

Grid Appliance Tutorial



Welcome to the Grid Appliance Tutorial! The purpose of this tutorial is to show you how to join our Condor pool in minutes, and even how to setup your own private pool with no sweat. This tutorial is simple and user-friendly and contains plenty of visual aids to help you along the way.

The tutorial contains the following sections:

- [Introduction](#)
- [System Requirements](#)
- [Running the Grid Appliance](#)
- [Checking to see if you are part of the Condor pool](#)
- [Transferring files to and from the Grid Appliance](#)
- [Submitting a Condor job through the Grid Appliance](#)
- Further interactions with the appliance:
 - [Changing your password](#)
 - [Gaining root access to the Grid Appliance](#)
 - [How to SSH into the Grid Appliance](#)

Introduction

The Grid Appliance is a self-configuring Virtual Machine appliance, that is used to create ad-hoc pools of computer resources over wide area networks to execute high-throughput, long-running jobs. Appliances are connected to each other through a peer-to-peer virtual network using private IP addresses. Upon starting the appliance, it is automatically connected to a pool of resources and is capable of submitting and executing jobs using the Condor Grid scheduler.

The purpose of this tutorial is to go into the details of the features provided by the appliance. This involves installing VMware, submitting Condor jobs, and accessing files on the appliance. Documents describing the design and configuration of the appliance and the Condor Manual are packaged alongside the image of the appliance. More information about the appliance and other related projects, including links to the latest VM image and detailed documentation can be found at:

<http://grid-appliance.org>

Users are advised to visit this page for most up-to-date information and available updates.

We also have a group for discussions/questions and answers among users and developers. [Follow this link to join.](#)

System Requirements

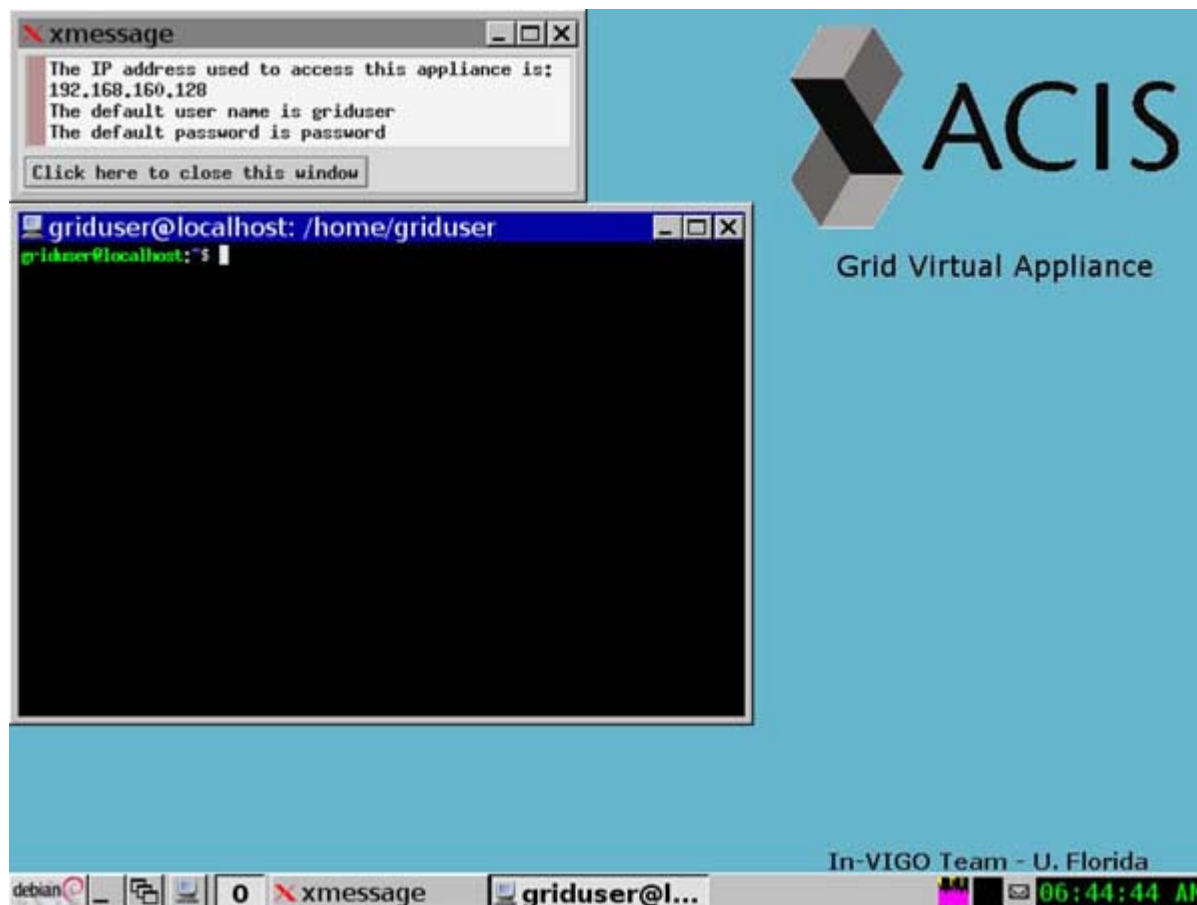
- Standard Intel/AMD x86-compatible PC
- 500 MHz or faster CPU (800 MHz recommended)
- 256 MB of RAM (512 recommended)
- 2G of free disk space available
- Internet connection
- A virtual machine monitor. [VMware](#), [KVM](#), [Qemu](#) are currently supported. Instructions on how to run the appliance on VirtualBox will be posted soon.

Running the Grid Appliance

Before you can run the Grid appliance, you will need to complete the following steps:

1. Download and install virtualization software (i.e. [VMware](#), [KVM](#), [Qemu](#)). VMware Player is recommended.
2. Download the [Grid Appliance](#), (zip files recommended for windows users, tgz files for Linux users)
3. Extract the zipped file
4. Start the Grid Appliance
 1. For VMware in windows, browse to directory and double-click on GridAppliance.vmx
 2. For VMware in Linux, cd to the extracted directory (type "vmplayer GridAppliance.vmx" if you are using VMware player, or type "vmware GridAppliance.vmx" if you are using VMware server.
 3. For Qemu on Windows, browse to directory and double-click on qemu.bat
 4. For Qemu on Linux, cd to extracted directory, and type "qemu.sh"
5. (For Qemu and KVM users only: when the Grid Appliance boots up, a boot menu will show up; use the keyboard down arrow to select the option Qemu Release)

Upon successful completion of the previous steps, the Grid Appliance will boot and display a terminal. See below.



There are some key things to notice about this boot page, here is a summary:

xmessage box (top left): shows the local IP address (which can be used to access files in the appliance as explained later in this tutorial), and the default user name and password for logging into the system. It is highly recommended that you change your password for future use.

Terminal window: The current Grid Appliance is a stripped-down version of Debian Linux 4. The terminal window provides a Linux shell, where you can execute your favorite commands (such as `ls`, `top`, `df`)

Checking to see if you are part of the Condor pool

Once you are done with all of the preliminary setup, you can check to see if you are part of the Condor pool. Type "condor_status" in the terminal window and you should see the following:

The terminal window displays the following output for the `condor_status` command:

```

C179196154 LINUX INTEL Unclaimed Idle 0.060 630 0+00:51:54
C180002072 LINUX INTEL Unclaimed Idle 0.100 630 0+00:44:17
C181103178 LINUX INTEL Unclaimed Idle 0.000 1773 0+03:29:40
C181119053 LINUX INTEL Unclaimed Idle 0.100 630 0+01:21:57
C181246119.ip LINUX INTEL Unclaimed Idle 0.070 1773 0+03:33:07
C183025147.ip LINUX INTEL Unclaimed Idle 0.000 1773 0+01:25:58
C183043036 LINUX INTEL Unclaimed Idle 0.000 1773 0+04:27:21
C184118124 LINUX INTEL Unclaimed Idle 0.130 630 0+01:23:45
C184137031 LINUX INTEL Unclaimed Idle 0.060 630 0+01:23:13
C186148211 LINUX INTEL Unclaimed Idle 0.120 630 0+01:22:56
C187195034.ip LINUX INTEL Unclaimed Idle 0.040 630 0+01:09:12
C188062122 LINUX INTEL Unclaimed Idle 0.010 630 0+01:20:54
C188253253 LINUX INTEL Unclaimed Idle 0.250 250[?????]
C190117206 LINUX INTEL Unclaimed Idle 0.040 1773 0+03:37:10
C190227127.ip LINUX INTEL Unclaimed Idle 0.050 1773 0+03:29:38
C191054199 LINUX INTEL Unclaimed Idle 0.000 630 0+01:23:30
C191119226.ip LINUX INTEL Unclaimed Idle 0.020 630 0+00:22:38

Total Owner Claimed Unclaimed Matched Preempting Backfill
INTEL/LINUX 74 2 1 71 0 0 0
Total 74 2 1 71 0 0 0
griduser@localhost:~$

```

In the terminal window, you will see a list of machines (which are grid appliances just like yours) connected to form a Condor pool. These machines are connected through a virtual network. With Condor and the Grid Appliance, all of these machines work together to provide more computing power to each grid appliance user. More information about Condor can be found [here](#).

** At times it may take several minutes to properly connect to the Condor pool, so if a warning message appears as a result of the "condor_status" command, wait a few moments, and try again.

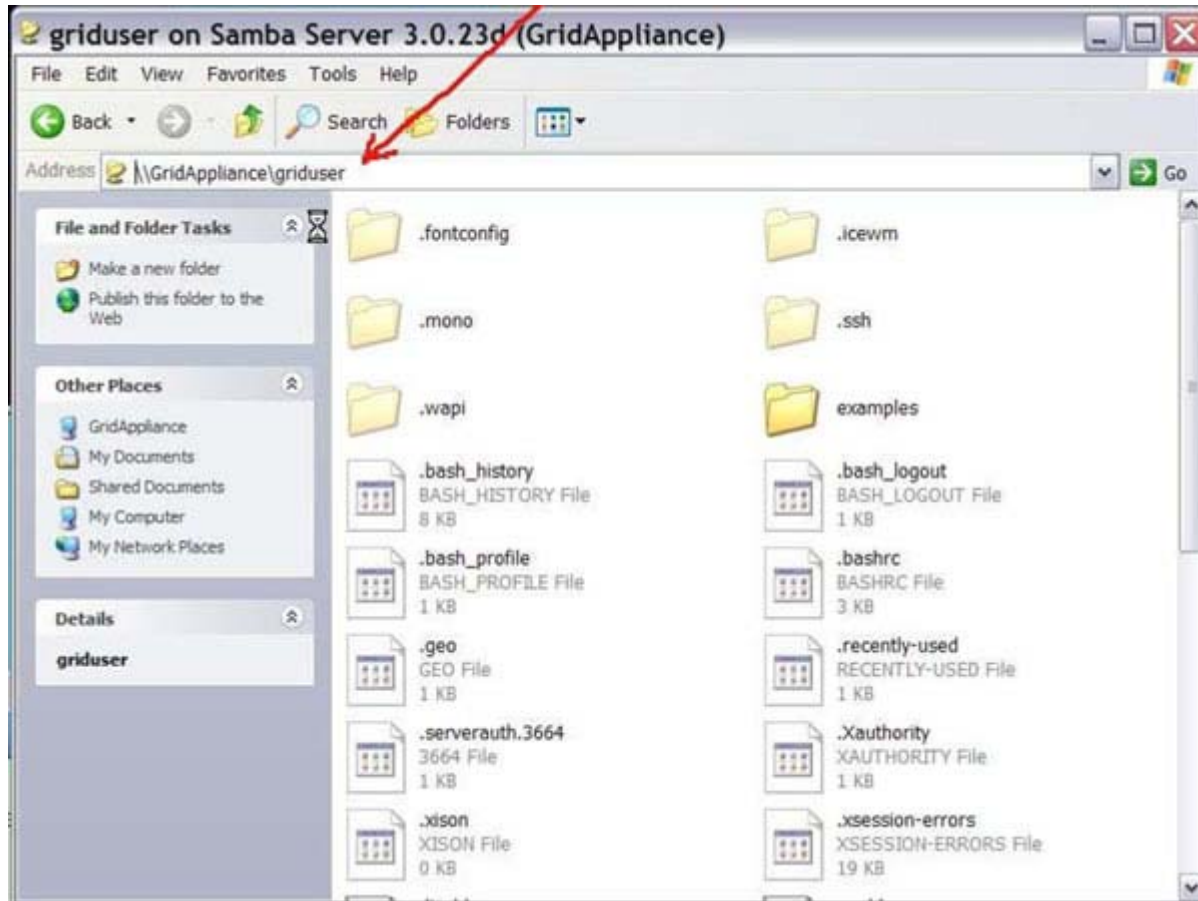
Transferring files to and from the Grid Appliance

There are two ways to transfer files to and from the Grid Appliance (over Samba or over SCP/SFTP). Samba is the recommended approach since it's much more user friendly, faster, and is built into Windows. SCP/SFTP is recommended for Linux users that do not have access to a Samba client.

SAMBA

To access files using Samba, simply type the following address into your internet explorer address bar:

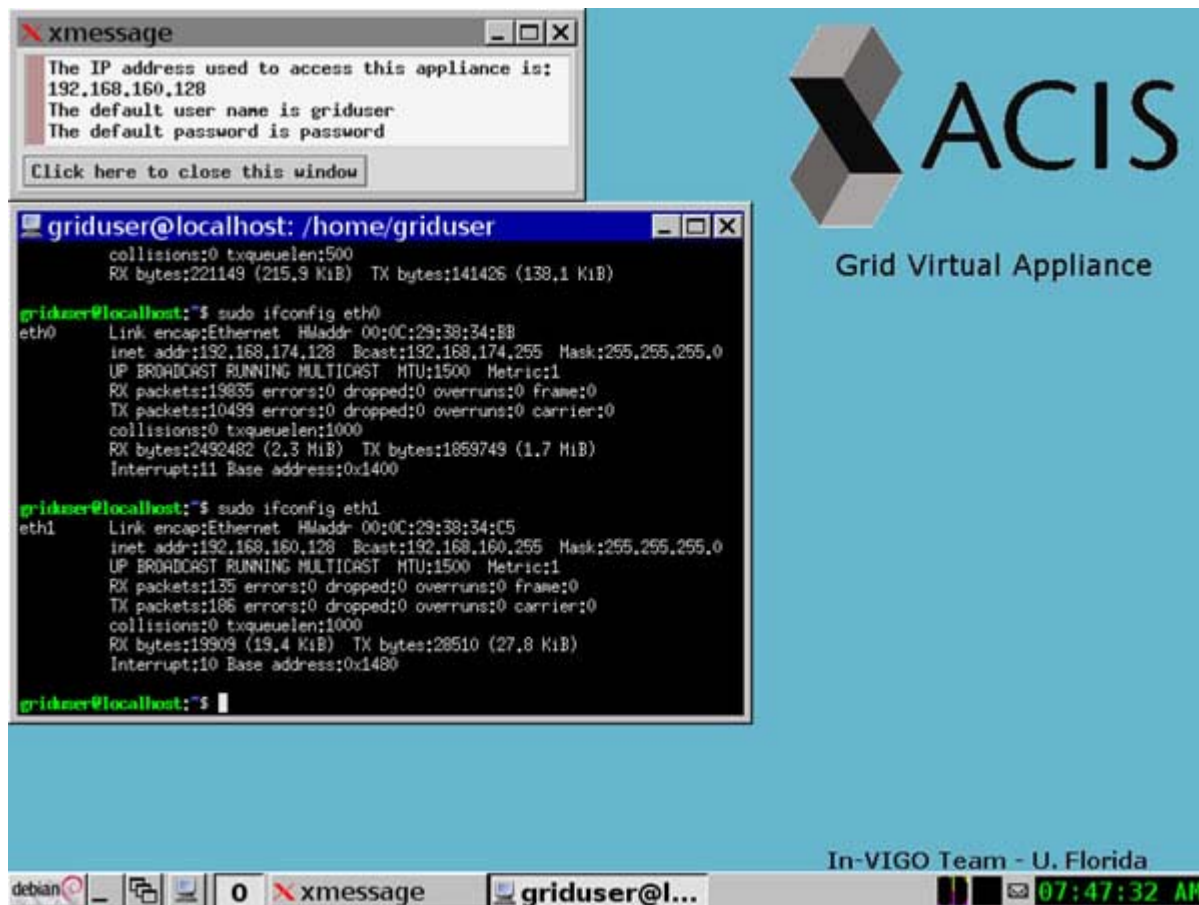
<file://gridappliance/griduser>, a new window will pop-up with the files in your home directory inside of the virtual appliance (/home/griduser). You will be able to drag-and-drop content to and from the grid appliance in the manner:



** If that does not work, use the address "\\ipaddress\griduser", where IP address is the one given in the top left xmessage box. In the screenshot example shown above, this would be \\192.160.160.128\griduser

SCP/SFTP

To access files from your Linux command line, type "sftp griduser@ipaddress" (e.g. sftp griduser@192.168.160.128). Follow the command line instructions and provide your password and you should have access. Once logged in, you can use command such as ls, and cd to see files on the remote directory. Then you get use "put" and "get" to get or put a file. Again, the IP address is specified in the xmessage top left window.



Submitting a Condor job through the Grid Appliance

Most users will primarily interface with the job submission portion of condor. The primary commands that users need to familiarize themselves with are "condor_submit", "condor_q" and "condor_status".

To explain their purpose, we will use one of the provided examples on the Grid Appliance. This demonstration uses a [Monte Carlo method to estimate the value of Pi](#). This example illustrates a typical use of high-throughput computing to execute many different tasks using the Condor scheduler. This application takes a single argument (how many data points to generate). Based on this parameter, it generates random points in a unit square and compute the ratio of points that fall within a unit circle to estimate Pi. Check the file `montepi.c` within `examples/montepi` for the source code.

1. Type the following commands on a shell terminal:
2. `cd examples/montepi`
3. `condor_submit submit_montepi_vanilla`
4. `condor_q`
5. `condor_status`

```

griduser@localhost: /home/griduser/examples/montepi
1.61 griduser 10/19 08:02 0+00:00:00 I 0 9.8 montepi 20000
1.62 griduser 10/19 08:02 0+00:00:00 I 0 9.8 montepi 20000
1.63 griduser 10/19 08:02 0+00:00:00 I 0 9.8 montepi 20000
1.64 griduser 10/19 08:02 0+00:00:00 I 0 9.8 montepi 20000
1.65 griduser 10/19 08:02 0+00:00:00 I 0 9.8 montepi 20000
1.66 griduser 10/19 08:02 0+00:00:00 I 0 9.8 montepi 20000
1.67 griduser 10/19 08:02 0+00:00:00 I 0 9.8 montepi 20000
1.68 griduser 10/19 08:02 0+00:00:00 I 0 9.8 montepi 20000
1.69 griduser 10/19 08:02 0+00:00:00 I 0 9.8 montepi 20000
1.70 griduser 10/19 08:02 0+00:00:00 I 0 9.8 montepi 20000
1.71 griduser 10/19 08:02 0+00:00:00 I 0 9.8 montepi 20000
1.72 griduser 10/19 08:02 0+00:00:00 I 0 9.8 montepi 20000
1.73 griduser 10/19 08:02 0+00:00:00 I 0 9.8 montepi 20000
1.74 griduser 10/19 08:02 0+00:00:00 I 0 9.8 montepi 20000
1.75 griduser 10/19 08:02 0+00:00:00 I 0 9.8 montepi 20000
1.76 griduser 10/19 08:02 0+00:00:00 I 0 9.8 montepi 20000
1.77 griduser 10/19 08:02 0+00:00:00 I 0 9.8 montepi 20000
1.78 griduser 10/19 08:02 0+00:00:00 I 0 9.8 montepi 20000
1.79 griduser 10/19 08:02 0+00:00:00 I 0 9.8 montepi 20000
1.80 griduser 10/19 08:02 0+00:00:00 I 0 9.8 montepi 20000
1.81 griduser 10/19 08:02 0+00:00:00 I 0 9.8 montepi 20000
1.82 griduser 10/19 08:02 0+00:00:00 I 0 9.8 montepi 20000
1.83 griduser 10/19 08:02 0+00:00:00 I 0 9.8 montepi 20000
1.84 griduser 10/19 08:02 0+00:00:00 I 0 9.8 montepi 20000
1.85 griduser 10/19 08:02 0+00:00:00 I 0 9.8 montepi 20000
1.86 griduser 10/19 08:02 0+00:00:00 I 0 9.8 montepi 20000
1.87 griduser 10/19 08:02 0+00:00:00 I 0 9.8 montepi 20000
1.88 griduser 10/19 08:02 0+00:00:00 I 0 9.8 montepi 20000
1.89 griduser 10/19 08:02 0+00:00:00 I 0 9.8 montepi 20000
1.90 griduser 10/19 08:02 0+00:00:00 I 0 9.8 montepi 20000
1.91 griduser 10/19 08:02 0+00:00:00 I 0 9.8 montepi 20000
1.92 griduser 10/19 08:02 0+00:00:00 I 0 9.8 montepi 20000
1.93 griduser 10/19 08:02 0+00:00:00 I 0 9.8 montepi 20000
1.94 griduser 10/19 08:02 0+00:00:00 I 0 9.8 montepi 20000
1.95 griduser 10/19 08:02 0+00:00:00 I 0 9.8 montepi 20000
1.96 griduser 10/19 08:02 0+00:00:00 I 0 9.8 montepi 20000
1.97 griduser 10/19 08:02 0+00:00:00 I 0 9.8 montepi 20000
1.98 griduser 10/19 08:02 0+00:00:00 I 0 9.8 montepi 20000
1.99 griduser 10/19 08:02 0+00:00:00 I 0 9.8 montepi 20000

100 jobs: 99 idle, 1 running, 0 held
griduser@localhost:~/examples/montepi$

```

At 3, we use "condor_submit" which takes as its input a condor job configuration file. Use "more submit_montepi_vanilla" to inspect this file. Condor submit files can be much more complex than this one, but this example illustrates a configuration that is useful for many simple applications. At 4, we check the progress of our jobs - each job is assigned an identifier, and the status "I" means a job is idle, and "R" means a job is running. At 5, we check the status summary of the entire Condor pool and other jobs running in the system.

While the jobs run, take a look at the Condor submit configuration file:

```
more submit_montepi_vanilla
```

You will notice it specifies the Condor "Universe" used (vanilla supports unmodified executables), the name of the application executable (montepi), and the arguments used for the application. In this example, 50 jobs are queued to run with argument 10000, then 50 additional jobs are queued with argument 20000.

Once the 100 jobs have completed, the results will be placed in files with extensions "*.out". Type the following command to view all 100 estimates of Pi which have been computed:

```
grep pi *.out
```

The files with extension "*.log" show where each Condor job executed. Type the following command to see the virtual IP address of each appliance which executed a job:

```
grep executing *.log
```

Changing your password in the Grid Appliance

To change your password is pretty straightfoward, simply type:

```
passwd
```

in the command line. You will be asked to provide the current default password, then set your own password.

Gaining root access in the Grid Appliance

Since the Grid Appliance is based off Debian, it uses the famous sudo tool to provide a subset of root privileges to a user. There are two ways for getting root access:

1. Use the sudo before every command (i.e. sudo ifconfig), or
 2. Type "sudo bash" to gain root access
-

How to SSH into the Grid Appliance

The Grid Appliance allows ssh connections from the host machine only. You can use the address displayed in the intro screen to connect to the Grid Appliance. If you are using Windows, [PuTTY](#) (a free ssh client) is recommended.

- Once you have PuTTY, double-click on the executable,
 - type "gridappliance" or the actual IP address obtained from the display window, and click open.
 - A login window will show up, use "griduser" as the login name and "password" for the password (unless you changed it).
-

Last Updated (Thursday, 08 November 2007 02:51)