PolyServe

PolyServe Understudy

QuickStart Guide

underSTUDY™

Hi-Availability Server Clustering for E-Business

PolyServe Understudy QuickStart Guide

PolyServe Understudy QuickStart Guide

This QuickStart Guide is meant for customers and evaluation users to quickly install, activate, configure, and test Understudy in their server environment.   Although the user documentation covers most of the material in this guide, it acts more as a reference manual for users to refer back to if they have any questions on specific topics. This QuickStart Guide, on the other hand, provides a step-by-step installation, configuration, and testing guide for first time users to quickly get Understudy installed, configured and working. Understudy must be activated by a license code that we supply once you register or purchase the product. Instructions for acquiring and activating the license are given below.

Understudy Software Distribution & Registration

Understudy is distributed in a variety of ways, including electronic and CD distribution. You can acquire a copy of our software in a number of ways:

- Download an evaluation copy from the web
- Get a CD copy directly from PolyServe or any of our distributors
- Install an already bundled copy in your server or operating system

Regardless of which method you use to get a copy of Understudy, each copy is fully functional and may be locked and require an activation license code from PolyServe or our distributors. Purchased copies of Understudy that are sent to a customer via CD or electronically may already have the license code installed, and in this case the license is active immediately upon installation. Some bundled versions of Understudy may already be unlocked and ready to operate immediately "out of the box." All other copies of the software need to be unlocked with the license code.

Versions of Understudy that are locked include copies downloaded from our web site and bundled versions that are present in the server or operating system release. Understudy is bundled in some systems (such as FreeBSD 4.0) and can be installed directly from the release, but still need to be activated with a license code. Instructions in your system release will take you through this process. You can also get instructions on getting a license code for your bundled version by going to our web site at www.polyserve.com.

To get a license code you must either register for a 30-day evaluation license or purchase the product. Instructions on purchasing our product are provided near the end of this document. Instructions for registering for a 30-day evaluation license are provided below.

Downloading an Evaluation Copy of Understudy
To download a copy of the Understudy software for the particular version you are interested in testing, you must first register and request an evaluation copy of the software. Once you have registered for an evaluation copy, you will be able to download one of the versions of Understudy on any of the supported platforms (see our web site www.polyserve.com for information on supported platforms). The downloaded software is in the particular format, which is native to the platform you will be testing

Understudy upon. For example, on Red Hat Linux this is RPM format; Windows/NT this is a standard setup file; on FreeBSD this is a tarball.

After Understudy is downloaded, a license code will be returned to you via email. Retain this license code, as it will allow you to activate Understudy for a 30-day evaluation of the product. This will be discussed in more detail in the installation & activation section.

You should also download the Understudy documentation in PDF format. The documentation is ideal as a reference manual for using and configuring Understudy. Our support page provides links and technical information for the evaluation version you will be installing.

Registering a CD copy of Understudy
If you purchased Understudy and received it as a CD, then it is already activated with a permanent license code. If you received an evaluation copy of Understudy on CD, then you must register your copy on the Polyserve web site to receive your 30-day evaluation license. The license will be returned to you via email.

Registering a Bundled copy of Understudy
PolyServe is pre-bundled in some operating system and server platforms. If you have a bundled version of Understudy, you simply need to register it to get a 30-day evaluation license, or purchase it to get a permanent unlock license code. You can go to our web site for information about registering bundled versions of Understudy. The system or software vendor will also provide information on how to register Understudy.   The following software or servers are bundled with Understudy (and you can register by following the link):

- FreeBSD 4.0

Installing & Activating Understudy

This section describes the basics of installing Understudy on the platform desired, as well as activating Understudy with the license code. A brief overview of the Understudy Console Administration utility is also given. This utility allows you to configure and manage Understudy from a system different than the ones running the Understudy software.

Installing Understudy on your Servers
Once you have obtained a copy of Understudy, you can now install it on the appropriate platform. A set of release notes exists for each of the versions of Understudy. If you have a bundled version of Understudy (this means that it is bundled in the operating system or server release) the release notes, documentation, and other technical information for this bundled release are available on our website. The release notes contain information on:

- Minimum Requirements for Understudy (Prerequisites)
- Important Information (Known Problems)
- Quick Installation Guide

The Understudy software must be installed on a pair of servers to fully appreciate its value.

Activating the Understudy License Code
Once you have installed Understudy on the desired platform, you must activate Understudy by correctly installing the license code.   A license code will be returned to you after you have registered for an evaluation license or purchased a copy of Understudy. This license code must be installed in the proper location after Understudy is installed for it to perform its failover and load balancing operations successfully. This section describes what must be done to properly install the purchase or evaluation license. *Please be aware that for evaluation licenses, the 30-day evaluation period commences the day the product is registered and you receive the license, NOT when it is installed!*

A license code will be returned to you via email after you have registered or purchased the product. Purchased versions delivered by CD will already be unlocked and a license code is not required. The license code will look something like:

understudy 12-31-1999 e4e473c45aa69936379b793e857b1c3f

OR

understudy permanent e4e473c45aa69936379b563e857b1c3f

Where 12-31-1999 is the date the evaluation version of Understudy expires. 'Permanent' is for purchased versions. The character string at the end is the encrypted license key. Do not attempt to change any of the information in the string, as this will prevent Understudy from activating successfully. You must place the entire string in the license file (the word understudy, the date, and the encrypted key) for Understudy to activate successfully.

This string must be placed in an ascii file with the name *license* (lower case, no extension). Copy and paste this string into the file so that you will not mistakenly type any of the characters. The file must be placed in the proper directory according to the table below:

Platform	Directory location for '*licens*e' file
Cobalt Server	/etc/polyserve
Red Hat Linux	/etc/polyserve
Microsoft Windows NT	%systemroot% (directory where Windows NT is installed …typically the C:\WINNT directory)
Slackware Linux	/etc/polyserve
FreeBSD	/usr/local/etc/polyserve
Sun Solaris	/etc/polyserve
Debian Linux	/etc/polyserve

Understudy License Code Violations & Problems
Understudy must be activated by properly placing the license code in the license file under the correct directory for the operating system that it is installed upon. If the Understudy license is not properly activated or if the 30-day evaluation period expires, Understudy will generate a license code violation

error message in both the log file and on the Understudy console (if the admin GUI console program is running). If Understudy indicates the product will expire on Dec. 31, 1999 it means the default evaluation license file shipped with the product has not been replaced with the correct license file.

It is important to be aware of the fact that the Understudy evaluation license is good for 30-days and starts from the moment you receive the email from PolyServe containing the license code string. A purchased version of Understudy must have a license string with the word 'permanent' to activate permanently.

When Understudy detects a license violation, it first gives the following message on the Understudy console host:

> License Error on Server 'servername':
> expired on mm-dd-yyyy; spare will be
> terminated in 2 hours

It also places this message into the Understudy Log file. The log file is a file in which Understudy puts important messages about its operation. To read the log file, select the server you want to view the log file upon (from amongst the servers in the console portion of the admin GUI dialog box). Then select the menu option **SERVER LOG** from the Understudy menu options, and finally select the option **READ**, and the log file will appear. The Log file will contain the same message that was displayed on your console screen, but with the following license error code attached:

> (internal error code = 0 – 18, error code 417)

Every 15 minutes Understudy will generate another license violation message on the console screen and in the log file, until it finally times out. At this point the Understudy daemon, spare, will exit. You may start Understudy again, at which point you will then be given the license violation message again and will have 2 hours of time until the Understudy daemon fully times out.

If you attempt to connect to a cluster server with Understudy (or you attempt a command in the Understudy console) and a license violation is in progress, Understudy will generate the following error message on the console:

> Licensing Problems on the server 'servername'
> See log file for details

If you need assistance with the license code activation or have any problems that cannot be resolved please use the Understudy support line for help.

Installing & Using the Understudy Console Admin GUI
Understudy is configured through a Java based administration console that is referred to as the admin console. Through this Java based utility, the server/network manager can configure the server cluster to provide the desired failover and load balancing policies required for correct server cluster operation. The Understudy Users Guide describes how to start and use the Understudy console to configure the cluster. When Understudy is installed, the admin console is automatically installed along with the

daemon "spare." Spare is the process that is responsible for Understudy's failover operation. So normally the admin console would be used on one or more of the servers in the cluster. Understudy could then be configured directly from one of the servers in the cluster. This is the easiest route to configuring the cluster.

In some cases it may not be possible to run the administration GUI console from the servers that run Understudy. For example, Cobalt Servers do not come with monitors (they are server appliances). If you need or want to run the Understudy admin GUI console on a host separate from the server running Understudy, PolyServe provides a separate Administration utility for either a Windows 98/NT or a Red Hat Linux platform. The Understudy manual (Users Guide) can be used to guide you on how to use and operate the separate Administration utility. This utility must be downloaded and installed separately from the Understudy software.

If the Understudy Administration utility will be used to access servers through a firewall, you must configure the firewall to allow port 9050 to access the servers. This is the port that the Administration utility uses to access the Understudy daemon.

If you would like to use the Administration utility on a platform other than Red Hat or Windows, please contact PolyServe directly.

Configuring & Operating Understudy

Once Understudy is installed and activated, the server cluster can be configured for operation. The following sections describe how to set up the servers, the concept and configuration of a virtual host, and how to define and use service monitors. We also review how to configure a cluster for either Active-Passive operation (pimary/backup) or Active-Active operation (load balancing).

Review of Understudy Internal Operations

Understudy supports failover and backup operations for access to server services by using dynamic IP address takeover and monitoring server access and service operation. The Understudy daemon "spare" runs on each host in the cluster and monitors whether the servers are up and whether defined services (such as HTTP, FTP, SMTP, etc…) are functioning properly. The operation of the cluster is completely distributed, there is no single point of failure. When a server or service is deemed down or unavailable by Understudy, the virtual host address for that service is migrated over to another server through the dynamic IP addressing mechanism of spare.

A virtual host is defined for each service to be accessed by a client. Multiple virtual hosts can be defined per server cluster. Understudy separately monitors each virtual host. A server cluster consists of a pair of servers and at least one virtual host address. The virtual address is accessed by clients as if it were a real host name and address. For instance, an commerce site might use the access point www.commerce.com which would be a virtual host consisting of the two real servers commerce1 and commerce2. Understudy would then monitor commerce1 and commerce2 and if one or the other was down, all access by clients to www.commerce.com would go to the active host only.

Two modes of operation of Understudy failover are available: Active-Passive and Active-Active. The following sections should allow you to quickly setup your cluster for either mode.

Configuring Servers for Active-Passive Operation (Primary/Backup)

An Understudy cluster consists of two or more servers that will act as a single entity for client access purposes. Version 1.3 of Understudy supports a pair of servers, version 1.6 (available in February) supports up to 10 servers per cluster. The servers in the cluster must all run the Understudy software. Understudy supports completely heterogeneous clustering. This means that different operating systems and hardware may be used on different servers in the Understudy cluster, as long as Understudy runs on each host. This section describes how to configure an Understudy cluster to support Active-Passive operation. In Active-Passive mode, a primary server handles all client requests and a backup server handles requests only if the primary is down or inactive. Whenever the primary server is up, it is the active server, and the backup handles no client requests.

The first step in configuring an Understudy cluster for Active-Active operation is to start the Understudy console admin GUI. This step is defined in the previous section of this paper. Assume there are two (2) servers in the cluster {acme1, acme2}. These are the real server names and must have unique IP addresses. These servers must be defined as members of the Understudy cluster by selecting the Menu option *CLUSTER → ADD SERVER*, then adding each of the servers {acme1, acme2) in turn through the admin GUI. For each server in the cluster you must perform this simple operation.

As was mentioned earlier in this guide, Understudy uses the *virtual host* concept for its dynamic failover operation. Whereas a real server is a physical host, a *virtual host* is a virtual server that could correspond to any of the servers in the cluster. A *virtual host* is a hostname/IP address pair that is unique (no other hostname/IP address in the network can match it) but can correspond to different physical hosts depending upon the state of that host. An example using our two real servers will show why a virtual host is different than a (real) host. As in our example above, the two servers are {acme1, acme2}. Let's say that we want to build a web server cluster where both these hosts appear to the user as a single host. As long as acme1 is up, we want all requests to go to that host. If acme1 is down, we would like acme2 to take over for it without users knowing that acme1 is down. This is what we term Active-Passive or Primary/Backup operation (one host is active, the other is passive). To set this up with Understudy, it is necessary to first define a virtual host.

The virtual host is the name that users outside the cluster will use to access the service. Let's assume that {acme1, acme2} will be accessed as www.acme.com by users. We must then configure a virtual host named www.acme.com in the Understudy console.

Virtual Host	*Primary Server*	*Backup Server*
www.acme.com	acme1	acme2

Remember (and this is very important) the virtual host name and address must be unique. No real host or IP address can match it. It is just like a real host namename/IP address in everyway but one – it is not statically defined to a single host; it can be moved (by Understudy) from host to host in the cluster. Let's see how this happens in a real-life scenario.

Assume a user accesses www.acme.com in our example above. The host name www.acme.com is not a real hostname, it is a virtual host name. Understudy assigns this virtual host to the server acme1 since it

is the primary host in the cluster, and is up (answering network requests). As long as acme1 is up, the virtual host name www.acme.com will resolve to the physical machine acme1. If acme1 goes down, Understudy deactivates this machine in the cluster, and activates acme2 in the cluster. Understudy senses that acme1 is down and acme2 takes over for all client accesses to www.acme.com. It takes over by sending a gratuitous ARP message on the LAN to tell all hosts, routers, and switches on the LAN that www.acme.com (the virtual host) requests should go to acme2. As long as acme1 is down, acme2 will be the active host in the cluster. As soon as acme1 comes back up, it will again take over as the active host in the cluster. Because Understudy uses the ARP method to dynamically assign the virtual host IP address to a server, all hosts in an Understudy cluster must run on the same LAN collision domain.

This concept of a virtual host is very powerful and allows for Active-Passive operation, as well as Active-Active load balancing operation, as we will see shortly. Before describing how to set up load balancing, we will first look at configuring virtual hosts.

Configuring Virtual Hosts and Support for ISPs and Other Service Providers
Configuring a virtual host in Understudy is a very simple process. Go to the menu option *CLUSTER→ADD VIRTUAL HOST* and simply add the virtual host name you desire (in our example it would be www.acme.com). If Understudy won't let you add it, the most likely cause is that you did not define this as a valid host name and IP address on the network. You will be asked if you want the primary host to be on acme1 and the backup on acme2 (or vice-versa).

What if you are an ISP or service provider and would like to host multiple sites on a single server but also want high availability and failover? Understudy allows you to define multiple virtual hosts per cluster. This is ideal for ISPs or other service providers. An example will help you to set this up. Again, let's use the example above with the two-host cluster {acme1, acme2}. Assume that we have already configured the virtual host www.acme.com. Now we want to add a second web site to the cluster. Let's say this site will be called www.roadrunner.com. To add this we simply add another virtual host to the cluster by again going to the menu option *CLUSTER→ADD VIRTUAL HOST* and add the virtual host name www.roadrunner.com. We then must decide if we want the primary to be on acme1 and the backup on acme2 (or vice-versa). Our suggestion is that you choose the primary on acme2 and the backup on acme1 – this will help statically balance the load for www.acme.com and www.roadrunner.com.

The Understudy GUI will then show two virtual hosts (www.acme.com and www.roadrunner.com) under each of the servers. You guessed it, now we can add a third virtual host (www.coyote.com) to the cluster to support another customers web site. Understudy can support up to 128 virtual hosts per cluster. By using multiple virtual hosts and service monitors, ISPs and other service providers can support multiple web sites per server, with per website failover. The key to this is the service monitor concept which we will explore in the next section.

Configuring and Using Service & Application Monitors
Understudy can be configured to monitor not only whether the server is up, but also whether specific services are actually working. Understudy supports service monitors for HTTP, SMTP, FTP, and generic TCP/IP. These service monitors run specific tests to validate that the service is functioning properly. A service monitor is defined per virtual host. In this way, each virtual host can be separately

monitored and failover policies for each virtual host can be defined. For instance, if we define an HTTP service monitor for each of the virtual hosts
[www.acme.com, www.roadrunner.com, www.coyote.com] and the HTTP service fails on www.roadrunner.com, then failover for only that virtual host will occur. The other virtual hosts will not be affected.

The Service Monitors run locally on each host and validate the status of the service on that host only. In case of a service failure, Understudy on the host where the failure occurs notifies the other hosts in the cluster of the failure so that service failover can occur. The following table describes the function of each service monitor.

Service Monitor	Operation (how it validates service)
HTTP	Generates HTTP head record and waits for reply
SMTP	Connects to SMTP port and waits for reply
FTP	Connects to FTP port and waits for reply
TCP/IP	Connects to the defined TCP/IP port and validates that the connection was opened correctly

The generic TCP/IP port monitor allows you to build your own service monitor to monitor any application that is TCP/IP based. This is ideal for validating homegrown applications or well known TCP/IP applications that are not currently supported by the existing Understudy service monitors.

Configuring Understudy for Active-Active Operation (Load Balancing)
Understudy can be configured to support Active-Active Operation with load balancing, using DNS round-robin. DNS round-robin load balancing is an effective and simple way to share the load amongst multiple servers. Its main drawback is that if a server is down, DNS does not know this and continues to send requests to the failed server. With Understudy, DNS can be configured so that this never happens, thereby removing the main disadvantage of using DNS round-robin.

The white paper **DNS Round Robin High Availability Load Sharing** discusses in some detail how to configure DNS Round Robin to support load sharing in an Understudy cluster. The following short discussion is a review of this paper.

The following example is a short dual-server cluster example that shows what must be done to setup round robin load sharing with Understudy. Assume that there are two hosts {acme1, acme2} and that you desire client requests to be load balanced between {acme1, acme2} under the name www.acme.com. To setup load sharing for these two hosts do the following:

1. Define the virtual host 'virtual_acme1' as the virtual cluster consisting of acme1 as the primary and acme2 as the backup.
2. Define the virtual host 'virtual_acme2' as the virtual cluster consisting of acme2 as the primary and acme1 as the backup
3. Add DNS A-record entries for www.acme.com that specify the IP addresses for virtual_acme1 and virtual_acme2 as the round robin "hosts."

4. Set the DNS A-record TTL values low (60 seconds or less) so that intermediate name servers do not cache server addresses where the server is down.

Once this is complete, client requests will be rotated between virtual_acme1 and virtual_acme2. The virtual host virtual_acme1 will point to acme1 (as long as it is up), the virtual host virtual_acme2 will point to acme2 (as long as it is up). If either real host (acme1 or acme2) is down, then the virtual host will be redirected to the backup host. In this way, DNS will not know that the real host is down, but will send requests instead to the operational host. Refer to the white paper for more details on this configuration.

There are also ways to easily configure DNS round robin for ISPs that require multiple web sites per server. Contact PolyServe support for more info on these configurations.

Testing & Evaluating Understudy

After you have successfully installed and configured Understudy, we recommend a set of basic tests to validate the operation of Understudy in a failure or load balancing mode. After these tests are successfully run, then you may elect to run a more substantial test of your specific requirements to validate that Understudy is working in your environment. We especially recommend the first set of test (Testing Primary/Backup Operation & Failover) as this validates the basic configuration of Understudy.

Testing Primary/Backup Operation & Failover
The first step in validating the successful operation of Understudy on your system is to verify that automatic failover and recovery reintegration works. It would be best to run these tests on non-operational systems in a test mode. To run this test we suggest the following procedure:

> *TEST #1 – Server failure and reintegration test*
>
> *Step 1: Using the Understudy console, login to the backup Understudy server*
>
> *Step 2: Configure the Cluster for a single virtual host*
>
> *Step 3: Validate that both Servers are up and that the Primary server is the active host in the virtual cluster and the backup server is inactive in the cluster*
>
> *Step 4: Pull the network connection from the Primary server and validate that the Understudy console shows the primary server as down and inactive and that the backup server shows active*
>
> *Step 5: Put the network connection back into the Primary server and validate that it re-enters the cluster as the active host and the backup becomes inactive*

Once this is successful, you can move onto the next step, which is testing failover when a service monitor fails. To run the service monitor test, we recommend the following procedure:

> *TEST #2 – Service failure and reintegration test*

Step 1: Using the Understudy console, login to the backup Understudy server

Step 2: Add a service monitor (HTTP, FTP, SMTP, or a generic TCP/IP monitor) to the virtual host already defined in the system

Step 3: Validate that both servers are up (and that the service you are testing is up) and the Primary server is the active host in the virtual cluster, and the backup server is inactive in the cluster

Step 4: Stop the specific service you are testing on the Primary server (ie for HTTP bring down the HTTP daemon) and validate that Understudy detects the service failure and shows the virtual host as down and inactive; and the backup servers' virtual host is the active host in the virtual cluster

Step 5: Start the specific service you are testing on the Primary server (ie for HTTP bring the HTTP daemon back up) and validate that Understudy detects the service as becoming active and that Understudy re-activates the primary servers' virtual host and de-activates the backup servers' virtual host

Successfully running these tests is highly recommended before attempting more advanced tests. ISPs and hosting services should also attempt multiple virtual host failover and reintegration tests as well.

Testing Load Balancing Operation & Failover

Testing the correct functionality of the load balancing configuration is a little more difficult than the Primary/Backup failover configuration. There are more components to test, including the DNS Server and multiple clients. The easiest way to test the load balancing configuration and operation is to use a pair of clients accessing the same service on a pair of servers configured with Understudy. The DNS server must be setup according to the configuration guide on Active-Active Operation discussed earlier in this paper. To test load balancing requires at least two clients to validate that each client is going to a different server (in a round-robin fashion). If the data or website that each accesses is exactly the same, it will be difficult to determine which server is handling which client request. In order to test load balancing, each website or service must be slightly altered on each of the servers to verify which server is handling the client request.

TEST #1 – Validating Correct Load Balancing Operation

Step 1: Access the load balanced service using Client #1

Step 2: Access the load balanced service using Client #2

Step 3: Validate that each client is handled by a different server.

To determine that the client-server pair is different for each request, you must have slightly altered data on each web site (or whatever service you are testing).

The next test verifies that the load balanced cluster operates correctly when one of the servers is down (or removed from the cluster). The latter part of this test validates that upon reintegration back into the cluster, the server begins to handle its portion of the load balanced traffic.

> ### *TEST #2 – Validating Correct Load Balancing Operation when a Server is Down*
>
> *Step 1: Using the Understudy console login to one of the Servers in the load balanced cluster.*
>
> *Step 2: Pull the network connection from the Server that you are NOT logged into.*
>
> *Step 3: Validate that Understudy senses the Server is down and that the backup for this virtual host takes over operation from the Primary*
>
> *Step 4: Access the service you are load balancing from both clients separately and validate that both client requests are handled correctly. DNS round robin will still be used, but now both requests should go to the remaining server that is up.*
>
> *Step 5: Put the network connection back into the Server and again access the service from both clients and validate that a unique server is handling each client request.*

If either of the above tests fail, the most likely cause is that the virtual host IP addresses were not correctly placed into the DNS server A-records.

PolyServe Support Policy

PolyServe offers a basic support agreement as a package for purchasing customers. This package is available during the 30-day evaluation period of Understudy at no cost. Please refer to our support page for contact information, hours of service, and basic support contract information.

Converting an Evaluation Version to a Purchased Version

When you have completed an evaluation to your satisfaction and would like to purchase Understudy, you can follow the procedure shown in the next section to purchase Understudy. When you purchase Understudy you have three (3) options for delivery of the software:

1) Get a new copy on CD with a permanent license key already installed
2) Get a permanent license key and install it using the same procedure that was used to activate the 30-day license key
3) Get both a CD copy and the license key

The option you choose is up to you. Please discuss these options with PolyServe at the time of purchase.

Purchasing Understudy

There are a variety of ways to purchase Understudy and a support contract. The best method is to contact PolyServe directly at 877-4POLYSERVE.