Design and Implementation of High-Performance Embedded Processing System Based on DSM

Yang Feng
School of Information and Electronic
Beijing Institute of Technology
Beijing, China
gy_beyond@163.com

Shanqing Hu*
School of Information and Electronic
Beijing Institute of Technology
Beijing, China
hushanqing@bit.edu.cn

Teng long,Feng Liu
School of Information and Electronic
Beijing Institute of Technology
Beijing, China
longteng@bit.edu.cn
liufeng@bit.edu.cn

***Abstract*—In this paper, on the basis of detailed analysis about an extended speedup of the high-performance embedded parallel processing system, we propose a new parallel processing architecture based on distributed storage, and then design a generic real-time signal processing card with multi DSPs of TMS320C6678. Finally, we construct a real-time signal processing system with multi cards which can get the parallel processing effect of 90%. It validates the proposed architecture has standardized, modular, scalable, reconfigurable features.**

Keywords—Speedup; distributed storage; multi-level interconnect; TMS320C6678;

I. INTRODUCTION

With the improvement of the VLSI(very large-scale integration) technology, embedded parallel processing system has achieved a rapid progress, and was widely used in aerospace, communications, medical electronics and other fields. Meanwhile, these applications have increasing demands on the system processing ability, storage capacity and transmission bandwidth [1]. Using the more high-performance processor, bigger storage capacity and faster speed memory has become an inevitable choice for hardware system design [2]. Parallel processing between the multiple processors have become an effective mean to further enhance the system capacity [3]. But the traditional parallel processing system architecture based on hierarchical shared storage are subject to the load capacity and access competition of shared bus, cannot achieve high transmission rate which greatly limits the processing ability of the system[4][5]. Therefore, on the basis of detailed analysis about an extended acceleration-ratio, we propose a new distributed storage parallel processing architecture with high speed serial bus interconnection (such as SRIO, PCIE), and then design a generic real-time signal processing card with multi DSPs of TMS320C6678, and construct a real-time signal processing system with multi cards. The experiment results of parallel processing effect show that the system has powerful parallel processing performance.

II. ANALYSIS OF SPEEDUP

System speedup is a reflection of the degree of parallel processing capabilities. A fixed workload speedup equation is shown in (1) [6].

$$S(N)=\frac{T(1)}{T(N)} \tag{1}$$

N is the number of processors, $T(1)$ is the execution time of the sequential algorithm, and $T(N)$ is the execution time of the parallel algorithm with N processors. If $S(N)$=N, we call it linear speedup or ideal speedup. But, in a real system, this is impossible. We must take some other factors into account. So an extended speedup can be described in (2) [7], in which T_{sync} means the overall synchronization time and T_{com} stand for the overall memory accessing time.

$$S(N)=\frac{T(1)}{T(N)+T_{sync}+T_{com}} \tag{2}$$

Let's assumed:

I: The total number of floating-point operations of an algorithm;

P: The processing speed of one processor in the system;

δ : The serial processing bottleneck of the algorithm, which means this part just can be processed in serial.

M: The total number of synchronization between multi processors in the algorithm.

t_{sync} : The time of one time of synchronization.

W: The total number of exchanged data in synchronization.

B: The bandwidth of interconnected bus between processors.

α : The ratio of parallel data transmission.

Then:

$$T(N)=\frac{I\delta}{P}+\frac{I(1-\delta)}{NP} \tag{3}$$

$$T_{sync}=Mt_{sync} \tag{4}$$

$$T_{com}=(1-\alpha)MN(\frac{W}{B}+T_{delay}) \tag{5}$$

978-1-4799-6233-4/14/$31.00 ©2014 IEEE

$$S(N)=\frac{T(1)}{T(N)+T_{sync}+T_{com}}$$

$$=\frac{\frac{I}{P}}{\frac{I\delta}{P}+\frac{I(1-\delta)}{NP}+Mt_{sync}+(1-\alpha)MN(\frac{W}{B}+T_{delay})} \quad (6)$$

$$E(N)=\frac{S(N)}{N}$$

$$=\frac{1}{N\delta+(1-\delta)+[NMt_{sync}+(1-\alpha)MN^2(\frac{W}{B}+T_{delay})]\frac{P}{I}} \quad (7)$$

Considering traditional shared bus parallel processing system, processors just can access the bus in serial, so α =0.Then:

$$S(N)=\frac{1}{\delta+\frac{(1-\delta)}{N}+Mt_{sync}\frac{P}{I}+MN(\frac{W}{B}+T_{delay})\frac{P}{I}} \quad (8)$$

$$E(N)=\frac{1}{N\delta+(1-\delta)+MN^2(\frac{W}{B}+T_{delay})\frac{P}{I}+MNt_{sync}\frac{P}{I}} \quad (9)$$

From (8) and (9), we can see that efficiency of shared memory parallel processing system has inverse relationship with N^2 .Meanwhile, with the increase of N, limited by the load capability of shared bus, the bandwidth of data transmission will decline dramatically, $E(N)$ will have a further decline. This is seriously limited the scalability [8]. Therefore, the shared bus parallel processing architecture can't meet the system requirements well.

According to (6), in order to further improve speedup and efficiency of parallel processing systems, we should:

- minimize the serial bottleneck factors δ , mainly rely on better architecture design of algorithms;
- increase the ratio of parallel data transmission α ;
- increase the data transmission bandwidth B;
- reduce synchronization time overhead t_{sync} ;

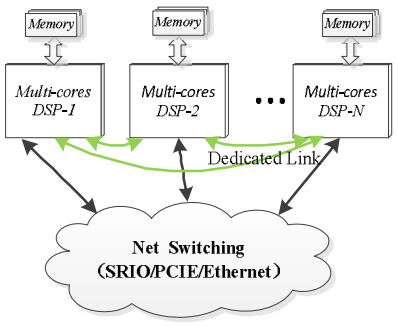

Fig. 1. Distributed storage parallel processing architecture.

Considering the above analysis, we proposed a distributed storage parallel processing architecture shown in Fig. 1. Each processor has its own memory, and exchange data through high-speed serial bus based on packet switching. With the well-designed switching net, It can increase data transmission bandwidth B and ratio of parallel data transmission α with very little influence of increase of N, which can get good parallel efficiency. And this architecture has good scalable, reconfigurable features.

III. Design of System

A. System Architecture Design

Based on the above analysis of distributed memory parallel processing architecture, we construct a high-performance embedded parallel processing system. As shown in Fig. 2.

The system mainly includes IO modules, network switching module, processing module, main control module.

- Main control module: This module has the function of enumeration PCIE EP devices within the system to achieve the formation of the system PCIE network, provide graphical user interfaces for the management and maintenance of the system and be responsible for the functions of task assignment, parameter initialization and auxiliary algorithm processing. The entity of the main control module is usually a SBC (Single Board Computer).
- I / O module: This module realizes interconnection with other external system. It is usually implemented by large-scale FPGAs. It mainly consists of I/O interface unit and I/O controller. IO interface unit implements various types of transmission interfaces, such as ADC, DAC, fiber input or output interface and so on. IO controller realizes data distribution and timing synchronization.
- Network switching module: This module realizes non-blocking exchange of data between multi kinds of high-speed serial bus based on package-switched protocol, such as SRIO, PCIE and Gigabit Ethernet, provides high-speed non-blocking data transmission channel for each processing node in the system.
- Processing module: Processing module is the core of high-performance embedded parallel processing system. It consists of many high-performance processors which interconnect with each other through high speed serial bus. The processors usually are DSPs or FPGAs which are suitable for the processing of embedded system.

As described above, a high-performance parallel processing system which is constructed based on different modular has standardized, scalable, reconfigurable features. Many kinds of systems with different performance can be constructed through increasing or decreasing the size of each module flexibly according to the specific needs of various applications [9].

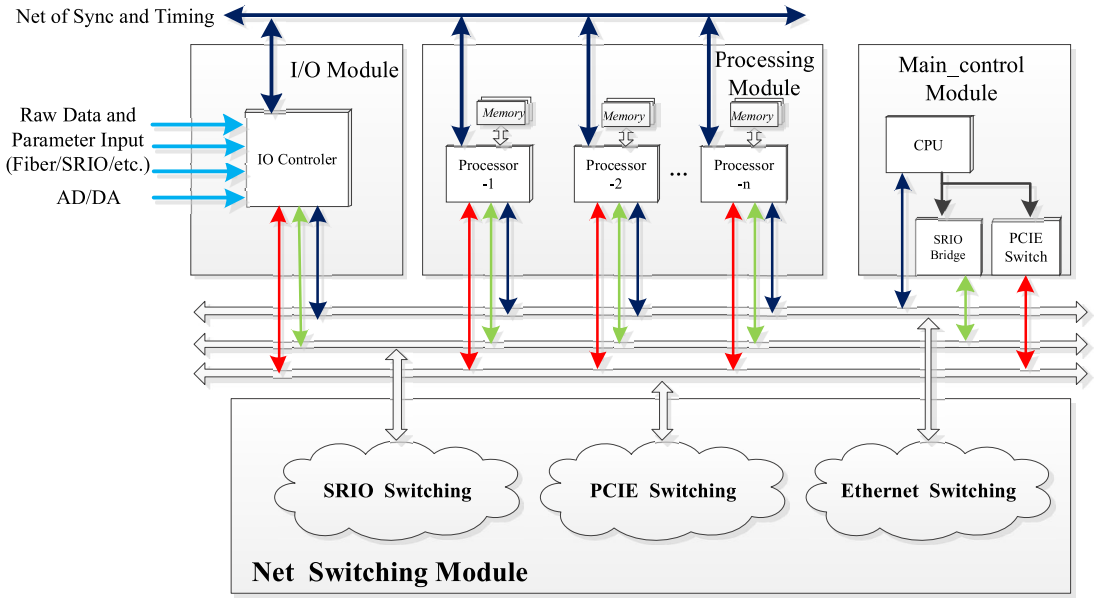

Fig. 2. Diagram of system.

B. *Characteristics of Architecture*

1) Multi-level network

The system architecture described above introduces three ways to interconnect each unit within the system:

- Interconnection based on high-speed serial bus(e.g. SRIO,PCIE) provide a large -bandwidth data transmission channel for each processing node;
- Gigabit Ethernet network realize low-speed, non-real-time data exchange ;
- Sync bus is used for the synchronization and timing among multi processors;
- Different interconnection is used to realize different data transmission. System can achieve non-blocking data exchange though multi-level network.

2) Flexible interconnection topology

Given the advantage of flexible topology building with SRIO network, we choose SRIO network as the main transmission path for data. With the assistance of SRIO switching module, it is easy to build different topologies shown in Fig. 3 between processing nodes through configuring different ID for the ports of SRIO switch [6]. So this processing module has highly scalable and reconfigurable characteristics.

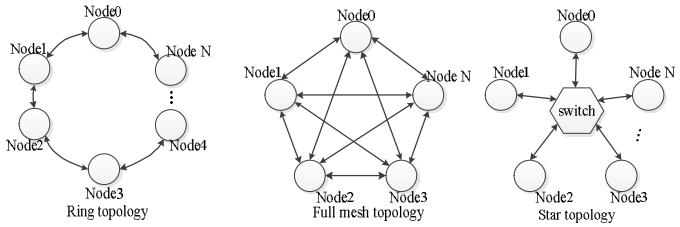

Fig. 3. Graphic of various topologies.

3) Large -capacity data cache

Because of distributed storage, each processor can own independent storage through SARAM controller. Therefore, the system capacity of data cache is in a linear increase with the number of processors. The high performance processors of the time usually support DDR3 SARAM controller, and with the growing space of DDR3 chip, processors can get more storage. So, system with multi processors based on distributed storage can easily get large capacity data cache.

IV. HARDWARE DESIGN AND IMPLEMENTATION

A. *Implementation of parallel processing unit based on TMS320C6678*

TMS320C6678 is a highest-performance fixed/floating-point DSP that is based on TI's KeyStone multicore architecture [10]. It contains eight TMS320C66x™ DSP cores whose speed can reach to 1.25GHz. Per core raw computational performance is an industry-leading 40 GMACS/core and 20GFLOPS/core (@1.25 GHz operating frequency). Meanwhile, this DSP provides abundant peripherals，such as four lanes of SRIO2.1, two lanes of PCIe Gen2, Hyperlink, Gigabit Ethernet, SPI, EMIF, I2C and so on. So it is a smart choice to choose this DSP as the main processor of HPEC system.

The hardware design of universal parallel processing unit is shown in Fig. 4. There are 4 C6678s with this unit. Each DSP gets 8GB DDR3 SDRAM so that the unit can achieve large-capacity cache of 32GB. There are four lanes SRIO links and two lanes PCIe links between DSPs and SRIO switch or PCIe switch. The hyperlink which connects two DSPs provides a high-speed data transmission channel. FPGA not only implements the interface conversion between SRIO, PCIe, Link interface and Rocket IO [11], but also realizes programming FPDP and synchronization timing bus [12]. The

main role of the CPLD is power monitor and management, reset management, other logic controls and so on [13].

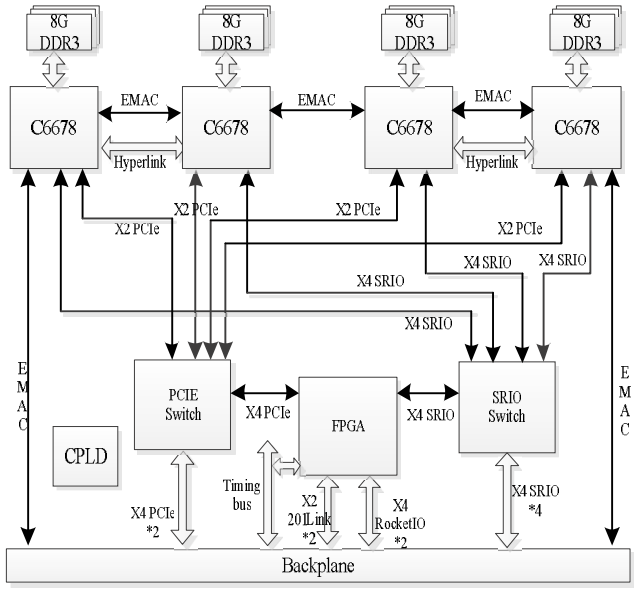

Fig. 4 Structure of parallel processing unit.

B. Implementation of Real-time Signal Processing System

Based on the C6678 parallel processing unit described above, we construct a high-performance embedded real-time signal processing system which is shown in Fig. 5.It consists of a IO card, 10 4DSP processing cards, a PowerPC card and a SRIO net switching card.

High-speed processing modular consists of 10 4DSP cards to complete complex signal processing algorithm computing. It can get peak processing capacity of 6.4TFLOPs and total data cache of 320GB. After a typical SAR imaging Specan algorithm validation, speedup ratios obtained exceed 3.6 and the efficiency is above 90% with the parallel processing algorithms. We can conclude that the processing module has powerful parallel processing capabilities.

V. CONCLUTUON

This paper aims to introduce a parallel processing architecture based on distributed storage with high speed serial bus interconnection on the help of the analysis of extended speedup formula. Then we design a high-performance embedded real-time signal processing system with powerful processing capacity and huge data cache, which validate that the proposed parallel processing architecture has standardized, modular, scalable, reconfigurable characteristics to meet the needs of large bandwidth, high-capacity storage and high processing performance, and can better adapt to a variety of signal processing algorithms.

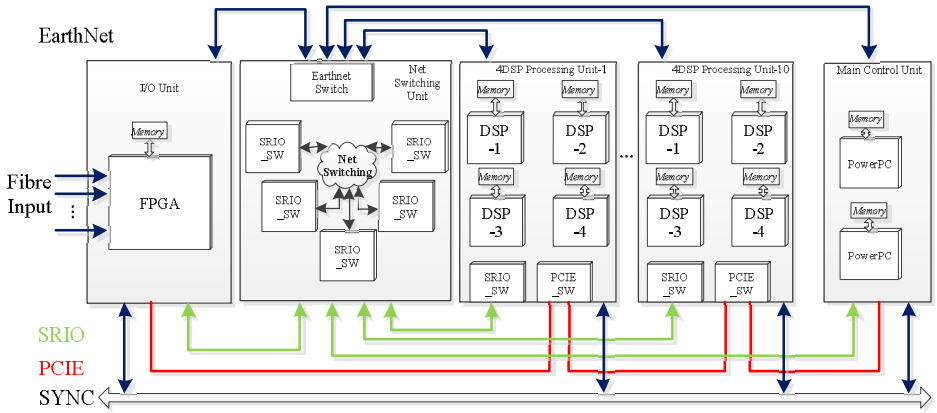

Fig. 5 Diagram of real-time signal processing system.

REFERENCES

[1] P. G. Paulin , C. Liem , M. Cornero , F. Nacabal and G. Grossens "Embedded software in real-time signal processing systems: application and architecture trends", Proceedings of the IEEE, vol. 85, no. 3, pp.419 -435 1997

[2] Mingzhi.He, Jiang.Zhu, Bo.Zhou, in: Research on missle-borne SAR real-time signal processing systmen[J], Journal of electronics and information, (2010).

[3] C Hansen "MicroUnity's mediaprocessor architecture", IEEE Micro, vol. 16, pp.34 -40 96

[4] J. S. Vetter, Contemporary High Performance Computing: From Petascale Toward Exascale. Chapman & Hall/CRC, 2013.

[5] K. Pretz, "The future of computing," The Institute, IEEE, 2013.

[6] Zhimin. Gu, Xianhe and Sun Parallel, "Parallel Processing System Architecture and Scalable Computing" Tsinghua University press ,2009.

[7] Yi. Deng, Shanqing. Hu, Teng. Long, in: A miniaturized universal architecture for radar signal processing systems[C], IET International Radar Conference, (2009).

[8] Y-T. S. Li and S. Malik "Performance Analysis of Embedded Software Using Implicit Path Enumeration", IEEE Trans on Computer-Aided Design, vol. 16, no. 12, pp.1477 -1487 1997

[9] Guoliang. Chen, Junmin. Wu, Feng. Zhang and Longbing. Zhang, in "Parallel Processing System Architecture", Higher Education Press

[10] TMS320C6678 Multicore Fixed and Floating-Point Digital Signal Processor Data Manual[Z], Texas Instruments corporation , (2013).

[11] J. Batlle, J. Marti, P. Ridao and J.Amat, "A New FPGA/DSP-Based Parallel Architecture for Real-Time Image Processing," Real-Time Imaging, 8, 345-356, 2002.

[12] Xiao. Ma, Lining. Gao, Xiaosong. Xie, in: Research of high-speed radar signal processing system based on OpenVPX standard[C], IET International Radar Conference, (2013).

[13] H. Le and V. Prasanna, "Scalable high throughput and power efficient Ip-lookup on FPGA," in 17th IEEE Symposium on Field Programmable Custom Computing Machines, 2009. FCCM '09., 2009