Research Brief

Parallelizing Applications for Multi-core Processors: The RapidMind Solution

Executive Summary

RapidMind is a maker of an application development environment that helps application programmers "parallelize" their applications. (Parallelization separates applications into segments that can be processed individually — the results of each calculation are then combined to achieve the desired outcome). By parallelizing applications, developers are better able to exploit the new generation of multiple core (multi-core) processors that have come to market — and, accordingly, greatly improve application performance.

At present, RapidMind's development environment is focused on:

- Windows and Linux operating environments;
- Graphical processing units (GPUs) from NVIDIA and ATI; and,
- Cell processors (Cell Broadband Engine or "Cell BE" processors).

And *Clabby Analytics* expects RapidMind to introduce support for commodity, multi-core, general-purpose processors from AMD and Intel by the end of 2007.

We also note that recent investments in RapidMind by Ventures West, EdgeStone, BDC Venture Capital should help the company grow rapidly (RapidMind just received $10M Canadian in investment in order to expand operations and broaden its product offerings — giving RapidMind a total of $11.3M to date). For those concerned with buying important products from small vendors, this investment should go a long way toward alleviating any vendor longevity fears.

In this *Research Brief, Clabby Analytics* examines the "RapidMind Development Platform". What we find is that this platform provides an easy-to-use set of programmatic interfaces designed to simplify parallelization tasks for application developers. And we find that the underlying RapidMind Development Platform engine not only helps improve application performance, it also allows for easy application portability between disparate processors.

We conclude that RapidMind's Development Platform should be examined by application developers who are writing custom code and wish to exploit GPU and/or Cell processors — as well as independent software vendors (ISVs) who are looking to greatly improve their software's performance on multi-core devices.

Introduction

A few years back, the ability to increase microprocessor performance simply by increasing clock rate and decreasing processor size through nanotechnology hit a wall.

What happened was that processors were becoming too hot to handle — and electrons were having trouble getting through their assigned gates. The laws of physics stepped-in, hampering the traditional "ratchet-up-the-clock-rate" designs.

To remedy this situation, microprocessor makers have turned to the use of multi-core processors — imprinting multiple processors on a single die. Using this design approach, microprocessor makers can link together a few, dozens, or ultimately hundreds or thousands of processors, and thus constantly increase processing power.

But, although processing power is increasing by leaps and bounds, software is not keeping pace. In order to exploit multi-core processors, software needs to be written in such a way as to be able to exploit unused processing cycles on available, unused or partially used processors on the die. In order to do this, software needs to be parallelized — separated into segments that can be processed across multi-core processors — and then reassembled to produce an end result. By architecting software in this manner, *application developers can expect to see improvements of up-to 30X over their current single-threaded (linear, serial processing) approaches on currently available Cell and GPU processors!* And this is where RapidMind's software comes in.

What RapidMind does is it offers a development environment that simplifies the task of parallelizing applications to take advantage of underlying, multi-core hardware. RapidMind does this by using a simple applications program interface (API) that enables application developers to present their code to RapidMind's Development Platform where it can be parallelized and driven across multiple, different hardware types. The following section takes a closer look at this development engine.

RapidMind's Development Platform

Probably the best way to look at RapidMind's Development Platform is to separate it into three parts: 1) RapidMind's API; 2) the RapidMind Development Platform; and, 3) the underlying hardware support (see Figure 1 — next page).

The RapidMind API

From an application development perspective, application programmers continue to write their applications in C++. Developers then need to identify the sections of their applications that could benefit from acceleration on multi-cores. Within these sections, developers need to replace numerical types representing floating point numbers and integers with the equivalent RapidMind platform types.

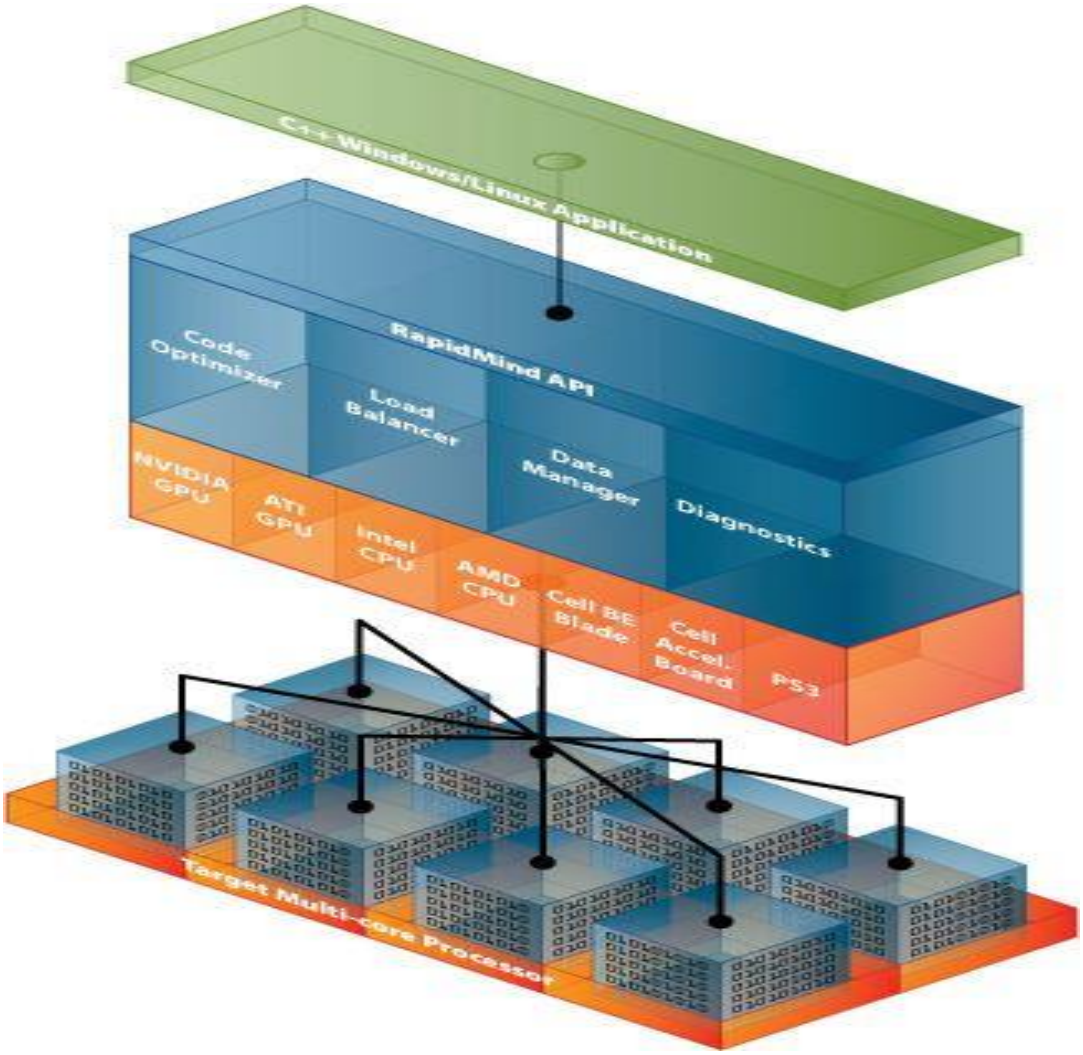
The RapidMind Development Platform

Once the developer replaces the numerical types with RapidMind platform types, the application can be run on RapidMind's development platform. As the application is run, sequences of numerical operations invoked by the user's application can be captured, recorded, and dynamically compiled into a parallelized program object. This platform layer also provides code optimization, load balancing, data management, and diagnostic services.

Underlying Hardware Support

The RapidMind platform runtime then drives managed parallel execution of program objects on the target hardware platform, which can be a GPU, the Cell processor, or eventually a multi-core CPU.

Figure 1 — The RapidMind Development Platform



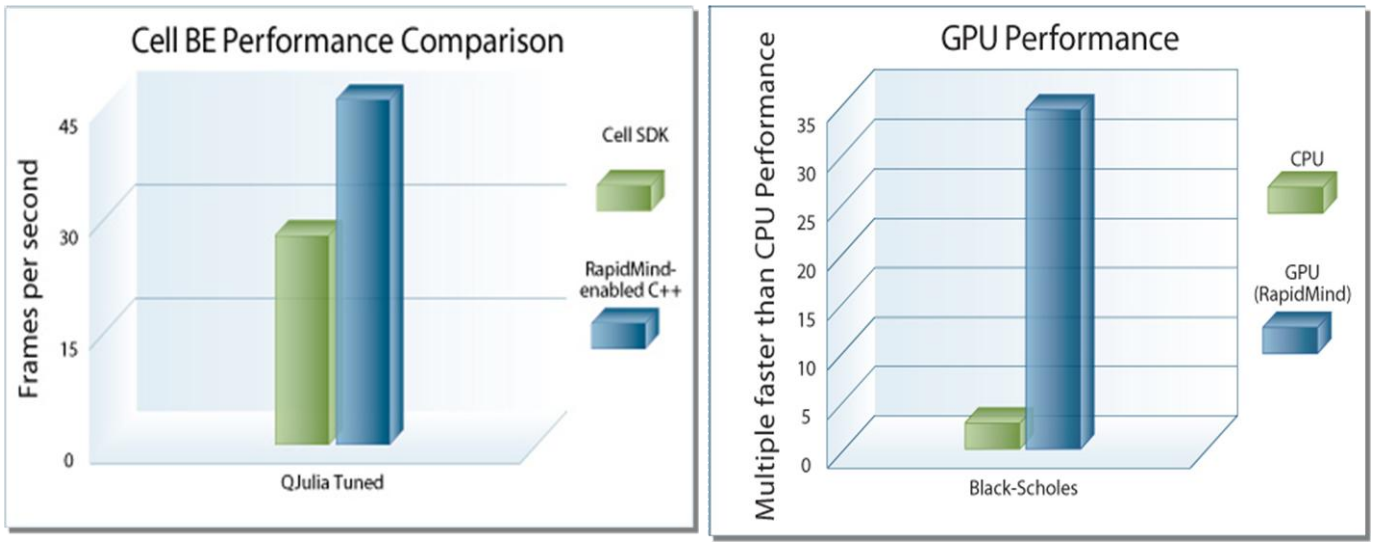
Source: RapidMind — May, 2007

As stated at the outset, RapidMind's approach to parallelizing applications is straightforward. Programmers write their applications in C++ language; they use an API to make calls to the RapidMind Development Platform. This platform parallelizes, compiles, and creates a run-time that drives underlying hardware. RapidMind's design is just this simple.

The Benefits of the RapidMind Approach

The biggest benefit of RapidMind's approach to application parallelization is the potential for massive performance gain. As stated previously, *RapidMind customers are seeing gains that range from 3X to 30X over traditional uniprocessor, single threaded approaches.* Figure 2 shows some of the performance gains that have been experienced using RapidMind's solutions.

Figure 2 — Performance Gains Using RapidMind Development Platform



Source: RapidMind, May—2007

Another leading benefit is RapidMind's simplicity of design. RapidMind's design makes its Development Platform:

- *Easy-to-learn* (programmers can write their applications in standard C++ — a language already familiar to the lion's share of today's application developers. To use RapidMind's Development Platform these developers need only learn a new set of application program interfaces in order to parallelize applications. The rest of the parallelization work is automatically done by the Development Platform); and,
- *Easy-to-use* (C++ developers can learn to use this product in a half-an-hour or less).

Another important aspect of RapidMind's approach is that the RapidMind Development Platform is processor agnostic. Application developers do not have to understand how to write low-level calls to a specific hardware environment or processor image. Instead, the RapidMind Platform handles application-to-underlying-hardware abstraction — making it possible to write an application once, but run it on many processor types. And because RapidMind's Development Platform is processor agnostic, application code that is written to drive over a GPU can automatically be compiled to run on a Cell processor — and ultimately RapidMind will also enable that code to drive general purpose AMD and Intel CPUs. Application portability is thus, another key benefit of RapidMind's Development Platform.

When Not to Use RapidMind's Development Platform

RapidMind's Development Platform is designed to help applications that can be parallelized exploit multi-core processors. So, obviously, applications that are written to execute in a linear, serial fashion are not candidates for RapidMind's platform solution.

Further, information technology (IT) executives who don't want to standardize on RapidMind APIs are not candidates for using this product set. (In the IT industry there is a constant fear of vendor lock-in; and standardizing on RapidMind's APIs means that custom code developers as well as ISVs could become dependent on RapidMind over time for application parallelization.

Clabby Analytics discounts this concern on a number of bases: 1) if something else better comes along later down the road, an enterprise should be able to easily switch to it simply by using a new set of APIs; 2) RapidMind could, over time, become the actual standard for application parallelization (and thus not be a lock-in threat); and, 3) RapidMind's recent investment funding assures the company's longevity for several years — reducing another major lock-in concern.

Which Applications Can Benefit Most From RapidMind's Development Platform

Parallelizable applications can readily be found in the high-performance computing (HPC) and enterprise application markets, in such disciplines as:

- 3D visualization;
- Broadcast-quality encoding;
- Medical imaging;
- Film and television content generation;
- Image and signal processing;
- Financial analysis;
- Seismic analysis;
- Database transactions; and,
- Enterprise search.

Clabby Analytics notes that as the cost of processing comes down due to the introduction of more powerful, less-expensive multi-core systems, new applications can and are being introduced into all of the above markets. Supercomputer-like applications are now wending their way into the business marketplace (as opposed to primarily the scientific and engineering communities). Further, only the surface has been scratched in the use of GPUs and Cell processors in the high-performance computing marketplace. There is a tremendous amount of up-side, and accordingly, significant growth potential for a company like RapidMind as both HPC and commercial multi-core use expands.

Summary Observations

RapidMind's solution aims to make it easy for application programmers to parallelize their applications — enabling application performance to improve by as little as 3X to as much as 30x over traditional single-threaded, uni-processor applications. And all indications from a massive, 1,000 developer beta test the company has just conducted, shows that the company is succeeding in this goal (incidentally, this test also shows how robust RapidMind's software actually is).

It should also be mentioned that RapidMind was awarded the IEEE (Institute of Electrical and Electronics Engineers) *Spectrum* magazine 'One of Five Best

Technologies of the Year' award this year. Excitement over this product is building throughout the development and engineering communities.

RapidMind's solution — the company's RapidMind Development Platform — is hitting the market at just the right time. There is a need by custom code writers as well as ISVs to exploit unused computing cycles on multi-core systems. RapidMind's product offering does just that. And with growing demand in both HPC and high-performance enterprise/business computing market segments, RapidMind's has the potential for strong and solid growth.

It must be noted that RapidMind's solution is not for everyone at this point. At present it runs on GPUs and Cell processors. Support for Intel and AMD dual-, quad- and beyond-processors is needed to broaden its market appeal. But support for Intel and AMD processors is forthcoming — and the existing Cell and GPU markets are also growing rapidly — so RapidMind is clearly well-positioned for growth. It should also be noted that not all applications can be parallelized — some run in a linear, serial fashion and would require major rewrites in order to exploit parallelism. But for those that can benefit from automatic parallelization, RapidMind offers the potential to deliver very large performance increases by exploiting multi-core processors.

In summary, RapidMind removes the complexity involved in writing parallelization routines into application code and completely abstracts the underlying hardware from the application programmer. It is easy-to-learn and; it enables application code to be readily migrated across multi-core processors (so it is highly portable); and this code is proven to be robust. Custom code developers and ISVs looking for a way to parallelize their applications should definitely evaluate RapidMind's Development Platform as a means to exploit multi-core processors.

Clabby Analytics
http://www.clabbyanalytics.com
Telephone: 001 (207) 846-0498

Clabby Analytics is an independent technology research and analysis organization that specializes in information infrastructure and business process integration/management. Other research and analysis conducted by Clabby Analytics can be found at: www.valleyviewventures.com.