SHARE
Technology • Connections • Results

Implementing the SUSE Linux Enterprise High Availability Extension on System z

Mike Friesenegger
Novell

Monday, February 28, 2011
Session Number: 8474

Agenda

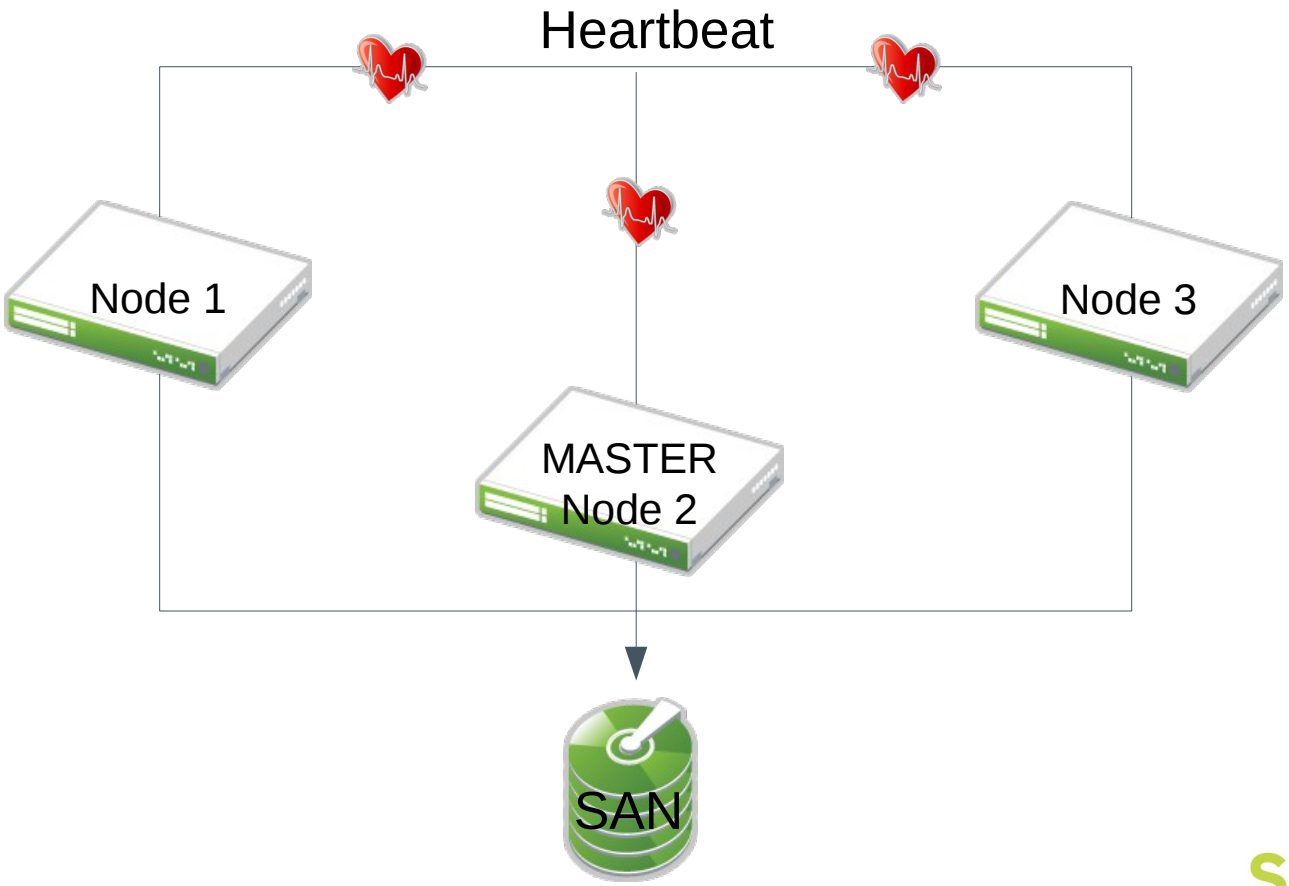
- What is a high availability (HA) cluster?
- What is required to build an HA cluster using SLES?
- Demoing the details
  - Managing a cluster with the GUI and CLI
  - Resources primitives and resource groups
  - Resource Constraints
  - STONITH
  - cLVM and OCFS2
- Wrap-up/Questions

If you attended this session at SHARE Boston in 2010...

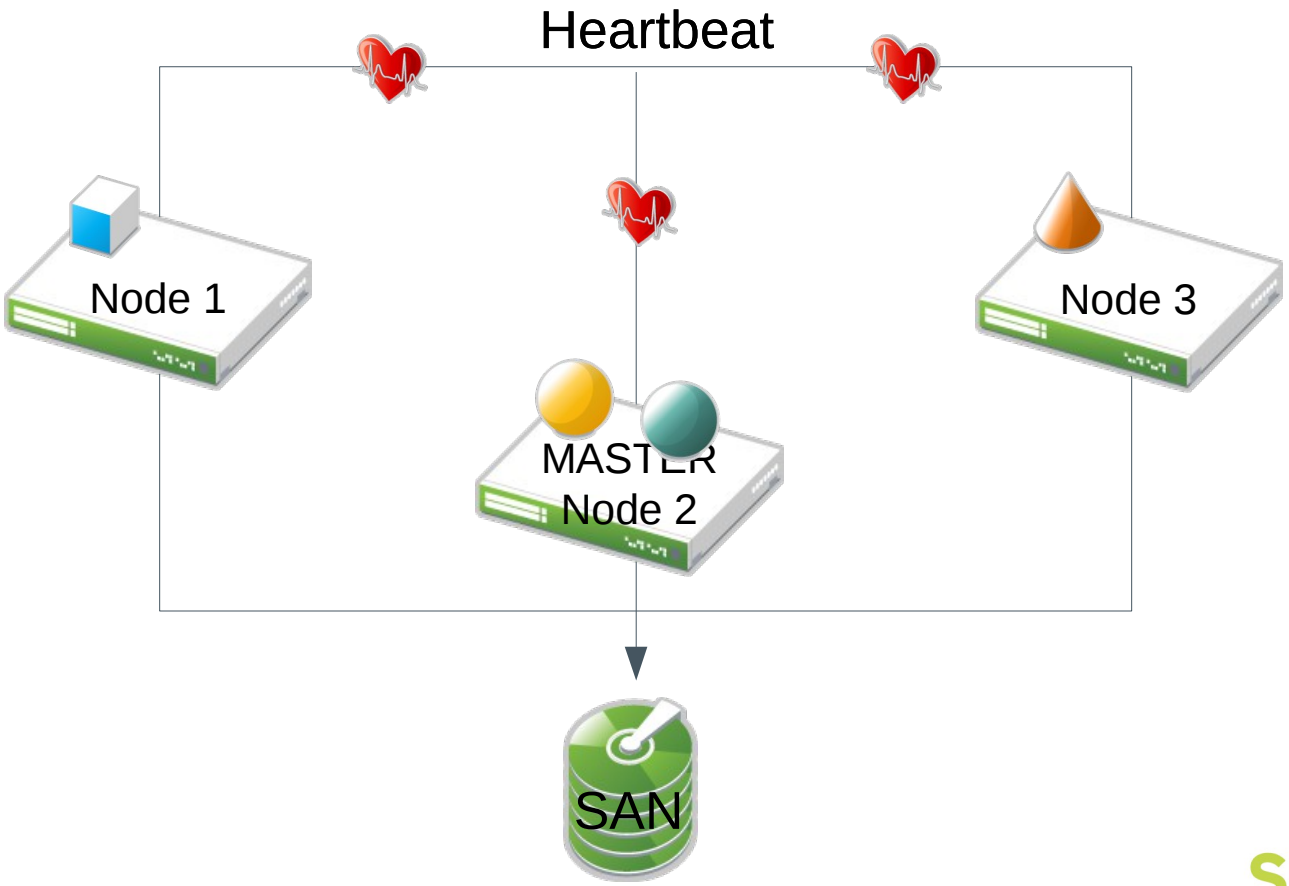
- Thank you for coming back!
- Did anyone setup a HA cluster?
  - How did it go?
- I have revamped the presentation and the demos
  - So no one should be bored!!
  - My goal is that everyone will learn new things about SLE HAE
  - Hopefully some attendees are interested in setting up SLE HAE on SLES when you get back!!
- If you want hands-on then attend the “SuSE Linux High Availability Extensions Hands-on Workshop”
  - Monday 3pm – 6pm (two sessions)
  - Hosted by Richard Lewis of IBM!

What is a high availability (HA) cluster?

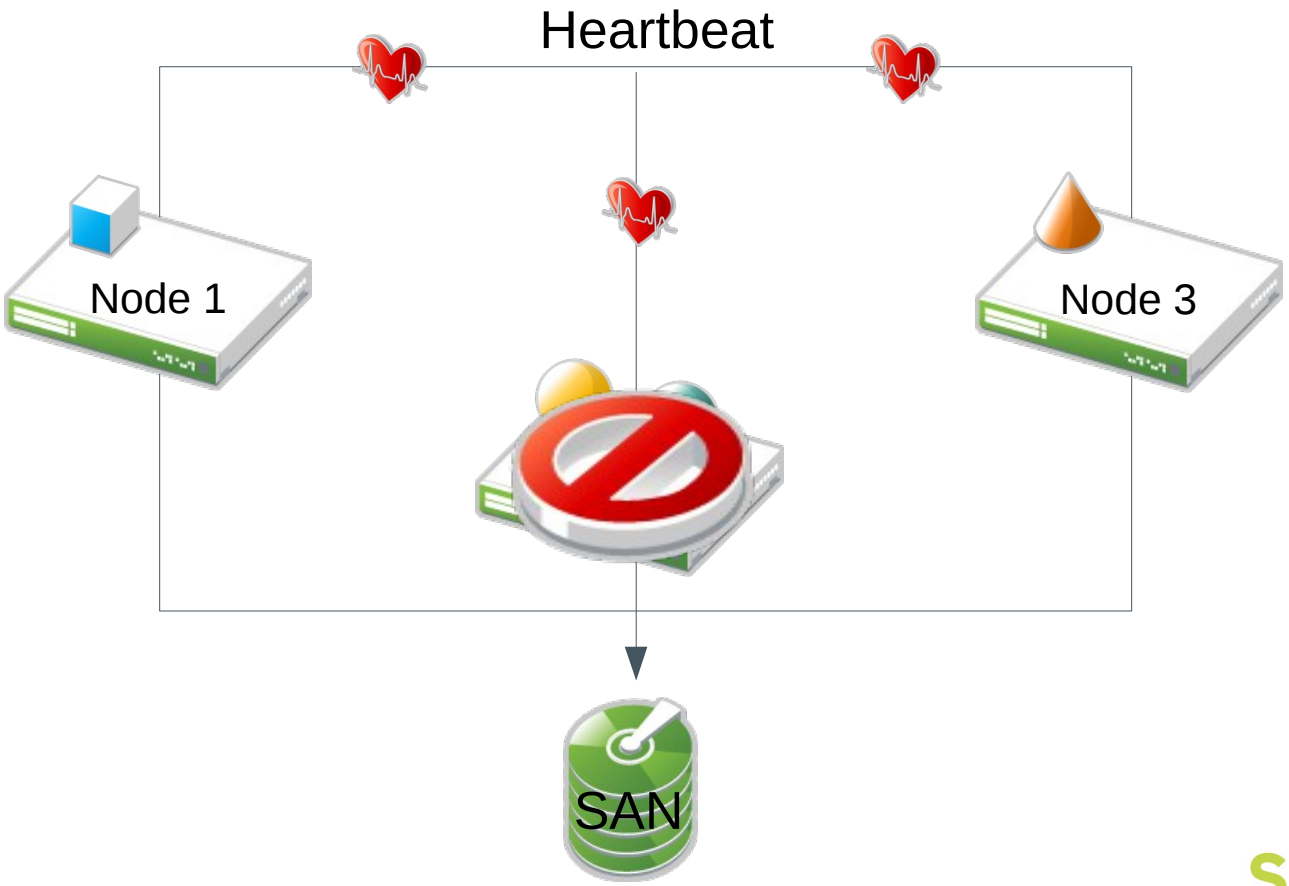
A Simple HA Cluster



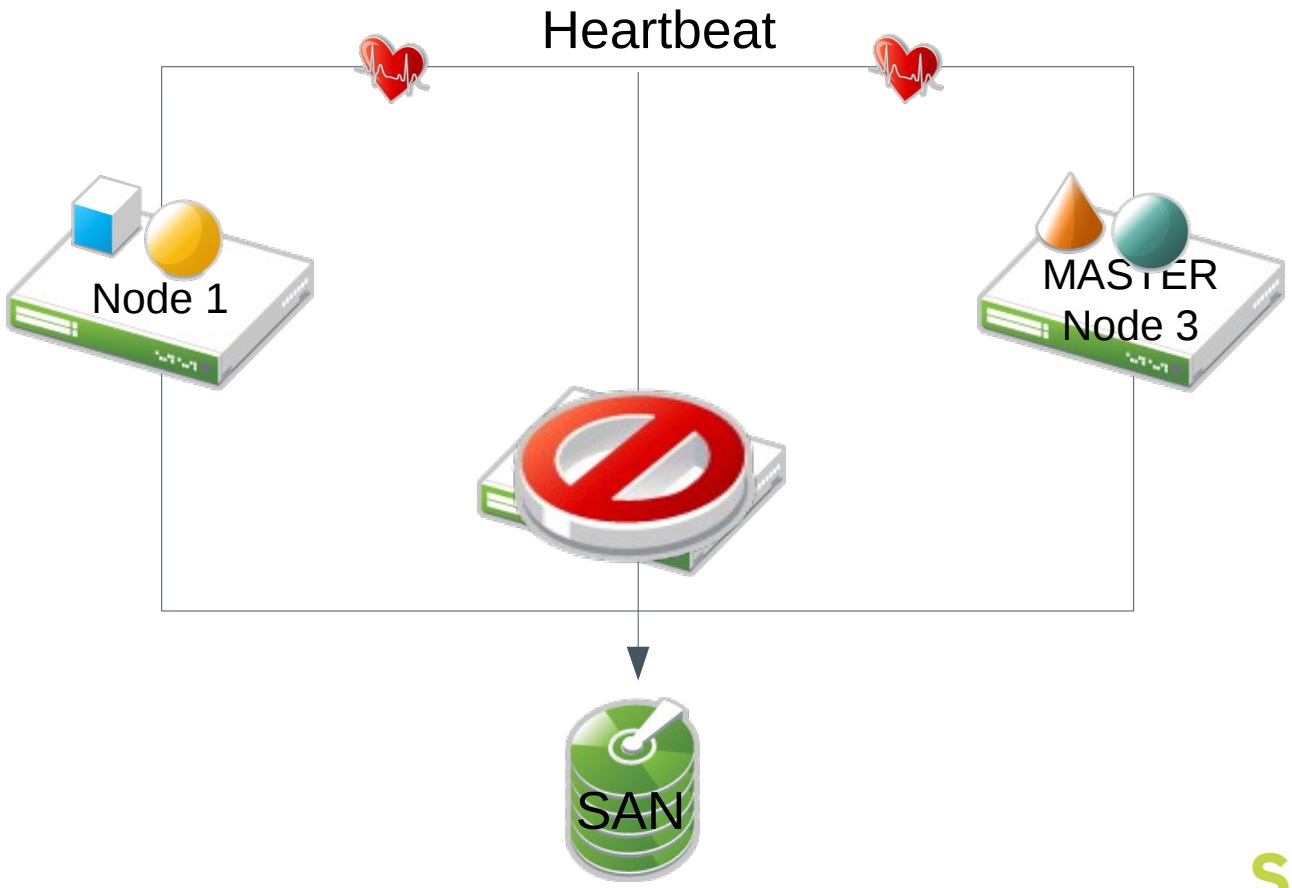
Resources Running in the Cluster



Node Failure in the Cluster



Services brought up on other nodes the Cluster

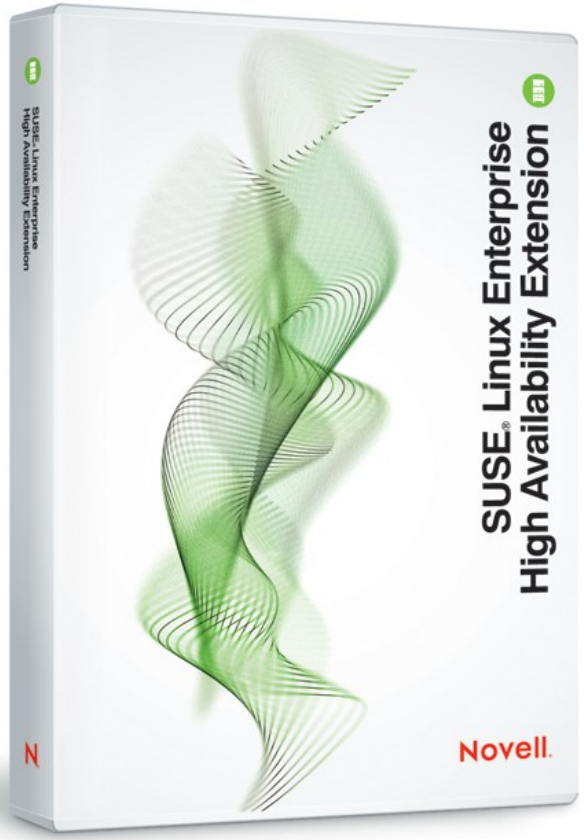


What is required to build an HA cluster using SLES?

SUSE® Linux Enterprise High Availability Extension

An affordable, integrated suite of robust open source clustering technologies that enables you to implement highly available physical and virtual Linux clusters.

Used with SUSE Linux Enterprise Server, it helps you maintain business continuity, protect data integrity, and reduce unplanned downtime for your mission critical Linux workloads



SHARE
Technology • Connections • Results
SHARE in Anaheim 2011

SUSE® Linux Enterprise High Availability Extension

Features

Flexible, policy-driven clustering solution

- OpenAIS
  - *Open Source initiative's certified implementation of the Service Availability Forum Application Interface Specification*
  - *Leading standards-based communication protocol for server and storage clustering*
  - *Messaging and membership layer*
- Pacemaker
  - *Cluster resource manager to continuously monitor resource health, manage dependencies, and automatically stop and start services*
  - *Configurable engine that uses rules and policies*
  - *Metro Area Cluster up to 20 miles*
  - *Clustered SAMBA (CIFS)*

SUSE® Linux Enterprise High Availability Extension

Features cont.

Resource Agents

- For popular third-party applications included at no extra charge
  - *SAP Instance and Database, IBM WebSphere Application Server, DB2, and Informix, Oracle and VMware*
- For popular open source applications included at no extra charge
  - *Apache, Ipv4 and IPv6, LVM, RAID, Pure-FTPd, Route, ServeRAID, Squid, VIPArip, Xen, Xinted, DRBD, Novell eDirectory™, iscsi, mysql, nfsserver, and postgres, sfex, tomcat, filesystems*
- For the most up to date list of resource agents, visit: www.novell.com/products/highavailability

SUSE® Linux Enterprise High Availability Extension

Features cont.

Cluster aware file system and volume manager

- OCFS2 (Oracle Cluster File System)
  - *Shared-disk POSIX-compliant generic cluster file system*
  - *Cluster-aware POSIX locking*
  - *Parallel I/O*
  - *Dual node use with DRBD*
- CLVM2 (Clustered Logical Volume Manager)
  - *Convenient, single, cluster-wide view of storage*
  - *Clustering extensions to the standard LVM2 toolset*
  - *Eliminates need to learn a new set of tools*

SUSE® Linux Enterprise High Availability Extension

Features cont.

Host Based Continuous data replication

- DRBD8 (Distributed Replicated Block Device)
  - *Leading open source networked disk management tool*
  - *Build single partitions from multiple disks that mirror each other*
  - *Fast data resynchronization capabilities*
  - *Supports both synchronous and asynchronous mirroring*
  - *Provides replicated storage area network (SAN) semantics, allowing cluster-aware file systems to be used without additional SANs*

SUSE® Linux Enterprise High Availability Extension

Features cont.

User-friendly tools

- Unified command line interface
  - *Powerful tool for installing, configuring and managing Linux clusters*
  - *For more experienced IT professionals*
- Graphical user interface
  - *Simple tool for monitoring and administering clustered environment*
  - *Does not require in-depth knowledge*
  - *Web Interface included*
- YaST modules
  - *DRBD*
  - *OpenAIS*
  - *Multipath*
  - *IP load balancer*

SUSE® Linux Enterprise High Availability Extension

x86 and x86_64

- Additional cost per year, per server
- Support level inherited by base SUSE Linux Enterprise Server

System z, Power, Itanium

- Bundled with base SUSE Linux Enterprise Server at no additional charge
- Support level inherited by base SUSE Linux Enterprise Server

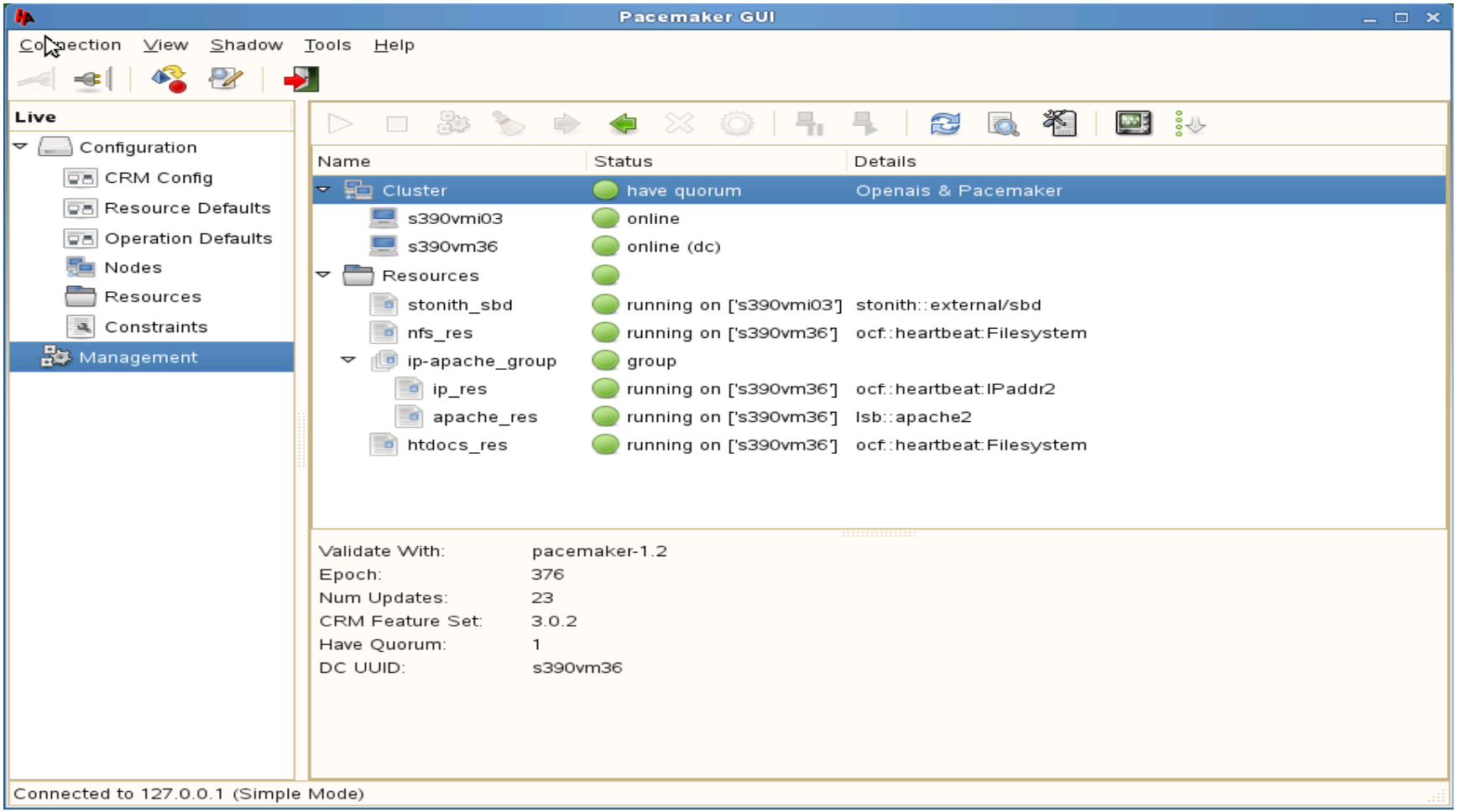
SHARE
Technology • Connections • Results

Demoing the details:

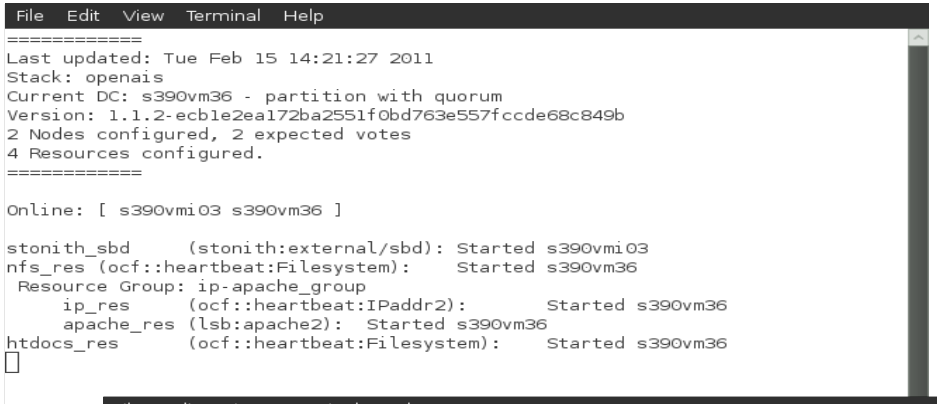
Managing a cluster with the GUI and CLI

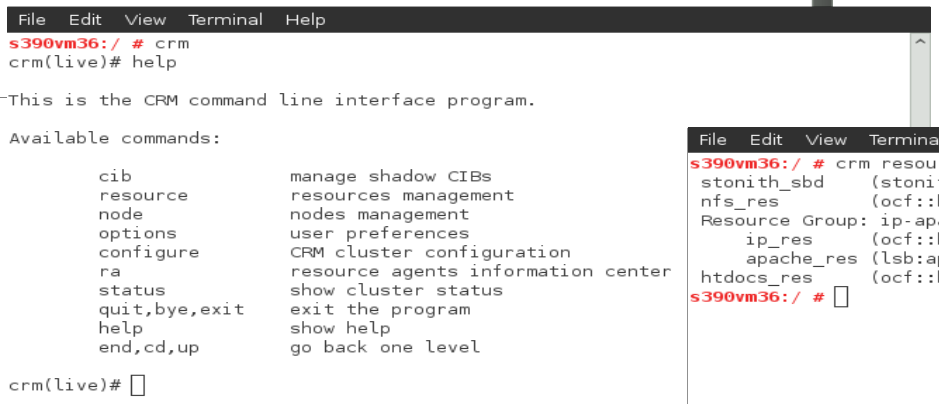
NOTE: Its hard to show a demo in slides so come to the session if you want to see the live demo! Several screenshots are provided to help visualize the demoed topic.

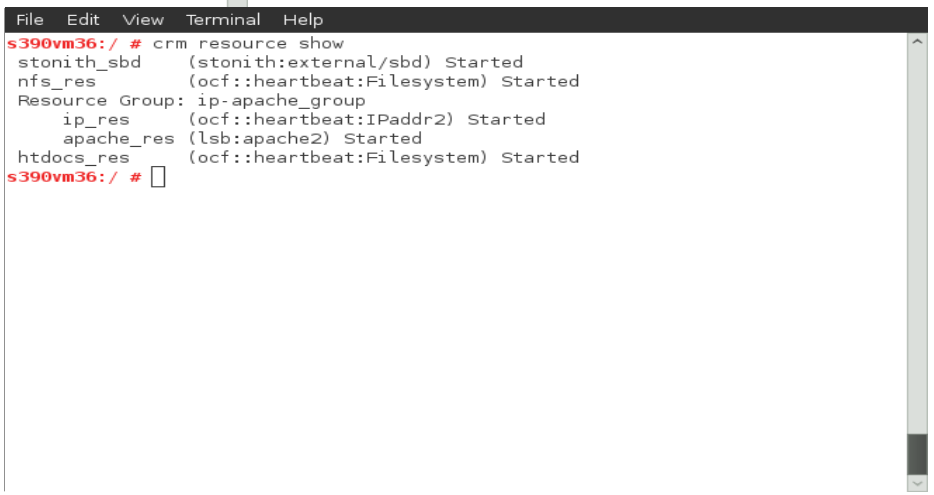
Start the GUI with crm_gui



Use crm_mon and crm for the CLI





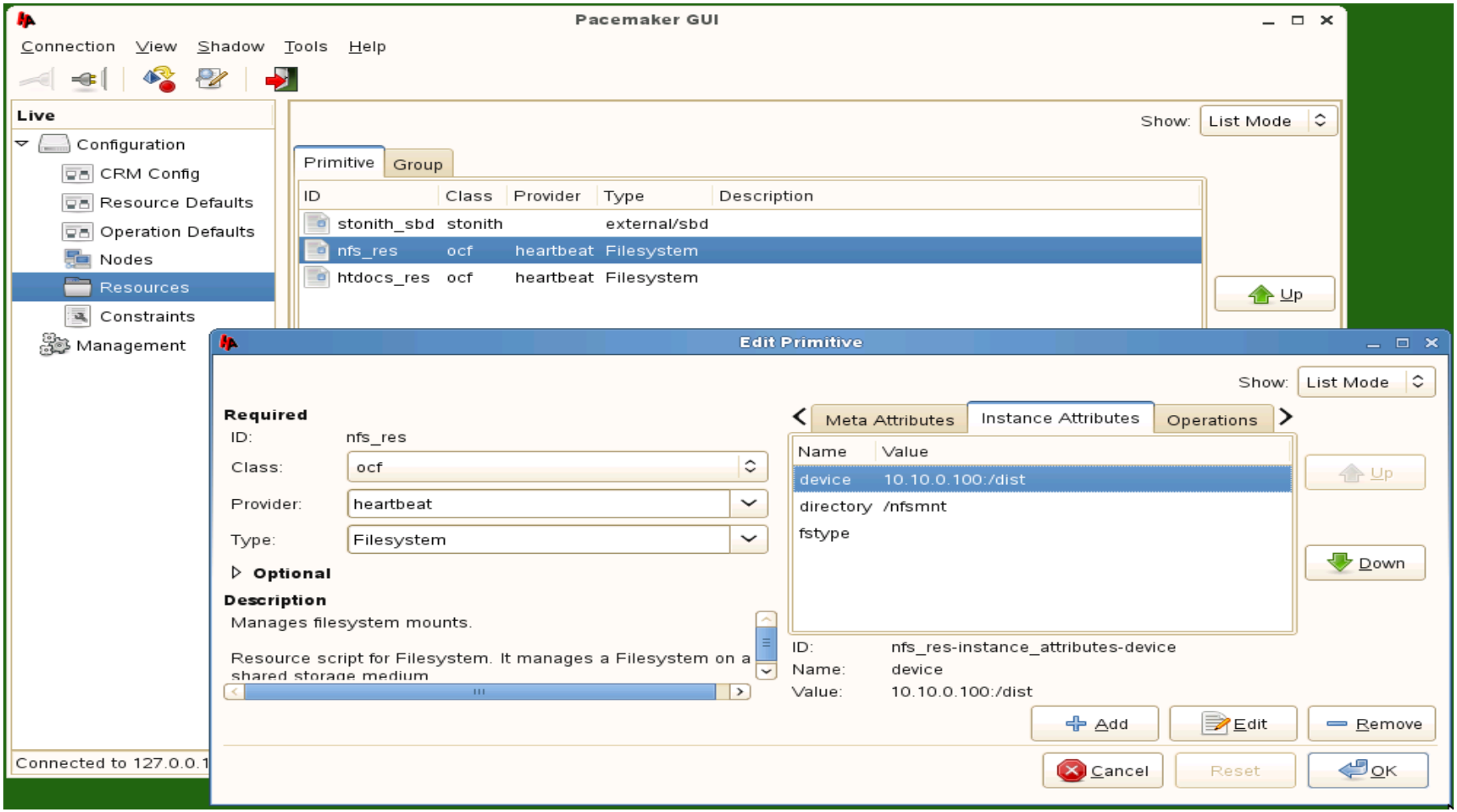


Demoing the details:

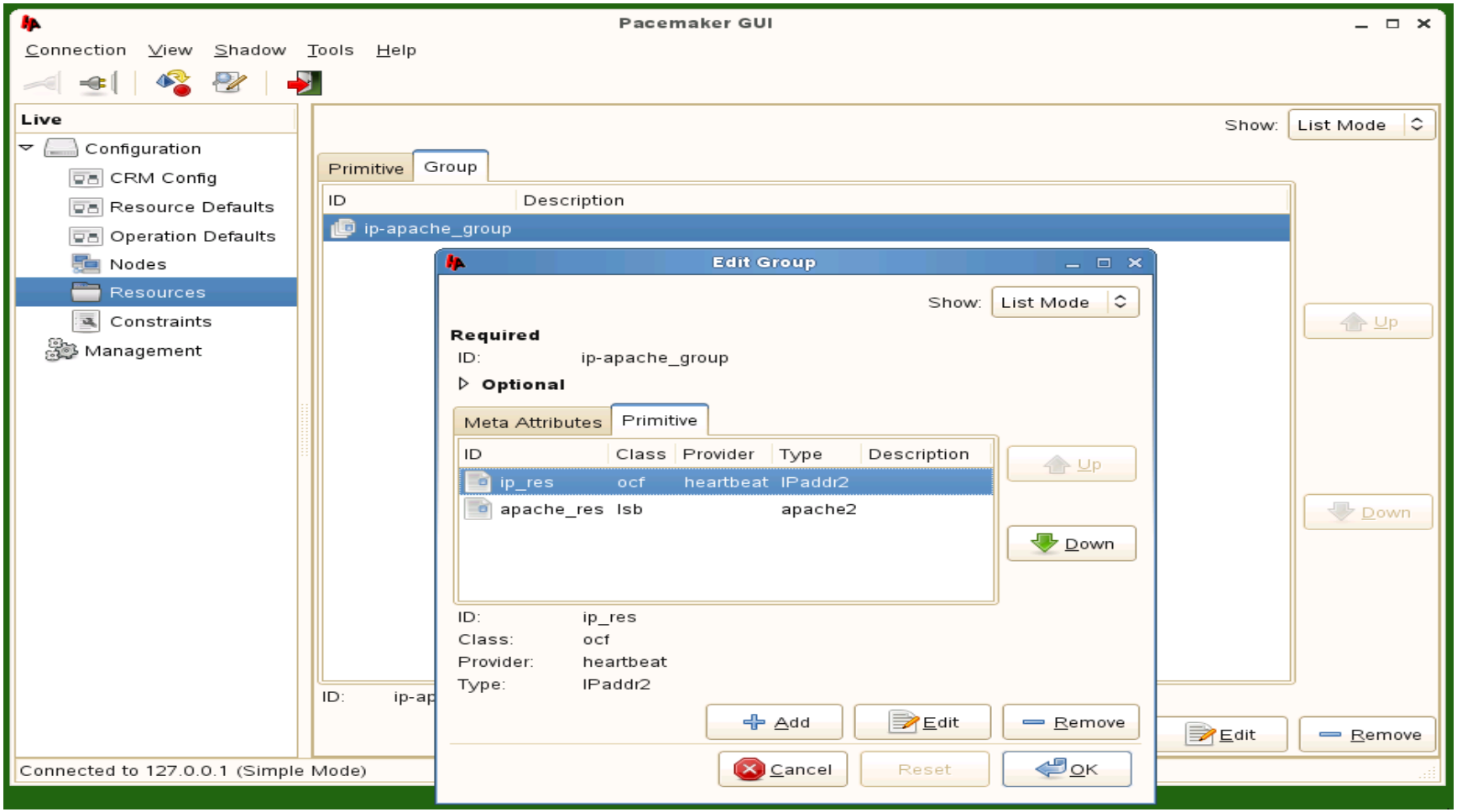
Resource primitives and resource groups

NOTE: Its hard to show a demo in slides so come to the session if you want to see the live demo! Several screenshots are provided to help visualize the demoed topic.

A resource primitive



A resource group

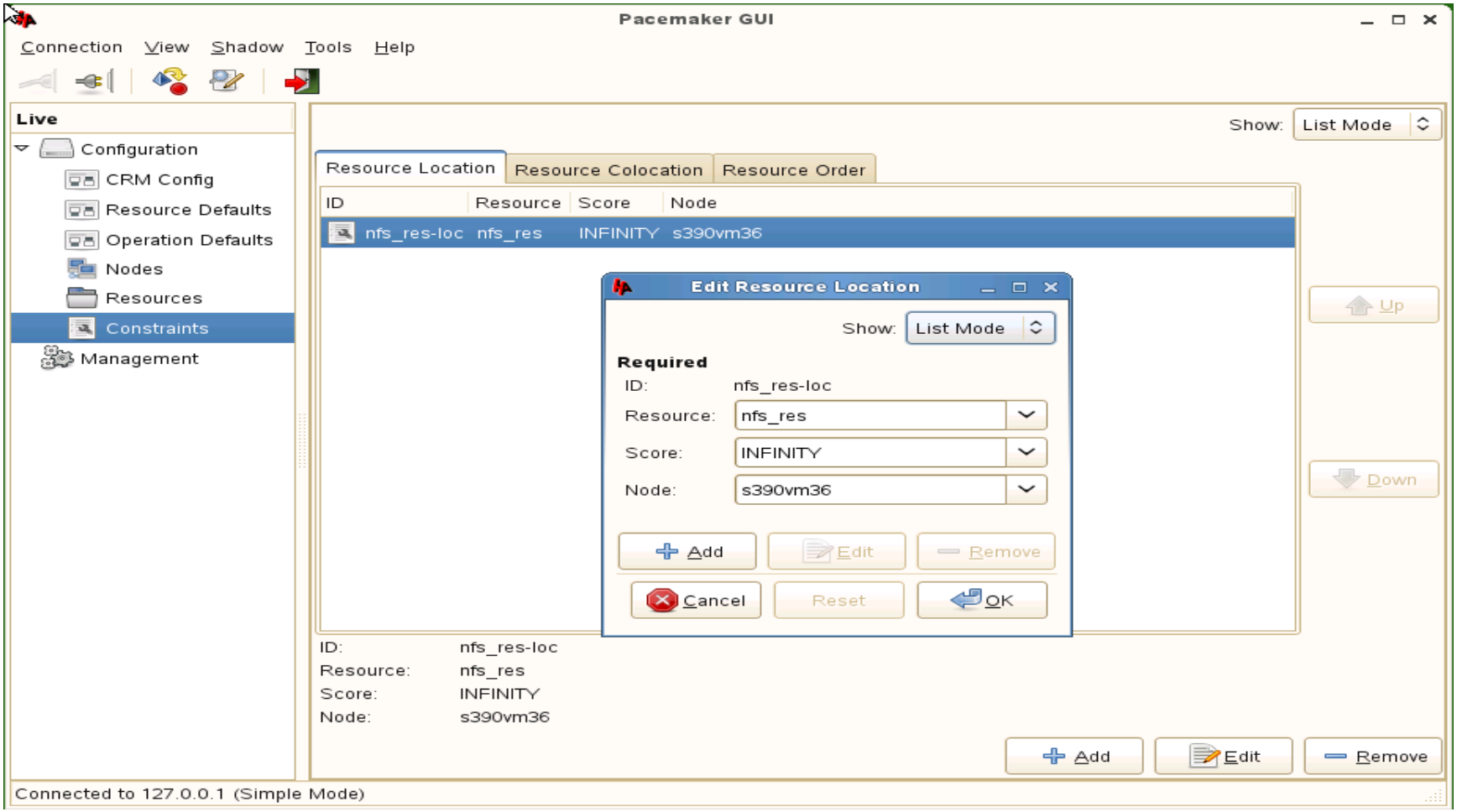


Demoing the details:

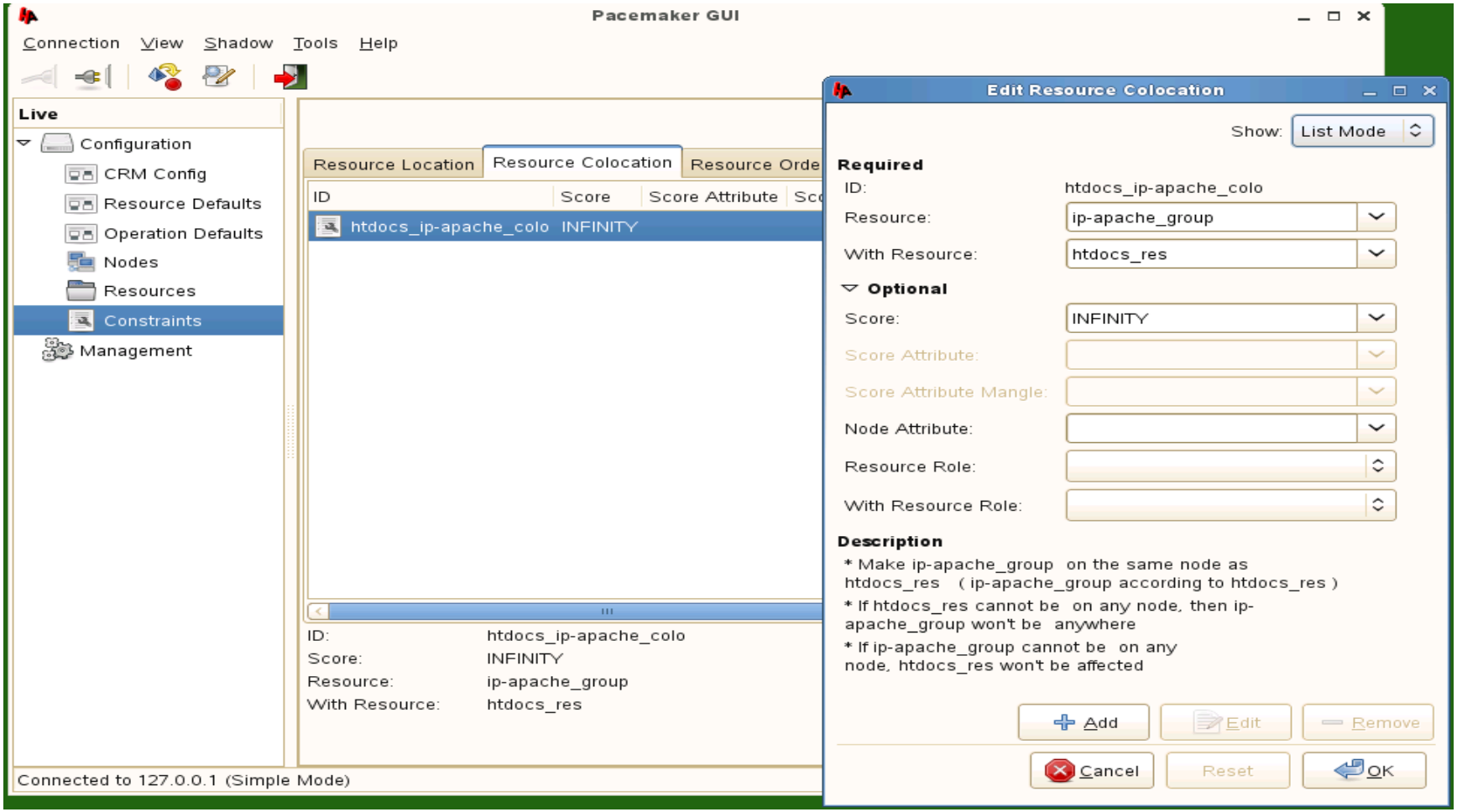
Resource Constraints

NOTE: Its hard to show a demo in slides so come to the session if you want to see the live demo! Several screenshots are provided to help visualize the demoed topic.

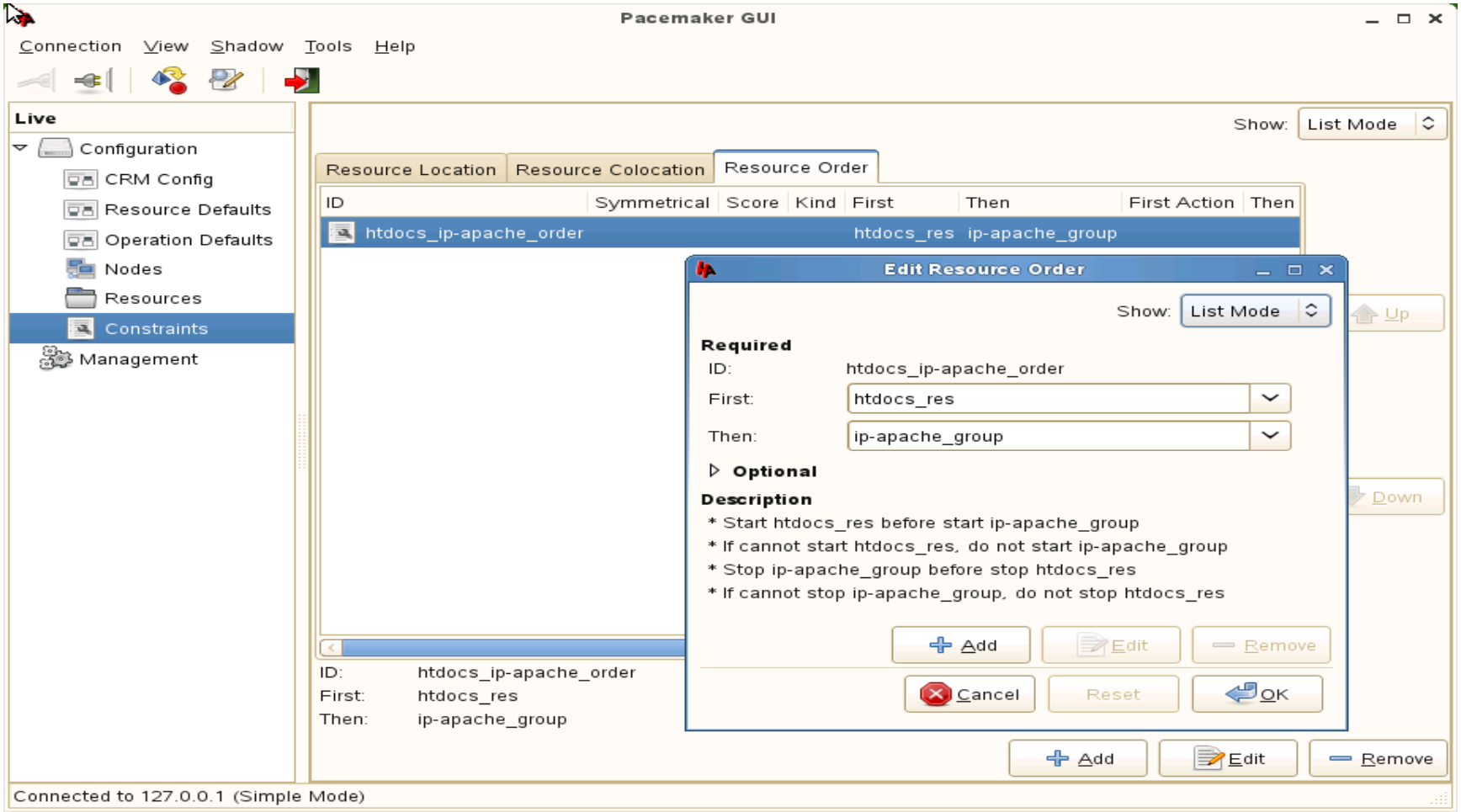
A resource location constraint



A resource colocation constraint



A resource order constraint



SHARE
Technology • Connections • Results

Demoing the details:

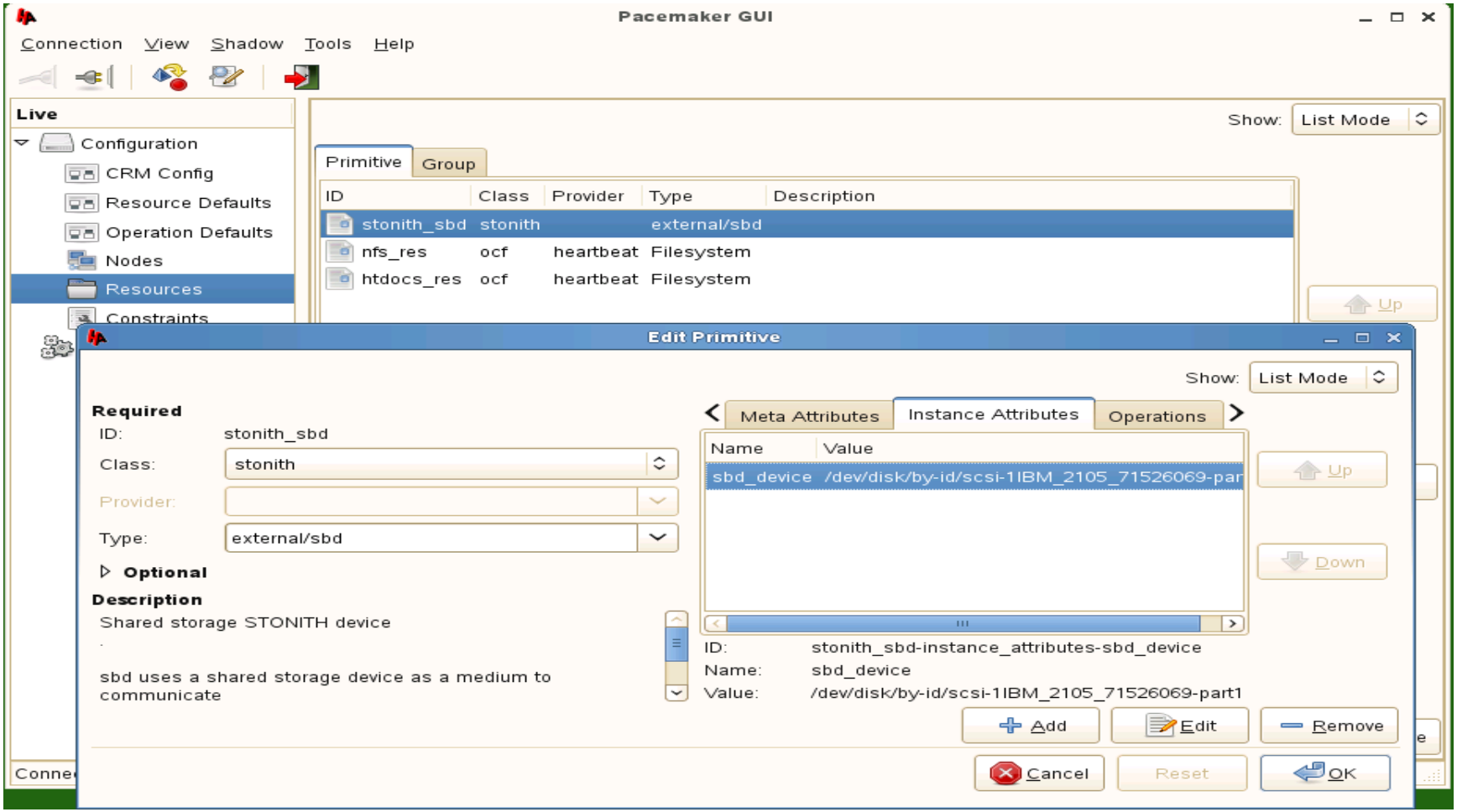
STONITH – Shoot The Other Node In The Head

NOTE: Its hard to show a demo in slides so come to the session if you want to see the live demo! Several screenshots are provided to help visualize the demoed topic.

What is STONITH?

- Simple concept
  - A machine in the cluster wants to make sure another machine in the cluster is dead
  - STONITH is used to remotely power down a node in the cluster
  - Simple and reliable, albeit admittedly brutal
- Fencing is another term but not as graphic!
- Modular and extensible
  - 33 STONITH modules included in SLE11 SP1 HAE
- SLE HAE requires a STONITH device by default
  - Recommended practice to have one configured!

A Split Brain Detector (SBD) STONITH resource



Demoing the details:

cLVM and OCFS2

NOTE: Its hard to show a demo in slides so come to the session if you want to see the live demo! Several screenshots are provided to help visualize the demoed topic.

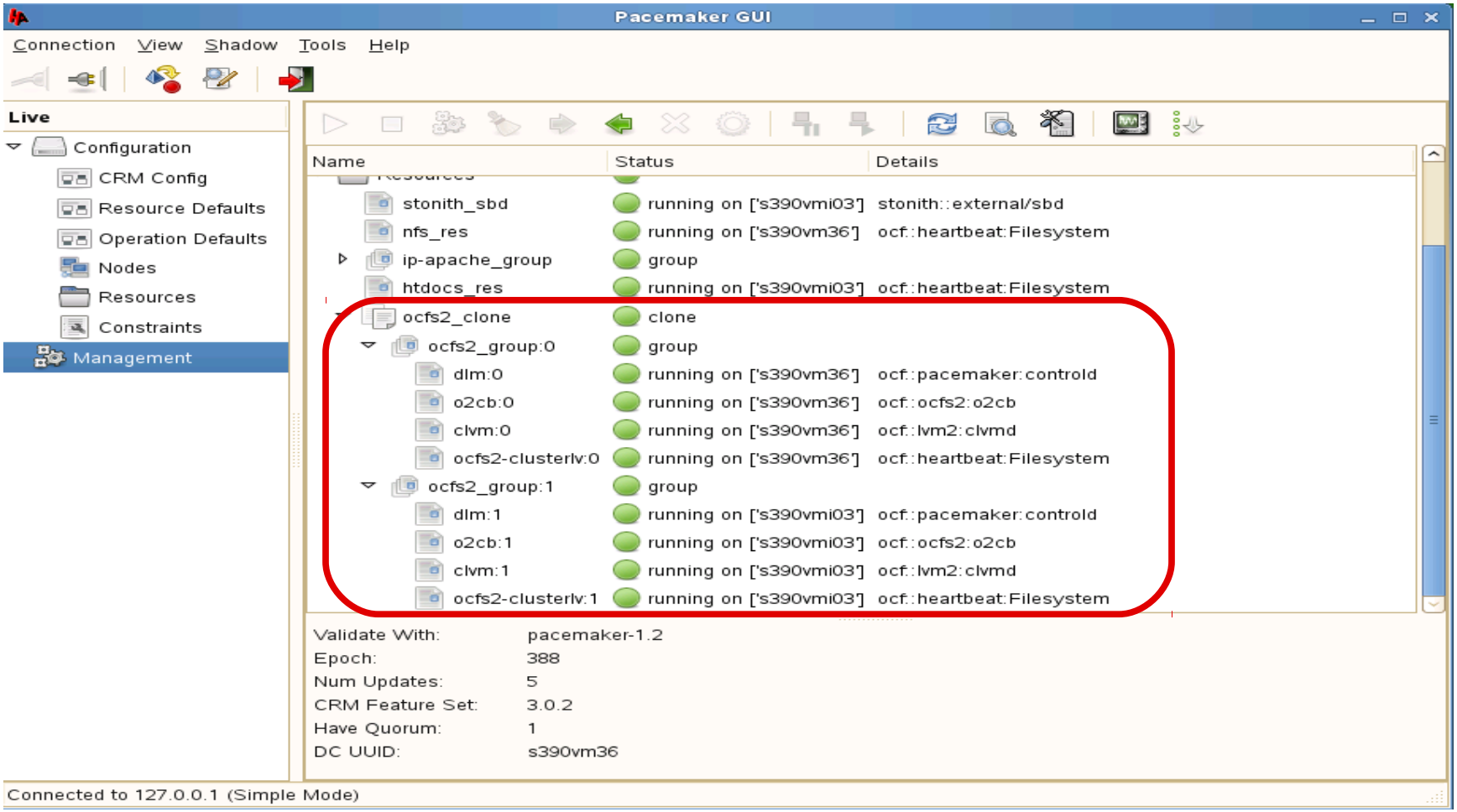
Understanding the definitions of cLVM and OCFS2 in the HA cluster

- cLVM
  - Cluster-aware logical volume manager uses the same LVM management tools to manage PVs, VGs and LVs
- OCFS2
  - Oracle Clustered File System v2
- dlm
  - Distributed Lock Manager manages locking within the cluster
- o2cb
  - OCFS2 cluster software stack
- Cloned resource
  - a resource or resource group that runs on all nodes in the cluster

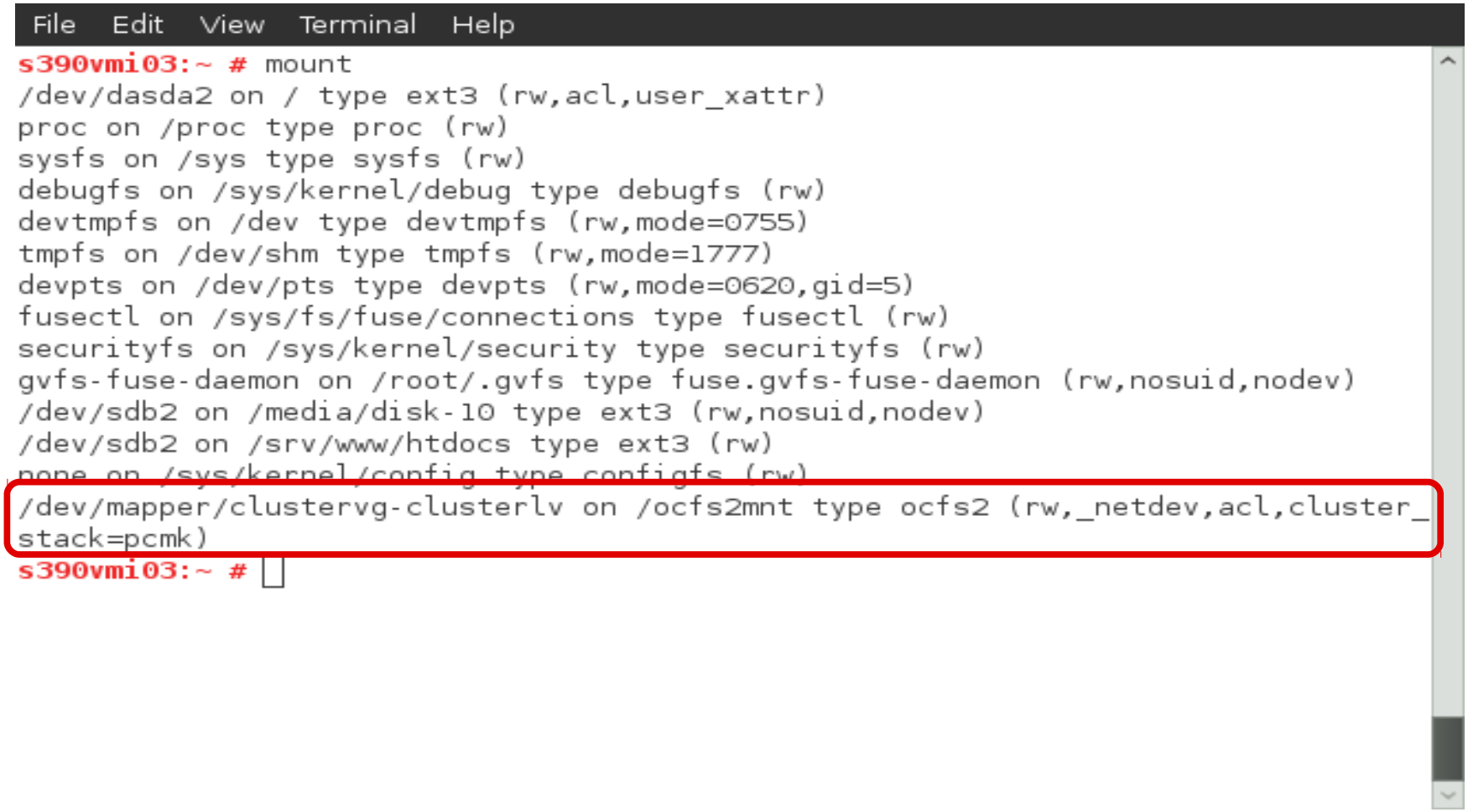
Understanding the configuration of cLVM and OCFS2 in the HA cluster

- Four resource primitives in a cloned resource group (primitive names are arbitrary)
  - dlm
  - o2cb
  - clvm
  - ocfs2-clusterlv
- Resource primitive start order is important
- The last resource primitive mounts the clustered filesystem on all nodes in the cluster

The cLVM and OCFS2 configuration



The mounted OCFS2 filesystem



Racing the space...

Grow the OCFS2 filesystem while online and being written to by all the clustered nodes!

NOTE: This demo is for attendees only!!

SHARE
Technology • Connections • Results

THANK YOU FOR ATTENDING!!