HPC2N @ Umeå University

Introduction to HPC2N and Kebnekaise

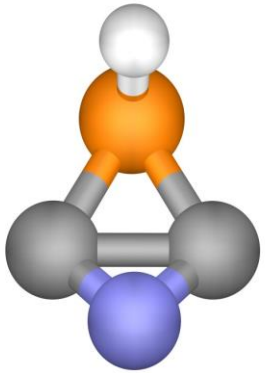
Jerry Eriksson, Pedro Ojeda-May, Birgitte Brydsö

HPC2N - “HPC to North"

- A national center for Parallel and Scientific Computing
- Five partners:
  - Luleå University of Technology
  - Mid Sweden University
  - Swedish Institute of Space Physics
  - Swedish University of Agricultural Sciences - SLU
  - Umeå University
- Funded by the **Swedish Research Council (VR)** and its Meta-Center **SNIC** togheter with the **partners.**

From macro scale to micro scale

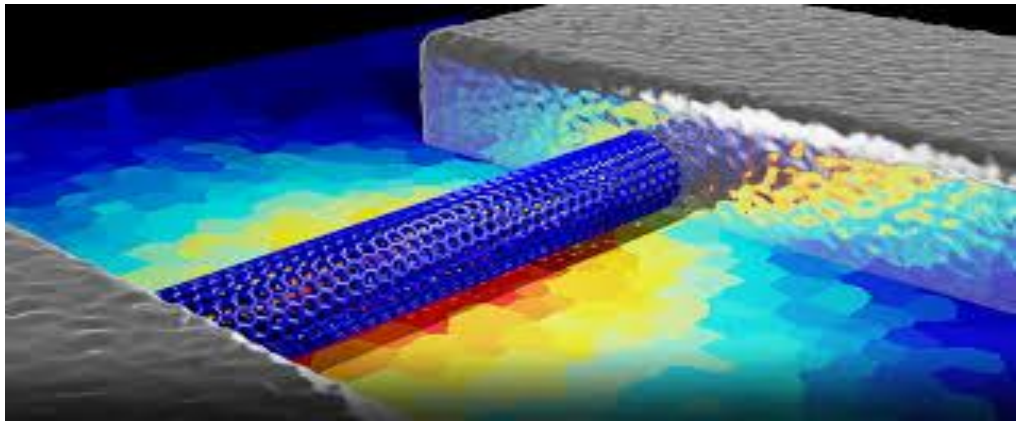
- Provides state-of-the-art resources and expertise for Swedish eScience
  - Scalable and parallel HPC
  - Large-scale storage facilities
  - Grid and cloud computing
  - Software and advanced support for eScience applications
  - International network for research and development

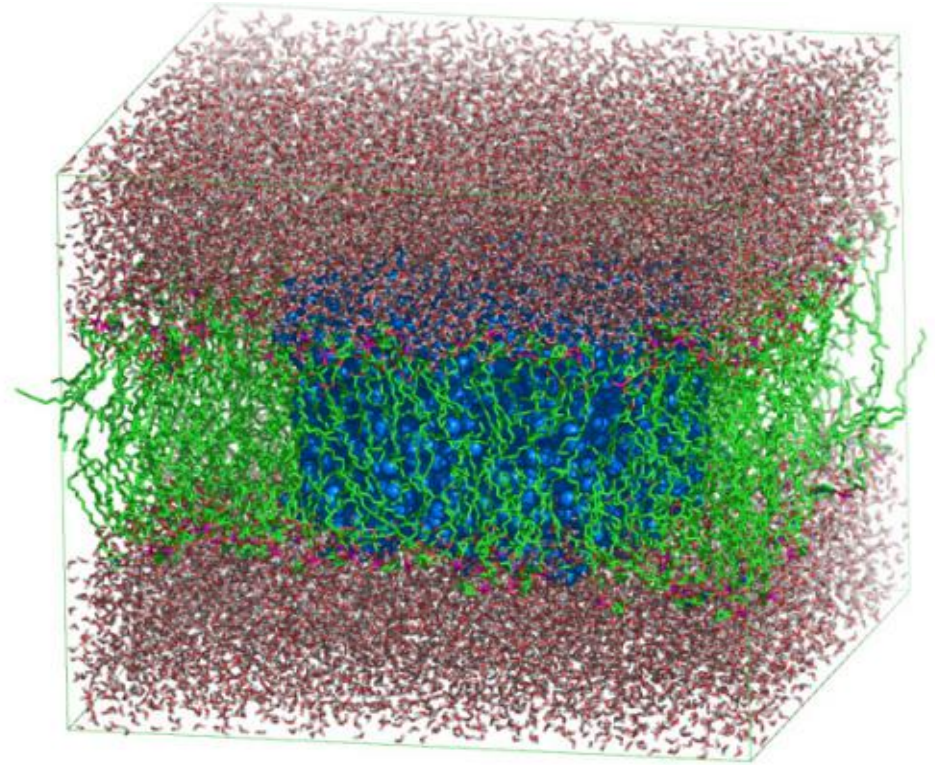


DFT computation, semi-stable, binding energy 15eV; Sven Öberg, LTU

Main areas of HPC2N users

- Biosciences and medicine
- Chemistry
- Computing science
- Engineering
- Materials science
- Mathematics and statistics
- Physics including space physics





Storage Levels @ HPC2N

Basically three types of storage are available at HPC2N:

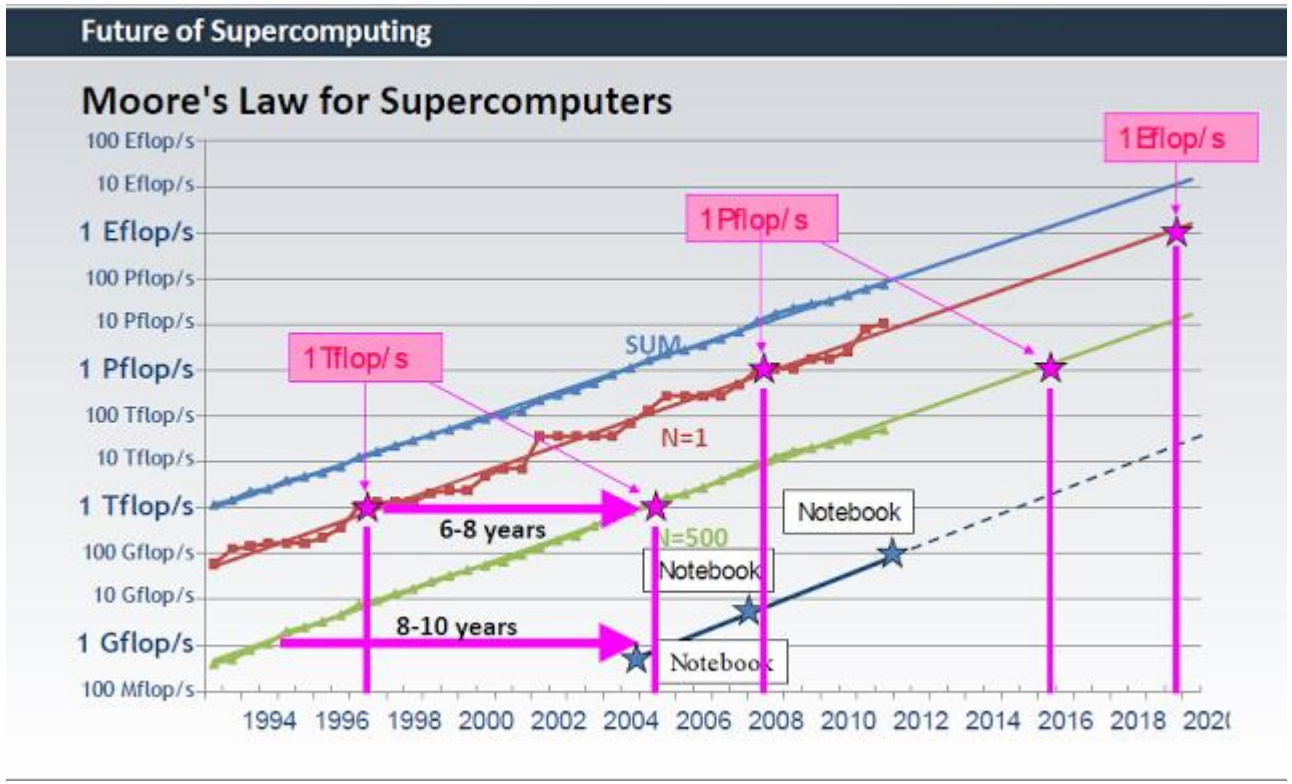
- Center Storage - Parallel file system (fast discs)
  - Closely connected to our computing resources; Abisko and Kebnekaise
- SweStore - disk based (dCache)
  - part of SNIC Storage, responsible for national accessible storage
- Tape Storage
  - Backup
  - Long term storage

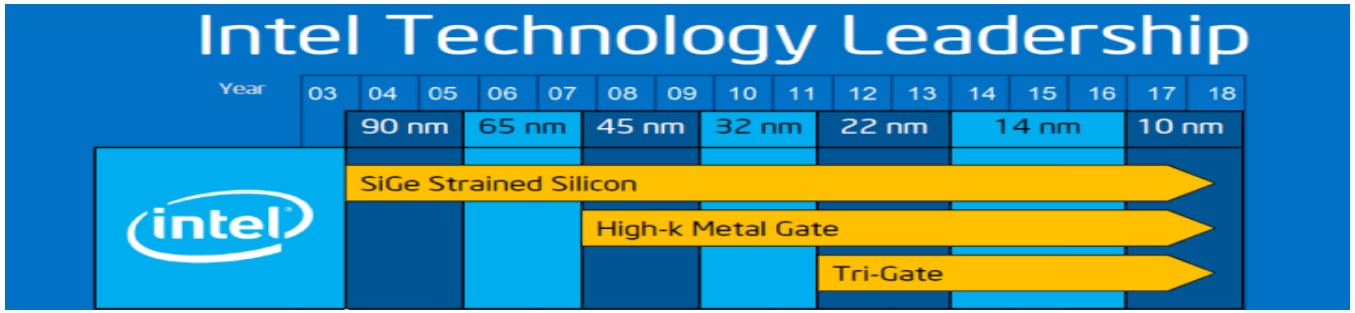
HPC2N Think Tank!

- User support (primary, advanced, tailored)
  - Research group meetings @ UmU
- User training and education program
- Workshops & Colloquia
- Research & Development - Technology transfer
- Provide various state-of-the-art HPC resources

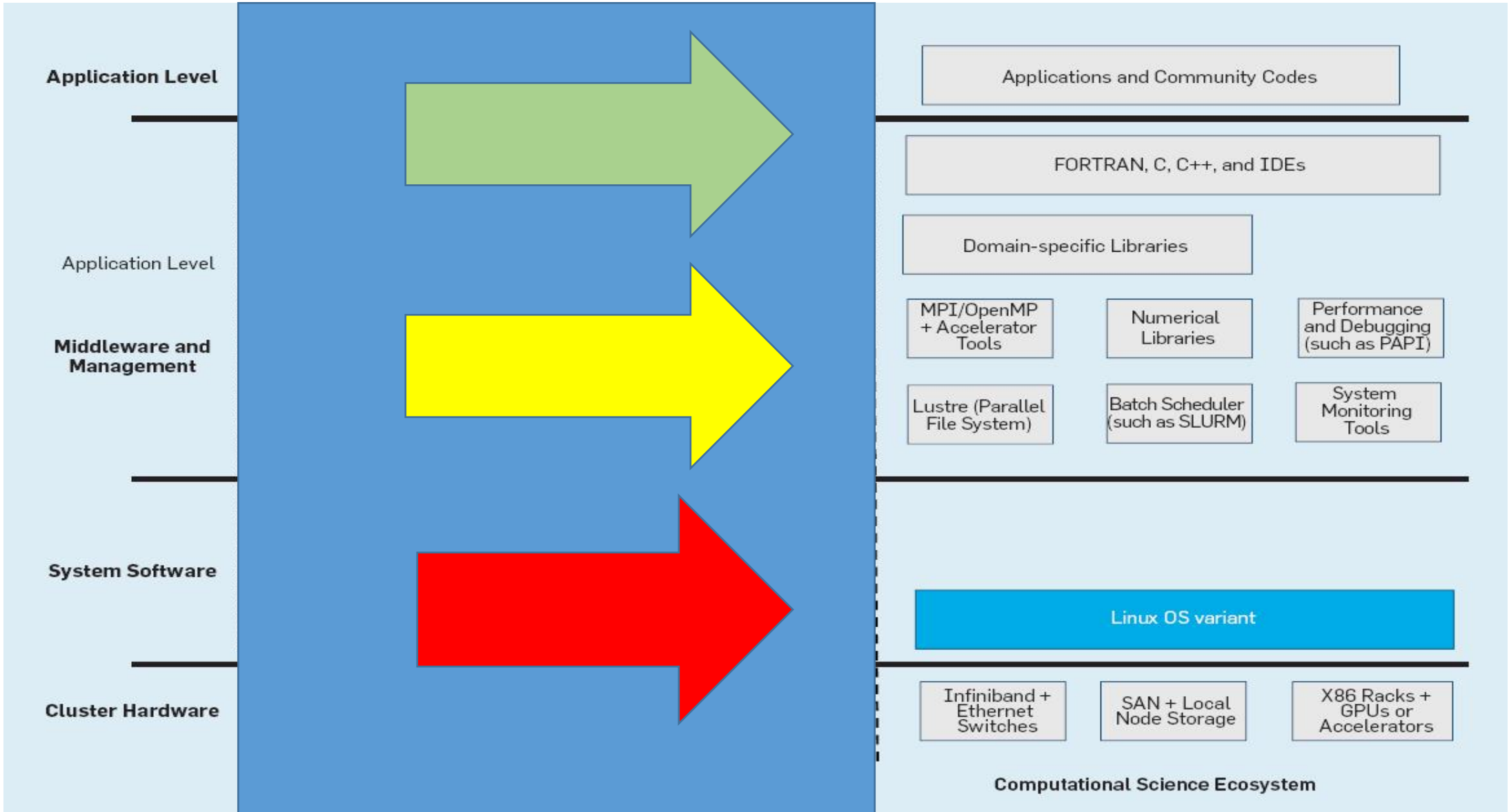
HPC – Towards Exascale Computing

- **Moore's law**: the number of transistors in a chip doubles every second years.
- Parallel Computing:
  - Increase number of cores.
- Heterogenous clusters
  - Different processors and memories.
- Power efficiency !





HPC EcoSystems



NLAFET

Parallel Numerical Linear Algebra for Future Extreme Scale Systems

Bo Kågström, Lennart Edblom, Lars Karlsson; Laura Grigori; Iain Duff, Jonathan Hogg; Jack Dongarra, and Nick Higham

Umeå University, Sweden; Inria Paris-Rocqueancourt, France; RAL—Science Technology Facilities Council, UK; and University of Manchester, UK

NLAFET—Aim and Main Research Objectives

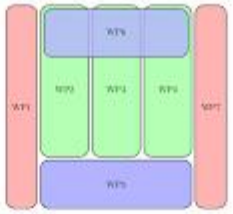
Aim: *Enable a radical improvement in the performance and scalability of a wide range of real-world applications relying on linear algebra software for future extreme-scale systems.*

- Development of novel *architecture-aware algorithms* that expose as much parallelism as possible, exploit heterogeneity, avoid communication bottlenecks, respond to escalating fault rates, and help meet emerging power constraints
- Exploration of *advanced scheduling strategies and runtime systems* focusing on the extreme scale and strong scalability in multi/many-core and hybrid environments
- Design and evaluation of novel strategies and software support for both *offline and online auto-tuning*
- Results will appear in the open source *NLAFET software library*

WP2, WP3 and WP4 at a glance!

- Linear Systems Solvers
- Hybrid BLAS
- Eigenvalue Problem Solvers
- Singular Value Decomposition Algorithms

NLAFET Work Package Overview



- WP1: *Management and coordination*
- WP5: *Challenging applications—a selection*
  Materials science, power systems, study of energy solutions, and data analysis in astrophysics
- WP7: *Dissemination and community outreach*
  Research and validation results; stakeholder communities

Research Focus—Critical set of fundamental LA operations

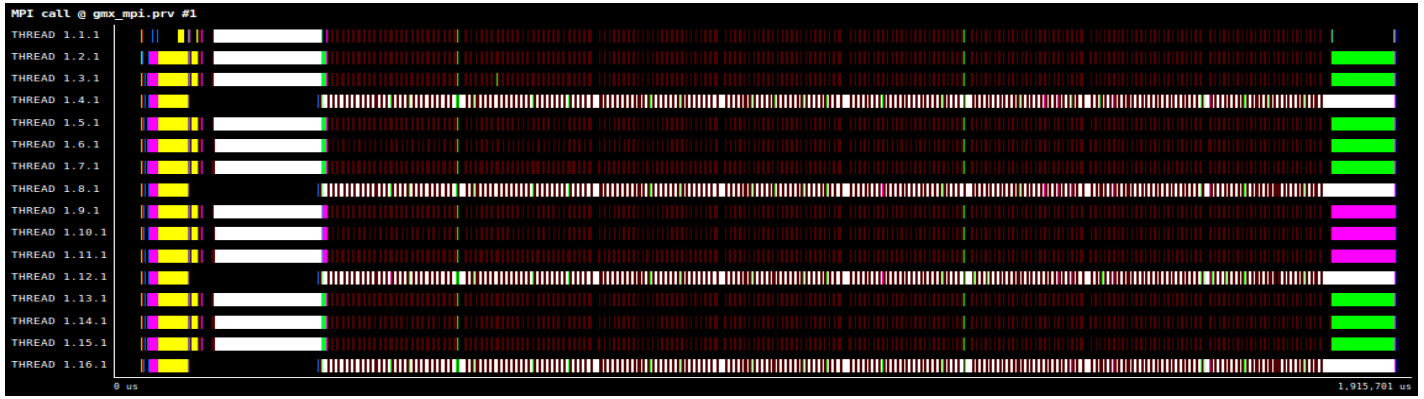
- WP2: *Dense linear systems and eigenvalue problem solvers*
- WP3: *Direct solution of sparse linear systems*
- WP4: *Communication-optimal algorithms for iterative methods*
- WP6: *Cross-cutting issues*

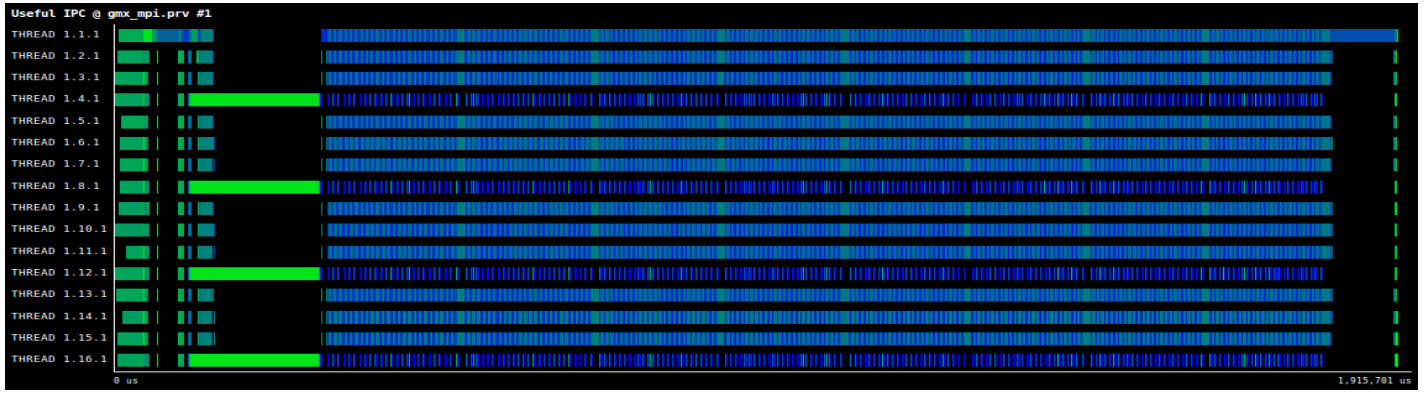
WP2, WP3 and WP4: research into extreme-scale parallel algorithms
WP6: research into methods for solving common cross-cutting issues

Avoid Communications—extreme-scale systems accentuate the need!

PRACE - Partnership for Advanced Computing in Europe

Tracing tools (GROMACS, 16 Cores)





Now to the clusters and programming models

A large amount of numbers and technical information will follow!!

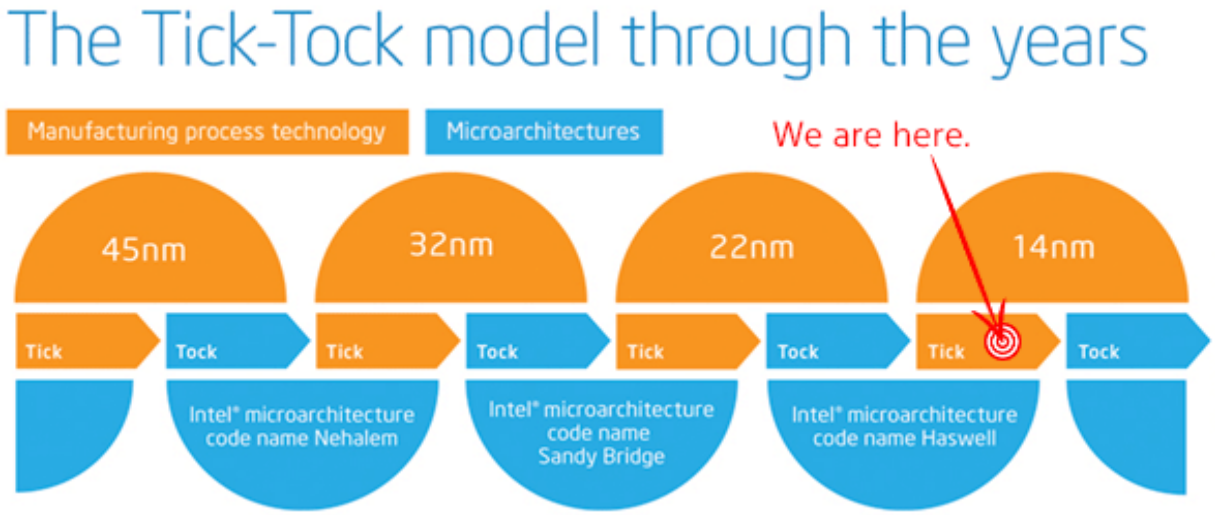
Relax, you do not need to now everything in detail, and we offer training for those things you should know.

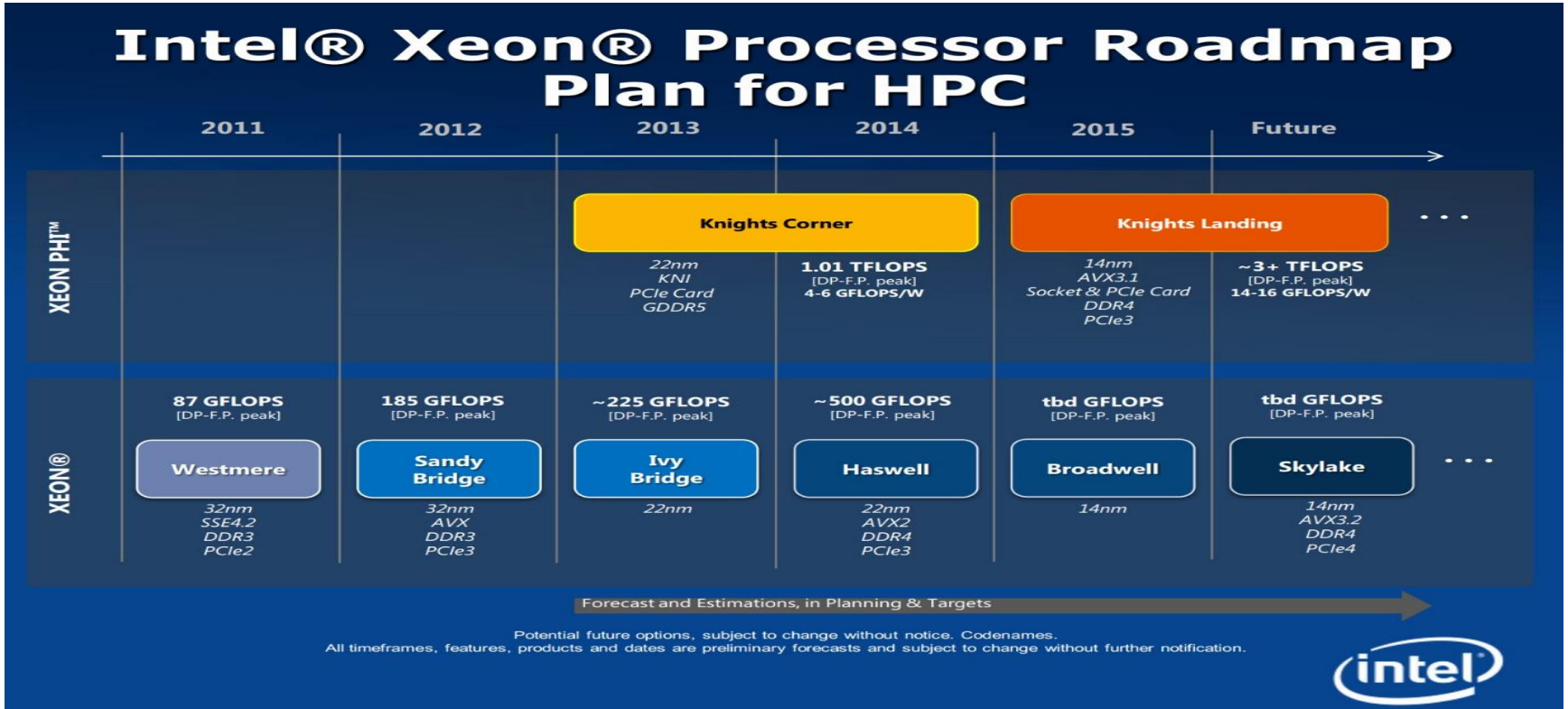
Abisko

- 332 nodes with a total of 15936 CPU cores.
- AMD Opteron 6238 (Interlagos)
- The 10 'fat' nodes have 512 GB RAM each, and the 322 'thin' nodes have 128 GB RAM each.
- (More details can be found on our web-pages)

Kebnekaise

Intels processors





Compute nodes

- 432 nodes
- Intel Broadwell (E5-2690v4)
- 2x14 cores/node
- 128GB memory
- Infiniband FDR

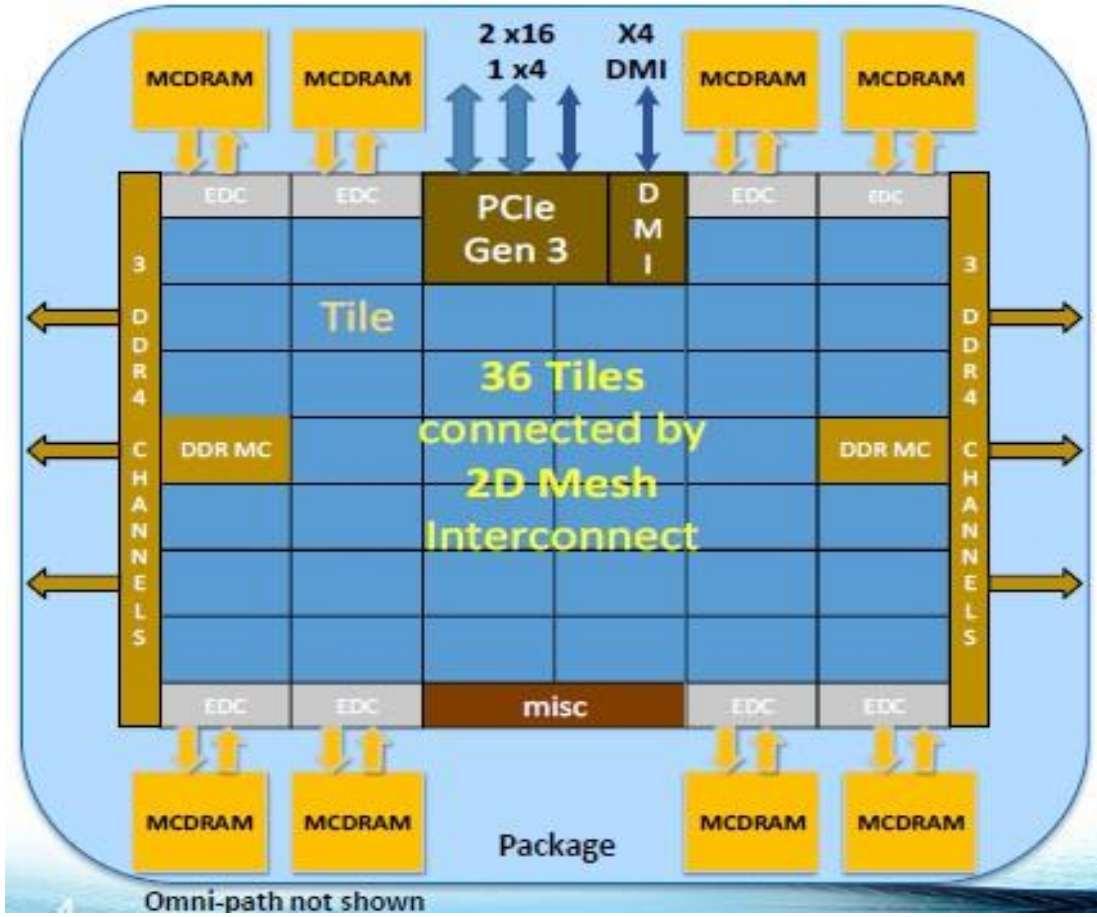
Large memory nodes

- 20 nodes
- Intel Broadwell (E7-8860v4)
- 4x18 cores/node
- 3TB memory
- Infiniband EDR

KNL - Intel Knights Landing

- 36 nodes
  - 68 cores
  - 1.4GHz (1.2GHz AVX)
- 192 GB memory - 16 GB MCDRAM
- Infiniband FDR
- *Installation in February*

Knights Landing Overview



Chip: 36 Tiles interconnected by **2D Mesh**
Tile: 2 Cores + 2 VPU/core + 1 MB L2

Memory: MCDRAM: 16 GB on-package; High BW
DDR4: 6 channels @ 2400 up to 384GB
IO: 36 lanes PCIe Gen3. 4 lanes of DMI for chipset
Node: 1-Socket only
Fabric: Omni-Path on-package (not shown)

Vector Peak Perf: 3+TF DP and 6+TF SP Flops
Scalar Perf: ~3x over Knights Corner
Streams Triad (GB/s): MCDRAM : 400+; DDR: 90+

Intel Xeon Phi

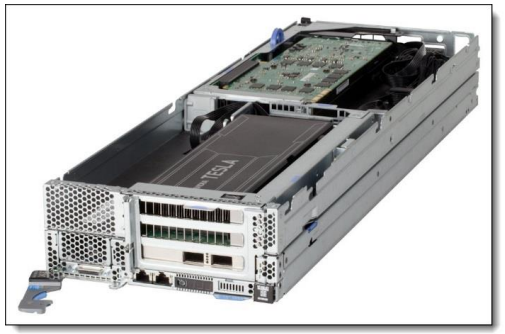
Xeon Phi	Clock Speed	Cores / Threads	Peak DP TFLOPS	DDR4 Memory	MCDRAM Memory Capacity	MCDRAM Memory Speed	TDP (Watts)	1K Tray Unit Price	$ / TFLOPS
Knights Landing									
7290	1.5 GHz	72 / 288	3.46	384 GB	16 GB	7.2 GT/sec	245	$6,254	$1,810
7250	1.4 GHz	68 / 272	3.05	384 GB	16 GB	7.2 GT/sec	215	$4,876	$1,601
7230	1.3 GHz	64 / 256	2.66	384 GB	16 GB	7.2 GT/sec	215	$3,710	$1,393
7210	1.3 GHz	64 / 256	2.66	384 GB	16 GB	6.4 GT/sec	215	$2,438	$916

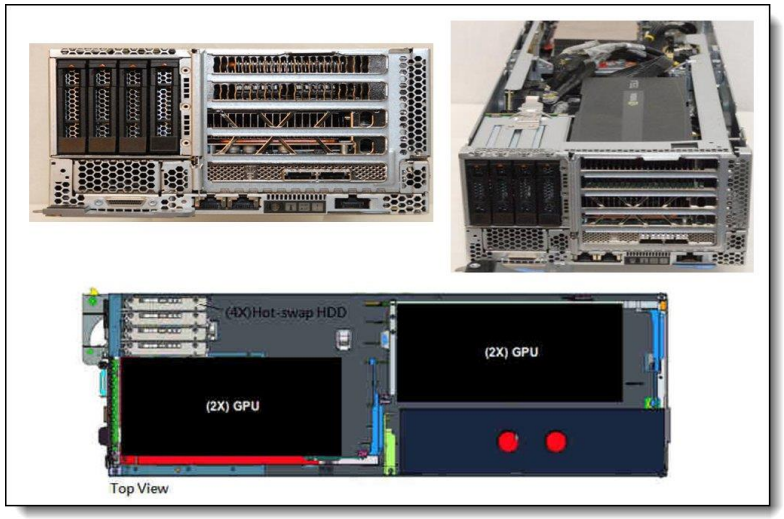
Xeon Phi	Clock Speed	Cores / Threads	Peak DP TFLOPS	Cache Memory	GDDR5 Memory Capacity	GDDR5 Memory Speed	TDP (Watts)	1K Tray Unit Price	$ / TFLOPS
Knights Corner									
7120P	1.24 GHz	61 / 61	1.21	30.5 MB	16 GB	5.5 GT/sec	300	$4,129	$3,412
7120X	1.24 GHz	61 / 61	1.21	30.5 MB	16 GB	5.5 GT/sec	300	$4,129	$3,412
5110P	1.05 GHz	60 / 60	1.01	30 MB	8 GB	5.0 GT/sec	225	$2,649	$2,623
5120D	1.05 GHz	60 / 60	1.01	30 MB	8 GB	5.5 GT/sec	245	$2,759	$2,732
3120A	1.10 GHz	57 / 57	1.0	28.5 MB	6 GB	5.0 GT/sec	300	$1,695	$1,695
3120P	1.10 GHz	57 / 57	1.0	28.5 MB	6 GB	5.0 GT/sec	300	$1,695	$1,695

General or special-purpose processor ?

GPU nodes

- 32 nodes with 2x NVidia K80
- 4 nodes with 4x NVidia K80
- Intel Broadwell 2x14 cores (E5-2690v4)
- 128 GB memory
- Infiniband FDR





High Speed Interconnect

- Infiniband
- Three level fat tree structure
- FDR cards in nodes (leafs)
- EDR cards in large memory nodes
- EDR in switches

Kebnekaise in numbers

- 13 racks
- 544 nodes
- 17552 cores (of which 2448 cores are KNL-cores)
- 399360 CUDA cores (80 * 4992 cores/K80)
- More than 125TB memory (20*3TB + (432 + 36) * 128GB + 36*192GB)
- 66 switches (Infiniband, Access network, Management network)

Kebnekaise in numbers

- 83% of the system are standard and Large Memory nodes
- 7% GPU-nodes
- 7% KNL-nodes
- 4% Other nodes (login and management nodes, LNET-routers etc)
- 728 TFlops/s Peak performance
- 629 TFLops/s HPL (all parts)
- HPL: 86% of Peak performance

Standard Nodes	374 TFlops/s
Large Memory Nodes	34 TFlops/s
2xGPU Nodes	129 TFlops/s
4xGPU Nodes	30 TFlops/s
KNL Nodes	62 TFlops/s
Total (All parts)	**629 Flops/s**