Mac OS X Server

Xgrid Administration
For Version 10.4 or Later

Contents

About This Guide

This guide describes the Xgrid components included in Mac OS X Server and tells you how to configure and use them in computational grids.

Xgrid in Mac OS X Server version 10.4 includes a controller for computational grids and an agent that allows the server's processor to work on jobs submitted to a grid. The agent is also available in computers using Mac OS X v10.3 or v10.4.

What's in This Guide

This guide includes the following chapters:

- Chapter 1, "Introducing Xgrid," highlights important concepts and introduces the tools you use to configure and manage grids.
- Chapter 2, "Setting Up Xgrid," explains how to set up the Xgrid agent and controller using the Server Admin application and how to plan a grid.
- Chapter 3, "Managing Grids and Jobs With Xgrid Admin," tells you how to manage agents, controllers, and jobs using the Xgrid Admin application.
- Chapter 4, "Planning and Submitting Xgrid Jobs," describes how to plan jobs and submit them using the Terminal application.

Note: Because Apple frequently releases new versions and updates to its software, images shown in this book may be different from what you see on your screen.

Using This Guide

The chapters in this guide are arranged in the order that you're likely to need them when setting up Mac OS X Server to configure Xgrid components and manage agents and jobs in a grid.

*   Review to acquaint yourself with Xgrid concepts.
*   Follow the instructions in Chapter 2 to configure the Xgrid components on the server.
*   Follow the instructions in Chapter 3 to manage agents, controllers, and jobs with Xgrid Admin.
*   Consult Chapter 4 for information about planning and submitting jobs to a grid.

Using Onscreen Help

You can view instructions and other useful information from this and other documents in the server suite by using onscreen help.

On a computer running Mac OS X Server, you can access onscreen help after opening Server Admin or Xgrid Admin. When Server Admin is running, from the Help menu, select one of these options:

*   *Server Admin Help* displays information about the application.
*   *Mac OS X Server Help* displays the main server help page, from which you can search or browse for server information.
*   *Documentation* takes you to www.apple.com/server/documentation, from which you can download server documentation.

When running Xgrid Admin:

*   *Xgrid Admin Help* displays information about the application.

You can also access onscreen help from the Finder or other applications on a server or on an administrator computer. (An administrator computer is a Mac OS X computer with server administration software installed on it.) Use the Help menu to open Help Viewer, then choose Library > Mac OS X Server Help.

To see the latest server help topics, make sure the server or administrator computer is connected to the Internet while you're using Help Viewer. Help Viewer automatically retrieves and caches the latest server help topics from the Internet. When not connected to the Internet, Help Viewer displays cached help topics.

The Mac OS X Server Suite

The Mac OS X Server documentation includes a suite of guides that explain the services and provide instructions for configuring, managing, and troubleshooting the services. All of the guides are available in PDF format from:

www.apple.com/server/documentation/

This guide...	tells you how to:
Mac OS X Server Getting Started for Version 10.4 or Later	Install Mac OS X Server and set it up for the first time.
Mac OS X Server Upgrading and Migrating to Version 10.4 or Later	Use data and service settings that are currently being used on earlier versions of the server.
Mac OS X Server User Management for Version 10.4 or Later	Create and manage users, groups, and computer lists. Set up managed preferences for Mac OS X clients.
Mac OS X Server File Services Administration for Version 10.4 or Later	Share selected server volumes or folders among server clients using these protocols: AFP, NFS, FTP, and SMB/CIFS.
Mac OS X Server Print Service Administration for Version 10.4 or Later	Host shared printers and manage their associated queues and print jobs.
Mac OS X Server System Image and Software Update Administration for Version 10.4 or Later	Use NetBoot and Network Install to create disk images from which Macintosh computers can start up over the network. Set up a software update server for updating client computers over the network.
Mac OS X Server Mail Service Administration for Version 10.4 or Later	Set up, configure, and administer mail services on the server.
Mac OS X Server Web Technologies Administration for Version 10.4 or Later	Set up and manage a web server, including WebDAV, WebMail, and web modules.
Mac OS X Server Network Services Administration for Version 10.4 or Later	Set up, configure, and administer DHCP, DNS, VPN, NTP, IP firewall, and NAT services on the server.
Mac OS X Server Open Directory Administration for Version 10.4 or Later	Manage directory and authentication services.
Mac OS X Server QuickTime Streaming Server Administration for Version 10.4 or Later	Set up and manage QuickTime streaming services.
Mac OS X Server Windows Services Administration for Version 10.4 or Later	Set up and manage services including PDC, BDC, file, and print for Windows computer users.
Mac OS X Server Migrating from Windows NT to Version 10.4 or Later	Move accounts, shared folders, and services from Windows NT servers to Mac OS X Server.

This guide...	tells you how to:
Mac OS X Server Java Application Server Administration For Version 10.4 or Later	Configure and administer a JBoss application server on Mac OS X Server.
Mac OS X Server Command-Line Administration for Version 10.4 or Later	Use commands and configuration files to perform server administration tasks in a UNIX command shell.
Mac OS X Server Collaboration Services Administration for Version 10.4 or Later	Set up and manage weblog, chat, and other services that facilitate interactions among users.
Mac OS X Server High Availability Administration for Version 10.4 or Later	Manage link aggregation, load balancing, and other hardware and software configurations to ensure high availability of Mac OS X Server services.
Mac OS X Server Xgrid Administration for Version 10.4 or Later	Manage computational Xserve clusters using the Xgrid application.
Mac OS X Server and Storage Glossary	Interpret terms used for server and storage products.

Getting Documentation Updates

Periodically, Apple posts new onscreen help topics, revised guides, and solution papers. The new help topics include updates to the latest guides.
• To view new onscreen help topics, make sure your server or administrator computer is connected to the Internet and click the Late-Breaking News link on the main Mac OS X Server help page.
• To download the latest guides and solution papers in PDF format, go to the Mac OS X Server documentation webpage: www.apple.com/server/documentation.

Getting Additional Information

For more information, consult these resources:

Read Me documents—important updates and special information. Look for them on the server discs.

Mac OS X Server website—gateway to extensive product and technology information.
www.apple.com/macosx/server/

AppleCare Service & Support website—access to hundreds of articles from Apple's support organization.
www.apple.com/support

Apple customer training—instructor-led and self-paced courses for honing your server administration skills.
train.apple.com/

Apple discussion groups—a way to share questions, knowledge, and advice with other administrators.
discussions.info.apple.com/

Apple mailing list directory—subscribe to mailing lists so you can communicate with other administrators using email.
www.lists.apple.com/

Preface About This Guide

Introducing Xgrid

1

Use Xgrid to create grids of multiple computers and distribute complex jobs among them for high-throughput computing.

Xgrid, a technology in Mac OS X Server and Mac OS X, simplifies deployment and management of computational grids. Xgrid enables administrators to group computers into grids or clusters, and allows users to easily submit complex computations to groups of computers (local, remote, or both), as either an ad hoc grid or a centrally managed cluster.

Xgrid Terminology

The Xgrid technology uses specific terms for its components and operations. These include:
- Grid: a group of computers that can collaboratively complete a job using the Xgrid technology in Mac OS X Server and Mac OS X.
- Job: a set of work submitted into a grid from the client to the controller.
- Task: a part of a job that one agent in the grid performs at one time.
- Controller: an Xgrid controller manages the grid and its work. It is built into Mac OS X Server.
- Agent: an Xgrid agent resides on one computer in a grid and runs tasks sent to it by the controller. Any computer running Mac OS X v10.3 or v10.4 can run the Xgrid agent.
- Client: any computer running Mac OS X v10.4 or Mac OS X Server v10.4 that submits a job to an Xgrid controller.

Xgrid Basics

Xgrid creates multiple tasks for each job and distributes those tasks among multiple nodes. These nodes can be desktop computers running Mac OS X v10.4 or v10.3, or systems running Mac OS X Server v10.4. Many desktop computers sit idle during the day, most evenings, and on weekends. The assembly of these systems into a computational grid is called "desktop recovery." This method of grid construction allows you to vastly improve your computational capacity without purchasing additional hardware, and Xgrid makes the software configuration a straightforward task.

For a server to function as a controller, Xgrid requires Mac OS X Server v10.4 or later, with a minimum of 256 MB of RAM. To operate as an agent in a grid, Xgrid requires Mac OS X v10.3 or later, with a minimum of 128 MB of RAM (256 MB recommended). All Xgrid participants must have a network connection. As always, the more RAM a system has, the better it will perform, particularly for high-performance computing applications.

A computational grid is a group of computers working together to solve a single problem. The systems in a grid can be loosely coupled, geographically dispersed, and, to some extent, heterogeneous. In contrast, systems in a cluster are often homogeneous, collocated, and strictly managed. Highly dispersed grids, such as SETI@Home, allow individuals to donate their spare processor cycles to a cause. In office environments, large rendering or simulation jobs may be distributed across all the systems left idle overnight. These can even be used to augment a dedicated computational cluster, which is available to Xgrid clients at all times. These three distinct grid configurations are discussed in more detail in "Three Ways to Use Xgrid" on page 13.

Xgrid has no real limitations on the amount of computational power it can support. The performance of the grid is dependent on the systems participating, the software running, and the network, among other factors. Individual applications, however, will strongly influence the performance of the grid. It is up to the user to determine how amenable a particular application is to being deployed on a computational grid. In the best case, application performance may scale linearly with the size of the grid. In the worst case, the addition of agents to a grid may cause the given job to complete in even more time than if there had been fewer agents. (In such a situation, tasks become so small that the overhead associated with distributing the increased number of tasks supersedes the performance gain of using more agents.) Users of the grid should be aware of these considerations.

There are many proprietary projects that allow you to participate in a large computational grid. Often these projects, such as SETI@Home and FightAids@home, are tied to a specific scientific purpose. They often have easy-to-install software that enables any volunteer to participate in that particular project, and they often take the form of a screen saver or background process.

But you don't need to think in terms of thousands or millions of seldom-used computers to see the significance of a computational grid. For example, computers used by university students and corporate employees often "work" fewer hours than the hours they sit idle at night or on weekends. These computers could contribute productively to the work of a grid without diminishing their usefulness to the students and employees.

Other grid projects are designed for large-scale computational grids, such as the Globus Alliance (a group founded by universities and researchers), with flexible resource management tools and more intelligent grid deployment methods. Instead of developing neatly packaged applications for a specific grid, such projects provide comprehensive frameworks for application deployment.

Xgrid allows users to participate in a computational grid of their choice, while still providing the flexibility of a more generic framework to grid developers in deploying their grid applications. Xgrid provides the primary benefits of both.

The advantages of the Xgrid technology include:
- Easy grid configuration and deployment.
- Straightforward yet flexible job submission.
- Automatic controller discovery by both agents and clients.
- Flexible architecture based on open standards.
- Support for the UNIX security model including Kerberos single sign-on or regular password authentication.
- Choice between a command-line interface or an API-based model for grid interaction.

Three Ways to Use Xgrid

Xgrid can be used in tightly coupled clusters, worldwide grids, and everything in between. This immense flexibility allows you to deploy grids of almost any nature. Three main topologies are commonly used, however, and these cover the vast majority of Xgrid deployments. They are:
- Xgrid clusters
- Local grids
- Distributed grids

Xgrid Clusters

Computational clusters are sets of systems entirely dedicated to computation. In a cluster, systems are typically collocated in a rack, connected via gigabit Ethernet or another high-performance network, and strictly managed for maximum performance. Cluster systems are often entirely homogeneous; their operating systems are the same versions, they have the same software installed, and they generally have the same processor, disk, and RAM configurations.

Xgrid allows administrators to easily configure the distributed resource management functionality of the cluster. Each server in the system runs the agent software, and the head node in the cluster runs the controller software. Xgrid distributes tasks across the cluster. In clusters, failure rates are generally very low. Systems are rarely, if ever, offline, and their resources are not shared with general user tasks. Clusters are the most efficient and most expensive model of distributed computing.

Local Grids

Systems that are under common administration in a company, university computer lab, or other managed environment can often be easily assembled into a grid for desktop recovery. Because these systems are often on a local area network, and because they are generally managed by a single organization, they provide the ability for good network performance and substantial manageability.

Because these systems are often also used as day-to-day workstations, users can easily interrupt grid tasks by moving the mouse, resetting the system, or even accidentally disconnecting the system from the network. In such cases, a task may fail as part of an Xgrid job. The Xgrid controller will eventually reassign the failed task to another agent, and the job will complete successfully. In local grids, performance is limited by such situations, as well as by the varying performance of any given agent on the grid.

Distributed Grids

The Xgrid agent allows a user to specify any IP address or host name for its desired controller. By specifying a grid, a user can dedicate his or her CPU time to that grid no matter where the controller is located. When any system is permitted to donate its time, a distributed grid is formed. The manager of the controller has no direct management control or knowledge of the agent system, but is nonetheless able to harness its CPU time.

Distributed grids have very high failure rates for jobs, but very low administrative burden for the grid administrator. With very, very large jobs, high task failure rates may not substantially affect the performance of the grid if they can be rapidly reassigned to another available agent. Network performance may also be a consideration, as data will be sent over the Internet, rather than over a local network, to agents connected to a grid. The monetary cost of such distributed grids is extremely low.

Xgrid Components

The Xgrid architecture consists of a controller, agents, and one or more client computers. Clients submit jobs to the controller.

An Example of Xgrid Components and Configuration

The illustration below shows the Xgrid components and the process of auto-configuration for a grid.

Distributed agents

4 Client submits using mDNS, DNS, or name/address

2 Agents locate Controller using mDNS, DNS, or name/address

1 Controller advertises via mDNS

Controller

Client

5 Clients and Controller mutually authenticate using passwords or single sign-on

3 Agents and Controller mutually authenticate using passwords or single sign-on

Dedicated Desktop

Dedicated Server

Part-time Desktop

The primary components of a computational grid perform these functions:

- An agent runs one task at a time per CPU; a dual-processor computer can run two tasks simultaneously.
- A controller queues tasks, distributes those tasks to agents, and handles task reassignment.
- A client submits jobs to the Xgrid controller in the form of multiple tasks. (A client can be any computer running Mac OS X v10.4 or Mac OS X Server v10.4.)

In principle, the agent, controller, and client can run on the same server, but it is often more efficient to have a dedicated controller node.

An Example of How Xgrid Works

The illustration below shows how a grid handles a job.

Distributed agents

1 Client submits job to Controller

2 Controller splits job into tasks, then submits tasks to Agents

3 Agents execute tasks

Dedicated Desktop

Dedicated Server

Part-time Desktop

Controller

Client

5 Controller collects tasks and returns job results to Client

4 Agents return tasks to Controller
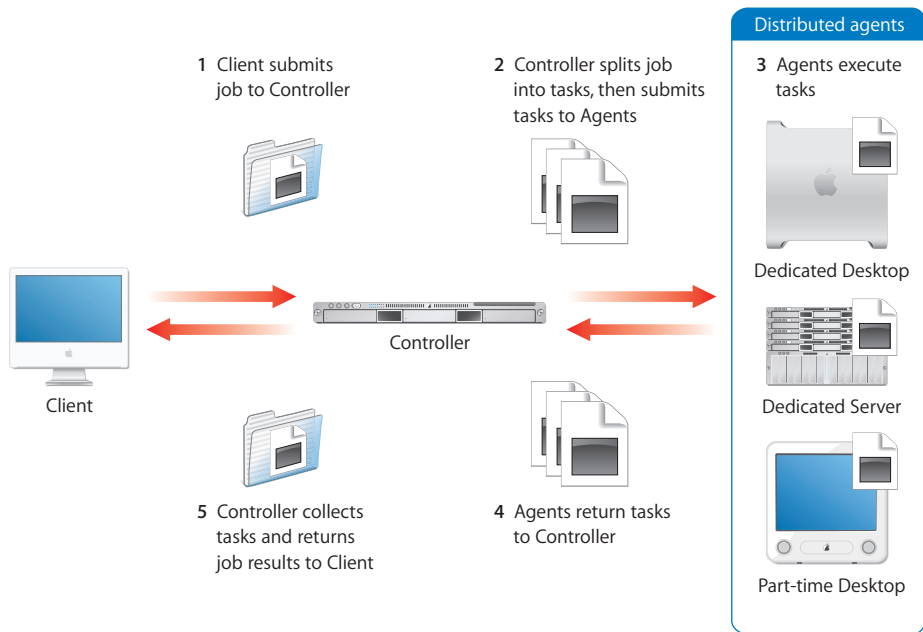
Agent

The Xgrid agents run the computational tasks of a job. In Mac OS X Server, the agent is turned off by default. When an agent has been turned on and becomes active at startup, it registers with a controller. (An agent can be connected to only one controller at a time.) The controller sends instructions and data to the agent as appropriate for the controller's jobs. Once it receives instructions from the controller, the agent executes its assigned tasks and sends the results back to the controller.

By default, agents seek to bind to the first available controller on the local network. Alternatively, you can specify that it bind to a specific controller.

You can also specify whether an agent is always available or is available only when the computer is idle. By default, the agent on Mac OS X Server is dedicated and the agent on a Mac OS X computer (not a server) is configured to accept tasks only when the system has had no user input for 15 minutes.

For details of configuring an agent, see "Configuring an Xgrid Agent on Mac OS X Server" on page 22. For information on managing agents, see "Managing a Grid" on page 31.

Client

Any system can be an Xgrid client provided it is running Mac OS X v10.4 or later, or Mac OS X Server v10.4, and has a network connection to the Xgrid controller system. In general, the client can connect to only a single controller at a time. Depending on how a controller is configured, the client must supply a password or be authenticated by Kerberos (single sign-on) before submitting a job to the grid.

A user submits a job to the controller from a system running the Xgrid client software, usually a command-line tool accessed with the Terminal application. The job can specify the controller or use multicast DNS (mDNS) to dynamically discover the first available controller. When the job is complete, the controller notifies the client and the client can retrieve the results of the job.

For information about client authentication to the controller, see "Setting Controller Passwords (for Agents and Clients)" on page 26.

Controller

The Xgrid controller manages the communications among the computational resources of a grid. The controller requires Mac OS X Server v10.4 or later. The controller accepts network connections from clients and agents. It receives job submissions from the clients, breaks the jobs up into tasks, dispatches tasks to the agents, and returns results to the clients.

Although there may be more than one Xgrid controller running on a subnet, there can be only one controller per logical grid. Each controller can have an arbitrary number of agents connected, but Apple has tested only 128 agents per controller. There is no software limitation on the number of agents, however, and users of Xgrid may choose to exceed 128 agents on a controller at their own risk with a theoretical maximum equal to the number of available sockets on the controller system.

For details of setting up an Xgrid controller, see "Configuring an Xgrid Controller" on page 24. For information about managing controllers and grids, see "Managing a Grid" on page 31.

Jobs

A job is a collection of execution instructions that may include data and executables. Xgrid can execute scripts, utilities, and custom software (anything that doesn't require user interaction).

A client submits a job to the grid. The controller accepts the job and its associated files, divides the job into tasks, then distributes the tasks to the agents. Agents accept the tasks, perform the calculations, and return the results to the controller, which aggregates them and returns them to the appropriate client.

For detailed information about jobs, see "Planning Jobs for Xgrid" on page 35.

Chapter 1   Introducing Xgrid

Setting Up Xgrid

2

Plan your grid and then use the Server Admin application to set up the Xgrid agent and controller.

Before configuring Xgrid, you need to make some decisions about the grid. For example, you should how to use the controller and agent in the server and any desktop agents available to you.

Planning Your Grid

Before configuring Xgrid, you need to understand what kind of grid environment you're trying to create. In particular, there are three major decisions you need to make:
*   The kind of authentication to use
*   Where to host your controller
*   How you will manage the controller

Authentication for the Grid

Xgrid requires controllers to mutually authenticate with both clients and agents. There are three authentication options.

Single Sign-On (SSO)
Single sign-on is the most powerful and flexible form of authentication. It leverages the Open Directory and Kerberos infrastructure in Mac OS X Server v10.4 to manage authentication behind the scenes, without direct user intervention. Each entity has a Kerberos principal, which passes the appropriate ticket to determine identity, which is then checked against the relevant group to determine privileges. Generally, you should use this option if any of the following conditions is true:
*   You already have SSO in your environment.
*   You have administrative control over all the agents and clients in use.
*   Jobs need to run with special privileges (such as for local, network, or SAN file system access).

Password-Based Authentication

You have the option of password authentication when you can't use single sign-on. You may not be able to use SSO if:

• Potential Xgrid clients are not trusted by your SSO domain (or you don't have one)
• You want to use agents across the public Internet or that are otherwise outside your control
• It is an ad hoc grid, without any ability to prearrange a web of trust

In these situations, your best option is to specify a password. Note that you have two distinct password settings: one for controller-client and one of controller-agent. Ideally these should be different phrases, for security reasons.

Note: It is possible to create hybrid environments, such as with client-controller authentication done via passwords, but controller-agent authentication done with SSO (or vice versa).

No Authentication

The option of no authentication should not be used. It creates a potential security hole (since anyone can then connect or run a job) and thus should never be used on any system exposed to the public Internet, especially when potentially sensitive data is involved. In fact, this option is appropriate only for testing a private network in a home or lab that is inaccessible from any untrusted machine, or when none of the jobs or the machines contain any sensitive or important information.

Where to Host Your Controller

The primary requirement for a controller is that it must be network-accessible to all potential clients and agents. In some cases this may mean the controller has to live outside an organizational firewall (or inside a buffer zone); otherwise you would need to open up port 4111 so the controller can be contacted.

It is much simpler (though not essential) for the controller to be on the same subnet as the agents and usual clients, so they can discover each other using mDNS. If that's not possible, you should host the controller on a server with a fixed IP address and fully qualified DNS name (or, alternatively via Dynamic DNS and a service lookup entry) so that agents and clients know where to find it.

How You Will Manage the Controller

In general, you will manage the controller like any other service running on Mac OS X Server, using either Server Admin or the Xgrid command-line tool to manage which processes are running, and either Xgrid Admin or the Xgrid command-line tool to manage the queues on the controller.

The amount of management required will also depend on how many queues you have and the number (and temperament) of the users who will be submitting jobs. The version of Xgrid in Mac OS X Server v10.4 uses a simple first-in, first-out (FIFO) queue for scheduling each individual grid, which means that as administrator you will need to get your colleagues' cooperation to ensure appropriate resource allocation among multiple users.

Note: On a server with multiple network interfaces, Mac OS X Server makes Xgrid service available over all interfaces. You can't configure Xgrid service separately for each interface.

Setting Up Xgrid in Server Admin

Using Server Admin, you can start Xgrid service and separately configure the controller and the agent on your server.
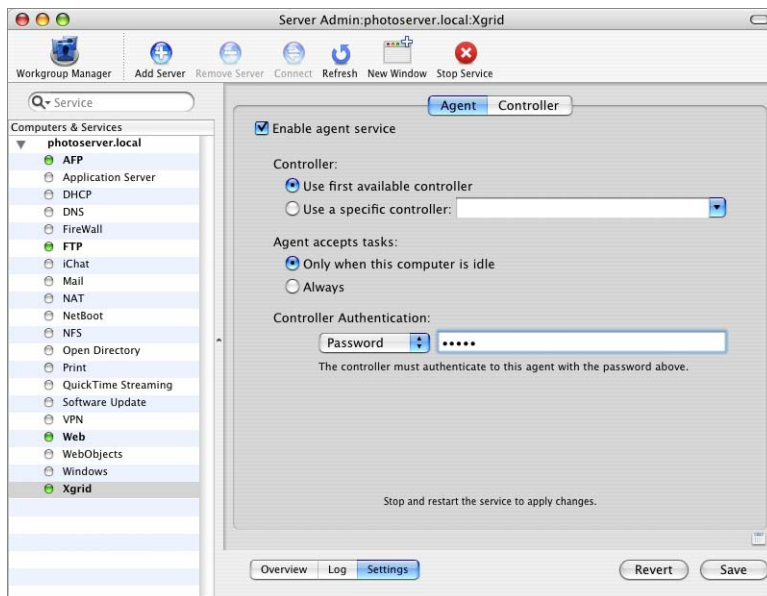
Starting and Stopping Xgrid

You start or stop the Xgrid service in Server Admin, and you also use Server Admin to enable the Xgrid controller and agent on your server. The Xgrid service must be running for your server to control a grid or participate in a grid as an agent. See "Configuring an Xgrid Agent on Mac OS X Server" on page 22 and "Configuring an Xgrid Controller" on page 24 for details on using the server as an agent and controller.

The Xgrid controller and agent are disabled by default.

To start or stop Xgrid:

1 Open Server Admin.

2 Click Xgrid in the list for the server you want.

3 Click Start Service or Stop Service.

 You must confirm your choice when stopping the service.

Configuring an Xgrid Agent on Mac OS X Server

You use Server Admin to set up the Xgrid agent on your server. In addition to enabling the agent, you can associate the agent with a specific controller or allow it to join any grid, specify when the agent accepts tasks, and set a password that the controller must recognize.

To configure an Xgrid agent:

1 Open Server Admin.

2 Click Xgrid in the list for the server you want.

3 Click Settings in the button bar at the bottom of the window.

4 Click Agent in the button bar at the top of the window.

5 Click "Enable agent service."

 Skip this step if the agent is already checked.

6 Choose the options you want for the agent.

 To specify a controller, choose its name in the pop-up menu or enter the controller name.

 Note: An agent can find a controller in one of three ways: the administrator specifies one by hostname or IP address; the agent binds to the first controller that advertises on mDNS on the local subnet; or by means of a service lookup against the domain name server for _xgrid._tcp._ip.

7 Choose an authentication option from the pop-up menu and enter the password. (See "Setting Passwords for Xgrid" on page 25 for details.)

 Note: No password is required if you choose None, but this option provides no protection from potentially malicious use of your system.

8 Click Save to save your changes and restart the service.

 Important: If you require authentication, both the agent and the controller must use the same password or must authenticate using Kerberos (single sign-on). See "Setting Passwords for Xgrid" on page 25 for details of authentication options.

 Your choices remain in effect until you change them.

Configuring an Xgrid Desktop Agent

While Server Admin provides a simple interface for enabling Xgrid services on one server or across a entire rack of Xserve systems, it doesn't provide a way to configure Xgrid on desktop machines running Mac OS X v10.3 or v10.4. If you are relying on volunteers to provide desktop agents, you can simply email them instructions for enabling Xgrid from the Sharing pane of System Preferences. (If they are using Mac OS X v10.3, they'll need to first download the Xgrid v1.2 agent from www.apple.com/acg/xgrid and then use the Xgrid pane of System Preferences.)

If you administer a group of computers and want them all to participate in a grid using Xgrid, there are several ways to do it. You can use these methods:
• Apple Remote Desktop
• SSH
• NetBoot or NetInstall

Apple Remote Desktop

Apple Remote Desktop (ARD) v2.1 is a separate product, available from Apple, that integrates many common administrative tasks across multiple machines (screen sharing, software installation, running UNIX scripts, and so on). You can use ARD to remotely run System Preferences on each machine, but it is usually simpler to change the preferences once and then "push" the new preferences file (/Library/Preferences/com.apple.xgrid.agent.plist) to all the relevant nodes. See the *Apple Remote Desktop Administration* guide at www.apple.com/server/documentation for more information about ARD.

SSH

If you don't have ARD but have set up SSH log-ins, you can do the same thing using SCP command-line tool (or even rsync, if you've set that up). Another simple way to accomplish the same thing is to use the xgridctl utility, with the following commands:

```
$ ssh root@remotehost xgridctl agent enable
$ ssh root@remotehost xgridctl agent on
```

For more details, see the man pages for SSH, SCP, SFTP, or rsync in the Terminal application.

NetBoot or NetInstall

For large networks, it often makes sense to use a common system image that is mounted or installed by each agent to configure the agents. While Xgrid in itself isn't reason enough to use NetBoot, it is a reason to consider whether using NetInstall would simplify your general administrative tasks. For more information, see the *Mac OS X Server System Image and Software Update Administration* guide at www.apple.com/server/documentation.

Configuring an Xgrid Controller

You use Server Admin to configure an Xgrid controller. When configuring the controller, you may also set a password for any agent using the grid and for any client that submits a job to the grid.
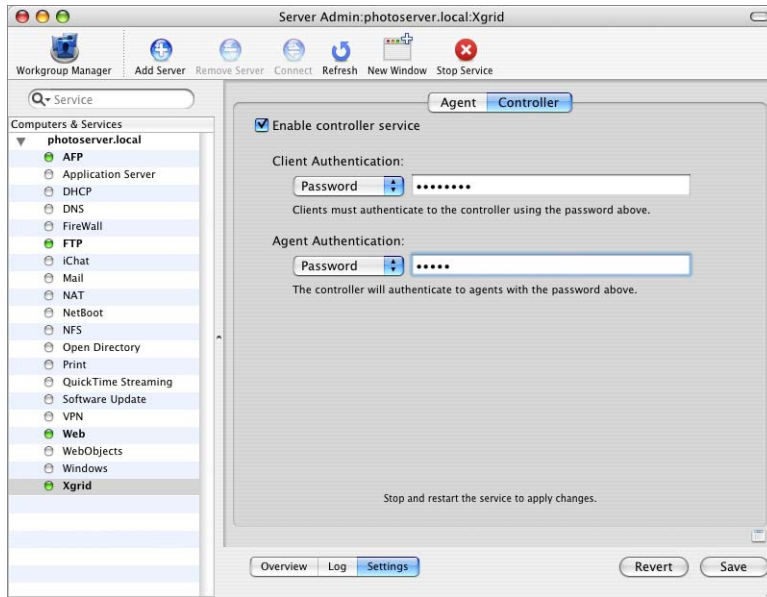
To configure an Xgrid controller:

1 Open Server Admin.

2 Click Xgrid in the list for the server you want.

3 Click Settings in the button bar at the bottom of the window.

4 Click Controller in the button bar at the top of the window.

5 Click "Enable controller service."

Skip this step if the agent is already checked.

6 Choose an authentication option for clients from the pop-up menu and enter the password. (Clients submit jobs to the controller.)

See "Setting Passwords for Xgrid" on page 25 for details on password options.

Note: No password is required if you choose None, but this option provides no protection from potentially malicious use of your system.

7 Choose an authentication option for agents from the pop-up menu and enter the password.

See "Setting Passwords for Xgrid" on page 25 for details on password options.

Note: No password is required if you choose None, but this option provides no protection from potentially malicious use of your system.

8   Click Save to save your changes and restart the service.

9   Confirm your changes by clicking Restart.

Important: If you require authentication, both the agent and the controller must use the same password or must authenticate using Kerberos (single sign-on). See "Setting Passwords for Xgrid" on page 25 for details of authentication options.

Xgrid and Multiple Network Interfaces
On a server with multiple network interfaces, Mac OS X Server makes Xgrid service available over all interfaces. You can't configure Xgrid service separately for each interface.

Setting Passwords for Xgrid
You can use passwords, Kerberos, or no authentication for Xgrid components.

You specify password options in Server Admin as part of configuring the agent and controller. See "Configuring an Xgrid Agent on Mac OS X Server" on page 22 and "Configuring an Xgrid Controller" on page 24.

About Xgrid Authentication
You set up an Xgrid controller in the Server Admin application. You can specify authentication for agents and clients, and the passwords entered in Server Admin for the controller must match those entered for each agent and client.

If you choose to have no authentication, agents can join the grid and clients can submit jobs to the grid without authenticating. This method is not recommended for any grid or job in which security is important, but it can be useful for testing a grid or running practice jobs.

Consider these points when establishing passwords for agents and clients:
* **Kerberos authentication (single sign-on).** If you use Kerberos authentication for agents or clients, the server that's the Xgrid controller must already be configured for Kerberos and in the same realm as the server running the KDC system and bound to the Open Directory master. The agent uses the host principal found in the /etc/krb5.keytab file. The controller uses the Xgrid service principal found in the /etc/krb5.keytab file.
* **Agents.** The agent determines the authentication method; the controller must conform to that method and password (if a password is used). When an agent is configured with a standard password (not single sign-on), you must use the same password for agents when you configure the controller. If the agent has specified single sign-on, the appropriate service principal and host principals must be available.

- **Clients.** If your server is the controller for a grid, you must be sure that administrators of Mac OS X and Mac OS X Server clients on your network use the correct authentication for the controller. A client cannot submit a job to the controller unless its password is the same as that for the controller, or the appropriate single sign-on (Kerberos) principals are available.

Setting the Password for a Server to Be an Agent

You use Server Admin to set the password for the server to participate in grids as an agent.

To set an agent password:

1 Open Server Admin.

2 Click Xgrid in the list for the server you want.

3 Click Settings in the button bar at the bottom of the window.

4 Click Agent in the button bar at the top of the window.

5 Choose an authentication option for the agent from the Controller Authentication pop-up menu.

Password requires that the agent and controller use the same password.

Kerberos uses the single sign-on authentication for the agent's administrator.

None does not require a password for the agent.

Note: Use the option of no password cautiously because it does not provide screening of participants in a grid.

6 Type the password for the authentication method you chose (except None).

7 Click Save to save your changes and restart the service.

8 Confirm your changes by clicking Restart.

Setting Controller Passwords (for Agents and Clients)

Passwords you set for the controller determine whether client computers can submit jobs and agents can join a grid. You use Server Admin to set controller passwords.

To set controller passwords:

1 Open Server Admin.

2 Click Xgrid in the list for the server you want.

3 Click Settings in the button bar at the bottom of the window.

4 Click Controller in the button bar at the top of the window.

5 Choose an authentication option for the client from the Client Authentication pop-up menu.

Password requires that the agent and controller use the same password.

Kerberos uses the single sign-on authentication for the agent's administrator.

None does not require a password for the agent.

Note: Use the option of no password cautiously because it does not provide screening of participants in a grid. With no authentication, a malicious agent could receive tasks and potentially access sensitive data.

6 Type the password for the client authentication method you chose (except None).

7 Choose an authentication option for the agent from the Agent Authentication pop-up menu.

Password requires that the agent and controller use the same password.

Kerberos uses single sign-on authentication for the agent's administrator.

Any allows the controller to use either the agent's specified password or the single sign-on password.

None does not require a password for the agent.

Note: Use the option of no password cautiously because it does not provide screening of participants in a grid.

8 Type the password for the agent authentication method you chose (except None).

9 Click Save to save your changes and restart the service.

10 Confirm your changes by clicking Restart.

Chapter 2   Setting Up Xgrid

Managing Grids and Jobs With Xgrid Admin

3

Use the Xgrid Admin application to manage grids, add controllers and agents, and work with jobs.

Once you have set up an Xgrid controller, you can use Xgrid Admin to manage a grid. You can use Xgrid Admin on the server or on a remote computer that is running Mac OS X v10.4.

About Xgrid Admin

Xgrid Admin is a tool you use to monitor one or more grids and manage agents and jobs.

With Xgrid Admin, you can:
- Check the status of a grid and its activity, including the number of agents working and available, processing power in use and available, and the number of jobs running and pending.
- Add or remove controllers and grids to manage.
- See a list of agents in a grid and the CPU power available and in use for each agent.
- Add or remove agents in a grid.
- See a list of jobs in a grid, the date and time each job was submitted, its progress, and the active CPU power for the job.
- Remove jobs in a grid.
- Stop a job in progress.
- Restart a job that was stopped or is complete.

Xgrid Admin provides controls in its graphical interface and menu commands for all of its options.

Note: You can also use the Xgrid command-line tool to perform all of these management tasks. See Chapter 4, "Planning and Submitting Xgrid Jobs," on page 35 for more information about using the command-line tool.

Status Indicators in Xgrid Admin
Small color bubbles indicate the status of controllers, agents, and jobs in Xgrid Admin.

The color indicators are:
- Clear = controller or agent is offline; job is pending
- Gray = job is submitting
- Green = controller is connected; agent is working; job is running
- Yellow = agent is available but not running
- Red = agent is unavailable; job is failed or canceled
- Blue = job is complete

Managing a Grid

You can manage one or more computational grids with Xgrid Admin. In this context, a grid is a fixed group of agents with a dedicated queue. There can be multiple grids per controller, but at present any given agent can belong to only a single grid. You cannot move an agent between grids while a job (or a task) is running.

Connecting to a Controller

You use Xgrid Admin to connect to a controller. The controller must simply be reachable on any network by the administrative computer running Xgrid Admin.

Once Xgrid Admin is connected to the controller, you can view the status of its grid and manage its agents and jobs.

To connect to an Xgrid controller:

1 Open Xgrid Admin and do one of the following:
   - Choose the controller from the pop-up menu or type its name and click Connect.
   - Select its name in the Controllers and Grids list and click Connect.
2 If necessary, select the appropriate authentication option and type a password, then click OK.

Adding a Controller to Xgrid Admin

You use Xgrid Admin to add a controller to the monitoring list.

To add a controller to the monitoring list:

1 Open Xgrid Admin.
2 Click Add Controller.
3 Choose a controller from the pop-up menu or type its name and click Connect.
4 If necessary, select the appropriate authentication option and type a password, then click OK.
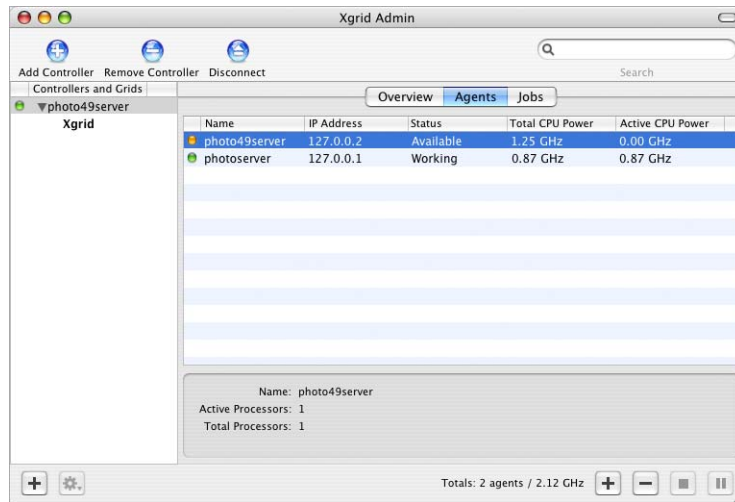
Removing a Controller

You can easily remove a controller from the monitoring list in Xgrid Admin.

To remove a controller from the monitoring list:

1 Open Xgrid Admin.
2 Select a controller in the Controllers and Grids list.
3 Click Remove Controller.

Viewing a List of Agents in a Grid

You can see a list of agents for a controller in Xgrid Admin.

To see a list of agents for a controller:

1 Open Xgrid Admin.

2 Select the grid in the Controllers and Grids list.

3 Click Agents in the button bar.

4 Select an agent in the list to display information about the CPU power and processors it uses.

The color bubble to the left of the name shows each agent's status. See "Status Indicators in Xgrid Admin" on page 30 for details.

Adding an Agent to a Grid

You can add an agent for a controller in Xgrid Admin. You can add agents that are currently offline and they will be available to the grid when their computers come online or their administrators make the agents active.

To add an agent:

1 Open Xgrid Admin.

2 Select the controller in the Controllers and Grids list.

3 Click Agents in the button bar.

4 Click the Add button (+) below the list of agents.

5 Type a name for the agent and click OK.

The agent is added to the list. The color bubble to the left of the name shows the agent's status. See "Status Indicators in Xgrid Admin" on page 30 for details.

Deleting an Agent
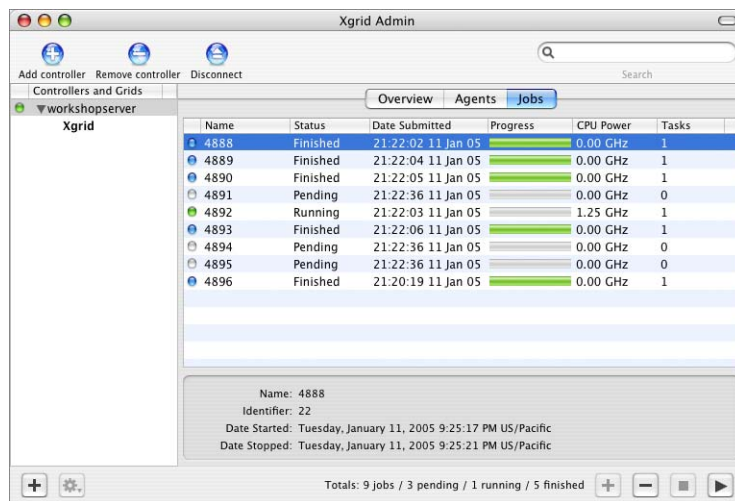
You can delete an agent for a controller in Xgrid Admin.

To delete an agent:

1 Open Xgrid Admin.

2 Select the controller in the Controllers and Grids list.

3 Click Agents in the button bar.

4 Click the Delete button (-) below the list of agents.

Note: If you mistakenly delete an agent that you know is on the local subnet and that is configured to attach to that controller, wait a few moments and it should reappear in the list. If the agent doesn't reappear, you can use the Add button and type its name to retrieve it.

Managing Jobs in the Grid

You use Xgrid Admin to manage jobs once they have been submitted by a client.

Note: It is not possible to move a job between grids.

Viewing a List of Jobs

You can see a list of jobs in Xgrid Admin.

To see a list of jobs:

1 Open Xgrid Admin.

2 Select the controller in the Controllers and Grids list.

3 Click Jobs in the button bar.

4   Select a job in the list to display details of that job.

Deciding how to structure a job may involve some experimentation to discover the best way to complete it. For instance, you might create a simple, small version of a job in two different styles, such as all tasks in one job or subdivided into multiple tiny jobs. Running both experimental jobs under similar conditions in the grid will give you a good idea of which job style is better suited to those conditions.

Repeating or Restarting a Job
You can repeat a job or restart a stopped job in Xgrid Admin.

To repeat or restart a job:
1   Open Xgrid Admin.
2   Select the controller in the Controllers and Grids list.
3   Click Jobs in the button bar.
4   Select the job you want to repeat or restart.
5   Click the Play button (the solid arrow below the list of jobs).

Deleting a Job
You can delete a job in Xgrid Admin.

To delete a job:
1   Open Xgrid Admin.
2   Select the controller in the Controllers and Grids list.
3   Click Jobs in the button bar.
4   Select the job you want to delete.
5   Click the Delete button (the minus sign below the list of jobs).

Monitoring Grid Activity
You can quickly check the activity of a grid in Xgrid Admin.

To monitor the activity of a grid:
1   Open Xgrid Admin.
2   Select the controller in the Controllers and Grids list.
3   Click Overview in the button bar, if necessary.

Planning and Submitting Xgrid Jobs

4

Use the Xgrid command-line tools and the Terminal application to submit jobs to a grid and to get information about jobs.

Once you've configured an Xgrid controller and added agents to a grid, you can use the Terminal application to send a job to the grid.

Planning Jobs for Xgrid

Planning and structuring a job carefully can result in efficient use of the grid. For example, the best structure for a job that requires multiple searches of a large database may be dividing the database into multiple sections and providing a section to each agent in the grid.

About Job Styles

Different styles of jobs often require different handling. Similarly, the way a job is structured influences how efficiently the grid completes it.

Among the job styles you should consider in your planning are:
- Everything in one single large job, with numerous small tasks.
- Everything broken into several medium-sized jobs, where each job has roughly as many tasks as there are nodes in the grid. (This type of job usually is created by a "meta-job" script, which divides the job into smaller chunks, each of which is a job in itself.)
- An entire workflow composed of several interrelated jobs.

Deciding how to structure a job may involve some experimentation to discover the best way to complete it. For instance, you might create a simple, small version of a job in two different styles, such as all tasks in one job or subdivided into multiple tiny jobs. Running both experimental jobs under similar conditions in the grid will give you a good idea of which job style is better suited to those conditions.

About Job Failure

Although not discussed in detail in this document, Xgrid jobs are permitted to rely on the message-passing interface (MPI) APIs. For jobs that rely on MPI, if a single task fails, the entire job fails and must be resubmitted. Therefore it is not recommended that MPI-based jobs be used on grids with high task failure rates.

Jobs that are embarrassingly parallel in nature are generally unaffected by occasional task failures. Tasks are typically reassigned to other available agents to complete the job. The vast majority of jobs falls into this category of problems.

Submitting a Job to the Grid

You submit jobs to a grid using the command-line tool and Terminal. There is also sample code on the Apple developer website (developer.apple.com) for alternate methods of submitting jobs.

The syntax and options for the Xgrid command-line tool are available in Terminal by entering the command:

```
man xgrid
```

Some developers and organizations may also offer specialized applications for submitting jobs to a grid. Or you can create such an application using Apple's developer tools for Xgrid.

When determining whether to use the Xgrid command-line tool or another method for submitting jobs, consider these points:
- If the job is simple, use the command-line tool.
- If you use a shell script, use the command-line tool.
- If you want to use Xgrid as part of an application with a graphical user interface (GUI), use the Xgrid API to create the GUI or incorporate it into an existing application. For more information about the API, see the *Xgrid Reference* at www.developer.apple.com/documentation.

Examples of Xgrid Job Submission and Results Retrieval

The Terminal commands that follow are examples of jobs a client can submit to the controller.

```
xgrid -h <controller. -p <password> -job submit /bin/echo "Hello,
    World!"
```

(This job copies /bin/echo to the controller and agent systems and executes it the "Hello, World!" parameter.)

```
xgrid -h <controller> -p <passworld> -job results -id <id>
```

(This command shows the results of the job with the id indicated.)

For an executable shell script "hello.sh"

```
!#/bin/sh
/bin/echo "Hello, World!"
xgrid -h <controller> -p <password> -job submit hello.sh
```

(This copies just the shell script "hello.sh" to the controller and agent systems and executes the script. /bin/echo must be installed on the agent system.)

Viewing Job Status

You can monitor jobs in Xgrid Admin (see "Managing Jobs in the Grid" on page 33 for details) or with the command-line tool.

The following commands in the Terminal application provide job status:

```
xgrid -h <controller> -p <password> -job list
Xgrid -h <controller> -p <password> -job attributes -id <job-id>
```

Chapter 4   Planning and Submitting Xgrid Jobs

Index