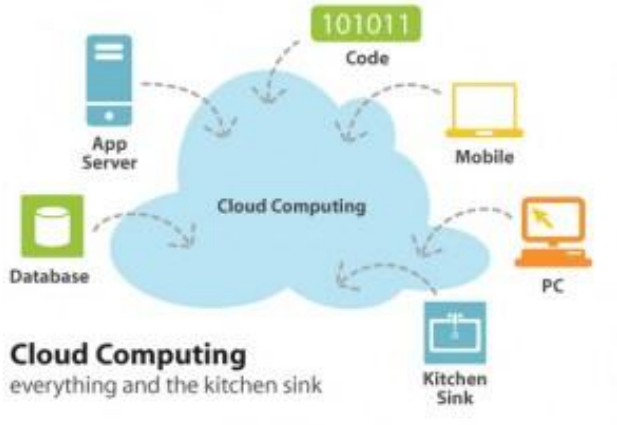
Harnessing PCIe Gen3 Capabilities for Storage Applications

Ashwin Matta
Engineering Director
Cadence Design Systems

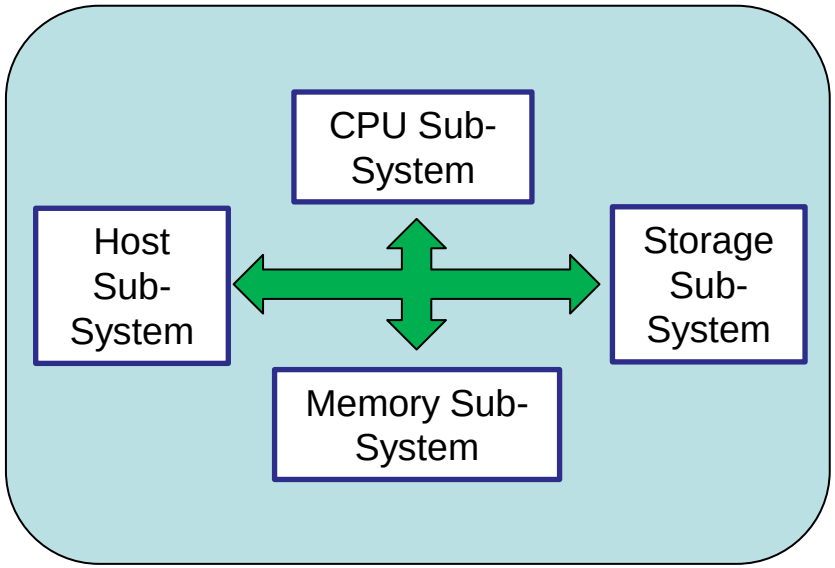
Growing Storage and Bandwidth Needs



Overview

- Generic storage system architecture
- PCIe Gen3 strengths for storage applications
  - Bandwidth and flexibility
  - Single Root IO Virtualization (SR-IOV)
  - Protocol extensions (ECNs)
- Example Configurable Flash Storage System
- Summary

Generic Storage System



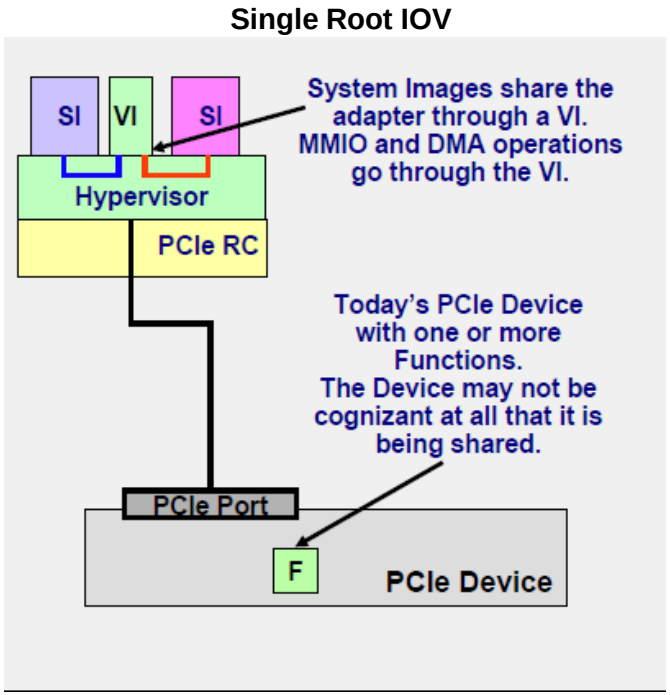
- Host Sub-System
  - SAS, SATA, PCIe, USB3, UFS, other
- CPU Sub-System
  - CPU, UART, Timer, GPIO, etc.
  - Multiple cores per socket
- Memory Sub-System
  - SRAM or DRAM
- Storage Sub-System
  - SAS/SATA-based HDD
  - Flash-based SSD

Advantages of PCIe Gen3 Based Host Interface – Bandwidth/Flexibility

- Higher raw bandwidth
  - Single Gen3 X4 link: 4GB/s
    - ✓ Compared to SATA/SAS 3.0: 600MB/s
  - PCIe with NVMExpress host interface can achieve close to maximum throughput unlike SAS/SATA
- With ONFI 3.0 and multiple flash channels PCIe Gen3 is the only protocol that can keep pace with data transfer rates
- PCIe with NVMExpress supports very large number of outstanding host commands
  - Necessary to support multiple controllers with multiple chip enables
  - SATA/SAS max out at 32 outstanding host commands

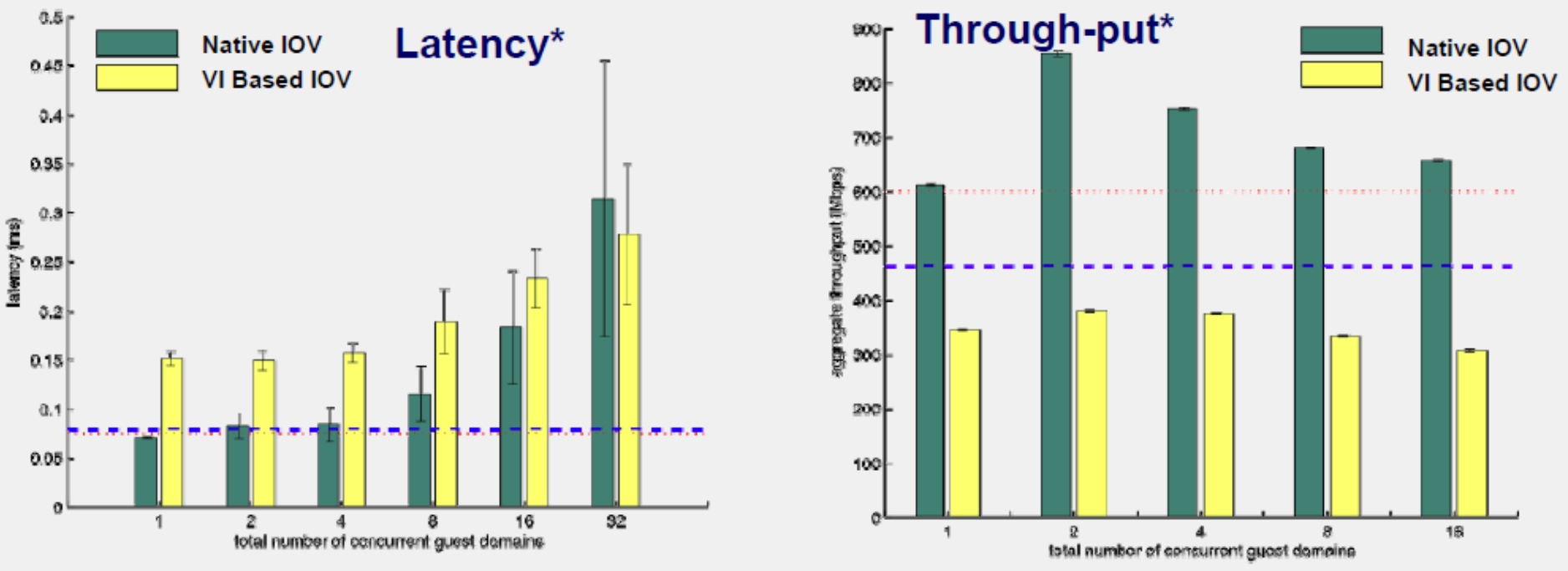
Advantages of PCIe Gen3 Based Host Interface – IO Virtualization (IOV)

- IOV enables multiple system images (SI) on a single storage system
  - SI is a real or virtual system of CPU, OS, I/O, etc. supported through Virtualization Intermediary (VI) or Hypervisor (eg. VMWare hosting Linux/Win32 in a PC)
  - Storage system may not be aware that it is being shared
- Significant improvement in IO performance without overhead of VI in every IO operation



Source: "IO Virtualization and Sharing: PCI-SIG Technical Seminar 2007" – Michael Krause (HP), Renato Recio (IBM)

IO Performance with Virtualization



Source: "IO Virtualization and Sharing: PCI-SIG Technical Seminar 2007" – Michael Krause (HP), Renato Recio (IBM)
Original Source: "Self-Virtualized I/O: High Performance, Scalable I/O Virtualization in Multi-core Systems" - R. Himanshu, I. Ganev, K. Schwan - Georgia Tech and J. Xenidis – IBM

- IO transactions through VI add significant latency through the path of every transaction
- Native IO virtualization nearly doubles throughput compared to VI-based IOV

Advantages of PCIe Gen3 Based Host Interface – ECNs

- Protocol extensions via ECNs
  - Extend native protocol with useful features for special applications
- Fully backwards compatible with PCI, PCI-X, Gen1/2
  - Eg. IOV un-aware software/firmware treats PCIe device with IOV support as base PCIe device

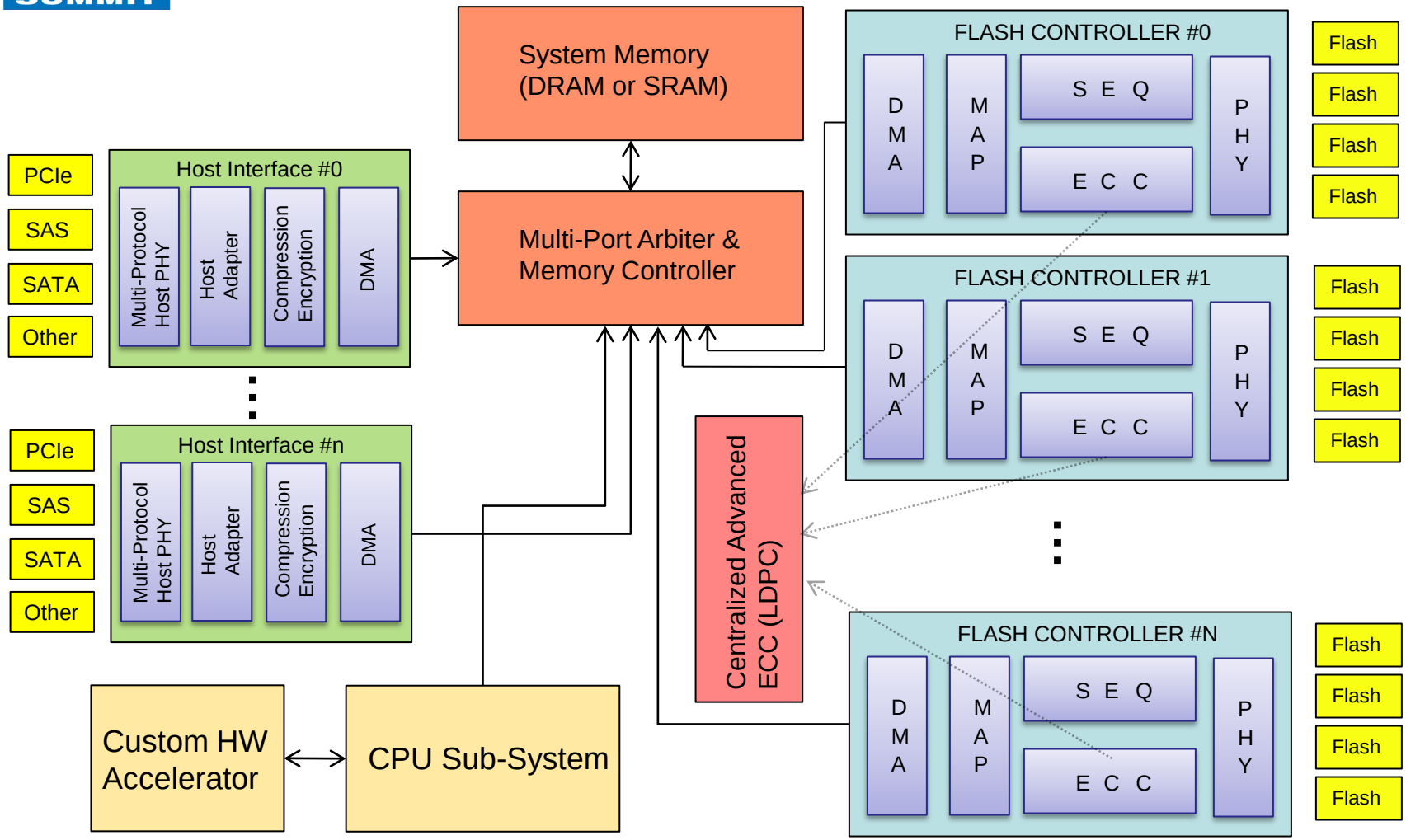
PCIe ECNs Applicable to Storage Systems

- ARI Support
  - Enables IOV with support of up to 256 functions (physical or virtual)
- Multicast
  - Mechanism to broadcast single data set or command to multiple receivers
  - Useful in sending data to RAID or mirrored storage
- TLP Processing Hints (TPH)
  - Hints for optimized PCIe packet (TLP) processing within host memory and system cache

PCIe ECNs Applicable to Storage Systems (contd.)

- Re-Sizable BARs (Base Address Registers)
  - Software selection of BAR aperture size based on system resources/constraints
  - Facilitates creation of adapters for high-end servers and low-end workstations with a wide span of memory requirements
- Optimized Buffer Flush and Fill (OBFF)
  - Based on premise that asynchronous device activity hampers power management of CPU and memory sub-systems
  - Mechanisms for devices to synchronize DMA activity for improved platform power management

Example Configurable Flash Storage System



Summary

- PCIe Gen3 doubles effective bandwidth of data transfer to nearly 1GB/s per lane
- PCIe protocol is ideal for storage transfer needs:
  - Practically unlimited number of outstanding host commands
  - IOV for single storage device to appear as multiple devices
  - Ease of use and optimized operation with new ECNs
- Flash Storage systems utilizing PCIe Gen3 can address growing needs for storage, bandwidth, performance and end-user flexibility