systems have the correct host names
   c. Create an exchange key list file with the hostname of each segment interface, master interface and standby interface.
   d. Run the `gpseginstall` utility to install the Greenplum Database binaries on the standby and segment servers

3. Verify the installation was successful

   a. Log in as the `gpadmin` user and source `/usr/local/greenplum-db/greenplum_path.sh`
   b. Run a command on all hosts using `gpssh` and verify you are not prompted for a password

Pivotal

Greenplum Database Installation Overview (Cont)

4 Create the data storage directories as `root` on all servers

a. Create `/data/master` on the master and standby servers and change ownership of directory to `gpadmin`
b. Create an exchange key file with segment server host names only
c. Create `/data/primary` and `/data/mirror` on all segment servers using the exchange key file and change ownership of directories to `gpadmin`

5 Synchronize system clocks across all servers

a. Configure NTP on the master server so that it points to the data center's NTP time server.
b. Configure NTP on the standby server so that it points first to the master server and then to the NTP time server.
c. Configure NTP on the segment servers so that they initially contact the master server if available. If not, they should synchronize with the standby server.

Pivotal

Hardware Verification and Testing

Test the limits of the environment by:

- Establishing baseline disk I/O, CPU performance, and network transfer rates
- Stress testing hardware

Perform the following tests on system components:

- `gpcheckperf`:
  - Test disk input and output rates
  - Test memory bandwidth
  - Test network transfer rates
- `gpcheck`: Validate the OS settings
- `bonnie++`: Stress test for the file system (download from the bonnie++ website)

Wrap Up

During this lesson the following topics were covered:

- Greenplum software and hardware solutions and requirements
- Reference architecture for Greenplum
- Verification steps and tools to prepare a system for Greenplum


Pivotal
A NEW PLATFORM FOR A NEW ERA