HA-OSCAR: unleashing HA Beowulf

HA-OSCAR: High Availability Open Source Cluster Application Resources

Team:
- Chokchai (Box) Leangsuksun, Louisiana Tech University
- Stephen L. Scott, Oakridge National Lab, DOE
- Ibrahim Haddad, Open System Lab, Ericsson
- Richard Libby, Intel Corporation
HA-OSCAR: unleashing HA Beowulf

- Goals:
  - COTS-based HPC solution towards non-stop services
  - Linux clustering production quality
  - Ease of build, operation, maintenance
HA-OSCAR: unleashing HA Beowulf

HA-OSCAR Goals

• Aims to drive downtime toward **ZERO** via infrastructure R&D
  – Unplanned downtime
  – Planned downtime
• COT-based HPC Beowulf cluster
  – Open source
  – Production quality
  – Solve HA issues with ease
  – Fault-Tolerant services (including computing nodes)
What is OSCAR?

- Framework for cluster installation configuration and management
- Common used cluster tools
- Wizard based cluster software installation
  - Operating system
  - Cluster environment
    - Administration
    - Operation
  - Automatically configures cluster components
- Increases consistency among cluster builds
- Reduces time to build / install a cluster
- Reduces need for expertise
HA-OSCAR: unleashing HA Beowulf
HA-OSCAR Beowulf
Innovation and information technology

HA-OSCAR: unleashing HA Beowulf
Adaptive recovery state diagram

- **Working**
  - Detect
  - Threshold reached
  - Previous state, # counter, recovery

- **Failure**
  - Switch over & take control at the standby

- **Alert**
  - After # retry
  - Previous state, # counter, recovery

- **Failover**
  - After the primary node repair, then optional Fallback
HA-OSCAR: unleashing HA Beowulf
Network monitoring (sample)
• HA-OSCAR 1.0 Beta release (March 2004)
  – active / hot-standby model for head node
  – The first known field-grade HA Beowulf cluster release
  – Self-configuration Multi-head Beowulf system
  – HA and HPC clustering techniques to enable critical HPC infrastructure
  – Self-healing with 3-5 sec automatic failover time
  – 1-1.5 hour to self-build failover headnodes w/o preloaded OS
  – Optional Image Server for disaster recovery
  – Support existing HPC App(e.g. MPI) without any modification
• How to build HA-OSCAR
  – Can retrofit an existing Linux Beowulf
  – Or start with OSCAR installation tool
  – HA-OSCAR GUI based installation tool
HA-OSCAR: unleashing HA Beowulf

HA-OSCAR installation

- Adopt ease of build and operation same as OSCAR concept
- ~30 min installation
- Initial HA build takes almost the same time as a disaster recovery (that is, each disaster recovery – providing you are prepared!)

**Step 1**

**Step 2 create head image**

**Step 3 clone image**

**Step 4 config**

**Step 5 web admin to add/config more services**
HA-OSCAR: unleashing HA Beowulf

1. HA-OSCAR Installation Wizard

Step 1: HA-OSCAR Packages Installation...

Step 2: Building Image for Standby server [Be patient!]

***In next step, make sure you only deal with OSCAR server nodes!

Step 3: Configuration for Standby server on [eth0]...

!!ATTENTION:
In the next step, please make sure you get the correct MAC address for HAOSCAR Standby Server, not for OSCAR clients!

Step 4: Network Setup & Make boot server on [eth0]...

When the previous steps over, ONLY network boot your Standby Server. Once it is over, reboot Standby Server from the hard drive.

Step 5: Complete installation

Now, switch to Standby server and do logout and login operation.

If you want to uninstall HA-OSCAR package, run command: # /usr/local/ha-oscar-1.0/installer/uninstall.sh

Innovation and information technology
HA-OSCAR: unleashing HA Beowulf

2. Fetch / Clone Server Image
3. Standby Server Initial Network Configuration

- **Image Name:**
  - serverimage
  - oscarimage

- **Primary server's Public Network Interface:** eth1

- **Standby server's Public Network Interface:** eth1

- **Standby server's Public Alias IP:** 138.47.21.199

- **Standby server's local IP:** 10.0.0.200

- **Subnet Mask:** 255.255.255.0

- **Default Gateway:** 10.0.0.1

- **Add Standby Server**
HA-OSCAR: unleashing HA Beowulf

4. Standby Server MAC Address Configuration

![Image of Standby Server MAC Address Collection interface]

- **Stop Collecting MACs**
- **Assign all MACs**
- **Close**
- **Assign MAC to Node**
- **Delete MAC from Node**
- **Configure DHCP Server**
- **Import MACs from file...**
- **Export MACs to file...**

Below are commands to create a boot environment. You can either boot from floppy or network:

- **Build Autoinstall Floppy...**
- **Setup Network Boot**
- **Dynamic DHCP update**

Note: Disable Alias IP of NIC first. MAC Address Collection Tool. When a new MAC address is received on the network, it will appear in the left column. To assign that MAC address to a machine highlight the address and the machine and click 'Assign MAC to Node'.
HA-OSCAR: unleashing HA Beowulf 4+. After MAC Address Collection
## HA-OSCAR: unleashing HA Beowulf

Add / Configure more services

### Edit Watch List

<table>
<thead>
<tr>
<th>Host group to watch</th>
<th>Service name</th>
<th>Monitor to use</th>
<th>Check every</th>
<th>Monitoring periods</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>process server</td>
<td>netsnmp-proc.monitor</td>
<td>3s</td>
<td>wd {Sun-Sat}</td>
</tr>
<tr>
<td></td>
<td>loadaverage server</td>
<td>netsnmp-loadaverage.monitor</td>
<td>20m</td>
<td>wd {Sun-Sat}</td>
</tr>
<tr>
<td></td>
<td>freespace server</td>
<td>netsnmp-freespace.monitor</td>
<td>2h</td>
<td>wd {Sun-Sat}</td>
</tr>
</tbody>
</table>

Save  | Delete

Innovation and information technology
HA-OSCAR: unleashing HA Beowulf

Architecture and Experiment

- 2 Head Nodes
  - dual Xeon 2.4 GHz
  - 1-GB RAM
  - 40 GB Disk
  - 2 NICs
- 4 Compute Nodes:
  - dual Xeon 2.4 GHz
  - 512-MB RAM
  - 40 GB Disk
  - 1 NIC
- 1 Switch 10/100 Mbps
Monitoring overheads

0.9% CPU usage at each monitoring interval

Comparison of network usages for HA-OSCAR different polling sizes
Roadmap

- Grid-aware HA-OSCAR
- Multi-head n+1 active-active
- Hardware abstraction and policy-based recovery management
- Hot-upgrade cluster (OS/CMS)
- Fault-tolerant applications/services and interface framework
- FCAPS Management
- Complete carrier grade
- Policy-based Access Controls (LDAP)
HA-OSCAR: unleashing HA Beowulf

Appeared in a front cover in two major Linux magazines, various technical papers, research exhibitions.
web site: http://xcr.cenit.latech.edu/ha-oscar