



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)



HA-OSCAR: unleashing HA Beowulf

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

Step 8 Done!

Step 1 Start...

Step 2

Step 3

Step 4

Step 5

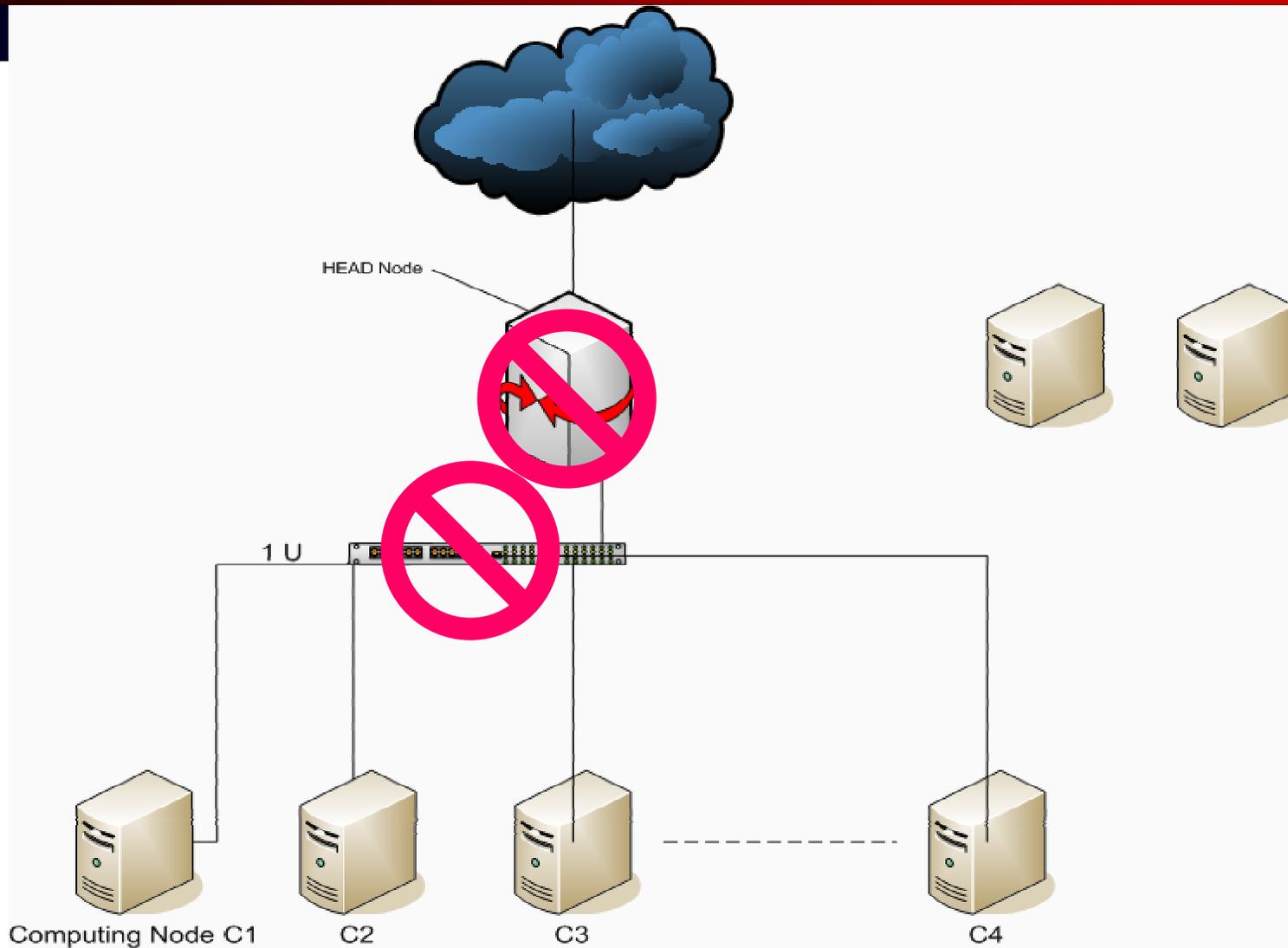
Step 6

Step 7





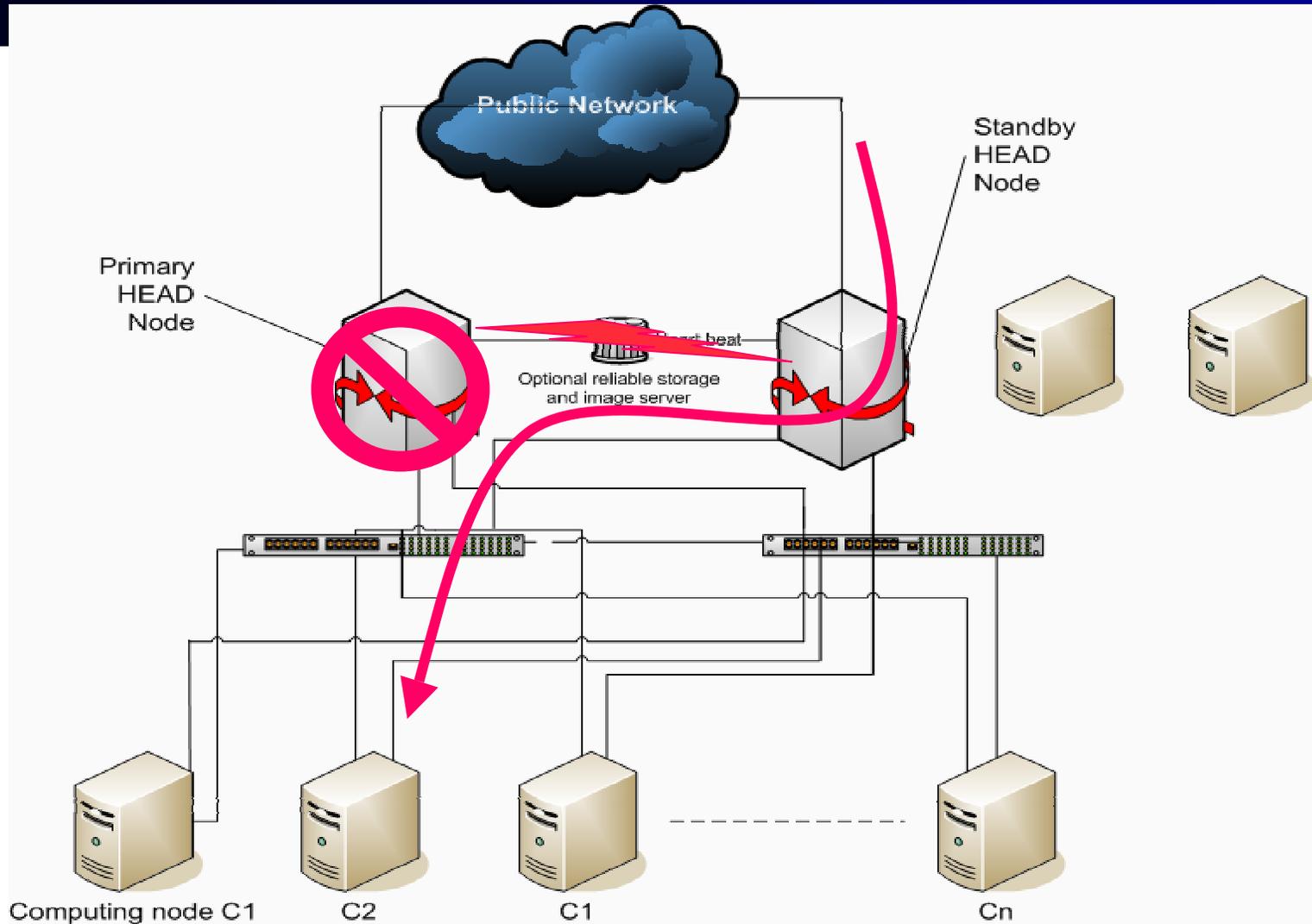
HA-OSCAR: unleashing HA Beowulf Cluster





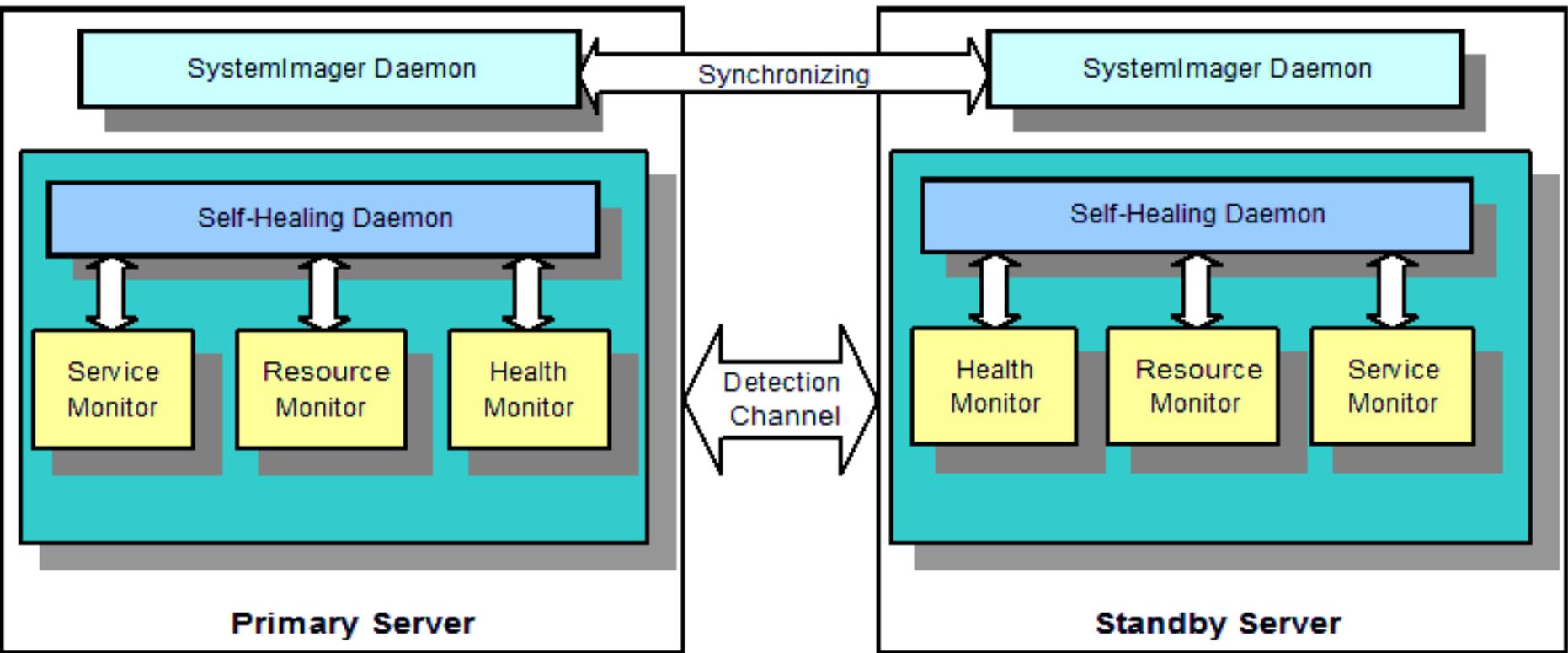
HA-OSCAR: unleashing HA Beowulf

HA-OSCAR Beowulf





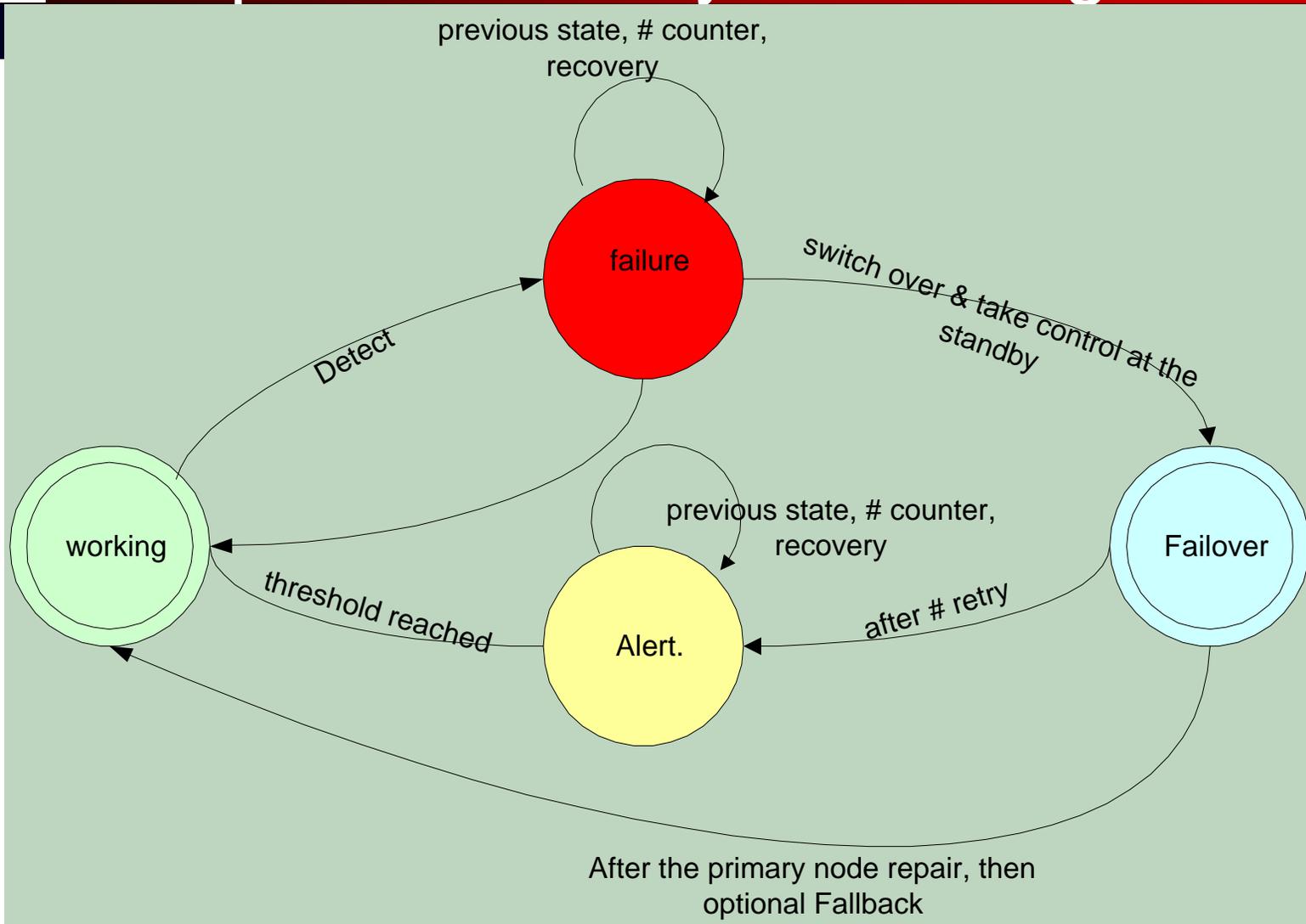
HA-OSCAR: unleashing HA Beowulf Self-healing Schemes





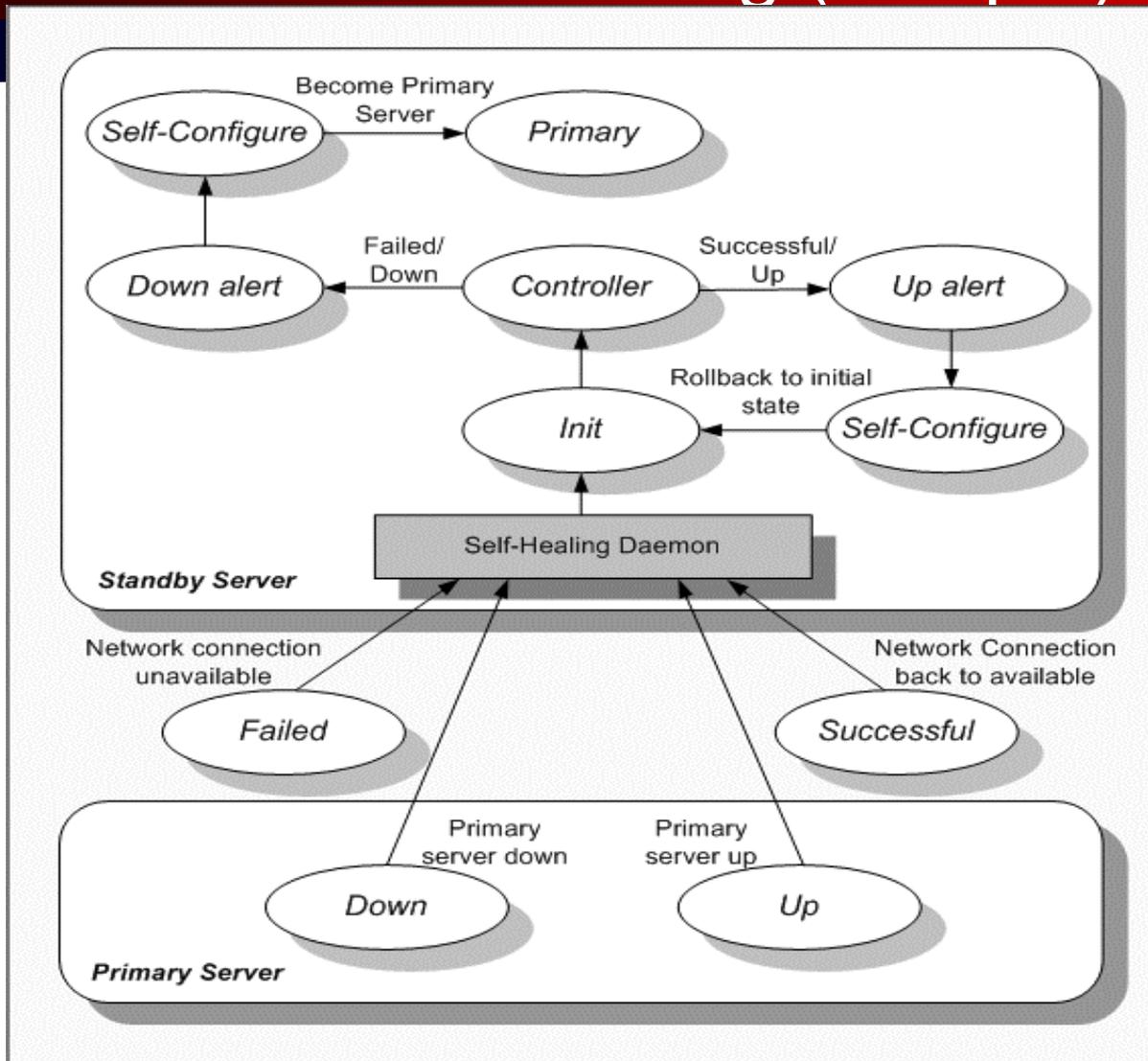
HA-OSCAR: unleashing HA Beowulf

Adaptive recovery state diagram





HA-OSCAR: unleashing HA Beowulf Network monitoring (sample)





HA-OSCAR: unleashing HA Beowulf

- 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



HA-OSCAR: unleashing HA Beowulf

- 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!)

Step2 create head image

step1

step 1: HA-OSCAR Packages Installation...
step 2: Building Image for Standby server [Be patient!]
step 3: Configuration for Standby server on [eth1]...
step 4: Network Setup & Make boot server on [eth1]...

Fetch a System Installation Suite Image

Fill out the following fields to fetch a System Installation Suite image. If you need help on any field, click the help button next to it

Image Name: serverimage Help
Client Name: haoscar Help
SSH User Name: Help
IP Assignment Method: static Help
Post Install Action: beep Help

Reset Fetch Image Close

Step3 clone image

Add Clients to a SIS Image

Image Name: oscarimage Help
Domain Name: oscaromain Help
Base Name: oscarserver Help
Number of Hosts: 1 Help
Starting Number: 1 Help
Padding: 0 Help
Starting IP: 10.0.0.1 Help
Subnet Mask: 255.255.255.0 Help
Default Gateway: 10.0.0.200 Help

Reset Addclients Close

Standby Server MAC Address Collection

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'.

Not Listening to Network. Click 'Collect MAC Addresses' to start.

Remove All

Assign all MACs Close

Node Delete MAC from Node Configure DHCP Server

Export MACs to file...

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

Floppy... Setup Network Boot # Dynamic DHCP update

Edit Watch List

Host group watch details

Host group to watch: ServiceMonitor

Services being watched

Service name	Monitor to use	Check every	Monitoring periods
process_server	netsmp-proc.monitor	3s	wd [Sun-Sat]
loadaverage_server	netsmp-loadaverage.monitor	20m	wd [Sun-Sat]
freepace_server	netsmp-freepace.monitor	2h	wd [Sun-Sat]

Add service... Save Delete

Step5 web admin to add/config more services

Step4 config Standby



HA-OSCAR: unleashing HA Beowulf

1. HA-OSCAR Installation Wizard

HAOSCAR Installation Wizard on eth0

HA-OSCAR Installation Wizard

Before beginning, make sure you have installed OSCAR, and Network Interface works properly, if you use alias IP, please disable them first!

- 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.

Exit

If you want to uninstall HA-OSCAR package, run command: # ./uninstall under ha-oscar-1.0/scripts.



HA-OSCAR: unleashing HA Beowulf

2. Fetch / Clone Server Image

Fetch a System Installation Suite Image

Fill out the following fields to fetch a System Installation Suite image. If you need help on any field, click the help button next to it

Image Name:	<input type="text" value="serverimage"/>	Help
Client Name:	<input type="text" value="hapcoscar"/>	Help
SSH User Name:	<input type="text"/>	Help
IP Assignment Method:	<input type="text" value="static"/>	Help
Post Install Action:	<input type="text" value="beep"/>	Help
<input type="button" value="Reset"/>	<input type="button" value="Fetch Image"/>	<input type="button" value="Close"/>



HA-OSCAR: unleashing HA Beowulf

3. Standby Server Initial Network Configuration

▼ Add Standby server to a SIS Image

Image Name:	<input type="text" value="serverimage"/> <input type="text" value="oscarimage"/>	<input type="button" value="Help"/>
Primary server's Public Network Interface:	<input type="text" value="eth1"/>	<input type="button" value="Help"/>
Standby server's Public Network Interface:	<input type="text" value="eth1"/>	<input type="button" value="Help"/>
Standby server's Public Alias IP:	<input type="text" value="138.47.21.199"/>	<input type="button" value="Help"/>
Standby server's local IP:	<input type="text" value="10.0.0.200"/>	<input type="button" value="Help"/>
Subnet Mask:	<input type="text" value="255.255.255.0"/>	<input type="button" value="Help"/>
Default Gateway:	<input type="text" value="10.0.0.1"/>	<input type="button" value="Help"/>
<input type="button" value="AddStandby Server"/>		<input type="button" value="Close"/>



HA-OSCAR: unleashing HA Beowulf

4. Standby Server MAC Address Configuration

Standby Server MAC Address Collection

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'.
Currently Scanning Network... Click 'Stop Collecting' to stop.

All Clients

- oscardmode1.oscardomain
 - eth0 mac = 00:e0:81:25:1c:7a
 - eth0 ip = 10.0.0.2
- oscarserver1.oscardomain
 - eth0 mac = 00:e0:81:25:1c:dc
 - eth0 ip = 10.0.0.3

Remove Remove All

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...

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

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

B A



HA-OSCAR: unleashing HA Beowulf 4+. After MAC Address Collection

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'.
Currently Scanning Network... Click 'Stop Collecting' to stop.

00:e0:81:25:1c:dc

All Clients

- oscamode1.oscardomain
 - eth0 mac = 00:e0:81:25:1c:7a
 - eth0 ip = 10.0.0.2
- oscarserver1.oscardomain
 - eth0 mac =
 - eth0 ip = 10.0.0.3

Remove Remove All

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

E F



HA-OSCAR: unleashing HA Beowulf

Add / Configure more services

Edit Watch List

Host group watch details

Host group to watch

ServiceMonitor ▾

Services being watched

Service name	Monitor to use	Check every	Monitoring periods
process server	netsnmp-proc.monitor	3s	wd { Sun-Sat }
loadaverage server	netsnmp-loadaverage.monitor	20m	wd { Sun-Sat }
freespace server	netsnmp-freespace.monitor	2h	wd { Sun-Sat }

[Add service..](#)

Save

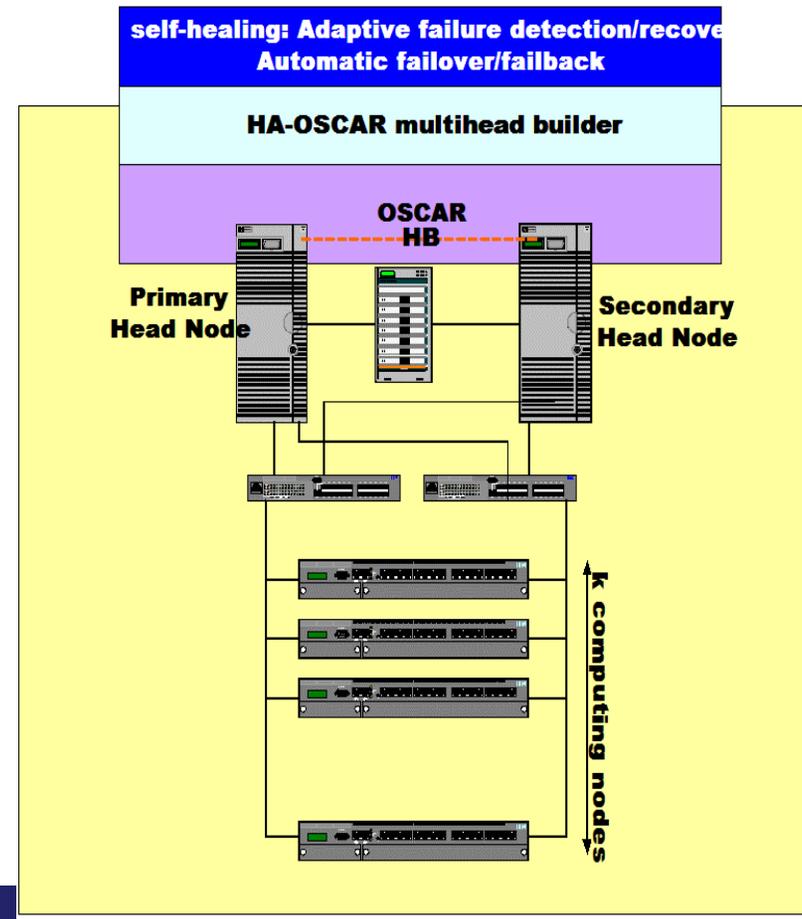
Delete



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

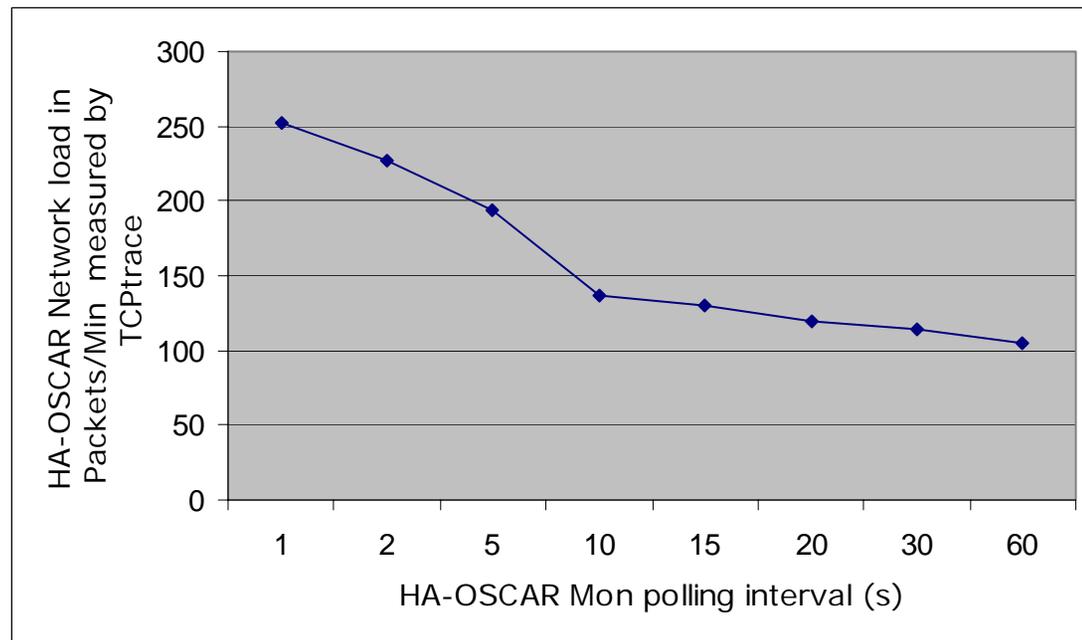




HA-OSCAR: unleashing HA Beowulf

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>

