HPP controller: a system controller for high performance computing

Fei CHEN (✉), Zheng CAO, Kai WANG, Xuejun AN, Ninhui SUN

Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China

© Higher Education Press and Springer-Verlag Berlin Heidelberg 2010

Abstract This paper introduces the design of a hyper parallel processing (HPP) controller, which is a system controller used in heterogeneous high performance computing systems. It connects several heterogeneous processors via HyperTransport (HT) interfaces, a commercial Infiniband HCA card with PCI-express interface, and a customized global synchronization network with self-defined high-speed interface. To accelerate intra-node communication and synchronization, global address space is supported and some dedicated hardware is integrated in the HPP controller to enable intra-node memory and shared I/O resources. On the prototype system with the HPP controller, evaluation results show that the proposed design achieves high communication efficiency, and obvious acceleration to synchronization operations.

Keywords hyper parallel processing (HPP), system controller, heterogeneous computing

1 Background

Heterogeneous computing has become a trend in building high performance computing systems as various kinds of accelerators are now being used. GPU and field programmable gate array (FPGA) are the most widely used accelerators, which can provide higher computing density and power efficiency than commercial general-purpose processors. For example, the Fermi GPU produced by nVidia can deliver 500 GFlops computing capacity on a single chip, while the latest 12-core AMD Opteron processor can only provide 132 GFlops. The power consumption of the Fermi GPU is 280 W, while 12-core AMD Opteron processor is 105 W. In this case, the GPU can provide more than twice the performance per watt of the latest AMD processor.

However, common GPUs and FPGAs cannot work without a general-purpose processor in a high performance computing system. GPU and FPGA integrated CPUs, such as Intel Clarkdale and Xilinx FPGA integrated with PowerPC, make it possible to change the common way of using these accelerators. Besides GPU and FPGA, some special processors can also be used as accelerators thanks to their high power efficiency or high performance for specific applications: the Godson processor for example.

To achieve high performance, more and more machines in the TOP10 List use accelerators. In June 2009, only one machine used an accelerator, however, by June 2010, their number had tripled. The architecture of machines using accelerators can be abstracted by the architecture shown in Fig. 1. They have many common characteristics: 1) multiple computing nodes with high performance interconnection; 2) architecture of computing node is preferred; 3) a few processors within a computing node; 4) accelerators (ACL in Fig. 1) connect to the general purpose processor via local bus; they are not directly connected.
connected with each other. Such connection between ACL and CPUs is suitable for some applications with strong space locality, such as gene-resequencing. However, for applications with much inter-node communication, such as MG, CG, and IS in NAS parallel benchmarks (NPB), the long communication path between accelerators becomes a significant drawback. Besides, the number of accelerators in a computing node is much limited or else the CPUs will become a bottleneck. Finally, the tightly-coupled node architecture prevents the computing node from using heterogeneous processors.

To use accelerators efficiently, we propose the hyper parallel processing (HPP) [1] architecture, which is shown in Fig. 2. The HPP architecture is suitable for using accelerator integrated CPUs or heterogeneous processors in a computing node. Referring to inter-node interconnection, there is no difference between HPP and former architectures (such as MPP/Cluster). However, to reduce the cost of the interconnection network, commercial networks are preferred in HPP. The primary contribution of HPP is the architecture of the computing node. A computing node of the HPP architecture, HPP node, has the following characteristics:

- Accelerators are logically and physically, directly connected to a global interconnect, which shortens the communication path between accelerators.
- Cache coherence is not maintained in an HPP node, so scalability is good in a computing node.
- At least one service processor, a commercial general-purpose processor, is used in an HPP node. A full-featured operating system runs on this processor and provides a single system image (SSI) to maintain compatibility with commercial software.
- Computing processors, play the role of accelerators. A light-weight operating system, which may only provide the runtime environment, is used to reduce operating system noise (the amount an OS interrupts user applications).

The above characteristics are directly related to the design of the system controller. More detailed descriptions on the HPP architecture appear in [1].

As shown in Fig. 3, a detailed implementation of HPP architecture named HyperDragon is introduced. The HPP node is the basic computing node of this heterogeneous computing system. There are two kinds of interconnection network between HPP nodes. One is a customized global synchronization network (G-Net), which is used for global communication and synchronization; such as Barrier and All-reduce operations. The other is a commercial Infiniband network (D_Net), which provides high bandwidth and low latency interconnection of nodes. There are two kinds of processor in the HPP architecture, one is a computing processor, and the other one is a service processor.

As shown in Fig. 3, the Godson processor is selected as