Photonic HPC Clusters for the Oil & Gas Industry

Key Benefits

- Accelerates a wide range of energy-industry workloads
- Meets entry-level, density-focused, and accelerator-focused requirements
- Fully disaggregated solution allowing for machines to be created and attached to HPC workloads
- Use of photonics removes physical rack locality requirements
- Better TCO by decoupling GPU resources from server upgrade path. Upgrade GPUs independent of CPU upgrades
- Lower latencies by using a direct connect photonic fabric versus stacked hierarchy of switching
- Efficiency – Grouping resources via direct connect resolves the stranded resource challenge
- Security – Workloads do not transgress over racks, spines and cores

Data centers are fast approaching an inflection point, where the performance gains from electrical interconnects diminish and the power consumption required to pump electrons through copper continue to escalate. We are now at the point where photonic interconnects can begin to be introduced to the data center. Over the next few years, the introduction of *N*xCPO chiplets will allow for the building of denser and lower cost photonic fabrics for the data center. Fortunately for HPC users, the initial system platform proofs are now available from Drut Technologies for testing and training. These photonic system platforms will not take the place of large super computers today, but over time they will find their place and grow in scale, in the same manner described by Clayton Christensen in the book *The Innovator's Dilemma*.

> "The early users of hydraulic excavators were, in a word, very different from the mainstream customers of the cable shovel manufactures – in size, in needs, and in the distribution channels through which they bought. They constituted a new value network for mechanical excavation…Hydraulics technology ultimately did progress to the point where it could address the needs of mainstream excavation contractors. That progression was achieved, however, by the entrant companies, who had first found a market for the initial capabilities of the technology, accumulated design and manufacturing experience in that market, and then used that commercial platform to attack the value networks above them. The established firms lost this contest." See pages 66 and 71.

Drut's photonic fabric allows for HPC solutions to be deployed to a broader group of users by providing a higher rate of resource utilization in a small footprint, versus a lower rate of utilization across a larger infrastructure footprint. Moving away from InfiniBand and stacked hierarchies of switches with a better architecture allows resources to be directly connected and grouped by machines. This is how organizations will begin to deploy photonic platforms for small HPC clusters today. By decoupling resources from the server with a fully disaggregated photonic fabric, users will add the ability to dynamically attach and detach resources to machines as well as upgrade resources independently of CPU upgrades.

200 Innovative Way, Suite 1360
Nashua, NH 03062

PCIe over Photonics

The introduction of PCIe over Photonics by Drut Technologies enables new design architectures for the data center from the edge to rack scale solutions. The underlying technology components of PCIe over Photonics are composed of (i) Fabric Interface Cards with 4xQSFP28 pluggable optics, (ii) software-based Fabric Manager (FIC), (iii) open-source based hypervisor Software Platform for applications, (iv) Photonic Resource Unit (PRU) for PCIe based resources cards and (v) an all-photonic, low-latency switch (PXC).

Latency, Distance, Complexity and Stranded Resources

The introduction of a photonic fabric at the server or machine level, provides several benefits for data center resource deployment. We begin by introducing a low-latency photonic network that contains a transport protocol that enables a software-based resource Fabric Manager to attach and detach resources from machines (i.e., servers) and a Photonic Resource Unit (PRU) to host resources. This is done independent of traditional networking solutions such as Ethernet, InfiniBand and Fibre Channel. By utilizing a secondary all photonic fabric dedicated to resource pools, several high benefits are realized:

- Latency: By directly connecting machine resources across racks, rows and pods, an all-photonic network absent of network hops provides the lowest possible latency.

- Distance: By implementing PCIe over Photonics, large data center facilities with infrastructure deployments in adjacent racks, rows, rooms, and adjoining facilities can now be unified avoiding the need for multiple network hops. For edge service providers with fiber networks, additional resources can be directly mapped into the photonic network resource pool and controlled by the software resource manager.

- Complexity: Direct allocation of data center resources into software defined bare metal resource pools is the simplest form of resource control and allocation. Important resources such as GPUs, FPGAs, IPUs and Storage can be built at the rack level and used to augment existing as well as new server machines. Resource racks can then be used across the data center to augment important workloads and provides an upgrade path for resources that is independent of a tightly coupled machine.

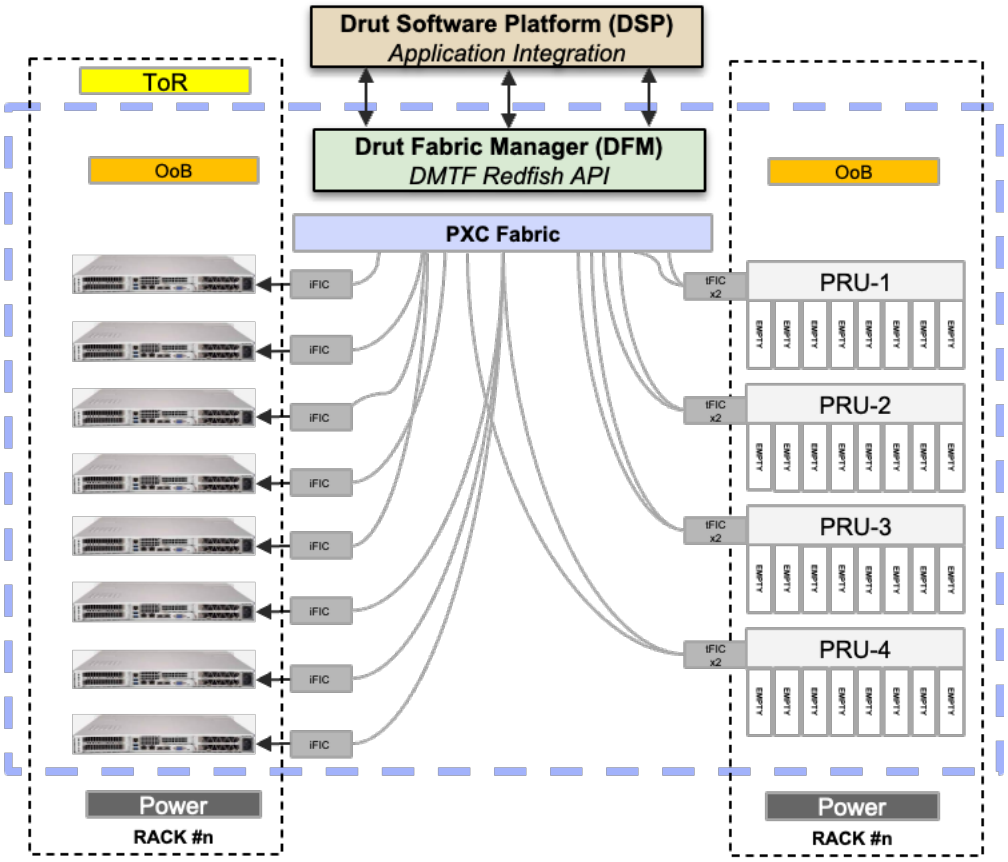
- Resource Utilization: By using a dynamic photonic fabric and a software defined Fabric Manager to allocate resources, workloads can now be connected to specific bare metal resources and defined for specific workloads. This method of resource allocation is far more efficient than using a traditional network fabric to connect resources in various machines across racks, rows and pods.

Ordering Information

Product:	FIC 1000
Code:	DRT-HW-iFIC-1000
	DRT-HW-tFIC-1000
Description:	Fabric Interface Cards
Product:	PRU 1000
Code:	DRT-HW-PRU-1000
Description:	Photonic Resource Unit (PRU)
Product:	PXC – 64 Port
Code:	DRT-HW-PXC-FABRIC-64P
Description:	Photonic fabric including switch, cables and optics.
Product:	Drut Fabric Manager (DFM)
Code:	DRT-SW-DFM-STD-RTU
Description:	Photonic fabric manager
Product:	Drut Software Platform DSP)
Code:	DRT-SW-DSP-PRO-RTU
Description:	Application Hypervisor Platform

Design Summary

The small two rack design below, illustrates how eight (8) Supermicro SuperServer 1019GP-TT single socket machines are assembled into cluster with a disaggregated pool of 32 double width slots or 64 single width slots. These resource slots can be filled with GPUs, FPGAs, DPUs, IPUs and storage.



Sample HPC Photonic BOM

- 8 x Supermicro SuperServer 1019GP-TT
- 8 x FIC 1000s for Compute Machines
- 4-8 FIC 1000s for PRUs
- 4 x PRU 1000
- 64 Single Width Slots
- 32 Double Width Slots
- Bring your own resources (BYOR)
- 1 x PXC (64 Port)
- DFM for 64 Ports
- DSP (Optional) for 8 x Machines
- *N*xOptics (4x25G to 100G QSFP28)
- ~38 RU (Design Flexibility)
- Bring your own ToR Switch

200 Innovative Way, Suite 1360
Nashua, NH 03062