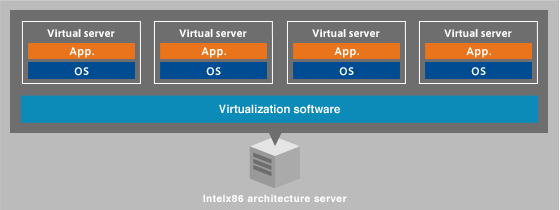
What is server virtualization?

Server virtualization is a technology for partitioning one physical server into multiple virtual servers. Each of these virtual servers can run its own operating system and applications, and perform as if it is an individual server. This makes it possible, for example, to complete development using various operating systems on one physical server or to consolidate servers used by multiple business divisions.



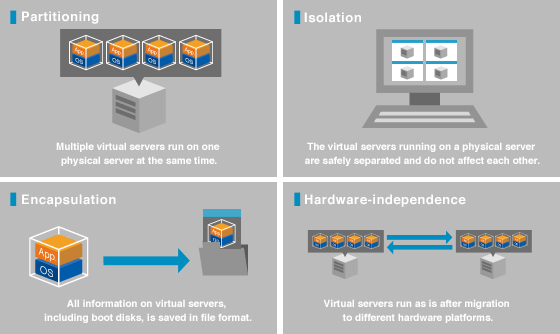
Among the various virtualization methods available, NEC primarily focuses on virtualization software solutions. Because the virtualization software, or hypervisor, used by NEC runs directly on bare hardware (physical servers), our virtualized environments have little overhead. NEC’s proven, reliable solutions are built upon years of experience with virtualization.

\* In addition to server virtualization solutions, NEC provides virtual desktop solutions.

[Find out more about virtual desktops solutions](http://www.nec.com/en/global/solutions/vpcc/virtualdesktop/index.html)

Server virtualization features

All the virtualization software provided by NEC has the following features.



Primary advantages of server virtualization

Reduce number of servers

**Partitioning and isolation**, the characteristics of server virtualization, enable simple and safe server consolidation.  
  
Through consolidating, the number of physical servers can be greatly reduced. This alone brings benefits such as reduced floor space, power consumption and air conditioning costs. However, it is essential to note that even though the number of physical servers is greatly reduced, the number of virtual servers to be managed does not change. Therefore, when virtualizing servers, installation of operation management tools for efficient server management is recommended.

Reduce TCO

Server consolidation with virtualization reduces costs of hardware, maintenance, power, and air conditioning. In addition, it lowers the Total Cost of Ownership (TCO) by increasing the efficiency of server resources and operational changes, as well as virtualization-specific features. As a result of today’s improved server CPU performance, a few servers have high resource-usage rates but most are often underutilized. Virtualization can eliminate such ineffective use of CPU resources, plus optimize resources throughout the server environment. Furthermore, because servers managed by each business division's staff can be centrally managed by a single administrator, operation management costs can be greatly reduced.

Improve availability and business continuity

One beneficial feature of virtualized servers not available in physical server environments is live migration. With live migration, virtual servers can be migrated to another physical server for tasks such as performing maintenance on the physical servers without shutting them down. Thus there is no impact on the end user. Another great advantage of virtualization technology is that its **encapsulation** and **hardware-independence** features enhance availability and business continuity.

Increase efficiency for development and test environments

At system development sites, servers are often used inefficiently. When different physical servers are used by each business division's development team, the number of servers can easily increase. Conversely, when physical servers are shared by teams, reconfiguring development and test environments can be time and labor consuming.  
  
Such issues can be resolved by using server virtualization to simultaneously run various operating system environments on one physical server, thereby enabling concurrent development and test of multiple environments. In addition, because development and test environments can be encapsulated and saved, reconfiguration is extremely simple.

General notes on server virtualization

Improve availability of your virtual environment

The impact of a failure in a physical server can increase after virtualizing and consolidating servers. To prevent such problems, NEC provides ExpressCluster and other products for improving the availability of virtual servers.

How To Consolidate Servers With Virtualization

**Using Microsoft's Virtual Server 2005 R2 Standard Edition lets system builders create virtual servers that better use high-power hardware, save customers money, and run multiple OSes.**

As time goes by, the relative cost of available CPU, RAM, and storage capacity continues to drop. As a result, many PC and server users have more computing horsepower than they can fully utilize. For system builders, this presents an opportunity to help customers make more complete use of this idle hardware. One solution that can help is virtualization technology.

By consolidating servers with virtualization, system builders can help their customers deploy applications more efficiently, using less hardware. System builders can set up a Windows server and a Linux server on the same physical machine, for instance, and allocate resources between them as necessary. In this way, virtualized servers can save customers a lot of money, and system builders can take credit for streamlining their customers' IT infrastructure. This should win more business opportunities in the long run. It should also reduce potential service calls by putting less hardware into the field to run the same applications.

among others, offers a product that addresses this need: Virtual Server 2005 R2 Standard Edition. In this Recipe, I'll discuss the ins and outs of server virtualization, and explain how to install Microsoft's Virtual Server 2005 R2 Standard Edition software to create a virtual server.

The proliferation of 64-bit computing hardware and software means system builders can now install extremely high-performance systems. For example, today's servers can address more than 4 GB of RAM. CPU clock speeds now approach 4 GHz, with multiple cores on each chip able to tackle additional computing demands. Storage capacity has skyrocketed, too: Terabyte storage arrays are now within the reach of system builders who serve the small and medium business (SMB) market. Until quite recently, these performance levels were limited to only the largest of enterprises. But no more.

Another phenomenon is contributing to the under-utilization of available hardware, too: Today's Commercial Off-The-Shelf (COTS) software is typically compiled for maximum compatibility, instead of for performance and resource utilization. In fact, these applications are typically optimized for use on a single machine with one or two CPUs and up to about 2 GB of RAM. But if your systems have more power than that, they're probably not being fully utilized by the software.

One way to combat this limitation is to make each application think it is running on just such a machine. For example, let's assume you have two key applications on the software side, and one server with two processors and 4 GB of RAM on the hardware. For this setup, you could create two virtual machines, one for each application. This way, the available hardware resources would be utilized by the software much more efficiently.

**Virtual Pitfalls**

But there are pitfalls to creating virtual machines. One is the added temptation to create more and more virtual machines. Also, the more virtual machines you have on one physical machine, the more problems you have with security vulnerabilities and time consumed with management duties. A hardware failure now affects not just one physical machine, but also several virtual ones.

Another pitfall: While CPU clock cycles, memory capacity, and storage are available in prolific quantity, I/O bottlenecks still exist. In fact, bottlenecks can be exacerbated if you create too many virtual machines on one physical machine. The risk is that all the virtual machines will attempt to access memory or storage simultaneously. If this happens, virtualization can actually reduce the server's performance.

There are other issues, too, most notably reliability, serviceability, and availability. As the number of virtual machines increases, the importance of these issues for the physical machine also increase, sometimes dramatically. For example, any situation that requires a re-boot will now affect multiple machines at once -- all the virtual machines, plus the physical machine. Also, any single hardware failure will now affect multiple servers, for the same reason.

Similar concerns apply to virtual machines from the standpoint of networking and security. Any time you connect a machine to a network, you have just created a security risk. Whether this risk is acceptable depends on the system owner and/or sys admin. But a security breach at one point could leave multiple servers vulnerable.

As a result, system builders must be careful when employing virtualization technology with servers. You need to strike a balance between financial savings on the one hand, and budget needed to manage, operate, and secure the virtualized servers on the other.

**HARDWARE:** To benefit from server consolidation through virtualization, you need a server that has plenty of hardware resources. This means a server with at least two CPUs (preferably multi-core), lots of RAM (at least 4 GB), and plenty of hard-drive capacity (at least 100 GB).

Also, you need to be able to allocate these resources across multiple virtual machines, each with its own operating system and application software. In fact, the Guest OS on the virtual machine does not have to be the same as the OS on the physical machine or host. You can even run multiple virtual machines, each with a different OS--another advantage of virtualization technology.

**SOFTWARE:** For this Recipe, you'll need Microsoft's Virtual Server 2005 R2 Standard Edition software. You will see how to install it, configure it, and create a virtual machine, as well as manage your virtual server remotely with built-in Web tools that are part of the Microsoft solution. Virtual Server is available as a free download, though the actual cost of running the software will depend on the licenses you purchase for your OS and application software. You can download Virtual Server 2005 from [this Microsoft page](http://www.microsoft.com/windowsserversystem/virtualserver/software/default.mspx). And for more information on licensing, check out [Microsoft Adapts Windows Server System Licensing to Virtualization Scenarios](http://www.microsoft.com/licensing/highlights/virtualization.mspx).

**How To Install And Configure Virtual Server 2005**

**Step 1:** Install the Microsoft Virtual Server 2005 R2 Standard Edition software. It is possible to install this software on Windows XP SP2; this is usually done for evaluation and testing purposes. But I recommend starting with a machine that uses any edition of the Windows 2003 Server OS, either SP1 or R2. For a more detailed list of system requirements, see this Microsoft page: [Virtual Server 2005 R2 System Requirements](http://www.microsoft.com/windowsserversystem/virtualserver/evaluation/sysreqs.mspx).

The Virtual Server software runs as a service within the host OS, and it can be started and stopped once installed. Also, the software creates a Web site within Internet Information Services (IIS) called the Virtual Server Administration site. This is essentially a browser-based Graphical User Interface (GUI) for your virtual machines.

If you don't have Microsoft's product release on CD, you can also obtain the software from this Microsoft page: [Download or Order Virtual Server 2005 R2](http://www.microsoft.com/windowsserversystem/virtualserver/software/default.mspx).

If you are installing Virtual Server 2005 for the first time, select the Complete installation option. This will give you access to documentation as well as the Virtual Server service and Administration site. In the future, if you are setting up multiple installations of the software across many physical machines, you can choose the Custom Installation option and install only the Virtual Server service.

As with any other Microsoft software, the default settings are adequate to get you up and running. They are usually best left unmodified for proper functionality. For example, while the default port number for the Administration Web site can be changed, if you change this port, you will need to make sure you account for this change when trying to connect to the Web site from another machine.

Also, if you ever need to remove the Virtual Server 2005 software, you can uninstall it without deleting your Virtual Machine files. This comes in handy if you want to move those virtual machines to another physical machine later.

**Step 2:** Configure the software and create a virtual machine. To do this, open a browser, and in the Address field, type this:

**http://FQDN:1024/VirtualServer/VSWebApp.exe**

Please note, FQDN stands for Fully Qualified Domain Name, so you will not actually type the letters "FQDN." Instead, type in the machine's domain name, which is the machine's name, sub-domain (optional; not every machine has one), and top-level domain. Here's an example, in which the FQDN is "ckb1.mail.app-com.com":

**http://ckb1.mail.app-com.com:1024/VirtualServer/VSWebApp.exe**

Also, if you have changed the port number, replace the "1024" with the port number you specified. After you enter your credentials, click OK. You will see a menu on the left-hand side of the screen. Under the Virtual Machines menu, click Create. This will get you started.

There are four parameters you will need to specify here to create your virtual machine:

* Name of the Virtual Machine (VM).
* How much physical RAM from the host machine will be allocated to this VM.
* Create and/or Attach Virtual Hard Disk (VHD).
* Type of Virtual Network Adapter to use.

Virtual Hard Drives are optional, as are Virtual Network Adapters. You can always create and attach these later, if you prefer. Once you've decided on the values for your parameters above, click the Create button to create the VM.

You now have the bare minimum necessary to get a VM running. But you may want to tweak a couple of settings. This may simplify things as you work with the VM.

For instance, you may want to enable the Virtual Machine Remote Control Server. It lets you connect directly to a VM via a client interface as well as the Web application. To do so, go to the Virtual Server menu. Click on an option called Server Properties. Then click on VMRC Server.

Also, you may want to specify your own folder location where your VM files are stored, to make them easier to find later. The default location is buried within the Documents and Settings folder -- not the best place. To change the default location for VM file storage, go to the Virtual Server menu and click on Server Properties. Then click Search Paths. You can enter both a folder location where your VM files will be stored, and custom search paths to look for related files.

To see other configuration options, follow the menu structure on the left-hand side of the screen. As you become more familiar with the software and use your VMs more, you will probably re-visit this menu often, tweaking settings to your liking. Also, all the parameters you configured at the beginning of this step can be modified here. You can also create additional VHDs and change your network connectivity, if desired.

## Reduce IT Costs and Increase Control

Eliminate over-provisioning, increase server utilization and limit the environmental impact of IT by consolidating your server hardware with [vSphere® with Operations Management™](http://www.vmware.com/products/vsphere/), VMware's virtualization platform.

Server consolidation lets your organization:

* Reduce hardware and operating costs by as much as 50 percent and energy costs by as much as 80 percent, saving more than $3,000 per year for each virtualized server workload.
* Reduce the time it takes to provision new servers by as much as 70 percent.
* Decrease downtime and improve reliability with [business continuity](http://www.vmware.com/business-continuity/disaster-recovery.html) and built-in disaster recovery.
* Deliver IT services on demand, independent of hardware, operating systems, applications or infrastructure providers.

## Revlon Customer Spotlight Video

[](http://www.vmware.com/consolidation/overview#1695973074001)

* [Submit Solutions Request](http://www.vmware.com/request.html)
* [Contact VMware Sales](http://www.vmware.com/company/contact_sales.html)
* [Use TCO Calculator](http://www.vmware.com/go/calculator)

### Consolidate Hardware with Desktop and Server Virtualization

By consolidating your server hardware with vSphere with Operations Management, your organization can increase existing hardware utilization from as low as 5 percent to as much as 80 percent. You can also reduce energy consumption by decreasing the number of servers in your data center. VMware server virtualization can reduce hardware requirements by a 15:1 ratio, enabling you to lessen the environmental impact of your organization's IT without sacrificing reliability or service levels.

Server and desktop hardware consolidation can also help you achieve a 20 to 30 percent lower cost per application, as well as defer data center construction costs by $1,000 per square foot. vSphere with Operations Management allows for a 50 to 70 percent higher virtual machine density per host than commodity offerings.

### Centralize Management of Your Virtual Infrastructure

Unlike vendors that only offer single-point solutions for server virtualizations, VMware lets you manage an entire virtual infrastructure from a single point of control. With vSphere with Operations Management, you can accelerate provisioning time by 50 to 70 percent, manage virtual machines from a central location and monitor the performance of all your virtual machines and their hosts.

### Automate the Virtual Infrastructure

A virtual infrastructure can deliver performance, scalability and availability levels that are impossible with a physical infrastructure. Using VMware® vCenter Server™ with VMware vSphere® Storage vMotion® and VMware vSphere® Distributed Resource Scheduler™, your organization can avoid planned downtime, enable dynamic, policy-based allocation of IT resources and eliminate repetitive configuration and maintenance tasks. vSphere also supports disaster recovery and can help you avoid unplanned downtime.

## Eliminate Over-Provisioning

### Increase Server Utilization

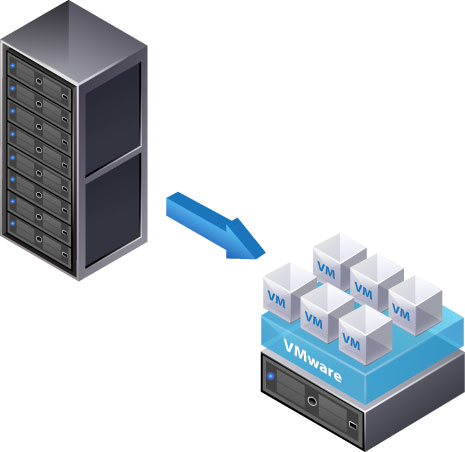
Get more out of your existing hardware by consolidating x86 servers with VMware vSphere®. In the traditional "one workload, one box" approach to server provisioning, most servers operate at just five to 15 percent of their total load capacity, which results in over-provisioning and under-utilization. You can decrease server sprawl and increase utilization by converting your x86 servers into virtual machines that run independently from the underlying hardware.

Each virtual machine represents a complete system that can run Windows, Linux, Solaris and NetWare operating systems and software applications; many virtual machines can run independently on the same physical server at the same time. Running multiple workloads on highly configured x86 servers allows you to increase the utilization of your server hardware from 10 to 15 percent to as much as 80 percent.

Download the free [VMware® vCenter™ Converter™](http://www.vmware.com/products/converter/) utility to convert physical machines to virtual machines as well as to convert virtual machine formats.

### Reduce Hardware Requirements

Using [VMware vSphere®](http://www.vmware.com/products/vsphere/) with Operations Management™, your organization can reduce the IT hardware in your data center. Each host server running vSphere can support more than 300 virtual machines. Many organizations run as many as 10 applications on a single piece of hardware, reducing their hardware requirements by a 15:1 ratio.



### Cut Hardware and Operating Costs

Consolidating your hardware means that you need fewer servers in your data center,which means you will spend less on hardware and maintenanceand less onenergy for power and cooling. VMware virtualization can help your organization save more than $3,000 annually for each server you virtualize.

You can also shift IT resources and budgets away from tactical maintenance to strategic projects. Through automation, VMware vSphere and the vCenter product family simplify tedious day-to-day tasks like provisioning, hardware maintenance and performance management. This reduces the cost and complexity of managing IT in your data center.

# Linux Set or Change User Password

by [nixCraft](http://www.cyberciti.biz/tips/about-us) on September 21, 2006 · [46 comments](http://www.cyberciti.biz/faq/linux-set-change-password-how-to/#comments)· LAST UPDATED May 27, 2010

in [Linux](http://www.cyberciti.biz/faq/category/linux/)

[](http://www.cyberciti.biz/faq/category/linux/)

How do I set or change Linux system password for any user account?  
  
Both Linux and UNIX use the passwd command to change user password. The passwd is used to update a user’s authentication token (password) stored in [shadow file](http://www.cyberciti.biz/faq/understanding-etcshadow-file/).

The passwd changes passwords for user and group accounts. A normal user may only change the password for his/her own account, the super user (or root) may change the password for any account. The administrator of a group may change the password for the group. passwd also changes account information, such as the full name of the user, user's login shell, or password expiry date and interval.

## Task: Set or Change User Password

Type passwd command as follows to change your own password:  
$ passwd  
Output:

Changing password for vivek

(current) UNIX password:

Enter new UNIX password:

Retype new UNIX password:

passwd: password updated successfully

The user is first prompted for his/her old password, if one is present. This password is then encrypted and compared against the stored password. The user has only one chance to enter the correct password. The super user is permitted to bypass this step so that forgotten passwords may be changed.

A new password is tested for complexity. As a general guideline, passwords should consist of 6 to 8 characters including one or more from each of following sets:

1. Lower case alphabetics
2. Upper case alphabetics
3. Digits 0 thru 9
4. Punctuation marks

## Task: Change Password For Other User Account

You must login as root user, type the following command to change password for user vivek:  
# passwd vivek  
Output:

Enter new UNIX password:

Retype new UNIX password:

passwd: password updated successfully

Where,

* vivek - is username or account name.

## Task: Change Group Password

When the -g option is used, the password for the named group is changed. In this example, change password for group sales:  
# passwd -g sales

The current group password is not prompted for. The -r option is used with the -g option to remove the current password from the named group. This allows group access to all members. The -R option is used with the -g option to restrict the named group for all users.

# 20 Linux System Monitoring Tools Every SysAdmin Should Know

by [nixCraft](http://www.cyberciti.biz/tips/about-us) on June 27, 2009 · [326 comments](http://www.cyberciti.biz/tips/top-linux-monitoring-tools.html#comments)· LAST UPDATED November 6, 2012

in [CentOS](http://www.cyberciti.biz/tips/category/centos), [Debian Linux](http://www.cyberciti.biz/tips/category/debian-linux), [fedora linux](http://www.cyberciti.biz/tips/category/fedora-linux)

[](http://www.cyberciti.biz/tips/category/linux)

Need to monitor Linux server performance? Try these built-in commands and a few add-on tools. Most Linux distributions are equipped with tons of monitoring. These tools provide metrics which can be used to get information about system activities. You can use these tools to find the possible causes of a performance problem. The commands discussed below are some of the most basic commands when it comes to system analysis and debugging server issues such as:

1. Finding out bottlenecks.
2. Disk (storage) bottlenecks.
3. CPU and memory bottlenecks.
4. Network bottlenecks.

## #1: top - Process Activity Command

The top program provides a dynamic real-time view of a running system i.e. actual process activity. By default, it displays the most CPU-intensive tasks running on the server and updates the list every five seconds.

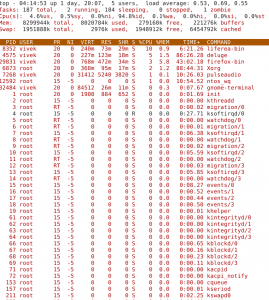
[](http://www.cyberciti.biz/tips/top-linux-monitoring-tools.html/top-output)

Fig.01: Linux top command

### Commonly Used Hot Keys

The top command provides several useful hot keys:

|  |  |
| --- | --- |
| **Hot Key** | **Usage** |
| t | Displays summary information off and on. |
| m | Displays memory information off and on. |
| A | Sorts the display by top consumers of various system resources. Useful for quick identification of performance-hungry tasks on a system. |
| f | Enters an interactive configuration screen for top. Helpful for setting up top for a specific task. |
| o | Enables you to interactively select the ordering within top. |
| r | Issues renice command. |
| k | Issues kill command. |
| z | Turn on or off color/mono |

**=> Related:** [How do I Find Out Linux CPU Utilization?](http://www.cyberciti.biz/tips/how-do-i-find-out-linux-cpu-utilization.html)

## #2: vmstat - System Activity, Hardware and System Information

The command vmstat reports information about processes, memory, paging, block IO, traps, and cpu activity.  
# vmstat 3  
Sample Outputs:

procs -----------memory---------- ---swap-- -----io---- --system-- -----cpu------

r b swpd free buff cache si so bi bo in cs us sy id wa st

0 0 0 2540988 522188 5130400 0 0 2 32 4 2 4 1 96 0 0

1 0 0 2540988 522188 5130400 0 0 0 720 1199 665 1 0 99 0 0

0 0 0 2540956 522188 5130400 0 0 0 0 1151 1569 4 1 95 0 0

0 0 0 2540956 522188 5130500 0 0 0 6 1117 439 1 0 99 0 0

0 0 0 2540940 522188 5130512 0 0 0 536 1189 932 1 0 98 0 0

0 0 0 2538444 522188 5130588 0 0 0 0 1187 1417 4 1 96 0 0

0 0 0 2490060 522188 5130640 0 0 0 18 1253 1123 5 1 94 0 0

### Display Memory Utilization Slabinfo

# vmstat -m

### Get Information About Active / Inactive Memory Pages

# vmstat -a  
**=> Related:** [How do I find out Linux Resource utilization to detect system bottlenecks?](http://www.cyberciti.biz/tips/linux-resource-utilization-to-detect-system-bottlenecks.html)

## #3: w - Find Out Who Is Logged on And What They Are Doing

w command displays information about the users currently on the machine, and their processes.  
# w username  
# w vivek  
Sample Outputs:

17:58:47 up 5 days, 20:28, 2 users, load average: 0.36, 0.26, 0.24

USER TTY FROM LOGIN@ IDLE JCPU PCPU WHAT

root pts/0 10.1.3.145 14:55 5.00s 0.04s 0.02s vim /etc/resolv.conf

root pts/1 10.1.3.145 17:43 0.00s 0.03s 0.00s w

### #4: uptime - Tell How Long The System Has Been Running

The uptime command can be used to see how long the server has been running. The current time, how long the system has been running, how many users are currently logged on, and the system load averages for the past 1, 5, and 15 minutes.  
# uptime  
Output:

18:02:41 up 41 days, 23:42, 1 user, load average: 0.00, 0.00, 0.00

1 can be considered as optimal load value. The load can change from system to system. For a single CPU system 1 - 3 and SMP systems 6-10 load value might be acceptable.

## #5: ps - Displays The Processes

ps command will report a snapshot of the current processes. To select all processes use the -A or -e option:  
# ps -A  
Sample Outputs:

PID TTY TIME CMD

1 ? 00:00:02 init

2 ? 00:00:02 migration/0

3 ? 00:00:01 ksoftirqd/0

4 ? 00:00:00 watchdog/0

5 ? 00:00:00 migration/1

6 ? 00:00:15 ksoftirqd/1

4881 ? 00:53:28 java

4885 tty1 00:00:00 mingetty

4886 tty2 00:00:00 mingetty

4887 tty3 00:00:00 mingetty

4888 tty4 00:00:00 mingetty

4891 tty5 00:00:00 mingetty

4892 tty6 00:00:00 mingetty

4893 ttyS1 00:00:00 agetty

12853 ? 00:00:00 cifsoplockd

12854 ? 00:00:00 cifsdnotifyd

14231 ? 00:10:34 lighttpd

14232 ? 00:00:00 php-cgi

54981 pts/0 00:00:00 vim

55465 ? 00:00:00 php-cgi

55546 ? 00:00:00 bind9-snmp-stat

55704 pts/1 00:00:00 ps

ps is just like top but provides more information.

### Show Long Format Output

# ps -Al  
To turn on extra full mode (it will show command line arguments passed to process):  
# ps -AlF

### To See Threads ( LWP and NLWP)

# ps -AlFH

### To See Threads After Processes

# ps -AlLm

### Print All Process On The Server

# ps ax  
# ps axu

### Print A Process Tree

# ps -ejH  
# ps axjf  
# pstree

### Print Security Information

# ps -eo euser,ruser,suser,fuser,f,comm,label  
# ps axZ  
# ps -eM

### See Every Process Running As User Vivek

# ps -U vivek -u vivek u

### Set Output In a User-Defined Format

# ps -eo pid,tid,class,rtprio,ni,pri,psr,pcpu,stat,wchan:14,comm  
# ps axo stat,euid,ruid,tty,tpgid,sess,pgrp,ppid,pid,pcpu,comm  
# ps -eopid,tt,user,fname,tmout,f,wchan

### Display Only The Process IDs of Lighttpd

# ps -C lighttpd -o pid=  
OR  
# pgrep lighttpd  
OR  
# pgrep -u vivek php-cgi

### Display The Name of PID 55977

# ps -p 55977 -o comm=

### Find Out The Top 10 Memory Consuming Process

# ps -auxf | sort -nr -k 4 | head -10

### Find Out top 10 CPU Consuming Process

# ps -auxf | sort -nr -k 3 | head -10

## #6: free - Memory Usage

The command free displays the total amount of free and used physical and swap memory in the system, as well as the buffers used by the kernel.  
# free   
Sample Output:

total used free shared buffers cached

Mem: 12302896 9739664 2563232 0 523124 5154740

-/+ buffers/cache: 4061800 8241096

Swap: 1052248 0 1052248

**=> Related:** :

1. [Linux Find Out Virtual Memory PAGESIZE](http://www.cyberciti.biz/faq/linux-check-the-size-of-pagesize/)
2. [Linux Limit CPU Usage Per Process](http://www.cyberciti.biz/faq/cpu-usage-limiter-for-linux/)
3. [How much RAM does my Ubuntu / Fedora Linux desktop PC have?](http://www.cyberciti.biz/tips/how-much-ram-does-my-linux-system.html)

## #7: iostat - Average CPU Load, Disk Activity

The command iostat report Central Processing Unit (CPU) statistics and input/output statistics for devices, partitions and network filesystems (NFS).  
# iostat   
Sample Outputs:

Linux 2.6.18-128.1.14.el5 (www03.nixcraft.in) 06/26/2009

avg-cpu: %user %nice %system %iowait %steal %idle

3.50 0.09 0.51 0.03 0.00 95.86

Device: tps Blk\_read/s Blk\_wrtn/s Blk\_read Blk\_wrtn

sda 22.04 31.88 512.03 16193351 260102868

sda1 0.00 0.00 0.00 2166 180

sda2 22.04 31.87 512.03 16189010 260102688

sda3 0.00 0.00 0.00 1615 0

**=> Related:** : [Linux Track NFS Directory / Disk I/O Stats](http://www.cyberciti.biz/faq/howto-linux-track-nfs-client-disk-metrics/)

## #8: sar - Collect and Report System Activity

The sar command is used to collect, report, and save system activity information. To see network counter, enter:  
# sar -n DEV | more  
To display the network counters from the 24th:  
# sar -n DEV -f /var/log/sa/sa24 | more  
You can also display real time usage using sar:  
# sar 4 5  
Sample Outputs:

Linux 2.6.18-128.1.14.el5 (www03.nixcraft.in) 06/26/2009

06:45:12 PM CPU %user %nice %system %iowait %steal %idle

06:45:16 PM all 2.00 0.00 0.22 0.00 0.00 97.78

06:45:20 PM all 2.07 0.00 0.38 0.03 0.00 97.52

06:45:24 PM all 0.94 0.00 0.28 0.00 0.00 98.78

06:45:28 PM all 1.56 0.00 0.22 0.00 0.00 98.22

06:45:32 PM all 3.53 0.00 0.25 0.03 0.00 96.19

Average: all 2.02 0.00 0.27 0.01 0.00 97.70

**=> Related:** : [How to collect Linux system utilization data into a file](http://www.cyberciti.biz/tips/howto-write-system-utilization-data-to-file.html)

## #9: mpstat - Multiprocessor Usage

The mpstat command displays activities for each available processor, processor 0 being the first one. mpstat -P ALL to display average CPU utilization per processor:  
# mpstat -P ALL  
Sample Output:

Linux 2.6.18-128.1.14.el5 (www03.nixcraft.in) 06/26/2009

06:48:11 PM CPU %user %nice %sys %iowait %irq %soft %steal %idle intr/s

06:48:11 PM all 3.50 0.09 0.34 0.03 0.01 0.17 0.00 95.86 1218.04

06:48:11 PM 0 3.44 0.08 0.31 0.02 0.00 0.12 0.00 96.04 1000.31

06:48:11 PM 1 3.10 0.08 0.32 0.09 0.02 0.11 0.00 96.28 34.93

06:48:11 PM 2 4.16 0.11 0.36 0.02 0.00 0.11 0.00 95.25 0.00

06:48:11 PM 3 3.77 0.11 0.38 0.03 0.01 0.24 0.00 95.46 44.80

06:48:11 PM 4 2.96 0.07 0.29 0.04 0.02 0.10 0.00 96.52 25.91

06:48:11 PM 5 3.26 0.08 0.28 0.03 0.01 0.10 0.00 96.23 14.98

06:48:11 PM 6 4.00 0.10 0.34 0.01 0.00 0.13 0.00 95.42 3.75

06:48:11 PM 7 3.30 0.11 0.39 0.03 0.01 0.46 0.00 95.69 76.89

**=> Related:** : [Linux display each multiple SMP CPU processors utilization individually](http://www.cyberciti.biz/faq/linux-mpstat-command-report-processors-related-statistics/).

## #10: pmap - Process Memory Usage

The command pmap report memory map of a process. Use this command to find out causes of memory bottlenecks.  
# pmap -d PID  
To display process memory information for pid # 47394, enter:  
# pmap -d 47394  
Sample Outputs:

47394: /usr/bin/php-cgi

Address Kbytes Mode Offset Device Mapping

0000000000400000 2584 r-x-- 0000000000000000 008:00002 php-cgi

0000000000886000 140 rw--- 0000000000286000 008:00002 php-cgi

00000000008a9000 52 rw--- 00000000008a9000 000:00000 [ anon ]

0000000000aa8000 76 rw--- 00000000002a8000 008:00002 php-cgi

000000000f678000 1980 rw--- 000000000f678000 000:00000 [ anon ]

000000314a600000 112 r-x-- 0000000000000000 008:00002 ld-2.5.so

000000314a81b000 4 r---- 000000000001b000 008:00002 ld-2.5.so

000000314a81c000 4 rw--- 000000000001c000 008:00002 ld-2.5.so

000000314aa00000 1328 r-x-- 0000000000000000 008:00002 libc-2.5.so

000000314ab4c000 2048 ----- 000000000014c000 008:00002 libc-2.5.so

.....

......

..

00002af8d48fd000 4 rw--- 0000000000006000 008:00002 xsl.so

00002af8d490c000 40 r-x-- 0000000000000000 008:00002 libnss\_files-2.5.so

00002af8d4916000 2044 ----- 000000000000a000 008:00002 libnss\_files-2.5.so

00002af8d4b15000 4 r---- 0000000000009000 008:00002 libnss\_files-2.5.so

00002af8d4b16000 4 rw--- 000000000000a000 008:00002 libnss\_files-2.5.so

00002af8d4b17000 768000 rw-s- 0000000000000000 000:00009 zero (deleted)

00007fffc95fe000 84 rw--- 00007ffffffea000 000:00000 [ stack ]

ffffffffff600000 8192 ----- 0000000000000000 000:00000 [ anon ]

mapped: 933712K writeable/private: 4304K shared: 768000K

The last line is very important:

* **mapped: 933712K** total amount of memory mapped to files
* **writeable/private: 4304K** the amount of private address space
* **shared: 768000K** the amount of address space this process is sharing with others

**=> Related:** : [Linux find the memory used by a program / process using pmap command](http://www.cyberciti.biz/tips/howto-find-memory-used-by-program.html)

## #11 and #12: netstat and ss - Network Statistics

The command netstat displays network connections, routing tables, interface statistics, masquerade connections, and multicast memberships. ss command is used to dump socket statistics. It allows showing information similar to netstat. See the following resources about ss and netstat commands:

* [ss: Display Linux TCP / UDP Network and Socket Information](http://www.cyberciti.biz/tips/linux-investigate-sockets-network-connections.html)
* [Get Detailed Information About Particular IP address Connections Using netstat Command](http://www.cyberciti.biz/tips/netstat-command-tutorial-examples.html)

## #13: iptraf - Real-time Network Statistics

The iptraf command is interactive colorful IP LAN monitor. It is an ncurses-based IP LAN monitor that generates various network statistics including TCP info, UDP counts, ICMP and OSPF information, Ethernet load info, node stats, IP checksum errors, and others. It can provide the following info in easy to read format:

* Network traffic statistics by TCP connection
* IP traffic statistics by network interface
* Network traffic statistics by protocol
* Network traffic statistics by TCP/UDP port and by packet size
* Network traffic statistics by Layer2 address

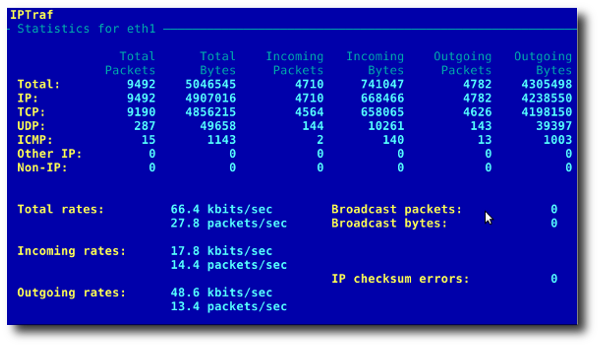
[](http://www.cyberciti.biz/tips/top-linux-monitoring-tools.html/iptraf3)

Fig.02: General interface statistics: IP traffic statistics by network interface

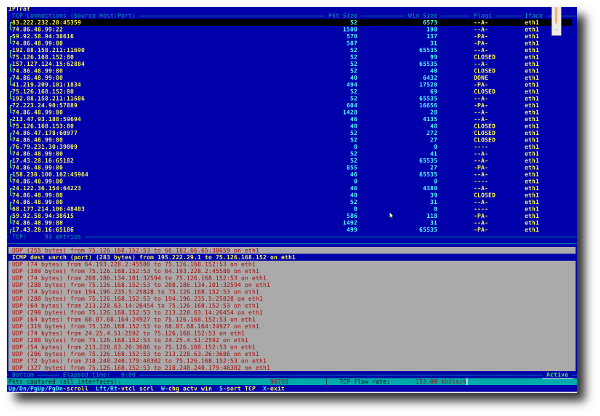
[](http://www.cyberciti.biz/tips/top-linux-monitoring-tools.html/iptraf2)

Fig.03 Network traffic statistics by TCP connection

## #14: tcpdump - Detailed Network Traffic Analysis

The tcpdump is simple command that dump traffic on a network. However, you need good understanding of TCP/IP protocol to utilize this tool. For.e.g to display traffic info about DNS, enter:  
# tcpdump -i eth1 'udp port 53'  
To display all IPv4 HTTP packets to and from port 80, i.e. print only packets that contain data, not, for example, SYN and FIN packets and ACK-only packets, enter:  
# tcpdump 'tcp port 80 and (((ip[2:2] - ((ip[0]&0xf)<<2)) - ((tcp[12]&0xf0)>>2)) != 0)'  
To display all FTP session to 202.54.1.5, enter:  
# tcpdump -i eth1 'dst 202.54.1.5 and (port 21 or 20'  
To display all HTTP session to 192.168.1.5:  
# tcpdump -ni eth0 'dst 192.168.1.5 and tcp and port http'  
Use [wireshark to view detailed](http://www.cyberciti.biz/faq/linux-unix-bsd-apache-tcpdump-http-packets-sniffing/) information about files, enter:  
# tcpdump -n -i eth1 -s 0 -w output.txt src or dst port 80

## #15: strace - System Calls

Trace system calls and signals. This is useful for debugging webserver and other server problems. See how to use to [trace the process and](http://www.cyberciti.biz/tips/linux-strace-command-examples.html) see What it is doing.

## #16: /Proc file system - Various Kernel Statistics

/proc file system provides detailed information about various hardware devices and other Linux kernel information. See [Linux kernel /proc](http://www.cyberciti.biz/files/linux-kernel/Documentation/filesystems/proc.txt) documentations for further details. Common /proc examples:  
# cat /proc/cpuinfo  
# cat /proc/meminfo  
# cat /proc/zoneinfo  
# cat /proc/mounts

## 17#: Nagios - Server And Network Monitoring

[Nagios](http://www.nagios.org/) is a popular open source computer system and network monitoring application software. You can easily monitor all your hosts, network equipment and services. It can send alert when things go wrong and again when they get better. [FAN is](http://fannagioscd.sourceforge.net/drupal/) "Fully Automated Nagios". FAN goals are to provide a Nagios installation including most tools provided by the Nagios Community. FAN provides a CDRom image in the standard ISO format, making it easy to easilly install a Nagios server. Added to this, a wide bunch of tools are including to the distribution, in order to improve the user experience around Nagios.

## 18#: Cacti - Web-based Monitoring Tool

Cacti is a complete network graphing solution designed to harness the power of RRDTool's data storage and graphing functionality. Cacti provides a fast poller, advanced graph templating, multiple data acquisition methods, and user management features out of the box. All of this is wrapped in an intuitive, easy to use interface that makes sense for LAN-sized installations up to complex networks with hundreds of devices. It can provide data about network, CPU, memory, logged in users, Apache, DNS servers and much more. See how [to install and configure Cacti network graphing](http://www.cyberciti.biz/faq/fedora-rhel-install-cacti-monitoring-rrd-software/) tool under CentOS / RHEL.

## #19: KDE System Guard - Real-time Systems Reporting and Graphing

KSysguard is a network enabled task and system monitor application for KDE desktop. This tool can be run over ssh session. It provides lots of features such as a client/server architecture that enables monitoring of local and remote hosts. The graphical front end uses so-called sensors to retrieve the information it displays. A sensor can return simple values or more complex information like tables. For each type of information, one or more displays are provided. Displays are organized in worksheets that can be saved and loaded independently from each other. So, KSysguard is not only a simple task manager but also a very powerful tool to control large server farms.

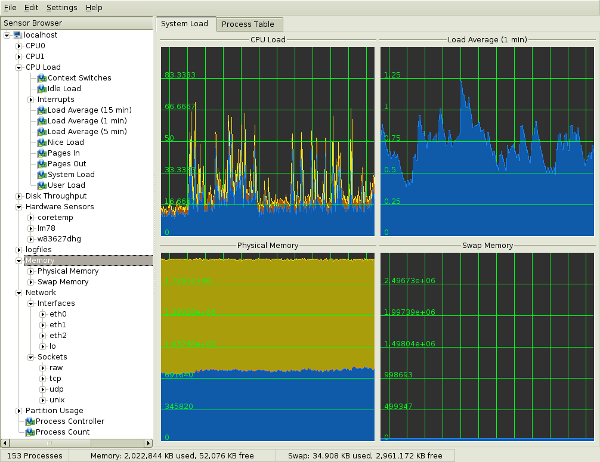
[](http://www.cyberciti.biz/tips/top-linux-monitoring-tools.html/kde-systemguard-screenshot)

Fig.05 KDE System Guard {Image credit: Wikipedia}

See [the KSysguard handbook](http://docs.kde.org/stable/en/kdebase-workspace/ksysguard/index.html) for detailed usage.

## #20: Gnome System Monitor - Real-time Systems Reporting and Graphing

The System Monitor application enables you to display basic system information and monitor system processes, usage of system resources, and file systems. You can also use System Monitor to modify the behavior of your system. Although not as powerful as the KDE System Guard, it provides the basic information which may be useful for new users:

* Displays various basic information about the computer's hardware and software.
* Linux Kernel version
* GNOME version
* Hardware
* Installed memory
* Processors and speeds
* System Status
* Currently available disk space
* Processes
* Memory and swap space
* Network usage
* File Systems
* Lists all mounted filesystems along with basic information about each.

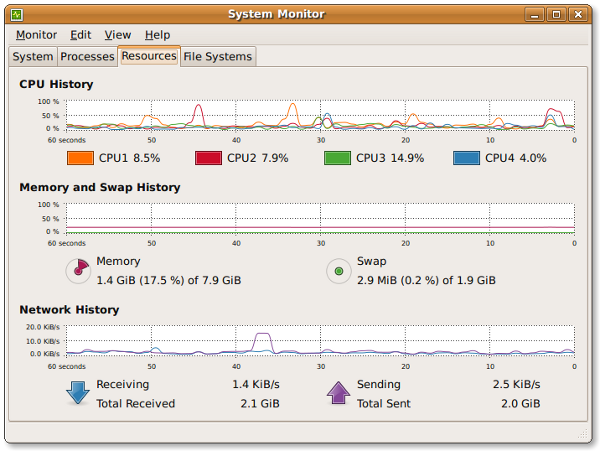
[](http://www.cyberciti.biz/tips/top-linux-monitoring-tools.html/gnome-system-monitor)

Fig.06 The Gnome System Monitor application

## Bonus: Additional Tools

A few more tools:

* [nmap](http://www.cyberciti.biz/tips/linux-scanning-network-for-open-ports.html) - scan your server for open ports.
* [lsof](http://www.cyberciti.biz/tips/tag/lsof-command) - list open files, network connections and much more.
* [ntop](http://www.cyberciti.biz/faq/debian-ubuntu-install-ntop-network-traffic-monitoring-software/) web based tool - ntop is the best tool to see network usage in a way similar to what top command does for processes i.e. it is network traffic monitoring software. You can see network status, protocol wise distribution of traffic for UDP, TCP, DNS, HTTP and other protocols.
* [Conky](http://conky.sourceforge.net/) - Another good monitoring tool for the X Window System. It is highly configurable and is able to monitor many system variables including the status of the CPU, memory, swap space, disk storage, temperatures, processes, network interfaces, battery power, system messages, e-mail inboxes etc.
* [GKrellM](http://members.dslextreme.com/users/billw/gkrellm/gkrellm.html) - It can be used to monitor the status of CPUs, main memory, hard disks, network interfaces, local and remote mailboxes, and many other things.
* [vnstat](http://www.cyberciti.biz/tips/keeping-a-log-of-daily-network-traffic-for-adsl-or-dedicated-remote-linux-box.html) - vnStat is a console-based network traffic monitor. It keeps a log of hourly, daily and monthly network traffic for the selected interface(s).
* [htop](http://htop.sourceforge.net/) - htop is an enhanced version of top, the interactive process viewer, which can display the list of processes in a tree form.
* [mtr](http://www.cyberciti.biz/tips/finding-out-a-bad-or-simply-overloaded-network-link-with-linuxunix-oses.html) - mtr combines the functionality of the traceroute and ping programs in a single network diagnostic tool.

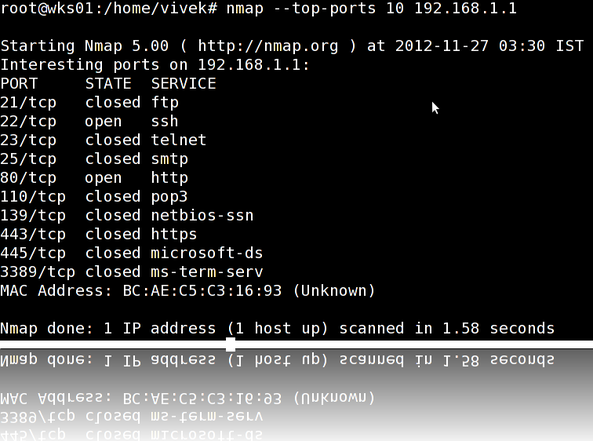
# Top 30 Nmap Command Examples For Sys/Network Admins

by [nixCraft](http://www.cyberciti.biz/tips/about-us) on November 26, 2012 · [8 comments](http://www.cyberciti.biz/networking/nmap-command-examples-tutorials/#comments)· Last updated December 11, 2012

in [Command Line Hacks](http://www.cyberciti.biz/topics/open-source/command-line-hacks/), [Howto](http://www.cyberciti.biz/topics/howto/), [Networking](http://www.cyberciti.biz/topics/networking/), [Security](http://www.cyberciti.biz/topics/security/)

Nmap is short for Network Mapper. It is an open source security tool for network exploration, security scanning and auditing. However, nmap command comes with lots of options that can make the utility more robust and difficult to follow for new users.

The purpose of this post is to introduce a user to the nmap command line tool to scan a host and/or network, so to find out the possible vulnerable points in the hosts. You will also learn how to use Nmap for offensive and defensive purposes.

[](http://www.cyberciti.biz/networking/nmap-command-examples-tutorials/attachment/welcome-nmap/)

nmap in action

## More about nmap

From the man page:

Nmap ("Network Mapper") is an open source tool for network exploration and security auditing. It was designed to rapidly scan large networks, although it works fine against single hosts. Nmap uses raw IP packets in novel ways to determine what hosts are available on the network, what services (application name and version) those hosts are offering, what operating systems (and OS versions) they are running, what type of packet filters/firewalls are in use, and dozens of other characteristics. While Nmap is commonly used for security audits, many systems and network administrators find it useful for routine tasks such as network inventory, managing service upgrade schedules, and monitoring host or service uptime.

It was originally written by Gordon Lyon and it can answer the following questions easily:

1. What computers did you find running on the local network?
2. What IP addresses did you find running on the local network?
3. What is the operating system of your target machine?
4. Find out what ports are open on the machine that you just scanned?
5. Find out if the system is infected with malware or virus.
6. Search for unauthorized servers or network service on your network.
7. Find and remove computers which don't meet the organization's minimum level of security.

## Sample setup (LAB)

Port scanning may be illegal in some jurisdictions. So setup a lab as follows:

+---------+

+---------+ | Network | +--------+

| server1 |-----------+ swtich +---------|server2 |

+---------+ | (sw0) | +--------+

+----+----+

|

|

+---------+----------+

| wks01 Linux/OSX |

+--------------------+

Where,

* wks01 is your computer either running Linux/OS X or Unix like operating system. It is used for scanning your local network. The nmap command must be installed on this computer.
* server1 can be powered by Linux / Unix / MS-Windows operating systems. This is an unpatched server. Feel free to install a few services such as a web-server, file server and so on.
* server2 can be powered by Linux / Unix / MS-Windows operating systems. This is a [fully patched server with firewall](http://www.cyberciti.biz/tips/linux-iptables-examples.html). Again, feel free to install few services such as a web-server, file server and so on.
* All three systems are connected via switch.

## How do I install nmap?

See:

1. [Debian / Ubuntu Linux: Install nmap Software For Scanning Network](http://www.cyberciti.biz/faq/install-nmap-debian-ubuntu-server-desktop-system/)
2. [CentOS / RHEL: Install nmap Network Security Scanner](http://www.cyberciti.biz/faq/howto-install-nmap-on-centos-rhel-redhat-enterprise-linux/)
3. [OpenBSD: Install nmap Network Security Scanner](http://www.cyberciti.biz/faq/installing-nmap-network-port-scanner-under-openbsd-using-pkg_add/)

## #1: Scan a single host or an IP address (IPv4)

*### Scan a single ip address ###*

**nmap** 192.168.1.1

*## Scan a host name ###*

**nmap** server1.cyberciti.biz

*## Scan a host name with* ***more*** *info###*

**nmap** -v server1.cyberciti.biz

Sample outputs:

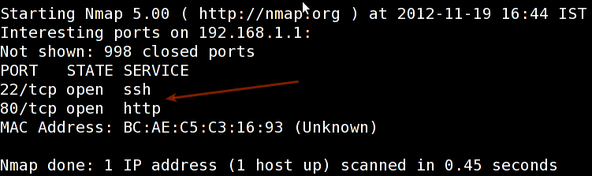
[](http://www.cyberciti.biz/faq/howto-install-nmap-on-centos-rhel-redhat-enterprise-linux/)

Fig.01: nmap output

## #2: Scan multiple IP address or subnet (IPv4)

nmap 192.168.1.1 192.168.1.2 192.168.1.3

## works with same subnet i.e. 192.168.1.0/24

nmap 192.168.1.1,2,3

You can scan a range of IP address too:

nmap 192.168.1.1-20

You can scan a range of IP address using a wildcard:

nmap 192.168.1.\*

Finally, you scan an entire subnet:

nmap 192.168.1.0/24

## #3: Read list of hosts/networks from a file (IPv4)

The -iL option allows you to read the list of target systems using a text file. This is useful to scan a large number of hosts/networks. Create a text file as follows:  
cat > /tmp/test.txt  
Sample outputs:

server1.cyberciti.biz

192.168.1.0/24

192.168.1.1/24

10.1.2.3

localhost

The syntax is:

nmap -iL /tmp/test.txt

## #4: Excluding hosts/networks (IPv4)

When scanning a large number of hosts/networks you can exclude hosts from a scan:

nmap 192.168.1.0/24 --exclude 192.168.1.5

nmap 192.168.1.0/24 --exclude 192.168.1.5,192.168.1.254

OR exclude list from a file called /tmp/exclude.txt

nmap -iL /tmp/scanlist.txt --excludefile /tmp/exclude.txt

## #5: Turn on OS and version detection scanning script (IPv4)

nmap -A 192.168.1.254

nmap -v -A 192.168.1.1

nmap -A -iL /tmp/scanlist.txt

## #6: Find out if a host/network is protected by a firewall

nmap -sA 192.168.1.254

nmap -sA server1.cyberciti.biz

## #7: Scan a host when protected by the firewall

nmap -PN 192.168.1.1

nmap -PN server1.cyberciti.biz

## #8: Scan an IPv6 host/address

The -6 option enable IPv6 scanning. The syntax is:

nmap -6 IPv6-Address-Here

nmap -6 server1.cyberciti.biz

nmap -6 2607:f0d0:1002:51::4

nmap -v A -6 2607:f0d0:1002:51::4

## #9: Scan a network and find out which servers and devices are up and running

This is known as host discovery or ping scan:

nmap -sP 192.168.1.0/24

Sample outputs:

Host 192.168.1.1 is up (0.00035s latency).

MAC Address: BC:AE:C5:C3:16:93 (Unknown)

Host 192.168.1.2 is up (0.0038s latency).

MAC Address: 74:44:01:40:57:FB (Unknown)

Host 192.168.1.5 is up.

Host nas03 (192.168.1.12) is up (0.0091s latency).

MAC Address: 00:11:32:11:15:FC (Synology Incorporated)

Nmap done: 256 IP addresses (4 hosts up) scanned in 2.80 second

## #10: How do I perform a fast scan?

nmap -F 192.168.1.1

## #11: Display the reason a port is in a particular state

nmap --reason 192.168.1.1

nmap --reason server1.cyberciti.biz

## #12: Only show open (or possibly open) ports

nmap --open 192.168.1.1

nmap --open server1.cyberciti.biz

## #13: Show all packets sent and received

nmap --packet-trace 192.168.1.1

nmap --packet-trace server1.cyberciti.biz

## 14#: Show host interfaces and routes

This is useful for debugging ([ip command](http://www.cyberciti.biz/faq/howto-linux-configuring-default-route-with-ipcommand/) or [route command](http://www.cyberciti.biz/faq/what-is-a-routing-table/) or [netstat command](http://www.cyberciti.biz/faq/linux-unix-open-ports/) like output using nmap)

nmap --iflist

Sample outputs:

Starting Nmap 5.00 ( http://nmap.org ) at 2012-11-27 02:01 IST

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*INTERFACES\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

DEV (SHORT) IP/MASK TYPE UP MAC

lo (lo) 127.0.0.1/8 loopback up

eth0 (eth0) 192.168.1.5/24 ethernet up B8:AC:6F:65:31:E5

vmnet1 (vmnet1) 192.168.121.1/24 ethernet up 00:50:56:C0:00:01

vmnet8 (vmnet8) 192.168.179.1/24 ethernet up 00:50:56:C0:00:08

ppp0 (ppp0) 10.1.19.69/32 point2point up

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*ROUTES\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

DST/MASK DEV GATEWAY

10.0.31.178/32 ppp0

209.133.67.35/32 eth0 192.168.1.2

192.168.1.0/0 eth0

192.168.121.0/0 vmnet1

192.168.179.0/0 vmnet8

169.254.0.0/0 eth0

10.0.0.0/0 ppp0

0.0.0.0/0 eth0 192.168.1.2

## #15: How do I scan specific ports?

map -p **[**port**]** hostName

*## Scan port 80*

**nmap** -p 80 192.168.1.1

*## Scan TCP port 80*

**nmap** -p T:80 192.168.1.1

*## Scan UDP port 53*

**nmap** -p U:53 192.168.1.1

*## Scan two ports ##*

**nmap** -p 80,443 192.168.1.1

*## Scan port ranges ##*

**nmap** -p 80-200 192.168.1.1

*## Combine all options ##*

**nmap** -p U:53,111,137,T:21-25,80,139,8080 192.168.1.1

**nmap** -p U:53,111,137,T:21-25,80,139,8080 server1.cyberciti.biz

**nmap** -v -sU -sT -p U:53,111,137,T:21-25,80,139,8080 192.168.1.254

*## Scan all ports with \* wildcard ##*

**nmap** -p "\*" 192.168.1.1

*## Scan top ports i.e. scan $number most common ports ##*

**nmap** --top-ports 5 192.168.1.1

**nmap** --top-ports 10 192.168.1.1

Sample outputs:

Starting Nmap 5.00 ( http://nmap.org ) at 2012-11-27 01:23 IST

Interesting ports on 192.168.1.1:

PORT STATE SERVICE

21/tcp closed ftp

22/tcp open ssh

23/tcp closed telnet

25/tcp closed smtp

80/tcp open http

110/tcp closed pop3

139/tcp closed netbios-ssn

443/tcp closed https

445/tcp closed microsoft-ds

3389/tcp closed ms-term-serv

MAC Address: BC:AE:C5:C3:16:93 (Unknown)

Nmap done: 1 IP address (1 host up) scanned in 0.51 seconds

## #16: The fastest way to scan all your devices/computers for open ports ever

nmap -T5 192.168.1.0/24

## #17: How do I detect remote operating system?

You can [identify a remote host apps and OS using the -O option](http://www.cyberciti.biz/faq/identify-remote-host-by-unix-linux-command/):

**nmap** -O 192.168.1.1

**nmap** -O --osscan-guess 192.168.1.1

**nmap** -v -O --osscan-guess 192.168.1.1

Sample outputs:

Starting Nmap 5.00 ( http://nmap.org ) at 2012-11-27 01:29 IST

NSE: Loaded 0 scripts for scanning.

Initiating ARP Ping Scan at 01:29

Scanning 192.168.1.1 [1 port]

Completed ARP Ping Scan at 01:29, 0.01s elapsed (1 total hosts)

Initiating Parallel DNS resolution of 1 host. at 01:29

Completed Parallel DNS resolution of 1 host. at 01:29, 0.22s elapsed

Initiating SYN Stealth Scan at 01:29

Scanning 192.168.1.1 [1000 ports]

Discovered open port 80/tcp on 192.168.1.1

Discovered open port 22/tcp on 192.168.1.1

Completed SYN Stealth Scan at 01:29, 0.16s elapsed (1000 total ports)

Initiating OS detection (try #1) against 192.168.1.1

Retrying OS detection (try #2) against 192.168.1.1

Retrying OS detection (try #3) against 192.168.1.1

Retrying OS detection (try #4) against 192.168.1.1

Retrying OS detection (try #5) against 192.168.1.1

Host 192.168.1.1 is up (0.00049s latency).

Interesting ports on 192.168.1.1:

Not shown: 998 closed ports

PORT STATE SERVICE

22/tcp open ssh

80/tcp open http

MAC Address: BC:AE:C5:C3:16:93 (Unknown)

Device type: WAP|general purpose|router|printer|broadband router

Running (JUST GUESSING) : Linksys Linux 2.4.X (95%), Linux 2.4.X|2.6.X (94%), MikroTik RouterOS 3.X (92%), Lexmark embedded (90%), Enterasys embedded (89%), D-Link Linux 2.4.X (89%), Netgear Linux 2.4.X (89%)

Aggressive OS guesses: OpenWrt White Russian 0.9 (Linux 2.4.30) (95%), OpenWrt 0.9 - 7.09 (Linux 2.4.30 - 2.4.34) (94%), OpenWrt Kamikaze 7.09 (Linux 2.6.22) (94%), Linux 2.4.21 - 2.4.31 (likely embedded) (92%), Linux 2.6.15 - 2.6.23 (embedded) (92%), Linux 2.6.15 - 2.6.24 (92%), MikroTik RouterOS 3.0beta5 (92%), MikroTik RouterOS 3.17 (92%), Linux 2.6.24 (91%), Linux 2.6.22 (90%)

No exact OS matches for host (If you know what OS is running on it, see http://nmap.org/submit/ ).

TCP/IP fingerprint:

OS:SCAN(V=5.00%D=11/27%OT=22%CT=1%CU=30609%PV=Y%DS=1%G=Y%M=BCAEC5%TM=50B3CA

OS:4B%P=x86\_64-unknown-linux-gnu)SEQ(SP=C8%GCD=1%ISR=CB%TI=Z%CI=Z%II=I%TS=7

OS:)OPS(O1=M2300ST11NW2%O2=M2300ST11NW2%O3=M2300NNT11NW2%O4=M2300ST11NW2%O5

OS:=M2300ST11NW2%O6=M2300ST11)WIN(W1=45E8%W2=45E8%W3=45E8%W4=45E8%W5=45E8%W

OS:6=45E8)ECN(R=Y%DF=Y%T=40%W=4600%O=M2300NNSNW2%CC=N%Q=)T1(R=Y%DF=Y%T=40%S

OS:=O%A=S+%F=AS%RD=0%Q=)T2(R=N)T3(R=N)T4(R=Y%DF=Y%T=40%W=0%S=A%A=Z%F=R%O=%R

OS:D=0%Q=)T5(R=Y%DF=Y%T=40%W=0%S=Z%A=S+%F=AR%O=%RD=0%Q=)T6(R=Y%DF=Y%T=40%W=

OS:0%S=A%A=Z%F=R%O=%RD=0%Q=)T7(R=N)U1(R=Y%DF=N%T=40%IPL=164%UN=0%RIPL=G%RID

OS:=G%RIPCK=G%RUCK=G%RUD=G)IE(R=Y%DFI=N%T=40%CD=S)

Uptime guess: 12.990 days (since Wed Nov 14 01:44:40 2012)

Network Distance: 1 hop

TCP Sequence Prediction: Difficulty=200 (Good luck!)

IP ID Sequence Generation: All zeros

Read data files from: /usr/share/nmap

OS detection performed. Please report any incorrect results at http://nmap.org/submit/ .

Nmap done: 1 IP address (1 host up) scanned in 12.38 seconds

Raw packets sent: 1126 (53.832KB) | Rcvd: 1066 (46.100KB)

See also: [Fingerprinting a web-server](http://www.cyberciti.biz/faq/find-out-remote-webserver-name/) and a [dns server](http://www.cyberciti.biz/tips/howto-remotely-determine-dns-server-version.html) command line tools for more information.

## #18: How do I detect remote services (server / daemon) version numbers?

nmap -sV 192.168.1.1

Sample outputs:

Starting Nmap 5.00 ( http://nmap.org ) at 2012-11-27 01:34 IST

Interesting ports on 192.168.1.1:

Not shown: 998 closed ports

PORT STATE SERVICE VERSION

22/tcp open ssh **Dropbear sshd 0.52 (protocol 2.0)**

80/tcp open http?

1 service unrecognized despite returning data.

## #19: Scan a host using TCP ACK (PA) and TCP Syn (PS) ping

If firewall is blocking standard ICMP pings, try the following host discovery methods:

nmap -PS 192.168.1.1

nmap -PS 80,21,443 192.168.1.1

nmap -PA 192.168.1.1

nmap -PA 80,21,200-512 192.168.1.1

## #20: Scan a host using IP protocol ping

nmap -PO 192.168.1.1

## #21: Scan a host using UDP ping

This scan bypasses firewalls and filters that only screen TCP:

nmap -PU 192.168.1.1

nmap -PU 2000.2001 192.168.1.1

## #22: Find out the most commonly used TCP ports using TCP SYN Scan

*### Stealthy scan ###*

**nmap** -sS 192.168.1.1

*### Find out the most commonly used TCP ports using TCP connect scan* ***(****warning: no stealth scan****)***

*### OS Fingerprinting ###*

**nmap** -sT 192.168.1.1

*### Find out the most commonly used TCP ports using TCP ACK scan*

**nmap** -sA 192.168.1.1

*### Find out the most commonly used TCP ports using TCP Window scan*

**nmap** -sW 192.168.1.1

*### Find out the most commonly used TCP ports using TCP Maimon scan*

**nmap** -sM 192.168.1.1

## #23: Scan a host for UDP services (UDP scan)

Most popular services on the Internet run over the TCP protocol. DNS, SNMP, and DHCP are three of the most common UDP services. Use the following syntax to find out UDP services:

nmap -sU nas03

nmap -sU 192.168.1.1

Sample outputs:

Starting Nmap 5.00 ( http://nmap.org ) at 2012-11-27 00:52 IST

Stats: 0:05:29 elapsed*; 0 hosts completed (1 up), 1 undergoing UDP Scan*

UDP Scan Timing: About 32.49% done*; ETC: 01:09 (0:11:26 remaining)*

Interesting ports on nas03 (192.168.1.12):

Not shown: 995 closed ports

PORT STATE SERVICE

111/udp open|filtered rpcbind

123/udp open|filtered ntp

161/udp open|filtered snmp

2049/udp open|filtered nfs

5353/udp open|filtered zeroconf

MAC Address: 00:11:32:11:15:FC (Synology Incorporated)

Nmap done: 1 IP address (1 host up) scanned in 1099.55 seconds

## #24: Scan for IP protocol

This type of scan allows you to determine which IP protocols (TCP, ICMP, IGMP, etc.) are supported by target machines:

nmap -sO 192.168.1.1

## #25: Scan a firewall for security weakness

The following scan types exploit a subtle loophole in the TCP and good for testing security of common attacks:

*## TCP Null Scan to fool a firewall to generate a response ##*

*## Does not* ***set*** *any bits* ***(****TCP flag header is 0****)*** *##*

**nmap** -sN 192.168.1.254

*## TCP Fin scan to check firewall ##*

*## Sets just the TCP FIN bit ##*

**nmap** -sF 192.168.1.254

*## TCP Xmas scan to check firewall ##*

*## Sets the FIN, PSH, and URG flags, lighting the packet up like a Christmas tree ##*

**nmap** -sX 192.168.1.254

See [how to block Xmas packkets, syn-floods and other conman attacks](http://www.cyberciti.biz/tips/linux-iptables-10-how-to-block-common-attack.html) with iptables.

## #26: Scan a firewall for packets fragments

The -f option causes the requested scan (including ping scans) to use tiny fragmented IP packets. The idea is to split up the TCP header over  
several packets to make it harder for packet filters, intrusion detection systems, and other annoyances to detect what you are doing.

nmap -f 192.168.1.1

nmap -f fw2.nixcraft.net.in

nmap -f 15 fw2.nixcraft.net.in

## Set your own offset size with the --mtu option ##

nmap --mtu 32 192.168.1.1

## #27: Cloak a scan with decoys

The -D option it appear to the remote host that the host(s) you specify as [decoys are scanning the target network too](http://www.cyberciti.biz/tips/nmap-hide-ipaddress-with-decoy-ideal-scan.html). Thus their IDS might report 5-10 port scans from unique IP addresses, but they won't know which IP was scanning them and which were innocent decoys:

nmap -n -Ddecoy-ip1,decoy-ip2,your-own-ip,decoy-ip3,decoy-ip4 remote-host-ip

nmap -n -D192.168.1.5,10.5.1.2,172.1.2.4,3.4.2.1 192.168.1.5

## #28: Scan a firewall for MAC address spoofing

*### Spoof your MAC address ##*

**nmap** --spoof-mac MAC-ADDRESS-HERE 192.168.1.1

*### Add other options ###*

**nmap** -v -sT -PN --spoof-mac MAC-ADDRESS-HERE 192.168.1.1

*### Use a random MAC address ###*

*### The number 0, means* ***nmap*** *chooses a completely random MAC address ###*

**nmap** -v -sT -PN --spoof-mac 0 192.168.1.1

## #29: How do I save output to a text file?

The syntax is:

nmap 192.168.1.1 > output.txt

nmap -oN /path/to/filename 192.168.1.1

nmap -oN output.txt 192.168.1.1

## #30: Not a fan of command line tools?

Try [zenmap the official network mapper](http://nmap.org/zenmap/) front end:

Zenmap is the official Nmap Security Scanner GUI. It is a multi-platform (Linux, Windows, Mac OS X, BSD, etc.) free and open source application which aims to make Nmap easy for beginners to use while providing advanced features for experienced Nmap users. Frequently used scans can be saved as profiles to make them easy to run repeatedly. A command creator allows interactive creation of Nmap command lines. Scan results can be saved and viewed later. Saved scan results can be compared with one another to see how they differ. The results of recent scans are stored in a searchable database.

You can install zenmap using the following [apt-get command](http://www.cyberciti.biz/tips/linux-debian-package-management-cheat-sheet.html):  
$ sudo apt-get install zenmap  
Sample outputs:

[sudo] password for vivek:

Reading package lists... Done

Building dependency tree

Reading state information... Done

The following NEW packages will be installed:

zenmap

0 upgraded, 1 newly installed, 0 to remove and 11 not upgraded.

Need to get 616 kB of archives.

After this operation, 1,827 kB of additional disk space will be used.

Get:1 http://debian.osuosl.org/debian/ squeeze/main zenmap amd64 5.00-3 [616 kB]

Fetched 616 kB in 3s (199 kB/s)

Selecting previously deselected package zenmap.

(Reading database ... 281105 files and directories currently installed.)

Unpacking zenmap (from .../zenmap\_5.00-3\_amd64.deb) ...

Processing triggers for desktop-file-utils ...

Processing triggers for gnome-menus ...

Processing triggers for man-db ...

Setting up zenmap (5.00-3) ...

Processing triggers for python-central ...

Type the following command to start zenmap:  
$ sudo zenmap  
Sample outputs

# RHEL: Linux Bond / Team Multiple Network Interfaces (NIC) Into a Single Interface

by [nixCraft](http://www.cyberciti.biz/tips/about-us) on April 3, 2006 · [116 comments](http://www.cyberciti.biz/tips/linux-bond-or-team-multiple-network-interfaces-nic-into-single-interface.html#comments)· LAST UPDATED September 4, 2011

in [fedora linux](http://www.cyberciti.biz/tips/category/fedora-linux), [Howto](http://www.cyberciti.biz/tips/category/howto), [Networking](http://www.cyberciti.biz/tips/category/networking)

[http://files.cyberciti.biz/cbzcache/3rdparty/rhlogo.gif](http://www.cyberciti.biz/tips/category/redhatfedora-linux)

Finally, today I had implemented NIC bounding (bind both NIC so that it works as a single device). Bonding is nothing but Linux kernel feature that allows to aggregate multiple like interfaces (such as eth0, eth1) into a single virtual link such as bond0. The idea is pretty simple get higher data rates and as well as link failover. The following instructions were tested on:

1. RHEL v4 / 5 / 6 amd64
2. CentOS v5 / 6 amd64
3. Fedora Linux 13 amd64 and up.
4. 2 x PCI-e Gigabit Ethernet NICs with Jumbo Frames (MTU 9000)
5. Hardware RAID-10 w/ SAS 15k enterprise grade hard disks.
6. Gigabit switch with Jumbo Frame

[](http://www.cyberciti.biz/tips/category/fedora-linux)

This server act as an heavy duty ftp, and nfs file server. Each, night a perl script will transfer lots of data from this box to a backup server. Therefore, the network would be setup on a switch using dual network cards. I am using Red Hat enterprise Linux version 4.0. But, the inductions should work on RHEL 5 and 6 too.

## Say Hello To bounding Driver

Linux allows binding of multiple network interfaces into a single channel/NIC using special kernel module called bonding. According to official bonding [documentation](http://www.cyberciti.biz/howto/question/static/linux-ethernet-bonding-driver-howto.php):

The Linux bonding driver provides a method for aggregating multiple network interfaces into a single logical "bonded" interface. The behavior of the bonded interfaces depends upon the mode; generally speaking, modes provide either hot standby or load balancing services. Additionally, link integrity monitoring may be performed.

## Step #1: Create a Bond0 Configuration File

Red Hat Enterprise Linux (and its clone such as CentOS) stores network configuration in /etc/sysconfig/network-scripts/ directory. First, you need to create a bond0 config file as follows:  
# vi /etc/sysconfig/network-scripts/ifcfg-bond0  
Append the following linest:

DEVICE=bond0

IPADDR=192.168.1.20

NETWORK=192.168.1.0

NETMASK=255.255.255.0

USERCTL=no

BOOTPROTO=none

ONBOOT=yes

You need to replace IP address with your actual setup. Save and close the file.

## Step #2: Modify eth0 and eth1 config files

Open both configuration using a text editor such as vi/vim, and make sure file read as follows for eth0 interface  
# vi /etc/sysconfig/network-scripts/ifcfg-eth0  
Modify/append directive as follows:  
DEVICE=eth0  
USERCTL=no  
ONBOOT=yes  
MASTER=bond0  
SLAVE=yes  
BOOTPROTO=none  
Open eth1 configuration file using vi text editor, enter:  
# vi /etc/sysconfig/network-scripts/ifcfg-eth1  
Make sure file read as follows for eth1 interface:  
DEVICE=eth1  
USERCTL=no  
ONBOOT=yes  
MASTER=bond0  
SLAVE=yes  
BOOTPROTO=none  
Save and close the file.

## Step # 3: Load bond driver/module

Make sure bonding module is loaded when the channel-bonding interface (bond0) is brought up. You need to modify kernel modules configuration file:# vi /etc/modprobe.conf  
Append following two lines:alias bond0 bonding  
options bond0 mode=balance-alb miimon=100  
Save file and exit to shell prompt. You can learn more about all bounding options by clicking [here](http://www.cyberciti.biz/howto/question/static/linux-ethernet-bonding-driver-howto.php#section_4)).

### Step # 4: Test configuration

First, load the bonding module, enter:  
# modprobe bonding  
Restart the networking service in order to bring up bond0 interface, enter:  
# service network restart  
Make sure everything is working. Type the following [cat command](http://www.cyberciti.biz/faq/howto-use-cat-command-in-unix-linux-shell-script/) to query the current status of Linux kernel bounding driver, enter:  
# [cat /proc/net/bonding/bond0](http://www.cyberciti.biz/faq/howto-use-cat-command-in-unix-linux-shell-script/)  
Sample outputs:

Bonding Mode: load balancing (round-robin)

MII Status: up

MII Polling Interval (ms): 100

Up Delay (ms): 200

Down Delay (ms): 200

Slave Interface: eth0

MII Status: up

Link Failure Count: 0

Permanent HW addr: 00:0c:29:c6:be:59

Slave Interface: eth1

MII Status: up

Link Failure Count: 0

Permanent HW addr: 00:0c:29:c6:be:63

To kist all network interfaces, enter:  
# ifconfig  
Sample outputs:

bond0 Link encap:Ethernet HWaddr 00:0C:29:C6:BE:59

inet addr:192.168.1.20 Bcast:192.168.1.255 Mask:255.255.255.0

inet6 addr: fe80::200:ff:fe00:0/64 Scope:Link

UP BROADCAST RUNNING MASTER MULTICAST MTU:1500 Metric:1

RX packets:2804 errors:0 dropped:0 overruns:0 frame:0

TX packets:1879 errors:0 dropped:0 overruns:0 carrier:0

collisions:0 txqueuelen:0

RX bytes:250825 (244.9 KiB) TX bytes:244683 (238.9 KiB)

eth0 Link encap:Ethernet HWaddr 00:0C:29:C6:BE:59

inet addr:192.168.1.20 Bcast:192.168.1.255 Mask:255.255.255.0

inet6 addr: fe80::20c:29ff:fec6:be59/64 Scope:Link

UP BROADCAST RUNNING SLAVE MULTICAST MTU:1500 Metric:1

RX packets:2809 errors:0 dropped:0 overruns:0 frame:0

TX packets:1390 errors:0 dropped:0 overruns:0 carrier:0

collisions:0 txqueuelen:1000

RX bytes:251161 (245.2 KiB) TX bytes:180289 (176.0 KiB)

Interrupt:11 Base address:0x1400

eth1 Link encap:Ethernet HWaddr 00:0C:29:C6:BE:59

inet addr:192.168.1.20 Bcast:192.168.1.255 Mask:255.255.255.0

inet6 addr: fe80::20c:29ff:fec6:be59/64 Scope:Link

UP BROADCAST RUNNING SLAVE MULTICAST MTU:1500 Metric:1

RX packets:4 errors:0 dropped:0 overruns:0 frame:0

TX packets:502 errors:0 dropped:0 overruns:0 carrier:0

collisions:0 txqueuelen:1000

RX bytes:258 (258.0 b) TX bytes:66516 (64.9 KiB)

Interrupt:10 Base address:0x1480

Linux Ethernet Bonding Driver HOWTO

Latest update: 21 June 2005

Initial release : Thomas Davis <tadavis at lbl.gov> Corrections, HA extensions : 2000/10/03-15 :

* Willy Tarreau <willy at meta-x.org>
* Constantine Gavrilov <const-g at xpert.com>
* Chad N. Tindel <ctindel at ieee dot org>
* Janice Girouard <girouard at us dot ibm dot com>
* Jay Vosburgh <fubar at us dot ibm dot com>

Reorganized and updated Feb 2005 by Jay Vosburgh

# Introduction

The Linux bonding driver provides a method for aggregating multiple network interfaces into a single logical "bonded" interface. The behavior of the bonded interfaces depends upon the mode; generally speaking, modes provide either hot standby or load balancing services. Additionally, link integrity monitoring may be performed.

The bonding driver originally came from Donald Becker's beowulf patches for kernel 2.0. It has changed quite a bit since, and the original tools from extreme-linux and beowulf sites will not work with this version of the driver.

For new versions of the driver, updated userspace tools, and who to ask for help, please follow the links at the end of this file.

# Table of Contents

1. Bonding Driver Installation
2. Bonding Driver Options
3. Configuring Bonding Devices 3.1 Configuration with sysconfig support 3.1.1 Using DHCP with sysconfig 3.1.2 Configuring Multiple Bonds with sysconfig 3.2 Configuration with initscripts support 3.2.1 Using DHCP with initscripts 3.2.2 Configuring Multiple Bonds with initscripts 3.3 Configuring Bonding Manually 3.3.1 Configuring Multiple Bonds Manually
4. Querying Bonding Configuration 5.1 Bonding Configuration 5.2 Network Configuration
5. Switch Configuration
6. 802.1q VLAN Support
7. Link Monitoring 8.1 ARP Monitor Operation 8.2 Configuring Multiple ARP Targets 8.3 MII Monitor Operation
8. Potential Trouble Sources 9.1 Adventures in Routing 9.2 Ethernet Device Renaming 9.3 Painfully Slow Or No Failed Link Detection By Miimon
9. SNMP agents
10. Promiscuous mode
11. Configuring Bonding for High Availability 12.1 High Availability in a Single Switch Topology 12.2 High Availability in a Multiple Switch Topology 12.2.1 HA Bonding Mode Selection for Multiple Switch Topology 12.2.2 HA Link Monitoring for Multiple Switch Topology
12. Configuring Bonding for Maximum Throughput 13.1 Maximum Throughput in a Single Switch Topology 13.1.1 MT Bonding Mode Selection for Single Switch Topology 13.1.2 MT Link Monitoring for Single Switch Topology 13.2 Maximum Throughput in a Multiple Switch Topology 13.2.1 MT Bonding Mode Selection for Multiple Switch Topology 13.2.2 MT Link Monitoring for Multiple Switch Topology
13. Switch Behavior Issues 14.1 Link Establishment and Failover Delays 14.2 Duplicated Incoming Packets
14. Hardware Specific Considerations 15.1 IBM BladeCenter
15. Frequently Asked Questions
16. Resources and Links

# 1. Bonding Driver Installation

Most popular distro kernels ship with the bonding driver already available as a module and the ifenslave user level control program installed and ready for use. If your distro does not, or you have need to compile bonding from source (e.g., configuring and installing a mainline kernel from kernel.org), you'll need to perform the following steps:

## 1.1 Configure and build the kernel with bonding

The current version of the bonding driver is available in the drivers/net/bonding subdirectory of the most recent kernel source (which is available on <http://kernel.org>). Most users "rolling their own" will want to use the most recent kernel from kernel.org.

Configure kernel with "make menuconfig" (or "make xconfig" or "make config"), then select "Bonding driver support" in the "Network device support" section. It is recommended that you configure the driver as module since it is currently the only way to pass parameters to the driver or configure more than one bonding device.

Build and install the new kernel and modules, then continue below to install ifenslave.

## 1.2 Install ifenslave Control Utility

The ifenslave user level control program is included in the kernel source tree, in the file Documentation/networking/ifenslave.c. It is generally recommended that you use the ifenslave that corresponds to the kernel that you are using (either from the same source tree or supplied with the distro), however, ifenslave executables from older kernels should function (but features newer than the ifenslave release are not supported). Running an ifenslave that is newer than the kernel is not supported, and may or may not work.

To install ifenslave, do the following:

# gcc -Wall -O -I/usr/src/linux/include ifenslave.c -o ifenslave # cp ifenslave /sbin/ifenslave

If your kernel source is not in "/usr/src/linux," then replace "/usr/src/linux/include" in the above with the location of your kernel source include directory.

You may wish to back up any existing /sbin/ifenslave, or, for testing or informal use, tag the ifenslave to the kernel version (e.g., name the ifenslave executable /sbin/ifenslave-2.6.10).

**IMPORTANT NOTE:**

If you omit the "-I" or specify an incorrect directory, you may end up with an ifenslave that is incompatible with the kernel you're trying to build it for. Some distros (e.g., Red Hat from 7.1 onwards) do not have /usr/include/linux symbolically linked to the default kernel source include directory.

# 2. Bonding Driver Options

Options for the bonding driver are supplied as parameters to the bonding module at load time. They may be given as command line arguments to the insmod or modprobe command, but are usually specified in either the /etc/modules.conf or /etc/modprobe.conf configuration file, or in a distro-specific configuration file (some of which are detailed in the next section).

The available bonding driver parameters are listed below. If a parameter is not specified the default value is used. When initially configuring a bond, it is recommended "tail -f /var/log/messages" be run in a separate window to watch for bonding driver error messages.

It is critical that either the miimon or arp\_interval and arp\_ip\_target parameters be specified, otherwise serious network degradation will occur during link failures. Very few devices do not support at least miimon, so there is really no reason not to use it.

Options with textual values will accept either the text name or, for backwards compatibility, the option value. E.g., "mode=802.3ad" and "mode=4" set the same mode.

The parameters are as follows:

arp\_interval

Specifies the ARP link monitoring frequency in milliseconds.

If ARP monitoring is used in an etherchannel compatible mode

(modes 0 and 2), the switch should be configured in a mode

that evenly distributes packets across all links. If the

switch is configured to distribute the packets in an XOR

fashion, all replies from the ARP targets will be received on

the same link which could cause the other team members to

fail. ARP monitoring should not be used in conjunction with

miimon. A value of 0 disables ARP monitoring. The default

value is 0.

arp\_ip\_target

Specifies the IP addresses to use as ARP monitoring peers when

arp\_interval is > 0. These are the targets of the ARP request

sent to determine the health of the link to the targets.

Specify these values in ddd.ddd.ddd.ddd format. Multiple IP

addresses must be separated by a comma. At least one IP

address must be given for ARP monitoring to function. The

maximum number of targets that can be specified is 16. The

default value is no IP addresses.

downdelay

Specifies the time, in milliseconds, to wait before disabling

a slave after a link failure has been detected. This option

is only valid for the miimon link monitor. The downdelay

value should be a multiple of the miimon value; if not, it

will be rounded down to the nearest multiple. The default

value is 0.

lacp\_rate

Option specifying the rate in which we'll ask our link partner

to transmit LACPDU packets in 802.3ad mode. Possible values

are:

slow or 0

Request partner to transmit LACPDUs every 30 seconds

fast or 1

Request partner to transmit LACPDUs every 1 second

The default is slow.

max\_bonds

Specifies the number of bonding devices to create for this

instance of the bonding driver. E.g., if max\_bonds is 3, and

the bonding driver is not already loaded, then bond0, bond1

and bond2 will be created. The default value is 1.

miimon

Specifies the MII link monitoring frequency in milliseconds.

This determines how often the link state of each slave is

inspected for link failures. A value of zero disables MII

link monitoring. A value of 100 is a good starting point.

The use\_carrier option, below, affects how the link state is

determined. See the High Availability section for additional

information. The default value is 0.

mode

Specifies one of the bonding policies. The default is

balance-rr (round robin). Possible values are:

balance-rr or 0

Round-robin policy: Transmit packets in sequential

order from the first available slave through the

last. This mode provides load balancing and fault

tolerance.

active-backup or 1

Active-backup policy: Only one slave in the bond is

active. A different slave becomes active if, and only

if, the active slave fails. The bond's MAC address is

externally visible on only one port (network adapter)

to avoid confusing the switch.

In bonding version 2.6.2 or later, when a failover

occurs in active-backup mode, bonding will issue one

or more gratuitous ARPs on the newly active slave.

One gratutious ARP is issued for the bonding master

interface and each VLAN interfaces configured above

it, provided that the interface has at least one IP

address configured. Gratuitous ARPs issued for VLAN

interfaces are tagged with the appropriate VLAN id.

This mode provides fault tolerance. The primary

option, documented below, affects the behavior of this

mode.

balance-xor or 2

XOR policy: Transmit based on the selected transmit

hash policy. The default policy is a simple [(source

MAC address XOR'd with destination MAC address) modulo

slave count]. Alternate transmit policies may be

selected via the xmit\_hash\_policy option, described

below.

This mode provides load balancing and fault tolerance.

broadcast or 3

Broadcast policy: transmits everything on all slave

interfaces. This mode provides fault tolerance.

802.3ad or 4

IEEE 802.3ad Dynamic link aggregation. Creates

aggregation groups that share the same speed and

duplex settings. Utilizes all slaves in the active

aggregator according to the 802.3ad specification.

Slave selection for outgoing traffic is done according

to the transmit hash policy, which may be changed from

the default simple XOR policy via the xmit\_hash\_policy

option, documented below. Note that not all transmit

policies may be 802.3ad compliant, particularly in

regards to the packet mis-ordering requirements of

section 43.2.4 of the 802.3ad standard. Differing

peer implementations will have varying tolerances for

noncompliance.

Prerequisites:

1. Ethtool support in the base drivers for retrieving

the speed and duplex of each slave.

2. A switch that supports IEEE 802.3ad Dynamic link

aggregation.

Most switches will require some type of configuration

to enable 802.3ad mode.

balance-tlb or 5

Adaptive transmit load balancing: channel bonding that

does not require any special switch support. The

outgoing traffic is distributed according to the

current load (computed relative to the speed) on each

slave. Incoming traffic is received by the current

slave. If the receiving slave fails, another slave

takes over the MAC address of the failed receiving

slave.

Prerequisite:

Ethtool support in the base drivers for retrieving the

speed of each slave.

balance-alb or 6

Adaptive load balancing: includes balance-tlb plus

receive load balancing (rlb) for IPV4 traffic, and

does not require any special switch support. The

receive load balancing is achieved by ARP negotiation.

The bonding driver intercepts the ARP Replies sent by

the local system on their way out and overwrites the

source hardware address with the unique hardware

address of one of the slaves in the bond such that

different peers use different hardware addresses for

the server.

Receive traffic from connections created by the server

is also balanced. When the local system sends an ARP

Request the bonding driver copies and saves the peer's

IP information from the ARP packet. When the ARP

Reply arrives from the peer, its hardware address is

retrieved and the bonding driver initiates an ARP

reply to this peer assigning it to one of the slaves

in the bond. A problematic outcome of using ARP

negotiation for balancing is that each time that an

ARP request is broadcast it uses the hardware address

of the bond. Hence, peers learn the hardware address

of the bond and the balancing of receive traffic

collapses to the current slave. This is handled by

sending updates (ARP Replies) to all the peers with

their individually assigned hardware address such that

the traffic is redistributed. Receive traffic is also

redistributed when a new slave is added to the bond

and when an inactive slave is re-activated. The

receive load is distributed sequentially (round robin)

among the group of highest speed slaves in the bond.

When a link is reconnected or a new slave joins the

bond the receive traffic is redistributed among all

active slaves in the bond by initiating ARP Replies

with the selected mac address to each of the

clients. The updelay parameter (detailed below) must

be set to a value equal or greater than the switch's

forwarding delay so that the ARP Replies sent to the

peers will not be blocked by the switch.

Prerequisites:

1. Ethtool support in the base drivers for retrieving

the speed of each slave.

2. Base driver support for setting the hardware

address of a device while it is open. This is

required so that there will always be one slave in the

team using the bond hardware address (the

curr\_active\_slave) while having a unique hardware

address for each slave in the bond. If the

curr\_active\_slave fails its hardware address is

swapped with the new curr\_active\_slave that was

chosen.

primary

A string (eth0, eth2, etc) specifying which slave is the

primary device. The specified device will always be the

active slave while it is available. Only when the primary is

off-line will alternate devices be used. This is useful when

one slave is preferred over another, e.g., when one slave has

higher throughput than another.

The primary option is only valid for active-backup mode.

updelay

Specifies the time, in milliseconds, to wait before enabling a

slave after a link recovery has been detected. This option is

only valid for the miimon link monitor. The updelay value

should be a multiple of the miimon value; if not, it will be

rounded down to the nearest multiple. The default value is 0.

use\_carrier

Specifies whether or not miimon should use MII or ETHTOOL

ioctls vs. netif\_carrier\_ok() to determine the link

status. The MII or ETHTOOL ioctls are less efficient and

utilize a deprecated calling sequence within the kernel. The

netif\_carrier\_ok() relies on the device driver to maintain its

state with netif\_carrier\_on/off; at this writing, most, but

not all, device drivers support this facility.

If bonding insists that the link is up when it should not be,

it may be that your network device driver does not support

netif\_carrier\_on/off. The default state for netif\_carrier is

"carrier on," so if a driver does not support netif\_carrier,

it will appear as if the link is always up. In this case,

setting use\_carrier to 0 will cause bonding to revert to the

MII / ETHTOOL ioctl method to determine the link state.

A value of 1 enables the use of netif\_carrier\_ok(), a value of

0 will use the deprecated MII / ETHTOOL ioctls. The default

value is 1.

xmit\_hash\_policy

Selects the transmit hash policy to use for slave selection in

balance-xor and 802.3ad modes. Possible values are:

layer2

Uses XOR of hardware MAC addresses to generate the

hash. The formula is

(source MAC XOR destination MAC) modulo slave count

This algorithm will place all traffic to a particular

network peer on the same slave.

This algorithm is 802.3ad compliant.

layer3+4

This policy uses upper layer protocol information,

when available, to generate the hash. This allows for

traffic to a particular network peer to span multiple

slaves, although a single connection will not span

multiple slaves.

The formula for unfragmented TCP and UDP packets is

((source port XOR dest port) XOR

((source IP XOR dest IP) AND 0xffff)

modulo slave count

For fragmented TCP or UDP packets and all other IP

protocol traffic, the source and destination port

information is omitted. For non-IP traffic, the

formula is the same as for the layer2 transmit hash

policy.

This policy is intended to mimic the behavior of

certain switches, notably Cisco switches with PFC2 as

well as some Foundry and IBM products.

This algorithm is not fully 802.3ad compliant. A

single TCP or UDP conversation containing both

fragmented and unfragmented packets will see packets

striped across two interfaces. This may result in out

of order delivery. Most traffic types will not meet

this criteria, as TCP rarely fragments traffic, and

most UDP traffic is not involved in extended

conversations. Other implementations of 802.3ad may

or may not tolerate this noncompliance.

The default value is layer2. This option was added in bonding

version 2.6.3. In earlier versions of bonding, this parameter does not exist, and the layer2 policy is the only policy.

# 3. Configuring Bonding Devices

There are, essentially, two methods for configuring bonding: with support from the distro's network initialization scripts, and without. Distros generally use one of two packages for the network initialization scripts: initscripts or sysconfig. Recent versions of these packages have support for bonding, while older versions do not.

We will first describe the options for configuring bonding for distros using versions of initscripts and sysconfig with full or partial support for bonding, then provide information on enabling bonding without support from the network initialization scripts (i.e., older versions of initscripts or sysconfig).

If you're unsure whether your distro uses sysconfig or initscripts, or don't know if it's new enough, have no fear. Determining this is fairly straightforward.

First, issue the command:

$ rpm -qf /sbin/ifup

It will respond with a line of text starting with either "initscripts" or "sysconfig," followed by some numbers. This is the package that provides your network initialization scripts.

Next, to determine if your installation supports bonding, issue the command:

$ grep ifenslave /sbin/ifup

If this returns any matches, then your initscripts or sysconfig has support for bonding.

## 3.1 Configuration with sysconfig support

This section applies to distros using a version of sysconfig with bonding support, for example, SuSE Linux Enterprise Server 9.

SuSE SLES 9's networking configuration system does support bonding, however, at this writing, the YaST system configuration frontend does not provide any means to work with bonding devices. Bonding devices can be managed by hand, however, as follows.

First, if they have not already been configured, configure the slave devices. On SLES 9, this is most easily done by running the yast2 sysconfig configuration utility. The goal is for to create an ifcfg-id file for each slave device. The simplest way to accomplish this is to configure the devices for DHCP (this is only to get the file ifcfg-id file created; see below for some issues with DHCP). The name of the configuration file for each device will be of the form:

ifcfg-id-xx:xx:xx:xx:xx:xx

Where the "xx" portion will be replaced with the digits from the device's permanent MAC address.

Once the set of ifcfg-id-xx:xx:xx:xx:xx:xx files has been created, it is necessary to edit the configuration files for the slave devices (the MAC addresses correspond to those of the slave devices). Before editing, the file will contain multiple lines, and will look something like this:

BOOTPROTO='dhcp'  
STARTMODE='on'  
USERCTL='no'  
UNIQUE='XNzu.WeZGOGF+4wE'  
\_nm\_name='bus-pci-0001:61:01.0'

Change the BOOTPROTO and STARTMODE lines to the following:

BOOTPROTO='none'  
STARTMODE='off'

Do not alter the UNIQUE or \_nm\_name lines. Remove any other lines (USERCTL, etc).

Once the ifcfg-id-xx:xx:xx:xx:xx:xx files have been modified, it's time to create the configuration file for the bonding device itself. This file is named ifcfg-bondX, where X is the number of the bonding device to create, starting at 0. The first such file is ifcfg-bond0, the second is ifcfg-bond1, and so on. The sysconfig network configuration system will correctly start multiple instances of bonding.

The contents of the ifcfg-bondX file is as follows:

BOOTPROTO="static"  
**BROADCAST="**[**10.0.2.255**](telnet://10.0.2.255)**"**  
**IPADDR="**[**10.0.2.10**](telnet://10.0.2.10)**"**  
**NETMASK="**[**255.255.0.0**](telnet://255.255.0.0)**"**  
**NETWORK="**[**10.0.2.0**](telnet://10.0.2.0)**"**  
**REMOTE\_IPADDR=""**  
STARTMODE="onboot"  
BONDING\_MASTER="yes"  
BONDING\_MODULE\_OPTS="mode=active-backup miimon=100" BONDING\_SLAVE0="eth0"  
BONDING\_SLAVE1="bus-pci-0000:06:08.1"

Replace the sample BROADCAST, IPADDR, NETMASK and NETWORK values with the appropriate values for your network.

The STARTMODE specifies when the device is brought online. The possible values are:

onboot: The device is started at boot time. If you're not

sure, this is probably what you want.

manual: The device is started only when ifup is called

manually. Bonding devices may be configured this

way if you do not wish them to start automatically

at boot for some reason.

hotplug: The device is started by a hotplug event. This is not

a valid choice for a bonding device.

off or ignore: The device configuration is ignored.

The line BONDING\_MASTER='yes' indicates that the device is a

bonding master device. The only useful value is "yes."

The contents of BONDING\_MODULE\_OPTS are supplied to the instance of the bonding module for this device. Specify the options for the bonding mode, link monitoring, and so on here. Do not include the max\_bonds bonding parameter; this will confuse the configuration system if you have multiple bonding devices.

Finally, supply one BONDING\_SLAVEn="slave device" for each slave. where "n" is an increasing value, one for each slave. The "slave device" is either an interface name, e.g., "eth0", or a device specifier for the network device. The interface name is easier to find, but the ethN names are subject to change at boot time if, e.g., a device early in the sequence has failed. The device specifiers (bus-pci-0000:06:08.1 in the example above) specify the physical network device, and will not change unless the device's bus location changes (for example, it is moved from one PCI slot to another). The example above uses one of each type for demonstration purposes; most configurations will choose one or the other for all slave devices.

When all configuration files have been modified or created, networking must be restarted for the configuration changes to take effect. This can be accomplished via the following:

# /etc/init.d/network restart

Note that the network control script (/sbin/ifdown) will remove the bonding module as part of the network shutdown processing, so it is not necessary to remove the module by hand if, e.g., the module parameters have changed.

Also, at this writing, YaST/YaST2 will not manage bonding devices (they do not show bonding interfaces on its list of network devices). It is necessary to edit the configuration file by hand to change the bonding configuration.

Additional general options and details of the ifcfg file format can be found in an example ifcfg template file:

/etc/sysconfig/network/ifcfg.template

Note that the template does not document the various BONDING\_ settings described above, but does describe many of the other options.

## 3.1.1 Using DHCP with sysconfig

Under sysconfig, configuring a device with BOOTPROTO='dhcp' will cause it to query DHCP for its IP address information. At this writing, this does not function for bonding devices; the scripts attempt to obtain the device address from DHCP prior to adding any of the slave devices. Without active slaves, the DHCP requests are not sent to the network.

## 3.1.2 Configuring Multiple Bonds with sysconfig

The sysconfig network initialization system is capable of handling multiple bonding devices. All that is necessary is for each bonding instance to have an appropriately configured ifcfg-bondX file (as described above). Do not specify the "max\_bonds" parameter to any instance of bonding, as this will confuse sysconfig. If you require multiple bonding devices with identical parameters, create multiple ifcfg-bondX files.

Because the sysconfig scripts supply the bonding module options in the ifcfg-bondX file, it is not necessary to add them to the system /etc/modules.conf or /etc/modprobe.conf configuration file.

## 3.2 Configuration with initscripts support

This section applies to distros using a version of initscripts with bonding support, for example, Red Hat Linux 9 or Red Hat Enterprise Linux version 3 or 4. On these systems, the network initialization scripts have some knowledge of bonding, and can be configured to control bonding devices.

These distros will not automatically load the network adapter driver unless the ethX device is configured with an IP address. Because of this constraint, users must manually configure a network-script file for all physical adapters that will be members of a bondX link. Network script files are located in the directory:

/etc/sysconfig/network-scripts

The file name must be prefixed with "ifcfg-eth" and suffixed with the adapter's physical adapter number. For example, the script for eth0 would be named /etc/sysconfig/network-scripts/ifcfg-eth0. Place the following text in the file:

DEVICE=eth0  
USERCTL=no  
ONBOOT=yes  
MASTER=bond0  
SLAVE=yes  
BOOTPROTO=none

The DEVICE= line will be different for every ethX device and must correspond with the name of the file, i.e., ifcfg-eth1 must have a device line of DEVICE=eth1. The setting of the MASTER= line will also depend on the final bonding interface name chosen for your bond. As with other network devices, these typically start at 0, and go up one for each device, i.e., the first bonding instance is bond0, the second is bond1, and so on.

Next, create a bond network script. The file name for this script will be /etc/sysconfig/network-scripts/ifcfg-bondX where X is the number of the bond. For bond0 the file is named "ifcfg-bond0", for bond1 it is named "ifcfg-bond1", and so on. Within that file, place the following text:

DEVICE=bond0  
**IPADDR=[192.168.1.1](telnet://192.168.1.1)**  
**NETMASK=[255.255.255.0](telnet://255.255.255.0)**  
**NETWORK=[192.168.1.0](telnet://192.168.1.0)**  
**BROADCAST=[192.168.1.255](telnet://192.168.1.255)**  
ONBOOT=yes  
BOOTPROTO=none  
USERCTL=no

Be sure to change the networking specific lines (IPADDR, NETMASK, NETWORK and BROADCAST) to match your network configuration.

Finally, it is necessary to edit /etc/modules.conf (or /etc/modprobe.conf, depending upon your distro) to load the bonding module with your desired options when the bond0 interface is brought up. The following lines in /etc/modules.conf (or modprobe.conf) will load the bonding module, and select its options:

alias bond0 bonding  
options bond0 mode=balance-alb miimon=100

Replace the sample parameters with the appropriate set of options for your configuration.

Finally run "/etc/rc.d/init.d/network restart" as root. This will restart the networking subsystem and your bond link should be now up and running.

## 3.2.1 Using DHCP with initscripts

Recent versions of initscripts (the version supplied with Fedora Core 3 and Red Hat Enterprise Linux 4 is reported to work) do have support for assigning IP information to bonding devices via DHCP.

To configure bonding for DHCP, configure it as described above, except replace the line "BOOTPROTO=none" with "BOOTPROTO=dhcp" and add a line consisting of "TYPE=Bonding". Note that the TYPE value is case sensitive.

## 3.2.2 Configuring Multiple Bonds with initscripts

At this writing, the initscripts package does not directly support loading the bonding driver multiple times, so the process for doing so is the same as described in the "Configuring Multiple Bonds Manually" section, below.

NOTE: It has been observed that some Red Hat supplied kernels are apparently unable to rename modules at load time (the "-o bond1" part). Attempts to pass that option to modprobe will produce an "Operation not permitted" error. This has been reported on some Fedora Core kernels, and has been seen on RHEL 4 as well. On kernels exhibiting this problem, it will be impossible to configure multiple bonds with differing parameters.

## 3.3 Configuring Bonding Manually

This section applies to distros whose network initialization scripts (the sysconfig or initscripts package) do not have specific knowledge of bonding. One such distro is SuSE Linux Enterprise Server version 8.

The general method for these systems is to place the bonding module parameters into /etc/modules.conf or /etc/modprobe.conf (as appropriate for the installed distro), then add modprobe and/or ifenslave commands to the system's global init script. The name of the global init script differs; for sysconfig, it is /etc/init.d/boot.local and for initscripts it is /etc/rc.d/rc.local.

For example, if you wanted to make a simple bond of two e100 devices (presumed to be eth0 and eth1), and have it persist across reboots, edit the appropriate file (/etc/init.d/boot.local or /etc/rc.d/rc.local), and add the following:

modprobe bonding mode=balance-alb miimon=100 modprobe e100  
ifconfig bond0 [192.168.1.1](telnet://192.168.1.1) netmask [255.255.255.0](telnet://255.255.255.0) up ifenslave bond0 eth0  
ifenslave bond0 eth1

Replace the example bonding module parameters and bond0 network configuration (IP address, netmask, etc) with the appropriate values for your configuration.

Unfortunately, this method will not provide support for the ifup and ifdown scripts on the bond devices. To reload the bonding configuration, it is necessary to run the initialization script, e.g.,

# /etc/init.d/boot.local

or

# /etc/rc.d/rc.local

It may be desirable in such a case to create a separate script which only initializes the bonding configuration, then call that separate script from within boot.local. This allows for bonding to be enabled without re-running the entire global init script.

To shut down the bonding devices, it is necessary to first mark the bonding device itself as being down, then remove the appropriate device driver modules. For our example above, you can do the following:

# ifconfig bond0 down  
# rmmod bonding  
# rmmod e100

Again, for convenience, it may be desirable to create a script with these commands.

## 3.3.1 Configuring Multiple Bonds Manually

This section contains information on configuring multiple bonding devices with differing options for those systems whose network initialization scripts lack support for configuring multiple bonds.

If you require multiple bonding devices, but all with the same options, you may wish to use the "max\_bonds" module parameter, documented above.

To create multiple bonding devices with differing options, it is necessary to load the bonding driver multiple times. Note that current versions of the sysconfig network initialization scripts handle this automatically; if your distro uses these scripts, no special action is needed. See the section Configuring Bonding Devices, above, if you're not sure about your network initialization scripts.

To load multiple instances of the module, it is necessary to specify a different name for each instance (the module loading system requires that every loaded module, even multiple instances of the same module, have a unique name). This is accomplished by supplying multiple sets of bonding options in /etc/modprobe.conf, for example:

alias bond0 bonding  
options bond0 -o bond0 mode=balance-rr miimon=100

alias bond1 bonding  
options bond1 -o bond1 mode=balance-alb miimon=50

will load the bonding module two times. The first instance is named "bond0" and creates the bond0 device in balance-rr mode with an miimon of 100. The second instance is named "bond1" and creates the bond1 device in balance-alb mode with an miimon of 50.

In some circumstances (typically with older distributions), the above does not work, and the second bonding instance never sees its options. In that case, the second options line can be substituted as follows:

install bond1 /sbin/modprobe --ignore-install bonding -o bond1 \

mode=balance-alb miimon=50

This may be repeated any number of times, specifying a new and unique name in place of bond1 for each subsequent instance.

# 5. Querying Bonding Configuration

## 5.1 Bonding Configuration

Each bonding device has a read-only file residing in the /proc/net/bonding directory. The file contents include information about the bonding configuration, options and state of each slave.

For example, the contents of /proc/net/bonding/bond0 after the driver is loaded with parameters of mode=0 and miimon=1000 is generally as follows:

Ethernet Channel Bonding Driver: 2.6.1 (October 29, 2004)

Bonding Mode: load balancing (round-robin)

Currently Active Slave: eth0

MII Status: up

MII Polling Interval (ms): 1000

Up Delay (ms): 0

Down Delay (ms): 0

Slave Interface: eth1

MII Status: up

Link Failure Count: 1

Slave Interface: eth0

MII Status: up

Link Failure Count: 1

The precise format and contents will change depending upon the

bonding configuration, state, and version of the bonding driver.

## 5.2 Network configuration

The network configuration can be inspected using the ifconfig command. Bonding devices will have the MASTER flag set; Bonding slave devices will have the SLAVE flag set. The ifconfig output does not contain information on which slaves are associated with which masters.

In the example below, the bond0 interface is the master (MASTER) while eth0 and eth1 are slaves (SLAVE). Notice all slaves of bond0 have the same MAC address (HWaddr) as bond0 for all modes except TLB and ALB that require a unique MAC address for each slave.

# /sbin/ifconfig

bond0 Link encap:Ethernet HWaddr 00:C0:F0:1F:37:B4

inet addr:XXX.XXX.XXX.YYY Bcast:XXX.XXX.XXX.255 Mask:255.255.252.0

UP BROADCAST RUNNING MASTER MULTICAST MTU:1500 Metric:1

RX packets:7224794 errors:0 dropped:0 overruns:0 frame:0

TX packets:3286647 errors:1 dropped:0 overruns:1 carrier:0

collisions:0 txqueuelen:0

eth0 Link encap:Ethernet HWaddr 00:C0:F0:1F:37:B4

inet addr:XXX.XXX.XXX.YYY Bcast:XXX.XXX.XXX.255 Mask:255.255.252.0

UP BROADCAST RUNNING SLAVE MULTICAST MTU:1500 Metric:1

RX packets:3573025 errors:0 dropped:0 overruns:0 frame:0

TX packets:1643167 errors:1 dropped:0 overruns:1 carrier:0

collisions:0 txqueuelen:100

Interrupt:10 Base address:0x1080

eth1 Link encap:Ethernet HWaddr 00:C0:F0:1F:37:B4

inet addr:XXX.XXX.XXX.YYY Bcast:XXX.XXX.XXX.255 Mask:255.255.252.0

UP BROADCAST RUNNING SLAVE MULTICAST MTU:1500 Metric:1

RX packets:3651769 errors:0 dropped:0 overruns:0 frame:0

TX packets:1643480 errors:0 dropped:0 overruns:0 carrier:0

collisions:0 txqueuelen:100

Interrupt:9 Base address:0x1400

# 6. Switch Configuration

For this section, "switch" refers to whatever system the bonded devices are directly connected to (i.e., where the other end of the cable plugs into). This may be an actual dedicated switch device, or it may be another regular system (e.g., another computer running Linux),

The active-backup, balance-tlb and balance-alb modes do not require any specific configuration of the switch.

The 802.3ad mode requires that the switch have the appropriate ports configured as an 802.3ad aggregation. The precise method used to configure this varies from switch to switch, but, for example, a Cisco 3550 series switch requires that the appropriate ports first be grouped together in a single etherchannel instance, then that etherchannel is set to mode "lacp" to enable 802.3ad (instead of standard EtherChannel).

The balance-rr, balance-xor and broadcast modes generally require that the switch have the appropriate ports grouped together. The nomenclature for such a group differs between switches, it may be called an "etherchannel" (as in the Cisco example, above), a "trunk group" or some other similar variation. For these modes, each switch will also have its own configuration options for the switch's transmit policy to the bond. Typical choices include XOR of either the MAC or IP addresses. The transmit policy of the two peers does not need to match. For these three modes, the bonding mode really selects a transmit policy for an EtherChannel group; all three will interoperate with another EtherChannel group.

# 7. 802.1q VLAN Support

It is possible to configure VLAN devices over a bond interface using the 8021q driver. However, only packets coming from the 8021q driver and passing through bonding will be tagged by default. Self generated packets, for example, bonding's learning packets or ARP packets generated by either ALB mode or the ARP monitor mechanism, are tagged internally by bonding itself. As a result, bonding must "learn" the VLAN IDs configured above it, and use those IDs to tag self generated packets.

For reasons of simplicity, and to support the use of adapters that can do VLAN hardware acceleration offloading, the bonding interface declares itself as fully hardware offloading capable, it gets the add\_vid/kill\_vid notifications to gather the necessary information, and it propagates those actions to the slaves. In case of mixed adapter types, hardware accelerated tagged packets that should go through an adapter that is not offloading capable are "un-accelerated" by the bonding driver so the VLAN tag sits in the regular location.

VLAN interfaces must be added on top of a bonding interface only after enslaving at least one slave. The bonding interface has a hardware address of 00:00:00:00:00:00 until the first slave is added. If the VLAN interface is created prior to the first enslavement, it would pick up the all-zeroes hardware address. Once the first slave is attached to the bond, the bond device itself will pick up the slave's hardware address, which is then available for the VLAN device.

Also, be aware that a similar problem can occur if all slaves are released from a bond that still has one or more VLAN interfaces on top of it. When a new slave is added, the bonding interface will obtain its hardware address from the first slave, which might not match the hardware address of the VLAN interfaces (which was ultimately copied from an earlier slave).

There are two methods to insure that the VLAN device operates with the correct hardware address if all slaves are removed from a bond interface:

1. Remove all VLAN interfaces then recreate them
2. Set the bonding interface's hardware address so that it matches the hardware address of the VLAN interfaces.

Note that changing a VLAN interface's HW address would set the underlying device -- i.e. the bonding interface -- to promiscuous mode, which might not be what you want.

# 8. Link Monitoring

The bonding driver at present supports two schemes for monitoring a slave device's link state: the ARP monitor and the MII monitor.

At the present time, due to implementation restrictions in the bonding driver itself, it is not possible to enable both ARP and MII monitoring simultaneously.

## 8.1 ARP Monitor Operation

The ARP monitor operates as its name suggests: it sends ARP queries to one or more designated peer systems on the network, and uses the response as an indication that the link is operating. This gives some assurance that traffic is actually flowing to and from one or more peers on the local network.

The ARP monitor relies on the device driver itself to verify that traffic is flowing. In particular, the driver must keep up to date the last receive time, dev->last\_rx, and transmit start time, dev->trans\_start. If these are not updated by the driver, then the ARP monitor will immediately fail any slaves using that driver, and those slaves will stay down. If networking monitoring (tcpdump, etc) shows the ARP requests and replies on the network, then it may be that your device driver is not updating last\_rx and trans\_start.

## 8.2 Configuring Multiple ARP Targets

While ARP monitoring can be done with just one target, it can be useful in a High Availability setup to have several targets to monitor. In the case of just one target, the target itself may go down or have a problem making it unresponsive to ARP requests. Having an additional target (or several) increases the reliability of the ARP monitoring.

Multiple ARP targets must be separated by commas as follows:

# example options for ARP monitoring with three targets alias bond0 bonding  
options bond0 arp\_interval=60 arp\_ip\_target=[192.168.0.1](telnet://192.168.0.1),[192.168.0.3](telnet://192.168.0.3),[192.168.0.9](telnet://192.168.0.9)

For just a single target the options would resemble:

# example options for ARP monitoring with one target alias bond0 bonding  
options bond0 arp\_interval=60 arp\_ip\_target=[192.168.0.100](telnet://192.168.0.100)

## 8.3 MII Monitor Operation

The MII monitor monitors only the carrier state of the local network interface. It accomplishes this in one of three ways: by depending upon the device driver to maintain its carrier state, by querying the device's MII registers, or by making an ethtool query to the device.

If the use\_carrier module parameter is 1 (the default value), then the MII monitor will rely on the driver for carrier state information (via the netif\_carrier subsystem). As explained in the use\_carrier parameter information, above, if the MII monitor fails to detect carrier loss on the device (e.g., when the cable is physically disconnected), it may be that the driver does not support netif\_carrier.

If use\_carrier is 0, then the MII monitor will first query the device's (via ioctl) MII registers and check the link state. If that request fails (not just that it returns carrier down), then the MII monitor will make an ethtool ETHOOL\_GLINK request to attempt to obtain the same information. If both methods fail (i.e., the driver either does not support or had some error in processing both the MII register and ethtool requests), then the MII monitor will assume the link is up.

# 9. Potential Sources of Trouble

## 9.1 Adventures in Routing

When bonding is configured, it is important that the slave devices not have routes that supercede routes of the master (or, generally, not have routes at all). For example, suppose the bonding device bond0 has two slaves, eth0 and eth1, and the routing table is as follows:

Kernel IP routing table

Destination Gateway Genmask Flags MSS Window irtt Iface

[10.0.0.0 0](telnet://10.0.0.0:0).0.0.0 [255.255.0.0](telnet://255.255.0.0) U 40 0 0 eth0

[10.0.0.0 0](telnet://10.0.0.0:0).0.0.0 [255.255.0.0](telnet://255.255.0.0) U 40 0 0 eth1

[10.0.0.0 0](telnet://10.0.0.0:0).0.0.0 [255.255.0.0](telnet://255.255.0.0) U 40 0 0 bond0

[127.0.0.0 0](telnet://127.0.0.0:0).0.0.0 [255.0.0.0](telnet://255.0.0.0) U 40 0 0 lo

This routing configuration will likely still update the

receive/transmit times in the driver (needed by the ARP monitor), but may bypass the bonding driver (because outgoing traffic to, in this case, another host on network 10 would use eth0 or eth1 before bond0).

The ARP monitor (and ARP itself) may become confused by this configuration, because ARP requests (generated by the ARP monitor) will be sent on one interface (bond0), but the corresponding reply will arrive on a different interface (eth0). This reply looks to ARP as an unsolicited ARP reply (because ARP matches replies on an interface basis), and is discarded. The MII monitor is not affected by the state of the routing table.

The solution here is simply to insure that slaves do not have routes of their own, and if for some reason they must, those routes do not supercede routes of their master. This should generally be the case, but unusual configurations or errant manual or automatic static route additions may cause trouble.

## 9.2 Ethernet Device Renaming

On systems with network configuration scripts that do not associate physical devices directly with network interface names (so that the same physical device always has the same "ethX" name), it may be necessary to add some special logic to either /etc/modules.conf or /etc/modprobe.conf (depending upon which is installed on the system).

For example, given a modules.conf containing the following:

alias bond0 bonding  
options bond0 mode=some-mode miimon=50  
alias eth0 tg3  
alias eth1 tg3  
alias eth2 e1000  
alias eth3 e1000

If neither eth0 and eth1 are slaves to bond0, then when the bond0 interface comes up, the devices may end up reordered. This happens because bonding is loaded first, then its slave device's drivers are loaded next. Since no other drivers have been loaded, when the e1000 driver loads, it will receive eth0 and eth1 for its devices, but the bonding configuration tries to enslave eth2 and eth3 (which may later be assigned to the tg3 devices).

Adding the following:

add above bonding e1000 tg3

causes modprobe to load e1000 then tg3, in that order, when bonding is loaded. This command is fully documented in the modules.conf manual page.

On systems utilizing modprobe.conf (or modprobe.conf.local), an equivalent problem can occur. In this case, the following can be added to modprobe.conf (or modprobe.conf.local, as appropriate), as follows (all on one line; it has been split here for clarity):

install bonding /sbin/modprobe tg3; /sbin/modprobe e1000;

/sbin/modprobe --ignore-install bonding

This will, when loading the bonding module, rather than performing the normal action, instead execute the provided command. This command loads the device drivers in the order needed, then calls modprobe with --ignore-install to cause the normal action to then take place. Full documentation on this can be found in the modprobe.conf and modprobe manual pages.

## 9.3. Painfully Slow Or No Failed Link Detection By Miimon

By default, bonding enables the use\_carrier option, which instructs bonding to trust the driver to maintain carrier state.

As discussed in the options section, above, some drivers do not support the netif\_carrier\_on/\_off link state tracking system. With use\_carrier enabled, bonding will always see these links as up, regardless of their actual state.

Additionally, other drivers do support netif\_carrier, but do not maintain it in real time, e.g., only polling the link state at some fixed interval. In this case, miimon will detect failures, but only after some long period of time has expired. If it appears that miimon is very slow in detecting link failures, try specifying use\_carrier=0 to see if that improves the failure detection time. If it does, then it may be that the driver checks the carrier state at a fixed interval, but does not cache the MII register values (so the use\_carrier=0 method of querying the registers directly works). If use\_carrier=0 does not improve the failover, then the driver may cache the registers, or the problem may be elsewhere.

Also, remember that miimon only checks for the device's carrier state. It has no way to determine the state of devices on or beyond other ports of a switch, or if a switch is refusing to pass traffic while still maintaining carrier on.

# 10. SNMP agents

If running SNMP agents, the bonding driver should be loaded before any network drivers participating in a bond. This requirement is due to the interface index (ipAdEntIfIndex) being associated to the first interface found with a given IP address. That is, there is only one ipAdEntIfIndex for each IP address. For example, if eth0 and eth1 are slaves of bond0 and the driver for eth0 is loaded before the bonding driver, the interface for the IP address will be associated with the eth0 interface. This configuration is shown below, the IP address [192.168.1.1](telnet://192.168.1.1) has an interface index of 2 which indexes to eth0 in the ifDescr table (ifDescr.2).

interfaces.ifTable.ifEntry.ifDescr.1 = lo

interfaces.ifTable.ifEntry.ifDescr.2 = eth0

interfaces.ifTable.ifEntry.ifDescr.3 = eth1

interfaces.ifTable.ifEntry.ifDescr.4 = eth2

interfaces.ifTable.ifEntry.ifDescr.5 = eth3

interfaces.ifTable.ifEntry.ifDescr.6 = bond0

ip.ipAddrTable.ipAddrEntry.ipAdEntIfIndex.10.10.10.10 = 5

ip.ipAddrTable.ipAddrEntry.ipAdEntIfIndex.192.168.1.1 = 2

ip.ipAddrTable.ipAddrEntry.ipAdEntIfIndex.10.74.20.94 = 4

ip.ipAddrTable.ipAddrEntry.ipAdEntIfIndex.127.0.0.1 = 1

This problem is avoided by loading the bonding driver before

any network drivers participating in a bond. Below is an example of loading the bonding driver first, the IP address [192.168.1.1](telnet://192.168.1.1) is correctly associated with ifDescr.2.

interfaces.ifTable.ifEntry.ifDescr.1 = lo

interfaces.ifTable.ifEntry.ifDescr.2 = bond0

interfaces.ifTable.ifEntry.ifDescr.3 = eth0

interfaces.ifTable.ifEntry.ifDescr.4 = eth1

interfaces.ifTable.ifEntry.ifDescr.5 = eth2

interfaces.ifTable.ifEntry.ifDescr.6 = eth3

ip.ipAddrTable.ipAddrEntry.ipAdEntIfIndex.10.10.10.10 = 6

ip.ipAddrTable.ipAddrEntry.ipAdEntIfIndex.192.168.1.1 = 2

ip.ipAddrTable.ipAddrEntry.ipAdEntIfIndex.10.74.20.94 = 5

ip.ipAddrTable.ipAddrEntry.ipAdEntIfIndex.127.0.0.1 = 1

While some distributions may not report the interface name in

ifDescr, the association between the IP address and IfIndex remains and SNMP functions such as Interface\_Scan\_Next will report that association.

# 11. Promiscuous mode

When running network monitoring tools, e.g., tcpdump, it is common to enable promiscuous mode on the device, so that all traffic is seen (instead of seeing only traffic destined for the local host). The bonding driver handles promiscuous mode changes to the bonding master device (e.g., bond0), and propagates the setting to the slave devices.

For the balance-rr, balance-xor, broadcast, and 802.3ad modes, the promiscuous mode setting is propagated to all slaves.

For the active-backup, balance-tlb and balance-alb modes, the promiscuous mode setting is propagated only to the active slave.

For balance-tlb mode, the active slave is the slave currently receiving inbound traffic.

For balance-alb mode, the active slave is the slave used as a "primary." This slave is used for mode-specific control traffic, for sending to peers that are unassigned or if the load is unbalanced.

For the active-backup, balance-tlb and balance-alb modes, when the active slave changes (e.g., due to a link failure), the promiscuous setting will be propagated to the new active slave.

# 12. Configuring Bonding for High Availability

High Availability refers to configurations that provide maximum network availability by having redundant or backup devices, links or switches between the host and the rest of the world. The goal is to provide the maximum availability of network connectivity (i.e., the network always works), even though other configurations could provide higher throughput.

## 12.1 High Availability in a Single Switch Topology

If two hosts (or a host and a single switch) are directly connected via multiple physical links, then there is no availability penalty to optimizing for maximum bandwidth. In this case, there is only one switch (or peer), so if it fails, there is no alternative access to fail over to. Additionally, the bonding load balance modes support link monitoring of their members, so if individual links fail, the load will be rebalanced across the remaining devices.

See Section 13, "Configuring Bonding for Maximum Throughput" for information on configuring bonding with one peer device.

## 12.2 High Availability in a Multiple Switch Topology

With multiple switches, the configuration of bonding and the network changes dramatically. In multiple switch topologies, there is a trade off between network availability and usable bandwidth.

Below is a sample network, configured to maximize the availability of the network:

| |

|port3 port3|

+-----+----+ +-----+----+

| |port2 ISL port2| |

| switch A +--------------------------+ switch B |

| | | |

+-----+----+ +-----++---+

|port1 port1|

| +-------+ |

+-------------+ host1 +---------------+

eth0 +-------+ eth1

In this configuration, there is a link between the two

switches (ISL, or inter switch link), and multiple ports connecting to the outside world ("port3" on each switch). There is no technical reason that this could not be extended to a third switch.

## 12.2.1 HA Bonding Mode Selection for Multiple Switch Topology

In a topology such as the example above, the active-backup and broadcast modes are the only useful bonding modes when optimizing for availability; the other modes require all links to terminate on the same peer for them to behave rationally.

active-backup: This is generally the preferred mode, particularly if

the switches have an ISL and play together well. If the

network configuration is such that one switch is specifically

a backup switch (e.g., has lower capacity, higher cost, etc),

then the primary option can be used to insure that the

preferred link is always used when it is available.

broadcast: This mode is really a special purpose mode, and is suitable

only for very specific needs. For example, if the two

switches are not connected (no ISL), and the networks beyond

them are totally independent. In this case, if it is

necessary for some specific one-way traffic to reach both

independent networks, then the broadcast mode may be suitable.

## 12.2.2 HA Link Monitoring Selection for Multiple Switch Topology

The choice of link monitoring ultimately depends upon your switch. If the switch can reliably fail ports in response to other failures, then either the MII or ARP monitors should work. For example, in the above example, if the "port3" link fails at the remote end, the MII monitor has no direct means to detect this. The ARP monitor could be configured with a target at the remote end of port3, thus detecting that failure without switch support.

In general, however, in a multiple switch topology, the ARP monitor can provide a higher level of reliability in detecting end to end connectivity failures (which may be caused by the failure of any individual component to pass traffic for any reason). Additionally, the ARP monitor should be configured with multiple targets (at least one for each switch in the network). This will insure that, regardless of which switch is active, the ARP monitor has a suitable target to query.

# 13. Configuring Bonding for Maximum Throughput

## 13.1 Maximizing Throughput in a Single Switch Topology

In a single switch configuration, the best method to maximize throughput depends upon the application and network environment. The various load balancing modes each have strengths and weaknesses in different environments, as detailed below.

For this discussion, we will break down the topologies into two categories. Depending upon the destination of most traffic, we categorize them into either "gatewayed" or "local" configurations.

In a gatewayed configuration, the "switch" is acting primarily as a router, and the majority of traffic passes through this router to other networks. An example would be the following:

+----------+ +----------+

| |eth0 port1| | to other networks

| Host A +---------------------+ router +------------------->

| +---------------------+ | Hosts B and C are out

| |eth1 port2| | here somewhere

+----------+ +----------+

The router may be a dedicated router device, or another host

acting as a gateway. For our discussion, the important point is that the majority of traffic from Host A will pass through the router to some other network before reaching its final destination.

In a gatewayed network configuration, although Host A may communicate with many other systems, all of its traffic will be sent and received via one other peer on the local network, the router.

Note that the case of two systems connected directly via multiple physical links is, for purposes of configuring bonding, the same as a gatewayed configuration. In that case, it happens that all traffic is destined for the "gateway" itself, not some other network beyond the gateway.

In a local configuration, the "switch" is acting primarily as a switch, and the majority of traffic passes through this switch to reach other stations on the same network. An example would be the following:

+----------+ +----------+ +--------+

| |eth0 port1| +-------+ Host B |

| Host A +------------+ switch |port3 +--------+

| +------------+ | +--------+

| |eth1 port2| +------------------+ Host C |

+----------+ +----------+port4 +--------+

Again, the switch may be a dedicated switch device, or another

host acting as a gateway. For our discussion, the important point is that the majority of traffic from Host A is destined for other hosts on the same local network (Hosts B and C in the above example).

In summary, in a gatewayed configuration, traffic to and from the bonded device will be to the same MAC level peer on the network (the gateway itself, i.e., the router), regardless of its final destination. In a local configuration, traffic flows directly to and from the final destinations, thus, each destination (Host B, Host C) will be addressed directly by their individual MAC addresses.

This distinction between a gatewayed and a local network configuration is important because many of the load balancing modes available use the MAC addresses of the local network source and destination to make load balancing decisions. The behavior of each mode is described below.

## 13.1.1 MT Bonding Mode Selection for Single Switch Topology

This configuration is the easiest to set up and to understand, although you will have to decide which bonding mode best suits your needs. The trade offs for each mode are detailed below:

balance-rr: This mode is the only mode that will permit a single

TCP/IP connection to stripe traffic across multiple

interfaces. It is therefore the only mode that will allow a

single TCP/IP stream to utilize more than one interface's

worth of throughput. This comes at a cost, however: the

striping often results in peer systems receiving packets out

of order, causing TCP/IP's congestion control system to kick

in, often by retransmitting segments.

It is possible to adjust TCP/IP's congestion limits by

altering the net.ipv4.tcp\_reordering sysctl parameter. The

usual default value is 3, and the maximum useful value is 127.

For a four interface balance-rr bond, expect that a single

TCP/IP stream will utilize no more than approximately 2.3

interface's worth of throughput, even after adjusting

tcp\_reordering.

Note that this out of order delivery occurs when both the

sending and receiving systems are utilizing a multiple

interface bond. Consider a configuration in which a

balance-rr bond feeds into a single higher capacity network

channel (e.g., multiple 100Mb/sec ethernets feeding a single

gigabit ethernet via an etherchannel capable switch). In this

configuration, traffic sent from the multiple 100Mb devices to

a destination connected to the gigabit device will not see

packets out of order. However, traffic sent from the gigabit

device to the multiple 100Mb devices may or may not see

traffic out of order, depending upon the balance policy of the

switch. Many switches do not support any modes that stripe

traffic (instead choosing a port based upon IP or MAC level

addresses); for those devices, traffic flowing from the

gigabit device to the many 100Mb devices will only utilize one

interface.

If you are utilizing protocols other than TCP/IP, UDP for

example, and your application can tolerate out of order

delivery, then this mode can allow for single stream datagram

performance that scales near linearly as interfaces are added

to the bond.

This mode requires the switch to have the appropriate ports

configured for "etherchannel" or "trunking."

active-backup: There is not much advantage in this network topology to

the active-backup mode, as the inactive backup devices are all

connected to the same peer as the primary. In this case, a

load balancing mode (with link monitoring) will provide the

same level of network availability, but with increased

available bandwidth. On the plus side, active-backup mode

does not require any configuration of the switch, so it may

have value if the hardware available does not support any of

the load balance modes.

balance-xor: This mode will limit traffic such that packets destined

for specific peers will always be sent over the same

interface. Since the destination is determined by the MAC

addresses involved, this mode works best in a "local" network

configuration (as described above), with destinations all on

the same local network. This mode is likely to be suboptimal

if all your traffic is passed through a single router (i.e., a

"gatewayed" network configuration, as described above).

As with balance-rr, the switch ports need to be configured for

"etherchannel" or "trunking."

broadcast: Like active-backup, there is not much advantage to this

mode in this type of network topology.

802.3ad: This mode can be a good choice for this type of network

topology. The 802.3ad mode is an IEEE standard, so all peers

that implement 802.3ad should interoperate well. The 802.3ad

protocol includes automatic configuration of the aggregates,

so minimal manual configuration of the switch is needed

(typically only to designate that some set of devices is

available for 802.3ad). The 802.3ad standard also mandates

that frames be delivered in order (within certain limits), so

in general single connections will not see misordering of

packets. The 802.3ad mode does have some drawbacks: the

standard mandates that all devices in the aggregate operate at

the same speed and duplex. Also, as with all bonding load

balance modes other than balance-rr, no single connection will

be able to utilize more than a single interface's worth of

bandwidth.

Additionally, the linux bonding 802.3ad implementation

distributes traffic by peer (using an XOR of MAC addresses),

so in a "gatewayed" configuration, all outgoing traffic will

generally use the same device. Incoming traffic may also end

up on a single device, but that is dependent upon the

balancing policy of the peer's 8023.ad implementation. In a

"local" configuration, traffic will be distributed across the

devices in the bond.

Finally, the 802.3ad mode mandates the use of the MII monitor,

therefore, the ARP monitor is not available in this mode.

balance-tlb: The balance-tlb mode balances outgoing traffic by peer.

Since the balancing is done according to MAC address, in a

"gatewayed" configuration (as described above), this mode will

send all traffic across a single device. However, in a

"local" network configuration, this mode balances multiple

local network peers across devices in a vaguely intelligent

manner (not a simple XOR as in balance-xor or 802.3ad mode),

so that mathematically unlucky MAC addresses (i.e., ones that

XOR to the same value) will not all "bunch up" on a single

interface.

Unlike 802.3ad, interfaces may be of differing speeds, and no

special switch configuration is required. On the down side,

in this mode all incoming traffic arrives over a single

interface, this mode requires certain ethtool support in the

network device driver of the slave interfaces, and the ARP

monitor is not available.

balance-alb: This mode is everything that balance-tlb is, and more.

It has all of the features (and restrictions) of balance-tlb,

and will also balance incoming traffic from local network

peers (as described in the Bonding Module Options section,

above).

The only additional down side to this mode is that the network

device driver must support changing the hardware address while

the device is open.

## 13.1.2 MT Link Monitoring for Single Switch Topology

The choice of link monitoring may largely depend upon which mode you choose to use. The more advanced load balancing modes do not support the use of the ARP monitor, and are thus restricted to using the MII monitor (which does not provide as high a level of end to end assurance as the ARP monitor).

## 13.2 Maximum Throughput in a Multiple Switch Topology

Multiple switches may be utilized to optimize for throughput when they are configured in parallel as part of an isolated network between two or more systems, for example:

+-----------+

| Host A |

+-+---+---+-+

| | |

+--------+ | +---------+

| | |

+------+---+ +-----+----+ +-----+----+

| Switch A | | Switch B | | Switch C |

+------+---+ +-----+----+ +-----+----+

| | |

+--------+ | +---------+

| | |

+-+---+---+-+

| Host B |

+-----------+

In this configuration, the switches are isolated from one

another. One reason to employ a topology such as this is for an isolated network with many hosts (a cluster configured for high performance, for example), using multiple smaller switches can be more cost effective than a single larger switch, e.g., on a network with 24 hosts, three 24 port switches can be significantly less expensive than a single 72 port switch.

If access beyond the network is required, an individual host can be equipped with an additional network device connected to an external network; this host then additionally acts as a gateway.

## 13.2.1 MT Bonding Mode Selection for Multiple Switch Topology

In actual practice, the bonding mode typically employed in configurations of this type is balance-rr. Historically, in this network configuration, the usual caveats about out of order packet delivery are mitigated by the use of network adapters that do not do any kind of packet coalescing (via the use of NAPI, or because the device itself does not generate interrupts until some number of packets has arrived). When employed in this fashion, the balance-rr mode allows individual connections between two hosts to effectively utilize greater than one interface's bandwidth.

## 13.2.2 MT Link Monitoring for Multiple Switch Topology

Again, in actual practice, the MII monitor is most often used in this configuration, as performance is given preference over availability. The ARP monitor will function in this topology, but its advantages over the MII monitor are mitigated by the volume of probes needed as the number of systems involved grows (remember that each host in the network is configured with bonding).

# 14. Switch Behavior Issues

## 14.1 Link Establishment and Failover Delays

Some switches exhibit undesirable behavior with regard to the timing of link up and down reporting by the switch.

First, when a link comes up, some switches may indicate that the link is up (carrier available), but not pass traffic over the interface for some period of time. This delay is typically due to some type of autonegotiation or routing protocol, but may also occur during switch initialization (e.g., during recovery after a switch failure). If you find this to be a problem, specify an appropriate value to the updelay bonding module option to delay the use of the relevant interface(s).

Second, some switches may "bounce" the link state one or more times while a link is changing state. This occurs most commonly while the switch is initializing. Again, an appropriate updelay value may help.

Note that when a bonding interface has no active links, the driver will immediately reuse the first link that goes up, even if the updelay parameter has been specified (the updelay is ignored in this case). If there are slave interfaces waiting for the updelay timeout to expire, the interface that first went into that state will be immediately reused. This reduces down time of the network if the value of updelay has been overestimated, and since this occurs only in cases with no connectivity, there is no additional penalty for ignoring the updelay.

In addition to the concerns about switch timings, if your switches take a long time to go into backup mode, it may be desirable to not activate a backup interface immediately after a link goes down. Failover may be delayed via the downdelay bonding module option.

## 14.2 Duplicated Incoming Packets

It is not uncommon to observe a short burst of duplicated traffic when the bonding device is first used, or after it has been idle for some period of time. This is most easily observed by issuing a "ping" to some other host on the network, and noticing that the output from ping flags duplicates (typically one per slave).

For example, on a bond in active-backup mode with five slaves all connected to one switch, the output may appear as follows:

# ping -n [10.0.4.2](telnet://10.0.4.2)  
PING [10.0.4.2](telnet://10.0.4.2) ([10.0.4.2](telnet://10.0.4.2)) from [10.0.3.10](telnet://10.0.3.10) : 56(84) bytes of data. 64 bytes from [10.0.4.2](telnet://10.0.4.2): icmp\_seq=1 ttl=64 time=13.7 ms 64 bytes from [10.0.4.2](telnet://10.0.4.2): icmp\_seq=1 ttl=64 time=13.8 ms (DUP!) 64 bytes from [10.0.4.2](telnet://10.0.4.2): icmp\_seq=1 ttl=64 time=13.8 ms (DUP!) 64 bytes from [10.0.4.2](telnet://10.0.4.2): icmp\_seq=1 ttl=64 time=13.8 ms (DUP!) 64 bytes from [10.0.4.2](telnet://10.0.4.2): icmp\_seq=1 ttl=64 time=13.8 ms (DUP!) 64 bytes from [10.0.4.2](telnet://10.0.4.2): icmp\_seq=2 ttl=64 time=0.216 ms 64 bytes from [10.0.4.2](telnet://10.0.4.2): icmp\_seq=3 ttl=64 time=0.267 ms 64 bytes from [10.0.4.2](telnet://10.0.4.2): icmp\_seq=4 ttl=64 time=0.222 ms

This is not due to an error in the bonding driver, rather, it is a side effect of how many switches update their MAC forwarding tables. Initially, the switch does not associate the MAC address in the packet with a particular switch port, and so it may send the traffic to all ports until its MAC forwarding table is updated. Since the interfaces attached to the bond may occupy multiple ports on a single switch, when the switch (temporarily) floods the traffic to all ports, the bond device receives multiple copies of the same packet (one per slave device).

The duplicated packet behavior is switch dependent, some switches exhibit this, and some do not. On switches that display this behavior, it can be induced by clearing the MAC forwarding table (on most Cisco switches, the privileged command "clear mac address-table dynamic" will accomplish this).

# 15. Hardware Specific Considerations

This section contains additional information for configuring bonding on specific hardware platforms, or for interfacing bonding with particular switches or other devices.

## 15.1 IBM BladeCenter

This applies to the JS20 and similar systems.

On the JS20 blades, the bonding driver supports only balance-rr, active-backup, balance-tlb and balance-alb modes. This is largely due to the network topology inside the BladeCenter, detailed below.

## JS20 network adapter information

All JS20s come with two Broadcom Gigabit Ethernet ports integrated on the planar (that's "motherboard" in IBM-speak). In the BladeCenter chassis, the eth0 port of all JS20 blades is hard wired to I/O Module #1; similarly, all eth1 ports are wired to I/O Module #2. An add-on Broadcom daughter card can be installed on a JS20 to provide two more Gigabit Ethernet ports. These ports, eth2 and eth3, are wired to I/O Modules 3 and 4, respectively.

Each I/O Module may contain either a switch or a passthrough module (which allows ports to be directly connected to an external switch). Some bonding modes require a specific BladeCenter internal network topology in order to function; these are detailed below.

Additional BladeCenter-specific networking information can be found in two IBM Redbooks ([www.ibm.com/redbooks):](http://www.ibm.com/redbooks):)

"IBM eServer BladeCenter Networking Options" "IBM eServer BladeCenter Layer 2-7 Network Switching"

## BladeCenter networking configuration

Because a BladeCenter can be configured in a very large number of ways, this discussion will be confined to describing basic configurations.

Normally, Ethernet Switch Modules (ESMs) are used in I/O modules 1 and 2. In this configuration, the eth0 and eth1 ports of a JS20 will be connected to different internal switches (in the respective I/O modules).

A passthrough module (OPM or CPM, optical or copper, passthrough module) connects the I/O module directly to an external switch. By using PMs in I/O module #1 and #2, the eth0 and eth1 interfaces of a JS20 can be redirected to the outside world and connected to a common external switch.

Depending upon the mix of ESMs and PMs, the network will appear to bonding as either a single switch topology (all PMs) or as a multiple switch topology (one or more ESMs, zero or more PMs). It is also possible to connect ESMs together, resulting in a configuration much like the example in "High Availability in a Multiple Switch Topology," above.

## Requirements for specific modes

The balance-rr mode requires the use of passthrough modules for devices in the bond, all connected to an common external switch. That switch must be configured for "etherchannel" or "trunking" on the appropriate ports, as is usual for balance-rr.

The balance-alb and balance-tlb modes will function with either switch modules or passthrough modules (or a mix). The only specific requirement for these modes is that all network interfaces must be able to reach all destinations for traffic sent over the bonding device (i.e., the network must converge at some point outside the BladeCenter).

The active-backup mode has no additional requirements.

## Link monitoring issues

When an Ethernet Switch Module is in place, only the ARP monitor will reliably detect link loss to an external switch. This is nothing unusual, but examination of the BladeCenter cabinet would suggest that the "external" network ports are the ethernet ports for the system, when it fact there is a switch between these "external" ports and the devices on the JS20 system itself. The MII monitor is only able to detect link failures between the ESM and the JS20 system.

When a passthrough module is in place, the MII monitor does detect failures to the "external" port, which is then directly connected to the JS20 system.

## Other concerns

The Serial Over LAN (SoL) link is established over the primary ethernet (eth0) only, therefore, any loss of link to eth0 will result in losing your SoL connection. It will not fail over with other network traffic, as the SoL system is beyond the control of the bonding driver.

It may be desirable to disable spanning tree on the switch (either the internal Ethernet Switch Module, or an external switch) to avoid fail-over delay issues when using bonding.

16. Frequently Asked Questions

1. Is it SMP safe?

Yes. The old 2.0.xx channel bonding patch was not SMP safe. The new driver was designed to be SMP safe from the start.

2. What type of cards will work with it?

Any Ethernet type cards (you can even mix cards - a Intel EtherExpress PRO/100 and a 3com 3c905b, for example). For most modes, devices need not be of the same speed.

3. How many bonding devices can I have?

There is no limit.

4. How many slaves can a bonding device have?

This is limited only by the number of network interfaces Linux supports and/or the number of network cards you can place in your system.

5. What happens when a slave link dies?

If link monitoring is enabled, then the failing device will be disabled. The active-backup mode will fail over to a backup link, and other modes will ignore the failed link. The link will continue to be monitored, and should it recover, it will rejoin the bond (in whatever manner is appropriate for the mode). See the sections on High Availability and the documentation for each mode for additional information.

Link monitoring can be enabled via either the miimon or arp\_interval parameters (described in the module parameters section, above). In general, miimon monitors the carrier state as sensed by the underlying network device, and the arp monitor (arp\_interval) monitors connectivity to another host on the local network.

If no link monitoring is configured, the bonding driver will be unable to detect link failures, and will assume that all links are always available. This will likely result in lost packets, and a resulting degradation of performance. The precise performance loss depends upon the bonding mode and network configuration.

6. Can bonding be used for High Availability?

Yes. See the section on High Availability for details.

7. Which switches/systems does it work with?

The full answer to this depends upon the desired mode.

In the basic balance modes (balance-rr and balance-xor), it works with any system that supports etherchannel (also called trunking). Most managed switches currently available have such support, and many unmanaged switches as well.

The advanced balance modes (balance-tlb and balance-alb) do not have special switch requirements, but do need device drivers that support specific features (described in the appropriate section under module parameters, above).

In 802.3ad mode, it works with with systems that support IEEE 802.3ad Dynamic Link Aggregation. Most managed and many unmanaged switches currently available support 802.3ad.

The active-backup mode should work with any Layer-II switch.

8. Where does a bonding device get its MAC address from?

If not explicitly configured (with ifconfig or ip link), the MAC address of the bonding device is taken from its first slave device. This MAC address is then passed to all following slaves and remains persistent (even if the first slave is removed) until the bonding device is brought down or reconfigured.

If you wish to change the MAC address, you can set it with ifconfig or ip link:

# ifconfig bond0 hw ether 00:11:22:33:44:55

# ip link set bond0 address 66:77:88:99:aa:bb

The MAC address can be also changed by bringing down/up the device and then changing its slaves (or their order):

# ifconfig bond0 down ; modprobe -r bonding # ifconfig bond0 .... up  
# ifenslave bond0 eth...

This method will automatically take the address from the next slave that is added.

To restore your slaves' MAC addresses, you need to detach them from the bond (`ifenslave -d bond0 eth0'). The bonding driver will then restore the MAC addresses that the slaves had before they were enslaved.

# 20 Linux Server Hardening Security Tips

by [nixCraft](http://www.cyberciti.biz/tips/about-us) on October 30, 2009 · [120 comments](http://www.cyberciti.biz/tips/linux-security.html#comments)· LAST UPDATED December 5, 2009

in [Debian Linux](http://www.cyberciti.biz/tips/category/debian-linux), [fedora linux](http://www.cyberciti.biz/tips/category/fedora-linux), [Gentoo Linux](http://www.cyberciti.biz/tips/category/gentoo-linux)

[](http://www.cyberciti.biz/tips/category/linux)

Securing your Linux server is important to protect your data, intellectual property, and time, from the hands of crackers (hackers). The system administrator is responsible for security Linux box. In this first part of a Linux server security series, I will provide 20 hardening tips for default installation of Linux system.

## #1: Encrypt Data Communication

All data transmitted over a network is open to monitoring. Encrypt transmitted data whenever possible with password or using keys / certificates.

1. Use [scp, ssh](http://openssh.org/), rsync, or sftp for file transfer. You can also mount [remote server file system](http://www.cyberciti.biz/tips/rhel-centos-mounting-remote-filesystem-using-sshfs.html) or your own home directory using special sshfs and fuse tools.
2. [GnuPG](http://www.gnupg.org/) allows to encrypt and sign your data and communication, features a versatile key managment system as well as access modules for all kind of public key directories.
3. [Fugu](http://rsug.itd.umich.edu/software/fugu/) is a graphical frontend to the commandline Secure File Transfer application (SFTP). SFTP is similar to FTP, but unlike FTP, the entire session is encrypted, meaning no passwords are sent in cleartext form, and is thus much less vulnerable to third-party interception. Another option is [FileZilla](http://filezilla-project.org/) - a cross-platform client that supports FTP, FTP over SSL/TLS (FTPS), and SSH File Transfer Protocol (SFTP).
4. [OpenVPN](http://openvpn.net/) is a cost-effective, lightweight SSL VPN.
5. [Lighttpd SSL (Secure Server Layer) Https](http://www.cyberciti.biz/tips/how-to-install-ssl-lighttpd-https-configuration.html) Configuration And Installation
6. [Apache SSL (Secure Server Layer) Https](http://www.cyberciti.biz/faq/rhel-apache-httpd-mod-ssl-tutorial/) (mod\_ssl) Configuration And Installation

### #1.1: Avoid Using FTP, Telnet, And Rlogin / Rsh

Under most network configurations, user names, passwords, FTP / telnet / rsh commands and transferred files can be captured by anyone on the same network using a packet sniffer. The common solution to this problem is to use either [OpenSSH](http://www.cyberciti.biz/tips/linux-unix-bsd-openssh-server-best-practices.html) , [SFTP, or FTPS](http://www.cyberciti.biz/tips/configure-vsfptd-secure-connections-via-ssl-tls.html) (FTP over SSL), which adds SSL or TLS encryption to FTP. Type the following command to delete NIS, rsh and other outdated service:  
# yum erase inetd xinetd ypserv tftp-server telnet-server rsh-serve

## #2: Minimize Software to Minimize Vulnerability

Do you really need all sort of web services installed? Avoid installing unnecessary software to avoid vulnerabilities in software. Use the RPM package manager such [as yum](http://www.cyberciti.biz/faq/rhel-centos-fedora-linux-yum-command-howto/) or [apt-get and/or dpkg to review](http://www.cyberciti.biz/tips/linux-debian-package-management-cheat-sheet.html) all installed set of software packages on a system. Delete all unwanted packages.  
# yum list installed  
# yum list packageName  
# yum remove packageName  
OR  
# dpkg --list  
# dpkg --info packageName  
# apt-get remove packageName

## #3: One Network Service Per System or VM Instance

Run different network services on separate servers or VM instance. This limits the number of other services that can be compromised. For example, if an attacker able to successfully exploit a software such as Apache flow, he / she will get an access to entire server including other services such as MySQL, e-mail server and so on. See how to install Virtualization software:

* [Install and Setup XEN Virtualization Software on CentOS Linux 5](http://www.cyberciti.biz/tips/rhel-centos-xen-virtualization-installation-howto.html)
* [How To Setup OpenVZ under RHEL / CentOS Linux](http://www.cyberciti.biz/faq/openvz-rhel-centos-linux-tutorial/)

## #4: Keep Linux Kernel and Software Up to Date

Applying security patches is an important part of maintaining Linux server. Linux provides all necessary tools to keep your system updated, and also allows for easy upgrades between versions. All security update should be reviewed and applied as soon as possible. Again, use the RPM package manager such [as yum](http://www.cyberciti.biz/faq/rhel-centos-fedora-linux-yum-command-howto/) and/or [apt-get and/or dpkg to](http://www.cyberciti.biz/tips/linux-debian-package-management-cheat-sheet.html) apply all security updates.  
# yum update   
OR  
# apt-get update && apt-get upgrade  
You can configure Red hat / CentOS / Fedora Linux to send yum package [update notification via email](http://www.cyberciti.biz/faq/red-hat-centos-fedora-send-package-update-notification-via-email/). Another option is to apply [all security updates](http://www.cyberciti.biz/faq/fedora-automatic-update-retrieval-installation-with-cron/) via a cron job. Under Debian / Ubuntu Linux you can use [apticron](http://www.cyberciti.biz/faq/apt-get-apticron-send-email-upgrades-available/) to send security notifications.

## #5: Use Linux Security Extensions

Linux comes with various security patches which can be used to guard against misconfigured or compromised programs. If possible use [SELinux and other Linux security](http://www.cyberciti.biz/tips/selinux-vs-apparmor-vs-grsecurity.html) extensions to enforce limitations on network and other programs. For example, SELinux provides a variety of security policies for Linux kernel.

### #5.1: SELinux

I strongly recommend using SELinux which provides a flexible Mandatory Access Control (MAC). Under standard Linux Discretionary Access Control (DAC), an application or process running as a user (UID or SUID) has the user's permissions to objects such as files, sockets, and other processes. Running a MAC kernel protects the system from malicious or flawed applications that can damage or destroy the system. See the official [Redhat](http://www.redhat.com/docs/en-US/Red_Hat_Enterprise_Linux/5.4/html/Deployment_Guide/selg-overview.html) documentation which explains SELinux configuration.

## #6: User Accounts and Strong Password Policy

Use the useradd / usermod commands to create and maintain user accounts. Make sure you have a good and strong password policy. For example, a good password includes at least 8 characters long and mixture of alphabets, number, special character, upper & lower alphabets etc. Most important pick a password you can remember. Use tools such as "[John the ripper](http://www.cyberciti.biz/faq/unix-linux-password-cracking-john-the-ripper/)" to find out weak users passwords on your server. Configure [pam\_cracklib.so to](http://www.cyberciti.biz/tips/linux-check-passwords-against-a-dictionary-attack.html) enforce the password policy.

### #6.1: Password Aging

The [chage command](http://www.cyberciti.biz/faq/tag/chage-command/) changes the number of days between password changes and the date of the last password change. This information is used by the system to determine when a user must change his/her password. The [/etc/login.defs file](http://www.cyberciti.biz/tips/linux-set-default-password-expiry-for-all-new-users.html) defines the site-specific configuration for the shadow password suite including password aging configuration. To disable password aging, enter:  
chage -M 99999 userName  
To get password expiration information, enter:  
chage -l userName  
Finally, you can also edit the [/etc/shadow file](http://www.cyberciti.biz/faq/understanding-etcshadow-file/) in the following fields:

{userName}:{password}:{lastpasswdchanged}:{Minimum\_days}:{Maximum\_days}:{Warn}:{Inactive}:{Expire}:

Where,

1. **Minimum\_days**: The minimum number of days required between password changes i.e. the number of days left before the user is allowed to change his/her password.
2. **Maximum\_days**: The maximum number of days the password is valid (after that user is forced to change his/her password).
3. **Warn** : The number of days before password is to expire that user is warned that his/her password must be changed.
4. **Expire** : Days since Jan 1, 1970 that account is disabled i.e. an absolute date specifying when the login may no longer be used.

I recommend chage command instead of editing the /etc/shadow by hand:  
# chage -M 60 -m 7 -W 7 userName  
Recommend readings:

* [Linux: Force Users To Change Their Passwords Upon First Login](http://www.cyberciti.biz/faq/rhel-debian-force-users-to-change-passwords/)
* [Linux turn On / Off password expiration / aging](http://www.cyberciti.biz/tips/setting-off-password-aging-expiration.html)
* [Lock the user password](http://www.cyberciti.biz/faq/linux-locking-an-account/)
* [Search for all account without password and lock them](http://www.cyberciti.biz/tips/search-for-all-account-without-password-and-lock-them.html)
* [Use Linux groups to enhance security](http://www.cyberciti.biz/faq/protect-command-by-configuring-linux-unix-group-permissions/)

### #6.2: Restricting Use of Previous Passwords

You can prevent all users from using or reuse same old passwords under Linux. The [pam\_unix module parameter remember](http://www.cyberciti.biz/tips/how-to-linux-prevent-the-reuse-of-old-passwords.html) can be used to configure the number of previous passwords that cannot be reused.

### #6.3: Locking User Accounts After Login Failures

Under Linux you can use the [faillog command to](http://www.cyberciti.biz/tips/rhel-centos-fedora-linux-log-failed-login.html) display faillog records or to set login failure limits. faillog formats the contents of the failure log from /var/log/faillog database / log file. It also can be used for maintains failure counters and limits.To see failed login attempts, enter:  
faillog  
To unlock an account after login failures, run:  
faillog -r -u userName   
Note you can use passwd command to lock and unlock accounts:  
# lock account  
[passwd -l userName](http://www.cyberciti.biz/faq/linux-locking-an-account/)  
# unlocak account  
[passwd -u userName](http://www.cyberciti.biz/faq/linux-locking-an-account/)

### #6.4: How Do I Verify No Accounts Have Empty Passwords?

Type the following command  
# awk -F: '($2 == "") {print}' /etc/shadow  
Lock all empty password accounts:  
# passwd -l accountName

### #6.5: Make Sure No Non-Root Accounts Have UID Set To 0

Only root account have UID 0 with full permissions to access the system. Type the following command to display all accounts with UID set to 0:  
# awk -F: '($3 == "0") {print}' /etc/passwd  
You should only see one line as follows:

root:x:0:0:root:/root:/bin/bash

If you see other lines, delete them or make sure other accounts are authorized by you to use UID 0.

## #7: Disable root Login

Never ever login as root user. You should [use sudo to](http://www.cyberciti.biz/tips/allow-a-normal-user-to-run-commands-as-root.html) execute root level commands as and when required. sudo does greatly enhances the security of the system without sharing root password with other users and admins. sudo provides simple [auditing and tracking](http://www.cyberciti.biz/faq/sudo-send-e-mail-sudo-log-file/) features too.

## #8: Physical Server Security

You must protect Linux servers physical console access. Configure [the BIOS](http://www.cyberciti.biz/tips/tips-to-protect-linux-servers-physical-console-access.html) and disable the booting from external devices such as DVDs / CDs / USB pen. Set BIOS and grub [boot loader password](http://www.cyberciti.biz/tips/how-do-i-secure-grub-boot-loader.html) to protect these settings. All production boxes must be locked in IDCs (Internet Data Center) and all persons must pass some sort of security checks before accessing your server. See also:

* [9 Tips To Protect Linux Servers Physical Console Access](http://www.cyberciti.biz/tips/tips-to-protect-linux-servers-physical-console-access.html).

## #9: Disable Unwanted Services

Disable all unnecessary services and daemons (services that runs in the background). You need to remove all unwanted services from the system start-up. Type the following [command to list](http://www.cyberciti.biz/faq/check-running-services-in-rhel-redhat-fedora-centoslinux/) all services which are started at boot time in run level # 3:  
# chkconfig --list | grep '3:on'  
To disable service, enter:  
# service serviceName stop  
# chkconfig serviceName off

### #9.1: Find Listening Network Ports

Use the following command to list all open ports and associated programs:  
[netstat -tulpn](http://www.cyberciti.biz/tips/linux-display-open-ports-owner.html)  
OR  
[nmap -sT -O localhost](http://www.cyberciti.biz/tips/linux-scanning-network-for-open-ports.html)  
nmap -sT -O server.example.com  
Use iptables to close open ports or stop all unwanted network services using above service and chkconfig commands.

#### #9.2: See Also

* [update-rc.d like command on Redhat Enterprise / CentOS Linux](http://www.cyberciti.biz/faq/rhel5-update-rcd-command/).
* [Ubuntu / Debian Linux: Services Configuration Tool to Start / Stop System Services](http://www.cyberciti.biz/tips/how-to-controlling-access-to-linux-services.html).
* [Get Detailed Information About Particular IP](http://www.cyberciti.biz/tips/netstat-command-tutorial-examples.html) address Connections Using netstat Command.

## #10: Delete X Windows

X Windows on server is not required. There is no reason to run X Windows on your dedicated mail and Apache web server. You can disable and remove X Windows to improve server security and performance. Edit [/etc/inittab](http://www.cyberciti.biz/tips/linux-changing-run-levels.html) and set run level to 3. Finally, remove X Windows system, enter:  
# [yum groupremove "X Window System"](http://www.cyberciti.biz/faq/rhel-centos-fedora-linux-yum-command-howto/)

## #11: Configure Iptables and TCPWrappers

[Iptables](http://www.cyberciti.biz/faq/category/iptables/) is a user space application program that allows you to configure the firewall (Netfilter) provided by the Linux kernel. Use [firewall](http://bash.cyberciti.biz/firewall/linux-iptables-firewall-shell-script-for-standalone-server/) to filter [out traffic and allow only](http://www.cyberciti.biz/tips/block-spamming-scanning-with-iptables.html) necessary traffic. Also use the [TCPWrappers a host-based](http://www.cyberciti.biz/faq/tcp-wrappers-hosts-allow-deny-tutorial/) networking ACL system to filter network access to Internet. You can prevent many denial of service attacks with the help of Iptables:

* [Lighttpd Traffic Shaping: Throttle Connections Per Single IP (Rate Limit)](http://www.cyberciti.biz/tips/lighttpd-set-throughput-connections-per-ip.html).
* [How to: Linux Iptables block common attack](http://www.cyberciti.biz/tips/linux-iptables-10-how-to-block-common-attack.html).
* [psad: Linux Detect And Block Port Scan Attacks In Real Time](http://www.cyberciti.biz/faq/linux-detect-port-scan-attacks/).

## #12: Linux Kernel /etc/sysctl.conf Hardening

/etc/sysctl.conf file is used to [configure kernel parameters](http://www.cyberciti.biz/faq/linux-kernel-etcsysctl-conf-security-hardening/) at runtime. Linux reads and applies settings from /etc/sysctl.conf at boot time. Sample [/etc/sysctl.conf](http://www.cyberciti.biz/faq/tag/etcsysctlconf/):

# Turn on execshield

kernel.exec-shield=1

kernel.randomize\_va\_space=1

# Enable IP spoofing protection

net.ipv4.conf.all.rp\_filter=1

# Disable IP source routing

net.ipv4.conf.all.accept\_source\_route=0

# Ignoring broadcasts request

net.ipv4.icmp\_echo\_ignore\_broadcasts=1

net.ipv4.icmp\_ignore\_bogus\_error\_messages=1

# Make sure spoofed packets get logged

net.ipv4.conf.all.log\_martians = 1

## #13: Separate Disk Partitions

Separation of the [operating system files](http://www.cyberciti.biz/tips/the-importance-of-linux-partitions.html) from user files may result into a better and secure system. Make sure the following filesystems are mounted on separate partitions:

* /usr
* /home
* /var and /var/tmp
* /tmp

Create septate partitions for Apache and FTP server roots. Edit /etc/fstab file and make sure you add the following configuration options:

1. **noexec** - Do not set execution of any binaries on this partition (prevents execution of binaries but allows scripts).
2. **nodev** - Do not allow character or special devices on this partition (prevents use of device files such as zero, sda etc).
3. **nosuid** - Do not set SUID/SGID access on this partition (prevent the setuid bit).

Sample [/etc/fstab](http://www.cyberciti.biz/faq/tag/etcfstab/) entry to to limit user access on /dev/sda5 (ftp server root directory):

/dev/sda5 /ftpdata ext3 defaults,nosuid,nodev,noexec 1 2

### #13.1: Disk Quotas

Make sure disk quota is enabled for all users. To implement disk quotas, use the following steps:

1. Enable quotas per file system by modifying the /etc/fstab file.
2. Remount the file system(s).
3. Create the quota database files and generate the disk usage table.
4. Assign quota policies.
5. See [implementing disk quotas](http://www.redhat.com/docs/en-US/Red_Hat_Enterprise_Linux/5.4/html/Deployment_Guide/ch-disk-quotas.html) tutorial for further details.

## #14: Turn Off IPv6

Internet Protocol version 6 (IPv6) provides a new Internet layer of the TCP/IP protocol suite that replaces Internet Protocol version 4 (IPv4) and provides many benefits. Currently there are no good tools out which are able to check a system over network for IPv6 security issues. Most Linux distro began enabling IPv6 protocol by default. Crackers can send bad traffic via IPv6 as most admins are not monitoring it. Unless network configuration requires it, disable IPv6 or configure Linux IPv6 firewall:

* [RedHat / Centos Disable IPv6 Networking](http://www.cyberciti.biz/faq/redhat-centos-disable-ipv6-networking/).
* [Debian / Ubuntu And Other Linux Distros Disable IPv6 Networking](http://www.cyberciti.biz/tips/linux-how-to-disable-the-ipv6-protocol.html).
* [Linux IPv6 Howto - Chapter 19. Security](http://mirrors.deepspace6.net/Linux+IPv6-HOWTO/chapter-security.html).
* [Linux IPv6 Firewall configuration and scripts are](http://www.cyberciti.biz/faq/redhat-fedora-ip6tables-firewall-configuration/) [available here](http://bash.cyberciti.biz/script/ipv6-firewall/).

## #15: Disable Unwanted SUID and SGID Binaries

All SUID/SGID bits enabled file can be misused when the SUID/SGID executable has a security problem or bug. All local or remote user can use such file. It is a good idea to find all such files. Use the find command as follows:  
#See all set user id files:  
find / -perm +4000  
# See all group id files  
find / -perm +2000  
# Or combine both in a single command  
find / \( -perm -4000 -o -perm -2000 \) -print  
find / -path -prune -o -type f -perm +6000 -ls  
  
You need to investigate each reported file. See reported file man page for further details.

### #15.1: World-Writable Files

Anyone can modify world-writable file resulting into a security issue. Use the following command to find [all world writable](http://www.cyberciti.biz/faq/find-all-world-writable-directories-have-stickybitsset-on/) and sticky bits set files:  
find /dir -xdev -type d \( -perm -0002 -a ! -perm -1000 \) -print  
You need to investigate each reported file and either set correct user and group permission or remove it.

### #15.2: Noowner Files

Files not owned by any user or group can pose a security problem. Just find them with the following command which do not belong to a valid user and a valid group  
find /dir -xdev \( -nouser -o -nogroup \) -print  
You need to investigate each reported file and either assign it to an appropriate user and group or remove it.

## #16: Use A Centralized Authentication Service

Without a centralized authentication system, user auth data becomes inconsistent, which may lead into out-of-date credentials and forgotten accounts which should have [been deleted](http://www.cyberciti.biz/faq/former-employees-keep-accessing-linux-unix-server/) in first place. A centralized authentication service allows you maintaining central control over Linux / UNIX account and authentication data. You can keep auth data synchronized between servers. Do not use the NIS service for centralized authentication. Use [OpenLDAP](http://www.openldap.org/) for clients and servers.

### #16.1: Kerberos

[Kerberos](http://web.mit.edu/kerberos/) performs authentication as a trusted third party authentication service by using cryptographic shared secret under the assumption that packets traveling along the insecure network can be read, modified, and inserted. Kerberos builds on symmetric-key cryptography and requires a key distribution center. You can make remote login, remote copy, secure inter-system file copying and other high-risk tasks safer and more controllable using Kerberos. So, when users authenticate to network services using Kerberos, unauthorized users attempting to gather passwords by monitoring network traffic are effectively thwarted. See how to setup and use [Kerberos](http://www.redhat.com/docs/en-US/Red_Hat_Enterprise_Linux/5.4/html/Deployment_Guide/ch-kerberos.html).

## #17: Logging and Auditing

You need to configure logging and auditing to collect all hacking and cracking attempts. By default syslog stores data in /var/log/ directory. This is also useful to find out software misconfiguration which may open your system to various attacks. See the following logging related articles:

1. [Linux log file locations](http://www.cyberciti.biz/faq/linux-log-files-location-and-how-do-i-view-logs-files/).
2. [How to send logs to a remote loghost](http://www.cyberciti.biz/tips/log-all-logs-to-central-linux-unix-loghost.html).
3. [How do I rotate log files?](http://www.cyberciti.biz/faq/how-do-i-rotate-log-files/).
4. man pages syslogd, syslog.conf and logrotate.

### #17.1: Monitor Suspicious Log Messages With Logwatch / Logcheck

Read your logs using [logwatch](http://nixcraft.com/linux-software/477-howto-linux-monitor-logfiles.html) or [logcheck](http://logcheck.org/). These tools make your log reading life easier. You get detailed reporting on **unusual items** in syslog via email. A sample syslog report:

################### Logwatch 7.3 (03/24/06) ####################

Processing Initiated: Fri Oct 30 04:02:03 2009

Date Range Processed: yesterday

( 2009-Oct-29 )

Period is day.

Detail Level of Output: 0

Type of Output: unformatted

Logfiles for Host: www-52.nixcraft.net.in

##################################################################

--------------------- Named Begin ------------------------

\*\*Unmatched Entries\*\*

general: info: zone XXXXXX.com/IN: Transfer started.: 3 Time(s)

general: info: zone XXXXXX.com/IN: refresh: retry limit for master ttttttttttttttttttt#53 exceeded (source ::#0): 3 Time(s)

general: info: zone XXXXXX.com/IN: Transfer started.: 4 Time(s)

general: info: zone XXXXXX.com/IN: refresh: retry limit for master ttttttttttttttttttt#53 exceeded (source ::#0): 4 Time(s)

---------------------- Named End -------------------------

--------------------- iptables firewall Begin ------------------------

Logged 87 packets on interface eth0

From 58.y.xxx.ww - 1 packet to tcp(8080)

From 59.www.zzz.yyy - 1 packet to tcp(22)

From 60.32.nnn.yyy - 2 packets to tcp(45633)

From 222.xxx.ttt.zz - 5 packets to tcp(8000,8080,8800)

---------------------- iptables firewall End -------------------------

--------------------- SSHD Begin ------------------------

Users logging in through sshd:

root:

123.xxx.ttt.zzz: 6 times

---------------------- SSHD End -------------------------

--------------------- Disk Space Begin ------------------------

Filesystem Size Used Avail Use% Mounted on

/dev/sda3 450G 185G 241G 44% /

/dev/sda1 99M 35M 60M 37% /boot

---------------------- Disk Space End -------------------------

###################### Logwatch End #########################

(Note output is truncated)

### #17.2: System Accounting with auditd

The auditd is provided for system auditing. It is responsible for writing audit records to the disk. During startup, the rules in /etc/audit.rules are read by this daemon. You can open /etc/audit.rules file and make changes such as setup audit file log location and other option. With auditd you can answers the following questions:

1. System startup and shutdown events (reboot / halt).
2. Date and time of the event.
3. User respoisble for the event (such as trying to access /path/to/topsecret.dat file).
4. Type of event (edit, access, delete, write, update file & commands).
5. Success or failure of the event.
6. Records events that Modify date and time.
7. Find out who made changes to modify the system's network settings.
8. Record events that modify user/group information.
9. See who made changes to a file etc.

See our [quick tutorial which](http://www.cyberciti.biz/tips/linux-audit-files-to-see-who-made-changes-to-a-file.html) explains enabling and using the auditd service.

## #18: Secure OpenSSH Server

The SSH protocol is recommended for remote login and remote file transfer. However, ssh is open to many attacks. See how to secure OpenSSH server:

* [Top 20 OpenSSH Server Best Security Practices](http://www.cyberciti.biz/tips/linux-unix-bsd-openssh-server-best-practices.html).

## #19: Install And Use Intrusion Detection System

A network intrusion detection system (NIDS) is an intrusion detection system that tries to detect malicious activity such as denial of service attacks, port scans or even attempts to crack into computers by monitoring network traffic.

It is a good practice to deploy any integrity checking software before system goes online in a production environment. If possible install AIDE software before the system is connected to any network. [AIDE is a host-based intrusion detection system (HIDS)](http://www.cyberciti.biz/faq/debian-ubuntu-linux-software-integrity-checking-with-aide/) it can monitor and analyses the internals of a computing system.

[Snort is](http://www.snort.org/) a software for intrusion detection which is capable of performing packet logging and real-time traffic analysis on IP networks.

## #20: Protecting Files, Directories and Email

Linux offers excellent protections against unauthorized data access. [File](http://www.cyberciti.biz/faq/unix-linux-bsd-chmod-numeric-permissions-notation-command/) [permissions](http://www.cyberciti.biz/faq/linux-write-protecting-a-file/) and MAC prevent unauthorized access from accessing data. However, permissions set by the Linux are irrelevant if an attacker has physical access to a computer and can simply move the computer's hard drive to another system to copy and analyze the sensitive data. You can easily protect files, and partitons under Linux using the following tools:

* To encrypt and decrypt files with a password, use [gpg command](http://www.cyberciti.biz/tips/linux-how-to-encrypt-and-decrypt-files-with-a-password.html).
* [Linux or UNIX password](http://www.cyberciti.biz/tips/linux-or-unix-password-protecting-files.html) protect files with openssl and other tools.
* See how to [encrypting directories](https://help.ubuntu.com/community/EncryptedPrivateDirectory) with ecryptfs.
* [TrueCrypt is](http://www.truecrypt.org/) free open-source disk encryption software for Windows 7/Vista/XP, Mac OS X and Linux.
* [Howto: Disk and partition encryption in Linux for mobile devices](http://www.cyberciti.biz/tips/file-system-partition-encryption-in-linux.html).
* How to setup [encrypted Swap](https://www.antagonism.org/privacy/encrypted-swap-linux.shtml) on Linux.

### #20.1: Securing Email Servers

You can use SSL certificates and gpg keys to secure email communication on both server and client computers:

* [Linux Securing Dovecot IMAPS / POP3S Server with SSL Configuration](http://www.cyberciti.biz/faq/unix-dovecot-ssl-tls-server-configuration/).
* [Linux Postfix SMTP (Mail Server) SSL Certificate Installations and Configuration](http://www.cyberciti.biz/tips/postfix-smtp-ssl-certificate-csr-installation-guide.html).
* [Courier IMAP SSL Server Certificate Installtion and Configuration](http://www.cyberciti.biz/tips/ssl-certificate-installation-courier-imap-server.html).
* [Configure Sendmail SSL encryption for sending and receiving email](http://www.cyberciti.biz/faq/howto-configure-sendmail-ssl-certificate-email/).
* [Enigmail: Encrypted mail with Mozilla thunderbird](http://enigmail.mozdev.org/home/index.php).

## Other Recommendation:

* [Backups](http://www.cyberciti.biz/faq/redhat-cetos-linux-remote-backup-snapshot-server/) - It cannot be stressed enough how important it is to make a backup of your Linux system. A proper offsite backup allows you to recover from cracked server i.e. an intrusion. The traditional UNIX backup programs are [dump and restore](http://bash.cyberciti.biz/backup/freebsd-dump-filesystem-shell-script/) are also recommended.
* How to: [Looking for Rootkits](http://www.cyberciti.biz/faq/howto-check-linux-rootkist-with-detectors-software/).
* Howto: [Enable ExecShield Buffer Overflows Protection](http://www.cyberciti.biz/faq/what-is-rhel-centos-fedora-core-execshield/).
* Subscribe to [Redhat](http://www.redhat.com/security/) or [Debian](http://www.debian.org/security/) Linux security mailing list or RSS feed.

#### Recommend readings:

1. [Red Hat Enterprise Linux](http://www.redhat.com/docs/manuals/enterprise/RHEL-4-Manual/en-US/Security_Guide/) - Security Guide.
2. [Linux security cookbook](http://www.amazon.com/gp/product/0596003919?ie=UTF8&tag=cyberciti-20&linkCode=as2&camp=1789&creative=390957&creativeASIN=0596003919)- A good collections of security recipes for new Linux admin.
3. [Snort 2.1 Intrusion Detection, Second Edition](http://www.amazon.com/gp/product/1931836043?ie=UTF8&tag=cyberciti-20&linkCode=as2&camp=1789&creative=390957&creativeASIN=1931836043) - Good introduction to Snort and Intrusion detection under Linux.
4. [Hardening Linux](http://www.amazon.com/gp/product/1590594444?ie=UTF8&tag=cyberciti-20&linkCode=as2&camp=1789&creative=390957&creativeASIN=1590594444) - Hardening Linux identifies many of the risks of running Linux hosts and applications and provides practical examples and methods to minimize those risks.
5. [Linux Security](http://tldp.org/HOWTO/html_single/Security-HOWTO/) HOWTO.

In the next part of this series I will discuss how to secure specific applications (such as Proxy, Mail, LAMP, Database) and a few other security tools. Did I miss something? Please add your favorite system security tool or tip in the comments.

# My 10 UNIX Command Line Mistakes

Anyone who has never made a mistake has never tried anything new. -- Albert Einstein.

Here are a few mistakes that I made while working at UNIX prompt. Some mistakes caused me a good amount of downtime. Most of these mistakes are from my early days as a UNIX admin.

## userdel Command

The file /etc/deluser.conf was configured to remove the home directory (it was done by previous sys admin and it was my first day at work) and mail spool of the user to be removed. I just wanted to remove the user account and I end up deleting everything (note -r was activated via deluser.conf):  
userdel foo

## Rebooted Solaris Box

On Linux [killall command](http://www.cyberciti.biz/tips/kill-frozen-crashed-unix-linux-desktop.html) kill processes by name (killall httpd). On Solaris it kill all active processes. As root I killed all process, this was our main Oracle db box:  
killall process-name

## Destroyed named.conf

I wanted to append a [new zone](http://bash.cyberciti.biz/domain/create-bind9-domain-zone-configuration-file/) to /var/named/chroot/etc/named.conf file., but end up running:  
./mkzone example.com > /var/named/chroot/etc/named.conf

## Destroyed Working Backups with Tar and Rsync (personal backups)

I had only one backup copy of my QT project and I just wanted to get a directory called functions. I end up deleting entire backup (note -c switch instead of -x):  
cd /mnt/bacupusbharddisk  
tar -zcvf project.tar.gz functions  
I had no backup. Similarly I end up running rsync command and deleted all new files by overwriting files from backup set (now I’ve switched to [rsnapshot](http://www.cyberciti.biz/faq/linux-rsnapshot-backup-howto/))  
rsync -av -delete /dest /src  
Again, I had no backup.

## Deleted Apache DocumentRoot

I had [sym links](http://www.cyberciti.biz/faq/unix-creating-symbolic-link-ln-command/) for my web server docroot (/home/httpd/http was symlinked to /www). I forgot about symlink issue. To save disk space, I ran rm -rf on http directory. Luckily, I had full working backup set.

## Accidentally Changed Hostname and Triggered False Alarm

Accidentally changed the current hostname (I wanted to see current hostname settings) for one of our cluster node. Within minutes I received an alert message on both mobile and email.  
hostname foo.example.com

## Public Network Interface Shutdown

I wanted to shutdown VPN interface eth0, but ended up shutting down eth1 while I was logged in via SSH:  
ifconfig eth1 down

## Firewall Lockdown

I made changes to sshd\_config and changed the ssh port number from 22 to 1022, but failed to update firewall rules. After a quick kernel upgrade, I had rebooted the box. I had to call remote data center tech to reset firewall settings. (now I use [firewall reset script](http://www.cyberciti.biz/faq/linux-iptables-firewall-flushout-configuration-every-5minutes/) to avoid lockdowns).

## Typing UNIX Commands on Wrong Box

I wanted to shutdown my local Fedora desktop system, but I issued halt on remote server (I was logged into remote box via SSH):  
halt  
service httpd stop

## Wrong CNAME DNS Entry

Created a wrong DNS CNAME entry in example.com zone file. The end result - a few visitors went to /dev/null:  
echo 'foo 86400 IN CNAME lb0.example.com' >> example.com && rndc reload

## Failed To Update Postfix RBL Configuration

In 2006 [ORDB went](http://www.cyberciti.biz/tips/ordborg-rbl-anti-spam-service-going-offline.html) out of operation. But, I failed to update my Postfix RBL settings. One day ORDB was re-activated and it was returning every IP address queried as being on its blacklist. The end result was a disaster.

## Conclusion

All men make mistakes, but only wise men learn from their mistakes -- Winston Churchill.

From all those mistakes I’ve learnt that:

1. Backup = ( Full + Removable tapes (or media) + Offline + Offsite + Tested )
2. The clear choice for preserving all data of UNIX file systems is dump, which is only tool that guaranties recovery under all conditions. (see [Torture-testing Backup and Archive Programs](http://www.coredumps.de/doc/dump/zwicky/testdump.doc.html) paper).
3. Never use rsync with single backup directory. Create a snapshots using rsync or rsnapshots.
4. Use CVS to store configuration files.
5. Wait and read command line again before hitting the dam [Enter] key.
6. Use your well tested perl / shell scripts and open source configuration management software such as puppet, Cfengine or Chef to configure all servers. This also applies to day today jobs such as creating the users and so on.

Mistakes are the inevitable, so did you made any mistakes that have caused some sort of downtime? Please add them into the comments below

# Linux: 20 Iptables Examples For New SysAdmins

Linux comes with a host based firewall called Netfilter. According to the official project site:

netfilter is a set of hooks inside the Linux kernel that allows kernel modules to register callback functions with the network stack. A registered callback function is then called back for every packet that traverses the respective hook within the network stack.

This Linux based firewall is controlled by the program called iptables to handles filtering for IPv4, and ip6tables handles filtering for IPv6. I strongly recommend that you first read our [quick tutorial that explains how to configure a host-based firewall called Netfilter](http://www.cyberciti.biz/faq/rhel-fedorta-linux-iptables-firewall-configuration-tutorial/) (iptables) under CentOS / RHEL / Fedora / Redhat Enterprise Linux. This post list most common iptables solutions required by a new Linux user to secure his or her Linux operating system from intruders.

## IPTABLES Rules Example

* Most of the actions listed in this post are written with the assumption that they will be executed by the root user running the bash or any other modern shell. Do not type commands on remote system as it will disconnect your access.
* For demonstration purpose I've used RHEL 6.x, but the following command should work with any modern Linux distro.
* This is NOT a tutorial on how to set iptables. See [tutorial here](http://www.cyberciti.biz/faq/rhel-fedorta-linux-iptables-firewall-configuration-tutorial/). It is a quick cheat sheet to common iptables commands.

## #1: Displaying the Status of Your Firewall

Type the following command as root:  
# iptables -L -n -v  
Sample outputs:

Chain INPUT (policy ACCEPT 0 packets, 0 bytes)

pkts bytes target prot opt in out source destination

Chain FORWARD (policy ACCEPT 0 packets, 0 bytes)

pkts bytes target prot opt in out source destination

Chain OUTPUT (policy ACCEPT 0 packets, 0 bytes)

pkts bytes target prot opt in out source destination

Above output indicates that the firewall is not active. The following sample shows an active firewall:  
# iptables -L -n -v  
Sample outputs:

Chain INPUT (policy DROP 0 packets, 0 bytes)

pkts bytes target prot opt in out source destination

0 0 DROP all -- \* \* 0.0.0.0/0 0.0.0.0/0 state INVALID

394 43586 ACCEPT all -- \* \* 0.0.0.0/0 0.0.0.0/0 state RELATED,ESTABLISHED

93 17292 ACCEPT all -- br0 \* 0.0.0.0/0 0.0.0.0/0

1 142 ACCEPT all -- lo \* 0.0.0.0/0 0.0.0.0/0

Chain FORWARD (policy DROP 0 packets, 0 bytes)

pkts bytes target prot opt in out source destination

0 0 ACCEPT all -- br0 br0 0.0.0.0/0 0.0.0.0/0

0 0 DROP all -- \* \* 0.0.0.0/0 0.0.0.0/0 state INVALID

0 0 TCPMSS tcp -- \* \* 0.0.0.0/0 0.0.0.0/0 tcp flags:0x06/0x02 TCPMSS clamp to PMTU

0 0 ACCEPT all -- \* \* 0.0.0.0/0 0.0.0.0/0 state RELATED,ESTABLISHED

0 0 wanin all -- vlan2 \* 0.0.0.0/0 0.0.0.0/0

0 0 wanout all -- \* vlan2 0.0.0.0/0 0.0.0.0/0

0 0 ACCEPT all -- br0 \* 0.0.0.0/0 0.0.0.0/0

Chain OUTPUT (policy ACCEPT 425 packets, 113K bytes)

pkts bytes target prot opt in out source destination

Chain wanin (1 references)

pkts bytes target prot opt in out source destination

Chain wanout (1 references)

pkts bytes target prot opt in out source destination

Where,

* **-L** : List rules.
* **-v** : Display detailed information. This option makes the list command show the interface name, the rule options, and the TOS masks. The packet and byte counters are also listed, with the suffix 'K', 'M' or 'G' for 1000, 1,000,000 and 1,000,000,000 multipliers respectively.
* **-n** : Display IP address and port in numeric format. Do not use DNS to resolve names. This will speed up listing.

### #1.1: To inspect firewall with line numbers, enter:

# iptables -n -L -v --line-numbers  
Sample outputs:

Chain INPUT (policy DROP)

num target prot opt source destination

1 DROP all -- 0.0.0.0/0 0.0.0.0/0 state INVALID

2 ACCEPT all -- 0.0.0.0/0 0.0.0.0/0 state RELATED,ESTABLISHED

3 ACCEPT all -- 0.0.0.0/0 0.0.0.0/0

4 ACCEPT all -- 0.0.0.0/0 0.0.0.0/0

Chain FORWARD (policy DROP)

num target prot opt source destination

1 ACCEPT all -- 0.0.0.0/0 0.0.0.0/0

2 DROP all -- 0.0.0.0/0 0.0.0.0/0 state INVALID

3 TCPMSS tcp -- 0.0.0.0/0 0.0.0.0/0 tcp flags:0x06/0x02 TCPMSS clamp to PMTU

4 ACCEPT all -- 0.0.0.0/0 0.0.0.0/0 state RELATED,ESTABLISHED

5 wanin all -- 0.0.0.0/0 0.0.0.0/0

6 wanout all -- 0.0.0.0/0 0.0.0.0/0

7 ACCEPT all -- 0.0.0.0/0 0.0.0.0/0

Chain OUTPUT (policy ACCEPT)

num target prot opt source destination

Chain wanin (1 references)

num target prot opt source destination

Chain wanout (1 references)

num target prot opt source destination

You can use line numbers to delete or insert new rules into the firewall.

### #1.2: To display INPUT or OUTPUT chain rules, enter:

# iptables -L INPUT -n -v  
# iptables -L OUTPUT -n -v --line-numbers

## #2: Stop / Start / Restart the Firewall

If you are using CentOS / RHEL / Fedora Linux, enter:  
# service iptables stop  
# service iptables start  
# service iptables restart  
You can use the iptables command itself to stop the firewall and delete all rules:  
# iptables -F  
# iptables -X  
# iptables -t nat -F  
# iptables -t nat -X  
# iptables -t mangle -F  
# iptables -t mangle -X  
# iptables -P INPUT ACCEPT  
# iptables -P OUTPUT ACCEPT  
# iptables -P FORWARD ACCEPT  
Where,

* **-F** : Deleting (flushing) all the rules.
* **-X** : Delete chain.
* **-t table\_name** : Select table (called nat or mangle) and delete/flush rules.
* **-P** : Set the default policy (such as DROP, REJECT, or ACCEPT).

## #3: Delete Firewall Rules

To display line number along with other information for existing rules, enter:  
# iptables -L INPUT -n --line-numbers  
# iptables -L OUTPUT -n --line-numbers  
# iptables -L OUTPUT -n --line-numbers | less  
# iptables -L OUTPUT -n --line-numbers | grep 202.54.1.1  
You will get the list of IP. Look at the number on the left, then use number to delete it. For example delete line number 4, enter:  
# iptables -D INPUT 4  
OR find source IP 202.54.1.1 and delete from rule:  
# iptables -D INPUT -s 202.54.1.1 -j DROP  
Where,

* **-D** : Delete one or more rules from the selected chain

## #4: Insert Firewall Rules

To insert one or more rules in the selected chain as the given rule number use the following syntax. First find out line numbers, enter:  
# iptables -L INPUT -n --line-numbers  
Sample outputs:

Chain INPUT (policy DROP)

num target prot opt source destination

1 DROP all -- 202.54.1.1 0.0.0.0/0

2 ACCEPT all -- 0.0.0.0/0 0.0.0.0/0 state NEW,ESTABLISHED

To insert rule between 1 and 2, enter:  
# iptables -I INPUT 2 -s 202.54.1.2 -j DROP  
To view updated rules, enter:  
# iptables -L INPUT -n --line-numbers  
Sample outputs:

Chain INPUT (policy DROP)

num target prot opt source destination

1 DROP all -- 202.54.1.1 0.0.0.0/0

2 DROP all -- 202.54.1.2 0.0.0.0/0

3 ACCEPT all -- 0.0.0.0/0 0.0.0.0/0 state NEW,ESTABLISHED

## #5: Save Firewall Rules

To save firewall rules under CentOS / RHEL / Fedora Linux, enter:  
# service iptables save  
In this example, drop an IP and save firewall rules:  
# iptables -A INPUT -s 202.5.4.1 -j DROP  
# service iptables save  
For all other distros use the iptables-save command:  
# iptables-save > /root/my.active.firewall.rules  
# cat /root/my.active.firewall.rules

## #6: Restore Firewall Rules

To restore firewall rules form a file called /root/my.active.firewall.rules, enter:  
# iptables-restore < /root/my.active.firewall.rules  
To restore firewall rules under CentOS / RHEL / Fedora Linux, enter:  
# service iptables restart

## #7: Set the Default Firewall Policies

To drop all traffic:  
# iptables -P INPUT DROP  
# iptables -P OUTPUT DROP  
# iptables -P FORWARD DROP  
# iptables -L -v -n  
#### you will not able to connect anywhere as all traffic is dropped ###  
# ping cyberciti.biz  
# wget http://www.kernel.org/pub/linux/kernel/v3.0/testing/linux-3.2-rc5.tar.bz2

### #7.1: Only Block Incoming Traffic

To drop all incoming / forwarded packets, but allow outgoing traffic, enter:  
# iptables -P INPUT DROP  
# iptables -P FORWARD DROP  
# iptables -P OUTPUT ACCEPT  
# iptables -A INPUT -m state --state NEW,ESTABLISHED -j ACCEPT  
# iptables -L -v -n  
### \*\*\* now ping and wget should work \*\*\* ###  
# ping cyberciti.biz  
# wget http://www.kernel.org/pub/linux/kernel/v3.0/testing/linux-3.2-rc5.tar.bz2

## #8:Drop Private Network Address On Public Interface

IP spoofing is nothing but to stop the following IPv4 address ranges for private networks on your public interfaces. Packets with non-routable source addresses should be rejected using the following syntax:  
# iptables -A INPUT -i eth1 -s 192.168.0.0/24 -j DROP  
# iptables -A INPUT -i eth1 -s 10.0.0.0/8 -j DROP

### #8.1: IPv4 Address Ranges For Private Networks (make sure you block them on public interface)

* 10.0.0.0/8 -j (A)
* 172.16.0.0/12 (B)
* 192.168.0.0/16 (C)
* 224.0.0.0/4 (MULTICAST D)
* 240.0.0.0/5 (E)
* 127.0.0.0/8 (LOOPBACK)

## #9: Blocking an IP Address (BLOCK IP)

To block an attackers ip address called 1.2.3.4, enter:  
# iptables -A INPUT -s 1.2.3.4 -j DROP  
# iptables -A INPUT -s 192.168.0.0/24 -j DROP

## #10: Block Incoming Port Requests (BLOCK PORT)

To block all service requests on port 80, enter:  
# iptables -A INPUT -p tcp --dport 80 -j DROP  
# iptables -A INPUT -i eth1 -p tcp --dport 80 -j DROP

To block port 80 only for an ip address 1.2.3.4, enter:  
# iptables -A INPUT -p tcp -s 1.2.3.4 --dport 80 -j DROP  
# iptables -A INPUT -i eth1 -p tcp -s 192.168.1.0/24 --dport 80 -j DROP

## #11: Block Outgoing IP Address

To block outgoing traffic to a particular host or domain such as cyberciti.biz, enter:  
# host -t a cyberciti.biz  
Sample outputs:

cyberciti.biz has address 75.126.153.206

Note down its ip address and type the following to block all outgoing traffic to 75.126.153.206:  
# iptables -A OUTPUT -d 75.126.153.206 -j DROP  
You can use a subnet as follows:  
# iptables -A OUTPUT -d 192.168.1.0/24 -j DROP  
# iptables -A OUTPUT -o eth1 -d 192.168.1.0/24 -j DROP

### #11.1: Example - Block Facebook.com Domain

First, find out all ip address of facebook.com, enter:  
# host -t a www.facebook.com  
Sample outputs:

www.facebook.com has address 69.171.228.40

Find CIDR for 69.171.228.40, enter:  
# whois 69.171.228.40 | grep CIDR  
Sample outputs:

CIDR: 69.171.224.0/19

To prevent outgoing access to www.facebook.com, enter:  
# iptables -A OUTPUT -p tcp -d 69.171.224.0/19 -j DROP  
You can also use domain name, enter:  
# iptables -A OUTPUT -p tcp -d www.facebook.com -j DROP  
# iptables -A OUTPUT -p tcp -d facebook.com -j DROP

From the iptables man page:

... specifying any name to be resolved with a remote query such as DNS (e.g., facebook.com is a really bad idea), a network IP address (with /mask), or a plain IP address ...

## #12: Log and Drop Packets

Type the following to log and block IP spoofing on public interface called eth1  
# iptables -A INPUT -i eth1 -s 10.0.0.0/8 -j LOG --log-prefix "IP\_SPOOF A: "  
# iptables -A INPUT -i eth1 -s 10.0.0.0/8 -j DROP  
By default everything is logged to /var/log/messages file.  
# tail -f /var/log/messages  
# grep --color 'IP SPOOF' /var/log/messages

## #13: Log and Drop Packets with Limited Number of Log Entries

The -m limit module can limit the number of log entries created per time. This is used to prevent flooding your log file. To log and drop spoofing per 5 minutes, in bursts of at most 7 entries .  
# iptables -A INPUT -i eth1 -s 10.0.0.0/8 -m limit --limit 5/m --limit-burst 7 -j LOG --log-prefix "IP\_SPOOF A: "  
# iptables -A INPUT -i eth1 -s 10.0.0.0/8 -j DROP

## #14: Drop or Accept Traffic From Mac Address

Use the following syntax:  
# iptables -A INPUT -m mac --mac-source 00:0F:EA:91:04:08 -j DROP  
## \*only accept traffic for TCP port # 8080 from mac 00:0F:EA:91:04:07 \* ##  
# iptables -A INPUT -p tcp --destination-port 22 -m mac --mac-source 00:0F:EA:91:04:07 -j ACCEPT

## #15: Block or Allow ICMP Ping Request

Type the following command to block ICMP ping requests:  
# iptables -A INPUT -p icmp --icmp-type echo-request -j DROP  
# iptables -A INPUT -i eth1 -p icmp --icmp-type echo-request -j DROP  
Ping responses can also be limited to certain networks or hosts:  
# iptables -A INPUT -s 192.168.1.0/24 -p icmp --icmp-type echo-request -j ACCEPT  
The following only accepts limited type of ICMP requests:  
### \*\* assumed that default INPUT policy set to DROP \*\* #############  
iptables -A INPUT -p icmp --icmp-type echo-reply -j ACCEPT  
iptables -A INPUT -p icmp --icmp-type destination-unreachable -j ACCEPT  
iptables -A INPUT -p icmp --icmp-type time-exceeded -j ACCEPT  
## \*\* all our server to respond to pings \*\* ##  
iptables -A INPUT -p icmp --icmp-type echo-request -j ACCEPT

## #16: Open Range of Ports

Use the following syntax to open a range of ports:  
iptables -A INPUT -m state --state NEW -m tcp -p tcp --dport 7000:7010 -j ACCEPT

## #17: Open Range of IP Addresses

Use the following syntax to open a range of IP address:  
## only accept connection to tcp port 80 (Apache) if ip is between 192.168.1.100 and 192.168.1.200 ##  
iptables -A INPUT -p tcp --destination-port 80 -m iprange --src-range 192.168.1.100-192.168.1.200 -j ACCEPT

## nat example ##  
iptables -t nat -A POSTROUTING -j SNAT --to-source 192.168.1.20-192.168.1.25

## #18: Established Connections and Restaring The Firewall

When you restart the iptables service it will drop established connections as it unload modules from the system under RHEL / Fedora / CentOS Linux. Edit, /etc/sysconfig/iptables-config and set IPTABLES\_MODULES\_UNLOAD as follows:

IPTABLES\_MODULES\_UNLOAD = no

## #19: Help Iptables Flooding My Server Screen

Use the crit log level to send messages to a log file instead of console:  
iptables -A INPUT -s 1.2.3.4 -p tcp --destination-port 80 -j LOG --log-level crit

## #20: Block or Open Common Ports

The following shows syntax for opening and closing common TCP and UDP ports:

Replace ACCEPT with DROP to block port:

*## open port* ***ssh*** *tcp port 22 ##*

iptables -A INPUT -m state --state NEW -m tcp -p tcp --dport 22 -j ACCEPT

iptables -A INPUT -s 192.168.1.0/24 -m state --state NEW -p tcp --dport 22 -j ACCEPT

*## open cups* ***(****printing service****)*** *udp/tcp port 631* ***for*** *LAN* ***users*** *##*

iptables -A INPUT -s 192.168.1.0/24 -p udp -m udp --dport 631 -j ACCEPT

iptables -A INPUT -s 192.168.1.0/24 -p tcp -m tcp --dport 631 -j ACCEPT

*## allow* ***time******sync*** *via NTP* ***for*** *lan* ***users******(****open udp port 123****)*** *##*

iptables -A INPUT -s 192.168.1.0/24 -m state --state NEW -p udp --dport 123 -j ACCEPT

*## open tcp port 25* ***(****smtp****)******for*** *all ##*

iptables -A INPUT -m state --state NEW -p tcp --dport 25 -j ACCEPT

*# open dns server ports* ***for*** *all ##*

iptables -A INPUT -m state --state NEW -p udp --dport 53 -j ACCEPT

iptables -A INPUT -m state --state NEW -p tcp --dport 53 -j ACCEPT

*## open http/https* ***(****Apache****)*** *server port to all ##*

iptables -A INPUT -m state --state NEW -p tcp --dport 80 -j ACCEPT

iptables -A INPUT -m state --state NEW -p tcp --dport 443 -j ACCEPT

*## open tcp port 110* ***(****pop3****)******for*** *all ##*

iptables -A INPUT -m state --state NEW -p tcp --dport 110 -j ACCEPT

*## open tcp port 143* ***(****imap****)******for*** *all ##*

iptables -A INPUT -m state --state NEW -p tcp --dport 143 -j ACCEPT

*## open access to Samba* ***file*** *server* ***for*** *lan* ***users*** *only ##*

iptables -A INPUT -s 192.168.1.0/24 -m state --state NEW -p tcp --dport 137 -j ACCEPT

iptables -A INPUT -s 192.168.1.0/24 -m state --state NEW -p tcp --dport 138 -j ACCEPT

iptables -A INPUT -s 192.168.1.0/24 -m state --state NEW -p tcp --dport 139 -j ACCEPT

iptables -A INPUT -s 192.168.1.0/24 -m state --state NEW -p tcp --dport 445 -j ACCEPT

*## open access to proxy server* ***for*** *lan* ***users*** *only ##*

iptables -A INPUT -s 192.168.1.0/24 -m state --state NEW -p tcp --dport 3128 -j ACCEPT

*## open access to mysql server* ***for*** *lan* ***users*** *only ##*

iptables -I INPUT -p tcp --dport 3306 -j ACCEPT

## #21: Restrict the Number of Parallel Connections To a Server Per Client IP

You can use connlimit module to put such restrictions. To allow 3 ssh connections per client host, enter:  
# iptables -A INPUT -p tcp --syn --dport 22 -m connlimit --connlimit-above 3 -j REJECT

Set HTTP requests to 20:  
# iptables -p tcp --syn --dport 80 -m connlimit --connlimit-above 20 --connlimit-mask 24 -j DROP  
Where,

1. **--connlimit-above 3** : Match if the number of existing connections is above 3.
2. **--connlimit-mask 24** : Group hosts using the prefix length. For IPv4, this must be a number between (including) 0 and 32.

## #22: HowTO: Use iptables Like a Pro

For more information about iptables, please see the manual page by typing man iptables from the command line:  
$ man iptables  
You can see the help using the following syntax too:  
# iptables -h  
To see help with specific commands and targets, enter:  
# iptables -j DROP -h

### #22.1: Testing Your Firewall

Find out if ports are open or not, enter:  
# netstat -tulpn  
Find out if tcp port 80 open or not, enter:  
# netstat -tulpn | grep :80  
If port 80 is not open, start the Apache, enter:  
# service httpd start  
Make sure iptables allowing access to the port 80:  
# iptables -L INPUT -v -n | grep 80  
Otherwise open port 80 using the iptables for all users:  
# iptables -A INPUT -m state --state NEW -p tcp --dport 80 -j ACCEPT  
# service iptables save  
Use the telnet command to see if firewall allows to connect to port 80:  
$ telnet www.cyberciti.biz 80  
Sample outputs:

Trying 75.126.153.206...

Connected to www.cyberciti.biz.

Escape character is '^]'.

^]

telnet> quit

Connection closed.

You can use nmap to probe your own server using the following syntax:  
$ nmap -sS -p 80 www.cyberciti.biz  
Sample outputs:

Starting Nmap 5.00 ( http://nmap.org ) at 2011-12-13 13:19 IST

Interesting ports on www.cyberciti.biz (75.126.153.206):

PORT STATE SERVICE

80/tcp open http

Nmap done: 1 IP address (1 host up) scanned in 1.00 seconds

# 20 Examples: Make Sure Unix / Linux Configuration Files Are Free From Syntax Errors

In Linux and UNIX in general, services are configured using various text files,in the /etc/ directory trees. A typical server system could have dozens of configuration files. It is important that you check the validity of the configuration file. In some cases it is possible to to check the sanity of the special data (such as keys) or directories (such as /var/lib/cache/). Text files are easier to manage remotely. You can use ssh and a text editor. If there is an error in configuration, server may not start. This may result into a disaster. In this article I will explains howto find out a syntax error for popular servers and test configuration file for syntax errors.

## Stop! Test Your Server Configuration Before Restarting Unix / Linux Services

1. The following option(s) will **not run server (or stop running server)**, it will just **test the configuration file** and then exit.
2. It will check configuration for **correct syntax and then try to open files** referred in configuration.
3. The **config file is parsed** and checked for syntax errors, **along with all files included config file(s)** by server.
4. In most case you can also **specify which configuration file server should use** instead of the default.
5. Once you've **verified your configuration files** and fixed any errors you can **go ahead and reload or restart** required services.

### A Note About Reloading Servers

The syntax is as follows under Linux:  
/sbin/service SERVICE-NAME [reload|restart]  
OR  
/etc/init.d/SERVICE-NAME [reload|restart]  
The reload option reloads the config file without interrupting pending operations. For example the following command will reload Apache web server after the config file changes:  
# /sbin/service httpd reload  
However, most Unix / Linux daemon programs sometimes use [SIGHUP as a signal](http://bash.cyberciti.biz/guide/Sending_signal_to_Processes) to restart themselves, the most common reason for this being to re-read a configuration file that has been changed. The syntax is as follows:  
kill -HUP $(cat /var/run/SERVICE.pid)  
OR  
kill -HUP `cat /var/run/SERVICE.pid`

## #1: OpenSSH Server

You can use the [following syntax to test OpenSSH config file](http://www.cyberciti.biz/tips/checking-openssh-sshd-configuration-syntax-errors.html), type:  
# /usr/sbin/sshd -t && echo $?  
A sample configuration error session:  
# usr/sbin/sshd -t  
Sample outputs:

/etc/ssh/sshd\_config line 26: Bad yes/without-password/forced-commands-only/no argument: Naa

To print line # 26, enter:  
# sed -n '26p' /etc/ssh/sshd\_config  
Sample outputs:

PermitRootLogin Naa

Use a text editor such as vi to edit the file, enter:  
# vi +26 etc/ssh/sshd\_config  
Finally update the syntax, enter:

PermitRootLogin No

Save and close the file. Test it again:  
# /usr/sbin/sshd -t

### OpenSSH Extended Test Mode

Use the -T option to check the validity of the configuration file, output the effective configuration to stdout (screen) and then exit:  
# /usr/sbin/sshd -T  
See also: [Openssh server security best practices](http://www.cyberciti.biz/tips/linux-unix-bsd-openssh-server-best-practices.html) for more information.

## #2: Apache Web Server

The syntax is as follows to run syntax tests for configuration files only:  
# /usr/sbin/apache2 -t  
Sample error reporting:

apache2: Syntax error on line 50 of /etc/apache2/apache2.conf: ServerRoot must be a valid directory

On RHEL and friend, enter:  
# /usr/sbin/httpd -t  
Sample outputs:

Syntax OK

You can also use the apachectl command (pass the configtest or -t option). It will run a configuration file syntax test. It parses the configuration files and either reports Syntax Ok or detailed information about the particular syntax error:  
# apachectl configtest  
OR  
# apachectl -t  
[Reload Apache server](http://www.cyberciti.biz/faq/apache-making-changes-config-file-without-restarting/), enter:  
# apachectl -k graceful

## #3: Nginx Web Server

To run syntax tests for nginx configuration files, enter:  
# /usr/local/nginx/sbin/nginx -t  
# /usr/local/nginx/sbin/nginx -t -c /usr/local/nginx/conf/nginx.conf  
Sample outputs:

nginx: the configuration file /usr/local/nginx/conf/nginx.conf syntax is ok

nginx: configuration file /usr/local/nginx/conf/nginx.conf test is successful

Where,

* **-c /path/to/file** : Specify which configuration file Nginx should use instead of the default.
* **-t** : Just test the configuration file.

See also: Howto [set nginx as a reverse proxy server](http://www.cyberciti.biz/faq/rhel-centos-fedora-keepalived-lvs-cluster-configuration/) and [top 20 nginx best security practices](http://www.cyberciti.biz/tips/linux-unix-bsd-nginx-webserver-security.html) for more information.

## #4: Lighttpd Web Server

To test the config-file, and exit, enter:  
# /usr/local/sbin/lighttpd -t -f /usr/local/etc/lighttpd/cyberciti.biz/lighttpd.conf  
Sample outputs:

Syntax OK

Where,

* **-f filename** : Use filename of the config-file.
* **-t** : Test config-file.

## #5: BIND (named) DNS Server

Use [named-checkconf command to check the syntax](http://www.cyberciti.biz/tips/howto-linux-unix-check-dns-file-errors.html), but not the semantics. The file is parsed and checked for syntax errors, along with all files included by it  
# named-checkconf /etc/named.conf  
You can also [check bind zone files](http://www.cyberciti.biz/faq/howto-linux-unix-zone-file-validity-checking/), enter:  
# named-checkzone cyberciti.biz /var/named/zone.cyberciti.biz

## #6: Squid Proxy Server

To parse and test configuration file, enter:  
# /usr/sbin/squid -k check  
# /usr/sbin/squid -k parse  
Sample outputs:

2012/03/30 07:44:35| Processing Configuration File: /etc/squid/squid.conf (depth 0)

2012/03/30 07:44:35| Initializing https proxy context

## #7: MySQL (mysqld) Database Server

Type the following command:  
# mysqld --verbose --help  
The above will procduce too much output. I recommend redirecting output to /dev/null and only display error/warning on the screen  
# /usr/libexec/mysqld --verbose --help 1>/dev/null  
Sample outputs:

120330 7:52:43 [Warning] '--log\_slow\_queries' is deprecated and will be removed in a future release. Please use ''--slow\_query\_log'/'--slow\_query\_log\_file'' instead.

You can specify a new configuration file such as /root/test-my.cnf  
# mysqld --defaults-file=/root/test-my.cnf --verbose --help 1>/dev/null

## #8: Postfix Mail Server (MTA)

Use the following syntax. To warn about bad directory/file ownership or permissions, and create missing directories, enter:  
# postfix check  
OR  
# postfix -vvv  
Sample outputs:

postfix: dict\_register: mail\_dict 1

postfix: dict\_update: config\_directory = /etc/postfix

postfix: dict\_update: queue\_directory = /var/spool/postfix

postfix: dict\_update: command\_directory = /usr/sbin

postfix: dict\_update: daemon\_directory = /usr/libexec/postfix

postfix: dict\_update: data\_directory = /var/lib/postfix

postfix: dict\_update: mail\_owner = postfix

postfix: dict\_update: inet\_interfaces = localhost

postfix: dict\_update: inet\_protocols = all

postfix: dict\_update: mydestination = $myhostname, localhost.$mydomain, localhost

postfix: dict\_update: unknown\_local\_recipient\_reject\_code = 550

postfix: fatal: /etc/postfix/main.cf, line 385: missing '=' after attribute name: "sss"

You can see errors in maillog log file, enter:  
# tail -f /var/log/maillog  
Sample outputs:

And it'll run mysqld (or drizzled), parse tMar 30 08:01:34 mx421 postfix[2284]: dict\_update: command\_directory = /usr/sbin

Mar 30 08:01:34 mx421 postfix[2284]: dict\_update: daemon\_directory = /usr/libexec/postfix

Mar 30 08:01:34 mx421 postfix[2284]: dict\_update: data\_directory = /var/lib/postfix

Mar 30 08:01:34 mx421 postfix[2284]: dict\_update: mail\_owner = postfix

Mar 30 08:01:34 mx421 postfix[2284]: dict\_update: inet\_interfaces = localhost

Mar 30 08:01:34 mx421 postfix[2284]: dict\_update: inet\_protocols = all

Mar 30 08:01:34 mx421 postfix[2284]: dict\_update: mydestination = $myhostname, localhost.$mydomain, localhost

Mar 30 08:01:34 mx421 postfix[2284]: dict\_update: unknown\_local\_recipient\_reject\_code = 550

Mar 30 08:01:34 mx421 postfix[2284]: fatal: /etc/postfix/main.cf, line 385: missing '=' after attribute name: "sss"

Mar 30 08:01:42 mx421 postfix[2285]: fatal: /etc/postfix/main.cf, line 385: missing '=' after attribute name: "sss"he config, report any problems, print help, and exit without initializing storage engines or trying to

## #9: Samba (SMB/CIFS) File Server

Type the following command:  
# testparm -v

## #10: tcpd

The tcpd program can be set up to monitor incoming requests for telnet, finger, ftp, exec, rsh, rlogin, tftp, talk, comsat and other services that have a one-to-one mapping onto executable files. The tcpdchk command examines your tcp wrapper configuration and reports all potential and real problems it can find:  
# tcpdchk  
# tcpdchk -a  
# tcpdchk -d  
# tcpdchk -i /path/to/inetd.conf  
# tcpdchk -v  
Where,

* **-a** : Report access control rules that permit access without an explicit ALLOW keyword.
* **-d** : Examine hosts.allow and hosts.deny files in the current directory instead of the default ones.
* **-i** inet\_conf : Specify this option when tcpdchk is unable to find your inetd.conf network configuration file, or when you suspect that the program uses the wrong one.
* **-v** : Display the contents of each access control rule. Daemon lists, client lists, shell commands and options are shown in a pretty-printed format; this makes it easier for you to spot any discrepancies between what you want and what the program understands.

## #11: dhcpd Server

The Internet Systems Consortium DHCP Server, dhcpd, implements the Dynamic Host Configuration Protocol (DHCP) and the Internet Bootstrap Protocol (BOOTP). DHCP allows hosts on a TCP/IP network to request and be assigned IP addresses, and also to discover information about the network to which they are attached. BOOTP provides similar functionality, with certain restrictions. To test syntax pass the following option:  
# dhcpd -t  
OR  
# dhcpd -t -cf /path/to/dhcpd.testing.conf  
OR  
# dhcpd -T   
OR  
# dhcpd -T -lf /path/to/dhcpd.lease.file

Where,

* **-t** : The -t flag is specified, the server will simply test the configuration file for correct syntax, but will not attempt to perform any network operations. This can be used to test the a new configuration file automatically before installing it.
* **-T** : This flag can be used to test the lease database file in a similar way.
* **-cf /path/to/dhcpd.testing.conf** : Use an alternate configuration file called /path/to/dhcpd.testing.conf.
* **-lf /path/to/dhcpd.lease.file** : an alternate lease file called -lf /path/to/dhcpd.lease.file

## #12: vsftpd FTP Server

vsftpd is the Very Secure File Transfer Protocol daemon. Use the following command to make sure configuration file is free from syntax errors:  
# vsftpd  
OR  
# vsftpd -olisten=NO /path/to/vsftpd.testing.conf   
That example overrides vsftpd's built-in default for the "listen" option to be NO, but then loads settings from /path/to/vsftpd.testing.conf.

## #13: Nagios

Nagios is a popular open source computer system monitor, network monitoring and infrastructure monitoring software application. Use the following syntax to run a sanity check on nagios.cfg as follows:  
# nagios -v /path/to/testing/nagios.cfg  
Where,

* **-v** : Verify your configuration.

## #14: Openntpd Server

The [ntpd daemon synchronizes](http://www.cyberciti.biz/faq/openntpd-on-centos-rhel-fedora-linux/) the local clock to one or more remote NTP servers or local timedelta sensors. [ntpd can also act as an NTP server](http://www.cyberciti.biz/faq/freebsd-install-configure-openntpd-server-to-sync-time/) itself, redistributing the local time. Use the following syntax to run a sanity check on ntpd.conf:  
# ntpd -n  
# ntpd -f /usr/local/etc/ntpd.conf -n  
# ntpd -d -f /usr/local/etc/ntpd.conf -n  
Where,

* **-n** : Only check the configuration file for validity.
* **-f /usr/local/etc/ntpd.conf** : Use /usr/local/etc/ntpd.conf file as the configuration file, instead of the default /etc/ntpd.conf.
* **-d** : Do not daemonize and ntpd will run in the foreground and log to [stderr](http://bash.cyberciti.biz/guide/Input_and_Output).

## #15: Xorg - The X11 Server

Linux and Unix like operating systems uses X11 to provide users with a powerful graphical user interface. X11 is a freely available version of the X Window System that is implemented in Xorg. The default xorg.conf is located in /etc/X11 directory. You can build an initial configuration file by simply running the following command:  
# Xorg -configure  
To test the existing configuration to verify that Xorg can work with the graphics hardware on the target system, enter:  
# Xorg -config /path/to/xorg.conf.new -retro  
Please note that the new Xorg is largely self-configuring and doesn't need one. But, if you use proprietary (say Nvidia) or other drives you need to test Xorg syntax using the above described method.

## #16: syslogd / rsyslogd

syslogd is Unix / Linux system logging server. rsyslogd is reliable and extended syslogd for modern Linux distros. Rsyslogd is derived from the sysklogd package which in turn is derived from the stock BSD sources. To check for syntax error, type:  
# syslogd -f /etc/rsyslog.testing.conf -d  
OR  
rsyslogd -c4 -f /etc/rsyslog.testing.conf -N 1  
Sample outputs:

rsyslogd: version 4.6.4, config validation run (level 1), master config /etc/rsyslog.conf

rsyslogd: invalid or yet-unknown config file command - have you forgotten to load a module? [try http://www.rsyslog.com/e/3003 ]

**rsyslogd: the last error occured in /etc/rsyslog.conf, line 11:"$FilesOnwer root"**

rsyslogd: CONFIG ERROR: could not interpret master config file '/etc/rsyslog.testing.conf'. [try http://www.rsyslog.com/e/2124 ]

A sanity check without any errors:

rsyslogd: version 4.6.4, config validation run (level 1), master config /etc/rsyslog.testing.conf

rsyslogd: End of config validation run. Bye.

Where,

* **-c4** : Selects the desired backward compatibility mode (in this example it is # 4).
* **-f /etc/rsyslog.testing.conf** : Specify an alternative configuration file instead of /etc/rsyslog.conf, which is the default.
* **-d** : Debug mode (only use with syslogd)
* **-N 1** : Do a config file check. Do NOT run in regular mode, just check configuration file correctness. This option is meant to verify a config file. The level argument modifies behaviour. Currently, 0 is the same as not specifying the -N option at all (so this makes limited sense) and 1 actually activates the code. Later, higher levels will mean more verbosity (this is a forward-compatibility option).

## #17: CUPS Printing System

CUPS is the standards-based, open source printing system developed by Apple, for Mac OS X and other UNIX/Linux-like operating systems. cupsd is the scheduler for CUPS. It implements a printing system based upon the Internet Printing Protocol, version 2.1. To test config files for error type:  
# cupsd -f -c /path/to/cupsd.testing.conf -t  
Sample outputs:

**Unknown directive Loggslevel on line 6.**

/etc/cups/cupsd.conf is OK

A sanity check without any errors:

/etc/cups/cupsd.conf is OK

Where,

* **-f** : Run cupsd in the foreground; the default is to run in the background as a "daemon".
* **-c /path/to/cupsd.testing.conf** : Uses the /path/to/cupsd.testing.conf configuration file.
* **-t** :  
  Test the configuration file for syntax errors.

## #18: slapd Stand-alone LDAP Daemon

OpenLDAP Software is a free, open source implementation of the Lightweight Directory Access Protocol (LDAP) developed by the OpenLDAP Project. To test whether the configuration file is correct or not, type:  
# slapd -Tt   
OR  
# slapd -Tt -f /path/to/slapd-config.conf  
Where,

* **-Tt** : Test for config file errors.
* **-f /path/to/slapd-config.conf** : Specifies the slapd configuration file. The default is /etc/openldap/slapd.conf.

## #19: varnishd - HTTP Accelerator Daemon

To test varnishd vlc syntax, enter:  
# varnishd -C -f /path/to/wordpress.vlc  
Where,

* **-C** : Print VCL code compiled to C language and exit. Specify the VCL file to compile with the -f option.
* **-F /path/to/wordpress.vlc** : Use the specified VCL configuration file instead of the builtin default.

## #20: exim MTA

Pass the -bV option to exim. If there are any errors in the configuration file, Exim outputs error messages. Otherwise it outputs the version number and build date, the DBM library that is being used, and information about which drivers and other optional code modules are included in the binary.  
# exim -bV  
Some simple routing tests can be done by using the address testing option. For example,  
# exim -bt <local username>  
Admin users can test the malware scanning configuration (in Exim >= 4.73) with the -bmalware option:  
# exim -bmalware <filename>  
Another source of information is running Exim with debugging turned on, by specifying the -d option. If a message is stuck on Exim's spool, you can force a delivery with debugging turned on by a command of the form  
# exim -d -M <exim-message-id>

## #21: Bonus tips

Some additional tips:

### Bash / KSH Shell Scripts

It is possible to check a bash script syntax without executing it:  
$ bash -n ./myscript  
Sample outputs:

./myscript<: line 16: syntax error near unexpected token `fi'

./myscript<: line 16: `fi'

OR  
$ ksh -n /path/to/backup.ksh  
See [how to debug a shell script under Linux or UNIX](http://www.cyberciti.biz/tips/debugging-shell-script.html) for more information.

### FreeBSD/OpenBSD/Apple OS X pf Firewall

The [pfctl command used to control the pf firewall and NAT devices](http://www.cyberciti.biz/faq/freebsd-openbsd-reset-pf-firewall-automatically/) under \*BSD operating system including Apple OS X. The syntax is as follows to test for errors:  
# pfctl -nf /etc/pf.conf

Where,

* **-n** : Do not load rules from /path/to/pf.testing.conf, just parse them.
* **-f /path/to/pf.testing.conf** : Load the firewall rules contained in a file called /path/to/pf.testing.conf.

## Command Summary

|  |  |
| --- | --- |
| **Daemon** | **Command** |
| OpenSSH | /usr/sbin/sshd -t && echo $? /usr/sbin/sshd -T |
| Apache | /usr/sbin/apache2 -t apachectl configtest |
| nginx | /usr/local/nginx/sbin/nginx -t /usr/local/nginx/sbin/nginx -t -c /usr/local/nginx/conf/nginx.conf |
| lighttpd | /usr/local/sbin/lighttpd -t -f /usr/local/etc/lighttpd/cyberciti.biz/lighttpd.conf |
| Bind (named server config) | named-checkconf /etc/named.conf |
| Bind (zone syntx) | named-checkzone cyberciti.biz /var/named/zone.cyberciti.biz |
| Squid proxy | /usr/sbin/squid -k check /usr/sbin/squid -k parse |
| MySQL server | mysqld --verbose --help /usr/libexec/mysqld --verbose --help 1>/dev/null |
| Postfix MTA | postfix check postfix -vvv |
| Samba SMB/CIFS | testparm -v |
| tcpd | tcpdchk tcpdchk -v |
| dhcpd (DHCP / BOOTP) server | dhcpd -t -cf /path/to/dhcpd.testing.conf |
| vsftpd server | vsftpd -olisten=NO /path/to/vsftpd.testing.conf |
| nagios | nagios -v /path/to/testing/nagios.cfg |
| Openntpd NTPD server | ntpd -d -f /usr/local/etc/ntpd.conf -n |
| Xorg (X11 Server) | Xorg -config /path/to/xorg.conf.new -retro |
| syslogd / rsyslogd | rsyslogd -c4 -f /etc/rsyslog.testing.conf -N 1 |
| CUPS Printing System | cupsd -f -c /path/to/cupsd.testing.conf -t |
| slapd (OpenLDAP) | slapd -Tt |
| varnishd | varnishd -C -f /path/to/wordpress.vlc |
| exim MTA | exim -bV |
| Bash/Ksh scripts | bash -n ./myscript ksh -n /path/to/script.ksh |
| BSD pf firewall | pfctl -nf /etc/pf.conf |
| proftpd | proftpd -t -c /path/to/proftpd.testing.conf |
| Perl scripts | perl -c /path/to/script.pl perl -wc /path/to/script.pl |

I hope you find these tips useful. Have a favorite command to check the Unix / Linux / BSD server config file for any syntax errors and/or warnings? Let's hear about it in the comments

# Linux: Force Users To Change Their Passwords Upon First Login

by [nixCraft](http://www.cyberciti.biz/tips/about-us) on September 11, 2008 · [6 comments](http://www.cyberciti.biz/faq/rhel-debian-force-users-to-change-passwords/#comments)· LAST UPDATED September 11, 2008

in [CentOS](http://www.cyberciti.biz/faq/category/centos/), [Debian / Ubuntu](http://www.cyberciti.biz/faq/category/debian-ubuntu/), [Fedora Linux](http://www.cyberciti.biz/faq/category/fedora-linux/)

Q. How can I force my Linux users to change their passwords upon the first login under CentOS / Debian Linux?  
  
A. You can use any one of the following command to change user passwords upon the first login:  
[a] **usermod command** - Modify various user account properties including user password expiry information.  
Where,

[b] **chage command** - Change user password expiry information

## Task: Use chage command to force users to chage their password upon first login

Use the following syntax:  
chage -d 0 {user-name}  
# chage -d 0 tom

* **-d 0** : Set the number of days since January 1st, 1970 when the password was last changed. The date may also be expressed in the format YYYY-MM-DD. By setting it to zero, you are going to force user to change password upon first login

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | |  |  |  | | --- | --- | --- | | * [HR](http://www.careerride.com/Interview-Questions-Human-Resources.aspx) * [Job interviews](http://www.careerride.com/Job-Interview-Tips-Question.aspx) * [Soft skills](http://www.careerride.com/Soft-Skills.aspx) * [Group discussion](http://www.careerride.com/Group-Discussion.aspx) * [Working from home](http://www.careerride.com/Work-from-home.aspx) | | | | http://www.careerride.com/images/44.gif | http://www.careerride.com/images/transperent.gif | http://www.careerride.com/images/33.gif | | | |  | | --- | | Linux interview questions and answers | | |  |  |  | | --- | --- | --- | | |  |  | | --- | --- | | Linux interview questions posted by [Rupali Kapil](http://www.careerride.com/profile.aspx?Name=Rupali%20kapil&Id=3678)  [Which command is used to check the number of files and disk space used and the each user’s defined quota?](http://www.careerride.com/Linux-Interview-Questions.aspx#linux1) [What is the name and path of the main system log?](http://www.careerride.com/Linux-Interview-Questions.aspx#linux2) [How secured is Linux? Explain.](http://www.careerride.com/Linux-Interview-Questions.aspx#linux3)  [Can Linux computer be made a router so that several machines may share a single Internet connection? How?](http://www.careerride.com/Linux-Interview-Questions.aspx#linux4) [What is the minimum number of partitions you need to install Linux?](http://www.careerride.com/Linux-Interview-Questions.aspx#linux5) [Which command is used to review boot messages?](http://www.careerride.com/Linux-Interview-Questions.aspx#linux6) [Which utility is used to make automate rotation of a log?](http://www.careerride.com/Linux-Interview-Questions.aspx#linux7) [What are the partitions created on the mail server hard drive?](http://www.careerride.com/Linux-Interview-Questions.aspx#linux8) [What are the fields in the/etc/passwd file?](http://www.careerride.com/Linux-Interview-Questions.aspx#linux9) [Which commands are used to set a processor-intensive job to use less CPU time?](http://www.careerride.com/Linux-Interview-Questions.aspx#linux10) [How to change window manager by editing your home directory?](http://www.careerride.com/Linux-Interview-Questions.aspx#linux11) [How documentation of an application is stored?](http://www.careerride.com/Linux-Interview-Questions.aspx#linux12) [How shadow passwords are given?](http://www.careerride.com/Linux-Interview-Questions.aspx#linux13) [How do you create a new user account?](http://www.careerride.com/Linux-Interview-Questions.aspx#linux14) [Which password package is installed for the security of central password?](http://www.careerride.com/Linux-Interview-Questions.aspx#linux15) [Which shell do you assign to a POP3 mail-only account?](http://www.careerride.com/Linux-Interview-Questions.aspx#linux16) [Which daemon is responsible for tracking events on Linux system?](http://www.careerride.com/Linux-Interview-Questions.aspx#linux17) [Which daemon is used for scheduling of the commands?](http://www.careerride.com/Linux-Interview-Questions.aspx#linux18) [How environment variable is set so that the file permission can be automatically set to the newly created files?](http://www.careerride.com/Linux-Interview-Questions.aspx#linux19) |  |   [**Test your linux knowledge with our multiple choice questions!**](http://www.careerride.com/linux-multiple-choice-questions.aspx)  [Linux part 1 (27 questions)](http://www.careerride.com/test.aspx?type=Linux) [Linux part 2 (25 questions)](http://www.careerride.com/test.aspx?type=Linux-1) [Linux part 3 (25 questions)](http://www.careerride.com/test.aspx?type=Linux-2)  [**Download Linux interview questions and answers**](http://www.careerride.com/Career-Jobs-Interview-Updates.aspx)  Linux interview questions - posted on June 27, 2013 at 16:25 PM by Kshipra Singh 1. Which account is created on Linux installation? - With the installation of Linux, a super user account is created called as ‘root’. 2. Which daemon tracks events on your system? - The syslogd daemon tracks the system information and saves it to specified log files. 3. Which command would you use if you want to remove the password assigned to a group? - gpasswd – r removes the password from the group.  - Here, the gpasswd changes the password of the group and when it is accompanied by –r, the password gets removed. 4. You wish to print a file ‘draft’ with 60 lines to a page. What command would you use? - The command that I would use is: pr -l60 draft - The default page length when using pr is 66 lines.  - The -l option specifies a different length. 5. Which file would you examine to determine the levels of messages written to system log files? - kernel.h 6. You are logged on as a regular user. Without logging off and logging on as root, you are required to create a new user account immediately. How would you do it? - This can be achieved by issuing the su command.  - This will prompt you for the password of the root account. - Providing the password, logs you in as root. Now, you can perform any administrative duties. . 7. You are required to restore the file memo.ben. It was backed up in the tar file MyBackup.tar. Which command would you use to do it? - The command that we would use is: tar xf MyBackup.tar memo.ben - It uses the x switch to extract a file. 8. What is partial backup? - When you select only a portion of your file hierarchy or a single partition to back up, it is called partial back up. 9. What is the fastest way to enter a series of commands from the command-line? - Write the commands, each separated by a semi-colon. Press enter after the last command.  - The semi-colon would inform the shell that multiple commands are being entered at the command line, to be executed serially. 10. What are the qualities of soft links? a.) Soft link files have different inode numbers than source file b.) The soft link file will be of no use if original file is deleted.  c.) Soft links are not updated d.) They can create links between directories e.) They can cross file system boundaries 11. Differentiate between Cron and Anacron. a.) Minimum granularity with Cron is minute while it is in days with Anacron.  b.) Cron job can be scheduled by any normal user while Anacron can be scheduled only by the super user.  c.) Cron expects the system to be up and running while the Anacron doesn’t expect the system to be up and running all the time. In case of anacron if a job is scheduled and the system is down that time, it will execute the job as soon as the system is up and running.  d.) Cron is ideal for servers while Anacron is ideal for desktops and laptops.  e. ) Cron should be used when you want a job to be executed at a particular hour and minute while Anacron should be used in when the job can be executed irrespective of the hour and minute. 12.) What is an INODE? - It is a structure which has the description of all the files and pointers to the data blocks of file stored in it.  - The information contained is file-size, access and modification time, permission and so on. Linux interview - August 21, 2008 at 22:00 pm by Rajmeet Ghai [**What is Linux and why is it so popular?**](http://www.careerride.com/Linux-Defined.aspx)  **Answer -** Linux is an operating system that uses UNIX like Operating system.......  [**Unix interview questions with answers**](http://www.careerride.com/Unix-Interview-Questions.aspx)  Discuss the mount and unmount system calls, What are the process states in Unix?, What is use of sed command?, What is 'inode'?,What are the Unix system calls for I/O?, How are devices represented in UNIX?, Brief about the directory representation in UNIX ......  [**What is LILO?**](http://www.careerride.com/Linux-Loader-LILO.aspx)  **Answer -** LILO is Linux Loader is a boot loader for Linux. It is used to load Linux into the memory and start the Operating system.......  [**What is the difference between home directory and working directory?**](http://www.careerride.com/Linux-home-working-directory.aspx)  **Answer -** Home directory is the default working directory when a user logs in. On the other hand, working directory is the user’s current directory.......  [**What is the difference between internal and external commands?**](http://www.careerride.com/Linux-internal-external-commands.aspx)  **Answer -** Internal commands are commands that are already loaded in the system. They can be executed any time and are independent.......  [**Explain the difference between a static library and a dynamic library.**](http://www.careerride.com/Linux-Static-dynamic-library.aspx)  **Answer -** Static libraries are loaded when the program is compiled and dynamically-linked libraries are loaded in while......  [**What is LD\_LIBRARY\_PATH?**](http://www.careerride.com/Linux-LD_LIBRARY_PATH.aspx)  **Answer -** LD\_LIBRARY\_PATH is an environment variable. It is used for debugging a new library or a non standard library.......  [**What is the file server in Linux server?**](http://www.careerride.com/Linux-file-server.aspx)  **Answer -** File server is used for file sharing. It enables the processes required fro sharing....... [What is NFS? What is its purpose?](http://www.careerride.com/Linux-NFS.aspx)  **Answer -** NFS is Network File system. It is a file system used for sharing of files over a network....... [How do I send email with linux?](http://www.careerride.com/Linux-send-email.aspx)  **Answer -** Email can be sent in Linux using the mail command. ...... [Explain RPM (Red Hat Package Manager) features.](http://www.careerride.com/Linux-RPM.aspx)  **Answer -** RPM is a package managing system (collection of tools to manage software packages)....... [What is Kernel? Explain the task it performs.](http://www.careerride.com/Linux-Kernel.aspx)  **Answer -** Kernel is used in UNIX like systems and is considered to be the heart of the operating system....... [What is Linux Shell? What is Shell Script?](http://www.careerride.com/Linux-Shell.aspx)  **Answer -** Linux shell is a user interface used for executing the commands. Shell is a program the user...... [What are Pipes? Explain use of pipes.](http://www.careerride.com/Linux-pipes.aspx)  **Answer -** A pipe is a chain of processes so that output of one process (stdout) is fed an input (stdin) to another....... [Explain trap command; shift Command, getopts command of linux.](http://www.careerride.com/Linux-command.aspx)  **Answer -** Trap command: controls the action to be taken by the shell when a signal is received. ...... [What Stateless Linux server? What feature it offers?](http://www.careerride.com/Linux-Stateless-Server.aspx)  **Answer -** A stateless Linux server is a centralized server in which no state exists on the single workstations. ...... [What does nslookup do? Explain its two modes.](http://www.careerride.com/Linux-nslookup.aspx)  **Answer -** Nslookup is used to find details related to a Domain name server. Details like IP addresses of a machine, MX records,...... [What is Bash Shell?](http://www.careerride.com/Linux-Bash-Shell.aspx)  **Answer -** Bash is a free shell for UNIX. It is the default shell for most UNIX systems. It has a combination of the C and Korn shell features. ...... [Explain some Network-Monitoring Tools in Linux: ping, traceroute, tcpdump, ntop](http://www.careerride.com/Linux-Network-Monitoring-Tools.aspx)  **Answer -** Network monitoring tools are used to monitor the network, systems present on the network, traffic etc....... [How does the linux file system work?](http://www.careerride.com/Linux-file-system.aspx)  **Answer -** Linux file structure is a tree like structure. It starts from the root directory, represented by '/', and then expands into sub-directories....... [What are the process states in Linux?](http://www.careerride.com/Linux-process-states.aspx)  **Answer -** Process states in Linux....... [What is a zombie?](http://www.careerride.com/Linux-zombie.aspx)  **Answer -** Zombie is a process state when the child dies before the parent process. In this case the structural information of the process is still in the process table....... [Explain each system calls used for process management in linux.](http://www.careerride.com/Linux-process-management.aspx)  **Answer -** System calls used for Process management...... Which command is used to check the number of files and disk space used and the each user’s defined quota? repquota command is used to check the status of the user’s quota along with the disk space and number of files used. This command gives a summary of the user’s quota that how much space and files are left for the user. Every user has a defined quota in Linux. This is done mainly for the security, as some users have only limited access to files. This provides a security to the files from unwanted access. The quota can be given to a single user or to a group of users. What is the name and path of the main system log? By default the main system log is /var/log/messages. This file contains all the messages and the script written by the user. By default all scripts are saved in this file. This is the standard system log file, which contains messages from all system software, non-kernel boot issues, and messages that go to 'dmesg'. dmesg is a system file that is written upon system boot. How secured is Linux? Explain. Security is the most important aspect of an operating system. Due to its unique authentication module, Linux is considered as more secured than other operating systems. Linux consists of PAM. PAM is Pluggable Authentication Modules. It provides a layer between applications and actual authentication mechanism. It is a library of loadable modules which are called by the application for authentication. It also allows the administrator to control when a user can log in. All PAM applications are configured in the directory "/etc/pam.d" or in a file "/etc/pam.conf". PAM is controlled using the configuration file or the configuration directory. Can Linux computer be made a router so that several machines may share a single Internet connection? How? Yes a Linux machine can be made a router. This is called "IP Masquerade." IP Masquerade is a networking function in Linux similar to the one-to-many (1: Many) NAT (Network Address Translation) servers found in many commercial firewalls and network routers. The IP Masquerade feature allows other "internal" computers connected to this Linux box (via PPP, Ethernet, etc.) to also reach the Internet as well. Linux IP Masquerading allows this functionality even if the internal computers do not have IP addresses. The IP masquerading can be done by the following steps:  1. The Linux PC must have an internet connection and a connection to LAN. Typically, the Linux PC has two network interfaces-an Ethernet card for the LAN and a dial-up PPP connection to the Internet (through an ISP).  2. All other systems on your LAN use the Linux PC as the default gateway for TCP/IP networking. Use the same ISP-provided DNS addresses on all systems.  3. Enable IP forwarding in the kernel. By default the IP forwarding is not enabled. To ensure that IP forwarding is enabled when you reboot your system, place this command in the /etc/rc.d/rc.local file.  4. Run /sbin/iptables-the IP packet filter administration program-to set up the rules that enable the Linux PC to masquerade for your LAN. What is the minimum number of partitions you need to install Linux? Minimum 2 partitions are needed for installing Linux. The one is / or root which contains all the files and the other is swap. Linux file system is function specific which means that files and folders are organized according to their functionality. For example, all executables are in one folder, all devices in another, all libraries in another and so on. / or ‘root’ is the base of this file system. All the other folders are under this one. / can be consider as C: .Swap is a partition that will be used as virtual memory. If there is no more available RAM a Linux computer will use an area of the hard disk, called swap, to temporarily store data. In other words it is a way of expanding your computers RAM. Which command is used to review boot messages? dmesg command is used to review boot messages. This command will display system messages contained in the kernel ring buffer. We can use this command immediately after booting to see boot messages. A ring buffer is a buffer of fixed size for which any new data added to it overwrites the oldest data in it. Its basic syntax is  dmesg [options]  Invoking dmesg without any of its options causes it to write all the kernel messages to standard output. This usually produces far too many lines to fit into the display screen all at once, and thus only the final messages are visible. However, the output can be redirected to the less command through the use of a pipe, thereby allowing the startup messages to be viewed on one screen at a time dmesg | less Which utility is used to make automate rotation of a log? logrotate command is used to make automate rotation of log. Syntax of the command is: logrotate [-dv] [-f|] [-s|] config\_file+ It allows automatic rotation, compression, removal, and mailing of log files. This command is mainly used for rotating and compressing log files. This job is done every day when a log file becomes too large. This command can also be run by giving on command line. We can done force rotation by giving –f option with this command in command line. This command is also used for mailing. We can give –m option for mailing with this command. This option takes two arguments one is subject and other is recipient name. What are the partitions created on the mail server hard drive? The main partitions are done firstly which are root, swap and boot partition. But for the mail server three different partitions are also done which are as follows: 1. /var/spool- This is done so that if something goes wrong with the mail server or spool than the output cannot overrun the file system. 2. /tmp- putting this on its own partition prevents any user item or software from overrunning the system files. 3. /home- putting this on its own is useful for system upgrades or reinstalls. It allow not to wipe off the /home hierarchy along with other areas. What are the fields in the/etc/passwd file? It contains all the information of the users who log into the system. It contains a list of the system's accounts, giving for each account some useful information like user ID, group ID, home directory, shell, etc. It should have general read permission as many utilities, like ls use it to map user IDs to user names, but write access only for the superuser (root). The main fields of /etc/passwd file are: 1. Username: It is used when user logs in. It should be between 1 and 32 characters in length. 2. Password: An x character indicates that encrypted password is stored in /etc/shadow file. 3. User ID (UID): Each user must be assigned a user ID (UID). UID 0 (zero) is reserved for root and UIDs 1-99 are reserved for other predefined accounts. Further UID 100-999 are reserved by system for administrative and system accounts/groups. 4. Group ID (GID): The primary group ID (stored in /etc/group file) 5. User ID Info: The comment field. It allow you to add extra information about the users such as user's full name, phone number etc. This field use by finger command. 6. Home directory: The absolute path to the directory the user will be in when they log in. If this directory does not exists then users directory becomes / 7. Command/shell: The absolute path of a command or shell (/bin/bash). Typically, this is a shell. Which commands are used to set a processor-intensive job to use less CPU time? nice command is used for changing priority of the jobs. Syntax: nice [OPTION] [COMMAND [ARG]...] Range of priority goes from -20 (highest priority) to 19 (lowest).Priority is given to a job so that the most important job is executed first by the kernel and then the other least important jobs. This takes less CPU times as the jobs are scheduled and are given priorities so the CPU executes fast. The priority is given by numbers like -20 describe the highest priority and 19 describe the least priority. How to change window manager by editing your home directory? /.xinitrc file allows changing the window manager we want to use when logging into X from that account. The dot in the file name shows you that the file is a hidden file and doesn't show when you do a normal directory listing. For setting a window manager we have to save a command in this file. The syntax of command is: exec windowmanager.After this, save the file. Next time when you run a startx a new window manager will open and become default. The commands for starting some popular window managers and desktop environments are: -KDE = startkde -Gnome = gnome-session -Blackbox = blackbox -FVWM = fvwm -Window Maker = wmaker -IceWM = icewm How documentation of an application is stored? When a new application is installed its documentation is also installed. This documentation is stored under the directory named for application. For example if my application name is App1 then the path of the documentation will be /user/doc/App1. It contains all the information about the application. It contains date of creating application, name of application and other important module of the application. We can get the basic information of application from the documentation. How shadow passwords are given? pwconv command is used for giving shadow passwords. Shadow passwords are given for better system security. The pwconv command creates the file /etc/shadow and changes all passwords to ‘x’ in the /etc/passwd file. First, entries in the shadowed file which don't exist in the main file are removed. Then, shadowed entries which don't have `x' as the password in the main file are updated. Any missing shadowed entries are added. Finally, passwords in the main file are replaced with `x'. These programs can be used for initial conversion as well to update the shadowed file if the main file is edited by hand. How do you create a new user account? useradd command is used for creating a new user account. When invoked without the -D option, the useradd command creates a new user account using the values specified on the command line and the default values from the system. The new user account will be entered into the system files as needed, and initial files copied, depending on the command line options. This command uses the system default as home directory. If –m option is given then the home directory is made. Which password package is installed for the security of central password? Shadow password packages are used for security of central passwords. Security is the most important aspect of every operating system. When this package is not installed the user information including passwords is stored in the /etc/passwd file. The password is stored in an encoded format. These encoded forms can be easily identified by the System crackers by randomly encoding the passwords from dictionaries. The Shadow Package solves the problem by relocating the passwords to another file (usually /etc/shadow). The /etc/shadow file is set so that it cannot be read by just anyone. Only root will be able to read and write to the /etc/shadow file. Which shell do you assign to a POP3 mail-only account? POP3 mail only account is assigned to the /bin/false shell. However, assigning bash shell to a POP3 mail only gives user login access, which is avoided. /bin/nologin can also be used. This shell is provided to the user when we don’t want to give shell access to the user. The user cannot access the shell and it reject shell login on the server like on telnet. It is mainly for the security of the shells. POP3 is basically used for downloading mail to mail program. So for illegal downloading of emails on the shell this account is assigned to the /bin/false shell or /bin/nologin. These both shells are same they both do the same work of rejecting the user login to the shell. The main difference between these two shells is that false shell shows the incorrect code and any unusual coding when user login with it. But the nologin shell simply tells that no such account is available. So nologin shell is used mostly in Linux. Which daemon is responsible for tracking events on Linux system? syslogd is responsible for tracking system information and save it to the desired log files. It provides two system utilities which provide system logging and kernel message trapping. Internet and UNIX domain sockets support enable this utility package to support both local and remote logging. Every logged message contains at least a time and a hostname field, normally a program name field, too. So to track these information this daemon is used. syslogd mainly reacts to the set of signals given by the user. These are the signals given to syslogd: SIGHUP: This lets syslogd perform a re-initialization. All open files are closed, the configuration file (default is /etc/syslog.conf) will be reread and the syslog facility is started again. SIGTERM: The syslogd will die. SIGINT, SIGQUIT: If debugging is enabled these are ignored, otherwise syslogd will die. SIGUSR1: Switch debugging on/off. This option can only be used if syslogd is started with the - d debug option. SIGCHLD: Wait for Childs if some were born, because of waiting messages. Which daemon is used for scheduling of the commands? The crontab command is used for scheduling of the commands to run at a later time. SYNTAX crontab [ -u user ] file crontab [ -u user ] { -l | -r | -e }  Options -l List - display the current crontab entries.  -r Remove the current crontab.  -e Edit the current crontab using the editor specified by the VISUAL or EDITOR environment variables. When user exits from the editor, the modified crontab will be installed automatically. Each user can have their own crontab, and though these are files in /var, they are not intended to be edited directly. If the –u option is given than the crontab gives the name of the user whose crontab is to be tweaked. If it is given without this then it will display the crontab of the user who is executing the command. How environment variable is set so that the file permission can be automatically set to the newly created files? umask command is used to set file permission on newly created files automatically.  Syntax umask [-p] [-S] [mode] It is represented in octal numbers. We can simply use this command without arguments to see the current file permissions. To change the permissions, mode is given in the arguments. The default umask used for normal user is 0002. The default umask for the root user is 0022. For calculating the original values, the values shown by the umask must be subtracted by the default values. It is mainly used for masking of the file and directory permission. The /etc/profile script is where the umask command is usually set for all users. The –S option can be used to see the current default permissions displayed in the alpha symbolic format.  For example, umask 022 ensures that new files will have at most 755 permissions (777 NAND 022). The permissions can be calculated by taking the NAND of original value with the default values of files and directories. |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | ***[Write your comment - Share Knowledge and Experience](http://www.careerride.com/add-comment.aspx?Page=Linux-Interview-Questions.aspx)***   |  | | --- | | **Discussion Board** | | **Fresher forward their resumes @ djohn6292@gmail.com**   Hi All,  Request all fresher RHCE certified candidates to share your resumes @ djohn6292@gmail.com.   John 09-11-2013 02:59 AM | | **Nice article**   very nice article.  Thanks, shashikant   shashikant 08-16-2013 05:23 AM | | **JOB**   hello,  I am sandeep rawat, i have completed my bca then i have RED HAT CERTIFIED ENG.. I am searching job in linux admin. I am fresher.  please call me 9871190569 (DELHI , NCR)   Sandeep Rawat 07-24-2013 01:24 PM | | **interview question on linux by tcs**   On telephonic interview they asking about your current job and some question in respective field. for linux admin they asked 1. how to change sshd port. 2. how to configure router.   sushil 07-10-2013 04:56 AM | | **correction to a question**   What is the minimum number of partitions you need to install Linux? Anwer is / and /boot (and not /swap which is optional). Nice article.   Amit 05-5-2013 12:58 PM | | **LInux**   POST does not tell me what CPU I use, dmesg does (Specifically "dmesg | grep CPU)  On a RAID 5 array with 6 20 GB drives, 100 GB are available, not 120. one disk worth is used parity (6x20=120 120-20=100)  SAMBA is designed to be used for Windows File sharing and is what Windows uses by default, NFS is Linux's default and making Windows use it takes work)  When using ssh-keygen to create the public/private key pairs, it is BEST to then disable the password for logging on as they are no longer needed and serve as an extra attack surface.    iceman 03-18-2013 12:17 AM | | **Looking For a Job**   Hello Friends,  I am Darshan Jadhav from Nasik. I have Completed My Diploma In Computer Technology From KKW COE,Nashik I have Completed My CCNA Certification from Router School,Pune Recently I have completed RHCE & RHCSA Exam from SEED InfoTech,Pune  plz do share if u have any openings in LINUX or Networking   DARSHAN JADHAV 03-15-2013 08:40 AM | | **Good article**   This is one is very helpful article. Thanks for your work and posting this useful information.   prasad 01-12-2013 12:41 PM | | **aaa**   supap yar  it is very usefull linux question   yasir khan 11-30-2012 11:30 AM | | **hai**   you r giving very help full questions for interview  Thank you   koti 11-29-2012 12:10 AM | | **comment**   nice questions it is really helpfull for me    gajanan bhure 11-8-2012 07:26 AM | | **Hi**   Thanks alot for sharing such kind of information,its really helpfull.....   pavan 11-5-2012 05:35 AM | | **Looking for Linux Administrator Job**   Hi friends, I did MCA & RHCE, currently Looking for Linux Administration Job. Let you Knw, please send me at adhibanraj@gmail.com   Raj 10-19-2012 04:39 AM | | **Redhat administration**   Hi I am searching redhat admin job , cause of I am diploma in etct, working exprnc 3 years & finally completed rhce certification.........Please help me.   Gour Ch Ghosh 09-20-2012 07:26 AM | | **hi**   It is good site for going to attend interview.   rajanna 08-30-2012 05:08 AM | | **job search**   i am jaikey kumar b-tech computer sciencs and rhce cretified i search jobs for rhce fresher   jay kumar 07-24-2012 12:03 PM | | **job**   Dear Friends,' My name is vipin jagwan completed B.A, i need Linux Admin Job.I have completd REDHAT Certification   Please help me    vipin jagwan 07-17-2012 11:37 AM | | **Requesting for Job**   Dear Friends,' My name is Arumugam completed BE(CS), i need Linux Admin Job.I have completd REDHAT Certification and MCSE too  Please help me out   Arumugam 07-3-2012 01:51 AM | | **Linux job in mumbai**   hi guys i want a ob in Linux administration(mumbai) ,could any one help out my emai id is diwadkar30@gmail.com   prasad diwadkar 03-17-2012 12:48 PM | | **Looking for Sr .Linux Developer**   • Location: Hyderabad • Experience : 6+ Years • Key Skills: LINUX, Kernel Programming, Device Drivers • Must hold a BE/BTech/ME/MTech or higher degree in any Specialization.  Expectations:  • Atleast 4+ years experience in Linux Kernel Programming • Mandatory hands-on linux device driver development experience • Worked on any of the these drivers - NAND, SDIO, USB, Accelerometer, GPS, Battery. • Experience on SNMP, TR69, WiFi, IPv6, MobileIP • Strong programming experience in C/C++ Email: hr.sushma09@gmail.com   sushma 03-16-2012 01:28 AM | | **very thankful**   very very thank you for help me about linux.   yogesh chavan 03-1-2012 02:20 PM | | **Comments**   Thanks for the great work...good for confidence boosting   Rajkumar 02-7-2012 09:35 AM | | **About Linux Question with Answer**   Great Work !!!   sanjeet mishra 01-29-2012 10:34 PM | | **Hi**   Good work   Saroj 12-28-2011 10:49 PM | | **comment**   Good work,smart answers...   SINGARAVELAN 12-21-2011 11:02 AM | | **Good work**   Nice work, keep posting.   sivasubramaniam.T 12-8-2011 05:52 AM | | **"bugs" on the test**   the expected answer for the raid5 question is wrong   maleas 11-18-2011 06:30 PM | | **LINUX Interview Qwestions**   Hi,  Thanks a lot for updating linux interview questions Could you please update some more important questions   p.karunakar reddy 11-15-2011 02:11 AM | | **Rgds Interview Question**   Thanks Teams  Gud Questionn of Linux Topics....   Sunny Kumar 11-14-2011 04:21 AM | | **Thanks Alot**   This is nice material for linux interviews. can you update more on clustering, migration, recovery and virtualization  Thanks   sandeep 11-11-2011 02:27 AM | |  | | | | |  |

# Linux system admin interview questions & answers

i

9 Votes

Quantcast

**Linux interview questions & answers**.

1. Who owns the data dictionary?

-> The SYS user owns the data dictionary. The SYS and SYSTEM users are created when the database is created.

2. You routinely compress old log files. You now need to examine a log from two months ago. In order to view its contents without first having to decompress it ?

zcat -> The zcat utility allows you to examine the contents of a compressed file much the same way that cat displays a file.

3. You suspect that you have two commands with the same name as the command is not producing the expected results. What command can you use to determine the location of the command being run?

which -> The which command searches your path until it finds a command that matches the command you are looking for and displays its full path.

4. You locate a command in the /bin directory but do not know what it does. What command can you use to determine its purpose.

whatis ->The whatis command displays a summary line from the man page for the specified command.

5. When you issue the command ls -l, the first character of the resulting display represents the file’s?

type ->The first character of the permission block designates the type of file that is being displayed.

6. What utility can you use to show a dynamic listing of running processes?  top -> The top utility shows a listing of all running processes that is dynamically updated.

7. Where is standard output usually directed? -> to the screen or display.

8. What daemon is responsible for tracking events on your system?       syslogd ->The syslogd daemon is responsible for tracking system information and saving it to specified log files.

9. You have a file called phonenos that is almost 4,000 lines long. What text filter can you use to split it into four pieces each 1,000 lines long?               split ->The split text filter will divide files into equally sized pieces. The default length of each piece is 1,000 lines.

10. You would like to temporarily change your command line editor to be vi. What command should you type to change it?                                                      set -o vi ->The set command is used to assign environment variables. In this case, you are instructing your shell to assign vi as your command line editor. However, once you log off and log back in you will return to the previously defined command line editor.

11. What account is created when you install Linux?                                              root ->Whenever you install Linux, only one user account is created. This is the superuser account also known as root.

12. What command should you use to check the number of files and disk space used and each user’s defined quotas?                                                          repquota ->The repquota command is used to get a report on the status of the quotas you have set including the amount of allocated space and amount of used space.

13. In order to run fsck on the root partition, the root partition must be mounted as ?                                                                                                                 readonly ->You cannot run fsck on a partition that is mounted as read-write.

14. In order to improve your system’s security you decide to implement shadow passwords. What command should you use?                                             pwconv ->The pwconv command creates the file /etc/shadow and changes all passwords to ‘x’ in the /etc/passwd file.

15. The top utility can be used to change the priority of a running process? Another utility that can also be used to change priority is \_\_\_\_\_\_\_\_\_\_\_? nice ->Both the top and nice utilities provide the capability to change the priority of a running process.

16. What file defines the levels of messages written to system log files? kernel.h ->To determine the various levels of messages that are defined on your system, examine the kernel.h file.

17. What can you type at a command line to determine which shell you are using?                                                                                                                           echo $SHELL ->The name and path to the shell you are using is saved to the SHELL environment variable. You can then use the echo command to print out the value of any variable by preceding the variable’s name with $. Therefore, typing echo $SHELL will display the name of your shell.

18. When you create a new partition, you need to designate its size by defining the starting and ending ?                                                                   cylinders ->When creating a new partition you must first specify its starting cylinder. You can then either specify its size or the ending cylinder.

19. How are devices represented in UNIX ? –> All devices are represented by files called special files that are located in/dev directory. Thus, device files and other files are named and accessed in the same way. A ‘regular file’ is just an ordinary data file in the disk. A ‘block special file’ represents a device with characteristics similar to a disk (data transfer in terms of blocks). A ‘character special file’ represents a device with characteristics similar to a keyboard (data transfer is by stream of bits in sequential order).

20. What is ‘inode’? –>All UNIX files have its description stored in a structure called ‘inode’. The inode contains info about the file-size, its location, time of last access, time of last modification, permission and so on. Directories are also represented as files and have an associated inode. In addition to descriptions about the file, the inode contains pointers to the data blocks of the file. If the file is large, inode has indirect pointer to a block of pointers to additional data blocks (this further aggregates for larger files). A block is typically 8k. Inode consists of the following fields: File owner identifier File type File access permissions File access times Number of links File size Location of the file data

21. Brief about the directory representation in UNIX ? –>A Unix directory is a file containing a correspondence between filenames and inodes. A directory is a special file that the kernel maintains. Only kernel modifies directories, but processes can read directories. The contents of a directory are a list of filename and inode number pairs. When new directories are created, kernel makes two entries named ‘.’ (refers to the directory itself) and ‘..’ (refers to parent directory). System call for creating directory is mkdir (pathname, mode).

22. What are the Unix system calls for I/O? –> open(pathname,flag,mode) – open file creat(pathname,mode) – create file close(filedes) – close an open file read(filedes,buffer,bytes) – read data from an open file write(filedes,buffer,bytes) – write data to an open file lseek(filedes,offset,from) – position an open file dup(filedes) – duplicate an existing file descriptor dup2(oldfd,newfd) – duplicate to a desired file descriptor fcntl(filedes,cmd,arg) – change properties of an open file ioctl(filedes,request,arg) – change the behaviour of an open file The difference between fcntl anf ioctl is that the former is intended for any open file, while the latter is for device-specific operations.

23. How do you change File Access Permissions? –>Every file has following attributes: owner’s user ID ( 16 bit integer ) owner’s group ID ( 16 bit integer ) File access mode word ‘r w x -r w x- r w x’ (user permission-group permission-others permission) r-read, w-write, x-execute To change the access mode, we use chmod(filename,mode). Example 1: To change mode of myfile to ‘rw-rw-r–’ (ie. read, write permission for user – read,write permission for group – only read permission for others) we give the args as: chmod(myfile,0664) . Each operation is represented by discrete values ‘r’ is 4 ‘w’ is 2 ‘x’ is 1 Therefore, for ‘rw’ the value is 6(4+2). Example 2: To change mode of myfile to ‘rwxr–r–’ we give the args as: chmod(myfile,0744).

24. What are links and symbolic links in UNIX file system? —>A link is a second name (not a file) for a file. Links can be used to assign more than one name to a file, but cannot be used to assign a directory more than one name or link filenames on different computers. Symbolic link ‘is’ a file that only contains the name of another file.Operation on the symbolic link is directed to the file pointed by the it.Both the limitations of links are eliminated in symbolic links. Commands for linking files are: Link ln filename1 filename2 Symbolic link ln -s filename1 filename2

25. What is a FIFO? –> FIFO are otherwise called as ‘named pipes’. FIFO (first-in-first-out) is a special file which is said to be data transient. Once data is read from named pipe, it cannot be read again. Also, data can be read only in the order written. It is used in interprocess communication where a process writes to one end of the pipe (producer) and the other reads from the other end (consumer).

26. How do you create special files like named pipes and device files? –> The system call mknod creates special files in the following sequence. 1. kernel assigns new inode, 2. sets the file type to indicate that the file is a pipe, directory or special file, 3. If it is a device file, it makes the other entries like major, minor device numbers. For example: If the device is a disk, major device number refers to the disk controller and minor device number is the disk.

27. Discuss the mount and unmount system calls The privileged mount system call is used to attach a file system to a directory of another file system; the unmount system call detaches a file system. When you mount another file system on to your directory, you are essentially splicing one directory tree onto a branch in another directory tree. The first argument to mount call is the mount point, that is , a directory in the current file naming system. The second argument is the file system to mount to that point. When you insert a cdrom to your unix system’s drive, the file system in the cdrom automatically mounts to /dev/cdrom in your system.

28. How does the inode map to data block of a file? Inode has 13 block addresses. The first 10 are direct block addresses of the first 10 data blocks in the file. The 11th address points to a one-level index block. The 12th address points to a two-level (double in-direction) index block. The 13th address points to a three-level(triple in-direction)index block. This provides a very large maximum file size with efficient access to large files, but also small files are accessed directly in one disk read.

29. What is a shell? A shell is an interactive user interface to an operating system services that allows an user to enter commands as character strings or through a graphical user interface. The shell converts them to system calls to the OS or forks off a process to execute the command. System call results and other information from the OS are presented to the user through an interactive interface. Commonly used shells are sh,csh,ks etc.

30. Brief about the initial process sequence while the system boots up. While booting, special process called the ‘swapper’ or ‘scheduler’ is created with Process-ID 0. The swapper manages memory allocation for processes and influences CPU allocation. The swapper inturn creates 3 children: the process dispatcher, vhand and dbflush with IDs 1,2 and 3 respectively. This is done by executing the file /etc/init. Process dispatcher gives birth to the shell. Unix keeps track of all the processes in an internal data structure called the Process Table (listing command is ps -el).

31. What are various IDs associated with a process? Unix identifies each process with a unique integer called ProcessID. The process that executes the request for creation of a process is called the ‘parent process’ whose PID is ‘Parent Process ID’. Every process is associated with a particular user called the ‘owner’ who has privileges over the process. The identification for the user is ‘UserID’. Owner is the user who executes the process. Process also has ‘Effective User ID’ which determines the access privileges for accessing resources like files. getpid() -process id getppid() -parent process id getuid() -user id geteuid() -effective user id

32. Explain fork() system call. The `fork()’ used to create a new process from an existing process. The new process is called the child process, and the existing process is called the parent. We can tell which is which by checking the return value from `fork()’. The parent gets the child’s pid returned to him, but the child gets 0 returned to him.

33. Predict the output of the following program code

34. 35. main() { fork(); printf(“Hello World!”); } Answer: Hello World!Hello World! Explanation: The fork creates a child that is a duplicate of the parent process. The child begins from the fork().All the statements after the call to fork() will be executed twice.(once by the parent process and other by child). The statement before fork() is executed only by the parent process.

36. Predict the output of the following program code main() { fork(); fork(); fork(); printf(“Hello World!”); } Answer: “Hello World” will be printed 8 times. Explanation: 2^n times where n is the number of calls to fork()

37. List the system calls used for process management: System calls Description fork() To create a new process exec() To execute a new program in a process wait() To wait until a created process completes its execution exit() To exit from a process execution getpid() To get a process identifier of the current process getppid() To get parent process identifier nice() To bias the existing priority of a process brk() To increase/decrease the data segment size of a process

38. How can you get/set an environment variable from a program?: Getting the value of an environment variable is done by using `getenv()’. Setting the value of an environment variable is done by using `putenv()’.

39. How can a parent and child process communicate? A parent and child can communicate through any of the normal inter-process communication schemes (pipes, sockets, message queues, shared memory), but also have some special ways to communicate that take advantage of their relationship as a parent and child. One of the most obvious is that the parent can get the exit status of the child.

40. What is a zombie? When a program forks and the child finishes before the parent, the kernel still keeps some of its information about the child in case the parent might need it – for example, the parent may need to check the child’s exit status. To be able to get this information, the parent calls `wait()’; In the interval between the child terminating and the parent calling `wait()’, the child is said to be a `zombie’ (If you do `ps’, the child will have a `Z’ in its status field to indicate this.)

41. What are the process states in Unix? As a process executes it changes state according to its circumstances. Unix processes have the following states: Running : The process is either running or it is ready to run . Waiting : The process is waiting for an event or for a resource. Stopped : The process has been stopped, usually by receiving a signal. Zombie : The process is dead but have not been removed from the process table.

42. What Happens when you execute a program? When you execute a program on your UNIX system, the system creates a special environment for that program. This environment contains everything needed for the system to run the program as if no other program were running on the system. Each process has process context, which is everything that is unique about the state of the program you are currently running. Every time you execute a program the UNIX system does a fork, which performs a series of operations to create a process context and then execute your program in that context. The steps include the following: Allocate a slot in the process table, a list of currently running programs kept by UNIX. Assign a unique process identifier (PID) to the process. iCopy the context of the parent, the process that requested the spawning of the new process. Return the new PID to the parent process. This enables the parent process to examine or control the process directly. After the fork is complete, UNIX runs your program.

43. What Happens when you execute a command? When you enter ‘ls’ command to look at the contents of your current working directory, UNIX does a series of things to create an environment for ls and the run it: The shell has UNIX perform a fork. This creates a new process that the shell will use to run the ls program. The shell has UNIX perform an exec of the ls program. This replaces the shell program and data with the program and data for ls and then starts running that new program. The ls program is loaded into the new process context, replacing the text and data of the shell. The ls program performs its task, listing the contents of the current directory.

44. What is a Daemon? A daemon is a process that detaches itself from the terminal and runs, disconnected, in the background, waiting for requests and responding to them. It can also be defined as the background process that does not belong to a terminal session. Many system functions are commonly performed by daemons, including the sendmail daemon, which handles mail, and the NNTP daemon, which handles USENET news. Many other daemons may exist. Some of the most common daemons are: init: Takes over the basic running of the system when the kernel has finished the boot process. inetd: Responsible for starting network services that do not have their own stand-alone daemons. For example, inetd usually takes care of incoming rlogin, telnet, and ftp connections. cron: Responsible for running repetitive tasks on a regular schedule.

45. What is ‘ps’ command for? The ps command prints the process status for some or all of the running processes. The information given are the process identification number (PID),the amount of time that the process has taken to execute so far etc.

46. How would you kill a process? The kill command takes the PID as one argument; this identifies which process to terminate. The PID of a process can be got using ‘ps’ command.

47. What is an advantage of executing a process in background? The most common reason to put a process in the background is to allow you to do something else interactively without waiting for the process to complete. At the end of the command you add the special background symbol, &. This symbol tells your shell to execute the given command in the background. Example: cp \*.\* ../backup& (cp is for copy)

48. How do you execute one program from within another? The system calls used for low-level process creation are execlp() and execvp(). The execlp call overlays the existing program with the new one , runs that and exits. The original program gets back control only when an error occurs. execlp(path,file\_name,arguments..); //last argument must be NULL A variant of execlp called execvp is used when the number of arguments is not known in advance. execvp(path,argument\_array); //argument array should be terminated by NULL

49. What is IPC? What are the various schemes available? The term IPC (Inter-Process Communication) describes various ways by which different process running on some operating system communicate between each other. Various schemes available are as follows: Pipes: One-way communication scheme through which different process can communicate. The problem is that the two processes should have a common ancestor (parent-child relationship). However this problem was fixed with the introduction of named-pipes (FIFO). Message Queues : Message queues can be used between related and unrelated processes running on a machine. Shared Memory: This is the fastest of all IPC schemes. The memory to be shared is mapped into the address space of the processes (that are sharing). The speed achieved is attributed to the fact that there is no kernel involvement. But this scheme needs synchronization. Various forms of synchronisation are mutexes, condition-variables, read-write locks, record-locks, and semaphores.

50. What is the difference between Swapping and Paging? Swapping: Whole process is moved from the swap device to the main memory for execution. Process size must be less than or equal to the available main memory. It is easier to implementation and overhead to the system. Swapping systems does not handle the memory more flexibly as compared to the paging systems. Paging: Only the required memory pages are moved to main memory from the swap device for execution. Process size does not matter. Gives the concept of the virtual memory. It provides greater flexibility in mapping the virtual address space into the physical memory of the machine. Allows more number of processes to fit in the main memory simultaneously. Allows the greater process size than the available physical memory. Demand paging systems handle the memory more flexibly.

51. What is major difference between the Historic Unix and the new BSD release of Unix System V in terms of Memory Management? Historic Unix uses Swapping – entire process is transferred to the main memory from the swap device, whereas the Unix System V uses Demand Paging – only the part of the process is moved to the main memory. Historic Unix uses one Swap Device and Unix System V allow multiple Swap Devices.

52. What is the main goal of the Memory Management? It decides which process should reside in the main memory, Manages the parts of the virtual address space of a process which is non-core resident, Monitors the available main memory and periodically write the processes into the swap device to provide more processes fit in the main memory simultaneously.

53. What is a Map? A Map is an Array, which contains the addresses of the free space in the swap device that are allocatable resources, and the number of the resource units available there. This allows First-Fit allocation of contiguous blocks of a resource. Initially the Map contains one entry – address (block offset from the starting of the swap area) and the total number of resources. Kernel treats each unit of Map as a group of disk blocks. On the allocation and freeing of the resources Kernel updates the Map for accurate information.

54. What scheme does the Kernel in Unix System V follow while choosing a swap device among the multiple swap devices? Kernel follows Round Robin scheme choosing a swap device among the multiple swap devices in Unix System V.

55. What is a Region? A Region is a continuous area of a process’s address space (such as text, data and stack). The kernel in a ‘Region Table’ that is local to the process maintains region. Regions are sharable among the process.

56. What are the events done by the Kernel after a process is being swapped out from the main memory? When Kernel swaps the process out of the primary memory, it performs the following: Kernel decrements the Reference Count of each region of the process. If the reference count becomes zero, swaps the region out of the main memory, Kernel allocates the space for the swapping process in the swap device, Kernel locks the other swapping process while the current swapping operation is going on, The Kernel saves the swap address of the region in the region table.

57. Is the Process before and after the swap are the same? Give reason. Process before swapping is residing in the primary memory in its original form. The regions (text, data and stack) may not be occupied fully by the process, there may be few empty slots in any of the regions and while swapping Kernel do not bother about the empty slots while swapping the process out. After swapping the process resides in the swap (secondary memory) device. The regions swapped out will be present but only the occupied region slots but not the empty slots that were present before assigning. While swapping the process once again into the main memory, the Kernel referring to the Process Memory Map, it assigns the main memory accordingly taking care of the empty slots in the regions.

58. What do you mean by u-area (user area) or u-block? This contains the private data that is manipulated only by the Kernel. This is local to the Process, i.e. each process is allocated a u-area.

59. What are the entities that are swapped out of the main memory while swapping the process out of the main memory? All memory space occupied by the process, process’s u-area, and Kernel stack are swapped out, theoretically. Practically, if the process’s u-area contains the Address Translation Tables for the process then Kernel implementations do not swap the u-area.

60. What is Fork swap? fork() is a system call to create a child process. When the parent process calls fork() system call, the child process is created and if there is short of memory then the child process is sent to the read-to-run state in the swap device, and return to the user state without swapping the parent process. When the memory will be available the child process will be swapped into the main memory.

61. What is Expansion swap? At the time when any process requires more memory than it is currently allocated, the Kernel performs Expansion swap. To do this Kernel reserves enough space in the swap device. Then the address translation mapping is adjusted for the new virtual address space but the physical memory is not allocated. At last Kernel swaps the process into the assigned space in the swap device. Later when the Kernel swaps the process into the main memory this assigns memory according to the new address translation mapping.

62. How the Swapper works? The swapper is the only process that swaps the processes. The Swapper operates only in the Kernel mode and it does not uses System calls instead it uses internal Kernel functions for swapping. It is the archetype of all kernel process.