Overview: Clustering MySQL with DRBD & Pacemaker

Trent Lloyd <trent.lloyd@oracle.com>

Copyright 2010 Trent Lloyd, Oracle

Overview

- Software Packages (OpenAIS, Pacemaker/CRM, DRBD)
- Concepts
- Setup & configuration
  - Installing packages
  - Configure OpenAIS
  - Configure DRBD
  - Configure MySQL
  - Configuring Pacemaker CRM
  - Confirm Cluster Status
  - Basic Troubleshooting
- Questions & Discussion
- Live Demo

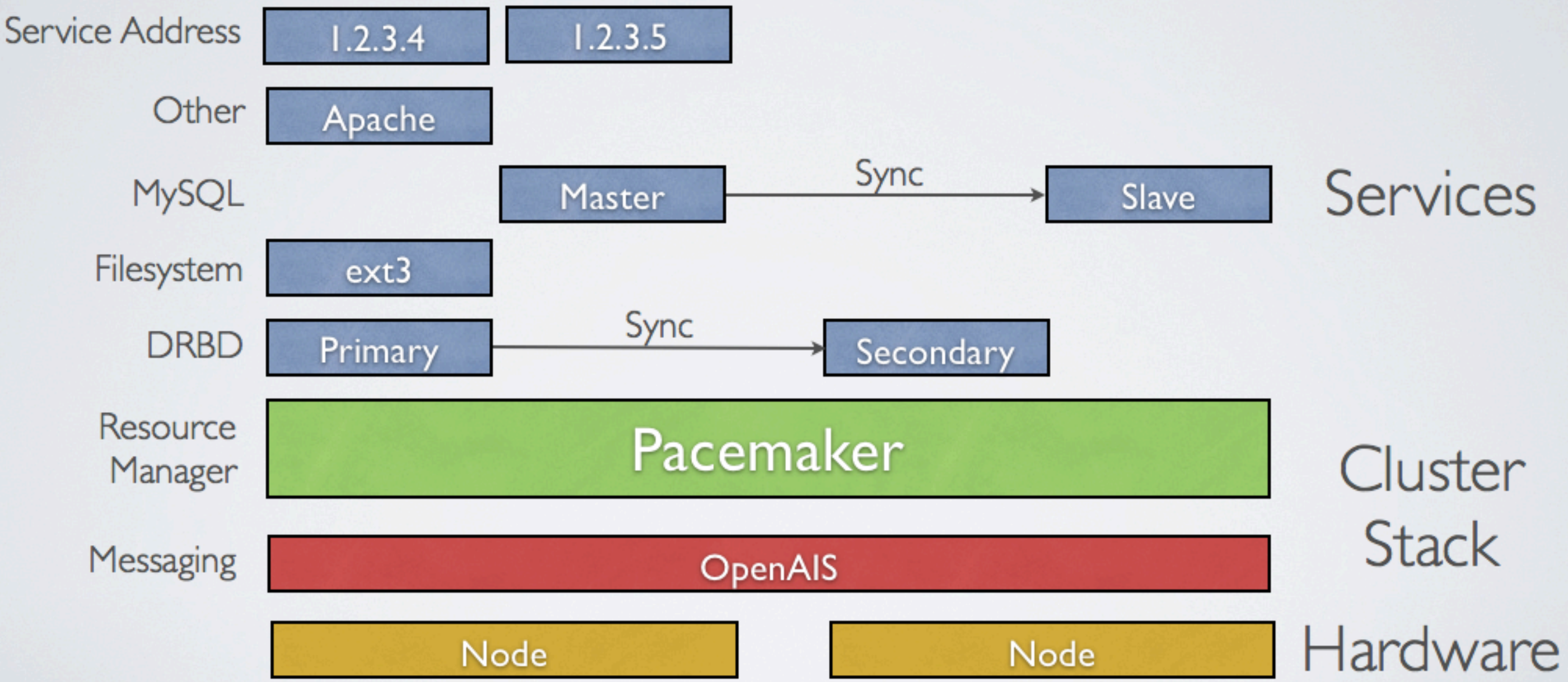
Copyright 2010 Trent Lloyd, Oracle

Software Packages

- OpenAIS ("Cluster manager") - replaces ha.cf
  - Manages the discovery of nodes, their status and transmission of cluster state information between them. "Heartbeat 3.0" is also available, but it is deprecated in favour of OpenAIS.
- Pacemaker ("Resource Manager") - replaces haresources
  - Communicates over OpenAIS to co-ordinate the starting, stopping and migration of resources between nodes. Also deals with policy decisions such as where to place a resource, what order they should be started in and if they should run on the same node.
- DRBD
  - Network-based disk replication, creates a "shared disk" between 2 nodes that is also copied to both nodes, for purposes of redundancy.
- MySQL

Copyright 2010 Trent Lloyd, Oracle

ACTIVE/PASSIVE



CC-BY-SA Source: http://www.clusterlabs.org/doc/en-US/Pacemaker/1.0/html/Pacemaker_Explained/s-intro-redundancy.html

Copyright 2010 Trent Lloyd, Oracle

Setup & Configuration - Packages

- RHEL 5 / CentOS 5
- RPM

http://www.clusterlabs.org/wiki/Install#Installing_on_EPEL_Compatible_Distributions:_RHEL.2C_CentOS.2C_etc

  - *rpm -Uvh http://download.fedora.redhat.com/pub/epel/5/i386/epel-release-5-3.noarch.rpm*
  - *wget -O /etc/yum.repos.d/pacemaker.repo http://clusterlabs.org/rpm/epel-5/clusterlabs.repo*
  - *yum install -y pacemaker corosync drbd83 kmod-drbd83 drbd-pacemaker drbd-utils*

- *MySQL*
  - *RPM download from website*
  - *MySQL-server, MySQL-shared, MySQL-client*
  - *rpm -Uvh MySQL-**

Copyright 2010 Trent Lloyd, Oracle

Setup & Configuration - Alternate Packages

- *Clusterlabs Sources:*

– *RPM: Fedora, OpenSuse, CentoS, RHEL5, etc*
– *DEB: Debian, Ubuntu*
– *Mac OS X / Darwin*
– *Source*

http://www.clusterlabs.org/wiki/Install

Copyright 2010 Trent Lloyd, Oracle

Setup & Configuration - Startup Changes

- *Disable cluster-managed services from startup*
  - *chkconfig --del heartbeat*
  - *chkconfig --del drbd*
  - *chkconfig --del mysql*

- *Disable firewall*

Copyright 2010 Trent Lloyd, Oracle

Setup & Configuration - OpenAIS

- Ensure the firewall is disabled
- Configure /etc/corosync/corosync.conf

```
totem {
  version: 2
  token: 3000
  token_retransmits_before_loss_const: 10
  join: 60
  consensus: 1500
  vsftype: none
  max_messages: 20
  clear_node_high_bit: yes
  secauth: on
  threads: 0
  rrp_mode: passive
  interface {
    ringnumber: 0
    bindnetaddr: 10.0.1.0
    mcastaddr: 239.94.1.1
    mcastport: 5405
  }
}

logging {
  to_stderr: yes
  debug: on
  timestamp: on
  to_file: no
  to_syslog: yes
  syslog_facility: daemon
}

service {
  ver:       0
  name:      pacemaker
  use_mgmtd: yes
}

aisexec {
  user:   root
  group:  root
}
```

Setup & Configuration - OpenAIS

- Ensure the firewall is disabled
- Configure /etc/corosync/corosync.conf

```
totem {
  version: 2
  token: 3000
  token_retransmits_before_loss_const:
  join: 60
  consensus: 1500
  vsftype: none
  max_messages: 20
  clear_node_high_bit: yes
  secauth: on
  threads: 0
  rrp_mode: passive
  interface {
    ringnumber: 0
    bindnetaddr: 10.0.1.0
    mcastaddr: 239.94.1.1
    mcastport: 5405
  }
}
```

```
interface {
  ringnumber: 0
  bindnetaddr: 10.0.1.0
  mcastaddr: 239.94.1.1
  mcastport: 5405
}
```

```
logging {
  name:      pacemaker
  use_mgmtd: yes
}

aisexec {
  user:   root
  group:  root
}
```

Setup & Configuration - OpenAIS

- Ensure the firewall is disabled
- Configure /etc/corosync/corosync.conf

```
totem {
  version: 2
  token: 3000
  token_retransmits_before_loss_const:
  join: 60
  consensus: 1500
  vsftype: none
  max_messages: 20
  clear_node_high_bit: yes
  secauth: on
  threads: 0
  rrp_mode: passive
  interface {
    ringnumber: 0
    bindnetaddr: 10.0.1.0
    mcastaddr: 239.94.1.1
    mcastport: 5405
  }
}
```

```
logging {
  name:      pacemaker
  group:  root
}
```

```
interface {
  ringnumber: 0
  bindnetaddr: 10.0.1.0
  mcastaddr: 239.94.1.1
  mcastport: 5405
}
```

This is the -network- address, i.e. if your IP is 10.0.1.1 and netmask 255.255.255.0 (/24) - network is 10.0.1.0 - first address

Setup & Configuration - DRBD

- Before starting corosync for the first time, you must get DRBD ready

- Create drbd.conf
  - Note in these examples, I expect the resource to be named ‘mysql’ and not ‘r0’ or other

- */etc/init.d/drbd start*
- *drbdadm -- --overwrite-data-of-peer primary mysql*
- *mkfs.ext3 /dev/drbd/by-res/mysql*

Copyright 2010 Trent Lloyd, Oracle

Setup & Configuration - Configure MySQL

- *rm -rf /var/lib/mysql/* # Remove installed version*
- *mount /dev/drbd/by-res/mysql /var/lib/mysql*

- */etc/my.cnf*
- *mysql_install_db --user=mysql # or restore backup*

- *restorecon -R /var/lib/mysql/*

- *umount /var/lib/mysql*
- *drbdadm secondary mysql*

Copyright 2010 Trent Lloyd, Oracle

Setup & Configuration - Status

- *All nodes: /etc/init.d/corosync start*
- *crm_mon*

============

Last updated: Wed Apr 21 05:23:00 2010
Stack: openais
Current DC: pacemaker1 - partition with quorum
Version: 1.0.8-9881a7350d6182bae9e8e557cf20a3cc5dac3ee7
2 Nodes configured, 2 expected votes
2 Resources configured.
============

*Online: [**pacemaker1 pacemaker2**]*

Copyright 2010 Trent Lloyd, Oracle

Setup & Configuration - Configure Pacemaker

- *CRM Configuration*

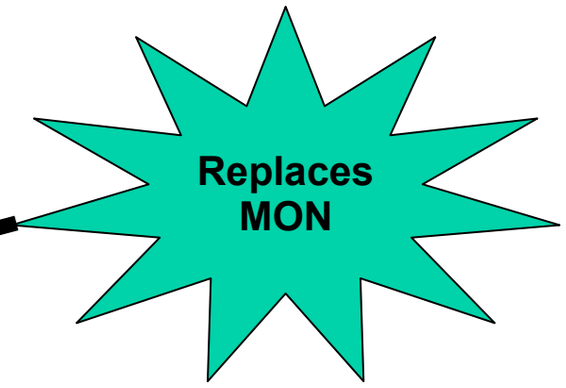
```
root@pacemaker1# crm

crm(live)# configure
crm(live)configure# primitive drbd_mysql ocf:linbit:drbd \
                    params drbd_resource="mysql" \
                    op monitor interval="15s"
crm(live)configure# ms ms_drbd_mysql drbd_mysql \
                    meta master-max="1" master-node-max="1" \
                         clone-max="2" clone-node-max="1" \
                         notify="true"
crm(live)configure# primitive fs_mysql ocf:heartbeat:Filesystem \
                    params device="/dev/drbd/by-res/mysql" directory="/var/lib/mysql" fstype="ext3"
crm(live)configure# primitive ip_mysql ocf:heartbeat:IPaddr2 \
                    params ip="10.9.42.1" nic="eth0"
crm(live)configure# primitive mysqld lsb:mysqld
crm(live)configure# group mysql fs_mysql ip_mysql mysqld
crm(live)configure# colocation mysql_on_drbd inf: mysql ms_drbd_mysql:Master
crm(live)configure# order mysql_after_drbd inf: ms_drbd_mysql:promote mysql:start
crm(live)configure# commit
crm(live)configure# exit
```

Copyright 2010 Trent Lloyd, Oracle

Setup & Configuration - Configure Pacemaker

- *CRM - "Primitives"*




```
crm(live)configure# primitive drbd_mysql ocf:linbit:drbd \
                    params drbd_resource="mysql" \
                    op monitor interval="15s"
crm(live)configure# primitive fs_mysql ocf:heartbeat:Filesystem \
                    params device="/dev/drbd/by-res/mysql" directory="/var/lib/mysql" fstype="ext3"
crm(live)configure# primitive ip_mysql ocf:heartbeat:IPaddr2 \
                    params ip="10.1.0.254" nic="eth0"
crm(live)configure# primitive mysqld lsb:mysqld
```

- primitive command creates a resource
- Uses either "OCF" or standard /etc/init.d (lsb) scripts
- A list of OCF handlers can be found in /usr/lib/ocf

- Beware of init.d scripts that are not truly LSB compliant

Copyright 2010 Trent Lloyd, Oracle

Setup & Configuration - Configure Pacemaker

- *CRM - "Master Slave"*

```
                                id          rsc
crm(live)configure# ms ms_drbd_mysql drbd_mysql \
                    meta master-max="1" master-node-max="1" \
                         clone-max="2" clone-node-max="1" \
                         notify="true"
```

options - specifies 2 node, 1 master - 1 slave

- ms makes a resource into a "Master-Slave" setup
- This means it is active on both nodes at once
- One node is a master, one node is a slave, roles switch
- Alternative to simply "started" or "stopped" with normal resource

Copyright 2010 Trent Lloyd, Oracle

Setup & Configuration - Configure Pacemaker

- *CRM - "Groups"*

```
crm(live)configure# group mysql fs_mysql ip_mysql mysqld
                    id      rsc1 rsc2 rsc3 ...
```

- group creates a group "mysql" which sub-resources
- Groups can be started and stopped as a whole, and used in other directives such for co-location and order

Setup & Configuration - Configure Pacemaker

- *CRM - "Co-location Constraints"*

```
crm(live)configure# colocation mysql_on_drbd inf: mysql ms_drbd_mysql:Master
                                  id      score  rsc          with-rsc
```

- Co-location is a constraint that says resources A and B must run on the same server at the same time

- Example: MySQL Server depends on the shared MySQL filesystem running on the same server

Setup & Configuration - Configure Pacemaker

- *CRM - "Order Constraints"*

```
crm(live)configure# order mysql_after_drbd inf: ms_drbd_mysql:promote mysql:start
```

id score first:action then:action

- order defines an order in which resources must be started

- Example: MySQL group cannot start until DRBD has become Master

Copyright 2010 Trent Lloyd, Oracle

Setup & Configuration - Configure Pacemaker

- *CRM Configuration*

```
root@pacemaker1# crm

crm(live)# configure
crm(live)configure# primitive drbd_mysql ocf:linbit:drbd \
                    params drbd_resource="mysql" \
                    op monitor interval="15s"
crm(live)configure# ms ms_drbd_mysql drbd_mysql \
                    meta master-max="1" master-node-max="1" \
                         clone-max="2" clone-node-max="1" \
                         notify="true"
crm(live)configure# primitive fs_mysql ocf:heartbeat:Filesystem \
                    params device="/dev/drbd/by-res/mysql" directory="/var/lib/mysql" fstype="ext3"
crm(live)configure# primitive ip_mysql ocf:heartbeat:IPaddr2 \
                    params ip="10.1.0.254" nic="eth0"
crm(live)configure# primitive mysqld lsb:mysqld
crm(live)configure# group mysql fs_mysql ip_mysql mysqld
crm(live)configure# colocation mysql_on_drbd inf: mysql ms_drbd_mysql:Master
crm(live)configure# order mysql_after_drbd inf: ms_drbd_mysql:promote mysql:start
crm(live)configure# commit
crm(live)configure# exit
```

Copyright 2010 Trent Lloyd, Oracle

Setup & Configuration - Status

- *crm_mon*

```
============
Last updated: Wed Apr 21 05:23:00 2010
Stack: openais
Current DC: pacemaker1 - partition with quorum
Version: 1.0.8-9881a7350d6182bae9e8e557cf20a3cc5dac3ee7
2 Nodes configured, 2 expected votes
2 Resources configured.
============

Online: [ pacemaker1 pacemaker2 ]

 Master/Slave Set: ms_drbd_mysql
     Masters: [ pacemaker1 ]
     Slaves: [ pacemaker2 ]
 Resource Group: mysql
     fs_mysql    (ocf::heartbeat:Filesystem):    Started pacemaker1
     ip_mysql    (ocf::heartbeat:IPaddr2):  Started pacemaker1
     mysqld      (lsb:mysql):    Started pacemaker1
```

Setup & Configuration - Troubleshooting

- *crm_mon*

  mysqld (lsb:mysql): Started pacemaker1 (unmanaged) FAILED

  Failed actions:

  mysqld_stop_0 (node=pacemaker1, call=86, rc=1, status=complete): unknown error

- *Fix problems*
  - *Check /var/log/messages*
- *crm(live)# resource cleanup mysqld*
- *Other actions*
  - *crm(live)# resource stop mysqld*
  - *crm(live)# resource start mysqld*
  - *crm(live)# configure show*

Copyright 2010 Trent Lloyd, Oracle

Resources

- DRBD User's Guide
  - http://www.drbd.org/docs/applications/
- Cluster Labs (Pacemaker)
  - http://www.clusterlabs.org/

Copyright 2010 Trent Lloyd, Oracle

END OF PREPARED PRESENTATION

QUESTIONS & DISCUSSION

Overview: Clustering MySQL with DRBD & Pacemaker

Trent Lloyd <trent.lloyd@oracle.com>

Copyright 2010 Trent Lloyd, Oracle