GPGPU AND MIC IN ACCELERATED CLUSTER FOR REMOTE SENSED IMAGE PROCESSING SOFTWARE

Olivier MELET, Ground Segment Architect, CNES, olivier.melet@cnes.fr

Toon Huysmans, PhD at University of Antwerp, toon.huysmans@ua.ac.be
Michel Hummel, HPC expert at Thales Service, michel.hummel@thalesgroup.com
Pierre-Marie Brunet, HPC expert at CNES, pierre-marie.brunet@cnes.fr

ABSTRACT
Processing of Earth observation remotely sensed images requires more and more powerful computing facilities. Since a few years, GPGPU (General Purpose processing on Graphics Processing Units) technology has been used to perform massively parallel calculations. The French Space Agency (CNES) has then made a portage of some IAS to assess their performance using this type of technology. Based on the very encouraging results and in order to improve operations during the in orbit commissioning phases of Earth observation satellites, CNES is developing, in partnership with Thales Services, an Expertise and Image Calibration Center based on a computer cluster accelerated by such technologies. Starting from an overall description of the image calibration objectives and GPGPU applications, the paper focuses on the study leaded by CNES in partnership with the Antwerp University and Thales Services to identify the most appropriate technologies to define a full computing cluster architecture.

Index Terms—GPGPU, MIC, Accelerated Computing Cluster, Image Algorithms Software, Calibration Center.

1. INTRODUCTION
Processing of Earth observation remotely sensed images requires more and more powerful computing facilities due to the increasing volume of images (higher spatial and spectral resolutions, repetitiveness of acquisitions,...), the timeliness requirements imposed by end users and the increasing complexity of Image Algorithms Software (IAS). Some of these algorithms can be parallelized (radiometric processing, resampling, equalization,...) thus image processing uses today's parallel capabilities of computer resources to execute [1] [2].

Since a few years, GPGPU technology has been used to perform massively parallel calculations. CNES has then made a portage of some IAS to assess their performance using this type of technology (portages for Cuda and OpenCL and measures on Nvidia GPU - Graphics Processing Units). The performance benchmarks showed gains in term of execution speed between 1.5 and 40 by comparison with the CPU version of the code depending on the image size and algorithm precision.

Based on these very encouraging results and in order to improve operations during the in orbit commissioning phases of Earth observation satellites, CNES is developing, in partnership with Thales Services, an Expertise and Image Calibration Center based on a computer cluster accelerated by such technologies.

The challenge for calibration centers is to be able to conduct the most exhaustive calibration campaign during the short in orbit test phase by increasing the quantity of processed images and the performance of the image quality algorithms. Increasing the quantity of processed images is useful to reduce the noise of the measures and therefore to improve the accuracy of the calibrations.

2. AN ACCELERATED COMPUTING CLUSTER ARCHITECTURE
The design of this cluster was preceded by a study to identify the most appropriate technologies to define a full computing cluster architecture including the hardware components (such as server platform, accelerator, node interconnect, and storage solutions) and the software components (such as the development tools, the parallelization frameworks, and the DRM - distributed resource managers).

Concerning development tools, the study provided an opportunity to compare the low-level framework such as CUDA (Compute Unified Device Architecture) and OpenCL (Open Computing Language) as well as high-level programming standard such as OpenMP (Open Multi-Processing), HMPP (Hybrid Multicore Parallel Programming) and OpenACC (Open ACCelerator). The most known DRMs have also been compared as well, in particular about their capabilities to manage GPGPU or MIC resources and the possibility to run a code in CPU mode or in accelerated mode depending on resource availability on the cluster.

The main considered high level criteria were performance, sustainability, scalability, maturity, maintenance effort and cost.
Those criteria have been taken in consideration in the study regarding specificities of an Image Calibration Center. Indeed, Image Calibration Centers are used during the in orbit commissioning phases but also during all the exploitation phases of Earth observation satellites which can be up to 10 years long. Also, the technologies used in the architecture must be sustainable over the lifetime of satellite operations and/or any necessary evolutions during this period have to be made at a lower cost. Furthermore, it is important that the IAS developed to be executed on this architecture are as generic as possible to allow their re-use on other Earth observation projects or other kind of center (then other kind of architecture). Finally, the architecture must be scalable to allow the introduction of new methods of image calibration during the life of the satellite.

It was a three stages study: the first step consisted in a market analysis, the second step was the appropriation of these technologies to the needs of a calibration center and the final step was the prototyping, benchmarking and the architecture definition.

The study showed that two technologies are competing on the market for massively parallel computation: GPGPU and more recently MIC (Many Integrated Core) implemented by Intel as Xeon Phi technology. The study also showed that the choice between one and the other of these two technologies has a significant impact on the ability to meet the demands of an Image Calibration Center and that choice was strongly linked with the programming language choice. So both technologies were analyzed and compared, the paper makes a focus on this trade-off.

2.1. GPGPU overview

GPGPU is now a wide-spread and proven technology to perform massively parallel calculations. The most popular GPUs are Nvidia’s ones with currently the Kepler architecture. AMD is present on the market too with GPU based currently on the GCN (Graphics Core Next) architecture.

The Theoretical Peak Performance for those accelerators are currently between 5 TFlops and 6 TFlops in single precision and more than 1 TFlops in double precision.

The frameworks available for GPGPU are:
- CUDA which is the NVIDIA’s C/C++ programming language extension for programming NVIDIA GPUs. The use of CUDA is spread widely through the academic world and industry. A clear advantage of CUDA is that it is the most feature-rich parallelization framework specifically designed for the most popular accelerator architecture. CUDA also comes with a significant number of high-quality accelerated libraries. The main disadvantage of CUDA is that it is only supported on Nvidia GPUs and not on AMD GPUs or INTEL MICs.
- OpenCL is an open standard for programming on CPU and GPGPU (compatible with Nvidia and AMD GPU). The main advantage of OpenCL is that it is an open standard and that its programs are portable between architectures. The main disadvantage of OpenCL that it comes with a relatively poor development tools environment. The other drawback is unlike the portability of binaries, the performances are not portable (necessity to adapt the code to obtain good performance on new architecture).
- OpenACC is an API (Application Programming Interface) for offloading work from a CPU to an attached accelerator. It is based on the same principles as OpenMP and uses simple compiler directives, interleaved in C/C++ or Fortran code, to direct the parallelization. OpenACC support parallelization on the host CPU and offloading to Nvidia and AMD accelerators. In terms of code portability, Nvidia seems to be geared more towards OpenACC than to OpenCL. The main advantage is the simplicity of the API, resulting in good development productivity. The disadvantage is a 10% performance loss over the CUDA (average generally noticed in reports).

In synthesis, the GPGPU leader is Nvidia, it is a wide-spread and proven technology, however the programming languages extensions are either proprietary (Nvidia) with full development environment, or open source but with few development tools and available libraries.

2.2. MIC overview

MIC is a recent technology developed by Intel as Xeon Phi with widespread commercial availability since end of 2013. The architectures of the Intel MIC Xeon Phi are quite different from GPGPU. The Xeon MIC has currently about 60 (depending the models) light-weight x86 cores that can each run 4 hardware threads resulting in a total of 240 threads and also each comes with a 512 bit vector unit allowing concurrent operations.

The Theoretical Peak Performance for those accelerators is currently more than 1 TFlops in double precision.

The frameworks available for MIC are:
- OpenCL with the same advantages and disadvantages than for GPGPU,
- OpenMP which is an API for shared-memory multi-threading in C/C++ and Fortran. OpenMP is widespread in the HPC (High Performance Computing) world and is supported on many platforms and in many compilers. The main advantages of OpenMP are the number of development tools available and that the code developed
with OpenMP for MIC also runs on CPU architectures, so it makes the code very generic regarding the computing facilities. The other advantage is the effort assumed to optimize the OpenMP performance on MIC resulting in an increase of performance on CPU (architecture very close).

In summary, only Intel proposes the MIC technology via Xeon Phi. The Theoretical Peak Performance of Xeon Phi is currently lower than the GPGPU in simple precision but quite equivalent in double precision. MIC allows the use of OpenMP API which is a very wide-spread language coming with number development tools and which allows to generate code for CPU and MIC.

2.3. GPGPU and MIC performances

Performances of a resampling IAS and elementary functions (Discrete Fourier Transform (DFT), dense matrix multiplication, matrix arithmetic) have been benchmarked on CPU, GPGPU and MIC architectures in order to consolidate the comparison.

First, a resampling CNES IAS (MARIO) already ported in OpenCL has been benchmarked on the cluster of the CNES data center. The benchmark was made on a node of the cluster to compare the execution time of the IAS on Nvidia K20 GPU and on Xeon Phi 5110P MIC. It was a quick benchmark, so the code initially optimized for execution on Nvidia GTX 480 has not been modified. The benchmark shows that the execution time is quite the same between K20 and Xeon Phi in the case of using with a large resampling filter (certainly because all the computations are made by the IAS in double precision).

In a second time, a comparative benchmarking of the Nvidia Kepler K20X GPU and Intel Xeon Phi 5110P MIC has been made on the facilities of the Antwerp University through popular linear algebra and image processing operations:
- DFT: using Intel MKL library (CPU/MIC) and cufft library (GPGPU),
- DGEMM: dense matrix-matrix multiplication using Intel MKL library (CPU/MIC) and cublas library (GPGPU),
- Matrix multiplication, image resampling, rotation, and histogram using ArrayFire library (CPU/MIC/GPGPU).

For the implementation, libraries available from Intel, Nvidia, and Accelereyes have been used.

The results of the benchmarks are:
- for small data sizes, the CPU remains the best choice, even when data-transfer, to and from the accelerator, is not taken into account,
- for larger data sizes both the MIC and the GPGPU outperform the host CPUs up to 7 times for MIC and 10 times for GPGPU. However, this was not true for the DFT when data transfer times were included. The overhead of transferring data to the MIC could be largely eliminated by asynchronous data transfer, at least when multiple data sets have to be processed.

In addition ten recent scientific publications (from [3] to [12]) on accelerating code with the Xeon Phi have been reviewed to achieve the comparison. The more valuable insights that have been learned from these publications about performance is that the GPU outperforms the Xeon Phi in most cases, but not dramatically (e.g. 30%). There are, however, three categories of algorithms for which the performance gap is significantly larger (e.g. 2-3 times):
- Code which relies heavily on bilinear interpolation. GPUs can compute interpolation in hardware while on the MIC this has to be done in software.
- Code which relies heavily on atomic operations. Much of the performance of the MIC comes from vectorization, but unfortunately the MIC does not support vector atomic operations.
- Code that has an irregular data access pattern. This is mainly due to the fact that the GPGPU does a better job at latency hiding due to more efficient scheduling of larger amounts of threads (no cost associated with context switching).

2.4. Trade-off between GPGPU and MIC

Finally, it was concluded that the MIC technology is best suited for a CNES Image Calibration Center. Indeed, the performance of GPGPU technology in single precision is not useful because the CNES IAS do calculations only in double precision. In addition, the ability to program with the OpenMP standard allows on one hand to generate generic CPU / MIC code that is compatible with current and future multi-core platforms and on the other hand the fact that OpenMP is a pragma API and is widespread in the HPC world would allow to reduce maintenance effort. Moreover, in-socket versions of the accelerator are planned by Intel that could eliminate the data transfer bandwidth bottleneck.

3. SOFTWARE EVOLUTION

To take advantage of this new technology, the portage of the following CNES IAS, which are the more CPU intensive processes, to OpenMP has begun:
- MARIO: geometric transformations on images, with or without modification of axes [13],
- ASTRIDZ: image restoration and image zoom by Fast Fourier Transform (FFT) [14],
- EGALSTAT: image radiometric statistic equalization,
- REECHFUSION: image resampling and image pan-sharpening [1] [2].
All the calculations in those IAS are done in double precision.
The design phases determined that only the part of the code which is the most time consumer and the most parallelizable and vectorizable will be ported in off-load mod. Also, the asynchronous data transfer to or from the MIC will be used to mask the transfer time with calculation time.

The development will be iterative with an analysis of the optimizations that have to be done for each iteration. Each IAS has different use cases depending on the filter dimension, the geometric transformation, etc. The first iteration will focus on functionnal porting and should give the following range of speedup:
- REECHFUSION: between 3.4 and 5.4,
- MARIO: between 5.7 and 9,
- ASTRIDZ: between 5.3 and 7.3,
- EGALSTAT: at least 4.6.
The next iterations will focus on IAS performances.

4. CONCLUSION AND OUTLOOK

Following the analysis of the different components, the final architecture of an operational cluster for an Image Calibration Centre was defined. The main characteristics identified are the following.

4.1. Hardware Platform

The main hardware component for the cluster is the dual-socket Intel Xeon E5-26xx v2 processor platform. It is by far the most popular HPC-platform with superior bandwidths between the system components, i.e. between processors, memory, accelerators, and storage. This dual-socket CPU platform can provide a double precision theoretical peak performance of more than 0.5 TFlops. The Xeon E5 platform has support for both the latest Intel many-core and Nvidia GPU accelerators.

Based on the benchmarking and the study of the recent literature, it was concluded that the Intel Xeon Phi 5110P accelerator platform is the most suitable for Image Calibration Center. It has a favorable double precision theoretical peak performance of 1.01 TFlops per accelerator. It is highly energy efficient at 4.5 GFlops/W, while still reasonably cost efficient at 0.37 GFlops/USD. In practice, a performance benefit is observed of 2-3 times per accelerator over a dual-socket Xeon E5-system. Taking the constraints of the conventional server solution into account, the solution adopted is each CPU-socket is paired with a single Xeon Phi accelerator.

4.2. Software Platform

OpenMP with vectorization to fine-scaled parallelization is recommended as previously presented.

For workload management on the cluster, the MOAB/Torque or OpenGridScheduler distributed resource and workload managers are recommended. It supports a good set of standard features, such as job priorities, reservations, check-pointing, and pre-emption, and has out-of-the-box support for scheduling accelerated workloads for Nvidia GPUs and Intel Xeon Phi.

4.3. Outlook

Using this cluster-proof architecture during the flight acceptance for the image calibration will help to validate this new approach of computing cluster and suggests the possibility of deploying these technologies in operational image processing centers.

5. REFERENCES