Dataflow-centric Warehouse-scale Computing

Hidemoto Nakada, Tomohiro Kudoh, Ryousei Takano, Isao Kojima,
Tsutomu Ikegami, Satoshi Sekiguchi
National Institute of Advanced Industrial Science and Technology (AIST)

Requirements for bigdata processing is keep expanding in its velocity, variety and volume, due to the rapid increasing and improvement of the data detection devices, including high-resolution monitoring cameras, smart phones, monitoring sensors. Furthermore, multimodal analysis that incorporate data streams from multiple source, such as analysis using geo-information and SNS data, will require further processing power. Next generation supercomputers have to handle large and heterogeneous data with minimum power-consumption.

While bigdata analysis is performed with ordinary server computer in ordinary data center, they are essentially not designed for bigdata analysis . Let us introduce a measure called IPR (I/O and Processing ratio) which denotes balance between data I/O performance and server processing power. IPR is defined as follows:

IPR = I/O performance (Bytes/sec) / Processing Performance (Bytes /s)

As an example, consider a conventional server computer that is capable of processing 240 GFLOP in double precision. The Processing performance is 5760 Gbytes/s. The I/O performance of the server will be around 5 GBytes/s. The IPR is less than 1/1000. Clearly, the I/O performance is too low for processing incoming data.

This is because the conventional computer architecture is not suitable for large amount of data. The incoming data go through I/O bus such as PCI Express and will be once stored in the memory. Then it will be copied to the cache and then passed to the ALU. These excessive steps slows down the processing. On the other hand, network speed in data center could be drastically speed up with using WDM (Wavelength Division Multiplex) technology, as shown in Figure 1. This means we have to device datacenter architecture better suited to bigdata stream processing.

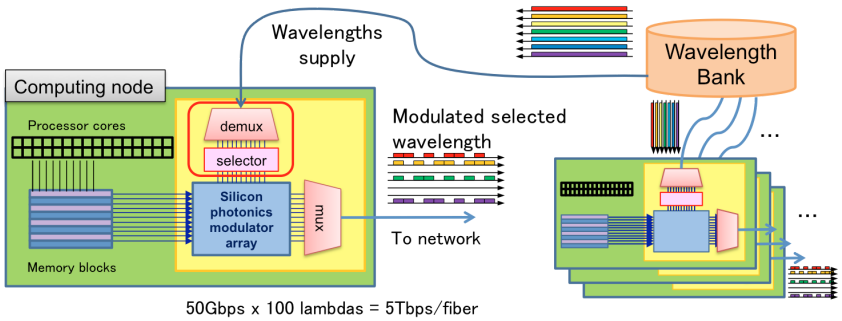

Figure 1. Wavelength Bank under development in AIST.

We propose a data-centric architecture for datacenter. The main idea is to **explicitly define data flow** in a datacenter and set up computation modules along the data flow. The data flow will be established as dedicated optical network path. The architecture is composed of DPCs (Data Processing Components) connected with congestion-free optical network (Figure 2). The red lines in the figure stand for the data flow. Each DPCs could be implemented in several way; it could be specially designed hardware such as a server with direct cache data injection, FPGA attached inline data processing storage unit, or it could be implemented with conventional server with dedicated software.

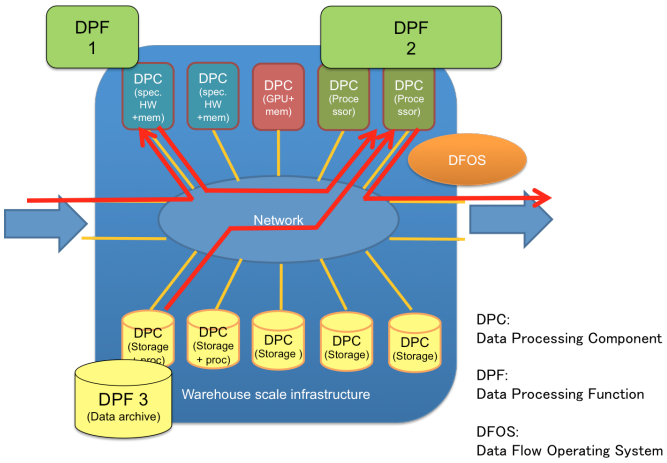

Figure 2. Overview of the Data-centric Architecture

Logical data processing capability is called Data Processing Function (DPF) . DPFs are managed and mapped onto DPCs by the resource management layer called DFOS (dataflow OS). DFOS selects DPC based on required IPR.

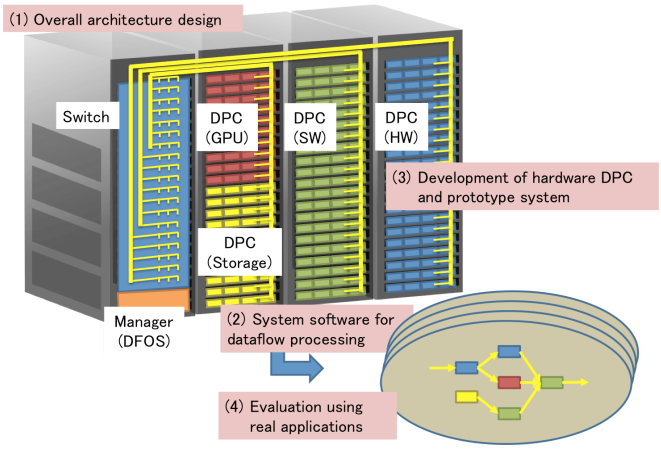

Figure 3. System Overview.

Figure 3 shows the implementation image of the proposed architecture. The leftmost rack stores optical switches and the management node. The other three racks houses DPCs.