European Exascale Projects

- **European Exascale Software Initiative (EESI)**
  - Co-funded by the European Commission is to build a European vision and roadmap to address the Exascale challenges
  - 120 experts contributing to 8 working groups

- **Intel, imec and Five Flemish Universities Open Flanders ExaScience Lab**
  - Lab to Develop Solar Flare Prediction as Driver for Intel's Exascale Roadmap

- **Cray Exascale Initiative in Europe**
  - Team includes EPCC, Swiss National Supercomputing Center

1b. Mont-Blanc (Contribution of BSC)

European scalable and power efficient HPC platform based on low-power embedded technology

Energy efficiency is the key to Exascale performance - A 30-fold improvement needed in FLOPS/Watt (Based on a 20 MW power budget, exaflop requires an <u>energy efficiency</u> of 50 GFLOPS/Watt).

MB believes that HPC systems developed from today's energy-efficient solutions used in embedded and mobile devices are the most likely to succeed in achieving exaflop performance

<u>Project Participants</u> – a strong European consortium:

- Barcelona Supercomputing Centre, Spain
- Bull, France
- ARM Ltd, UK - Leader in energy-efficient processors
- Gnodal Ltd, UK
- Forschungszentrum Julich GmbH (FZJ), Germany
- Leibniz-Rechenzentrum der Bayerischen Akademie der Wissenschaften (LRZ), Germany
- Grand Equipement National de Calcul Intensif (GENCI), France
- CINECA, Italy

1c. CRESTA – Collaborative Research into Exascale Systemware, Tools and Applications (Contribution by HLRS)
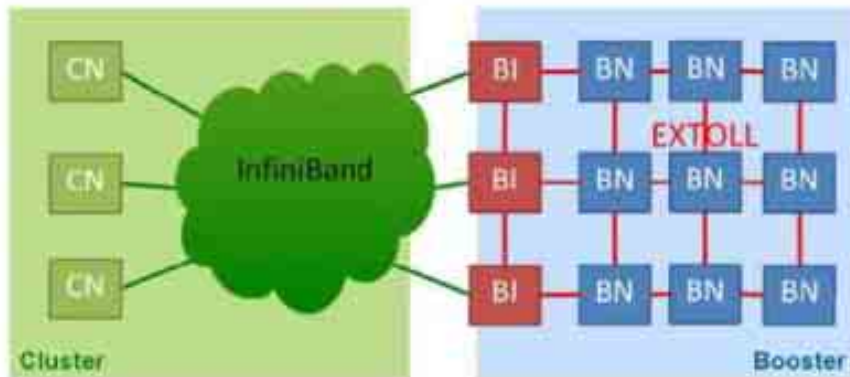
- CRESTA will deliver major advances in the techniques and technologies required to build Exascale computers in Europe this decade

- Consortium
  - EPCC, HLRS, CSC, PDC
  - Cray
  - TUD (Vampir), Allinea (DDT)
  - ABO, JYU, UCL, ECMWF, ECP, DLR

- Projects start 1st October 2011 – currently negotiating contracts

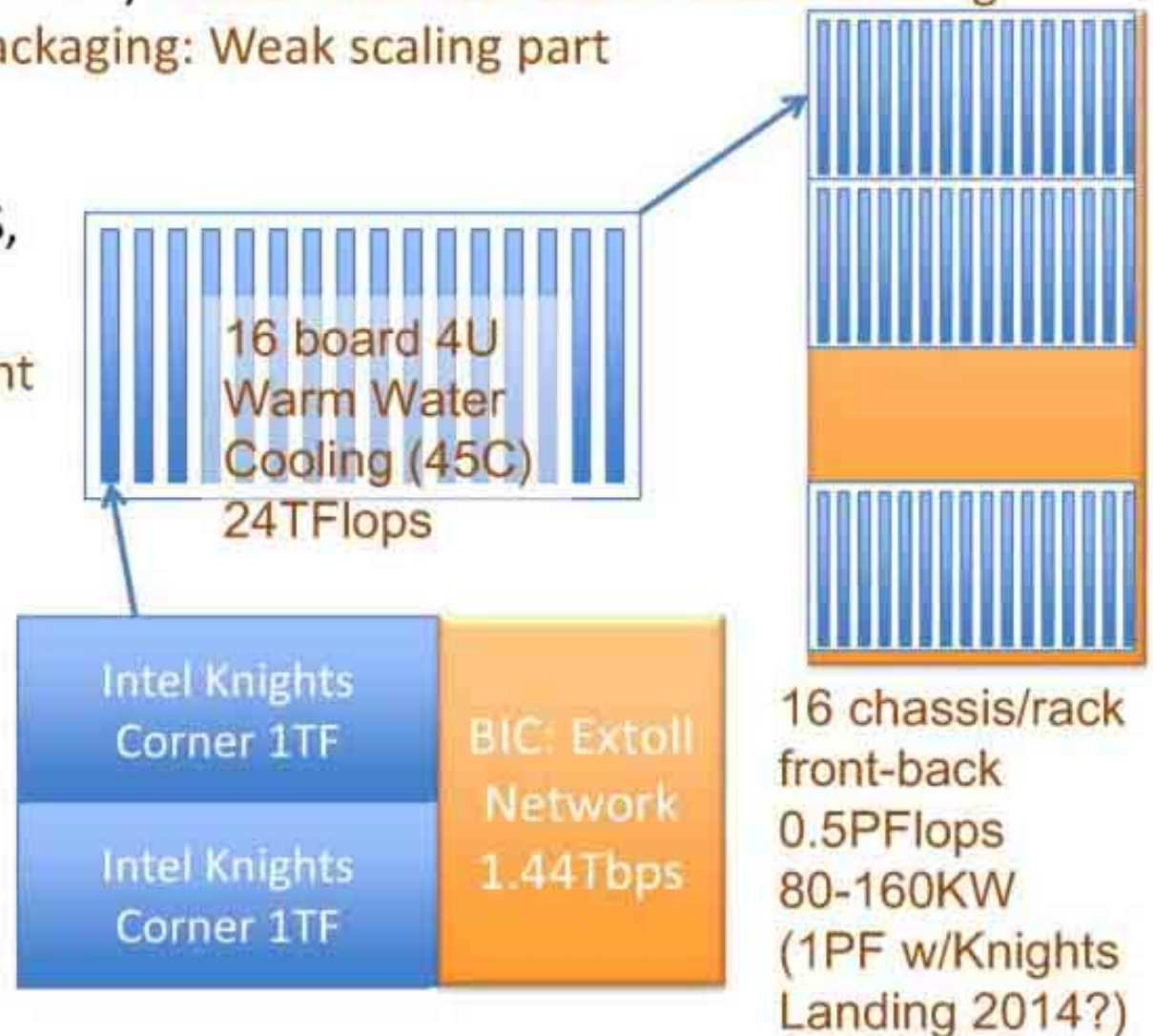- 3 year project, 13 partners, €12 million costs, €8.5 million funding

CRESTA

Julich-Intel-others DEEP Project

- **HW: Loosely-coupled hybrid system**
  - Cluster: Infiniband+standard x86, Strong Scaling part
  - BOOSter Intel MIC (KnightsXXX)+Extoll Network+Warm water Cooling (45 degrees C) + dense packaging: Weak scaling part

- **SW: For Hybrid System**
  - System Software (MIC-OS, ParaStation, etc.)
  - Programming Environment
  - Energy Efficiency

16 board 4U Warm Water Cooling (45C) 24TFlops

Intel Knights Corner 1TF

Intel Knights Corner 1TF

BIC: Extoll Network 1.44Tbps

16 chassis/rack front-back 0.5PFlops 80-160KW (1PF w/Knights Landing 2014?)

CN

CN

CN

InfiniBand

BI — BN — BN — BN

BI — BN — BN — BN

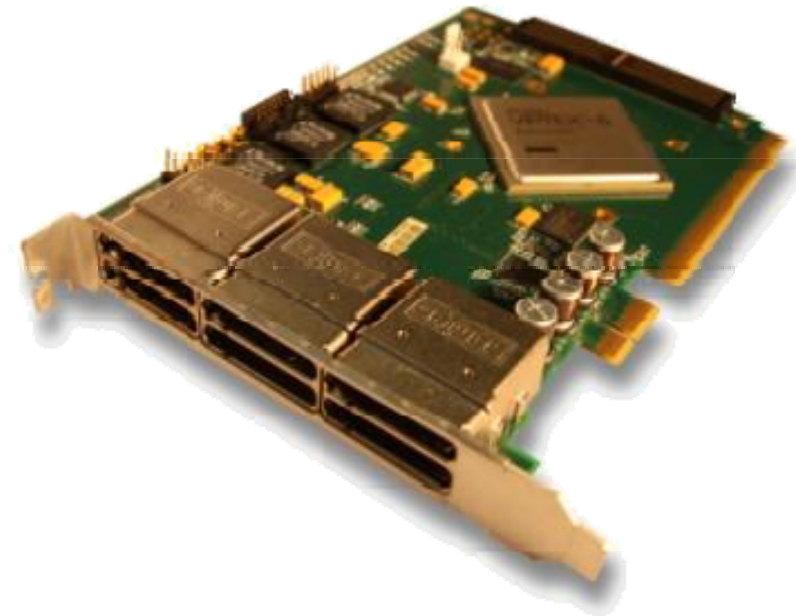BI — BN — BN — BN

EXTOLL

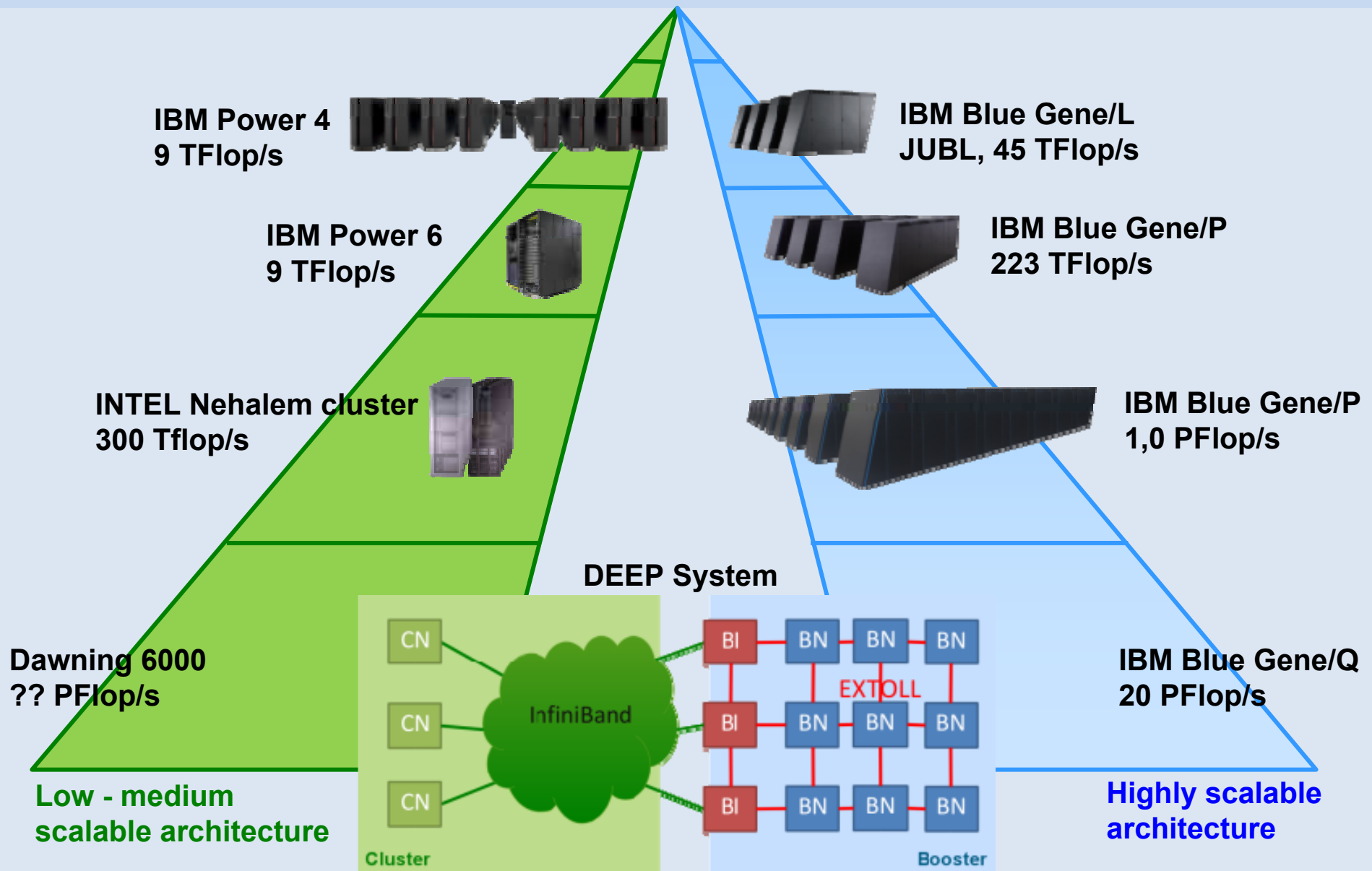Cluster

Booster

The DEEP Architecture sketch

Technology Components

- **Intel Knights Corner**

- **EXTOLL (for booster)**

  - 120 Gbit per link unidir

  - 1440 Gbit/card bidir, 3d

  - 0.4 µs latency

- **Mellanox IB (for cluster)**

  - State-of-the-art interconnect

- **ParaStation cluster OS**

- **Intel Compiler and Tools**

Positioning DEEP: A fusion of general purpose and high scalability supercomputers

IBM Power 4
9 TFlop/s

IBM Blue Gene/L
JUBL, 45 TFlop/s

IBM Power 6
9 TFlop/s

IBM Blue Gene/P
223 TFlop/s

INTEL Nehalem cluster
300 Tflop/s

IBM Blue Gene/P
1,0 PFlop/s

DEEP System

Dawning 6000
?? PFlop/s

IBM Blue Gene/Q
20 PFlop/s

Low - medium
scalable architecture

Highly scalable
architecture

Motivations and organization of EESI

Coordinate the European contribution to IESP

Enlarge the European community involved in the software roadmapping activity

Build and consolidate a vision and roadmap at the European Level, **including applications**, both from academia and industry

EESI is a coordination and support action – FP7/Infrastructures www.eesi-project.eu
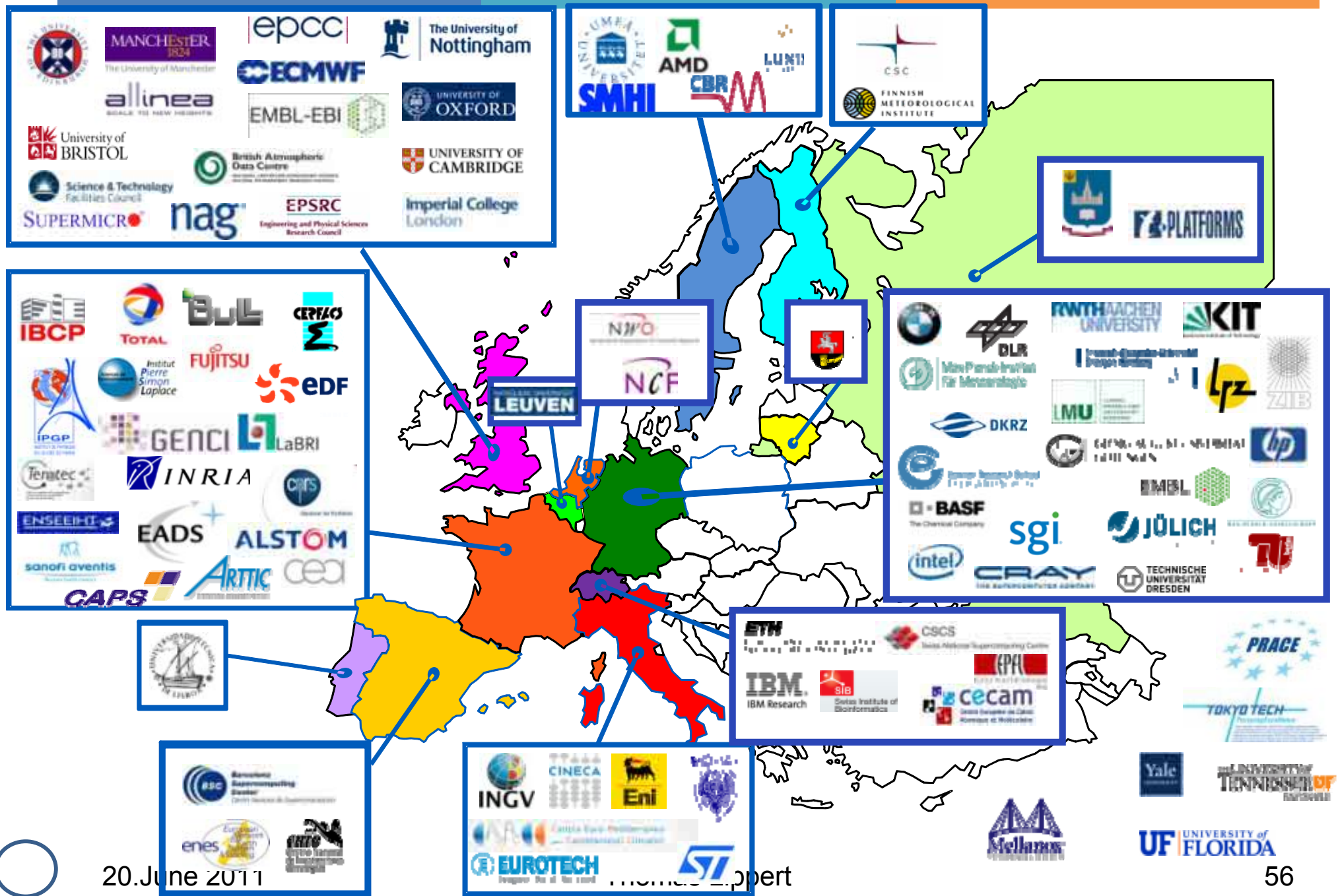
Coordinator: EDF R&D, **Jean-Yves Berthou**

Starting date : 1st of June 2010, for 18 months

Requested EC contribution : 640 000 €

Consortium :   **8 contractual partners**,17 associated participants, 11 contributing participants

EESI participants

European Exascale Software Initiative AGENDA

Enabling technologies for Exaflop computing

- Hardware roadmap, links with vendors
- Software eco-systems
- Numerical libraries, solvers and algorithms
- Scientific software engineering

Initial International workshop

Internal workshop: presentation of each working group results and roadmaps

Final conference: public presentation of project result

Application Grand Challenges

- Industrial and Engineering Applications (Transport, Energy)
- Weather, Climatology and Earth Sciences
- Fundamental Sciences (Chemistry, Physics)
- Life science, Health, BPM

Constitution of WG, Setup of guidelines, organisation modes

Initial cartography of existing HPC projects, initiatives in Europe, US and ASIA

Synthesis of all contributions and production of a set of recommendations

Presentation of EESI results to the EC

Updated cartography of existing HPC projects, initiatives in Europe, US and ASIA

T0	T0+5	T0+12	T0+13	T0+16	T0+17	T0+18
5 months	8 months		1 month	3 months	1 month	1 month

June 1, 2010   October 2010   June 2011   November 2011

WG4.2 Testbed needs in 2015 (draft)

━━━ < 100 Petaflops enough ━━━ 100 Petaflops needed ━

Dedicated/reconfigurable	
-**Prog. Models and Runtime** (interaction of runtime with kernel scheduler) -**Performance tools**: (interaction with resilience / power components) -**OS** (scheduling, memory management) -Measure noise generation (most of the syst fotware)? -**I/O-File system** (may be done at lower scale) (need root access for reconfiguration) -**Job and resource manager** (need root access for reconfiguration)	-**Performance tools** (system level): measure and minimize overheads induced at the full scale by low-level performance monitoring infrastructure tightly integrated with OS -**Prog. Models and Runtime** (interaction of runtime with kernel scheduler)? -...

| Production	
-**Compilers** (node level) -**Programming models** (node level API) -**Performance tools** (node level) -**Power management**? -**Validation and correctness checking**?	-**Resilience** (FT protocols, ABFT, NFTA, execution state storage) -**Parallel debuggers**? (scalability test) -**Performance tools** (scalability of data collection and analysis) -**Runtime** (scalability test) -**Prog. models** (system level scalability) -**Performance modeling at scale**

China Exascale Plans

- ## 12th 5-year Plan (2011-15)
  - **Several petascale HPCs**
    - E..g. Chinese Academy of Sciences – 10PF (CNY 700mil)
    - Godson Processor R&D?
  - **At least one 50-100 PFLOPS**
  - **Budget: CNY 4 Billions**
    - MOST: 60%, Local Government: 40%

- ## 13th 5-year Plan (2016-20)
  - **1~10 ExaFLOPS HPC**
  - **Budget: unknown yet**

China National Grid (CNGrid)
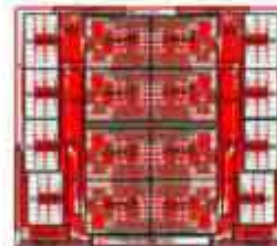
Godson-3B

IESP

- International Exascale Software Project
  - formed with the realization that current software used for terascale and petascale computing is inadequate for exascale computing.
  - Industry – Academia – National Labs working synergistically
  - 6 workshops held (latest in San Francisco)
- X-Stack : Co-design efforts with the software and vendor segments of the IESP community
  - Key Application Areas:
  - Chemistry, Nuclear Energy, Fusion, Materials, High Energy Density Physics, Astronomy

International Exascale Software Project
Software Topics in the Current Roadmap

4.1 Systems Software

4.1.1 Operating systems

4.1.2 Runtime Systems

4.1.3 I/O systems

4.1.4 Systems Management

4.1.5 External Environments

4.2 Development Environments

4.2.1 Programming Models

4.2.2 Frameworks

4.2.3 Compilers

4.2.4 Numerical Libraries

4.2.5 Debugging tools

4.3 Applications

4.3.1 Application Element: Algorithms

4.3.2 Application Support: Data Analysis and Visualization

4.3.3 Application Support: Scientific Data Management

4.4 Crosscutting Dimensions

4.4.1 Resilience

4.4.2 Power Management

4.4.3 Performance Optimization

4.4.4 Programmability
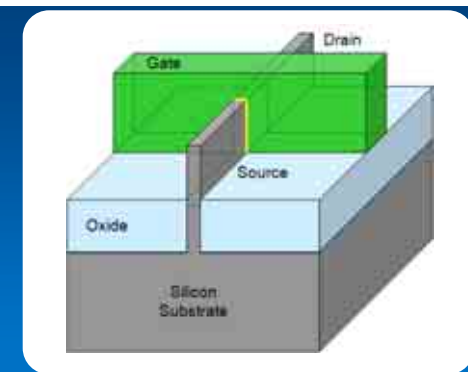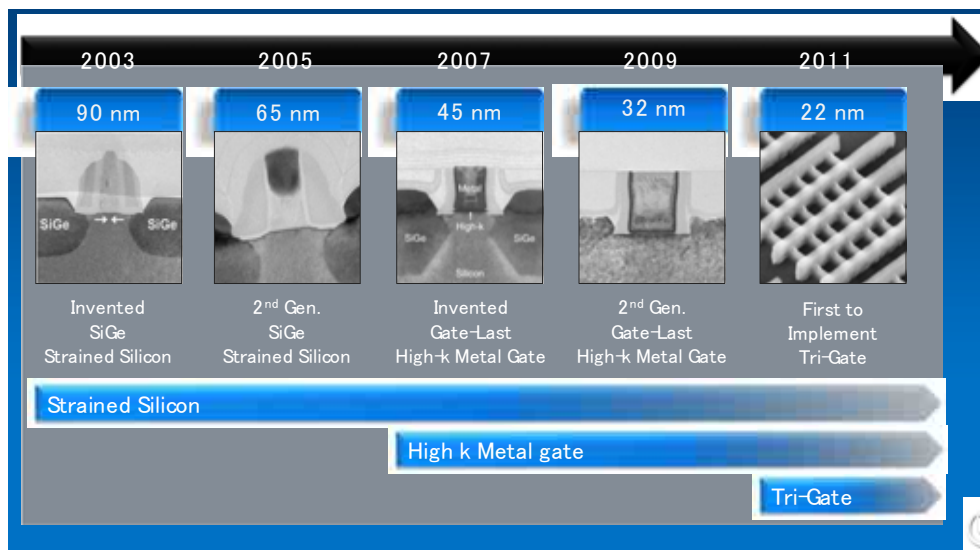
INTERNATIONAL EXASCALE ROADMAP 1.0
SOFTWARE PROJECT

Intel 22nm 3D TriGate Transistor Technology

- Intel announces a major breakthrough and historic innovation in microchips: the world's first 3-D transistors in mass production

- Intel's 3-D Tri-Gate transistors enable chips to operate at lower voltage with lower leakage, providing a combination of improved performance and energy efficiency compared to previous state-of-the-art transistors.

- The 3-D Tri-Gate transistor will be implemented in Intel's upcoming 22nm manufacturing process

- Intel demonstrates a 22nm microprocessor -- code-named Ivy Bridge -- that will be the first high-volume chip to use 3-D transistors.
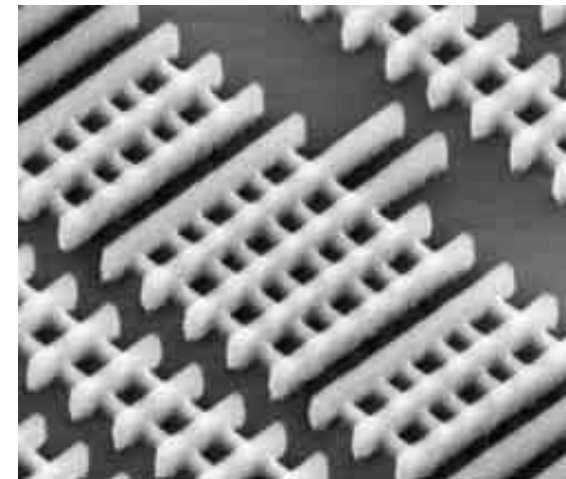
This image shows the vertical fins of Intel's revolutionary tri-gate transistors passing through the gates. The image shows the vertical fins of Intel's revolutionary tri-gate transistors passing through the gates.
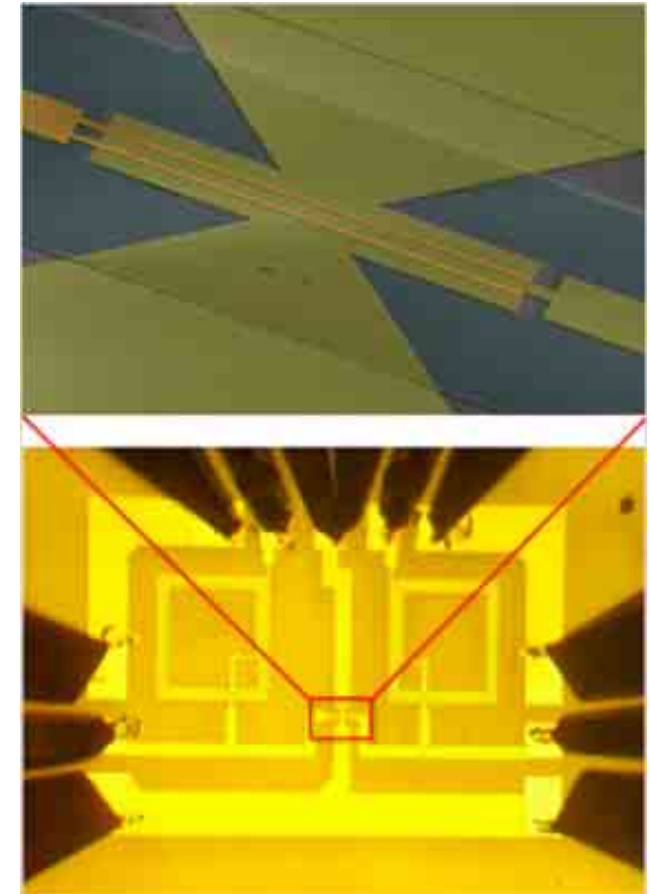
2003	2005	2007	2009	2011
90 nm	65 nm	45 nm	32 nm	22 nm
Invented SiGe Strained Silicon	2nd Gen. SiGe Strained Silicon	Invented Gate-Last High-k Metal Gate	2nd Gen. Gate-Last High-k Metal Gate	First to Implement Tri-Gate

Strained Silicon

High k Metal gate

Tri-Gate

3-D Tri-Gate transistors form conducting channels on three sides of a vertical fin structure, providing "fully depleted" operation

Transistors have now entered the third dimension!

IBM: Graphene based Integrated Circuits

- First integrated circuit based on Graphene transistor developed at IBM T.J. Watson Research Center

- Graphene was "grown" on the silicon wafer and then coated with a common polymer PMMA.

- Prototype Graphene based transistors created at IBM run at 100 GHz

- The integrated circuit operates as a broadband radio-frequency mixer at frequencies up to 10 GHz

- Graphene is being explored as a substitute for materials like Gallium Arsenide


Graphene based integrated circuit

http://www.nytimes.com/2011/06/10/technology/10chip.html?hpw
http://spectrum.ieee.org/semiconductors/devices/first-graphene-integrated-circuit

3D Stacking

- IBM Zurich Research Labs

  - Developing techniques to resolving exponential amount of heat developed by 3d stacked circuits

  - The 50 micron channels between individual chip layers are performing at 180 watt/cm^2 per layer for a stack with a typical footprint of 4 cm^2

- Xilinx FPGA

  - Large-capacity Virtex 7 FPGAs using 28nm process technology

  - Xilinx has chosen "2.5D" interconnect technology

  - The silicon interposer is thermally matched to the silicon FPGA slices and it offers 20x the I/O density of ceramic interposers
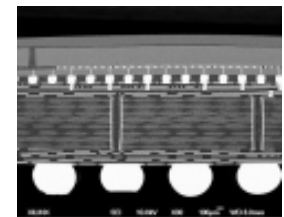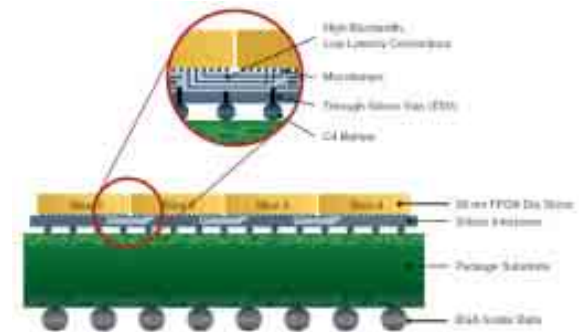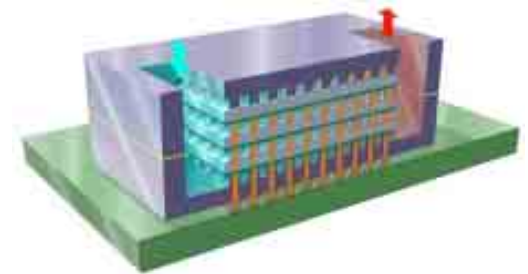
IBM SC vs. IT Leverage Strategy

- NCSA Blue Waters

- LRZ Cluster(温水冷却)

- LANL Roadrunner

- LLNL/ANL/Julich BlueGene

コンポーネンツや
IT技術の直接的な
レバレジ

それぞれの
マーケット
セグメントに
対応

共通のソフトウェア
エコシステム

我が国の
スパコン開発は
このような
大域的な視点が
伝統的に
欠けている

- Mainframes

- Mid-Low x86
Servers & PCs

- Gaming&Multimedia
(PS3/Cell, XBOX360)

- Embedded
(PowerPC)

今後:エクサプロジェクトへ向けて

- グローバリゼーション時代、日本の産業の強みは実は垂直的ものづくりとエンド製品ではなく、中間部品とそのIPに移りつつある
  - 今回の震災:「日本の工場が止まると世界が止まる」
- 上から下まで「国産技術」で垂直的にエクサのスパコンを開発するのはその流れには必ずしも一致しない
  - むしろ世界中のスパコンやIT製品に独占的に使われる技術の開発の方が技術戦略としては好ましい?
- スパコンを構成する全ハード・ソフト・アプリの部品で、どこを国内技術として死守・国際共同開発・他の技術の利用をきちんと評価すべき
  - EESIのように「何を研究開発すべきか」「ロードマップは」を検討するWG等の早期立ち上げ・報告が必須
  - 中間的なシステムで段階的進化と技術レビュー
  - 技術的ディスラプションの為の基礎研究も
  - ハード・ソフト・アプリのCo-Design体制の確立