



Urchin® 3.4

**Installation and
Administration Guide**
for Cobalt™ RaQ3/RaQ4 Platforms

Urchin 3 Website Reporting
Industrial strength web traffic analysis

Industry-leading speed,
scalability, and efficiency

Advanced e-commerce
reporting allows true
ROI analysis

Intuitive, browser-based reports

Copyright 2002 Quantified Systems Inc., 2165 India Street, San Diego, CA 92101 USA. All rights reserved.

This product and document is protected by copyright and distributed under licenses restricting its use, copying, distribution, and decompilation. No part of this product or document may be reproduced in any form by any means without prior written authorization of Quantified and its licensors, if any. Quantified, Quantified Systems, Inc., Quantified Systems, Urchin, Urchin Pro, Urchin Dedicated, Urchin Enterprise, the Quantified logo, and the Urchin logo are trademarks, registered trademarks, or service marks of Quantified Systems, Inc. Sun, Sun Microsystems, Solaris, and the Sun logo are trademarks, registered trademarks, or service marks of Sun Microsystems, Inc. The Netscape Communications Corporation logo, Netscape, and Netscape Navigator are trademarks of AOL Time Warner Corporation. All other trademarks are property of their respective holders.

DOCUMENTATION AND SOFTWARE ARE PROVIDED "AS IS." QUANTIFIED'S LICENSOR, QUANTIFIED SYSTEMS, INC. ("QUANTIFIED"), MAKES NO WARRANTIES, EXPRESSED OR IMPLIED, INCLUDING WITHOUT LIMITATION THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE, REGARDING THE SOFTWARE. QUANTIFIED DOES NOT WARRANT, GUARANTEE OR MAKE ANY REPRESENTATIONS REGARDING THE USE OR THE RESULTS OF THE USE OF THE SOFTWARE IN TERMS OF ITS CORRECTNESS, ACCURACY, RELIABILITY, CURRENTNESS, OR OTHERWISE. THE ENTIRE RISK AS TO THE RESULTS AND PERFORMANCE OF THE SOFTWARE IS ASSUMED BY YOU. IN NO EVENT WILL QUANTIFIED, AND ITS DIRECTORS, OFFICERS, EMPLOYEES, OR AGENTS, COLLECTIVELY QUANTIFIED, BE LIABLE TO YOU FOR ANY CONSEQUENTIAL, INCIDENTAL, OR INDIRECT DAMAGES INCLUDING DAMAGES FOR LOSS OF BUSINESS PROFITS, BUSINESS INTERRUPTION, LOSS OF BUSINESS INFORMATION, AND THE LIKE, ARISING OUT OF THE USE OR INABILITY TO USE THE SOFTWARE EVEN IF QUANTIFIED HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. PLEASE DON'T SUE US, IT ISN'T NICE.

Document Version: sysadmin.cobalt.raq34.3.401.rel.us.1.0.rel

Table of Contents

1. INTRODUCTION.....	4
1.1 REVERSE DNS	4
1.2 PRIVACY AND SECURITY	5
1.3 LOG DESTINY	5
1.4 END-USER ACCESS	5
1.5 SCHEDULE.....	6
2. INSTALLATION	6
2.1 DOWNLOAD URCHIN FOR COBALT	6
2.2 OPEN THE COBALT ADMINISTRATOR	7
2.3 INSTALL URCHIN PACKAGE.....	7
3. CONFIGURATION.....	9
3.1 ACTIVITY	10
3.2 CUSTOMERS	11
3.3 REPORTS	12
3.4 USERS	15
3.5 SETTINGS	18
4. FILTERING AND USING DYNAMIC URL.....	19
4.1 FILTER IN AND FILTER OUT.....	20
4.2 DYNAMIC URL FILTERING	21
5. THE E-URCHIN MODULE (E-COMMERCE REPORTING).....	22
5.1 ARCHITECTURE	22
5.2 CONFIGURATION	23
5.3 ELF, IN DETAIL.....	24
6. TROUBLESHOOTING	26
6.1 THE URCHIN BUTTON	26
6.2 INSTALLATION FAILURES	26
6.3 REPORT DATA IS MISSING	26
6.4 ADDITIONAL TECHNICAL HELP	28

1. Introduction

Urchin Dedicated 3 is a software product targeted to hosting providers and businesses that serve one or more websites on a single server. Urchin 3 is designed to run directly on the webserver. Urchin processes log files generated from the webserver and creates easy-to-read, web-based reports that your company and/or user community will find immensely valuable and useful.

Urchin for Cobalt webservers is designed to be completely automated, running each night to keep log files in check and reports current. Once configured, Urchin requires virtually no maintenance and very little computer resources.

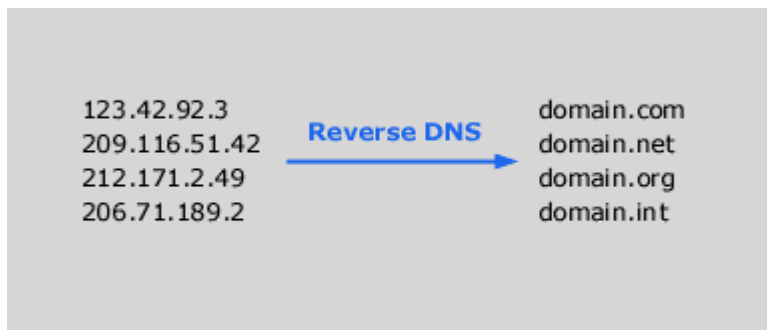
This Installation and Administration Guide contains instructions and information for installing and configuring Urchin software on Cobalt platforms. This guide is intended for experienced webmasters and system administrators who:

- Are familiar with Cobalt administration.
- Understand the webserver configuration on the system.
- Have administrator access on the system.

Chapter 2 illustrates the installation process. Chapter 3 describes how to configure Urchin for your system. Chapter 4 outlines maintenance tasks including how to add new reports to the Urchin configuration. Chapter 5 covers custom installations including e-commerce integrations. Chapter 6 has troubleshooting tips.

1.1 Reverse DNS

Reverse DNS allows you to take the raw numeric addresses of the visitors to your site and convert them into meaningful domain and network names such as aol.com or xerox.co.jp. This information can be useful in determining the source and type of visitors to your website.



Urchin 3 contains a built-in accelerated reverse-DNS module that quickly processes all visitor IP addresses. In order to use Urchin's reverse-DNS module, you will need to configure Urchin to point to a local caching capable, fully operating name server. The DNS server configured in the network control panel is usually a good choice. If you have

a particularly busy webserver, you should notify the network administrator of your intention to use a particular name server.

1.2 Privacy and Security

Urchin has a very clear privacy policy. Urchin does not use cookies; it does not aggregate data or make any connections outside your network; it does not profile users or record any personal or private information. Urchin simply reports on webserver activity. Urchin does count visitors and keeps track of which networks and browsers are being used, as well as referring sites. But all of these parameters are part of the main HTTP protocol and are passed to the server by all standard browsers.

Urchin for Cobalt includes an administrative interface that enables password controls on report access. Each website report can be individually protected. Users can be added and permissions can be assigned to each report and user individually. Urchin uses the industry standard “.htaccess” method of authentication.

If you are using the reverse-DNS module during log processing, Urchin will use UDP network sockets to connect to the DNS name server specified in the configuration. If you are operating behind a firewall, you will need to connect to a DNS server behind the firewall. Please check with your administrator for details.

1.3 Log Destiny

In order to fully automate Urchin, you will need to use the Log Destiny feature to either delete or archive logs after they are processed. Once processed, Urchin stores all pertinent information in its internal database. There is no need to keep the logs for future reporting needs. Using the delete setting will keep disk space usage to a minimum. However, if you need to archive logs for other applications or for backup purposes, you can use the archive setting.

1.4 End-user Access

The Urchin Administrative Control Panel allows you to create users that have access to different reports in the configuration. This allows you to provide web-based reporting to your user community without giving them admin access to the machine.

Once the users and permissions are created in the Urchin configuration, you can give them a central URL with which they can access the reports using their username and password. The Urchin Reporting server runs on port 888. Therefore, the URL for accessing reports is

`http://www.yourdomain.com:888/`

1.5 Schedule

Once installed, Urchin is automatically configured to run once per day. This coincides with the built-in log rotation on the Cobalt machine. The next time Urchin will run is displayed in the top of the “Activity” screen in the Urchin Administrative Control Panel .

2. Installation

The installation process on Cobalt systems uses the built in administrative interface to add the Urchin package to the system. The following steps should be followed to install the package. Once installed, refer to Chapter 3 for configuration.

2.1 Download Urchin for Cobalt

If you haven't already downloaded the Urchin package, you should do so at this time from either a web browser or via ftp at the following addresses:

`http://www.urchin.com/cobalt`
`ftp://ftp.urchin.com/pub/`

Save the package on your local machine (the desktop is usually a good place). From here it will be uploaded through the Cobalt administration interface.

2.2 Open the Cobalt Administrator

Using a browser, open your Cobalt Administration Control Panel. You will most likely be prompted for the admin password. Click on the “Maintenance” button on the left (see Fig. 1) and then the “Install Software” button at the top.

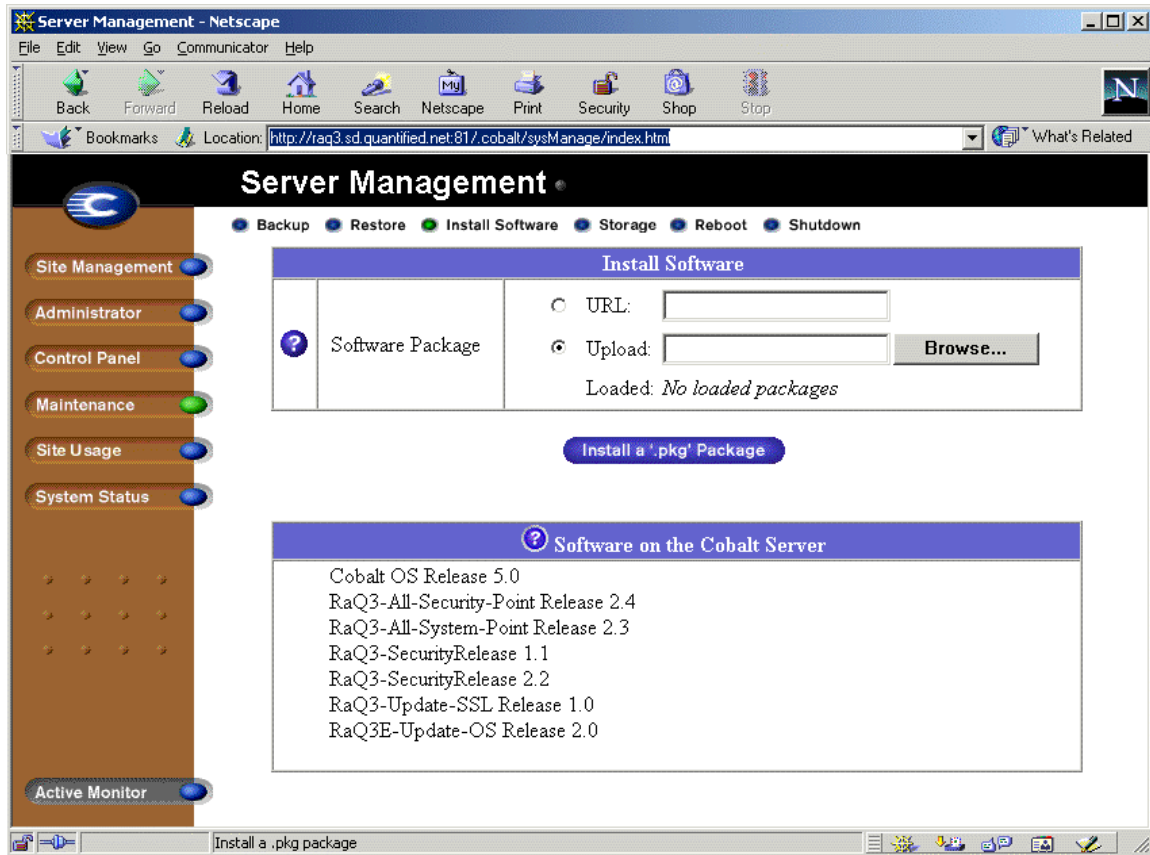


Figure 1

2.3 Install Urchin Package

In the “Install Software” window of the Cobalt Admin Control Panel shown above, click on the “Upload” radio button and the “Browse...” button to its right. Locate the Urchin package on your desktop using the browse window. The path is automatically entered into the Upload window. Click the “Install a ‘.pkg’ Package” button to begin the installation process.

When the installation has successfully completed, an entry for the Urchin package will appear in the “Software on the Cobalt Server” box (see Fig.2).

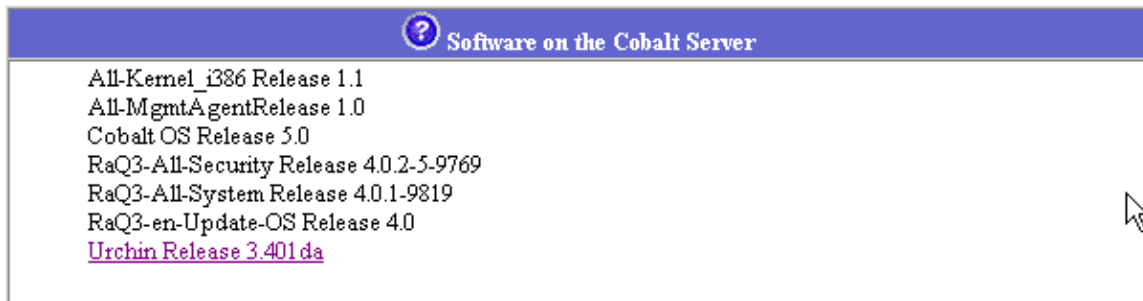


Figure 2

On RaQ3 and RaQ4 systems, this entry will be an active hyperlink that when clicked will take you to the main menu of the Urchin Control Panel. If you refresh your screen at this point, you should see the Urchin Admin button appear in the left hand frame of your screen. This button also takes you to the main menu of the Urchin Control Panel (see Fig. 3).

3. Configuration

The sections of the Urchin Administration Main Menu (Fig 3) are:

- Urchin Administration Control Panel – the center for all of Urchin’s configuration options
- Urchin Site Reports – a list of all reports available to you
- What is Urchin – an overview of Urchin’s features and capabilities
- Urchin Dedicated Cobalt Manual – an HTML version of the manual
- Troubleshooting Urchin – some advice on how to deal with installation, configuration, and management problems
- Uninstall Urchin Dedicated – an easy way to uninstall the entire Urchin program

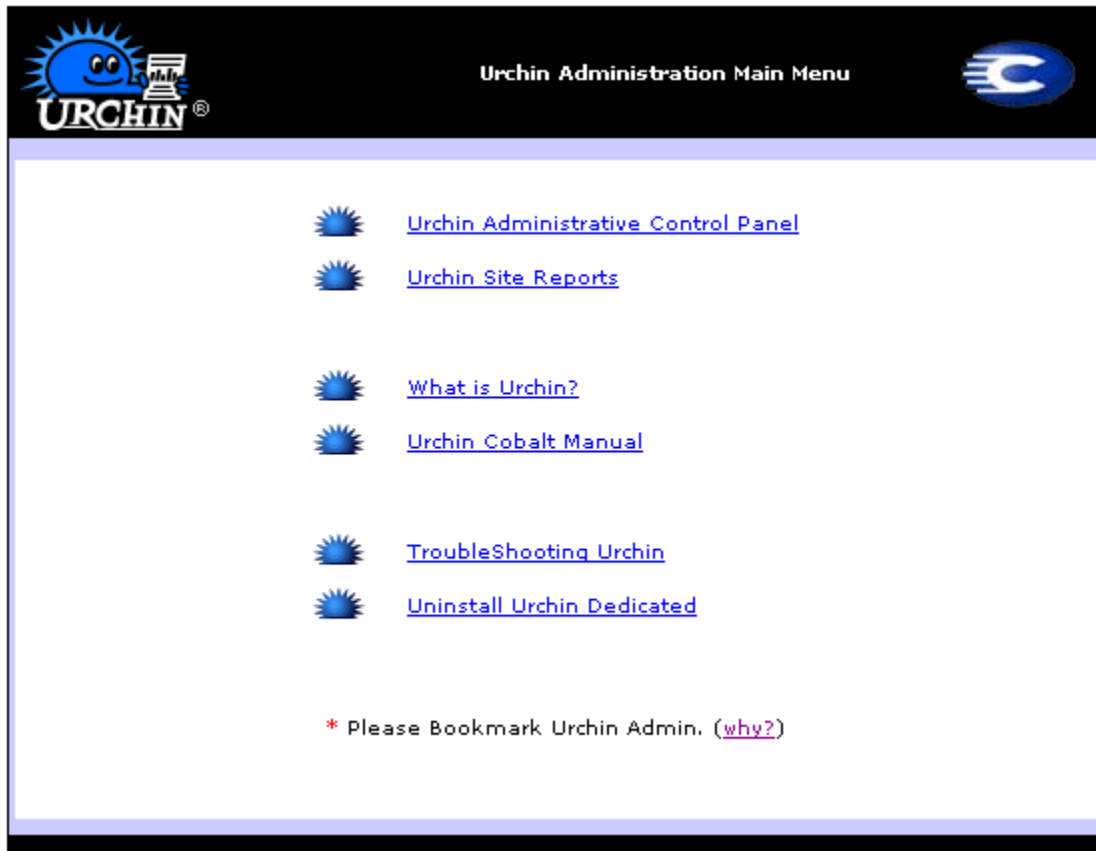


Figure 3

In the Urchin Administration main menu click on “Urchin Administrative Control Panel” to open the Urchin Administration window. You will be prompted to re-enter the admin password.

Each area of this control panel is accessed via tabs at the top of the initial screen. The following sections describe the primary configuration screens that are available in the Administration Control Panel: Activity, Customers, Reports, Users, and Settings.

3.1 Activity

The Activity tab allows you to check the status of Urchin operations and performance for each site in the configuration. The date fields at the top can be used to view historical performance of any day recorded. When you first install Urchin, there will be no activity data present.

The Begin Time and End Time are displayed for each report indicating how long it took for Urchin to process the data for that report (see Fig. 4). The Status field indicates whether the processing completed successfully. The status can be clicked on for more detailed information.

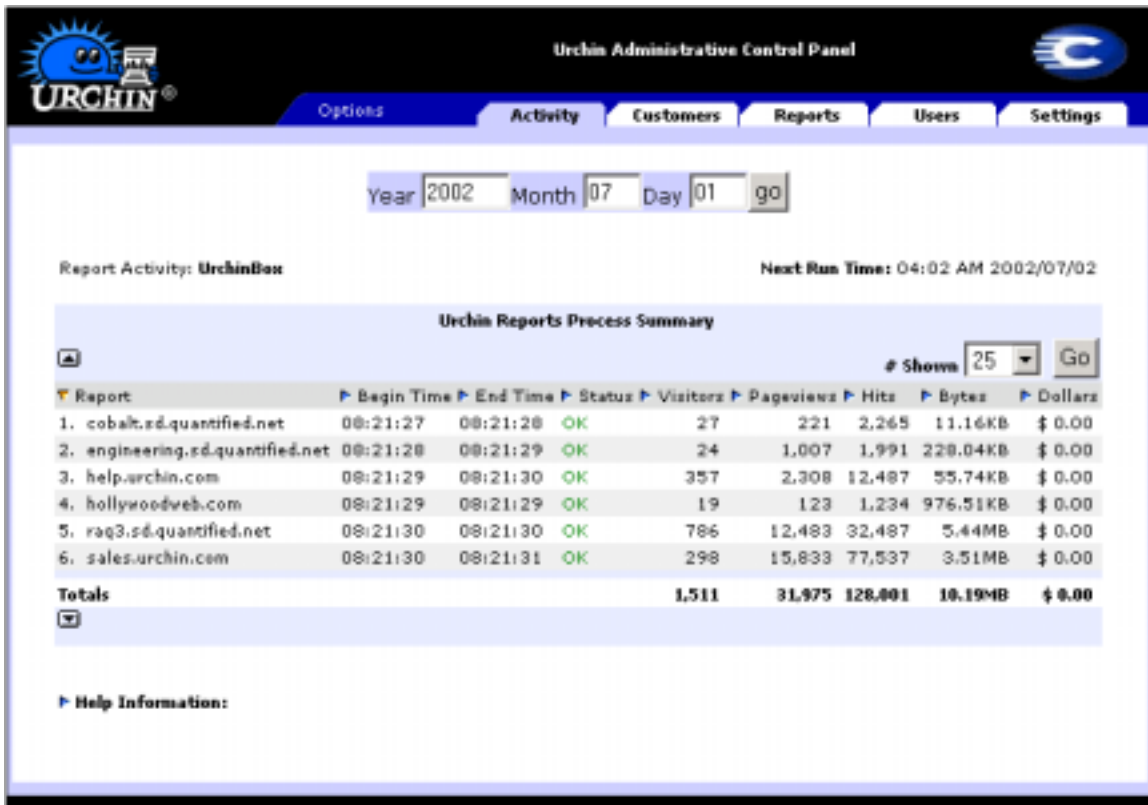


Figure 4

The number of Visitors, Pageviews, Hits, Bytes, and Revenue processed for that day are shown as well. This information can be used to monitor the overall balance and performance of Urchin processing as well as give a quick glance at how each site is doing. The Urchin Process Summary list can be sorted by any of the fields by clicking on the arrow next to the field name at the top of each column.

Note the up and down arrows at the left hand side of the Activity list. These controls will appear in any Administrative Control Panel screen that contains lists of entries in a table. They allow you to navigate forwards and backwards one screen at a time if you have a large number of entries. By default the number of table entries visible in one screen is set to 25. This setting is controlled by the box labeled # **Shown** just above the list in the upper right. To change the # **Shown** setting, choose from the choices in the drop down menu and then click the Go button.

3.2 Customers

The Customers tab (see Fig. 5) allows you to add, edit, and delete customers from the Urchin configuration. Customers are logical groups defined by the Urchin administrator. A customer group consists of reports and users who may view those reports. Each customer can have multiple reports and users, but a user may only belong to one customer. This is a convenient way to group users and reports from the same "billing" customer. Even if each of your customers only has one user and one report, you must enter the customer names here. This will ensure that users who belong to one customer group will not have access to other customers' reports.

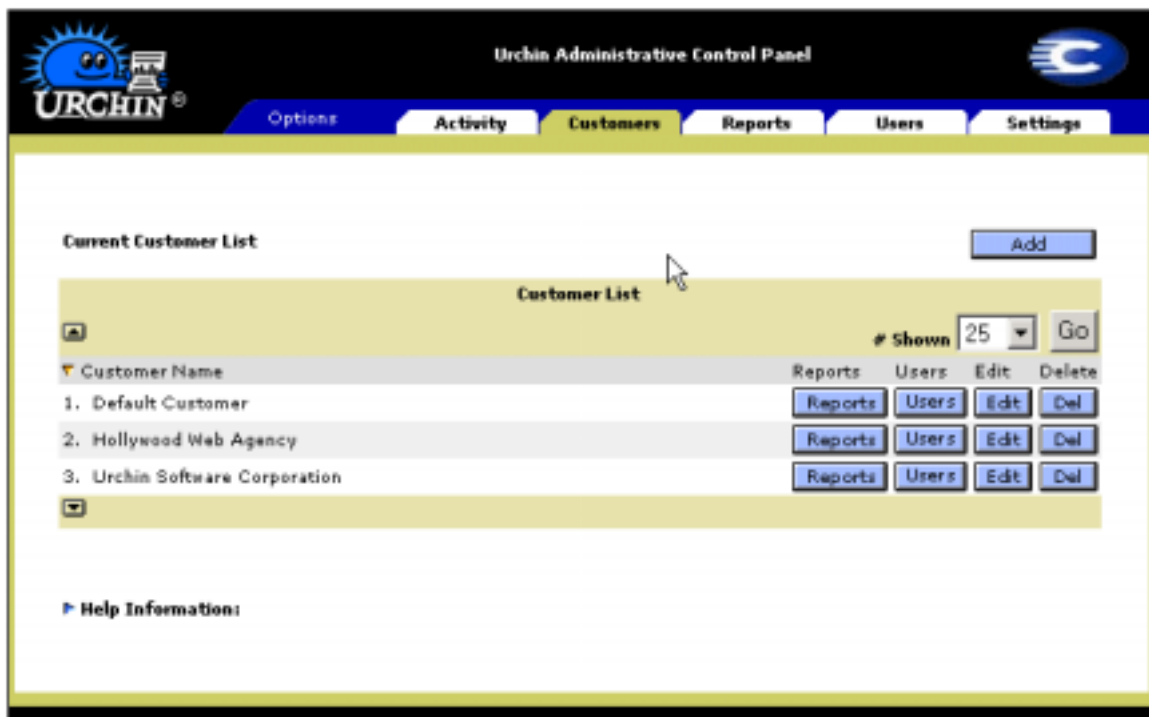


Figure 5

To add a customer to the list, simply click the "Add" button in the upper right. You will be asked for a customer name (which can include spaces and any other characters you wish). For each customer in the list, additional control buttons are provided including "Reports", "Users", "Edit", and "Del". Use the "Edit" and "Del" buttons to edit the customer name or delete a customer from the list. The "Reports" button provides quick access to the report list for that particular customer. The "Users" list provides quick

access to the user list for that particular customer, including information about which users have permission to view the various reports of the customer. When deleting a customer, all users and reports assigned to that customer must be deleted first.

3.3 Reports

The Reports tab allows you to add, edit, and delete Urchin reports for your website. Each "report" provides various statistical views for a specific website. Reports are associated with a specific customer. By default only the administrator can view the reports. However, users may be granted permission to view a particular report.



Figure 6

The Report list (see Fig. 6) displays several fields, including the name of the report and the associated customer. To sort by a particular field, click on the arrow next to that field name.

For each report in the list, buttons are provided for changing permissions, editing attributes, and deleting the report from the list. The permissions button will allow you to specify which users have access to a particular report.

The Customer pull-down menu at the top of the Reports screen is used to filter the list of reports displayed by customer name.

To add a report to the list, click the "Add" button on the upper right. Report configurations can be instantly imported from the Cobalt configuration. After clicking on the Add Report button you will see the following screen (see Fig 7):



Figure 7

Urchin provides two separate methods for adding reports to the configuration. Use the "Import Reports from Webserver Configuration" option to automatically detect existing sites in the Cobalt configuration. Urchin will allow you to import the sites as Urchin reports into the Urchin configuration.

By selecting this option you will be taken to the Import Report List table (see Fig. 8), which lists all sites currently in the Cobalt configuration. Sites that have already been added will be indicated by a check mark. To add a report to the Urchin configuration check the import box to the right of the corresponding site. Be sure to select the correct customer for each site. You may need to go back to the Customers Tab to add customers to the configuration first. Click the "Import Selected" button to add the sites you've checked to the configuration.

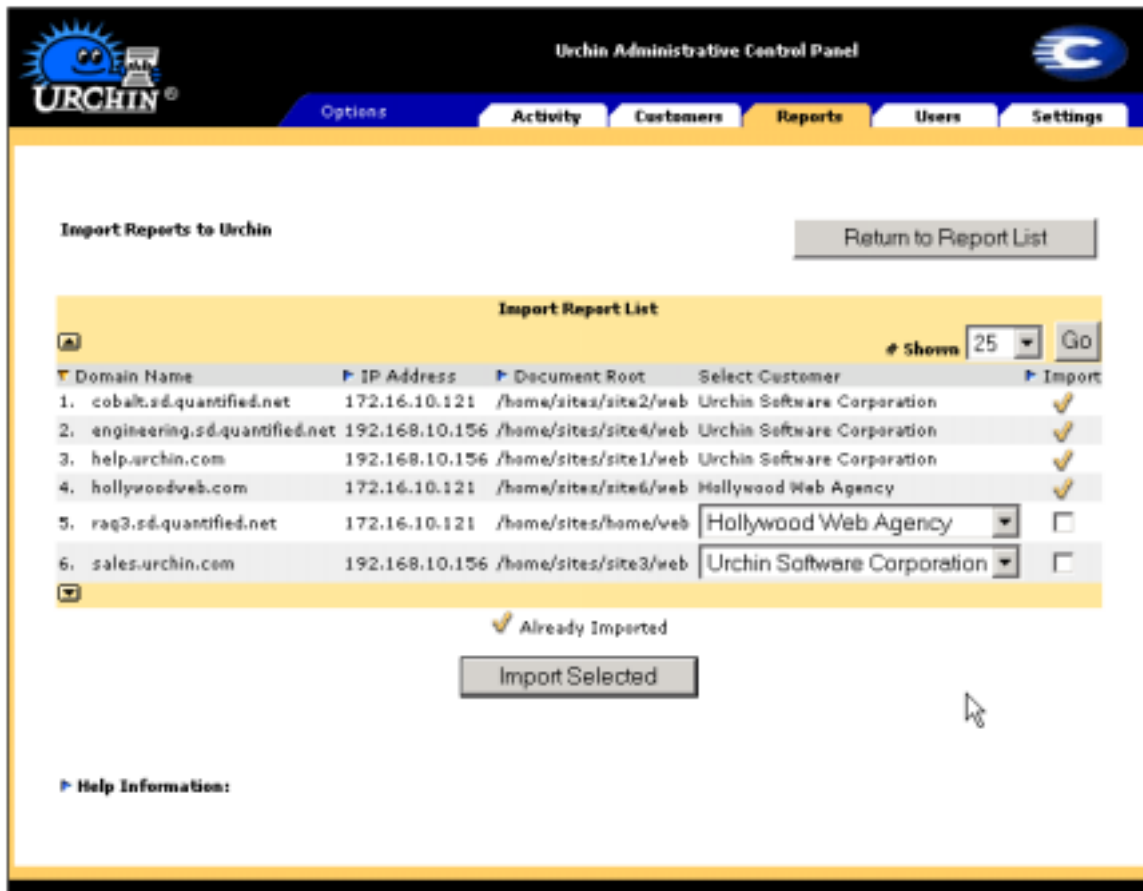


Figure 8

Use the "Manually Add Report" option to manually configure Urchin reports. If you have log files from other sources, this option will give you the flexibility to custom configure Urchin reports. Unlike the Import Reports function described in the last section, when you manually add reports Urchin does not automatically fill in most of the fields. So you will have to know the exact data to enter into the various boxes. Here are some guidelines for determining what to enter:

- **Report Name** - should always be the complete DNS name of the website referred to in the logfile
- **ReportDirectory** - a writeable, web-accessible area where Urchin can store the database files it builds when analyzing the log and the cgi program it uses for reporting
- **ReportURL** - the URL that refers to the web-accessible directory specified in ReportDirectory

• **Logs & Filters** - should at least contain a TransferLog directive with the full path to the logfile. A LogFormat directive indicating that the logfile is "combined" will automatically be entered.

You can view proper examples of these entries by clicking the **Reports** tab and selecting **Edit** next to any site that you have already configured for reporting. If you are automatically copying a new version of the logfile from a remote location every night, you must assure the logfile is in place at the location specified in the TransferLog directive before Urchin runs. Urchin typically runs at approximately 4:00AM.

3.4 Users

The Users tab allows you to add, edit, and delete individual users and set their report-viewing permissions. Each user should be associated with a customer. Individual users will have their own username and password. The admin user by default has access to all reports, and therefore does not appear in the user list.

To add a user to the list, click the "Add" button on the upper right. The list of users displays several fields, including the real name of the user, the login username, and the associated customer. The customer pull-down menu at the top can be used to filter the list of users (see Fig. 9).



Figure 9

For each user in the list, buttons are provided for changing permissions, editing attributes and deleting the user from the list.

To add a user, click on the Add button in the upper right. When adding a User, you will be asked to provide the username, password, full name and the Customer they belong to. The Edit button allows you to change these attributes.

After creating a user, click on the “Permissions” button to specify which reports that particular user has access to (see Fig. 10). The left box shows the reports to which the user does not have access, and the right box shows the reports to which the user does have access. Only reports assigned to the customer with whom the user is associated will appear.



Figure 10

To add access to a report, select the report names in the left box and click the >> button to move the report into the right box (Access Granted). To remove access click the << button to move the report into the left box, removing it from the list of accessible reports.

The report access configuration is immediately saved with each change. No "Update" button is needed. The users can access the reports through the following URL:

`http://www.yoursite.com:888/`

where "www.yoursite.com" is the IP address or domain name of the Cobalt machine. The ":888" specifies that the reporting server is running on port 888 which is the default. Users will be prompted for their username and password, and based on their username will be presented with a list of reports for which they have access. They can then select the report they want to view by clicking on it.

3.5 Settings

The Update Settings form (see Fig. 11) allows you to modify attributes of the Urchin configuration. The current values (if any) of the fields are displayed in the form. To change any of the attributes, modify the form and click the "Update Settings" button at the bottom.

Log Destiny specifies what to do with logs after they are processed. Archive will save the log file in a compressed format, and append a timestamp to the saved file. The DNS resolution section specifies which machine to use to resolve IP addresses.



The screenshot shows the 'Update Settings' form in the Urchin Administrative Control Panel. The form is titled 'Update Settings: UrchinBox' and contains several configuration options:

- Log Destiny:** Radio buttons for 'archive' and 'delete'. The 'delete' option is selected.
- Enable DNS:** Radio buttons for 'on' and 'off'. The 'off' option is selected.
- DNS Speed/Performance:** Radio buttons for 'slow', 'fast', and 'default'. The 'default' option is selected.
- DNS Resolver IP:** A text input field with a placeholder '(eg. 127.0.0.1)'. The field is currently empty.
- Current Serial/License:** The text 'Urchin Currently Not Licensed' is displayed.
- License Key Code:** The text 'ZAGP-FTSV-BG4R (Purchase License)' is displayed.
- Add Serial Number:** An empty text input field.
- Add License Code:** An empty text input field.

At the bottom of the form is a button labeled 'Update Settings'. Below the form is a link for 'Help Information:'.

Figure 11

To enable DNS resolution (recommended), click the "on" radio button next to the "Enable DNS" setting. Enter the IP address of the DNS server you wish to use as your

local resolution server. This can be the same server as listed in your Cobalt administration interface under Control Panel -> Network -> Primary DNS Server.

The SerialNumber and LicenseCode fields hold the license information. After evaluating the software, you can license the software by clicking on the words “Purchase New License”. This link will take you to our on-line licensing center and automatically enter your Key Code. Licenses are IP based and valid for 25 sites.

4. Filtering and Using Dynamic URLs

Urchin 3 has a couple of different filtering capabilities. It can filter out stuff from a particular log file that you don't want to get into a report (filter out). It can require that all log lines contain some pattern in order to make it into the reports (filter in). And, Urchin 3 can filter Dynamic URLs, such as those produced by CGI scripts, in order to determine a better *pagename* for reporting.

All of the filtering mechanisms covered in this chapter use *POSIX extended regular expressions*. Regular expressions allow you to use wildcards in order to perform pattern matches used in the filtering. If you have never used regular expressions before, you might need to consult with another administrator or with our technical support group if you have trouble setting it up. It's not really that complicated. Here is a list of some of the basic components of the expressions that will help you get started.

Regular Expressions	
wildcard	meaning
.	match any single character
*	match zero or more of the previous item
+	match one or more of the previous item
?	match zero or one of the previous item
()	remember contents of parenthesis as item
[]	match one item in this list
-	create a range in a list
	or
^	match to the beginning of the line
\$	match to the end of the line
\	escapes any of the above

The regular expression is compiled and applied according to each purpose outlined below. Multiple filters can be applied simultaneously, although each one will incur a

performance hit. For each of the types of filters below, examples are provided to help understand the basic usage.

4.1 Filter In and Filter Out

The FilterIn and FilterOut directives allow you to control which hits in a log file actually make it into a report. This can be used to filter unwanted hits or only include hits that match a certain format. It is perfectly acceptable to have different <Report> entries use the same log files. Urchin will recognize this fact and reuse the log files for each occurrence. This allows you to create multiple reports from the same log file. Using filters, these reports can be tailored to a specific need. Please note that for each report you will need a separate <Report> entry and a separate ReportDirectory. Also, each filter directive applies only to the next "TransferLog" entry. If you have multiple logs for one site, you will need to have a Filter command *before* each one.

The FilterIn directive requires that each hit matches the specified filter. If a hit does not match the requirement, then it is ignored. If you wanted to create an entire report based on a subdirectory within the website, you could use this directive to require that each matches that subdirectory name.

On the opposite side of the fence, the FilterOut directive can be used to block certain hits from getting into the report. Anything that matches the specified pattern will be ignored and not counted in the reporting.

Here is an example that will help illustrate the usage of the two filters. The main website in question is <http://www.serious.com> and is located in the directory `/www/serious/`. Within this directory we have the main corporate website, and a fun subsite located in a subdirectory called `smurfs`. Since both sites are logged together, we may want to use the filtering capabilities in order to create separate reports for the serious site and the smurfs site. To do this, you would make entries by editing the report in the log file box similar to:

```
ReportName:      serious.com
ReportDirectory: /www/serious/urchin/
FilterOut:       smurfs
Transferlog:     /www/serious/logs/access-log
```

```
ReportName:      serious.com/smurfs
ReportDirectory: /www/serious/smurfs/urchin/
FilterIn:        smurfs
Transferlog:     /www/serious/logs/access-log
```

Since each hit for the fun site will contain "smurfs" in the path, we can use the FilterOut directive to *exclude all entries that contain the name*; and use the FilterIn

directive to *only include entries that contain the name*. We used the FilterOut directive for the serious corporate site, and the FilterIn directive for the fun site.

4.2 Dynamic URL Filtering

Many sites today will use a CGI, ASP or other scripting mechanism to provide dynamic content. Often, a single script is used to deliver multiple pages of information. While this can be a handy way to track users sessions or provide “live” content, it poses an additional challenge for meaningful reporting.

Under normal operation the “Pages” reporting section reports on the actual object, page or script requested and discards any parameters that may appear in the request line as arguments to the script or page. The `DynamicURL` directive allows you to use regular expressions to capture this potentially valuable information. Let’s use an example to illustrate the concept.

In this example, a CGI script is used to deliver information about all products in a catalog. The script draws from database, and uses parameters passed through the request (known as the GET method), to determine which product to display. A particular hit or request may look like:

```
/cgi-bin/showProduct.cgi?sessionId=123456789&productId=knobs  
|_____||_____||
```

Under normal operation, Urchin reports on the first part of the request only. That is everything before the question mark in the above example. And while it is nice to know people are looking at products, it would be real nice if we knew which ones. By using a `DynamicURL` filter we can grab this valuable information embedded in the second part of the request.

Now in this example, we don’t necessarily want to capture the entire second part of the request because of the “sessionId.” Let’s assume that this parameter changes for each visit and we get 30,000 visits per day. Including this piece of information would render the reports useless as they would be much to granular. Instead we just want to capture the “productId” and report on that information.

```
/cgi-bin/showProduct.cgi?sessionId=123456789&productId=knobs
```

We may still want to know which script was used as well as which product was implicated in the request. By using a `DynamicURL` filter, we can capture multiple parts of the request and recombine them into a new, formatted request ready for reporting. Here is an example of a filter that could be used in this example:

```
DynamicURL: (/cgi-bin/showProduct.cgi\?).*productId=(.*)
```

This regular expression will match the above request no matter what the value of the sessionId or productId was. And the parenthesis capture the parts of the request that we want to keep for reporting. The effective request of the above example would look like:

```
/cgi-bin/showProduct.cgi/knobs
```

Up to 5 sets of parenthesis can be used. And, multiple filters can be applied. If a request does not match the DynamicURL filter, it is left unmodified, but still included in the reporting. This allows you to use multiple DynamicURL filters for each area of a site. Keep in mind there is a slight performance hit for each filter used. If you need assistance getting this to work correctly, please contact our technical support. Be sure to send them an example from a log and what pieces you want to capture.

5. The e-Urchin Module (e-Commerce Reporting)

ELF is an e-commerce log format that allows in-depth analysis of shopping behavior on e-commerce enabled websites. Combining the webserver's standard access log with ELF, online merchants gain unprecedented insight into the relationship between all website visitation parameters and actual money spent - Return on Investment Reporting is here!

ELF is produced by several popular e-commerce systems, including Miva Merchant 2.0. If your e-commerce package does not support ELF, please consult the [urchin.com](http://www.urchin.com) site for the latest supported formats, or request that your vendor include ELF support.

5.1 Architecture

Urchin allows multiple log files to be processed and merged before creating reports. Visitors from one log file are compared and correlated with the other log files. With the addition of e-commerce reporting, this feature can be used to correlate data from the normal webserver log with the ELF e-commerce log. In order to take advantage of this feature, there are two preparatory steps:

1. Create the ELF log file with your e-commerce system.
2. Make both the ELF log and the normal webserver access log available on the same system.

The first step is to create the ELF log file from your on-line transaction system (if you use Miva Merchant 2.0 you have already got the ability). If your e-commerce system is proprietary or uses simple cgi scripts, you should be able to have your site administrator add the ELF log functionality to the system. A complete description of the log file format follows below.

If your system is off the shelf or provided by an ISP or hosting company, please forward them to <http://www.urchin.com/> for more information. We have many resources available

and are currently working with many merchant software companies to provide seamless integration.

Once you have the ELF log file format created, you will need to have the file accessible along with the normal webserver access log. This can be trickier than it seems as often the main webserver can be a completely different machine than the secure machine running the e-commerce system. You will need to work with your system administrator to see which solution makes the most sense for your system. If the log files are on different machines, you may need to use a script, nfs, ftp or some other way to transfer the log from one machine to another.

5.2 Configuration

Once both log files are available to the Urchin program, you will need to edit the report entry using the control panel. Select the Reports tab, then click the Edit box next to the site for which you want to do e-commerce reporting. On the Update Report screen in the Logs& Filters box, you will see these lines:

```
LogFormat:          combined
TransferLog:        /usr/local/apache/logs/access-log
```

You should add the following e-commerce directives after these lines:

```
EcommerceFormat:   elf
EcommerceLog:       /usr/local/commerce/logs/elf-log
```

as shown in Fig15. Make sure that the LogFormat directive always comes before any actual log file entry.



Figure 12

5.3 ELF, in Detail

The E-Urchin Logfile Format (ELF) format is a tab-separated, multiline format. The transaction begins with the "!" exclamation mark character (which is thusly prohibited from the rest of the data). The first line contains the geographic and overall information on the transaction. Subsequent lines contain details on individual products. The basic format of the file will look something like:

```
!transfield1transfield2....
productfield1productfield2...
productfield1productfield2...
.
.
.
!transfieldtransfield2...
etc.
```

blank fields should contain a "-" dash character. The format of the transaction line is as follows:

```
\!% {ORDERID}%h% {STORE}% {SESSIONID}%t% {TOTAL}% {TAX}% {SHIPPING}% {BILL_CITY}% {BILL_STATE}% {BILL_ZIP}% {BILL_CNTRY}
```

where

% {ORDERID}	is the order number
%h	is the remote host (see apache.org)
% {STORE}	is the name/id of the storefront
% {SESSIONID}	is the unique session identifier of the customer
%t	is time in the common log format (see apache.org)
% {TOTAL}	is the transaction total including tax and shipping. (decimal only, no "\$" signs)
% {TAX}	is the amount of tax charged to the subtotal
% {SHIPPING}	is the amount of shipping charges
% {BILL_CITY}	is the billing city of the customer
% {BILL_STATE}	is the billing state of the customer
% {BILL_ZIP}	is the billing zip of the customer
% {BILL_CNTRY}	is the billing country of the customer

the product line format is:

```
% {ORDERID}% {PRODUCTCODE}% {PRODUCTNAME}% {VARIATION}% {PRICE}% {QUANTITY}% {UPSOLD}
```

where

% {ORDERID}	is the order number.
% {PRODUCTCODE}	is the identifier of the product.
% {PRODUCTNAME}	is the name of the product.
% {VARIATION}	is an optional variation of the product for colors, sizes, etc.
% {PRICE}	is the unit price of the product (decimal only, no "\$" signs).
% {QUANTITY}	is the quantity ordered of the product.
% {UPSOLD}	is a boolean (1 0) if the product was on sale.

Here are some example entries:

```
!12313 ppp-46.mia-tc-2.netrox.net StuffStore 1102323131
[27/Jul/1999:11:43:02 -0700] 198.12 8.12 10.00 Cedar Rapids Iowa 52403 US
12313 102 T Shirt XL 10.00 10 0
12313 103 Boxers L 9.00 10 0
!12314 213.12.54.123 - 110123413 [27/Jul/1999:11:43:02 -0700] 11.75
0.75 1.00 Santa Ana CA 92705 US
12314 102 T Shirt S 10.00 1 0
```

Currently Supported Shopping/e-Commerce Systems:

Miva Merchant 2.0

6. Troubleshooting

6.1 The Urchin Button

The Urchin for Cobalt interface includes a button to access Urchin reports, which is normally located in the left navigation bar under the standard Cobalt system buttons. Periodic improvements by Cobalt to the RaQ operating systems ("patches") may result in this button's disappearance upon installation of the patch(es). In the course of Urchin's standard operation it will normally replace this button automatically. In the unlikely event it is missing for more than 24 hours, you can restore it manually from the Urchin Administrative Control Panel Troubleshooting Page.

To access the Urchin Administrative Control Panel when the Urchin button is not available, do the following:

On RaQ3 and RaQ4 systems, go to the main Cobalt Administration Control Panel, select Maintenance, then select Install Software in the top line of links. This will give you a page that lists currently loaded software. Urchin will appear in the list as an active link you can click that will take you to the main Urchin Administration Control Panel. Select Troubleshooting Urchin and then click the Restore Urchin Menu Button icon.

You may also want to bookmark the URL to your Urchin Administrative Main Menu page for future reference. Netscape users should right click within the page and select Add Bookmark. Internet Explorer users should right click within the page and select Add to Favorites.

Then select Troubleshooting Urchin and click the Restore Urchin Menu Button icon.

6.2 Installation Failures

There are some cases where you will attempt to install the Urchin .pkg image from the Cobalt Install Software page and it will fail. The error messages are displayed in a popup dialog box at the bottom of the software installation screen.

Installation failed because .pkg file was corrupt – This error can occur when using either the URL or Upload method. Most likely the .pkg file got damaged while downloading it via your browser. So far the one case where this happens consistently is when using Netscape from a Windows machine to download and save your .pkg image. Netscape on a Unix system does not have the problem. If you are downloading to a Windows machine, then it is advisable to use Internet Explorer.

6.3 Report data is missing

You may notice that some (or all) of your reports have no data. There are several

possible reasons why your reports may not be updating:

Your server may be running a demo copy of Urchin and the license has expired

You will need to go to the Urchin Administrative Control Panel and do the following:

- Click on the "Settings" tab
- Copy your License Key Code
- Click on "Purchase New License"
- Enter your License Key Code into the appropriate blank on the Urchin Licensing Center page
- Complete the licensing payment process
- Enter the Serial Number and License Code into the appropriate blanks in the Settings area of the Urchin Administrative Control Panel.

Your Cobalt log rotation mechanism may need to be adjusted

Cobalt servers typically log activity for all web sites in a single log file, which is then split nightly into individual site log files by a program called logrotate. Under certain circumstances logrotate will perform log rotation activities that cause data to be missed by Urchin. To find out if the reason you have missing report data is because Urchin is unable to find or process logs, you should examine Urchin's own runtime log output. You can do this by going to the Urchin Administration Control Panel and selecting the Activity Tab. Fill in the Year/Month/Day boxes with a date that you are missing a report for, then click Go. The Urchin Reports Process Summary will change to reflect the day you are interested in. In the Status column for each site will be a status message, which when clicked will give you a new window which will have the Urchin runtime output for that day. You want to examine this output for Urchin errors. Errors concerning missing log files will appear like so-

```
+R Processing report "somesite.domain.net" starting: 20010119 09:40:53
```

```
+E WARNING(5001): Logfile does not exist.  
/home/sites/site12/logs/web.log  
Continuing...
```

This shows the log file was moved before Urchin was able to process it. To alleviate this problem, increase the disk quota for the site. The logrotation size parameter is calculated based on the site's disk quota, so you can cause it to be adjusted indirectly by changing the disk quota for a site that exhibits this problem.

Note: On RaQ3 and RaQ4 systems, a patch from Cobalt is required to enable your Cobalt Administration GUI to make this adjustment. Without the patch, adjusting the disk quota will have no effect on the logrotate size setting. The appropriate patches are:

- RaQ3: RaQ3-All-System-4.0.1-9447.pkg
- RaQ4: RaQ4-All-System-1.0.1-9447.pkg

These patches are available on Cobalt's website at <http://www.cobalt.com/support/download/>.

To change the disk quota:

1. In the Cobalt Administration interface, select **Site Management**. You are presented with a list of all the virtual sites on your system.
2. Click the wrench symbol to the right of a site that does not have report data. The Site Management screen for the site appears.
3. Click **Site Settings**.
4. Increase the **Max. Allowed Disk Space (MB)** setting. The logrotation size parameter is calculated as 10% of this disk quota. Your current logrotation setting will be 10% of whatever is in this field. Determining how much to adjust the number is case dependent and should be based on the amount of traffic your site gets. A good place to start is to increase your logrotation setting by 50%. Increase the number in this field by 50% to accomplish this.
5. Click the **Save Changes** button when you are finished.

Your Cobalt log rotation mechanism has halted completely

Sometimes processing and rotation of the logs on your Cobalt system will halt completely. This is not the result of the Urchin installation or Urchin processing. This happens when the internal site usage statistics extracted by the Cobalt system fail in some way, which causes the rest of the logrotate mechanism to become inactive. Since Urchin has to wait until logrotate finishes its job, there is no data for Urchin to process in this case. Logrotate failures of this type are best fixed by making sure you have all the latest Cobalt patches installed on your system. In particular, the OS update patches that are pertinent to your Cobalt platform can alleviate this problem.

6.4 Additional Technical Help

If you have any questions that are not addressed in this manual then check our Cobalt specific technical web page http://www.urchin.com/support/v3/faq/faq_cobalt.html for tech notes and FAQs. To request technical support go to the technical support request web page at http://www.urchin.com/support/techsupport_request.html