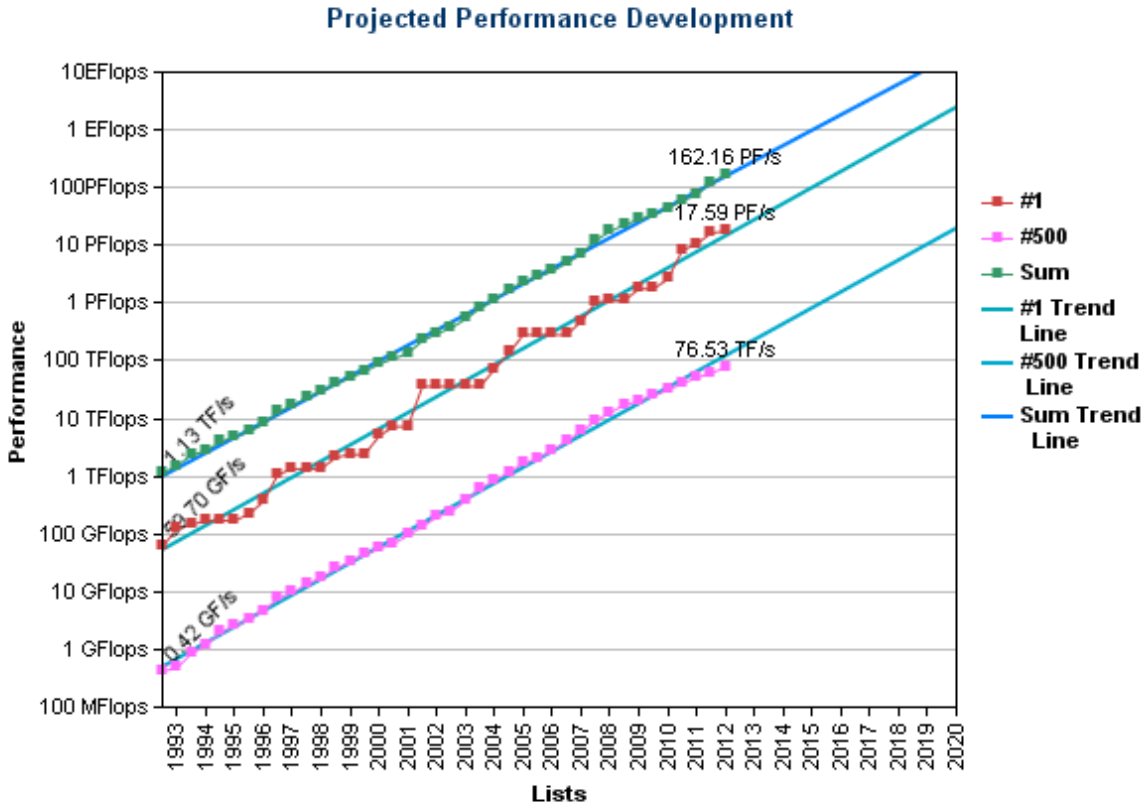
Performance Analysis of HPC Applications on Low-Power Embedded Platforms

Luka Stanisic, Brice Videau, Johan Cronsioe,
Augustin Degomme, Vania Marangozova,
Arnaud Legrand, Jean-François Méhaut
(CEA, CNRS, INRIA, UJF, Grenoble, France)
Edson Padoin, Laercio Pilla, Pedro Velho,
Marcio Castro, Alexandre Carissimi,
Philippe Navaux
(UFRGS, Porto Alegre, Brazil)

HPC: “We have a dream...”
“Reaching the Exaflop before 2020”

- Current efficiency: ~2 GFLOPS/W
- 20 MW Budget
- Target efficiency: 50 GFLOPS/W
- 25x better efficiency needed



Potential Solution: the Embedded Way

- Use components designed and produced by the embedded (european) industry (ARM, STMicroelectronics, Kalray,...)
- Power efficiency is excellent
- Use off-the-shelf components to keep costs down
- The Mont-Blanc European Project is working on the use of these components for HPC

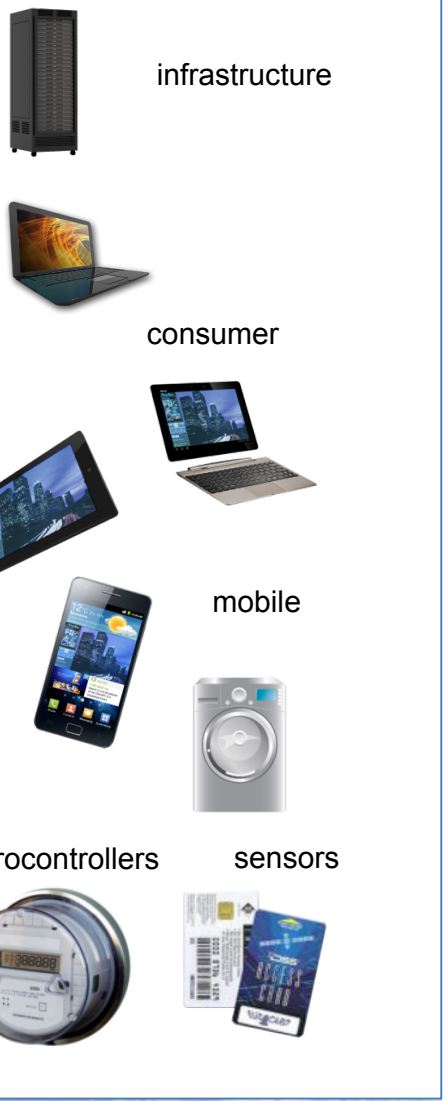
Mont-Blanc 1 - Objectives

- Develop prototypes of HPC clusters using low power (ARM) commercially available embedded technology.
- Design the next generation in HPC systems based on embedded technologies and experiments on the prototypes.
- Develop a portfolio of existing applications to test these systems and optimize their efficiency, using BSC's OmpSs programming model (11 existing applications were selected for this portfolio).

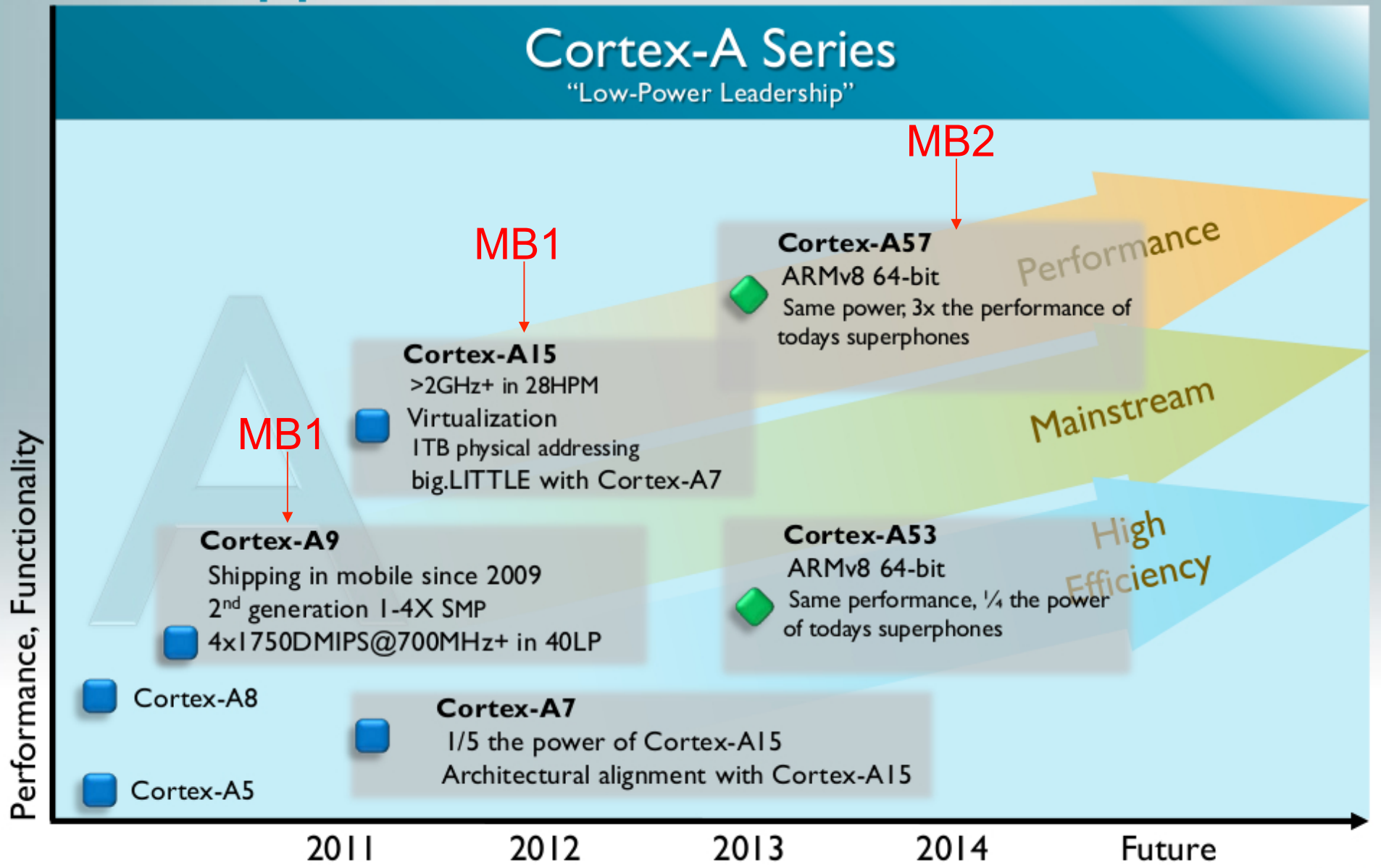
MONT-BLANC

Scalability of the ARM Architecture

- Low power foundation enables ARM processor to scale from sensors to servers
  - Software compatibility makes it easier to scale up
  - Lower power processors enables a more integrated system-on-chip
- Power efficiency is important for all applications
  - More attractive system design – thinner and lighter
  - Lower cost – plastic chip packaging, no heat sink or fan, cheaper components
  - Lower energy use – reduced running costs and
  - carbon footprint



ARM Applications Processors



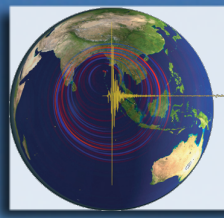
Cortex-A57: Implementation for High Performance

- Maximum performance in smartphone power budget
  - 3x performance of 2012 super phones, in 32-bit mode
  - Performance per clock comparable to today's PC's
- Driving advanced computing
  - 5x power-efficiency for tomorrow's tablets and notebooks
  - Mobile-level power consumption for enterprise and cloud
- Enhanced capabilities for enterprise and servers
  - 64-bit support for full range of applications
  - High performance IEEE compliant DP floating point / SIMD
  - Solutions already scaling up to a 16x SMP compute node
  - **SoC flexibility** to scale **efficiently** to 100K's or more nodes

Mont-Blanc 1 – Applications Beyond Linpack

Code	Scientific Domain	Contact	Institution
YALES2	Combustion	V. Mouveau	CNRS/CORIA
EUTERPE	Fusion	T. Akgun	BSC
SPECFEM3D	Wave propagation	D. Komatitsch	CNRS
MP2C	Multi-particle collision	G. Sutmann, A. Schiller	JSC
BigDFT	Elect. Structure	T. Deutsch	CEA
Quantum Expresso	Elect. Structure	C. Cavazzoni	CINECA
PEPC	Coulomb + gravitational forces	P. Gibbon, L. Arnold	JSC
SMMP	Protein folding	J. Meinke	JSC
ProFASI	Protein folding	S. Mohanty	JSC
COSMO	Weather forecast	M. Culpo	CINECA
BQCD	Particle physics	M. Allalen	LRZ

MONT-BLANC



SPECFEM3D (INRIA, CNRS,...)

- acoustic wave propagation for geophysics (earthquakes) or industrial problems (non destructive testing, ocean acoustics…)
- Implementation details
  - 80,000 lines of Fortran 90 – MPI, CUDA; a mini-app port in OMPSs + MPI exists; External libs : none
  - Portability : runs on almost all types of existing systems
  - Scalability : excellent, 0.7 PF on 149 784 cores of Jaguar and up to 1152 GPUs (TSUBAME 2.0)

MONT-BLANC

BigDFT (CEA INAC, UJF,...)

- DFT code using Daubechies wavelets for electronic structure calculation in nano-sciences and biology
- Implementation details
  - 200,000 lines of Fortran 90, C
  - Supports MPI, OpenMP, CUDA and OpenCL
  - External libs : BLAS
  - Portability : wide
  - Scalability : up to 4000 cores of BG/P and 288 GPUs

MONT-BLANC

Snowball Board

- Full embedded system by Calao Systems
  - Dual Core ST-Ericsson ARM A9 1GHz SoC
  - Neon Floating point Unit
  - Integrated GPU
  - 1 Gbyte RAM
  - HDMI output
  - Ethernet
  - Runs Linux (Linaro Ubuntu) or Android
  - Igloo community for development and support
  - 2.5W maximum energy consumption

MONT-BLANC

Experiments on Snowball

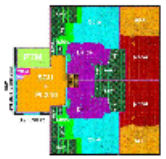
- Benchmarks
  - Linpack
    - Standard benchmark for HPC
  - CoreMark
    - Aims at becoming and industry standard for embedded platforms
    - Lots of results available on the website, allows to compare quickly CPU efficiency
  - StockFish
    - Open Source chess engine with benchmarking capabilities

Snowball results

- Compared to an Intel Xeo²n x5550 (4 Nehalem cores, TDP:95W)
  - Ratio : Xeon perf / Snowball perf
  - Perf/Watt ratio : (Xeon perf/95)/(Snowball perf/2.5)

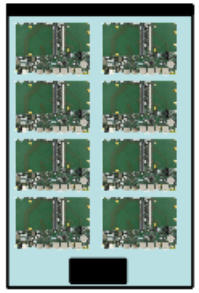
Bench	Snowball	Xeon 4 cores	Ratio	Perf / Watt Ratio
Linpack (MFlops)	620	24000	38.7	1.0
CoreMark (ops/s)	5877	41950	7.1	0.2
StockFish (ops/s)	224,113	4,521,733	20.2	0.5
Specfem3D (s)	186.8	23.5	7.9	0.2
BigDFT (s)	420.4	18.1	23.2	0.6

Tibidabo cluster



Tegra2 SoC:
2x ARM Corext-A9 Cores
2 GFLOPS
0.5 Watt

Tegra2 Q7 module:
1x Tegra2 SoC
2x ARM Corext-A9 Cores
1 GB DDR2 DRAM
2 GFLOPS
~4 Watt
1 GbE interconnect



1U Multi-board container:
1x Board container
8x Q7 carrier boards
8x Tegra2 SoC
16x ARM Corext-A9 Cores
8 GB DDR2 DRAM
16 GFLOPS
~35 Watt

Rack:
32x Board container
10x 48-port 1GbE switches
256x Q7 carrier boards
256x Tegra2 SoC
512x ARM Corext-A9 Cores
256 GB DDR2 DRAM
512 GFLOPS
~1.7 Kwatt

300 MFLOPS / W

32 blades of 8 harmony Tegra 2 boards already deployed

- 35W/blade
- 16 ARM Cortex A9 1GHz cores/blade

MONT-BLANC

Tibidabo Benchmarks – Single Node

- Slightly quicker than snowball board, but given for 4Watts

	Tegra 2	Snowball
Coremark	5952	5877
Specfem3D 2 cores	178.9s	186.8s
Linpack 2 cores peak	698 MFlops	620 MFlops

MONT-BLANC

Tibidabo Scaling

Impossible d'afficher l'image. Votre ordinateur manque peut-être de mémoire pour ouvrir l'image ou l'image est endommagée. Redémarrez l'ordinateur, puis ouvrez à nouveau le fichier. Si le x rouge est toujours affiché, vous devrez peut-être supprimer l'image avant de la réinsérer.

Impossible d'afficher l'image. Votre ordinateur manque peut-être de mémoire pour ouvrir l'image ou l'image est endommagée. Redémarrez l'ordinateur, puis ouvrez à nouveau le fichier. Si le x rouge est toujours affiché, vous devrez peut-être supprimer l'image avant de la réinsérer.

Impossible d'afficher l'image. Votre ordinateur manque peut-être de mémoire pour ouvrir l'image ou l'image est endommagée. Redémarrez l'ordinateur, puis ouvrez à nouveau le fichier. Si le x rouge est toujours affiché, vous devrez peut-être supprimer l'image avant de la réinsérer.

Tibidabo - scaling

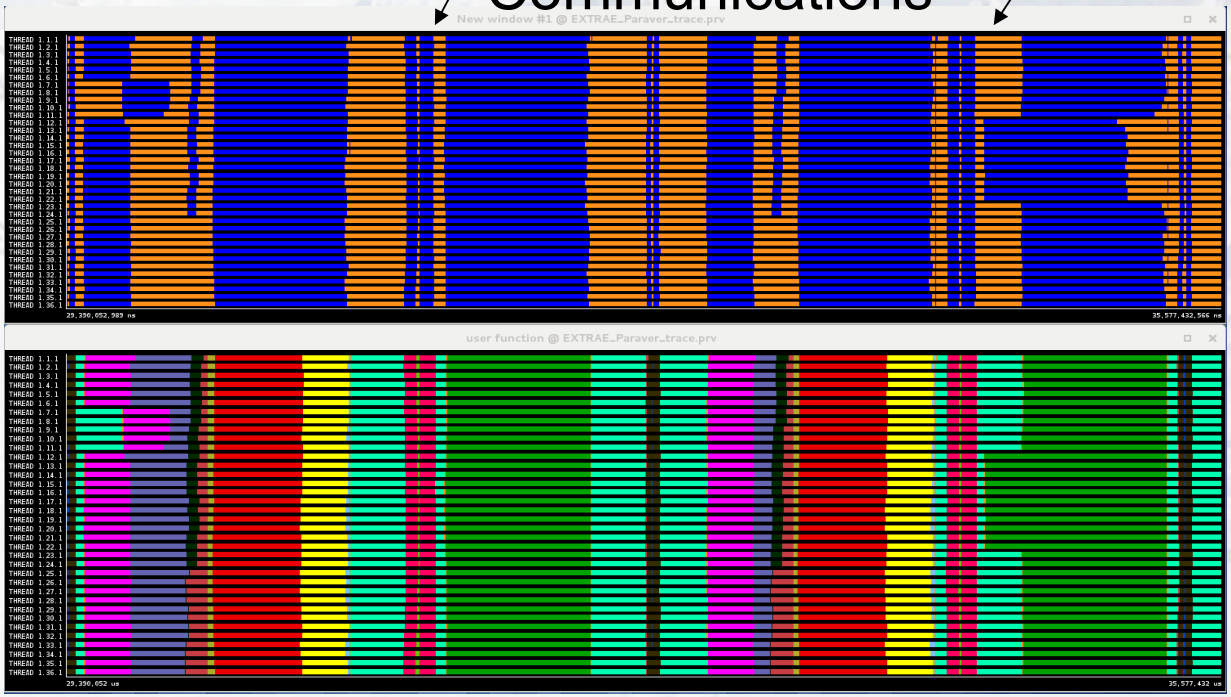
Poor scaling due to collective communications problems: speedup 15 for 36 cores

Most of BIGDFT communicatio ns are delayed

Normal communication (small communications)

Delayed communication

Communications



Computations

Optimization is Never-ending

- HPC environment is ever-changing, Multiple paradigms coexist
- Architectures have to be efficiently used

Archihtecture	AMD/Intel CPU	ARM	GPU	MIC
Cores	4-12	2-4	512-1536	64
Cache	3l coherent	2l coherent	1l incoherent	2l coherent
Memory (GiB)	2-4 (per core)	1 (per core)	2-6	2-?
Vector Size	2-4	1-2	1-2	8
Peak GFLOPS	10-20(per core)	2-4 (per core)	500-1000	1000

Counter-Intuitive Results

Seemingly harmless optimization can prove harmful on different architectures

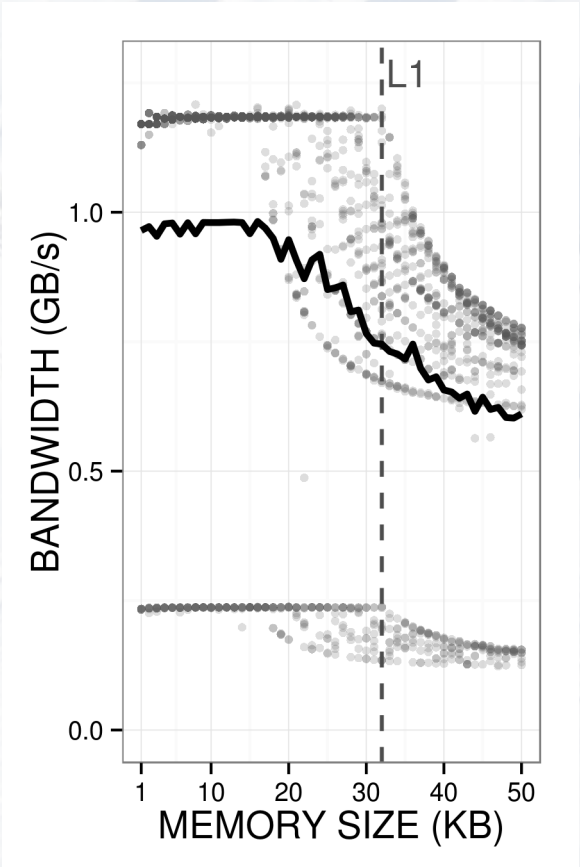
Intel Xeon 5500

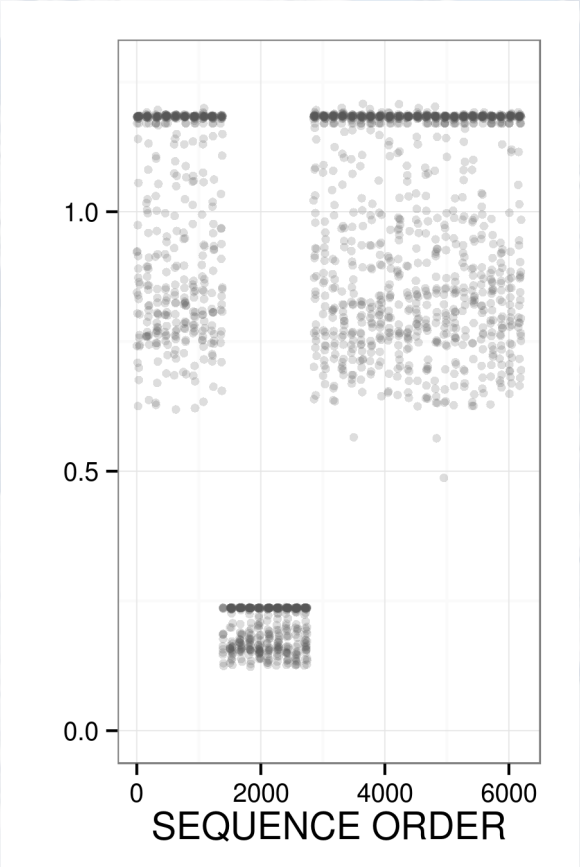
ARM A9500

Toward Practical Optimization of HPC Applications on ARM

Memory alignment and scheduling policies are important:

- Real-time is counter productive
- Variations due to memory continuity





MB2: Use Auto-Tuning in Applications

- Test case: BigDFT
- Build a parametrized generator that can optimize BigDFT building blocks: convolutions
- Those operations are badly optimized and vectorized by compilers (Intel Nehalem, also ARM):
  - Simple: 0.3 GFLOPS
  - Hand unrolled: 3 GFLOPS
  - Hand vectorized: 6.5 GFLOPS

MB2: Harness the Developer Knowledge

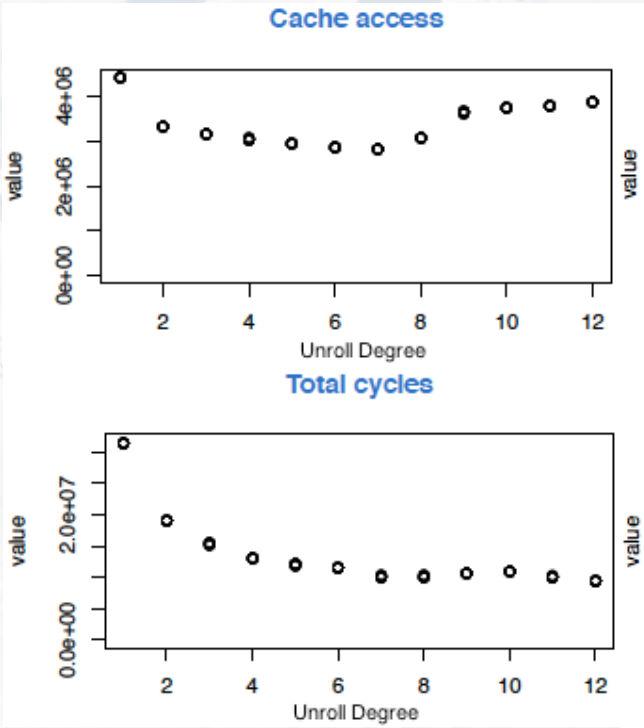
- The developer often knows the kind of optimization that would be beneficial
- Even when asked nicely compilers don't always comply
- Give developers tool to meta-program optimizations in their code

MB2: Source-to-Source Transformation

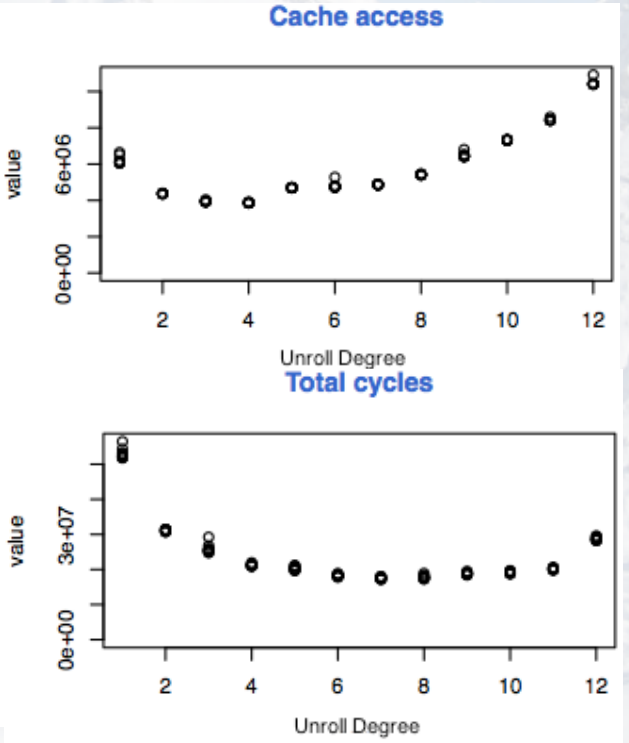
- Bringing Optimization through Automatic Source-to-Source Transformations
- Used to generate multiple version of the reference convolution
- Versions are automatically benchmarked using performance counters.
- Selection of best versions

Architecture Dependant Optimization Parameters

Different optimal unroll degree and resource usage pattern, depending on the architecture



Intel T7500



Tegra 2

Conclusions – Mont-Blanc 1

- Porting to Tibidabo is a success
- Most limitation come from the prototype (network) and should be lifted in the next version
- Second round of porting is to begin
- SPECFEM3D and BigDFT kernels have been selected specifically for optimization

Perspectives – Mont-Blanc 2

- Generator is a success
- Porting to the second Mont-Blanc prototype: an opportunity to generalize our generator?
- Apply generator to OpenCL kernels instead of using C++
- Can we generate vectorized convolutions?