DPU Management

Yuval Itkin, Distinguished architect, and
Yigal Edery, Senior Director of Products, Nvidia
DPU overview

- **NVIDIA BlueField DPUs** are SoC data center infrastructure-on-a-chip devices, optimized for traditional enterprises’ modern cloud workloads and high-performance computing
- NVIDIA DPUs are combining NVIDIA’s ConnectX® network adapters with an array of Arm® cores and infrastructure-specific offloads
- Combining a network adapter and Arm® cores mandate managing each subsystem individually
DPU Management interfaces

- Standard-based NIC management sideband interfaces
  - RBT
  - MCTP over SMBus
  - MCTP over PCIe VDM
- Arm® cores management interfaces
  - USB – a Vendor defined interface for managing the Arm® cores
  - UART – a console interface to the embedded OS

All these interfaces are already part of OCP NIC 3.0 specification
  - MCTP over PCIe is an optional requirement in OCP NIC 3.0 specification
DPU Management methods

- NIC management
  - NC-SI
  - PLDM for FW Update
  - PLDM for monitoring and control
  - PLDM for FRU
  - RDE
- Arm® cores management interfaces
  - USB – a Vendor defined interface for managing the Arm® cores
  - UART – a console interface to the embedded OS
  - Redfish Host Interface
- SoC security management
  - SPDM
Arm® cores management

- **OCP NIC 3.0** Pins A68 & A69 are used to provide USB 2.0
- USB → UART is needed on SFF cards
DPU management - Platform Level

- Embedded server OS boot time is always slower than the currently defined OCP NIC 3.0 timing specifications
- Provisioning a bare-metal platform mandates configuring the DPU embedded OS with the right security settings before the tenant OS starts on the host server
- Recovering/re-provisioning a DPU on a bare-metal server shall not happen from the external host
- Trusting the platform BMC for DPU configuration on bare-metal server mandates extra isolation between the tenant on the host and the platform BMC which is not always possible
- While a platform BMC may be untrusted, it is still the owner of the platform’s thermal management, mandating allowing it to monitor the DPU temperature
Arm® cores management

- The Arm® cores cluster is in effect an embedded server within the DPU
- Managing an embedded server mandates:
  - OS Provisioning/Update
  - Full OS recovery
  - OS parameters configuration
  - Boot parameters configuration
  - Security parameters configuration
  - Health and operational state monitoring
  - Reset/NMI the OS from a trusted BMC
  - SOL and CLI
Product/Facility Info

https://docs.mellanox.com/m/view-rendered-page.action?abstractPageId=39257476
Call to Action

• While the HW interfaces are available and defined, there is no standard-defined protocol for managing an embedded server through USB interface.

• **Redfish Host Interface** protocol is defined over network interface. There is a need to complement this definition with new transport using **MCTP over USB**.

• Allowing an OCP server to track the DPU OS state requires a new TBD standard method to allow a BMC to query it over the existing sideband interfaces.

• OCP NIC 3.0 WiKi with latest specifications: [http://www.opencompute.org/wiki/Server/Mezz](http://www.opencompute.org/wiki/Server/Mezz)
Thank you!