insideHPC

The insideHPC Guide to

HPC in Life Sciences

BROUGHT TO YOU BY

sgi | intel

Introduction

The term next generation sequencing (NGS) is really a misnomer. NGS implies a single methodology, but the fact is that over the past 10 to 15 years there have been multiple generations and the end is nowhere in sight. Technological advances in the field are continuing to emerge at a record setting pace.

Early sequencing efforts used a methodology known as Sanger chemistry, which employs specifically labeled nucleotides to read through a DNA template during DNA synthesis. Despite technical innovations, after 25 years the Sanger method had reached the limits of its capability to handle large genomic datasets.

NGS, an important new approach to sequencing, removed these roadblocks. The technology achieved a major milestone in 2000 when Craig Venter, founder of Celera Genomics, and Francis Collins of the National Institute of Health and the U.S. Public Genome Project used the newly developed NGS technology to accomplish the complete mapping of the human genome some three years ahead of the expected timeline set by the Public Genome Program. The NGS solution they employed, known as shotgun sequencing, uses chemical and mechanical methods that allow researchers to sequence DNA and RNA more quickly and cheaply than the Sanger method. Because the speed of sequencing has been ramped by many orders of magnitude, scientists can now sequence vast amounts of data.

One of the results is that today's NGS workflows are more computer science than biology. It is now possible to do a complete human genome — about three billion nucleotides — in a few days for around one thousand dollars. Fifteen years ago this would have been a billion dollar project that took multiple man-years of effort.

Contents

Genomics and Big Data

Advances in computational biology as applied to NGS workflows have led to an explosion of sequencing data. All that data has to be sequenced, transformed, analyzed, and stored. The machines capable of performing these computations at one point cost millions of dollars, but today the price tag has dropped into the hundreds of thousands of dollars range. A large sequencing center might have 20 or 30 of these machines; a few larger organizations might have over 100. The aggregate cost is still so large it's no wonder that labs both large and small want to feed these NGS systems a constant stream of new samples for sequencing on a 24x7 basis.

In 2013 the market size was about $4.6 billion, a 25% increase over the previous year. Continued growth in the field will boost the sequencing market to more the $10 billion by 2017.

According to a Global and China Industry Research Report[1], with the advances in sequencing technology and a "dramatic" decline in sequencing costs, the sequencing market is growing rapidly. The report states that in 2013 the market size was about $4.6 billion, a 25% increase over the previous year. Continued growth in the field will boost the sequencing market to more the $10 billion by 2017.

But despite this and other equally optimistic sequencing market forecasts[2], there are still many problems to be overcome, many of them computational. NGS computational requirements have been outpacing computational advances as forecasted by Moore's Law[3] and the rapid rise of hard disk capacity as observed by Kryder's Law[4].

[1] See: http://www.prnewswire.com/news-releases/dna-sequencing-market-to-hit-10-billion-by-2017-says-a-global-and-china-industry-2014-research-report-279012231.html
[2] http://www.transparencymarketresearch.com/dna-sequencing-market.html
[3] http://en.wikipedia.org/wiki/Moore%27s_law
[4] http://en.wikipedia.org/wiki/Mark_Kryder
[5] A computer architecture in which each processor is connected to *n* others based on the geometry of a hypercube of *n* dimensions. See: http://www.yourdictionary.com/hypercube

From the Lab to the Datacenter

In the late 1980s, genomic sequencing began to shift from wet lab work to a computationally intensive science; by end of the 1990s this trend was in full swing. The application of computer science and high performance computing (HPC) to these biological problems became the normal mode of operation for many molecular biologists.

Most of these researchers did not have a history of working with expensive supercomputers — typically they relied on commodity Linux workstations and small HPC clusters to meet their sequencing needs.

The problem is that for many genomic workflows, this kind of distributed computing model was, and still is, not the best one to use. When you are dealing with thousands of cores in a very large system with distributed nodes, each with its own, unshared memory, node-to-node communications becomes an issue. The limited amount of RAM available to each CPU can also become a bottleneck. Both issues can be addressed with globally shared memory.

Solutions such as hypercube architecture[5] are being implemented to improve the design of large clusters. Also, given the terabytes of data being generated by today's NGS machines — and the need to sequence and analyze genomic data quickly and cost-effectively — parallel processing is essential. However much of the sequencing code in use today is not parallelizable or is embarrassingly parallel at best. Very little code optimization has occurred with NGS applications over the past fifteen years, certainly as compared to other computational solutions in other research disciplines.

Even if the code is modernized, conventional distributed Linux HPC clusters are difficult to scale to keep up with the torrents of genomics data being processed. Just throwing more cores at the problem is not the answer. For example, consider the computational requirements of processing an entire genomic dataset containing some 10TB of data. A standard Linux cluster or a cloud solution such as Amazon EC2 can't squeeze a 10TB dataset into memory. Today even a terabyte of memory on a fat node is a luxury, and the task (and cost) of transferring that amount of data over the Internet to remote storage is still a major challenge.

The growing capabilities and complexities of today's NGS tools are adding to extensive and unpredictable flows of data that threaten to swamp conventional server and storage solutions. In fact, many current complex genomics problems are beyond the reach of these architectures. In many instances, code written to run on distributed commodity Linux clusters uses algorithms that cannot handle the massive amounts of data that have become the norm in the latest NGS incarnations.

Also, as NGS systems and applications continue to rapidly evolve, storage is a key consideration. The size and complexity of the datasets produced by these systems, sometimes running to petabytes, continue to place high demands on the capacity and throughput of the IT infrastructure's storage systems, whether it is part of an HPC cluster or a supercomputer.

Fortunately there is a solution. It is under these challenging circumstances that a unique architectural approach like SGI's NUMA (non-uniform memory access) and a large shared-memory machine, like the SGI UV system, really come into their own.

Beyond Commodity Clusters

Commodity clusters, mentioned above, are typically based on a number of thin rack-mounted or blade servers with limited memory provided on each server. When an application running on one node needs to access the memory on a different node, the application has to communicate with the node where the memory resides and issue special commands to retrieve the desired information. In Big Data situations, this can lead to serious bottlenecks, resulting in excessive processing time or, in some cases, making it impossible to run jobs composed of large, distributed datasets.

SGI addresses this and other computational and storage problems inherent in NGS applications by providing a broad technology portfolio that includes coherent shared memory (CSM) and distributed memory compute platforms, along with a full range of integrated storage solutions. Included is the NUMA architecture, a system that links multiple nodes of processors and/or memory over a special high-speed network. Under NUMA a processor can access its own local memory faster than non-local memory residing on other processors or shared between processors.

Specifically the SGI solution allows all memory that resides on any node or blade to be completely visible to any of the cores in the system. All data within memory is available using direct access mechanisms, regardless of where the memory actually resides. The SGI UV family of servers, featuring Intel® Xeon® E5-4600 processors, is based on the NUMA architecture and provides the full range of CSM capabilities without modifying the existing IT infrastructure. The SGI UV uses standard Intel processors, and runs standard Linux — no need to rewrite code or recompile to achieve top performance as well as compatibility with older applications.

The SGI UV 2000

The SGI UV 2000, featuring the Intel® Xeon® processor E5-4600 product family, is the sixth generation of the company's scalable global shared memory architecture. SGI began shipping systems with CSM capabilities with the 64-core Origin systems in 1996. The newest GSM (global shared memory) platform — the SGI UV 2000 — scales up to 256 CPU sockets and 64TB of cache-coherent shared memory in a single system.

The platforms are built using the SGI NUMAlink interconnect, which provides the high-bandwidth, low latency and coherence-optimized functionality required by GSM platforms. The NUMAlink interconnect fabric can also be used for efficient communications between OS instances. This supports scaling up to many thousands of CPU cores for shared memory applications, as well as for MPI applications and applications developed using partitioned global address space compilers like OpenMP or Unified Parallel C.

The SGI UV 2000 features a modular chassis design that allows users to grow their system without adding complexity. A 10U chassis contains up to 16 sockets and 24 threads coupled with an All-to-All NUMAlink network topology. By adding additional chassis — up to four per standard 19" rack — and using Enhanced Hypercube topology, the system can scale up to 256 sockets and 4,000 threads, all operating as a single system.

A New Solution

The SGI UV system allows computational biologists to take a different, far more effective approach to dealing with the huge data sets generated by contemporary NGS sequencing and analytic applications. Traditionally they have relied on Hadoop-style solutions — break the problem up into small chunks, send the chunks out across a distributed architecture for processing, and then reassemble the results. This approach has limited utility in genomics. Scientists and computational biologists are more likely to achieve breakthrough insights when looking at the entire dataset at once. This is not about finding a needle in a haystack — it's looking at the entire haystack in order to define new needles. And that can only happen if you can hold the entire haystack in immediately accessible memory, which is what a CSM architecture allows you to do.

Another reason to use the SGI UV for genomics is that although it is made up of modular chassis, the system looks to the user and IT like one giant machine with 32 PCI slots available in a single partition. Physically the SGI UV is architected as a blade server and each blade has its own PCI slot. Intel Xeon Phi coprocessor cards can be inserted in these slots.

For example, a 5,000 core Xeon cluster that also contains another 2,500 Intel® Xeon Phi™ cores all looks like one memory footprint. The result is a very high end, scalable machine; no other system on the marketplace comes close to providing the total aggregate RAM and RAM per core access that the SGI UV does.

Maximizing Throughput

SGI's UV systems are complemented by a variety of capabilities including the free SGI High Throughput Computing (HTC) Wrapper Program for Bioinformatics. HTC helps maximize throughput on HPC systems — a major issue when thousands of jobs are running, each carrying its own load of management overhead.

HTC is a wrapper that presorts the data to ensure that the largest and longest jobs are run first and the smaller jobs last. These are distributed in a balanced way over the system without the overhead associated with using a third party scheduler like LSF or PBS.

> Because the SGI UV is based on industry standards, the system is easily integrated into a hetero-geneous datacenter environment.

Solving the Storage Problem

In addition to scheduling issues, storage is a major component of effectively dealing with the massive amounts of structured and unstructured data typical of genomics and life sciences applications, including the huge datasets flowing from NGS machines operating on a 24/7 basis.

To meet this challenge, SGI has developed ArcFiniti, a disk-based active archive solution. Based on the Intel® Xeon® E5 processor family, this is an integrated hardware and software platform designed specifically to handle large amounts of unstructured file-based data, which constitutes the bulk of the data being generated by the constantly evolving, advanced sequencing machines.

The solution includes patented SGI technology that significantly reduces power consumption and ensures data integrity. ArcFiniti is available in five different configurations, ranging from 156TB to 1.4PB of usable storage in a single rack before compression. Not only does this density result in significant infrastructure savings, but also allows users to quickly and easily access the archived data.

Plug and Play

One of the SGI UV's features much appreciated by IT departments is its simple deployment. The system arrives at its destination fully configured and ready to plug in. Unless the customer has specified special customization, the system can be running in minutes.

And, because the SGI UV is based on industry standards — Intel processors, Linux OS, and standard I/O and management interfaces — the system is easily integrated into a heterogeneous datacenter environment.

Key Benefits

The key benefits of the SGI UV 2000 architecture include:

- **Massive In-Core Computation**
  The SGI UV 2000 allows much larger and more detailed models and simulations of physical systems, or any large data set, to be entirely memory resident.
- **Massively Memory Mapped I/O**
  For applications that are bound by random I/O accesses on large data sets, the SGI UV with the Intel® Xeon® processor E5-4600 product family offers up to a 1,000x performance increase by enabling entire datasets to be brought into main memory.
- **Highly Efficient Application Scaling and Message Passing**
  SGI UV uses an array of advanced hardware and software features to offload thread synchronization, data sharing and message passing overhead from CPUs accelerating critical tasks by up to 100x.
- **Greatly Simplified Application Load Balancing**
  In cluster computing environments, each node completes its own threads and then waits until all other nodes complete their assigned tasks. The global shared memory available with the SGI UV with the Intel® Xeon® processor E5-4600 product family allows processors that finish early to also work on other threads, since each processor has access to all data and synchronization points through globally shared memory.
- **Smooth Scaling of Application Size and Complexity**
  In most cluster environments, applications run on a fixed number of nodes, each with a fixed amount of CPU cores and memory. Applications run into a "wall" when they exceed the fixed amount of memory per core or node in the cluster. Applications running on an SGI UV scale smoothly by drawing on additional memory distributed throughout the system.
- **Petascale System and Application Scalability**
  In addition to global shared memory support where all resources are shared by a single copy of the operating system, SGI UV provides an even larger Globally Addressable Memory (GAM) which enables systems to be built that extend beyond the shared memory reach of any given processor or OS implementation.

The SGI UV system is uniquely suited for bioinformatics and genomics by providing the computational capabilities and global shared memory architecture needed for even the most demanding sequencing and analytic tasks, including post sequencing and other data intensive workflows.

Because of the systems outstanding speed and throughput, genomics researchers can perform very large jobs in less time, realizing a dramatically accelerated time-to-solution.

And best of all, they can explore avenues of research that were computationally beyond the reach of HPC systems lacking the power and in-memory capabilities of the SGI UV.

To learn more about how SGI improves large storage problems in Life Sciences please visit: http://www.sgi.com/solutions/research/