OPEN POSSIBILITIES.

DC-XPI
Datacenter-ready eXtended Peripheral Interface
DC-XPI
Datacenter-ready eXtended Peripheral Interface

Mike Branch, H/W Engineer, Google
Nilesh Dattani, H/W Engineer, Microsoft
2019: OCP Summit

For a successful Modular Building Block Architecture, we need:

- Compute Modules (CPU/Memory/IO) (CMIO-M)
- IO & Accelerator Add-in Card Modules (AIC)
- Security, Control, and Management (SCM)
- Data-plane Control
- An Interconnect

2021: The MBA has evolved to:
Open Accelerator Infrastructure (OAI) and Datacenter-ready Modular Hardware System (DC-MHS)
Datacenter-ready Modular Hardware System

An overview from: [OCP Server Project Monthly Call Presentation on DC-Stack](5/26/2021) for Enterprise, Hyperscale, and Edge datacenter

Hyperstack Hardware Modules:
Logical Blocks overlaid on Physical Blocks for a Datacenter-ready Integrated System (DC-Stack)

DC Environmental Requirements

- **DC-SCM** (Out-of-Band Control)
  - BMC
  - RoT

- **HPM** (baseboard)
  - Compute
  - Security & Control sidebands
  - DC-SCI
  - USB / 13C / 1xPCIe

- **DIMMs**
  - Partners go here

- **CPUs or GPUs, TPUs, xPPUs**
  - OCP tracks go here
  - 1S, 2S, 4S, 8S CPUs
  - Xeon, EPYC, ARM64, ...
  - xPPU Expansion Chassis
  - ... others

- **DC-MIO** (modular IO)
  - Interconnect
  - Form Factors (details here)
  - NVMe requirements
  - Cables & Interfaces
  - DC-XPI (eXtended Peripheral Interconnect)
  - RoT requirements
  - OCP tracks go here
  - SSD
  - IB NICs
  - Accelerators
  - ... others

- **SmartNIC**
  - Dataplane Control
  - Partners go here

- **Others**
  - e.g., Blue Field, Stingray
  - AWS Nitro, MSFT FPGA
  - ... others
Why I/O Modularity?

- Interface speeds have been increasing
  - Increasing mobo material costs and/or
  - Increasing need for re-timers
- Higher power peripherals (requiring additional cabling)
- Increasing # of peripheral shapes to support (CEM, U.2, EDSFF, custom, …)
- Desire for “pay-as-you-go” addition of peripherals
- Increasing # of server platforms; desire to reduce validation time & effort
Datacenter-ready Modular I/O (DC-MIO)

- Packaging approach that separates the motherboard (HPM\(^1\)) from the I/O peripherals
- Allows high-speed I/O connector(s) near the CPU(s)
- Uses I/O Adapters to connect peripherals to the HPM
- System cost reduction opportunities:
  - Reduces motherboard size & cost
  - Allows for cabled and riser-style I/O Adapters
  - Cabled I/O adapters may eliminate need for retimers
- Accommodates multiple peripheral form factors
- I/O Adapters can be installed as-needed based on tray config

\(^1\) Host Processor/Memory Module
Implementation Goals for DC-XPI 1.0

How should this modular interface be implemented?

Goals:

● A high-speed (PCIe Gen5+), high-density connector
● A high-volume connector with multiple sources
● Re-use existing high-volume connector and pinout if possible
● Cable and riser-card support
● Support for x16 (not too concerned with optimizing for smaller width connectors)
● Support (12V) higher-power peripherals without additional cables
● Support a flexible set of sideband interfaces, supporting a wide range of standard peripherals
● Support flexible mounting orientations: vertical/horizontal/coplanar (1U/2U/…)

OPEN POSSIBILITIES.
An Implementation

Datacenter-ready eXtended Peripheral Interface (DC-XPI 1.0)
- SFF-TA-1002 4C+ connector provided the desired speed, density and pin count
  - PCIe Gen6, 0.6mm/<3” length, x16 + sidebands
- Connector already has volumes being driven by OCP NIC & DC-SCM
- Allows for cabled and riser-style I/O adapters
- Created a pinout that supports high power (150W) peripheral(s)
  - Supports 2x 75W CEM cards
- Optional (separate) auxiliary power block to support up to 400W peripheral(s)
- Rich set of sideband interfaces including USB2, USB3, UART, I2C
- Supports individual Presence Detect for I/O Adapter and Peripheral
A New Pinout for 4C+?

Several existing pinout/connector options, including:

- EDSFF / PECFF (4C)
- PECFF-HP-12V (4C)
- OCP NIC 3.0 / PECFF (4C+)

4C+ connector meets most goals, but existing pinouts don’t support:

- High power (150W) peripherals without additional power cables
- A rich set of sideband interfaces including USB2, USB3, and UART
Implementations

OPEN POSSIBILITIES.
Implementations (cont’d)

Example of a front I/O server using Modular I/O w/ vertical DC-XPI connectors (and DC-SCM).

HPM (mobo) PCB pulled back from front of chassis.

Front volume has been divided into four I/O “bays”.

DC-SCM (vertical style)
Implementations (cont’d)

1x16 CEM cabled I/O Adapter

2x8 CEM cabled I/O Adapter

Allows for riser-based I/O Adapters, as well

Multiple vertical DC-XPI connectors across front of HPM

(top view)
Implementations (cont’d)

Two 1x16 Cabled CEM I/O Adapters in an I/O Module
(top view)
DC-XPI Status

The DC-XPI 1.0 spec has been largely completed for productization in 2022.

Similar to DC-SCM 1.0, we hope to gather support and feedback from OCP members which could lead to a second iteration of the spec, i.e., DC-XPI 2.0.

We are targeting the DC-XPI 2.0 spec for use in 2023+ servers, coincident with the DC-SCM 2.0 and DC-MHS 1.0 specs for DC-Stack 1.0
Call to Action

• Adopt the Modular Building Block Architecture using DC-SCM and DC-XPI as the base. They are enabling high-volume designs going into production; take advantage of them in your new designs.
  ○ DC-XPI specification is available at: [DC-XPI rev. 0.9 specification](https://www.opencompute.org/wiki/DC-XPI) (1.0 soon to be released)
  ○ DC-SCM 1.0 specification is available at: [DC-SCM 1.0 Specification Released to OCP](https://www.opencompute.org/wiki/DC-SCM)

• DC-SCM 2.0 specification is currently in revision 0.7; provide feedback to make it better for 2023 products.
  Find it at Hardware Management Module Subgroup: [https://www.opencompute.org/wiki/Hardware_Management/Hardware_Management_Module](https://www.opencompute.org/wiki/Hardware_Management/Hardware_Management_Module)

• Stay tuned for Datacenter-ready Modular Hardware System (DC-MHS) and the Datacenter-ready Integrated System (DC-Stack) built around DC-SCM
Thank you!