

## Microsoft News

Product News ▶

Consumer News

International Contacts

Legal News

Security &amp; Privacy News

Events

News Archive

## Corporate Information

Microsoft Executives ▶

Fast Facts About Microsoft

Image Gallery ▶

## Related Sites

Analyst Relations

Community Affairs

Essays on Technology

Executive E-Mail

Global Citizenship

Investor Relations

Microsoft Research



## Q&A: Taking High-Performance Computing Mainstream

Customers report that Microsoft Windows Compute Cluster Server 2003 makes it easier and more cost effective to develop, deploy and use HPC systems, and they report that a Windows-based HPC platform integrates smoothly with existing IT resources.

**REDMOND, Wash., Aug. 1, 2006** -- Today Microsoft announced the general availability of Microsoft Windows Compute Cluster Server 2003, the company's first product designed specifically for high-performance computing (HPC). With Windows Compute Cluster Server 2003, Microsoft aims to make it easier to create, integrate and operate HPC clusters within organizations, thereby expanding the technology beyond traditional supercomputing centers by bringing the value of computational clusters within reach of more people.

To understand the impact of today's milestone, PressPass convened a roundtable of customers who have been test driving Microsoft Windows Compute Cluster Server 2003 in demanding applications, including biomedical research and scientific modeling. Providing their insight are:

- **Ron Elber**, professor of computer science at Cornell University
- **John Michalakes**, senior software engineer at the National Center for Atmospheric Research (NCAR) in Boulder, Colo.
- **Matt Wortman**, director of computational biology and IT at the Genome Research Institute, University of Cincinnati

**PressPass: Would each of you begin by briefly describing the work you're doing as it relates to Microsoft Windows Compute Cluster Server?**



**Ron Elber, Professor of Computer Science at Cornell University**

**Elber:** At Cornell, we have a core facility called the Computational Biology Service Unit (CBSU) that's dedicated to computational biology and bioinformatics for Cornell researchers. We provide both research and computational support to biology groups. The cluster serves as a platform for computational biology applications used in a range of research activities in bioinformatics. We support many popular applications for sequence-based datamining, population genetics and protein structure prediction. Many of the projects require lengthy calculations, and massively parallel computing helps shorten the clock time and obtain results in a reasonable period. We have developed a Web-based interface that allows biologists to access the applications without any prior knowledge of cluster computing.

**Michalakes:** About eight years ago, NCAR and a number of partner organizations involved in atmospheric research and operational forecasting began working on a next-generation community weather model and data assimilation system to eventually replace aging model codes in use for forecasting and research. This new model, called the Weather Research and Forecast (WRF) model, is basically all new software, designed from the outset for HPC systems. WRF is maintained and

## Related Links

## Quote Sheet:

- [Windows CCS Partner Quote Sheet](#)

## Feature Stories:

- [Microsoft Flexes Muscle with the High-Performance Computing Community](#) – June 26, 2006
- [Microsoft Showcases Excel Services and Windows Compute Cluster Server 2003](#) – June 20, 2006

## Archives:

- [Video interview: HP and Windows Compute Cluster Server 2003 video interview](#)
- [Video interview: Intel, Ciara Technologies and Microsoft Partner to Bring HPC to the Masses](#)

## Press Release:

- [Microsoft Releases Windows Compute Cluster Server 2003](#) – June 9, 2006

## Microsoft Resources

- [Windows Compute Cluster Server 2003 Web Site](#)
- [Windows HPC Community](#)

freely distributed as a community model and is being run at hundreds of institutions across the range of systems, from individual workstations to large supercomputers. Thus, portability and portable performance has been a key concern in the design and implementation of WRF.

**Wortman:** One of our key focus areas at the Genome Research Institute is drug discovery. Early in the drug-discovery process, millions of chemical compounds are screened against disease targets to identify classes of molecules whose properties and activities guide researchers toward the discovery of new drugs. Our research focuses on applying computational tools to this process to reduce costs and save time. Specifically, we perform virtual *in silico* screening experiments that simulate the interactions between a disease target and those millions of chemical compounds to predict which compounds participate in desired interactions. The compounds predicted to have the most favorable properties are selected from the chemical library, and then proceed to *in vitro* testing to confirm the computational predictions. This combination of *in silico* and *in vitro* screening is much faster and less expensive than *in vitro* screening alone because the number of chemicals that need to be tested is reduced by several orders of magnitude. A typical job on our cluster begins when the disease target is sent to the scheduler along with a list of chemicals to be used during the simulation. The head node sends a copy of the disease target and a portion of the chemicals to each node where simulations occur independently. The head node analyzes and ranks the results of each simulation.

**PressPass: What made you decide to use the Microsoft Windows Compute Cluster Server 2003, and what benefits do you think it offers to your organization and your work?**

**Michalakes:** We strive to maintain WRF [NCAR's Weather Research and Forecast model] on as many systems deployed in our user community as possible. Until now, that meant systems running some flavor of UNIX or Linux. With the emergence of Microsoft Windows as a viable HPC operating system, and given that we receive on average one user request per month asking if WRF will work on Windows, we see Windows CCS as an opportunity for further broadening the range of computational resources available to the WRF user community.



**John Michalakes, Senior Software Engineer at the National Center for Atmospheric Research**

**Wortman:** Our decision to use Windows Compute Cluster Server was motivated by the need to lower costs by reducing the complexity of our infrastructure. Windows Compute Cluster Server has several advantages to an organization like ours that uses Active Directory for identity management. First, our Windows technicians could apply their knowledge of Windows-based servers to it. This was evidenced by the fact that individuals with no HPC experience set up a Windows-based Compute Cluster Server HPC cluster without guidance or supervision. Second, using Active Directory and the Microsoft job scheduler enables our users to submit jobs from their workstations and reduces the number of user accounts.

**Elber:** Upgrading to Windows Compute Cluster Server was a natural step for us. We have been using a Windows-based HPC platform since the computational biology unit was started in 2001. Until recently, we used Windows-based systems adapted by the Cornell Theory Center (CTC) for HPC. We use Microsoft SQL Server for our database needs and Windows-based servers for hosting our Web interfaces. Therefore, Windows Compute Cluster Server allows for a homogeneous and easy-to-develop environment. Our experience with the CTC's Windows-based HPC systems is very positive, and we expect Windows Compute Cluster Server to be even better.

**PressPass: What business needs are you solving with high-performance computing?**

**Elber:** Due to high computational demands, many research projects are plainly impossible to pursue without an HPC platform, or they would take an unreasonable amount of time to complete. For example, a typical datamining operation with several thousand input sequences will take several hours to run on a parallel machine; otherwise, it will take several days. Preparing data for learning scoring functions for protein folding might take a month or two on a massively parallel machine of suitable size; otherwise, it would take several years, making the project impractical. An HPC machine is also a convenient tool for serving the computational needs of many small projects. It's an easy-to-manage and uniform platform, where installing the software, updating databases and other such tasks are much easier to do than they would be on a set of separate computers.

**Michalakes:** Although the WRF model is used primarily in public sector institutions -- atmospheric research departments and government-run research and forecast centers -- a growing number of

commercial weather forecast companies use WRF as well. These companies make specialized weather forecast products for customers in construction, agriculture, energy and other businesses.

**Wortman:** For us, the key business need was reducing costs by eliminating complexity. We did that by eliminating Linux support costs.

**PressPass: A couple of you have brought up Linux. Based on your experience, how does a Windows-based HPC platform compare to Linux-based HPC clusters in areas like development, setup, maintenance, interoperability, scalability and applications?**

**Wortman:** The setup and management of Windows Compute Cluster Server vs. a Linux cluster is worlds apart for us. The proof of this is the fact that a Windows technician with no HPC experience can set up a cluster from scratch in a matter of hours. Linux clusters simply take more care and feeding, and substantial knowledge of Linux, which nearly all biomedical researchers lack.

**Elber:** From our perspective, large-scale Linux clusters are difficult to set up and then to appropriately tune up, whereas a Windows-based cluster seems to be easier to set up, even considering the fact that we've been running a beta version of Windows Compute Cluster Server. Also, a Linux cluster is less friendly for an average user who is not computer-oriented, for example, a biologist who's in need of a computing environment. And, because we use the Windows platform for databases, file storage and Web interfacing, a Windows-based HPC cluster integrates much better, and it's easier to develop software with.



**Matt Wortman, Director of Computational Biology and IT at the Genome Research Institute, University of Cincinnati**

**PressPass: As HPC becomes more of a mainstream technology, what key opportunities and challenges do you foresee for commercial and research application developers?**

**Michalakes:** One shift we've seen as HPC has matured is from thinking about the performance-at-any-cost of HPC systems to thinking more about cost-performance of such systems. I believe operational numerical weather prediction is solidly terascale, but it remains to be seen whether operational centers will move to petascale systems for their day-to-day, real-time forecasting production schedules. The issue will be cost -- the cost of petascale systems themselves, as well as the cost to operate such systems and the cost to retool and maintain modeling software to run on such systems, weighed against some hoped-for improvement in forecast quality. In the near term, I believe petascale computing will be used for non-real-time, very-high-resolution simulations for research to improve understanding of atmospheric processes that will, in turn, provide improvements to lower-resolution operational real-time prediction runs.

Having said all this about petascale computing, my sense in the context of this discussion is that Microsoft is not currently targeting frontier computing systems for Windows Compute Cluster Server, but focusing instead on small- to mid-level clusters more widely deployed in the research and commercial areas of weather modeling. From this perspective, the challenges and opportunities are efficient integration and management of computing and data systems to allow for more seamless coordination and management of workflows for end-to-end computing and analysis of weather and climate applications.

**Elber:** The biggest challenge -- and an opportunity -- is to develop easy-to-use solutions with intuitive interfaces that allow users to access software on an HPC cluster without prior knowledge of the cluster operating system or scheduling intricacies. Here at Cornell, we've developed an interface to computational biology applications that is very popular among biologists and separates them from our particular hardware/software implementation. This is especially true for small or medium-sized clusters that will be used in small research groups whose members have no experience in HPC or parallel programming.

**Wortman:** I foresee HPC changing from a niche occupied by financial and scientific technical experts to a mainstream blackbox affair with many small group or department clusters outnumbering the large HPC centers. Application developers will provide plug-and-play devices that integrate into your infrastructure via USB or Ethernet. These devices will be simple and capable of a small variety of very high-speed calculations. For example, a standalone bioinformatics server will store and analyze sequence data, or a drug discovery appliance will screen chemical compounds. These simple "unitaskers" will be made and supported for integration into your existing Windows-

based environment.

[↑ Top of page](#)



[Printer-Friendly Version](#)



[Send This Page](#)



[Add to Favorites](#)

---

[Manage Your Profile](#) | [Contact Us](#) | [Newsletter](#)

© 2006 Microsoft Corporation. All rights reserved. [Terms of Use](#) | [Trademarks](#) | [Privacy Statement](#)