

# Research on DSP-GPU Heterogeneous Computing System

Li Xiangzhen

Jiangsu Automation Research Institute of CSIC,  
Lianyungang, China, 18248966608

Hu Jingying

Jiangsu Automation Research Institute of CSIC,  
Lianyungang, China

Lv Weiguo

Jiangsu Automation Research  
Institute of CSIC,  
Lianyungang, China

Wang Guiqiang

Jiangsu Automation Research  
Institute of CSIC,  
Lianyungang, China

Cai Xinrong

Jiangsu Automation Research  
Institute of CSIC,  
Lianyungang, China

**Abstract**—In this paper, DSP-GPU heterogeneous computing system is studied and the system architecture is designed. The task scheduling model is analyzed and the discrete particle swarm optimization algorithm is used for the DSP-GPU heterogeneous computing. The communication framework between DSP and GPU is designed for the heterogeneous computing.

**Keywords**—DSP; GPU; Heterogeneous Computing; Task Scheduling

## I. INTRODUCTION

With rapid development and wider application, general purpose computing by GPU became a research focus in recent years. With going deep into the research, some shortcomings of GPU are found. For example, GPU is weak in integer computing and logic computing. If control statement is existed in program, the running efficiency of GPU is deeply affected. For this reason, GPU is not suitable in condition control. GPU can not work well for Random memory accessing and dynamic jumping because the most kinds of GPU communicate with host only by bus. To solve GPU's limitation, Intel, AMD and some laboratories start to study the cooperation computing between CPU and GPU. This technology was called heterogeneous computing and the open computing language (OpenCL) was designed to meet this important need[1-4].

In the domain of real-time digital signal processing, the DSP with harvard architecture is used widely. However, the difficulty of real-time digital signal processing increased with the more information detections. Only DSP is not suitable for need in engineer. Multi-core DSP abated this difficulty in a certain extent. The multi-core technology is suitable for task parallel processing and GPU is suitable for data parallel computing[5]. Thus it is necessary to study DSP-GPU heterogeneous computing system.

## II. DSP-GPU HETEROGENEOUS COMPUTING SYSTEM

Heterogeneous computing system is defined as follows: systems which are assembled from different subsystems, each of them optimized to achieve different optimization points or to address different workloads. The heterogeneous computing

models of different equipments include two kinds[4,6]: master-slave model and parallel model. Parallel model usually used for condition that two computing equipments capability and tasks are similar. Master-slave model is now used widely. In CPU-GPU heterogeneous computing system, CPU is usually used as master core and GPU is slave core. Master core takes charge control and series computing and slave core do parallel computing[7].

With separate program and data memories, DSP can access instruction and data at one time. DSP has special hardware multiplication and complete one multiplication and addition in one instrument cycle. To accelerate processing, DSP has special instructions for SIMD operations. According to aforementioned characteristic, DSP can separate signals and complete control computing such as analysis and feedback for the result of signal processing. Be different from DSP, GPU complete data parallel computing by multi shader processors and is suitable for data intensive computing task. However, GPU is not work well for integer computing and logic computing. Thus master-slave model is designed for DSP-GPU heterogeneous computing system. The model structure is shown as Fig.1.

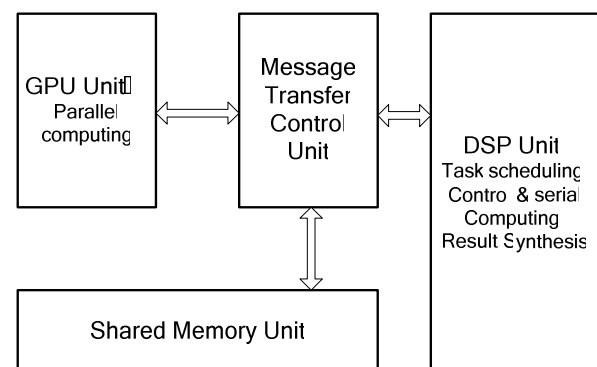


Figure1. DSP-GPU heterogeneous computing system structure

In DSP-GPU heterogeneous computing system, DSP is used as master core and GPU is slave one. In DSP, The

parallel characteristic of computing task is analyzed and assigned. The independent parallel computing task is transferred to GPU unit from message transfer control unit. At the same time, the dependent computing tasks and serial computing tasks are processed in DSP unit. Finally, the computing results from master core and slave core are synthesized and the final processing result is presented.

According to the characteristic of heterogeneous computing system, the selected DSP should has powerful logic and data computing capability. The high speed bus of DSP should be provided for high speed communication with GPU. The processor TMS320C6678 is a processor based on Keystone framework. Eight C66x cores are integrated in the processor and include fixed and floating-point computing. The process rate of each core is up to 1.25GHz and suffices for complex application. C6678 supports PCIe x2 and can communicate with GPU unit. The DSP unit model is shown in Fig.2.

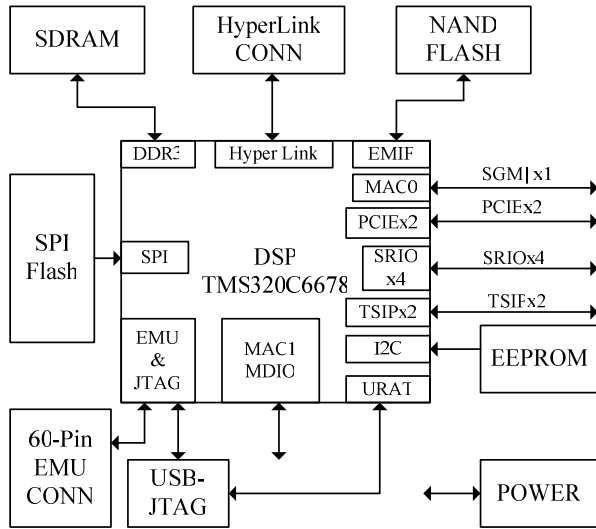


Figure 2. DSP unit model structure

DSP unit completes task analysis and scheduling. SDRAM is used as shared memory for heterogeneous computing system. FLASH is the DSP program memory and the

initialization information is saved in it. After the task is separated by DSP, the independent attached task is transmitted to GPU unit through shared memory and the dependent computing is completed.

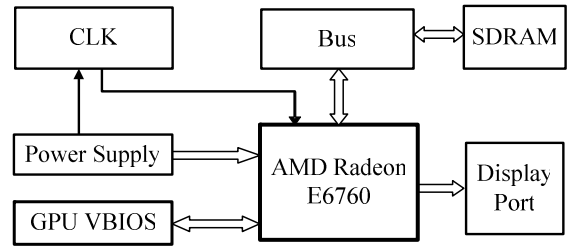


Figure 3. GPU unit model structure

In heterogeneous computing system, GPU unit mainly complete data intensive parallel computing. Thus the selected GPU should have general programmable computing capability. The Mobility Radeon E6760 provided by AMD company supports OpenCL 1.1 and the number of shader processor is up to 480. The processes capability of it is up to 576 GFLOPS SPFP and suffices for the parallel computing requirement of heterogeneous computing system. The model structure of GPU unit is shown in Fig.3.

During the process of heterogeneous computing, some key questions such as the parallel analysis of computing task、task scheduling algorithm、communication between master and slave core, et al, should be taken into account. This questions influence the efficiency of heterogeneous computing and some analyses are given as follows.

### III. TASK SCHEDULING

In DSP-GPU heterogeneous computing system, the task process time in different units is different. To improve the system efficiency, the task scheduling should be optimized. The discrete particle swarm optimization algorithm is used to optimize maximum and minimum complete times and load stabilization of each process unit for heterogeneous computing system<sup>[7]</sup>. First, the random element is initialized. Second, the result is estimate according to automatic variable. Finally, next selection is decided to run or not. The flow detail is shown in Fig.4.

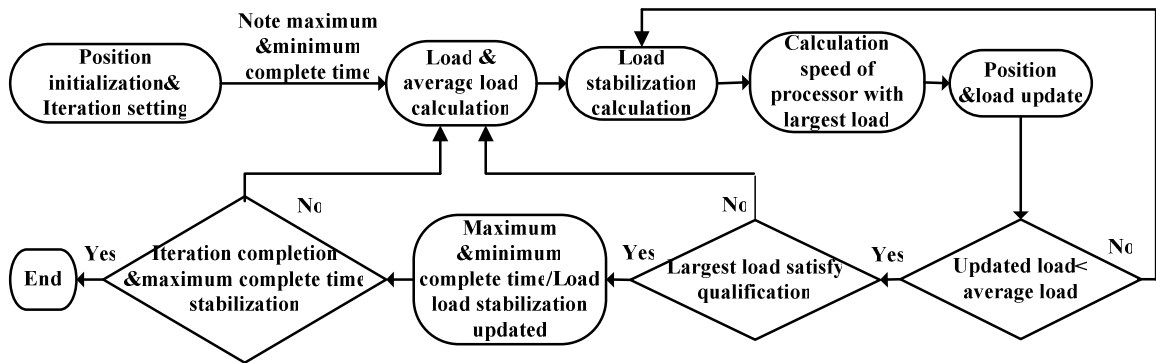


Figure 4. Task scheduling algorithm flow chart

In DSP-GPU heterogeneous computing system, the load stabilization of different processor units is improved by setting a free resource list in DSP. To optimize the parallel process, the tasks are assigned by message transferring. The transferred message includes control command and pointer. The control command triggers task and the pointer points to task data.

#### IV. COMMUNICATION BETWEEN DSP UNIT AND GPU UNIT

In DSP-GPU heterogeneous computing system, "Master-slave" communication structure is adopted and shown in Fig.5[8]. In which the message control unit is independent and can run parallel with DSP and GPU. Data intensive task is assigned by DSP and this procedure is called data transfer communication. In our communication model, this procedure is separated into two steps. First, DSP transfers the task characteristic information to message control unit. Second, message control unit transfers the computing task to GPU from shared memory according to the characteristic information. At the same time, the control information of result given by GPU is transferred to message control unit and the computed result is transferred to shared memory.

To accomplish the first step, the mail box is defined as a message receiver to receive the message transferred from DSP and GPU. To accomplish the second step, the special data transfer module is set in the message control unit. Except for the two steps, few message communication and data in phase need real-time transfer and are usually communicated by PCIE bus.

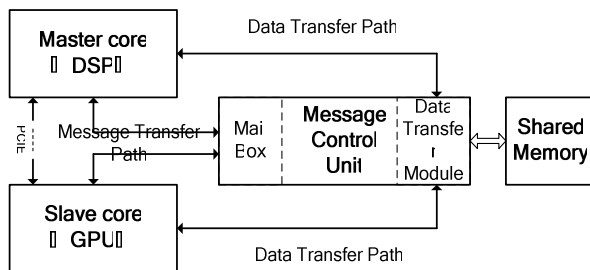


Figure 5. Model of communication between DSP and GPU

#### V. CONCLUSION

In this paper, DSP-GPU heterogeneous computing system is presented. Master-slave model is designed for heterogeneous computing system. DSP and GPU are defined as master core and slave core separately. The computing task parallel analysis and assignment are processed in DSP and the independent parallel computing task is transferred to GPU unit by message control unit. At the same time, the dependent and serial computing tasks are processed by DSP. Finally, the computed data of DSP and GPU are integrated by DSP and the final result is given. To improve the communication speed, master-slave communication model is designed. The message control unit can run parallel with DSP and GPU so that the heterogeneous computing system work with great efficiency and speed.

#### REFERENCES

- [1] CUDA-Zero: a framework for porting shared memory GPU applications to multi-GPUs. Science China(Information Science), Vol.03, pp.663-676, 2002.
- [2] Chuntao Hong, Dehao Chen, Yubei Chen, et al. "Providing Source Code Level Portability Between CPU and GPU with MapCG," Journal of Computer Science & Technology, Vol.01, pp.42-56, 2012.
- [3] Yang Xin, Wang Tianming, Xu Duanqing, "Fast BVH construction on GPU," Journal of Zhejiang University(Engineering Science). Vol.01, pp.84-89, 2012.
- [4] Benedict Gaster, Lee Howes, David R. Kaeli, Perhaad Mistry, Dana Schaa, Heterogeneous Computing with OpenCL, Morgan Kaufmann, 2011.
- [5] Huang keming, Wang Guocheng, Wang Yang, "Multi-Source Image Fusion System Based on DSP. Ordnance Industry Automation," Vol.02, pp.61-63, 2012.
- [6] Lu Fengshun, Song Junqiang, Yin Fukang, Zhang Lilun. "Survey of CPU/GPU Synergetic Parallel Computing," Computer Science. Vol.03, pp.45-49, 2011.
- [7] Jiang Jianchun, Wang Tongqing, Zeng Suhua. "Solving task assignment problem of heterogeneous parallel systems by hybrid DPSO algorithm," Control and Decision. Vol.26, pp.1315-1320, 2011.
- [8] Chen Guobing On-chip Communication of Embedded Heterogeneous Multicore Architecture. Zhejiang University Master's Degree Paper, 2007.