**Facebook Voyager Program**

**Transponder Software Requirements**

**Document Number: TBD**

# Document Revision History

|  |  |  |  |
| --- | --- | --- | --- |
| Version | Comment | Author | Date |
| V0.10 | Initial Version | Michael Bortz | 05-10-2016 |
| V0.20 | Many changes | Michael Bortz | 05-16-2016 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

# Document Approval History

|  |  |  |
| --- | --- | --- |
| Person | Signature | Date |
| Michael Bortz - Engineering |  |  |
| Ilya Lyubomirsky - Engineering |  |  |
| Cathie Deal/Amit Ajuha - Supplier Management |  |  |
| Sridharan Rajagopal/Ramesh Padmanabhan - Program Management |  |  |
| Hans-Juergen Schmidkte - Director |  |  |

**Table of Contents**

Document Revision History 2

Document Approval History 2

Executive Summary 5

System Overview 5

Software Description and Block Diagram 5

Major Operational Modes 6

Major Software Functions 7

Software Block Diagram 8

User Classes 8

User Interfaces 9

Hardware Context 9

Control Plane Requirements 9

Development and Implementation Constraints 11

Meta-Requirements 12

Detailed Transponder Functional Requirements 13

Configuration Management 14

System Configuration Requirements 14

Client Interface Configuration Requirements 15

Broadcom Tomahawk ASIC Configuration Requirements 17

Network Interface (AC400) Configuration Requirements 18

Performance Monitoring Management 20

System Performance Management 20

Client Interface Performance Management 21

Broadcom Tomahawk Performance Management 22

Network Interface (AC400) Performance Management 22

Platform Performance Management 23

Fault Management 23

Alarm Severity 24

System Fault Management 24

Platform Fault Management 25

Client Interface Fault Management 26

Network Interface (AC400) Fault Management 27

Non-Functional Requirements 28

System Requirements 28

Firmware-Software Upgrade Downgrade Requirements 29

Flight Recorder Requirements 30

Security Requirements 30

Regulatory and Compliance Requirements 30

Reliability and Quality Assurance Requirements 31

Appendices 31

A. State Diagram for Transponder Module 31

B. Transponder Alarm Management 35

C. Detailed LLDP Snoop and Drop Requirements 38

D. LED stream format 40

References 55

# Executive Summary

This software requirements specification captures the application software requirements for the Facebook Voyager DWDM transponder. This transponder is a component of the Facebook Voyager optical DWDM transport system, which is a high capacity point-to-point WAN optical transport system to be used within the Facebook metro network over distances of roughly 10 km to 100 km. A block diagram of this system is shown in Figure 1. The transponder contains a layer 2 switch as well as an optical module to enable both layer 1 (i.e., layer 2 transparent) transport functions as well as advanced layer 2 packet-optical capabilities. In future expansions Facebook is considering to use the same transponder for long haul applications.

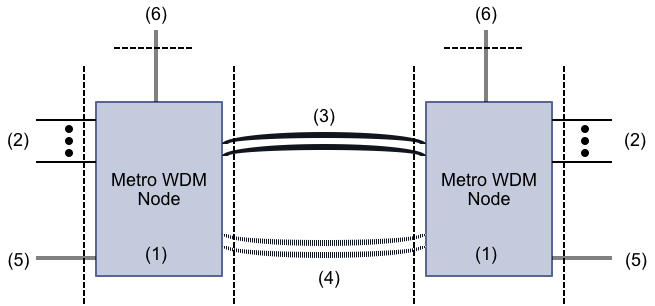


Figure . Block diagram of a metro optical WDM system. The numbers denote the following: 1) the Network element, which consists of multiple transponders, 2) the client interfaces, 3) the primary fiber link, 4) the backup fiber link, 5) the local management interfaces and 6) the northbound DCN management interface

# System Overview

## Software Description and Block Diagram

The Voyager transponder block diagram is shown below in Figure 2. The major components of the data plane are the 1) client optical interfaces, 2) the Broadcom Tomahawk Ethernet switching ASIC, and 3) the Acacia AC400 optical module. The transponder application software runs on the Com-E microprocessor and controls all aspects of the transponder, including major operational modes (described below) and FCAPS functionality. The application software shall have sufficient configuration capabilities and intelligence to enable semi-automated configuration of the major operational modes and simplified provisioning for circuits. The application software will also have simplified performance and fault management diagnostics to reduce operations complexity.

## 

Figure . Block diagram of the Voyager optical transponder

## Major Operational Modes

The transponder may be operated in 3 major operational modes, dominated by the configuration of the Broadcom Ethernet switching ASIC.

1. **Wire Mode (Layer 2 transparent):** In this mode the switch is configured in a port-mapped mode and is operated in an non-oversubscribed manner, and the transponder is effectively transparent to layer 2. In this mode there are a maximum of eight (8) 100GE client inputs and 800G of output capacity. In this mode the configuration and port mapping is static and may be fixed for all transponders. This is the basic mode of operation for the transponder. In this mode the transponder may be able to be operated 1) as a transponder in which the client inputs are enabled and data flows from client interfaces to the AC400 network interface, and b) as a regenerator in which the data is looped back within each AC 400 and the client interfaces are essentially inactive. These details are described in the document.
2. **Oversubscribed Wire Mode:** In this mode the transponder is operated with 12 active client interfaces and arbitrary oversubscription and VLAN tagging mechanisms are used to support identification of the data streams. In this mode of operation 4x10GE client interfaces may be used and the tagging can be standards complaint to either IEEE 802.1BR or 802.1Q. In this mode the Broadcom Tomahawk switching ASIC is used as a statistical multiplexing device but more advanced routing and switching features are not used.
3. **Packet-Optical Mode:** In this mode the full complement of switching and routing protocols is made available to support packet optical integration. In this mode the switch ports connected to the AC 400 are routed ports. The full description of the layer 2 and 3 capabilities in this mode are not described in detail in this document, but are contained in other FB and vendor documents.
4. **Maintenance Mode**: There shall also be an out of service maintenance mode to permit a richer diagnostic capability of the transponder than what is needed in operation. In this mode of operation live traffic may or may not be carried, although test traffic may be present on the transponder. This mode is used to define and store specific transponder configurations.

In all operational cases the FCAPS capabilities of the transponder application software are comparable. However, in maintenance mode various alarm and performance monitoring capabilities can be selectively deactivated. The operation of the AC 400 network interface is largely independent of the configuration and operation of the Broadcom Tomahawk switching ASIC.

## Major Software Functions

The major functions of the application software include:

1. Provide for selection of the transponder operational mode or maintenance mode
2. Develop a structured schema for the data model for the transponder that is compliant with YANG standards
3. Provide for northbound interfaces (API’s) to permit management of the transponder, with protocols including Netconf, Restful API, Thrift API, and CLI.
4. Provide an interface to the Broadcom Tomahawk Ethernet switching ASIC using the Broadcom SDK, and simplify the necessary set of commands based on the chosen transponder operational mode.
5. Provide an interface to the AC 400 using the AC400 SDK, and use a simplified set of commands to manage this subsystem
6. Provide for an interface to the client interfaces and permit the management of these interfaces
7. Permit configuration of the transponder, including selection of host mediated or autonomous turn up of the AC 400.
8. Permit fault management of the transponder with sufficient intelligence to take corrective action and simplify required actions by maintenance personnel
9. Permit advanced health monitoring of the transponder though judicious use of performance management capabilities of the underlying hardware
10. Monitor the health of the transponder by measuring the board voltages, currents, temperatures and other telemetry indicators

## Software Block Diagram

***Need OS and Open BMC description and requirements***

## 

Figure . SW stack

## User Classes

There shall be two classes of uses for the transponder

Operations Personnel: This class of user must be able to configure and operate the transponder. This class of user does not require access to or a complete understanding of all the capabilities of the maintenance mode and other than simple troubleshooting (i.e., confirming that the transponder is shut down) need not diagnose underlying faults.

Engineers: the class of user requires access to all capabilities of th transponder, including the full complement of capabilities in maintenance mode. This class of user can perform detailed troubleshooting of the transponder. This class of user may also use the SW to test the transponder during various design verification activities.

## User Interfaces

There shall be two user interfaces for the transponder

**Command Line Interface:** There shall be a CLI that enables setting and monitoring all transponder parameters

**Graphical User Interface:** There shall be a GUI that enables setting and monitoring a critical subset of transponder parameters. It is expected that the GUI will evolve over time and that initial version will have limited functionality.

## Hardware Context

The electrical block diagram of the transponder is shown below in Figure 4.

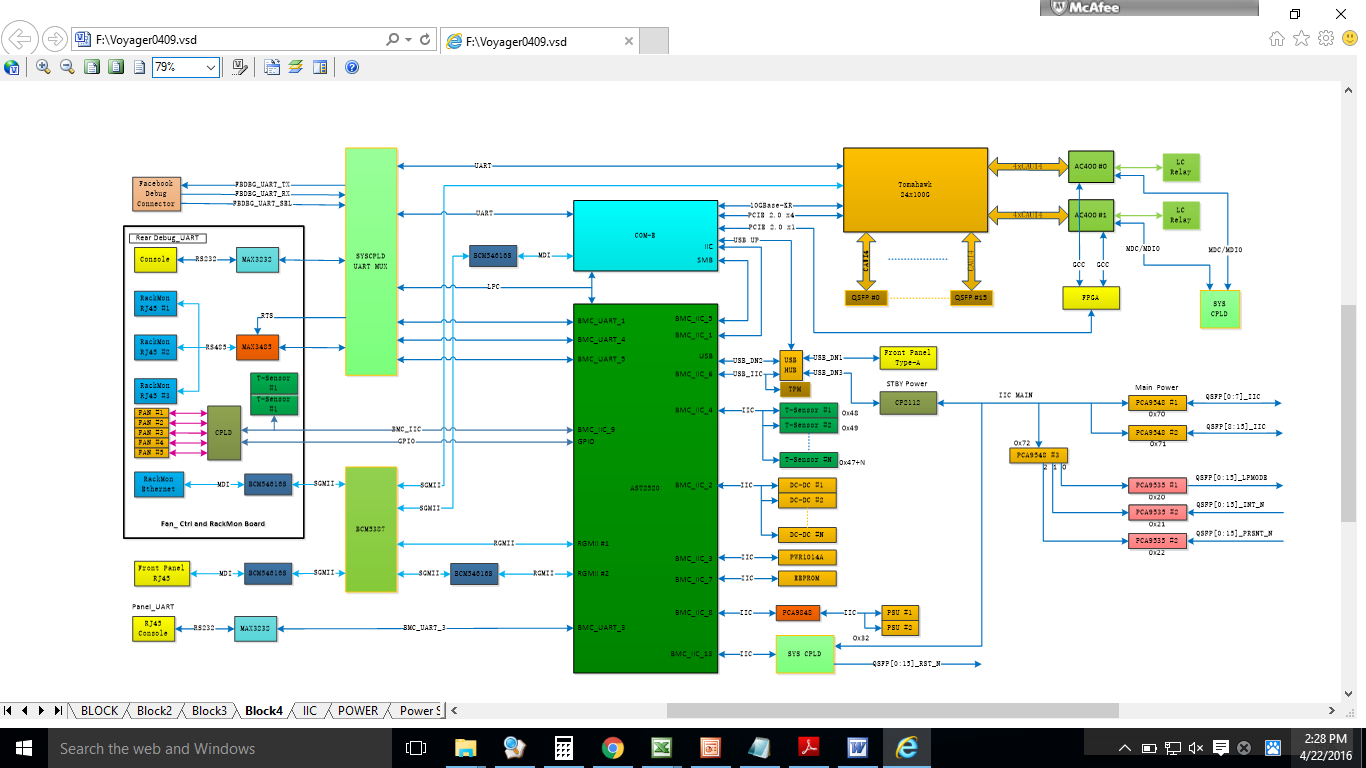
****

Figure . Electrical block diagram of transponder. Note that the CPLD is the control interface to the AC400

## Control Plane Requirements

Transponder application software will run on the COM-E micro-server, which is a Portwell PCOM-B634. The block diagram of this device is shown below in Figure 5. The user guide and specification are described in Appendix A.

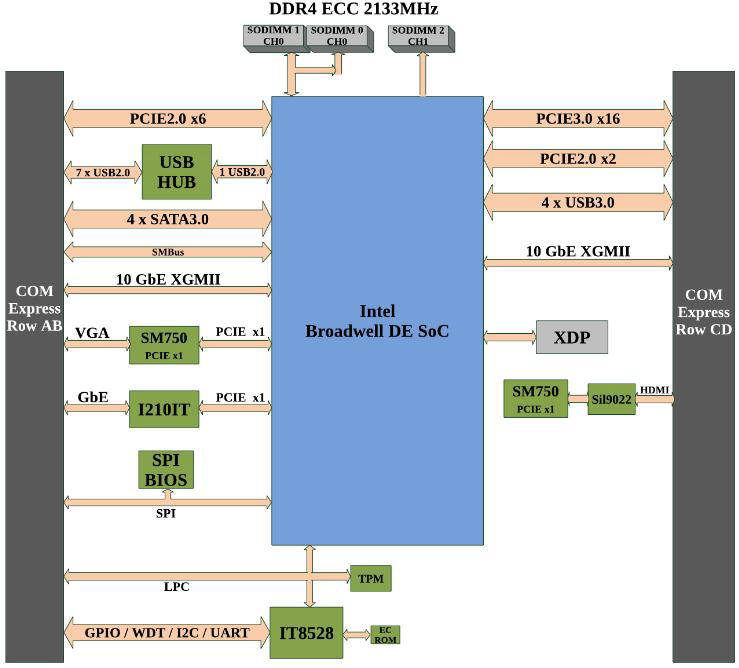


Figure . Block diagram of Portwell COM-E micro-server

The platform is also managed with the BMC chip and OpenBMC. Below are the high level requirements of the COM-E micro-server and the BMC chip

|  |  |  |  |
| --- | --- | --- | --- |
| ID | L3 Requirement | Rational/Comment | Verification |
| : COM-E Server | Transponder shall use a COME server design based on Intel Broadwell-DE processor |  | Design |
| : COM-E Server Part Number | COME server design shall use the following part  Intel part # D1517/25W/4 core/8thread/1.6GHz base Frequency/2.2GHZ Turbo Frequency/LLC 6 MB |  | Design |
| : COM-Server Form Factor | COM-E server design shall be able to use either the Basic or Compact form factor |  | Design |
| : COM-E SSD | The COME Server design shall use a pluggable SSD, part number M.2, 128GB SLC/TLC SSD |  | Design |
| : COM-E Server TPM Interface | COM-E server design shall have an external TPM interface with I2C bus SLB9645 |  | Design |
| 9-2-1: BMC Chip | The BMC design shall use the Aspeed AST2520 | Same as used in Galaxy CMM | Design |
| 9-2-2: BMC TPM Interface | The BMC design shall have an external TPM interface with I2C bus SLB9645 |  | Design |
|  | The Open BMC software shall be used for board management | Available on GitHub |  |

## Development and Implementation Constraints

This document is not a development or project plan and does not address and programmatic or development issues. This document describes the application software requirements for the transponder. However, there are certain implementation requirements that are addressed in this section

It is expected that all key components of the transponder will be based on Facebook Wedge100 Open Compute Project (OCP) switch or will be commercially available components. The transponder will use two Acacia AC400 coherent optical subsystems, and hence the use of two AC400 subsystems per transponder module is considered as a requirement. The specification of the AC400 are incorporated into this requirements document by reference, and specifications of the AC400, including performance, management, and configuration, are implicitly included in this requirements document.

In the case of the aggregation subsystem, the Broadcom Tomahawk chip shall be used in a manner as similar as possible to Wedge100.

The transponder will strongly leverage the Open Compute Project (OCP) Wedge100 switch platform. That means that the mechanical form factor, the power supplies, thermal management, COME-E micro-server, and Board Management Controller (BMC) should replicate the Wedge100. Wedge100 utilizes FBOSS as FW-SW platform. Clearly changes will be required to use this overall platform, but leveraging the platform should be considered a design implementation goal.

The following table describes the characteristics of the different development phases of modules described in this requirements document

|  |  |
| --- | --- |
| Development Phase | First SW Code drop- estimated at 8-9 weeks duration   1. Sw developed in order to meet all necessary requirements 2. Waivers or omissions to requirements documents |
| Completion of test phase | Completion of testing phase- estimated at 6-8 weeks duration   1. Requirements verified 2. No critical bugs or defects 3. Tabulation of defects |
| Revision Phase | Completion of second iteration of code   1. Additional requirements coded and tested/verified 2. No critical bugs or defects |

The transponder software shall be designed to meet the requirements described in this document.

# Meta-Requirements

Meta-requirements are higher-level requirements that serve as “parents” to the underlying detailed software functional requirements

|  |  |  |  |
| --- | --- | --- | --- |
| MR: Transponder Software Meta-Requirements | | | |
| ID | L2 Requirement | Rational/Comment | Verification |
| MR-1 | The transponder network operating system shall be Centos 6 with the Facebook Kernel, Facebook drivers and Facebook SW modules | Facebook to supply the Kernel, drivers and FB modules |  |
| MR-2 | The transponder shall support operational (in-service) and nonoperational (also termed maintenance, administrative, or out of service states). |  |  |
| MR-2 | The transponder shall support the following Operational Modes satisfying different network use cases   1. Wire mode (L2 transparent) 2. Oversubscribed wire mode 3. Packet optical mode |  |  |
| MR-3 | The transponder shall support the following Wire-mode configurations  Transponder configuration: Data plane subtends client port and network ports  Regenerator configuration: Client ports are not used and network port is looped back at AC400 | Configuration and power state of client ports and Broadcom Tomahawk ASIC TBD |  |
| MR-4 | Fault, alarm, and performance management shall be activated in any operational mode. Fault, alarm, and performance management shall be able to be de-activated in a operational mode | Need maintenance mode to be able to set various fault, alarm, and performance management parameters |  |
| MR-5 | There shall be local system management capabilities that can control and monitor all system parameters including   1. Configuration management 2. Provisioning management 3. Fault management 4. Performance management 5. Maintenance management 6. Security management |  | Design |
| MR-6 | There shall be remote system management capabilities access through the DCN that can control and monitor all system parameters including   1. Configuration management 2. Provisioning management 3. Fault management 4. Performance management 5. Maintenance management 6. Security management |  | Design |
| MR-7 | Local management access shall not be required for performing any system function. |  | Design |
| MR-8 | There shall be a one-for-one correspondence between the local and remote management capabilities |  | Test |
| MR-9 | Maintenance mode shall permit maximum flexibility for using internal transponder test capabilities, such as internal PRBS generation, internal BER detection, and various loopbacks within the component subsystems | Minimize the use of external test equipment for system testing and performance assessment |  |
| MR-10 | Network element data models shall conform to YANG models and standards | Need specification |  |
| MR-11 | Network element northbound interfaces shall be API’s that conform to the following protocols   1. Netconf 2. Thrift API 3. Restful API 4. Open-Config API | Need specifications and a matrix of interfaces vs. end-use applications | Design/Test |
| MR-12 | Voyager equipment shall be used in facilities with DCN connectivity and shall be connected to the existing management network. No tunneling of the DCN to remote locations shall be required. | TBC |  |
| MR-13 | All software-firmware upgrades shall be non-service affecting |  | Test |

# Detailed Transponder Functional Requirements

Shown below in Figure 6 is a block diagram of the transponder module. The functional requirements in this section are organized according to the FCAPS model, or Fault, Configuration, Accounting, Performance, and Security. In the case of Voyager there are no accounting requirements. Provisioning is considered part of Configuration. For each major functional requirement there are also sub-requirements on the individual components of the transponder

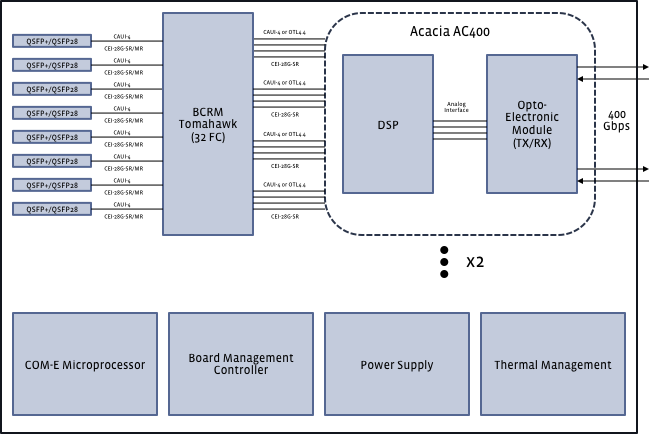


Figure . System block diagram of the transponder module

## Configuration Management

A proper configuration management system performs the following functions:

* Gather and store configurations from network devices (this can be done locally or remotely).
* Simplify the configuration of the device
* Track changes that are made to the configuration
* Configure ('provision') circuits or paths through non-switched or switched networks
* Plan for future expansion and scaling

### System Configuration Requirements

|  |  |  |  |
| --- | --- | --- | --- |
| System Configuration Requirements | | | |
| ID | L3 Requirement | Rational/Comment | Verification |
|  | Transponder module inventory/configuration management parameters shall include: Name, Vendor, PN, SN, HW version, FW Version FPGA version, and any other necessary parameters | Transponder Module shall have inventory fields to denote location in chassis, shelf, rack, row, section of data center (overall location) |  |
|  | The transponder shall be able to be configured into one of the following modes   1. Wire mode (L2 transparent) 2. Transponder configuration 3. Regenerator configuration 4. Oversubscribed wire mode 5. Packet optical mode 6. Maintenance mode 7. Standby Mode |  |  |
|  | While the maintenance mode is considered OOS (out of service) it may be enabled while the system is connected to real traffic sources |  |  |
|  | The maintenance mode may be used to set up specific configuration (also termed a profile) which may be saved to define specific operational configuration states (1,2,3) above . |  |  |
|  | In maintenance mode all parameters from all subsystems are available for configuration |  |  |
|  | Specific configuration states defined by a single command may be used to place the transponder into one of the following operational modes described above (1,2,3). In this case the SW shall not require each individual subcomponent to be individually configured | Individual component configuration described in subsequent sections |  |
|  | Standby mode is used to transition between other modes. In standby mode the AC400 line side optical signals are turned off | Need to understand the AC 400 state sequencing in each transponder mode |  |

### Client Interface Configuration Requirements

|  |  |  |  |
| --- | --- | --- | --- |
| ID | L3 Requirement | Rational/Comment | Verification |
|  | The shall be 12 client ports that can be individually managed and configured |  |  |
|  | Each client optic shall demonstrate management compliance to SFF-8636, “SM-R” (Separable Module Required) requirements only |  |  |
|  | Client ports shall be able to be configured in the following states   1. disabled (power off) 2. enabled   in-service (provisioned) |  |  |
|  | When a client port is enabled the software shall automatically query the client optic and retrieve the inventory/configuration management parameters which shall include: populated/not populated, client optic type, Name, Vendor, PN, SN, HW version, FW Version and any other necessary parameters. This process shall iterate to monitor and retrieve the inventory information if the client optic is inserted after the port is enabled. |  |  |
|  | In wire mode, transponder configuration   1. If two (2) AC400 are populated then client ports 1-8 may be disabled, enabled, or provisioned and 9-12 remain disabled 2. If only one (1) AC400 is populated in position A then client ports 1-4 may be disabled, enabled, or provisioned and client ports 5-12 remain disabled 3. If only one (1) AC400 is populated in position B then client ports 5-8 may be disabled, enabled, or provisioned and client ports 1-4 and 9-12 remain disabled |  |  |
|  | In wire mode, regenerator configuration the client ports are all disabled |  |  |
|  | In Oversubscribed wire mode, packet optical mode, and maintenance mode any client port may be disabled, enabled, or provisioned |  |  |
|  | In enabled state the client optic alarm and performance management is not operational and all historical readings are set to zero. Removal of a client optic when the port is in an enabled state shall not cause alarms. |  |  |
|  | When a client port is moved to a provisioned or in-service state all alarm, fault, and performance management functions shall be activated. |  |  |
|  | For 100GE CWDM-4 client optics, 802.3bj FEC shall be enabled for all Broadcom Tomahawk ASIC SERDES CAUI-4 interfaces |  |  |
|  | For 100GE LR4 client optics FEC shall be disabled for all Broadcom Tomahawk ASIC SERDES CAUI-4 interfaces |  |  |
|  | Software shall permit any client port to be configured as 100GE or 4x10GE |  |  |
|  | Each client port shall be independently configurable to a maximum frame size between 64 and 9212 bytes |  |  |

### Broadcom Tomahawk ASIC Configuration Requirements

These are an incomplete set of requirements for the Broadcom Tomahawk ASIC (abbreviated TH ASIC in the detailed requirements).

|  |  |  |  |
| --- | --- | --- | --- |
| ID | L3 Requirement | Rational/Comment | Verification |
|  | Upon power-up the software shall automatically query the TH ASIC and retrieve the inventory/configuration management parameters which shall include: populated/not populated, chip type, Name, Vendor, PN, SN, HW version, FW Version and any other necessary parameters. |  |  |
|  | In wire mode, transponder configuration the following 100G host (client) and network (line) ports shall be mapped across the TH ASIC SERDES according to the following table   |  |  |  | | --- | --- | --- | | 100GE | TH Host SERDES | TH Network SERDES | | A | 0 | 2 | | B | 1 | 3 | | C | 8 | 10 | | D | 11 | 12 | | E | 16 | 18 | | F | 17 | 19 | | G | 24 | 26 | | H | 25 | 27 | | Needs to be confirmed based on internal layout constraints |  |
|  | Wire mode, regenerator configuration the TH ASIC is not used and is effectively disabled | Assumes loopback in AC400. Tomahawk ASIC should be disabled or powered down to conserve power |  |
|  | In wire mode the port macros shall all be set to LR (line rate mode) | See Broadcom Theory of Operations (TOO) page 501 |  |
|  | In wire mode if 4x10GE interfaces are used the 10GE data streams shall be VLAN tagged with either IEEE 802.1BR or IEEE 802.1Q |  |  |
|  | In oversubscribed wire mode the 100GE and 10GE data streams shall be VLAN tagged with either IEEE 802.1BR or IEEEE 802.1Q |  |  |
|  | In maintenance mode the following BRCM SERDES functions shall be able to be enabled or disabled for each SERDES   1. Unframed PRBS-7. PRBS-23, and PRBS-31 generation and detection 2. Line (or facility) loopbacks in which a received signal can be looped back towards its corresponding transmitter. 3. Diagnostic loopbacks in which a transmitted signal can be looped back towards its corresponding receiver. |  |  |
|  | In any operational mode or the maintenance mode the TH ASIC shall be able to be configured to support one of 3 LLDP snooping modes for each SERDES port   1. Deactivated 2. Snoop and Forward 3. Snoop and Drop | See Appendix C for detailed Snoop and Drop Requirements | Design/Test |

### Network Interface (AC400) Configuration Requirements

|  |  |  |  |
| --- | --- | --- | --- |
| ID | L3 Requirement | Rational/Comment | Verification |
|  | The transponder shall use either one (1) or two (2) Acacia AC400 Modules. Configuration details are described in the following documents   1. Acacia AC400 Hardware Specification Version AC400-001-290 2. Acacia AC400 Software Specification 1.0 | Transponder shall be able to be deployed with one (1) AC 400 module | Design |
|  | Upon power-up the software shall automatically query each AC400 and retrieve the inventory/configuration management parameters which shall include: populated/not populated, optic type, Name, Vendor, PN, SN, HW version, FW Version and any other necessary parameters. |  | Design |
|  | The software shall permit all functional requirements to be met regardless if one (1) or two (2) AC400 modules are populated |  | Design/Test |
|  | The software shall permit a selectable option for either full host control for AC400 turn-up or autonomous control (no host control) for AC400 turn-up. This applies for all operational modes as well as maintenance mode | See Appendix A | Design/Test |
|  | For any operational mode or maintenance mode, the turn up sequence should follow the recommended sequence described in the AC400 SIS | See AC400 SIS Table 8. This requirement may require some timing control |  |
|  | Each AC400 input line shall be hard configured for 100GE input | 100GE client inputs only, so no need to ever configure for OTN input |  |
|  | Each AC400 input lane shall be able to be configured for 100GE client input either with or without IEEE 802.3bj FEC. |  | Design/Test |
|  | Each AC400 shall be user configurable to be in one of the following operational states   1. Two independent 100G QPSK waves 2. Two coupled 150G 8 QAM waves 3. Two independent 200G 16-QAM waves | See table 2 of Acacia SW SIS. This is a restricted set of operational states |  |
|  | For all operational modes except wire mode- regenerator configuration the AC400 crossbar switch shall be configured in the same fixed state. | No reason for operational states to adjust the internal AC400 crossbar switch |  |
|  | For wire-mode, regenerator configuration the AC 400 crossbar switch shall be set in an internal loopback mode so that line side optical I/O A is looped back to line side optical I/O B | Need to reconcile this with regenerator mode. Regenerator mode can be access using BRCM loopback or using AC 400 crossbar |  |
|  | In maintenance mode the crossbar switch configuration shall be user accessible |  |  |
|  | In maintenance mode the loopback capabilities of the AC 400 shall be able to be enabled or disabled. |  |  |
|  | In maintenance mode the following AC400 line side parameters shall be adjustable   1. FEC type 2. Encoding (diff vs non-diff) 3. TX pulse shaping 4. Enable/disable CD pre-compensation 5. Enable/disabled CD post compensation 6. Set auto or manual dispersion search | These parameters all have default values and are used to set up a configuration for an operational state |  |
|  | The following AC400 line side parameters shall be able to be adjusted in standby mode   1. Wavelength: selectable ITU grid wavelength between 1530nm-1565nm on a 50GHz grid as per ITU-T G.694.1. 2. Optical power: See AC400 hardware specifications | Each transponder needs to have wavelength and power adjusted during test and turn-up |  |
|  | The AC 400 line side optical power may be adjusted in any operational mode | Power may able to be adjusted while the system is in operation |  |

## Performance Monitoring Management

Performance monitoring management provides for early detection of service degradation. For each performance monitor (PM), performance thresholds are set, predetermined limits that form the basis for the detection of service degradation, and once these limits are cross a Threshold Crossing Alert (TCA) is generated by the network element. Performance monitoring management include the following functions

* Enable or disable PM’s
* Identify useful PM’s fort the network use case
* Set up PM collection intervals and TCA’s
* Collect historical PM information
* Reset the PM counters
* Export the PM data

For all PM’s the PM parameter name must be defined (precise naming is a implementation detail, not a requirement), and the default TCA must also be defined. For all PM’s, the class must be defined or configured in software.

### System Performance Management

|  |  |  |  |
| --- | --- | --- | --- |
| System Performance Management | | | |
|  | Any PM shall be able to be enabled or disabled |  |  |
|  | There shall be 3 classes of Performance Management parameters with different measurement interval, measurement reporting, and storage requirements  PM Class A (real time) shall support   1. Measurement interval: 1 second 2. Measurement reporting: Max, Min, Ave 3. Storage: real time or 24 hours   PM Class B shall support   1. Measurement interval: 15 Minutes 2. Measurement reporting: Max, Min, Ave 3. Storage: 24 hours   PM Class C shall support   1. Measurement interval: 24 hours 2. Measurement reporting: Max, Min, Ave 3. Storage: 365 days |  | Design |
|  | Each PM shall permit a PM class to be set, but not all PM’s permit all 3 classes |  |  |
|  | Each PM shall permit a TCA to be set. There shall be a default value for each TCA | Default TCA’s TBD |  |
|  | PM’s shall be set to zero when the transponder enters the maintenance mode |  |  |
|  | Historical PM data shall be able to be exported from the NE |  |  |

### Client Interface Performance Management

Client interface performance management is performed on the actual client pluggable optic for some layers and also on the Ethernet switch SERDES for the Ethernet layer parameters. It is important to note that there may be up to 16 client interfaces and they can consist of different client optics (CWDM4 or LR4).

|  |  |  |  |
| --- | --- | --- | --- |
| ID | L3 Requirement | Rational/Comment | Verification |
| Equipment | Client optic equipment PM’s shall include   1. Temperature: TCA H/L, Class A,B,C 2. Voltage: TCA H/L, Class A,B,C |  |  |
| Physical Layer | Client optic physical layer PM’s shall include   1. Tx optical power: TCA H/L, Class A,B,C 2. RX optical Power: TCA H/L, Class A,B,C 3. Laser Bias current, TCA H/L, Class A,B,C 4. Laser Temp, TCAS H/L, Class A,B,C |  |  |
| Ethernet layer | Ingress and egress client performance monitoring shall make per channel statistics counters available in integrated Ethernet MACs including   1. IEEE 802.3 clause 30, Basic, mandatory, and recommended options 2. RFC 2665 (MIB) 3. RFC 2819 (RMON1, Statistics group) |  | Design/Test |
|  | Client layer Ethernet performance monitoring PM’s shall include the following:   * TX and RX undersize packet count * TX and RX Oversized packet count * TX and RX fragment packet count * TX and RX cyclical redundancy check align error count * TX and RX Jabber count * TX and RX jabber seconds * TX and RX total packets * TX and RX Ethernet packets * TX and RX errored octets * TX and RX errored packets * TX and RX multicast packets * TX and RX broadcast packets * TX and RX frame counters * TX and RX packets of 64 or less octets * TX and RX packets of 65-127 octets * TX and RX packets of 128-255 octets * TX and RX packets of 256-511 octets * TX and RX packets of 512-1023 octets * TX and RX packets of 1024-1518 octets * TX and RX packets of > 1519 octets * TX and RX Pause MAC control frames * TX and RX BER Count * TX and RX PCS layer errored seconds * TX and RX PCS layer severely errored seconds * TX and RX PCS layer severely errored sync seconds | TBD if these will all be implemented in initial version of SW. TCA and Class TBD |  |

### Broadcom Tomahawk Performance Management

|  |  |  |  |
| --- | --- | --- | --- |
| ID | L3 Requirement | Rational/Comment | Verification |
|  | Physical layer PM’s shall include   1. BRCM internal temperature: TCA H/L, Class A, B, C 2. BRCM internal voltages: TCA H/L, Class A, B, C |  |  |
|  | For each SERDES interface the 802.3bj FEC shall be monitored and the following PM’s made available   1. FEC corrected errors 2. FEC uncorrected errors |  |  |

### Network Interface (AC400) Performance Management

Network Interface performance management is performed independently for each AC 400 module. It is important to note that there may be one (1) or two (2) AC400’s per transponder. For all PM’s the PM parameter name must be defined, as well as the default TCA. For all PM’s, the class shall be defined in SW

|  |  |  |  |
| --- | --- | --- | --- |
| ID | L3 Requirement | Rational/Comment | Verification |
| Equipment | For each network interface (each AC400) the equipment PM’s shall include   1. Temperature: TCA H/L, Class A,B,C 2. Voltage: TCA H/L, Class A,B,C |  |  |
| Physical Layer | For each network interface physical layer PM’s shall include:   1. TX optical power: TCA H/L, Class A, B, C 2. TX laser bias current: TCA H/L, Class A, B, C 3. TX laser temperature: TCA H/L, Class A, B, C 4. RX optical power: TCA H/L, Class A, B, C 5. RX chromatic dispersion: Class B, C 6. RX differential group delay: Class B, C 7. RX carrier frequency offset: TCA H/L, Class A, B, C 8. RX LO laser bias current: TCA H/L, Class A, B, C 9. RX LO laser temperature: TCA H/L, Class A, B, C |  |  |
| OTN layer | For each network interface the OTN PM’s shall include:   1. FEC corrected errors 2. FEC uncorrected errors | To be confirmed that this is sufficient | Design/Test |

### Platform Performance Management

|  |  |  |  |
| --- | --- | --- | --- |
| ID | L3 Requirement | Rational/Comment | Verification |
| Power Supply | For each Power Supply the PM’s shall include   1. Input voltage: TCA H/L, Class A, B, C 2. Output Voltage: TCA H/L, Class A, B, C |  |  |
| Fans | For each Fan the PM’s shall include revolutions per minute: TCA H/L, Class A, B, C |  |  |
| Temp | The platform PM’s shall include temperature monitoring for each temperature sensor |  | Design/Test |

## Fault Management

Fault management consists of the mechanism to identify, correlate, and correct faults in equipment or the network leading to problem resolution. Fault management consists of

* Alarm indication of equipment faults
* Correlation of alarms to simplify troubleshooting
* Time tagging of alarms
* Alarm clearing

In more traditional telecom equipment the fault management hierarchy is extremely rich and complicated. Given that the Voyager system has few field replaceable parts, , it is expected that a failed unit will be replaced and RMA’ed, simplifying the fault management requirements

### Alarm Severity

A default severity level will define each Voyager alarm. The severity levels are

Critical - this indicates a service affecting condition that usually results in loss of traffic and requires immediate corrective action.

Major - This indicate a service affecting condition that, while not resulting in traffic loss upon assertion, has resulted in service degradation and may result in traffic loss in a very short time frame. A PM that has asserted a TCA will often be a major fault condition.

Minor - this indicates a non-service affecting condition that requires corrective action to prevent the situation from evolving to a major or critical alarm. A PM that has asserted a TCA will almost always be in a minor fault condition.

### System Fault Management

The following general fault and alarm management attributes shall apply

|  |  |  |  |
| --- | --- | --- | --- |
| ID | Requirement | Rational/Comment | Verification |
|  | Alarms shall be defined by the following severity levels   1. Critical- Service affecting and resulting in loss of traffic 2. Major- service affecting and resulting in service degradation and impending loss of traffic 3. Minor- non service affecting but requiring corrective action |  |  |
|  | In any operational mode any alarming associated with and specific fault condition may be able to be deactivated | Default state is all fault management is activated |  |
|  | Fault management shall be able to be deactivated in the maintenance mode, either in a wholesale (global manner) or selectively |  | Design |
|  | A fault or failure shall be detected continuously for 3 seconds or 3 polling periods before the alarm is asserted |  |  |
|  | Once asserted, a fault or failure shall not be detected for 3 second or 3 polling periods before the alarm is de-asserted |  |  |
|  | Fault management thresholds shall have the following trigger thresholds   1. Maximum Major 2. Maximum Minor 3. Minimum Minor 4. Minimum Major   Crossing a minor threshold shall result in asserting a minor alarm. Crossing a major threshold shall result in asserting a major alarm | All alarm thresholds shall have default values. For faults with no minimum value (i.e., temperature) the defaults may be set arbitrarily low (i.e., -99C) |  |
|  | Alarm thresholds can be set so that   1. Minor Maximum and Major Maximum can have the same value, and in this case crossing the threshold shall result in asserting a major alarm 2. Minor Minimum and Major Minimum can have the same value and in this case crossing the threshold shall result in asserting a major alarm |  |  |
|  | Alarms shall be reported with the following general information   1. Object that has the fault or failure 2. Alarm severity 3. Alarm description 4. Alarm counter and timing information |  |  |

### Platform Fault Management

Platform fault management is concerned with the chassis and general HW infrastructure. Listed below are specific supported platform faults and alarms

|  |  |  |  |
| --- | --- | --- | --- |
| ID | Requirement | Rational/Comment | Verification |
|  | The following Power Supply Unit (PSU) alarms shall be made available for each power supply   1. PSU failed 2. PSU feed failed 3. PSU missing | Two (2) PSU’s per system |  |
|  | The following Fan unit alarms shall be made available for each fan unit   1. Fan unit failed 2. Fan unit failed 3. Fan unit speed fault | Five (5) fan units per system |  |
|  | The following software upgrade alarms shall be made available   1. SW upgrade in progress 2. SW upgrade download failed 3. SW upgrade installation failed | TBC. More alarms might be needed for SW upgrade process |  |
|  | The Following DCN alarms shall be supported   1. DCN link failure 2. DCN link configuration error |  |  |
|  | The following facilities alarms shall be supported   1. Client port down 2. Line port down 3. Remote client port down 4. Remote link port down 5. Local-remote port mismatch |  |  |
|  | The following general platform alarms shall be supported   1. Hardware failure- the transponder is not operational [critical] 2. Hardware temperature out of range- the transponder is either too cold or too hot [Major-if major threshold is crossed, Minor- if minor threshold is crossed] 3. AC400 failure- one AC 400 has failed [Critical- if provisioned, Major if not provisioned] 4. SW failure- a critical software module has failed [critical] |  |  |

### Client Interface Fault Management

|  |  |  |  |
| --- | --- | --- | --- |
| ID | Requirement | Rational/Comment | Verification |
|  | The following physical layer client interface alarms shall be made available: [severity level]   1. Transceiver not supported- transceiver is not supported by the Voyager platform [Major] 2. Transceiver communications failure- the transceiver I2C is defective [Major] 3. Transceiver failure- the transceiver has failed [Critical] 4. Transceiver missing- a transceiver is not present on a provisioned port [Major] 5. Transceiver temperature limits exceeded- temperature of transceiver exceeds limits [Major-if major threshold is crossed] [Minor- if minor threshold is crossed] 6. Transceiver provisioning mismatch- A transceiver has been replaced with a new transceiver that is mismatched to the provisioned transceiver settings [Major] 7. Transceiver power limits exceeded- A provisioned transceiver has a received optical power outside of its limits [Major-if major threshold is crossed] [Minor- if minor threshold is crossed] |  |  |
|  | The following client interface signal (facility) alarms shall be made available   1. Loss of signal [critical] 2. Loss of FEC alignment [critical] 3. Loss of RX sync [critical] 4. RX PCS high BER [critical] 5. Loss of RX alignment [critical] 6. Loss of TX sync [critical] 7. TX PCS high BER [critical] 8. Loss of TX alignment [critical] 9. TX local fault [critical] 10. TX remote fault [critical] | Some of these do not have minimum values |  |

### Network Interface (AC400) Fault Management

|  |  |  |  |
| --- | --- | --- | --- |
| ID | Requirement | Rational/Comment | Verification |
|  | Alarm and fault management shall apply separately and independently for each AC400 |  |  |
|  | AC400 fault management shall follow the global alarm logic and hierarchy described in the Acacia SIS | TBD on precise registers to activate |  |
|  | The following HW and physical layer alarms shall be made available for each AC400   1. Voltage [Major-if major threshold is crossed, Minor- if minor threshold is crossed] 2. Temperature [Major-if major threshold is crossed, Minor- if minor threshold is crossed] 3. TX optical power [Major-if major threshold is crossed, Minor- if minor threshold is crossed] 4. RX optical power [Major-if major threshold is crossed, Minor- if minor threshold is crossed] 5. Loss of Signal [Critical] |  |  |
|  | The following facilities alarms associated with the OTN (OTU4) overhead shall be made available for each wavelength   1. Loss of Frame (LOF) [Major] 2. Loss of Multiframe (LOM) [Major] 3. Degraded Defect (DEG) [Major] 4. Alarm Indication Signal (AIS) [Major] 5. Backward Defection Insertion (BDI) [Major] |  |  |
|  | There shall be a BER alarm for each wavelength [Major-if major threshold is crossed, Minor- if minor threshold is crossed] |  |  |
|  | There shall be a FEC uncorrected error alarm for each wavelength [Major-if major threshold is crossed, Minor- if minor threshold is crossed] |  |  |

# Non-Functional Requirements

## System Requirements

|  |  |  |  |
| --- | --- | --- | --- |
| 8: System Requirements | | | |
| ID | L3 Requirement | Rational | Verification |
| 8-1: Cold Start | The transponder module shall meet all requirements within 2 minutes after power-up from cold start |  | Design |
| 8-2: Reset Timing | The transponder module shall have an active reset capability and all requirements shall be met within 60 seconds of de-assertion of a reset event |  | Design/Test |
| 8-3: Firmware Upgrade | The transponder module firmware shall be in-service upgradable with no impact to the data bearing signal traffic. Necessary alarm masking and OAMP (FCAPS) functionality must be disabled during the FW upgrade process. |  | Design/Test |
| 8-4: AC400 Firmware Upgrade | AC400 shall permit firmware-software upgrade. This may or may not be service affecting. Necessary alarm masking and OAMP (FCAPS) functionality must be disabled during the FW upgrade process. |  | Design/Test |
| 8-5: Broadcom Tomahawk Firmware Upgrade | Broadcom shall permit firmware-software upgrade. This may or may not be service affecting. Necessary alarm masking and OAMP (FCAPS) functionality must be disabled during the FW upgrade process. |  | Design/Test |
| 8-6: Firmware Update Timing | All software-firmware updates shall take place within 2 minutes |  | Test |
| 8-7; Internal Polling | Internal polling of all monitoring and control points shall exceed the control actuator requirements by at least 4x |  | Design/Test |
| 8-8: Diagnostic States | The transponder shall be able to placed into a diagnostic state for debugging that selectively disables various control loops and alarm functionality. This is not intended to be used as an operational mode, but for troubleshooting and test and turn-up |  | Design/Test |

## Firmware-Software Upgrade Downgrade Requirements

|  |  |  |  |
| --- | --- | --- | --- |
| ID | L3 Requirement | Rational/Comment | Verification |
| 10-2-1 | FW and SW upgrades shall be able to be performed in operational and administrative states |  | Design |
| 10-2-2 | FW and SW upgrades shall be conducted through local management port and northbound interface |  | Design |
| 10-2-3 | Module shall always store two (2) complete SW package revision images |  | Design |
| 10-2-4 | Any new FW-SW image may or may not include AC400 and/or BRCM Tomahawk changes |  | Design |
| 10-2-5 | Module shall confirm proper download of new SW image. Module shall confirm that SW image is compatible with module |  | Design |
| 10-2-6 | Module shall permit automatic and manual upgrade options for good download image |  | Design |
| 10-2-7 | Module shall support manual downgrade to last known good image if problem with upgrade |  | Design |
| 10-2-8 | Upgrade process shall take < 2 minutes |  | Design |
| 10-2-9 | Module shall have ability to restore to factory defaults |  | Design |

## Flight Recorder Requirements

|  |  |  |  |
| --- | --- | --- | --- |
| ID | L3 Requirement | Rational/Comment | Verification |
| 10-8-1 | The transponder shall support a Flight Recorder log in which entries by users, configuration changes, and all PM’s and alarms are recorded |  | Design |
| 10-8-2 | The Flight Recorder log shall store up to 365 days of data and shall be downloadable via the local console or northbound interface |  | Design |

## Security Requirements

|  |  |  |  |
| --- | --- | --- | --- |
| 10-3. Security | | | |
| 10-3-1 | The transponder shall provide 3 tiers of access, with configurable user rights and capabilities   1. System administrator 2. Privileged User access 3. User Access |  | Design |
| 10-3-2 | The transponder shall provide secure access to its applications, including but not limited to User access profiles, User ID/s and Passwords, and a mechanism to restrict users to the data and transactions necessary |  | Design |

## Regulatory and Compliance Requirements

These requirements are a small subset of the regulatory requirements. These are only relevant to the application software

|  |  |  |  |
| --- | --- | --- | --- |
| 11: Regulatory and Compliance Requirements | | | |
| ID | L3 Requirement | Rational | Verification |
| 11-4: Laser Safety | The Transponder module shall meet the following requirements:   1. Laser Safety: the overall system shall be classified as Class 1M and compliant with the following laser safety regulations 2. IEC/EN 60825-1: 2007 3. IEC/EN 60825-2: 2004 + A1 + A2 4. ANSI Z136.2 5. FDA 21 CFR 1040: Performance Standard for Light Emitting Products |  | Test |
|  | The transponder shall employ eye-safety shutdown mechanisms to comply with these regulatory requirements  This shall be implemented by monitoring the RX\_LOS1 and RX\_LOS2 alarm pin on each AC 400 and employing interrupts to enable SW controlled shutdown of the laser using the TX\_DIS1 and TX\_DIS2 control pins within 500 milliseconds of assertion of LOS. The threshold for this function shall be settable in SW | SW shall employ a multi-reading de-bounce strategy to ensure no false shutdowns. 500 msec TBD |  |

## Reliability and Quality Assurance Requirements

These requirements are a subset of the reliability and Q requirements and these are aimed at transponder application software

|  |  |  |  |
| --- | --- | --- | --- |
| 12: Reliability and Quality Assurance Requirements | | | |
| ID | L3 Requirement | Rational | Verification |
|  | Failure of one AC400 module shall not affect the other AC 400 module | Failures must be isolated | Design/Test |
|  | The vendor shall develop a software quality assurance plan including a defect tracking process to enable the test and verification process | TBD | Analysis |

# Appendices

## State Diagram for Transponder Module

The state diagram for the Transponder Module is shown below in Figure 7. The circles represent transition states and the squares represent static states. The blue lines represent transitions that are autonomous to the module and the black lines represent transitions activated by either human or higher-level management protocol interaction. See Reference 1 for more details.



Figure . Transponder module state diagram

**Power-up and Resets**

The Transponder Module main board must support power-on and reset modes. Furthermore, each coherent module must also support power on and reset modes. Hence there are effectively nested loops that denote the status of the overall module main board and its two AC400 coherent optical subsystems.

**Module Power-up**

Upon module power up the coherent optical modules remain off. Once the main board is initialized it enters the Module OOS state. In this state the module is ready to support all functions, but the AC400’s remain powered off.

**Cold Restart**

A cold restart initializes all hardware on the module and is, thus, service-affecting. The following is a list of conditions that result in a cold restart:

* Module power-up
* Manually initiated cold restart by the user
* A software download

After a cold restart, the boot loader performs the necessary tests and initialization of microprocessor, Flash, SDRAM and support devices. When completed the operating system and applications are loaded. Any other start-up or self-test procedures are performed. A cold restart of the transponder module performs the following actions:

* Clears all alarms and logs
* Clears all collected PM data (current and historical)
* Clears all active loopbacks and forced laser settings
* Releases any active PRBS testing or latency measurement
* Resets and initializes all hardware
* Loads and provisions all hardware

Warm Restart

A warm restart initializes all software on the modules and is not considered service-affecting, however, some functions may be interrupted during a warm restart. The following is a list of conditions that results in a warm restart:

* Manually initiated warm restart by the user
* Non-service effecting software downloads
* Software watchdog timeout

A warm restart restarts and initializes the microprocessor and restarts the software. Since no re-loading, re-configuration, or testing of the HW devices are performed, a warm restart is not service effecting. A warm restart of the transponder module performs the following actions:

* Clears all alarms
* Clears all collected PM data (current and historical)
* All TCAs are cleared as a result of clearing the PM data
* Clears all active loopbacks and forced laser settings
* Releases any active PRBS testing or latency measurement
* Resets and initializes the microprocessor
* Loads and provisions the microprocessor



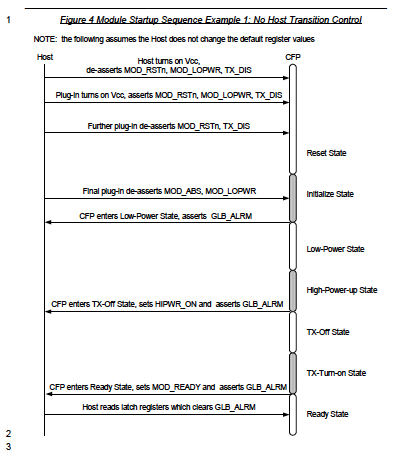


Figure . Host to AC400 module control sequence. This demonstrates full Host control

## Transponder Alarm Management





* ACT: Activation (in the TCM ACT byte)
* AIS: Alarm Indication Signal
* APS: Automatic Protection Switching
* BDI: Backward Defect Indication
* BEI: Backward Error Indication
* BIAE: Backward Incoming Alignment Error
* BIP: Bit Interleaved Parity
* BMP: Bit-synchronous Mapping Procedure
* CAUI: (Chip to) 100Gb/s Attachment Unit Interface
* CSF: Client Signal Fail
* DEG: Degraded Defect
* DM: Delay Measurement
* EXP: Experimental
* FAS: Frame Alignment Signal
* FEC: Forward Error Correction
* FTFL: Fault Type and Fault Location
* GCC: General Communication Channel
* GMP: Generic Mapping Procedure
* IAE: Incoming Alignment Error
* LCK: Locked Defect
* LF: Local Fault
* LFOS: Local Fault Ordered Set
* LOAM: Loss of Alignment Marker
* LOBL: Loss of Block Lock
* LOF: Loss of Frame
* LOFLOM: Loss of Frame and Loss of Multiframe
* LOM: Loss of Multiframe
* LOS: Loss of Signal
* LTC: Loss of Tandem Connection
* MFAS: MultiFrame Alignment Signal
* MSI: Multiplex Structure Identifier
* OCI: Open Connection Indication
* ODU: Optical Channel Data Unit
* OMFI: OPU Multi-Frame Identifier
* OPU: Optical Channel Payload Unit
* OTN: Optical Transport Network
* OUT: Optical Channel Transport Unit
* OTUJ: Optical Channel Transport Unit-J
* PCC: Protection Communication Channel
* PCS: Physical Coding Sublayer
* PLM: Payload Mismatch
* PM: Path Monitoring
* PSI: Payload Structure Identifier
* PT: Payload Type
* RES: Reserved for future international standardization
* RF: Remote Fault
* RFOS: Remote Fault Ordered Set
* RS: Reed-Solomon
* SM: Section Monitoring
* STAT: Status Field
* TCM: Tandem Connection Monitoring
* TIM: Trace Identifier Mismatch
* TTI: Transmitted Trace Identifier

## Detailed LLDP Snoop and Drop Requirements

The aggregation subsystem shall have the ability to process LLDP packets in order to assess what client devices (routers) are attached to each optical port. The LLDP protocol is described in IEEE 802.1AB Link layer Discovery Protocol, Reference 3. LLDP agents typically operate in a one of three modes

1. *Transmit-only mode:* The agent can only transmit the information about the capabilities and the current status of the local system.
2. *Receive-only mode:* The agent can only receive information about the capabilities and the current status of the remote systems.
3. *Transmit and receive mode:* The agent can transmit the local system capabilities and status information and receive remote system's capabilities and status information.

Facebook has two required operational modes for LLDP

1. Snoop and Forward
2. Snoop and Drop

A third operational mode is deactivated (or transparent), in which the optical node ignores the LLDP packet. These modes are described in Figure 9.

Snoop and Forward

In this mode a Facebook Router A sends an LLDP announcement, Facebook optical node A will inspect it, store the data for query and pass it on to Facebook Optical Node Z. Facebook Optical Node Z will forward it on to Facebook Router Z. In the reverse direction, Facebook Router Z sends an LLDP announcement, Facebook optical node Z will inspect it, store the data for query and pass it on to Facebook Optical Node A. Facebook Optical Node A will forward it on to Facebook Router A.

Snoop and Drop

In this mode a Facebook Router A sends an LLDP announcement, Facebook optical node A will inspect it, store the data for query and pass it on to Facebook Optical Node Z. Facebook Optical Node Z will drop the LLDP packet. The LLDP packet payload (TLV parameters) will not be forwarded to the 3rd party router Z. In the reverse direction, the 3rd Party Router Z sends an LLDP announcement, Facebook optical node Z will inspect it, store the data for query and pass it on to Facebook Optical Node A. Facebook Optical Node A will forward it on to Facebook Router A.

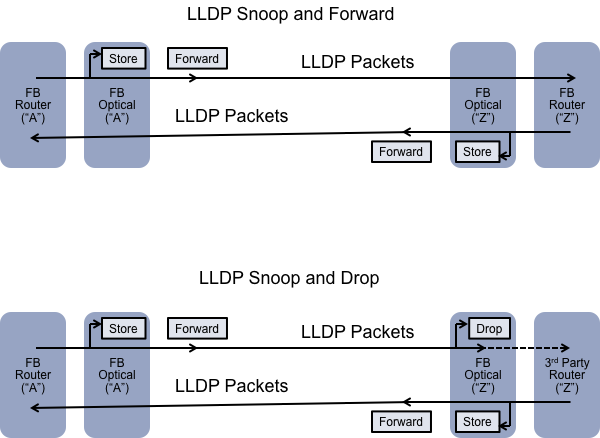


Figure . LLDP packet snooping, showing LLDP Snoop and Forward and LLDP Snoop and Drop

In the case of Snoop and Forward, the network ports are operated in a transparent, or deactivated mode. The host port interrogates the LLDP packet, stores the value in the local optical node, and forwards the LLDP packets to the network port. This information does not need to be stored by the remote optical node Z. However, in a specific design implementation, it remote information from optical node A could be stored in optical node z and vis-a-versa (not shown).

In the case of Snoop and Drop, the network ports are operated in a transparent, or deactivated mode. However, the FB optical Node Z (attached to the 3rd party router) host port is operated in a different state. Now, the in the transmit direction the port is operated in a state where the LLDP packets are dropped, or in a specific design implementation where the TLV payload is set to null values. However, in the receive direction, the host port interrogates the LLDP packet, stores the values in the local optical node, and forwards the LLDP packets to the network port.

LLDP information includes mandatory and *optional* fields listed below:

* System name and description
* Port name and description
* IP management address
* *VLAN name*
* *System capabilities (switching, routing, etc.)*
* *MAC/PHY information*
* *MDI power*
* *Link aggregation*

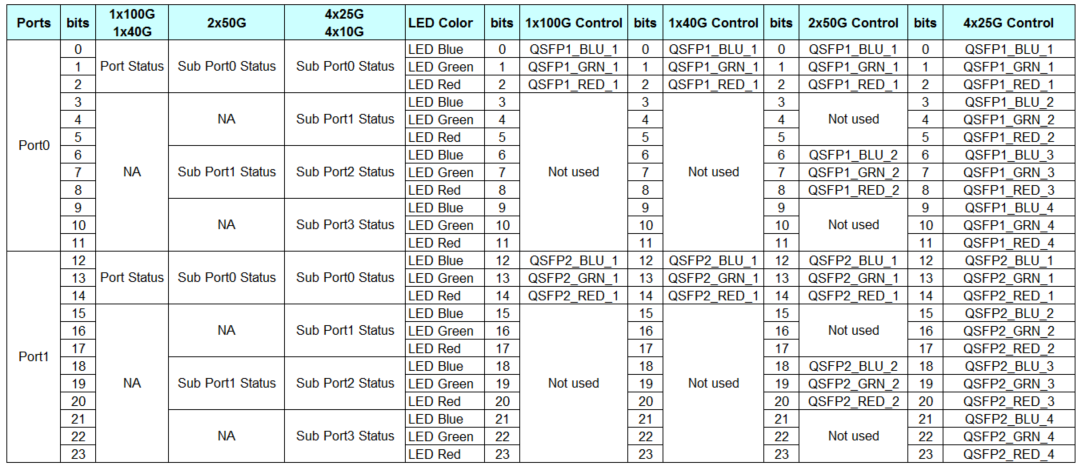
## LED stream format

**LED stream format from Tomahawk- Simplified**

Each QSFP28 port has 12-bit LED stream format, the following table shows the definition. Each line has 3 bit to indicate the RGB status of tri-color LED.

|  |  |  |  |
| --- | --- | --- | --- |
| Port | Bit | Name | Definition |
| QSFP-1  BOT | [0:2] | Lane\_1\_led[0:2] | B-G-R LED status  Blue 100  Green 010  Aqua 110  Yellow 011  Purple 101  OFF 000  Red 001 |
| [3:5] | Lane\_2\_led[0:2] | B-G-R LED status  Blue 100  Green 010  Aqua 110  Yellow 011  Purple 101  OFF 000  Red 001 |
| [6:8] | Lane\_3\_led[0:2] | B-G-R LED status  Blue 100  Green 010  Aqua 110  Yellow 011  Purple 101  OFF 000  Red 001 |
| [9:11] | Lane\_4\_led[0:2] | B-G-R LED status  Blue 100  Green 010  Aqua 110  Yellow 011  Purple 101  OFF 000  Red 001 |
| QSFP-0  TOP | [0:2] | Lane\_1\_led[0:2] | B-G-R LED status  Blue 100  Green 010  Aqua 110  Yellow 011  Purple 101  OFF 000  Red 001 |
| [3:5] | Lane\_2\_led[0:2] | B-G-R LED status  Blue 100  Green 010  Aqua 110  Yellow 011  Purple 101  OFF 000  Red 001 |
| [6:8] | Lane\_3\_led[0:2] | B-G-R LED status  Blue 100  Green 010  Aqua 110  Yellow 011  Purple 101  OFF 000  Red 001 |
| [9:11] | Lane\_4\_led[0:2] | B-G-R LED status  Blue 100  Green 010  Aqua 110  Yellow 011  Purple 101  OFF 000  Red 001 |

Attention: LSB shift out of Tomahawk first



Please note that lane\_1\_led, lane\_2\_led, lane\_3\_led1, and lane\_4\_led are four LED of one QSFP28 port, the status of these four LEDs are from tomahawk LED stream. Each QSFP28 port has four logic LED associated to it. Software can chose to output led status based on port configuration and port status.

For QSFP28 physical port, there are four LED for two QSFP28 port. LED\_A, LED\_B, LED\_C, and LED\_D are real tri-color LED built in QSFP28 2x1 cage. Each QSFP28 port only has two tri-color LED. External LED select button (LED\_SEL) is used to select LED to display. The following are mapping between LED[1:4] and LED\_A/B/C/D

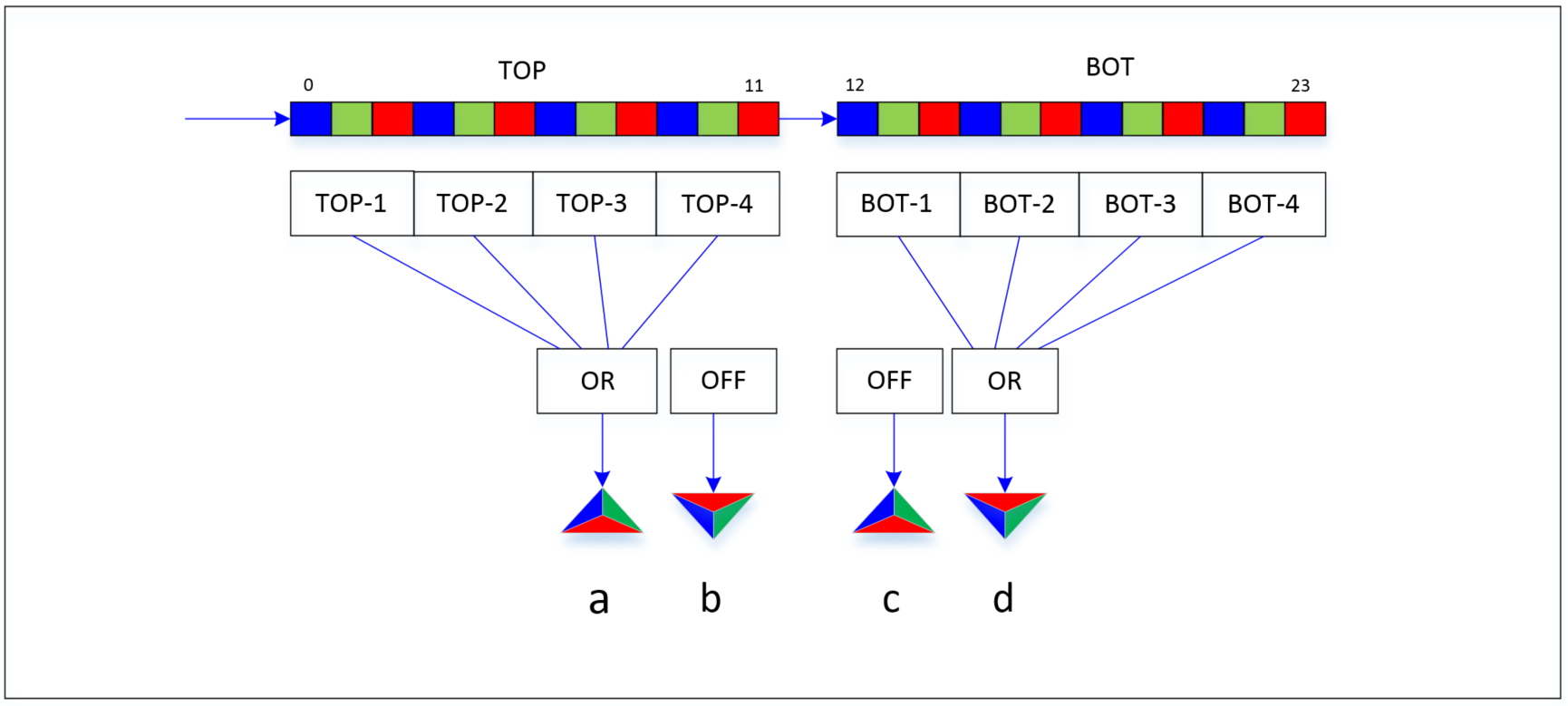
In order to simplify LED micro-code design, only LED A is used for TOP QSFP28 port, and LED D is used for BOT QSFP28 port. LED B and C are OFF

* For 100G mode, LED1 indicates the status of 100G, LED 2/3/4 are OFF, physical LED A is used to display the status of LED1. Color is Blue
* For 40G mode, LED1 indicates the status of 40G, LED 2/3/4 are OFF, physical LED A is used to display the status of LED1. Color is Blue
* For 2x50G mode, LED1 indicates the status of 50G sub-port 0, LED3 indicates the status of 50G sub-port 1, LED 2/4 are OFF, physical LED A is used to display the status of **LED1 OR LED3**. Color is Aqua
* For 4x25G mode, LED1 indicates the status of 25G sub-port 0, LED2 indicates the status of 25G sub-port 1, LED3 indicates the status of 25G sub-port 2, LED4 indicates the status of 25G sub-port 3, physical LED A is used to display the status of **LED1 OR LED2 OR LED3 OR LED4**. Color is Green
* For 4x10G mode, LED1 indicates the status of 10G sub-port 0, LED2 indicates the status of 10G sub-port 1, LED3 indicates the status of 10G sub-port 2, LED4 indicates the status of 10G sub-port 3, physical LED A is used to display the status of **LED1 OR LED2 OR LED3 OR LED4**. Color is Green

CPLD is used to OR all four logic LED together to display on physical led A for top QSFP28 port and on physical LED D for bottom QSFP28 port



Figure : 2x1 stacked QSFP28 connector LED definition



**LED stream format from Tomahawk- Complicated**

Each QSFP28 port has 12-bit LED stream format, the following table shows the definition. Each line has 3 bit to indicate the RGB status of tri-color LED.

|  |  |  |  |
| --- | --- | --- | --- |
| Port | Bit | Name | Definition |
| QSFP-1  BOT | [0:2] | Lane\_1\_led[0:2] | B-G-R LED status  Blue 100  Green 010  Aqua 110  Yellow 011  Purple 101  OFF 000  Red 001 |
| [3:5] | Lane\_2\_led[0:2] | B-G-R LED status  Blue 100  Green 010  Aqua 110  Yellow 011  Purple 101  OFF 000  Red 001 |
| [6:8] | Lane\_3\_led[0:2] | B-G-R LED status  Blue 100  Green 010  Aqua 110  Yellow 011  Purple 101  OFF 000  Red 001 |
| [9:11] | Lane\_4\_led[0:2] | B-G-R LED status  Blue 100  Green 010  Aqua 110  Yellow 011  Purple 101  OFF 000  Red 001 |
| QSFP-0  TOP | [0:2] | Lane\_1\_led[0:2] | B-G-R LED status  Blue 100  Green 010  Aqua 110  Yellow 011  Purple 101  OFF 000  Red 001 |
| [3:5] | Lane\_2\_led[0:2] | B-G-R LED status  Blue 100  Green 010  Aqua 110  Yellow 011  Purple 101  OFF 000  Red 001 |
| [6:8] | Lane\_3\_led[0:2] | B-G-R LED status  Blue 100  Green 010  Aqua 110  Yellow 011  Purple 101  OFF 000  Red 001 |
| [9:11] | Lane\_4\_led[0:2] | B-G-R LED status  Blue 100  Green 010  Aqua 110  Yellow 011  Purple 101  OFF 000  Red 001 |

Attention: LSB shift out of Tomahawk first



Figure : Front panel Port Numbering

LED stream from Tomahawk, bit-0 is shift out first, and bit-191 is shifted out last.

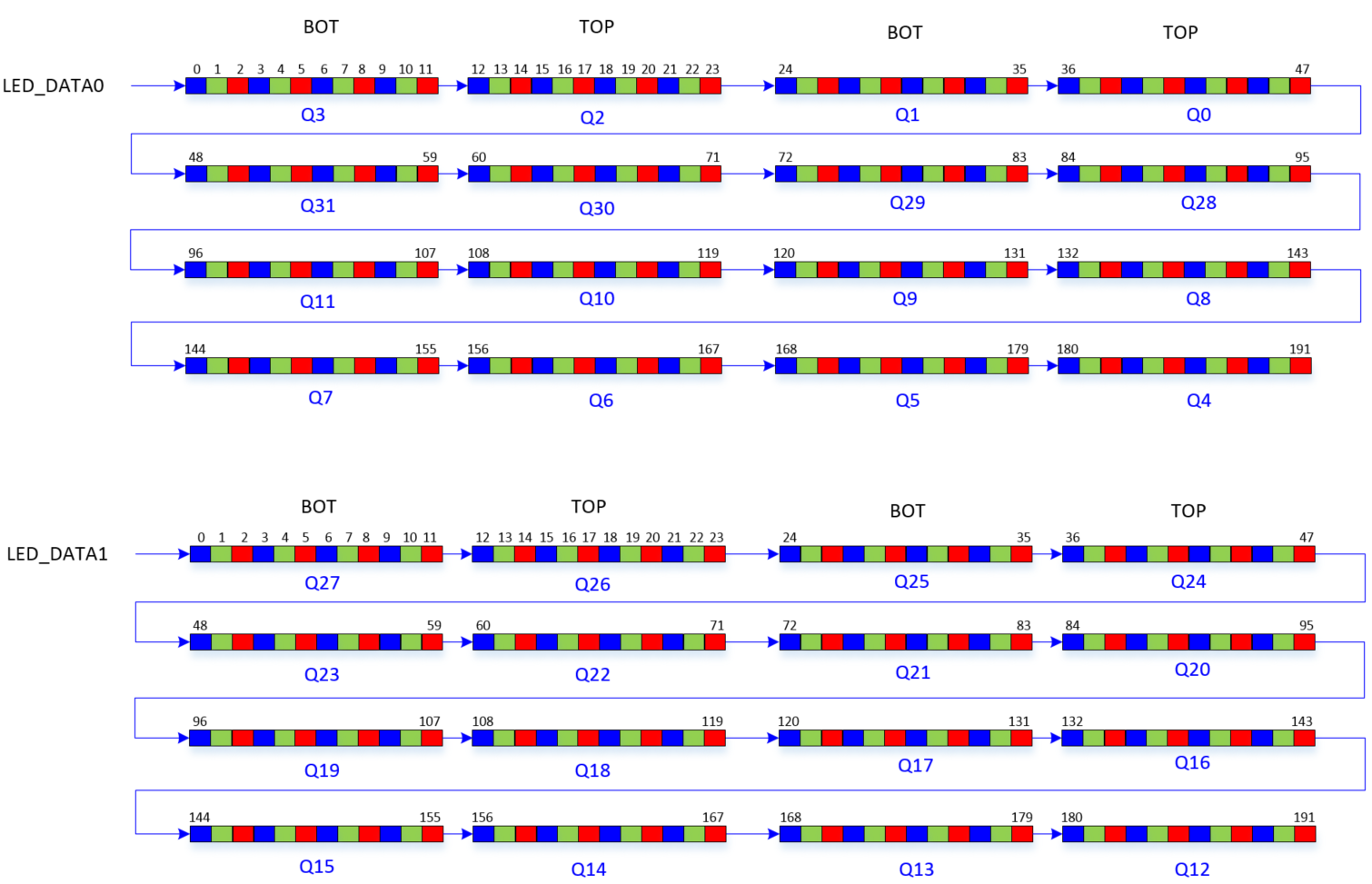


Figure : LED bit stream from tomahawk

Please note that lane\_1\_led, lane\_2\_led, lane\_3\_led1, and lane\_4\_led are four LED of one logic 100G port, the status of these four LEDs are from tomahawk LED stream. Each 100G port has four logic LED associated to it. Software can chose to output led status based on port configuration and port status.

For QSFP28 physical port, there are four LED for two QSFP28 port. LED\_A, LED\_B, LED\_C, and LED\_D are real tri-color LED built in QSFP28 2x1 cage. Each QSFP28 port only has two tri-color LED. External LED select button(LED\_SEL) is used to select LED to display. The following are mapping between LED[1:4] and LED\_A/B/C/D

The CPLD shouldn't care about the mode. It shouldn't ever try to display info for both top and bottom ports at the same time, regardless of the port modes. It should just display either the top or bottom info depending on the button status.

If both top and bottom ports are in 1x100G mode and we want the LEDs to show info about both, then we will handle this in software. Software will just emit a bitstream that has info for both ports in

For instance, if ports 1 and 2 are both in 1x100G mode, software can emit info for both port 1 and 2 in the 12-bit value for port 1, and it can replicate the same data in the 12-bit value for port 2. Then regardless of the button state the LEDs displayed will have info for both ports.

To make things clear, here are some examples showing the bitstreams that would be sent for the first two QSFPs in various modes:

Both in 1x100G mode, both link up:

010 000 000 010 010 000 000 010

Both in 2x50G mode, with link up on eth1/1, eth1/2, and eth2/2, but errors on eth2/1:

010 100 010 010 010 100 010 010

(Assuming that LEDs 1 and 3 are the up arrows and LEDs 2 and 4 are the down arrows.)

Top port in 2x50G mode, with both links up, bottom port in 4x25G mode, with only the last 3 ports up:

010 000 010 000 000 010 010 010



Figure : 2x1 stacked QSFP28 connector LED definition

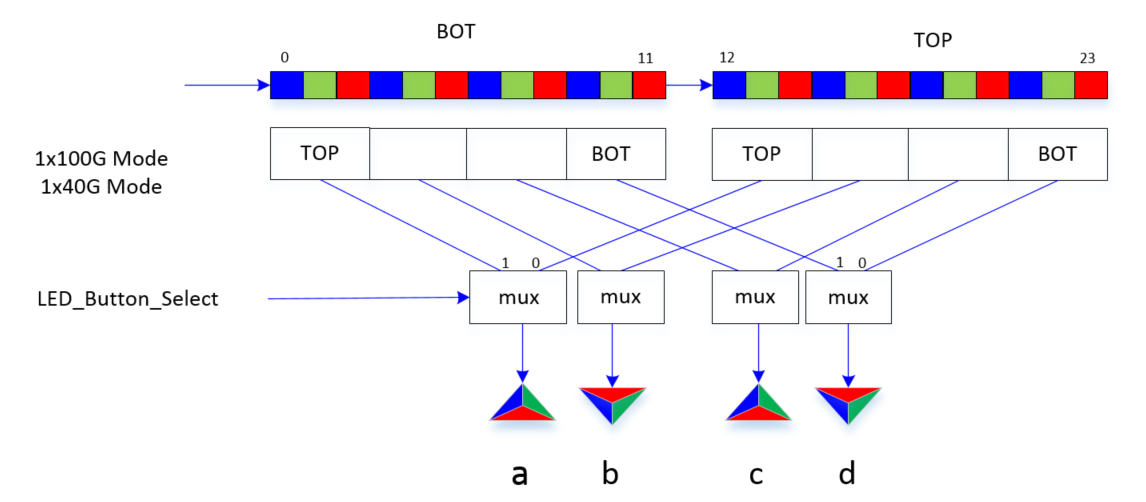


Figure : LED Mux in 1x100G mode or 1x40G mode

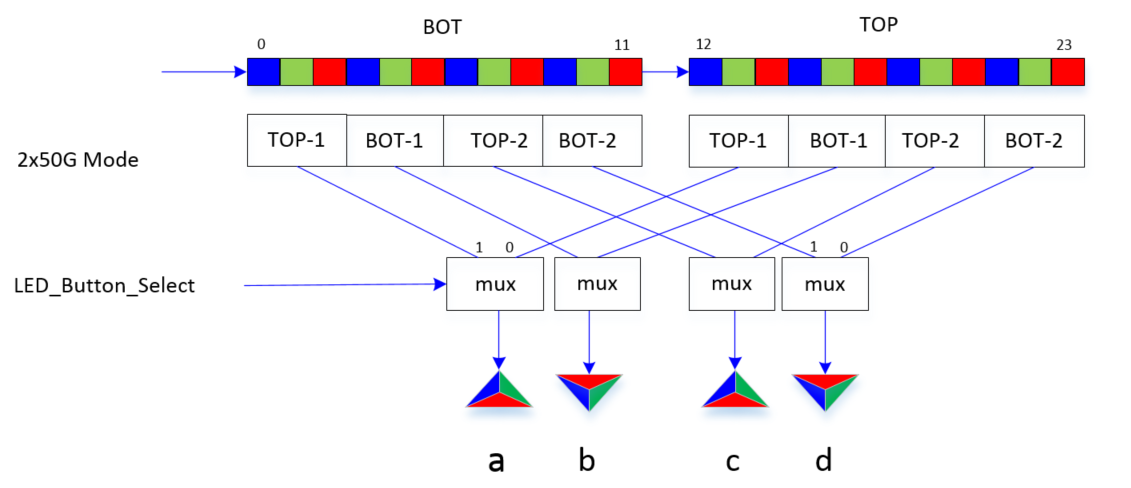


Figure : LED Mux in 2x50G mode

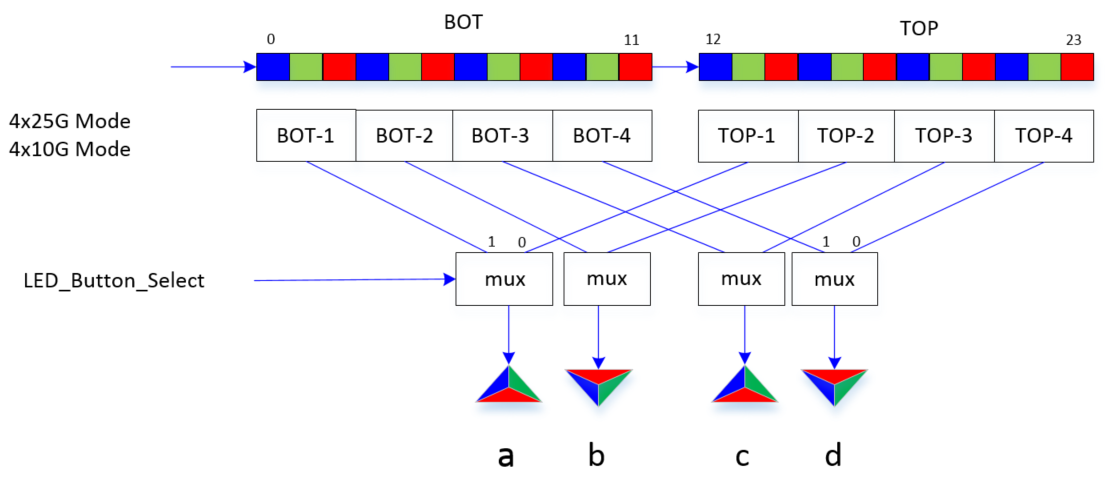


Figure : LED Mux in 4x25G mode or 4x10G mode

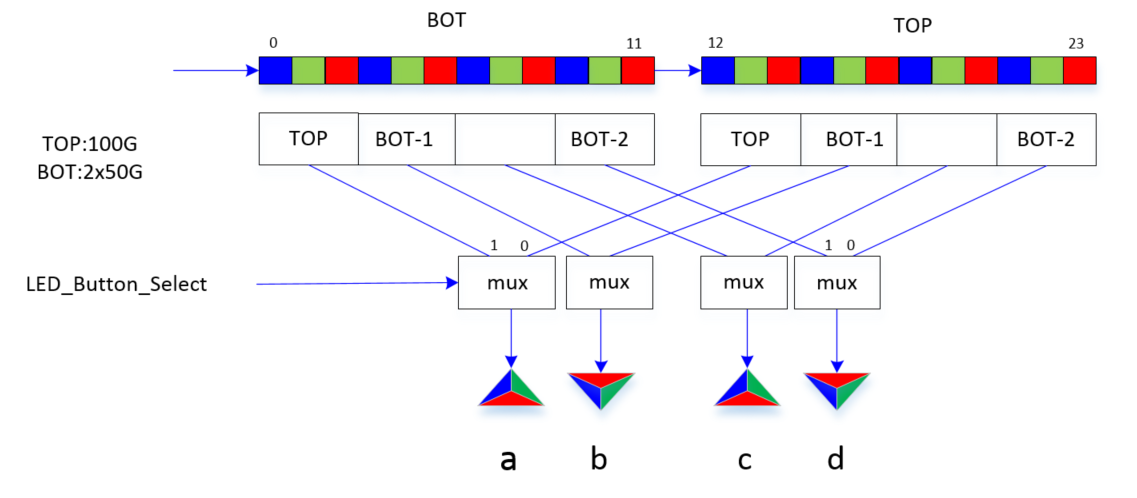


Figure : LED Mux in 1x100G mode and 2x50G mode

**Wedge100 shift register structure**

Please note that the following diagram show the physical shift sequence, not the timing sequence. From shift time point of view, MSB shift out first, for example, LED\_DATA0, Q19-2[7] is shift out first, and Q4-1[0] is shift out last.



**Wedge100 front panel port**



Wedge100 CPLD can support different LED mode

* Port LEDs are directly driven by tomahawk stream: reg\_0x3C[2:0] = 3’b010
* Port LEDs use FB 12bit stream format: reg\_0x3c[2:0] =3’b110
* Port LEDs are driven by CPLD test logic: reg\_0x3c[2:0] =3’b000

Software need to set both bit1 TH\_LED\_EN and bit2 FB\_STREAM\_EN to high to enable the 12-bit format.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Tomahawk LED Control register** | | | | |
| 0x3C | D[0] | 0 | R/W | TH\_LED\_CLR  0: Tomahawk LED stream 0/1 normal  1: Tomahawk LED stream 0/1 reset clear |
| D[1] | 0 | R/W | TH\_LED\_EN  0: Tomahawk LED stream 0/1 disabled  1: Tomahawk LED stream 0/1 enabled |
| D[2] | 0 | R/W | FB\_STREAM\_EN  0: direct stream from tomahawk  1: 12-bit FB stream format used |
| D[3] | 0 | R/W | WALK\_TEST\_EN  0: LED walk test disabled  1: LED walk test enabled |
| D[5:4] | 10 | R/W | 00: stream-0 single led check, led number and color is decided by register 0x3d, constant  01: stream-1 single led check, led number and color is decided by register 0x3d, constant  10: stream-0 all led check, led color is decided by register 0x3d, constant  11: stream-1 all led check, led color is decided by register 0x3d, constant |
| D[6] | 1 | R/W | LED\_TEST\_BLINK\_EN  0: normal constant  1: Blinking during LED test |
| D[7] | 1 | R/W | LED\_TEST\_MODE\_EN  0: normal  1: Port LED test mode enabled |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **TEST LED Control register** | | | | |
| 0x3D | D[4:0] | 00000 | R/W | LED\_TEST\_NUM[4:0]  Indicate one of 32 tri-color LED in one stream  Each QSFP28 port has two tri-color LED, and 16 QSFP28 are in one LED stream. |
| D[5] | 1 | R/W | LED\_TEST\_GREEN  0: green led is disabled in LED test mode  1: green led is enabled in LED test mode |
| D[6] | 1 | R/W | LED\_TEST\_BLUE  0: blue led is disabled in LED test mode  1: blue led is enabled in LED test mode |
| D[7] | 1 | R/W | LED\_TEST\_RED  0: red led is disabled in LED test mode  1: red led is enabled in LED test mode |

LED selection button status is register 0x1B bit-0

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Debug status register 1** | | | | |
| 0x1B | D[0] | 0 | R | TOP\_LED\_ACTIVE\_N 0: TOP LED active 1: Normal |
| D[1] | 1 | R | BOT\_LED\_ACTIVE\_N 0: TOP LED active 1: Normal |
| D[2] | 0 | R | reserved |
| D[3] | 0 | R | reserved |
| D[4] | 0 | R | reserved |
| D[5] | 0 | R | reserved |
| D[6] | 0 | R | reserved |
| D[7] | 0 | R | reserved |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **LED\_TEST MODE** | | | | |
| 0x27 | D[2:0] | 0 | RW | RGB\_DATA\_0\_TOP[2:0] |
| D[5:3] | 0 | RW | RGB\_DATA\_1\_TOP[2:0] |
| D[6] | 0 | RW | reserved |
| D[7] | 0 | RW | CAGE\_LED\_TEST\_EN |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **LED\_TEST Register 0** | | | | |
| 0x40 | D[2:0] | 0 | RW | RGB\_DATA\_0\_TOP[2:0] |
| D[5:3] | 0 | RW | RGB\_DATA\_1\_TOP[2:0] |
| D[6] | 0 | RW | reserved |
| D[7] | 0 | RW | TOP\_LED\_TEST\_EN |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **LED\_TEST Register 1** | | | | |
| 0x41 | D[2:0] | 0 | RW | RGB\_DATA\_2\_TOP[2:0] |
| D[5:3] | 0 | RW | RGB\_DATA\_3\_TOP[2:0] |
| D[6] | 0 | RW | reserved |
| D[7] | 0 | RW | reserved |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **LED\_TEST Register 2** | | | | |
| 0x42 | D[2:0] | 0 | RW | RGB\_DATA\_0\_BOT[2:0] |
| D[5:3] | 0 | RW | RGB\_DATA\_1\_BOT[2:0] |
| D[6] | 0 | RW | reserved |
| D[7] | 0 | RW | BOT\_LED\_TEST\_EN |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **LED\_TEST Register 3** | | | | |
| 0x43 | D[2:0] | 0 | RW | RGB\_DATA\_2\_BOT[2:0] |
| D[2:0] | 0 | RW | RGB\_DATA\_3\_BOT[2:0] |
| D[6] | 0 | RW | reserved |
| D[7] | 0 | RW | reserved |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Stream 0 LED swap Register 0** | | | | |
| 0x44 | D[7:0] | 0x88 | RW | ST0\_LED\_SWAP[7:0] |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Stream 0 LED swap Register 1** | | | | |
| 0x45 | D[7:0] | 0x8C | RW | ST0\_LED\_SWAP[15:8] |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Stream 1 LED swap Register 0** | | | | |
| 0x46 | D[7:0] | 0x88 | RW | ST1\_LED\_SWAP[7:0] |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Stream 1 LED swap Register 1** | | | | |
| 0x47 | D[7:0] | 0x8C | RW | ST1\_LED\_SWAP[15:8] |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Flip Button Register** | | | | |
| 0x48 | D[0] | 0 | W | BUTTON\_FLIP write 1 to this bit will flip the LED selection button |
| D[1] | 1 | RW | reserved |
| D[2] | 0 | RW | reserved |
| D[3] | 0 | RW | reserved |
| D[4] | 0 | RW | reserved |
| D[5] | 0 | RW | reserved |
| D[6] | 0 | RW | reserved |
| D[7] | 0 | RW | reserved |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **LED Stream Control** | | | | |
| 0x49 | D[0] | 1 | RW | SWAP\_UP\_DOWN  0: UP port led status output first  1: Down port LED status output first  In wedge100, SW team choose to output down port status first. |
| D[1] | 0 | RW | reserved |
| D[2] | 0 | RW | reserved |
| D[3] | 0 | RW | reserved |
| D[4] | 0 | RW | reserved |
| D[5] | 0 | RW | reserved |
| D[6] | 0 | RW | reserved |
| D[7] | 0 | RW | reserved |

## References

1. CFP MSA Management Interface Specification, Version 2.4 r06b, June 8th 2015
2. SFF-8636 Specification for Management Interface for Cabled Environments, Revision 2.6, June 19, 2015.
3. Acacia AC400 HW specification: AC400-001-290\_ Product\_Facebok\_Config\_Rev0p1, 2016
4. Acacia AC400 Software Specification: AC400\_Software\_Interface\_Specifiction, Document Number 120-0012-00, Nov 5th, 2015