



Linux-HA Clustering, no SAN

Prof Joe R. Doupnik

jrd@netlab1.oucs.ox.ac.uk

Ingotec, Univ of Oxford, MindworksUK

Ingotec offers IT to charities

A charity is receiving a new server, having GroupWise, POSIX email, eGroupWare suite, and portions of Novell's OES11 (the Nows suite)

Active server, plus hot spare server as backup

Cost must be the least possible, thus no SAN, no shared disk etc, yet no single points of failure

Service needs to be active with only minutes gaps

The situation thus far

No SAN or equivalent, but two local disk drives
GroupWise, database copies cannot be merged
Short outage times, automatic failover
One movable IP number, not lots of them

Taken together these indicate
mirroring disk blocks over the net
cluster software to perform failover
grouping apps to move as a large unit

The general idea

Avoid single points of failure where possible

Two servers: active and hot spare

Only local disk drives, no SAN, no iSCSI etc

Must automatically fail-over within a minute or two

GroupWise database holds mail, cannot merge copies from two machines, thus need mirrored disks

Other apps (MySQL, Apache, eGroupWare, Mailman, home directories...) need to move as a unit from one machine to another

Active/Active clusters are complicated

I chose Active/Passive fail-over clustering

The general idea

Real time disk mirroring by DRBD (distributed replicated block device) is a vital component and was the most difficult to get working properly

Apps must be moved as a whole, and in a specified order as they are interrelated. Up in order, down in the reverse order.

One movable IP number, which is that seen by users

Servers are remote and are not shut down except for maintenance

Settling on a plan

After much experimenting I selected

OES11 (NOWS bundle including GroupWise)

SLES 11 SP2

SLES11 HA: pacemaker, openais/cororsync, DRBD

DRBD kernel module and user level controls have stabilised in only recent kernels, hence SLES 11 SP2 (+ HA) but not earlier

[OES11 SP1 will enable all software on SLES 11 SP2]

The HA players

Pacemaker is the decision maker and control unit, where we configure the rules of behaviour

Start/stop/status scripts interface apps/resources into pacemaker

Corosync/openais communicates status over the net to mates, using UDP in my case (not multicast rubbish)

DRBD complex subsystem mirrors the two disks

Getting started: HA add-on

The screenshot shows the YaST2 installation configuration window, specifically the 'Patterns' tab. The 'High Availability' pattern is selected, indicated by a red arrow and the text 'HA' to its left. The 'Primary Functions' list on the left includes 'High Availability' at the bottom, which is checked. The main window displays a list of packages with their summaries, installed versions, and sizes.

Package	Summary	Installed (Available)	Size
<input checked="" type="checkbox"/> cmirrord	Clustered RAID 1 support using device-mapper and ...	2.02.84-0.7.37	124.0 KiB
<input checked="" type="checkbox"/> contrack-tools	Userspace tools for interacting with the Connection ...	1.0.0-0.7.42	388.0 KiB
<input checked="" type="checkbox"/> corosync	The Corosync Cluster Engine and Application Progra...	1.4.1-0.13.1	412.0 KiB
<input checked="" type="checkbox"/> crmsh	Pacemaker/CRM shell	1.1.0-0.15.6	1.5 MiB
<input checked="" type="checkbox"/> csync2	Cluster synchronization tool	1.34-0.4.34	239.0 KiB
<input checked="" type="checkbox"/> ctdb	Clustered TDB	1.0.114.2-0.1.1	1.2 MiB
<input checked="" type="checkbox"/> drbd	Distributed Replicated Block Device	8.4.1-0.11.6	55.0 KiB
<input checked="" type="checkbox"/> hawk	HA Web Konsole	0.5.1-0.35.5	12.3 MiB
<input checked="" type="checkbox"/> ldirectord	The Heartbeat Subsystem for High-Availability Linux	3.9.2-0.25.5	163.0 KiB
<input checked="" type="checkbox"/> lvm2-clvm	Clustered LVM2	2.02.84-3.25.30	623.0 KiB
<input checked="" type="checkbox"/> ocfs2-tools	Oracle Cluster File System 2 Core Tools	1.6.4-0.3.5	2.2 MiB
<input checked="" type="checkbox"/> openais	The OpenAIS Standards-Based Cluster Framework ...	1.1.4-5.6.3	491.0 KiB
<input checked="" type="checkbox"/> pacemaker	The Pacemaker scalable High-Availability cluster res...	1.1.6-1.27.26	13.5 MiB
<input checked="" type="checkbox"/> pacemaker-mgmt	Management Tools for Pacemaker based HA solution	2.1.0-0.8.74	265.0 KiB
<input checked="" type="checkbox"/> pacemaker-mgmt-client	Management Tools for Pacemaker based HA solution	2.1.0-0.8.74	479.0 KiB
<input checked="" type="checkbox"/> pssh	Parallel SSH to control large numbers of computers ...	2.2.2-0.9.51	31.0 KiB
<input checked="" type="checkbox"/> python-dateutil	A Python Datetime Library	1.4.1-1.20	470.0 KiB
<input checked="" type="checkbox"/> python-pssh	Parallel SSH to control large numbers of computers ...	2.2.2-0.9.51	88.0 KiB
<input checked="" type="checkbox"/> release-notes-hae	The Most Important Changes for This SUSE Linux E...	10.2.0.3-9.24.2	59.0 KiB
<input checked="" type="checkbox"/> resource-agents	The Heartbeat Subsystem for High-Availability Linux	3.9.2-0.25.5	1.3 MiB
<input checked="" type="checkbox"/> sle-hae-release	SLE HASI release file	11.2-1.66	2.0 KiB
<input checked="" type="checkbox"/> sleha-bootstrap	SLE HA Bootstrap Tool	0.3-0.9.9	40.0 KiB
<input checked="" type="checkbox"/> yast2-cluster	Configuration tool for HA clusters	2.15.0-8.33.29	197.0 KiB
<input checked="" type="checkbox"/> yast2-drbd	YaST2 - DRBD Configuration	2.13.1-217.39.19	179.0 KiB
<input checked="" type="checkbox"/> yast2-iplb	Configuration tool for IP Load Balance	2.15.0-0.15.3	157.0 KiB

HA



Many goodies, I need just a few

Simplified operational steps

If ping of the IP gateway fails (or machine taken off line) then shut down all services on this machine

Else failure to communicate with the other machine implies start services on this machine

Determine which machine will be the master. Add a preference for one machine as a personal bias.

Move DRBD master to active machine, follow with apps in sequence

Permit manual migration of app group (with DRBD)

Pacemaker view (development gear)

Name	Status	Details	
Cluster	● have quorum	Openais & Pacemaker	
sles112a	● online		Nodes, a & b
sles112b	● online (dc)		
Resources	●		
MS-DRBD	● master		Disk mirroring active/passive
DRBD:0	● running (Slave) on [sles112b]	ocf::linbit:drbd	
DRBD:1	● running (Master) on [sles112a]	ocf::linbit:drbd	
APP-GROUP	● group		Items which must move as a group, in this order
ClusterIP	● running on [sles112a]	ocf::heartbeat:IPaddr2	
Files	● running on [sles112a]	ocf::heartbeat:Filesystem	
Dovecot	● running on [sles112a]	lsb::dovecot	
MySQL	● running on [sles112a]	lsb::mysql	
Apache2	● running on [sles112a]	lsb::apache2	
GW-GWDVA	● running on [sles112a]	ocf::novell:GroupWise	
GW-PO	● running on [sles112a]	ocf::novell:GroupWise	
GW-MTA	● running on [sles112a]	ocf::novell:GroupWise	
GW-GWIA	● running on [sles112a]	ocf::novell:GroupWise	
GW-IM-agent	● running on [sles112a]	lsb::novell-nmma	
GW-IM-archiver	● running on [sles112a]	lsb::novell-nmaa	
Resource_Tracker	● running on [sles112a]	lsb::rt	
PING-GATEWAY-CLONE	● clone		Connectivity ping testers
PING-GATEWAY:0	● running on [sles112b]	ocf::pacemaker:ping	
PING-GATEWAY:1	● running on [sles112a]	ocf::pacemaker:ping	

CRM command line view

```
    meta is-managed="true" allow-migrate="true"
primitive GW-GWIA ocf:novell:GroupWise \
    operations $id="GW-GWIA-operations" \
    op monitor interval="10" timeout="80" \
    params object_name="gwia.domain"
primitive GW-MTA ocf:novell:GroupWise \
    operations $id="GW-MTA-operations" \
    op monitor interval="10" timeout="10" \
    params object_name="domain"
primitive GW-PO ocf:novell:GroupWise \
    operations $id="GW-PO-operations" \
    op monitor interval="10" timeout="40" \
    params object_name="po.domain" \
    meta is-managed="true"
primitive MySQL lsb:mysql \
    operations $id="MySQL-operations" \
    op monitor interval="15" timeout="15"
primitive PING-GATEWAY ocf:pacemaker:ping \
    operations $id="PING-GATEWAY-operations" \
    op monitor interval="10" timeout="60" \
    params host_list="10.0.0.254" attempts="2" \
    meta is-managed="true" target-role="Started"
group APP-GROUP ClusterIP Files Dovecot MySQL Apache2 GW-GWDVA GW-PO GW-MTA GW-GWIA \
    meta is-managed="true" target-role="Started"
ms MS-DRBD DRBD \
    meta master-max="1" master-node-max="1" clone-max="2" clone-node-max="1" notify="true" target-role="
"Started" is-managed="true"
clone PING-GATEWAY-CLONE PING-GATEWAY \
    meta target-role="Started" interleave="true"
location MS-DRBD-ON-CONNECTED MS-DRBD \
    rule $id="MS-DRBD-ON-CONNECTED-rule" $role="master" -inf: not_defined pingd or pingd lte 0
location PREFER-sles112a MS-DRBD 100: sles112a
location cli-prefer-APP-GROUP APP-GROUP \
    rule $id="cli-prefer-rule-APP-GROUP" inf: #uname eq sles112a
colocation APP-GROUP-WITH-DRBD inf: MS-DRBD:Master APP-GROUP:Started
order MS-DRBD-BEFORE-APP-GROUP inf: MS-DRBD:promote APP-GROUP:start symmetrical=true
order PING-GATEWAY-BEFORE-MS-DRBD 1000: PING-GATEWAY-CLONE:start MS-DRBD:promote symmetrical=true
```

A portion of the rule sets
as seen by command
line utility CRM

Not all rules can be
created via the GUI or by
Hawk web interface

Group members, in order

Required
ID: APP-GROUP
Optional

Show: List Mode

Meta Attributes Primitive

ID	Class	Provider	Type	Template	Description
ClusterIP	ocf	heartbeat	IPAddr2		
Files	ocf	heartbeat	Filesystem		
Dovecot	lsb		dovecot		
MySQL	lsb		mysql		
Apache2	lsb		apache2		
GW-GWDVA	ocf	novell	GroupWise		
GW-PO	ocf	novell	GroupWise		
GW-MTA	ocf	novell	GroupWise		
GW-GWIA	ocf	novell	GroupWise		

ID: ClusterIP
Class: ocf
Provider: heartbeat
Type: IPAddr2

+ Add Edit - Remove
Cancel Reset OK

Assign cluster IP (user visible),
then mount mirrored file system.
After that start apps in sequence.

Shut down items from bottom up

Thus no open files to worry about

Down

Control of DRBD by pacemaker

Required

ID: DRBD

Class: ocf

Provider: linbit

Type: drbd

Template:

Optional

Description

Manages a DRBD device as a Master/Slave resource.

This resource agent manages a DRBD resource as a master/slave resource.

DRBD is a shared-nothing replicated storage device.

Note that you should configure resource level fencing in DRBD this cannot be done from this resource agent.

See the DRBD User's Guide for more information.

<http://www.drbd.org/docs/applications/>

Meta Attributes

Name	Value
migration-threshold	2
is-managed	true
target-role	Started

ID: DRBD-meta_attributes-migration-threshold

Name: migration-threshold

Value: 2

+ Add Edit - Remove

Cancel Reset OK

Use Linbit's OCF start/stop/status script for DRBD, else trouble

Educating Pacemaker

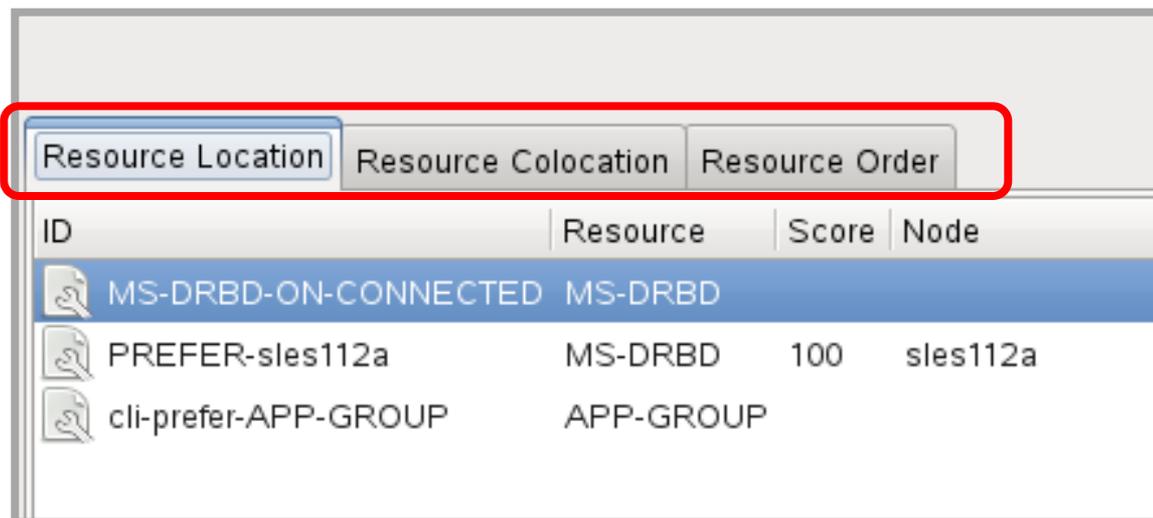
Location rule: DRBD master requires a valid gateway connection (ping does the testing)

Co-location rule: application group goes with DRBD master (the scout brings along its friends)

Ordering rule: ping gateway, choose DRBD master, follow it with application group (please form an orderly queue)

Then it decides which side to use if both are eligible

The rule book, three chapters



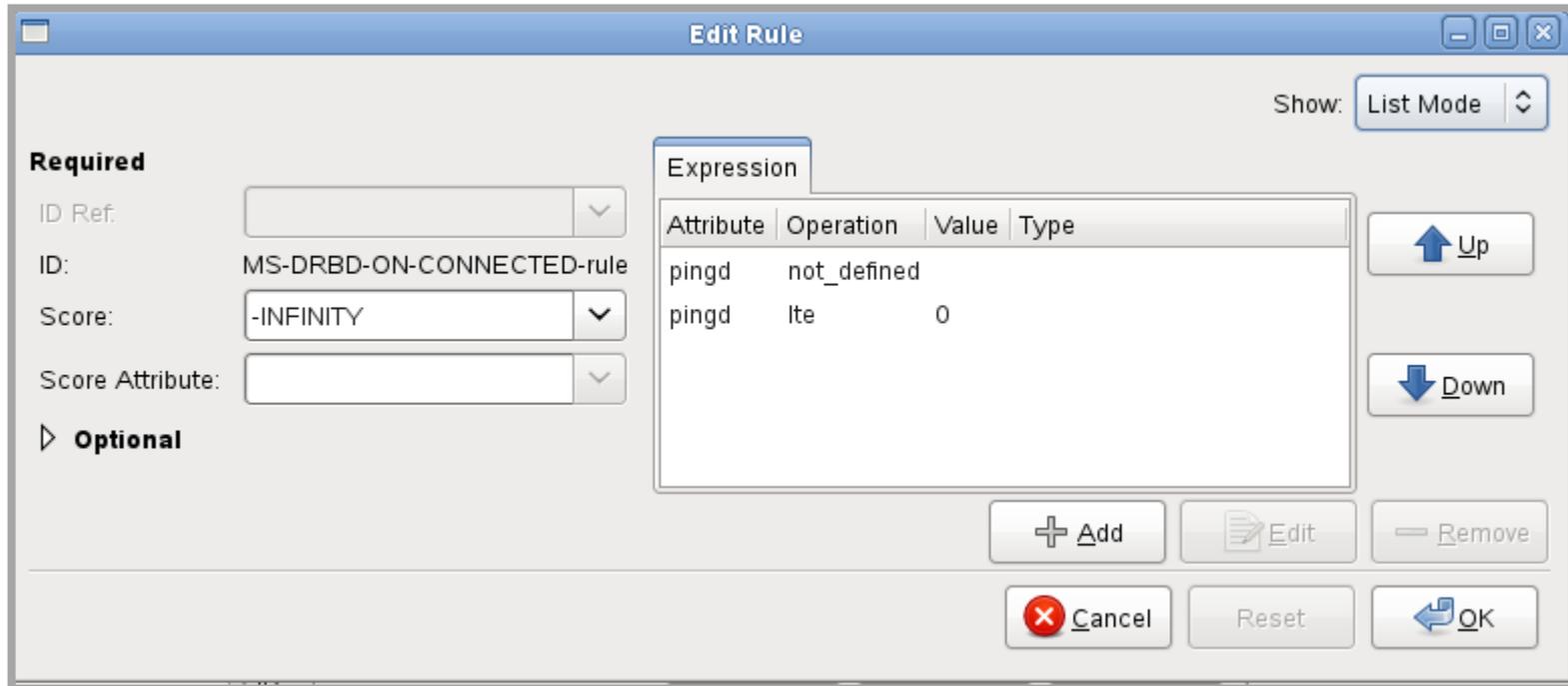
ID	Resource	Score	Node
MS-DRBD-ON-CONNECTED	MS-DRBD		
PREFER-sles112a	MS-DRBD	100	sles112a
cli-prefer-APP-GROUP	APP-GROUP		

DRBD master need gateway connectivity. If no connectivity then don't run apps on this machine.

Thumb on the scales: prefer sles112a ,if given choice

Remember last voluntary migration and put things there

Detail on DRBD needs gateway



Strangely, this says

if there is no ping daemon or
its result is less than or equal to 0 successful pings
then assign a score of -infinity (don't run DRBD master here)
It is a veto rule

Pacemaker colocation rule

The screenshot displays the Pacemaker configuration interface. On the left, a table lists resource colocation rules. The rule 'APP-GROUP-WITH-DRBD' is selected, showing a score of 'INFINITY'. The main window, titled 'Edit Resource Colocation', provides a detailed view of this rule's configuration.

Required

- ID: APP-GROUP-WITH-DRBD
- Resource: APP-GROUP
- With Resource: MS-DRBD

Optional

- Score: INFINITY
- Score Attribute: (empty)
- Score Attribute Mangle: (empty)
- Node Attribute: (empty)
- Resource Role: Started
- With Resource Role: Master

Description

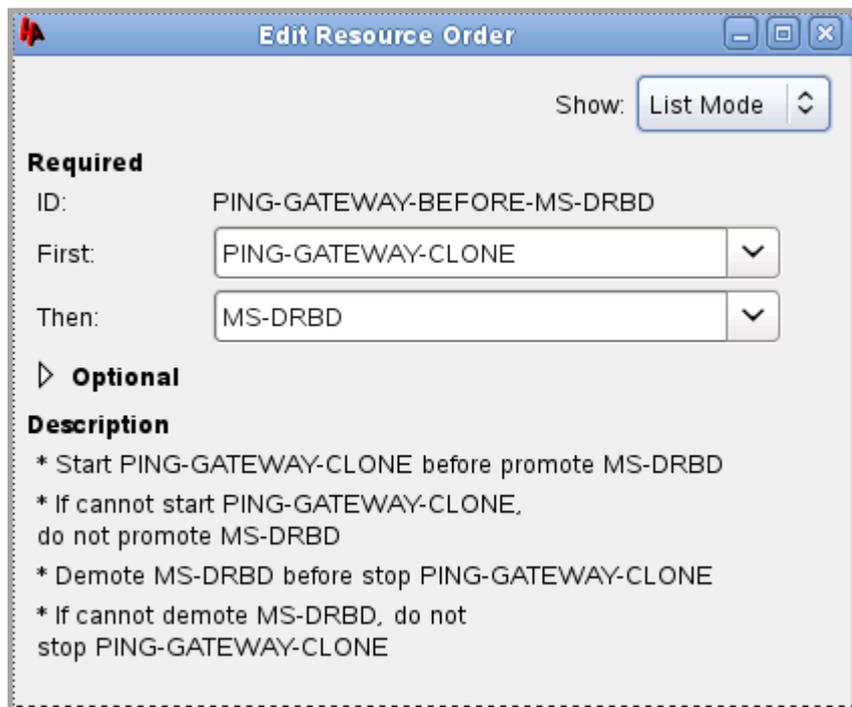
- * Make APP-GROUP Started on the same node as MS-DRBD Master (APP-GROUP according to MS-DRBD)
- * If MS-DRBD cannot be Master on any node, then APP-GROUP won't be Started anywhere
- * If APP-GROUP cannot be Started on any node, MS-DRBD won't be affected

ID	Score
APP-GROUP-WITH-DRBD	INFINITY

ID:	APP-GROUP-WITH-DRBD
Score:	INFINITY
Resource:	APP-GROUP
With Resource:	MS-DRBD
Resource Role:	Started

App-group runs only on the machine having DRBD master

Pacemaker ordering of events



Edit Resource Order

Show: List Mode

Required

ID: PING-GATEWAY-BEFORE-MS-DRBD

First: PING-GATEWAY-CLONE

Then: MS-DRBD

Optional

Description

- * Start PING-GATEWAY-CLONE before promote MS-DRBD
- * If cannot start PING-GATEWAY-CLONE, do not promote MS-DRBD
- * Demote MS-DRBD before stop PING-GATEWAY-CLONE
- * If cannot demote MS-DRBD, do not stop PING-GATEWAY-CLONE



Edit Resource Order

Show: List Mode

Required

ID: MS-DRBD-BEFORE-APP-GROUP

First: MS-DRBD

Then: APP-GROUP

Optional

Description

- * Promote MS-DRBD before start APP-GROUP
- * If cannot promote MS-DRBD, do not start APP-GROUP
- * Stop APP-GROUP before demote MS-DRBD
- * If cannot stop APP-GROUP, do not demote MS-DRBD

**Ping of gateway must work
else stop all apps on this machine
Decide which side is DRBD master
App group follows DRBD master**

“crm configure edit” view of this

```
location MS-DRBD-ON-CONNECTED MS-DRBD \  
    rule $id="MS-DRBD-ON-CONNECTED-rule" $role="master" -inf: not_defined pingd or pingd lte 0  
location PREFER-sles112a MS-DRBD 100: sles112a  
location cli-prefer-APP-GROUP APP-GROUP \  
    rule $id="cli-prefer-rule-APP-GROUP" inf: #uname eq sles112a  
colocation APP-GROUP-WITH-DRBD inf: MS-DRBD:Master APP-GROUP:Started  
order MS-DRBD-BEFORE-APP-GROUP inf: MS-DRBD:promote APP-GROUP:start symmetrical=true  
order PING-GATEWAY-BEFORE-MS-DRBD 1000: PING-GATEWAY-CLONE:start MS-DRBD:promote symmetrical=true
```

Location, location, location

Co-location

Order, order I say!

Yes, the GUI is much easier to use for these items

DRBD status via /proc/drbd

```
# cat /proc/drbd
version: 8.4.1 (api:1/proto:86-100)
GIT-hash: 91b4c048c1a0e06777b5f65d312b38d47abaea80 build by phil@fat-tyre, 2011-12-20 12:43:15
1: cs:Connected ro:Primary/Secondary ds:UpToDate/UpToDate C r-----
   ns:2918 nr:2892 dw:5810 dr:11545 al:13 bm:11 lo:0 pe:0 ua:0 ap:0 ep:1 wo:b oos:0
# █
```

Shows communication status (Connected), which side has which role (Primary/Secondary), data status on each (UpToDate), etc. Useful

***drbdadm* is the command line controller**

/usr/lib/ocf/resource.d/novell/GroupWise

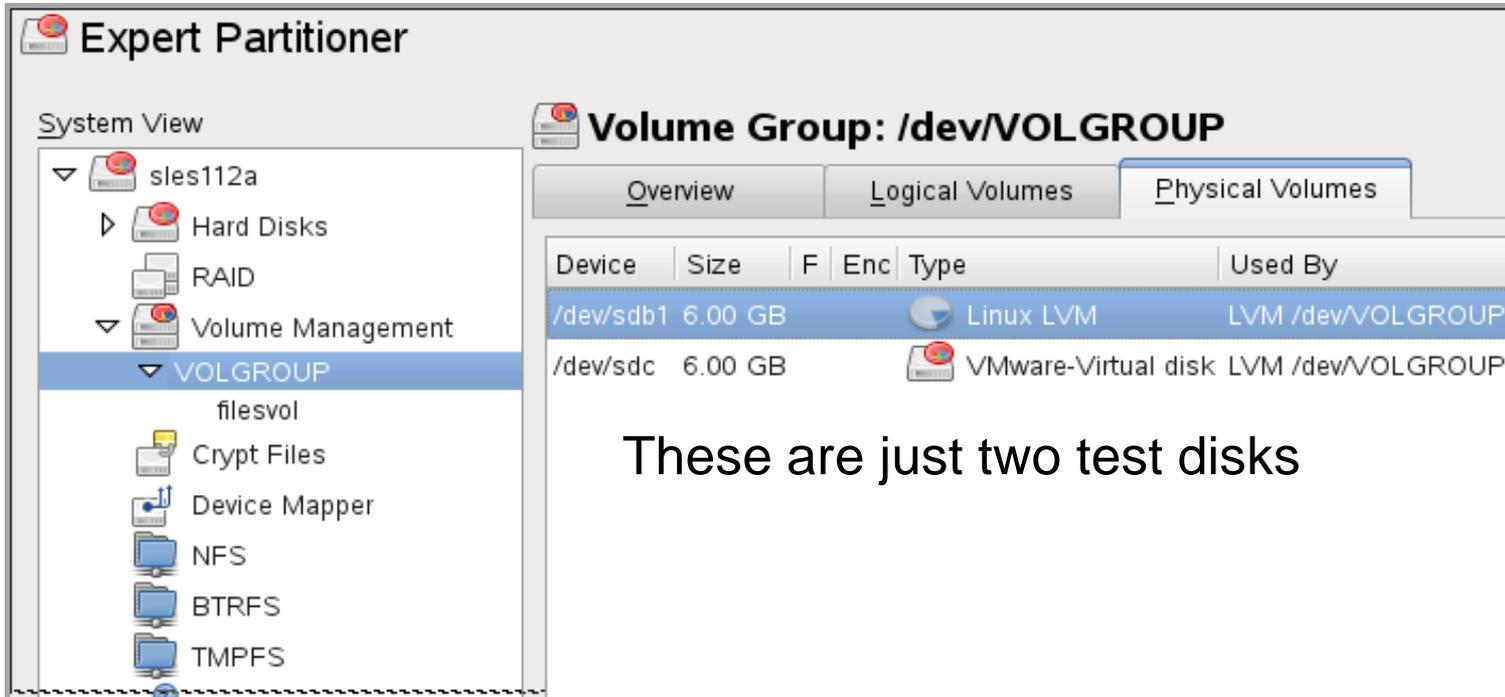
```
GroupWise_start()
{
##JRD zap uid.run files
find /home/shared/groupwise -name uid.run -print | xargs rm -f
##JRD end
    if ! [ -f ${pid_file} ] || ! checkproc -p ${pid_file} ${agent[1]} ; then
        startproc -u root -f -t ${agent[4]} -p ${pid_file} ${agent[1]} "@${agent[3]}" >& /dev/null
    fi
    return $?
}

#
# GroupWise_stop
# Desc: Stop the agent
#
GroupWise_stop()
{
    local error

    if [ -f ${pid_file} ] ; then
##JRD WAS          killproc -p ${pid_file} -TERM -t ${agent[5]} ${agent[1]} >& /dev/null
## Avoid -TERM as a -KILL will not follow automatically
        killproc -p ${pid_file} -t ${agent[5]} ${agent[1]} >& /dev/null
        error=$?
##JRD add proper cleanup of pid file
rm ${pid_file} >& /dev/null
        case $error in
```

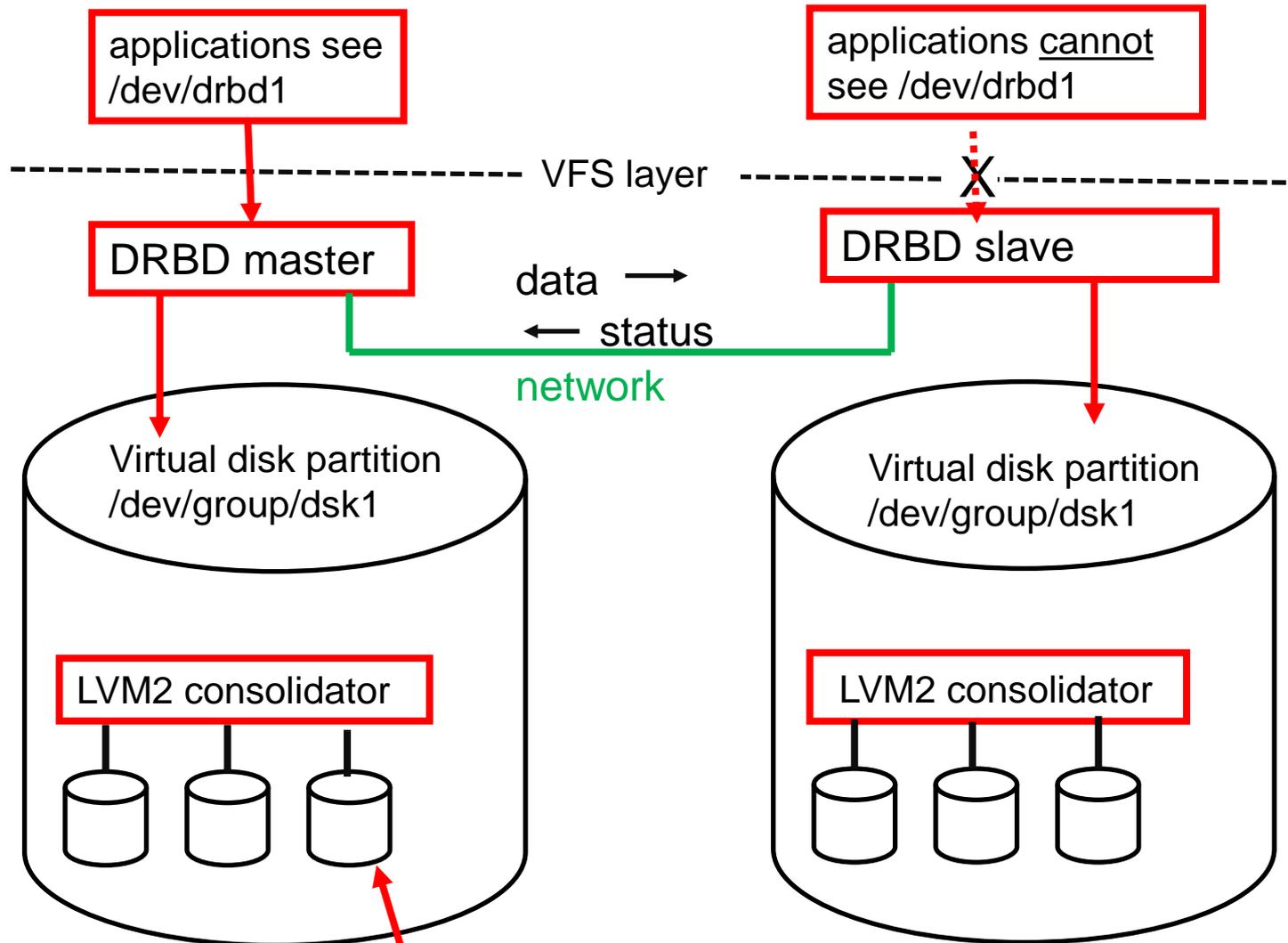
Modify original script to remove weird behaviour, else it fails

Build large virtual disk from parts



LVM2 at work here. DRBD sees the resulting large virtual disk.

DRBD disk mirroring



VMware .vmdk as "physical" disks

<http://www.drbd.org/>, DRBD site

The screenshot shows the DRBD website interface. At the top left is the DRBD logo with the tagline "Software Development for High Availability Clusters" and "A PRODUCT BY LINBIT". On the top right, there are language selection options (UK and JP flags) and a "Choose a language" dropdown menu. Below the logo, there are navigation links: "download", "docs", "site", and social media icons for Facebook and Twitter. A "60" badge is also visible.

The main navigation bar includes: HOME, DOCUMENTATION, DOWNLOAD, SUPPORT & TRAINING, MANAGEMENT CONSOLE, and USAGE.

The left sidebar under "HOME" contains a list of links:

- » What is DRBD
- » What is HA
- » Mirroring
- » Recovery
- » DRBD is open source
- » Feature List
- » Releases
- » Roadmap
- » Publications
- » External Resources
- » Mailing lists
- » IRC

The central diagram illustrates the DRBD architecture. It shows two nodes connected via a network. Each node has a "SERVICE" layer at the top. Below it is the "FILE SYSTEM", which connects to a "BUFFER CACHE". The "BUFFER CACHE" connects to the "DRBD" component. The "DRBD" component has two interfaces: "RAW DEVICE" and "TCP/IP". The "RAW DEVICE" connects to "DISK SCHED", which then connects to the "DISK DRIVER" and finally to the "DISK". The "TCP/IP" interface connects to the "NIC DRIVER", which connects to the "NIC" (Network Interface Card). The two "NIC" components are connected to each other via a network line, with two computer icons representing the nodes. A red arrow points from a green callout bubble to the "TCP/IP" interface in the left node.

Linux mainline since 2.6.33 click here

This means SLES 11 SP2, not SP1

DRBD's very own cluster GUI

The screenshot displays the Linux Cluster Management Console 1.3.9 interface. The top bar shows 'Session Help' and 'Cluster 1'. Below this, there are buttons for 'Cluster Wizard' and 'Disconnect', and an 'Operating Mode' dropdown set to 'Administrator'. The left sidebar contains a tree view with categories like Cluster Hosts, Networks, Storage (DRBD, LVM), CRM (Pacemaker), Available Services, and Services. The main area shows two nodes, sles112a and sles112b, with their respective disk and service status. A 'shared' arrow points from sles112a to sles112b. The right panel shows configuration options for global, common handlers, common startup, common net-options, and common disk-options.

Java .jar file, free, downloadable

DRBD resource configuration

```
$ more /etc/drbd.d/shared.res
resource shared {
  protocol C;
  disk {
    on-io-error pass_on;
    fencing resource-only;
  }
  handlers {
    split-brain "/usr/lib/drbd/notify-split-brain.sh root";
    fence-peer "/usr/lib/drbd/crm-fence-peer.sh";
    after-resync-target "/usr/lib/drbd/crm-unfence-peer.sh";
  }
  syncer {
    rate 80M;
  }
  net {
    after-sb-0pri discard-least-changes;
    after-sb-1pri discard-secondary;
    after-sb-2pri call-pri-lost-after-sb;
    data-integrity-alg sha1;
    use-rle;
    verify-alg sha1;
    csums-alg sha1;
  }
}
```

```
startup {
  degr-wfc-timeout 30;
  wfc-timeout 20;
}
on sles112a {
  device /dev/drbd1 ;
  address 10.0.0.21:7788;
  meta-disk internal;
  disk /dev/VOLGROUP/filesvol;
}
on sles112b {
  device /dev/drbd1 ;
  address 10.0.0.22:7788;
  meta-disk internal;
  disk /dev/VOLGROUP/filesvol;
}
}
```

Note various split-brain recovery procedures for 0, 1, and 2 primaries

Hawk, via lighttpd web server

The screenshot shows the Hawk Cluster Status web interface. The browser address bar displays `https://10.0.0.21:7630/main/status`. The page title is "Cluster Status". On the left, there is a vertical toolbar with icons for home, refresh, search, and other actions. The main content area is divided into three columns. The first column lists resources for two sets: "Master/Slave Set: MS-DRBD" and "Clone Set: PING-GATEWAY-CLONE". The second column shows the status of these resources. The third column is labeled "Inactive Resources".

Resource Set	Resource Name	Status
Master/Slave Set: MS-DRBD	sles112a	Online
	sles112b	Online
Clone Set: PING-GATEWAY-CLONE	DRBD:0	Master
	DRBD:1	Slave
Other Resources	PING-GATEWAY:0	Started
	PING-GATEWAY:1	Started
	ClusterIP	Started
	Files	Started
	Dovecot	Started
	MySQL	Started
	Apache2	Started
	GW-GWDVA	Started
	GW-PO	Started
	GW-MTA	Started
GW-GWIA	Started	

This screenshot shows a summary view of the Hawk Cluster Status. The browser address bar is the same. The page title is "Cluster Status". The main content area shows a summary of resources:

- 2 nodes configured
 - sles112a: Online
 - sles112b: Online
- 13 resources configured
 - Master/Slave Set: MS-DRBD
 - Group: APP-GROUP
 - Clone Set: PING-GATEWAY-CLONE

A red arrow points to the "Clone Set: PING-GATEWAY-CLONE" resource, which has a context menu open. The menu options are:

- Start
- Stop
- Move...
- Drop Relocation Rule
- Clean Up
- View Recent Events...

Is this different than Novell Cluster Services?

Yes, a lot

NCS uses an IP number for each resource, plus a redirector to route NCP-level requests to the proper server.

NCP resources are normally tied to NCP server objects in eDir. NCS abstracts them into eDir cluster objects.

NCS runs in Active/Active mode with an SBD partition in shared storage to coordinate access to storage. No explicit notion of resource groups.

Full OES11 support over SLES 11 SP2 is in closed beta

Ingotec version of things

This system runs:

active/passive mode mirrored disks

no shared storage (mirrored pair instead)

no SBD partition nor STONITH (no locks needed)

one IP number for all accesses by users

no NCP access (alas)

entire application group moves together

This creates a fail-over pair, not a distributed computing environment

What are the hard parts?

Finding version of DRBD, with matching kernel module, which works reliably

I tried many versions prior to SLES 11 SP2, without satisfactory results. Version 8.4.1 is best to-date.

Figuring out what pacemaker/corosync want and do. Their documentation is, er, “obscure.”

Correcting service start/stop scripts and similar to work properly. Sloppy code found.

Does it work in real life?

Amazingly, yes it does

Failover occurs as intended

Manual migration works as intended

We can mount the slave disk if its node is offline

The system is robust

The weak part is starting up both servers together and getting the right one to take over. Some hand clicking in pacemaker GUI may be required.

The costs are reasonably low (NOWS, SLES 11 HA)

What do I think of all this?

Very impressive material, needing better docs

DRBD is a gem, but its states are complicated

**Pacemaker is a marvel, if only it were documented
for mere mortals**

Corosync gets its job done, with unicast

Suggestion for others: try simple experiments

HA is a large tasty pie; we nibble around the edges

One of many guides to HA:

http://doc.opensuse.org/products/draft/SLE-HA/SLE-ha-guide_sd_draft/book.sleha.html