SGI
INNOVATION FOR RESULTS™

White Paper

Scientific Workflow Solutions: A Multi-application, Multi-workflow Approach to High-performance Computing

Table of Contents

1.0 The Gap: Advancing Technology and Diverging User Requirements

Rapid advancements in technology continue to broaden the range of choices in the high-performance computing market:

- Clusters, blade servers, shared memory systems
- EPIC, proprietary RISC, x86 architectures
- Single-core, dual-core, quad-core implementations
- FPGAs, vector processors, computing and storage appliances

The market acceptance of Linux and open source software as viable development and deployment platforms has simplified application portability among these various architectures, and expanded their market reach. At the same time, scientific workflows have become more complex as multiple applications are linked together to automate production aspects of the research process. Supporting a single workflow now requires supporting multiple applications, and most medium to large size research organizations must support multiple distinct workflows with a single computing solution. But, how do you select the best computing platform? How do you optimize short-term results without limiting long-term flexibility?

Good decision making requires weighing a broad range of criteria and selecting the platform that best matches the application and workflow requirements. The total cost of the hardware, software, implementation, and operation are critical decision criteria, as are the amount and uniqueness of scientific results that can be generated. A solution's ability to solve complex problems, evaluate millions of possible scenarios, and its flexibility match changing application mixes must be matched with the vendor's ability to provide the delivery, reliability, and service quality required in production environments.

Today, selecting a new computing platform has often become a frustrating, time-consuming process. The ubiquitous application of computing technology has caused an explosion in the diversity of requirements that is matched only by the inventiveness of the computing industry in creating new computing platforms. In the past, one could turn to a few key R&D leaders who would often define scientific requirements for high-performance computing, but the open-source driven democratization of scientific research means that everyone has a the opportunity to create the next big breakthrough, and computing solutions must be tailored to the specific mix of applications run at a particular site.

A widening gap between computing architectures and many technical users' requirements has opened up in the industry. Several trends have resulted:

- The diversity of technical computing applications continues to grow. For example, within pharmaceutical industry, high-throughput computing is used to screen millions of potential compounds in silico using automated structure prediction and docking techniques, while single simulations of complex ligand-protein interactions can run for days or weeks on even the most powerful of computers
- Leading edge research is often at the boundary between two or more disciplines, and the simulation of complex physical phenomena that cross these disciplines are routinely undertaken. Global warming research already combines simulations of the atmosphere, oceans, land and ice and increasingly includes atmospheric chemistry and dynamic biological processes
- A single computing architecture is decreasingly able to optimize both the high-throughput needs of a scientific workflow and its requirement for in-depth detailed simulations. Dramatically different technologies are required to accelerate different portions of a single workflow

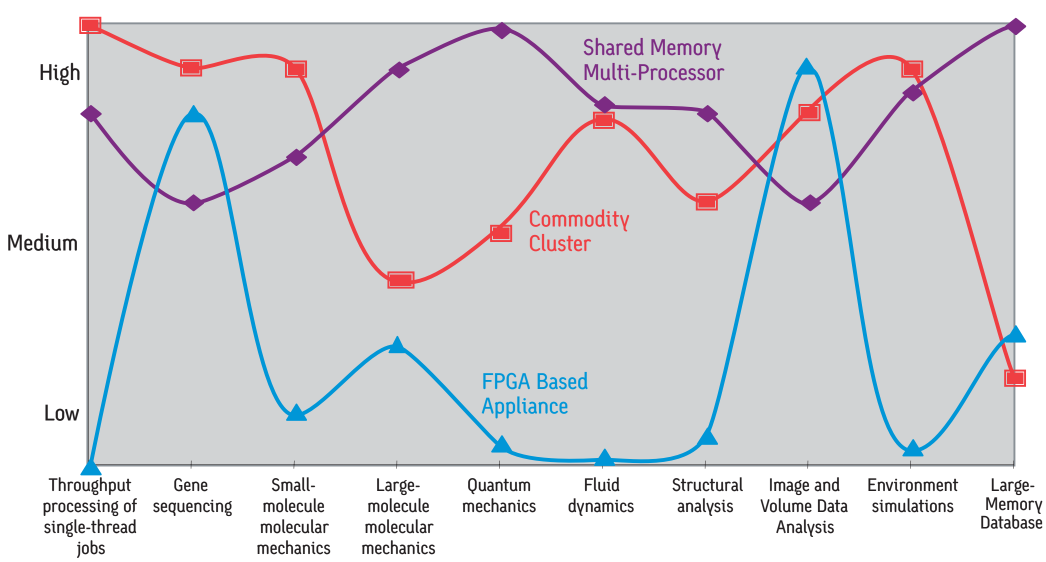

Figure 1. Applicability of Hardware Architectures

- User productivity is often declining as non-expert users are forced to navigate a sea of different computer systems and data storage solutions. At the same time, a new generation of workflow management tools are emerging with will simplify, if not eliminate the need for scientists to worry about the computers they use
- Data sets continue to grow exponentially, and their exploitation has become a leading indicator of the success of a particular organization. For instance, regularly used proteomics data bases are 8 terabytes in size, and are rapidly expanding. Working with full data sets instead of partial problems would improve the productivity of individual researchers

Divergence between current computing platforms and application requirements present in single- or multi-workflow environment translates into losses of productivity. Still, many high-performance computing solution suppliers continue to offer a single architecture for all situations, arguing that workflows must be adjusted as part of the performance optimization process.

2.0 A Better Approach: A Scientific Workflow Solution

The typical scientific computing environment must support a hierarchy of different applications and research workflows that together add up to the organization's scientific enterprise workload. Rather than force-fitting all applications and workflows onto

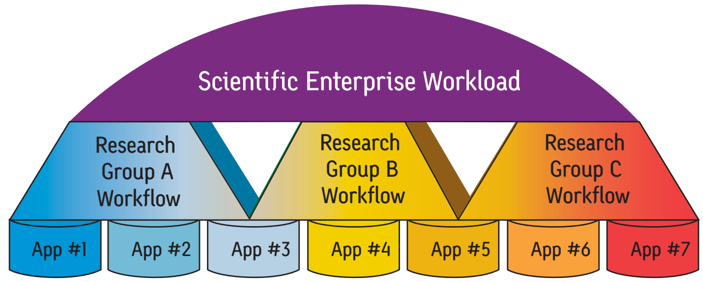

Figure 2. The Components of a Scientific Enterprise Workload

a single computing architecture, it is possible to develop a high-performance computing solution which provides a uniform end-user view but is customized to the needs of each particular customer. To be successful, a scientific workflow solution must be able to optimally support each application of a specific workflow and be able to adapt to the unpredictable nature of scientific research. To illustrate, consider the following two workflows, each of which may be present within a single scientific enterprise workload:

Selecting a Scientific Workflow Solution is not an either-or decision, but rather one that maximizes research results. The first step in evaluating a solution is to characterize each application within each workflow, and then aggregate similar requirements

	High-performance on Insight Workflow	High-throughput or Screening Workflow
Results Accelerated By:	Running a single job faster or with more detail. This requires many processors working together to shorten time to insight.	Quickly running a large number of independent jobs, and not from how long it takes to run a single job.
Applications Benefit From:	Shared memory architecture that offers highly scalable amounts of memory, I/O and storage and delivers low memory latency and high memory bandwidth across the system.	Low-cost, industry standard nodes that offer small amounts of internal scaling but share a common queuing system and a uniform, but not necessarily fast, access to storage.
Workflows Require:	A high level of human involvement to define inputs and analyze results from complex applications, and then transform them into inputs to the next stage of the research workflow.	A minimum level of human involvement. Applications often generate results from databases or statistical means, with the output of one application used as the input to another in an automated research workflow.
Resource Demands Are:	Non-deterministic since different applications are used in different stages of the pipelines, and each requires different amounts of resources. The size and frequency of jobs are not fully predictable so they are difficult to schedule.	Highly deterministic since each job uses a consistent amount of resources and runs in a predictable amount of time. The workflow is easy to schedule, and may take advantage of dedicated computing systems or appliances.
Costs Are Driven By:	4-6 hour turn-around for "typical" jobs, and the ability of the system to solve infrequent but critical tasks over the weekend.	Industry standard price/performance of the complete compute, storage and system management environment.

Table 1: Workload Characteristics

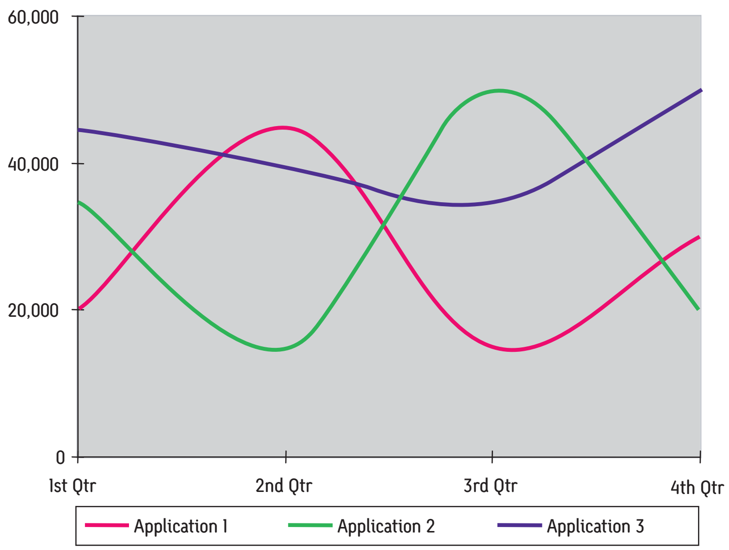

Figure 3a: Total 4 Quarter Application Demand (Hours)

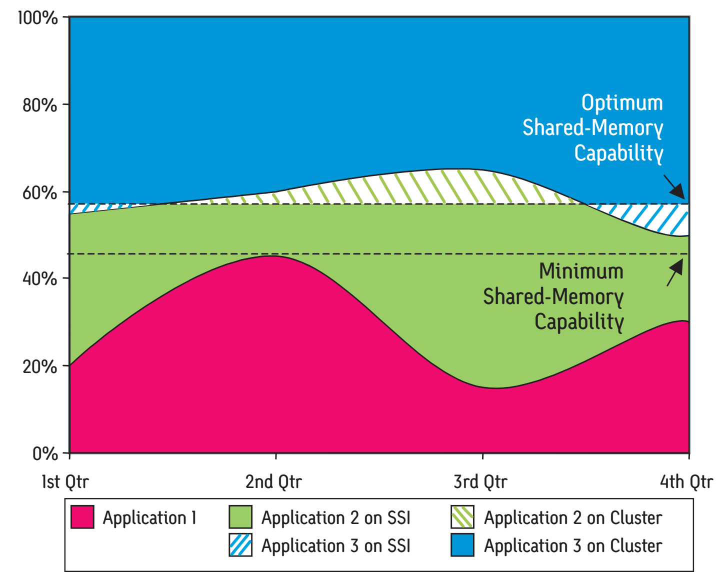

Figure 3b: Min-Max Aggregate Application Demand Analysis

across the organization being served by the solution. Most evaluations result in two or three different groups of similar applications, each one of which may point towards a different "best of breed" platform that optimally supports those needs. As a result, the typical Scientific Workflow Solution will be made up of a multiple platforms.

It is important to note, however, that the mix of platforms in the Scientific Workflow Solution may not be a one-to-one match to the loads defined in the previous analysis. This is because research workflows change over time and it is critical to evaluate how well each platform supports other classes of applications in case the organization needs to redirect its research focus more rapidly than the procurement cycle allows. This can be done with a "Max-Min" evaluation that provides a way of minimizing the downside risk of an investment – or at least making the decision explicit.

A visual example of a Max-Min analysis is shown in Figure 3a, where a scientific enterprise needs to run three different groups of applications – one that only runs on a high-performance shared-memory system, one that runs best on the high-performance shared-memory system, and one that runs best on a high-throughput cluster system. The upsides and downsides are listed in Table 1. By seeking to maximize the minimum result, we can see that it is critical that 100% of the peak shared-memory need for Application 1, beyond that we need to make a choice. Since it is slightly better to run Application 2 on the cluster than it is to run Application 1 on the shared memory system, we should have a slight bias towards the cluster. One option would be to minimize the size of the shared memory system so that all of Application 3 runs on the cluster, while Application 2 meets 60% of its peak requirements and 50% of it's aggregate requirement is met with the minimal shared memory system. However, if the size of the shared memory system is increased slightly, as shown in Figure 3b, with "excess requirements" of both Application 2 and Application 3 running on a second system, then the overall efficiency can be increased. Thus, the "best" solution is to have a multi-platform workflow system which has slightly more than the minimum amount of shared-memory capability and the rest complemented with a high-throughput cluster. When the workflow grows or changes, it will be easy for this organization to meet its short term needs while the budget catches up.

To be successful, the Scientific Workflow Solution must also provide users with the look and feel of a single system by integrating multiple physical systems into a single virtual system with common methods of accessing computational power and data. Without adequate integration at the systems level, users would be required to understand and deal with multiple systems, operating systems, and queuing environments. The integration process must also meet the market's demands in terms of reliability, supportability, and provide efficient data with the end-user's desktop environment.

By considering current and future workflow needs and then combining a series of compute and storage platforms that optimally meet compute, memory, I/O, networking and storage requirements of critical applications, a multi-application, multi-workflow solution provides significant value to the many users and organizations seeking innovation and breakthrough results.

3.0 Implementing Scientific Workflow Solutions

Over the past 20+ years, SGI has become known for solving the most challenging customer computing and data management problems. SGI has used this experience to define a modular solution platform that specifically address the variable requirements of mixed workflow environments. This platform integrates SGI's full line of scalable shared-memory systems, cluster compute systems and storage capabilities in a way that can be tuned to match specific requirements.

High-throughput and Screening Workflows. SGI recently extended its server portfolio with the introduction of a new line of servers that efficiently address the high-throughput and screening workflows with a value-priced cluster solution. The new SGI® Altix® XE platform is based on the dual-core Intel Xeon® 5100 sequence processor with up to 4 CPU cores per server that can be used to build high-performance cluster computing solutions. This new Altix XE cluster server line complements the SGI® Altix® scalable shared memory product line by offloading throughput oriented tasks that do not benefit from the high levels of compute, memory and I/O scalability. SGI Altix XE is a customizable cluster solution that ships to the customer fully-integrated from SGI's factory. This enables customers to easily buy and deploy a full-featured cluster solution, without sacrificing functionality or performance.

High-performance and Insight Workflows. The SGI Altix family of scalable, shared memory servers was introduced in 2003, and leads the industry in single system scalability and performance. The newest SGI Altix 4700 and 450 servers run the Linux® operating system and use dual-core Intel® Itanium® 2 9000 sequence processors to increase insight generated by solving large, complex problems. SGI Altix servers are built using a "plug and solve" modular blade architecture that scales to 1000s of processors and terabytes of memory. System resources, such as compute, memory, I/O, and even FPGA technology can be independently configured to suit the needs of any workload. The current Altix 4700 and 450 systems effectively double performance density and scalability, while consuming less power. They also introduce peer I/O technology, which further enhances performance by enabling application-specific co-processors like FPGAs to tap into the shared memory address space at full system bandwidth. This innovative new architecture not only improves performance today but paves the way for the incorporation of future computing technologies in a highly efficient manner.

The SGI server platform support the entire spectrum of computational requirements, yet are easy to mix and match to support a single multi-application workflow or multiple workflows that span a scientific enterprise. They use a unified end-user access and system management infrastructure that is based on industry-standard Linux operating system, a comprehensive storage offering, and common development and administrative toolsets across the platform.

Using this approach, the breadth of hardware platforms and technologies does not require users to master multiple computing architectures. Common system-level characteristics make it possible to move from one platform to another, and to manage projects that span multiple systems and technologies:

- A common Linux operating system across all platforms so that users need learn only one environment

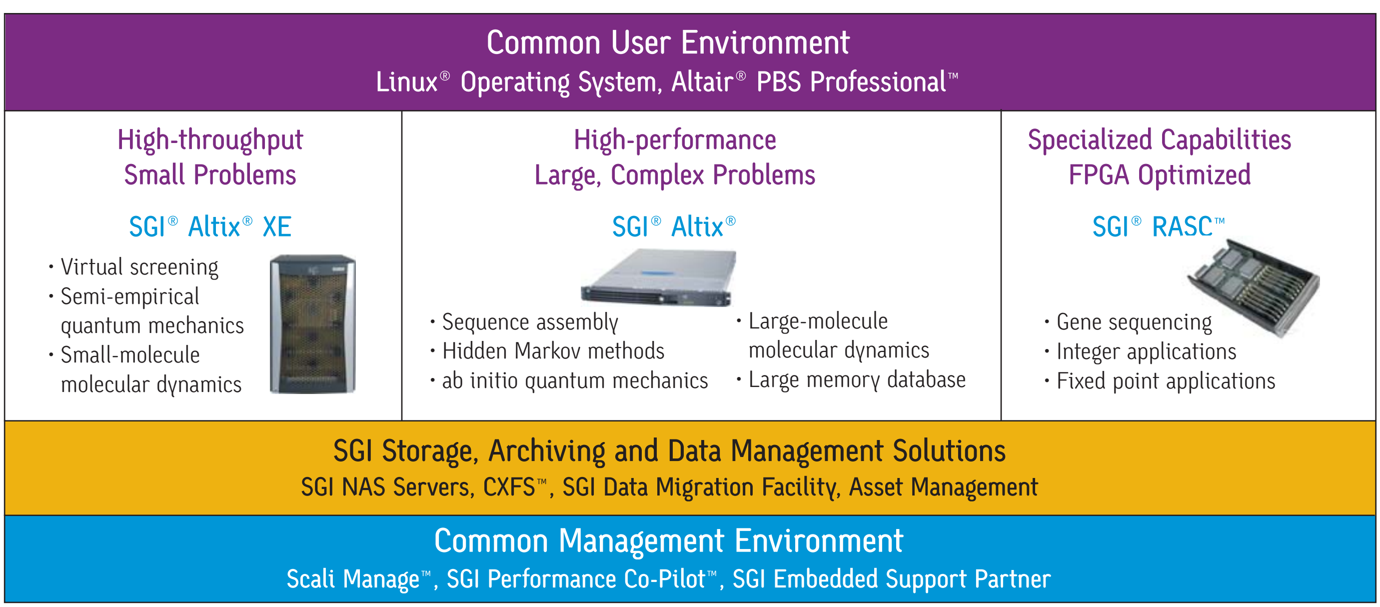

Figure 4: Scientific Workflow Solution for BioScience

- Common compilers and development tools that enable application portability
- A single job scheduling/workload management system to manage workflow across all systems
- A centralized systems management interface and operational dashboard for updating, provisioning, configuring, and monitoring
- Platform-wide storage solutions to simplify data sharing, access, movement, and management across systems
- Optional system specific performance and tuning enhancements through SGI® ProPack® for Linux®

The common access and management environment available with the SGI Scientific Workflow Solution enables customers to deploy the best point-solutions that target specific workflow requirements without having to expose the presence of multiple independent systems to end-users. This environment bridges the gap between divergent application requirements and system architecture, making it realistic to deploy high-performance computing solutions in mainstream scientific and research environments. And deliver both scientific breakthroughs and new levels of resource utilization and efficiency.

4.0 Extending Scientific Workflow Solutions

Today, advances in the performance of processor cores and cluster technologies increasingly enable the use of mass market processors for improved levels of application performance. But, as high-performance solutions become limited by the combination of Moore's Law for CPUs and Amdahl's law for cluster-based parallelism, alternate strategies will increasingly factor into advanced architectures to drive performance and scalability in the multi-application, multi-workflow environments. To that end, SGI is once again leading the industry with a multi-paradigm computing strategy to advance its shared memory architecture by supporting a variety of tightly coupled alternative processing elements. This approach is based on the concept of attaching devices to the core SGI scalable shared-memory architecture, to deliver the best-performing processing elements for particular applications.

SGI has introduced this capability to the market in its Altix 4700 and 450 platforms, which integrate SGI's peer I/O technology, to provide direct, high-speed access to SGI's large shared memory for all system components. Through peer I/O, SGI Altix systems are able to support new computing paradigms with high efficiency. In the sciences, an early example of this environment is the use of reconfigurable FPGA based computing elements for genome sequencing applications. Because FPGAs directly implement algorithms, they offer the opportunity to break through the limits of Moore's Law.

5.0 Long-Term Business Value

Within the Scientific Workflow Solution, individual platforms can independently evolve and incorporate new innovations at the earliest possible moment. The common software environment minimizes the impact on end users and applications already deployed on the platforms, while easily integrating new systems or upgrading old ones. Some of the key business values delivered with the SGI Scientific Workflow Solution are derived from:

- End-user focus: Scientists can spend more time focusing on science because they only have to learn a single user environment, a single set of system access tools, and a single way of accessing data
- Maximum efficiency: Applications run on platforms that are highly tuned to their needs, reducing time to insight and scientific results
- Fully integrated platforms: Systems and clusters ship pre-configured from the SGI factor so you can quickly start generating results
- Common software: Industry-standard operating environments, and common development and workload management solutions to provide a smooth path to future platforms
- Ease of administration: Flexible system configuration options (blade servers, clusters, resource scaling) and common toolsets across comprehensive environment

This type of "next generation" high-performance computing solution effectively applies multiple technologies to the job at hand, and enables real-world workflows to be optimally matched with the most suitable technology or combination of technologies. The best technology "fit" can then be applied to meet specific workflow and business requirements, but in a unified compute strategy.

The freedom to choose the best platform, without adding complexity to the user computing environment, provides users with a superior solution and the ability to maximize productivity for each computing project. The consistent software and system environment across all hardware technologies also protects user investments: software can be easily migrated to new systems, and hardware resources can be reallocated if project requirements change. The multi-workflow architecture makes the underlying technology transparent to the users, applications, and workflows, introducing operational efficiencies at every level of the organization.

6.0 Conclusions

Today, users recognize the benefit that high-performance computing applications can bring to the organization in all types of industries. Scientists are applying high-throughput and high-performance computing platforms to an increasingly broad range of uses. The SGI Scientific Workflow Solution helps bridge the gap between divergent user requirements and computing architectures, helping unlock the value of computing for scientific enterprises.

Multi-application, multi-workflow solutions like the SGI Scientific Workflow Solution, is an exciting new approach that enables organizations to deploy the best computing platform for each application workload. The SGI Scientific Workflow Solution helps improve efficiency with better, more transparent access to computing resources, better data flow, and better system and data management.

SGI
1140 E. Arques Ave.
Sunnyvale, CA 94085-4602
650.960.1980
www.sgi.com

North America +1 800.800.7441
Latin America +55 11.5185.2860
Europe +44 118.912.7500
Japan +81 3.5488.1811
Asia Pacific +1 650.933.3000

3967 [02.01.2007]
J15193