Scalability of Transient CFD on Large-Scale Linux Clusters with Parallel File Systems

Stan Posey[1] and Bill Loewe[2]
Panasas Inc., 6520 Kaiser, Dr, Fremont, CA 94555, USA
sposey@panasas.com; bloewe@panasas.com

Srinivas Kodiyalam[3] and Mark Kremenetsky[4]
Silicon Graphics Inc., 1140 East Arques Ave, Sunnyvale, CA 94085, USA
skodiyal@sgi.com; mdk@sgi.com

Hughes Mathis[5]
Intel Corporation (UK) Ltd., Pipers Way, Swindon SN3 1RJ, UK
hughes.a.mathis@intel.com

Paul Calleja[6]
University of Cambridge, Trinity Lane, Cambridge CB2 1TN, UK
pjc82@cam.ac.uk

Abstract

This work examines the parallel scalability characteristics of commercial CFD software FLUENT and STAR-CD for up to 256 processing cores, and research CFD software CDP from Stanford University for up to 512 cores – for transient CFD simulations that heavy in IO relative to numerical operations. In three independent studies conducted with engineering contributions from the University of Cambridge, Intel, SGI, and Panasas, each Linux HPC environment combined Intel Xeon clusters with a Panasas parallel file system and shared storage. The motivation for these studies was to quantify the performance and scalability benefits of parallel IO in CFD software on a parallel file system (PanFS) verses a conventional serial NFS file system for a variety of transient CFD cases.

Keywords: High Performance Computing, Transient CFD, Linux Clusters, Parallel File Systems, IO.

Introduction

CFD parallel efficiency and simulation turn-around times continue to be an important factor behind engineering and scientific decisions to develop models at higher fidelity. Most parallel CFD simulations use scalable Linux clusters for their demanding HPC requirements, but for certain classes of CFD models, data IO can severely degrade overall job scalability and limit CFD effectiveness. As CFD model sizes grow and the number of processing cores are increased for a single simulation, it becomes critical for each thread on each core to perform IO operations in parallel, rather than rely on the master compute thread to collect each IO process in serial.

Examples of IO demanding simulations include ~100-million-cell steady state models on a very large (> 64) number of cores, and moderate-sized and moderately parallel transient CFD models that require frequent and multiple solution files writes in order to collect time history data for subsequent post-processing. In the case of frequent time history files being written in a serial process, any parallel benefit from a CFD solver is soon overwhelmed as more processing cores (and therefore more serial IO threads) are added to a simulation.

[1] Director, Industry and Applications Development, Member AIAA
[2] Senior Applications Engineer, Ph.D.
[3] Director, HPC Business Development, Associate Fellow AIAA, Ph.D.
[4] Principal Scientist, CFD Applications, Associate Fellow AIAA, Ph.D.
[5] Senior Applications Engineer, Ph.D.
[6] Director, HPC Services Center, Ph.D.

For this reason, commercial CFD software developers such as ANSYS and CD-adapco, and research CFD developers such as Stanford University, offer parallel IO schemes that scale with the CFD solvers in each, but just as parallel solvers require scalable clusters, parallel IO requires that the cluster be configured with a parallel file system and shared storage, also sometimes referred to as parallel network attached storage (NAS). Panasas offers a parallel NAS capability that scales IO to overcome IO bottlenecks, enabling IO-bound CFD to scale to its full potential.

Parallel File Systems and Shared Storage

A new class of parallel file system and shared storage technology has developed that scales IO in order to extend overall scalability of CFD simulations on clusters. For most implementations, entirely new storage architectures were introduced that combine key advantages of legacy shared storage systems, yet eliminate the drawbacks that have made them unsuitable for large distributed cluster deployments. Parallel NAS can achieve both the high-performance benefits of direct access to disk, as well as data-sharing benefits of files and metadata that Linux clusters require for CFD scalability.

Panasas offers a parallel NAS technology with an object-based storage architecture that overcomes serial IO bottlenecks. Object-based storage enables two primary technological breakthroughs vs. conventional block-based storage. First, since an object contains a combination of user data and metadata attributes, the object architecture is able to offload IO directly to the storage device instead of going through a central file server to deliver parallel IO capability. That is, just as a cluster spreads the work evenly across compute nodes, the object-based storage architecture allows data to be spread across objects for parallel access directly from disk. Secondly, since each object has metadata attributes in addition to user-data, the object can be managed intelligently within large shared volumes under a single namespace.

Object-based storage architectures provide virtually unlimited growth in capacity and bandwidth, making them well-suited for handling CFD run-time IO operations and large files for post-processing and data management. With object-based storage, the cluster has parallel and direct access to all data spread across the shared storage, meaning a large volume of data can be accessed in one simple step by the cluster for computation and visualization to improve speed in the movement of data between storage and other tasks in the CFD workflow. Panasas provides this architecture by offering finely tuned hardware components that optimize the parallel file system software architecture capabilities.

Transient CFD with Parallel IO

The combination of scalable CFD application software, Linux clusters and a parallel file system and storage, also referred to as parallel network attached storage (NAS) – can provide engineers and scientists with new and significant performance advantages for transient CFD simulations. In recent studies, the advantages of FLUENT 12, STAR-CD v4, and CDP 2.4 with their parallel IO capability, demonstrated efficient scalability, total job turn-around time improvements of 2x and greater, and the ability to conduct high-fidelity yet cost-effective transient solutions that were previously impractical for consideration.

The FLUENT, STAR-CD, and CDP models for these studies comprised cases that were relevant in size and flow features to current commercial and research CFD practice:

- FLUENT *capability* study of a large external aerodynamics case of 111 million cells, modelled as transient with a detached eddy simulation (DES) model, and the writing of 20 GB of output data per every 5 time steps. [Only 100 iterations and 5 time steps were run]

- STAR-CD transient LES study of a 17 million cell turbomachinery case received from an undisclosed customer for performance evaluation on increasing Intel Xeon cores, with frequent time history writes of 48 GB of total output data.

- CDP transient LES study of a 30 million cell internal flow case of a Pratt & Whitney combustor single injector geometry , run on up to 512 cores of an SGI ICE Xeon cluster for scalability evaluation of a 50 iteration restart with a solution write of 12 GB of total data.

FLUENT: Collaboration between ANSYS and Panasas application engineers that began in 2005 has developed a parallel IO scheme based on MPI-IO under the MPI-2 standard. FLUENT parallel IO will be introduced in the FLUENT 12 release scheduled for 1H-2009 and will leverage several parallel file systems including Panasas' PanFS. This development will permit FLUENT scalability for even heavy IO applications such as large steady state models with frequent checkpoints, or even more challenging, transient CFD that may require 100 or more solution writes per single simulation.

The benefits of parallel IO for transient CFD were demonstrated with a production case of a FLUENT aerodynamics model of 111M cells, provided by an industrial truck vehicle manufacturer. Details of the model and results are provided in Figure 1. FLUENT 12 with parallel IO in the way of concurrent writes of local (partition) solutions to the global solution data file on PanFS demonstrated a ~ 2x performance advantage in total time vs. FLUENT 6.3 on NFS and a conventional NAS system.

In the case of 64 cores, the solver advantage of FLUENT 12 over 6.3 is only 4% and the total time benefit of 1.8x shown in Figure 1 is a result of parallel IO speed-up. The FLUENT 12 solver advantage grows to 9% at 128 cores, and 24% at 256 cores, which contribute to the growing benefits of total time improvements of 2.0x on 128 and 2.3x on 256 cores for FLUENT 12 on PanFS vs. 6.3 on NFS. It is important to note that the performance of CFD solvers and numerical operations are not affected by the choice of file system, which only improves IO operations. That is, a CFD solver will perform the same on a given cluster regardless of whether a parallel or NFS file system is used.

STAR-CD: Indeed this is the case observed with STAR-CD v4.6 which uses the same CFD software on the same cluster, but compares two different file systems. The model and Intel HPC environment details and the results are provided in Figure 2. From 64 to 256 cores, solver times are the same regardless of file system, and all the total time gains come from parallel IO speed-up of STAR-CD on PanFS. STAR-CD uses an IO scheme whereby each core writes to its own output file of growing time history data independent of other cores. The advantage of parallel IO from 64 to 256 cores grows from 24% to 85% and consistent with results observed in the case of FLUENT from 64 to 256 cores.

CDP: This case also compares the performance of parallel PanFS vs. serial NFS on a serial NAS system for the same CFD software on the same cluster, and for up to 512 cores on an SGI ICE cluster. Details of the model, SGI HPC environment and results are provided in Figure 3. Similar to results observed with the commercial CFD software, CDP improves both overall performance and parallel scalability in this case to the full 512 cores utilized in this study. Total time speed-ups range from a 36% advantage on 128 cores to 82% on 512 cores.

Conclusions

Joint studies conducted between research and industry organizations demonstrate that CFD software with parallel IO on a parallel file system can show full parallel benefit for transient simulations that are heavy in IO relative to numerical operations. The favourable results were conclusive for a range of commercial and research CFD software on a variety of Intel-based Linux clusters with a Panasas parallel file system. Benefits to industry include an expanded and more common use of transient CFD in applications such as aerodynamics, aeroacoustics, reacting flows and combustion, multi-phase and free surface flows, and DES/LES turbulence, among other advanced CFD modeling and practices.

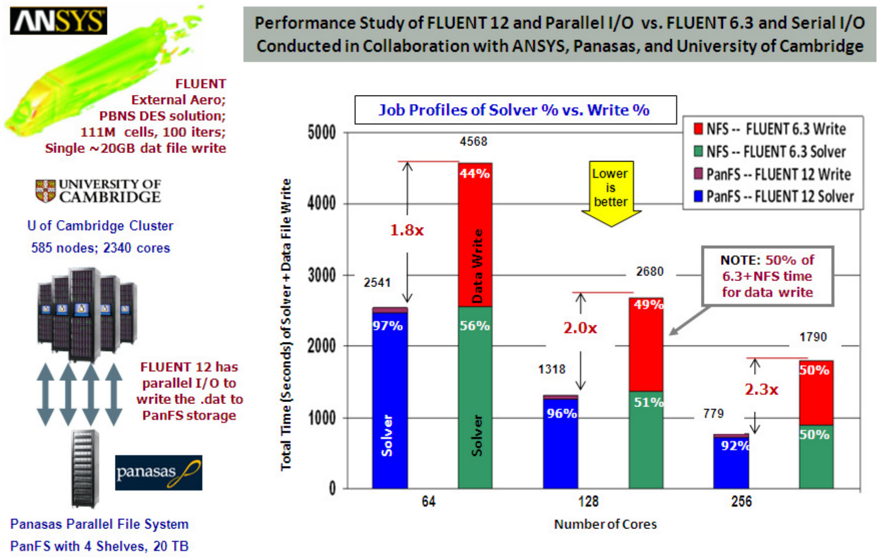


Figure 1: Comparison of FLUENT 12 using parallel IO on PanFS vs. FLUENT 6.3 on NFS

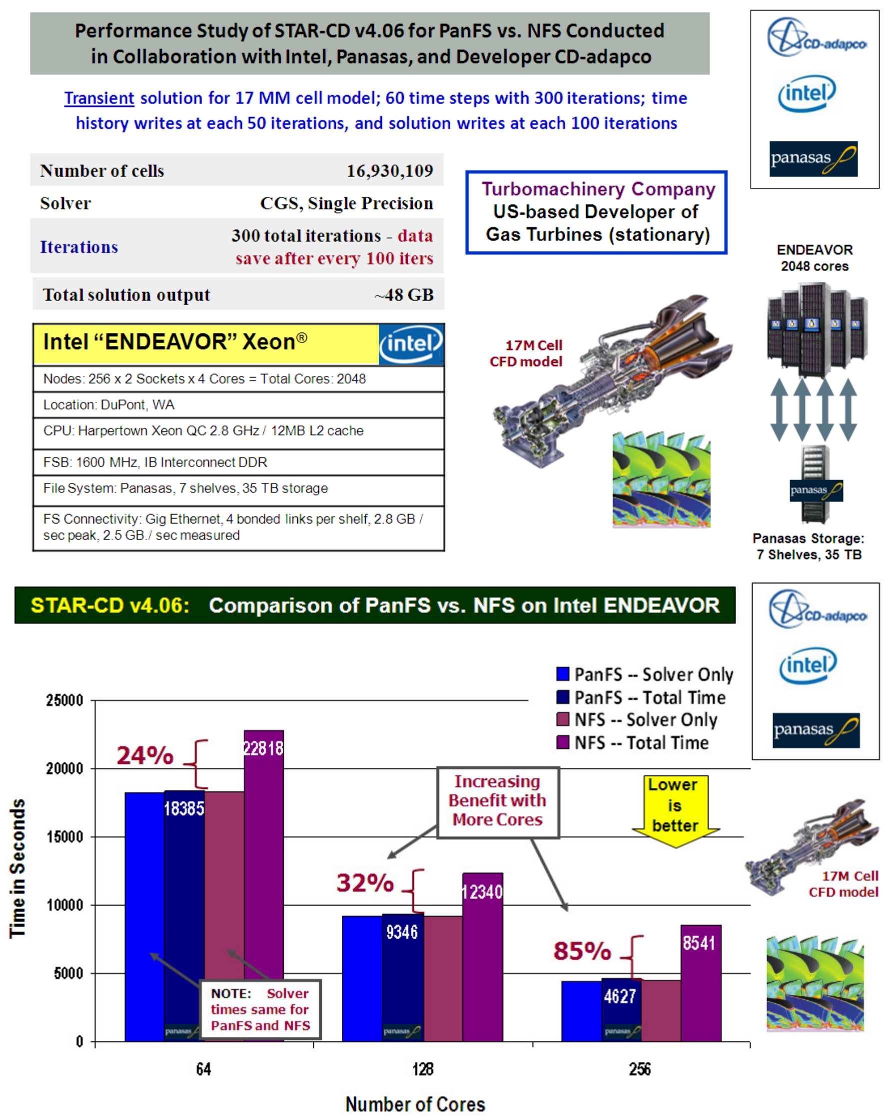
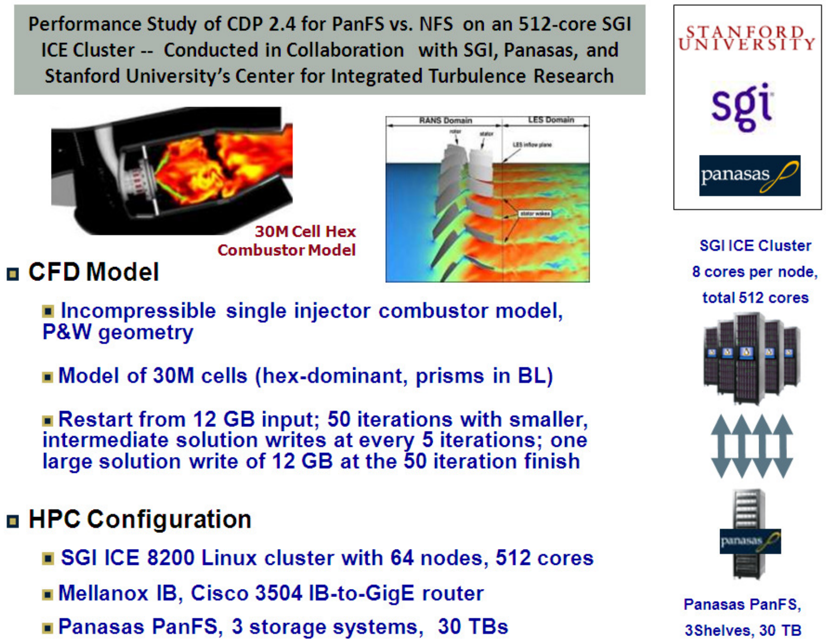


Figure 2: Comparison of STAR-CD v4.6 using parallel IO on PanFS vs. serial IO on NFS



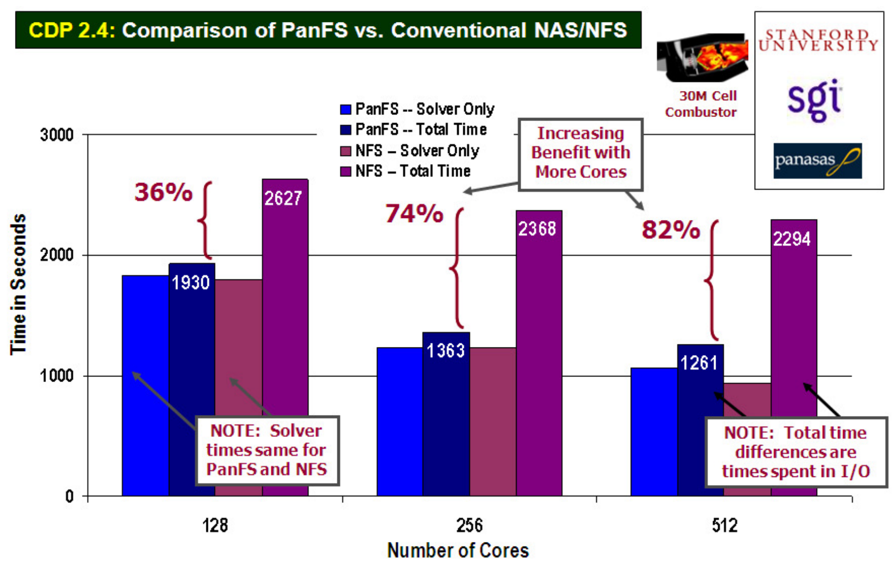


Figure 3: Comparison of CDP 2.4 using parallel IO on PanFS vs. serial IO on NFS

References

1. S. Kodiyalam, M. Kremenetsky and S. Posey, "Balanced HPC Infrastructure for CFD and associated Multidiscipline Simulations of Engineering Systems," Proceedings, 7th Asia CFD Conference 2007, Bangalore, India, November 26 – 30, 2007.

2. Gibson, G.A., Van Meter, R., "Network Attached Storage Architecture," Communications of the ACM, Vol. 43, No. 11, November 2000.

3. FLUENT Software IO Features: http://www.fluent.com/software/fluent/index.htm

4. STAR-CD Software IO Features: http://www.cd-adapco.com/products/STAR-CD

5. Discussions with Dr. Frank Ham, CDP development, Stanford University, Feb – Aug, 2008, and CDP Unstructured LES Code: http://www.stanford.edu/group/cits/research/combustor/cdp.html