Broadcom J2/Ramon White Boxes components in **Distributed Disaggregated Chassis Routing System**



AT&T Peng WB HW Group 6/5/2019 DRAFT1-

Compute Project



Large WAN SP-Class Routers Use Cases



Draft1 - 6/4/2019



Traditional Modular Chassis Design for Large WAN SP-Class Routers

Physical Chassis View



Mechanical and Physical Challenge

- **Port Density** \square
- **Cabling Density**
- **Scale Out Growth**
- **Power Density**
- **Cooling Density**

Functional Block View



Other Considerations:

- System Timing Distribution and Synchronization **Via Electrical PCB traces**
- **RP to LC Control and Management path via** internal Ethernet Switched path or can be via the internal Switch Fabric or other methods → Depend on technology and implementation.

Draft1 - 6/4/2019





Broadcom J2 and Ramon White Box Components For DDC

WB1: 2 RU [40x100G QSFP28] I/0 Ports + [13 QSFP-DD 400G] for Fabric Ports – **Need to support 4x10G Breakout mode.**

WB2: 2 RU [10x400G QSFP-DD] I/0 Ports + [13 QSFP-DD 400G] Cell Fabric Connections

WB3: 2 RU [48x400G QSFP-DD] Cell Fabric Connections

WB4: 1 RU [24x400G QSFP-DD] Cell Fabric Connections.



DDC J2/Ramon WB Elements: [WB1, WB2, WB3, WB4]





Major Requirements

- requirements for use in Most of the World (MOW).
 - Physical Dimension: 1RU or 2RU, EIA-19", Max-Depth = 30 in.
 - Front Flush Mount, Rear Mount Sliding adjustable For Post Depth from [26-30in]
 - Front to Back Air flow (Port to Power Direction)
 - Hot-swappable Fans and PSU modules. 1+1 Redundant/ Loadsharing AC or DC PSU High Efficiency 80Plus Gold or better.
 - Circuitry to support 1588V2 and SyncE, T-TC, T-OC, T-BC support within each WB. ** (DDC Synch). Target Class B.
 - OnBoard BMC with dual flash for recoverable remote field upgrade of BMC. BMC is an optional factory installable option.
 - Intel x86 CPU with TPM (No Coin Cell Battery) and dual flash for recoverable remote field upgrade. Factory Option for different CPU models. Broadwell DE {Model1 8 cores, Model2 6 cores, Model3 4 cores}
 - Broadcom [J2 + HBM Deep buffer] + [OP2 for high route Scale] designed to support line rate forwarding on all ports at published minimum packet size.
 - **Broadcom Ramon for Fabric Whitebox.**

The White Boxes in a DDC Routing system are designed to meet TP76200/TP76450 Level 3 requirements operating in a Central Office/Telco environment. Must pass homologation









DDC WBx mounting – Allowing clearance for Cabling & Power

Min 6"

Front

Mounting

NOTE: Cabinet Illustration is to provide a physical mounting context for the application of these WB in AT&T which do not necessarily apply to other customers of service providers.

> The key requirement is that the rear mounting needs to be adjustable.



Side View

Back



High-Level Block Diagram: WB1 [40x100G] I/O + [13x400G] CFC





Silk Screen Numbering Convention:

- Within a grouping
- **1.** {0,1,2,....i}
- **2. Top to bottom**
- 3. Left to right

I/O Ports in Group 0 support 4x10G break out

I/O Ports Numbering Silk Screen

0	2	4	6	8	10	12	14	16	18
1	3	5	7	9	11	13	15	17	19
20	22	24	26	28	30	32	34	36	38

Fabric Ports Numbering

0	2	4	6	8	10	12
1	3	5	7	9	11	

Example of Associated Port Blocked when a Port is operating as 4x10G

0	2	4	6	8	10	12	14	16	18
1	3	5	7	9	11	13	15	17	19
20	22	24	26	28	30	32	34	36	38
X	23	25	27	29	31	33	35	37	39



AT&T - OCP

High-Level Block Diagram: WB2 [10x400G] I/O + [13x400G] CFC





orts Optical	2-digit
er & LED	7 LED

Silk Screen Numbering Convention:

- Within a grouping
- 1. {0,1,2,...i}
- 2. Top to bottom
- 3. Left to right

I/O Ports Numbering Silk Screen

0	2	4	6	8
1	3	5	7	9

Fabric Ports Numbering

0	2	4	6	8	10	12
1	3	5	7	9	11	



AT&T - OCP Draft1 - 6/4/2019

High-Level Block Diagram: WB3 [48x400G] CFC







- Within a grouping
- 1. {0,1,2,...i}
- 2. Top to bottom
- 3. Left to right

Fabric Ports Numbering Silk Screen

0	2	4	6	8	10	12	14	16	18	20
1	3	5	7	9	11	13	15	17	19	21
24	26	28	30	32	34	36	38	40	42	44
25	27	29	31	33	35	37	39	41	43	45





High-Level Block Diagram: WB4 [24x400G] CFC



Craft Console	
OOB/SFP+	24 X 400G QSFP-DD Fabric Ports Grou
Section	



Silk Screen Numbering Convention:

- Within a grouping
- 1. {0,1,2,...i}
- 2. Top to bottom
- 3. Left to right

Fabric Ports Numbering Silk Screen

0	2	4	6	8	10	12	14	16	18	20
1	3	5	7	9	11	13	15	17	19	21





Craft/Console Ports

Craft Type	QTY	Purpose
Micro USB Serial	1	Console
RJ-45 USB Serial	1	Console
USB Port	1	USB data access
RJ-45 10M/100M/1G	1	Ethernet OOB Powered on Standby Pov
10G SFP+ ports	2	For Internal Management Communicat

Console:

Only one Serial input can be active for the Console. Micro USB will have higher priority than RJ45 when something is connected to Micro USB such as a terminal from a technician. Otherwise, the default is RJ45 which would be connecting to a terminal server by default. The Serial console needs to support default selectable between the BMC or the X86 CPU. (customer-provisioned choice).

OOB Management:

The RJ-45 OOB Ethernet management port needs to be operational even when the system is in the shutdown mode. As such it needs to be designed using the standby power rail. It also needs to provide simultaneous connectivity to the X86 CPU and the BMC. The Intel I-210 NIC is specified to use for the RJ-45 OOB Ethernet Management. Design should allow to shutdown Ethernet OOB access to the BMC as needed.

USB port:

USB port can be used to provide access to external USB drive for initial system setup or system rebuild. However, for security reasons, once the system is up and running under the control of the NOS, the hardware/firmware of the box needs to provide a mechanism to turn off access to this external USB port. This lock needs to persist even after a cold boot, warm boot, other reset mechanism. To unlock this would require authentication to access the utility to set the registers to unlock it.

2x10G SFP+:

These two 10G SFP+ ports provide connections to the X86 Host CPU. This will be used for communication with the Route-Engine Controller of the DDC-RS when this is one component of the larger DDC-RS.

wer rail	
ion for the Modules in a large DDC-RS system	



Notes on Cell Fabric Connections (CFC)

1 Cell Fabric Connection = 400G QSFP-DD Optics + Fiber (100m Max) + 400G QSFP-DD Optics

* 1 Cell Fabric Connection = 400G QSFP-DD DAC (1-3m Max)

* 1 Cell Fabric Connection = 400G QSFP-DD AOC (100m Max)

400G QSFP-DD Optic

Fiber (100m Max)

Traffic over Fabric Cell Connections are Fixed sized Cells As opposed to the I/O Port traffic which are Ethernet variable length <u>Packetized format.</u>

An ingress Ethernet Packet from one I/O port on WBx needs to egress an I/O port on WBy, it is broken up into a number of cells and sprayed over multiple Fabric interfaces toward the Egress WBy where it is reassembled before egressing an I/O port on WBy.



400G QSFP-DD Optic

Note:

100m Max is a Broadcom Fabric Latency Constraint





Draft1 - 6/4/2019

Sample Illustration for a Larger Cluster Configuration: Different Sized "DDC Router"





Q&A:

Next session: Timing & Synchronization across DDC



• Q&A:

• Future Discussion: Timing & Synchronization across DDC



Q&A:

Next session: Timing & Synchronization across DDC





Compute Project

