



With the BullSequana X1000 range of supercomputers, Atos demonstrates its strategic commitment to the development of innovative high performance computing systems - the systems needed to meet the major challenges of the 21st century.

Designed by the Atos R&D in close cooperation with major customers, the BullSequana X1000 supercomputer leverages the latest technological advances, so as to guarantee maximum performance for a minimized operation cost. The race to exascale calls for technological breakthroughs. With BullSequana, Atos delivers an innovative solution that matches the exascale technological challenges.

## Open for future technologies

BullSequana X1000 is highly modular, flexible and scalable. It is designed to support several types and several generations of compute nodes, equipped either with conventional processors or with accelerators. The following types of compute nodes are available or currently planned for future availability

- Intel Xeon Phi (Knights Landing)
- Intel Xeon processors scalable family (Skylake SP & Next Gen)
- Nvidia Pascal GPUs + Intel Xeon processors scalable family (Skylake SP & Next Gen)
- ARM processors Cavium ThunderX2

Sequana X1000 supports fast state-of-the-art interconnects, such as:

- Bull eXascale Interconnect (BXI),
- InfiniBand EDR.

## The BullSequana cell

In BullSequana 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 dual node for distributed management and diskless support.

This packaging consisting of large building blocks facilitates high-scale deployment - up to 64.000 nodes, and optimizes density, scalability and cost-effectiveness.

The BullSequana X1000 cell is organized across three cabinets.

- the two side cabinets contain compute nodes,
- the central cabinet houses the interconnect switches.

## Limiting energy consumption

Controlling energy consumption is the main roadblock on the path to exascale. Sequana is ultra energy-efficient: it targets a PUE very close to 1.

100% of the components of BullSequana - both compute nodes and switches - are cooled using an enhanced version of the Bull patented Direct Liquid Cooling (DLC) solution. DLC is a proven cooling technology that minimizes the global energy consumption of a system by using warm water up to 40°C.

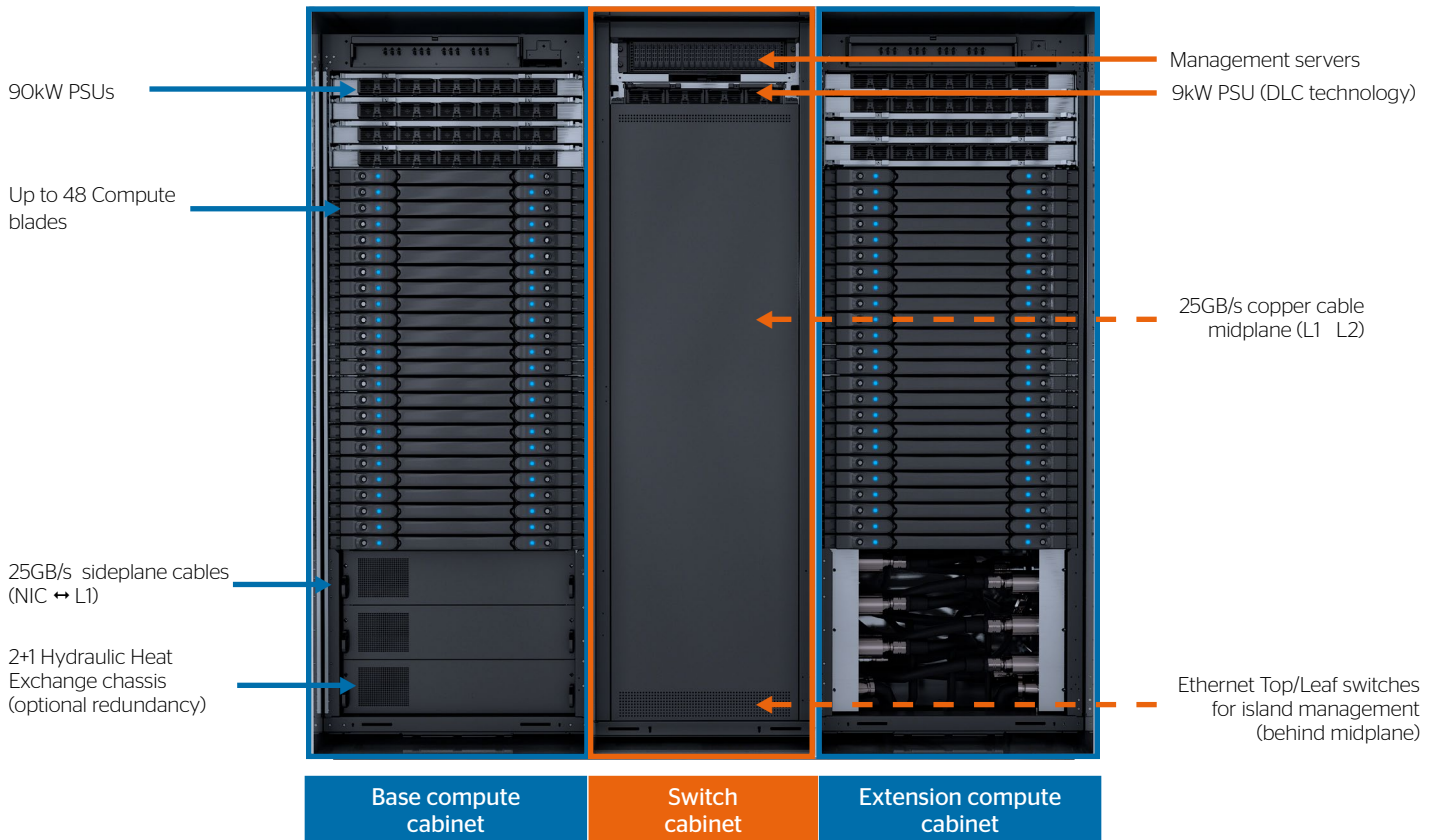
The cooling system of BullSequana X1000 is sized to evacuate the heat generated even by the most extreme configurations - with very powerful and energy-hungry nodes.

## Unchanged serviceability

For easy maintenance, all BullSequana X1000 components (compute nodes, interconnect switches, PSUs, heat exchangers...) are hot-swappable and can be serviced without interrupting system production. In the compute blades, the processors and DIMMs can be replaced without removing the cold plate.

In a BullSequana X1000 cell, all maintenance operations are just as easy as in a standard air-cooled system.

# The BullSequana X1000 cell



There are two types of compute cabinets:

- The base compute cabinet provides liquid cooling to the compute blades it houses and also to the switch cabinet.
- The extension compute cabinet is optional and has its own cooling sub-system.

Three types of interfaces connect the compute and switch cabinets together:

- Liquid interface with two flexible pipes from Base compute cabinet hydraulic chassis to switch cabinet manifolds.
- Interconnect interface: each node is connected through one or two interconnect links, one per NIC.
- Management interface for all management controllers in the compute cabinet (PMC, BMC, HMC): one 1Gb/s Ethernet link per node plus 4

sidebands signals for platform management.

The compute nodes are connected to the switch cabinet (L1 switches) via 12 copper cable octopuses named:

- Sideplane Octopus Double for 2 NICs per node.
- Sideplane Octopus Single for 1 NIC per node.

## Compute cabinet

Each BullSequana X1000 compute cabinet houses up to 48 horizontal compute blades, with the associated power modules at the top of the cabinet and the redundant hydraulic modules for cooling at the bottom of the cabinet.

24 blades are mounted on the front side of the cabinet, while the 24 other blades are mounted on the rear side.

Each 1U blade contains three compute nodes, except for GPU-based blades which contain one node.

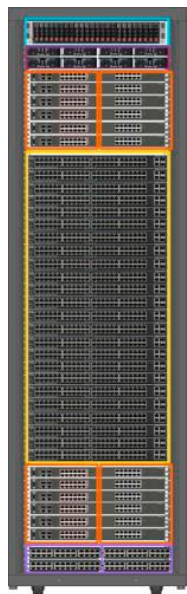
Each cell can therefore contain up to:

- 288 Intel nodes (single-socket Xeon Phi OR dual-socket Xeon)
- 96 dual-socket Intel Xeon nodes with 384 Nvidia Pascal GPUs
- 288 ARM nodes (dual-socket Cavium)

<b>Dimensions /Weight</b>	Mounting Capacity: 42U - Size (HxWxD) without UCM: 2020x750x1270 mm - With UCM (max): 2400x750x1270 mm Weight without compute blade: 1000 kg - Packing weight (without compute blade, with pallet):1110 kg Total weight of cabinet on slab floor (fully equipped): 2200 Kg ( 2300Kg with UCM)
<b>Compute Power Group</b>	75 U form factor at the top of the cabinet <ul style="list-style-type: none"> <li>▶ Up to 4 x 15kW air cooling PSU shelves at the front of the cabinet with at PSU level (N+X)</li> <li>▶ Up to 4 x 15kW air cooling PSU shelves at the rear of the cabinet with at PSU level (N+X)</li> <li>▶ 1 Power Management Controller Module</li> <li>▶ 2 x2 64A tri-phased 400V AC EU input power lines or 2 x4) 64A tri-phased 600V AC US input power lines</li> <li>▶ Option: 2 Ultra Capacity Module chassis supporting power outage of 300ms at full load or 800ms at 45% load in case of a 2 or 3 phases AC power outage</li> </ul> Ready to support liquid cooling PSU shelf in Q3 2017 (FCS TBC).

<b>Cooling sub-system</b>	3 Hydraulic chassis (HYCs) at the bottom of the cabinet The primary manifold system connects the customer loop to HYCs primary water inlets The secondary manifold system connects HYCs outlets to each blade in the compute cabinet The manifolds of Base compute cabinet provide input / output pipes for the liquid cooling of the switch cabinet
<b>Interconnect between compute &amp; switch cabinet</b>	Connection between compute node and L1 switch via 12 copper cable octopuses Interconnect interface: each node is connected with one or two interconnects links, one per NIC Management interface for all management controllers in compute cabinet (PMC, BMC, HMC): one 1Gb/s Ethernet link per node plus 4 sidebands signals for platform management.
<b>Regulatory compliance</b>	<b>Safety</b> (EC, IEC, UL, CSA certifications); <b>Electromagnetic Compatibility</b> (EC, FCC, ICES-03, VCCI certifications), <b>Environment</b> (RoHS II & WEEE directives, REACH regulation)
<b>Warranty</b>	Standard warranty: 1 year - Extended Warranty: consult us

## Switch cabinet



**BXI configuration with 2 NICs**

- ← ISMA
- ← Switch power group
- ← 12 L2 DLC switches
- ← 24 L1 DLC switches
- ← 12 L2 DLC switches
- ← Ethernet L1 switch
- ← Ethernet L2 switch

The central BullSequana X1000 switch cabinet houses:

- ▶ two levels of interconnect switches in full fat-tree topology,
- ▶ their associated DC power generator (SWPG: SWitch Power Group) and
- ▶ management components such as Ethernet switches and one redundant server for Island management and administration (ISMA).

The switch cabinet provides internal connections (within the cell) through copper cables and external connections through optical cables.

288 external ports are available to interconnect each cell with the rest of the system to form different topologies, such as Fat-tree, All-to-all or other directly connected configurations.

External nodes (such as for example input/output gateways, services nodes...) can be plugged into the L1 switches replacing compute nodes.



**InfiniBand EDR or configuration with 1 NIC**

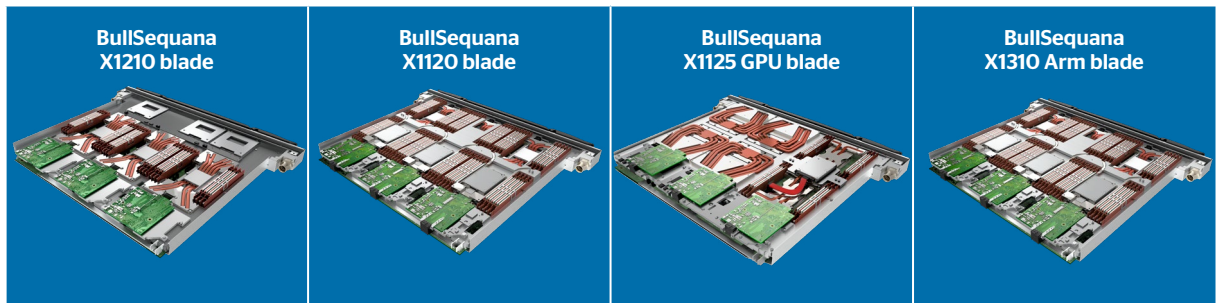
<b>Dimensions /Weight</b>	Mounting capacity: 42U - Size (HxWxD) without UCM: 2020x600x1270 mm - With UCM (max): -2400x600x1270mm Weight on castors w/ SW & W/o UCM: up to 745 kg - Packing weight (with pallet) w/ SW & W/o UCM: up to 800 kg Total weight of cabinet on slab floor (fully equipped & W/ UCM): up to 810 kg
<b>L1 switches</b>	<ul style="list-style-type: none"> <li>▶ Up to <b>24 BXI Direct Liquid Cooled switches</b> (48 ports module in 1U format): <ul style="list-style-type: none"> <li>– without optical ports and with a 1 Gb/s/s Ethernet switch</li> <li>– or with 3 optical ports and with a 1 Gb/s Ethernet switch</li> </ul> </li> <li>▶ OR up to <b>12 InfiniBand EDR 36 ports Direct Liquid Cooled modules</b> in 1U format: <ul style="list-style-type: none"> <li>– without optical ports and with a 1 Gb/s/s Ethernet switch</li> <li>– or with 3 optical ports and with a 1 Gb/s Ethernet switch</li> </ul> </li> </ul>
<b>L2 switches</b>	<ul style="list-style-type: none"> <li>▶ Up to 12 1U L2 Direct Liquid Cooled switch Modules with two 48 ports <b>BXI</b> L2 switches per module</li> <li>▶ OR up to 12 1U L2 Direct Liquid Cooled switch Modules with one 36 ports <b>IB EDR</b> L2 switch per module</li> </ul>
<b>Interconnect between L1 &amp; L2 switch</b>	<ul style="list-style-type: none"> <li>▶ Backplane Octopus Double (BOD) providing all-to-all connections between 24 L1 switch modules and 12 L2 switch modules (24 L2 switches) for configurations with 2 NIC per compute node</li> <li>▶ OR Backplane Octopus Single (BOS) providing all-to-all connections between 12 L1 switch modules and 6 L2 switch modules (12 L2 switches) for configurations with 1 NIC per compute node</li> </ul> <p>The management backplane implements the Ethernet and sideband management connections between L1/L2/Leaf/Top switch modules with:</p> <ul style="list-style-type: none"> <li>▶ A Backplane Octopus Management Base (BOMB) that connects left L2 switches</li> <li>▶ A Backplane Octopus Management Extension (BOME) that connects right L2 switches</li> </ul>
<b>Island Monitoring and Administration</b>	<ul style="list-style-type: none"> <li>▶ Up to 2 x 24 ports 1Gb/s Ethernet Leaf switch Management Modules in 1U format</li> <li>▶ Up to 2 x Ethernet Top switch Management Modules in 1U format.</li> </ul> <p>Each switch provides 24 downlink 1Gb/s Ethernet ports and 4 uplink 10 Gb/s Ethernet ports.</p> <ul style="list-style-type: none"> <li>▶ 2U Island Monitoring and Administration (ISMA) built with 2 redundant servers</li> </ul>
<b>Switch Power Group</b>	<ul style="list-style-type: none"> <li>▶ 2 air-cooled PSU shelves (1+1 redundancy) with 4 Power supply Units (PSU) of 3KW</li> <li>▶ 1 Power Distribution unit (PDU) with 2 x 32A tri-phased 230/400Vac input power lines</li> <li>▶ Power Management Controller Module</li> <li>▶ Option: Ultra Capacity Modules to support an AC power outage of up to 900ms at full load</li> </ul> <p>Ready to support liquid cooling PSU shelf in Q3 2017 (FCS TBC)</p>
<b>Regulatory compliance</b>	<b>Safety</b> (EC, IEC, UL, CSA certifications); <b>Electromagnetic Compatibility</b> (EC, FCC, ICES-03, VCCI certifications), <b>Environment</b> (RoHS II & WEEE directives, REACH regulation)
<b>Warranty</b>	Standard warranty: 1 year - Extended Warranty: consult us

# Compute blades

In each 1U blade, a cold plate with active liquid flow cools all hot components by direct contact – the BullSequana compute blades contain no fan.

The following compute blades are initially available:

- The 1U BullSequana X1110 blade integrates 3 compute nodes, each powered by 2 Intel® Xeon® E5-2600 V4 processors (code named Broadwell-EP).
- The 1U BullSequana X1120 blade integrates 3 compute nodes, each powered by 2 Intel® Xeon® Processor Scalable Family (Skylake SP).
- The 1U BullSequana X1125 blade includes a 2 Intel® Xeon® Processor Scalable Family (Skylake SP) compute node equipped with 4 NVIDIA® Tesla Pascal P100 16GB GPUs.



<b>Design</b>	1U blade comprising 3 compute nodes side-by-side	1U blade comprising 3 compute nodes side-by-side	1U blade with 1 accelerated compute node	1U blade comprising 3 compute nodes side-by-side
<b>Processors</b>	3 x 1 Intel® Xeon Phi™ KNL processor	3 x 2 Intel® Xeon® Processor Scalable Family (Skylake SP)	2 Intel® Xeon® processors Scalable Family <b>4 Nvidia® Pascal P100 16GB GPUs</b>	3 x 2 Cavium® ThunderX2™ Armv8 processors with up to 32 cores
<b>Architecture</b>	3 x 1 Intel® C610 chipset	3 x 1 Intel® C620 chipset	1 Intel® C620 chipset	3 x 1 motherboard compatible with Cavium Borg reference platform
<b>Memory</b>	3 x 6 DDR4 memory slots (max 192 GB with 32 GB DIMMs)	3 x 12 DDR4 memory slots (max 384 GB with 32 GB DIMMs) + 4 optional NVRAM DIMMS (NVRAM availability TBC)	12 DDR4 memory slots (max 384 GB with 32 GB DIMMs)	3 x 16 DDR4 memory slots (max 1024 GB with 64 GB DIMMs)
<b>I/O slots</b>	<ul style="list-style-type: none"> <li>▶ InfiniBand EDR 1 port mezzanine board</li> <li>▶ or BX1 1 or 2 ports mezzanine board</li> </ul>	<ul style="list-style-type: none"> <li>▶ InfiniBand EDR 1 port mezzanine board</li> <li>▶ or BX1 1 or 2 ports mezzanine board</li> </ul>	<ul style="list-style-type: none"> <li>▶ InfiniBand EDR 1 port mezzanine board</li> <li>▶ or BX1 1 port mezzanine board</li> </ul>	<ul style="list-style-type: none"> <li>▶ InfiniBand EDR 1 port mezzanine board</li> </ul>
<b>Storage</b>	3 x 1 optional SATA drive 3 x 1 optional NVMe SSD drive via PCIe switch	3 x 1 optional SATA drive 3 x 1 optional NVMePCIe SSD drive via PCIe switch	1 optional SATA SSD drive	1 optional SATA SSD drive
<b>Power supply</b>	In cabinet			
<b>Cooling</b>	Cooling by direct contact with DLC coldplate or through heat spreaders for DIMMS			
<b>Physical specifications</b>	44.45 x 600 x 540 mm (H x W x D) Weight: 25 kg max			
<b>OS and software</b>	Red Hat Enterprise Linux & Bull Supercomputer Suite support			
<b>Warranty</b>	Standard warranty: 1 year - Extended warranty: consult us			
<b>Regulatory compliance</b>	<b>Safety</b> (EC, IEC, UL, CSA certifications); <b>Electromagnetic Compatibility</b> (EC, FCC, ICES-O3, VCCI certifications), <b>Environment</b> (RoHS II & WEEE directives, REACH regulation)			

Contact: [hpc@atos.net](mailto:hpc@atos.net)

All trademarks are the property of their respective owners. Atos, the Atos logo, Atos Codex, Atos Consulting, Atos Worldgrid, Bull, Canopy, equensWorldline, Unify, Worldline and Zero Email are registered trademarks of the Atos group. Atos reserves the right to modify this document at any time without notice. Some offerings or parts of offerings described in this document may not be available locally. Please contact your local Atos office for information regarding the offerings available in your country. This document does not represent a contractual commitment.  
June 2018 © 2018 Atos