Image courtesy of Wolfram Research, Inc.

Microsoft Windows Compute Cluster Server 2003
Partner Solution Brief

gridMATHEMATICA™

On-Demand Supercomputing Multiplies the Possibilities

Microsoft® Windows® Compute Cluster Server 2003 and grid*Mathematica*™

Overview

Region: Global
Industry: Technical Computing

Partner Profile

Wolfram Research, Inc. is a global leader in technical computing software. Wolfram Research is the maker of *Mathematica*—the tool of choice for computational work at the frontiers of science, engineering and technology.

Business Situation

The challenge was to make the powerful computational resources of *Mathematica* available in a flexible system configuration and enable *Mathematica* to distribute complex calculations across clusters, multiprocessor and multicore machines, or computing grids.

Solution

Wolfram Research released grid*Mathematica*, and the Cluster Integration Package, which deliver an optimized parallel *Mathematica* environment for multiprocessor and multicore machines, clusters, grids and supercomputers. Combined with Microsoft® Windows® Compute Cluster Server 2003, grid*Mathematica* provides exceptionally high-performance computing on a platform that is cost effective, simple to deploy, easy to maintain and familiar to use.

Benefits

- Access to the world's largest algorithm collection in one integrated system
- Rapidly gain insight into compute-intensive problems
- Seamless, scalable processing from the desktop to the supercomputer
- Familiar Windows environment lessening the learning curve
- Leverage existing Windows infrastructure
- High-performance computing at lower cost

These two programs are good partners because they allow you to breeze through the housekeeping and get to the science you really want to do.

Dr. Yuko Matsuda, Assistant Professor, Global Scientific Information and Computing Center, Tokyo Institute of Technology

Grid and cluster technologies have become part of mainstream computing. While the availability of grid and cluster technologies has brought considerable capabilities to users, the increased administration efforts and programming complexity required by these technologies have significantly reduced their ease of use. Today's scientists, engineers and industry experts rely on the ability to rapidly process complex data sets, and compute-intensive problems. For most organizations, the challenge is to make the powerful computational resources and capacity of grid and cluster computing available to their users. Moreover, only a small number of the users are actual programmers. Instead they are research students, engineers, scientists, or business analysts who are unfamiliar with Linux or UNIX and have little or no experience deploying or administering a cluster, or they are experts that would rather spend time gaining insight into the problem at hand instead of writing programs to run their simulations. Microsoft Windows Compute Cluster Server 2003 and grid*Mathematica* provide a highly scalable and easily managed rapid development and deployment environment alternative to traditional parallel programming languages on clusters.

Windows Compute Cluster Server 2003

As access to grid and cluster technologies increases, fewer of those who would use the grid or cluster resources are actual programmers. The real challenge is to make high-performance computing useful for these folks, as well as the programmers.

Prior to the release of Windows Compute Cluster Server, setting up grid*Mathematica* on a cluster of desktops was complicated and required the use of non-native tools, none of which provided the coherent experience to which Windows users are accustomed.

Schoeller Porter, Senior Software Engineer, Wolfram Research, Inc

Situation

Grid and cluster technologies have become mainstream and their use across industries is continually increasing. The availability of grid and cluster computing brings considerable capability to its users, however its administration efforts and programming complexity have also increased, and its ease of use has decreased.

Many of the organizations envisioning the capacity of high-performance computing (HPC) have little or no experience in deploying and managing a grid or cluster technology. Moreover, only a small number of the potential users of the grid or cluster are actual programmers. Instead the users are research students, engineers, scientists, or business analysts. Many of these users are unfamiliar with Linux or UNIX, or they are scientists and engineers that would rather spend time gaining insight into the problem at hand instead of programming their simulations. "The real challenge is to make high-performance computing useful for these folks, as well as the programmers," says Schoeller Porter, Senior Software Engineer, Wolfram Research, Inc.

Wolfram Research, Inc. is the maker of *Mathematica*—the tool of choice for computational work at the frontiers of science, engineering and technology. *Mathematica* combines a numeric and symbolic computational engine, a graphics and visualization system, a programming language, and a documentation system to form a single, coherent and flexible computing platform.

In today's competitive markets, scientific insights, medical advancements, product innovations, and technological developments increasingly depend on the ability to rapidly analyze complex data and models. An increase in the complexity of the data has a corresponding increase on the turnaround time for computation. Many problems cannot be reasonably solved on a desktop computer, and high-performance computing (HPC) has become critical for reducing the turnaround time of compute-intensive problems. Recognizing the need for increased access to HPC capacity for its customers, the challenge for Wolfram Research was to make the powerful computational resources of *Mathematica* available in a flexible system configuration and enable *Mathematica* to distribute compute-intensive problems across clusters, multiprocessor and multicore machines, or computing grids. In response, grid*Mathematica* was released. grid*Mathematica* takes the capabilities of *Mathematica* and extends them to compute-intensive problems that can be solved by using a cluster or grid.

Microsoft Windows-based customers comprise a large part of Wolfram Research's overall customer base. Prior to the release of Microsoft Windows Compute Cluster Server 2003, "configuring grid*Mathematica* on a cluster of desktops was complicated and required the use of non-native tools, none of which provided the coherent experience to which Windows users are accustomed," says Porter. With the release of Windows CCS, the challenge for Wolfram Research was to enable a seamless user experience for its customers running grid*Mathematica* on a Windows CCS cluster.

Customers running grid*Mathematica* in a Windows environment require a cluster solution that delivers complete access to the *Mathematica* algorithms in a single interface, the ability to leverage their existing Windows infrastructure and expertise, easy cluster deployment and administration, and seamless scalable processing that is cost effective.

This is a big advantage for our Microsoft Windows customers and system administrators. It reduces the setup cost, time, and effort required to establish a cluster to a level that is very reasonable and attractive to more mainstream customers.

Roger Germundsson, Director of Research and Development, Wolfram Research, Inc.

Windows Compute Cluster Server provides our customers with an environment that minimizes the amount of work necessary to maintain a grid*Mathematica* installation.

Schoeller Porter, Senior Software Engineer, Wolfram Research, Inc.

Solution

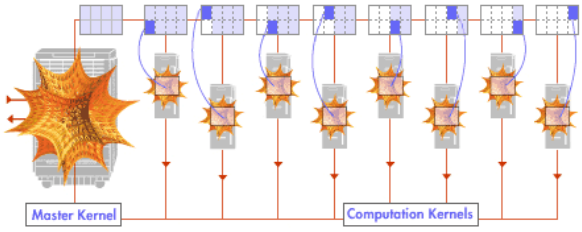
Microsoft Windows Compute Cluster Server 2003

Windows CCS is a high-performance computing platform for processing large-scale, complex computing problems. Windows CCS runs on commodity x64-based computers, and provides an inexpensive and highly scalable platform for HPC. Windows CCS is based on Windows Server 2003 Standard x64 Edition.

Windows CCS provides a familiar environment for accessing and managing a cluster. Administrative tasks such as cluster deployment, job submission and status monitoring are all simplified in Windows CCS. Windows CCS includes prescriptive setup procedures, a complete suite of management tools and an integrated Job Scheduler. Windows CCS leverages existing Windows infrastructure, so it works with Microsoft Active Directory directory service, can be monitored with Microsoft Management Console (MMC) and Microsoft Operations Manager (MOM), and uses tools such as Microsoft Visual Studio 2005 to support parallel job development and debugging.

grid*Mathematica*

grid*Mathematica* is a parallel computing environment that comprises a collection of *Mathematica* kernels (or compute kernels). The Master Kernel handles all input, output and scheduling functions and delegates *Mathematica* expressions to a pool of computational kernels, see diagram below.



The ease of development in grid*Mathematica* lowers the barrier to effectively using grid resources and provides a conduit for nonprogrammers to realize the gains in performance and capacity of grid technology. It provides an affordable, easy-to-use way to take full advantage of grid-computing hardware such as Windows CCS. “The strengths of grid*Mathematica* are its ability to interact with technologies common in a business environment—it can work with information stored in SQLServer or an Access database, access information in Excel or available through web services, interoperate with information systems in .NET, and interface with legacy applications. Because of this, grid*Mathematica* gives you the ability to implement innovative solutions within your existing Windows infrastructure,” says Porter.

Cluster Integration Package for grid*Mathematica*

Wolfram Research’s new Cluster Integration Package enables affordable, out-of-the-box operation of grid*Mathematica* on Microsoft Windows Compute Cluster Server 2003. The Cluster Integration Package for grid*Mathematica* makes it possible for Microsoft Windows users to immediately take advantage of HPC provided by Windows CCS. The Cluster Integration Package provides a direct interface to cluster management systems from within grid*Mathematica*. It enables the administration of the grid*Mathematica* compute kernels from within Windows CCS. grid*Mathematica* is an essential application for HPC, and with the Cluster Integration Package for Windows CCS, customers have a zero-configuration, seamless solution that meets their needs.

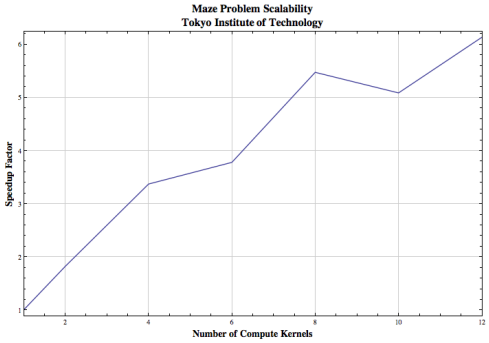


Figure 1 – Maze Problem Scalability. The solution is about 6 times faster as 12 kernels are used.

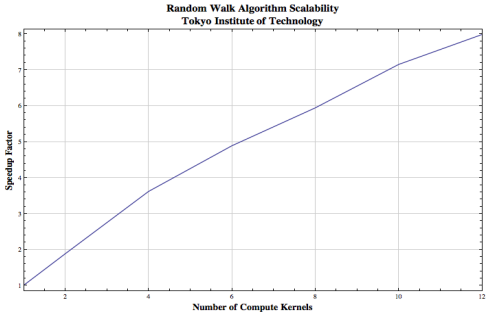


Figure 2 – Random Walk Algorithm Scalability. Results show the same algorithm as in Figure 1 when applied to a Walk problem. Speedup is approximately 8 times with 12 kernels. This problem type has less interdependency between the different kernels and so allows for more scalability. As a result speedup happens in a nearly straight line versus the deviations that are visible in Figure 1.

Success Story

The Tokyo Institute of Technology (Tokyo Tech) is a premier academic and research institute in Japan, and is one of the world's leading science and technology universities. The Global Scientific Information and Computing Center (GSIC) at Tokyo Tech, hosts the Problem Solving Environment Group. The mission of this group is to research, construct, and manage the new generation computational infrastructure for large-scale computational science.

Tokyo Tech hosts the fastest supercomputer in Asia, the 47.38 teraflop TSUBAME Grid Cluster, which is utilized extensively in grand challenge problem solving in a variety of research areas. However, an institutional system on such a scale can have a lot of overhead administration and bureaucracy associated with it as people compete to use its resources.

To expedite their research, GSIC Assistant Professor, Dr. Yuko Matsuda and team chose to deploy their own departmental cluster for HPC capacity. At the workgroup level, they administer their own cluster running Windows Compute Cluster Server 2003 and grid*Mathematica*2.

Dr. Matsuda and team are using grid*Mathematica* for multiple research projects. One project uses graph clustering techniques, specifically the Markov cluster algorithm, to identify associations between words in the Japanese language. The computational requirements of a project on this scale are intensive. A second project aims to calculate a complete model of quantum physics. As Dr. Matsuda explains, this project "investigates the roots of the coefficients of high-order polynomials representing physical systems, which can be used in applications such as simulations of biological ecosystems. These polynomials may be up to degree 300 and necessitate high-precision computation to solve the equations within the desired error tolerance. For a problem of this scale, computation time and memory requirements grow exponentially with the system size."

A single iteration requires 16GB of memory and one CPU-hour of computational time. To solve the system to the desired accuracy requires hundreds of thousands of iterations. Figures 1 and 2 display results related to a third project led by Dr. Matsuda to develop a parallel algorithm. The results demonstrate that as compute kernels are added to a grid*Mathematica* session running on a Windows CCS cluster, significant speedup is achieved.

The Tokyo Tech Windows cluster is a mix of different systems from different vendors using different processors, but in total it has 270 processor cores. A combination of systems and processors is fairly typical of most organizations that don't have the advantage of buying all new machines.

The installation of grid*Mathematica* on Windows CCS is simple to perform, explains Matsuda, "the packages load in easy and logical steps, allowing the user to be up and running in no time. By combining Windows CCS and grid*Mathematica*, users can launch remote math kernels in only two steps. There isn't intensive code and programming to coordinate the interface. The Compute Cluster Job Scheduler provides a smooth interface to submit intensive jobs and allows for easy selection of CPUs. The Job Scheduler also provides the right information to accurately check the load balancing among kernels."

"These two programs are good partners because they allow you to breeze through the housekeeping and let you get to the work you really want to do."

Dr. Yuko Matsuda, Assistant Professor, Global Scientific Information and Computing Center, Tokyo Institute of Technology

The Compute Cluster Job Scheduler provides a smooth interface to submit intensive jobs and allows for easy selection of CPUs. The Job Scheduler also provides the right information to accurately check the load balancing among kernels.

Dr. Yuko Matsuda, Assistant Professor, Global Scientific Information and Computing Center, Tokyo Institute of Technology

Solution Architecture

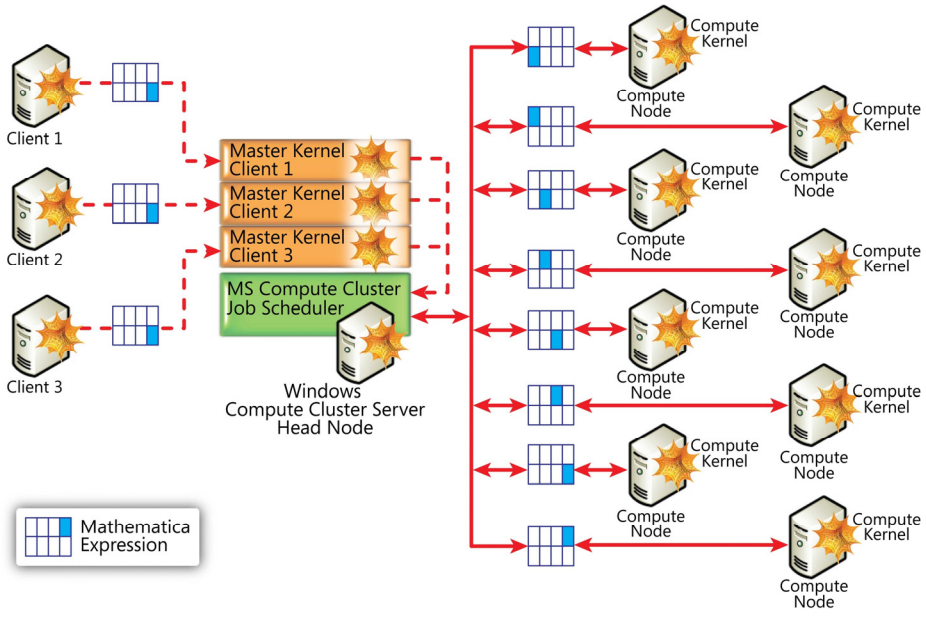
A Windows CCS cluster of servers includes a single head node and one or more compute nodes. The head node controls and mediates all access to the cluster resources and is the single point of management, deployment, and job scheduling for the compute cluster.

grid*Mathematica* consists of a collection of *Mathematica* kernels that form a coherent parallel technical computing environment. grid*Mathematica* implements a distributed manager-worker parallel model and integrates with Windows CCS.

Referring to the diagram below, grid*Mathematica* manages all *Mathematica* expressions itself, and uses Windows CCS as a hardware resource manager to allocate machines running grid*Mathematica* kernels on Windows CCS compute nodes.

A client machine running grid*Mathematica* and the Cluster Integration Package uses a single start command to access grid*Mathematica* and to submit *Mathematica* expressions to the grid*Mathematica* Master Kernel. In the diagram, the Master Kernel is running on the Windows CCS Head Node—this is not a requirement and is for illustrative purposes only.

The grid*Mathematica* Master Kernel manages all *Mathematica* expressions within grid*Mathematica* and uses the Microsoft Compute Cluster Job Scheduler to manage hardware resources for computing and tracking each *Mathematica* expression as it flows through the architecture. Notice that there are multiple Master Kernels on the Windows CCS Head Node, because each grid*Mathematica* client has its own instance of a Master Kernel.



Windows CCS System Requirements

CPU Requirement: 64-bit architecture computer Intel Pentium, or Xeon family with Intel Extended Memory 64 Technology (EM64T) processor architecture, or AMD Opteron family, AMD Athlon family, or compatible processor(s).
Minimum RAM: 512 MB
Maximum RAM: 32 GB
Multiprocessor Support: Up to 4 processors per node.
Disk Space for Setup: 4 GB
Disk Volumes: Head node requires a minimum of two volumes (C:\ and D :\ ). For additional roles, additional partitions are recommended. Compute node requires a single volume. RAID 0/1/5 may be used, but is not required.
Network Interface Cards: All nodes require at least one network interface card. Each node may require additional network interface cards as appropriate for the network topology, for public network access or in support of an MPI network.

Windows CCS Architecture

Windows CCS leverages the functionality of several components to provide authentication and authorization mechanisms, simple and familiar interfaces for managing and administering the cluster, and tools for cluster setup, deployment, job management, CPU efficiency and more.

Job Scheduling

The Microsoft Compute Cluster Job Scheduler is a core job management component in the Windows CCS architecture. Job scheduling ensures that the operating system allocates the necessary resources to the simulations, tracks the processors associated with the job, partitions the large scaled-out system to the size required by the simulation job, and deallocates the resources after producing the result.

Microsoft Message Passing Interface (MS MPI)

The Microsoft Message Passing Interface (MS MPI) acts as the communication software layer between compute nodes on the cluster. Windows CCS provides end-to-end security over secure and encrypted channels throughout the job process when using MS MPI. As the node manager schedules and assigns the job, the job always runs in the context of the scheduling users. MS MPI runs over Gigabit Ethernet, InfiniBand, or any network that provides a WinsSock Direct-enabled driver. MS MSPI is based on and compatible with the Argonne National Labs MPICH2 implementation of MPI2.

Microsoft Active Directory

Windows CCS uses the existing corporate infrastructure and Microsoft Active Directory® to provide authorization and authentication services. Each node of the cluster must be a member of an Active Directory domain. The Active Directory domain can be independent of the cluster, or run within the cluster.

Head Node

The head node provides deployment and administration user interfaces (UIs) as well as management services for the compute cluster. The UIs provided by the head node include the Compute Cluster Administrator, the Compute Cluster Manager and the Command Line Interface (CLI). The management services provided by the head node include job scheduling as well as job and resource management.

Compute Node

Any computer configured to provide computational resources as part of the compute cluster is a compute node. Compute nodes allow users to run computational jobs. These nodes must run a supported operating system, but they do not require the same operating system or even the same hardware configuration. Optimally, compute nodes include a similar configuration to simplify deployment, administration, and resource management.

For More Information

For more information about Microsoft products and services, call the Microsoft Sales Information Center at (800) 426-9400. In Canada, call the Microsoft Canada Information Centre at (877) 568-2495. Customers who are deaf or hard-of-hearing can reach Microsoft text telephone (TTY/TDD) services at (800) 892-5234 in the United States or (905) 568-9641 in Canada. Outside the 50 United States and Canada, please contact your local Microsoft subsidiary. To access information using the World Wide Web, go to: www.microsoft.com.

For more information about Windows Compute Cluster Server 2003, please visit: http://www.microsoft.com/hpc

To join the HPC Community, please visit http://www.windowshpc.net

For more information about Wolfram Research, Inc. or purchasing *Mathematica* and grid*Mathematica* products, please visit https://www.wolfram.com

For information about purchasing Microsoft Windows Compute Cluster Server 2003, please email hpcinfo@microsoft.com

Microsoft Corporation acknowledges WinHPC.org for contributing source material for this solution brief. Please visit http://www.winhpc.org

Document published April, 2007

Benefits

Windows Compute Cluster Server 2003, combined with grid*Mathematica* provides you with a highly scalable and easily managed rapid development and deployment environment. This combined solution delivers *depth*, *focus*, *scalability*, *access* and more.

Depth

Windows CCS provides the grid management system for grid*Mathematica* so that users can access the world's largest algorithm collection and supercomputing capacity in one integrated system.

Focus

The details of parallelism in grid*Mathematica* are largely abstracted, allowing the developers to focus completely on creating solution.

Scalability

grid*Mathematica* applications can be deployed on a Windows XP laptop or desktop and be immediately deployed on Windows Compute Cluster Server 2003.

Access

grid*Mathematica* is able to read and write many data formats, including HDF5 and Microsoft Excel spreadsheets, and is able to access data from legacy applications through C/C++, Java and .NET.

Simple Deployment and Administration

Windows CCS includes prescriptive setup procedures, a complete suite of management tools and an integrated Job Scheduler.

Familiarity and Integration with Existing Windows Infrastructure

Windows CCS leverages existing Windows infrastructure, so it works with Microsoft Active Directory directory service, can be monitored with Microsoft Management Console (MMC) and Microsoft Operations Manager (MOM), and uses tools such as Microsoft Visual Studio 2005 to support parallel job development and debugging.

Immediate Access to HPC Capacity from Within a Single User Interface

The Cluster Integration Package for grid*Mathematica* enables Windows CCS to administer grid*Mathematica* compute kernels by providing users with an interface to cluster management systems from within grid*Mathematica*.

Scientists and engineers are increasingly forced to do more with less. By using grid*Mathematica* on Windows CCS, they now have a seamless way to upscale their technical workstation by tapping into clustered compute power.

Shawn Hansen, Director of High Performance Computing, Microsoft Corporation.

Microsoft®