High-Availability on Linux

The Future

Lars Marowsky-Brée <lmb@suse.de>
Principal Engineer
SUSE Labs

Novell®

Agenda

- Introduction
- Linux-HA
- Pacemaker
- Cluster Developer Summit
  - Prague, 2008-09-30 – 2008-09-02
- Q&A

© Novell, Inc. All rights reserved.

Linux-HA project

- New governance structure
- The Cluster Resource Manager is now maintained as Pacemaker
- New release 3.0.0:
  - In beta as 2.99.x
  - Removal of lots of unmaintained code
    - `1176 files changed, 5346 insertions(+), **213851** deletions(-)`
  - Resource Agent updates
  - Focuses on infrastructure components again

© Novell, Inc. All rights reserved.

Pacemaker

- Not a fork, but a split from heartbeat
- Next release: 1.0.0
  - Currently in beta as 0.7.x
  - Focus on clean-ups and stability
- Features:
  - Supports both openAIS and heartbeat-2.99.x
  - New CLI
  - Timeout of failures
  - CIB syntax improved
  - Better documentation

© Novell, Inc. All rights reserved.

Pacemaker CIB changes

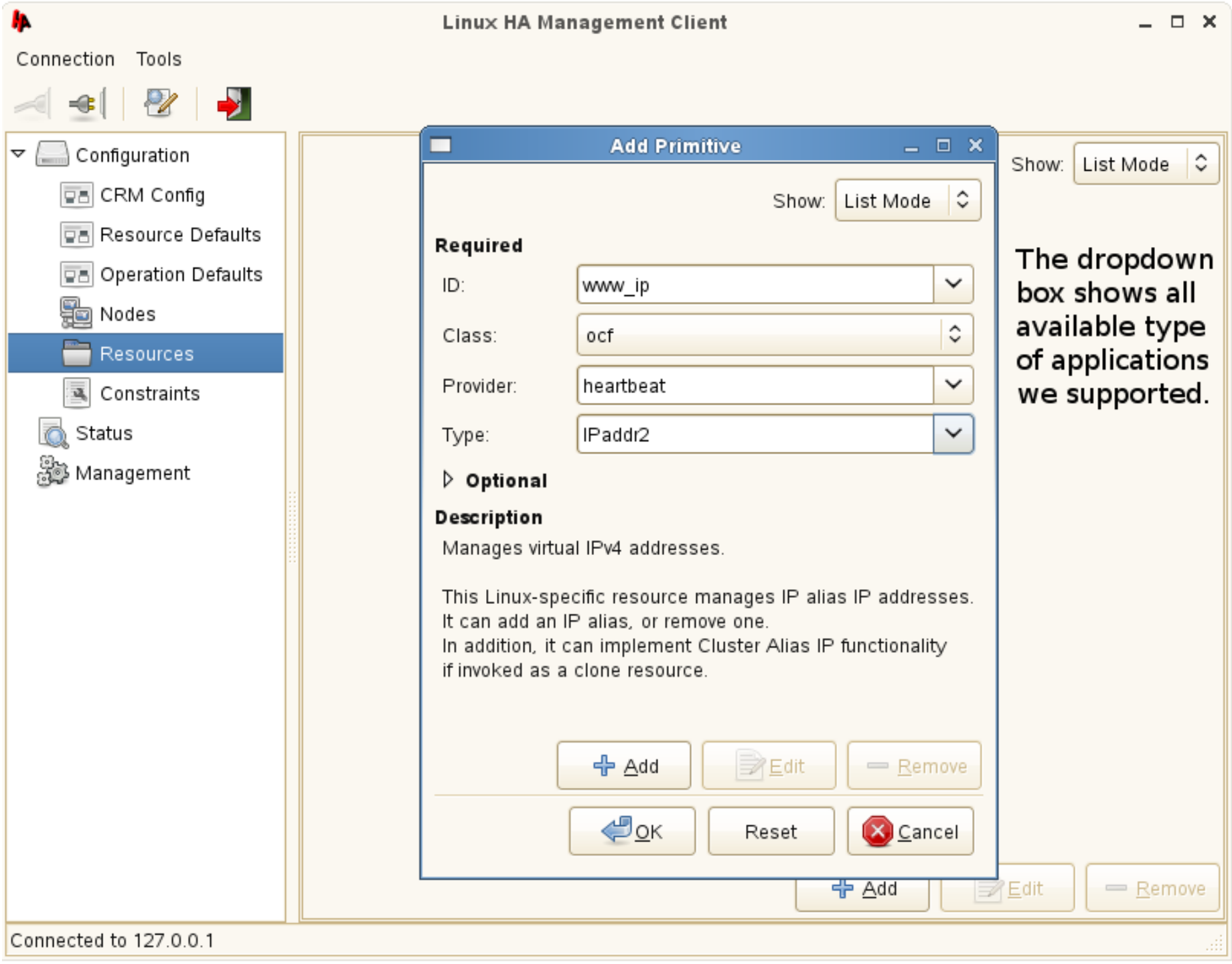
- CIB Shadow support
- RelaxNG Schema support
  - Allows better validation (IDREFs etc)
- Element references
  - Rules, instance_attributes, meta_attributes, operations
- M:N constraints
- Failure migration thresholds
- Syntax scaffolding reduced
- Smooth migration supported
- Supports remote connects

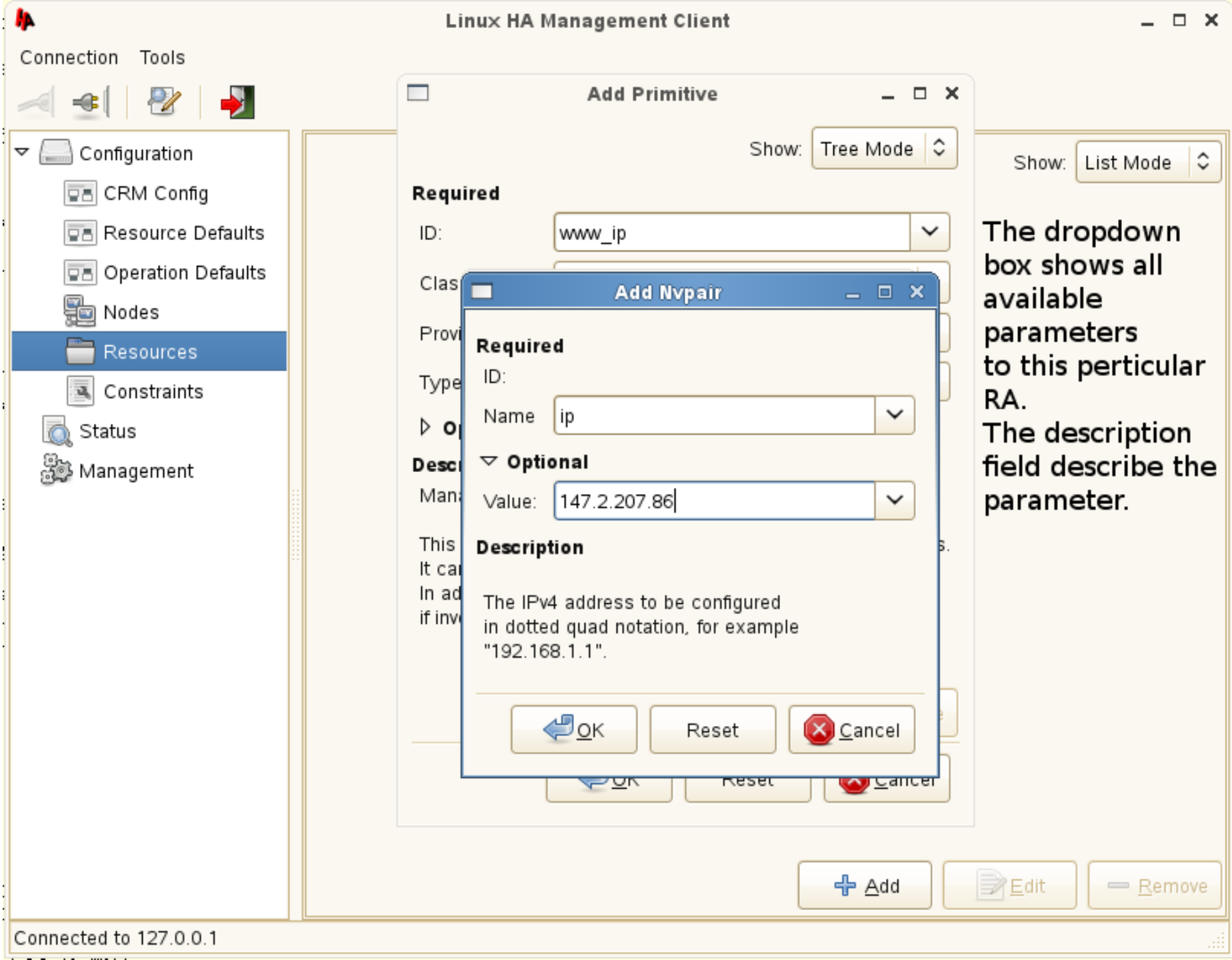
© Novell, Inc. All rights reserved.

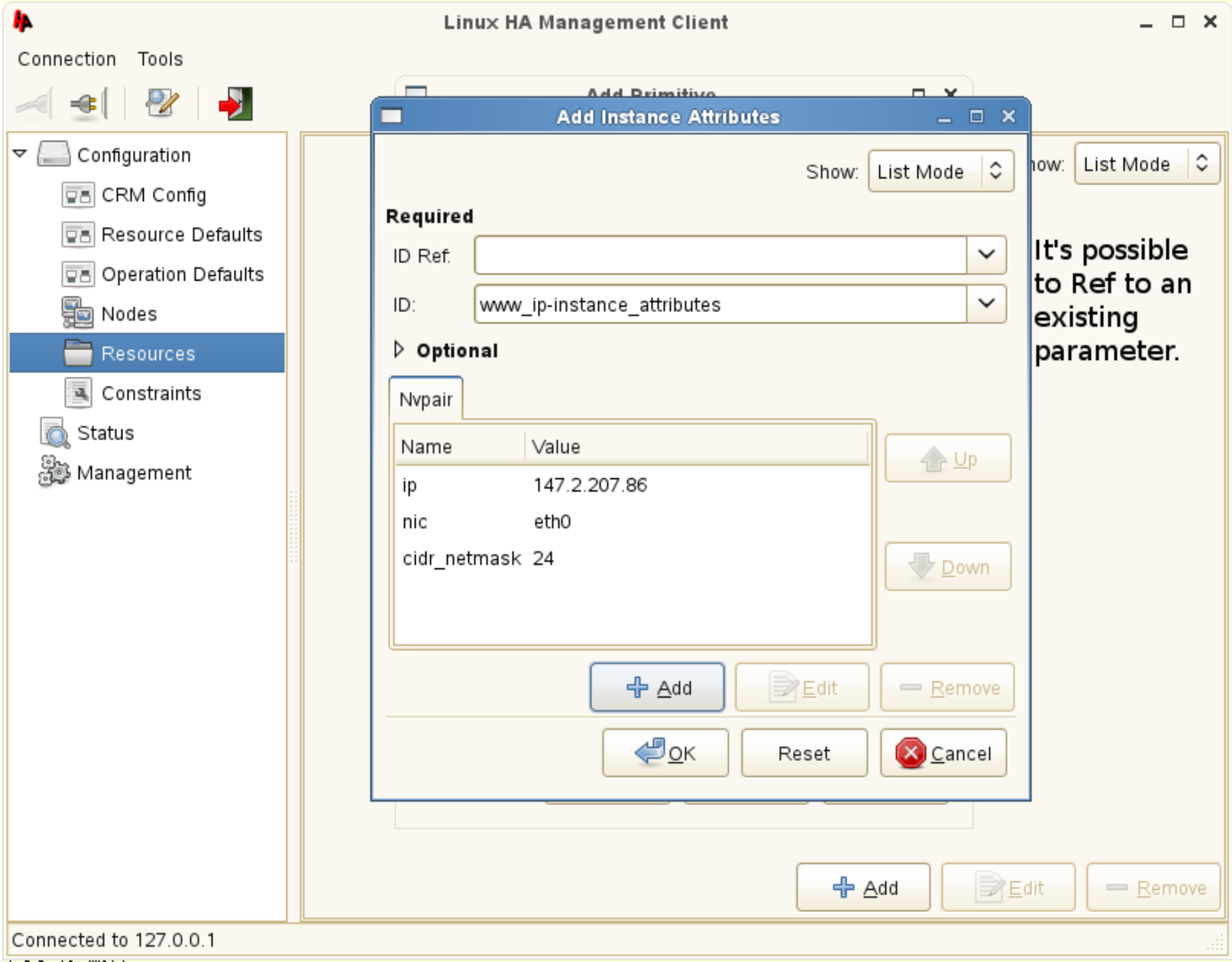
Improved GUI for Pacemaker

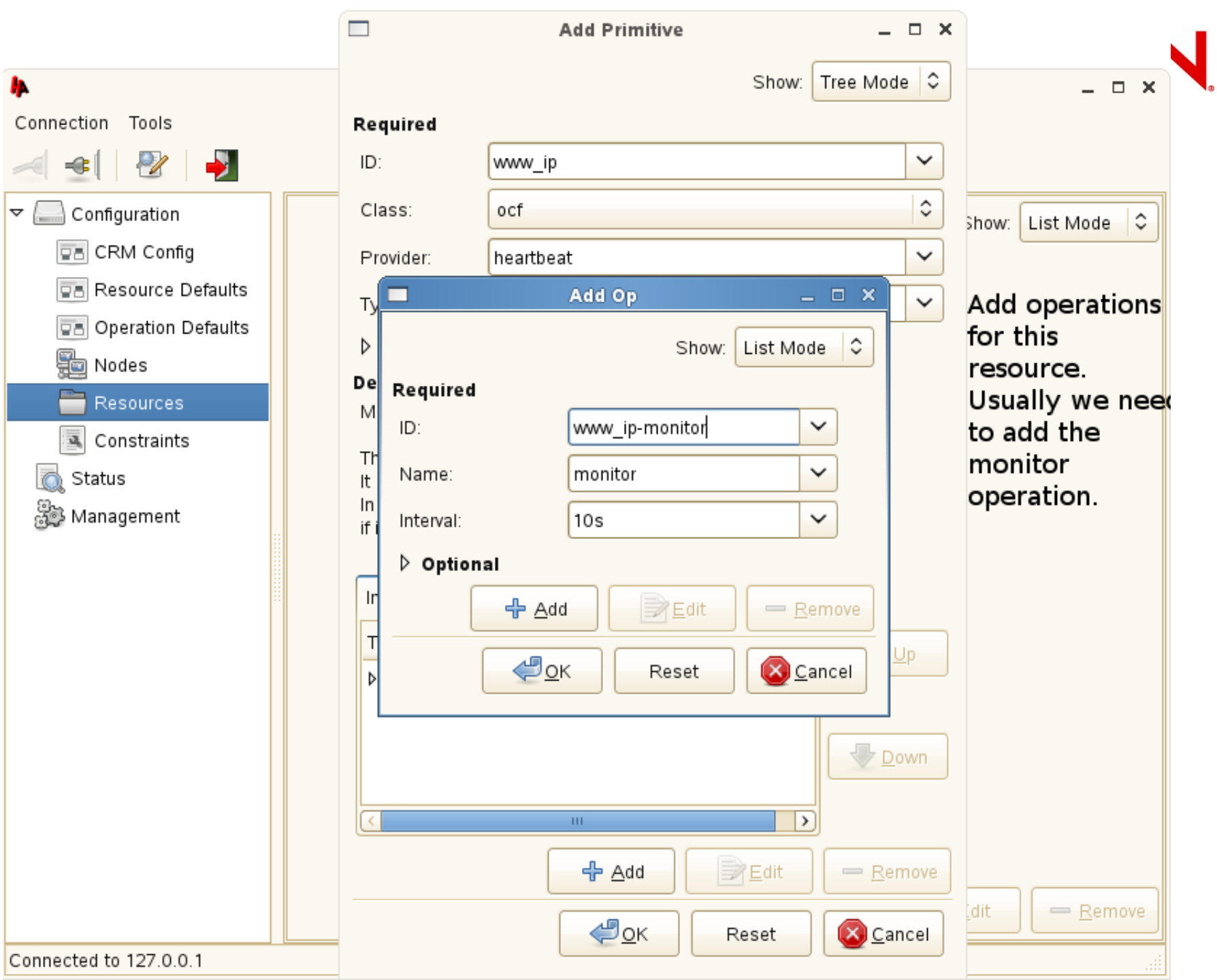
- Compatible with the new cluster stack including pacemaker-1.0, heartbeat-3.0 and openAIS
- Supports Relax-NG scheme/DTD
- Expert mode for providing full features: adding and modifying kinds of objects (resources, constraints, configurations of resource manager...)
- Support viewing/manually editing/importing/exporting CIB XML
- Management functionalities (Start/Stop/Migrate/Cleanup a resource, Standby a node)
- Generate cluster report

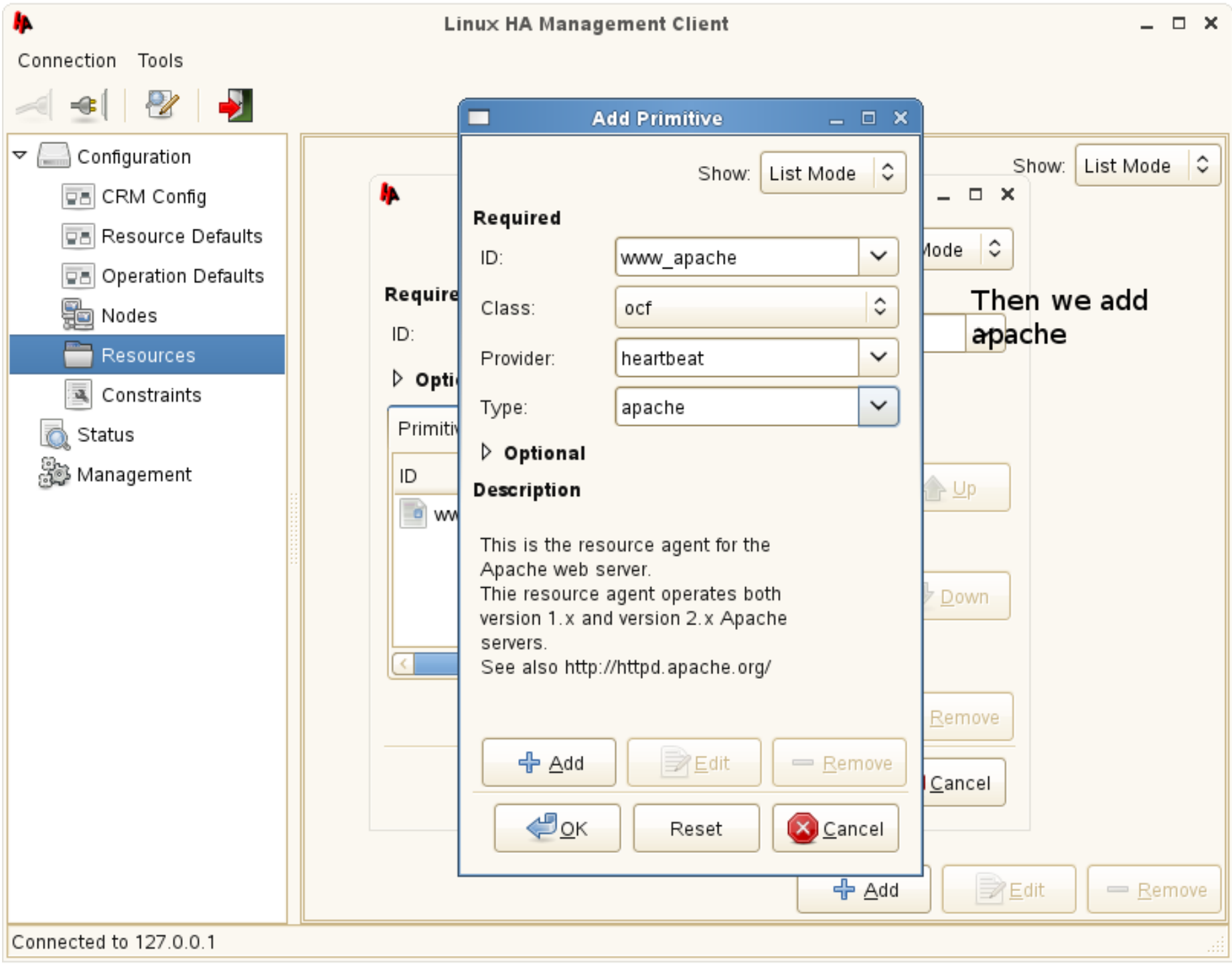
© Novell, Inc. All rights reserved.

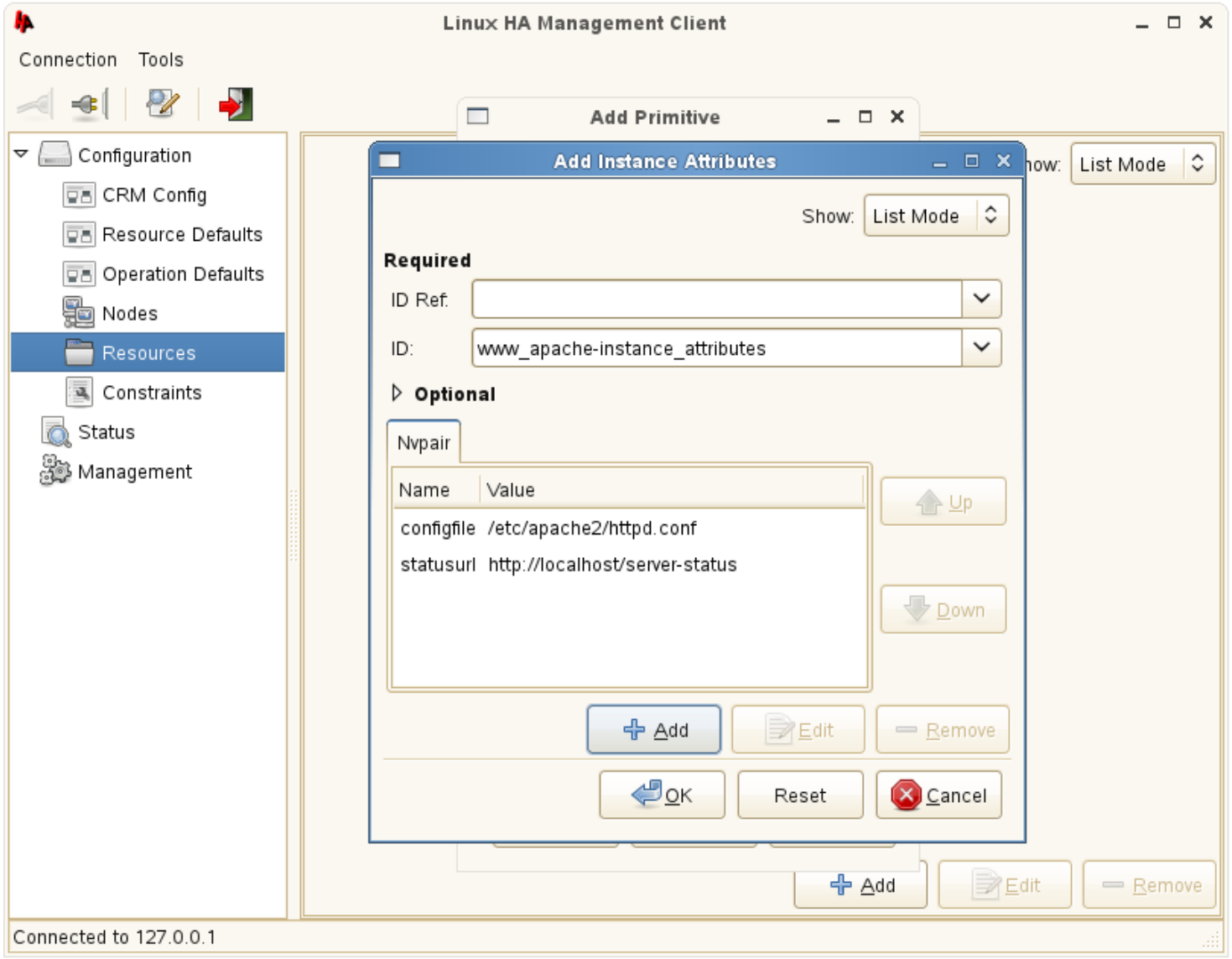

Linux HA Management Client
Connection
Tools
Configuration
CRM Config
Resource Defaults
Operation Defaults
Nodes
Resources
Constraints
Status
Management
Add Primitive
Show:
List Mode
Required
ID:
www_ip
Class:
ocf
Provider:
heartbeat
Type:
IPaddr2
Optional
Description
Manages virtual IPv4 addresses.
This Linux-specific resource manages IP alias IP addresses.
It can add an IP alias, or remove one.
In addition, it can implement Cluster Alias IP functionality
if invoked as a clone resource.
Add
Edit
Remove
OK
Reset
Cancel
Show:
List Mode
The dropdown box shows all available type of applications we supported.
Add
Edit
Remove
Connected to 127.0.0.1


Linux HA Management Client
Connection
Tools
Configuration
CRM Config
Resource Defaults
Operation Defaults
Nodes
Resources
Constraints
Status
Management
Add Primitive
Show:
Tree Mode
Required
ID:
www_ip
Add Nvpair
Required
ID:
Name
ip
Optional
Value:
147.2.207.86
Description
The IPv4 address to be configured
in dotted quad notation, for example
"192.168.1.1".
OK
Reset
Cancel
Show:
List Mode
The dropdown box shows all available parameters to this perticular RA.
The description field describe the parameter.
Add
Edit
Remove
Connected to 127.0.0.1

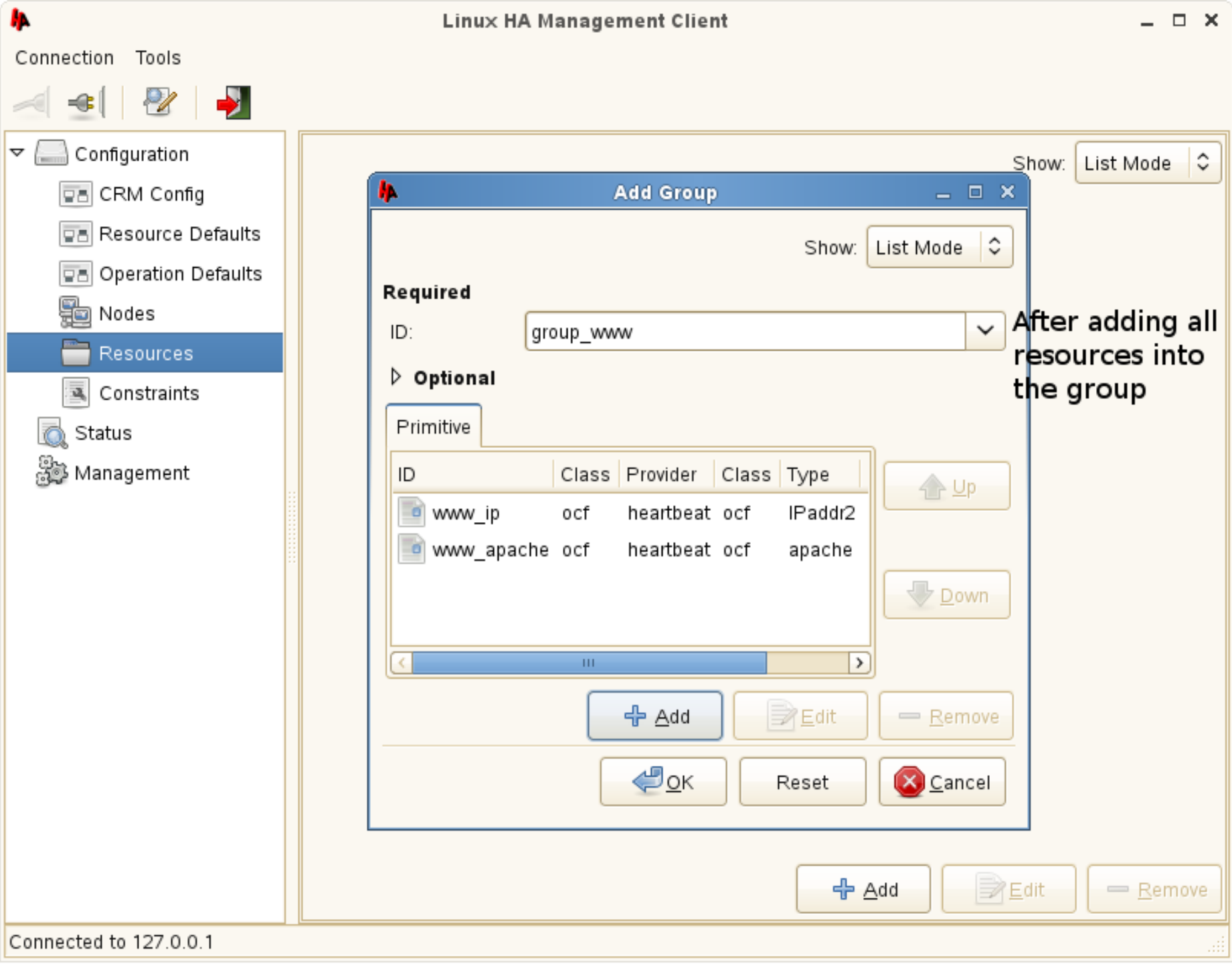

Linux HA Management Client
Connection
Tools
Configuration
CRM Config
Resource Defaults
Operation Defaults
Nodes
Resources
Constraints
Status
Management
Add Instance Attributes
Show:
List Mode
Required
ID Ref:
ID:
www_ip-instance_attributes
Optional
Nvpair
Name
Value
ip
147.2.207.86
nic
eth0
cidr_netmask
24
Up
Down
Add
Edit
Remove
OK
Reset
Cancel
It's possible to Ref to an existing parameter.
Add
Edit
Remove
Connected to 127.0.0.1

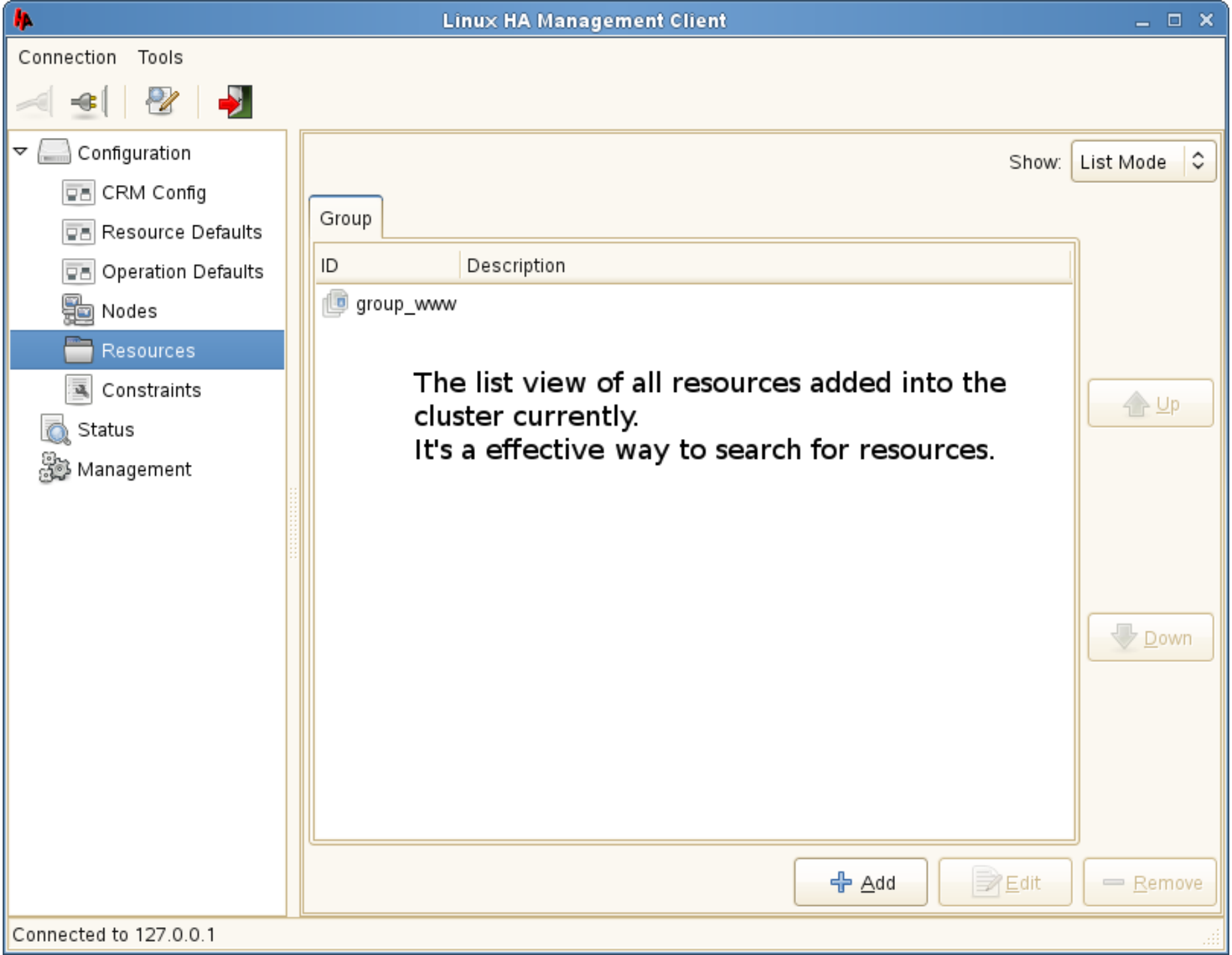

Add Primitive
Show: Tree Mode
Required
ID: www_ip
Class: ocf
Provider: heartbeat
Add Op
Show: List Mode
Required
ID: www_ip-monitor
Name: monitor
Interval: 10s
Optional
Add
Edit
Remove
OK
Reset
Cancel
Connection
Tools
Configuration
CRM Config
Resource Defaults
Operation Defaults
Nodes
Resources
Constraints
Status
Management
Show: List Mode
Add operations for this resource. Usually we nee to add the monitor operation.
Up
Down
Add
Edit
Remove
OK
Reset
Cancel
Connected to 127.0.0.1

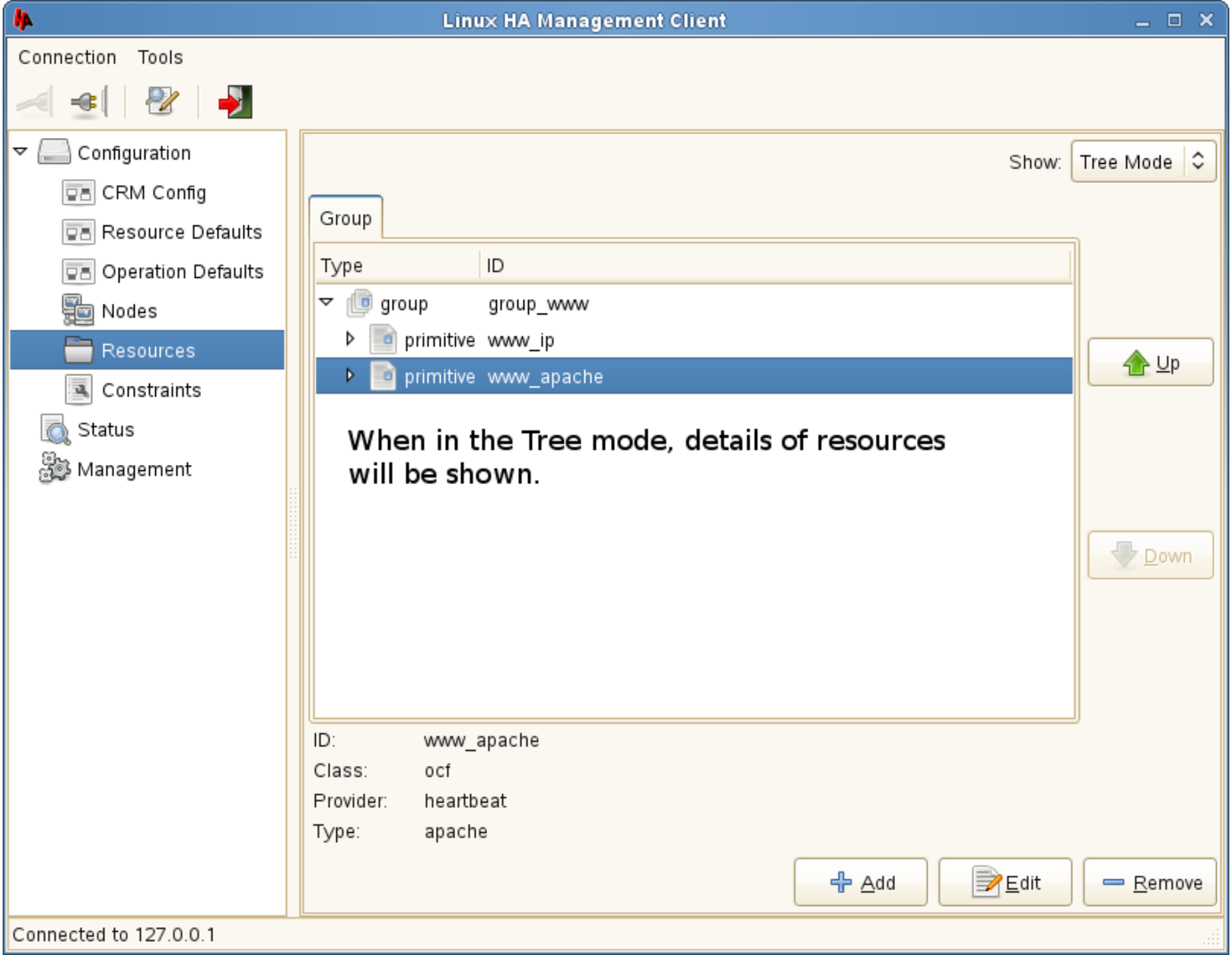

Linux HA Management Client
Connection
Tools
Configuration
CRM Config
Resource Defaults
Operation Defaults
Nodes
Resources
Constraints
Status
Management
Add Primitive
Show:
List Mode
Required
ID:
www_apache
Class:
ocf
Provider:
heartbeat
Type:
apache
Optional
Description
This is the resource agent for the
Apache web server.
Thie resource agent operates both
version 1.x and version 2.x Apache
servers.
See also http://httpd.apache.org/
Add
Edit
Remove
OK
Reset
Cancel
Then we add
apache
Up
Down
Remove
Cancel
Add
Edit
Remove
Connected to 127.0.0.1

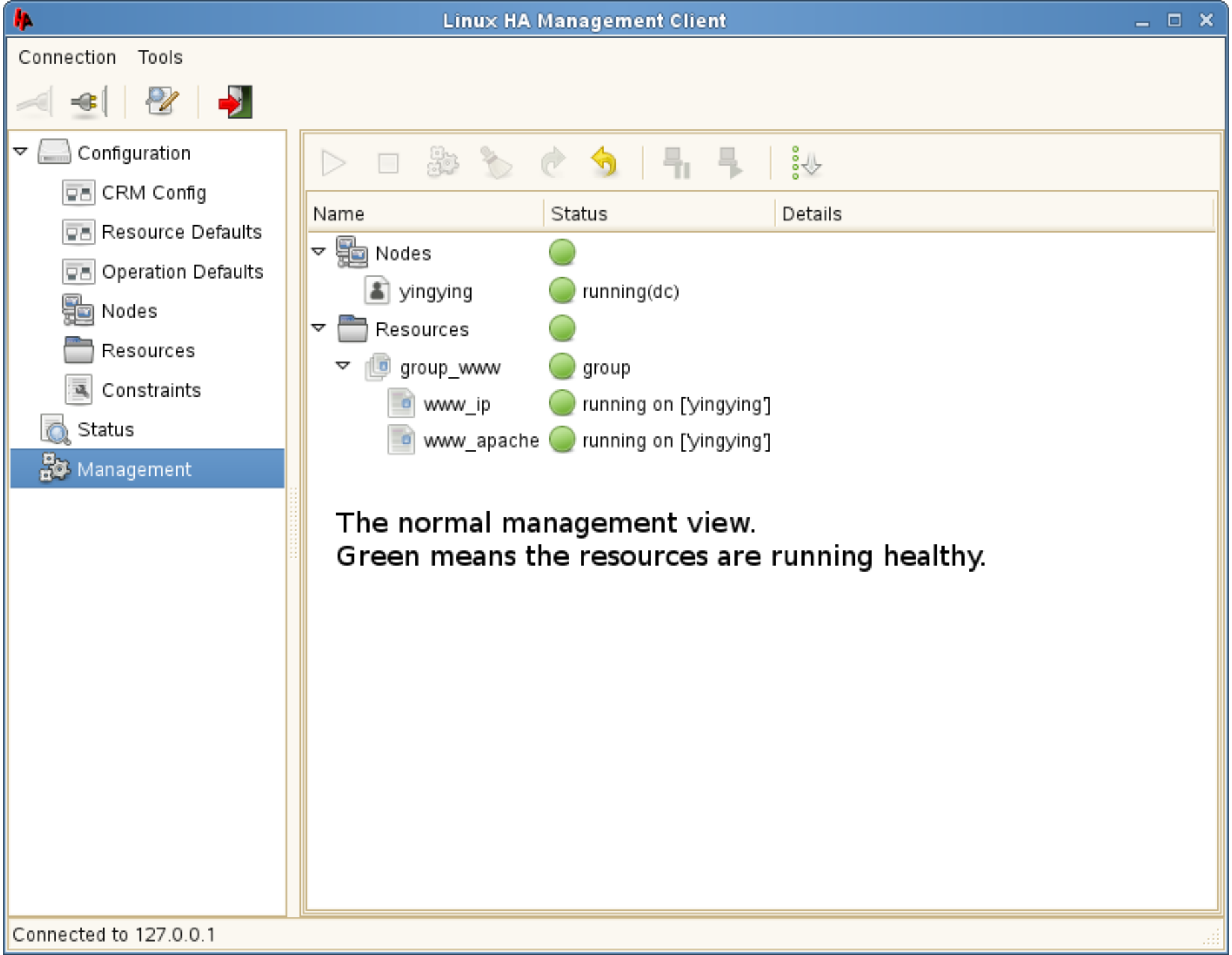


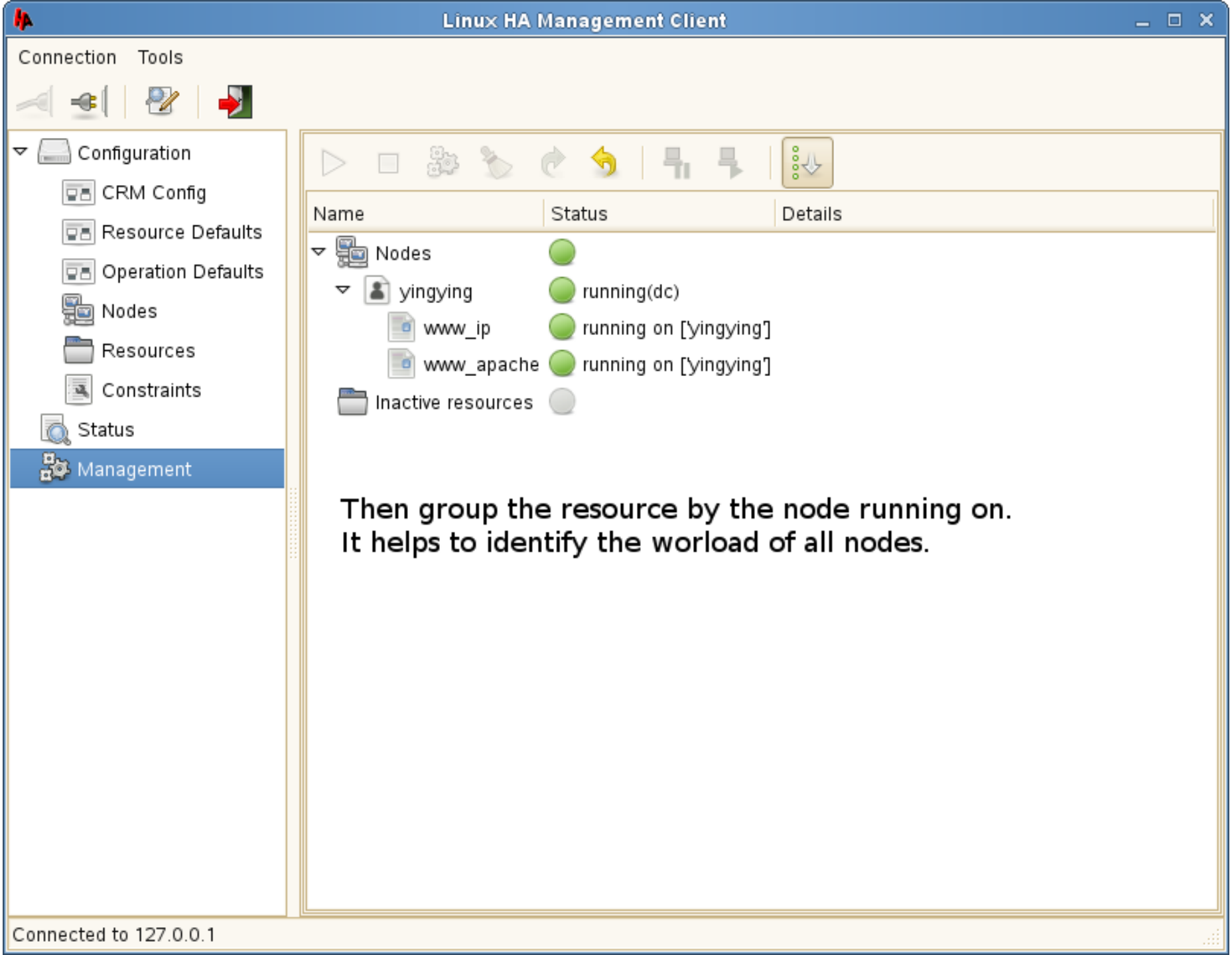
© Novell, Inc. All rights reserved.

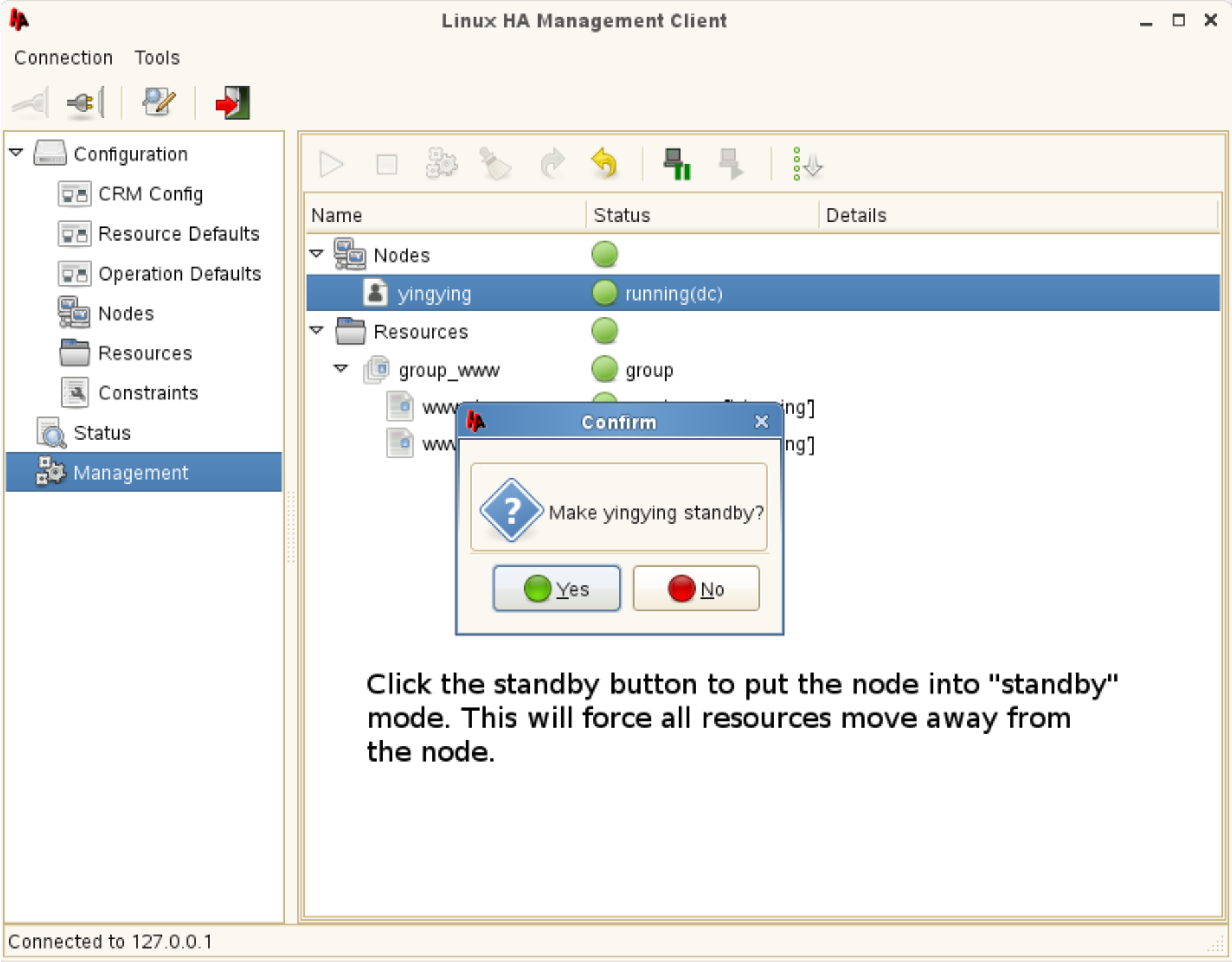

Linux HA Management Client
Connection
Tools
Configuration
CRM Config
Resource Defaults
Operation Defaults
Nodes
Resources
Constraints
Status
Management
Show:
List Mode
Add Group
Show:
List Mode
Required
ID:
group_www
Optional
Primitive
ID
Class
Provider
Class
Type
www_ip
ocf
heartbeat
ocf
IPaddr2
www_apache
ocf
heartbeat
ocf
apache
Up
Down
Add
Edit
Remove
OK
Reset
Cancel
After adding all resources into the group
Add
Edit
Remove
Connected to 127.0.0.1

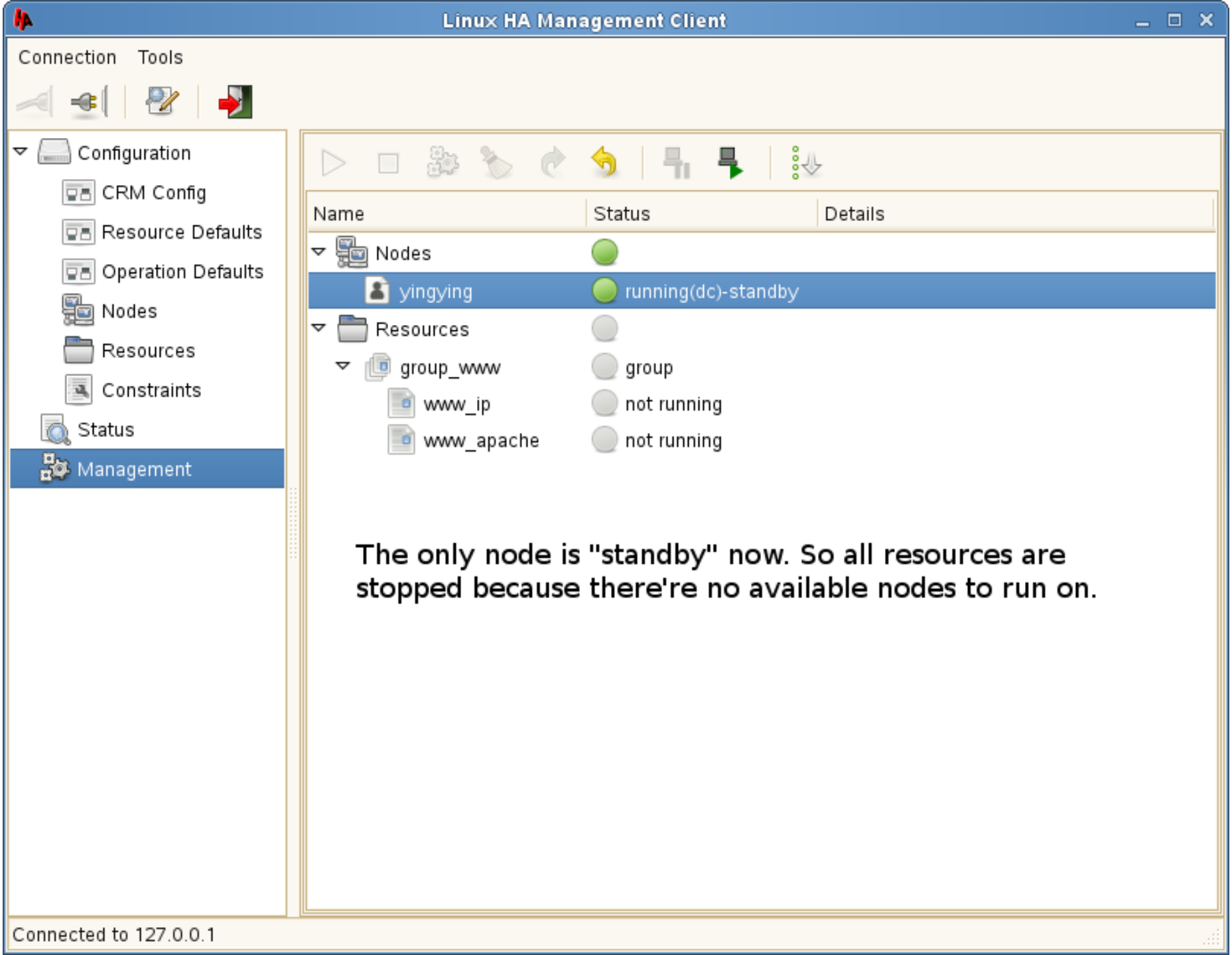

Linux HA Management Client
Connection
Tools
Configuration
CRM Config
Resource Defaults
Operation Defaults
Nodes
Resources
Constraints
Status
Management
Show:
List Mode
Group
ID
Description
group_www
The list view of all resources added into the cluster currently.
It's a effective way to search for resources.
Up
Down
Add
Edit
Remove
Connected to 127.0.0.1

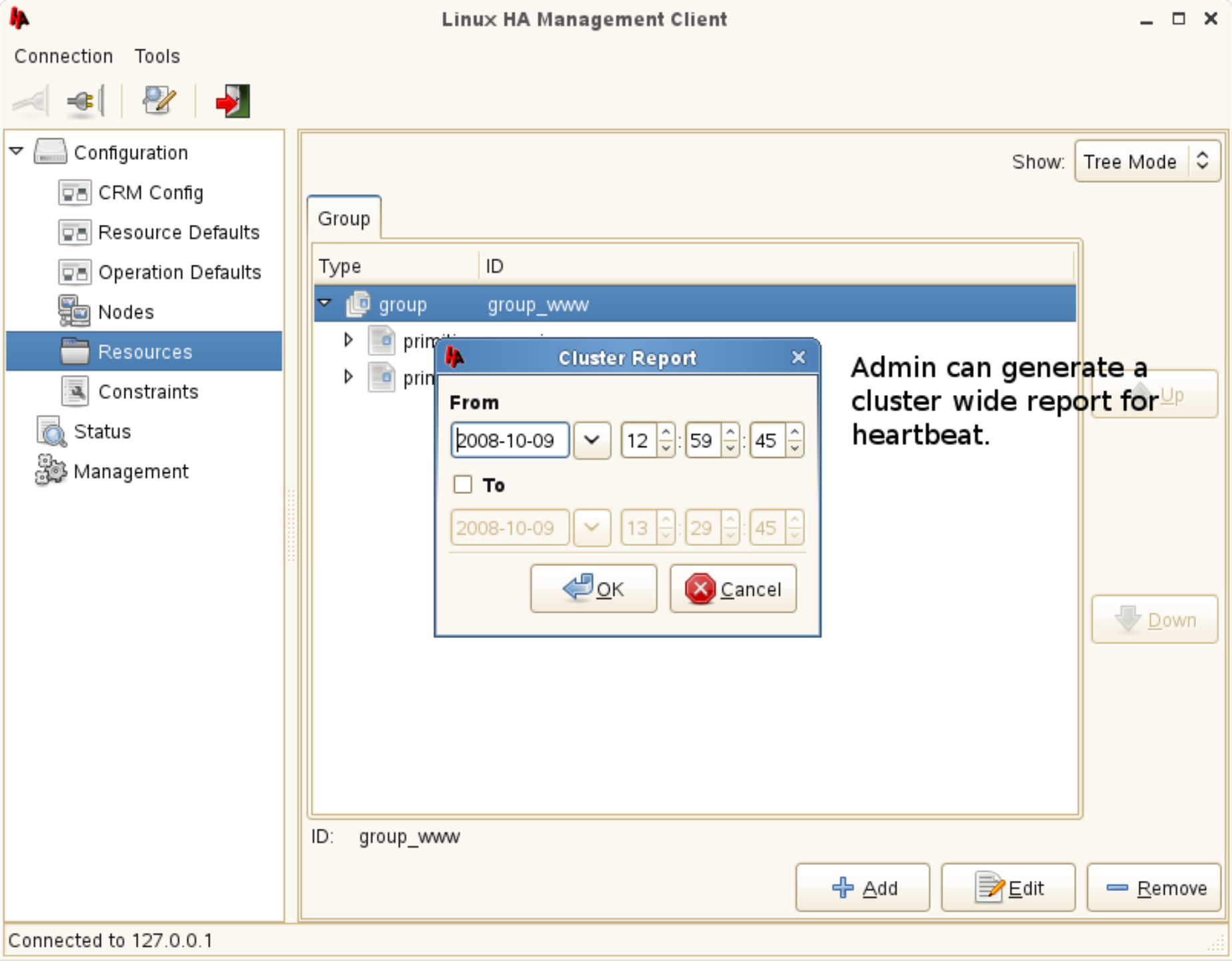

Linux HA Management Client
Connection
Tools
Configuration
CRM Config
Resource Defaults
Operation Defaults
Nodes
Resources
Constraints
Status
Management
Show:
Tree Mode
Group
Type
ID
group
group_www
primitive
www_ip
primitive
www_apache
When in the Tree mode, details of resources will be shown.
Up
Down
ID:
www_apache
Class:
ocf
Provider:
heartbeat
Type:
apache
Add
Edit
Remove
Connected to 127.0.0.1

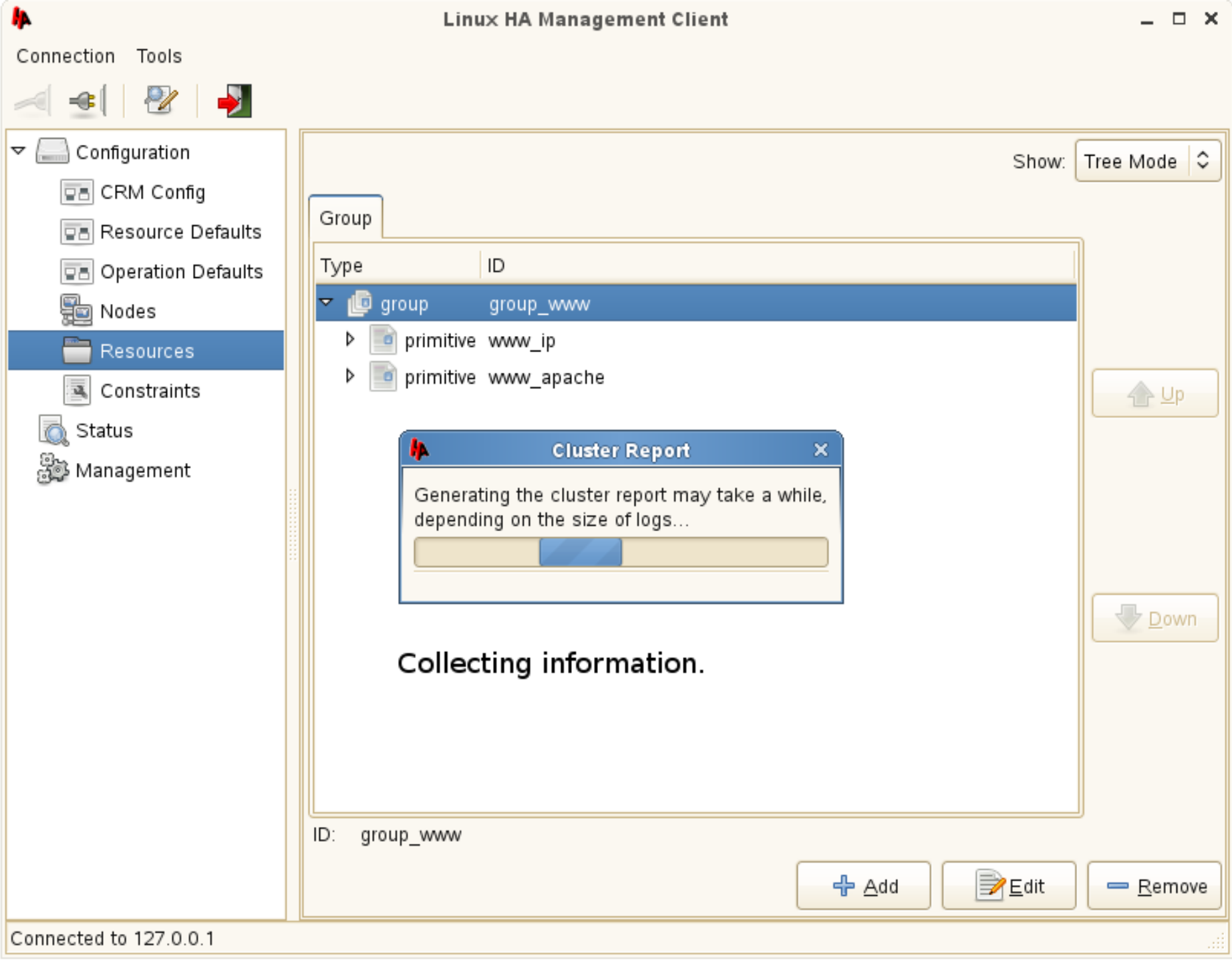

Linux HA Management Client
Connection
Tools
Configuration
CRM Config
Resource Defaults
Operation Defaults
Nodes
Resources
Constraints
Status
Management
Name
Status
Details
Nodes
yingying
running(dc)
Resources
group_www
group
www_ip
running on [yingying]
www_apache
running on [yingying]
The normal management view.
Green means the resources are running healthy.
Connected to 127.0.0.1

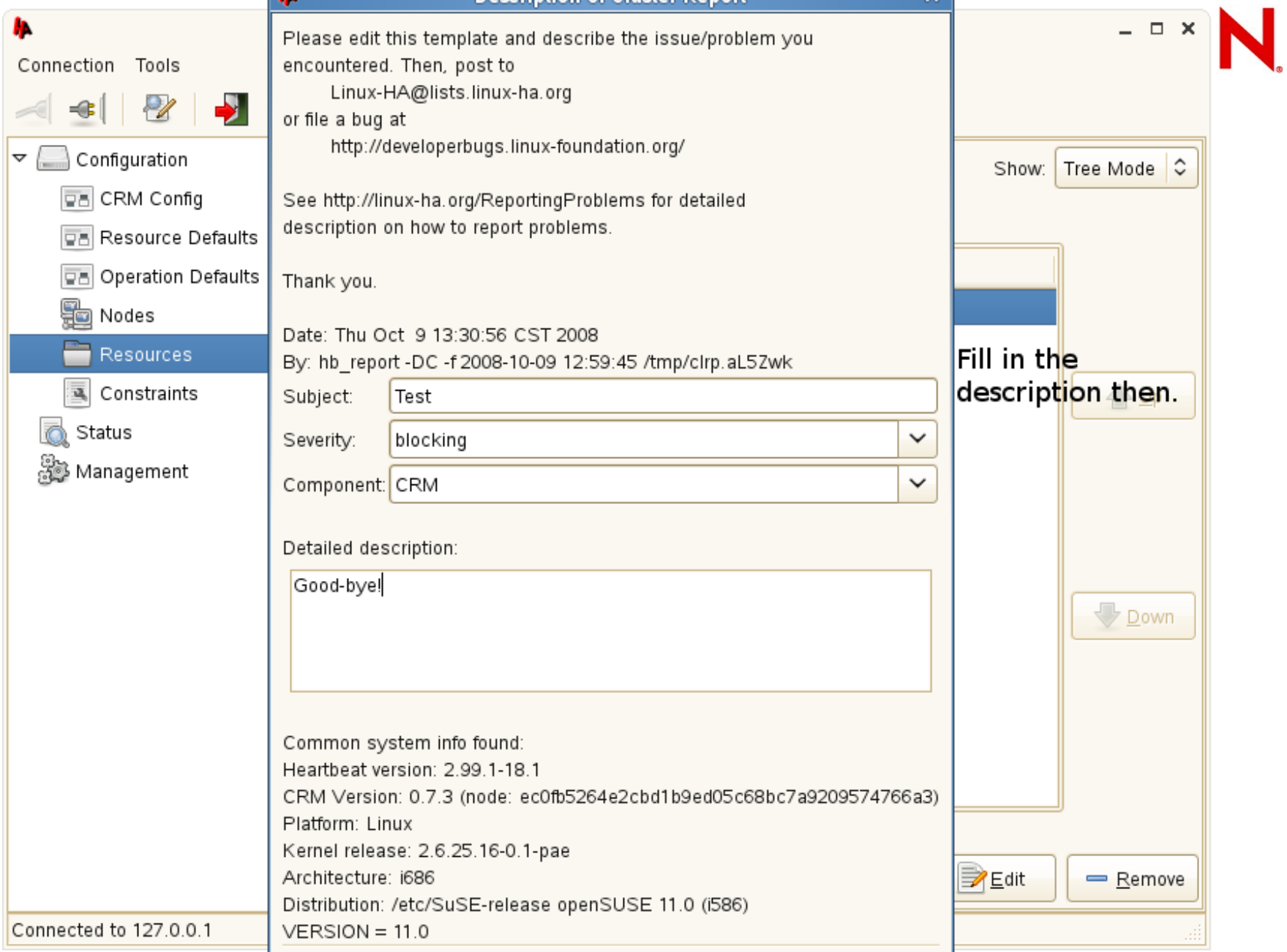

Linux HA Management Client
Connection
Tools
Configuration
CRM Config
Resource Defaults
Operation Defaults
Nodes
Resources
Constraints
Status
Management
Name
Status
Details
Nodes
yingying
running(dc)
www_ip
running on ['yingying']
www_apache
running on ['yingying']
Inactive resources
Then group the resource by the node running on.
It helps to identify the worload of all nodes.
Connected to 127.0.0.1

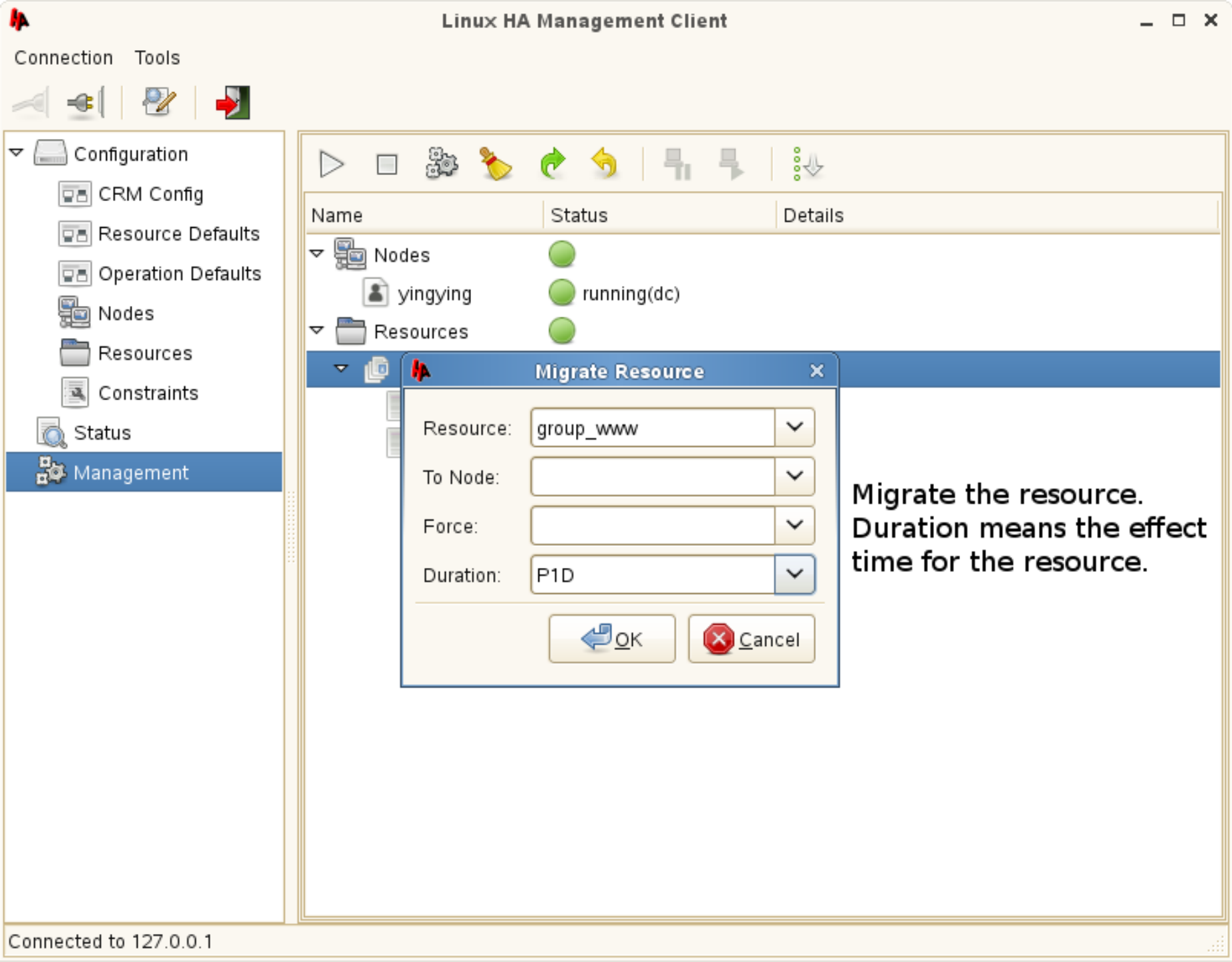

Linux HA Management Client
Connection
Tools
Configuration
CRM Config
Resource Defaults
Operation Defaults
Nodes
Resources
Constraints
Status
Management
Name
Status
Details
Nodes
yingying
running(dc)
Resources
group_www
group
Confirm
Make yingying standby?
Yes
No
Click the standby button to put the node into "standby" mode. This will force all resources move away from the node.
Connected to 127.0.0.1

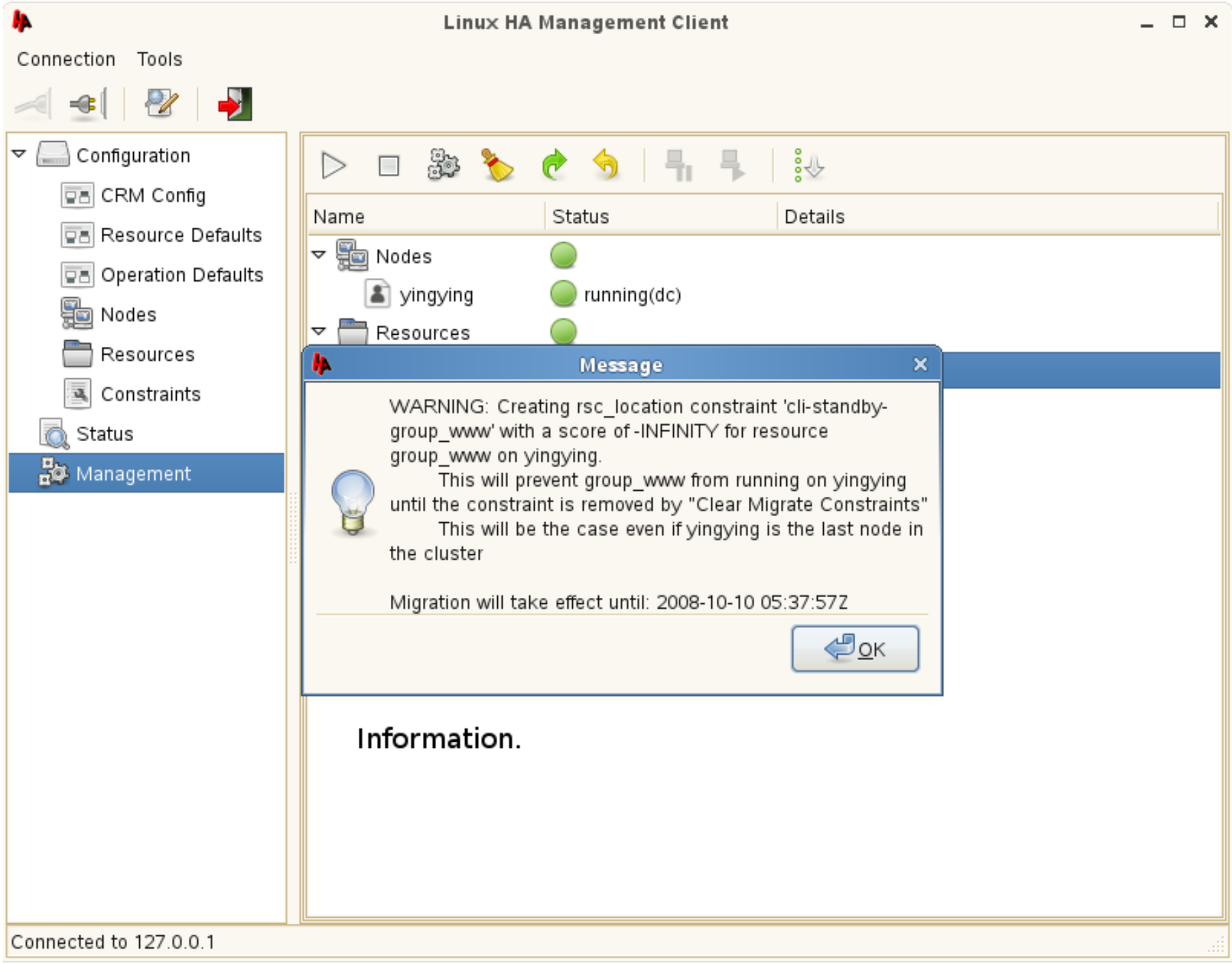

Linux HA Management Client
Connection
Tools
Configuration
CRM Config
Resource Defaults
Operation Defaults
Nodes
Resources
Constraints
Status
Management
Name
Status
Details
Nodes
yingying
running(dc)-standby
Resources
group_www
group
www_ip
not running
www_apache
not running
The only node is "standby" now. So all resources are stopped because there're no available nodes to run on.
Connected to 127.0.0.1

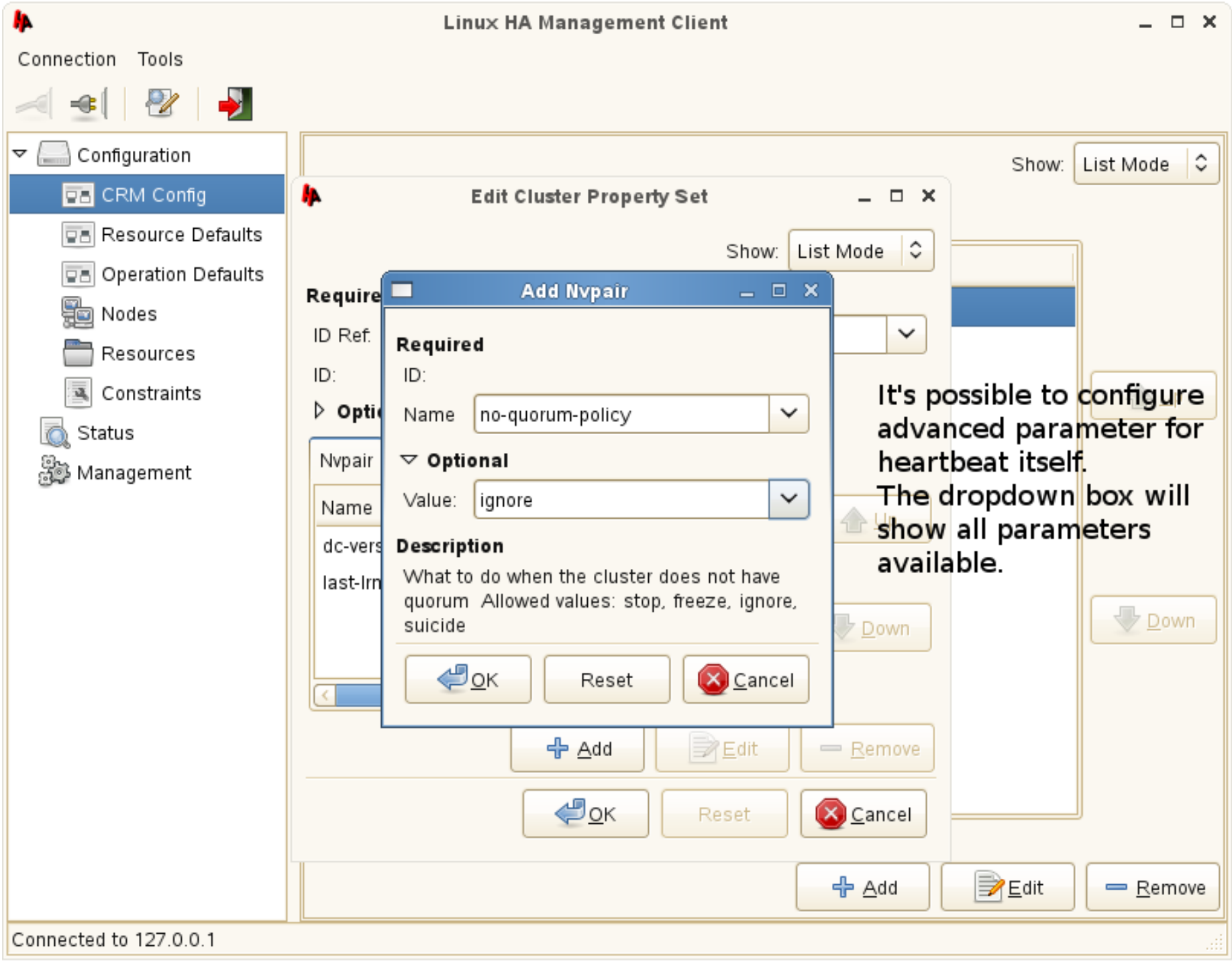

Linux HA Management Client
Connection
Tools
Configuration
CRM Config
Resource Defaults
Operation Defaults
Nodes
Resources
Constraints
Status
Management
Show:
Tree Mode
Group
Type
ID
group
group_www
Cluster Report
From
2008-10-09
12
59
45
To
2008-10-09
13
29
45
OK
Cancel
Admin can generate a cluster wide report for heartbeat.
Up
Down
ID:
group_www
Add
Edit
Remove
Connected to 127.0.0.1

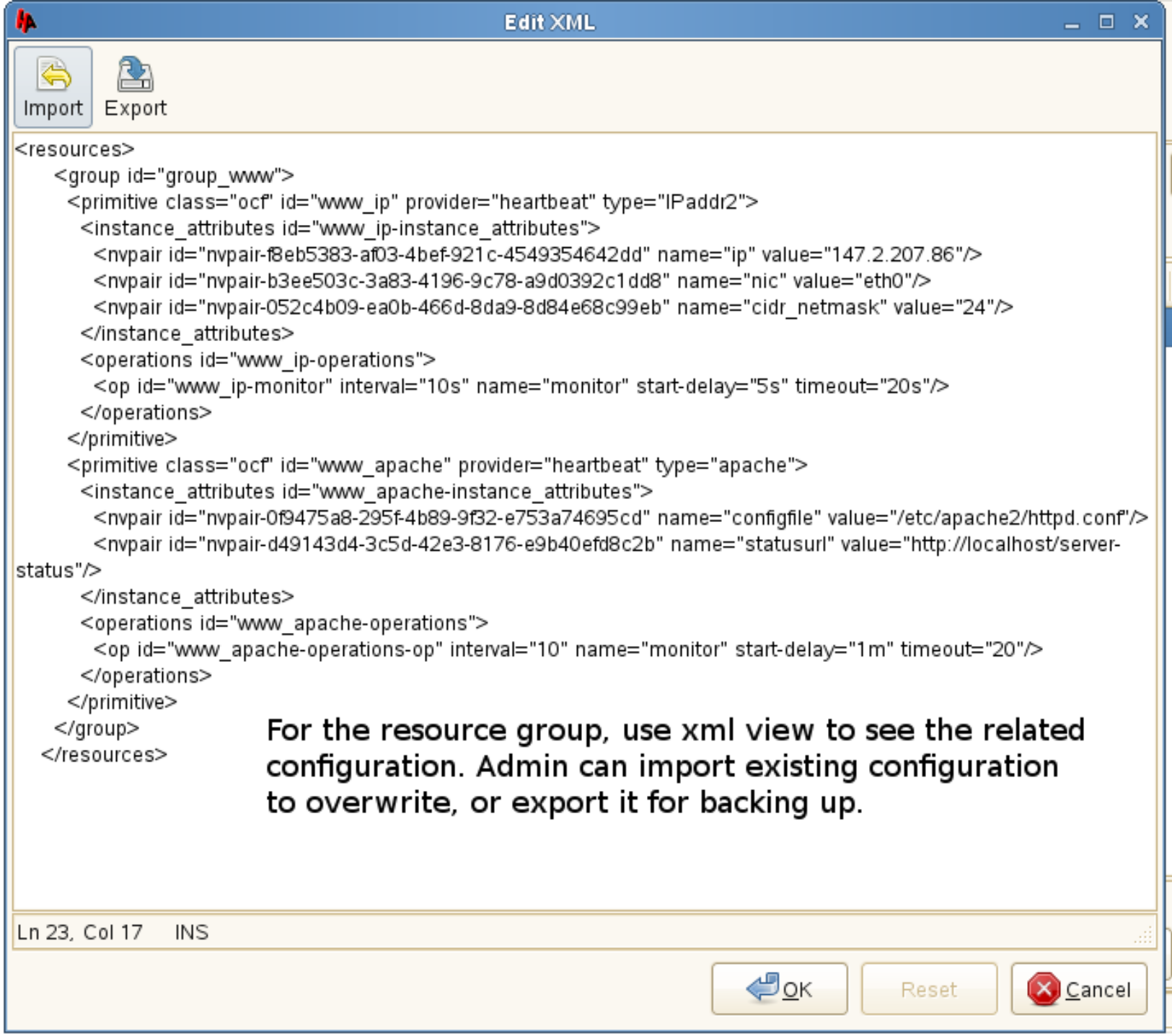

Linux HA Management Client
Connection
Tools
Configuration
CRM Config
Resource Defaults
Operation Defaults
Nodes
Resources
Constraints
Status
Management
Show:
Tree Mode
Group
Type
ID
group
group_www
primitive
www_ip
primitive
www_apache
Up
Down
Cluster Report
Generating the cluster report may take a while, depending on the size of logs...
Collecting information.
ID:
group_www
Add
Edit
Remove
Connected to 127.0.0.1

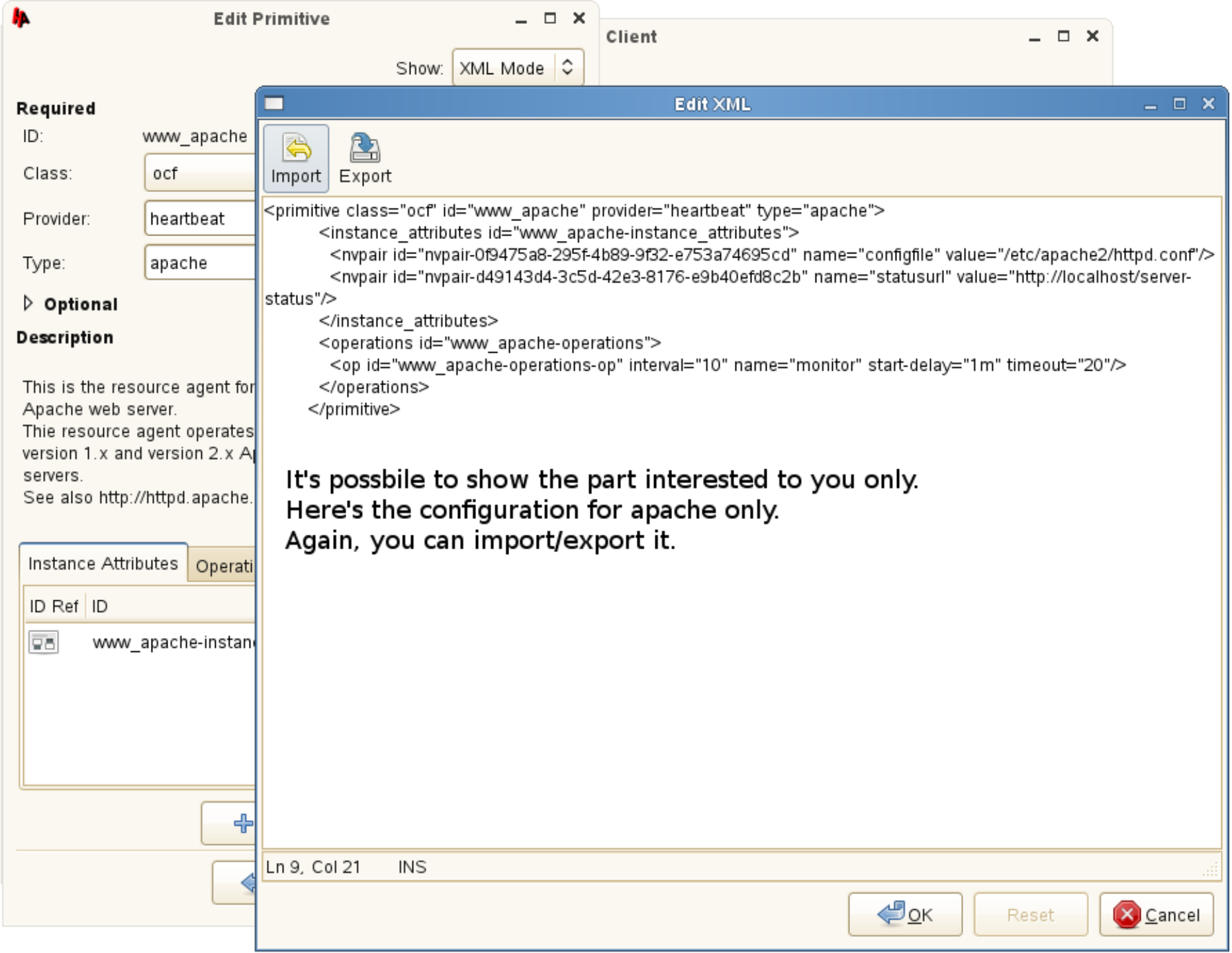

Description of Cluster Report
Please edit this template and describe the issue/problem you encountered. Then, post to
Linux-HA@lists.linux-ha.org
or file a bug at
http://developerbugs.linux-foundation.org/
See http://linux-ha.org/ReportingProblems for detailed description on how to report problems.
Thank you.
Date: Thu Oct 9 13:30:56 CST 2008
By: hb_report -DC -f 2008-10-09 12:59:45 /tmp/clrp.aL5Zwk
Subject: Test
Severity: blocking
Component: CRM
Detailed description:
Good-bye!
Common system info found:
Heartbeat version: 2.99.1-18.1
CRM Version: 0.7.3 (node: ec0fb5264e2cbd1b9ed05c68bc7a9209574766a3)
Platform: Linux
Kernel release: 2.6.25.16-0.1-pae
Architecture: i686
Distribution: /etc/SuSE-release openSUSE 11.0 (i586)
VERSION = 11.0
OK
Cancel
Connection
Tools
Configuration
CRM Config
Resource Defaults
Operation Defaults
Nodes
Resources
Constraints
Status
Management
Connected to 127.0.0.1
Show:
Tree Mode
Fill in the description then.
Down
Edit
Remove

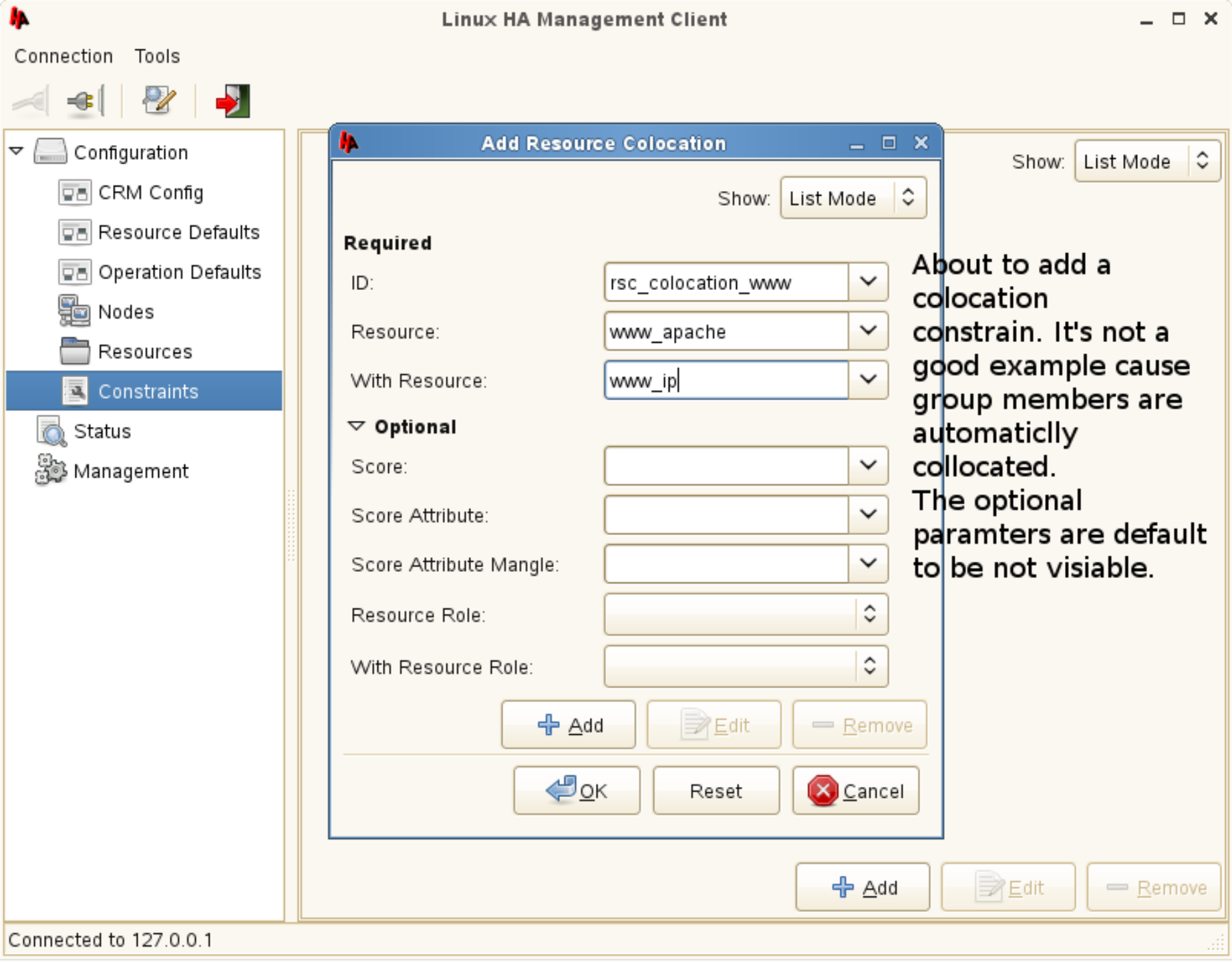

Linux HA Management Client
Connection
Tools
Configuration
CRM Config
Resource Defaults
Operation Defaults
Nodes
Resources
Constraints
Status
Management
Name
Status
Details
Nodes
yingying
running(dc)
Resources
Migrate Resource
Resource:
group_www
To Node:
Force:
Duration:
P1D
OK
Cancel
Migrate the resource.
Duration means the effect
time for the resource.
Connected to 127.0.0.1

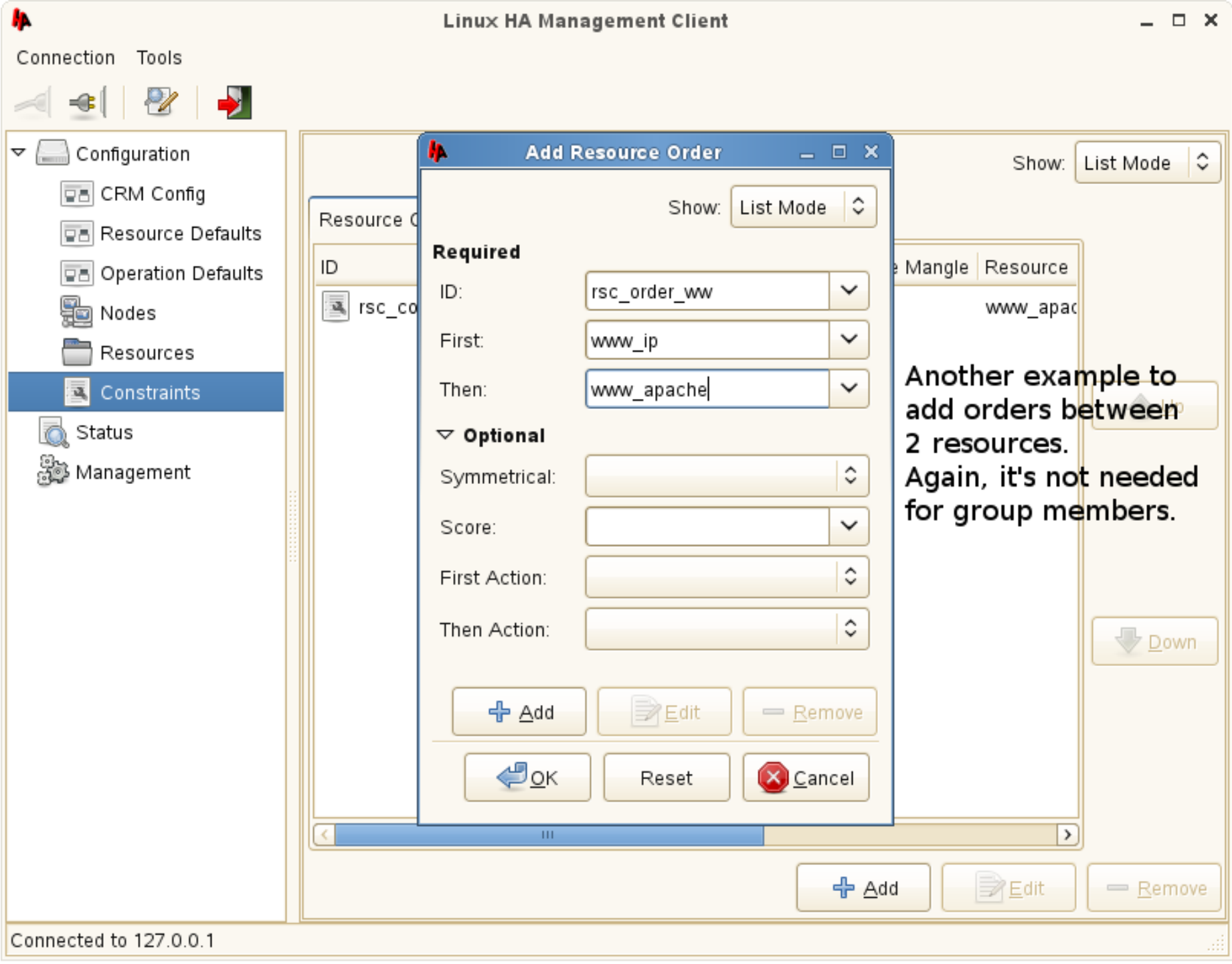

Linux HA Management Client
Connection
Tools
Configuration
CRM Config
Resource Defaults
Operation Defaults
Nodes
Resources
Constraints
Status
Management
Name
Status
Details
Nodes
yingying
running(dc)
Resources
Message
WARNING: Creating rsc_location constraint 'cli-standby-group_www' with a score of -INFINITY for resource group_www on yingying.
This will prevent group_www from running on yingying until the constraint is removed by "Clear Migrate Constraints"
This will be the case even if yingying is the last node in the cluster
Migration will take effect until: 2008-10-10 05:37:57Z
OK
Information.
Connected to 127.0.0.1

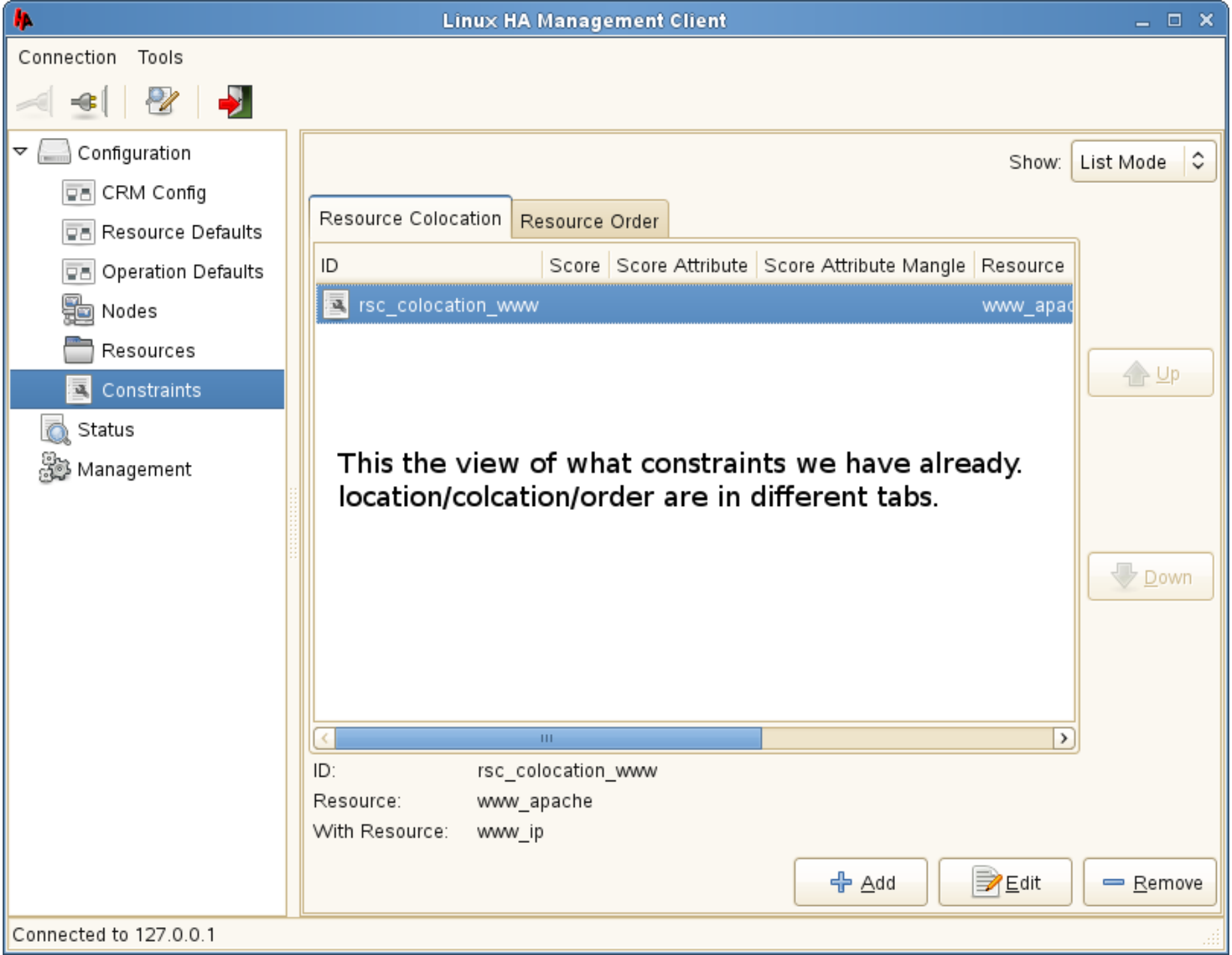

Linux HA Management Client
Connection
Tools
Configuration
CRM Config
Resource Defaults
Operation Defaults
Nodes
Resources
Constraints
Status
Management
Show:
List Mode
Edit Cluster Property Set
Show:
List Mode
Require
ID Ref:
ID:
Opti
Nvpair
Name
dc-vers
last-lrm
Add Nvpair
Required
ID:
Name
no-quorum-policy
Optional
Value:
ignore
Description
What to do when the cluster does not have quorum Allowed values: stop, freeze, ignore, suicide
OK
Reset
Cancel
Add
Edit
Remove
OK
Reset
Cancel
Down
Down
It's possible to configure advanced parameter for heartbeat itself.
The dropdown box will show all parameters available.
Add
Edit
Remove
Connected to 127.0.0.1


Edit XML
Import
Export
<resources>
<group id="group_www">
<primitive class="ocf" id="www_ip" provider="heartbeat" type="IPaddr2">
<instance_attributes id="www_ip-instance_attributes">
<nvpair id="nvpair-f8eb5383-af03-4bef-921c-4549354642dd" name="ip" value="147.2.207.86"/>
<nvpair id="nvpair-b3ee503c-3a83-4196-9c78-a9d0392c1dd8" name="nic" value="eth0"/>
<nvpair id="nvpair-052c4b09-ea0b-466d-8da9-8d84e68c99eb" name="cidr_netmask" value="24"/>
</instance_attributes>
<operations id="www_ip-operations">
<op id="www_ip-monitor" interval="10s" name="monitor" start-delay="5s" timeout="20s"/>
</operations>
</primitive>
<primitive class="ocf" id="www_apache" provider="heartbeat" type="apache">
<instance_attributes id="www_apache-instance_attributes">
<nvpair id="nvpair-0f9475a8-295f-4b89-9f32-e753a74695cd" name="configfile" value="/etc/apache2/httpd.conf"/>
<nvpair id="nvpair-d49143d4-3c5d-42e3-8176-e9b40efd8c2b" name="statusurl" value="http://localhost/server-status"/>
</instance_attributes>
<operations id="www_apache-operations">
<op id="www_apache-operations-op" interval="10" name="monitor" start-delay="1m" timeout="20"/>
</operations>
</primitive>
</group>
</resources>
For the resource group, use xml view to see the related configuration. Admin can import existing configuration to overwrite, or export it for backing up.
Ln 23, Col 17 INS
OK
Reset
Cancel
XML Mode
Up
Down
Remove


Edit Primitive
Show: XML Mode
Required
ID: www_apache
Class: ocf
Provider: heartbeat
Type: apache
Optional
Description
This is the resource agent for
Apache web server.
Thie resource agent operates
version 1.x and version 2.x A
servers.
See also http://httpd.apache.
Instance Attributes
Operati
ID Ref
ID
www_apache-instan
Client
Edit XML
Import
Export
<primitive class="ocf" id="www_apache" provider="heartbeat" type="apache">
<instance_attributes id="www_apache-instance_attributes">
<nvpair id="nvpair-0f9475a8-295f-4b89-9f32-e753a74695cd" name="configfile" value="/etc/apache2/httpd.conf"/>
<nvpair id="nvpair-d49143d4-3c5d-42e3-8176-e9b40efd8c2b" name="statusurl" value="http://localhost/server-status"/>
</instance_attributes>
<operations id="www_apache-operations">
<op id="www_apache-operations-op" interval="10" name="monitor" start-delay="1m" timeout="20"/>
</operations>
</primitive>
It's possbile to show the part interested to you only.
Here's the configuration for apache only.
Again, you can import/export it.
Ln 9, Col 21
INS
OK
Reset
Cancel


Linux HA Management Client
Connection
Tools
Configuration
CRM Config
Resource Defaults
Operation Defaults
Nodes
Resources
Constraints
Status
Management
Add Resource Colocation
Show:
List Mode
Required
ID:
rsc_colocation_www
Resource:
www_apache
With Resource:
www_ip
Optional
Score:
Score Attribute:
Score Attribute Mangle:
Resource Role:
With Resource Role:
Add
Edit
Remove
OK
Reset
Cancel
About to add a colocation constrain. It's not a good example cause group members are automaticlly collocated.
The optional paramters are default to be not visiable.
Add
Edit
Remove
Connected to 127.0.0.1


Linux HA Management Client
Connection
Tools
Configuration
CRM Config
Resource Defaults
Operation Defaults
Nodes
Resources
Constraints
Status
Management
Add Resource Order
Show:
List Mode
Required
ID:
rsc_order_ww
First:
www_ip
Then:
www_apache
Optional
Symmetrical:
Score:
First Action:
Then Action:
Add
Edit
Remove
OK
Reset
Cancel
Another example to add orders between 2 resources.
Again, it's not needed for group members.
Down
Connected to 127.0.0.1


Linux HA Management Client
Connection
Tools
Configuration
CRM Config
Resource Defaults
Operation Defaults
Nodes
Resources
Constraints
Status
Management
Show:
List Mode
Resource Colocation
Resource Order
ID
Score
Score Attribute
Score Attribute Mangle
Resource
rsc_colocation_www
www_apac
Up
This the view of what constraints we have already.
location/colcation/order are in different tabs.
Down
ID:
rsc_colocation_www
Resource:
www_apache
With Resource:
www_ip
Add
Edit
Remove
Connected to 127.0.0.1

Cluster Developer Summit
Prague, 2008-09-29 – 2008-10-02

Dogs, cats, and mice ...

© Novell, Inc. All rights reserved.

Topics discussed:

- Merging of stacks
  - Tree splitting
- Community release versus “Enterprise” distributions
- Quorum
- Fencing
- Resource Agents + common library
- Standard logging
- Standard configuration(?)
- Shared root clusters
- Clustered Samba

© Novell, Inc. All rights reserved.

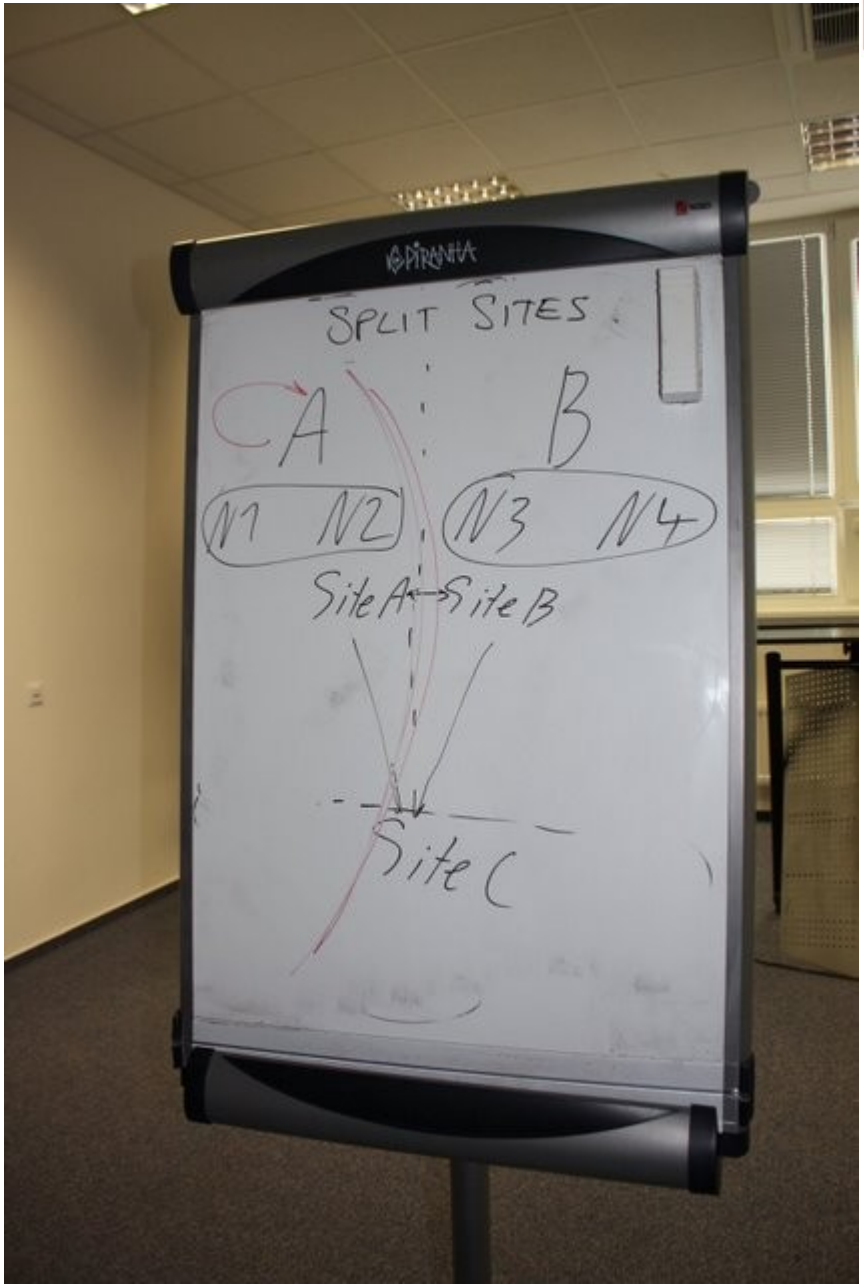
The future stack:

- ~~heartbeat~~ openAIS (corosync)
- Local Resource Manager
- ~~rgmanager~~ Pacemaker
- OCF Resource Agents
- Fencing component
- ~~C-EVMS2~~ cLVM2
- OCFS2
- GFS2

© Novell, Inc. All rights reserved.

Split-site clusters

- Wide-area:
  - High latency
  - Low bandwidth
  - Redundancy expensive
- Metro-area:
  - Low latency
  - High bandwidth
  - Redundancy expensive
- Quorum?
- Fencing?
- Storage?



© Novell, Inc. All rights reserved.

~~Questions?~~ Answers!

Novell®

Unpublished Work of Novell, Inc. All Rights Reserved.

This work is an unpublished work and contains confidential, proprietary, and trade secret information of Novell, Inc. Access to this work is restricted to Novell employees who have a need to know to perform tasks within the scope of their assignments. No part of this work may be practiced, performed, copied, distributed, revised, modified, translated, abridged, condensed, expanded, collected, or adapted without the prior written consent of Novell, Inc. Any use or exploitation of this work without authorization could subject the perpetrator to criminal and civil liability.

General Disclaimer

This document is not to be construed as a promise by any participating company to develop, deliver, or market a product. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. Novell, Inc. makes no representations or warranties with respect to the contents of this document, and specifically disclaims any express or implied warranties of merchantability or fitness for any particular purpose. The development, release, and timing of features or functionality described for Novell products remains at the sole discretion of Novell. Further, Novell, Inc. reserves the right to revise this document and to make changes to its content, at any time, without obligation to notify any person or entity of such revisions or changes. All Novell marks referenced in this presentation are trademarks or registered trademarks of Novell, Inc. in the United States and other countries. All third-party trademarks are the property of their respective owners.