# Overview of Cygnus:
# a new supercomputer at CCS

# About Cygnus

At Center for Computational Sciences (CCS), University of Tsukuba, we have been developed and deployed several supercomputers in PC cluster architecture with accelerators such as GPU or many-core processor. As accelerated supercomputer with many-core architecture, we have been operating COMA cluster with Intel Xeon Phi (Knights Corner) which will be shut down on March 2019. Programming and utilization technology of Intel Xeon Phi is transferred to Oakforet-PACS system which is operated under collaboration with University of Tsukuba and University of Tokyo. On the other hand, we had been operating another supercomputer HA-PACS and HA-PACS/TCA with GPU accelerators, which were shut down on March and October 2018, respectively.

At CCS, we will introduce the next generation accelerated supercomputer, based on our experience to use GPU clusters, with additional accelerator technology of FPGA, which is named Cygnus. All computation nodes of Cygnus are equipped with the most advanced GPUs as four devices per node. Moreover, about half of nodes are additionally equipped with advanced FPGAs where two FPGA devices and four GPUs work together in each node.

Merging two different type of accelerators, GPU and FPGA, in each node to dedicate to parallel processing, we will challenge applications where GPU acceleration is not sufficient by several reasons to be compensated by FPGA technology, toward the new generation of strong scaling supercomputing. Cygnus is the world first GPU-FPGA equipped supercomputer to be opened for public use by various fields of application users.

# Configuration of Cygnus (1)

- Cygnus is a supercomputer which will be started to use at Center for Computational Sciences, University of Tsukuba
- System configuration
  - The system consists of two groups of computation nodes: “Deneb nodes” with CPU and GPU only, and “Albireo nodes” with CPU, GPU and FPGA.
    - Common specification for Deneb and Albireo nodes
      - CPU: Intel Xeon Gold 6126 2.6GHz (12 core) x 2 socket
      - GPU: NVIDIA Tesla V100 PCIe x 4台
      - Memory (CPU): 192 GiB, 255.9 GB/s
      - Memory (GPU): 32 GiB x 4 = 128 GiB, 900 GB/s x 4 = 3600 GB/s
      - Network: InfiniBand HDR100 x 4 channel (400Gbps x 4)
      - Node Storage: NVMe 3.2TB
    - Specification for Albireo node only
      - FPGA: Nallatech 520N (Intel Stratix10 with 100GBps x 4 channel ext. port)
      - Inter-FPGA Network: 100Gbps 2D Torus (8x8)

# Configuration of Cygnus (2)

- Albireo nodeのFPGA
  - Intel Stratix10 GX2800 H-Tile
  - LE: 2753 Kgates (8-in 4-out Lookup Table + 1bit register)
  - Memory
    - BRAM: M20K 229Mbit
    - MLAB: 15Mbit
    - External memory: 32 GiB, 76.8 GB/s
  - External network: 100Gbps x 4 channel

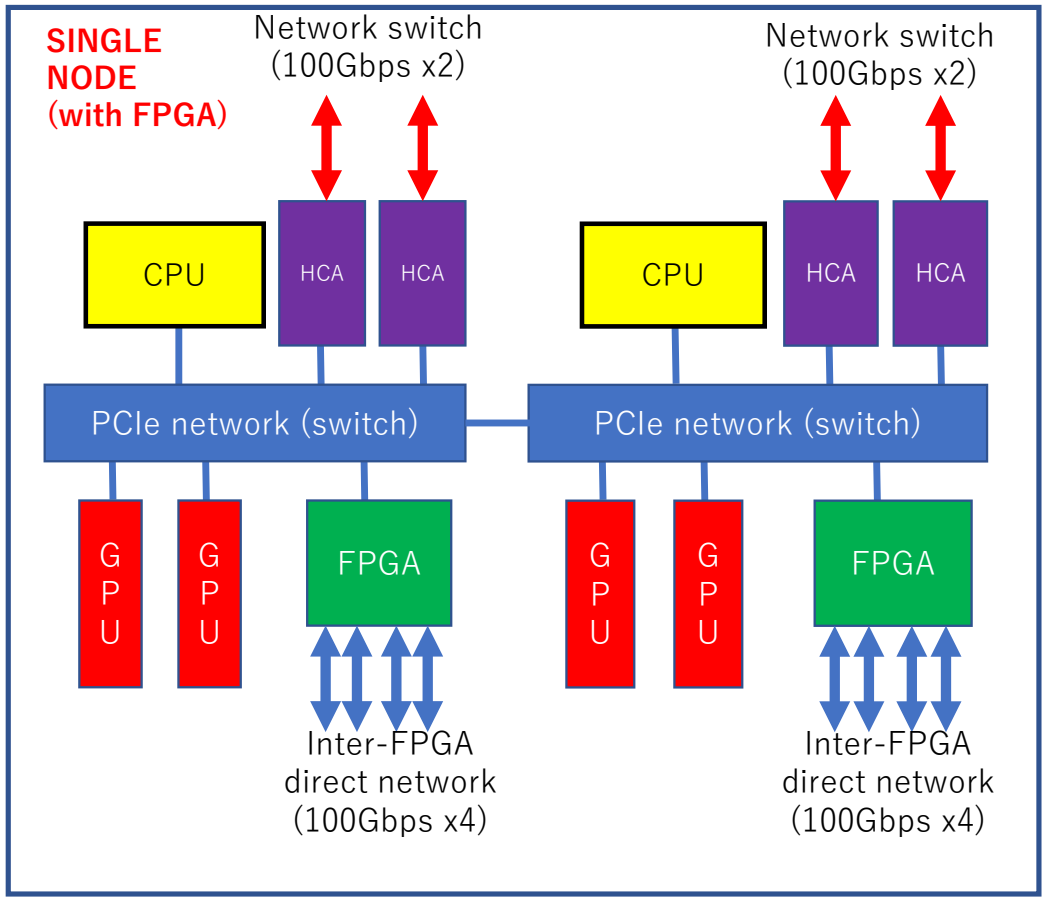
# Configuration of Cygnus (3)

- Shared file system
  - A large capacity of shared file system which can be equally accessed from any computation node, with 2.5 PB (user space) supported by Lustre cluster file technology by RAID6 configuration
  - It can be referred as “/work” file system which is used for main computation data storage and archive after computation
- Theoretical peak performance
  - Number of nodes： Deneb x 48 nodes, Albireo x 32 nodes = 80 nodes
  - CPU performance (double precision)：2 TFLOPS x 80 = 160 TFLOPS
  - GPU performance (double precision)：28 TFLOPS x 80 = 2240 TFLOPS
  - FPGA performance (single precision): 10 TFLOPS x 32 = 320 TFLOPS
  - Total system performance (double precision): 2.4 PFLOPS
    Total system performance (single precision): 5.12 PFLOPS
- Theoretical peak performance of interconnection network
  - 100Gbps x 4 x 80 = 4GB/s (full bisection bandwidth)

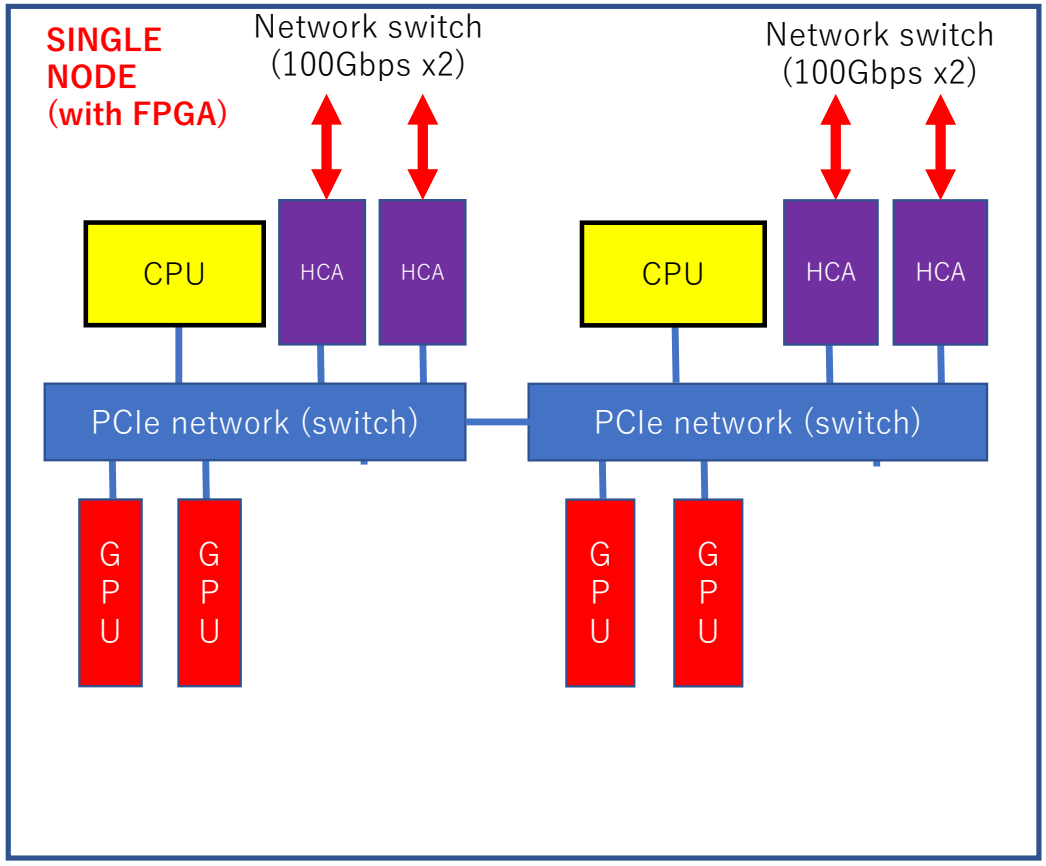
# Configuration of Cygnus (4)

- Home file server
  - NFS shared file server over InfiniBand
  - It can be referred as “/home” , and works as each user’s home directory at login time
- External network
  - 10Gbps x 4 port L3 router
- Login node
  - 3 login nodes for accessing to Cygnus system by ssh
  - Special nodes for FPGA compilation are prepared beside login nodes

# Block diagram of Albireo node

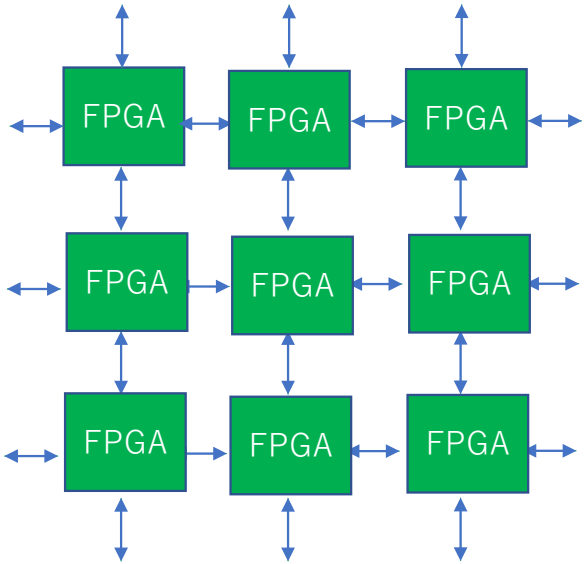


# Block diagram of Deneb node



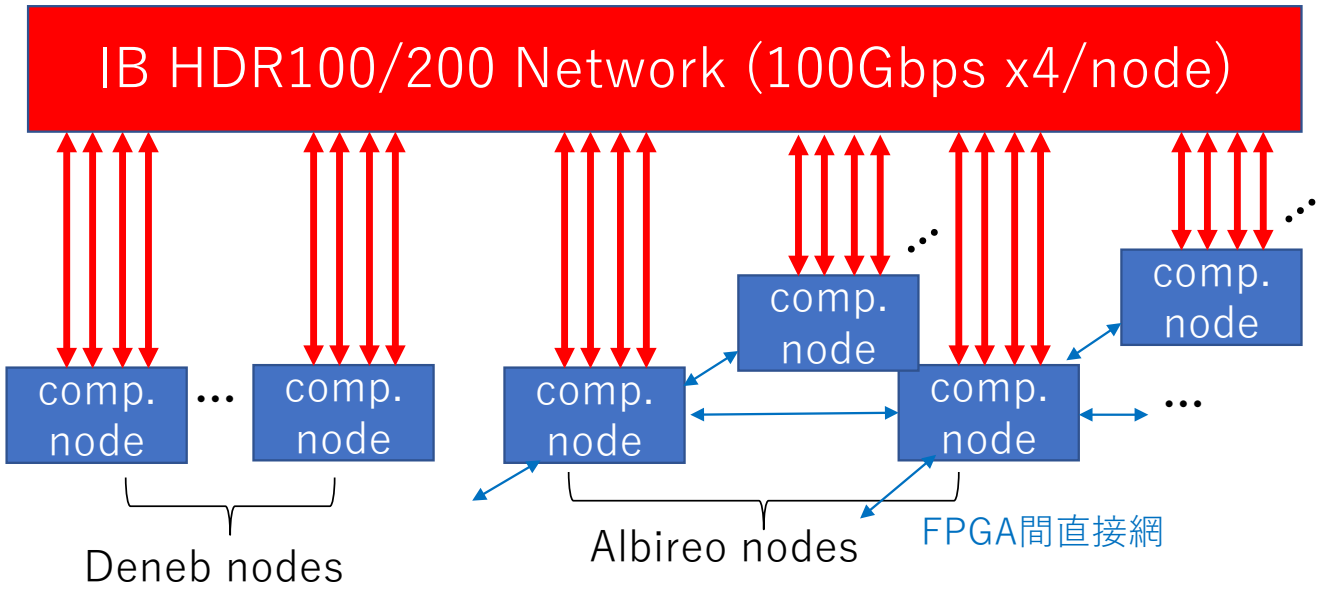
# Two types of interconnection network

Inter-FPGA direct network
(only for Albireo nodess)




64 of FPGAs on Albireo nodes (2 FPGAS/node) are connected by 8x8 2D torus network without switch

InfiniBand HDR100/200 network for parallel processing
communication and shared file system access from all nodes




For all computation nodes (Albireo and Deneb) are connected by full-bisection Fat Tree network with 4 channels of InfiniBand HDR100 (combined to HDR200 switch) for parallel processing communication such as MPI, and also used to access to Lustre shared file system.

# Overall specification of Cygnus

| Item | Specification |
|---|---|
| Peak performance | 2.4 PFLOPS DP<br>(GPU: 2.2 PFLOPS, CPU: 0.2 PFLOPS, FPGA: 0.6 PFLOPS SP)<br>⇨ enhanced by mixed precision and variable precision on FPGA |
| # of nodes | 78 (32 Albireo (GPU+FPGA) nodes, 46 Deneb (GPU-only) nodes)<br>⇨ 2 additional nodes will come, in total 80 |
| Memory | 192 GiB DDR4-2666/node = 256GB/s, 32GiB x 4 for GPU/node = 3.6TB/s |
| CPU / node | Intel Xeon Gold (SKL) x2 sockets |
| GPU / node | NVIDIA V100 x4 (PCIe) |
| FPGA / node | Intel Stratix10 x2 (each with 100Gbps x4 links/FPGA and x8 links/node) |
| Global File System | Lustre, RAID6, 2.5 PB |
| Interconnection Network | Mellanox InfiniBand HDR100 x4 (two cables of HDR200 / node)<br>4 TB/s aggregated bandwidthj |
| Programming Language | CPU: C, C++, Fortran, OpenMP, GPU: OpenACC, CUDA<br>FPGA: OpenCL, Verilog HDL |
| System Vendor | NEC |