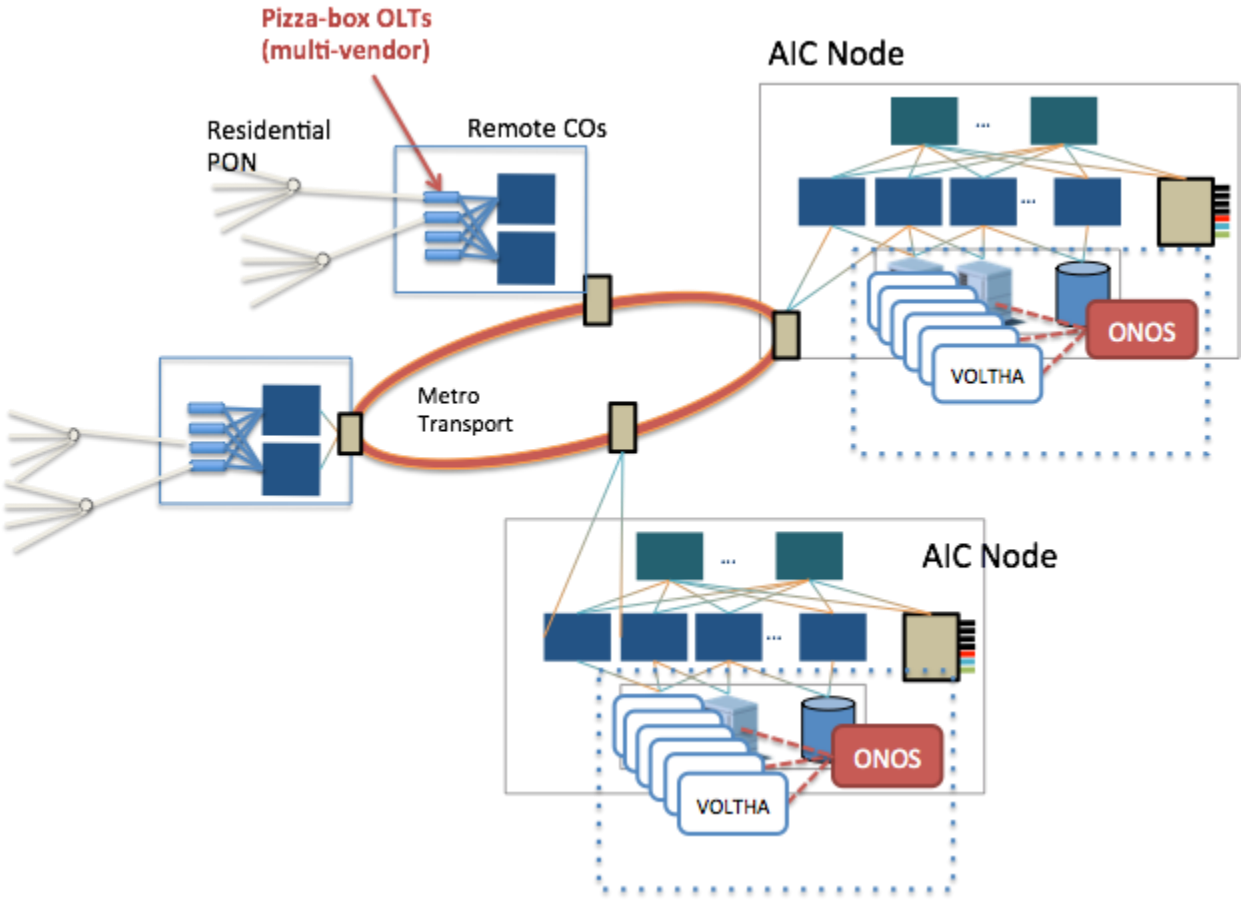
Understanding PON Epics

This document is a collection of questions regarding the EPICs contributed by AT&T towards the development of an open-source SDN based OLT/ONT control mechanism. It's goal is to help us understand the requirements in the EPICs so that they can be converted into actionable software development stories tracked in the OpenCORD Jira. This is a working document and will be edited after every meeting between ON.Lab and AT&T teams.



Connectivity in AIC Questions

1. The pizzabox based OLTs will be deployed in remote COs together with Access Leaf's from a traditional networking vendor.
2. The Access leafs will communicate using EVPN tunnels with AIC nodes. The control and management of the OLTs will be done by VOLTHA instances in the AIC cloud. The management traffic will reach the VOLTHA instance on a management VLAN (vlan 4093

through the EVPN tunnel) tagged. Data plane traffic will go to a traditional BNG or vBNG.

3. For Micro-OLT - what does that plug into? Too difficult to control the olt in a vendor AIC remote leaf -- maybe into some other box (eg Whitebox) and then these plug into the vendor AIC remote leaf - “blue” box.
4. Touch free installation like G.Fast? Who is assigning the management vlan- -- console port management vlan and system IP address is manual. Standard for touch free in G.fast not agreed on yet - so it is a goal in this project, not a requirement (PMA-like persistence/management/application function).
5. Would be good to have bootstrap and control flow writeup.

OLT Device Function Questions

These set of questions are related to the OLT device itself, and have been grouped into various categories for ease of understanding.

Other Features

1. NTP, BITS, 1588 timing - is this a device requirement? Yes. Does NAL have to manage timing in the device? Yes via get/set functions - in first phase NTP will be running in VOLTHA. 1588/BITS for other implementations, NTP starting point
2. Configure: Who develops the generic Driver Abstraction Layer API for these? Do these need to be configurable via Netconf/Yang from an external body? If so, then what YANG models do they use? Talking to Sumithra … have initial list of available YANG models from BBF (requesting access to the actual models) … needed contributions for the gaps identified in the models, hopefully from open-source -- as much as possible try to make medium independent for configuring these parameters.
    a. PON FEC on/off - always on in 10G PON as it limits the range if you turn it off
    b. MTU - jumbo frames
    c. ONT fiber range MIN/MAX
3. XGSPON - who’s hardware? It's all 10G PON (10up, 10down) ONTs are the question mark. Several vendors. Field trial focus is on 10G PON (not GPON)
4. TWDM - NGpon2 (future - 2 years down the road) - administrative domains, G.989 ONT related? Not a focus right now
5. Business Ethernet - transparent cross-connect mode (pass L2CP, ARP, IGMP) - what configuration on OLT? port/vlan -- always mapped 1:1 to upstream port/vlan & turn mac learning off. (different product - evolutionary approach over PON)

VLAN/Bridging Features

Is the assumption being made here that there is always going to be a bridging chip? Or is the pon chip expected to have this functionality? Currently it is expected that there is a bridging chip together with the PON asic or fpga.

1. Functionality
    - LAG or is it M-LAG? What kind? - active/passive and active/active - is there a standard? Ask the A-Leaf folks. Implemented in OLT pizzabox software or in VOLTHA proxy (avoid in VOLTHA if possible)? Broadcom may be able to do it in the box with their chipsets, but what about other boxes or the whitebox (with micro-OLT)? Is it better to have common software for all boxes?
        i. Blaine: The pair of leafs use the same LAG Bundle ID and work as a team using the EVPN VPWS in an Active/Active path method.
        ii. 2 groups of LAGs in a Fast Failover - - Blaine will ask A-Leaf for LACP necessity
    - 5 VLAN tags? DPI for 2 tags? → The packet may have many tags (for example business services may already have 3 tags before OLT/ONT sees it), but the OLT/ONT has to deal with pushing and popping at most 2 vlan tags, and potentially MPLS label
    - PAUSE between ONOS and OLT? -- not needed
    - 802.3Qbb priority based flow control -- in Yang model? May or may not be used.
    - MAC ACL and Ethertype filters (pppoe, ipv4,v6, arp) → setup at beginning - eg. say 4 macs you can see from consumer - very static -- may be either OF or Yang
    - How do the following relate to current ONOS app? Saurav action-item - see JIRA stories in this doc
        i. S and C Vlan tag configuration and manipulation in both directions -- via ONOS app and OF
        ii. XGEM port-id or multiple such ids to single vlan config multiple GEM ids assigned to a single vlan - where GEM ids identify the traffic class. GEM ids assigned via OMCI - VOLTHA? Example from Shawn
        iii. Ethertype and “acceptable frame types” per port, translations and priority markings
        iv. TLS enablement and Transparent ports? Translation table & priority marking? Is Transparent same as “admit all”?
    - N:1 S-VLAN and MAC learning (disablement for 1:1)
    - Duplicate MAC blocking? Bridging MACs among XGEM ports? Not done in CORD due to vSG --- not for residential, maybe for business service → not done today, not sure of need? Maybe future - will check on this - 1.1.1.82

- Add MPLS in the OLT itself - eth-dst-src-mac/outervlan/innervlan/mpls/<ethernet> -- alleviates scaling issues in the A-Leaf; requires OLT being part of control/management plane as well. --- ISIS, or MP-BGP or more.

2. Protocols
    - LACP - what is really used for M-LAG?
    - LLDP? Need with M-LAG?
    - Loop detection? Where is the loop? No STP.
    - What else is needed to interface with AIC remote-access-leaf boxes?
    - Who implements these protocols in the box?

QoS Features

1. Per vlan per service queuing/scheduling on the egress of uplinks? Upstream?
    - Does this mean per subscriber, per traffic class (or gem port)?
2. How are service classes determined? Do they need to be configured (marked or re-marked) according to
    - User port (what is physical user port? What is logical - is it GEM?)
    - Ethertype
    - Received Ethernet priority bits
    - IP proto for IGMP
3. Mark/remark for OLT initiated traffic? What is IP protocol port number?
4. GEM portids → to 8 priority levels Eth p bits → 8 traffic classes - upstream or downstream? Any combination of p-bit values to egress queues

The GEM port assignments are as follows:

- GEM port id (bidirectional ports) for the subscriber ports to carry traffic of different priorities. These GEM port ids are used to carry both the data service traffic and the multicast IGMP control message traffic.
- GEM port id (bidirectional port) to carry POTS signaling and bearer traffic.
- GEM port-id 4095 (unidirectional) to carry the downstream multicast stream. This GEM port id is considered a broadcast port in the downstream direction to deliver a single copy of the multicast group stream.
- GEM port id (bidirectional port) to carry OMCI traffic

Trusted entity like VOLTHA PON manager or ONOS app has to mark p bits to reach Radius server or ecomp DCAE → does the AIC network use these p bits and identify the management traffic and give it the right priority? Yes. AIC network will mark ONOS traffic all the same.

4 classes of traffic

- 1st 2 queques are strict priority
- Other 2 are WFQ 80/20 share

Hierarchical queuing done too - but requires sophisticated OLT - left out of OCP design

5. Ingress policing by classified flow? Which direction? 8 policers by subscriber physical port?

From Shawn:

802.1p Priority	Traffic Type and Description	Queue	Schedule Weight
7	(NC) Network Control (not used)	3	strict
6	(H1) Voice – includes CVoIP Bearer and Signaling	3	strict
5	(EF) Video - Generated IGMP traffic	2	strict
4	(H2) Controlled Load – includes VoD and IPTV SMS/RDP	1	Weighted 75%
3	(L1) Excellent Effort – includes RG management	1	Weighted 75%
2	(AF) not currently used	0	Weighted 25%
1	(L2) Background includes ICC and R-UDP	1	Weighted 75%
0	(BE) Best Effort – HSIA and default traffic	0	Weighted 25%

Upstream P-Bit and Queue Mapping

802.1p Priority	(Forwarding Class) Traffic Type and Description	Queue	Schedule Weight
7	(NC) Network Control	3	strict
6	(H1) Voice – includes CVoIP Bearer and Signaling	3	strict
5	(EF) Video – includes Multicast Video and IGMP traffic	2	strict

4	**(H2) Video – includes VoD and IPTV SMS/RDP**	**1**	**Weighted 90%**
3	**(L1) Excellent Effort – includes RG management traffic**	**1**	**Weighted 90%**
2	**(AF) - not currently used for downstream traffic**	**0**	**Weighted 10%**
1	**(L2) - includes ICC and R-UDP traffic**	**1**	**Weighted 90%**
0	**(BE) Best Effort – includes HSIA traffic**	**0**	**Weighted 10%**

Downstream QoS Queue Mapping

Multicast Features

1. 64 active per sub (configure max), and 4096 active mcast groups in total?
2. IGMP snooping function-- who is writing the software? Is it local or SDN based? Will matter for RFC 3376, 4541, 4604 (transparent snooping, proxy reporting) --- IGMP challenge - too much channel changes - wish to IGMP proxy. But have it both ways - forward in data plane as well. Standards based proxying
3. Multicast VLAN? Vlan 4000 - User ports or multicast groups are mapped to multicast VLAN? Jim will send control/data flow for multicast traffic - see below
4. Multicast CAC? Bandwidth control as per whitelist and channel rates -- not important in 10G PON, even with 4k channels.
5. Configurable treatment of IGMP per subs-port/vlan + whitelist groups - drop/transparently forward IGMP --- IGMP throttling rate limit per PON port (eg. 16 per second) - high priority
6. Rate limit or stop user ports injecting mcast traffic? Or just IGMP msgs - not about datatraffic
7. Why the need for configuring users in IP multicast groups? Users may not have bought service -- drop IGMP
8. Marking IGMP traffic that is initiated by the OLT - snoop and propagate

From Jim Brunetti

I tried to modify how we do things now to match what should happen with the Pizza box (or Tibit plugged into a White Box switch that is not the Leaf). Some items are left to the vendor. Some vendors may decide not to implement an IGMP snooper in the PON MAC/PHY and simply forward the IGMP requests to the internal Ethernet switch to be translated to VLAN 4000 (step 4) or some will be looking for IGMP packets coming from the subscriber and will do the translation to VLAN 4000 at the plug.

1. STB sends out IGMP request towards RG
2. IGMP proxy on RG forwards request out Ethernet port if request is for a new channel
3. The GPON MAC ONU will snoop the IGMP requests to determine which IGMP channels are present. The IGMP join message will be sent to the OLT via the User Data GEM port on the Subscribers assigned VLAN.
4. On the PON MAC/PHY, the IGMP proxy will receive the IGMP join on the Users VLAN. The IGMP join message will be sent towards the Ethernet switch on VLAN 4000.
5. VLAN 4000 is configured on the OLT uplink as the Multicast VLAN for IGMP and streaming traffic.
6. VLAN 4000 is configured on the ALeaf as the Multicast VLAN for IGMP and streaming traffic.
7. The ALeaf will replicate multicast traffic onto each port with a valid IGMP join for this channel on VLAN 4000. The Ethernet switch card will replicate the multicast traffic to each PON MAC/PHY in the distribution list for this channel on VLAN 4000. The PON MAC/PHY will replicate the multicast traffic onto the PON to each ONU in the distribution list for this channel using VLAN tag 4000.
8. The GPON MAC ONU will replicate the multicast traffic to its Ethernet port in the distribution list for this channel towards the RG as priority tagged traffic.

Each ONT is configured with filters to allow a maximum of 64 multicast channels and 16 IGMP messages per second per sub.

IP Features

1. DHCP handling -- where is it done today - BNG? What is expected in XGS-PON deployment? Still on the vBNG
2. ARP requirement -- for other home devices? For gateway? How is it done today? is this some kind of downstream filtering? 1.1.1.180 -- RG arps for gateway -- ensure it does not get flooded back downstream to all PON ports
3. IP addr spoof from where - home users or internet?- home! Mitigation via ingress or egress filtering, uRPF? **Where is it done today - BNG**? What is expected in deployment? Still on the vBNG
4. V4 and v6 simultaneous - performance degradation? Software ONOS and other control plane apps (dual stack)
5. IPv6 scenarios
    a. Multicast -- everything we do for v4 multicast? IPTV not done today on v6 multicast. More of a future requirement.
    b. ICMPv6 -- supporting RS/RA, ND? More for business than residential? RG will have v6 addr, and vBNG will be handling ND, RS?
    c. Supporting autoconfiguration?
    d. And DHCPv6? Again more vBNG requirements

6. NAT in vBNG?
7. 6to4? dual-stack?

OAM Features

1. CFM for business services - Business services (not priority over PON)
2. Under what scenarios are Ethernet OAM (802.1ag) Continuity Check, LinkTrace and Loopback protocols used? How does it apply in current XGS-PON plans?
3. CFM - end to end -- NTE connects to ONT --- to NTE on other business localtion. MEP is end to end, OLT spared from being MIP - also leafs (MIP), also ONTs -- today it's the ONT as MIP
4. MEPs for every vlan for a customer

Alarms

1. Threshold alarms for packet forwarding latency and jitter in the OLT? -- not a priority. Maybe approximated via queue length measurements
2. Alarm for IGMP join and leave processing latency - how to measure
3. Rogue interference detection and alarm? -- ONT registration, rogue ONT is different and is misbehaving -- link layer detection -- reported to controller/voltha supporting action at higher layer, OLt can do mitigation via shutting down ONT - PLOAM (DBA assignment) one level below OMCI, may leverage OMCI.
4. Alarm for Tx/Rx optical power min/max thresholds
5. Trap events from OLT when IP address spoofing occurs? - eg. Mac learning of 1, so if olt gets more, it is discarded there and alarm generated (does not happen at BNG); monolithic OLT may look at IP layers. In our case, OLT-MAC will not be looking at IP?
6. What other alarms or async event propagation are needed?

IDs

1. Access node identifier?
2. Agent circuit ID per individual access loop & logical port?
    a. Shawn will send representation
    b. More discussion and conclusion needed // or just a string to the controller app.

ONU/ONT Management Features

1. Set of techniques for rogue ONT detection and isolation/mitigation? Nothing more than standards based rogue ONT detection and mitigation in VOLTHA
2. Exposing loopback capability if available

3. ONT bootstrap & software upgrade
4. AES security, watchful sleep, watchdog?
5. Data rates?

Clustering Features

G.Fast support in VOLTHA

1. Gfast DPU today is transparent in OLT -- just passes through

JIRA Stories

1. VOLTHA VM support for receiving management/control traffic from (remote) OLT tagged (for example on VLAN 4093) -- already done
2. Support for time synchronization method - NTP - what does it need to send down to OLT?
3. **CORD-934**: Support for pushing, popping or swapping up to 2 VLAN tags, on traffic that may already have other vlan tags (need to test). This includes
    a. S-VID (0x88a8) and C-VID (ox8100) pushing (one or both) - done, but need to verify 0x88a8
    b. Translating existing C-VID to S-VID (with same or different vlan-id)
    c. Rewriting existing C-VID vlan-id and pushing S-VID
    d. Mapping multiple user ports at the ONU to the same S-VID (N:1 mode)
4. **CORD-935:** Support mapping of X-GEM Port Ids on the PON side to 8 Ethernet priority levels and at least 4 queues (up to 8 queues) - see tables above
    a. requires marking/re-marking Eth priority bits
    b. 2 of the queues are strict priority and the other 2 WFQ
5. **CORD-935:** Support classification of traffic based on Ethernet p-bits, ethtype, IP proto, L4 port numbers to queues - see tables above
6. **CORD-936:** Support MAC layer ACLs - permit/deny src/dst/both MAC addresses
7. **CORD-937:** Support Ethertype ACLs - permit/deny ppoe/ipoe/arp/ipv6
8. Multicast stories
    a. **CORD-938:** Ability to configure and support a max number of active mcast groups per sub (eg 64)
    b. **CORD-939:** Ability to configure and support a total of 4096 active mcast groups in the PON (a hardware requirement but also a software scalability requirement)
    c. **CORD-940:** Support standards based IGMP proxying so IGMP messages propagate to upstream nodes (in addition to IGMP snooping at the SDN controller)
    d. **CORD-941:** Ability to support a configurable multicast vlan (eg 4000) and mark (vlan priority) for IGMP traffic proxied by the OLT
    e. **CORD-942:** Ability to rate limit IGMP messages from a user port (eg. 16 messages max per second) -- rate should be configurable - needs hardware support
    f. **CORD-943:** Ability to drop IGMP messages from a user port (eg. user may not have bought service)
    g. **CORD-940:** Support IGMPv2 and v3 - today we do v3
9. **CORD-944:** Take changes to olt, aaa, and multicast apps and move to CORD repos, in a way that they work for voltha and traditional CORD.

10. **CORD-939:** Evaluate scalability of mcast RIB in ONOS to meet max rates and no. of active groups in PON
11. **CORD-945:** Alarm for Tx/Rx optical power min/max thresholds
12. **CORD-946:** Alarm for more than 1 MAC detected at user port
13. **CORD-947:** Alarm for rogue interference detection
14. **CORD-948:** Alarm for IGMP join/leave processing if latency exceeds threshold
15. **CORD-949:** Support 2 LAG groups in Active/Standby mode. May require LACP support as well.