

**10<sup>18</sup>**

**Bull  
exascale  
program**

**Statement  
Of Direction**

**November 2014**



# Contents

- Executive summary ..... 1
- Exascale challenges and benefits..... 2
- SEQUANA: the foundation for exascale ..... 3
- Software comes into play..... 3
- Communication overhead is the hurdle ..... 4
- It's time to merge HPC and Big Data ..... 5
- Understanding challenges and providing expertise..... 6
- Long-standing experience in the HPC field ..... 6
- Milestones ..... 6
- Glossary ..... 6
- About Bull, an Atos company..... 7
- About Atos ..... 7
- Acknowledgement ..... 7

## Executive summary

Bull, an Atos company, is a leader in Big Data, HPC and cyber-security with a worldwide market presence. Bull has extensive experience in implementing and running petaflops-scale supercomputers. The exascale program is a new step forward in Bull's strategy to deliver exascale supercomputers capable of addressing the new challenges of science, industry and society.

The Bull exascale program aims to address technologies and processes to meet the requirements of High-Performance Computing (HPC) and Big Data: i.e. the rapid computation and analysis of massive data sets. In a first step, it involves developing five main elements:

- An open exascale platform (code-named SEQUANA)
- A complete software stack (named bullx supercomputer suite)
- A new generation of fast interconnect (code-named BXI)
- A range of servers with ultra-high memory capacity (named bullx S6000 series)
- A complete set of services to optimize customer applications and make the most of exascale.

Future scalability for all applications depends on the system's ability to deliver data to the processing elements. For compute-intensive HPC workloads, the BXI interconnect is the foundation that will enable the next generation of application scalability, while the Bull S6000 series incorporates in-memory computing capabilities that accelerate the new breed of analytics and data intensive applications. The SEQUANA system will bring the merger of HPC and Big Data applications to exascale.

The Bull exascale program is a key enabling element for Europe's competitiveness and sovereignty. HPC development in Europe is driven by the need to address societal and economic challenges that require extreme computational resources, as well as by industry's need to innovate using computer simulation and intensive data processing.

With this in mind, the European Commission has set an ambitious goal for Europe to become a leader in the HPC field. The Bull exascale program leverages this initiative. It is also aligned with the Strategic Research Agenda proposed by the ETP4HPC. The ETP4HPC technology platform is one of the cornerstones of European HPC. It brings together leading industrial and academic players carrying out R&D in HPC technologies in Europe. Bull is one of the ETP4HPC members along with the CEA, a long-time R&D partner of Bull.

The development of an exascale computing capability – with machines capable of executing a billion billion of operations per second - is characterized by significant and disruptive changes in computing hardware architecture, related to compute resource and data concurrency, data movement, energy efficiency and resilience.

The Bull exascale program makes possible the simulation and analysis of the even most complex phenomena. It directly benefits from Bull's mission and synergies when it comes to delivering Big Data, HPC and security solutions. It is the result of major investments in R&D and of the long-standing experience and expertise of Bull's HPC teams in designing, implementing and running large HPC systems.

It will consist of a set of development programs to deliver new solutions with new technologies, combined with professional services and the expertise needed to reengineer applications and secure the transition. The Bull Center for Excellence in Parallel Programming is a cornerstone of this program to support users, reengineer their HPC applications and secure their evolution to exascale.

## Exascale challenges and benefits

There are natural synergies between HPC and Big Data, and both are essential for next-generation breakthroughs and innovation. Besides enabling computer simulations, HPC is also becoming the chosen means for analyzing large amounts of data. The development of sensors, measurement systems and human data production will generate an overwhelming amount of data that can only be modelled and analyzed with the help of HPC.

All these challenges are currently at the crossroads where extreme computing meets extreme data. In 2013 the overall volume of data reached 4.4 Zettabytes and by 2020 it is expected to soar up to 44 Zettabytes, according to recent IDC research carried out for EMC.

The best example is genomics, where the detailed genomes of millions of species or human beings will soon be available. It is only possible to analyze this data and turn it into useful information for understanding biological mechanisms with large-scale computing resources. HPC applications in this sector will lead to better risk assessment of diseases, to customized therapies and to more efficient drugs. Complex transportation, the optimization of energy grids, environmental risk prevention and better weather forecasting are other examples requiring huge computing power and data bandwidth.

### New technologies and solutions are required

Today, new technologies and solutions are required; otherwise progress in the key fields will falter. The Bull exascale program meets this essential need.

The program encompasses the new challenges of fundamental science, applied sciences, manufacturing, finance, oil and gas, environmental science, defense, healthcare, life sciences and mainstream IT business. It will increase productivity and competitiveness, facilitate time-constrained processes and speed up decision-making.

The Bull exascale program will enable organizations to design new products faster and more efficiently, to better anticipate multi-dimensional phenomena such as climate change and simulate them to a finer level of detail, to analyze massive data flows in real time, and to protect society from threats, reduce their impact and manage crisis situations as they unfold.

Tight integration between extreme computing and extreme data is the cornerstone of the Bull exascale program. It unifies theory, experimentation and simulation with large-scale multidisciplinary data. It combines more flops, greater data capacity, faster access to data, global data access, on-going data management, energy efficiency, resilience, ease of use and scalable packaging. It will enable applications to capture, store, compute, analyze and visualize very massive data models as fast as possible.

### A never-ending need for more performance

The Bull exascale program aims to ensure that HPC performance will continue to increase in the foreseeable future. That means a complete change in scale compared to today's petascale systems, but with more constraints on size and power footprints, while improving system resiliency and usability for current and new categories of users.

*"The most efficient system needs one to two megawatts per petaflop/s. Multiply that by 1,000 to get to exascale and the power is simply unaffordable".*  
Horst Simon, Deputy Director of Lawrence Berkeley National Laboratory

## SEQUANA: the foundation for exascale

### SEQUANA at a glance

SEQUANA is designed as an open exascale platform to easily integrate and make the most of current and future technologies. It is designed to be compatible with successive generations of different technologies (CPUs and accelerators). It supports fast state-of-the-art interconnects, such as BXI.

SEQUANA addresses the balance between processor performance, memory capacity and access, as well as data movement, both between compute nodes and into or out of storage. It supports large configurations of tens of thousands of nodes grouped into large building blocks, the most efficient direct liquid cooling technologies, and the most demanding data management

### An innovative design targeting performance and efficiency

SEQUANA is designed for exascale configurations. It is highly modular, cost-effective and scales from a few hundred nodes to tens of thousands of nodes. The computing resources are grouped into cells. Each cell tightly integrates compute nodes, interconnect switches, redundant power supply units, redundant liquid cooling heat exchangers, a distributed management and diskless support.

The packaging has been designed to facilitate high-scale deployment: optimizing density, scalability and cost-effectiveness.

A SEQUANA cell is organized across three cabinets: two cabinets containing the compute nodes and the central cabinet which houses the interconnect switches.

Each cell may contain up to 288 compute nodes equipped either with CPUs (such as Intel® Xeon® EP processors) or HPC accelerators (e.g. Intel® Xeon Phi™ or NVIDIA® GPUs).

The interconnect components located in the central cabinet form the first two levels of a fat-tree. External nodes (such as I/O nodes and service nodes) plug directly into the system fabric at the cell level

All SEQUANA components are hot-swappable and can be serviced without interrupting production. The SEQUANA checkpoint/restart service enhances resilience.

The SEQUANA components are cooled using an enhanced version of the Bull Direct Liquid Cooling (DLC) solution. DLC is a proven technology currently used in the bullx DLC B700 racks. DLC minimizes the global energy consumption of a system by using warm water up to 40°C. The SEQUANA DLC solution is sized to evacuate all the heat generated by compute nodes and BXI switching components, even in the most extreme configurations.

## Software comes into play

### **bullx supercomputer suite: the operating system for exascale**

Part of the exascale program, the bullx supercomputer suite addresses a broad range of needs, from supercomputer management to data management, application and performance management. It is equally effective for small-scale supercomputers as it is on massive systems with tens of thousands of cores, where ultra-high availability, reliability, automated event and health management, an optimized energy footprint, very high data throughput, high performance and a fine-tuned MPI library are all essential features.

The suite is based on best-of-breed open source software and leading open standards enhanced with Bull's added value features.

Bullx supercomputer suite components are integrated through common access to the Super Computer Information Repository, which provides a centralized and consistent view of the super computer and its application states for all components. This results in a dramatic increase in reliability, as well as delivering a whole new set of features.

### **Applications: more than a challenge**

The Bull Center for Excellence in Parallel Programming is a cornerstone of this program to support HPC users and ISVs. It helps to reengineer their applications and improve their application performance, and enables HPC users to make the most of new-generation supercomputers, including new processor technologies such as accelerators and graphics-based processors.

The Center is a key enabler when it comes to ensuring a successful transition to extreme computing and extreme data. Today, the main limitation lies within the code itself. Most applications cannot handle the new performance challenges brought about by advances in hardware technology, most notably the need to achieve better performance for less power consumption. So programmers are struggling to extract more parallelism, handle increasingly hybrid configurations and support greater heterogeneity.



## Getting rid of the communications overhead

### A new generation of interconnect is required

Exascale entails an explosion of performance, data volumes, power consumption and data movement. So it will be increasingly critical to ensure that CPUs are fully dedicated to computation. Today, with current interconnects, CPUs are also responsible for communications, at the expense of performance. Getting rid of this overhead would immediately and significantly free up CPU performance.

As a result, one of the cornerstones of Bull's exascale program is the development of a new-generation interconnect.

### The Bull exascale Interconnect unleashes CPU performance

Bull is developing a new-generation interconnect, the Bull exascale Interconnect: code-named BXI. BXI represents a paradigm shift when it comes to boosting performance, scalability, efficiency, reliability and quality of service for extreme workloads.

The core feature of BXI is a full hardware-encoded communication management system, which enables CPUs to be dedicated to computational tasks while communications are independently managed by BXI.

BXI hardware primitives map directly to communication libraries such as MPI (Message Passing Interface) and PGAS (Partitioned Global Address Space). Thanks to this hardware acceleration, BXI will deliver the highest level of communication performance for HPC applications, at full scale, characterized by high bandwidth, low latency and high message rates.

The BXI architecture is based on the Portals 4 communication library. This enables full optimization for all MPI communication types, including the latest MPI-2 and MPI-3 extensions and PGAS. The Portals 4 non-connected protocol guarantees a minimum constant memory footprint, irrespective of system size.

BXI quality of service (QoS) enables the definition of several virtual networks and will ensure, for example, that bulky I/O messages do not impede small data message flow. In addition, BXI adaptive routing capabilities will dynamically avoid communication bottlenecks.

End-to-end error checking and link level retry have been implemented to enhance communication reliability and resilience without jeopardizing communication performance.



## It's time to merge HPC and Big Data

Bull is committed to a full-scale move towards Big Data for enterprise IT and HPC, covering security, Fast Data (in-memory databases and containers), data management, search and analytics, and event processing. The Bull exascale program derives immediate benefits from the Bull's Big Data program.

### Intensive data processing for search and analytics

Thanks to its data-oriented conception, SEQUANA implements a tight coupling of data and simulation, and makes it possible to compute, explore and analyze massive multi-disciplinary datasets, thanks to intensive data processing and high throughput event processing. SEQUANA is designed to unlock a new generation of complex simulations and analyses.

SEQUANA benefits from bullx data management offering for HPC, as well as from Bull's project experience and expertise in a variety of storage management, storage configurations and data management scenarios based on various parallel file systems, Lustre and GPFS appliances, and HSMs.

SEQUANA is the foundation for High-Performance Data Analytics, and will leverage the upcoming Bull Hadoop-based search and analytics appliances and discovery appliances.

### In-memory data processing: no longer an option for many applications

Some applications such as some genomics applications require in-memory data processing. Bull has invested in this technology and is addressing this challenge today with a new range of ultra-high memory capacity servers: the bullx S6000 series.

In-memory computation opens up new opportunities for several market sectors, but is very hardware intensive. In-memory applications require very powerful systems with a huge memory to handle massive data volumes of tens of TBs. Beyond that, the server must guarantee a perfect quality of service and security.

The bullx S6000 series has been designed specifically for that purpose, with unique features to meet in-memory requirements and to converge business-critical computing and HPC. It is designed to support massive in-memory databases, pre-processing, post-processing and visualization operations.

The bullx S6000 series relies on two major innovations: a new generation of Bull BCS (Bull Coherence Switch) to inter-connect up to 16 CPUs, and the connecting box to assemble modules together with no apparent cabling, enabling compute resources to be easily added. This helps to adjust the memory and compute resources to the application needs and facilitates the scaling of Big Data applications. Another distinctive feature is

*"The next generation radio telescope, the Square Kilometre Array (SKA), would be generating just under 400 exabytes per year. This would appear to make it the world's largest single data generator – but medical imaging, social data, or the internet of things could well be larger by 2022."*  
Gary M. Johnson

the 100% reliability of its memory. In-memory applications require non-stop service and high memory reliability. The bullx S6000 series features RAS functions, full memory protection, and memory and I/O hot swap functions.

The bullx S6130, the first available model of the bullx 6000 series, is fully scalable up to 16 CPUs/24 TB. It supports full memory protection and hot swappable memory and I/O capabilities. The bullx S6130 is the HPC version of the bullion S16 enterprise server which is the fastest x86 server on the market (SPECint\_rate2006 for the 16 socket configuration).

## Long-standing experience in the HPC field

Bull has been in the HPC business for more than a decade, with major contracts in Europe and worldwide and significant investments in R&D. Today, Bull has long-standing experience in supporting customers, from design to delivery and service operations. This makes Bull a unique partner to understand the new challenges.

## Milestones

The Bull exascale program will be rolled out starting now. The bullx S6130 will be available for shipment starting Q2 2015.

Bull will disclose functionalities and will demo BXI and SEQUANA along 2015 with a general availability of the first version planned by 2016.

## Glossary

SOD: Statement of Direction

HPC: High-Performance Computing (close to scientific computing or intensive computing or technical computing).

Exascale: domain of exaflops (see exaflops)

Exaflops: one billion billion operations (i.e. floating operations) per second

Petaflops: one million billion operations (i.e. floating operations) per second

Zettabytes: one thousand billion billion bytes

## About Bull, an Atos company

Bull is the trusted partner for enterprise data. The Group, which is firmly established in the Cloud and in Big Data, integrates and manages high-performance systems and end-to-end security solutions. Bull's offerings enable its customers to process all the data at their disposal, creating new types of demand. Bull converts data into value for organisations in a completely secure manner.

Bull currently employs around 9,200 people across more than 50 countries, with over 700 staff totally focused on R&D. In 2013, Bull recorded revenues of €1.3 billion.

## About Atos

Atos SE (Societas Europaea) is a leader in digital services with 2013 pro forma annual revenue of €10 billion and 86,000 employees in 66 countries. Serving a global client base, the Group provides Consulting & Systems Integration services, Managed Services & BPO, Cloud operations, Big Data & Security solutions, as well as transactional services through Worldline, the European leader in the payments and transactional services industry. With its deep technology expertise and industry knowledge, the Group works with clients across different business sectors: Defence, Health, Manufacturing, Media & Utilities, Public Sector, Retail, Telecommunications and Transportation.

Atos is focused on business technology that powers progress and helps organizations to create their firm of the future. The Group is the Worldwide Information Technology Partner for the Olympic & Paralympic Games and is listed on the Euronext Paris market. Atos operates under the brands Atos, Atos Consulting, Atos Worldgrid, Bull, Canopy, and Worldline.

For more information, visit: [www.atos.net](http://www.atos.net)

