

Skulker2 Version 1.2.2 Administrator Guide

Document Version: 2.3
Updated: 6th August, 2012.

Version History

Version	Date	Author	Changes
1.0	12 th December, 2006.	Simon Edwards	Original.
1.1	8 th June, 2007.	Simon Edwards	Improvement documentation.
1.2	20 th December, 2007.	Simon Edwards	Support and use HTML reporting.
1.3	10 th January, 2009.	Simon Edwards	Use of --var setting.
2.0	9 th July, 2010.	Simon Edwards	Updated to cover 1.0.0 release features.
2.1	March, 2012.	Simon Edwards	Review and include all recent minor changes since last document release.
2.2	April, 2012.	Simon Edwards	Review and update complete document.
2.3	August, 2012.	Simon Edwards	Added listing and checksum functionality documentation.

Contents

1	INTRODUCTION.....	3
1.1	WHAT IS SKULKER?	3
1.2	PURPOSE OF ADMINISTRATION GUIDE.....	3
2	BASIC COMMAND EXECUTION	5
2.1	ENVIRONMENT CONFIGURATION	5
2.2	DISPLAY USAGE OPTIONS	6
2.3	OVERVIEW OF COMMAND ARGUMENT SYNTAX	8
3	SUMMARY OF SUPPORTED ACTIONS.....	10
3.1	AVAILABLE ACTIONS	10
3.2	ORDER OF ACTIONS PERFORMED	11
4	SUPPORTED COMMAND LINE ARGUMENTS	13
4.1	MANDATORY ARGUMENTS	13
4.2	OPTIONAL ARGUMENTS	14
4.3	COMMONLY USED OPTIONS AND EXAMPLES	17
4.3.1	<i>Generating a HTML Report</i>	<i>18</i>
4.3.2	<i>Listing Files that will be processed.....</i>	<i>18</i>
5	UNDERSTANDING SKULKER OUTPUT	19
5.1	THE RULE PARSING OUTPUT	19
5.2	THE LIST RULES OUTPUT	20
5.3	THE SKULK ACTION OUTPUT	21
5.3.1	<i>Command output</i>	<i>21</i>
5.3.2	<i>Rotate output</i>	<i>21</i>
5.3.3	<i>Compress output.....</i>	<i>21</i>
5.3.4	<i>Chksumadd output</i>	<i>21</i>
5.3.5	<i>Chksum output</i>	<i>22</i>
5.3.6	<i>Setperms output.....</i>	<i>23</i>
5.3.7	<i>Delete output</i>	<i>23</i>
5.3.8	<i>List output.....</i>	<i>23</i>
5.3.9	<i>Tail output</i>	<i>23</i>
5.3.10	<i>Organise output.....</i>	<i>23</i>
5.3.11	<i>Archive output.....</i>	<i>24</i>
5.3.12	<i>Append output.....</i>	<i>24</i>
5.3.13	<i>Move output.....</i>	<i>24</i>
5.3.14	<i>Chmod output.....</i>	<i>24</i>
5.4	THE REPORT ACTION OUTPUT	24
5.5	THE SUMMARY ACTION LOG.....	25
5.6	THE TOTALS SUMMARY ACTION LOG.....	26
5.7	THE "HTML REPORT" OPTION OUTPUT	27
5.7.1	<i>Rule Details</i>	<i>27</i>
5.7.2	<i>Runs over 31 Days</i>	<i>27</i>
	<i>Summary</i>	<i>28</i>

1 Introduction

1.1 *What is Skulker?*

“Skulker” is a tool that has been designed to automatically manage many files that either grow in size or number over a period of time whilst the Operating System runs and applies a series of actions against these files to ensure the disk space they occupy is automatically managed.

Skulker works by applying a series of rules that define the files to apply a particular rule again, the frequency that a rule can be run and optional logic that must be matched for a particular file is considered a candidate for application by the rule.

Skulker is entirely written in Perl and has been written to work across any number of UNIX variants, including HP-UX, AIX, Solaris and Linux. It supports multiple levels of logging, accurately evaluates (and reports on) space saved, and offers a simple syntax based on XML files. The very latest versions of Skulker also supports Windows natively too (i.e. no requirement for a UNIX emulation layer such as Cygwin).

Skulker also provides a “preview” mode allowing the administrator peace of mind regarding what files will be impacted by the default rules. The environment is extensible - the administrator can change the existing rules or add further rules if they wish.

Commercial environments expect information on results, and so Skulker supports basic report generation on performance (amount of storage saved; time taken to run; files processed) and can also generate per-rule emails to users if necessary.

Skulker is designed to help you not hinder; and hence it provides the concept of “safe areas” which it will not manipulate files; such as “/usr” on UNIX platforms for example.

1.2 *Purpose of Administration Guide*

This document aims to provide a summary of the command line options that Skulker supports and typical methods of running the utility. It also describes possible automation frequencies and how to interpret the output generated by the various options.

Since Skulker provides statistics reporting it describes not only how to view this information, (and reset it), but also common things to look for which might indicate problems.

This document does not describe the process of installing Skulker - that information can be found in the document “Skulker Version 2 - Installation Guide”. Further this document does not describe how the software environment can be configured - that can be found in the document “Skulker Version 2 - Configuration Guide”.

The power of this utility comes from the handling of rules. Although the default rules ensure that the standard log files for a support UNIX variant are managed additional rules to manage application logs/files for a particular host are recommended. For

information on rule configuration the document "Skulker Version 2 - Rule Configuration Guide" should be consulted.

2 Basic Command Execution

2.1 Environment Configuration

Typically Skulker is invoked using the complete command path, for example:

UNIX:

```
# /opt/skulker2/bin/skulkerv2
```

Windows:

```
> c:  
> cd "c:\program files\skulkerv2"  
> perl skulkerv2
```

Unless a suitable environment is configured the output generated will likely be similar to the following;

```
2010/07/09 23:36:28 LOG      Successfully loaded internal module logging  
2010/07/09 23:36:28 LOG      Successfully loaded internal module unix_magic  
2010/07/09 23:36:28 LOG      Successfully loaded internal module skutils  
2010/07/09 23:36:28 LOG      Successfully loaded internal module skulkercfg  
2010/07/09 23:36:28 LOG      Successfully loaded internal module match_parse  
2010/07/09 23:36:28 LOG      Successfully loaded internal module skulkerrules  
2010/07/09 23:36:28 LOG      Successfully loaded internal module skcompression  
2010/07/09 23:36:28 LOG      Successfully loaded internal module stats  
2010/07/09 23:36:28 LOG      Successfully loaded internal module stats_old  
2010/07/09 23:36:28 LOG      Successfully loaded internal module histstats  
2010/07/09 23:36:28 LOG      Successfully loaded internal module histstats_old  
2010/07/09 23:36:28 LOG      Successfully loaded internal module histgraphs  
2010/07/09 23:36:28 ERROR    No --cfg argument and SKULKERV2_CFG not defined.  
2010/07/09 23:36:28 ERROR    Please set either one to a valid Skulker v2  
2010/07/09 23:36:28 ERROR    configuration file.
```

Hence for Skulker to function it must understand the configuration file that is to be used. There are two ways of defining this - via the command line and via the SKULKERV2_CFG environment variable. To use the command line option use;

```
# /opt/skulker2/bin/skulkerv2 \  
--cfg /opt/skulker2/cfg/cfg.xml
```

If you intend to use Skulker on the command line rather than just via schedules it is recommended that the environment variable is set automatically during login. For example the following instead could have been added to a ".profile" file (or equivalent):

UNIX:

```
SKULKERV2_CFG=/opt/skulker2/cfg/cfg.xml  
export SKULKERV2_CFG
```

On windows it should be added to the user's environment variables via the "System Properties -> advanced" button on the "system" or "my computer" settings depending on the version of Windows in question.

- For the rest of the example commands it is assumed that the "SKULKERV2_CFG" environment variable has been set as shown to reduce the length of the command lines.
- For Windows the rest of the commands assume the commands are run from the "bin" directory of the installation directory – for example "c:\program files\skulkerv2\bin".

2.2 Display Usage Options

Once Skulker understands where to find the configuration file it should no longer complain. However unless any meaningful arguments are given on the command line it will not actually do anything!

Due to the number of command line arguments a brief argument summary screen is available. To access this screen, use either of the following commands;

UNIX:

```
# /opt/skulker2/bin/skulkerv2 -h
```

or;

```
# /opt/skulker2/bin/skulkerv2 -?
```

Windows:

```
> perl skulkerv2 -h
```

Due to the way Skulker works prior to the actual help message some log output may be shown (this can be ignored). As of the current version, the help screen looks similar to the following;

```
Usage: skulkerv2 [--cfg=F] [--logfile=LF] [--loglevel=N] [rulesfile=F]
      [--rulesdir=D] [--statusfile=F] [--actions=A[,...]]
      [--ignoreintervals] [--testfiles=F] [--historyfile=F]
      [--historykeep] [--historydays D] [--reporttitle T]
      [--reportname X] [--top N] [--reportdir X]
      [--rule[s] file:num[-num][, ...] [--var name=value ...]
      [--days | --start YYYYMMDD --end YYYYMMDD]
```

skulkerv2(1m) is an extensible and flexible temporary and log file management tool designed for UNIX platforms. It is typically used on a nightly basis to run a series of rules to manage storage. The rules specify minimum intervals between processing to ensure certain actions do not occur too often.

The tool also supports statistic gathering and reporting for the last and all runs of the current rule set, which in practise is especially useful.

All command line arguments are optional - though in most cases at one or two tend to be given to influence how the tool runs.

```
--cfg          The path to the configuration file if the default in the
                environment variable SKULKERV2_CFG is not set or the user
                wishes to override it.
--var          Creates a variable and sets the value as specified. This can
                then be used in rules to add further flexibility. This
                can be used any number of times to set variables.
--logfile      The name of the file to log information and errors to.
                If it is defined here it overrides the value in the
                configuration file.
--loglevel     The amount of logging to produce, ranging from 0 ... 5.
                0 is no logging apart from errors and warnings to 4 which
                is maximum verbosity. A log level of 5 produces even more
                information, but makes no actual changes - like a 'dry run'
                function.
--rules        Specifies a specific series of rules that should be considered
                to run. The rule file [without the .xml extension must be
                followed by a ':' and then a number or a range of rules to
                run. For example 'rules1:10-100'.
--rulesfile    The name of the rules file to process if the user wishes to
                override the value in the configuration file.
                No directory component is necessary.
--rulesdir     The directory where the rules files can be found. If is
                given separately since it is used by rules files to include
                other rules files.
--functionsdir The directory to use to look for functions to load,
                overriding config file value.
--functionslist A comma-separated list of functions that should be loaded -
                others in the directory will be ignored.
--generatorsdir The directory to use to look for generators to load,
                overriding config file value.
--generatorslist A comma-separated list of generators that should be loaded -
                others in the directory will be ignored.
--statusfile    The path to the status file to use. The file generated or
                maintained typically two files depending on the platform.
--[no]threading Turns on threading off [nothreading] or on [threading] -
                overriding the settings in the configuration file.
--actions       The list of actions to perform. This is given as a
                comma-separated list. Supported actions include:
                * parse      Parse the rules files, (pre-requisite for most
                             activities).
                * generate   Generate additional statistic entries for missing
```

```

rules. If this is not present only rules with
existing statistics entries will be handled.
* list      Simply list the files that will be impacted by
the rules to do - do not perform any changes.
* reset    Reset space statistics.
* skulk    Actually perform the work in the rules being
evaluated.
* test     Perform the system self test - usually given as
the only action if used.
* report   Report the results of the last run of the rules
specified.
* summary  Produce a summary report of the last run of the
rules specified.
* tsummary Produce a summary report of the totals of all
runs of the rules specified.
* htmlreport Generates a html report [with graphs for the
rules which save the most disk space].
--ignoreintervals Usually a rule can only be run only after a rule specified
period of time has passed. This ignores the intervals for
all rules defined to be run.
--testfiles  When running in 'test' mode it allows just specified tests
to be performed. Typically only used when a new version
is installed. The test files can be found in the 'tests'
sub-directory.
--historykeep A flag to turn on history recording if it is not enabled
in the configuration file.
--historyfile The file where to keep the history - defaults are in the
main configuration file.
--historydays The number of days of history to keep - older items will
be removed.
--reportname  The prefix to use for the reportname in the given directory.
--reportdir   The directory to create the report and associated graphics.
--reporttitle The title to put down on the report.
--top         Maximum number of rules to report on defaulting to 10 if
not specified.
--days       Number of days to check over - will default to 30 if
not specified.
--start       Start date to show information for the 'tsummary' or
'html_report' actions. YYYYMMDD format.
--end         End date to show information for the 'tsummary' or
'html_report' actions. YYYYMMDD format.

```

Written by Simon Edwards, (C) 2005-2012, advantsys.co.uk

2.3 Overview of Command Argument Syntax

Skulker uses the GNU style of argument handling - that is all arguments use "long" names (rather than just single characters) to indicate the option. Each of these long options is prefixed with two dashes, for example;

```
--ignoreintervals
```

If multiple arguments are required they should be white space separated. Since some of the argument names are long another feature that is example is "shortest unique" prefix support. In this case the complete argument is not required just enough characters to uniquely identify the argument in question. For example all of the below are equivalent.

```
-i
--i
```

```
--ig  
--ignore  
--ignoreintervals
```

Note - if shortening the name to a single character (if unique), then only a single dash is actually necessary, though two are still supported.

Some arguments expect a value (such as "--loglevel") and this can be specified in either of the following ways;

```
--loglevel 3  
--loglevel=3
```

Of course "--loglevel" could be shortened to "--logl" if necessary (since there is also a "--logfile" option).

3 Summary of Supported Actions

The Skulker command can work in various ways, and the actual work it performs is determined by the "--action" argument. Although the actions are described in the help text they are shown here in more detail.

3.1 Available Actions

There are a significant number of options, and if more than one is required (which is common), then they should be comma-separated.

Action	Purpose
parse	<p>Simply parse the rules specified and exit with a true status (return code 0) if all rules are valid or false (return code > 0) if any of the rules contain a syntax error.</p> <p>When running in this manner it is recommended that a verbosity level of at least three is chosen to provide context information for any failures.</p>
generate	<p>Skulker must keep information about each rule that is run, and it does this using a "statistics" file. If this file does not contain anything for a particular rule that is scheduled to run Skulker will abort.</p> <p>When this option is specified it will automatically add entries to the statistics file if deemed necessary.</p>
list	<p>List the files that would be processed by each of rules that have been defined. This "list" option does not omit rules that are not scheduled to run as yet, but for each rule summary information does include the frequency, time of last run, and how much space last running the rule saved.</p> <p>The list of files shown adheres to any pattern specified and any attribute logic that has been defined for the rule.</p>
reset	<p>As indicated Skulker keeps statistics about each of the rules that it manages. The statistics cover the time it was last run and the totals of all runs of the rule.</p> <p>The "reset" actions "zeros" the totals for the "all runs" details. Typically unlikely to be used, but handy following significant application or file system changes.</p>
skulk	<p>Actually perform the work specified in the rules file. Each rule has a frequency which means that not all rules configured might run depending on the current time/date and when the rules were last processed.</p>

Action	Purpose
test	<p>Run Skulker in self-test mode. It is strongly recommended that this be used at least once prior to automating Skulker to run on a regular basis.</p> <p>The self-test is not exhaustive; however it does give a good degree of confidence that Skulker is working as expected on the current platform.</p>
report	<p>Produce a detailed report of the last run of the rules specified. The output is shown to the log file - not the screen (unless that is configured as the log file!).</p> <p>The information is quite detailed and is meant to be human readable, though should be straightforward to parse by other tools if necessary.</p>
summary	<p>Produce a summary report for the rules specified. As for the "report" option please note that the report is shown to the log device - so it will only appear on the screen if the log device is set appropriately.</p> <p>This action displays a single line for each rule, showing the number of files processed the last time the rule was run, and also the amount of space saved for that particular run.</p>
tsummary	<p>This provides a report in the same format as "summary" - but shows the details of the rules against all rule runs. As stated above this information can be zeroed using the "reset" action if desired.</p>
htmlreport	<p>This will generate a HTML-based report including graphics of the amount of disk space saved. This functionality is optional and depends on additional Perl modules not distributed with Skulker - see the installation guide for details.</p> <p>This can also be specified as "htmlreport+tsummary" to add a table at the end of the report showing the details of the rule usage in that period.</p>

3.2 Order of Actions Performed

When several actions are specified the order they are given in on the command line is not important - Skulker will always perform them in the order it expects. Thus for example, "summary" is always performed after a "skulk", if "skulk" is provided as an action.

It should also be noted that "parse" is always performed - since none of the other actions can actually occur until that step has been completed.

The reporting options always occur at the end of the run - to take account of any changes that might just have run.

4 Supported Command Line Arguments

Skulker supports many command line arguments, though many of them are only necessary in certain conditions. It should be noted that Skulker has been designed to run as an automated process rather than interactively and this means that running it on the command line might result in a significant amount of typing to get it to work in the expected manner!

4.1 Mandatory Arguments

Skulker has no mandatory arguments (if the optionally `SKULKERV2_CFG` variable has been set)! However unless the following are included it will only parse the configuration and do nothing else.

Argument	Purpose
<code>--cfg</code>	Actually this is only mandatory if the environment variable <code>SKULKERV2_CFG</code> is not defined. It is used to determine the location of the configuration file. It can be an absolute pathname or relative to the current directory, for example; <code>--cfg=../cfg/cfg.xml</code>
<code>--action</code>	Must be specified if you want Skulker to do anything apart from parse the specified configuration. An example might be; <code>--action=generate,list</code> If more than a single action is given each must be separated by a comma - not white space!

4.2 Optional Arguments

There are many optional arguments.

Argument	Purpose
<code>--logfile</code>	<p>The location where the logs should be written to. If not specified the information from the configuration file is used (which is typically set to a file in <code>/var/adm</code> for example. Another possible value is <code>-</code> which indicates the log should be written to standard output. An example might be:</p> <pre>--logfile="/var/adm/skulker-%D.log"</pre> <p>In the above case <code>%D</code> will be replaced by the current date in format <code>YYYYMMDD</code>. The details of the available <code>%</code> macros can be found in the "Skulker Rules" document.</p>
<code>--loglevel</code>	<p>The amount of logging to perform. If not specified the setting in the configuration file will be used. However the default logging level, <code>5</code>, does not actually perform the work - just shows the work it would perform. Typically this is set to <code>3</code> which gives a reasonable amount of output. Setting it to <code>0</code> will produce no output - apart from any warnings or errors. An example usage might be;</p> <pre>--loglevel=3</pre>
<code>--rulesfile</code>	<p>The file containing the rules to process. If not specified it will take the details from the configuration file specified. This is the file name in the <code>rules directory</code> - and is not expected to contain a path element - see <code>--rulesdir</code> entry for that setting.</p> <p>Hence an example of this setting might be;</p> <pre>--rulesfile=test.xml</pre> <p>[The extension must be specified if present, though the file does not need to be suffixed with <code>.xml</code>]</p> <p>The default and recommended setting for performing all rules is <code>MAIN.xml</code>.</p>
<code>--rulesdir</code>	<p>Overrides the directory where rules files can be found compared to the setting in the configuration file. This option is useful when the ruleset in question references other rule sets (via the <code>include</code> statement).</p> <p>However in most invocations this does not need to be specified.</p>
<code>--rules</code>	<p>A list of actual rules that should be processed. If multiple rules are required then they should be comma-separated. For example:</p> <pre>--rules hpux-tmplogs:1020,1030</pre>

Argument	Purpose
--statusfile	<p>The path to the status file (which records details of when each rule was last run and the current and historical statistics). As usual this will override the similar setting in the main configuration file.</p> <p>The value specified is relative to the current working directory, and so relative names can be used if desired, for example;</p> <pre>--statusfile="../var/status.info"</pre> <p>The file name in question will be created if it does not exist. It will actually create two files - with ".dir" and ".pag" extensions. This extension is omitted when specifying the name on this argument.</p>
--ignoreintervals	<p>Usually when a series of rules are to be performed the time the rule was last run will be checked. If less time has elapsed than as specified in the rule frequency or the date/time settings are not met, then the rule is silently ignored.</p> <p>This option will always run the specified rules ignoring the time they were last run.</p>
--[no]threading	Turns off or on threading – overriding the setting in the used configuration file.
--testfiles	<p>When running in test mode this will indicate which test files should be run. The values specified should be comma separated (if more than one) and be the complete file names with any path component. An example might be;</p> <pre>--testfiles delete.0001.test,tail.0001.test</pre> <p>It is also possible to specify patterns, for example to perform all "delete" tests:</p> <pre>--testfiles delete.*</pre>
--historykeep	This will force history to be kept and will override a "keep=0" setting in the main configuration file.
--historyfile	The name of the file to store history information in. Again this will override the settings in the main configuration file.
--historydays	The maximum age of any history statistics to keep - as indicated this is a number of days. Older entries will be deleted from the history file.
--reportname	The name of the report to generate - this should not include a directory component - see the next argument for that. This argument is mandatory if the series of actions to run include the "htmlreport" option. If this does not include a ".html" extension one will be added automatically. The action will be aborted if a file with this name already exists.

Argument	Purpose
--reportdir	The directory to generate the files in for the "htmlreport" action. Please note several files will be generated here - one for the main html report and one for each graph that is added to the report. All graphs are prefixed with the name of the main report file.
--reporttitle	The title to put at the top of the report and in the title of the window when displaying in a web browser. This is optional - a suitable default will be used if not present.
--top	The number of rules to show details for in the "htmlreport" action. If not specified it defaults to 10.
--days	The number of days over which to gather the history to work out which rules have saved the most space - if not specified it defaults to 30.
--var name=value	<p>This can be used to set a variable which can be used in file handling rules ... either setting new values, or overriding the values of variables that are predefined in the used configuration file.</p> <p>This can be used any number of times to set as many variables as necessary. A variable should begin with a letter and consist of letters and numbers.</p>
--start YYYYMMDD	<p>Used for the "tsummary" action to indicate the start date for the history statistics to use for the table. If specified must also include the "--end" option.</p> <p>If the "--start" and "--end" option are not present then the "tsummary" action is performed against all runs known.</p>
--end YYYYMMDD	The end date for the "tsummary" table to be generated from.

Note: The separate use of directory and file options for rules is necessary it helps to determine the location for rules files that are referenced via the support "include" statement.

4.3 Commonly Used Options and Examples

The most common invocation of Skulker is probably the following;

UNIX:

```
/opt/skulker2/bin/skulker2 \  
  --cfg /opt/skulker2/cfg/cfg.xml \  
  --loglevel=3 --actions=generate,skulk,summary \  
  --rulesfile=MAIN.xml
```

Windows (assuming SKULKERV2_CFG set):

```
perl skulkerv2 --loglevel=3 --actions=generate,skulk,summary
```

Typically such a line is scheduled automatically to occur at least once a day; commonly during the early hours of the morning or any other time where the server in question is expected to be least busy.

Skulker can be run at any time and some sites run it every six hours - which rules get run obviously is determined by the frequency given for the rules compared to the time when each particular rule was run.

In the above the logging level is set since the default of "5" would not actually make any changes - essentially it would work as a "preview mode"!

The configuration file is referenced as a full path; as is the name of the program itself, since it is unlikely it will exist in the standard PATH setting for a scheduling tool.

By specifying the actions "generate,skulk,summary" the invocation will automatically add new rules to the statistics database if necessary, actually run the rules and generate a summary report of what changes have been performed.

4.3.1 Generating a HTML Report

A HTML report based on the contents of any history contains summary details both in terms of text and simple graphs. The generation of the graphs requires the Perl module "GD::Graph". If this module is not present then an error will be shown indicating the fact and the below commands will not function.

A typical command line to generate a html report might be:

UNIX:

```
/opt/skulker2/bin/skulkerv2 \  
  --cfg /opt/skulker2/cfg/cfg.xml --loglevel=3 --actions=htmlreport \  
  --rulesfile=MAIN.xml \  
  --reportdir=/tmp --reportname=skulker_report --reporttitle "Test Report"
```

Windows:

```
perl skulkerv2 --cfg ../cfg/cfg-windows.xml --action=htmlreport --reportdir=Z:\tmp --  
reportname=skulker_report --reporttitle "report title"
```

4.3.2 Listing Files that will be processed

The list option is very useful to check the impact of a change to a particular rule:

```
/opt/skulker2/bin/skulkerv2 \  
  --cfg /opt/skulker2/cfg/cfg.xml --loglevel=3 --actions=list \  
  --rulesfile=MAIN.xml \  
  --reportdir=/tmp --reportname=skulker_report --reporttitle "Test Report"
```

5 Understanding Skulker Output

All samples should here are generated when "--loglevel" set to 3 - which is the recommended setting since it provides plenty of feedback without excessive levels of "noise". In most case the output should be self-explanatory of course.

5.1 The Rule Parsing Output

The rules parsing can generate significant amounts of output since if it needs to default a missing setting it will say so. It will also indicate any changes that it has made to any patterns since it assumes a shell-based pattern has been specified and changes are necessary to convert this to a Perl equivalent.

Some sample output lines might be:

```
2010/07/12 23:08:12 WARN Added $ to end of pattern "/tmp/skulkerv2-tests-5179/files/xx2".
2010/07/12 23:08:12 WARN Added ^ to start of file pattern "/tmp/skulkerv2-tests-5179/files/xx2$".
```

The two lines are the most common type of warnings the administrator is likely to see. It will add "^" to the start of patterns and "\$" to the end - this will ensure that only the file named "xx2" in the specified directory is matched. Without it "sxx2" and "bxx2" and "xx2s" would also be matched if the existed in the directory, for example.

```
2006/05/08 02:20:01 WARN Added . to prefix '*' at start of file pattern "/tmp/.../*.Z$".
```

A comma shell pattern to match is "*.something" to match any file with a ".something" extension. Unfortunately this is not a valid Perl pattern - and this warning has indicating that it has added a "." prefix to approximate the pattern you might expect to use from the shell.

```
2006/05/08 02:20:01 WARN Rotate "format" defaulted to "%d/%f.%N"
2006/05/08 02:20:01 WARN Rotate "keep" defaulted to "0"
```

Lines similar the above are shown whenever a rule requires certain parameters that have not been provided. Where possible defaults from the current configuration file are used rather than Skulker failing to run.

5.2 The List Rules Output

When the "list" action is used then following the "parse" log messages each of the rules specified will be shown along with the files that match the rule in question.

In this instance (since nothing is actually being changed) the interval setting is ignored - all rule details are shown. The output generated for each rule is formatted as follows;

```
2010/07/12 23:08:12 LOG      Rule t:10 Rule action   : append
2010/07/12 23:08:12 LOG                        Match type    : file
2010/07/12 23:08:12 LOG                        Match pattern : /tmp/skulkerv2-tests-5179/files/^xx$
2010/07/12 23:08:12 LOG                        Match by     : ALWAYS
2010/07/12 23:08:12 LOG                        Run interval  : 0D
2010/07/12 23:08:12 LOG                        Last run     : Never
2010/07/12 23:08:12 LOG                        Last run info: Files: 0, Duration: 00:00:00, Space: 0 Kb
2007/03/10 00:06:42 LOG      /var/log/cups/error_log
2007/03/10 00:06:42 LOG      Matched 1 file in 1 directory.
```

The "rule action" line shows the type of action to be performed ("delete" in the above example), and any arguments that are to be used ("handleopen" in this case). The "Match type" is typically "file" - i.e. only regular files are processed by the rule other file system objects are ignored.

The "Match pattern" is the pattern that is used to generate matches. This is the pattern after the parser has processed the output and so it might be slightly different than that which was entered as part of the rule definition.

The "Match by" string is an expression that Skulker uses to check the attributes of each file that matches the pattern. The Skulker document "Skulker version 2 - Rule Configuration Guide" describes the available syntax for this option.

The last run is the date (format is "YYYY/MM/DD") and time when the rule was actually last processed via the "skulk" action. The rule interval is the minimum time that must elapse between two "skulk" actions against this rule being performed.

Finally the last line of the header is a terse summary of the characteristics of the last run of the rule, for example;

```
[Files: 4, Duration: 00:00:00, Space: 422.2 Kb]
```

The above indicates that on the last run the rule processed 4 files in less than a second (time is HH:MM:SS format) and is reclaimed 422 Kb.

The output generated for each rule also includes the filenames of all files matched - and so can be very lengthy if a rule matches many files.

5.3 The Skulk Action Output

The "Skulk" output shows summary details of each rule that is processed (ignoring those for which the elapsed time between invocations has not been met). The summary output is the same as the "list" action shown above.

After that, assuming that loglevel 3 is in use each type of rule will produce some output regarding any file that is processed. Examples of the more common rule types are shown below.

5.3.1 Command output

```
2010/07/12 23:29:11 LOG      Running command : /usr/bin/compress /tmp/skulkerv2-tests-5900/files/xx2
2010/07/12 23:29:11 LOG      Running command : mv /tmp/skulkerv2-tests-5900/files/xx2.Z /tmp/skulkerv2-tests-5900/files/simon.Z
2010/07/12 23:29:11 LOG      Processed 1 file in 1 directory.
```

For each command the command line is shown along with the return code of each command.

5.3.2 Rotate output

```
2006/05/09 02:20:06 LOG      Rotated [copy/truncated] file /opt/hpws/apache/logs/error_log successfully.
2006/05/09 02:20:06 LOG      Rotated [copy/truncated] file /opt/hpws/apache/logs/access_log successfully.
```

For rotate the names of the files that have been rotated are given. Optionally the lines may contain the text ["copy/truncated"] which is a method Skulker uses when the rule indicates that it should work with the possibility of the files being open for writing.

5.3.3 Compress output

```
2006/05/09 02:20:06 LOG      Compress: /tmp/sh/.mikemm.1092 -> /tmp/sh/.mikemm.1092.Z [ 56.02%].
```

When compressing files it shows details for each file - the original name and the newly compressed name. The figure in brackets the relative size of the file compared to the original - so 56% means the compressed file is just over half the size of the uncompressed one.

5.3.4 Chksumadd output

```
2012/08/06 21:50:38 LOG      Added file /tmp/simon1/auth2.log checksum 3921b38a087ca958866e8d3131fa98de12518c7d to DB /opt/skulker2/var/chksum.db.
2012/08/06 21:49:55 WARN      Record for /tmp/simon1/boot.log already exists.
```

The above output shows two typical lines that the "chksumadd" might generate. The first indicates that a new file (with the specified checksum) is being added to the specified database.

The second line indicates that the specified file is already in the database - but with a different checksum. The fact that the warning is shown indicates that the checksum in the database has been configured not to be updated by the rule.

5.3.5 Chksum output

```
2012/08/06 21:50:38 LOG Checksum failure for /tmp/simon1/dpkg.log: Should be  
da39a3ee5e6b4b0d3255bfef95601890afd80709 but is 8121e921473c734d3fc5c75347e95a6c708781a0.
```

If the file in the database does not match the checksum of the file in the file system then the output indicates the filename and the old (what the database stores) and new (what is in the file system currently) checksums.

5.3.6 Setperms output

```
2012/03/16 00:21:26 LOG      Chmod /tmp/skulkerv2-tests-10144/files/xx5 from '0744' to '0754'.
2012/03/16 00:21:26 LOG      Chmod /tmp/skulkerv2-tests-10144/files/xx5 from '0754' to '0750'.
```

If the setperms function changes file attributes, (such as the permissions bits, owner or group of the matched files), a line will be shown indicating the changes that have been made.

5.3.7 Delete output

```
2006/05/09 02:20:07 LOG      Delete: Truncated - /var/adm/btmp [user:group]
2006/05/09 02:20:07 LOG      Delete: Removed - /tmp/xxwewe [user:group]
```

The messages shown for delete will include "Truncated" if the file has not actually been removed and just reduced to 0 bytes. This occurs if the "truncate" or "handleopen" flags have been specified.

5.3.8 List output

```
2012/08/06 21:55:03 LOG      Matched /tmp/simon1/auth2.log [root:root, 2330 Bytes]
2012/08/06 21:55:03 LOG      Matched /tmp/simon1/boot.log [root:root, 804 Bytes]
```

Shows the names, owner, group and size of files matching the criteria for the rule - no actual changes are made to the file system.

5.3.9 Tail output

```
2006/05/09 12:55:24 LOG      /tmp/881/files/xx2 - 15 lines removed.
```

When tail removes data from a file a line similar to the above is shown - indicating the number of lines that were removed. If no lines need to be removed that will also be indicated.

5.3.10 Organise output

Organise is a facility that moves files to specified directories based on the characteristics of each file. When a file is moved output similar to the above is shown - which indicates the name of the directory the particular file was moved to.

```
2010/07/12 23:40:08 LOG      Moved file '/tmp/skulkerv2-tests-6073/files/xx' to '/tmp/skulkerv2-
tests-6073/files/2010-06-10'.
2010/07/12 23:40:08 LOG      Moved file '/tmp/skulkerv2-tests-6073/files/xx2' to '/tmp/skulkerv2-
tests-6073/files/2010-06-10'.
```

5.3.11 Archive output

The facilities currently offered by the "archive" module simply allow files to be copied into "ar" archives - further archive formats will be added later.

```
2010/07/14 23:26:29 LOG      Added file '/tmp/skulkerv2-tests-16002/files/xx' to archive
'/tmp/skulkerv2-tests-16002/files/myarchive'.
2010/07/14 23:26:29 LOG      Added file '/tmp/skulkerv2-tests-16002/files/xx2' to archive
'/tmp/skulkerv2-tests-16002/files/myarchive'.
```

5.3.12 Append output

Remember that the addition of the contents from one file to another works on compressed files to (both source and destination), if compression format supported.

```
2010/07/14 23:29:19 LOG      Added 5 lines [40 bytes] from /tmp/skulkerv2-tests-16045/files/xx3
to /tmp/skulkerv2-tests-16045/files/xx3.log.gz.
```

5.3.13 Move output

The differences between organising and moving files is subtle. The "organise" module allows you to simply move the files to a different directory (creating them as necessary). However the "move" can also move the files, but rename them as well.

```
2010/07/14 23:22:24 LOG      Moved file /tmp/skulkerv2-tests-15965/files/xx => /tmp/skulkerv2-
tests-15965/files/destdir1/xx
2010/07/14 23:22:24 LOG      Processed 1 file in 1 directory.
```

5.3.14 Chmod output

When permissions on a file are changed as a result of running the "chmod" function against a rule all files/directories which are changed are shown at log level 3. The format of the output for change files shows the original and the new permissions in octal.

```
2012/04/03 22:58:52 LOG      Chmod /tmp/skulkerv2-tests-2604/files/xx5 from '0644' to '0744'.
2012/04/03 22:58:52 LOG      Chmod /tmp/skulkerv2-tests-2604/files/xx5 from '0744' to '0754'.
```

5.4 The Report Action Output

The "report" action gives a long listing for each of the rules specified showing all statistics available for each rule. The format of the output generated is;

```
2006/05/09 14:18:10 LOG      Rule hostonly-myhost:1100      Rule action      : compress
handleopen type=compress
2006/05/09 14:18:10 LOG      Match type        : file
2006/05/09 14:18:10 LOG      Match pattern     : /a/b/.../*\.(log|spdslog)$
2006/05/09 14:18:10 LOG      Match by          : MTIME OLDER_THAN 1D
2006/05/09 14:18:10 LOG      Last run          : 2006/05/09 12:54
2006/05/09 14:18:10 LOG      Run interval      : 1D
2006/05/09 14:18:10 LOG      Last run Info     : Files: 0, Duration: 00:00:00,
Space: 0 Kb
2006/05/09 14:18:10 LOG      Total run count   : 37
```


5.6 The Totals Summary Action Log

The totals summary log is similar to the normal summary log but instead it shows the results for all invocations of the rules in question, not just the latest one.

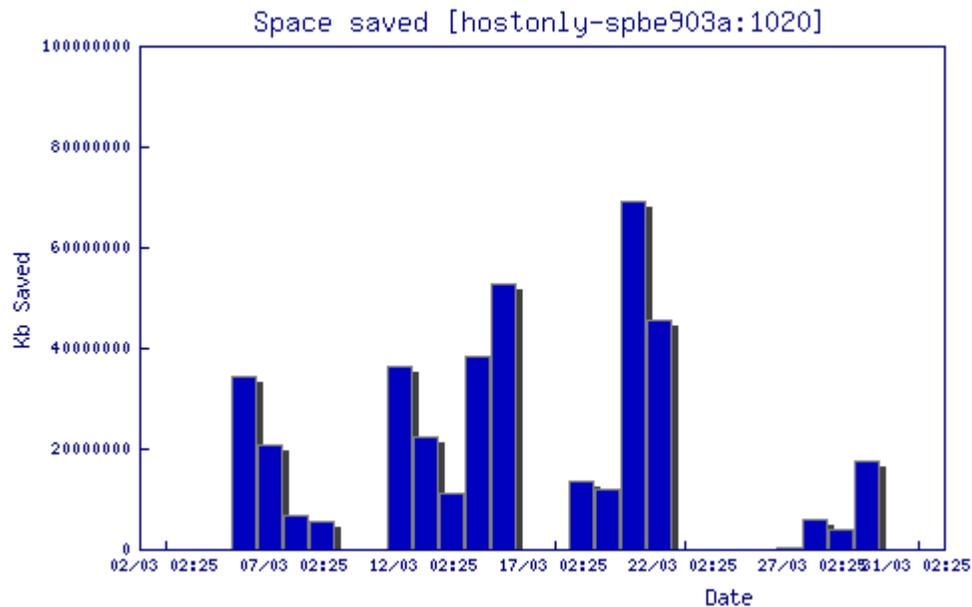
```
2007/03/10 00:41:09 LOG LAST RUN SUMMARY REPORT
2007/03/10 00:41:09 LOG =====
2007/03/10 00:41:09 LOG
2007/03/10 00:41:09 LOG Rule Files Duration Saved
2007/03/10 00:41:09 LOG linux-cups:10 0 00:00:00 0.00 Kb
2007/03/10 00:41:09 LOG linux-cups:20 0 00:00:00 0.00 Kb
2007/03/10 00:41:09 LOG linux-httpd:10 0 00:00:00 0.00 Kb
2007/03/10 00:41:09 LOG linux-httpd:20 0 00:00:00 0.00 Kb
2007/03/10 00:41:09 LOG linux-ppp:10 0 00:00:00 0.00 Kb
2007/03/10 00:41:09 LOG linux-ppp:20 0 00:00:00 0.00 Kb
2007/03/10 00:41:09 LOG linux-psacct:10 1 00:00:00 0.00 Kb
2007/03/10 00:41:09 LOG linux-psacct:20 2 00:00:00 72.60 Kb
2007/03/10 00:41:09 LOG linux-samba:10 0 00:00:00 0.00 Kb
2007/03/10 00:41:09 LOG linux-samba:20 0 00:00:00 0.00 Kb
2007/03/10 00:41:09 LOG linux-samba:30 0 00:00:00 0.00 Kb
2007/03/10 00:41:09 LOG linux-squid:10 0 00:00:00 0.00 Kb
2007/03/10 00:41:09 LOG linux-squid:20 0 00:00:00 0.00 Kb
2007/03/10 00:41:09 LOG linux-squid:30 0 00:00:00 0.00 Kb
2007/03/10 00:41:09 LOG linux-varcache:1000 740 00:00:07 511609.30 Kb
2007/03/10 00:41:09 LOG linux-varlogs:10 1 00:00:00 0.00 Kb
2007/03/10 00:41:09 LOG linux-varlogs:20 9 00:00:00 319.30 Kb
2007/03/10 00:41:09 LOG linux-varlogs:30 0 00:00:00 0.00 Kb
2007/03/10 00:41:09 LOG linux-varlogs:40 0 00:00:00 0.00 Kb
2007/03/10 00:41:09 LOG linux-varlogs:50 0 00:00:00 0.00 Kb
2007/03/10 00:41:09 LOG linux-varspool:1000 0 00:00:00 0.00 Kb
2007/03/10 00:41:09 LOG linux-varspool:1010 1 00:00:00 66.10 Kb
2007/03/10 00:41:09 LOG linux-varspool:1020 0 00:00:00 0.00 Kb
2007/03/10 00:41:09 LOG linux-varspool:1030 0 00:00:00 0.00 Kb
2007/03/10 00:41:09 LOG skulker2:1000 0 00:00:00 0.00 Kb
2007/03/10 00:41:09 LOG skulker2:1010 0 00:00:00 0.00 Kb
2007/03/10 00:41:09 LOG
2007/03/10 00:41:09 LOG TOTALS 754 00:00:07 512,067.30 Kb
2007/03/10 00:41:09 LOG
```

Notice for this feature the values are given in Kb, Mb or Gb depending on the volumes of space actually saved.

5.7 The "Html Report" Option Output

The html report option allows the user to generate a graphical output showing the largest space saving rules over a period of time. This is only possible if historical information is being kept. A typical output will have entries similar to the following:

Top 10 Skulker Rules for December 2007



5.7.1 Rule Details

```
Rule action : saswork handleopen
Match type : directory
Match pattern : /sas/work/saswork/.../^SAS_work.*myhost$
Match by : ALWAYS
Last run : 2007/12/20 02:25
Run interval : 6H
[Files: 49, Duration: 00:00:00, Space: 3221380.5 Kb]
```

5.7.2 Runs over 31 Days

```
Runs : 4
Files: 84
Space: 4,115,729 Kb
Spent: 00:00:00
```

The distributed utility "util/gen_html_report" gives an example of how the above output was generated. It should be noted that two formats of the option can be passed to Skulker:

```
--action=htmlreport
```

Produces the specified number of rules in the format shown on the previous page.

```
--action=htmlreport+tsummary
```

An example call using the above format would also include a summary table of all rules currently known at the end of the report, similar to the following:

Summary

Rule	Run Occurrences	Files Processed	Space Save
testrules-21021:30	14	1017	
testrules-20595:30	9	1008	
testrules-21021:100	1	1	
testrules-20595:100	1	1	
testrules-20595:90	1	1	
testrules-21021:90	1	1	
testrules-21021:1044	1	1	
testrules-21021:1043	1	1	
testrules-20595:10	12	521	
testrules-21021:10	17	523	
testrules-21021:1042	1	1	
testrules-21021:20	17	223	
testrules-20595:20	12	215	
testrules-21021:1040	1	1	
testrules-10144:10	1	1	
testrules-9595:40	1	0	
Total Space Saved (MB)	106		