information technology & management

viabiliitjit

INTRO TO OPEN SOURCE OPERATING SYSTEMS

ILLINOIS INSTITUTE OF TECHNOLOGY

ITMO456

Linux Server Deployment

Sean Hughes-Durkin

ITMO/IT-O 456 Fall 2017

Information Technology & Management Programs

School of Applied Technology

Objectives

At the end of this lesson students should be able to:

- Identify the types of hardware present in most server systems
- Describe the configuration of SCSI devices
- Explain the different levels of RAID and types of RAID configurations
- Configure the ZFS filesystem

Objectives

At the end of this lesson students should be able to:

- Install a Linux server distribution
- Troubleshoot the Linux server installation process
- Access an installed system using system rescue

Understanding Server Hardware

- Nearly all standard server hardware is supported by Linux
- If your system has specialized hardware
  - First verify with the hardware vendor that it has an adequate driver for your Linux distribution
- Rack-mount servers: a thin form factor used to house server hardware installed in a server rack
  - Storage Area Network (SAN) and uninterruptible power supply devices may also be contained in a server rack

Understanding Server Hardware

- Minimum height of a rackmount server is 1.75 inches
  - Called a 1U server
  - A 2U server is twice as high as a 1U server and often contains up to four CPUs and eight hard disks
- Rackmount servers rarely exceed 4U

Understanding Server Hardware

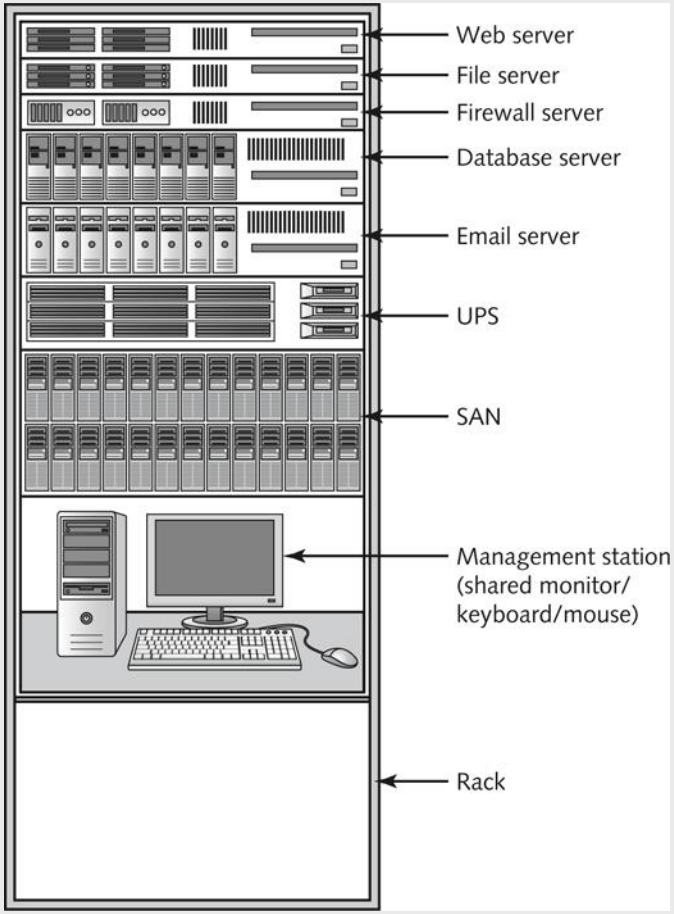


Figure 6-1 A sample server rack

Configuring Server Storage

- One of the most important configuration tasks involves configuring the permanent storage devices that will host the OS
- This involves:
  - Selecting the storage devices
  - Creating partitions and filesystems

SCSI Hard Drive Configuration

- SCSI (Small Computer System Interface): a way to connect multiple peripherals to the system
- Three types of SCSI disk configurations:
  - Parallel SCSI
  - Serial Attached SCSI
  - iSCSI

Parallel SCSI Configuration

- In Parallel SCSI, ribbon cables transmit information between hard disk and SCSI controller

Parallel SCSI Configuration

- Terminator
  - Device used to terminate an electrical conduction medium to absorb the transmitted signal
  - Prevents signals from bouncing back and forth on a cable
  - In many cases SCSI devices require a terminator at the end of the connection chain

Parallel SCSI Configuration

- SCSI ID
  - Uniquely identifies devices attached to a SCSI controller
  - Also called **target ID**
  - Identifies priority
  - Can have up to 15 devices on one controller

Parallel SCSI Configuration

SCSI Type	Speed (MB/s)	Bus Width (bits)	Connector	Number of Devices Supported
SCSI-1 (narrow, slow)	5	8	50-pin centronics or 50-pin LPT (Line Port Terminal) type	7
SCSI-2 (fast)	10	8	50-pin LPT type	7
SCSI-2 (wide)	20	16	68-pin LPT type	15
SCSI-3 (ultra)	40	16	68-pin LPT type or 80-pin SCA (Single Connector Adapter) type	15
SCSI-3 (ultra2 wide)	80	16		15
SCSI-3 (ultra3 wide)	160	16		15
SCSI-3 (ultra320)	320	16		15
SCSI-3 (ultra640)	640	16		15

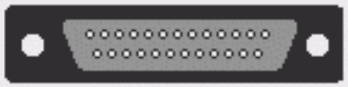
Table 6-1: Common SCSI standards

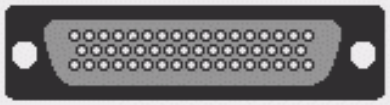
Parallel SCSI Configuration

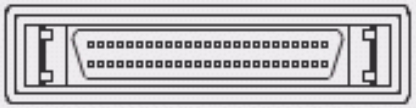
External Parallel SCSI Connectors

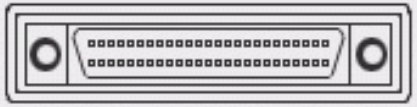
1

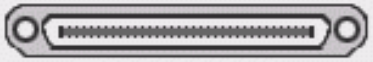
Centronics 50-Pin

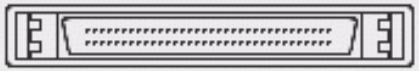
2

DB 25-Pin

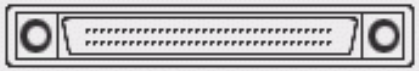
3

DB 50-Pin

4

High Density DB 50-Pin (Clip Type)

5

High Density DB 50-Pin (Screw Type)

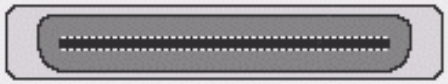
6

VHDCI 68-Pin

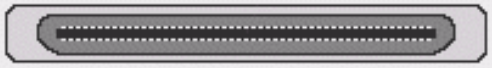
7

High Density DB 68-Pin (Clip Type)

8

High Density DB 68-Pin (Screw Type)

9

High Density Centronics 50-Pin

10

High Density Centronics 60-Pin

11

High Density Centronics 68-Pin

12

HDI30 (Apple/Mac)

Parallel SCSI Configuration

- Configuration steps:
  - Verify all SCSI components support the same technology
  - Ensure that components are connected properly
  - Make sure system recognizes hard drives at startup

Serial-Attached SCSI (SAS) Devices

- Uses same command set as “traditional” SCSI
- SAS bus operates point-to-point while SCSI bus is multidrop
- Has no termination issues
  - Does not require terminators
- Supports up to 65,535 devices (via expander cards – lots of them)
- Higher transfer speeds (3 or 6 Gbit/s)
- SAS controllers may support connecting to SATA devices

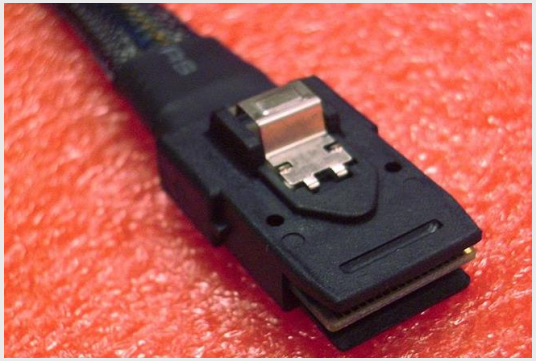
Serial Attached SCSI Configuration

- Serial attached hard disks connected to SCSI controller via serial cables with small serial connectors
- Configuration steps:
  - Connect hard disk to controller via correct serial cable
  - Ensure hard disk is detected by system or SCSI BIOS
  - The rest is performed automatically by controller

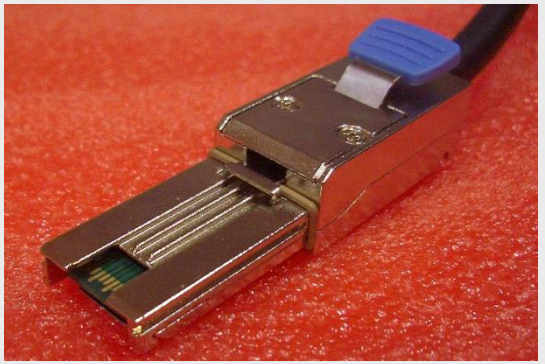
SAS Drive Connectors

SF8484 32-pin hi-density internal connector

SFF 8470 Infiniband 32-pin connector; also used as an internal connector



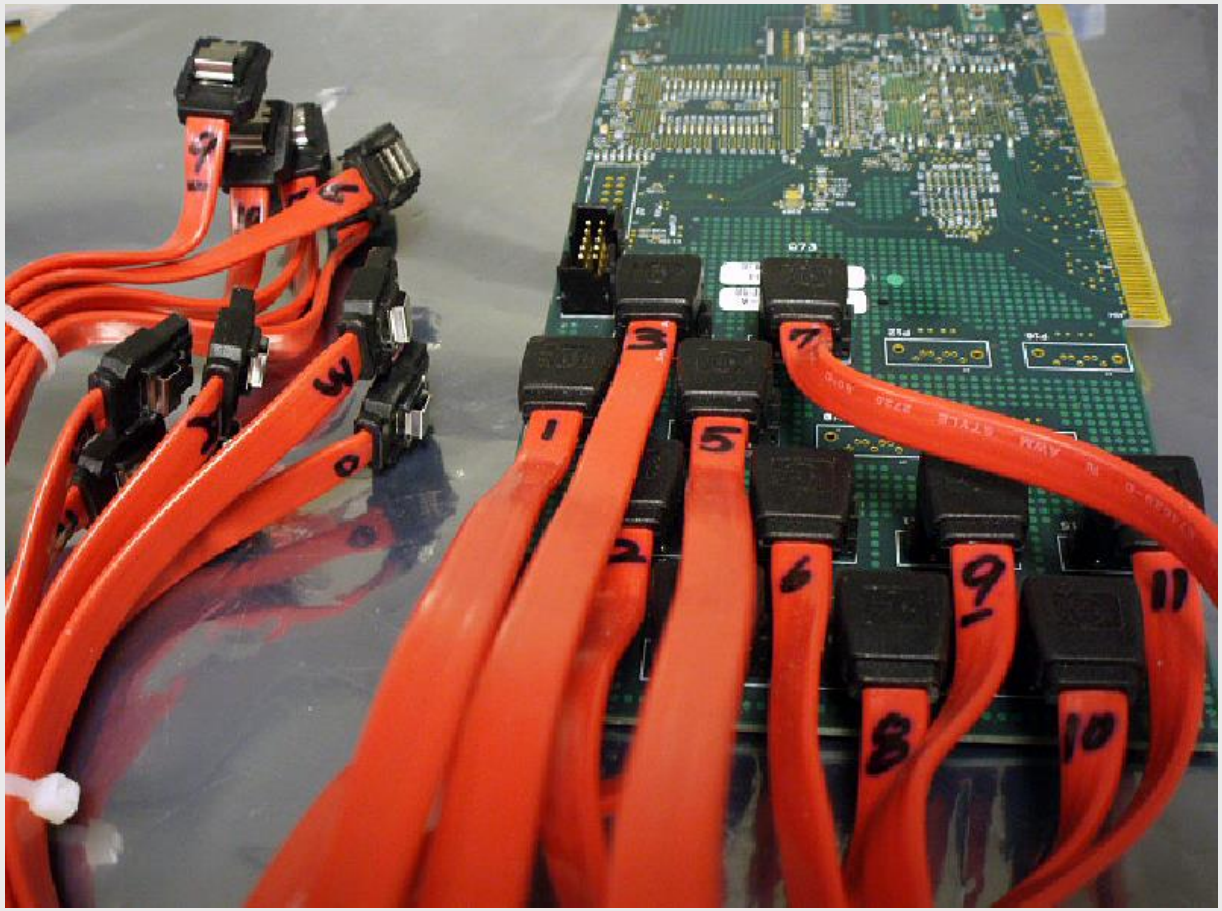
SFF 8087 36-pin internal mini-SAS with future 10 Gbit/s support



SFF 8088 26-pin external mini-SAS with future 10 Gbit/s support

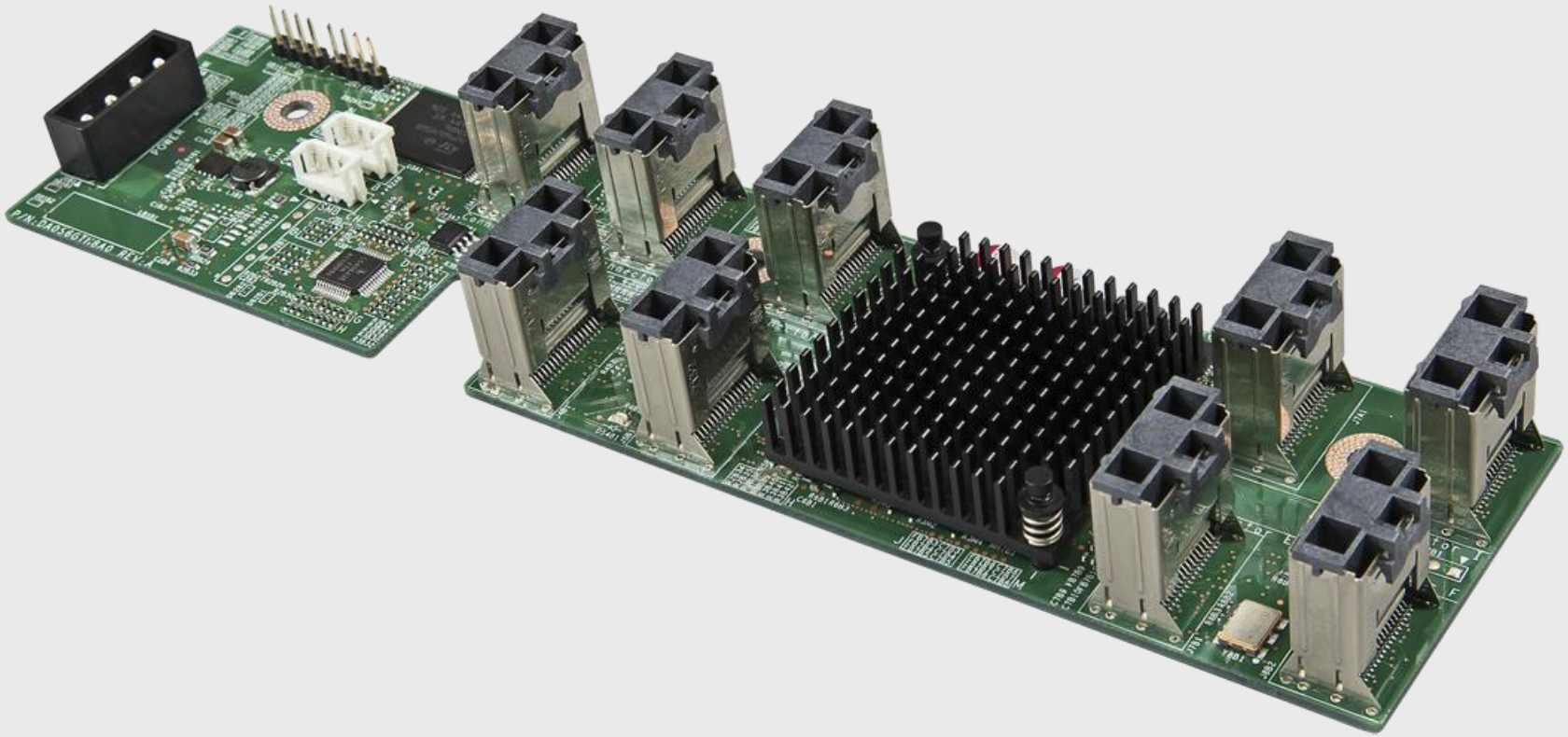
Serial-Attached SCSI connectors

SAS Drive Connectors



Serial-Attached SCSI connectors

SAS Drive Connectors



Serial-Attached SCSI expander card

iSCSI Configuration

- Internet SCSI (iSCSI): uses network cables to transfer data to/from remote hard disks
  - iSCSI initiator: computer connected to remote hard disk
    - Can be software or hardware
  - iSCSI target: remote hard disk
    - Contained within remote network attached device

Serial ATA Devices

- Successor to legacy Advanced Technology Attachment standard (aka IDE)
  - IDE retroactively renamed Parallel ATA (PATA)
- 1st Generation: 1.5 Gb SATA
  - Data transfer rate 1.2 Gb/s or 150 MB/s
  - Allows use of longer drive cables
    - Up to 1 meter long

Serial ATA Devices

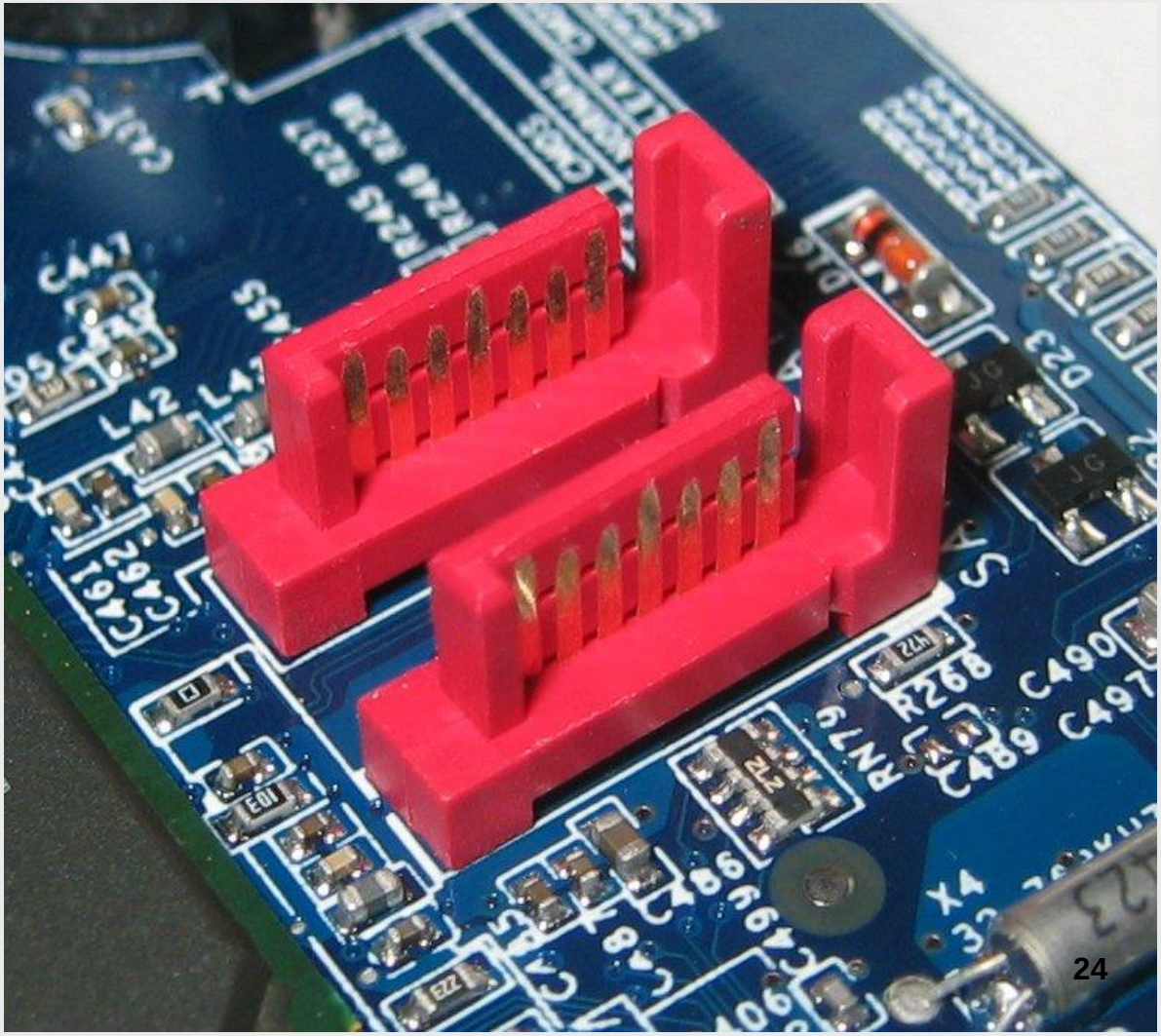
- Rev 2: 3 Gbit/sec SATA
  - Actual data transfer rate 2.4 Gb/s or 300 MB/s
- Each device has a dedicated cable and dedicated bandwidth
- Features include hot-swapping and native command queueing
- 7 wire/7pin data cable
- 15 pin power connector

Serial ATA Devices

- Rev 3: 6 Gbit/sec SATA
- Can use existing SATA cables and connectors
  - Some OEMs will upgrade connectors for higher speeds
  - Adds a small Low Insertion Force (LIF) connector for 1.8-inch devices and also a connector for 7 mm optical disc drives
- Backwards compatible with 3 Gbit/sec SATA

Serial ATA Devices

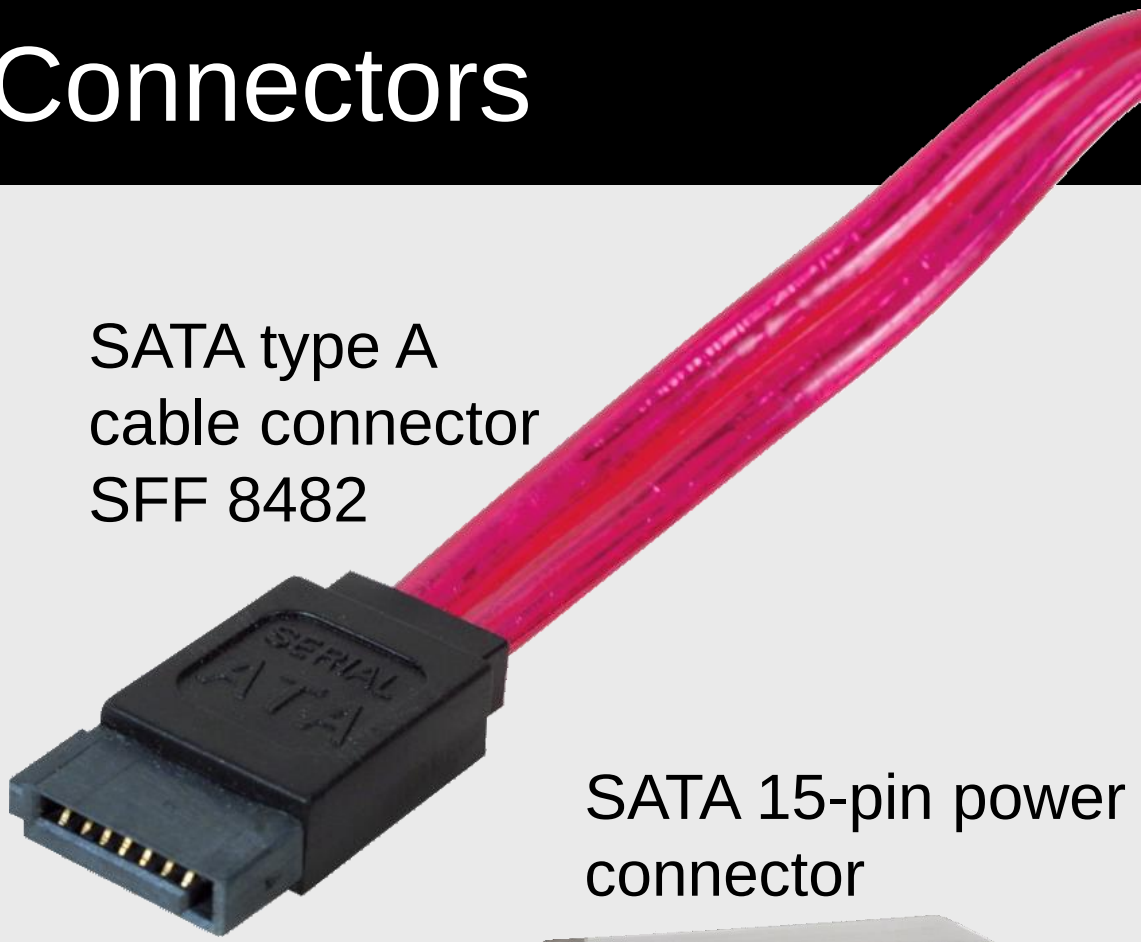
Serial ATA connectors on a mainboard



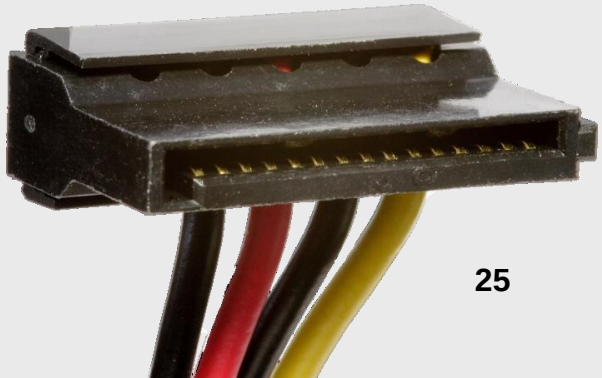
Serial ATA Connectors

Serial ATA 6-pin power connector (mostly for notebook PCs)

SATA type A cable connector SFF 8482



SATA 15-pin power connector



Note: all types also used for SAS

SATA/SAS Compatibility

- There is a SAS backplane standard allowing connection of SATA devices
  - Uses SATA-style connectors
- SAS drives *cannot* be connected to a SATA controller (even if they have compatible connectors)

RAID Configuration

- Fault tolerance
  - The measure of downtime a device exhibits in the event of a failure
- Fault tolerant
  - Device exhibiting minimum downtime after failure

RAID Configuration

- Redundant Array of Inexpensive Disks (RAID)
  - Typical implementation of fault tolerant disk
  - Can be used to speed up access to hard disk
  - Combines several HDDs' storage space
  - Seven raid configurations
  - Hardware-, software-, or firmware-based
    - Different configuration process for each type

RAID Configuration

- Spanning
  - Two HDDs seen as one volume
  - RAID level 0, not fault tolerant
  - Good when need large amount of space in single volume
- Disk striping
  - Write separate information to different HDDs
  - RAID level 0, not fault tolerant
  - Decreases read/write time

RAID Configuration

- Disk mirroring
  - Two identical hard disks
  - RAID level 1, fault tolerant
  - Personal note: Never RAID 1 two disk from the same disk lot
- Disk Striping with Parity
  - Write separate information to HDDs & maintain parity information

RAID Configuration

- RAID level 5, fault tolerant
  - Parity bits indicate what data is where
    - Intermixed on HDDs that contain the data
    - Used to re-generate data when HDD fails
  - Requires three HDDs minimum
  - Most common RAID configuration
  - Improves on RAID levels 3 and 4
  - If any of the HDDs fail, the information can be recovered from the other HDDs

RAID Configuration

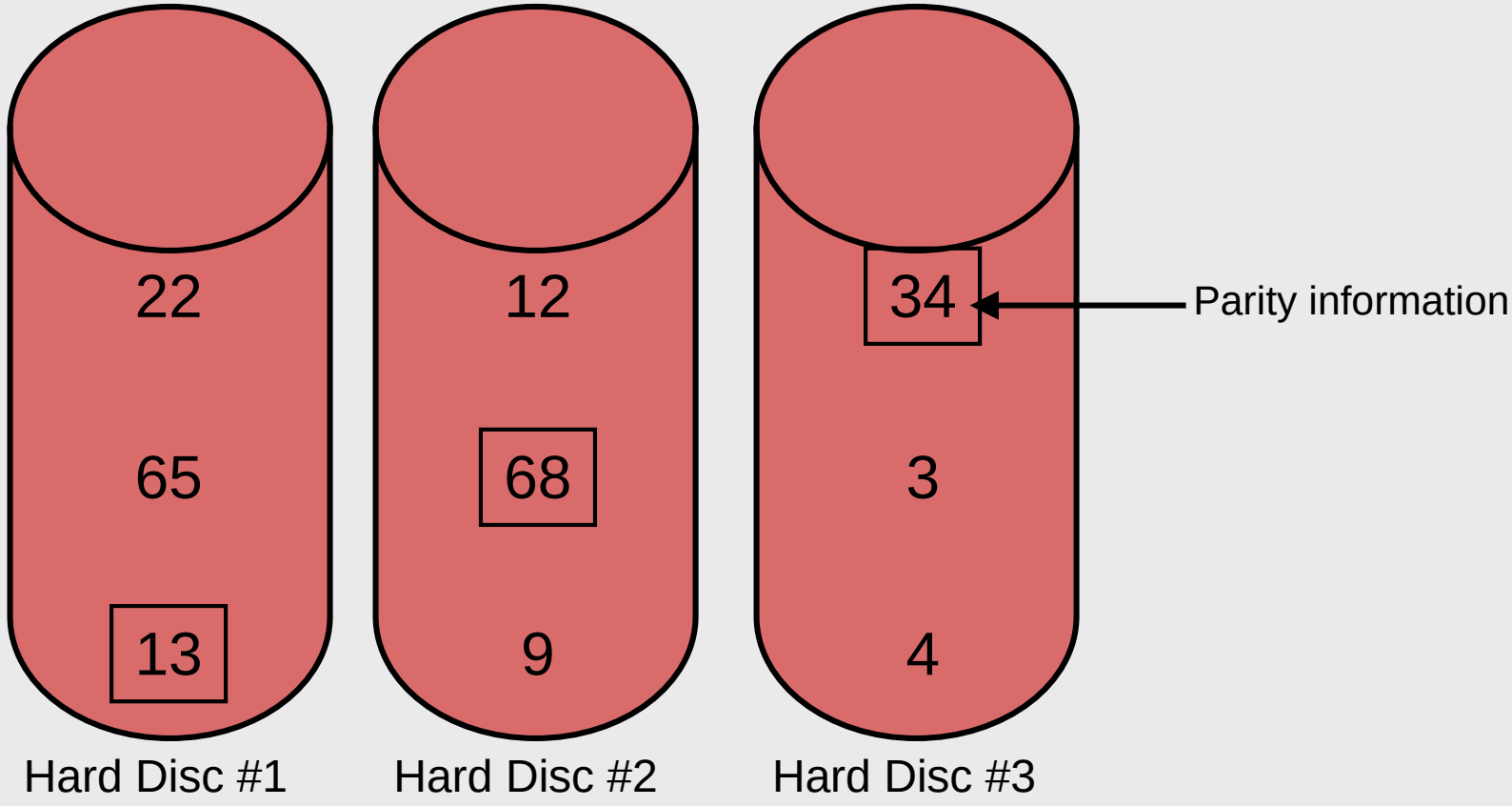


Figure 6-6: Organization of data on RAID level 5

RAID Configuration

- RAID level 6
  - Uses two sets of parity bits for added fault tolerance
  - Allows 2 HDDs to fail simultaneously while remaining fault tolerant
- RAID 10
  - Combination of 0 and 1 (striping and mirroring)

SATA and RAID

- SATA RAID drive controllers often use software/BIOS-dependent RAID instead of hardware RAID
  - Often called “fakeraid”
  - Normally require proprietary drivers
- If this is true, probably better to mount each drive as a block device and use Linux software RAID
  - Likely to be faster & more reliable

ZFS Configuration

- Zettabyte File System (ZFS) - a high-performance filesystem and volume management software
  - Designed for large-scale Linux systems that need to store data on multiple disks, SANS, and remote systems
  - Detects and repairs data errors automatically as data is read and written

ZFS Configuration

- ZFS protects against problems commonly seen on large systems that write large amounts of data to a non-ZFS filesystem, including:
  - Silent data corruption
  - Bit rot
  - Disk firmware bugs
  - Phantom writes
  - Misdirected writes
  - Driver and kernel buffer errors
  - Accidental driver overwrites

ZFS Configuration

- ZFS:
  - Supports new storage technologies such as PCIe SSDs and battery-backed RAM disk devices
  - Caches frequently accessed information in RAM and SSDs to provide ultra-fast performance
  - Primarily used on large Linux systems
    - Also used within small and medium-sized organizations that require flexible volume management for data that is easy to configure
  - Has similar features to RAID

ZFS Configuration

- ZFS pools: groups of physical disks that ZFS can manage
- ZFS volumes: ZFS-managed filesystems that are created from ZFS pools
- Use the **`zpool`** command to create a ZFS pool
  - Use with the **`list`** option to view details for your ZFS volume
  - Use with the **`status`** option to view the status of the disks that are part of a ZFS pool

ZFS Configuration

- Use the **`zfs`** command to manage specific features of the ZFS filesystem stored within ZFS volumes
  - Allows you to set directory size quotas and file- and directory-specific features and performance options
- On Linux systems, ZFS volumes are mounted by the ZFS system
  - Not via entries within the **`/etc/fstab`** file by default
  - Must modify
    - **/etc/init/zpool-import.conf**
    - **/etc/rc.local**

Installing a Linux Server Distribution

- Linux server distributions do not have a GUI environment installed by default
  - Most distributions ship with a set of packages that are commonly used on Linux servers, such as:
    - Web server, database server, file server, and email server software
- Linux server distribution installation differs from a standard or live Linux installation
  - No GUI environment is loaded and installation program often prompts for additional information

Installing a Linux Server Distribution

- You will need to know the following:
  - The host name and IP configuration of the server
  - Whether to allow for automatic updating
  - Package selection
  - Server service configuration
  - Boot loader configuration
    - GRUB (GRand Unified Boot loader) is the standard boot loader used on Linux systems

Dealing with Problems during Installation

- Installation may end abnormally with a "fatal signal 11" error message displayed on the screen
  - Indicates a segmentation fault, in which a program access a certain area of RAM that was not assigned
  - First check the RAM for errors by running the memtest86 utility
    - If no RAM errors are found, check with the system hardware manufacturer to determine if an updated driver is available for your Linux distribution

Dealing with Problems during Installation

- Problems may arise if the installation program failed to detect the computer's hardware properly
  - Check the installation log files after installation and verify system settings to ensure all hardware was detected with correct values
- Installation log files are created by the installation program
  - Record events that occurred during the installation process, including errors
- Fedora stores these files in /var/log/anaconda

Dealing with Problems during Installation

- To verify hardware settings
  - Examine the content of the /proc directory or boot-up log files
  - The /proc directory lists system information made available by the Linux kernel (readable only by the root user)
- To view information that Linux has detected regarding a computer's CPU
  - View the contents of the cpuinfo file in the /proc directory

Dealing with Problems during Installation

- To ensure Linux has detected the correct amount of RAM
  - View the contents of the /proc/meminfo file
- To see a list of modules currently inserted into the Linux kernel, view the /proc/modules file
  - If a hardware device is not working, consult the hardware vendor's Web site to find the name of the kernel module
  - If the module is not present in the modules file, the driver for that hardware is not loaded into the kernel

Dealing with Problems during Installation

- To verify hardware settings, examine contents of **`/proc`** directory or boot-up log files
  - For CPU info, view **`/proc/cpuinfo`**
  - For RAM info, view **`/proc/meminfo`**
  - For list of modules, view **`/proc/modules`**
- To view hardware detected at boot time, use **`dmesg`** command

Common Files in `/proc`

Filename	Contents
cmdline	Current location of the Linux kernel
cpuinfo	Information regarding the processors in the computer
devices	List of the character and block devices that are currently in use by the Linux kernel
execdomains	List of execution domains for processes on the system; execution domains allow a process to execute in a specific manner
fb	List of framebuffer devices in use on the Linux system; typically, these include video adapter card devices
filesystems	List of filesystems supported by the Linux kernel
interrupts	List of IRQs in use on the system
iomem	List of memory addresses currently used
ioports	List of memory address ranges reserved for device use
kcore	A representation of the physical memory inside the computer; this file should not be viewed
kmsg	Temporary storage location for messages from the kernel
loadavg	Statistics on the performance of the processor
locks	List of files currently locked by the kernel
mdstat	Configuration of multiple-disk RAID hardware
meminfo	Information regarding physical and virtual memory on the Linux system
misc	List of miscellaneous devices (major number=10)
modules	List of currently loaded modules in the Linux kernel
mounts	List of currently mounted filesystems
partitions	Information regarding partition tables loaded in memory on the system
swaps	Information on virtual memory utilization
version	Version information for the Linux kernel and libraries

Common Files in /proc

```
[root@localhost /]# cat /proc/cpuinfo
processor       : 0
vendor_id       : GenuineIntel
cpu family      : 6
model           : 61
model name      : Intel(R) Core(TM) i7-5600U CPU @ 2.60GHz
stepping        : 4
microcode       : 0x1d
cpu MHz         : 2594.030
cache size      : 4096 KB
physical id     : 0
siblings        : 4
core id         : 0
cpu cores       : 4
apicid          : 0
initial apicid  : 0
fpu             : yes
fpu_exception   : yes
cpuid level     : 20
wp              : yes
flags           : fpu vme de pse tsc msr pae mce cx8 apic sep mtrr pge mca cmov pat pse
36 clflush dts mmx fxsr sse sse2 ss ht syscall nx pdpe1gb rdtscp lm constant_tsc arch_p
erfmon pebs bts nopl xtopology tsc_reliable nonstop_tsc aperfmperf eagerfpu pni pclmulq
dq ssse3 fma cx16 pcid sse4_1 sse4_2 x2apic movbe popcnt tsc_deadline_timer aes xsave a
vx f16c rdrand hypervisor lahf_lm abm 3dnowprefetch epb fsgsbase tsc_adjust bmi1 hle av
x2 smep bmi2 invpcid rtm rdseed adx smap xsaveopt dtherm ida arat pln pts
bugs            :
bogomips        : 5188.06
```

Common Files in `/proc`

```
[root@localhost /]# cat /proc/meminfo
MemTotal:        4028808 kB
MemFree:         1407436 kB
MemAvailable:    2989524 kB
Buffers:          248512 kB
Cached:          1509692 kB
SwapCached:            0 kB
Active:           904784 kB
Inactive:        1340684 kB
Active(anon):     488340 kB
Inactive(anon):     1400 kB
Active(file):     416444 kB
Inactive(file):  1339284 kB
Unevictable:          16 kB
Mlocked:              16 kB
SwapTotal:       2097148 kB
SwapFree:        2097148 kB
Dirty:                88 kB
Writeback:             0 kB
AnonPages:        487440 kB
Mapped:           205644 kB
Shmem:              2468 kB
Slab:             125776 kB
SReclaimable:      62904 kB
SUnreclaim:        62872 kB
KernelStack:        8384 kB
PageTables:        29312 kB
NFS_Unstable:          0 kB
```

Common Files in /proc

```
[root@localhost /]# cat /proc/modules
rfcomm 69632 2 - Live 0xffffffffa0562000
fuse 98304 3 - Live 0xffffffffa0549000
nf_conntrack_netbios_ns 16384 0 - Live 0xffffffffa0544000
nf_conntrack_broadcast 16384 1 nf_conntrack_netbios_ns, Live 0xffffffffa053f000
ip6t_rpfilter 16384 1 - Live 0xffffffffa053a000
ip6t_REJECT 16384 2 - Live 0xffffffffa0535000
nf_reject_ipv6 16384 1 ip6t_REJECT, Live 0xffffffffa0530000
xt_conntrack 16384 21 - Live 0xffffffffa051c000
ip_set 36864 0 - Live 0xffffffffa0526000
nfnetlink 16384 1 ip_set, Live 0xffffffffa0517000
ebtable_broute 16384 1 - Live 0xffffffffa0521000
bridge 126976 1 ebtable_broute, Live 0xffffffffa04f7000
ebtable_nat 16384 1 - Live 0xffffffffa04f2000
ip6table_mangle 16384 1 - Live 0xffffffffa04ed000
ip6table_raw 16384 1 - Live 0xffffffffa04e8000
ip6table_security 16384 1 - Live 0xffffffffa04e3000
ip6table_nat 16384 1 - Live 0xffffffffa04de000
nf_conntrack_ipv6 20480 12 - Live 0xffffffffa04d4000
nf_defrag_ipv6 36864 1 nf_conntrack_ipv6, Live 0xffffffffa04c6000
nf_nat_ipv6 16384 1 ip6table_nat, Live 0xffffffffa04c1000
iptable_mangle 16384 1 - Live 0xffffffffa04bc000
iptable_raw 16384 1 - Live 0xffffffffa04b7000
iptable_security 16384 1 - Live 0xffffffffa04b2000
iptable_nat 16384 1 - Live 0xffffffffa04ad000
nf_conntrack_ipv4 16384 11 - Live 0xffffffffa04a4000
nf_defrag_ipv4 16384 1 nf_conntrack_ipv4, Live 0xffffffffa0498000
nf_nat_ipv4 16384 1 iptable_nat, Live 0xffffffffa0493000
```

Dealing with Problems during Installation

- To view the system processes that started successfully or unsuccessfully during boot
  - View the contents of the **/var/log/boot.log** file
- For more detailed information regarding system processes during startup
  - View the contents of **/var/log/messages** or **/var/log/syslog** files

Dealing with Problems during Installation

- In Fedora 20, use the **`journalctl -b`** command to view the same information that would normally be stored within **`/var/log/boot.log`**, **`/var/log/messages`**, or **`/var/log/syslog`** files
  - The **`journalctl`** command displays entries for each boot process, therefore, it is important to narrow down the time frame

System Rescue

- System Rescue: A small bootable Linux kernel and virtual filesystem used to fix problems
  - Used to fix:
    - The boot loader
    - Filesystems and partitions
    - The configuration file
    - Drivers
- Can select Rescue installed system at Fedora installation welcome screen
  - Many options regarding how to rescue

System Rescue

- Can configure the Linux kernel loaded by the live OS to use the **/** (root) filesystem on the hard disk
  - Mount the **/** (root) filesystem on your local disk under an empty directory in the live OS
    - Also boot and proc filesystems
  - Type **`chroot /mnt`** at the command prompt
  - The **`chroot`** command will change the root of the live OS to the **`/mnt`** directory
  - You will have root user access to any commands on your **/** (root) filesystem

System Rescue

```
[liveuser@localhost ~]$ su -
[root@localhost ~]# mount /dev/mapper/fedora-root /mnt
[root@localhost ~]# mount /dev/sda1 /mnt/boot/
[root@localhost ~]# mount -t proc proc /mnt/proc/
[root@localhost ~]# chroot /mnt
[root@localhost /]#
```

Figure 6-13: Using the System Rescue BASH shell

Summary

- Linux servers typically use far more hardware than desktops and is usually installed in a rack using a rackmount form factor
- Parallel SCSI HDDs are uniquely identified by a SCSI ID and attach to a controller via a terminated cable
- SAS SCSI HDDs transfer information to hard disks using a serial cable

Summary

- iSCSI is a SAN technology used to transfer information from iSCSI initiators to iSCSI targets across a TCP/IP network
- RAID is used in Linux servers to combine several hard disks into one for speed or fault tolerance
- RAID can be implemented by software, hardware, or the system BIOS
- Different levels of RAID determine how disks are combined and written to

Summary

- ZFS is a high-performance, fault tolerant filesystem that is commonly installed on Linux servers
- Linux server distributions do not contain a GUI environment and are often administered remotely
- Unsupported hardware, defective RAM, overclocked CPUs, and improper RAM settings can cause a Linux installation to fail
- The **`/proc`** directory contains information regarding detected hardware and is useful when verifying whether an installation was successful

Summary

- You can use the bootable OS found on standard and live Linux installation media to access and repair a damaged Linux installation

The End…

- Questions?