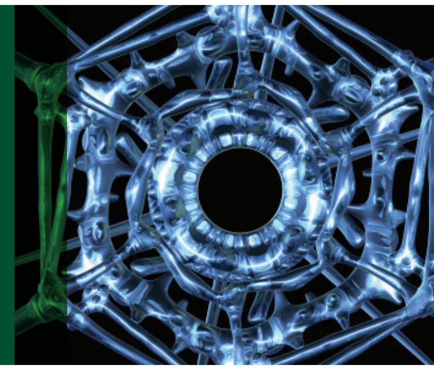
HPC in Manufacturing – Computer Aided Engineering (CAE)

The Benefits of Effortless Storage in CAE

Computer Aided Engineering (CAE) is a broad term in the manufacturing industry to describe a set of compute programs, encompassing many disciplines. These programs are used to meet design criteria for parts and assemblies such as strength to weight ratios, fluid flow and electrical and magnetic transmissions. CAE programs also allow for modeling of complete systems for failure analysis prediction and the interaction of designs with an operating environment.



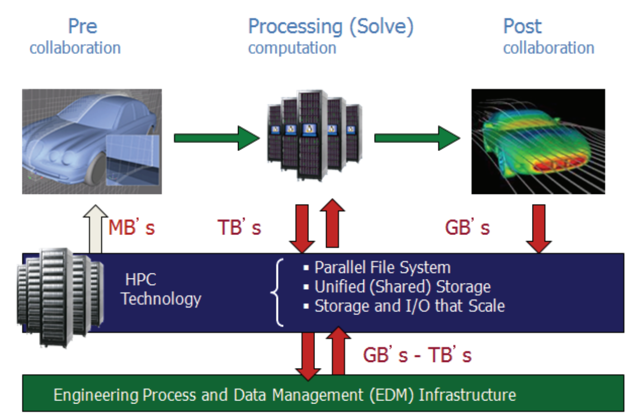
Examples of engineering disciplines found in the CAE industry include:

- Computational Structural Mechanics (CSM)
  - Strength
  - Fatigue
  - Noise, Vibration and Harshness (NVH)
  - Impact and contact testing
  - Failure modes
  - Crash testing
- Computational Fluid Dynamics (CFD)
  - Aerodynamics
  - Heat transfer
  - Fluid flow and transfer
- Computational Electromagnetics (CEM)
  - Electromagnetic compatibility (EMC) for sensors
  - Control systems
  - Antenna design
  - Safety systems
- Multi-Disciplinary Optimization (MDO)
  - Coupling of simulation codes to model interaction within assembled systems
    - Vehicle crash testing with occupant modeling
    - Engine electronic controls with engine combustion modeling

The Challenges

The table below highlights storage issues in technical computing.

Cluster	Client	Backup & Restore
Cluster waiting for I/O complete	Clients waiting to access job results	Backup interfering with other processes
Reduced cluster utilization	Scalability problems as users increase	Backup window pressures
Scalability problems as nodes increase	Hindrance to collaboration and sharing data	Inability to do fast backups and restores



Bottlenecks from I/O in the CAE Workflow

I/O bottlenecks in the CAE workflow:

The following are examples illustrate I/O bottlenecks that commonly occur in the CAE workflow, and how storage plays a crucial role in datacenter performance.

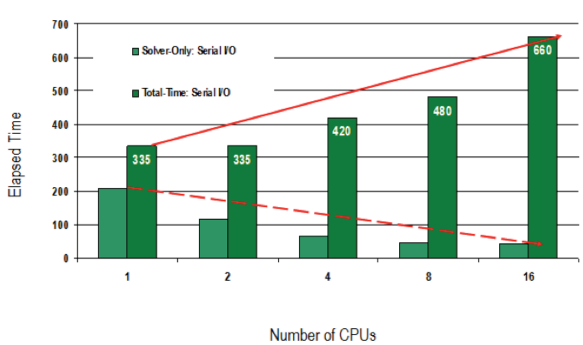
The engineer begins by the modeling of a component or system targeted for analysis. The model may describe the surface of a vehicle or airplane wing, or it could be the inside of an engine cylinder, to analyze combustion of a fuel-air mixture. Storage requirements for individual models or systems will typically fall into the MB to 100s of MB range, and are not critical I/O performance burdens.

After modeling is complete, the engineer will submit the simulation to a computational resource. At this stage, the compute resource will read the model and associated input files and compute a solution. Output during and upon completion of computation typically totals in the GB to TB of data range. I/O performance during this stage is paramount to overall job and compute resource performance.

Once the compute-phase is complete, where large amounts of data in the TB range are generated, we then enter the post-collaboration phase. During this phase, single-threaded programs ingest data generated during the compute-phase, manipulate that data and then save relevant state information, which is now in the GB-range of data volume. Data ingest is the flow of data from the storage medium to the CPU. Once the data is manipulated, it goes back to storage, but in a heavily processed form. Instead of purely raw numbers, the out-put of this post-collaboration phase could be formatted into graphics, slide shows, post-processing parameter files or even MP4 video streams. Data transfer from storage to the cluster becomes critical from a performance and data integrity perspective.

Job Scaling and I/O Performance:

The value that defines the total run time for a job in the compute stage is called wall clock time. Wall clock time can be broken down into two distinct categories. Solver time (or CPU time) is the amount of time the job spends computing actual results within the applications solver. I/O wait time is the amount of time that the job spends performing or waiting for data transactions in and out of the CPU, such as reading and writing to disk.



I/O for Transient CFD (transient is defined as non-steady flow)

In this case, as the job is scaled across an increasing number of CPUs, the number of threads performing I/O to the storage increases. Solver time per core is decreasing, but due to contention at the storage subsystem, I/O wait time is increasing. The result is, as the job is scaled to more cores, wall clock time for the job actually increases.

The solution

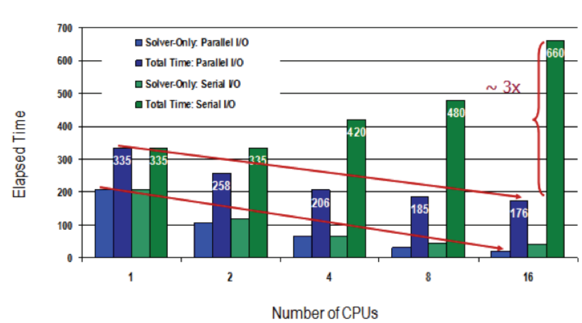
Parallel Clustered Storage vs. Serial NFS Storage:

Based up on the previous example, the architecture of a compute complex's I/O infrastructure is critical to job performance. A parallel clustered storage configuration based upon SGI IS5600, IS5500 or IS5000 storage arrays coupled with SGI Infinite Storage System Manager can effectively solve I/O bottlenecks. In the above example, we are able to construct a storage system that significantly reduces I/O wait times and reduces job wall-clock time.

The advantage to reduced job wall-clock time is speeding up the engineering design iteration cycle. By minimizing I/O wait times in a parallel clustered storage solution, job wall-clock time can be reduced to less than eight hours. At this rate, multiple design iterations can be performed in a single work day.

As an example, an engineer could review the results of an overnight run in the morning. Then, based on those results, make changes to the model, and resubmit the job for processing. The resubmitted job now runs for the next eight hours and is ready for review at the end of the day. If more adjustments are made, the job is modified, and resubmitted for over-night processing. In the morning, the engineer has a fresh set of data to review and continue or conclude the process.

The following chart graphically shows the serial and the parallel I/O scenarios and the resulting advantages. Note: in serial mode, while it is possible to tune the solver portion to reduce wait time dramatically by adding CPUs, the I/O wait time dramatically increases proportional to the CPUs added. Adding CPUs does not address the bottleneck between the server and the storage controller and array. In parallel mode, the solver wait-time is reduced by adding CPUs, and I/O wait time is reduced because of high-performance, dual-pathed controllers in the SGI IS5600 storage array.



Comparing serial and parallel I/O for transient CFD

Key technology benefits:

The four key tenets of the InfiniteStorage System Manager Software with dynamic disk pool (DDP) technology are:

- Simplified array management – no RAID to configure
- Active sparing – all drives working
- Streamlined capacity expansion – three steps to expand capacity
- Predictable performance over time – under all conditions

Key business benefits:

- Efficient resource utilization reduces cost of development
  - Improved margins
- Able to do more designs
  - Individually
  - In parallel
- Improved performance on compute investment
  - No-wait computer analysis

Data Protection:

Protecting data from hardware failure is always at the forefront of any storage design. In the past, this has been accomplished with RAID arrays using dedicated parity devices. If a device was to fail, the overall storage system can be healed using the remaining parity devices. Conventional RAID 5 storage uses one parity device and can sustain a single device failure without losing data. Likewise, RAID 6 storage uses two parity devices and can sustain two device failures.

With the increase in capacity of available disk devices, RAID rebuilds have become problematic. A RAID 5 device created with 4TB disks will take more than 72 hours to complete a rebuild. During that time, I/O performance to the storage is degraded. As part of the rebuild process, all data blocks on the RAID device must be read. If an additional data error is encountered in this process, the storage system will fail with data loss. RAID 6 attempts to mitigate this problem by adding a second parity device.

The InfiniteStorage System Manager with Dynamic Disk Pools:

Dynamic Disk Pools = Effortless Data Protection

Patent pending Dynamic Disk Pools (DDP) greatly simplifies traditional RAID management by distributing data parity information and spare capacity across a pool of drives, enabling easier future capacity expansion and greater protection. DDP offers improved data protection by quickly recovering a failed drive up to 8X faster than traditional RAID while maintaining greater performance.

A key concept of DDP is the dynamic rebalancing of data during changes in the number of drives, whether adding drives or in the case of drive failure. Unlike a traditional RAID volume group's rigid configuration with a specific number of drives, Dynamic Disk Pools can optimize from a minimum of 11 to the maximum of 384 drives supported by the IS5600 system. The result is protection of data in minutes vs. days.

A storage device constructed with DDP can lose multiple disks and still keep a file system up and running. If a disk drive failure within a DDP occurs, spare capacity configured into the pool masks the event. Unlike RAID, spare capacity is striped across the pool, and contributes to the performance of the storage device. Spares in DDP do not stand idle. Data in a DDP is then rebalanced across all remaining disks utilizing the incorporated spare capacity, restoring the storage device. Time to rebalance across the DDP is significantly less than rebuilding a typical RAID device.

IS5000, IS5500 & IS5600 Storage Arrays:

The IS5000, IS5500 & IS5600 storage arrays perpetuate flexible, modular architecture. The IS5000 storage system supports data-intensive bandwidth applications that benefit from the sustained high read and write throughput, while database-driven transactional applications benefit from its responsiveness and linear scalability. The IS5500 is equally adept at supporting high-performance file systems and bandwidth-intensive streaming applications. And its fully redundant I/O paths, advanced protection features, and extensive diagnostic capabilities deliver the highest levels of availability, integrity, and security. The IS5600 is a 8th generation storage system delivering industry leading bandwidth performance efficiency and data throughput per footprint.

The IS5X00 platform series offers multiple form factors and drive technology options to best meet requirements. The ultra-dense 60-drive disk shelf supports up to 240TB in just 4U and is perfect for environments with vast amounts of data and limited floor space. Its 24-drive shelf combines low power consumption and exceptional performance density with its cost-effective 2.5" drives. And the 12-drive shelf is a great fit for cost-conscious organizations that need to deploy both performance and capacity drives. All three shelves support the IS5X00 controllers, or can be used for expansion, enabling optimized configurations that best meet performance, capacity, or cost requirements.

The latest software release, InfiniteStorage System Manager (ISSM) 10.86, adds the foundation for SGI's AutoSupport services by Event-based or Time-based (weekly, daily, other) service criteria to be automatically gathered.

Global Sales and Support: sgi.com/global