

Configuring the HPCC Systems® Platform

Boca Raton Documentation Team



Configuring the HPCC Systems® Platform

Boca Raton Documentation Team

Copyright © 2025 HPCC Systems®. All rights reserved

We welcome your comments and feedback about this document via email to <docfeedback@hpccsystems.com>

Please include **Documentation Feedback** in the subject line and reference the document name, page numbers, and current Version Number in the text of the message.

LexisNexis and the Knowledge Burst logo are registered trademarks of Reed Elsevier Properties Inc., used under license.

HPCC Systems® is a registered trademark of LexisNexis Risk Data Management Inc.

Other products, logos, and services may be trademarks or registered trademarks of their respective companies.

All names and example data used in this manual are fictitious. Any similarity to actual persons, living or dead, is purely coincidental.

2025 Version 9.14.24-1

Using Configuration Manager	4
Running the Configuration Manager	5
Configuration Manager Advanced View	12
Using ConfigMgr in Advanced Mode	13
Hardware Section	17
Software Section	20

Using Configuration Manager

Configuration Manager is the utility with which we configure the HPCC Systems platform. The HPCC Systems platform's configuration is stored in an XML file named **environment.xml**. When you install a package, a default single-node environment.xml is generated. After that, you can use the Configuration Manager to modify it and add nodes and configure components.

The Configuration Manager Wizard creates a similar file, but after it is generated, you must rename it and put it into place on each node.

Configuration Manager also offers an **Advanced View** which allows you to add instances of components or change the default settings for components. Even if you plan to use Advanced View, it is a good idea to start with a wizard generated configuration and use Advanced View to finish it.

This document will guide you through configuring an HPCC Systems environment using the Configuration Manager.

Running the Configuration Manager

This document will guide you through configuring an HPCC Systems environment using the Configuration Manager.

The HPCC Systems package should already be installed on ALL nodes.

You can use any tool or shell script you choose.

1. SSH to a node in your environment and login as a user with sudo privileges. We would suggest that it would be the first node, and that it is a support node, however that is up to your discretion.
2. Start the Configuration Manager service on the node (again we would suggest that it should be on a support node, and further that you use the same node to start the Configuration Manager every time, but this is also entirely up to you).

```
sudo /opt/HPCCSystems/sbin/configmgr
```

A screenshot of a terminal window with a blue title bar. The terminal shows the command 'sudo /opt/HPCCSystems/sbin/configmgr' being executed. The output text is as follows:

```
node219008 ~]$ sudo /opt/HPCCSystems/sbin/configmgr
Using default filename /etc/HPCCSystems/source/environment.xml and default port
"8015"
Validating environment file /etc/HPCCSystems/source/environment.xml using config
mgr ... Success
Verifying configmgr startup ... Success
Exit by pressing ctrl-c...
```

A green cursor is visible on the line following the last message.

3. Using a Web browser, go to the Configuration Manager's interface:

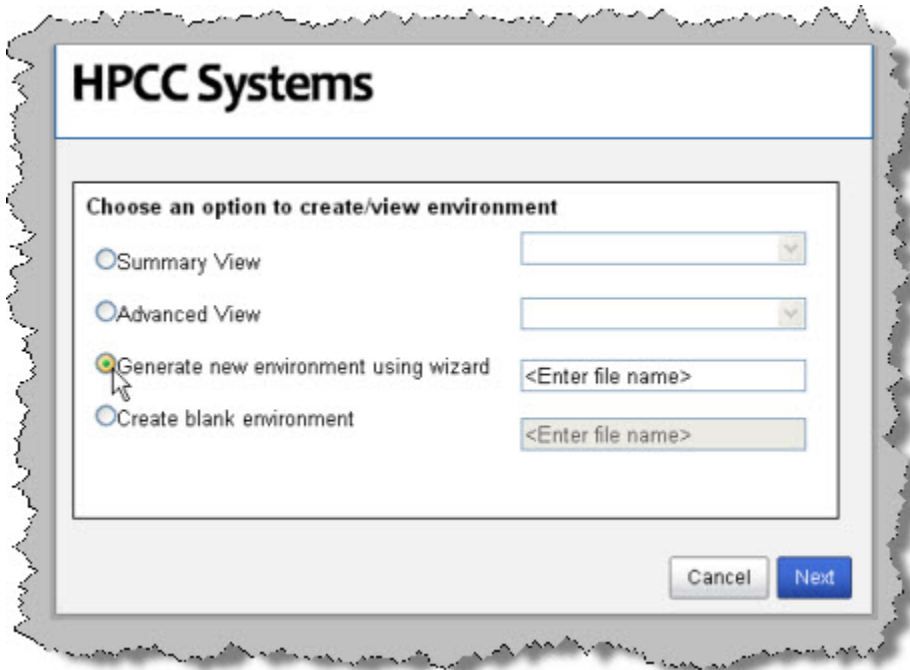
```
http://<ip of installed system>:8015
```

The Configuration Manager startup wizard displays.

There are different ways to configure your HPCC Systems platform. You can use the **Generate environment wizard** and use that environment or experienced users can then use the **Advanced View** for more specific customization. There is also the option of using **Create blank environment** to generate an empty environment that you could then go in and add only the components you would want.

Environment Wizard

1. To use the wizard select the **Generate new environment using wizard** button.



2. Provide a name for the environment file.

This will then be the name of the configuration XML file. For example, we will name our environment *NewEnvironment* and this will produce a configuration XML file named *NewEnvironment.xml* that we will use.

3. Press the Next button.

Next you will need to define the IP addresses that your HPCC Systems platform will be using.

4. Enter the IP addresses or hostname(s).

IP Addresses can be specified individually using semi-colon delimiters. You can also specify a range of IPs using a hyphen (for example, nnn.nnn.nnn.x-y). In the image below, we specified the IP addresses 10.239.219.1 through 10.239.219.100 using the range syntax, and also a single IP 10.239.219.111. Alternatively, you can enter the hostnames.



HPCC Systems

Environment setup

Welcome to wizard mode!

Define IP Addresses and/or hostnames for the environment being configured.
IP Address format: X.X.X.X; X.X.X.X-XXX;

192.168.56.1-125;192.168.56.128;MyHostName;

Cancel Back Next

5. Press the Next button.

Now you will define how many nodes to use for the Roxie and Thor clusters.

6. Enter the appropriate values as indicated.

HPCC Systems

Environment setup

Enter number of nodes for Roxie and Thor clusters. No Roxie/Thor cluster will be generated for zero (0) number of nodes.

Number of support nodes	7
Number of nodes for Roxie cluster	20
Number of slave nodes for Thor cluster (A Thor Master will be added to the cluster and assigned to a support node)	100
Number of Thor slaves per node (default 1)	1
Enable Roxie on demand	<input checked="" type="checkbox"/>

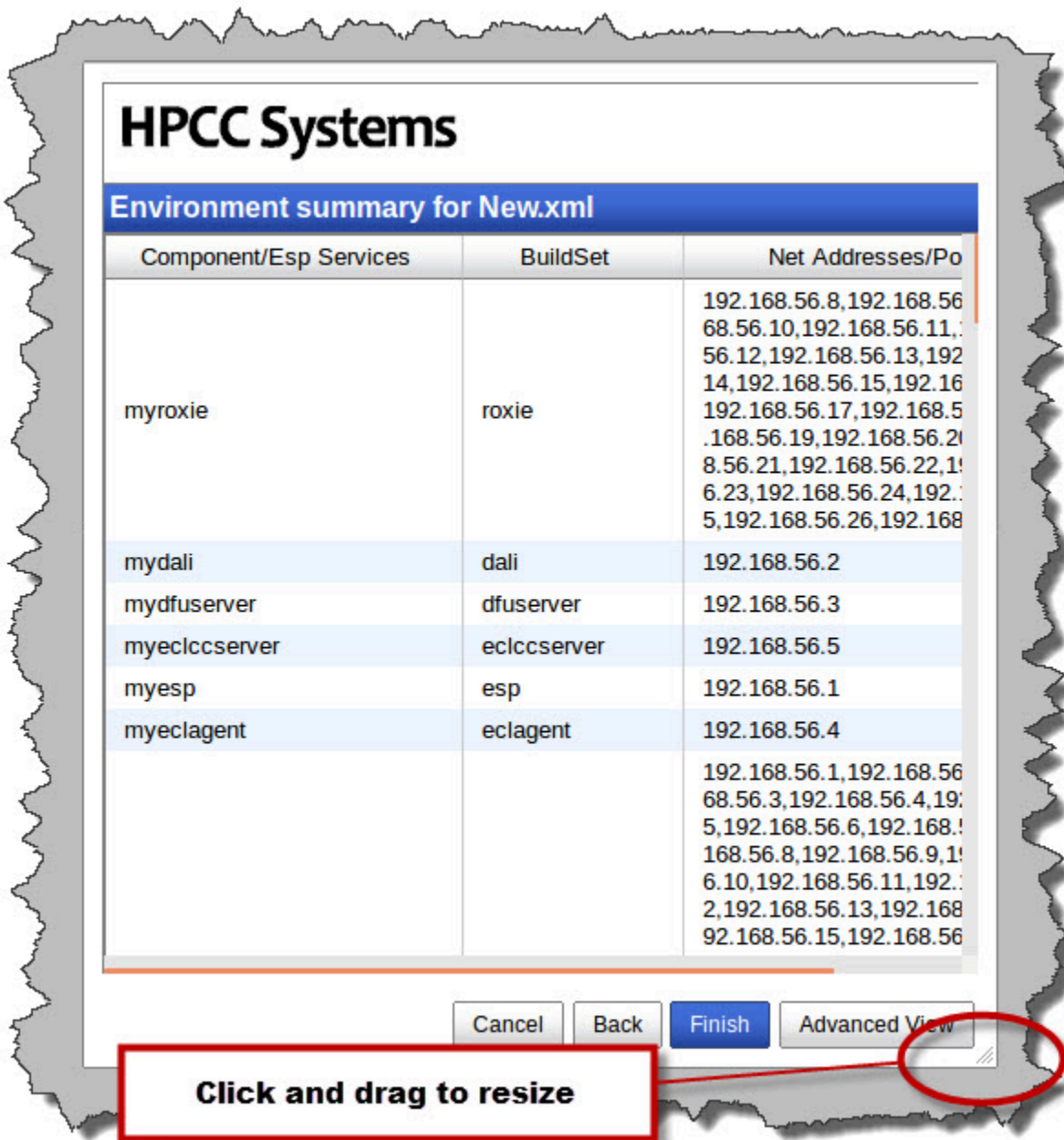
Cancel Back Next

Number of support nodes:	Specify the number of nodes to use for support components. The default is 1.
Number of nodes for Roxie cluster:	Specify the number of nodes to use for your Roxie cluster. Enter zero (0) if you do not want a Roxie cluster.
Number of slave nodes for Thor cluster	Specify the number of slave nodes to use in your Thor cluster. A Thor master node will be added automatically. Enter zero (0) if you do not want any Thor slaves.
Number of Thor slaves per node (default 1)	Specify the number of Thor slave processes to instantiate on each slave node. Enter zero (0) if you do not want a Thor cluster.
Enable Roxie on demand	Specify whether or not to allow queries to be run immediately on Roxie. (Default is true)

7. Press the **Next** button

The wizard displays the configuration parameters.

8. Press the **Finish** button to accept these values or press the **Advanced View** button to edit in advanced mode.



You will now be notified that you have completed the wizard.



At this point, you have created a file named NewEnvironment.xml in the **/etc/HPCCSystems/source** directory



Keep in mind, that your HPCC Systems configuration may be different depending on your needs. For example, you may not need a Roxie or you may need several smaller Roxie clusters. In addition, in a production [Thor] system, you would ensure that Thor and Roxie nodes are dedicated and have no other processes running on them. This document is intended to show you how to use the configuration tools. Capacity planning and system design is covered in a training module.

Distribute the Configuration

1. Stop the HPCC Systems platform.

If it is running stop the HPCC Systems platform (on every node), using a command such as this:

```
sudo systemctl stop hpccsystems-platform.target
```

Note: You may have a multi-node system and a custom script such as the one illustrated in Appendix of the [Installing and Running the HPCC Systems Platform](#) document to start and stop your system. If that is the case please use the appropriate command for stopping your system on every node.



Be sure the HPCC Systems platform is stopped before attempting to copy the environment.xml file.

2. Back up the original environment.xml file.

```
# For example
sudo -u hpcc cp /etc/HPCCSystems/environment.xml /etc/HPCCSystems/source/environment-date.xml
```

Note: The live environment.xml file is located in your **/etc/HPCCSystems/** directory. ConfigManager works on files in **/etc/HPCCSystems/source** directory. You must copy from this location to make an environment.xml file active.

You can also choose to give the environment file a more descriptive name, to help differentiate any differences.

Having environment files under source control is a good way to archive your environment settings.

3. Copy the new .xml file from the source directory to the /etc/HPCCSystems and rename the file to *environment.xml*

```
# for example
sudo -u hpcc cp /etc/HPCCSystems/source/NewEnvironment.xml /etc/HPCCSystems/environment.xml
```

4. Copy the **/etc/HPCCSystems/environment.xml** to the **/etc/HPCCSystems/** on to every node.

You may want to use a script to push out the XML file to all nodes. See the *Example Scripts* section in the Appendix of the [Installing and Running the HPCC Systems Platform](#) document. You can use the scripts as a model to create your own script to copy the environment.xml file out to all your nodes.

5. Restart the HPCC Systems platform on all nodes.

Configuration Manager Advanced View

For the advanced user, the Advanced View offers access to adding additional instances of components or making configuration settings for individual components.

Using ConfigMgr in Advanced Mode

This section shows some of the configuration options in Advanced Mode. There are a few different ways to configure your system. If you are not an experienced user you can use the Generate environment wizard discussed in the previous section. The following steps will detail the Advanced set up.

1. SSH to the first box in your environment and login as a user with sudo privileges.
2. If it is running, stop the HPCC Systems platform using this command on every node:

```
sudo systemctl stop hpccsystems-platform.target
```

Note: If you have a large system with many nodes, you may want to use a script to perform this step. See the *Example Scripts* section in the Appendix of the [Installing and Running the HPCC Systems Platform](#) document.



You can use this command to confirm HPCC Systems processes are stopped:

```
sudo systemctl status hpccsystems-platform.target
```

3. Start the Configuration Manager service on one node (usually the first node is considered the head node and is used for this task, but this is up to you).

```
sudo /opt/HPCCSystems/sbin/configmgr
```

```
node219008 ~]$ sudo /opt/HPCCSystems/sbin/configmgr
Using default filename /etc/HPCCSystems/source/environment.xml and default port
"8015"
Validating environment file /etc/HPCCSystems/source/environment.xml using config
gen ... Success
Verifying configmgr startup ... Success
Exit by pressing ctrl-c...
```

4. Using a Web browser, go to the Configuration Manager's interface:

```
http://<ip of installed system>:8015
```

The Configuration Manager startup wizard displays.

5. Select **Advanced View**, then press the **Next** button.

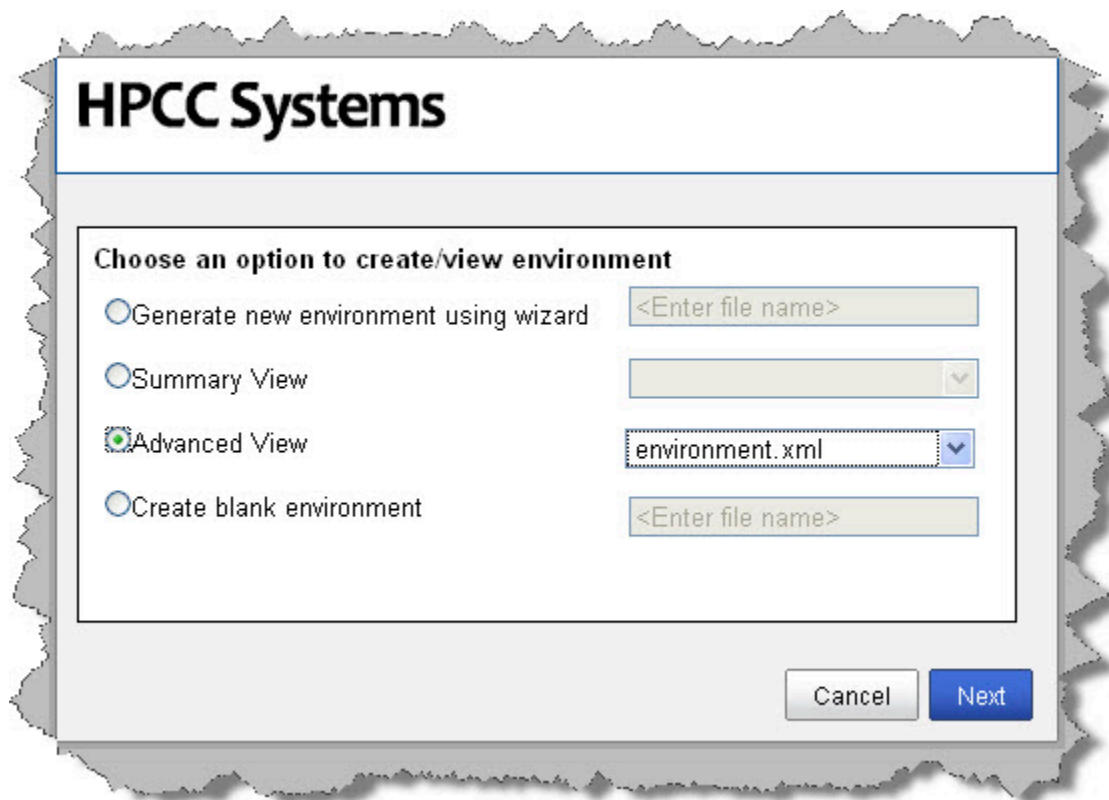
There are a few different ways to configure your system. If you are not an experienced user you can use the Generate environment wizard discussed in the previous section.

6. Select an XML file from the drop list.

This list is populated from versions of an environment XML file in your server's /etc/HPCCSystems/source/ directory.

The system will check the current environment file and if a match is found here it will highlight in blue the current environment file being used.

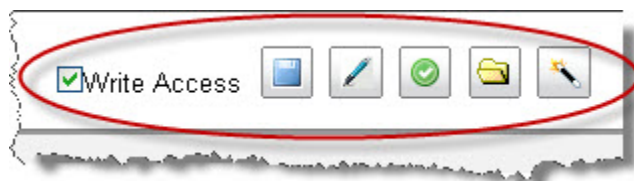
7. Press the **Next** button.





8. The Configuration Manager interface displays.


	<p>Default access is read-only. Many options are only available when write-access is enabled.</p> <p>Gain write access by checking the Write Access checkbox.</p> <p>Unchecking this box returns the environment to read-only mode. All menu items are disabled in read-only mode.</p> <p>Closing the web page automatically removes any write-access locks.</p>
--	---


9. Check the **Write Access** box.




The **Save** button  validates and saves the environment.

The **Save Environment As** button  validates and lets you specify the environment filename to save.

The **Validate Environment** button  just validates the current environment including any changes that have not yet been saved.

The **Open Environment** button  allows you to open a new environment file to work on.

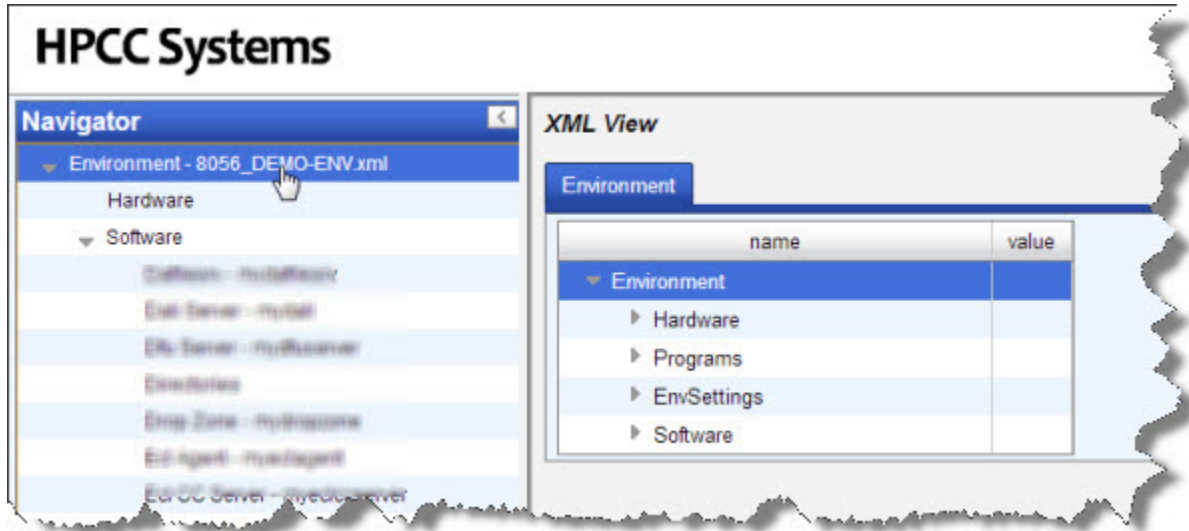
The **Wizard** button  will bring up the Configuration Manager chooser form which will allow you to create or view an environment file where you can also launch the configuration wizard.

These buttons are only enabled in Write Access mode.

XML View

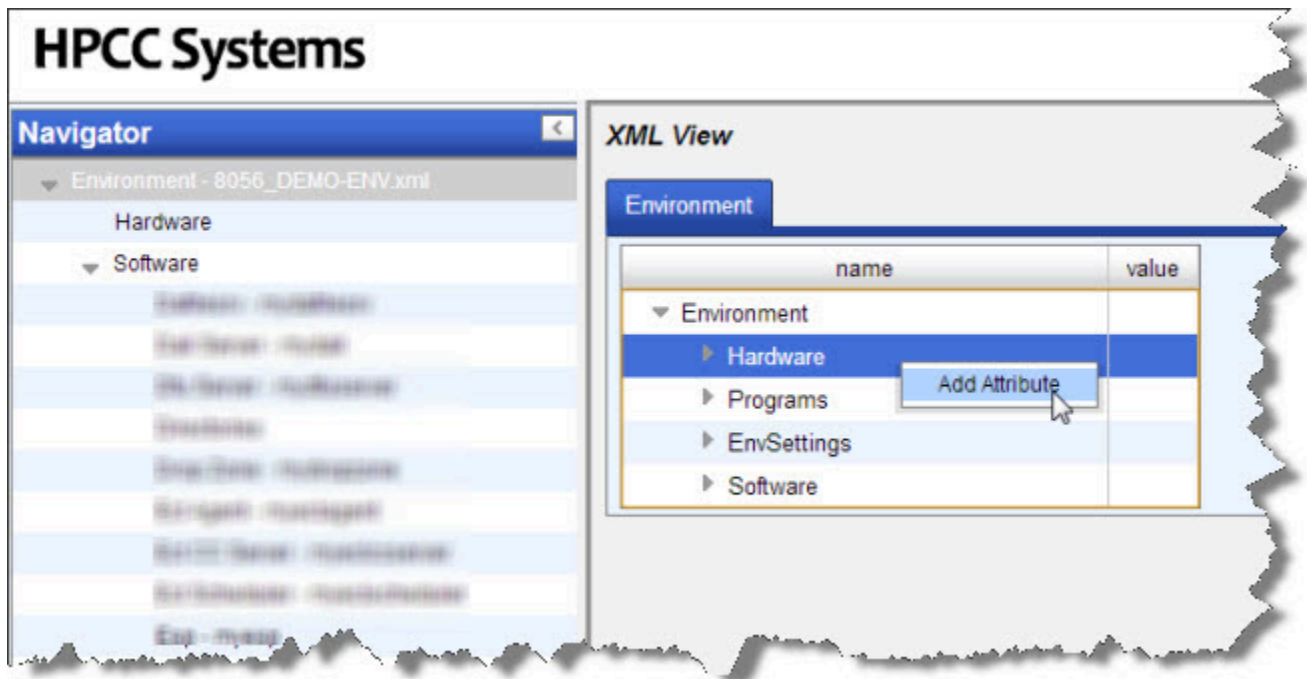
In the advanced view of Configuration Manager, you can optionally choose to work with the XML View.

To see the configuration in XML View, click on the Environment heading in the Navigator panel on the left side.



You can access all attributes through the XML view.

If you wish to add an attribute that does not exist, right-click on one of the components then you can choose to add an attribute.



Hardware Section

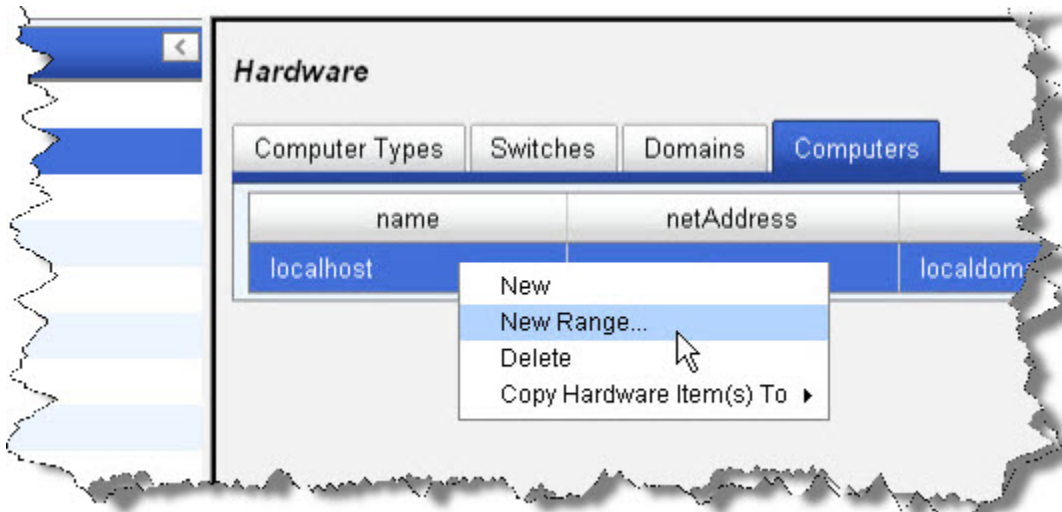
This section allows you to define your list of servers. When defining instances of components, you will choose from servers in this list.

1. Select **Hardware** in the Navigator panel on the left side.



2. Select the **Computers** tab.

3. right-click on one of computers listed, then select New Range.



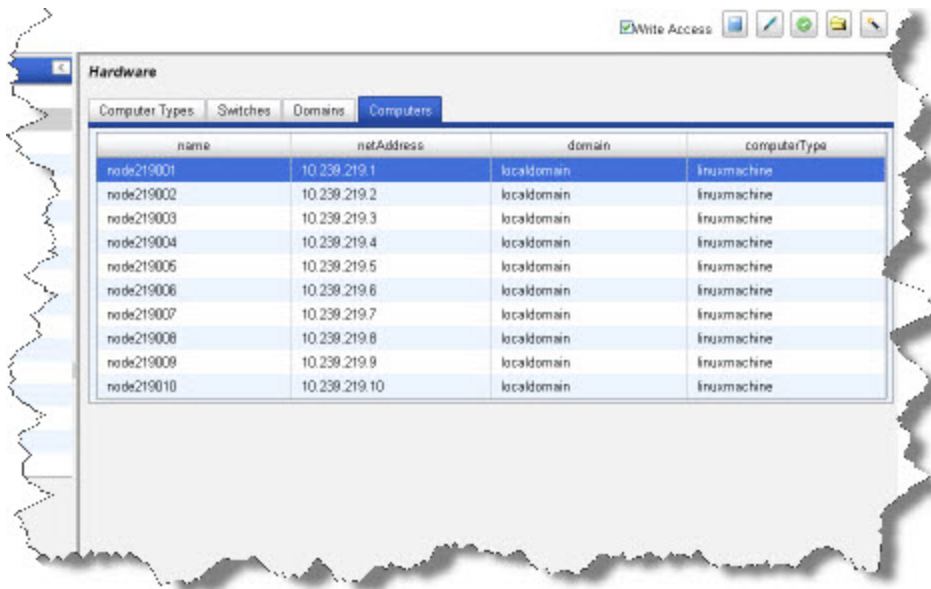
4. Specify the following:

- Name Prefix - any name that will help you to identify the node or range
- Start IP Address
- Stop IP Address


The IP Addresses can be specified in a range if all your host IP addresses are consecutively numbered. If the IP addresses are not sequential you should repeat the process for each individual IP address and just add the IP address in both the start and stop IP address field. You will then need to repeat the process for each node.

5. Press the **OK** button.

The list of nodes now displays with the nodes that you just added.



Next, edit each System Server component instance and set it to a newly defined node.

6.  Click the disk icon to save
7. Expand the **Software** section, if necessary, in the Navigator panel on the left side, by clicking on the ► button.

Software Section

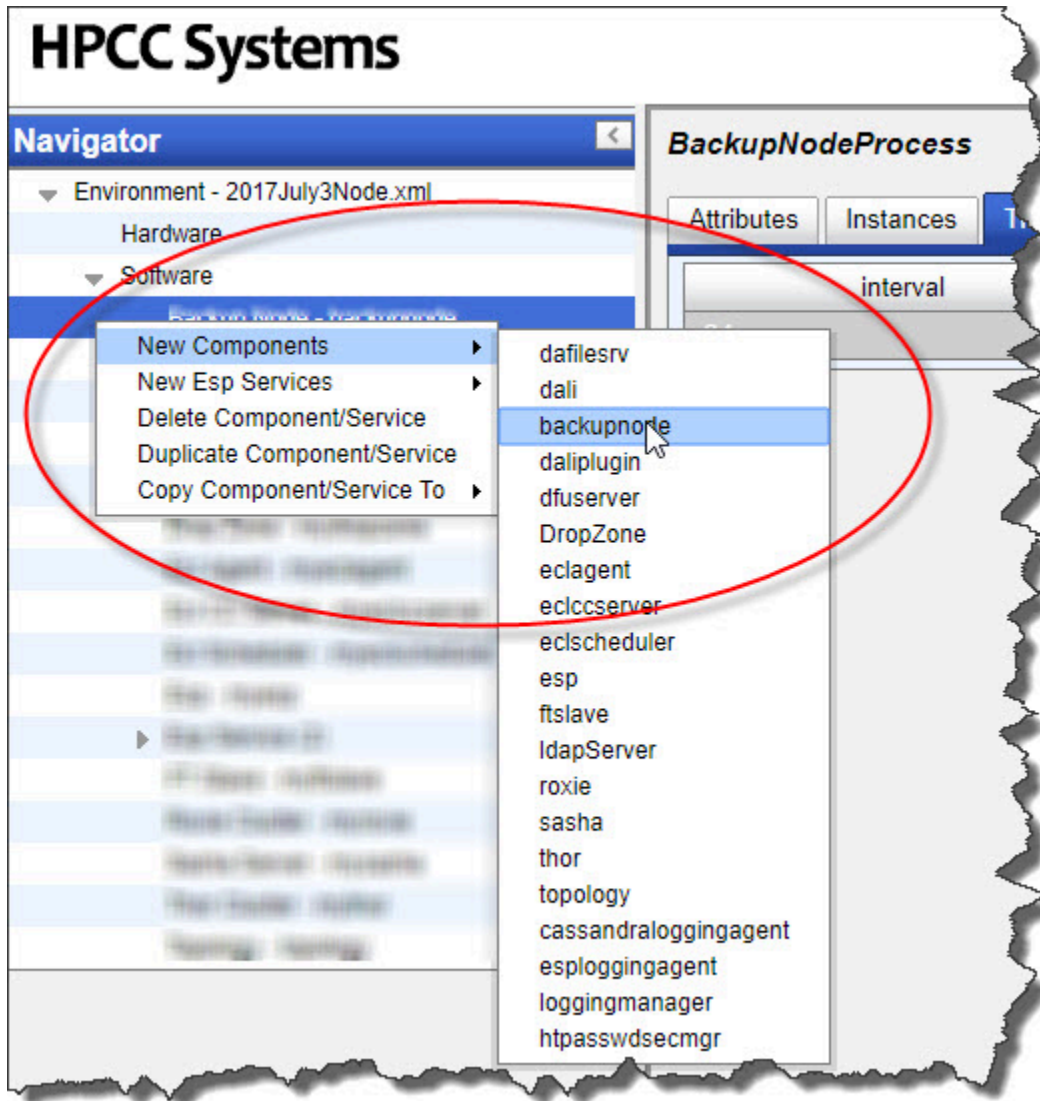
Use the software components section to configure software components of the HPCC Systems platform. Most software components are actual running processes; however, some are just definitions used by the system. These definitions are used by the configuration generator.

Items that appear in **red** indicate optional values. They are only written to the environment if you add to or change that value. If untouched, they will not appear in the environment XML file.

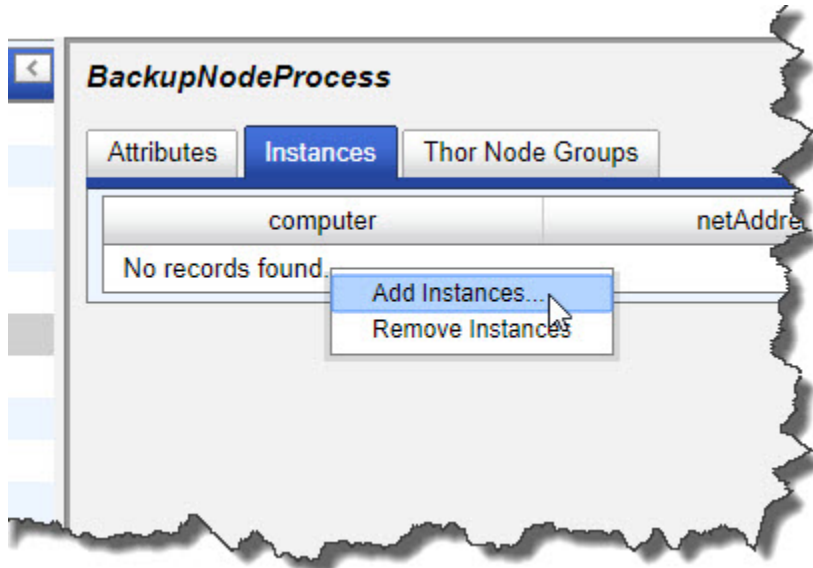
Backupnode


Backupnode allows you to back up Thor clusters at regular intervals. The Backupnode component is a way to allow administrators to manage the backupnode process without using a cron job.

To configure scheduled Thor node backups, add the backupnode component, choose the hardware instance to run it on and then add Thor groups to it.



1. Right click on the **Software** component in the **Navigator** panel (on the right side), choose **New Components** then **backupnode** from the drop list.
2. From the tabs on the right side, select the **Instances** tab.
3. Right click on the **computer** column and choose **Add Instances...**



4. Select the computer for the backupnode component, or press **Add Hardware** to add a new computer instance. You would always want to run backupnode on the Thor master of the cluster.
5. Select the **Thor Node Groups** tab.
6. Right click on the the **Interval** column and choose the interval and/or Thor group to back up.
7.  Click the disk icon to save

The default backup locations are:

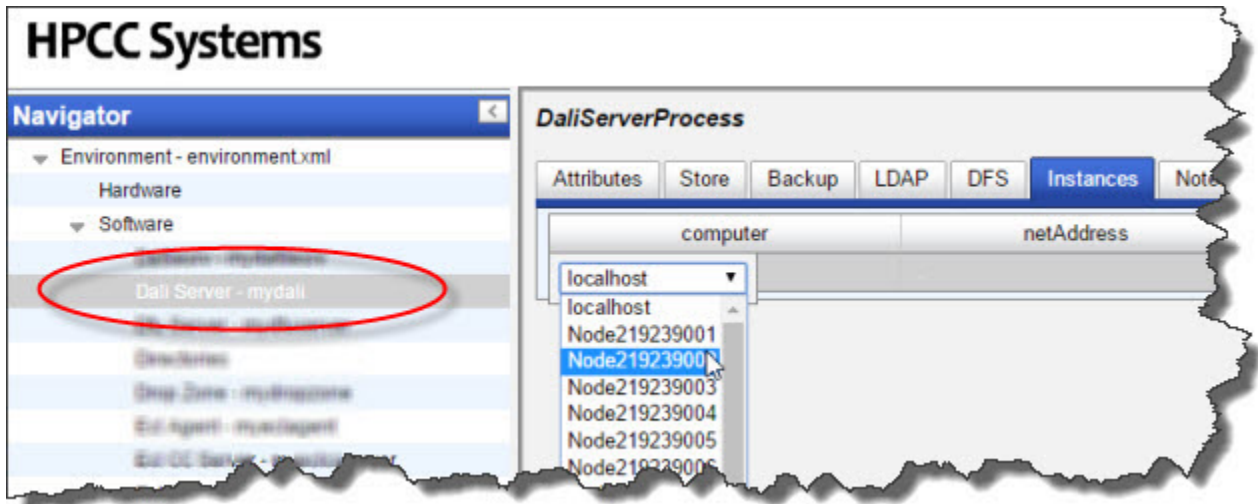
```
/var/lib/HPCCSystems/hpcc-data/backupnode/<thorname>/last_backup
```

The interval attribute of the backupnode component determines how many hours between backups.

Dali

Instances

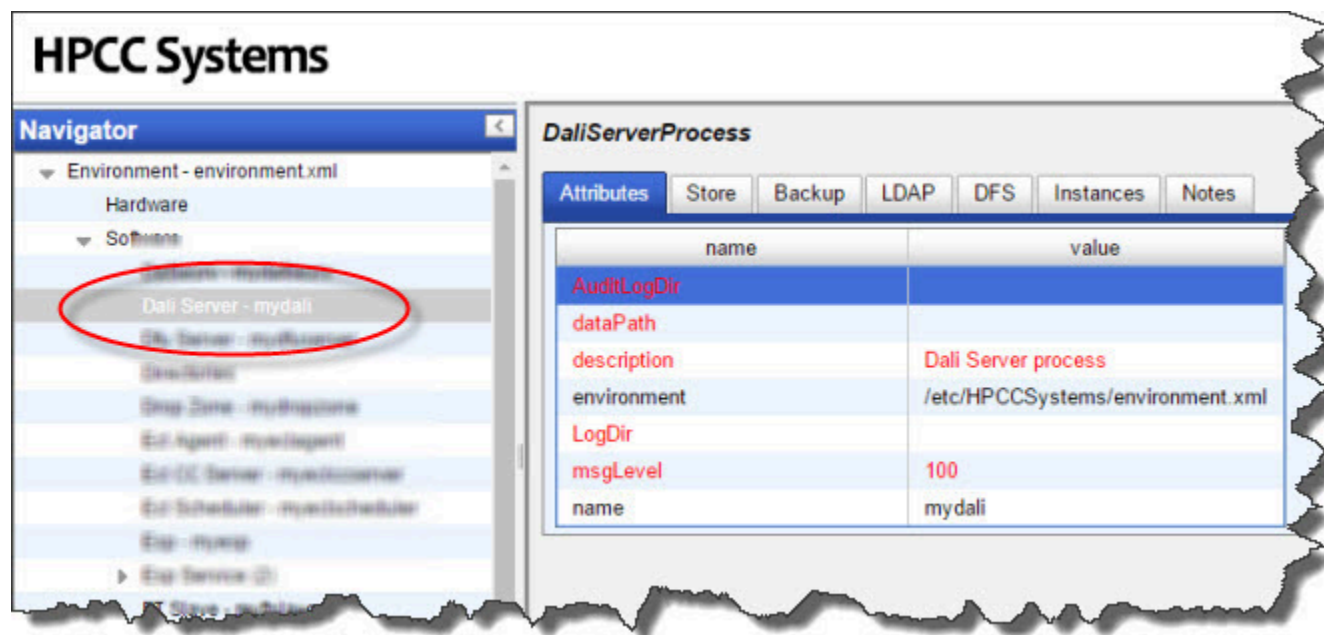
1. Select **Dali Server** in the Navigator panel on the left side.
2. Select the Instances tab.
3. In the computer column, choose a node from the drop list as shown below:



4. Click the  disk icon to save

DaliServer attributes

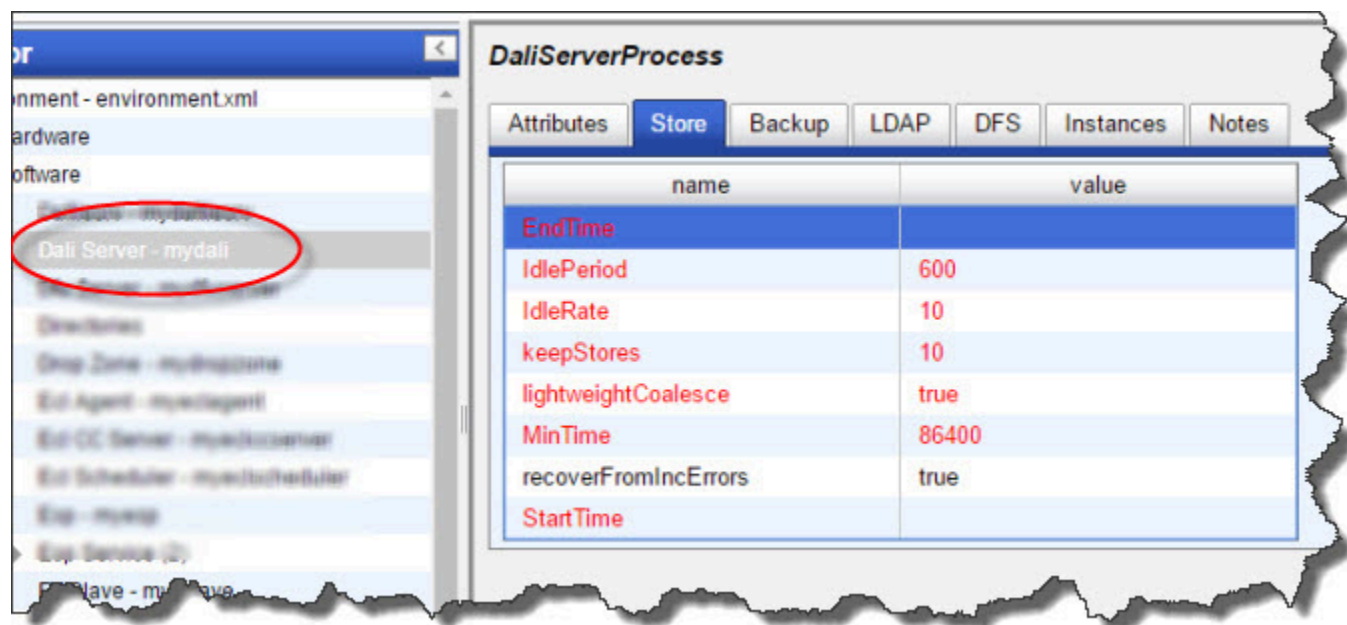
This section describes the DaliServer attributes.



attribute	description	values	required
name	Name for this process		required
description	Description for this process	Dali Server process	optional
msgLevel	Severity threshold for reporting errors in log file	100	optional
dataPath	Directory in which dali's data files will be written to		optional
LogDir	Directory in which to store server log files		optional
AuditLogDir	Directory in which to store audit log files		optional
environment	Path to an xml file containing an Environment to use		optional

DaliServer store

This section describes the attributes configuring how Dali handles the system data store.



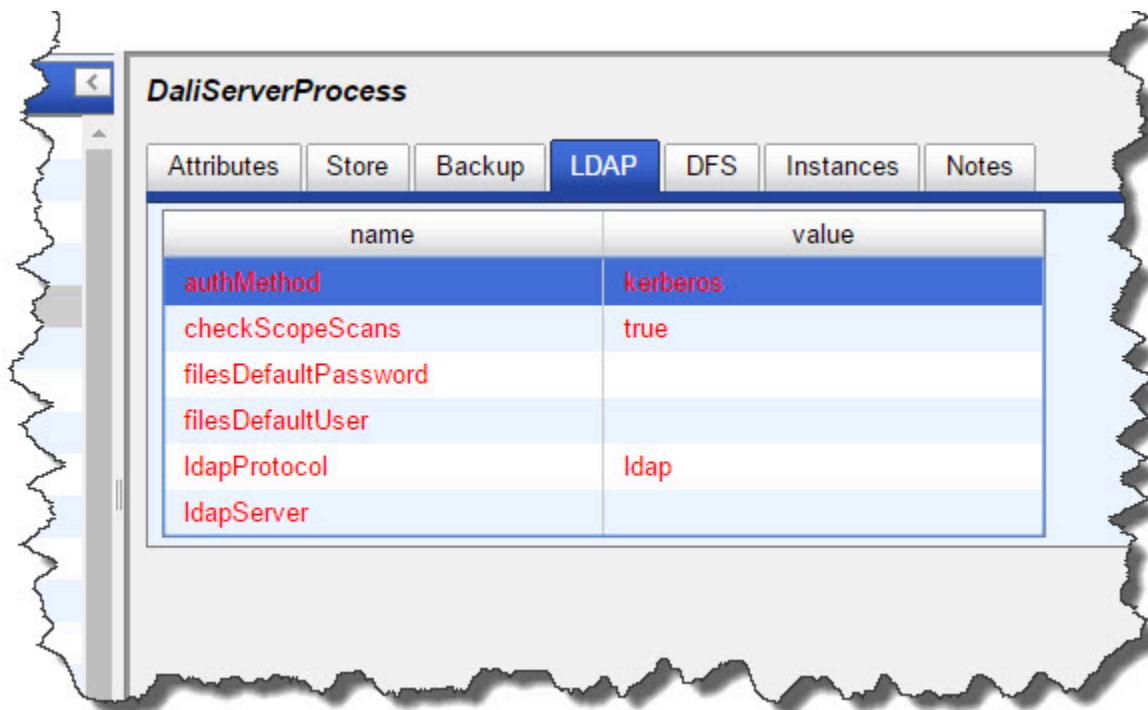
attribute	description	values	required
lightweightCoalesce	Enable non memory loaded consolidation of store	true	optional
IdlePeriod	Period of client to server quiet time to trigger store save	600	optional
IdleRate	Number of transaction per minute to be considered quiet time	10	optional
MinTime	Minimum amount of time between lightweight store saves	86400	optional
StartTime	Start time of lightweight coalesce checking		optional
EndTime	End time of lightweight coalesce checking		optional
keepStores	Number of old saved stores to keep	10	optional
recoverFromIncErrors	Switch on to auto recover from corruption to delta files on load	true	optional
saveBinary	Switch on saving the store in binary format	false	optional
deltaSaveThresholdSecs	Seconds to hold off between transaction commits to disk (default off, will commit immediately)	0	optional
deltaTransactionQueueLimit	The max limit of pending uncommitted transactions. If exceeded a synchronous save will be forced	10000	optional

Configuring the HPCC Systems® Platform
Configuration Manager Advanced View

attribute	description	values	required
deltaTransactionMaxMemMB	The max total memory limit of pending uncommitted transactions. If exceeded a synchronous save will be forced	10000	optional

DaliServer LDAP options

This section describes the DaliServer LDAP tab.



attribute	description	values	required
ldapServer	The ldap server to be used for authentication.		optional
ldapProtocol	The protocol to use - standard ldap or ldap over SSL.	ldap	optional
authMethod	The method to use for authentication and authorization.		optional
authPluginType	Security Manager plugin name (when authentication method is secmgrPlugin).		optional
filesDefaultUser	The default username for Files access (ActiveDirectory).		optional
filesDefaultPassword	The default password for filesDefaultUser.		optional
checkScopeScans	Enable LDAP checking for all logical file listings	true	optional

DaliServer Notes

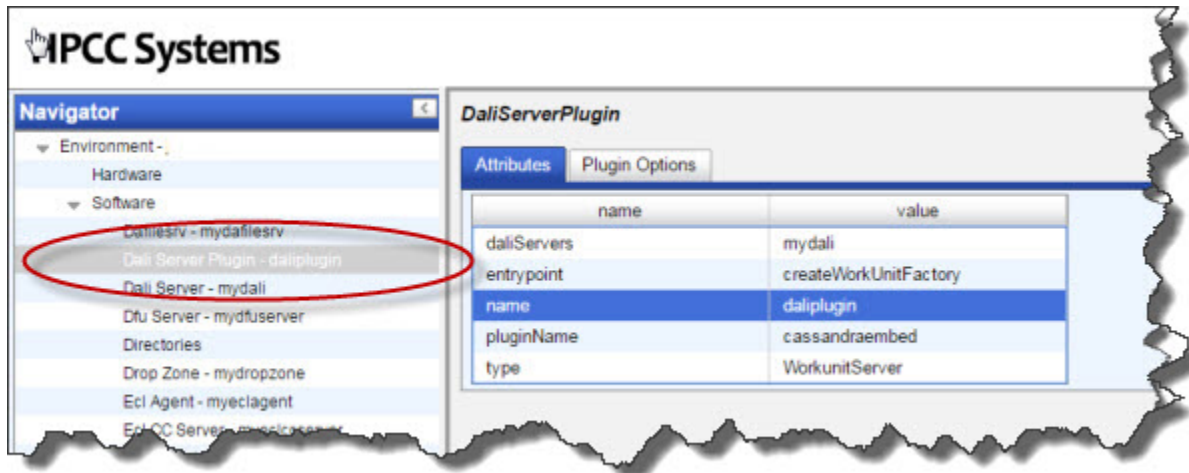
This tab allows you to add any notes pertinent to the component's configuration. This can be useful to keep a record of changes and to communicate this information to peers.

DaliServerPlugin

DaliServerPlugin allows you to add plugin functionality to a Dali server.

DaliServerPlugin attributes

This section describes the DaliServerPlugin attributes.



attribute	description	values	required
name	Name for this dali server plugin		required
entrypoint	Plugin entrypoint method	createWorkUnit-Factory	required
daliServers	Specifies the dali server to which this plugin is associated with.		required
pluginName	name of the plugin		required
type	Specifies the type of the plugin		optional

DaliServerPlugin Options

This section describes the DaliServerPlugin options

These options are available for the DaliServerplugin when configuring a Casandra server. See the System Administrator's Guide for more details about configuring a Cassandra server as a system datastore.

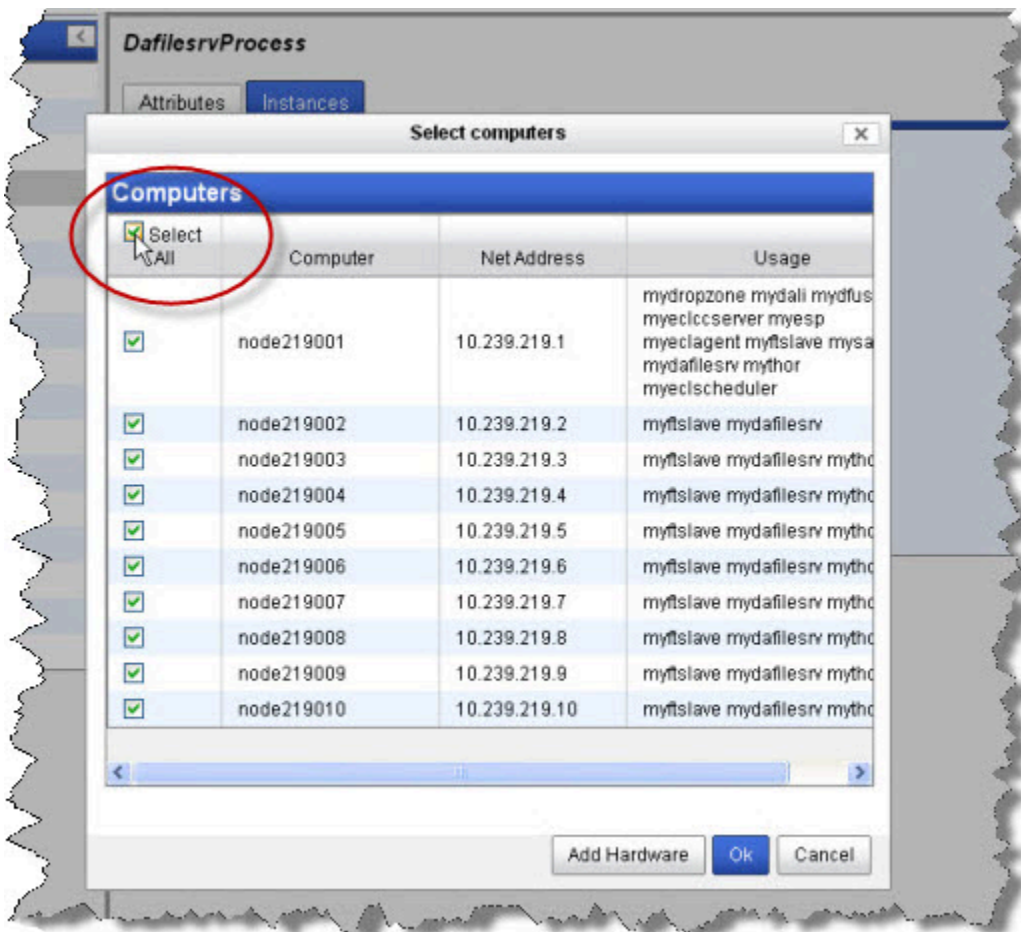
randomWuidSuffix	An integer value indicating how many randomized digits to append to workunits. Set this if you need to create workunits at a high rate to reduce the risk of collisions (which would slow down the process of creating a new unique workunit id).
traceLevel	An integer value indicating how much tracing to output from Cassandra workunit operations. Set to zero or do not set in normal usage.
partitions	An integer value indicating how many ways to partition the data on a Cassandra cluster. The default is 2. The value only takes effect when a new Cassandra workunit repository is created. Larger values permit scaling to a more distributed store but at the expense of some overhead on smaller stores where the scaling is not needed.
prefixsize	An integer value specifying the minimum number of characters that must be provided when wildcard searching in the repository. Larger values will be more efficient but also more restrictive on users. The default is 2. As with partitions, this value only takes effect when a new Cassandra workunit repository is created.
keyspace	The name of the Cassandra keyspace to use for the HPCC Systems data store. The default is <i>hpcc</i> .
user	The username to use if the Cassandra server is configured to require credentials.
password	The password to use if the Cassandra server is configured to require credentials.

Dafilesrv Process

Dafilesrv Instances

Dafilesrv is a helper process that every node needs.

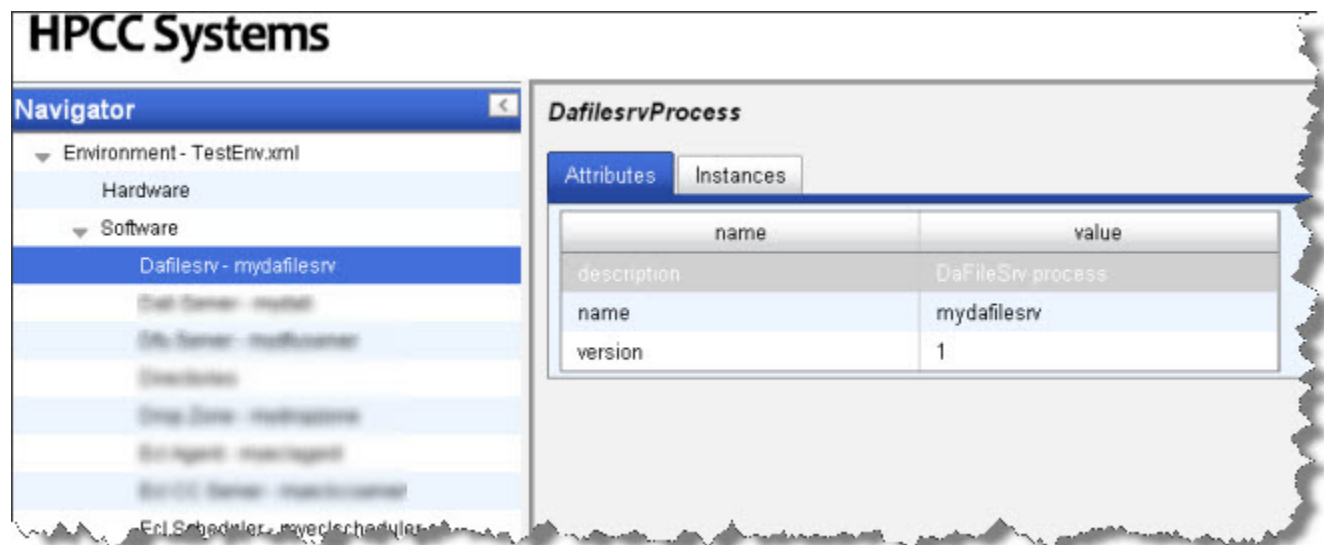
1. Select Dafilesrv in the Navigator panel on the left side.
2. Select the Instances tab.
3. Right-click on a computer in the computer column, and select Add Instance .
4. Select all computers in the list by checking the **Select All** box, then press the **OK** button.



5. Click the  disk icon to save

Dafilesrv attributes

This section describes the Dafilesrv attributes.

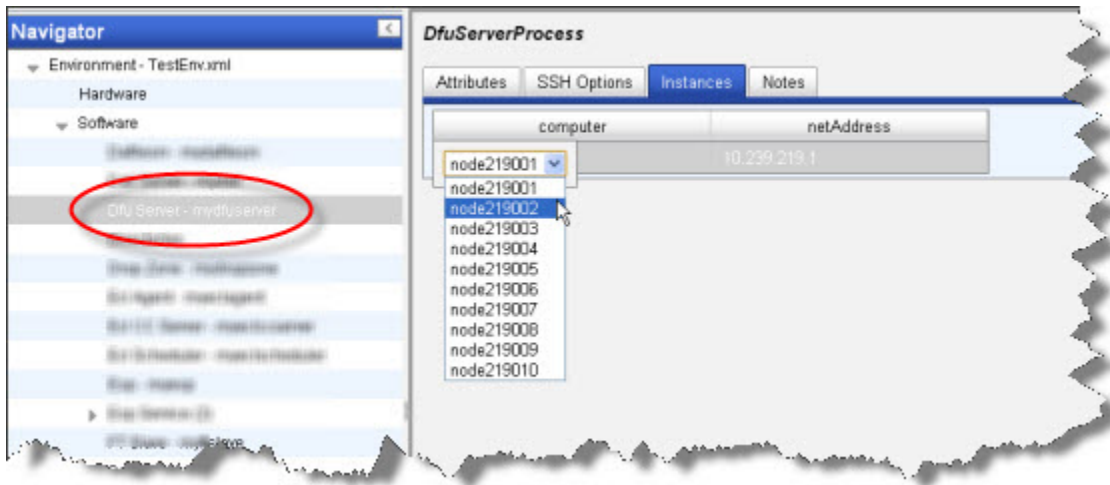


attribute	description	values	required
name	Name for this process		required
description	Description for this process	DaFileSrv process	optional
version	Version identifier used to select which process will be started	1	optional
parallelRequestLimit	Defines the maximum number of concurrent dafilesrv requests allowed. Requests that exceed the limit will be delayed. A value of 0 disables throttling. This a global setting.	20	optional
throttleDelayMs	Defines how many milliseconds delayed requests will be delayed by. This a global setting.	5000	optional
throttleCPULimit	If after the initial delay, the CPU % falls below this setting, the transaction will be allowed to continue, i.e. the limit can be exceeded this way. This a global setting.	75	optional

DFU Server

DfuServer Instances

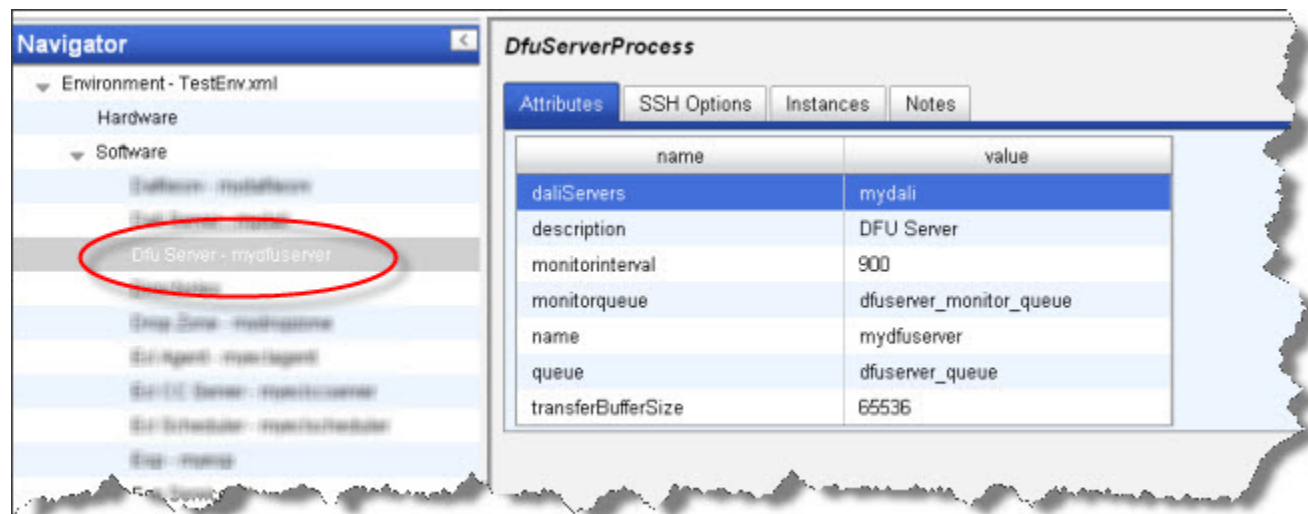
1. Select **DFU Server** in the Navigator panel on the left side.
2. Select the Instances tab.
3. In the computer column, choose a node from the drop list as shown below:



4. Click the  disk icon to save

DfuServer Attributes Tab

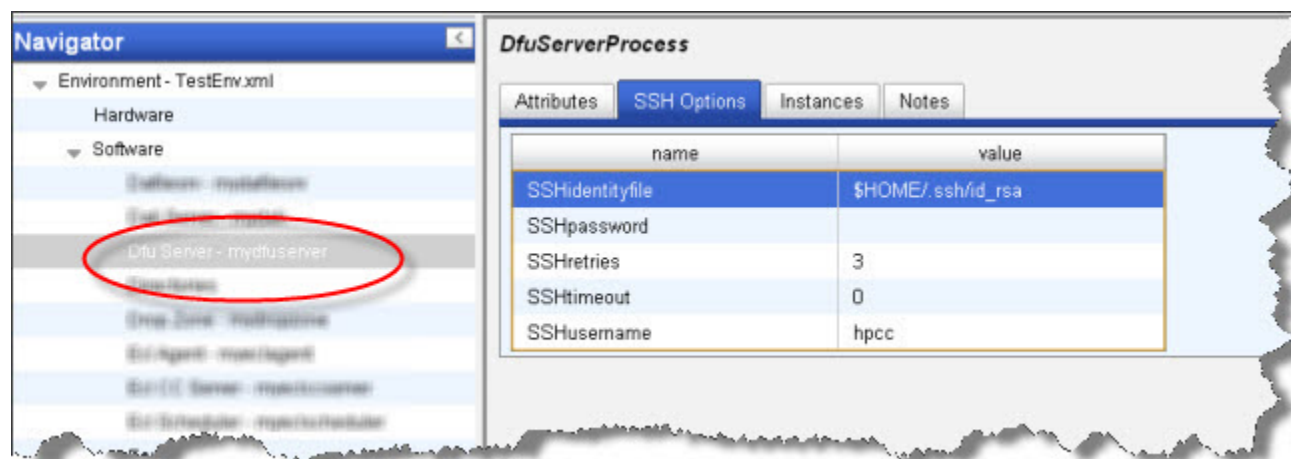
This section describes the DfuServer attributes.



attribute	description	values	required
name	Name for this process	dfuserver	required
description	Description for this process	DFU Server	optional
daliServers	Specifies the dali server to which this DFU server is attached.		required
queue	Specifies the queue name to send DFU Server jobs to.	dfuserver_queue	required
monitorqueue	Specifies the queue name to send DFU monitoring jobs to.	dfuserver_monitor_queue	required
monitorinterval	Specifies the polling interval for DFU monitoring (in seconds).	900	optional
transferBufferSize	Default buffer size used when transferring data.	1048576	optional

DfuServer SSH Options

This section describes the DfuServer SSH Options..



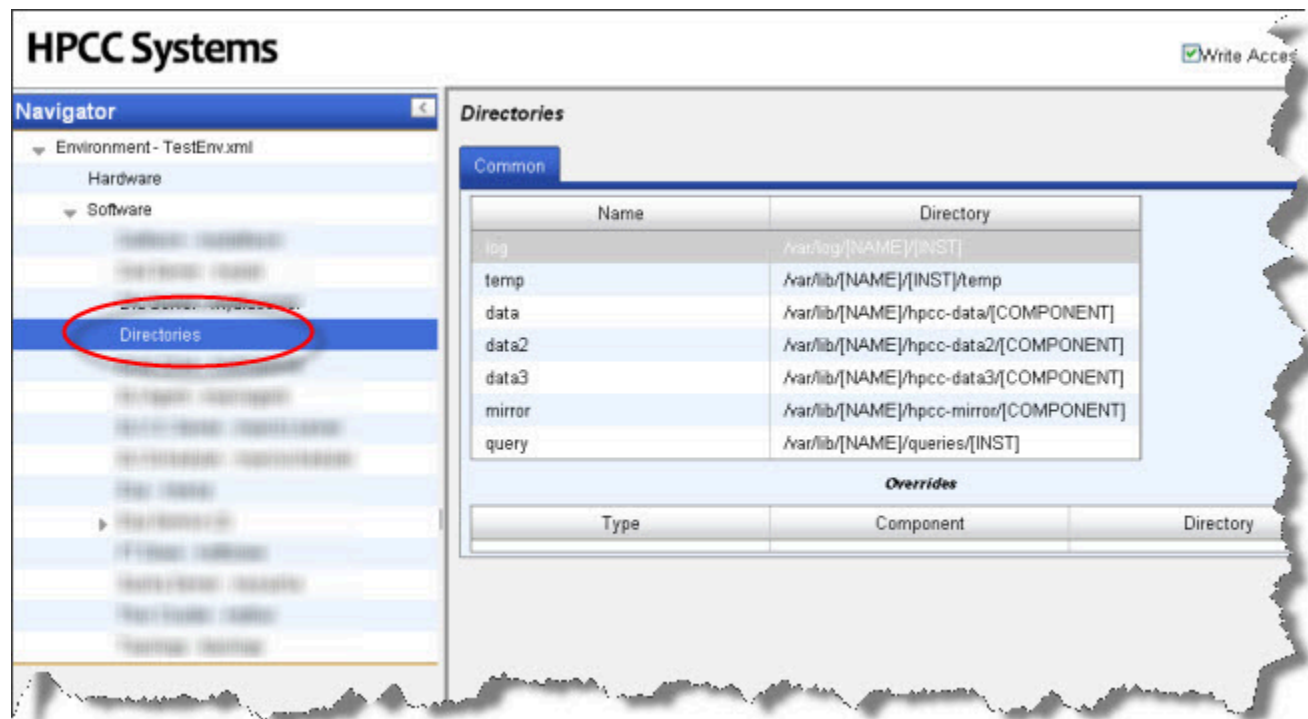
attribute	description	values	required
SSHidentityfile	location of identity file (private key) on Thor master	\$HOME/.ssh/id_rsa	optional
SSHusername	Username to use when running Thor slaves	hpcc	optional
SSHpassword	Fixed password - only required if no identity file present NB **insecure**		optional
SSHtimeout	Timeout in seconds for SSH connects	0	optional
SSHretries	Number of times to retry failed connect	3	optional

DfuServer Notes

This tab allows you to add any notes pertinent to the component's configuration. This can be useful to keep a record of changes and to communicate this information to peers.

Directories

The Directories component is a global definition used by other components to determine the directories they will use for various functions.



Name	Directory	Description
log	/var/log/[NAME]/[INST]	Location for Log files
temp	/var/lib/[NAME]/[INST]/temp	Location for temp files
data		Base Location for data files
data2		Base Location for 2nd copy of roxie data
data3		Reserved for future use
mirror		Base Location for mirror data files
query		Base Location for Queries

Drop Zone

A Drop Zone (or landing zone) is a location where files can be transferred to or from your HPCC Systems installation. The drop zone is a logical combination of a path and one or more servers.

Multiple drop zones allow you to configure different top level folders for one or more servers. Multiple servers for a single drop zone provides a logical grouping of distinct locations. Multiple drop zones are useful to allow different permissions for users or groups.

To add a drop zone:

1. Right-click on the Navigator panel on the left side and choose **New Components**
2. Select **Drop Zone**

Drop Zone Attributes

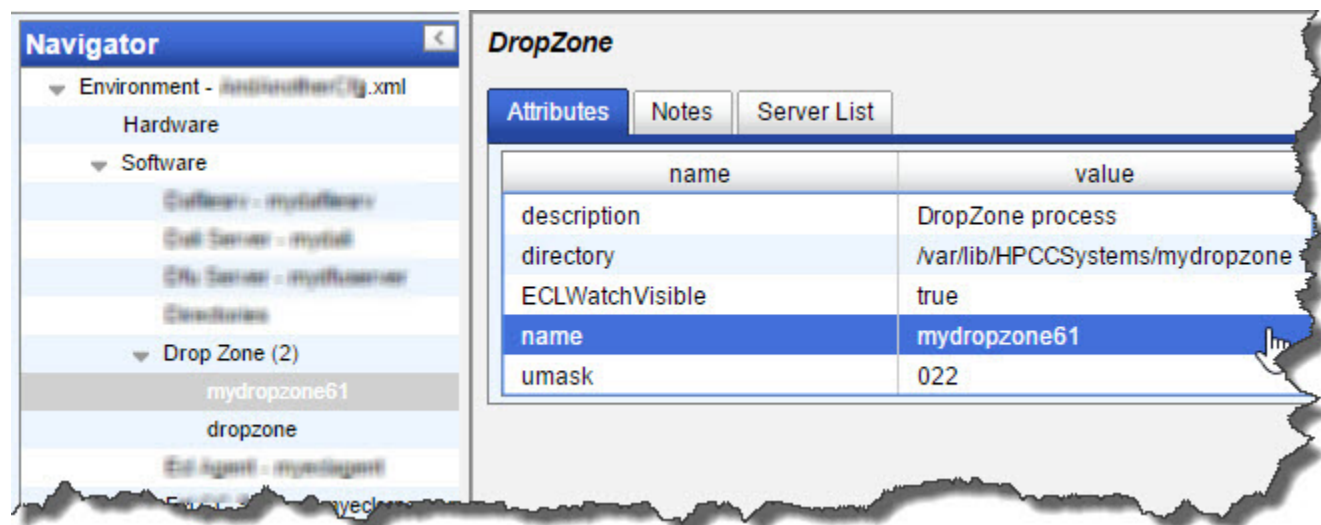
You can change the configuration of your drop zone using the attributes tab. If you have multiple drop zones, select the drop zone to configure from the Navigator panel on the left side.

To change the drop zone attributes:

1. On the **Attributes** tab, select the Attribute to modify.
2. Double-click on the value on the right side of the attribute table for the value you wish to modify.

For example, select the **name** attribute, double click on the **value** column and provide the drop zone with a more meaningful name.

3. Click the disk icon to save.



attribute	description	values	required
name	Name for this process	dropzone	required
description	Description for this process	DropZone process	optional
directory	The directory where the drop zone is located.	/var/lib/HPCCSystems/mydropzone	required

attribute	description	values	required
umask	umask for desprayed files	022	optional
ECLWatchVisible	Allow access from ECLWatch UI	true	optional

You can supply a list of preferred dropzone destinations and keep some destinations hidden using the *ECLWatchVisible* attribute. When set to true, a specified landing zone will be shown in the dropdown list of available dropzones in ECL Watch. When the *ECLWatchVisible* attribute is set to false, the dropzone is hidden. This may be useful in situations where you want to maintain multiple dropzones but hide and protect some locations from being used by certain users or groups. You can use the Notes tab to record the details about your dropzone configurations.

Drop Zone Server List

This tab allows you to add any servers that you wish to configure as a part of the selected drop zone.

To add a server to the current drop zone:

1. Select the **Drop Zone** to configure from the Navigator panel on the left side.
2. Select the **Server List** tab, right-click on the Server Address field and choose **Add**.
3. Enter the hostname or IP address of the server.
4. Click the disk icon to save.

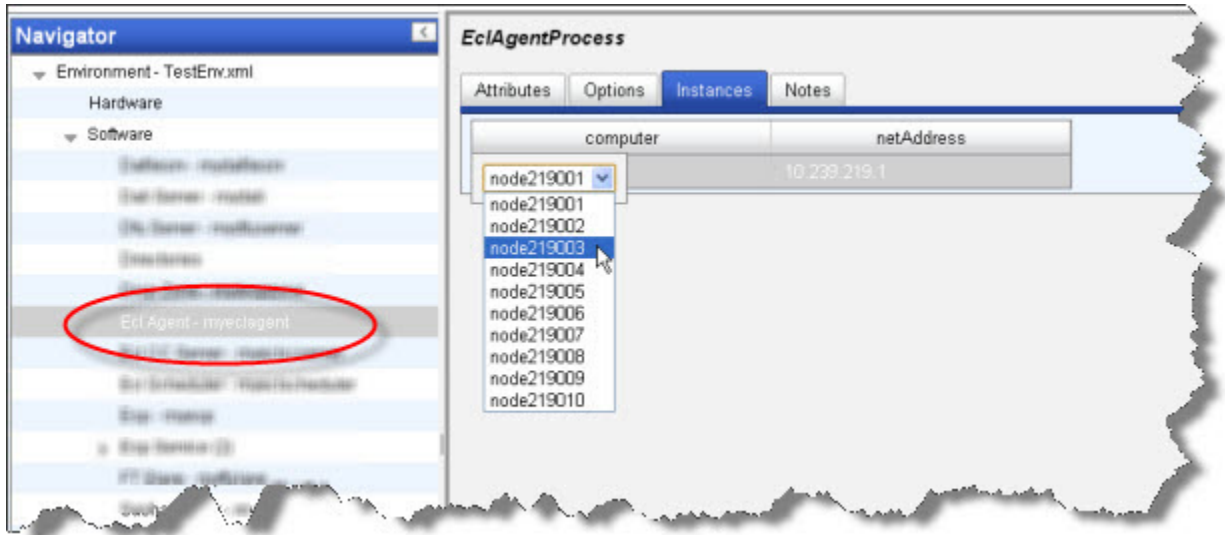
Drop Zone Notes

This tab allows you to add any notes pertinent to the component's configuration. This can be useful to keep a record of changes and to communicate this information to peers.

ECL Agent

Instances

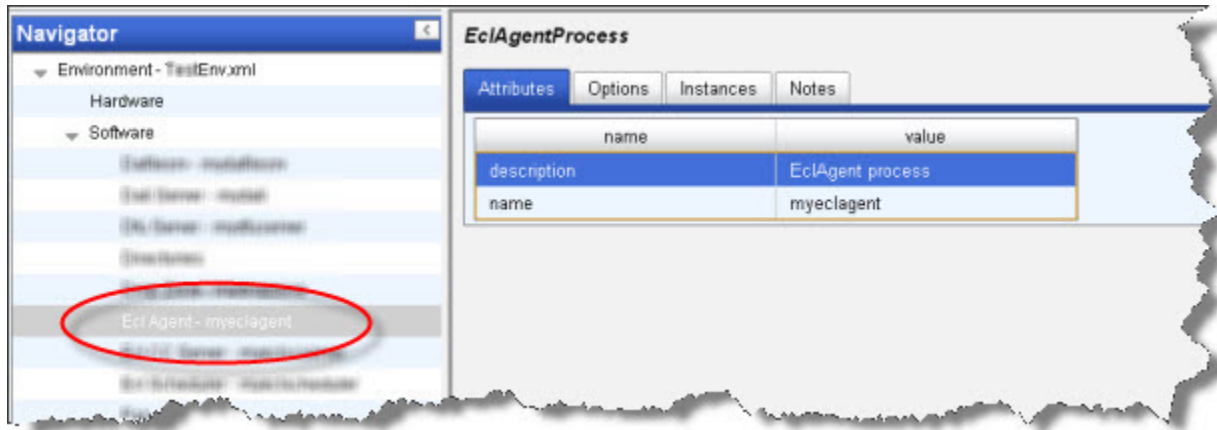
1. Select ECL Agent in the Navigator panel on the left side.
2. Select the Instances tab.
3. In the computer column, choose a node from the drop list as shown below:



4. Click the  disk icon to save

EclAgent Attributes Tab

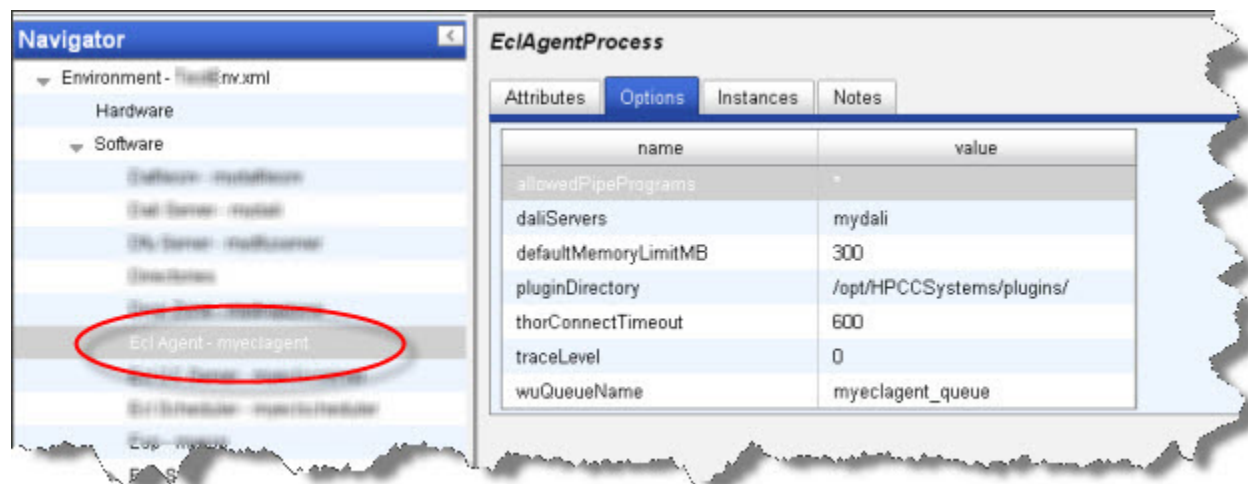
This section describes the EclAgent Attributes tab.



attribute	description	values	required
name	Name for this process		required
description	Description for this process	EclAgent process	optional

EclAgent Options Tab

This section describes the EclAgent Options tab.



attribute	description	values	required
allowedPipePrograms	Comma separated list of allowed PIPE programs (* for allow all)	*	optional
daliServers	Specifies the dali server to which this eclagent is attached.		required
defaultMemoryLimitMB	Default memory limit in MB for eclagent	300	optional
fieldTranslationEnabled	Enables translation (where possible) of mismatched file layouts on-the-fly. Specify 'payload' to attempt to translate payload fields only	payload-RemoveOnly	optional
heapUseHugePages	Use memory from huge pages if they have been configured.	false	optional
heapUseTransparentHugePages	Use memory from transparent huge pages.	true	optional
heapRetainMemory	Retain and do not return unused memory to the operating system.	false	optional
pluginDirectory	Directory where plugins are located	/opt/HPCCSystems/plugins/	optional
traceLevel	Trace level	0	optional
thorConnectTimeout	Default connection timeout when sending query to Thor	600	optional
wuQueueName	eclAgent Workunit Execution Queue Name		optional
analyzeWorkunit	eclAgent analyzes workunit post-execution	true	optional
httpCallerIdHeader	HTTP Header field to use for sending and receiving CallerId	HPCC-Caller-Id	optional

attribute	description	values	required
httpGlobalIdHeader	HTTP Header field to use for sending and receiving GlobalId	HPCC-Global-Id	optional
mapHttpCallUrlsToSecrets	In SOAPCALL and HTTPCALL check if URLs have been mapped to secrets	false	optional
warnIfUrlNotMappedToSecret	In SOAPCALL and HTTPCALL warn if URLs not mapped to secrets	false	optional
requireUrlsMappedToSecrets	Require SOAPCALL and HTTPCALL URLs are secrets or mapped to secrets	false	optional

Allowed Pipe Programs

In version 9.2.0 and greater, commands used in a PIPE action are restricted by default. However, for legacy reasons, the default stock behavior is different in bare-metal and containerized deployments. In both types of systems, if `allowedPipePrograms` is unset, then all but "built-in" programs are restricted (The only built-in program currently is 'roxiepipe').

In bare-metal, the default `environment.xml` includes a setting value of "*" for **`allowedPipePrograms`**. This means (by default) any PIPE program can be used.

In a secure system, this should be removed or edited to avoid arbitrary programs, including system programs, from being executed.

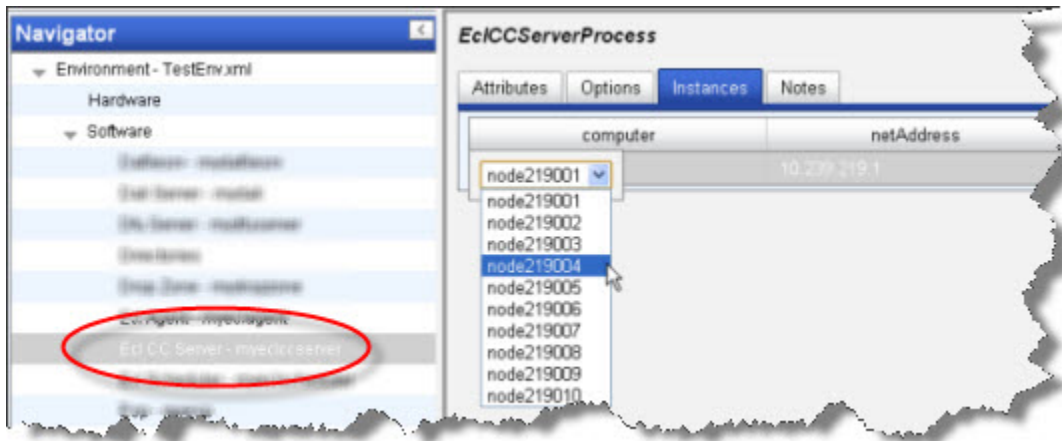
EclAgentProcessNotes

This tab allows you to add any notes pertinent to the component's configuration. This can be useful to keep a record of changes and to communicate this information to peers.

ECL CC Server Process

Ecl CC Server Instances

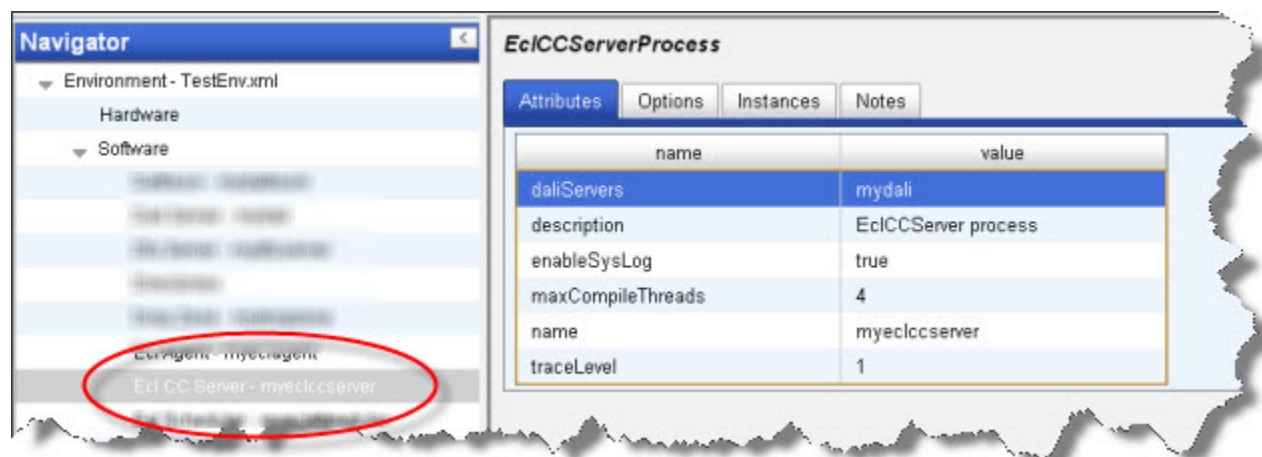
1. Select Ecl CC Server - myeclccserver in the Navigator panel on the left side.
2. Select the Instances tab.
3. In the computer column, choose a node from the drop list as shown below:



4.  Click the disk icon to save

Ecl CC Server Attributes Tab

This section describes the Ecl CC Server Attributes tab.



attribute	description	values	required
name	Name for this process	ecclccserver	required
description	Description for this process	EclCCServer process	optional
daliServers	Specifies the dali server to which this ecclccserver is attached.		required
enableEclccDali	Enables passing Dali address to eclcc for compile-time file resolution.	true	optional
enableSysLog	Enables syslog monitoring of the eclcc-server process.	true	optional
generatePrecompiledHeader	Generate precompiled header when ecclccserver starts.	true	optional
traceLevel		1	optional
maxEclccProcesses	Maximum number of instances of eclcc that will be launched in parallel.	4	optional
monitorInterval	Interval (in seconds) for reporting of memory usage stats. Set to 0 to disable.	60	optional

EclCC Server Process Options

To add a custom option, right-click and select add. These options are passed to the eclcc compiler.

See the ECL Compiler chapter in the [Client Tools](#) manual for details.

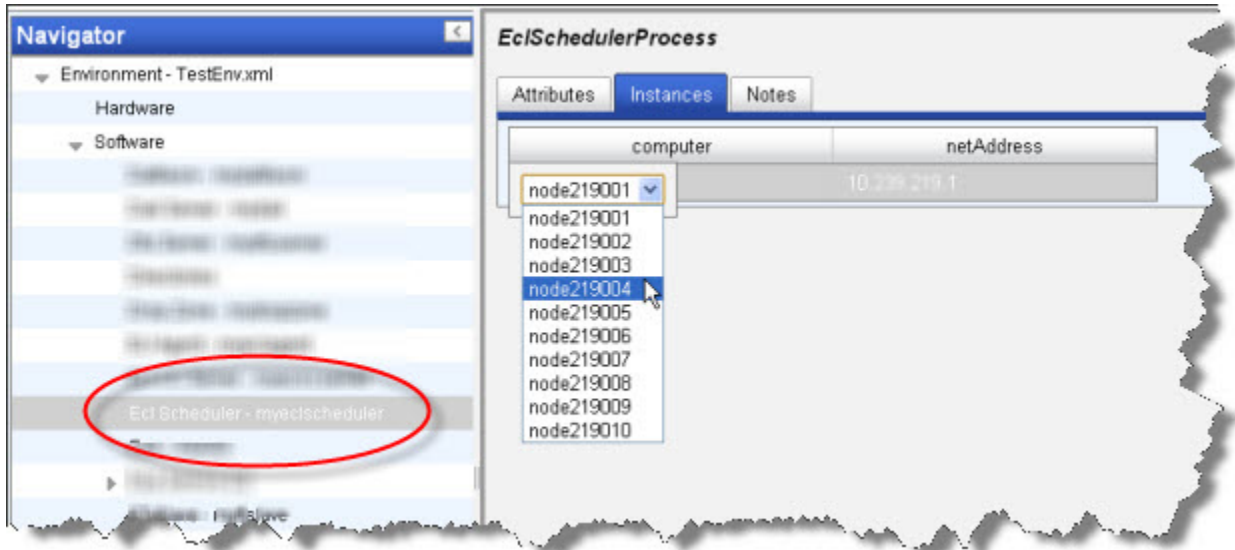
EclCC Server Process Notes


This tab allows you to add any notes pertinent to the component's configuration. This can be useful to keep a record of changes and to communicate this information to peers.

ECL Scheduler

Instances

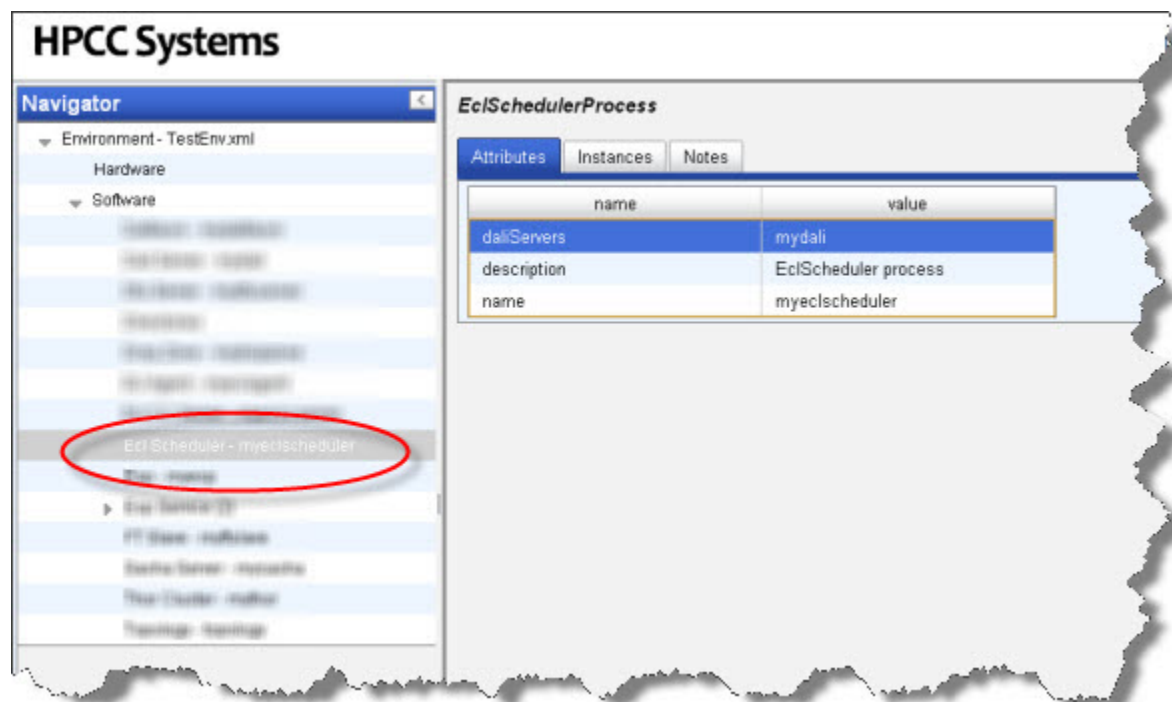
1. Select **ECL Scheduler** in the Navigator panel on the left side.
2. Select the Instances tab.
3. In the computer column, choose a node from the drop list as shown below:



4. Click the  disk icon to save

EclScheduler Attributes Tab

This section describes the EclScheduler Attributes tab.



attribute	description	values	required
name	Name for this process	eclscheduler	required
description	Description for this process	EclScheduler process	optional
daliServers	Specifies the dali server to which this eclscheduler is attached.		required

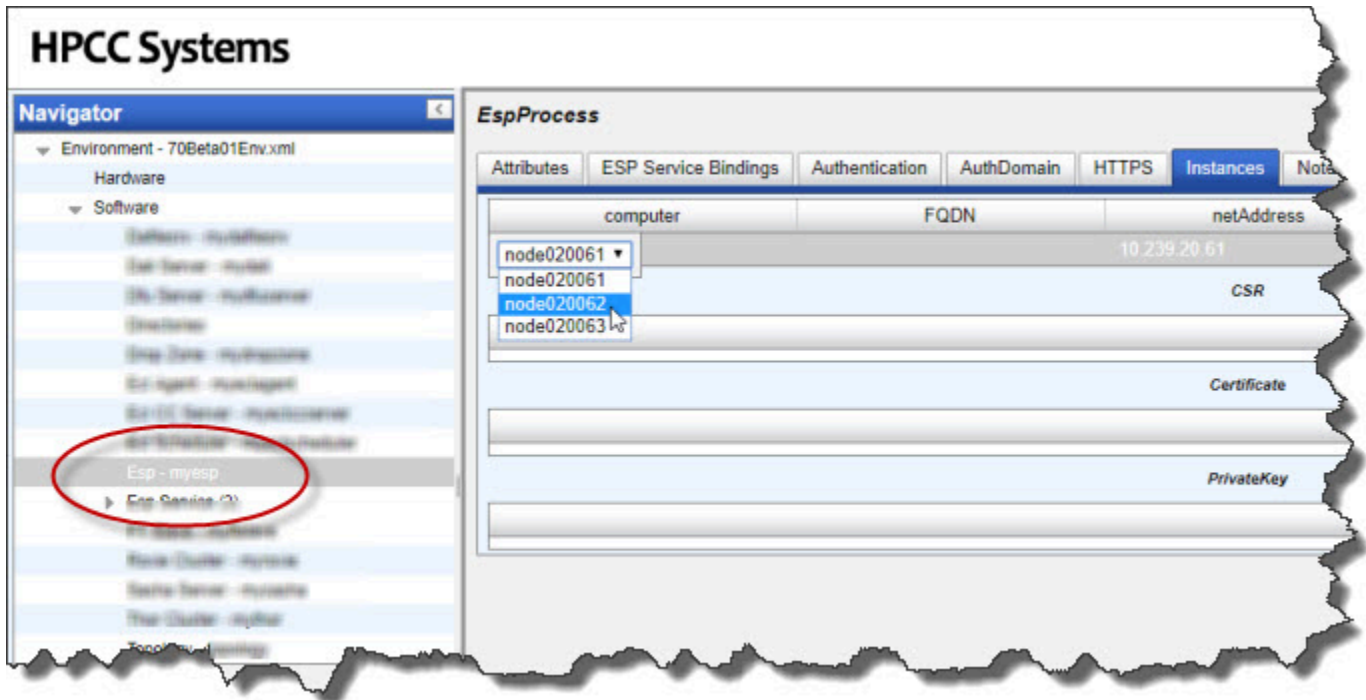
EclScheduler Notes

This tab allows you to add any notes pertinent to the component's configuration. This can be useful to keep a record of changes and to communicate this information to peers.

ESP Server

Esp Process Instances

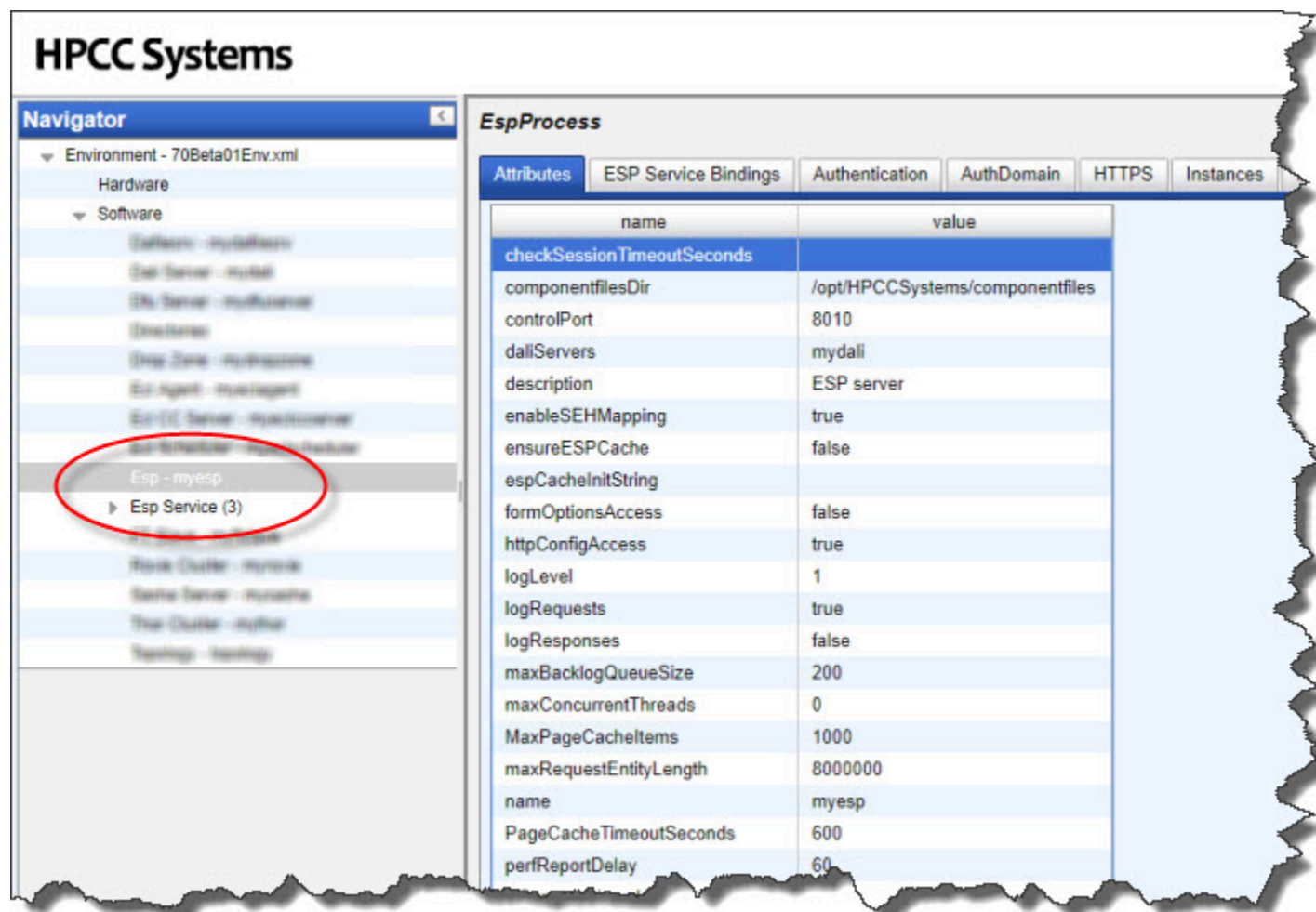
1. Select **ESP - myesp** in the Navigator panel on the left side.
2. Select the Instances tab.
3. In the computer column, choose a node from the drop list as shown below:



4. Click the  disk icon to save

Esp - myesp Attributes Tab

This section describes the Esp - myesp Attributes tab.



attribute	description	values	required
name	Name for this process		required
description	Description for this process	ESP server	optional
daliServers	Specifies the dali server to which this ESP is attached.		required
bindings	Path to esdl service bundles to load		optional
enableSEHMapping	Enables SEH to exception mapping.	true	optional
enableSysLog	Enables SysLog logging.	false	optional
httpConfigAccess	Allows esp config file to be viewed via a web browser.	true	optional
formOptionsAccess	Allows show Options in test form page.	false	optional
maxLogFileSize	The maximum log file size (0 for no size limit)	0	optional

Configuring the HPCC Systems® Platform
Configuration Manager Advanced View

attribute	description	values	required
maxRequestEntityLength	The maximum length of request entity allowed	8000000	optional
maxConcurrentThreads	The maximum number of concurrent threads. 0 means unlimited.	0	optional
maxBacklogQueueSize	Sets the sockets parameter for the maximum number of backlogged requests.	200	optional
perfReportDelay	Sets the frequency for logging resource usage stats.	60	optional
controlPort	Sets the network port for ESP control	8010	optional
logLevel	Sets the log level [0: none, 1: min, 5: noraml, 10: max]	1	optional
componentfilesDir	Sets the componentfiles directory.	/opt/HPCCSystems/componentfiles	optional
logRequests	never only-ones-with-issues all	never	optional
logResponses		false	optional
txSummaryLevel	Sets the TxSummary level [0: none, 1: min, 5: normal, 10: max]	1	optional
txSummaryStyle	Sets the style of TxSummary: [text json all]	text	optional
txSummaryGroup	Selects which groups of entries are output in the TxSummary: [core enterprise all]	core	optional
txSummaryResourceReq	Log TxSummary for Resource Requests	false	optional
espCacheInitString	String used for initializing ESP cache client.		optional
checkSessionTimeoutSeconds	Check Session Timeout in every given seconds.		optional
urlAlias	The URL alias for this ESP. This can be used to detect Cross-Origin Resource Sharing (CORS) access.		optional
PageCacheTimeoutSeconds	Paging data timeout in the given seconds.	600	optional
MaxPageCacheItems	The maximum number of cached items inside one page cache. 0 means no cache.	1000	optional
maxPersistentIdleTime	The maximum idle time in seconds (-1 for unlimited, 0 to disable).	60	optional
maxPersistentRequests	Maximum number of query requests per persistent http connection. (-1 for unlimited, 0 to disable).	100	optional

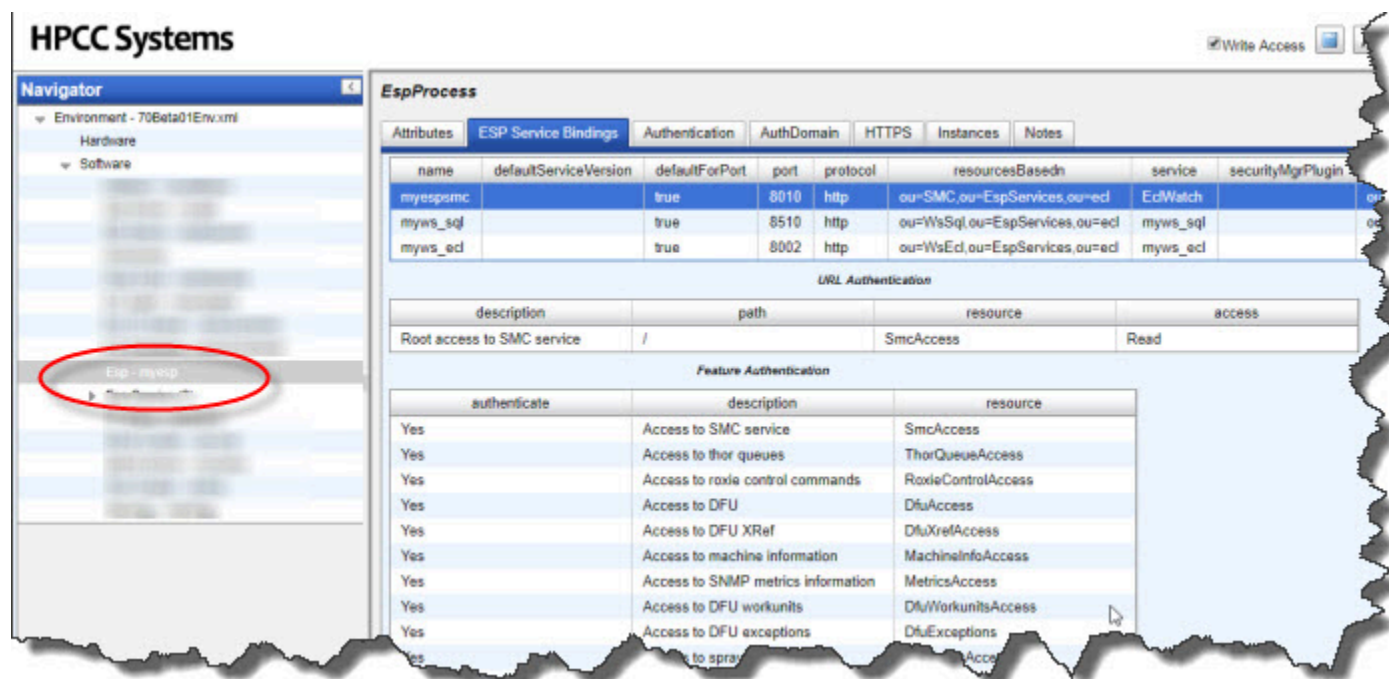
Configuring the HPCC Systems® Platform
Configuration Manager Advanced View

attribute	description	values	required
minCompressLength	Minumum content length in bytes for the content to be compressed.	1000	optional
namespaceScheme	The scheme used to generate dynamic service namespaces.	basic	optional
MaxFileAccessExpirySeconds	Maximum time a file access request is valid for	86400	optional
dafilesrvConnectTimeout	Remote file access connect timeout in seconds	10	optional
dafilesrvReadTimeout	Remote file access read timeout in seconds	10	optional

Persistent HTTP connections can cause backwards compatibility issues with non-standard HTTP clients. You can disable this feature by setting *maxPersistentRequests* to 0.

Esp - myesp Service BindingsTab

This section describes the Esp - myesp Service Bindings tab. This tab requires additional steps to configure the service bindings.



You must first add the service bindings in the first table (Right-click, add). Then you would configure the attributes in the other tables on that tab. The next table describes the **URL Authentication** table.

attribute	description	values	required
description			optional
path	The logical path of a resource used for authentication.	/	required
resource	The physical resource for which access is checked.		required
access	The access level required to the sepci-fied resource.	Read	optional

The following tables describe the ESPProcess Service Bindings, **Feature Authentications**.

attribute	description	values	required
authenticate	Validate access rights for this capability?	Yes	optional
description			optional
path	The logical path of a resource used for authentication.	/	required
resource	The physical resource for which access is checked.		required

To add feature access flags to a preexisting ECL Watch configuration:

If you update the platform, but are using a preexisting configuration, you could encounter a situation where Feature level access flags are not automatically created. Missing flags can deny access to users trying to access features in the system.

1. In Configuration Manager, open your working copy of the environment.xml and enable *Write Access*.
2. On the left, select the ESP that is hosting the ECL Watch service.
3. On the right, select the *ESP Service Bindings* tab.

Manually add the access flag for the new service in the Feature Authentication table.

4. Right-click on the Feature Authentication table, then select *Add*

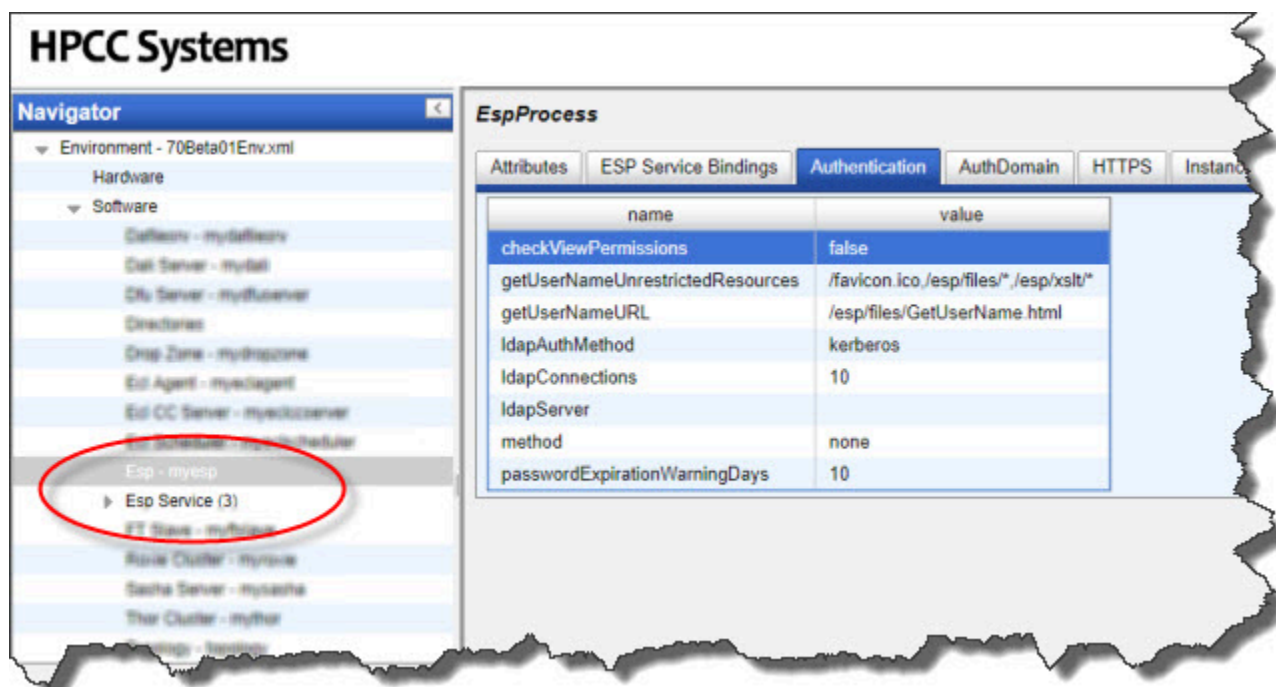
The screenshot shows the HPCC Configuration Manager interface. On the left, the 'Navigator' pane shows a tree view of the environment. The 'Esp - myesp' entry is selected and circled in red. On the right, the 'EspProcess' configuration window is open. The 'ESP Service Bindings' tab is selected and circled in red. It displays a table with columns: name, defaultServiceVersion, defaultForPort, port, protocol, and ou. Below this, there are sections for 'URL Authentication', 'Secure User Settings', and 'Feature Authentication'. The 'Feature Authentication' table has columns: authenticate, description, path, and resource. A right-click context menu is open over the 'Feature Authentication' table, with the 'Add' option highlighted and circled in red.

5. Provide the values for path and resource (usually the same value, for example, WsStoreAccess).
6. Click the disk icon to save
7. Copy the file to every node, then restart the system.

The OUs should get created in LDAP automatically upon restart.

Esp - myesp AuthenticationTab

This section describes the Esp - myesp Service Authentication tab.



attribute	description	values	required
method	The protocol to use for authenticating the service	none	required
ldapServer	The ldap server to be used for authentication.		optional
ldapConnections	The maximum number of connections to the LDAP server.	10	optional
passwordExpirationWarningDays	In this time period, ESP displays a warning about password expiration.	10	optional
checkViewPermissions	Enable file and column access permission checking for all view enabled queries	false	optional
loginLogoURL	URL to Login Logo	/esp/files/eclwatch/img/Loginlogo.png	optional
getUserNameURL	URL to getUserName	/esp/files/GetUserName.html	optional
getUserNameUnrestrictedResources	unrestricted resources for getUserNameURL	/favicon.ico,/esp/files/*,/esp/xslt/*	optional

Additional information about the available Authentication methods:

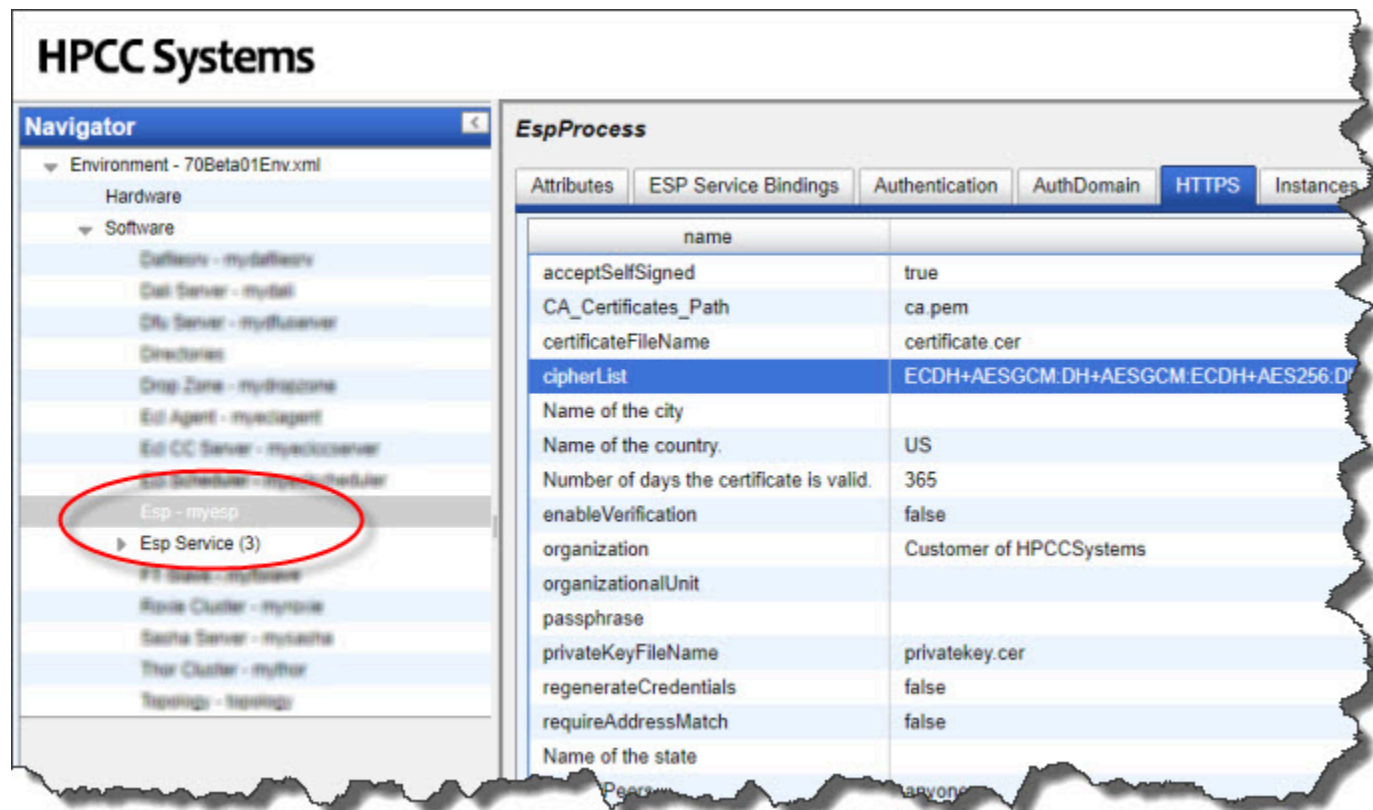
none	uses no authentication
ldap	uses Lightweight Directory Access Protocol for authentication
ldaps	similar to LDAP but uses a more secure (TLS) protocol
secmgrPlugin	uses the security manager plug-in

Esp - AuthDomain Tab

The **AuthDomain** attribute allows you to configure the settings used for session management.

Esp - myesp HTTPS Tab

This section describes the Esp - myesp HTTPS tab.



The **cipherList** attribute allows you to set the ordered list of available ciphers for use by openssl. See the documentation at [openssl.org](https://www.openssl.org) for more information about ciphers.

attribute	description	values	required
cipherList	Sets the ordered list of available ciphers for use by openssl. See openssl documentation on ciphers for information on use and formatting.	ECDH +AESGCM:DH +AESGCM:ECDH +AES256:DH +AES256:ECDH +AES128:DH +AES:ECDH +3DES:DH +3DES:RSA +AESGCM:RSA +AES:RSA +3DES:!aNULL:!MD5	optional

attribute	description	values	required
certificateFileName	Name of destination file in which the certificate will be written.	/home/hpcc/certificate/certificate.pem	optional
privateKeyFileName	Name of destination file in which the private key will be written.	/home/hpcc/certificate/key.pem	optional
passphrase	The passphrase used to generate the private key.		optional
enableVerification	whether to enable verification or not.	false	optional
requireAddressMatch	whether to require the source address of the request match that of the certificate.	false	optional
acceptSelfSigned	whether to accept self-signed certificates.	true	optional
CA_Certificates_Path	path to the file that contains CA certificates.	/home/hpcc/certificate	optional
trustedPeers	List of trusted peers, in smartsocket format.	anyone	optional
organization	Name of company or organization	Customer of HPC- CSystems	optional
organizationalUnit			optional
city			optional
state			optional
country		US	optional
daysValid		365	optional
regenerateCredentials	Set this to true to regenerate the private key, certificate and CSR.	false	optional

EspProcess Notes

This tab allows you to add any notes pertinent to the component's configuration. This can be useful to keep a record of changes and to communicate this information to peers.

ESP Services

ESP Services provide a means to add functionality to an ESP Server.

ECL Watch Service

Ecl Watch allows you to configure options for the ECL Watch utility.



ECL Watch Attribute definitions

attribute	description	values	required
name	Name for this ESP service	espsmc	required
description	Description for this process	ESP services for SMC	optional
syntaxCheckQueue	Queue Name of ECL server which is used for ECL Syntax Check		optional
pluginsPath	Path where plugin files are deployed	/opt/HPCCSystems/plugins	optional
viewTimeout	timeout for XXXX (in seconds).	1000	optional

Configuring the HPCC Systems® Platform
Configuration Manager Advanced View

attribute	description	values	required
clusterQueryStateThread-PoolSize	Default thread pool size for checking query state on clusters	25	optional
AWUsCacheTimeout	timeout for archived WU search cache (in minutes).	15	optional
NodeGroupCacheMinutes	timeout for node group cache (in minutes).	30	optional
ActivityInfoCacheSeconds	If user makes a request and the cache is older than this value (in seconds), ESP returns the cache and starts to rebuild a new cache.	10	optional
ActivityInfoCacheAutoRebuildSeconds	When idle, ESP automatically rebuilds the cache if this value (in seconds) exceeded.	120	optional
MachineUsageCacheMinutes	timeout for Component/TargetCluster/NodeGroup usage cache (in minutes).	3	optional
MachineUsageCacheAutoRebuildMinutes	When idle, ESP automatically rebuilds the cache if this value (in minutes) is exceeded.	10	optional
serverForArchivedECLWU	Specify Sasha server for archiving ECL workunits		optional
enableSystemUseRewrite	To disable ESP Service links for System Servers that use rewrite rules.	false	optional
defaultTargetCluster	Default target for published queries		optional
disableUppercaseTranslation	To disable upper case translation for filter values in ViewKeyFile function.	false	optional
ZAPEmailTo	Default email address to receive ZAP Report email		optional
ZAPEmailFrom	Default email address to send ZAP Report email from		optional
ZAPEmailServer	Default Email Server URL to handle ZAP Report email		optional
ZAPEmailServerPort	Default Email Server port to handle ZAP Report email		optional
ZAPEmailMaxAttachmentSize	Maximal attachment size in ZAP Report email		optional
ThorSlaveLogThreadPoolSize	Thread pool size for getting thor slave logs.		optional
WUResultMaxSizeMB	the size limit in megabyte for getting workunit result.		optional
WUResultDownloadFlushThreshold	When streamly downloading WUResult, WUResult buffer will be flushed out if its size reaches this value.		optional
enableLogDaliConnection	Enable ESP/Dali Connection ID to be logged into esp.xml.	false	optional

attribute	description	values	required
allowNewRoxieOnDemandQuery	allow new queries to be used by roxie on demand and roxie browser	false	optional

ECL Watch Monitoring attributes

attribute	description	values	required
monitorDaliFileServer	Warn if dafilesrv process is not running on computers	false	required
excludePartitions	Comma, space or semicolon delimited list of partitions not to be monitored for free space	/dev*,/sys,/proc/*	optional
warnIfCpuLoadOver	CPU load over this value is flagged as warning in monitoring output	95	optional
warnIfFreeStorageUnder	Available disk storage space under this value is flagged as warning in monitoring output	5	optional
warnIfFreeMemoryUnder	Available memory under this value is flagged as warning in monitoring output	5	optional

Configuring ELK Log Visualization integration

HPCC Systems provides a mechanism to integrate ELK (ElasticSearch, Logstash, and Kibana) based visualizations into ECL Watch. This feature can be configured and enabled through the HPCC Systems Configuration manager.

To configure the ELK Log Visualization component, click and expand the **ESP Service** link on the left side then select the **EclWatch** link. Then select the **Elk Log Visualization** tab from the right side.

The ELK Log Visualization tab in ECL Watch:

attribute	description	values	required
integrateKibana	Declare if Kibana dashboard is to be integrated in ECLWatch	false	required
kibanaAddress	Address (ip/hostname) of Kibana server	localhost	optional
kibanaPort	Port number of target Kibana server	5601	optional
kibanaEntryPointURI	Declares the ECLWatch integration entry point URI for the target Kibana dashboard		optional
reportLogStashHealth	Declare if connectivity to LogStash nodes is to be reported in ECLWatch	true	required
logStashAddresses	Declares the target LogStash nodes' addresses (semicolon separated)		optional
logStashPort	Declares the target LogStash nodes' port		optional

Configuring the HPCC Systems® Platform
Configuration Manager Advanced View

attribute	description	values	required
reportElasticHealth	Declare if connectivity to Elastic Search nodes is to be reported in ECLWatch	true	required
elasticSearchAddresses	Declares the target ElasticSearch nodes' addresses (semicolon separated)		optional
elasticSearchPort	Declares the target LogStash nodes' port	9200	optional

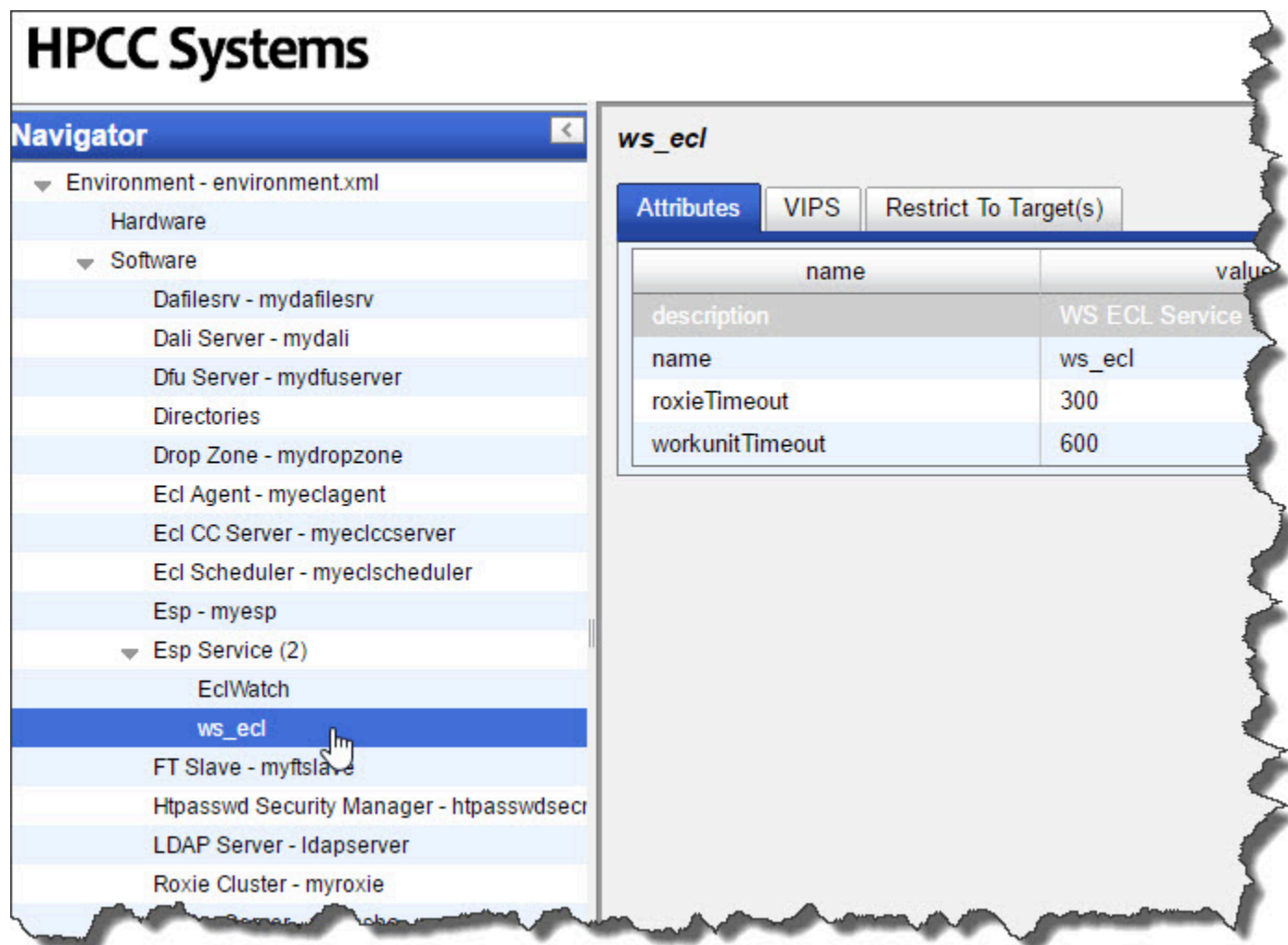
To configure the ELK visualization integration, provide the following information:

1. In the *kibanaAddress* field, supply the IP address for your Kibana component.. For example:
http://123.123.123.123
2. In the *kibanaPort* field, supply the Kibana port number. For example: 5601
3. In the *kibanaEntryPointURI* field, enter the Kibana URI pointing to your shared Kibana dashboard.

For example: /app/kibana#/dashboard/etc. (obtained from the "share" tab in Kibana)

WsECL Service

The WsECL service allows you to configure options for the WsECL utility.



The Ws ECL configuration attributes.

attribute	description	values	required
roxie			required
vip			required
dnsInterval	DNS lookup cache timeout in seconds. Set to 0 to resolve DNS for every transaction. Set to -1 (default) to keep DNS lookup cached indefinitely.	-1	optional
sendTargetToRoxie	Send roxie the target from which to run query (disable for backward compatibility issues)	true	optional
tls	Use TLS	false	optional

Ws ECL VIPS option attributes.

attribute	description	values	required
name	WsEcl will only display specified targets, if none specified WsEcl will display all targets.		required

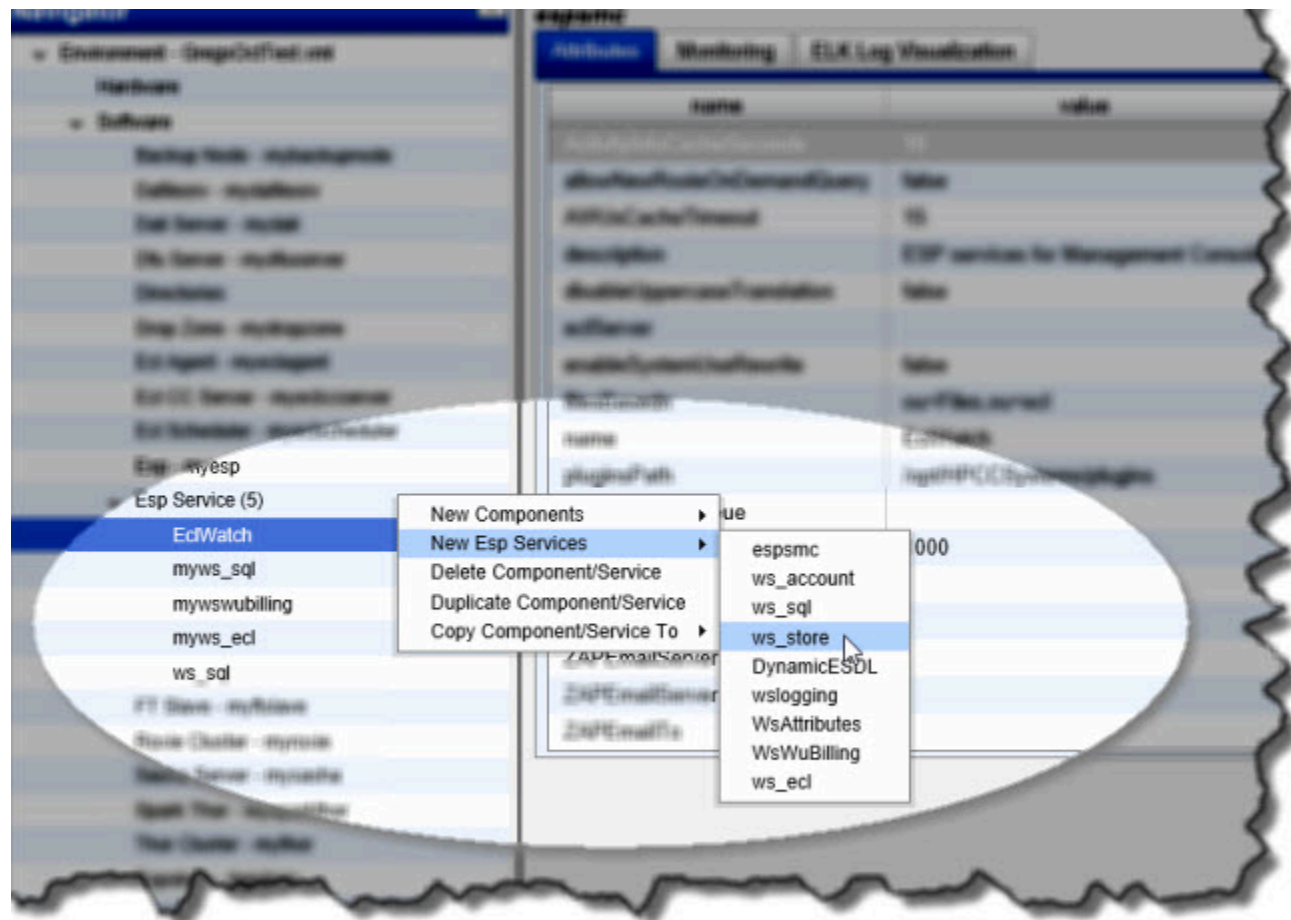
Ws ECL Target Restrictions table.


attribute	description	values	required
name	Name for this ESP service	ws_ecl_service	required
description	Allows creation of web services using ECL language	WS ECL Service	optional
roxieTimeout	Timeout (in seconds) for WsEcl connections to roxie (0 == wait forever)	300	optional
workunitTimeout	Timeout (in seconds), for WsEcl to wait for workunit to complete (0 == wait forever)	600	optional
xsltMaxDepth	Maximum libxslt stylesheet template depth (affects size and complexity of dataset that can be rendered)	100000	optional
xsltMaxVars	Maximum libxslt stylesheet template variables (affects size and complexity of dataset that can be rendered)	1000000	optional
httpCallerIdHeader	HTTP Header field to use for sending and receiving CallerId	HPCC-Caller-Id	optional
httpGlobalIdHeader	HTTP Header field to use for sending and receiving GlobalId	HPCC-Global-Id	optional

Ws_Store

One instance of ws_store is included with ECL Watch, but you can choose to add another ws_store ESP Service to your HPCC Systems platform.

To add the ws_store service.



1. Right click on the **Software** component in the **Navigator** panel (on the right side), choose **New ESP Services** then **ws_store** from the drop list.
2. Configure the ws_store attributes as required.
3.  Click the disk icon to save


The following values are the configurable attributes for ws_store:

attribute	description	values	required
name	Name for this WsStore instance	ws_store_service	required
description	Dali based WsStore		optional
StoreProviderLibraryName	The store provide library name lib'providername'.so		required
StoreProviderFactoryMethod	Method used to instantiate the store provider - default: newEspStore		optional

Ws_SQL

You can choose to add the ws_sql ESP Service to your HPCC Systems platform.

To add the ws_sql service:

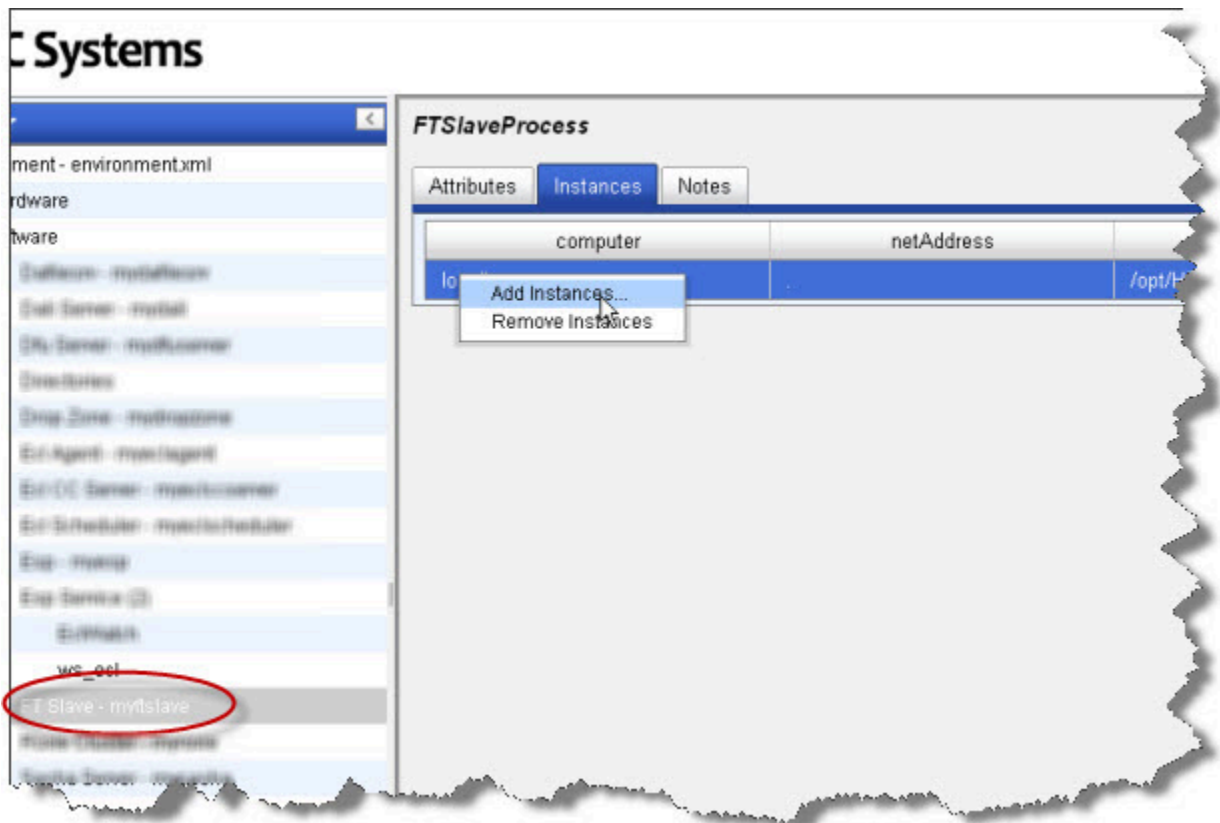
1. Right click on the **Software** component in the **Navigator** panel (on the right side), choose **New ESP Services** then **ws_sql** from the drop list.
2. Configure the ws_sql by providing the name of the service.
3.
Click the  disk icon to save

FTSlave Process

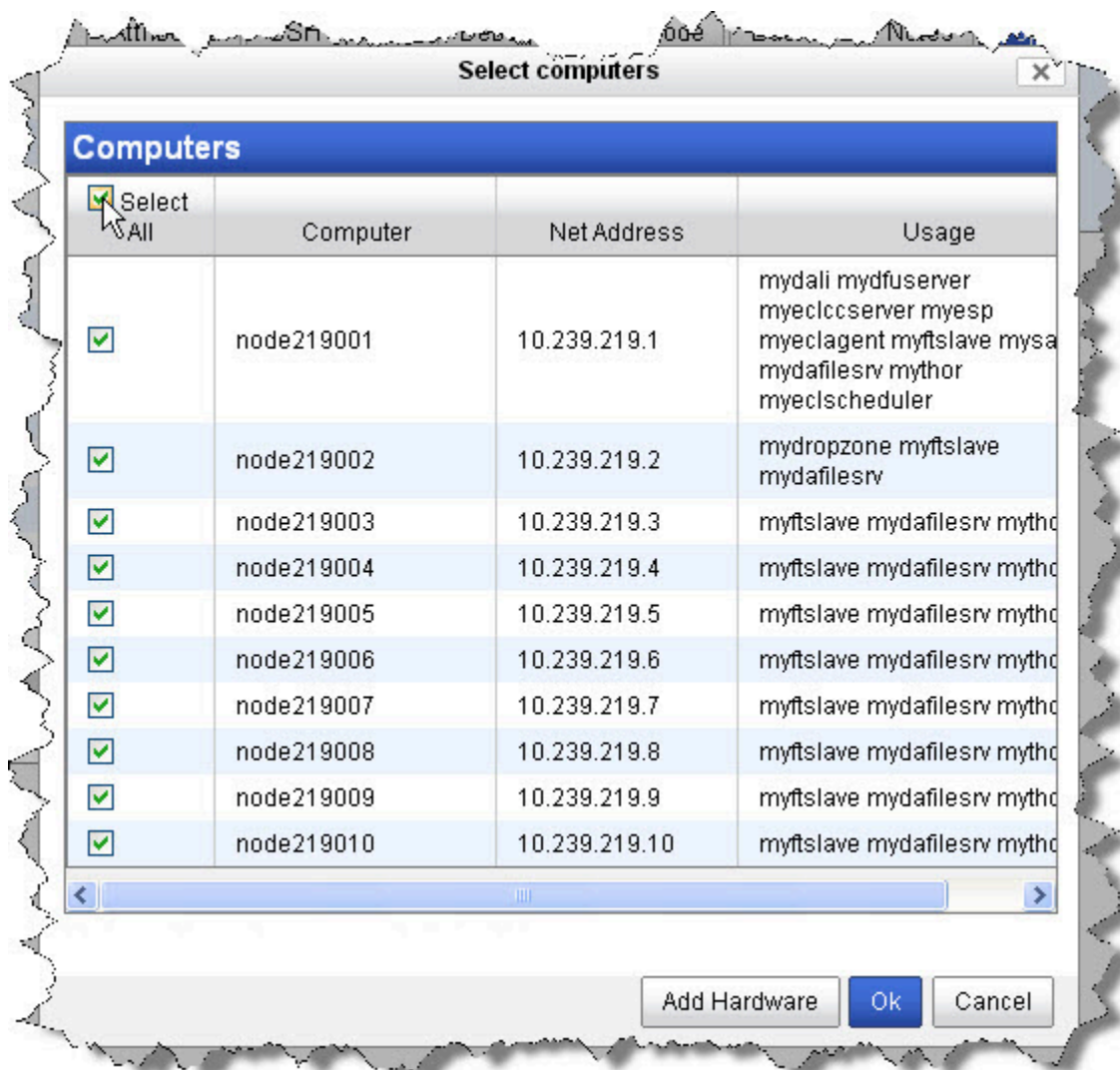
FTSlave is a helper process that every node needs. This section depicts an FTSlave installation.

Instances

1. Select FTSlave in the Navigator panel on the left side.
2. Select the Instances tab.
3. Right-click on a computer in the computer column, and select Add Instance.



4. Select all computers in the list, then press the **OK** button.



5. Click the  disk icon to save

FtSlave attributes

This section describes an FTSlaveProcess attributes tab.

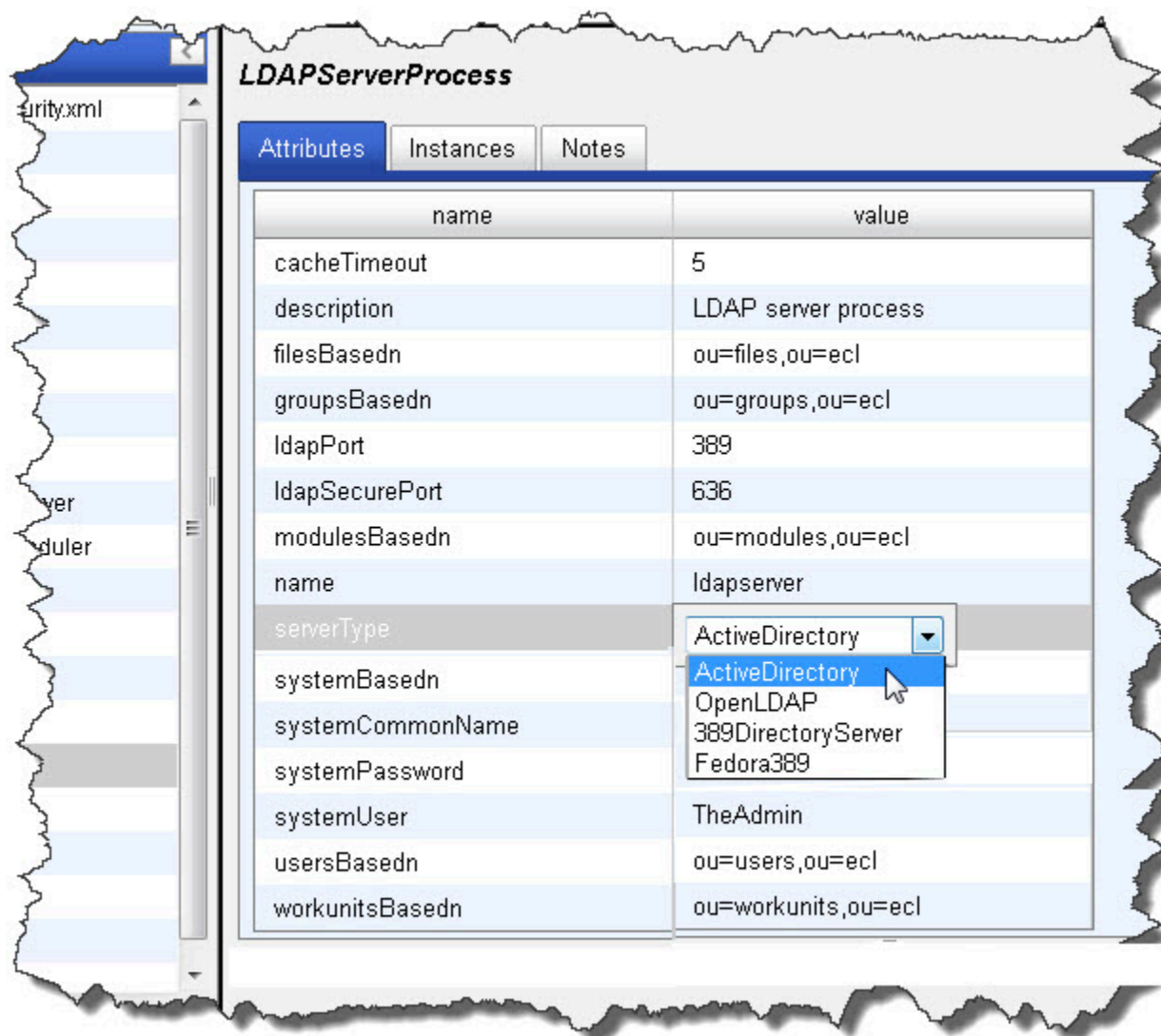
attribute	description	values	required
name	Name for this process		required
description	Description for this process	FTSlave process	optional
version	Version identifier used to select which process will be started	1	optional

FtSlave Process Notes

This tab allows you to add any notes pertinent to the component's configuration. This can be useful to keep a record of changes and to communicate this information to peers.

LDAP Server Process

This section describes the configuration attributes of an LDAPServer Installation in ConfigManager. For a complete description of how to add LDAP Authentication see *Using LDAP Authentication* section in the [Installing and Running The HPCC Systems Platform](#) document.



attribute	description	values	required
name	Name for this process		required
description	Description for this process	LDAP server process	optional
adminGroupName	The Active Directory group containing HPCC Administrators	Administrators	optional
ldapPort	The port of the ldap (Active Directory) server.	389	required
ldapSecurePort	The port of the ldap (Active Directory) server.	636	required

attribute	description	values	required
ldapTimeoutSecs	The maximum number of seconds to wait for most LDAP calls.	60	required
ldapCipherSuite	The optional Cipher Suite string to be applied to LDAPS calls.		optional
cacheTimeout	Time in minutes after which the cached security information should expire.	60	optional
sharedCache	Use a single, shared LDAP cache.	true	optional
systemUser	An LDAP administrator account id to be used by HPCC to create and manage HPCC-specific LDAP branches.		optional
systemPassword	The password for the systemUser.		optional
systemCommonName	Required if systemUser is specified. The LDAP Common Name (cn) for the systemUser account as specified on the LDAP server.		optional
systemBasedn	The ldap "base distinguished name" of the systemUser.	cn=Users	required
groupsBasedn	The ldap "base distinguished name" that ecl server should use when looking up groups in the ldap (Active Directory) server.	ou=groups,ou=ecl	required
viewsBasedn	The ldap "base distinguished name" that ecl server should use when looking up views in the ldap (Active Directory) server.	ou=views,ou=ecl	required
usersBasedn	The ldap "base distinguished name" that ecl server should use when looking up users in the ldap (Active Directory) server.	ou=users,ou=ecl	required
modulesBasedn	The ldap "base distinguished name" that ecl server should use when looking up modules in the ldap (Active Directory) server.	ou=modules,ou=ecl	required
workunitsBasedn	The ldap "base distinguished name" that ecl server should use when looking up workunit scopes in the ldap (Active Directory) server.	ou=workunits,ou=ecl	required
filesBasedn	The ldap "base distinguished name" that ecl server should use when looking up file scopes in the ldap (Active Directory) server.	ou=files,ou=ecl	required
serverType	LDAP Server Implementation Type	ActiveDirectory	required

LDAP Server Process Instances

This tab allows you to add instances to your LDAP Configuration. In order to add instances you would have previously added the LDAP computers in the Hardware section. For a complete description of how to add

LDAP Authentication see *Using LDAP Authentication* section in the [Installing and Running The HPCC Systems Platform](#) document.

1. On the **Instances** tab, right-click on the table on the right hand side, choose **Add Instances...**
2. Select the computer to use by checking the box next to it.

This is the computer you added in the **Hardware / Add New Computers** portion earlier.

The HPCC Systems Administrator's account

You can set up an HPCC Systems Administrator's account with limited administrative rights. The HPCC Systems Administrator does not have to be an LDAP Administrator's account. This allows the HPCC Systems Administrator to be able to set up users, groups, and set permissions in HPCC Systems without having rights to perform other LDAP administrative functions. To use this feature:

1. Create an LDAP group to contain all the HPCC Administrator users. For example: "HPCCAdminGroup"
2. In the HPCC Systems configuration manager navigate to this (LDAP Server Process) page and enter the HPCC Systems Administrator group name as the value in the **adminGroupName** field.
3. Add (HPCC Systems Administrator) users to this new group.
4. Save and deploy the new configuration file, then restart ESP to apply the new configuration.

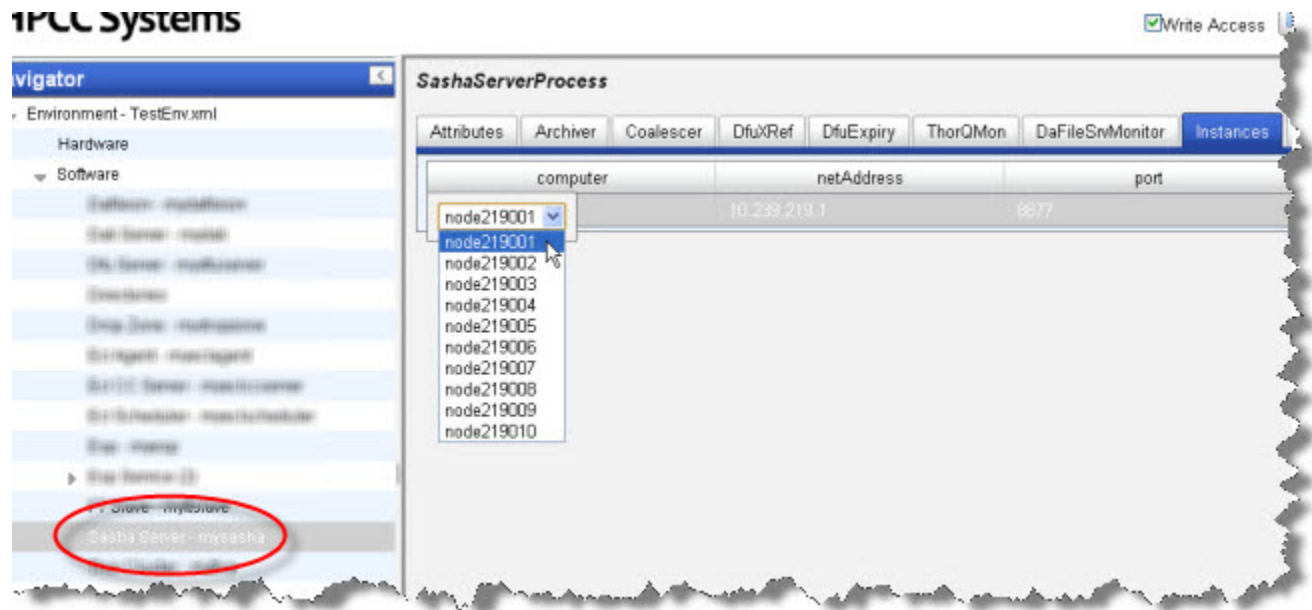
LDAP Server Process Notes

This tab allows you to add any notes pertinent to the component's configuration. This can be useful to keep a record of changes and to communicate this information to peers.

Sasha Server

Instances

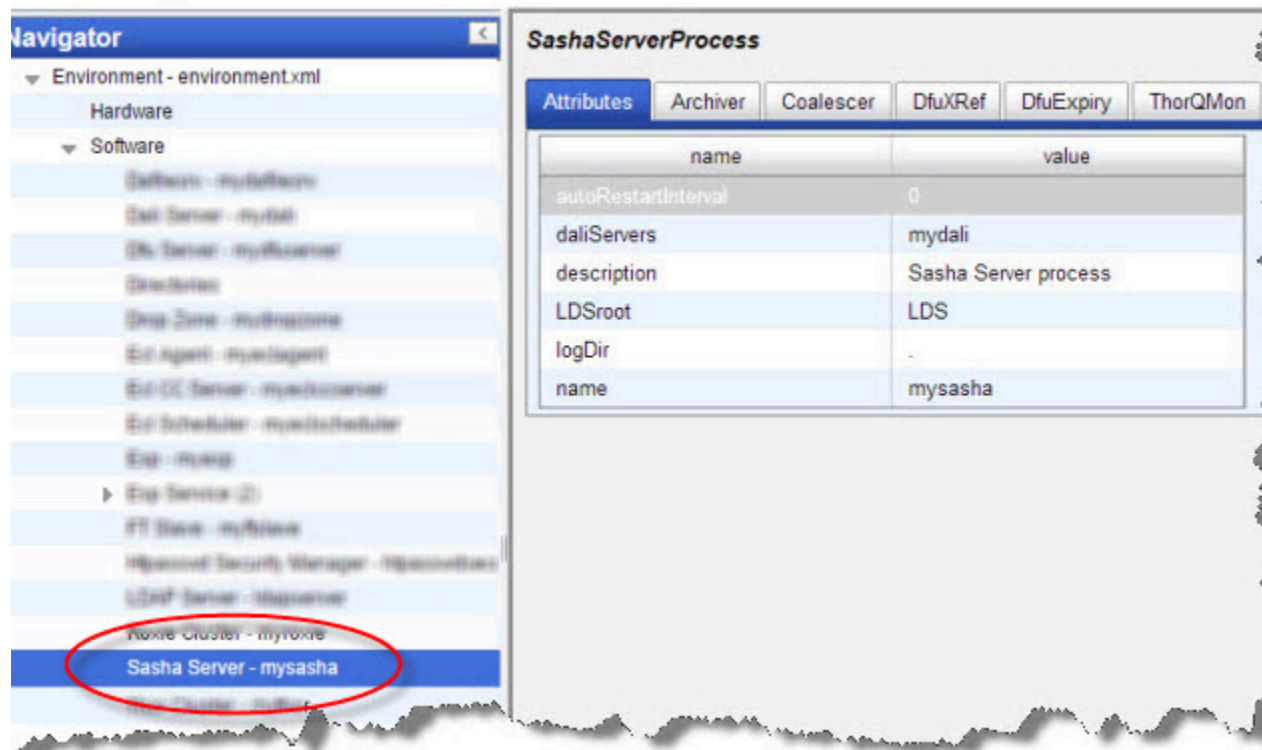
1. Select Sasha Server in the menu on the left side.
2. Select the Instances tab.
3. In the computer column, choose a node from the drop list as shown below:



Sasha Server Attributes

This section described the SashaServerProcess **Attribute** tab values.

HPCC Systems



attribute	description	values	required
name	Name for this process		required
description	Description for this process	Sasha Server process	optional
daliServers	Specifies the Dali server to which this Sasha server is attached.		required
logDir	Specifies the Sasha server log directory.	.	required
autoRestartInterval	Specifies interval between sasha server auto-restart (hours, 0 no auto restart (default))	0	optional
LDSroot	Specifies the Sasha data store root directory.	LDS	optional
sashaUser	Specifies the Sasha Username used for authorization.		optional

SashaServer Process Archiver

This section describes the SashaServer Process Archiver tab.

The screenshot shows the HPCC Systems Configuration Manager interface. On the left, the 'navigator' pane displays a tree view of the configuration environment. Under 'Environment - TestEnv.xml', the 'Software' section is expanded, and 'Sasha Server - mysasha' is selected. On the right, the 'SashaServerProcess' configuration pane is shown with the 'Archiver' tab active. This tab contains a table of attributes and their values.

name	value
cachedWUat	*****
cachedWUinterval	24
cachedWUlimit	100
DFUrecoveryAt	*****
DFUrecoveryCutoff	4
DFUrecoveryInterval	12
DFUrecoveryLimit	20
DFUWUat	*****
DFUWUcutoff	14
DFUWUduration	0
DFUWUinterval	24
DFUWUlimit	1000
DFUWUthrottle	0
keepResultFiles	false
WUat	*****
WUbackup	0
WUcutoff	8
WUduration	0
WUinterval	6
WUlimit	1000
WUretryinterval	7
WUthrottle	0

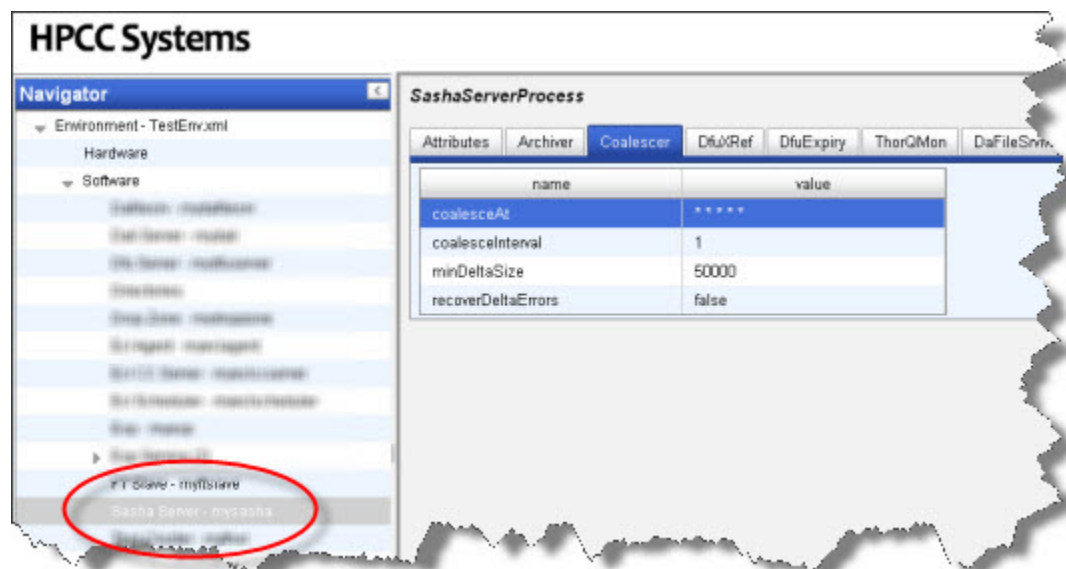
attribute	description	values	required
WUlimit	threshold number of workunits before archiving starts (0 disables).	1000	optional
WUcutoff	minimum workunit age to archive (days).	8	optional
WUbackup	minimum workunit age to backup (days, 0 disables).	0	optional
WUinterval	minimum interval between running WorkUnit archiver(in hours, 0 disables).	6	optional

Configuring the HPCC Systems® Platform
Configuration Manager Advanced View

attribute	description	values	required
WUat	schedule to run WorkUnit archiver (cron format).	* * * * *	optional
WUduration	maximum duration to run WorkUnit archiving session (hours, 0 unlimited).	0	optional
WUthrottle	throttle ratio (0-99, 0 no throttling, 50 is half speed).	0	optional
WUretryinterval	minimal time before retrying archive of failed WorkUnits (days).	7	optional
keepResultFiles	option to keep result files owned by workunits after workunit is archived	false	optional
DFUrecoveryLimit	threshold number of DFU recovery items before archiving starts (0 disables).	20	optional
DFUrecoveryCutoff	minimum DFU recovery item age to archive (days).	4	optional
DFUrecoveryInterval	minimum interval between running DFU recovery archiver(in hours, 0 disables).	12	optional
DFUrecoveryAt	schedule to run DFU recovery archiver (cron format).	* * * * *	optional
DFUWUlimit	threshold number of DFU workunits before archiving starts (0 disables).	1000	optional
DFUWUcutoff	minimum DFU workunit age to archive (days).	14	optional
DFUWUinterval	minimum interval between running DFU recovery archiver (in hours, 0 disables).	24	optional
DFUWUat	schedule to run DFU workunit archiver (cron format).	* * * * *	optional
DFUWUduration	maximum duration to run DFU WorkUnit archiving session (hours, 0 unlimited).	0	optional
DFUWUthrottle	throttle ratio (0-99, 0 no throttling, 50 is half speed).	0	optional
cachedWUlimit	threshold number of cached workunits before removal starts (0 disables).	100	optional
cachedWUinterval	minimum interval between running cached workunit removal(in hours, 0 disables).	24	optional
cachedWUat	schedule to run cached workunit removal (cron format).	* * * * *	optional

SashaServer Process Coalescer

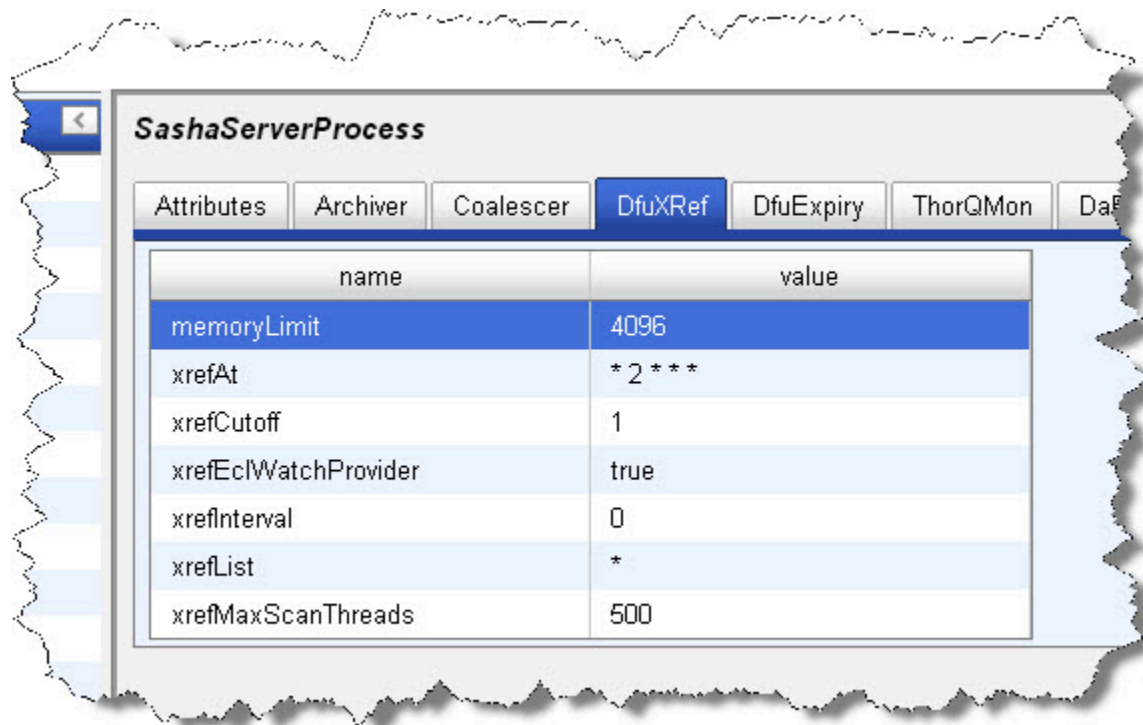
This section describes the SashaServer Process Coalescer tab.



attribute	description	values	required
coalesceInterval	minimum interval between running Dali datastore coalescer (in hours, 0 disables).	1	optional
coalesceAt	schedule to run Dali datastore coalescer (cron format).	* * * * *	optional
minDeltaSize	Coalescing will only begin, if the delta size is above this threshold (K)	50000	optional
recoverDeltaErrors	Switch on to auto recover from corruption to delta files on load	false	optional

SashaServer Process DfuXRef

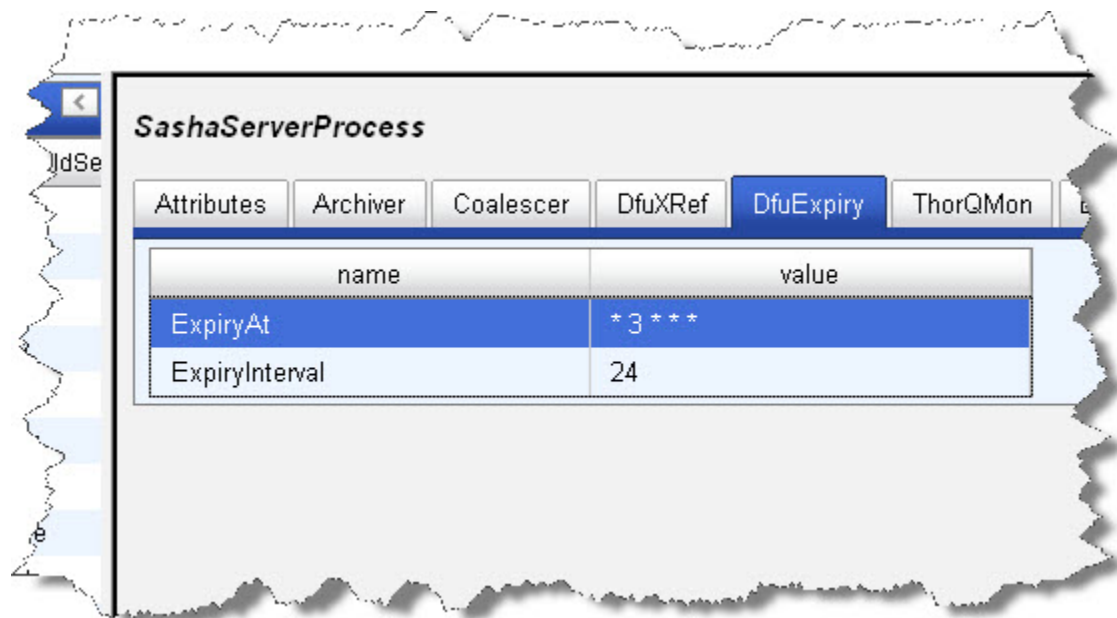
This section describes the SashaServer Process DfuXref tab.



attribute	description	values	required
xrefInterval	minimum interval between running DFU XREF (in hours, 0 disables).	672	optional
xrefAt	schedule to run DFU XREF (cron format).	* 2 * * *	optional
xrefList	comma separated list of clusters to xref (* for all clusters).	*	optional
xrefCutoff	cutoff (in days) to ignore recent files.	1	optional
xrefMaxScanThreads	maximum thread count for scanning directories	500	optional
xrefEcIWatchProvider	use sasha for EcIWatch initiated xref.	true	optional
xrefMaxMemory	The upper memory limit that xref can use.	4096	optional
suspendCoalescerDuringXref	Suspend the coalescer whilst xrefing.	true	optional

SashaServer Process DfuExpiry

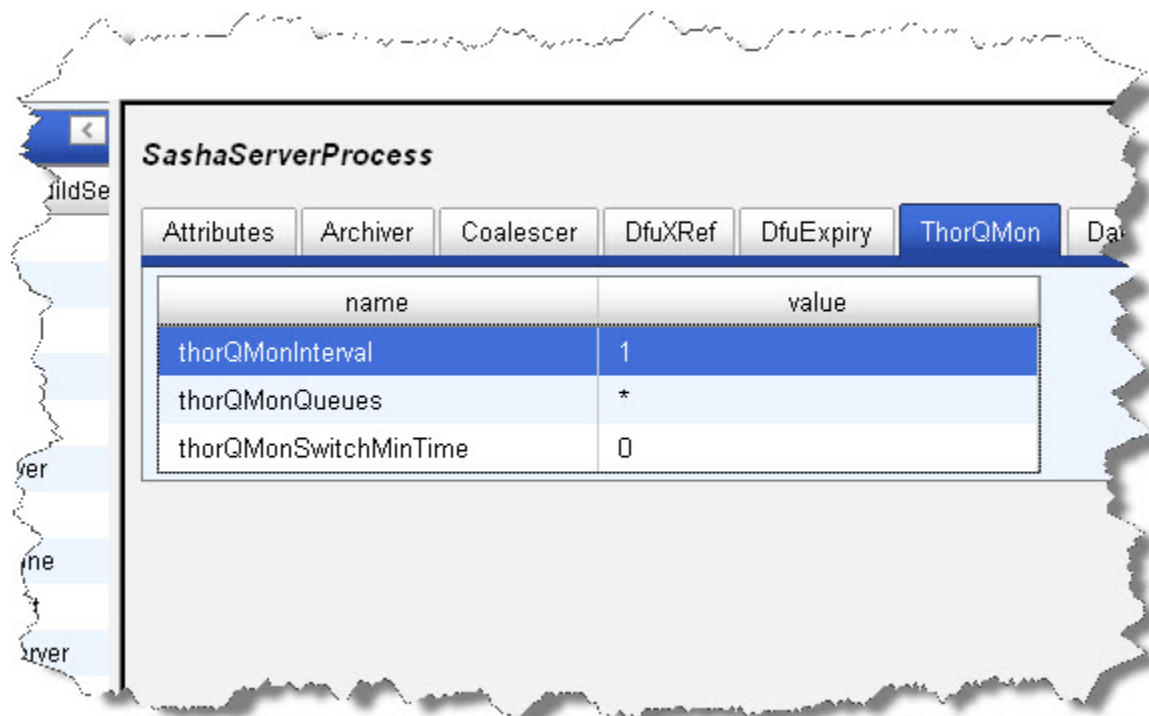
This section describes the SashaServer Process DfuExpiry tab.



attribute	description	values	required
ExpiryInterval	minimum interval between checking for distributed file expiry (in hours, 0 disables).	24	optional
ExpiryAt	schedule to check for distributed file expiry (cron format).	* 3 * * *	optional
PersistExpiryDefault	Default number of days to delete unused persist files	7	optional
ExpiryDefault	Default number of days to delete unused standard files that are flagged with EXPIRY	14	optional

SashaServer Process ThorQMon

This section describes the SashaServer Process ThorQMon tab.



attribute	description	values	required
thorQMonQueues	comma separated list of Thor queues to monitor (* for all).	*	optional
thorQMonInterval	interval to monitor and log queue activity (in minutes).	1	optional
thorQMonSwitchMinTime	minimum idle time before job will switch queues. (in minutes, 0 disables)	0	optional

SashaServer Process DaFileSrvMonitor

This section describes the SashaServer Process DaFileSrvMonitor tab.



attribute	description	values	required
dafsmonInterval	minimum interval between running DaFileSrv monitor (in hours, 0 disables).	0	optional
dafsmonAt	schedule to run DaFileSrv monitor (cron format).	*****	optional
dafsmonList	comma separated list of clusters and IPs to monitor (* for all clusters).	*	optional

SashaServer Process Notes

This tab allows you to add any notes pertinent to the component's configuration. This can be useful to keep a record of changes and to communicate this information to peers.

Thor

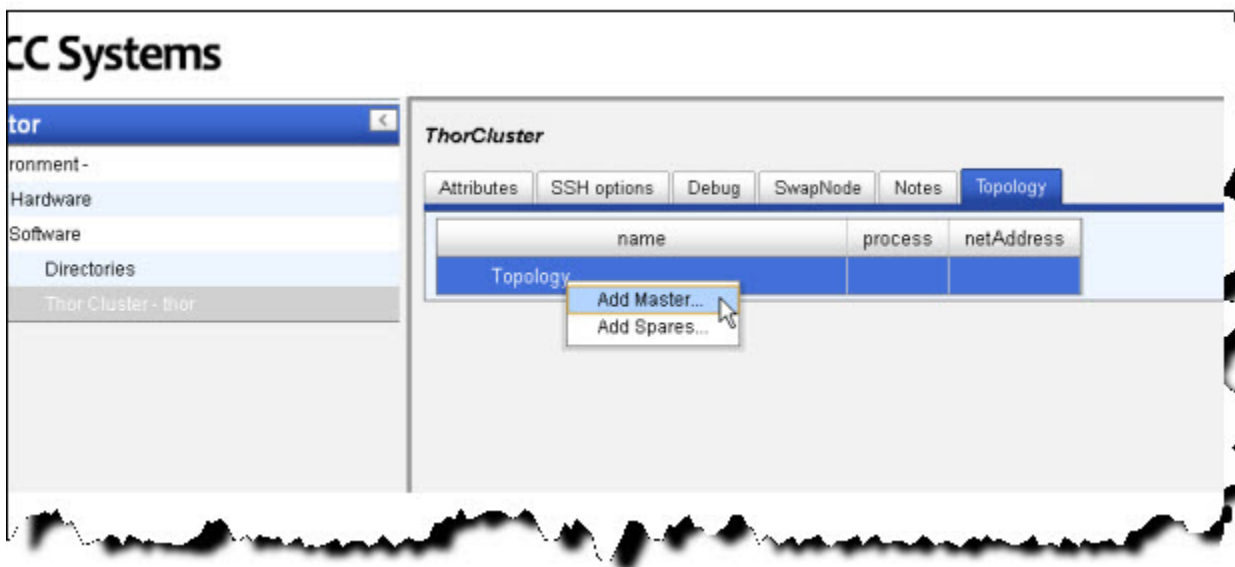
This section details how to define a Data Refinery (Thor) cluster. Before you begin, you should decide the width of the cluster (i.e., how many slave nodes will you have).

1. Select **Thor Cluster - mythor** in the Navigator panel on the left side.
2. Select the **Topology** tab.
3. Expand the Topology, if needed, then right-click the Master and select Delete.

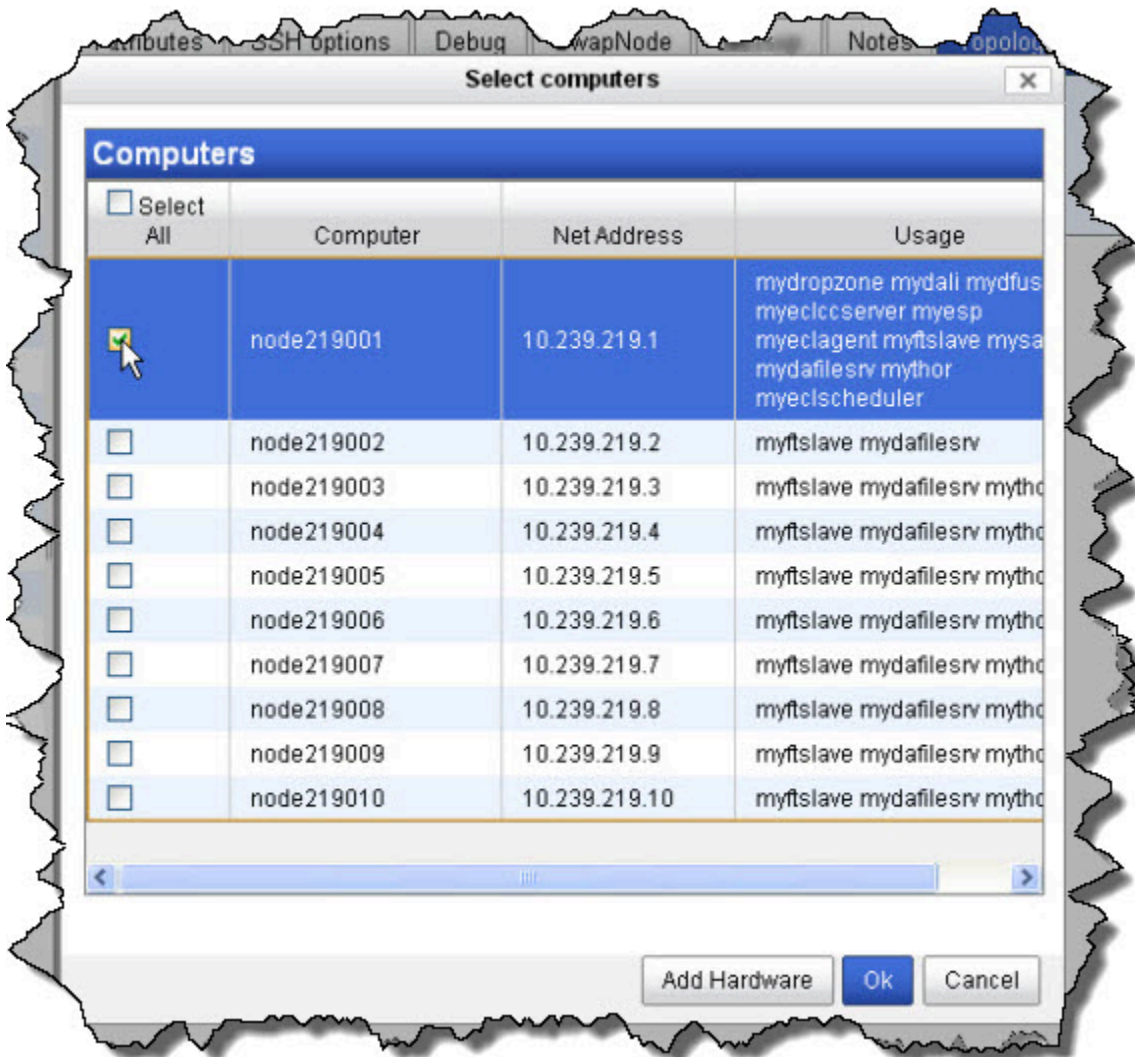
This deletes the sample one-node Thor.

You will replace this with a multi-node cluster.

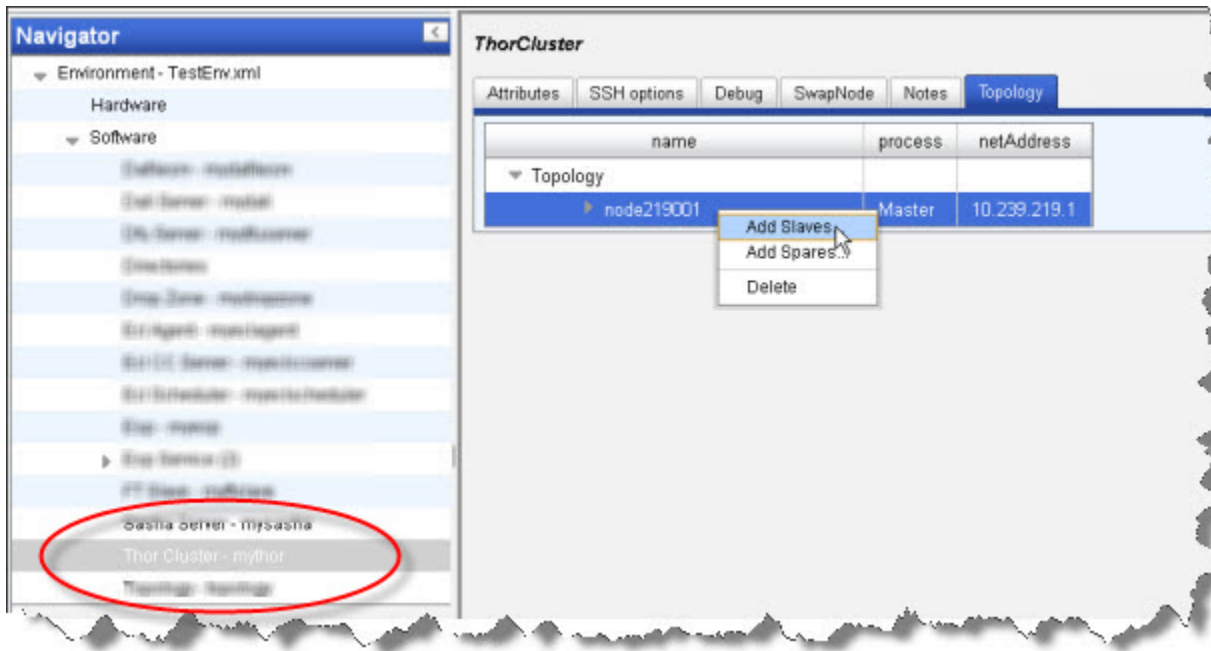
1. right-click on the Topology and select Add Master.



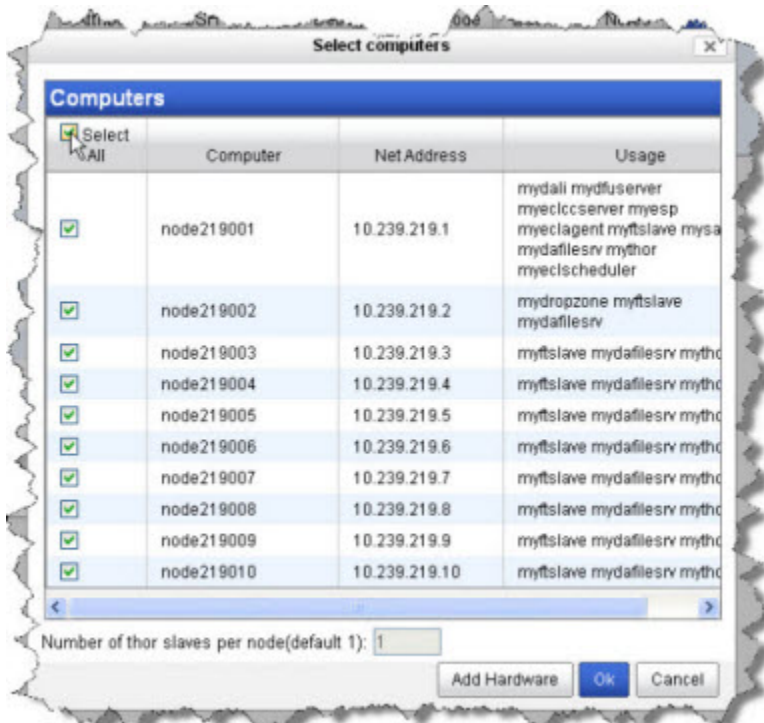
2. Select a computer from the list, then press the OK button.



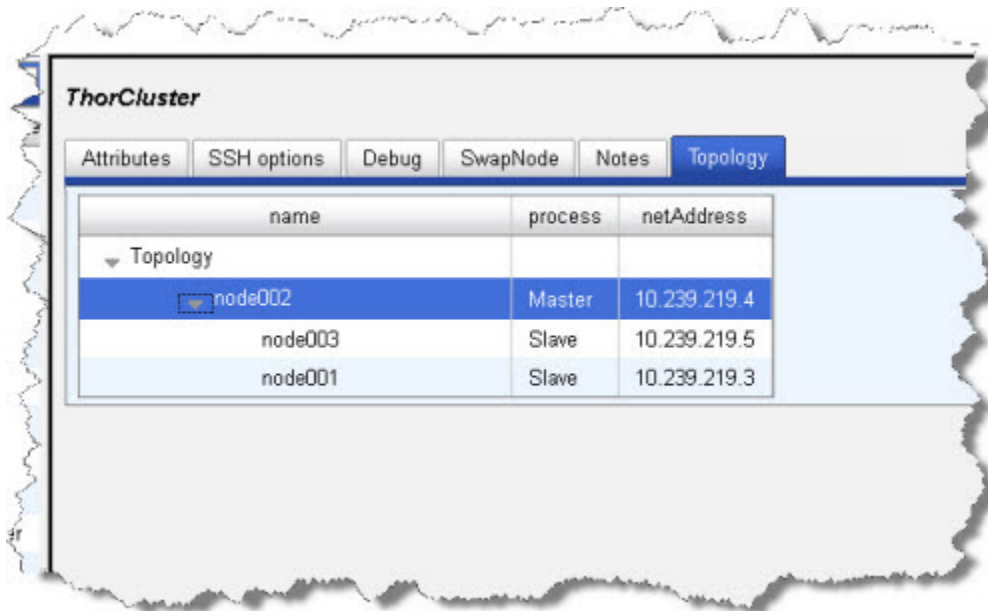
3. Right-click on the Master and select Add Slaves.



4. Select the computers to use as slaves from the list, then press the OK button. Use CTRL+CLICK to multi-select or SHIFT+CLICK to select a range.



The Nodes now display below the Thor Master node.



5. Select Thor Cluster - mythor in the Navigator panel on the left side.

6. Select the Attributes tab.

HPCC Systems	
Navigator	
Environment - TestEnv.xml	
Hardware	
Software	
CudaExec - mythor	
CudaServer - mythor	
DaliServer - mythor	
Directories	
Group Zone - mythor	
Hadoop Agent - mythor	
Hadoop Server - mythor	
Hadoop Scheduler - mythor	
Hadoop	
Hadoop Service (2)	
Hadoop Slave - mythor	
Sasha Server - mysasha	
Thor Cluster - mythor	
Toolbox - mythor	

allowedPipePrograms	*
autoCopyBackup	false
checkPointRecovery	false
daliServers	mydali
defaultOutputNodeGroup	
description	Thor process
externalProgDir	
globalMemorySize	
idleRestartPeriod	480
largeMemSize	
localThor	false
localThorPortBase	false
localThorPortInc	true
masterport	
maxActivityCores	0
monitorDaliFileServer	true
multiThorExclusionLockName	
multiThorMemoryThreshold	
multiThorPriorityLock	false
name	mythor
nodeGroup	
pluginsPath	/opt/HPCCSystems/plugins/
replicateAsync	true

7. Change the value of the localThor to **false**

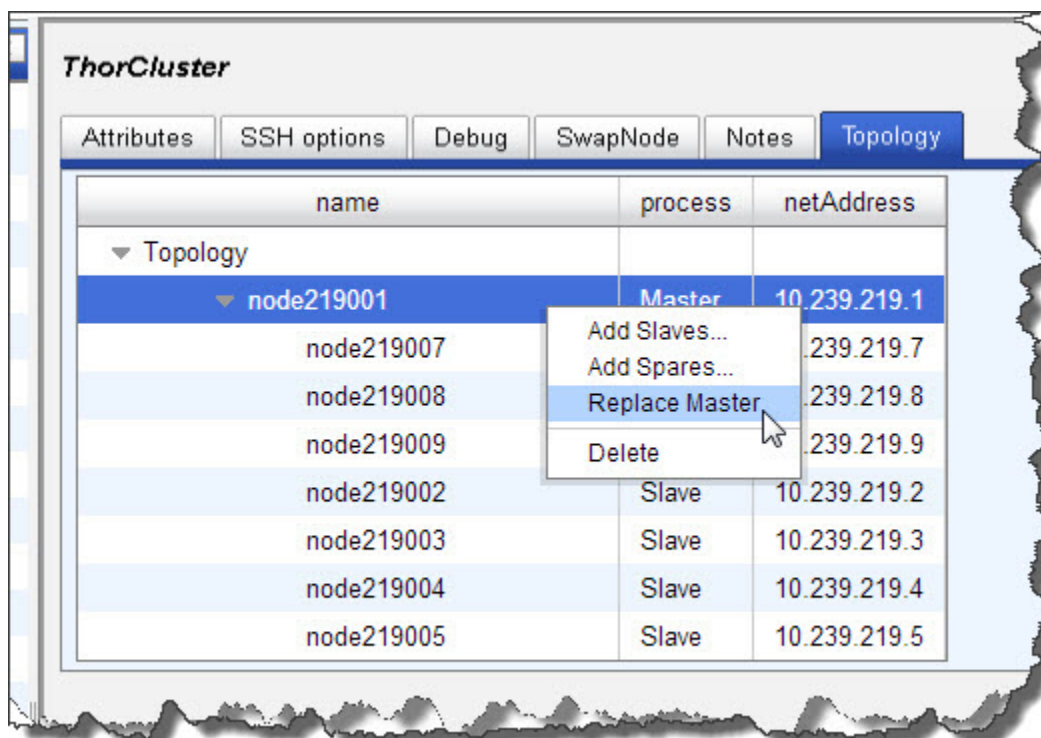
8.

Click the  disk icon to save

Changing Thor topology

If you want to designate a different node as the Thor master when setting up a multi-node system, follow these steps.

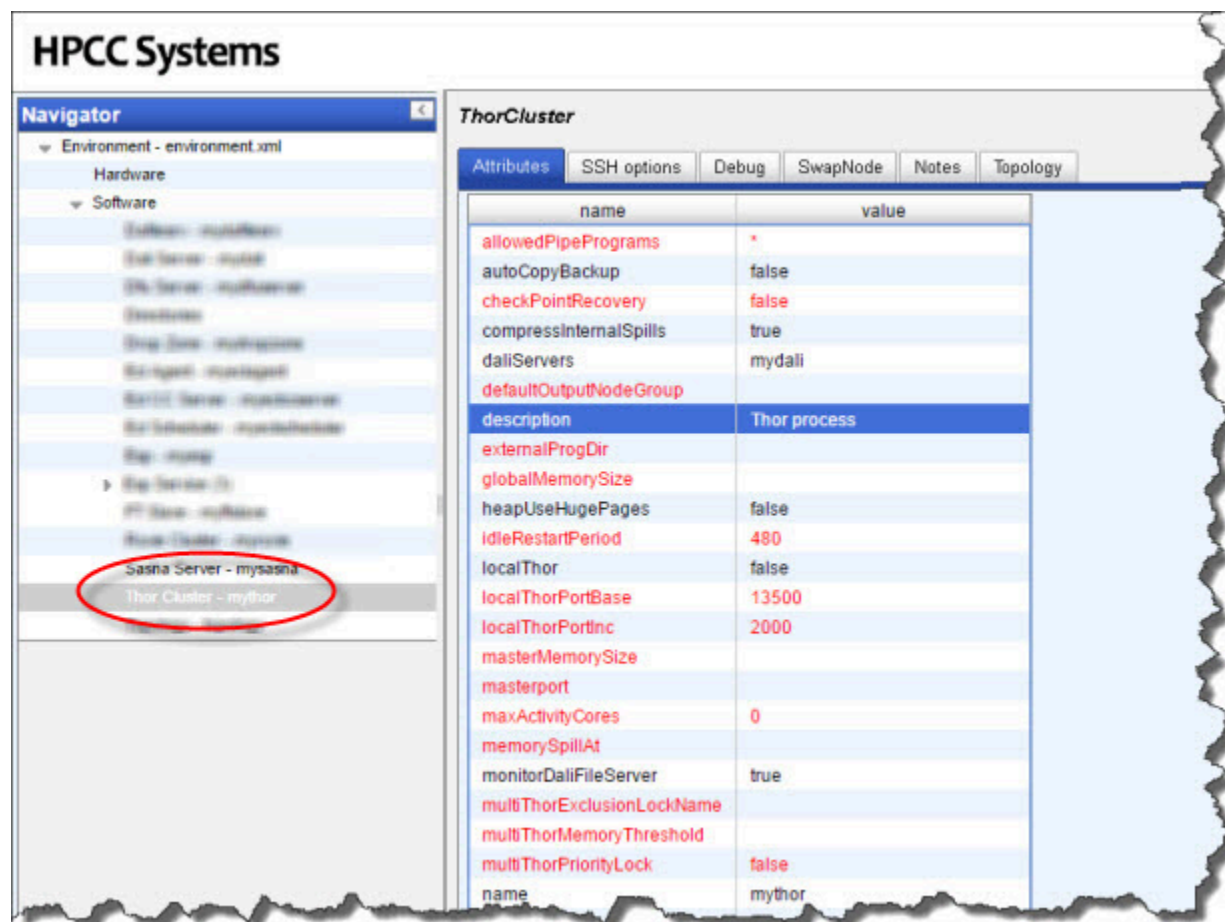
1. Select **Thor Cluster - mythor** in the Navigator panel on the left side.
2. Select the **Topology** tab.
3. Right-click on the Master node
4. Select the **Replace Master** option.



You should only use this feature when initially setting up your system. If there is data on the nodes when attempting to Swap Master, you run the risk of losing or corrupting some data.

ThorCluster Attributes

This section describes the Thor Cluster Attributes tab.



Thor Memory Settings

When the **globalMemorySize** is left unset, Thor[master] detects total physical memory and allocates 75% of it. If there are multiple slaves per node (`slavesPerNode>1`) it divides the total among the slaves. If `globalMemorySize` is defined, then it allocates that amount of memory to each slave. The `masterMemorySize` attribute allocates memory for the Thor master. If omitted, Thor master uses `globalMemorySize`, or the default 75% of memory.

On systems with a lot of memory, the default 75% of physical memory is probably too conservative and reserving total physical minus 2GB (for the OS and other processes) is sensible. You should then divide that number by the number of `slavesPerNode`.

If there are multiple Thors sharing the same nodes, then `globalMemorySize` must be configured to take that into account.

For example, if there are 2 Thors each with 2 slaves per box, that will mean there are 4 slaves per physical node. So you should use a formula similar to the following in your calculations when configuring `globalMemorySize`:

```
globalMemorySize = (total-physical-memory)-2GB / (2*2)
```


Without any specified setting, Thor assumes it has exclusive access to the memory and would therefore use too much (because each Thor is unaware of the other's configuration and memory usage).

If **localThor** is set to true and **masterMemorySize** and **globalMemorySize** are unspecified, then the defaults will be 50% for **globalMemorySize** (divided by **slavesPerNode**) and 25% for **masterMemorySize**.

Although a configuration may be set using upper memory limits that exceed total physical memory, Thor will not actually reserve the memory ahead of time and will only hit memory problems when and if your jobs use all of memory. So, for example, two Thors that are configured to use all available memory could peacefully co-exist until a query on each are simultaneously using more memory than the node has available.

attribute	description	values	required
name	Name for this process		required
description	Description for this process	Thor process	optional
daliServers	Specifies the dali server to which this thor is attached.		required
externalProgDir	If specified, external programs executed via PIPE will be started in the named directory		optional
masterMemorySize	Memory (in MB) to use for rows on thor master. It will default to globalMemorySize if unset		optional
fileCacheLimit	File Cache limit (in MB). It will default to 1800 if unset	1800	optional
globalMemorySize	Memory (in MB) to use for rows per Thor slave process. If unset, default = [75% of physical memory] / slavesPerNode		optional
memorySpillAt	Threshold that the memory manager should start requesting memory to be freed (percentage)		optional
heapUseHugePages	Allow Thor master and slave to use memory from huge pages if they have been configured.	false	optional
heapMasterUseHugePages	Allow heapUseHugePages to be overridden for the master. Useful to disable if running on a single machine.		optional
heapUseTransparentHugePages	Allow Thor master and slave to use memory from transparent huge pages.	true	optional
heapRetainMemory	Retain and do not return unused memory to the operating system.	false	optional
pluginsPath		/opt/HPCCSystems/plugins/	optional
nodeGroup	Name of a node group running Thor slaves. (if omitted uses same name as Thor)		optional
defaultOutputNodeGroup	Default group to output to. Leave blank if output defaults to local cluster.		optional

Configuring the HPCC Systems® Platform
Configuration Manager Advanced View

attribute	description	values	required
masterport	Base port to use for master	20000	optional
slaveport	Base port to use for slaves	20100	optional
localThor	Assume all slaves are local to this machine rather than started via agent	false	optional
slavesPerNode	Defines how many slave processes there are on each node	1	optional
channelsPerSlave	Defines how many slave channels per slave process	1	optional
localThorPortInc	Port increment between slaves on same node	20	optional
multiThorMemoryThreshold	Memory usage (in MB) beneath which multiple Thors will run in parallel. Leave blank if no limit.		optional
multiThorExclusionLockName	Prevents other thors (on any queue) sharing the same multiThorExclusionLockName name from running jobs at the same time		optional
replicateOutputs	Replicate output files	true	optional
replicateAsync	Perform output replication in the background, allowing thor to process next task	true	optional
autoCopyBackup	If files at primary location are missing, copy into place from backup location	false	optional
checkPointRecovery	Enable support for continuing failed workunits from point of failure	false	optional
watchdogEnabled	Enable/disable watchdog process (periodically checking slaves are still alive)	true	optional
watchdogProgressEnabled	Enable/Disable graph progress reporting in watchdog process	true	optional
watchdogProgressInterval	Graph progress reporting interval (seconds)	30	optional
slaveDownTimeout	A slave will be marked down after the specified elapsed time period (seconds)	300	optional
verifyDaliConnectionInterval	Interval between verification of thor to dali connection (seconds)	300	optional
idleRestartPeriod	Period of idle time after which to initiate an auto restart (minutes)	480	optional
smallSortThreshold	Sort size threshold for thor global sorting on multiple nodes (MB)	1024	optional
maxActivityCores	Maximum number of cores to use per activity (only currently used by sorting activities). Default equals all available	0	optional

attribute	description	values	required
monitorDaliFileServer	Warn if dafilesrv process is not running on computers	true	optional
allowedPipePrograms	Comma separated list of allowed PIPE program (* for allow all)	*	optional
compressInternalSpills	Compress internal writes to disk when spilling	true	optional
affinity	A comma separated list of cpu ids (and ranges) to bind all thor slaves to		optional
autoAffinity	Automatically bind slave processes to a single cpu socket, if multiple slaves are running on a multi socket machine	true	optional
numaBindLocal	Restrict allocations to memory attached to the cpu sockets the slave process is bound to	false	optional
httpCallerIdHeader	HTTP Header field to use for sending and receiving CallerId	HPCC-Caller-Id	optional
httpGlobalIdHeader	HTTP Header field to use for sending and receiving GlobalId	HPCC-Global-Id	optional

Allowed Pipe Programs

In version 9.2.0 and greater, commands used in a PIPE action are restricted by default. However, for legacy reasons, the default stock behavior is different in bare-metal and containerized deployments. In both types of systems, if `allowedPipePrograms` is unset, then all but "built-in" programs are restricted (The only built-in program currently is 'roxiepipe').

In bare-metal, the default `environment.xml` includes a setting value of "*" for **`allowedPipePrograms`**. This means (by default) any PIPE program can be used.

In a secure system, this should be removed or edited to avoid arbitrary programs, including system programs, from being executed.

XREF with multiple Thor clusters

XREF runs on the primary Thor cluster. For a configuration with more than one Thor cluster on the same physical node group, ECL Watch only displays the primary Thor and not any other Thor that shares the same node group. This is the Thor cluster whose name matches the *nodeGroup* in the configuration.

To allow XREF to run in an environment with multiple Thor clusters, set the `nodeGroup` to the same value for all Thor clusters.

ThorCluster SSH Options

This section describes the ThorCluster SSH Options tab.

name	value
SSHidentityfile	\$HOME/.ssh/id_rsa
SSHpassword	
SSHretries	3
SSHtimeout	0
SSHusername	hpcc

attribute	description	values	required
SSHidentityfile	location of identity file (private key) on Thor master	\$HOME/.ssh/id_rsa	optional
SSHusername	Username to use when running Thor slaves	hpcc	optional
SSHpassword	Fixed password - only required if no identity file present NB **insecure**		optional
SSHtimeout	Timeout in seconds for SSH connects	0	optional
SSHretries	Number of times to retry failed connect	3	optional

ThorCluster Debug

The debug tab is for internal use only

ThorCluster Swap Node

This section describes the ThorCluster Swap Node tab.

The screenshot shows the HPCC Systems Configuration Manager interface. On the left is the 'Navigator' pane with a tree view under 'Environment - A1NodeLDAP.xml'. The 'Software' section is expanded, and 'Thor Cluster - mythor' is highlighted with a red circle. On the right is the 'ThorCluster' configuration pane. It has tabs for 'Attributes', 'SSH options', 'Debug', 'SwapNode' (which is selected), 'Notes', and 'Topology'. The 'SwapNode' tab displays a table with configuration attributes and their values.

name	value
AutoSwapNode	false
CheckAfterEveryJob	false
SwapNodeCheckMirrorDrive	true
SwapNodeCheckPrimaryDrive	true
SwapNodeCheckScript	
SwapNodeCheckScriptTimeout	0
SwapNodeInterval	24
SwapNodeMaxConcurrent	1
SwapNodeRestartJob	false

attribute	description	values	required
AutoSwapNode	Failing nodes will be automatically swapped for spare nodes	false	optional
SwapNodeCheckPrimaryDrive	Primary drive is checked for read/write	true	optional
SwapNodeCheckMirrorDrive	Mirror drive is checked for read/write	true	optional
SwapNodeMaxConcurrent	Maximum number of concurrent automatic swaps (within the swap interval)	1	optional
SwapNodeInterval	Interval (in hours) for maximum concurrent swaps	24	optional
SwapNodeRestartJob	Restart job that failed prior to swap	false	optional
CheckAfterEveryJob	Check nodes after every job (if false, only check after failed jobs)	false	optional

ThorCluster Notes

This tab allows you to add any notes pertinent to the component's configuration. This can be useful to keep a record of changes and to communicate this information to peers.

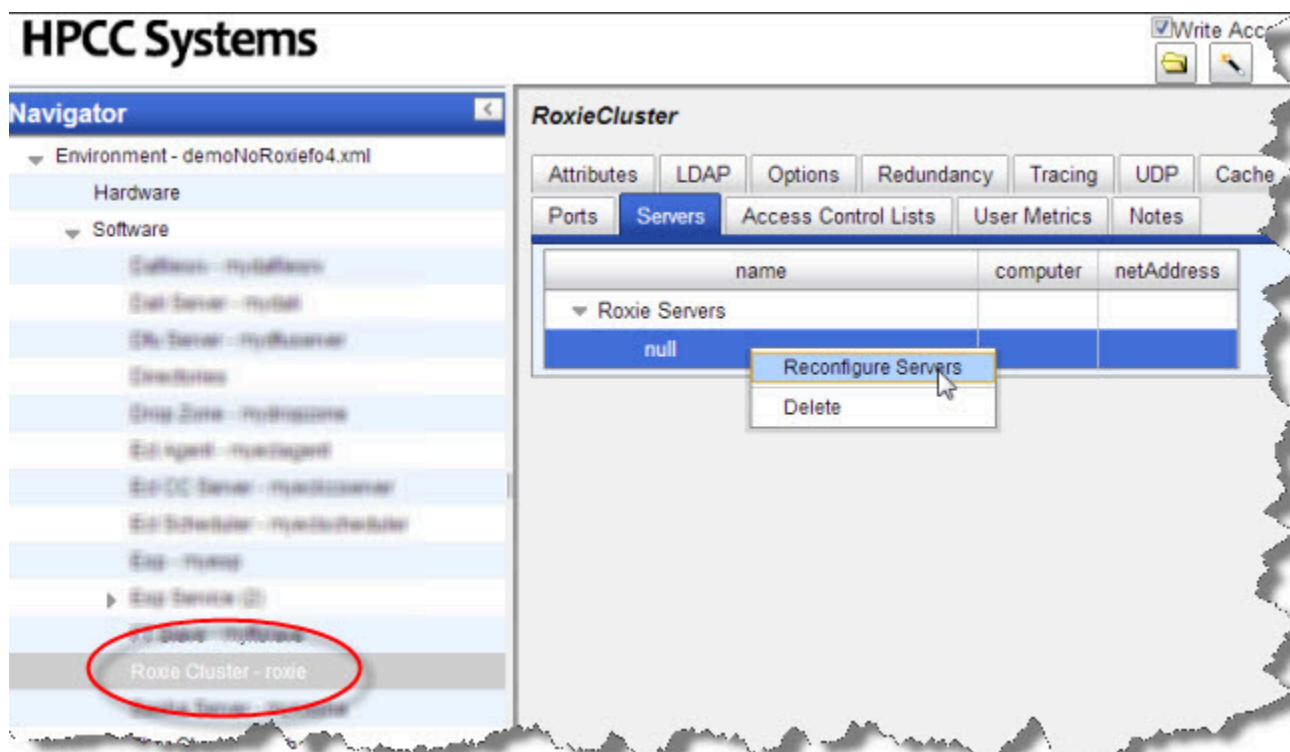
Roxie

This section details how to define a Rapid Data Delivery Engine (Roxie) cluster. Before you begin, you should decide the width of the cluster (i.e., how many agent nodes will you have).

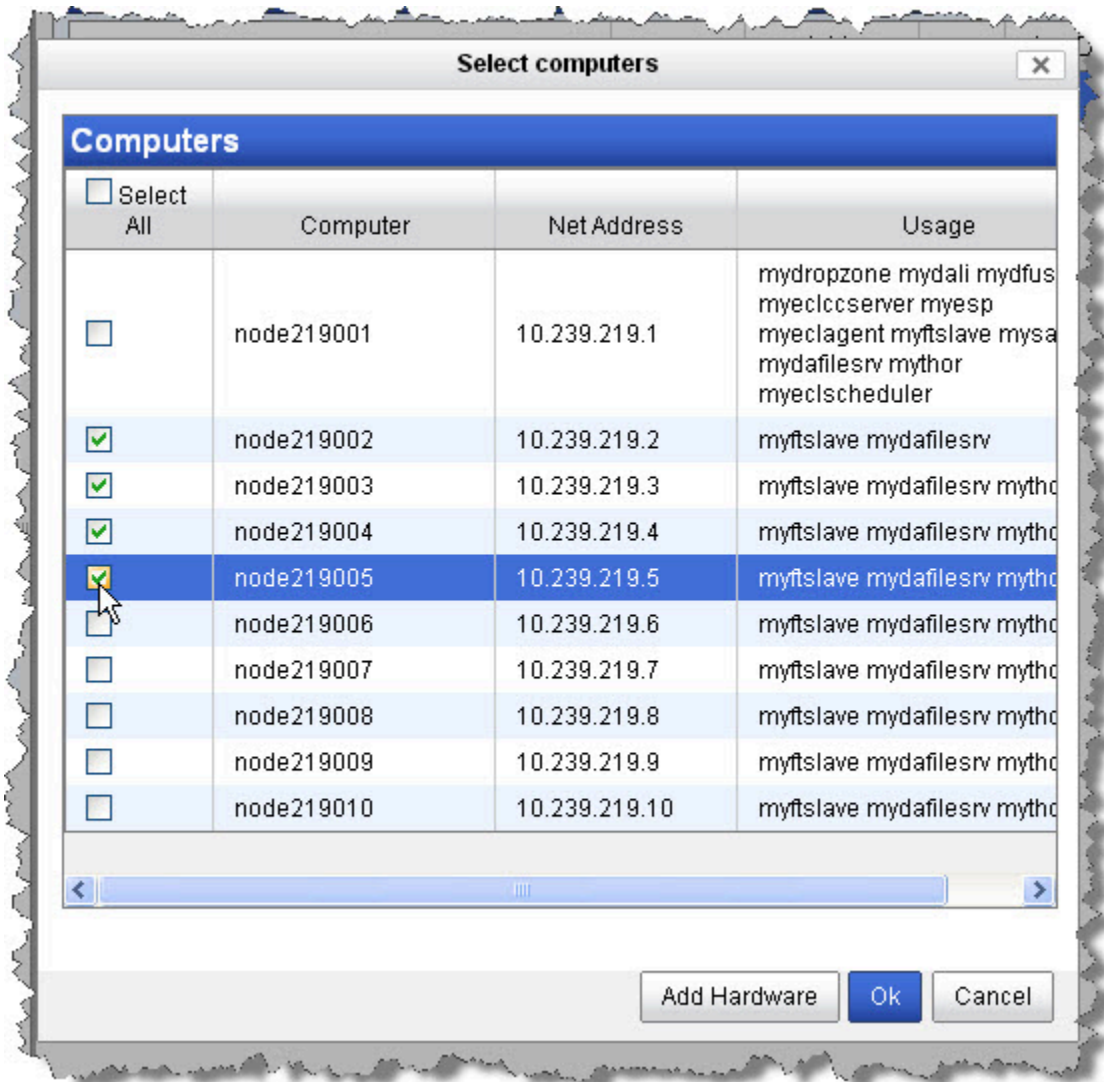
1. Select **Roxie Cluster** in the Navigator panel on the left side.

Note: If you did not specify a value in the *Number of nodes for Roxie cluster* field when you first set up your environment, you will not have a Roxie Cluster. To add a Roxie Cluster component: Right-click on the **Software** component in the Navigator Panel, then select **New Components** then **roxie** from the drop lists.

2. Select the **Servers** tab.
3. Right-click the Roxie Servers and select Reconfigure Servers.

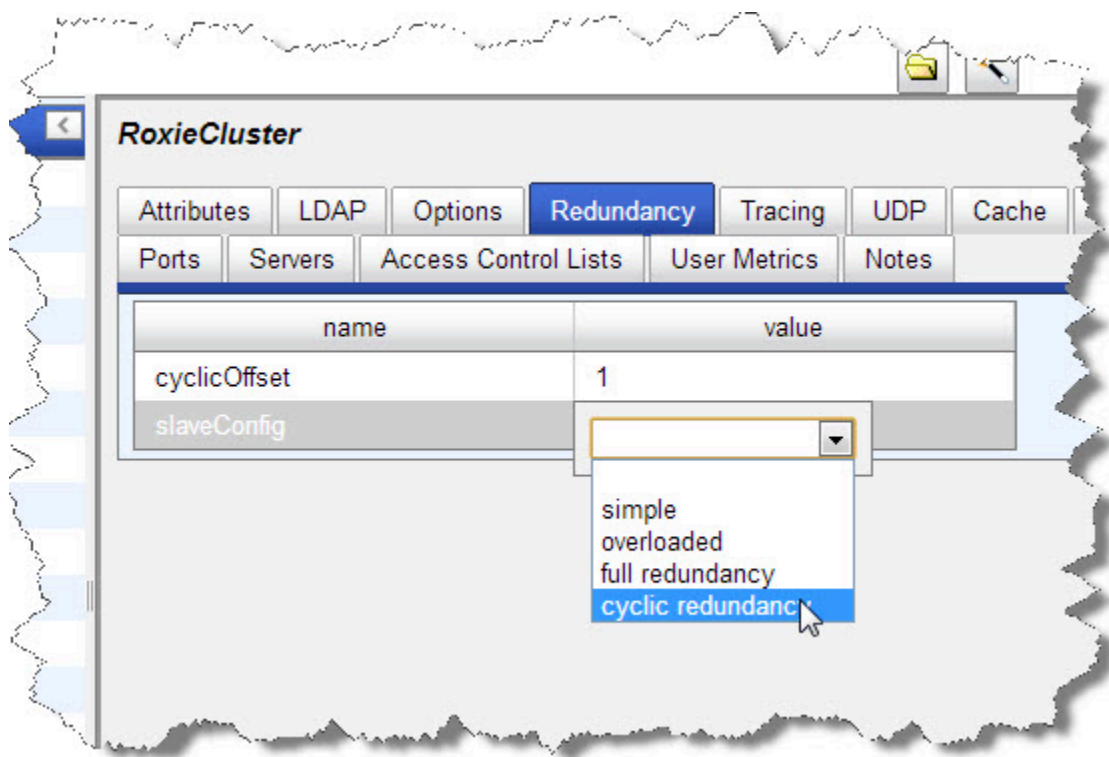


4. Select the computers to use as Servers from the list, then press the OK button.



5. Select the **Redundancy** tab.

6. Select the redundancy scheme to use. Typically, this is cyclic redundancy, as shown below.



7.  Click the disk icon to save

8. Close Configuration Manager by pressing ctrl+C in the command window where it is running.

Roxie Configuration Attributes

Roxie has many configurable attributes which can be used to for customizing and tuning to your specific needs. The following section expands on each of the Roxie tabs and the available attributes. There is additional Roxie configuration information in the section immediately following these tables.

RoxieCluster

RoxieCluster

Attributes

attribute	description	values	required
name	Name for this process		required
description	Description for this process	Roxie cluster	optional
daliServers	Specifies the dali server to which this roxie is attached.		optional
lockDali	If set, Roxie will use cached info from dali only, and will not connect to dali or refresh the cache.	false	optional
multicastBase	The base multicast IP for this roxie cluster. Multicast ranges must not overlap for any roxie clusters in the same multicast domain.	239.1.1.1	optional
multicastLast	The last multicast IP available for this roxie cluster. Multicast ranges must not overlap for any roxie clusters in the same multicast domain.	239.1.254.254	optional
multicastTTL	The multicast TTL (Time To Live) setting for this roxie cluster. Zero means do not explicitly set TTL, and use the default OS setting.	1	optional
directory	Specifies the directory to which the software will be deployed.	/var/lib/HPCCSystems/roxie/	optional
pluginsPath	Alternate path where plugin files are deployed (./plugins is assumed if not specified)	/opt/HPCCSystems/plugins	optional

RoxieCluster RoxieFarmProcess

Attributes

attribute	description	values	required
port	the network port on which the Roxie servers accept connections	9876	required
numThreads	Number of simultaneous queries Roxie servers will accept on this port	30	optional
listenQueue	Number of pending connections that can be accepted	200	optional
requestArrayThreads	Number of simultaneous queries Roxie servers will process using the MERGE option of SOAPCALL	5	optional
aclName	Name of any Access Control List to use		optional
protocol	Protocol to use	native	optional

attribute	description	values	required
passphrase	Pass phrase for cert		optional
certificateFileName	Path to certificate filename		optional
privateKeyFileName	Path to private key filename		optional

RoxieCluster RoxieServerProcess

Attributes

attribute	description	values	required
computer			required
netAddress			optional

RoxieCluster ACL

Attributes

attribute	description	values	required
name	Name of this Access Control List.	acl	required

RoxieCluster ACL BaseList

Attributes

attribute	description	values	required
name	Name of another Access Control List to extend		required

RoxieCluster ACL Access

Attributes

attribute	description	values	required
allow	Whether or not to allow the access	Yes	required
ip	I.P. Address	0.0.0.0	optional
mask	Internet address mask	255.255.255.255	optional
query	wildcard for queries to allow/disallow	.*	optional
errorCode	optional error code to associate with the query		optional
name	Name of this Access Control Rule.	ACLrule	required
error	optional error message to associate with the query		optional

RoxieCluster PreferredCluster

Attributes

attribute	description	values	required
name	Name of the cluster		required
priority	Priority (negative to disable)		required

RoxieCluster UserMetric

Attributes

attribute	description	values	required
name	Name of this metric.		required
regex	Expression to match.		required

RoxieCluster Notes

Attributes

attribute	description	values	required
severity	Significance of this note.	Minor	optional
date	Date and time this note was entered		optional
computer	Computer from which this note was entered		optional
user	User account from which this note was entered		optional

LDAP

attribute	description	values	required
ldapUser	Specifies the user name for LDAP file access checking.	roxie	optional
ldapPassword	Specifies the password for LDAP file access checking.		optional

Options

attribute	description	values	required
affinity	If non-zero, binds the roxie process to use the specified cores only (bitmask)	0	optional
allFilesDynamic	If enabled, files will be resolved per-query and not locked between queries	false	optional
backgroundCopyClass	Specify an IONICE class for the background copy thread.	none	optional
backgroundCopyPrio	Specify an IONICE value for the background copy thread, if background-CopyClass set to best-effort.	0	optional
callbackRetries	Number of retries before callbacks from agents to server are aborted	3	optional
callbackTimeout	Timeout (in ms) before callbacks from agents to server are resent	5000	optional
checkFileDate	Compare file dates of physical files with the information in DFS.	true	optional
collectFactoryStatistics	Accumulate summary statistics for all queries	true	optional

Configuring the HPCC Systems® Platform
Configuration Manager Advanced View

attribute	description	values	required
copyResources	Copies any missing data files/keys from the position they were in when query was deployed.	true	optional
coresPerQuery	If non-zero, binds each incoming query to use the specified number of cores only	0	optional
debugPermitted	Allow the ECL query debugger to attach to queries on this Roxie	true	optional
defaultHighPriorityTimeLimit	Maximum run time (in ms) for any single active high-priority query (if not overridden)	0	optional
defaultHighPriorityTimeWarning	Time (in ms) before generating SNMP warning for a high-priority query (if not overridden)	5000	optional
defaultLowPriorityTimeLimit	Maximum run time (in ms) for any single active low-priority query (if not overridden)	0	optional
defaultLowPriorityTimeWarning	Time (in ms) before generating SNMP warning for a low-priority query (if not overridden)	0	optional
defaultMemoryLimit	Maximum amount of memory available for row data in any single active query (if not overridden)	0	optional
defaultSLAPriorityTimeLimit	Maximum run time (in ms) for any single active SLA-high-priority query (if not overridden)	0	optional
defaultSLAPriorityTimeWarning	Time (in ms) before generating SNMP warning for a SLA-high-priority query (if not overridden)	5000	optional
defaultStripLeadingWhitespace	Default value for stripping leading whitespace in input XML values	1	optional
enableSysLog	Enable use of syslog for monitoring.	true	optional
encryptionInTransit	Encrypt traffic between Roxie nodes.	false	optional
flushJHtreeCacheOnOOM	Should the index node memory allocation flush the cache and retry if memory allocation fails	true	optional
fieldTranslationEnabled	Enables translation (where possible) of mismatched file layouts on-the-fly. Specify 'payload' to attempt to translate payload fields only	payload-RemoveOnly	optional
highTimeout	Timeout (in ms) before high priority requests are resent to agents	2000	optional
httpCallerIdHeader	HTTP Header field to use for sending and receiving CallerId	HPCC-Caller-Id	optional

Configuring the HPCC Systems® Platform
Configuration Manager Advanced View

attribute	description	values	required
httpGlobalIdHeader	HTTP Header field to use for sending and receiving GlobalId	HPCC-Global-Id	optional
mapHttpCallUrlsToSecrets	In SOAPCALL and HTTPCALL check if URLs have been mapped to secrets	false	optional
warnIfUrlNotMappedToSecret	In SOAPCALL and HTTPCALL warn if URLs not mapped to secrets	false	optional
requireUrlsMappedToSecrets	Require SOAPCALL and HTTPCALL URLs are secrets or mapped to secrets	false	optional
ignoreOrphans	Treat out-of-date local files as if they were not present.	true	optional
lazyOpen	Delay opening files until first use. Select smart to use lazy mode only after a restart	smart	optional
localFilesExpire	Period (in ms) of inactivity before a local datafile handle is closed	-1	optional
localSlave	All Roxie servers talk only to their embedded agent.	false	optional
lockSuperFiles	If enabled, superfiles will be locked while queries that use them are loaded	false	optional
lowTimeout	Timeout (in ms) before low priority requests are resent to agents	10000	optional
maxHttpConnectionRequests	Max number of query requests per persistent http connection	0	optional
maxHttpKeepAliveWait	Max number of milliseconds to wait for additional requests on a persistent http connection	5000	optional
maxLocalFilesOpen	Maximum number of local files to keep open	4000	optional
maxRemoteFilesOpen	Maximum number of remote files to keep open	1000	optional
minFreeDiskSpace	Minimum amount of disk space needed to be available for file copy to succeed	1073741824	optional
minLocalFilesOpen	Minimum number of local files to keep open	2000	optional
minRemoteFilesOpen	Minimum number of remote files to keep open	500	optional
monitorDaliFileServer	Warn if dafilesrv process is not running on computers	false	optional
parallelQueryLoadThreads	Use up to n multiple threads for initial query loading	0	optional
preferredSubnet	Preferred subnet to use for multi-NIC machines		optional
preferredSubnetMask	Preferred subnet mask to use for multi-NIC machines		optional

Configuring the HPCC Systems® Platform
Configuration Manager Advanced View

attribute	description	values	required
preloadOnceData	Evaluate : ONCE sections of queries at query load time	true	optional
prestartSlaveThreads	Prestart slave worker threads at start-up	true	optional
reloadRetriesSuspended	Retry loading of suspended queries whenever QuerySet reloads	true	optional
remoteFilesExpire	Period (in ms) of inactivity before a remote datafile handle is closed	3600000	optional
serverThreads	Default number of threads processing Roxie server requests (if not specified on Servers tab)	30	optional
siteCertificate	Name of the site certificate component that is used for security		optional
slaTimeout	Timeout (in ms) before SLA high priority requests are resent to agents	2000	optional
slaveQueryReleaseDelaySeconds	Delay before unregistering slave queries to allow in-flight to complete. Files are locked until query is unregistered.	60	optional
slaveThreads	Number of threads processing agent requests	30	optional
statsExpiryTime	Time (in seconds) that detailed reporting stats are kept	3600	optional
totalMemoryLimit	Maximum amount of memory available for row data in all active queries	1073741824	optional
heapUseHugePages	Allow roxie to use memory from huge pages if they have been configured.	false	optional
heapUseTransparentHugePages	Allow roxie to use memory from transparent huge pages.	true	optional
heapRetainMemory	Retain and do not return unused memory to the operating system.	false	optional
trapTooManyActiveQueries	should an SNMP trap get sent when too many active query error occurs	true	optional
useHardLink	If the data file exists on the current machine but in a different directory than roxie expects - create a hard link	false	optional
useMemoryMappedIndexes	Using memory-mapped files when merging multiple result streams from row-compressed indexes.	false	optional
useRemoteResources	Reads any missing data files/keys from the position they were in when deployed.	true	optional
heapLockMemory	Disable locking of heap memory	true	optional

Redundancy

Configuring the HPCC Systems® Platform
Configuration Manager Advanced View

attribute	description	values	required
cyclicOffset	Offset for cyclic redundancy mode	1	optional
channelsPerNode	Number of channels/data locations to use per node, in overloaded mode	1	optional
numDataCopies	Number of copies of the data in redundant modes	1	optional
slaveConfig	Roxie data redundancy mode		optional

Tracing

attribute	description	values	required
traceLevel	Level of detail in reporting (set to 0 for none, 1 for normal, > 1 or more for extended)	1	optional
logFullQueries	Log full text (unless blindLogging) and resource usage of all queries received	false	optional
blindLogging	Suppress all logging of any data or query text	false	optional
memTraceLevel	Level of detail in reporting mem mgr information(set to 0 for none, 1 for normal, >1 or more for extended)	1	optional
miscDebugTraceLevel	Level of miscellaneous debug tracing unrelated to all other tracing(set to 0 for none, 1 for normal, >1 or more for extended)	0	optional
soapTraceLevel	Level of detail in reporting SOAPCALL information(set to 0 for none, 1 for normal, >1 or more for extended)	1	optional
traceTranslations	Trace record layout translations to log file"	true	optional
traceEnabled	TRACE activity output enabled by default (can be overridden in workunit or query)	false	optional
traceLimit	Number of rows output by TRACE activity	10	optional
udpTraceLevel	Level of detail in reporting udp information(set to 0 for none, 1 for normal, >1 or more for extended)	1	optional
useLogQueue	Queue logs messages	true	optional
logQueueDrop	Specifies the number of log messages which will be dropped if the maximum length of the queue of unhandled messages is exceeded.	32	optional
logQueueLen	Specifies the maximum length of the queue of unhandled log messages. Messages will be dropped if this is exceeded.	512	optional

UDP

attribute	description	values	required
roxieMulticastEnabled	Controls whether multicast is used to communicate between nodes	true	optional
udpAdjustThreadPriorities	Should UDP socket reading threads operate at elevated priority	false	optional
udpFlowSocketsSize	Controls the read socket buffer size of the UDP layer flow control sockets	131071	optional
udpLocalWriteSocketSize	Controls the write socket buffer size of the local UDP sockets (Agent to Server on same node)	1024000	optional
udpMulticastBufferSize	Controls the read socket buffer size of the UDP multicast sockets	262142	optional
udpOutQsPriority	Turns on/off Priority weight-based for output queues (0 round-robin no priority - old logic, 1 round-robin new logic, 2 and higher is factor of priority)	0	optional
udpQueueSize	UDP transport layer receive queue size	100	optional
udpResendLostPackets	UDP transport layer resend lost packets	true	optional
udpSendQueueSize	UDP transport layer send queue size	50	optional
udpStatsReportInterval	UDP transport layer stats reporting interval	60000	optional

Cache

attribute	description	values	required
blobCacheMem	Size (in Mb) of blob index page cache	0	optional
leafCacheMem	Size (in Mb) of leaf index page cache	50	optional
nodeCacheMem	Size (in Mb) of non-leaf index page cache	100	optional
mysqlCacheCheckPeriod	Time to wait (ms) between checking if any cached MySQL connections can be closed	10000	optional
mysqlCacheTimeoutPeriod	Time to wait (ms) before closing a cached MySQL connection	60000	optional
mysqlConnectionCacheSize	Number of MySQL connections to hold in cache	10	optional

SSH

attribute	description	values	required
SSHidentityfile	location of identity file (private key) on Thor master	\$HOME/.ssh/id_rsa	optional

Configuring the HPCC Systems® Platform
Configuration Manager Advanced View

attribute	description	values	required
SSHusername	Username to use when running Thor slaves	hpcc	optional
SSHpassword	Fixed password - only required if no identity file present NB **insecure**		optional
SSHtimeout	Timeout in seconds for SSH connects	0	optional
SSHretries	Number of times to retry failed connect	3	optional

Ports

attribute	description	values	required
multicastPort	Port used for multicast (server->slave) data	8887	optional
serverFlowPort	Port used for UDP (slave->server) server flow control messages	9000	optional
dataPort	Port used for UDP (slave->server) data	9001	optional
clientFlowPort	Port used for UDP (slave->server) client flow control messages	9002	optional

Debug

attribute	description	values	required
alwaysFailOnLeaks	Fail any query that triggers internal consistency check or row leak errors	false	optional
checkCompleted	Check pending replies when agent gets a retry request	true	optional
dafilesrvLookupTimeout	Maximum time (in milliseconds) dafilesrv will wait before timing out the first time through the list	10000	optional
defaultConcatPreload	Default concat preload	0	optional
defaultFetchPreload	Default fetch preload	0	optional
defaultFullKeyedJoinPreload	Default full keyed join preload	0	optional
defaultKeyedJoinPreload	Default keyed join preload	0	optional
defaultParallelJoinPreload	Default parallel join preload	0	optional
defaultPrefetchProjectPreload	Default prefetch value for PROJECT,PREFETCH activity	10	optional
dolbytiDelay	Enables the IBYTI delay logic in the agents.	true	optional
enableHeartBeat	Enable HeartBeat messages to clients	true	optional
fastLaneQueue	special fast-lane queue for simple queries.	true	optional
forceStdLog	Force log output to stderr even when redirected to null	false	optional

attribute	description	values	required
ignoreMissingFiles	Ignore missing files	false	optional
indexReadChunkSize	Break up results from indexRead (and other remote activities) every N bytes	60000	optional
initlbytiDelay	Initial time (in milliseconds) a secondary agent will wait for an IBYTI packet from a primary peer.	50	optional
jumboFrames	Set to true if using jumbo frames (MTU=9000) on the network.	false	optional
linuxYield	Yield to scheduler in some tight loops. May help latency on uniprocessor machines	false	optional
maxBlockSize	Max size of block read from client socket	10000000	optional
maxLockAttempts	Number of retries to get lock for global queries	5	optional
memoryStatsInterval	Interval (in seconds) between reports on Roxie heap usage	60	optional
memTraceSizeLimit	Generate stacktrace whenever a request is made for a row larger than this threshold (0 to disable)	0	optional
parallelAggregate	Number of parallel threads to use for in-memory aggregate processing. Set to 0 to use one per CPU, 1 to disable parallel processing of in-memory aggregates	0	optional
perChannelFlowLimit	Number of pending queries permitted per channel (per active activity) before blocking	10	optional
preabortIndexReadsThreshold	Use seek to precheck keyed limits (i.e. assume ,COUNT) on index reads if limit greater than this value	100	optional
preabortKeyedJoinsThreshold	Use seek to precheck limits on keyed joins if limit greater than this value	100	optional
simpleLocalKeyedJoins	Enable single-threaded local keyed joins	true	optional
socketCheckInterval	Interval (in milliseconds) between checks that client socket is still open	5000	optional
systemMonitorInterval	How often to send an "alive" message	60000	optional

Additional Roxie Configuration items

Persistent Connections to Roxie

Persistent connections can cause backwards compatibility issues with non-standard HTTP clients. You can disable this feature by setting *maxHttpConnectionRequests* to either 0 or 1.

In addition, persistent connections can affect Roxie load balancing. If this is a concern, you can adjust *maxHttpConnectionRequests* and *maxHttpKeepAliveWait* values to optimize how long all the transactions from a particular persistent connection go to a single node rather than being distributed across the cluster.

Add Servers to Roxie Farm

To add servers to a Roxie farm

1. Select the **Roxie Cluster - myroxie** (default) from the Navigator window on the left side.
2. Select the **Servers** tab.
3. Right-click on **Roxie Servers**.
4. Select **Reconfigure Servers**.
5. Press the **Add Hardware** button.
6. Enter the values for the new server(s) in the dialog then press OK.

All configured servers are then used when you create a port to listen on.

NOTE If working with an older environment file this process has changed. You no longer have to specify for a server to use a specific port.

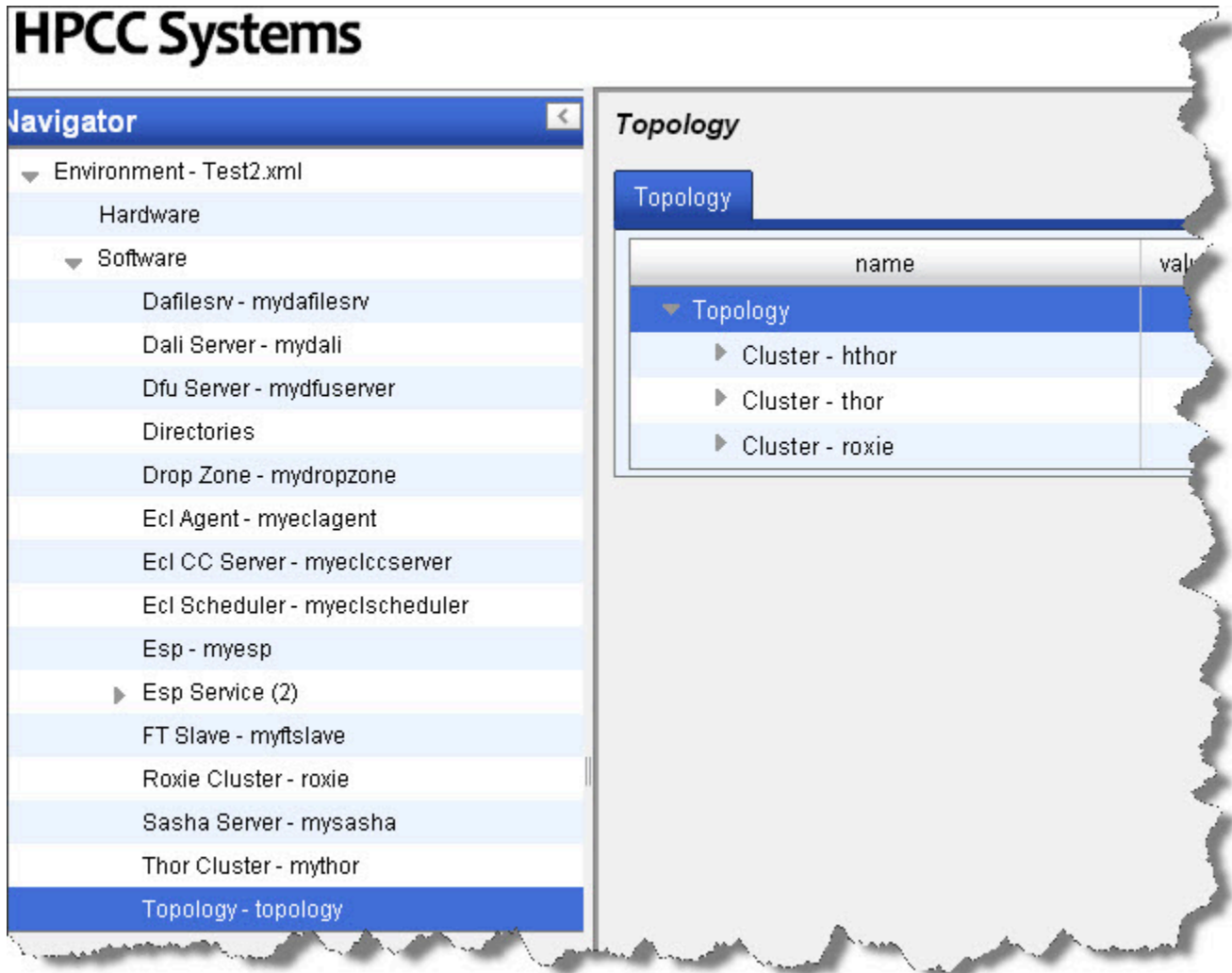
Redundancy

Roxie can be configured to utilize a few different redundancy models.

- Simple Redundancy - One channel per agent. Most commonly used for a single node Roxie.
- Full Redundancy - More agents than the number of channels. Multiple agents host each channel.
- Overloaded Redundancy - There are multiple channels per agent.
- Cyclic Redundancy - Each node hosts multiple channels in rotation. The most commonly used configuration.

Topology

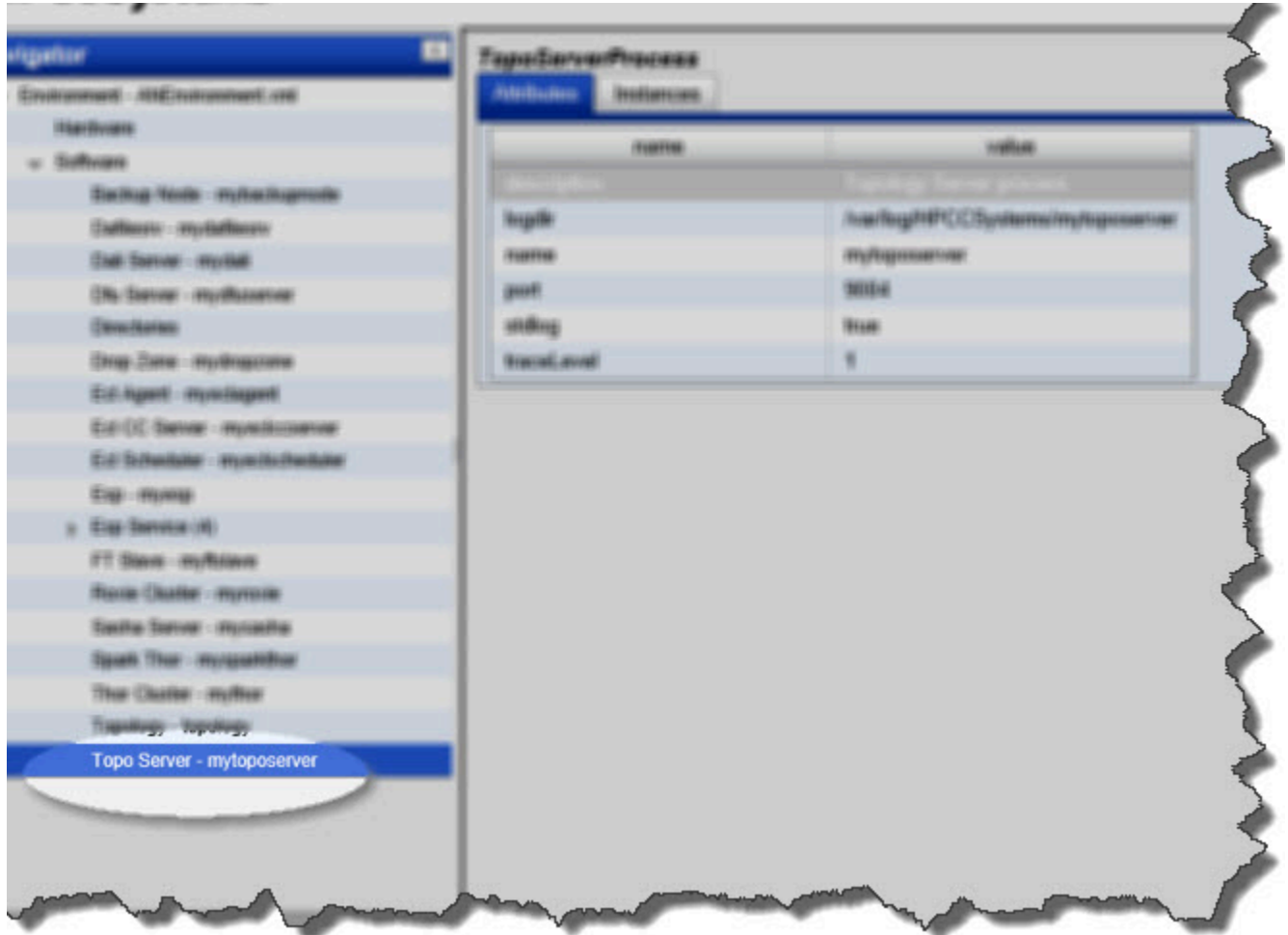
This section describes the topology tab.



Attribute name	Definition
<i>Topology</i>	describes the system topology
<i>Cluster - thor</i>	describes the Thor clusters
<i>Cluster - hthor</i>	describes the hthor clusters
<i>Cluster - roxie</i>	describes the Roxie clusters

Topo Server

This TopoServer Process is configured through the *Topo Server - mytoposerver* from configuration manager. You define the instances and then set the configuration attributes.



attribute	description	values	required
name	Name for this process	mytoposerver	required
description	Description for this process	Topology Server process	optional
port	Base port to use for master	9004	optional
traceLevel	Trace level for logging	1	optional
stdlog	Standard log	true	optional
logdir	Default log directory	/var/log/HPCCSystems/mytoposerver	optional

Distribute Configuration Changes to all Nodes

Once your environment is set up as desired, you must copy the configuration file to the other nodes.

1. If it is running,, stop the system



Be sure system is stopped before attempting to copy the Environment.xml file.

2. Back up the original environment.xml file

```
# for example
sudo -u hpcc cp /etc/HPCCSystems/environment.xml /etc/HPCCSystems/environment.bak
```

Note: the "live environment.xml file is located in your **/etc/HPCCSystems/** directory. ConfigManager works on files in **/etc/HPCCSystems/source** directory. You must copy from this location to make an environment.xml file active.

3. Copy the NewEnvironment.xml file from the source directory to the /etc/HPCCSystems and rename the file to environment.xml

```
# for example
sudo -u hpcc cp /etc/HPCCSystems/source/NewEnvironment.xml /etc/HPCCSystems/environment.xml
```

4. Copy the /etc/HPCCSystems/environment.xml to the /etc/HPCCSystems/ on every node.

5. Restart the HPCC Systems platform

You might prefer to script this process, especially if you have many nodes. See the Example Scripts section in the Appendix of the *Installing_and_RunningtheHPCCPlatform* document. You can use the scripts as a model to create your own script to copy the environment.xml file out to all your nodes.