



Linux Clusters Institute: Node Health Check (NHC)

Zhongtao Zhang, System Administrator, Holland Computing Center,
University of Nebraska-Lincoln

What is NHC?

- NHC is a tool for determining the health status of a node.
 - Node misconfiguration
 - Failure situations
 - Hardware failures
- NHC can mark “unhealthy” nodes offline so jobs don’t fail
- NHC helps to increase the reliability and throughput of jobs run on a cluster
- NHC has one function, to verify that a node is healthy to run a new job!
 - We don’t want any dazed and confused nodes that appear to still be working

Why NHC?

- NHC provides a framework for node monitoring
 - Get away from home grown scripts, which are not always portable or reliable.
 - Administration issues
- There are a large number of built-in checks. (check the website for them: <https://github.com/mej/nhc#built-in-checks>)
- Lawrence Berkeley National Laboratory (LBNL) Design Goals
 - Reliable
 - Flexible
 - Extensible
 - Should be fast
 - Code should be reusable and easy to port

Using NHC?

- Using NHC
 - From the command line
 - As an addition to the job scheduler, i.e. Slurm, PbsPro, LSF
 - As a cron job on a node
- Checks are based on node name
 - Matching checks are run
 - If a check fails, NHC exits
 - Prints a message with information about which check failed and why
 - If run from a job scheduler, can mark the node offline
 - Can log failure(s) to syslog

Installing NHC on a node

- You can download RPM's from the hosting Github site
 - NHC is installed into the OS standard paths
 - /usr/sbin/nhc (the NHC command)
 - /etc/nhc
 - /usr/libexec/nhc
 - Default configuration files will be installed in /etc/nhc
- You can also download the source and build the code locally
 - ./configure --prefix=/usr --sysconfdir=/etc --libexecdir=/usr/libexec
 - make test
 - make install

Testing NHC

- If you install from source, you can run a verification test suite
 - Run “make test”
 - Once you make changes to the config files it’s best to run the health checks
 - Run “/usr/sbin/nhc”
 - When you are satisfied with the results you can add it to your job scheduler

Configuring NHC

- NHC uses a configuration file “/etc/nhc/nhc.conf”
- Simple configuration
 - Default configuration file has multiple sections, broken down by
 - Configuration variables
 - Hardware checks
 - filesystems checks
 - Process checks
 - Scheduler checks
 - Other check, i.e. none of the above areas
- In the default configuration file most of the checks are commented out
- NHC uses scripts to handle failure check functions found in “/etc/nhc/scripts”
- Users can add their own checks to NHC
- User can auto generate the config file using “/etc/nhc-genconf” command

Configuration Example

```
#####  
###  
### Filesystem checks  
###  
# All nodes should have their root filesystem mounted read/write.  
# * || check_fs_mount_rw -f /  
  
# Assert that /tmp is a mounted filesystem of type "tmpfs."  
# * || check_fs_mount_rw -t tmpfs -f /tmp  
  
# Controlling TTYS are a good thing!  
# * || check_fs_mount_rw -t devpts -s '(none|devpts)!' -f /dev/pts  
  
# Make sure the root filesystem doesn't get too full.  
# * || check_fs_free / 3%  
  
# Free inodes are also important.  
# * || check_fs_ifree / 1k  
  
# The following illustrates how to assert an NFSv3 mount (or any other specific mount option).  
# * || check_fs_mount -s bluearc0:/home -t nfs -o '(^|,|vers=3(|$))!' -f /home
```


Write your own check

```
function hcc_check_lustre() {
    if (grep lustre /proc/mounts &> /dev/null); then
        continue
    else
        die 1 "$FUNCNAME: Lustre is not mounted"
        return 1
    fi

    if [[ $(cat /proc/fs/lustre/health_check) != "healthy" ]]; then
        die 1 "$FUNCNAME: Lustre not healthy"
        return 1
    fi

    return 0
}
```

Intregation with Job schedulers

- Torque
 - Add the following lines to your pbs_mom config files

```
$node_check_script /usr/sbin/nhc
$node_check_interval 5,jobstart,jobend
$down_on_error 1
```

This will run NHC every five minutes and at job start and job end, marking the node offline if NHC fails a check

- You will need to enable “operator” access on each node

```
qmgr -c “set server operators += root@*”
```
- NHC will add a note indicating the failure. Once the failure has been corrected the note will be remove.

Intregation with Job schedulers

- Slurm
 - Add the following lines to your `/etc/slurm/slurm.conf` config files
 - HealthCheckProgram=/usr/sbin/nhc
 - HealthCheckInterval=300
 - HealthCheckNodeState=ANY

This will run NHC every five minutes and at job start and job end, marking the node drain if NHC fails a check
- Other scheduler
 - Cron
 - nhc-wrapper
 - # /usr/sbin/nhc-wrapper -M root -X 12h

References

- NHC software site
 - <https://github.com/mej/nhc>
- RPM's location
 - <https://github.com/mej/nhc/releases/>