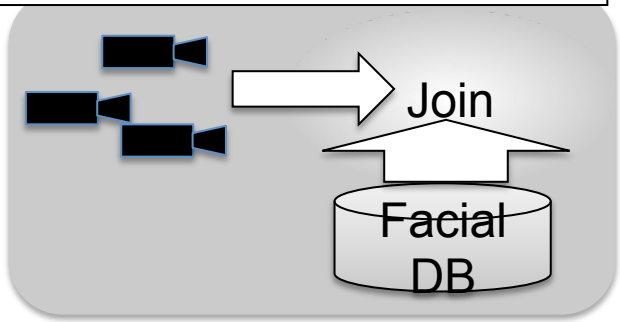
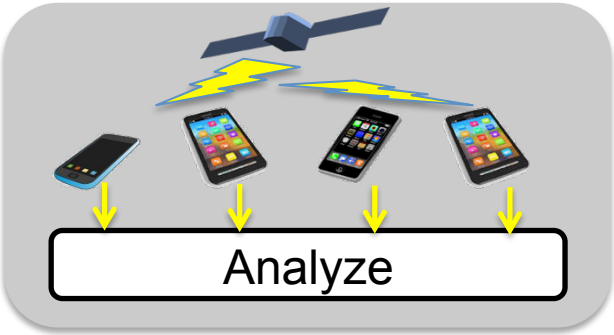
Dataflow-centric Warehouse-scale Computing

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

Real-time Bigdata Processing in 2020

- More data will be processed to get better results
  - Video Monitoring
    - # of cameras = 0.1M
    - Joining with Facial feature vector
      1KB x 1Billion (1TB)
  - Smartphone GPS Data Analysis
    - # of smartphones = 100M
    - Car / Train / Pedestrian Traffic analysis,
      Traffic jam prediction
- Multi-modal Analysis
  - Analyze Tweet stream with video stream
  - Analyze Tweet stream with smartphone stream





IPR: I/O and Processing Ratio

$$IPR = \frac{\text{I/O Performance (Bytes/sec)}}{\text{Processing Performance (Bytes/sec)}}$$

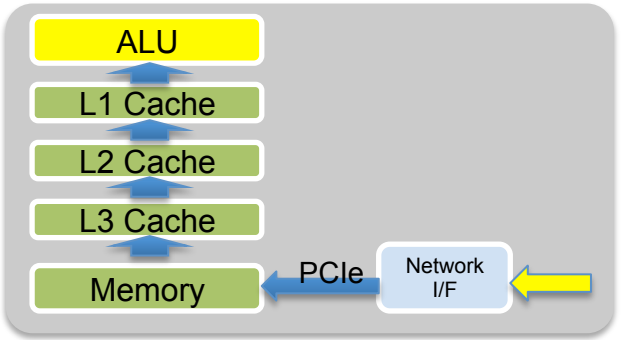
Typical CPU (Binary operation)
<< 5000

FPGA based processor
~= 1

Dataflow-centric node

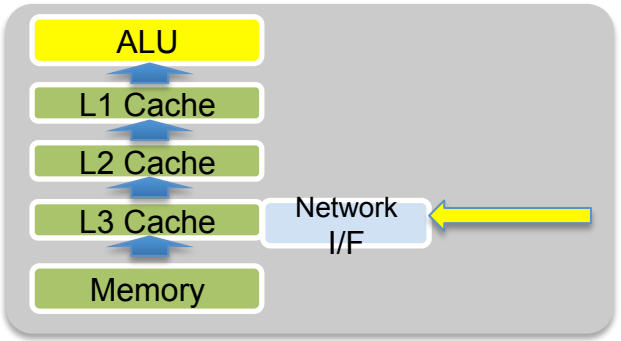
Current Architecture is not suitable

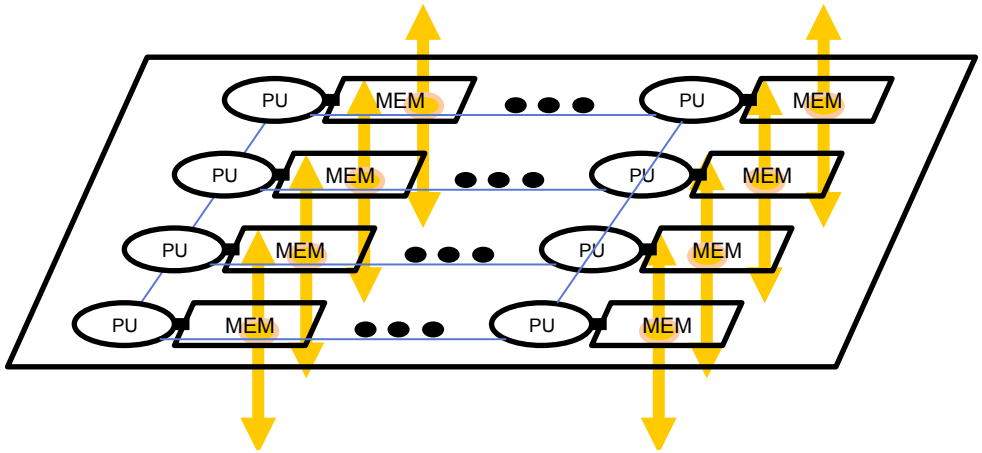
- Data transfer and processing are not tightly coupled
  - Poor performance for communication intensive applications
  - Processing performance depends on cache memory. Data transfer to/from cache is the bottleneck.



Ideal Architecture

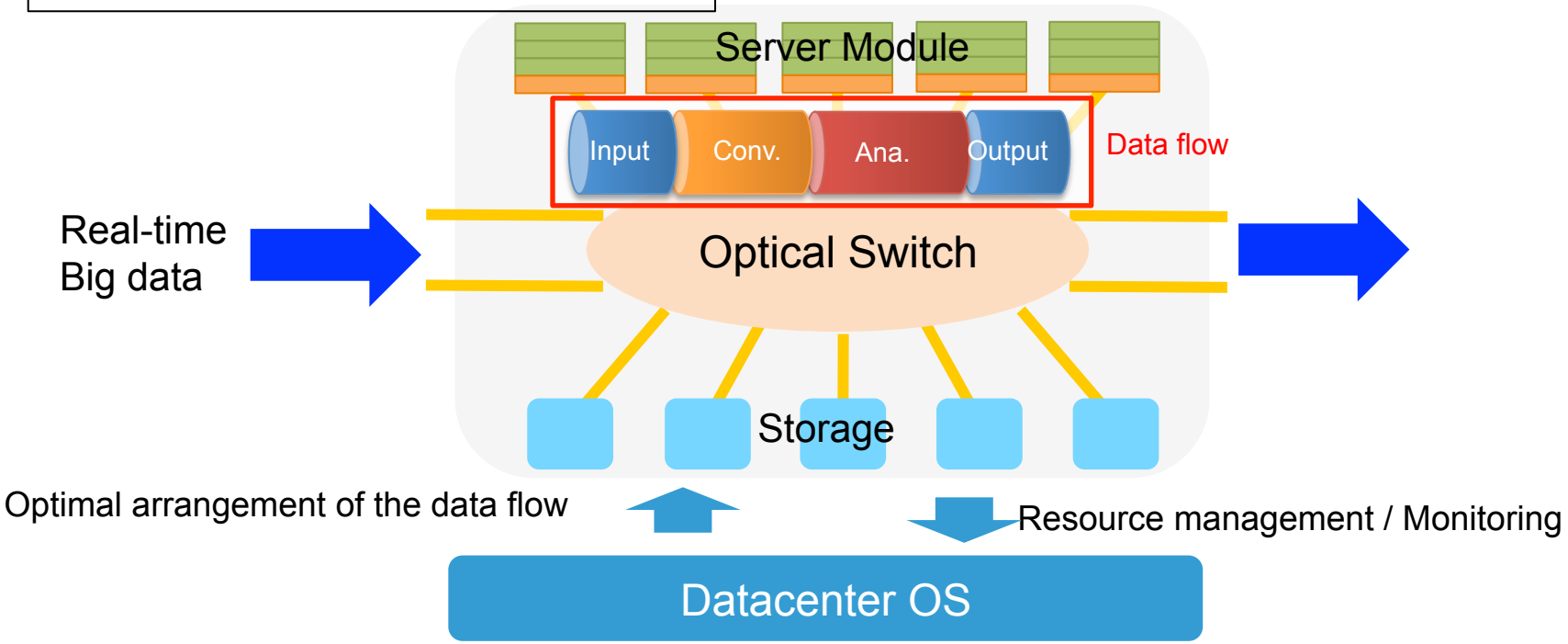
- Directly inject data from Optical I/O to the cache
  - Avoid memory bottleneck
- Integrate 10s of them into single chip
  - WDM (wavelength division multiplexing)





Dataflow-centric architecture overview

Nodes are tightly connected via Optical Switch forming Dataflow to process the incoming data in pipeline manner



A single OS called Datacenter OS governs the whole system.
It sets up the optical path, deploys the proper code on each node.

Optical Network Technology for Future Datacenters

- Large-scale silicon photonics based cluster switches
- DWDM, multi-level modulation, highly integrated “elastic” optical interconnects
- Ultra-low energy consumption network by making use of optical switches

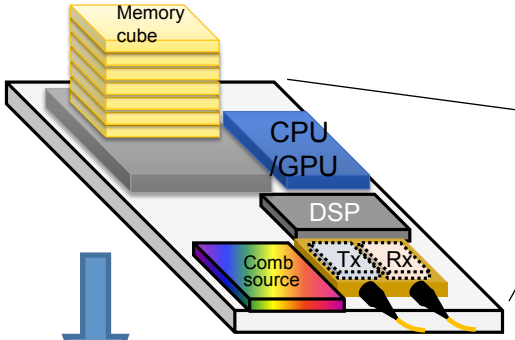
DWDM, multi-level modulation optical interconnects

Datacenter server racks

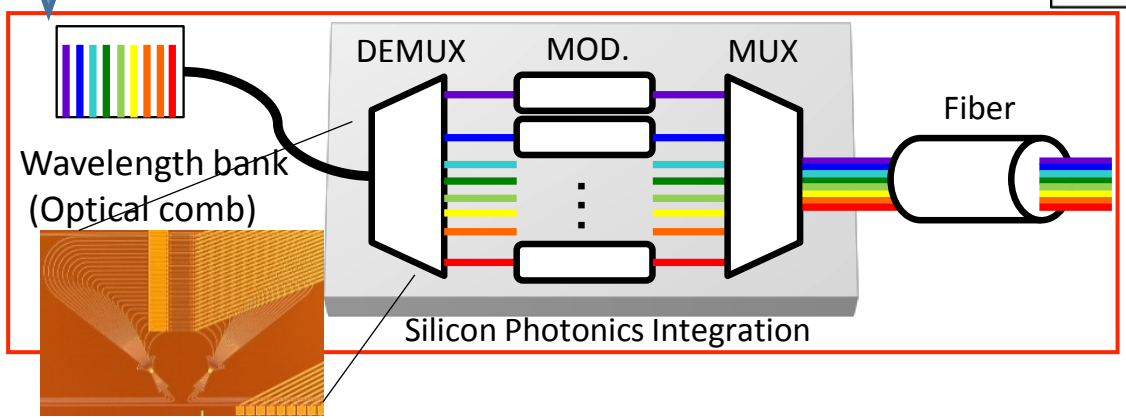
Silicon photonics cluster switches

- Ultra-compact switches based on silicon photonics
- 3D integration by amorphous silicon
- A new server architecture

Current electrical switches:
～130Tbps → ～500Pbps



2.5D-CPU Card



No of λs	Order of mod.	Bit rate
1	1	20 Gbps
4	8	640 Gbps
32	8	5.12 Tbps

Current state-of-the-art Tx
100Gbps → ～ 5.12Tbps