

Gretina Private Network - July 2015

Prepared by: LBLnet Services Group, Lawrence Berkeley National Laboratory
Contact: lblnet@lbl.gov

Table of Contents

- [1.0 - Initial Installation](#)
- [2.0 - Basic Configuration](#)
 - [2.1 - Management IP Address](#)
 - [2.1.1 - In-Band Virtual Interface](#)
 - [2.1.2 - Out-of-Band Physical Interface](#)
 - [2.2 - Changing Gretina Password](#)
 - [2.3 - Remote Syslog](#)
 - [2.4 - SNMP](#)
 - [2.5 - NTP](#)
 - [2.6 - Link Aggregation](#)
 - [2.6.1 - Static](#)
 - [2.6.2 - LACP](#)
- [3.0 - Basic Troubleshooting](#)
 - [3.1 - Management Interface Access](#)
 - [3.2 - Show Configuration](#)
 - [3.3 - Check Virtual-Chassis \(Stack\) Status](#)
 - [3.4 - Check VC-Port Status](#)
 - [3.5 - Check Light Levels](#)
 - [3.6 - Check Alarms](#)
 - [3.7 - Show Interface Status](#)
 - [3.8 - Show Interface Counters](#)
 - [3.9 - Show MAC address table](#)
 - [3.10 - Check Logging](#)
 - [3.11 - Save RSI and Logs to send to LBLnet / Juniper TAC](#)
 - [3.12 - Check LACP status](#)
 - [3.13 - Graceful Shutdown](#)
 - [3.13.1 - Shutdown via LED Front Panel](#)
 - [3.13.2 - Shutdown via SSH/Console](#)
- [4.0 - Replacing a failed unit](#)
- [5.0 - EMERGENCY: Minimally Operational Configuration](#)
- [6.0 - Reference](#)
 - [6.1 - Naming Conventions](#)
 - [6.2 - Management Login](#)
 - [6.3 - SNMP](#)
 - [6.4 - Syslog](#)
 - [6.5 - Reference Diagrams](#)

1.0 - Initial Installation

1. Physically install the switches in the rack, do not power them up yet.
 2. Review the diagram labeled "INITIAL INSTALLATION"
 3. Connect the 2x40G QSFP+ cables in the rear of the Switches 0 and 1
 4. Connect the 2x10G SFP+ SR links from Switches 0 and 1 to Switch 2
 - a. Switch 0 xe-0/2/3 to Switch 2 xe-2/2/2
 - b. Switch 1 xe-1/2/3 to Switch 2 xe-2/2/3
 5. These switches contain dual power supplies, so whenever possible, it is recommended that each power supply be fed from a different power source (i.e. one on Building/Utility power, the other on Generator-backed or UPS-backed power).
 6. Power up the switches. While not completely necessary, it is a good practice to boot them in the following order:
 - a. Switch 0
 - b. Switch 1
 - c. Switch 2
 7. The switches will take several minutes to boot. When they are fully online, you should see the following in the LED panel of each:
 - a. Switch 0: RE
 - b. Switch 1: BK
 - c. Switch 2: LC
 8. Either console into Switch 0 or SSH to the management address (10.0.0.6, subnet mask 255.255.0.0).
 9. Login with credentials supplied under [6.2 - Management Login](#)
 10. Confirm there are no alarms using the instructions in [3.6 - Check Alarms](#)
 11. Confirm Light Levels using the instructions in [3.5 - Check Light Levels](#)
 12. Confirm the Virtual-Chassis (Stack) status using the instructions found under section [3.3 - Check Virtual-Chassis \(Stack\) Status](#)
 13. Connect Switch 0 to the NetApp. Confirm pingability from a server or laptop on the Switch 0 to the NetApp. Test NFS, CIFS or the applicable storage protocol.
 14. Connect the remaining devices (Master nodes, Storage nodes, compute nodes, DAQ, sensors, etc).
-

2.0 - Basic Configuration

These switches ship pre-configured and should not require any additional configuration during installation. The below is optional.

2.1 - Management IP Address

You can configure either an in-band virtual interface for management or an out-of-band physical interface (located on the rear of the switch and labeled MGMT) for management. The example IP address used here is invalid, replace it with the actual IP address you wish to use. **Note: the units ship with a single virtual (in-band) management interface: 10.0.0.6, which can be SSH'd to from a device on any port.**

2.1.1 - In-Band Virtual Interface

1. Console into the Master Routing Engine
2. Login with the supplied credentials
3. Perform the following commands:

```
edit
```

```
set interfaces irb.1511 family inet address 256.0.0.6/16
```

```
set interfaces irb.1511 family inet filter input Mgmt-Firewall-In
```

```
set vlan vlan1511 13-interface irb.1511
```

4. Compare your changes

```
show | compare
```

5. Check the commit will be successful

```
commit check
```

6. Commit changes

```
commit
```

2.1.2 - Out-of-Band Physical Interface

1. Console into the Master Routing Engine
2. Login with the supplied credentials
3. Perform the following commands:

```
edit
```

```
set interfaces me0.0 family inet address 256.0.0.6/16
```

```
set interfaces me0.0 family inet filter input Mgmt-Firewall-In
```

4. Compare your changes

```
show | compare
```

5. Check the commit will be successful

```
commit check
```

6. Commit changes

```
commit
```

2.2 - Changing Gretina Password

1. Console into the Master Routing Engine or SSH to the Management IP

2. Login with the supplied credentials
3. Perform the following commands:

```
edit
```

```
set system login user gretina authentication plain-text-password
```

4. You will then be prompted to enter a New Password and confirm it.
5. Compare your changes

```
show | compare
```

6. Check the commit will be successful

```
commit check
```

7. Commit changes

```
commit
```

2.3 - Remote Syslog

1. Console into the Master Routing Engine or SSH to the Management IP
2. Login with the supplied credentials
3. Perform the following commands:

```
edit
```

```
set system syslog host 256.0.0.10 any error
```

4. You may wish to customize your syslog configuration further.
5. Compare your changes

```
show | compare
```

6. Check the commit will be successful

```
commit check
```

7. Commit changes

```
commit
```

2.4 - SNMP

1. Console into the Master Routing Engine or SSH to the Management IP
2. Login with the supplied credentials
3. Perform the following commands:

```
edit
```

```
set system syslog host 256.0.0.10 any error
```

4. You may wish to customize your syslog configuration further.
5. Compare your changes

```
show | compare
```

6. Check the commit will be successful

```
commit check
```

7. Commit changes

```
commit
```

2.5 - NTP

1. Console into the Master Routing Engine or SSH to the Management IP
2. Login with the supplied credentials
3. Perform the following commands:

```
edit
```

```
delete system ntp
```

```
set system ntp server 256.0.0.10
```

```
set system ntp boot-server 256.0.0.10
```

4. You can have multiple servers if preferred.
5. Compare your changes

```
show | compare
```

6. Check the commit will be successful

```
commit check
```

7. Commit changes

```
commit
```

2.6 - Link Aggregation

2.6.1 - Static

1. Console into the Master Routing Engine or SSH to the Management IP
2. Login with the supplied credentials
3. Perform the following commands to show the current number of aggregated ethernet interfaces:

```
edit
```

```
show chassis aggregated-devices
```

4. Now increase the device-count for the number of link aggregation groups you are adding. For example, to increase from 1 to 2:

```
set chassis aggregated-devices ethernet device-count 2
```

5. Now list the currently configured Aggregated Ethernet (ae) links:

```
show interface ae?
```

6. Configure a new Aggregated Ethernet link using a number not already in use. For example, if there is already an ae0, you would configure ae1. You will need to modify some of these values according to your application (for example, 10G links instead of 1G)

```
set interface ae1 aggregated-ether-options link-speed 1g
```

```
edit interface ae1.0 family ethernet-switching
```

```
set interface-mode access
```

```
set vlan members vlan1511
```

```
set filter input ACCESS-filter
```

```
set storm-control ACCESS
```

```
set recovery-timeout 30
```

```
exit
```

7. Now place the physical interface(s) you wish to use in the new Aggregated Ethernet link. The examples here are invalid, you will need to replace ge-0/0/98 and ge-0/0/99 with real values.

```
set interface ge-0/0/98 ether-options 802.3ad ae1
```

```
set interface ge-0/0/99 ether-options 802.3ad ae1
```

8. Remove the above physical interfaces from any interface-range EXCEPT "COPPER_PORTS". You will need to look at the member-range to see if they fall within it. For example, if you are aggregating ge-0/0/24 and ge-1/0/24, they fall within "member-range ge-0/0/0 to ge-0/0/47" AND "member-range ge-1/0/0 to ge-1/0/47". You must remove them from both. Below is an example to do so:

```
edit interface interface-range ACCESS_PORTS
```

```
delete member-range ge-0/0/0 to ge-0/0/47
```

```
delete member-range ge-1/0/0 to ge-1/0/47
```

```
set member-range ge-0/0/0 to ge-0/0/23
```

```
set member-range ge-0/0/25 to ge-0/0/47
```

```
set member-range ge-1/0/0 to ge-1/0/23
```

```
set member-range ge-1/0/25 to ge-1/0/47
```

```
exit
```

9. Compare your changes

```
show | compare
```

10. Check the commit will be successful

```
commit check
```

Note: If you receive an error that 'logical unit is not allowed on aggregated links' check that the physical links you are placing into an aggregated-ethernet link are NOT members of ACCESS_PORTS, or WIRED_ACCESS_PORTS, which would cause them to inherit 'unit 0' logical unit configuration.

11. Commit changes

```
commit
```

2.6.2 - LACP

1. Carry out all the same steps above under "2.7.1 Static"
2. Perform the additional configuration:

```
set interface ae1 aggregated-ether-options lacp active
```

```
set interface ae1 aggregated-ether-options lacp periodic slow
```

3. Compare your changes

```
show | compare
```

4. Check the commit will be successful

```
commit check
```

Note: If you receive an error that 'logical unit is not allowed on aggregated links' check that the physical links you are placing into an aggregated-ethernet link are NOT members of ACCESS_PORTS, or WIRED_ACCESS_PORTS, which would cause them to inherit 'unit 0' logical unit configuration.

5. Commit changes

commit

3.0 - Basic Troubleshooting

3.1 - Management Interface Access

- a. There is a management interface configured with IP address 10.0.0.6 on the same private network for SNMP, syslog, and remote access. This should be able to be SSH'd to from the MSU/NSCL SSH gateway.
- b. If the management interface is not reachable, you can console into the back of the switches. In general, you'll want to do so on the Master Routing Engine (identifiable by an RE in the LED panel and a solidly lit 3rd LED next to the LED panel. A blinking 3rd LED and the letters BK indicate the Backup Routing Engine which shouldn't be consoled into).

3.2 - Show Configuration

```
show configuration
```

It may be beneficial to display the configuration with inheritance:

```
show configuration | display inheritance no-comments
```

3.3 - Check Virtual-Chassis (Stack) Status

```
show virtual-chassis
```

Expected Output:

Member ID	Status	Serial No	Model	Mstr prio	Role	Mixed Mode	Route Mode	Neighbor ID	List Interface
0 (FPC 0)	Prsnt	PE3714490317	ex4300-48t	129	Master*	N	VC	1	vcp-255/1/0
								1	vcp-255/1/1
								2	vcp-255/2/3
1 (FPC 1)	Prsnt	PE3714490136	ex4300-48t	129	Backup	N	VC	0	vcp-255/1/0
								0	vcp-255/1/1
								2	vcp-255/2/3
2 (FPC 2)	Prsnt	PE3714080276	ex4300-48t	0	Linecard	N	VC	1	vcp-255/2/2
								0	vcp-255/2/3

3.4 - Check VC-Port Status

```
show virtual-chassis vc-port
```

Expected Output:

```
gretina@j43s-msu-gretinaprivate> show virtual-chassis vc-port
```


fpc0:

```
-----
```

Interface or PIC / Port	Type	Trunk ID	Status	Speed (mbps)	Neighbor ID Interface
1/0	Configured	5	Up	40000	1 vcp-255/1/0
1/1	Configured	5	Up	40000	1 vcp-255/1/1
2/3	Configured	-1	Up	10000	2 vcp-255/2/3
1/2	Configured		Absent		
1/3	Configured		Absent		

fpc1:

```
-----
```

Interface or PIC / Port	Type	Trunk ID	Status	Speed (mbps)	Neighbor ID Interface
1/0	Configured	5	Up	40000	0 vcp-255/1/0
1/1	Configured	5	Up	40000	0 vcp-255/1/1
2/3	Configured	-1	Up	10000	2 vcp-255/2/2
1/2	Configured		Absent		
1/3	Configured		Absent		

fpc2:

```
-----
```

Interface or PIC / Port	Type	Trunk ID	Status	Speed (mbps)	Neighbor ID Interface
2/2	Configured	-1	Up	10000	1 vcp-255/2/3
2/3	Configured	-1	Up	10000	0 vcp-255/2/3
1/3	Configured		Absent		
1/2	Configured		Absent		
1/1	Configured		Absent		
1/0	Configured		Absent		

3.5 - Check Light Levels

Note: each time you run these commands, you must "request" before "show".

```
request virtual-chassis vc-port diagnostics optics
```

```
show virtual-chassis vc-port diagnostics optics | grep "optical power"
```

Expected Output:

```
Receiver signal average optical power      : 0.0064 mW / -21.92 dBm
Rx loss of average optical power alarm     : Off
Receiver signal average optical power      : 0.0056 mW / -22.49 dBm
Rx loss of average optical power alarm     : Off
Receiver signal average optical power      : 0.0060 mW / -22.22 dBm
Rx loss of average optical power alarm     : Off
Receiver signal average optical power      : 0.0058 mW / -22.40 dBm
```

Rx loss of average optical power alarm : Off

NOTE: These values will vary, but should not be less than the thresholds you see when running the below command (note these are negative values):

```
request virtual-chassis vc-port diagnostics optics
show virtual-chassis vc-port diagnostics optics | grep "rx power low"
```

```
Laser rx power low alarm           : Off
Laser rx power low warning         : Off
Laser rx power low alarm threshold : 0.0001 mW / -40.00 dBm
Laser rx power low warning threshold : 0.0002 mW / -38.01 dBm
Laser rx power low alarm           : Off
Laser rx power low warning         : Off
Laser rx power low alarm threshold : 0.0001 mW / -40.00 dBm
Laser rx power low warning threshold : 0.0002 mW / -38.01 dBm
Laser rx power low alarm           : Off
Laser rx power low warning         : Off
Laser rx power low alarm threshold : 0.0001 mW / -40.00 dBm
Laser rx power low warning threshold : 0.0002 mW / -38.01 dBm
Laser rx power low alarm           : Off
Laser rx power low warning         : Off
Laser rx power low alarm threshold : 0.0001 mW / -40.00 dBm
Laser rx power low warning threshold : 0.0002 mW / -38.01 dBm
```

3.6 - Check Alarms

Command:

```
show system alarms
```

Expected Output:

```
No alarms currently active
```

Command:

```
show chassis alarms
```

Expected Output:

```
No alarms currently active
```

3.7 - Show Interface Status

```
show interfaces terse
```

Expected Output:

```

ge-0/0/0          up    up
ge-0/0/0.0       up    up    eth-switch
ge-0/0/1         up    up
ge-0/0/1.0       up    up    eth-switch
ge-0/0/2         up    up
ge-0/0/2.0       up    up    eth-switch
[...]

```

Note: This is just an example, ports that are not connected or don't have a host powered up with show down.

3.8 - Show Interface Counters

The output of this command can be very large, so it is useful to narrow it down to a specific interface, or if you're looking to compare a single value across all interfaces, to use grep to limit the output.

Examples using Gigabit interfaces (ge) and 10G interface (xe):

```
show interfaces ge-### extensive
```

```
show interfaces xe-### extensive
```

Example showing all CRC errors:

```
show interfaces extensive | grep CRC
```

Example showing all rates:

```
show interfaces | grep rate
```

3.9 - Show MAC address table

```
show ethernet-switching table
```

3.10 - Check Logging

```
show log messages
```

You can skip to the end of the log and just show the most recent:

```
show log messages | last 50
```

Juniper allows you to filter easily with grep, for example:

```
show log messages | grep "NTP"
```

3.11 - Save RSI and Logs to send to LBLnet / Juniper TAC

Replace DATE with today's date:

```
request support information | save /var/tmp/rsiDATE
```

```
start shell
```

```
tar czf /var/tmp/logsDATE.tar.gz /var/log
```

Then from a server that you can SSH to the switch from, SCP those files back to you. (See [2.1 - Management IP Address](#) for instructions on configuring a management IP address):

```
scp 256.0.0.50:/var/tmp/rsiDATE /var/tmp
```

```
scp 256.0.0.50:/var/tmp/logsDATE.tar.gz /var/tmp
```

3.12 - Check LACP status

If you are using LACP for Link Aggregation, you can check its status with:

```
show lacp interfaces
```

An example of working LACP is below:

Aggregated interface: ae0

LACP state:	Role	Exp	Def	Dist	Col	Syn	Aggr	Timeout	Activity
xe-0/2/2	Actor	No	No	Yes	Yes	Yes	Yes	Slow	Active
xe-0/2/2	Partner	No	No	Yes	Yes	Yes	Yes	Slow	Active
xe-1/2/2	Actor	No	No	Yes	Yes	Yes	Yes	Slow	Active
xe-1/2/2	Partner	No	No	Yes	Yes	Yes	Yes	Slow	Active

LACP protocol:	Receive State	Transmit State	Mux State
xe-0/2/2	Current	Slow periodic	Collecting distributing
xe-1/2/2	Current	Slow periodic	Collecting distributing

3.13 - Graceful Shutdown

Juniper switches should be gracefully shutdown whenever possible to prevent corruption of their file system. The switches do maintain 2 separate partitions in the event that the master becomes corrupted, but repairing and booting back onto the primary partition requires rebooting the switches and is a disruptive action.

You have two options for performing a shutdown:

3.13.1 - Shutdown via LED Front Panel

Note this must be performed on each unit!

1. Using the front panel on each switch, press the menu button (looks like 3 lines) until you reach MAINTENANCE MENU
2. Press the Enter button
3. The LED screen should display "SYSTEM HALT?"
4. Press the Enter button
5. You will be prompted "CONFIRM HALT?"
6. Press the Enter button
7. The LED screen should change to "SYSTEM HALTING..." though sometimes it will simply freeze and the status LEDs will no longer be illuminated. The screen will NOT change again to show that the system has fully halted. You should wait ~5 minutes before removing power from the switch to allow it time to fully halt.
8. Disconnect the power cables from the back of the switch. The switch should go dark.
 - a. When the cables are re-connected, the switches will automatically boot back up.

3.13.2 - Shutdown via SSH/Console

1. SSH or console into the Master switch unit.
2. Enter the command

```
request system power-off all-members in 0
```

3. Type "yes" to confirm.
4. The switch will likely wait 1 minute before initiating the power off.

- a. If you are consoled into the switch, you should eventually see a message that says "System is halted. Press any key to reboot". This indicates that you can remove power from the switch.
5. The LED screen should change to "SYSTEM HALTING..." though sometimes it will simply freeze and the status LEDs will no longer be illuminated. You should wait ~5 minutes before removing power from the switch to allow it time to fully halt.
6. Disconnect the power cables from the back of the switch. The switch should go dark.
 - a. When the cables are re-connected, the switches will automatically boot back up.

4.0 - Replacing a failed unit

1. Boot the spare unit up independently from any other connections (do not connect it to the other switches).
2. Console into the unit, login with root, there should be no password.
3. Type: **cli**
4. Now you must configure which SFP+ (fiber) ports will be used for Stacking (AKA Virtual-Chassis Ports, vc-ports):

When replacing unit 0 or 1 (compute cluster):

```
request virtual-chassis vc-port set pic-slot member 0 2 port 3
```

When replacing unit 2 (DAQ/electronics):

```
request virtual-chassis vc-port set pic-slot member 0 2 port 2
request virtual-chassis vc-port set pic-slot member 0 2 port 3
```

5. Now shut down the spare unit cleanly:


```
request system power-off
```
6. When it has fully halted (message will say press any key to reboot), disconnect power.
7. Now you should be able to transfer the optical modules and rear QSFP+ stacking cables (if applicable) from the failed unit to the spare unit. Do so using the same ports as the failed unit. DO NOT POWER ON SPARE UNIT YET.
8. Login to the current Master Routing Engine (this should be either unit 0 or 1) identified by a solid-lit green 3rd LED on the front panel. A blinking 3rd LED indicates the Backup Routing Engine.
9. At a CLI, type: **edit**
10. Type: **edit virtual-chassis member #**
 - a. Where # is the unit number you are replacing
11. Type: **set serial-number NEWSERIAL**
 - a. Where **NEWSERIAL** is the serial number (found on the back) of the spare unit

12. Type: **top**
 13. Type: **show | compare**
 - a. Verify that only the serial number is changing
 14. Type: **commit**
 15. Type: **exit**
 16. Now power on spare unit and wait several minutes for it to finish booting.
 17. From the command line, you can check the status by doing a "**show virtual-chassis**". When all members are up, their status will show "Prsnt".
-

5.0 - EMERGENCY: Minimally Operational Configuration

WARNING: These steps should only be taken in an emergency when all other troubleshooting has failed. Performing these steps will RESTORE TO FACTORY DEFAULTS and will reduce bandwidth, resiliency and remove safeguards against other issues (loops, broadcast storms, etc). Recovering from this will take much time and effort so this should only be used as a very last resort to restore minimal operations.

1. Disconnect all cables from all switches.
 2. Using the front panel on each switch, press the menu button (looks like 3 lines) until you reach MAINTENANCE MENU
 3. Press the Enter button
 4. Press the menu button until you reach "FACTORY DEFAULT?"
 5. Press the Enter button
 6. You will be prompted "RESTORE DEFAULT?"
 7. Press the Enter button
 8. The switch will reboot. Repeat this step on each switch.
 9. Now connect Switch 0 and Switch 2 with a single 10G SFP+ SR link. See diagram labeled "EMERGENCY - MINIMALLY OPERATIONAL CONFIGURATION"
 - a. Switch 0 xe-0/2/3 to Switch 2 xe-0/2/3
 - b. Switch 1 is NOT CONNECTED
 10. Connect the Master Node 1, Storage Node 1, and the NetApp via single Direct Attach Copper cables to Switch 0. Do not connect multiple links to the NetApp.
 11. Connect the compute nodes and the DAQ/sensors.
 12. When the switches have finished booting, they will each be operating independently of each other, not as a virtual-chassis. It is important that loops are not caused in this configuration or it may hinder network operation.
-

6.0 - Reference

6.1 - Naming Conventions

Juniper uses naming conventions that may be unfamiliar to non-Juniper staff. Below is a quick guide to help make sense of conventions used in this document:

1G Ethernet = ge
10G Ethernet = xe
40G Ethernet = et
Management Ethernet = me
Virtual Interface = irb
Aggregated Ethernet = ae
Virtual-Chassis Port = vcp

When used for virtual-chassis (stacking), the following ports show up with vcp (Virtual-Chassis Port) designations:

et-0/1/0 = vcp-255/1/0
et-0/1/1 = vcp-255/1/1
xe-0/2/3 = vcp-255/2/3
et-1/1/0 = vcp-255/1/0
et-1/1/1 = vcp-255/1/1
xe-1/2/3 = vcp-255/2/3
xe-2/2/2 = vcp-255/2/2
xe-2/2/3 = vcp-255/2/3

See attached diagram labeled "NAMING CONVENTIONS".

6.2 - Management Login

Username: _____

Password: _____

IP Address: 10.0.0.6
Subnet Mask: 255.255.0.0 or /16
Default Gateway: None

Firewall ACL restricted:
10.0.0.0/16

6.3 - SNMP

SNMP Community: _____

Restricted to:

10.0.0.0/8

172.16.0.0/12

192.168.0.0/16

6.4 - Syslog

None configured by default.

Firewall ACL restricted:

10.0.0.0/16

6.5 - Reference Diagrams

See attached.