Overview

These release notes support Cumulus Linux POC Branch with the lsb-release ending in 20130328-cumulus_linux_poc. These release notes describe supported HW platforms, currently available features, and known issues.

Supported HW platforms:

Cumulus Linux POC 03-28-2013 supports the following hardware platforms:

Configuration	Accton	DNI
48 x 10G + 4 x 40G	ES5652BT1 or BT2	ET-7448BF
48 x 1G + 4 x 10G		ET-6448R

Cumulus is currently qualifying other vendors and configurations, such as Quanta (LY2, LB9) and the Accton 48x1G platform. If there is a specific platform that is of interest please contact Cumulus.

Available memory per system on qualified platforms:

System	MTD Flash	Block Device
dni-7448	128MB	SD 8GB
dni-6448	64MB	8GB front-panel USB
Accton-bt1	8MB	USB 2GB

Features Supported:

Networking L2/L3 Features

Feature	Supported	Notes
LLDP/CDP (both rx/tx)	yes	patched lldpd
Bridging	yes	supported via brctl command in linux
VLAN 802.1q trunk	yes	
Control Plane ACL's	yes	cl-mgmtacltool is available in /usr/cumulus/bin - allows control of packets headed to CPU
Jumbo MTU	yes	
ECMP	yes	64 is supported on HW but Cumulus has tested to 16
OSPFv2	yes	
OSPFv3	yes*	*Still in testing, issues are listed in this release note.
v4/v6 Static Routes	yes	
BGP v4/v6	yes	

Management Interface/Trouble Shooting/Monitoring

Feature	Supported	Notes
SSH interactive or explicit command	yes	
FTP & TFTP	yes	
Scripting: Bash, Perl, Python, ruby	yes	
Ping & traceroute	yes	
syslog, rsyslog	yes	
logrotate	yes	
auditd	yes	
SCP	yes	Untested
SNMP v2 (via Net-SNMP)	yes	Untested
collectd, ganglia, monit*	Only Monit	Monit tested & preconfigured in the image

Known Issues:

Issues are categorized for easy review. Some issues are fixed but will be available in a later release. Future fixed issues are noted in the "fixed release" column with the branch name the fix will be available in.

If the "fixed release" is "mainline", this means the fix is in Cumulus Linux's internal mainline branch, but not yet allocated to a customer branch/release.

Layer 2 Issues:

Key	Summary	Description	Affected Release	Fixed Release
RN-32	Adding bridges will increase boot up time	If the "bridge_maxwait" parameter is not set, the system could take approximately 2x forwarding delay to bring the system up. Its best to set the "bridge_maxwait" to 1. e.g. CFG auto br1004 iface br1004 inet static address 14.0.0.37 netmask 255.255.0.0 bridge_ports regex (swp[1|6|7|8].1004) bridge_stp on bridge_bridgeprio 32768 bridge_maxwait 1 bridge_ageing 200 bridge_fd 30 down ip addr flush dev br1004	cumulus_linux_poc	
RN-33	bridge mac aging_time is preset to 5minutes	The bridge mac aging_time is preset to 5 minutes and cannot be changed. This issue is fixed in a later release	cumulus_linux_poc	mainline

RN-26	STP:Port ID election tie breaker is not predictable unless send port priority is set	Documentation note for linux behavior w.r.t STP PortID selection STP election: http://en.wikipedia.org/wiki/Spanning_tree_protocol ~~~~~~~~~~ In summary, the sequence of events to determine the best received BPDU (which is the best path to the root) is • Lowest root bridge ID - Determines the root bridge • Lowest cost to the root bridge - Favors the upstream switch with the least cost to root • Lowest sender bridge ID - Serves as a tie breaker if multiple upstream switches have equal cost to root • Lowest sender port ID - Serves as a tie breaker if a switch has multiple (non-Etherchannel) links to a single upstream switch, where: o Bridge ID = priority (16 bits) + ID [MAC address] (48 bits); the default bridge priority is 32768, and o Port ID = priority (4 bits) + ID [Interface number] (12 bits); the default port priority is 128. Linux Implementation: ~~~~~~~~~~~~~~~~~~~~ We elect the port ID based on the sequence the port is getting added in the given bridge. As result we may endup getting larger physical port lowest interface number. If user does not define the port priority at the sender side then topology predictability is lost. It is mandated for our system that we set appropriated sender port priority. consider following operations:	cumulus_linux_poc	
RN-5	forwarding will be broken if the bridge mac address is changed manually	The kernel bridge driver always makes a bridge device inherit its mac from one of its member ports, but user can also set arbitrary mac for the bridge device with ifconfig or equivalent. In case user does change it, the bridge driver doesn't enter the new mac in the fdb and doesn't make it a 'local' mac. So packets that are supposed to terminate on the bridge device or routed through the bridge device don't actually get delivered up the ip stack.	cumulus_linux_poc	mainline
RN-39	Bridge, VLAN and L3 routed interface limits on Cumulus Linux	Currently, Cumulus linux "tested" limits (on Trident+ systems) are: 200 bridges 200 L3 interfaces (this includes 64 used for physical ports as well as routed sub-interfaces) 200 VLANs	cumulus_linux_poc	mainline
RN-42	user vlans are restricted to the range of 1-1999	Currently, user can only configure vlans in the fixed range of 1-1999.	cumulus_linux_poc	mainline
RN-25	Management interface is down if a bridge with no ports is configured through the configuration file.	This only affects Cumulus Linux POC. It has been fixed in a later release yet to be released. If a bridge is configured with no ports on it. The management port eth0, will not be able to get an IP address from DHCP because the interface is down, even if eth0 is configured in the /etc/network file.	cumulus_linux_poc	mainline
RN-1	restarting switchd flaps all switchports	switchd is a user level process created by cumulus to provide an abstraction of the physical ports and the functionality provided by the switching ASIC sdk. Switchd maps physical ports on the switching ASIC to logical ports (tap ports) in the kernel and ensures that CPU bound packets are properly exposed on the proper logical objects to user level processes. These exposed tap ports in the kernel are considered "running" if their file descriptors are open. If switchd exists, its closes the tap fds, hence resulting in all links going down.	cumulus_linux_poc	

RN-12	The switch's forwarding of VLAN Tagged packets is different from Linux	In Cumulus Linux, if tagged packets are sent to a untagged port, they are dropped. This is similar to general switch functionality from most vendors. However in Linux, if a tagged packet is sent to a untagged port, it will be forwarded. Fixed in later release of cumulus linux	cumulus_linux_poc	mainline
RN-44	mac entry is not properly removed in libnl cache	A dynamic mac entry is stuck in libnl cache after a port is removed from a bridge. This causes switchd to repeatedly try and delete the entry from kernel. Note: switchd is a user level process created by cumulus to provide an abstraction of the physical ports and the functionality provided by the switching ASIC sdk. Switchd maps physical ports on the switching ASIC to logical ports (tap ports) in the kernel and ensures that CPU bound packets are properly exposed on the proper logical objects to user level processes. Observed behavior: - swp1 and swp2 are members of a bridge, macs learned on each port, total about 10 macs - swp1 is removed from the bridge, all it's learned macs are flushed from kernel and hardware. But one mac seems to be caught in the libnl cache. On every sync, switchd thinks it needs to be deleted from kernel - Following Message shows up continuously - "bridge: RTM_DELNEIGH swp1 not a bridge port" - after a forced resync, the mac is gone -> TO RESYNC - a) First find switchd PID b) assume PID ix X -> type: 'kill -SIGRTMIN X' Example: Right before the bridge port removal: root@dni-7448-26:~# brctl showmacs br0 port name mac addr is local? ageing timer swp2 00:02:00:00:00:08 no 10.66 swp2 00:02:00:00:00:09 no 3.93 swp1 44:38:39:00:12:9c yes 0.00 swp2 44:38:39:00:12:9d yes 0.00 swp2 90:e2:ba:04:ef:14 no 10.16 After bridge port removal (that mac is gone from both hw and sw): root@dni-7448-26:~# cl-bcmcmd l2 show mac=00:02:00:00:00:08 vlan=2000 GPORT=0x2 modid=0 port=2/xe1 mac=00:02:00:00:00:09 vlan=2000 GPORT=0x2 modid=0 port=2/xe1 Hit root@dni-7448-26:~# brctl showmacs br0 port name mac addr is local? ageing timer swp2 00:02:00:00:00:09 no 1.59 swp2 44:38:39:00:12:9d yes 0.00 root@dni-7448-26:~# br0: port 2(swp2) entering forwarding state THE FOLLOWING SHOWS UP ON THE SCREEN CONTINUOUSLY: bridge: RTM_DELNEIGH swp1 not a bridge port bridge: RTM_DELNEIGH swp1 not a bridge port bridge: RTM_DELNEIGH swp1 not a bridge port bridge: RTM_DELNEIGH swp1 not a bridge port	cumulus_linux_poc	mainline

Layer 3 Issues:

Key	Summary	Description	Affected Release	Fixed Release
RN-30	IP packets to a network broadcast address are routed	IP packets with the network broadcast address (the network with the host bit all set to 1) is being routed to different interfaces. As an example - assume the following networks 11.0.1.0/24 and 11.0.2.0/24 are configured on the device. A host, 11.0.1.5, sends a IP packet with destination address of 11.0.2.255. The hosts on the 11.0.2.0/24 network are receiving these packets. Default behavior will be to disable this and a command to enable/disable will be added in the future.	cumulus_linux_poc	
RN-20	software forwarding for IPv6 does not support multi path	Currently, the kernel does not support multipath forwarding for IPv6. Issue is being worked for the initial release of cumulus linux. It will allow multipath functions like trace route to properly find/expose the various paths in IPv6.	cumulus_linux_poc	mainline
RN-31	IP packets with illegal source addresses are forwarded	IP packets will the illegal source addresses of 127.0.0.1, 255.255.255.255, and the subnet broadcast are now forwarded as if the source addresses were legal. RFC1812 states that packets with these source addresses should not be forwarded.	cumulus_linux_poc	
RN-4	if site-local v6 address is present on an interface, all addresses are lost on that interface on a ifconfig down	Two scenarios shown below. 1) Adding a site-local scope address and shut down the interface. The interface v6 address was lost. 2) Same experiment with global scope addresses show that they are retained. 3) (not shown- but also breaks) If site-local and global-scope are both present, both are lost, when interface is brought down. root@dni-7448-09:/home/cumulus# ifconfig swp36 swp36 Link encap:Ethernet HWaddr 44:38:39:00:02:a5 UP BROADCAST MULTICAST MTU:1500 Metric:1 RX packets:0 errors:0 dropped:0 overruns:0 frame:0 TX packets:0 errors:0 dropped:0 overruns:0 carrier:0 collisions:0 txqueuelen:500 RX bytes:0 (0.0 B) TX bytes:0 (0.0 B) root@dni-7448-09:/home/cumulus# ip addr add fec0::1/128 dev swp36 root@dni-7448-09:/home/cumulus# ifconfig swp36 swp36 Link encap:Ethernet HWaddr 44:38:39:00:02:a5 inet6 addr: fec0::1/128 Scope:Site UP BROADCAST MULTICAST MTU:1500 Metric:1 RX packets:0 errors:0 dropped:0 overruns:0 frame:0 TX packets:0 errors:0 dropped:0 overruns:0 carrier:0 collisions:0 txqueuelen:500 RX bytes:0 (0.0 B) TX bytes:0 (0.0 B) root@dni-7448-09:/home/cumulus# ifconfig swp36 down root@dni-7448-09:/home/cumulus# ifconfig swp36 swp36 Link encap:Ethernet HWaddr 44:38:39:00:02:a5 BROADCAST MULTICAST MTU:1500 Metric:1	cumulus_linux_poc	

RX packets:0 errors:0 dropped:0 overruns:0 frame:0
TX packets:0 errors:0 dropped:0 overruns:0 carrier:0
collisions:0 txqueuelen:500
RX bytes:0 (0.0 B) TX bytes:0 (0.0 B)

Now with global-scope address:
root@dni-7448-09:/home/cumulus# ip addr add 2002::1/64 dev
swp36
root@dni-7448-09:/home/cumulus# ifconfig swp36
swp36 Link encap:Ethernet HWaddr 44:38:39:00:02:a5
inet6 addr: 2002::1/64 Scope:Global
BROADCAST MULTICAST MTU:1500 Metric:1
RX packets:0 errors:0 dropped:0 overruns:0 frame:0
TX packets:0 errors:0 dropped:0 overruns:0 carrier:0
collisions:0 txqueuelen:500
RX bytes:0 (0.0 B) TX bytes:0 (0.0 B)

root@dni-7448-09:/home/cumulus# ifconfig swp36 down
root@dni-7448-09:/home/cumulus# ifconfig swp36
swp36 Link encap:Ethernet HWaddr 44:38:39:00:02:a5
inet6 addr: 2002::1/64 Scope:Global
BROADCAST MULTICAST MTU:1500 Metric:1
RX packets:0 errors:0 dropped:0 overruns:0 frame:0
TX packets:0 errors:0 dropped:0 overruns:0 carrier:0
collisions:0 txqueuelen:500
RX bytes:0 (0.0 B) TX bytes:0 (0.0 B)

Address is not "lost"\!

RN-41	Route Summarization acts differently in OSPFv3 vs. OSPFv2	In OSPFv2, regardless of the order of the routes, the longer prefix is always chosen: area 0.0.0.0 range 11.0.0.0/16 area 0.0.0.0 range 11.0.0.0/8 In this case, 11.0.0.0/16 is always used regardless of the input order. In OSPFv3, this is not always the case. Different results occur depending on the order routes are entered. In the following case we get the longest match: area 0.0.0.0 range 2000:1000::/32 area 0.0.0.0 range 2000::/16 However, if we do it in the reverse order of: area 0.0.0.0 range 2000::/16 area 0.0.0.0 range 2000:1000::/32 We get both summarizations.	cumulus_linux_poc	
RN-34	system will not support 8000 /64 static IPv6 routes with ecmp	If ECMP is set up for 8000 static /64 IPv6 routes, the system will not support the configuration in that the routes will not be distributed from the kernel to the switch ASIC. This issue is fixed in future release.	cumulus_linux_poc	mainline
RN-22	OSPFv3: Current multipath limit is set to 4	OSPFv3 currently has a #define OSPF6_MULTI_PATH_LIMIT that is set to 4. Its been verified in running a Nx2 topology that the maximum ECMP computed by OSPFv3 is 4. For reference, OSPFv2 does not even have such a #define. Enabling greater than 4 is enabled in a future Cumulus Linux version, but not the Cumulus Linux POC version.	cumulus_linux_poc	mainline

RN-28	In OSPFv3, ABR will not advertise the prefix until the backbone area refreshes the LSA if the non-backbone prefix is aged out.	In an ABR if a prefix is learned from both the backbone and non-backbone area the ABR will not advertise the prefix until the backbone area refreshes the LSA if the non-backbone prefix is aged out.	cumulus_linux_poc	mainline
RN-8	SeqNum Wrapping Handling not supported in quagga OSPFv3	As specified in the OSPFv2/v3 RFC, when an LSA's sequence number reaches its maximum value, where the next increment would cause the number to wraparound, special handling must be triggered. Quagga OSPFv2 code has this handling, but OSPFv3 does not. This addition is being considered for future release of Quagga and Cumulus Linux	cumulus_linux_poc	
RN-6	Max-metric support needed in ospfv3 (quagga)	Quagga's ospf6d doesn't support max-metric command. Without this command, graceful withdrawal from the network due to maintenance and other scheduled downtime may result in more network disruption.	cumulus_linux_poc	
RN-23	OSPFv3: Support Stub and Totally Stubby Areas	OSPFv3 doesn't support stub areas and totally stubby areas. This issue is fixed in a later version of Cumulus Linux.	cumulus_linux_poc	mainline
RN-40	OSPFv3 needs support for the command "log-adjacency-changes detail"	Currently, OSPFv3 doesn't support the command "log-adjacency-changes detail". OSPFv2 already has support for this. Cumulus will try enable this for OSPFv3 in Quagga.	cumulus_linux_poc	

Configuration Management/Trouble Shooting/Monitoring Issues:

Key	Summary	Description	Affected Release	Fixed Release
RN-2	defunct processes (fan monitor and mgmtacl_check.s)	Due to the current version of monit, unwanted zombie processes will remain in cumulus linux. In particular - fan_monitor, and mgmtacl. root 30167 616 0 12:24 ? 00:00:00 [fan_monitor.py] <defunct> root 30168 616 0 12:24 ? 00:00:00 [mgmtacl_check.s] <defunct> From the following: http://mmonit.com/monit/documentation/monit.html#program_status_testing Requoted: The asynchronous nature of the program check allows for non-blocking behavior in the current Monit design, but it comes with a side-effect: when the program has finished executing and is waiting for Monit to collect the result, it becomes a so-called "zombie" process. A zombie process does not consume any system resources (only the PID remains in use) and it is under Monit's control; The zombie process is removed from the system as soon as Monit collects the exit status. This means that every "check program" will be associated with either a running process or a temporary zombie. This unwanted zombie side-effect will be removed in a later release of Monit.	cumulus_linux_poc	
RN-10	cl-phy-update doesn't support aggregated ports	Ports can be aggregated into a larger interface in Cumulus Linux. Unfortunately support for aggregated ports is not yet supported when running cl-phy-update. If there are any ganged ports during a SW upgrade it is recommended to ungang these ports	cumulus_linux_poc	

		Support is being planned for the initial Cumulus Linux release.		
RN-3	mgmtacltool needs to be extended to non-Trident chips and IPv6	The cl-mgmtacltool does not yet support IPv6 on any switch on the cumulus linux hardware compatibility list (HCL). The tool currently only supports the Trident family of products.	cumulus_linux_poc	
RN-14	cl-sfputil does not work for dni-7448 ganged ports.	If 4 ports are "ganged" as a 40G port, cl-spfutil will only list the sfp details of the first port.	cumulus_linux_poc	
RN-21	cl-sfputil does not pick up qsfp details on accton switches	cl-sfputil does not properly provide QSFP details on accton switches.	cumulus_linux_poc	
RN-11	IEEE 802.3 packets with length error are not included in ingress port packet counter	When sending a L2 frame into the system, the packet counters in "ifconfig" for the ingress port do not reflect the correct number of packets. However, the byte counters are reporting the correct values. Packet being sent: sudo mz -v swp1 -c 100 "FF:FF:FF:FF:FF:FF:00:00:00:00:00:FF:00: 00:00:00:00:00:00:00:00:00:00:00:00:00:00:00:00:00:00:00:00: 00:00:00:00:00:00:00:00:00:00:00:00:00:00:00:00:00:00:00:00: 00:00:00:00:00:00:00:00:00:00:00:00:00:00:00:00" Output: cumulus@dni-7448-05$ sudo ifconfig swp1 swp1 Link encap:Ethernet HWaddr 44:38:39:00:03:81 inet6 addr: fe80::4638:39ff:fe00:381/64 Scope:Link UP BROADCAST RUNNING MULTICAST MTU:1500 Metric:1 RX packets:143 errors:0 dropped:24 overruns:0 frame:0 TX packets:419 errors:0 dropped:0 overruns:0 carrier:0 collisions:0 txqueuelen:500 RX bytes:46284 (45.1 KiB) TX bytes:39663 (38.7 KiB) cumulus@dni-7448-05$ sudo ifconfig swp1 swp1 Link encap:Ethernet HWaddr 44:38:39:00:03:81 inet6 addr: fe80::4638:39ff:fe00:381/64 Scope:Link UP BROADCAST RUNNING MULTICAST MTU:1500 Metric:1 RX packets:155 errors:0 dropped:26 overruns:0 frame:0 TX packets:467 errors:0 dropped:0 overruns:0 carrier:0 collisions:0 txqueuelen:500 RX bytes:54862 (53.5 KiB) TX bytes:43953 (42.9 KiB) counters have incremented by 7K, but the packet counts have only increased by 12.	cumulus_linux_poc	
RN-37	debuginfo not available on all packages in cumulus linux	Unfortunately none of the Cumulus Linux packages have the symbols package included. If debugging and/or symbol packages are needed, please contact Cumulus Support	cumulus_linux_poc	