Extreme Computing

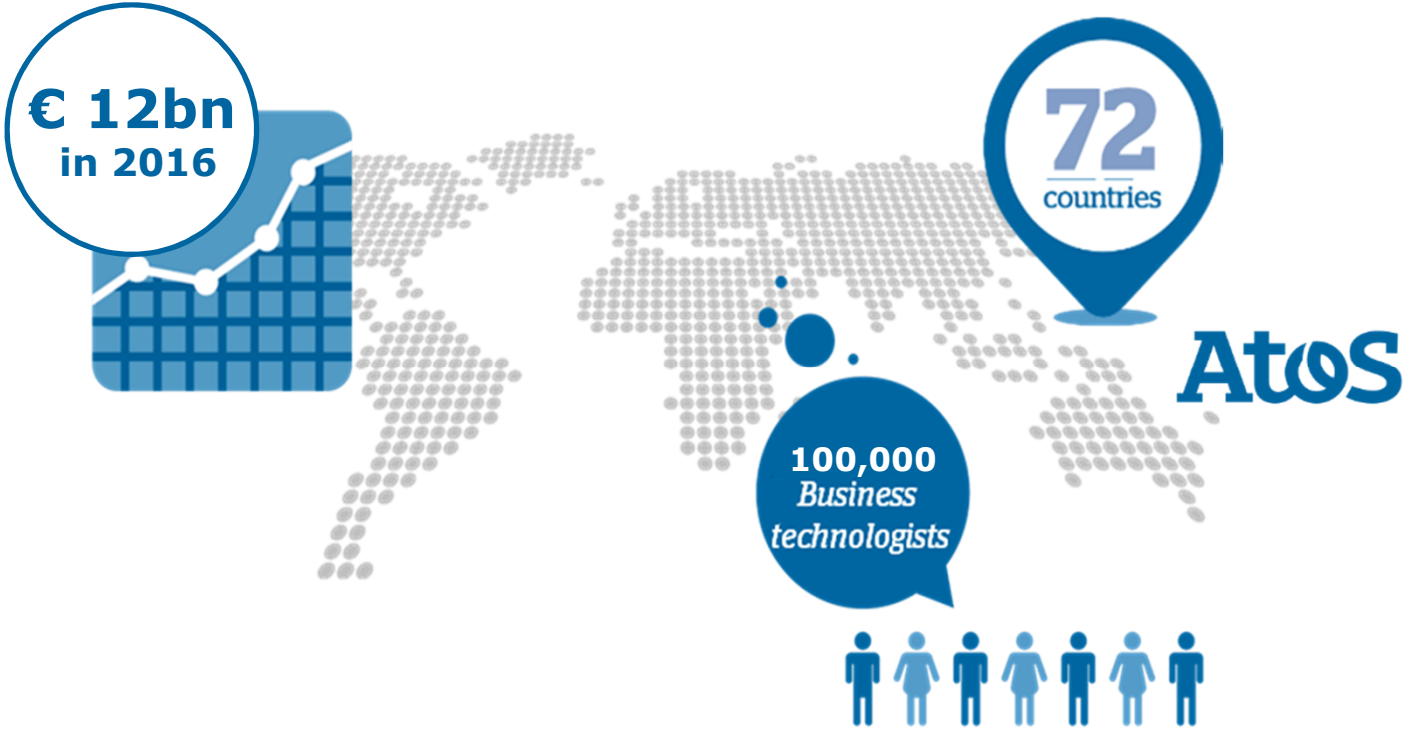
Strategic directions for the 2020s

04-07-2017

Bull
atos technologies

We are THE European IT Leader

and in the top 5 of worldwide Digital services players



worldline
e-payment services

Bull
atos technologies

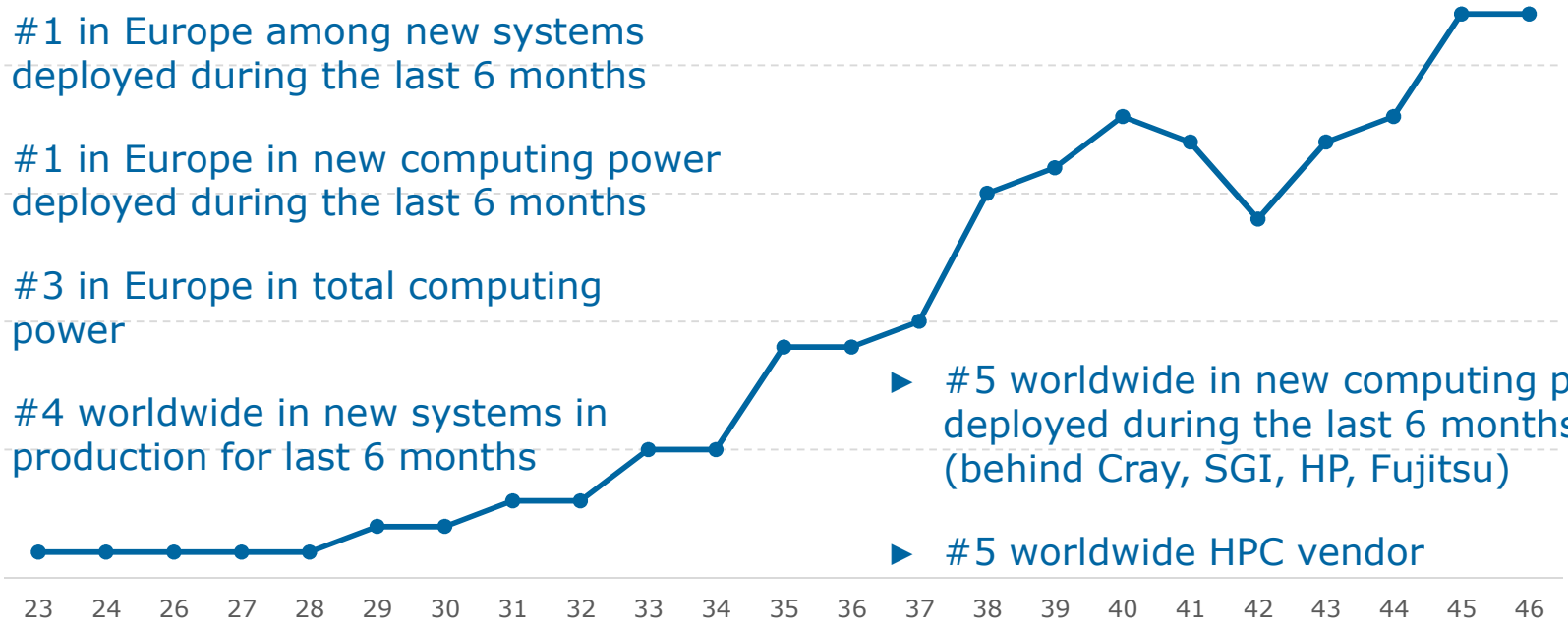
UNIFY

Bull
atos technologies

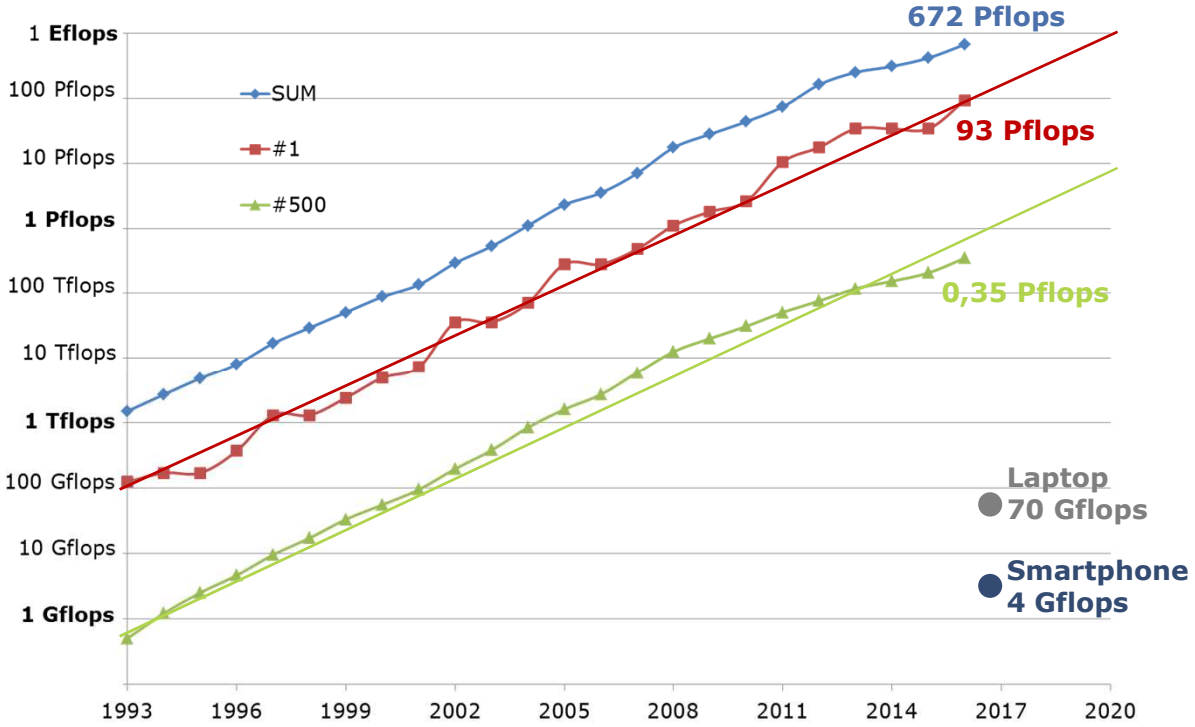
Top500 trend

For the last 10 years

- 22 systems in top500 of Nov. 2016
- #1 in Europe among new systems deployed during the last 6 months
- #1 in Europe in new computing power deployed during the last 6 months
- #3 in Europe in total computing power
- #4 worldwide in new systems in production for last 6 months
- #5 worldwide in new computing power deployed during the last 6 months (behind Cray, SGI, HP, Fujitsu)
- #5 worldwide HPC vendor



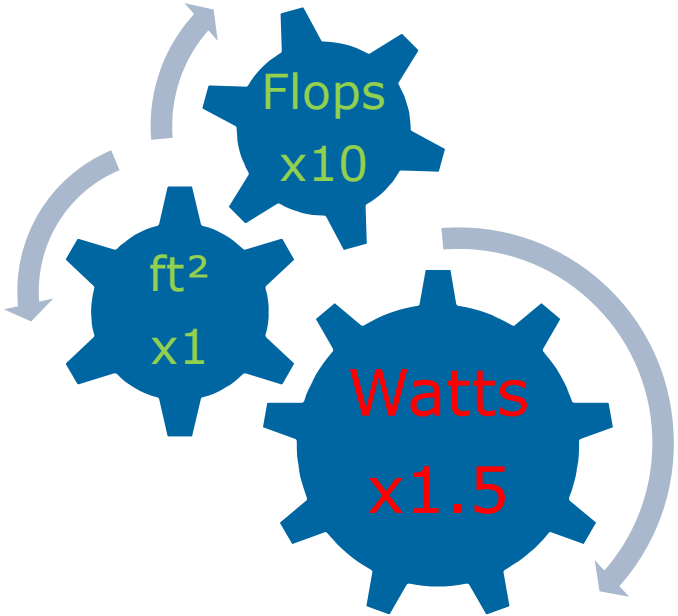
Are all lights green?

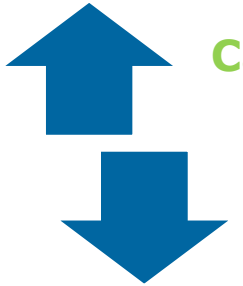


Exascale era still early 2020s

Bull atos technologies

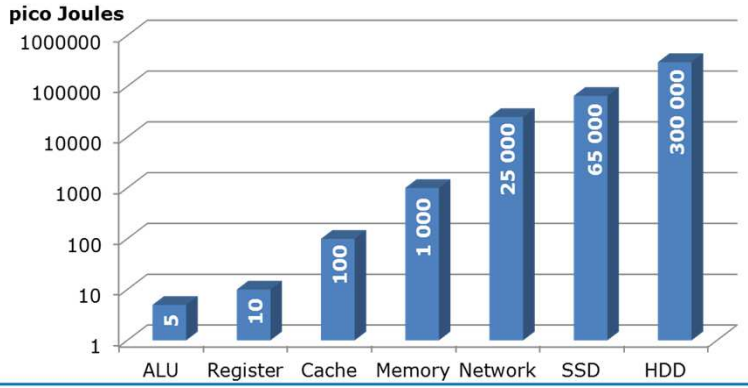
Some divergences





Capacity: x10

Throughput: x2

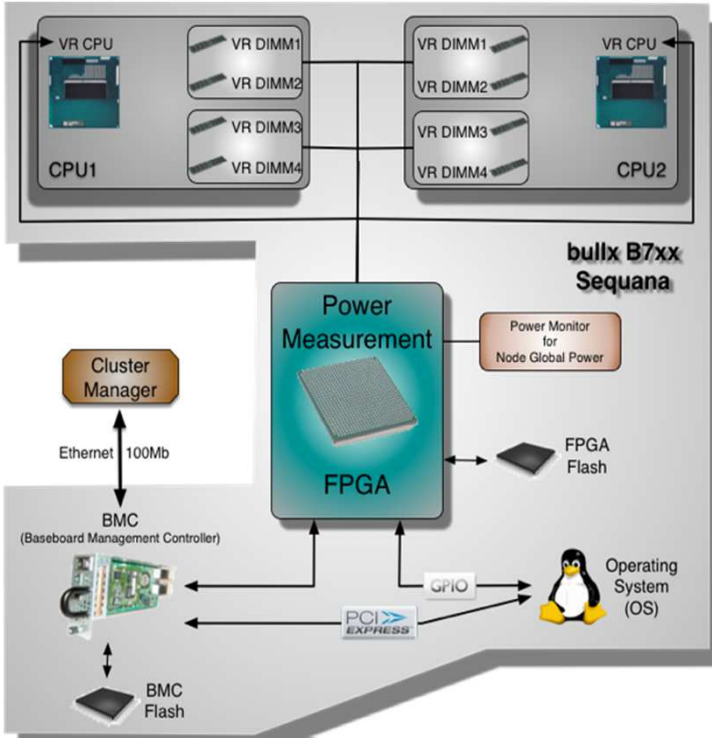


Ultra-efficient PUE

- Free cooling data centers
- 100% direct liquid cooling
- Up to 40°C for inlet water

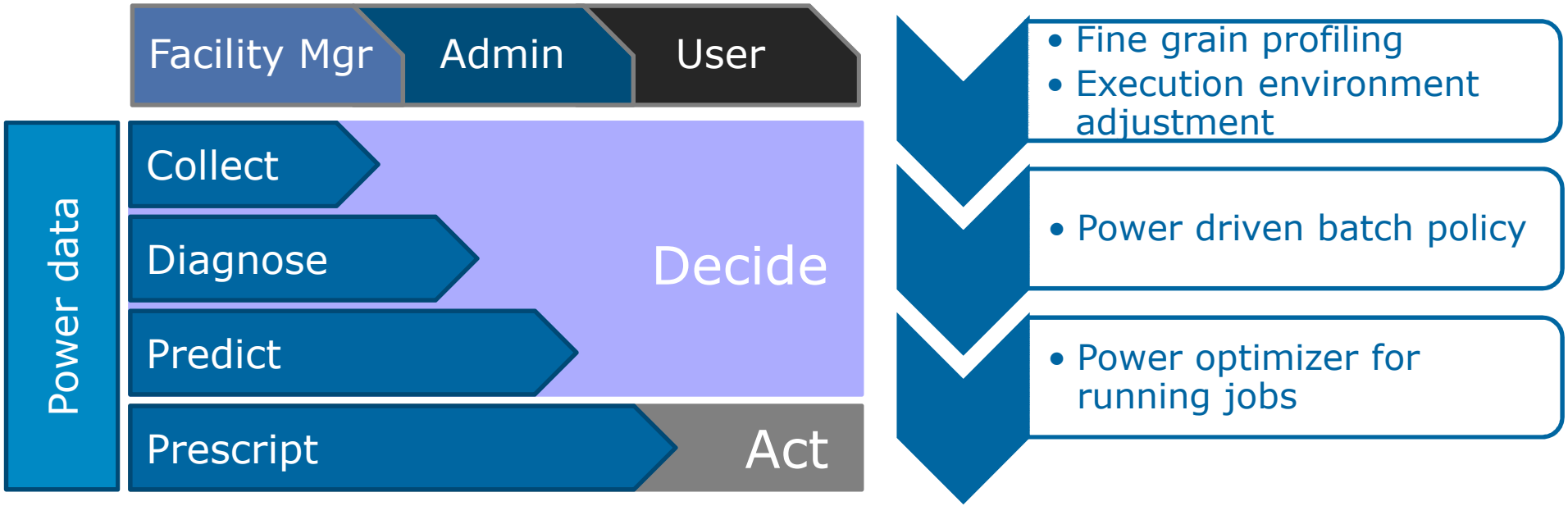
HDEEM: detailed power consumption

- 1,000 Hz Time stamped sampling
- CPUs
- DRAM
- Interconnect
- 8 hours ring buffer
- High accuracy



Bull
atos technologies

Prescriptive data analytics for power

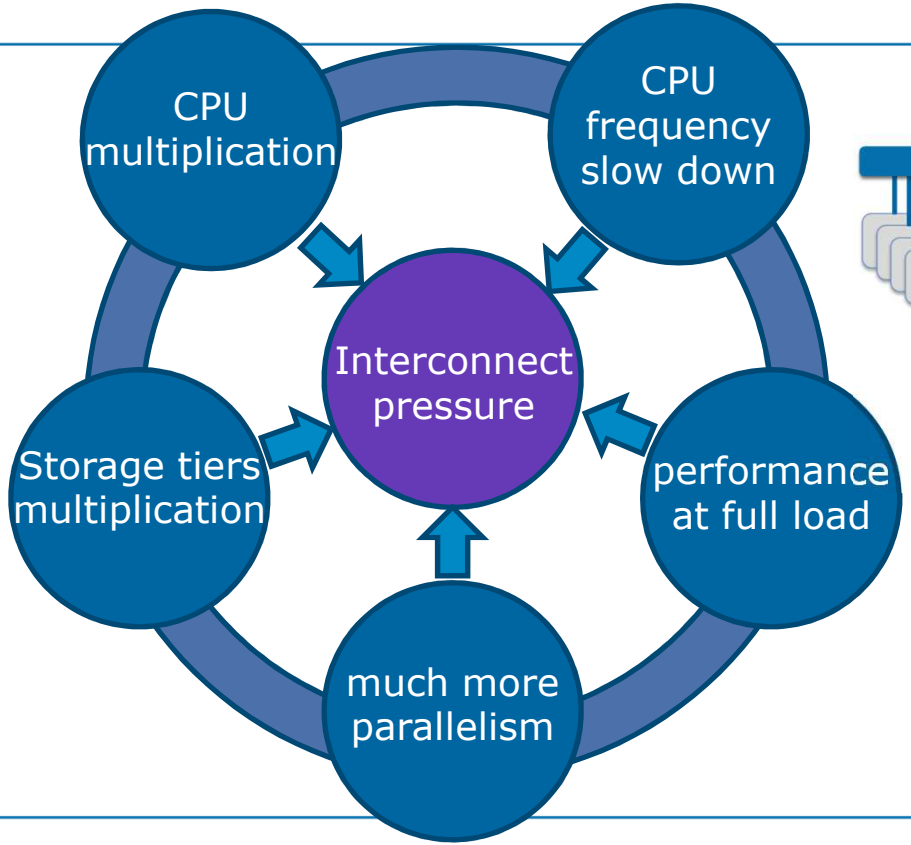


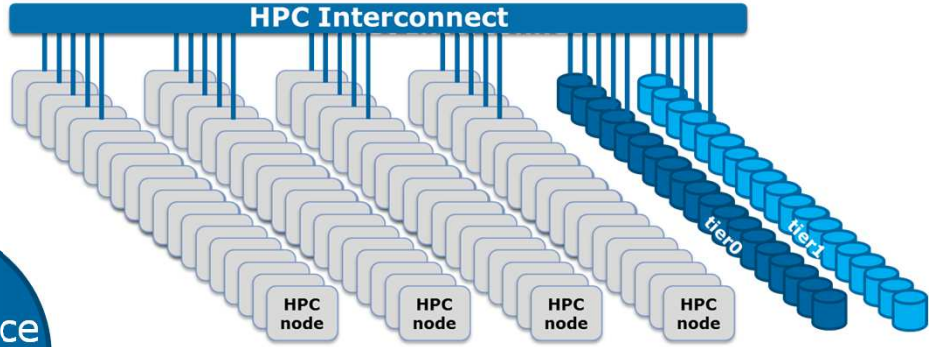
Make energy optimization happen

NEMO (**N**ucleus for **E**uropean **M**odelling of the **O**cean)

	Fixed frequency	Supervised frequency
Energy (J)	693 704	635 964
Execution time (s)	516	521
Energy saving (J)		**8.3%**
Execution penalty (s)		**1.0%**

Supercomputing trend

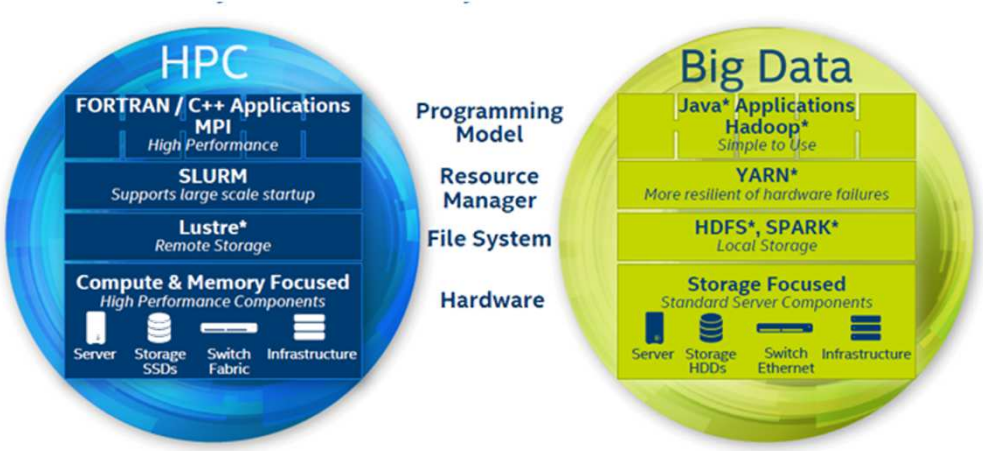


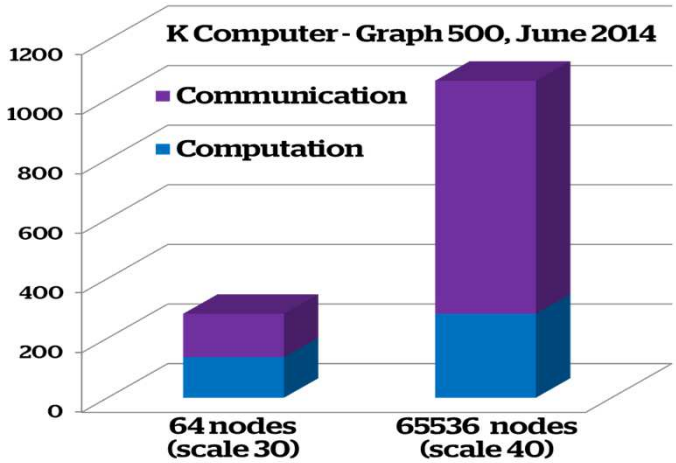


Bull atos technologies

HPDA: which gap do we have to close?

- Rationalize the infrastructure
- Algorithms will be boosted by data capacity and bandwidth, not Flops anymore





Bull atos technologies

Free up Supercomputer performances

BXI

Full acceleration in hardware for HPC applications

- HW coded Portals 4, a rich low level network API for message passing (MPI & PGAS)
- Ultra fast path inside the NIC for PGAS / MPI one-sided messaging
- HW offloaded collective operations
- Sustained performance under heavy load

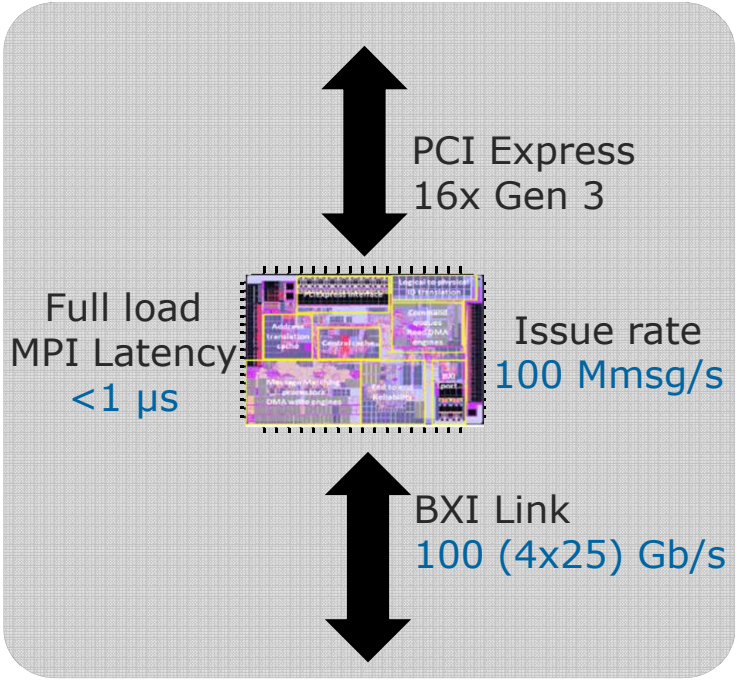
Highly scalable, efficient and reliable

- 64k nodes
- Small memory footprint
- Adaptive Routing
- Quality of Service
- End-to-end error checking + link level CRC + ASIC ECC

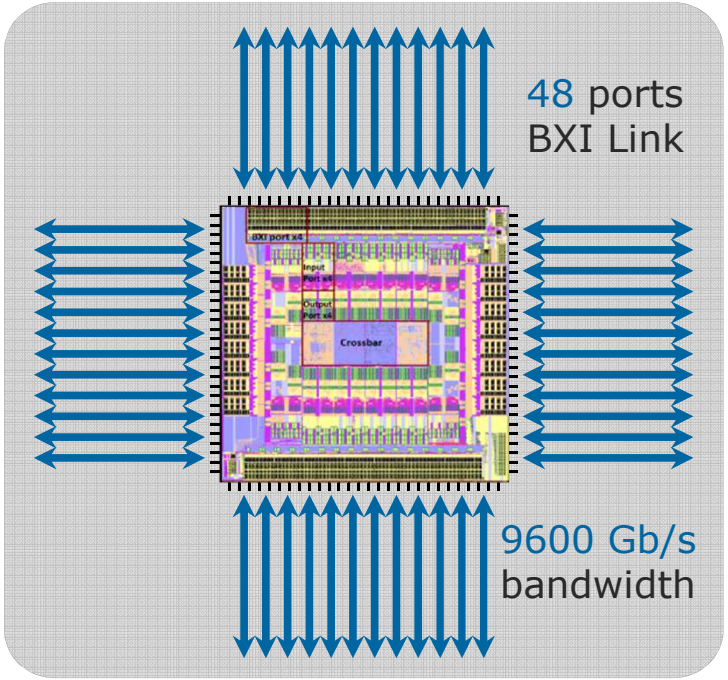
Designed with and for end users

- Co-design in collaboration with CEA

BXI Network is based on 2 in house ASICs



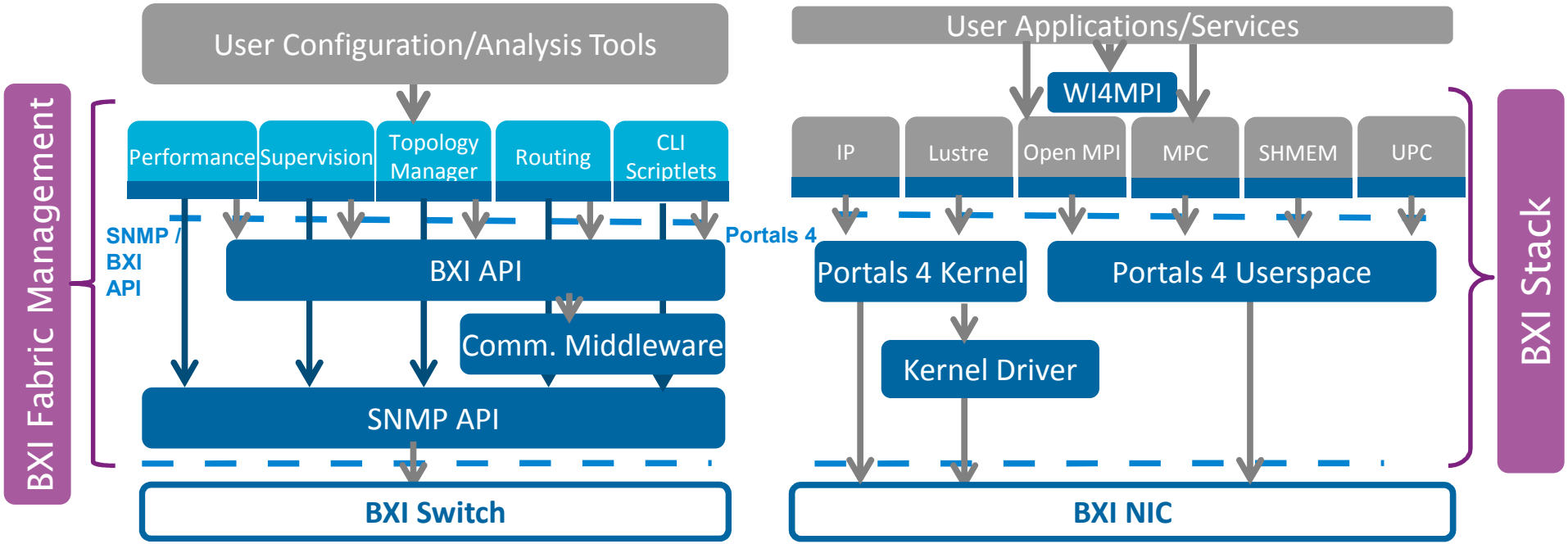
Lutetia



Divio

Bull
atos technologies

BXI Software Suite



Bull atos technologies

BXI:
offloading MPI communication in HW

address V2P

size

rank L2P

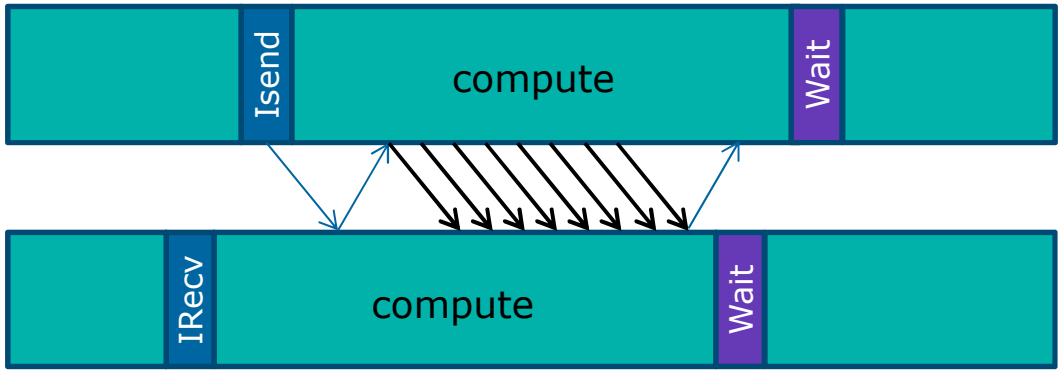
message order

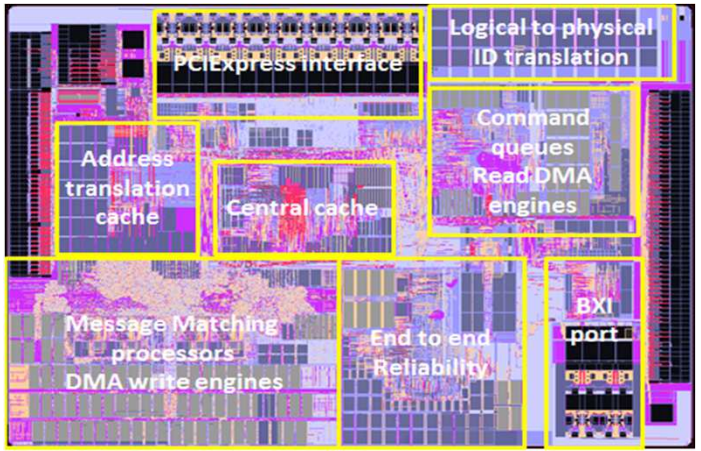
```
#include <mpi.h>

int MPI_Isend( const void *buf, int count, MPI_Datatype datatype, int dest,  int tag, MPI_Comm comm, MPI_Request *request)

int MPI_IRecv(void *buf, int count, MPI_Datatype datatype, int source, int tag, MPI_Comm comm, MPI_Request *request)

int MPI_Wait(MPI_Request *request, MPI_Status *status)
```

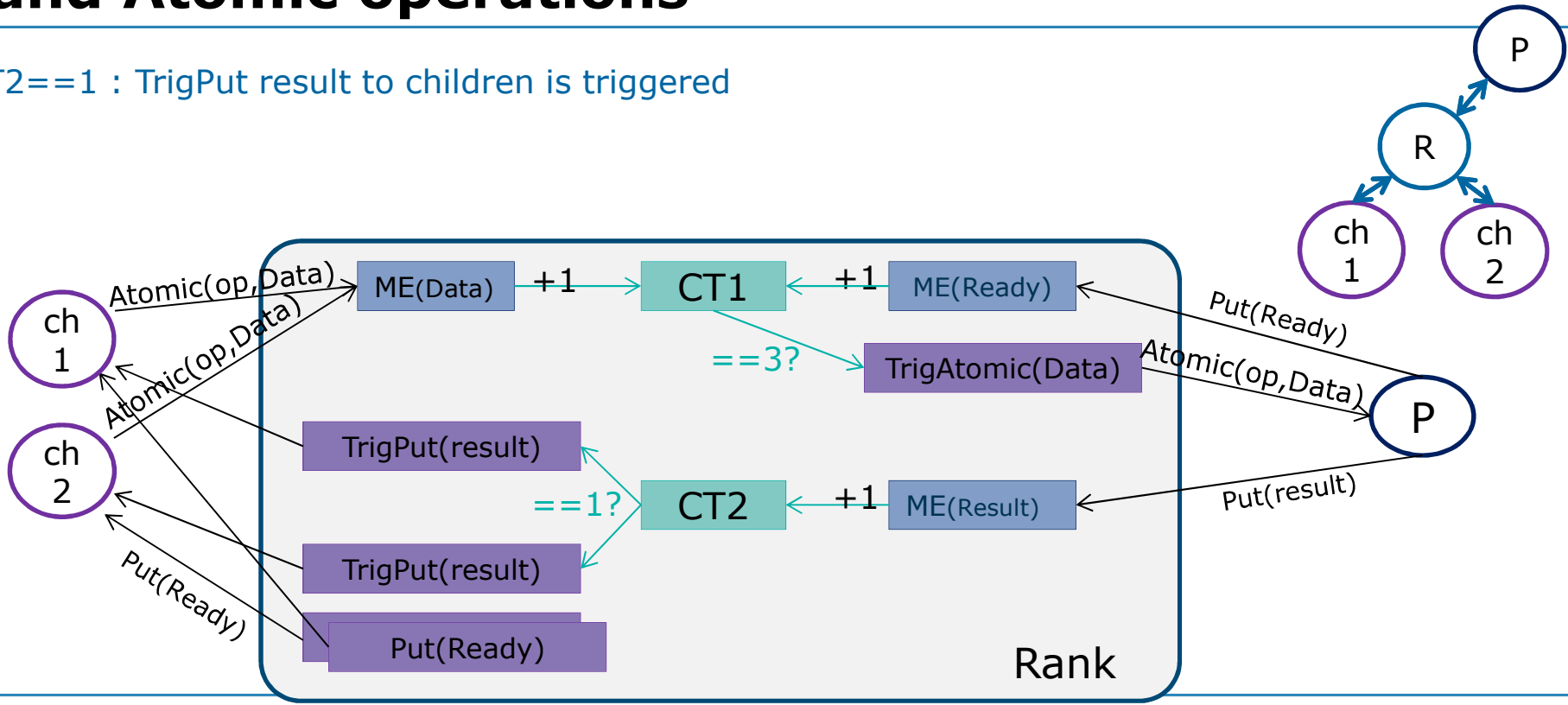




Bull atos technologies

Allreduce example using Triggered and Atomic operations

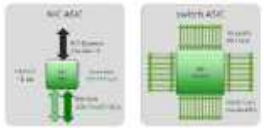
CT2==1 : TrigPut result to children is triggered



On Track for Exascale era

2016	2017	2018	2019	2020	2021	2022

ARM

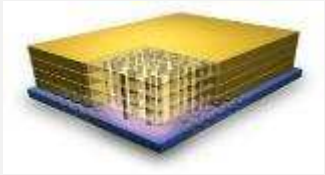


New Exascale BXI Interconnect

Sequana New flexible packaging

SCS

New integrated computing architecture



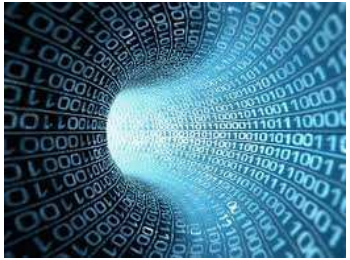
Compute + Memory + Network + Storage integration

Extreme scale packaging

SCS

Energy & performance oriented

10^{18} Exascale



Bull atos technologies

Thanks

For more information please contact:
T+ 33 4 76297270
F+ 33 4 76297607
M+33 6 80357914
eric.monchalin@atos.net

Bull
atos technologies