Microsoft's High Performance Computing Oil & Gas Forum
June 20, 2003

Introduction

Around 100 attended Microsoft's High Performance Computing for Oil and Gas Forum held in Houston. The launch of an initiative in high performance computing puts Microsoft in the unusual position of the underdog confronting the dominance of Linux – at least in upstream oil and gas. It is instructive to analyze the speakers in terms of their true affiliations. Cornell's CTC HPC unit is Microsoft-funded. Intel is agnostic in its marketing literature and made no real attempt to justify Windows as an HPC platform. HP again is agnostic in its marketing literature and was bent on selling its 'Agile' infrastructure rather than defending Windows. CMG's Erdle is a Windows enthusiast, but the record breaking reservoir simulation he refers to was run on a Unix machine. Finally, Modviz's software currently only runs on Linux (a Windows version is under development). But this begs the question where is Microsoft in the key area of upstream HPC – seismic processing. The answer seems to be nowhere at all! All in all this morning's meet seems to have missed the mark by a long chalk. No doubt there are reasons to consider Windows and the new .NET architecture as candidates for HPC – and some of these are developed on the toolkit CD. But none were explored in any depth by the speakers.

Microsoft in E&P – Marise Mikulis energy manager[1]

Microsoft is entering a new era in its dedication to the E&P industry. A Global Business Unit will focus on E&P industry. Microsoft wants to ensure that its technology meets E&P needs – both upstream and downstream

Microsoft's High Performance Computing – Greg Rankich HPC manager

The research community still develops in FORTRAN. Parallel programming hasn't changed in 10 years – it still requires specialized skills. The plethora of tools and approaches make HPC a fragmented market with many applications but a general lack of integrated solutions. Microsoft is working with customers to optimize deployments by investing in third party companies developing products to manage clusters. Microsoft is also investing internally on cluster optimization. Microsoft has been working with the Cornell Theory Center (see below), HP, Intel and other customers on HPC. Microsoft is giving away a toolkit[2] to help users get started with cluster-based computing. Rankich concluded – by asking 'why not Microsoft? An unusual soft sell approach to confronting Linux's dominance of the HPC market.

From Research to Enterprise Computing – Roger Lang, Cornell Theory Center[3]

Lang's vision is to make HPC simple enough to be used by 'the masses' through standard off-the-shelf tools. CTC has demonstrated the world's first Windows-based CAVE

[1] All company affiliations Microsoft unless otherwise stated.
[2] The Computational Clustering Preview toolkit includes evaluation versions of Windows XP Pro, 2000 Advanced Server, Visual Studio .NET and Windows Services for Unix and third party software.
[3] The CTC High-Performance Solutions unit is an alliance between CTC, Dell, Intel and Microsoft.

visualization system and is 'trying to make clusters easy to build'. CTC is working with Microsoft to 'push the envelope' using Web Services and Microsoft's SQL Server DBMS[4]. The recent move to Windows 2003 'reveals a new roadmap' for HPC with Enterprise Web Services making HPC resources transparent to anyone in the enterprise and promising 'excellent reliability and scalability'. Lang addressed the $ 64,000 question 'Why not Linux' by claiming that cross-platform enterprise Web Services requires complex systems integration[5] and the final project is a custom product and not easy to roll out. On the other hand, Microsoft's .NET framework offers 'end-to-end integration for scalable Web Services'. Windows-based clusters offer the following benefits

- Reduced staff size
- Reduced cost[6] with a 'drastic reduction in annual maintenance fees'
- Ease of installation
- Scalability
- Deployable enterprise-wide
- Reduced technical management training costs
- Computing 'by the slice' (easy to add servers on the fly)
- .NET technology provides an option for automatic load-balancing and 'near-linear scalability'.

Intel Architecture Overview for HPC David Barkai, HPC Architect, Intel

Around 1980, a gigaflop cost around $5 million. Today the same compute power costs $2,000. Clusters now offer around 10 teraflops (TF). By the end of the decade (2010) Intel foresees 30 GHz processors. The challenge for cluster designers is to keep up with the speed of memory. Barkai states that 56 of the top 500 supercomputers[7] sites are Intel-based (up from 3 in '99). Intel builds components or building blocks for HPC: architecture, processors, platforms, software, interconnects and offers services such as application tuning. Intel claims processor leadership for all HPC solutions with its Pentium 4, Xeon and 64 bit Itanium 2. Barkai believes HPC requires 'ecosystems' for industry and end users and Intel's HPC program office is 'more than silicon'. Intel's HPC involvement extends to joint solution centers, portfolio investments, Intel Web services and 'strategic collaboration'. HPC is differentiated by computational intensity large scale applications and random, dynamic data access patterns. The community today is composed of 'pre-early adopters' and expert users. Intel wants to make technical computing adoptable by a larger market. Choices in cluster components let users tailor their clusters to specific requirements. Common off-the-shelf (COTS) components make HPC clusters affordable at the department or project level.
But there are problems. Existing challenges will get worse as Moore's Law drives on. I/O and interconnect bandwidth will have a hard time keeping pace with processor speed and

[4] A handout from CTC "Post-Cluster computing and the next generation of scientific applications" described the use of 'an alternative approach' to finite element analysis using Microsoft's SQL Server database.
[5] This is surely putting the cart before the horse. Web Services' *raison d'être* is to solve 'complex systems integration' issues.
[6] The case for cost reduction over a free OS such as Linux might merit more explanation!
[7] The Top 500 list does not include many (most) clusters used in geophysical processing. Other commercial installations are likely missing from the list.

memory size. Cluster management also raises issues of performance monitoring and maintenance. Intel Itanium 2 processor 'Madison' is to launch this month. 'Madison-plus' will come out next year. In 2005, the 'Montecito' will be released Intel is also 'making great progress' with its 32-bit architecture. The new PCI Express interface will bypass some of the latency issues with high performance clusters. But improvements are not just silicon-based, there is innovation at all levels of the platform with interconnects and storage (Ethernet, Infiniband, PCI Express). COTS enables choices and offers many approaches to cluster construction. The trick is to consider the entire stack, i.e., what is the application and what are the tools needed. Don't just focus on what is the fastest processor.

4) Upstream Infrastructure Michele Isernia, Industry Director, Oil & Gas, HP

Isernia asks, "How can we really use and capitalize on technology?" HP's 'Agile Computing' initiative is the way forward. Agile Computing aims to 'loosen' the tight relationship that software has with hardware. Software should be 'owned' by the user, who can use whatever computers or appliances are available. By moving to this new model, users are freed from the burden of carrying any electronics with them at all. Computers will be able to 'recede in the background and become pervasively useful instead of a constant annoyance'.

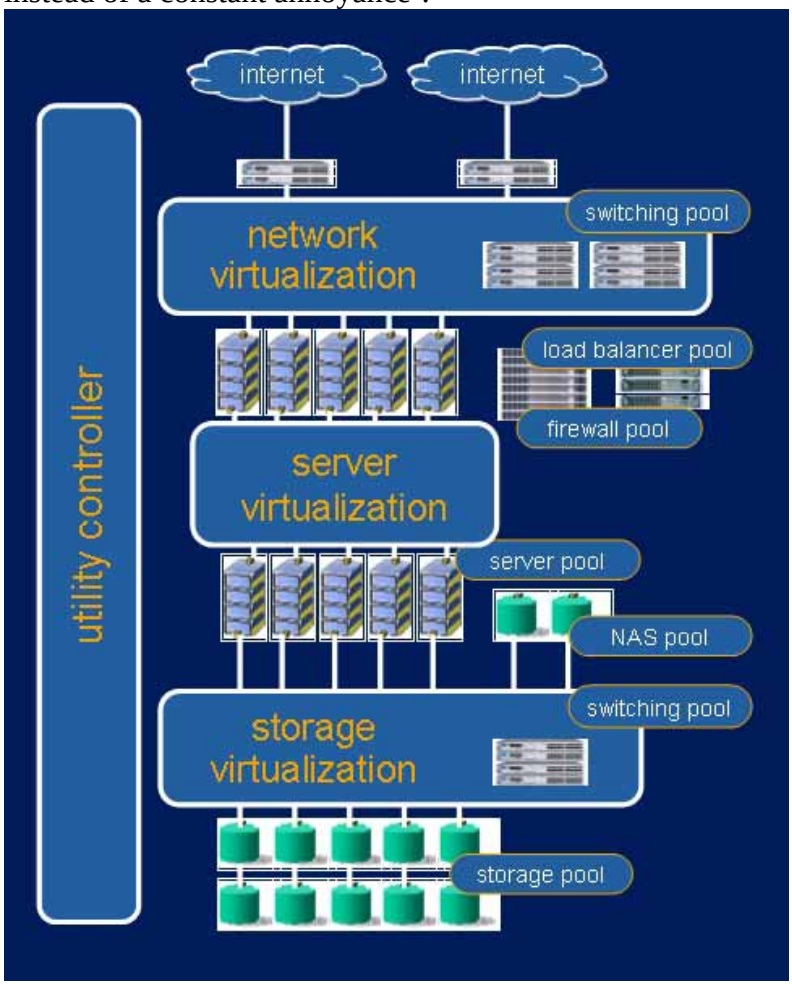

Figure 1 HP's Utility Data Center (UDC)

HP's Agile technology aims to 'refresh' technology every 6 months without disruption. Components include the Agile Client – an 'industry-standard' and the Agile Data Center built from industry standard servers and storage providing access to legacy data and applications. Infrastructure must be designed for use-on-demand by focusing on services, shares resources and pay for use. A pool of virtual servers can expand or shrink as required. HP technology to achieve this is the Utility Data Center a 'complete solution for virtualizing data center environments'. Isernia offered a lot more information on HP's

Fluid Flow Simulation Solutions for Itanium2/64 bit Windows – Jim Erdle, CMG

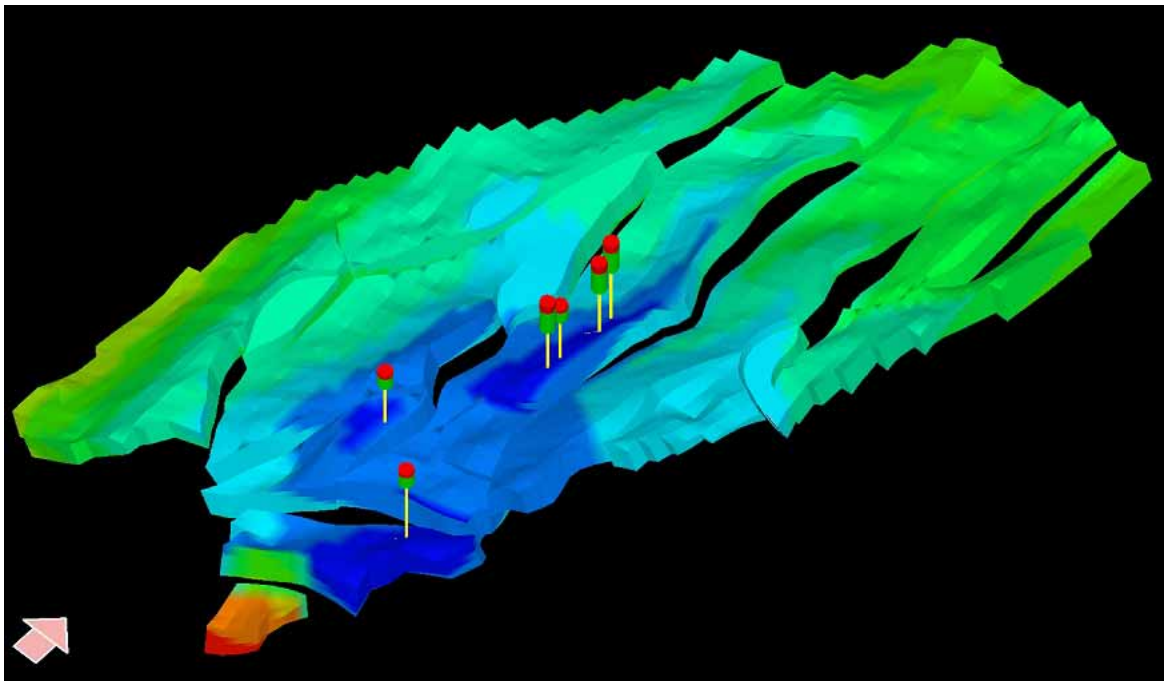

Figure 2 GMG HPC reservoir simulation[8].

Calgary-based Computer Modelling Group (CMG) provides reservoir simulation with its flagship IMEX II GEM and STARS products. One driver behind HPC in reservoir modeling is to avoid time consuming upscaling by using geocellular models. This is being achieved by a combination of 64-bit computing, shared memory parallel processing, and dynamic PEBI grids. Erdle believes that 1-5 million cell models will soon be commonplace. CMG is a Windows shop they appreciate the easy to use GUI and the speed of Windows on the Itanium. CMG claims a world record for its 112 million cell simulation – but this was achieved on an IBM cluster running IBM Unix (AIX)[9]!

6) ModViz Inc. Scalable Visualization Solution – Shing Pan, VP ModViz Inc.

Modviz (a Siemens spin-off) develops hardware and software to enable large scale visualization using PC technology. The goal is to provide a scalable software platform to enable high performance real time visualization. ModViz synchronizes 3D data across a cluster in real time. ModViz is partnering with HP partnership to 'make scalable visualization a reality'. Modviz's Renderizer Visualization Cluster software for Open

[8] Courtesy Computer Modelling Group (CMG).
[9] See Oil IT Journal Report.

Inventor is currently shipping in Linux and will be available for Windows ‘real sson now’.