THE CLIPPER GROUP

Navigator™

Navigating Information Technology Horizons℠

Published Since 1993 Report #TCG2006032 May 7, 2006

The Intelligent Person's Guide to Shrinking Application Process

Analyst: Anne MacFarland and Joseph De Natale

Management Summary

Time is the measure and often the enemy of business. There has always been a need to do things faster than what is generally available. We have become used to the demands of continuous, and yes, uncomfortable rapidity in our lives, but are now faced with a need for speed that, in some situations, once seemed impossible to achieve. This is particularly true in the world of Information Technology, where the pressures to respond to the demands of the enterprise have increased due to the global economy, competition, government, social pressures and an increased dependency on contingency planning and testing. In complex computational processes, such as those that occur in financial services, pharmaceutical, and earth sciences, the need for more rapid time-to-completion is paramount. It is neither wise nor practical to address this issue by adding expensive computational power, i.e., throwing hardware at the problem, to make those gains. Constantly modifying code to optimize the software for speed makes no practical sense, in particular when optimizing the application for one type of problem solution would be different for another within the same application area, because of a different set of input variables.

A practical alternative is available, one that has been proven in a range of complex applications. ASPEED Software of New York, NY, whose motto is *ASPEED - Makes Fast*, *Faster*, offers its application software, self-optimizing product called ACCELLERANT. The product can be applied with ASPEED SDK components and/or a self-optimizing run-time component. The SDK is used to identify the iterative areas and related data objects, which can be applied without changing the business logic, logical flow or the processes of the application. ACCELLERANT uses disciplines similar to grid computing, complemented by a real-time optimization engine, to parallelize and distribute the application at a fine granular level, keeping the logic of application process intact and secure. Use of ACCELERANT shrinks time-to-completion by a factor of the number of CPUs that can be dedicated to the task. The application code remains intact because of the manner in which ACCELLERANT has been designed. From the programming and directory standpoint there is only one copy of the program. It spawns clones of itself at run time.

By shielding the application code, enterprises can make use of the methodology of ACCELLERANT confident that application integrity is maintained and, equally important, substantial reductions in compute time are achieved without the addition of more hardware. As a result, a business can gain competitive advantage using ASPEED's ACCELLERANT product. Business advantage aside, when used in the drug industry, ACCELLERANT can markedly improve the development of new drugs, an important social factor in the world today where epidemics and pandemics are potential realities. To find out more on the ACCELLERANT product, read on.

IN THIS ISSUE

How ASPEED ACCELLERANT Works

First, the application must be analyzed. Some applications are simply a bounded collection of organized independent and dependent steps.

Some of these long running applications are designed to be run as a single thread. Some may have already been parallelized to run on multiple processors on the same system. Others may also have been designed to run across multiple processors and systems. These are referred to as distributed or, depending on the number of CPUs and systems involved, as *massively parallelized applications*. The technical challenge is that single thread applications need to be parallelized and/or distributed, parallelized applications need to be distributed, and distributed applications need to be re-designed to break the 32- to 64-bit system barriers, in order to achieve the economies of the Linux x86 technologies. Running the application in parallel across multi-processor servers or other forms of compute grids can cut time-to-completion substantially.

However, there are many applications (programs) that cannot be easily compartmentalized and scheduled. There are other applications (programs) that have already been compartmentalized but still take too long. These are ASPEED's targets.

ACCELLERANT provides the intelligent distribution and optimization management to achieve optimum performance across these configurations for an individual application by managing the performance of the parallel components of the application across the dedicated shared or grid provisioned engines.

The trick to accelerating the time-to-completion of these individual applications is to work at a more granular level. Even the most serial application will have recursive loops within its execution path. The trick is to parallelize these loops to reduce the time needed to execute the whole – without changing the application.

Touching the applications is risky, of course, particularly for legacy applications that have evolved over time. ASPEED's ACCELERANT does not change the business logic or the processes of the application, just the topography of where it is run. ACCELERANT identifies the areas of recursive computation – the loops – and uses disciplines similar to distributed computing to parallelize the application at a fine granular level, keeping the logic of application process intact ands secure.

The parts of the code where iterative processes occur are identified either by using a popular profiler or identified by the application analyst. Once identified, the contributions of the iterations are measured (or estimated) and best practices rules are applied to predict the run time reduction. The iteration(s) with the greatest benefit are chosen for the initial phase of computational improvement.

ASPEED does this in two ways. For simple loops with no dependencies in linear processes, it uses a *Linear Adapter*, basically wrapping the code so that it can run in parallel. For non-linear processes, where one calculation provides the basis for the next calculation, ASPEED has an *algorithm-specific* API (Application Protocol Interface). As a run-time component, this algorithmic-specific API interfaces with the ACCELLERANT run-time engine, which automatically deploys and launches the application clones (no infrastructure service prerequisites), distributes the ranges and the data it uses, monitors the loop progress, and dynamically rebalances the load to achieve optimum completion time. ACCELLERANT aggregates and returns the result to the original process location. The clones self-destruct and the originating application continues or completes.

ASPEED handles applications that use *mesh* or *matrices* with complex algorithms in a similar way, but with APIs that place more emphasis on the mapping of the data to the various parallel processes and to the placement of synchronization dependencies in the code and transparently "reaching" adjacent data cells. *Tree structures* also pose a challenge. ACCELLERANT dissects them into smaller trees at the major branches, to enable a controlled degree of parallel processing while avoiding the exposures introduced by recursively-invoked resources tying up resources and the network.

For customers who don't have the source code for their application, ASPEED offers an *Application Manager* – a script driven ACCELLERANT engine that runs cloned copies of the application controlling the elapsed time by appropriately portioning the input or control data each clone operates on. This provides a coarser grain parallelization but does not require any application program changes. ASPEED has also developed a *Workload Balancer*, a tool that can take advantage of additional machines (up to a set number) should they become available.

ASPEED is the dispatcher that makes the application run most efficiently. ASPEED's Kurt

Ziegler, executive VP of ASPEED likes to call what the product does an application grid, as opposed to a processor grid or a database grid, because the focus is on the individual application performance and its dependencies and requirements.

Use Cases

Pharmacology

In the drug development and approval phases, the researchers must identify relationships between their subject group, characteristics, and reactions as individuals and as a group. This means that the analysis models must use regression analysis algorithms where each point in the regression is dependent on the previous point. As such, it is inherently not parallelizable. The comprehensive analysis that underlies drug development today utilizing NONMEM is exhaustive and highly recursive. The time-to-completion of the drug discovery analysis, when you factor in the cost of the technology and of the scientists and doctors who must wait on the results of the analysis, is what adds a lot to these costs. Reducing the time-to-completion of the analysis reduces the development costs of the drug.

Financial Services

In financial services, the time challenge is somewhat different. Here, different investment scenarios can have fleeting validity in an environment where the values of the investment vehicles are subject to change (due to a large number of factors), at any time. Here you are not wringing out costs, but accelerating process to fit a window or relevance – rather like the velocity a plane needs to take off successfully.

Data Issues

ASPEED *ACCELERANT* uses the tools of serialization/deserialization, compression, and double buffering to move transparently the data to where it needs to be, but just arbitrarily splitting and cloning data adds to the overhead. The choreography must be tied closely to the infrastructure topography that is available to the process to be effective. For linear applications, ACELLERANT dynamically learns and adapts the workload in real-time to achieve optimum response time. For the more complex algorithms, ACCELLERANT offers pre-conditioning APIs that space the distributed memory to achieve a best fit for each run. Because of the abstraction that the API offers the application, the pre-conditioning algorithm can be changed without opening the application again. Hard coding the pre-conditioning was one of the primary limitations of early-distributed system application implementations. When the data is being captured in real time, predictive analysis doesn't work. In such cases, dynamic application algorithms are needed to adapt where loops are run, and move data as needed, as the process is running. ASPEED has experience with these scenarios.

The Challenge of Random Numbers and Distributed Processing

Many processes make use of random numbers. A centrally held table can be a memory problem, and if you distribute it, the inherent structure of the distribution process renders the numbers less than random. ASPEED recommends placing the random number table into a cached file that can be called as needed. They offer three random number algorithms, one repeatable, one that assures integrity, and one a traditional random.

The Acceleration to Expect

The overall acceleration factor provided by ASPEED is a function of actual processing. It is dependent on whether the loop identified for improvement is totally CPU bound, or a mixture of moderate CPU and moderate I/O, or moderate CPU and heavy I/O. Once a loop has been determined to be a candidate for acceleration, and the run time is calculated, the reduction in run time, if CPU bound, will be reduced by the number of CPUs assigned, that is:

$$\textit{accelerated run time} = \frac{\textit{original run time}}{\textit{\# of CPUs assigned}}$$

For those processes that are a mixture of CPU and moderate I/O, the divisor will range between 1.5 and two times the number of CPUs available. For a process that is heavily I/O bound and with light CPU requirements, the divisor can approach three times the number of CPUs available for processing.

If the process is non-linear, as with the NONMEM application used in pharmaceutical and bio-medical enterprises, the amount of acceleration to be garnered is specific to the application. The improvement comes in stair-steps, as more processors are used, not as a smooth curve.

How You Get ACCELERANT

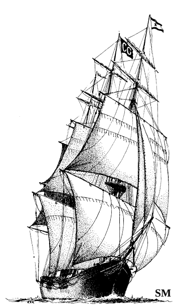
The procurement process should not be undertaken lightly or without cause – and it is priced accordingly. *ACCELLREANT* is priced by the instance – the process splits. Two instances will split the elapsed time needed to complete a process in half. Four instances will reduce the time to completion further. *ACCELERANT* is available for an entry fee of $100,000, which includes 50 instances. Additional instances incur an additional cost of $1000/instance. This tailors the cost to the complexity of what must be accelerated.

Most procurements start with a proof-of-concept engagement. The feasibility study can be done up front in the ASPEED lab, if the customer doesn't have excess resources. For the risk averse, a Professional Services contract can implement a complex solution, and the customer need only buy it, if it works.

Conclusion

Process acceleration is somewhat like dieting. There are the diets that just consist of very careful eating (optimization) which work for specific food addictions or sensitivities, but don't generalize to work for the majority of people who try them. Then there are diets that are good for taking off a few pounds, and do nothing substantive to further weight loss over time. Finally, there are the serious diets, diets that change your lifestyle and metabolism, and require a closer look at what and when you eat and how your body responds to food. These require more attention, and usually require in a permanent change of eating habits.

ASPEED ACCELERANT takes an approach like these serious diets. Therefore, if you have a problem with process duration that cannot be addressed by traditional hardware optimization, or coarse-grained parallelization, ASPEED may have precisely the approach you need.



About The Clipper Group, Inc.

The Clipper Group, Inc., is an independent consulting firm specializing in acquisition decisions and strategic advice regarding complex, enterprise-class information technologies. Our team of industry professionals averages more than 25 years of real-world experience. A team of staff consultants augments our capabilities, with significant experience across a broad spectrum of applications and environments.

- ***The Clipper Group can be reached at 781-235-0085 and found on the web at www.clipper.com.***

About the Author

Anne MacFarland **is Director of Data Strategies and Information Solutions for The Clipper Group.** Ms. MacFarland specializes in strategic business solutions offered by enterprise systems, software, and storage vendors, in trends in enterprise systems and networks, and in explaining these trends and the underlying technologies in simple business terms. She joined The Clipper Group after a long career in library systems, business archives, consulting, research, and freelance writing. Ms. MacFarland earned a Bachelor of Arts degree from Cornell University, where she was a College Scholar, and a Masters of Library Science from Southern Connecticut State University.

- ***Reach Anne MacFarland via e-mail at Anne.MacFarland@clipper.com or at 781-235-0085 Ext. 128. (Please dial "128" when you hear the automated attendant.)***

Joseph S. De Natale **is Director of Enterprise Systems Planning with The Clipper Group.** He brings more than forty years of experience in the data processing field with particular emphasis on systems management and application development on large-scale mainframes. Prior to joining The Clipper Group shortly after its founding, Mr. De Natale was an independent consultant with Ropes and Gray, Attorneys at Law, where he provided expert opinion on data center management for a civil case. He later joined International Data Corporation (IDC), as a senior consultant and analyst, covering banking systems, data center management software, and large systems computers and storage. Formerly, Mr. De Natale spent eleven years at Citicorp Information Resources (CIR) as CIO of the Boston Data Center, where he managed the support of over 200 outsourcing contracts for thrift institutions. Earlier, he was MIS Director for the Lahey Clinic, and prior to that was a Project Manager for Computer Sciences Corporation, where he was involved with NASA and FAA outsourcing contracts. Previously he was Director of AVCO Computer Services for fourteen years. At AVCO, in addition to being responsible for all internal data processing, he initiated the marketing and sales of computer services to commercial clients. Mr. De Natale began his career with Pratt and Whitney Aircraft as a programmer of nuclear physics and business applications. During his career, Mr. De Natale was involved in the evaluation, installation, and operation of large-scale mainframe systems and for the development of software application systems. Mr. De Natale earned a Bachelor's and Master's degree in Mathematics from Boston College. During his period at AVCO, he was selected by AVCO to attend the Northeastern University Management Development program, a co-op program covering an MBA curriculum.

- ***Reach Joe De Natale via e-mail at denatale@clipper.com or at 781-235-0085 Ext. 124. (Please dial "124" when you hear the automated attendant.)***

Regarding Trademarks and Service Marks

The Clipper Group Navigator, **The Clipper Group Explorer**, **The Clipper Group Observer**, **The Clipper Group** ***Captain's Log***, **The Clipper Group Voyager**, and *"clipper.com"* are trademarks of The Clipper Group, Inc., and the clipper ship drawings, *"Navigating Information Technology Horizons"*, and *"teraproductivity"* are service marks of The Clipper Group, Inc. The Clipper Group, Inc., reserves all rights regarding its trademarks and service marks. All other trademarks, etc., belong to their respective owners.

Disclosure

Officers and/or employees of The Clipper Group may own as individuals, directly or indirectly, shares in one or more companies discussed in this bulletin. Company policy prohibits any officer or employee from holding more than one percent of the outstanding shares of any company covered by The Clipper Group. The Clipper Group, Inc., has no such equity holdings.

Regarding the Information in this Issue

The Clipper Group believes the information included in this report to be accurate. Data has been received from a variety of sources, which we believe to be reliable, including manufacturers, distributors, or users of the products discussed herein. The Clipper Group, Inc., cannot be held responsible for any consequential damages resulting from the application of information or opinions contained in this report.