



Citadel Accelerates Compute Grid Processes with GPU Deployment through **Parallels Workstation Extreme**

Getting the Most from Existing Assets

Citadel is one of the world's leading financial institutions, with a diverse business platform that includes asset management, fund administration and investment banking. This work involves building complex financial models to represent the performance of a financial asset, portfolio or project, constantly requiring millions of complex calculations to be processed.

Performing these calculations requires a significant amount of compute processing power. Citadel could address this need by deploying dedicated servers in their datacenter, but this approach is very expensive and entails a considerable management burden for the system administrators. Instead, Citadel decided to create a compute grid or "carpet cluster" using the existing desktop computers being deployed across the company.

This is possible because today's desktop computers have a computational capacity that far exceeds the needs of a typical user. This excess capacity of workstation hardware resources is common across many industries. In Citadel's case, it provided an opportunity for the IT department to use Parallels' virtualization technology to effectively run two computers on each physical machine already deployed. The "host" computer would continue to be used in the traditional way by the end-user, providing email, internet access and use of various office and productivity applications.

The second computer – or "guest" – is a virtual machine that runs in the background without affecting the performance or user experience, so the end-user sees no impact on their regular workflow. By connecting these virtual machines to each other, and through the use of sophisticated proprietary job dispatch software, Citadel has created a powerful compute grid called the Message Dispatch Layer (MDL) that effectively farms out hundreds of millions of calculations a day across all of their workstations.

In this way, Citadel has been able to more fully utilize its existing hardware, and increase the return on the company's ongoing desktop investments, saving millions of dollars that would otherwise be spent to deploy a specialized server-side solution in the datacenter. As the calculations are spread across thousands of different machines, the time-to-results is also improved, increasing the company's efficiency.

Industry:

Financial Services

URL:

<http://www.citadelgroup.com/>

Challenges:

- Millions of complex financial calculations to process
- Server-side solutions are very expensive and management-intensive
- Lack of GPU support in traditional virtualization solutions

Solution:

- Proprietary Message Dispatch Layer (MDL) compute grid
- Parallels Workstation Extreme with support for Intel VT-d for full GPU acceleration
- Dedicated GPU for both host and guest workstations

Benefits:

- Near-native performance for GPU-based computing
- 80% savings on hardware investments
- Savings of \$1.9million per year, every year

Configuration:

- Dell Precision T5500 with NVIDIA Quadro FX 5800
- Parallels Workstation Extreme

Greatest Return on Investment

Since inception, the MDL compute grid has gone through many revisions. During the most recent revision, Scott Donovan, System Architect for the Global Architecture Group, started to investigate how Citadel can take advantage of the massive computing capabilities of graphics processing units (GPUs).

Donovan explains: “We knew that GPUs could significantly accelerate computations, but we did not want to have to purchase specialized GPU servers. Essentially we needed the ability to integrate the GPU into our grid, but traditionally GPUs and virtualization technologies have not been a good match. This is because most virtualization solutions are not able to recognize, let alone directly assign, a high end GPU.”

Having worked with Parallels’ virtualization solutions in the past, Donovan was very interested when his hardware vendors introduced him to Parallels Workstation Extreme as a cost-effective way of addressing the problem. Parallels Workstation Extreme is a powerful, next-generation virtualization platform that gives end-users dedicated graphic and networking resources for both host and guest workstation environments.

“Parallels’ solution allows us to do GPU-based computing with near-native performance at a fraction of the cost, while actually gaining more control over the system itself.”

*– Scott Donovan,
System Architect for the Global Architecture Group at Citadel*

Donovan explains, “Parallels Workstation Extreme leverages Intel’s VT-d technology and is able to directly assign a dedicated GPU to the virtual machine. Parallels’ solution allows us to do GPU-based computing with near-native performance at a fraction of the cost, while actually gaining more control over the system itself.”

Citadel has recently run its first pilot of the GPU-based computing approach, recording performance gains of 70x to 140x over a CPU-based approach, depending on the graphics card used. By integrating GPUs into their grid with Parallels Workstation Extreme, rather than using dedicated server-side CPU resources, Citadel stands to save more than 80% of traditional server deployment costs. In addition to buying the virtualization solution, they will need to purchase additional high-end graphics cards for each workstation, but this is a minor expense compared to starting from scratch with new, dedicated servers for computational intensive applications.

Banking on Parallels

“We’re effectively building supercomputers with Parallels Workstation Extreme and our existing workstation hardware,” says Donovan. “If you had 200 desktop workstations with Parallels Workstation Extreme and an additional NVIDIA Quadro 5000 in each, you would have 100 teraflops of computational capacity, landing you around number 70 in the top 500 supercomputers in the entire world. We aim to initially have 15 machines virtualized with Parallels Workstation Extreme and calculate that even this low number will save Citadel approximately \$2 million dollars per year, every year, over a CPU server-based deployment.”

“Parallels Workstation Extreme virtual machines are currently running at 98.7% efficiency of native performance because we can directly assign a GPU. That’s an amazing number for a virtualized environment,” says Donovan. “Given that Parallels software plus the NVIDIA Quadro GPU is less than 20% of the cost of deploying a server for the same task, I can definitely live with such a minor degradation! I would certainly recommend Parallels’ virtualization solutions; when we look at the costs associated with all the options, deploying Parallels is a no-brainer.”

Contact Parallels

For more information about Parallels Workstation Extreme, please visit www.parallels.com/extreme. Parallels offers volume licensing to business and educational institutions. Please contact your local Parallels sales representative for more information or contact us directly at:



Parallels
Phone: +1 425 282-6400
Fax: +1 425 282-6444
www.parallels.com