CRAY

BoF: eLogin Usability and Best Practices
Jeff Keopp, Cray Inc.

CUG.2017.CAFFEINATED COMPUTING

Redmond, Washington May 7-11, 2017

eLogin Usability and Best Practices

- **Goals of the BoF**
  - Provide direct conversation between customers and feature developers
  - Identify areas of improvement within the product
  - Provide a short FAQ
  - Don't forget the survey at:
    - https://www.surveymonkey.com/r/KMG657S
- **Usability**
  - Image Management
    - SMW and CMC
  - Image Configuration
    - SMW and CMC
  - CMC installation and configuration
  - eLogin administration

eLogin Usability and Best Practices

- **Best Practices**
  - Image Management
    - Clone recipes before making changes
    - Delete unused image_roots on the SMW
      - Images can always be rebuilt from the recipe
    - Delete unused images from glance on the CMC
  - Image Configuration
    - Clone config set before making changes
    - Delete unused config sets from the CMC (/var/opt/cray/imps/config/sets/ and swift)
  - CMC installation and configuration
    - Make all configuration changes in the site overrides file
    - Determine network configuration of the CMC up front
      - Changing later is not trivial
    - Logs are numerous, very verbose and therefore grow large quickly
      - Use logrotate and archive logs frequently
  - eLogin administration
    - Deploy and shutdown eLogin nodes using heat stack commands

COMPUTE | STORE | ANALYZE

eLogin Usability and Best Practices

- **Troubleshooting and Debugging eLogin node issues**
  - Remote console
    - Attach to the console with "`ironic_conman <hostname>`"
    - Console logs are under `/var/log/conman/ironic_<UUID>.log`
      - <UUID> is found by running "`ironic node-list`" and locating the hostname in the resulting listing of nodes
  - Systemd logging
    - Use "`journalctl –a`" to see all kernel messages and other available information
    - User "`journalctl –f`" to tail the logs like "tail –f" did

eLogin Usability and Best Practices

- ## Troubleshooting and Debugging eLogin node issues
  - /var/log
    - System log messages are located in `/var/log/messages`
      - eLogin forwards these messages to the CMC via rsyslog
    - Each OpenStack service logs messages under `/var/log/<service>`
      - There may be multiple logs per service
        - For example, `/var/log/ironic` has `ironic-api.log` and `ironic-conductor.log` files – most ironic error issues would be in the `ironic-conductor.log`
    - cray_dumpsys command
      - Used for debugging the CMC.
      - To include information from eLogin nodes, enable and configure the eLogin plugin

eLogin Usability and Best Practices

- **Troubleshooting and Debugging eLogin node issues**
  - Kdump and crash
    - Installed and enabled by default
  - OpenStack Service Diagnostics
    - Heat (orchestration and deployment of eLogin)
      ```
      # heat stack-show <stack_name>
      ```
    - Nova (Running node life cycle)
      ```
      # nova show <host_name>
      ```
    - Ironic (physical server)
      ```
      # ironic node-show <node_name>
      ```

COMPUTE | STORE | ANALYZE

eLogin FAQ

- **Intermittent boot/deploy failures. Sometimes the node comes up, but not always. What might cause this?**
  - You might have a rogue DHCP server. Have a look in `/var/log/messages` and see where the DHCP requests are coming and going.
  - If you don't see a message in the log that shows the CMC handing the elogin node an IP address that corresponds to the MAC address of the elogin node, that could be evidence of another DHCP server on your network answering the request when the elogin node boots.

eLogin FAQ

- **What happens if I mistype the hostname in the inventory or env file?**
  - In the nova-scheduler log, you'll see the following error message

    "`No hosts matched due to not matching`"

  - Fix `/etc/opt/cray/ansible/inventory.csv` or the `/etc/opt/cray/opentack/heat/templates/<env>` file to have the correctly matching hostname
  - Delete offending ironic node
  - Enroll the corrected node
  - Redeploy the heat stack

eLogin FAQ

- **"heat stack-show" of a failed deploy has the error "Message: Unknown, Code: Unknown"**
  - This often means that the RAM size specified in the inventory file doesn't match the size given in the nova flavor. (eloginflavor)
    - If anything, the ironic node RAM size should be larger than the number specified in the flavor.
  - Don't change the nova flavor; instead, adjust the RAM size in the inventory file if necessary, deleted and re-enroll the ironic node.

```
# cd /var/log
# grep DHCP messages

# cd /etc/opt/cray/openstack/ansible
# heat stack-delete <your-stack-name>
# ironic node-delete <your-node-name>
# vi inventory.csv

# ./csms_ironic_enrollment.sh
```

COMPUTE | STORE | ANALYZE

eLogin FAQ

- **ironic power state is wrong, preventing ironic enrollment**
  - If ironic enrollment fails with some errors about power states, have a look at the ironic.conductor log, and see if it contains something like:
    - `ERROR ironic.conductor.manager [-] During sync_power_state, max retries exceeded for node <IRONIC_NODE_UUID>, node state None does not match expected state 'None'. Updating DB state to 'None' Switching node to maintenance mode.`
  - Try resetting the BMC via ipmitool:

  `# ipmitool -I lanplus -H <BMC_IP_ADDR> -L ADMINISTRATOR -U root mc reset warm`

  - If that's not valid, try "cold" instead of "warm". Once that has completed, try checking the power status:

  `# ipmitool -I lanplus -H <BMC_IP_ADDR> -L ADMINISTRATOR -U root chassis power status`

COMPUTE | STORE | ANALYZE

eLogin FAQ

- **iptables error on `csms_install.sh` because of 3 networks vs. 2 networks problem (CSMS 1.1.2)**
  - If you encounter a situation in which the installer displays a screen full of iptables rules, then stops, it is probably caused by a known issue in which the number of networks configured differs from the number of networks actually present.
  - The fix for this is to change the number of interfaces the `firewall.yaml` play looks for in `/etc/opt/cray/openstack/ansible/firewall.yaml`. Change `(networks|length > 2)` to `(networks|length >3)`
  - This occurs in two places, one around line 210, and another around line 222:
  - After editing `firewall.yaml`, rerun the firewall play like so:

    ```
    # cd /etc/opt/cray/openstack/ansible
    # ansible-playbook -I hosts -ask-vault-pass firewall.yaml
    ```

- **Rerun the installer, and it should complete without issue.**

COMPUTE | STORE | ANALYZE

eLogin FAQ

- **FATAL: Could not find specified device or file /root/isos/Cray- CentOSbase**
  - You might have forgotten to symlink the centos iso to the cray-centos iso.

- **Some elogin nodes out of a group fail to deploy because of metadata or fuel timeout**
  - Symptoms of this problem include neutron logs containing errors like "`respawning metadata-proxy`", and one or more nodes fail to fully deploy because they time out waiting for metadata.
  - The workaround is to increase `deploy_timeout` in `/etc/ironic/ironic.conf` from the default of 15 to a higher value, such as 30, then restart ironic services and reattempt the deploy.

eLogin FAQ

- **CMC /var partition filling up**
  - There are a couple recommendations to help this situation
    - remove some kafka logs
    - add physical disks to increase the amount of available space.
  - If deleting some logs to clear space in the short term, that can be accomplished by changing the retention settings and restarting the kafka service, which will automatically age-out and clean the logs automatically
    - Kafka log retention setup is handled in `/etc/kafka/server.properties`, don't edit that file directly. Use the site overrides file to make changes

    ```
    # vi /etc/opt/cray/openstack/ansible/config/site/elogin-site-overrides.yaml
    ```

COMPUTE | STORE | ANALYZE

eLogin FAQ

- **CMC /var partition filling up**

    - Add values at the end of the overrides file
        - Each site will have their own local considerations for size and frequency, based on needs and policy. The example below will cause a fairly dramatic change in log storage.
        - These settings will keep the logs for three days, roll them over when logs reach 512MB, and check if the logs need rolling every five minutes.

    ```
    # Kafka log retention settings
    log_retention_hours: 72
    log_retention_bytes: 536870912
    log_retention_check_interval_ms: 300000
    ```

    - Rerun the kafka ansible play and restart the kafka service

    ```
    # cd /etc/opt/cray/openstack/ansible
    # python ./csms_common.py –ansible-playbook "kafka.yaml" "$@"
    # systemctl restart kafka
    ```

COMPUTE | STORE | ANALYZE

eLogin FAQ

- **How to set up ipmi in the iDRAC (provides remote console and power control)**

  Note: This needs to be checked on any system that had its motherboard replaced. This setting might also be unset during a firmware update.

  - Connect to the iDRAC UI via the Web interface
  - Go into the iDRAC Settings section
  - Click the Services tab
  - Under Remote RACADM, uncheck the checkbox to disable RACADM
  - Uncheck the checkbox at the top of the page under "Local Configuration" that says "Disable iDRAC Local Configuration using RACADM"
  - Click Apply button at bottom of window to save settings
  - Navigate back to main Network page for the iDRAC
  - Click on IPMI Settings tab and check the checkbox for "Enable IPMI Over LAN"
  - Click Apply button to save settings

COMPUTE | STORE | ANALYZE

Legal Disclaimer

Information in this document is provided in connection with Cray Inc. products. No license, express or implied, to any intellectual property rights is granted by this document.

Cray Inc. may make changes to specifications and product descriptions at any time, without notice.

All products, dates and figures specified are preliminary based on current expectations, and are subject to change without notice.

Cray hardware and software products may contain design defects or errors known as errata, which may cause the product to deviate from published specifications. Current characterized errata are available on request.

Cray uses codenames internally to identify products that are in development and not yet publically announced for release. Customers and other third parties are not authorized by Cray Inc. to use codenames in advertising, promotion or marketing and any use of Cray Inc. internal codenames is at the sole risk of the user.

Performance tests and ratings are measured using specific systems and/or components and reflect the approximate performance of Cray Inc. products as measured by those tests. Any difference in system hardware or software design or configuration may affect actual performance.

The following are trademarks of Cray Inc. and are registered in the United States and other countries: CRAY and design, SONEXION, and URIKA. The following are trademarks of Cray Inc.: APPRENTICE2, CHAPEL, CLUSTER CONNECT, CRAYPAT, CRAYPORT, ECOPHLEX, LIBSCI, NODEKARE, REVEAL, THREADSTORM. The following system family marks, and associated model number marks, are trademarks of Cray Inc.: CS, CX, XC, XE, XK, XMT, and XT. The registered trademark LINUX is used pursuant to a sublicense from LMI, the exclusive licensee of Linus Torvalds, owner of the mark on a worldwide basis. Other trademarks used in this document are the property of their respective owners.

COMPUTE | STORE | ANALYZE