ZyLAB Platform Cluster Installation and Configuration Manual

Contact Us

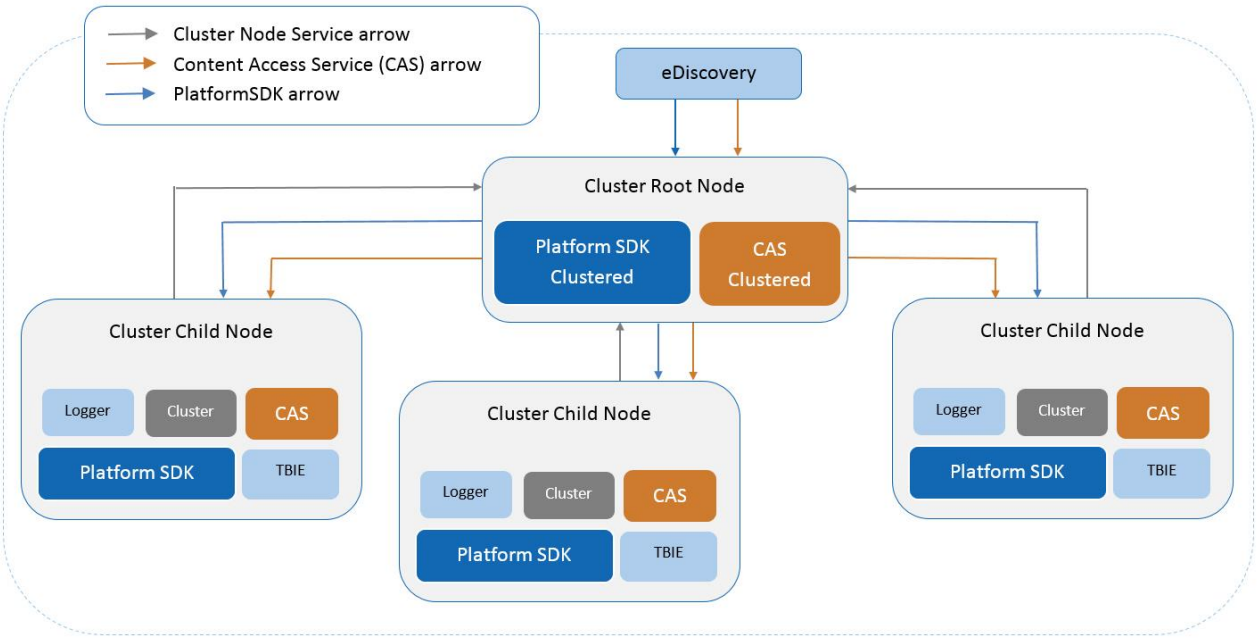
For full contact details, visit the ZyLAB website - https://www.zylab.com

For support, submit a ticket on - https://help.zylab.com

Contents

Overview ZyLAB Platform Cluster

The ZyLAB Platform Cluster provides distributed Indexing, Searching and Content Access. A cluster consists of one root node and one or more child nodes. On each node, the ZyLAB Information Management Platform will be installed. The root node hosts the clustered services (Platform SDK and Content Access).



Cluster Root Node

The cluster root node hosts the clustered services (Platform SDK and Content Access).

On the root node, all information about the cluster will be stored:

- Cluster configuration;
- Clustered index templates;
- All other configurations (CAS, TBIE, Platform templates)

ZyLAB Cluster Management Service

The ZyLAB Cluster Management Service enables you to distribute the workload of searching (PlatformSDK) and content access (CAS). eDiscovery uses the ClusterService host (root) to communicate with the PlatformSDK Service host (node) and WinService host (node).

The Cluster Management Service is responsible for handling federated index management, federated search requests and federated content access requests for cluster based systems. It should be used by clients of the cluster in order to make changes to the cluster configuration, and retrieve information about

the cluster, such as:

- State of the child nodes;
- Performance counters and other statistics, like number of executed queries, longest running query etc.

Cluster Node (Root) Service

The Cluster Node (Root) Service acts as a transparent proxy. The service merges, sorts or aggregates results from multiple child nodes. For example, when a search is performed across clustered indexes, all child nodes will be searched and the results will be merged, taking into account paging and sorting. However, the service will not modify the results from the child nodes, like changing field values, filtering documents, etc. In addition, for Content Access, extracted text or generated tiff files will not be modified in the clustered environment.

From the functional perspective it:

- Will establish the communication between root node and child nodes and provide:
  - Node registration and configuration;
  - Aggregate logs from child nodes;
  - Monitor and manage child nodes;
  - Assign roles to child nodes (some nodes might be allocated for indexing or searching or doing content access);
- Manage location of indexes and index replication;
- Coordinates how requests (search/indexing) will be distributed between child nodes (load distribution):
  - Round Robin;
  - Index size;
  - Based on document properties

Cluster Node Service

The root node uses the Cluster Node Service to control the child node(s):

- Discover child node(s) and perform automatic registration;
- Configure child node(s) and check prerequisites (version of installed software, disk space, etc.);
- Retrieve status, logs and other information from the child node(s).

Installation and Configuration

In this example installation, we will set up a cluster environment with two nodes: ZySystem1 (root node) and ZySystem2 (child node). The computer names should be ZySystem1 and ZySystem2.

> **Note**: Once the cluster environment has been installed, older EDRM matters cannot be opened any longer!
> **Note**: If you use a local or virtual machine as one of the cluster nodes, open a port for it.

For the ZyLAB clustered environment, the following ports are needed:

ZyLAB Information Management Service	Service	4112	tcp	
Platform SDK Service	Service	4115	tcp	
Platform Cluster	Service	4119	udp	Logging
Platform Cluster	Service	4117	tcp	Root node services and Cluster management service.

ZyCLUSTER Management Tool

Instructions

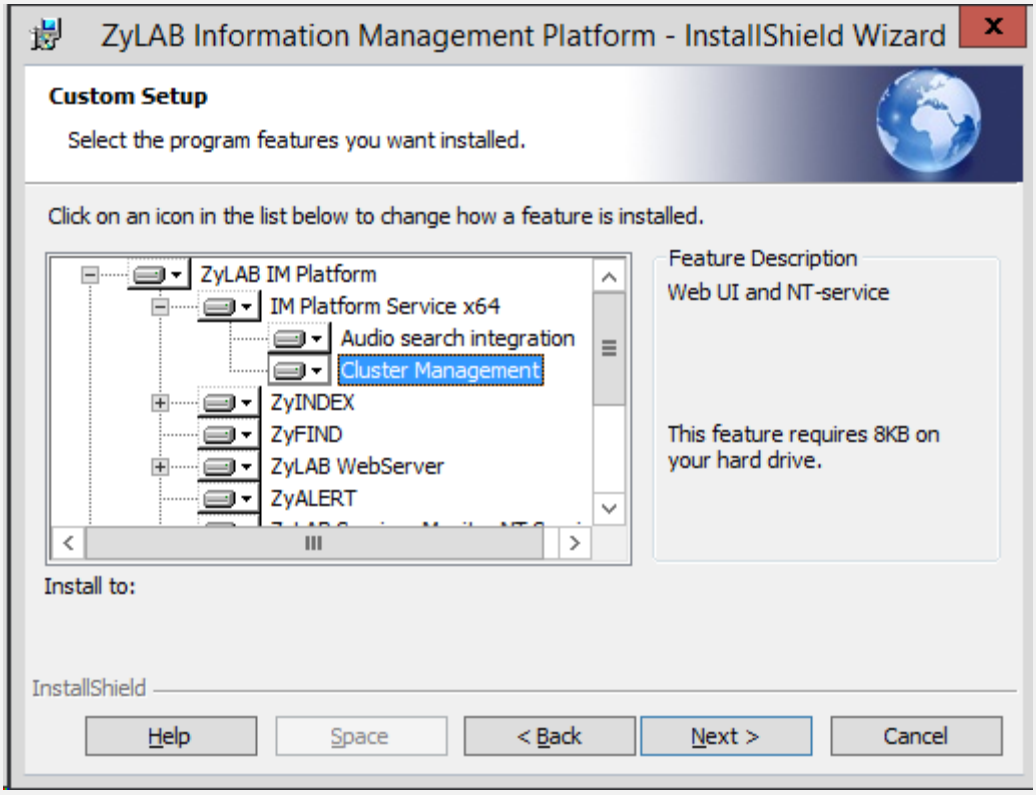
1. Install the ZyLAB Information Management Platform on each machine. For more information, see the ZyLAB Information Management Platform Installation and Upgrade Manual. The required cluster software is installed by default as part of the IM Platform Service. The service is not installed though and the features are not enabled.

   The installed parts for the cluster environment are located here:

   - DRIVE:\ZyLAB Data\ClusterService
   - DRIVE:\Program Files\ZyLAB\Information Management Platform\Services\Bin

> The feature Cluster Management will install the ZyLAB Platform ClusterService (ZyClusterManService) and the Cluster Management Web Interface.

The ZyLAB Cluster Management Service is responsible for handling federated index management, federated search requests and federated content access requests for cluster based systems.
The service should only be installed on the root machine (root node). The Cluster Management feature can be disabled, thus avoiding that (candidate) children get the service installed. The service will be installed on port 4117.



2. Install eDiscovery on ZySystem1. For more information, see the ZyLAB eDiscovery Installation and Upgrade Manual.
3. On ZySystem1, open C:\Program Files\ZyLAB\Information Management Platform\Services\Bin\ZyLAB.Platform.Management.Cluster.exe
   The *ZyCLUSTER Management Tool* (page 7) will appear.
   - Select Root to install the Cluster Service (only on the root node).
   - Define the Working folder. Use a network path that will be visible for all nodes.
   - Replace localhost with ZySystem1 for Cluster, CAS, SDK, and Management.
   - Select Child. Replace localhost with ZySystem1 for Root URI.
   - Click Apply.

Check that the ZyLAB Platform ClusterService is installed and started. During an upgrade of the Information Management Platform, the Cluster Service stops (if it is running), and restarts once the upgrade is done. Changes are logged. The tool preserves previous configuration files (copies are made with .bak extension).

4. On ZySystem2, open C:\Program Files\ZyLAB\Information Management Platform\Services\Bin\ZyLAB.Platform.Management.Cluster.exe
   The ZyCLUSTER Management Tool will appear.
   - Select Child. Replace localhost with ZySystem1.
   - Select Check (Connection status becomes 'live').
   - Click Apply.
5. On ZySystem1, open C:\Program Files\ZyLAB\Information Management Platform\Services\Bin\ZyLAB.Platform.Management.Cluster.exe
   - Select Tools.
   - Select Child nodes overview.

Check that the status of CAS and SDK is listed on both nodes.

Cluster Configuration

Instructions

1. Open the "C:\Program Files (x86)\ZyLAB\EDiscovery\LegalReview\Review_web\Web.config" file.
2. Modify the ContentCache to a shared location (make sure to share the current ContentCache which is in C:\ProgramData\Zylab\LegalReview\ContentCache (this folder is created after Legal Review Web has been opened once)) like this:

   ```
   <add key="ContentCachePath" value="\\Zysystem1\ContentCache" />
   ```
3. Modify the ContentAccessServiceUri setting:

   ```
   <add key="ContentAccessServiceUri"
   value="net.tcp://Zysystem1:4117/ContentAccess" />
   ```
4. Open the "C:\Program Files (x86)\ZyLAB\EDiscovery\LegalProductionService\ZyLAB.LegalProduction.WinService.exe.config" file.
5. Modify the ContentAccessServiceUri setting:

   ```
   <add key="ContentAccessServiceUri"
   value="net.tcp://Zysystem1:4117/ContentAccess" />
   ```
6. Change the port in the "C:\Program Files (x86)\ZyLAB\EDiscovery\LegalReview\Review_Standalone\ZyLAB.EDiscovery.Review.Standalone.WinService.exe.config" file from 4115 to 4117:

   ```
   <add key="PlatformSDKServiceUri"
   value="net.tcp://Zysystem1:4117/ZyIMAGE" />
   ```
7. The templates for ediscoveryLR and ediscoveryLRAudio have to be copied from the ZyINDEX Templates on the root node to the child node's ZyINDEX Templates folder.

8. Make sure the Audio Search software is installed on the child node(s). If it is not installed yet, do it now. For more information, see the ZyLAB Information Management Platform Installation and Upgrade Manual. After installing Audio Search, restart the Platform Service for changes to take effect (ZyLAB Information Management Service x64).
9. Before running a job in Legal Processing, provide in the Global Settings a proper Shared Folder Path.

   \\ProgramData\ZyLAB\EDiscovery\SharedData, or:
   \\ZySystem1\c$\ProgramData\ZyLAB\EDiscovery\SharedData, should be shared and used here, like:
   \\ZySystem1\SharedData
10. Restart ZyLAB Legal Review Standalone Service.
    Restart ZyLAB Legal Production Service 2.
    Restart ZyLAB EDiscoveryProcessing Service.

Child Node(s)

Configure all the child node(s) and the root node if it is used as a child node too.

Instructions

1. Create folders
   - \\ZyLAB Data\SearchEngine\Templates\ediscoveryLR
   - \\ZyLAB Data\SearchEngine\Templates\ediscoveryLRaudio
2. Copy files from \\ZyLAB Data\SearchEngine\Templates\Default to \\ZyLAB Data\SearchEngine\Templates\ediscoveryLR
3. Copy files from \\ZyLAB Data\SearchEngine\Templates\Audio to \\ZyLAB Data\SearchEngine\Templates\ediscoveryLRaudio
4. Restart the Information Management Service on each Child Node.
5. Restart the ClusterService on the Root Node.

Cluster Environment Verification

1. Open Legal Processing and run a matter with file system collector.
   Check the following:
   - ZySystem1 has index Cluster1.
   - ZySystem2 has index Cluster2.
2. Open the matter in Legal Review.
   Check the following:
   - In Legal Review, search for documents. Documents from both indexes will be shown.

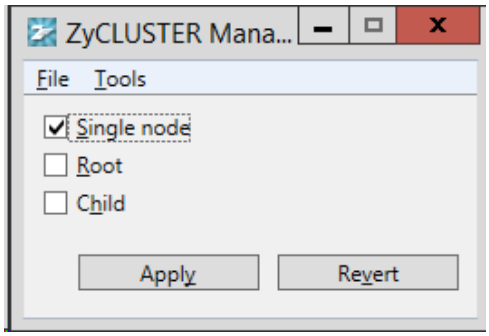
ZyCLUSTER Management Tool

Instructions

1. To enable the cluster you need to run the ZyCLUSTER Management Tool.

   "\\Program Files\ZyLAB\Information Management Platform\Services\Bin\ZyLAB.Platform.Management.Cluster.exe"

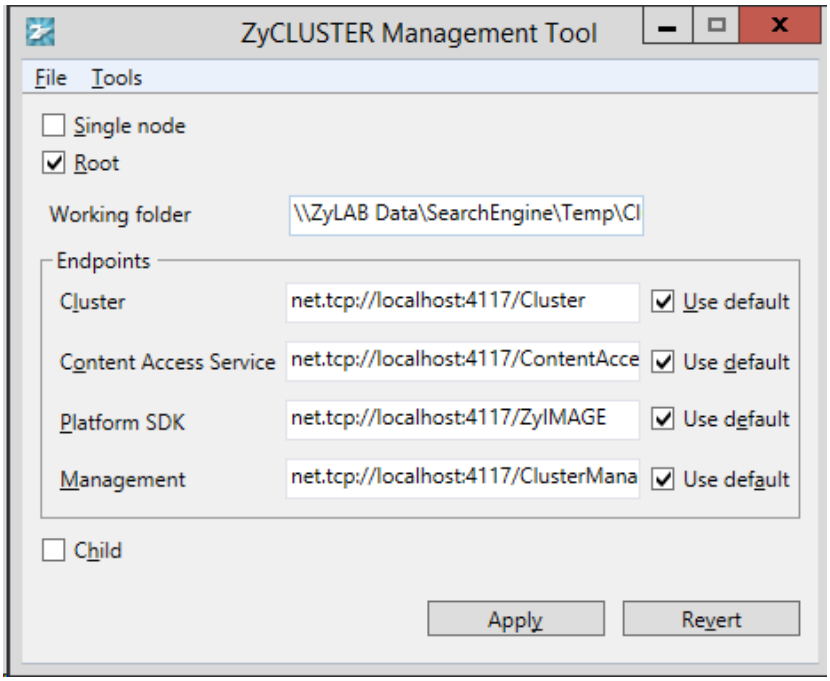
   By default, it looks like this:



   Single Node: This server is not part of a cluster.

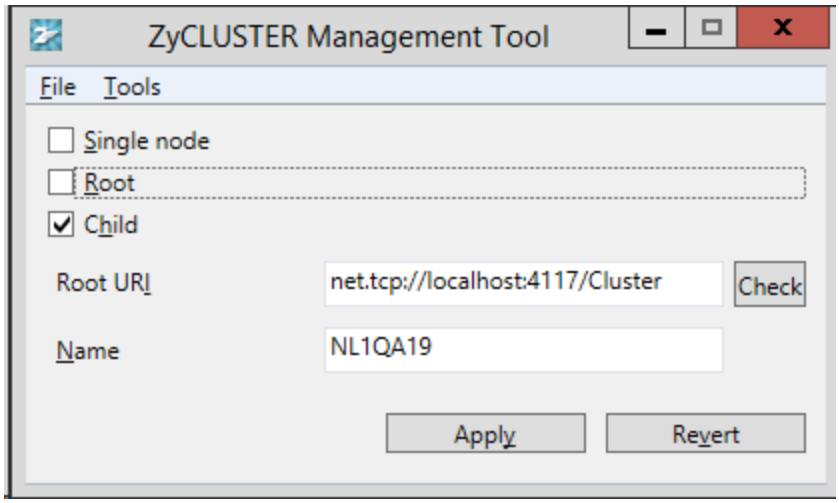
2. Select Root and configure the Endpoints.

   Root: This server hosts the Cluster Management Service (only 1 per cluster) on the root node.



3. Define the Working folder. Use a network path that will be visible for all nodes.
4. A root node will expose 4 APIs (endpoints):
   - **Cluster**: used for cluster related communication (between root and children). Children should point to this URI (the root can have 'localhost' there, the children should have the root server name or IP there).
   - **Content Access Service**: Entry point for applications that want to use clustered CAS (like Legal Review or Legal Production).
   - **Platform SDK**: Entry point for applications that want to use the PlatformSDK Service in a clustered environment (e.g. searching, hits, indexing, near duplicates). This can only be used by Legal Review at this moment.
   - **Management**: Entry point for the Cluster Management Web Interface that can be used for basic administrative tasks and diagnostics of the cluster.
5. Select Child and configure Root URI and Name.

   Child: This server is a child node (a server can be both a root and a child, but there can be only one root).


ZyCLUSTER Management Tool
File
Tools
Single node
Root
Child
Root URI
net.tcp://localhost:4117/Cluster
Check
Name
NL1QA19
Apply
Revert

Usage with ZyLAB eDiscovery

How to enable usage in Legal Review (Standalone)

Changes in the ZyLAB.EDiscovery.Review.Standalone.WinService.exe.config

Change the address and port in the PlatformSDKServiceUri:

```
<add key="PlatformSDKServiceUri"
value="net.tcp://SERVERNAME:4117/ZyIMAGE" />
```

How to enable usage in Legal Production

Config C:\Program Files (x86)\ZyLAB\EDiscovery\LegalProductionService\ZyLAB.LegalProduction.WinService.exe.config

Changes:

```
<add key="PlatformSDKServiceUri"
   value="net.tcp://SERVERNAME:4117/ZyIMAGE" />
<add key="ContentAccessServiceUri"
   value="net.tcp://Zysystem1:4117/ContentAccess" />
```

Things to keep in mind

1. Legal Review asks ContentAccessService to render documents to PNG directly to the ContentCache of Legal Review. This means that the nodes in the cluster must have write access to the ContentCache, e.g. like this:

   ```
   <add key="ContentCachePath" value="\\SERVERNAME\ContentCache" />
   ```

2. The indexing nodes need the eDiscovery Index Templates, which are installed by EDRM and therefore not present by default. These need to be copied to all nodes!

Cluster Management Web Interface

With the Cluster Management Web Interface you get an overview of the current state of the cluster. You can also perform simple administrative tasks such as running diagnostics, enabling or disabling specific nodes, and scheduling actions on indexes. This environment is intended for Administrators to monitor and manage the cluster and get an insight in its status.

Administrators can monitor:

Active nodes and their status (HD space, CPU usage, MEM usage etc.)

Activity per node

Active indexes on clusters

Current outstanding requests

Manage:

Enable/disable services (Machine tab)

Download Events

Download Configuration

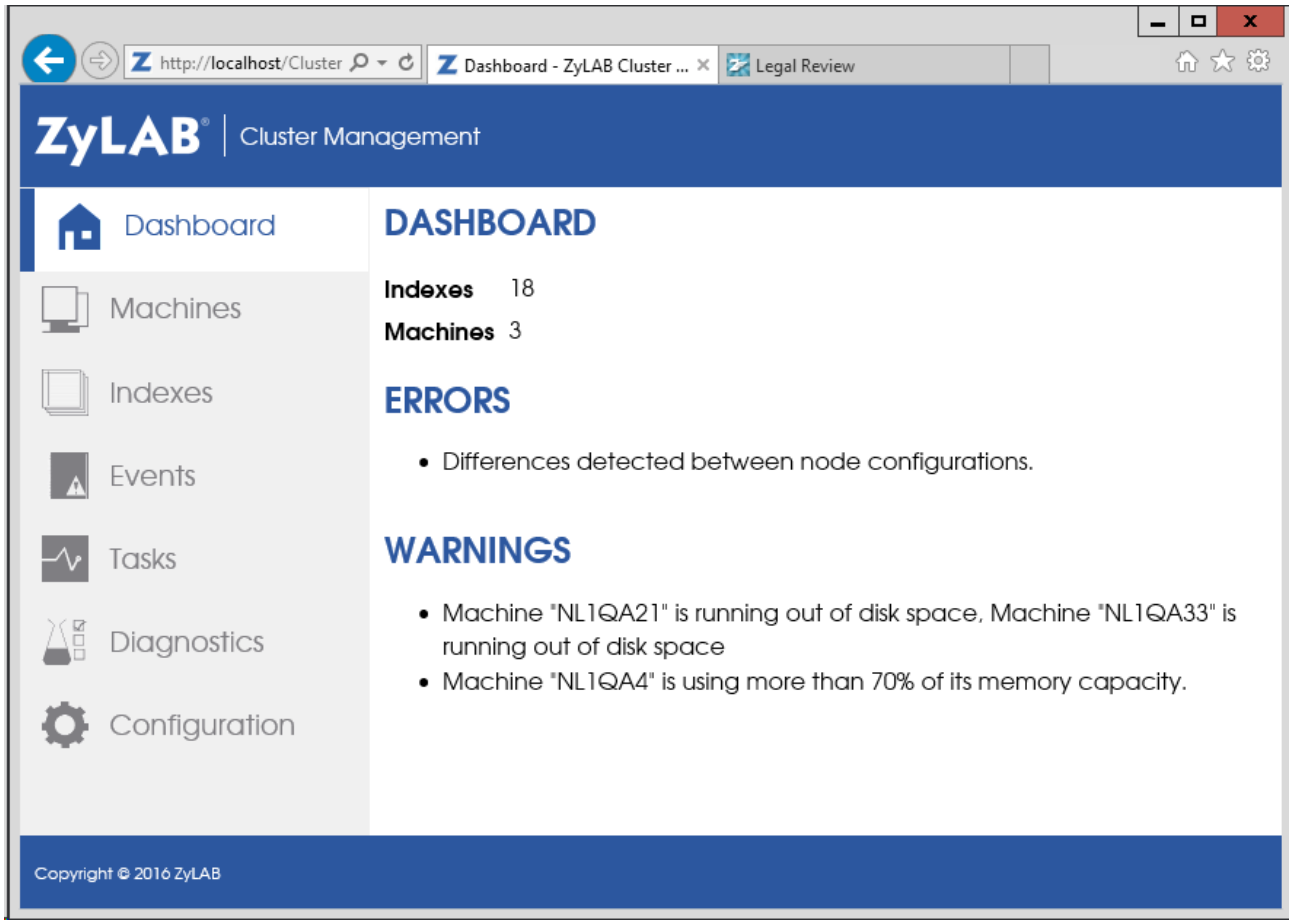
Open the Cluster Management Web Interface by navigating to:

http://host/ClusterWebUI

Dashboard

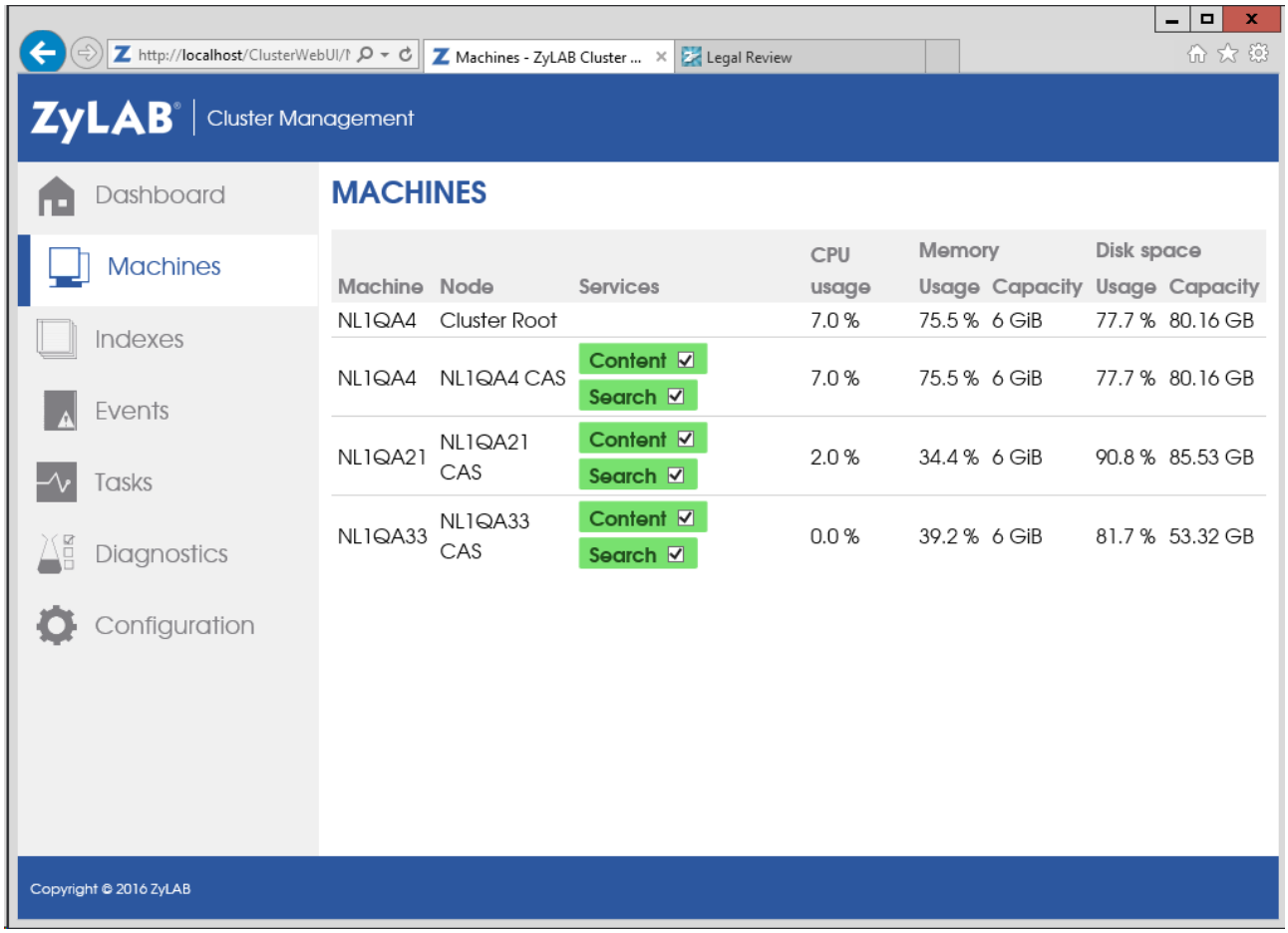
Dashboard

On this page, an overview of the number of machines and indexes is shown. Basic diagnostics are run and errors or warnings are shown (such as unreachable nodes, running out of disk space or high memory usage).



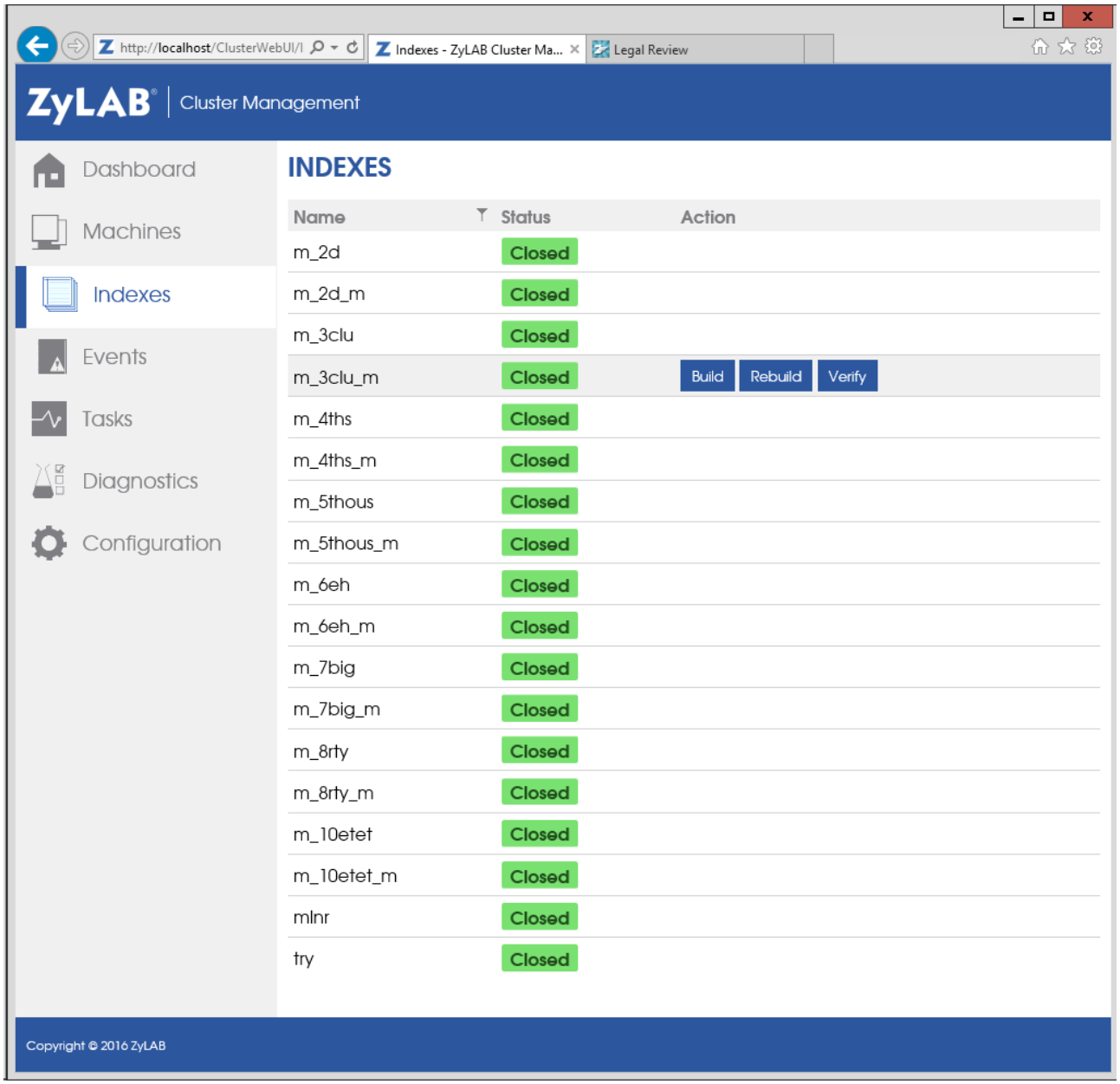
Machines

On this page, an overview of all the machines in the cluster is shown. For each machine the current CPU Usage, Memory (Usage and Capacity), and Disk Space (Usage and Capacity) are shown. It is also possible to disable services.



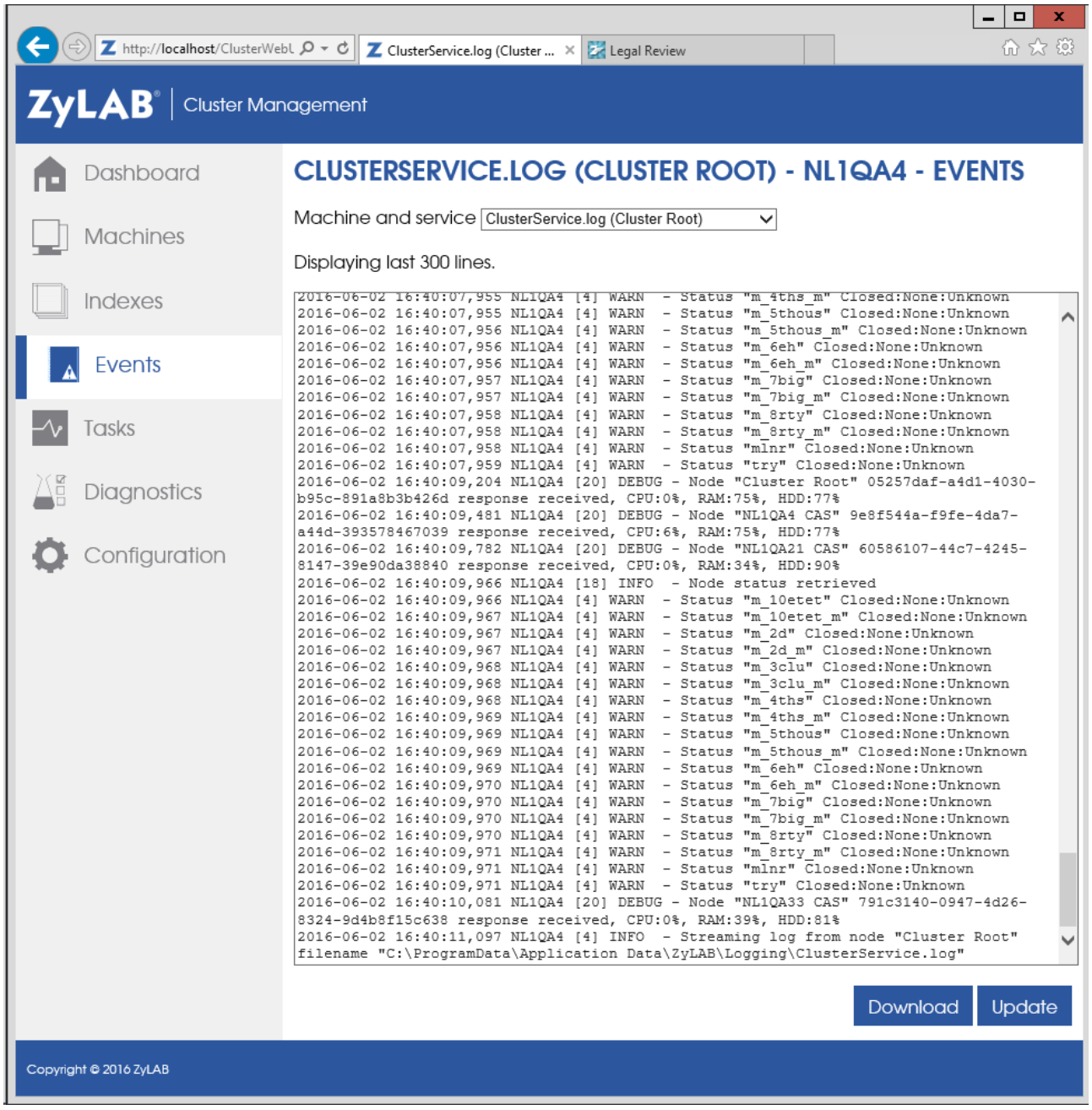
Indexes

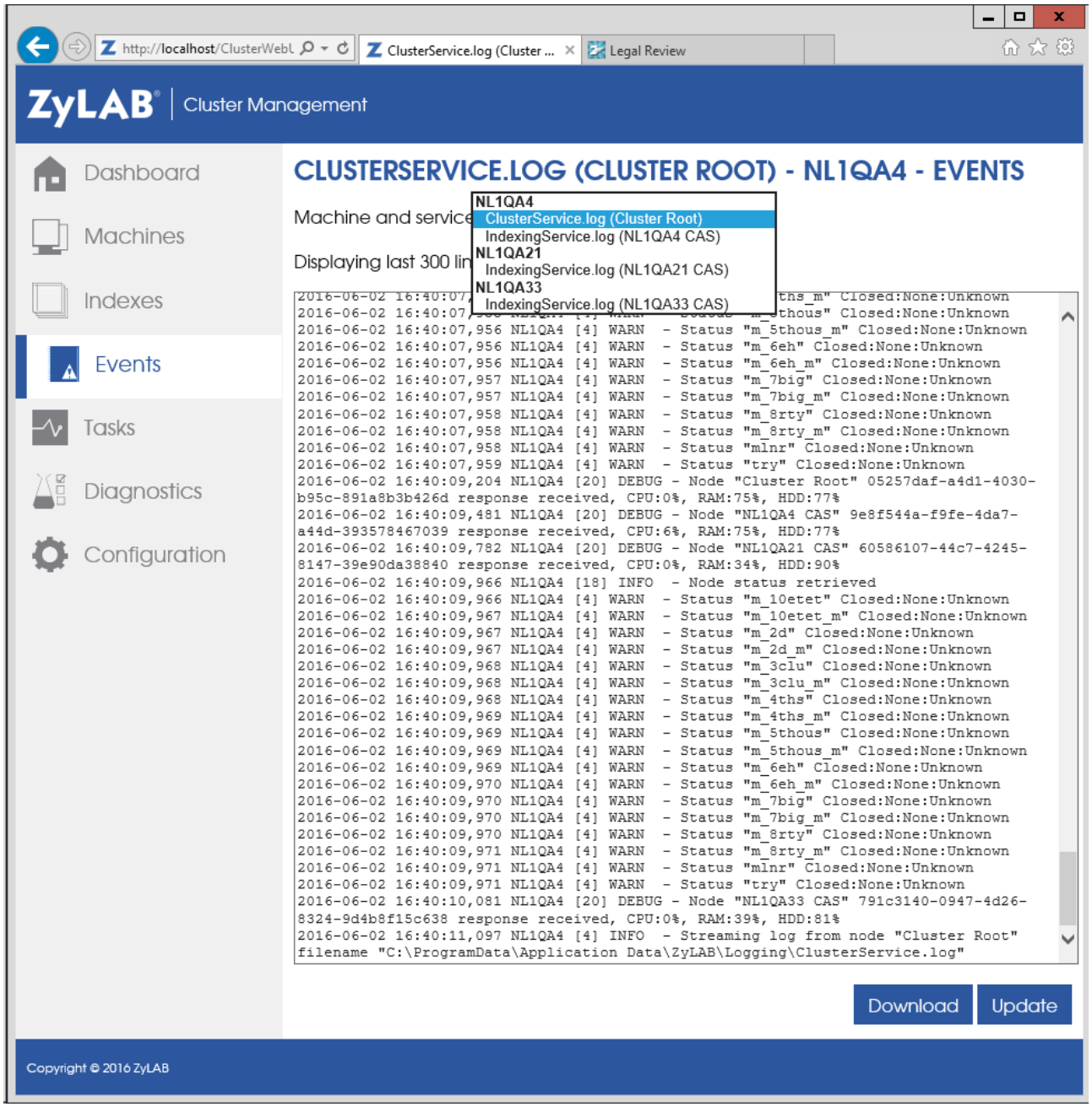
Shows a list of the indexes hosted by the cluster. Each index also shows its status and it is possible to schedule an action (Build, Rebuild or Verify). Results will be shown in the log.



Events

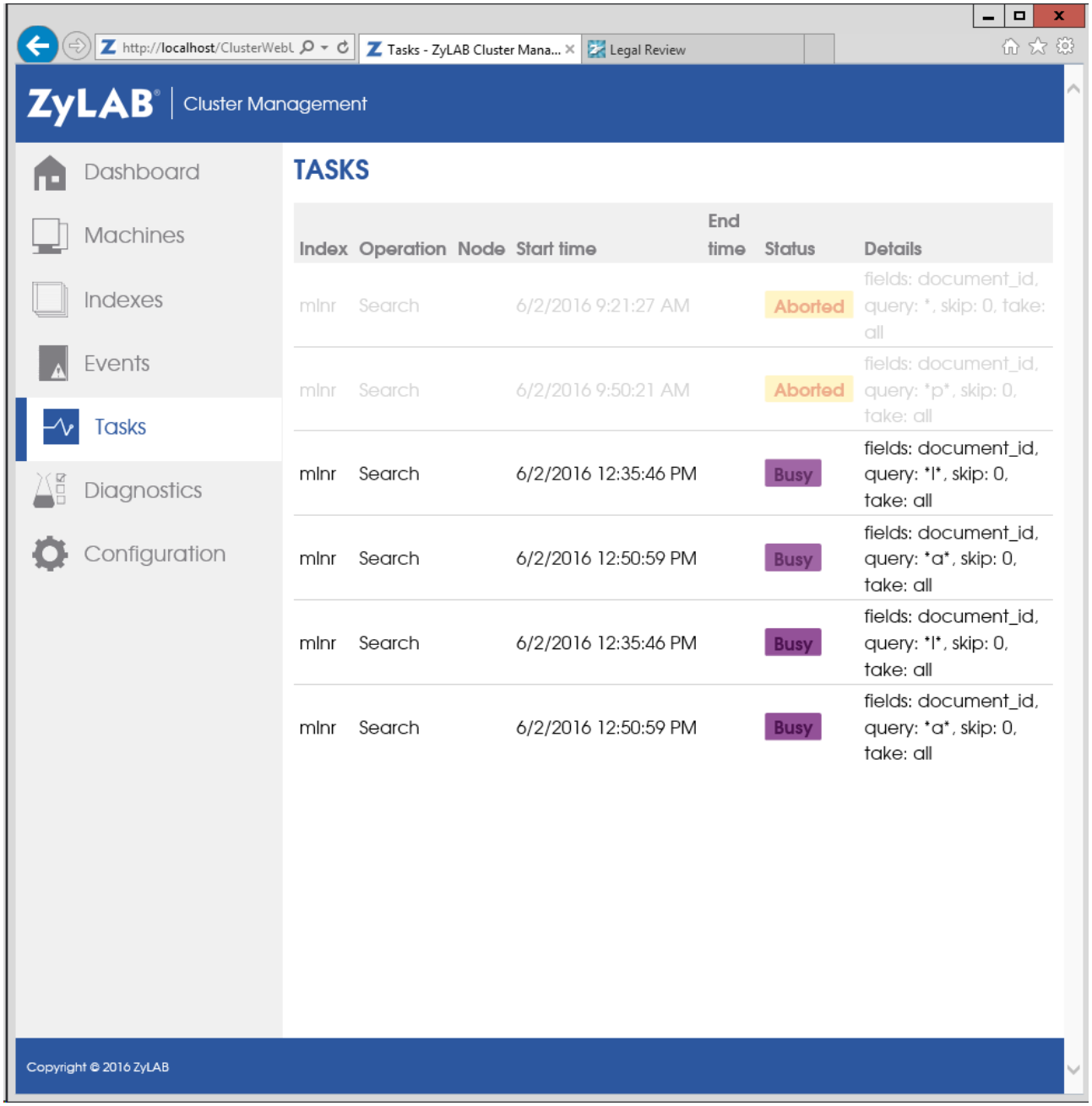
Shows the last 300 lines of the log file of a node. You can select a node and a log file to show from the dropdown list.




http://localhost/ClusterWebL
ClusterService.log (Cluster ...
Legal Review
ZyLAB | Cluster Management
Dashboard
Machines
Indexes
Events
Tasks
Diagnostics
Configuration
CLUSTERSERVICE.LOG (CLUSTER ROOT) - NL1QA4 - EVENTS
Machine and service
Displaying last 300 lin
NL1QA4
ClusterService.log (Cluster Root)
IndexingService.log (NL1QA4 CAS)
NL1QA21
IndexingService.log (NL1QA21 CAS)
NL1QA33
IndexingService.log (NL1QA33 CAS)
2016-06-02 16:40:07,956 NL1QA4 [4] WARN - Status "m_5thous_m" Closed:None:Unknown
2016-06-02 16:40:07,956 NL1QA4 [4] WARN - Status "m_6eh" Closed:None:Unknown
2016-06-02 16:40:07,956 NL1QA4 [4] WARN - Status "m_6eh_m" Closed:None:Unknown
2016-06-02 16:40:07,957 NL1QA4 [4] WARN - Status "m_7big" Closed:None:Unknown
2016-06-02 16:40:07,957 NL1QA4 [4] WARN - Status "m_7big_m" Closed:None:Unknown
2016-06-02 16:40:07,958 NL1QA4 [4] WARN - Status "m_8rty" Closed:None:Unknown
2016-06-02 16:40:07,958 NL1QA4 [4] WARN - Status "m_8rty_m" Closed:None:Unknown
2016-06-02 16:40:07,958 NL1QA4 [4] WARN - Status "mlnr" Closed:None:Unknown
2016-06-02 16:40:07,959 NL1QA4 [4] WARN - Status "try" Closed:None:Unknown
2016-06-02 16:40:09,204 NL1QA4 [20] DEBUG - Node "Cluster Root" 05257daf-a4d1-4030-b95c-891a8b3b426d response received, CPU:0%, RAM:75%, HDD:77%
2016-06-02 16:40:09,481 NL1QA4 [20] DEBUG - Node "NL1QA4 CAS" 9e8f544a-f9fe-4da7-a44d-393578467039 response received, CPU:6%, RAM:75%, HDD:77%
2016-06-02 16:40:09,782 NL1QA4 [20] DEBUG - Node "NL1QA21 CAS" 60586107-44c7-4245-8147-39e90da38840 response received, CPU:0%, RAM:34%, HDD:90%
2016-06-02 16:40:09,966 NL1QA4 [18] INFO - Node status retrieved
2016-06-02 16:40:09,966 NL1QA4 [4] WARN - Status "m_10etet" Closed:None:Unknown
2016-06-02 16:40:09,967 NL1QA4 [4] WARN - Status "m_10etet_m" Closed:None:Unknown
2016-06-02 16:40:09,967 NL1QA4 [4] WARN - Status "m_2d" Closed:None:Unknown
2016-06-02 16:40:09,967 NL1QA4 [4] WARN - Status "m_2d_m" Closed:None:Unknown
2016-06-02 16:40:09,968 NL1QA4 [4] WARN - Status "m_3clu" Closed:None:Unknown
2016-06-02 16:40:09,968 NL1QA4 [4] WARN - Status "m_3clu_m" Closed:None:Unknown
2016-06-02 16:40:09,968 NL1QA4 [4] WARN - Status "m_4ths" Closed:None:Unknown
2016-06-02 16:40:09,969 NL1QA4 [4] WARN - Status "m_4ths_m" Closed:None:Unknown
2016-06-02 16:40:09,969 NL1QA4 [4] WARN - Status "m_5thous" Closed:None:Unknown
2016-06-02 16:40:09,969 NL1QA4 [4] WARN - Status "m_5thous_m" Closed:None:Unknown
2016-06-02 16:40:09,969 NL1QA4 [4] WARN - Status "m_6eh" Closed:None:Unknown
2016-06-02 16:40:09,970 NL1QA4 [4] WARN - Status "m_6eh_m" Closed:None:Unknown
2016-06-02 16:40:09,970 NL1QA4 [4] WARN - Status "m_7big" Closed:None:Unknown
2016-06-02 16:40:09,970 NL1QA4 [4] WARN - Status "m_7big_m" Closed:None:Unknown
2016-06-02 16:40:09,970 NL1QA4 [4] WARN - Status "m_8rty" Closed:None:Unknown
2016-06-02 16:40:09,971 NL1QA4 [4] WARN - Status "m_8rty_m" Closed:None:Unknown
2016-06-02 16:40:09,971 NL1QA4 [4] WARN - Status "mlnr" Closed:None:Unknown
2016-06-02 16:40:09,971 NL1QA4 [4] WARN - Status "try" Closed:None:Unknown
2016-06-02 16:40:10,081 NL1QA4 [20] DEBUG - Node "NL1QA33 CAS" 791c3140-0947-4d26-8324-9d4b8f15c638 response received, CPU:0%, RAM:39%, HDD:81%
2016-06-02 16:40:11,097 NL1QA4 [4] INFO - Streaming log from node "Cluster Root" filename "C:\ProgramData\Application Data\ZyLAB\Logging\ClusterService.log"
Download
Update
Copyright © 2016 ZyLAB

Tasks

Displays the tasks in the cluster. This is done using the contents of the ClusterService log.



Diagnostics

On this page diagnostics can be run. The diagnostics and their current result is shown. Tests can be selected and executed. This is very similar to running unit tests in, for example, Visual Studio.

Diagnostics that can be run:

Are all nodes reachable?

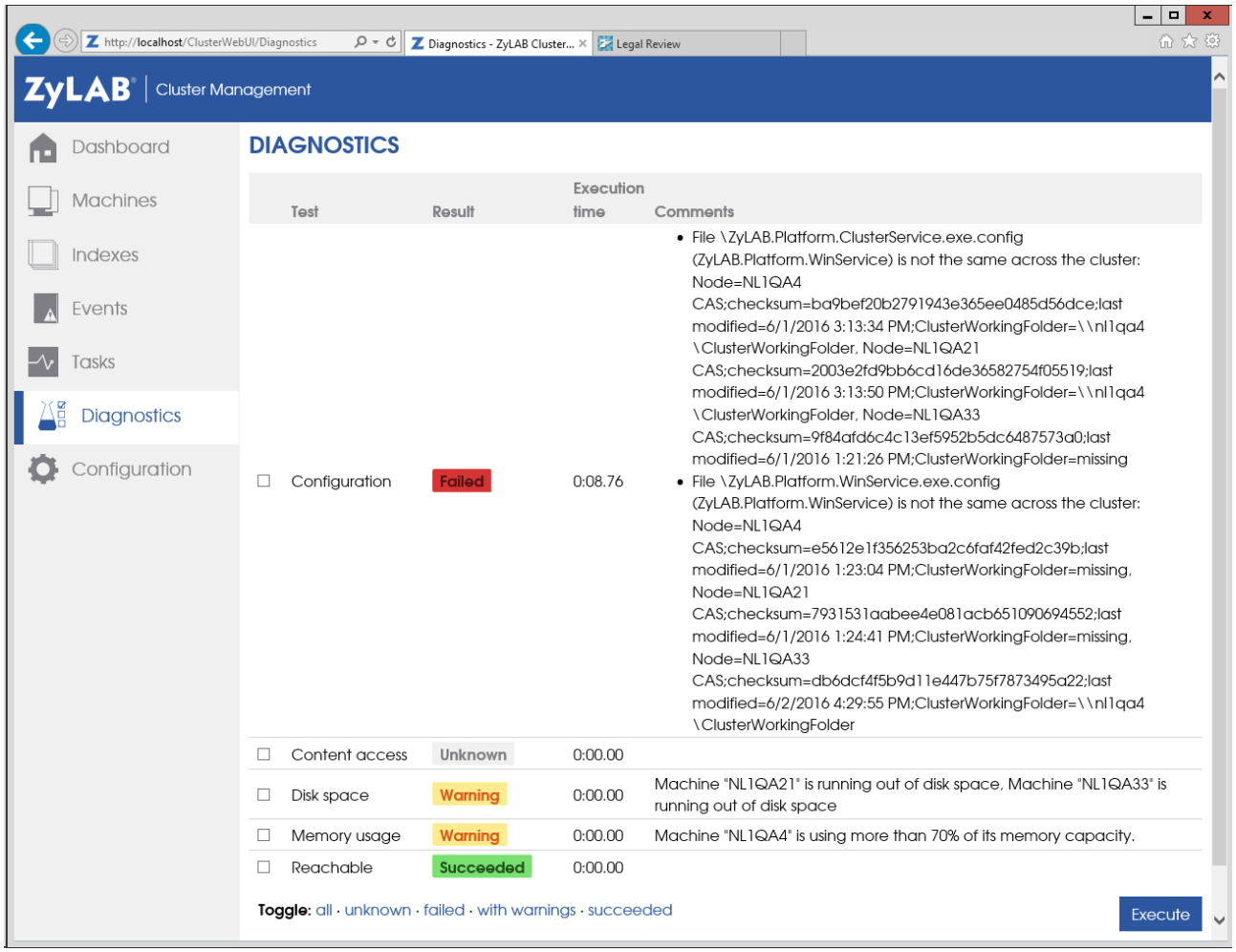
Is there enough disk space?

Is nothing running out of memory?

Is the configuration correct?

- Is the installed platform version the same on all machines?
- Are templates and workflows the same on all machines?
- Are screen resolution settings correct?
- Are date, time, time zone and regional settings correct?

Each check can return with success, warning or error.


http://localhost/ClusterWebUI/Diagnostics
Diagnostics - ZyLAB Cluster...
Legal Review
ZyLAB | Cluster Management
Dashboard
Machines
Indexes
Events
Tasks
Diagnostics
Configuration
DIAGNOSTICS
Test
Result
Execution time
Comments
Configuration
Failed
0:08.76
File \ZyLAB.Platform.ClusterService.exe.config (ZyLAB.Platform.WinService) is not the same across the cluster: Node=NL1QA4 CAS;checksum=ba9bef20b2791943e365ee0485d56dce;last modified=6/1/2016 3:13:34 PM;ClusterWorkingFolder=\\nl1qa4\ClusterWorkingFolder, Node=NL1QA21 CAS;checksum=2003e2fd9bb6cd16de36582754f05519;last modified=6/1/2016 3:13:50 PM;ClusterWorkingFolder=\\nl1qa4\ClusterWorkingFolder, Node=NL1QA33 CAS;checksum=9f84afd6c4c13ef5952b5dc6487573a0;last modified=6/1/2016 1:21:26 PM;ClusterWorkingFolder=missing
File \ZyLAB.Platform.WinService.exe.config (ZyLAB.Platform.WinService) is not the same across the cluster: Node=NL1QA4 CAS;checksum=e5612e1f356253ba2c6faf42fed2c39b;last modified=6/1/2016 1:23:04 PM;ClusterWorkingFolder=missing, Node=NL1QA21 CAS;checksum=7931531aabee4e081acb651090694552;last modified=6/1/2016 1:24:41 PM;ClusterWorkingFolder=missing, Node=NL1QA33 CAS;checksum=db6dcf4f5b9d11e447b75f7873495a22;last modified=6/2/2016 4:29:55 PM;ClusterWorkingFolder=\\nl1qa4\ClusterWorkingFolder
Content access
Unknown
0:00.00
Disk space
Warning
0:00.00
Machine "NL1QA21" is running out of disk space, Machine "NL1QA33" is running out of disk space
Memory usage
Warning
0:00.00
Machine "NL1QA4" is using more than 70% of its memory capacity.
Reachable
Succeeded
0:00.00
Toggle: all · unknown · failed · with warnings · succeeded
Execute

Configuration

Allows the download of a configuration file describing the entire configuration of a node. This can be used for manual comparison (for example, when the configuration diagnostics fail). The root node and all child nodes should be configured in the same way.