Technical Disclosure Commons

Defensive Publications Series

December 19, 2018

Asymmetric PCIe

Nishant Patil

Asymmetric PCIe

ABSTRACT

Many devices such as accelerators, storage appliances, video transcoding accelerators, etc. have asymmetric bandwidth requirements. For example, the ingress bandwidth for a machine learning accelerator can be more than ten times the egress bandwidth. However, this asymmetry in bandwidth is not reflected in the interconnect. For example, PCI-express typically has as many lanes for host-to-device communications as it does for device-to-host communications. This disclosure presents techniques for asymmetric links between host and device. By reflecting the relative magnitudes of to-and-fro traffic more accurately, the asymmetric PCIe link achieves greater efficiency of communication.

KEYWORDS

Peripheral component interconnect; PCI;, PCIe, PCI-express; OpenCAPI; asymmetric link; host-device communication

BACKGROUND

Many devices such as accelerators, storage appliances, video transcoding accelerators, etc. have asymmetric bandwidth requirements. For example, the ingress bandwidth for a machine learning accelerator can be more than ten times the egress bandwidth. However, this asymmetry in bandwidth is not reflected in the interconnect. For example, PCI-express typically has as many lanes for host-to-device communications as it does for device-to-host communications.

This is illustrated in Fig. 1, which shows a sketch of a PCI-express ribbon, with symmetric bidirectional bandwidth.

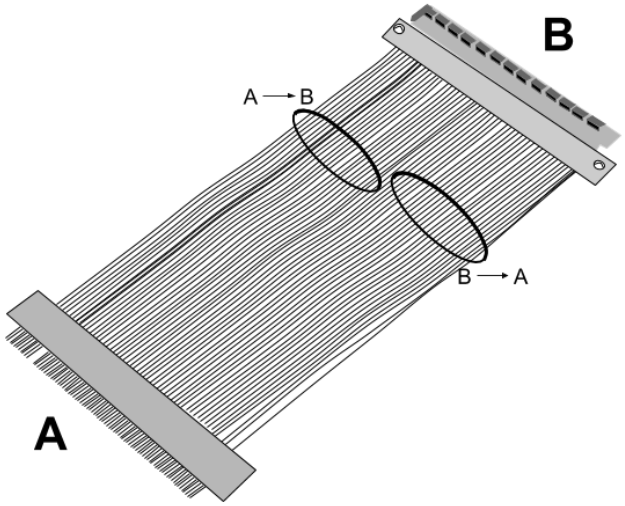

Fig. 1: PCIe link with symmetric bidirectional bandwidth

Due to link symmetry in the PCIe, it is often the case that a queue of data builds up at one end whereas lanes in the opposite direction carry no data.

DESCRIPTION

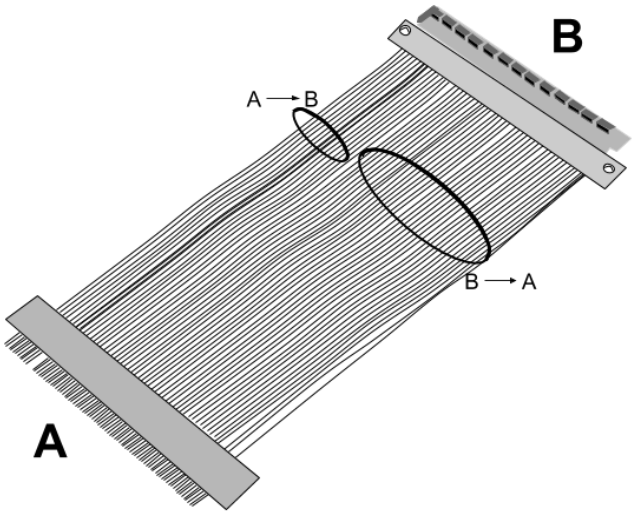

Fig. 2: Asymmetric PCIe

Per the techniques of this disclosure, and as illustrated in Fig. 2, PCIe (or other interconnect standards) with ends A and B, is configurable to have the number of A→B lanes

unequal to the number of B→A lanes. The number of lanes in a direction corresponds to the magnitude of traffic in that direction.

Asymmetric PCI is achieved by reconfiguring more than half the transceivers on one end as receivers, and reconfiguring the corresponding transceivers at the other end as transmitters. The reconfiguration can happen on a per-application basis, e.g., in a GPU application, there may be more data going from host to device, whereas in an FPGA application, it may be the other way around. The reconfiguration can happen at boot-time, at virtual machine migration time, etc. Reconfiguration can also occur dynamically, provided the link is re-trained (which typically takes several seconds).

In this manner, techniques of this disclosure improve the efficiency of bidirectional communication over a bus (and thereby the performance of applications) without requiring any increase the raw transfer rate of the link.

CONCLUSION

Many devices such as accelerators, storage appliances, video transcoding accelerators, etc. have asymmetric bandwidth requirements. For example, the ingress bandwidth for a machine learning accelerator can be more than ten times the egress bandwidth. However, this asymmetry in bandwidth is not reflected in the interconnect. For example, PCI-express typically has as many lanes for host-to-device communications as it does for device-to-host communications. This disclosure presents techniques for asymmetric links between host and device. By reflecting the relative magnitudes of to-and-fro traffic more accurately, the asymmetric PCIe link achieves greater efficiency of communication.