# *Building A Highly Available Linux Cluster*

Presentation by:
John Lange
john.lange@open-it.ca

OpenIT
OpenSource Technology

# *Goals*

Build a Linux cluster that:

- is highly-available
- is suitable for dynamic, data-intensive operations
- does not require specialized hardware
- is cost effective

OpenIT
OpenSource Technology

# *Requirements for High Availability*

- Both nodes must have identical information and be fully interchangeable
- Node monitoring / failure detection
- Automatic take-over
- A broken node can be reintroduced to the cluster easily

OpenIT
OpenSource Technology

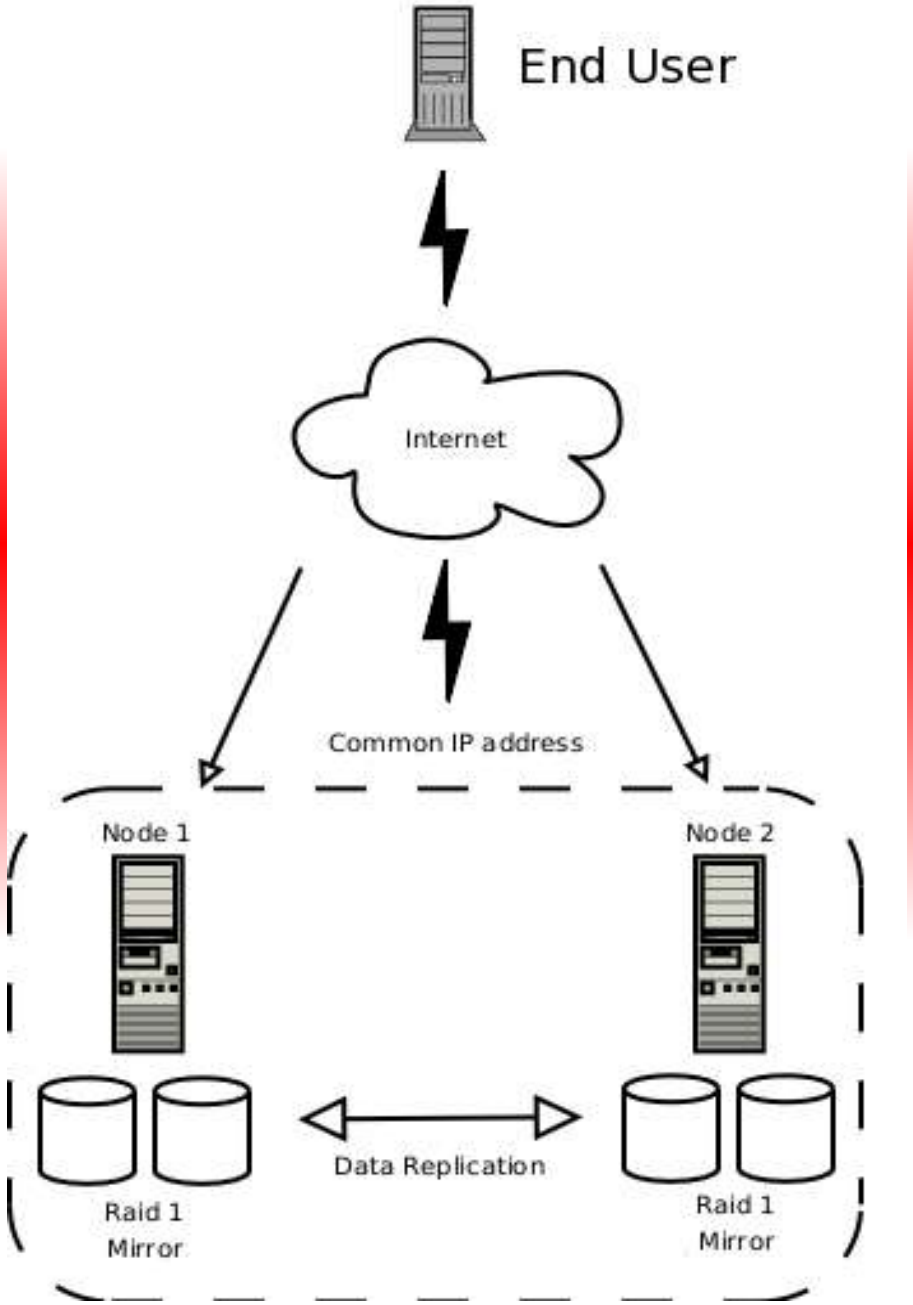
# *The Big Challenge: Real-time data replication*

Here are some potential solutions that do **NOT** meet our requirements:

- mySQL replication
  reason: one-way only
- Back-end data storage solutions
  reason: cost
- Front-end solutions (switches)
  reason: no data replication
- Commercial clustering
  reason: cost, too specialized

OpenIT
OpenSource Technology

# Network Topology




OpenIT
OpenSource Technology

# *Parts of the Solution*

- LAMP - Linux (2.4.22) Apache mySQL PHP (of course)
- drbd 0.6.12 (data replication)
- heartbeat 1.2.2 (monitoring and take-over)
- openMosix 2.4.22 (load balancing)

OpenIT
OpenSource Technology

# *Installation Overview*

- Setup your basic Linux build leaving a partition to use for replication.
- Setup drbd and get replication working
- Install Apache, PHP, mySQL etc. on the replicated partition and bring up a complete test environment
- Manually test the two nodes
- Install heartbeat

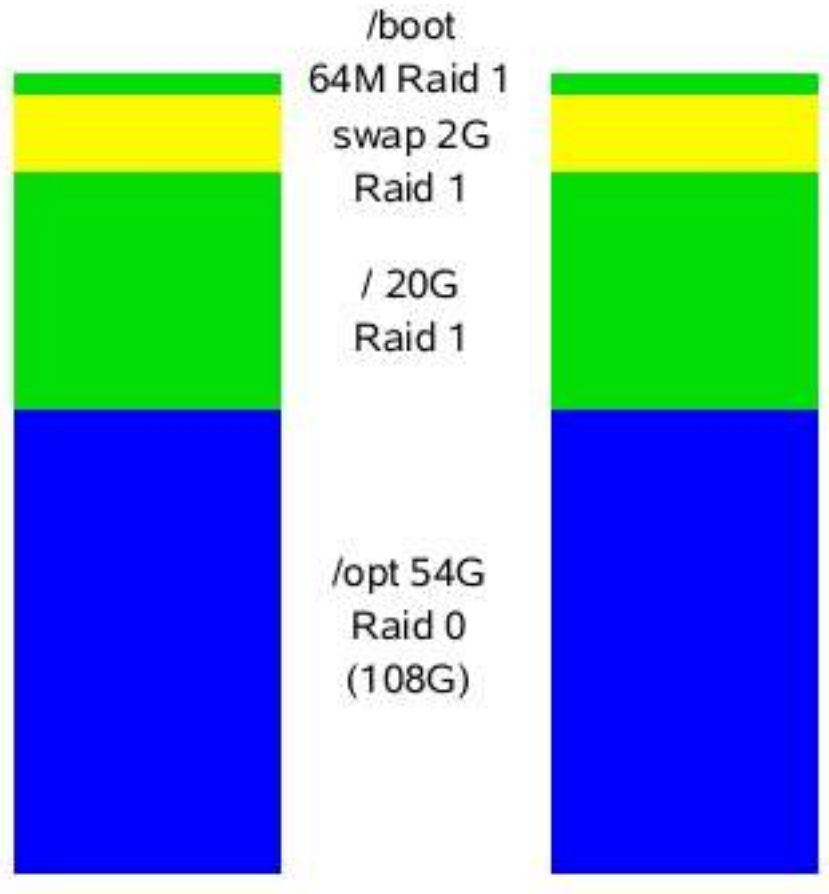
OpenIT
OpenSource Technology

# *Linux Build*

Setup is normal but you might have to resort to some trickery to get the raid setup in place before the install.

Raid 0 is probably a bad idea.

Don't actually create the /opt partition, it will be done later.




OpenIT
OpenSource Technology

# *drbd setup*

/etc/drbd.conf

```
resource drbd0 {
  protocol = C
  fsckcmd = /usr/bin/true

  on castor {
    device  = /dev/nb0
    disk    = /dev/md3
    address = 10.0.0.1
    port    = 7788
  }

  on pollux {
    device  = /dev/nb0
    disk    = /dev/md3
    address = 10.0.0.2
    port    = 7788
  }
}
```

/etc/rc.d/init.d/drbd start

OpenIT
OpenSource Technology

# *Apache, mySQL setup*

With the replication working, you may now format the device, mount it, and install anything that needs to be made redundant on it.

The drbd device is /dev/nb0.

When the system is working, manually test it on both nodes.

OpenIT
OpenSource Technology

# *heartbeat setup*

/etc/ha.d/haresources:

```
castor 209.177.102.55/24/eth0 datadisk::drbd0 mysql apache::/opt/apache/conf/httpd.conf
```

At startup, each of these are called left to right passing “start”.

At shutdown, each are called right to left passing “stop”.

Any service can be started here as long as it has a script that follows this convention.

OpenIT
OpenSource Technology

# *Demo*

Time for a demo…

OpenIT
OpenSource Technology