SC12
Salt Lake City, Utah
November 10-16, 2012

SC12 Report

Muhong Zhou

What's going around

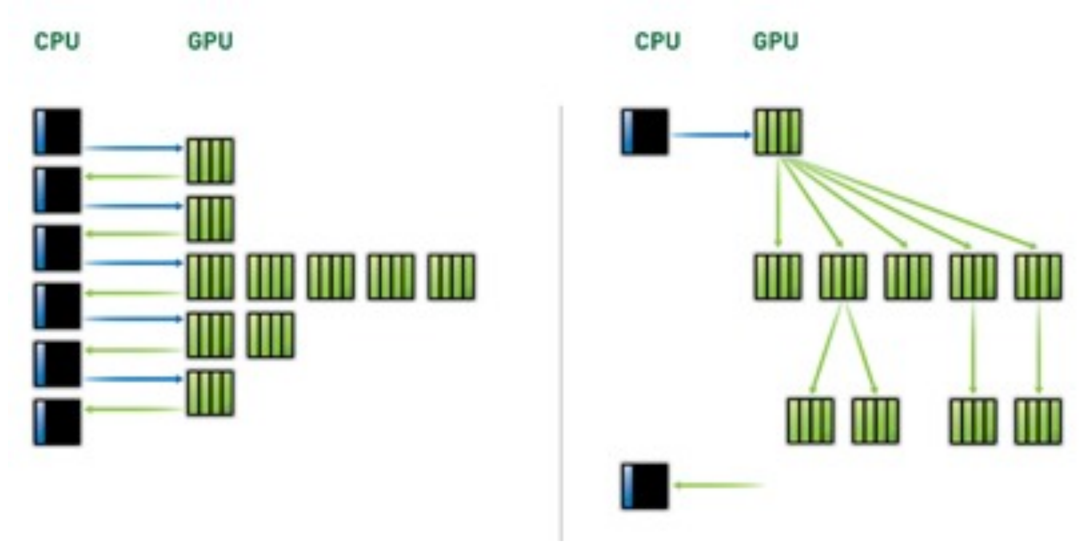
GPU to Accelerators

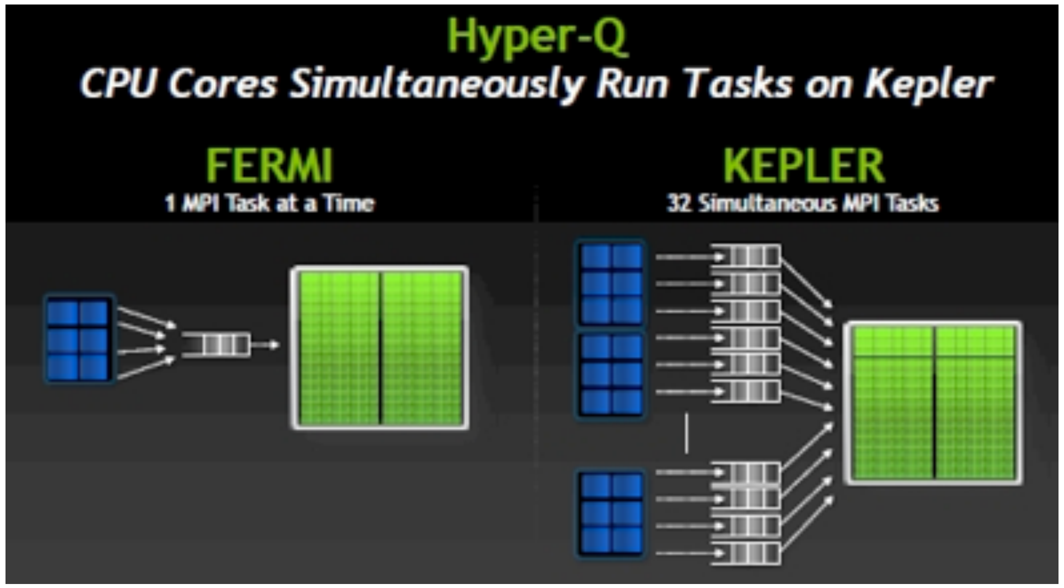
Nvidia Kepler K20X(Telsa)

Intel Xeon Phi Coprocessor

New features of CUDA 5 for its Kepler GPU

Dynamic Parallelism
Hyper-Q
GPU Direct





Top 500 Supercomputers

#	Site	Manufacturer	Computer	Country	Cores	Rmax [Pflops]	Power [MW]
1	Oak Ridge National Laboratory	Cray	Titan Cray XK7, Opteron 16C 2.2GHz, Gemini NVIDIA K20x	USA	560,640	17.6	8.21
2	Lawrence Livermore National Laboratory	IBM	Sequoia BlueGene/Q, Power BQC 16C 1.6GHz, Custom	USA	1,572,864	16.3	7.89
3	RIKEN Advanced Institute for Computational Science	Fujitsu	K Computer SPARC64 VIIIfx 2.0GHz, Tofu Interconnect	Japan	795,024	10.5	12.66
4	Argonne National Laboratory	IBM	Mira BlueGene/Q, Power BQC 16C 1.6GHz, Custom	USA	786,432	8.16	3.95
5	Forschungszentrum Juelich (FZJ)	IBM	JuQUEEN BlueGene/Q, Power BQC 16C 1.6GHz, Custom	Germany	393,216	4.14	1.97
6	Leibniz Rechenzentrum	IBM	SuperMUC iDataPlex DX360M4, Xeon E5 8C 2.7GHz, Infiniband FDR	Germany	147,456	2.90	3.52
7	Texas Advanced Computing Center/UT	Dell	Stampede PowerEdge C8220, Xeon E5 8C 2.7GHz, Intel Xeon Phi	USA	204,900	2.66	
8	National SuperComputer Center in Tianjin	NUDT	Tianhe-1A NUDT TH MPP, Xeon 6C, NVidia, FT-1000 8C	China	186,368	2.57	4.04
9	CINECA	IBM	Fermi BlueGene/Q, Power BQC 16C 1.6GHz, Custom	Italy	163,840	1.73	0.82
10	IBM	IBM	DARPA Trial Subset Power 775, Power7 8C 3.84GHz, Custom	USA	63,360	1.52	3.57

The Green500 List

Green500 Rank	MFLOPS/W	Site*	Computer*	Total Power (kW)
1	**2,499.44**	National Institute for Computational Sciences/University of Tennessee	Beacon - Appro GreenBlade GB824M, Xeon E5-2670 8C 2.600GHz, Infiniband FDR, Intel Xeon Phi 5110P	44.89
2	**2,351.10**	King Abdulaziz City for Science and Technology	SANAM - Adtech ESC4000/FDR G2, Xeon E5-2650 8C 2.000GHz, Infiniband FDR, AMD FirePro S10000	179.15
3	**2,142.77**	DOE/SC/Oak Ridge National Laboratory	Titan - Cray XK7 , Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x	8,209.00
4	**2,121.71**	Swiss Scientific Computing Center (CSCS)	Todi - Cray XK7 , Opteron 6272 16C 2.100GHz, Cray Gemini interconnect, NVIDIA Tesla K20 Kepler	129.00
5	**2,102.12**	Forschungszentrum Juelich (FZJ)	JUQUEEN - BlueGene/Q, Power BQC 16C 1.600GHz, Custom Interconnect	1,970.00
6	**2,101.39**	Southern Ontario Smart Computing Innovation Consortium/University of Toronto	BGQdev - BlueGene/Q, Power BQC 16C 1.600GHz, Custom Interconnect	41.09
7	**2,101.39**	DOE/NNSA/LLNL	rzuseq - BlueGene/Q, Power BQC 16C 1.60GHz, Custom	41.09
8	**2,101.39**	IBM Thomas J. Watson Research Center	BlueGene/Q, Power BQC 16C 1.60GHz, Custom	41.09
9	**2,101.12**	IBM Thomas J. Watson Research Center	BlueGene/Q, Power BQC 16C 1.60 GHz, Custom	82.19
10	**2,101.12**	Ecole Polytechnique Federale de Lausanne	CADMOS BG/Q - BlueGene/Q, Power BQC 16C 1.600GHz, Custom Interconnect	82.19

Heterogeneous Programming Tools

Directives:

OpenACC

- generate CUDA C/Fortran
- Portability
- Scalability(work on multiple GPUs)

```
change = tolerance + 1.0
!$acc data create(newa(1:m,1:n)) copy(a(1:m,1:n))
do while(change > tolerance)
  change = 0
  !$acc kernels reduction(max:change)
  do i = 2, m-1
    do j = 2, n-1
      newa(i,j) = w0*a(i,j) + &
       w1 * (a(i-1,j)+a(i,j-1)+a(i+1,j)+a(i,j+1)) + &
       w2 * (a(i-1,j-1)+a(i-1,j+1)+a(i+1,j-1)+a(i+1,j+1))
      change = max(change,abs(newa(i,j)-a(i,j)))
    enddo
  enddo
  do i = 2, m-1
    do j = 2, n-1
      a(i,j) = newa(i,j)
    enddo
  enddo
 !$acc end kernels
enddo
!$acc end data
```

Heterogeneous Programming Tools

Libs:

MAGMA(just released Xeon Phi coprocessor version)
CUBLAS
CUSPARSE
Thrust(hiding GPU operations)
....
for more libs, please check on *Keeneland*(KFS)

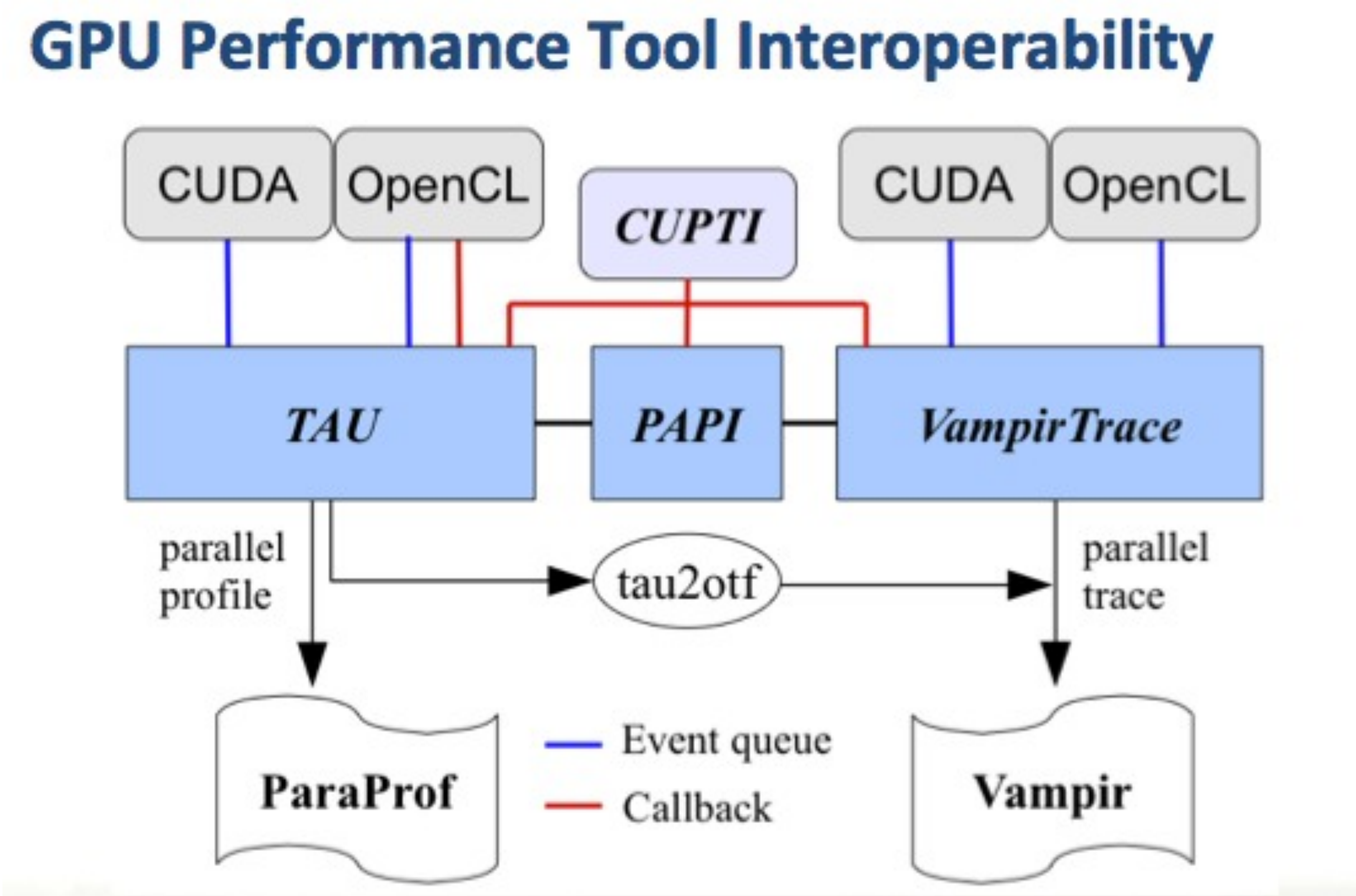
Compiling and debugging tools:

Ocelot(compile with PTX, enable CUDA program executed on different GPU architectures without recompilation)
....
Nvidia Nsight (Eclipse Version contains profiling function!)

Heterogeneous Programming Tools

Profiling Tools:

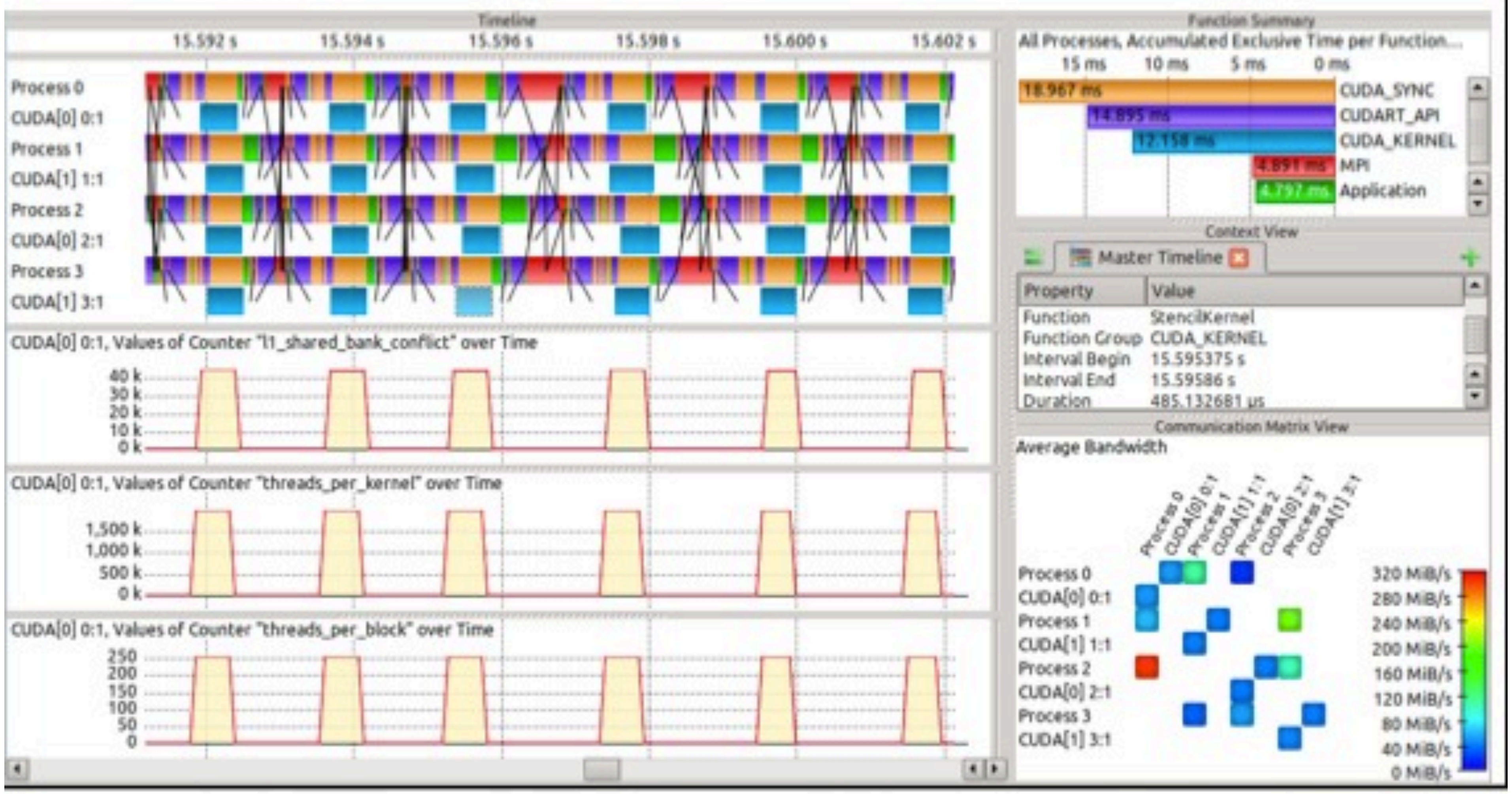
TAU performance system
VampirTrace Measurement and Vampir Analysis
PAPI CUDA
NVIDIA CUPTI



Heterogeneous Programming Tools

Stencil2D Trace (Vampir/VampirTrace)

- Four MPI processes each with one GPU
- VampirTrace measurements



Heterogeneous Programming Tools

Python for HPC

Easy Build
PyTrillinos

Chapel Parallel Programming language

Led by Cray Inc.
Go for productivity and portability
Talks:

Express inherent parallelism without Synchronization
(From Rice Compiler Group)
Auto-tuning by using Chapel tackle issues such as the cache size chunking(Maryland U)

Papers

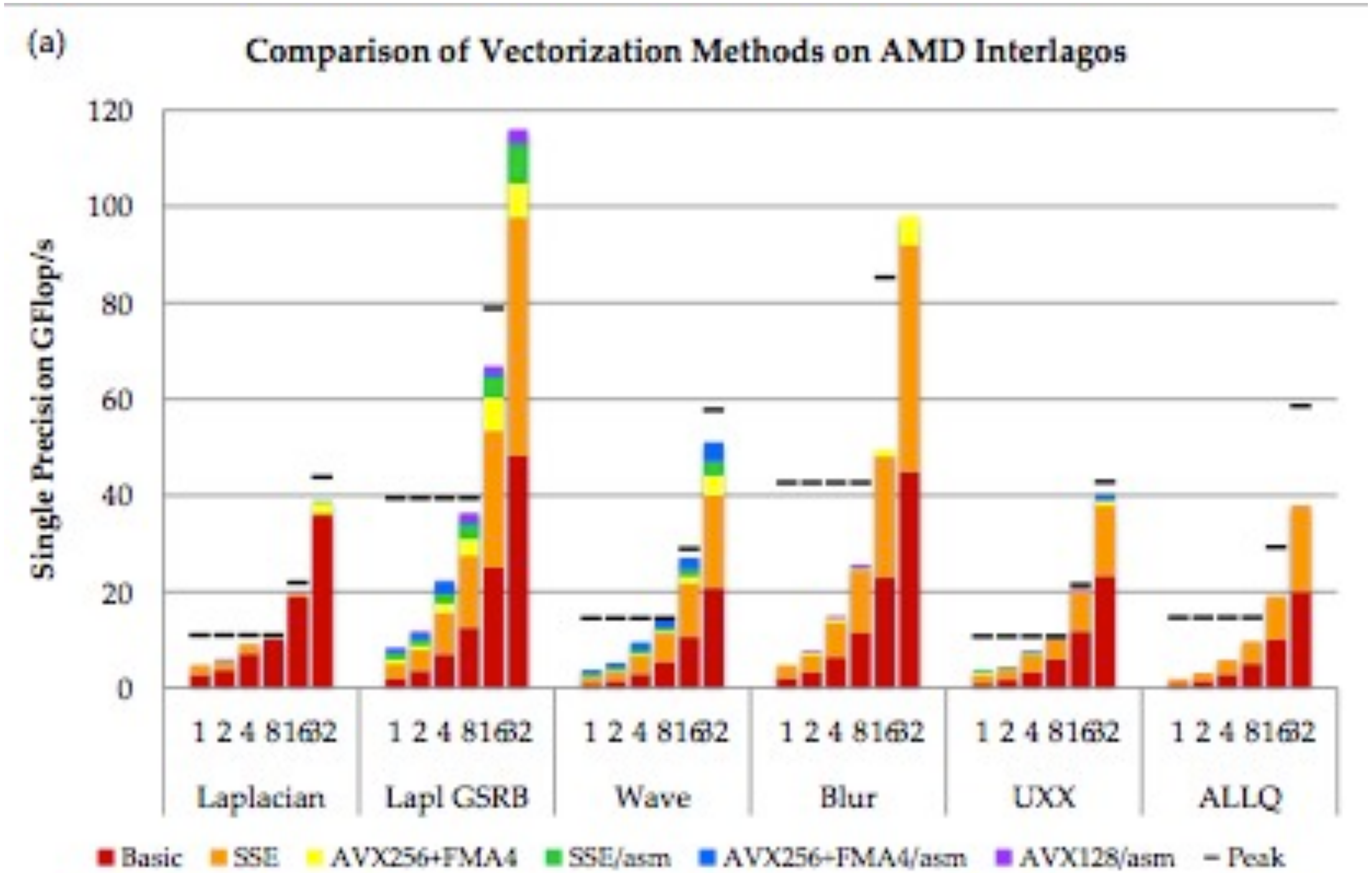
#1. PATUS for Convenient High-Performance Stencils: Evaluation in Earthquake Simulations

PATUS: code generation and auto-tuning frame-work for stencil computations

```
stencil wave {
  domainsize = (2..x_max-3, 2..y_max-3, 2..z_max-3);
  t_max = 1;
  operation (float grid u, float param dt_dx_sq) {
    float c1 = 2 - 15/2*dt_dx_sq;
    float c2 = 4/3*dt_dx_sq; float c3 = -1/12*dt_dx_sq;
    u[x,y,z; t+1] = c1*u[x,y,z; t] - u[x,y,z; t-1] +
      c2*(u[x+1,y,z; t] + u[x-1,y,z; t] + ...)+ //etc.for y,z
      c3*(u[x+2,y,z; t] + u[x-2,y,z; t] + ...);
  }
}
```

It utilized optimization techniques such as loop unrolling, SSE intrinsics. They pointed out that **vectorization** is the key optimization step.

#1. PATUS for Convenient High-Performance Stencils: Evaluation in Earthquake Simulations



#2. Tiling Stencil Computations to Maximize Parallelism

They devised a way for finding the optimal tiling plane to ensure concurrent start-up as well as perfect load balance.

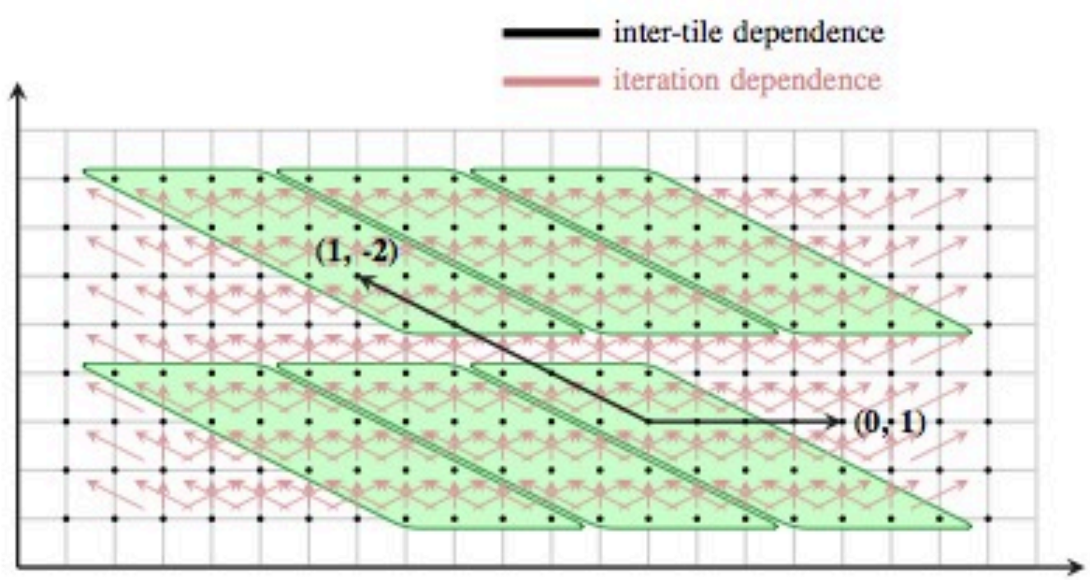


Fig. 6. Pipelined start-up

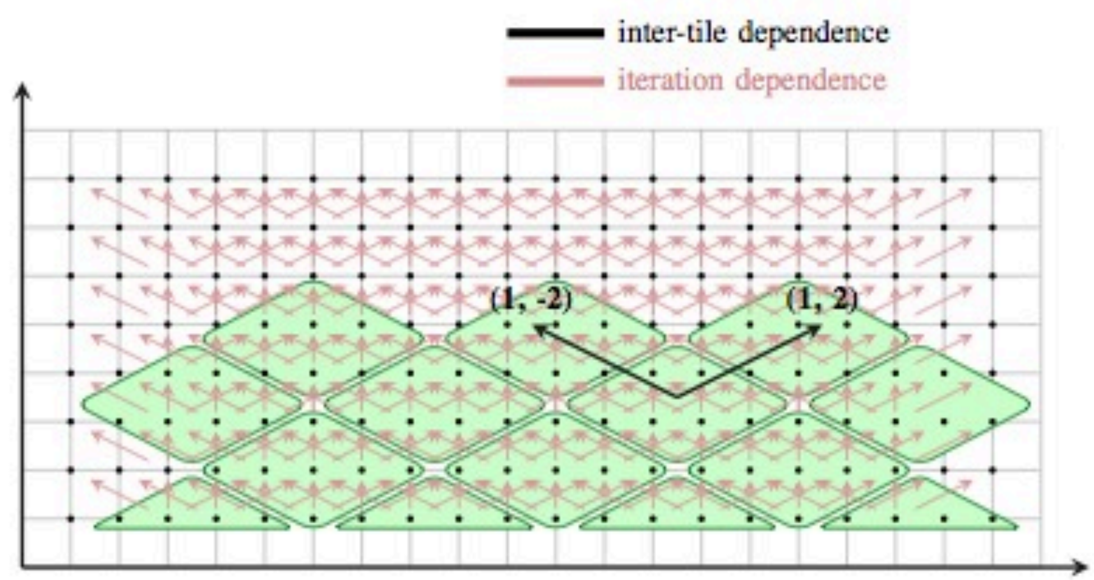
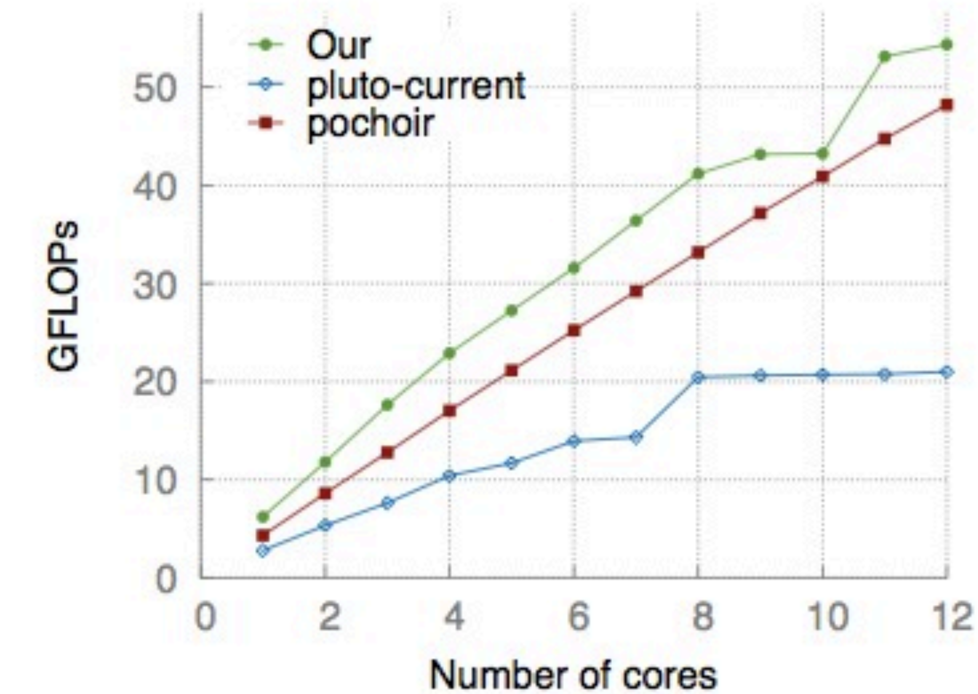
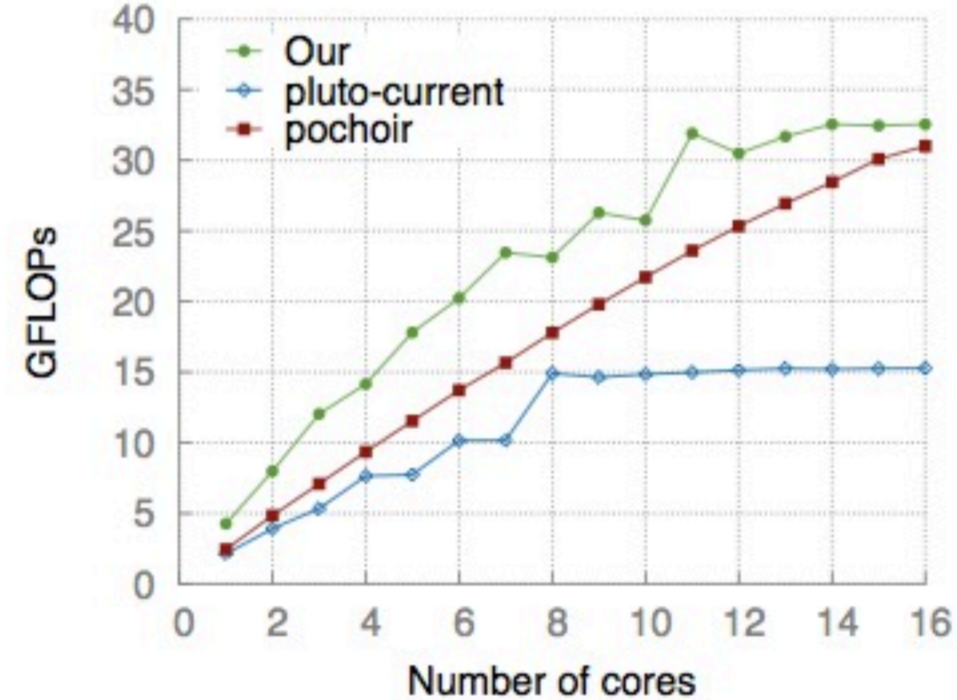


Fig. 7. Concurrent start-up

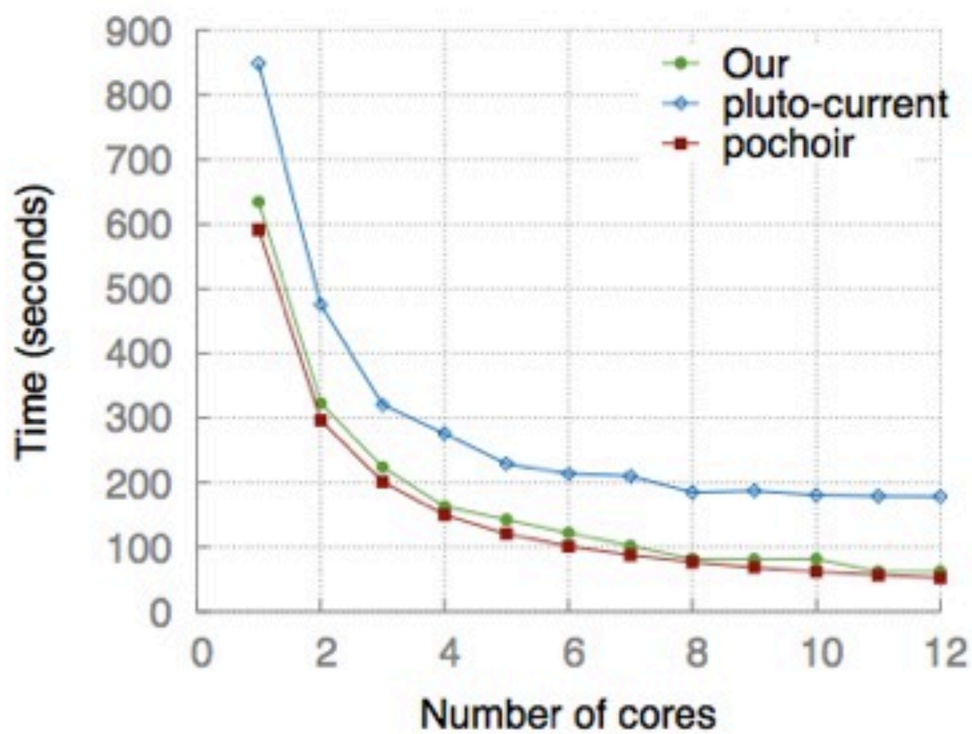
#2. Tiling Stencil Computations to Maximize Parallelism

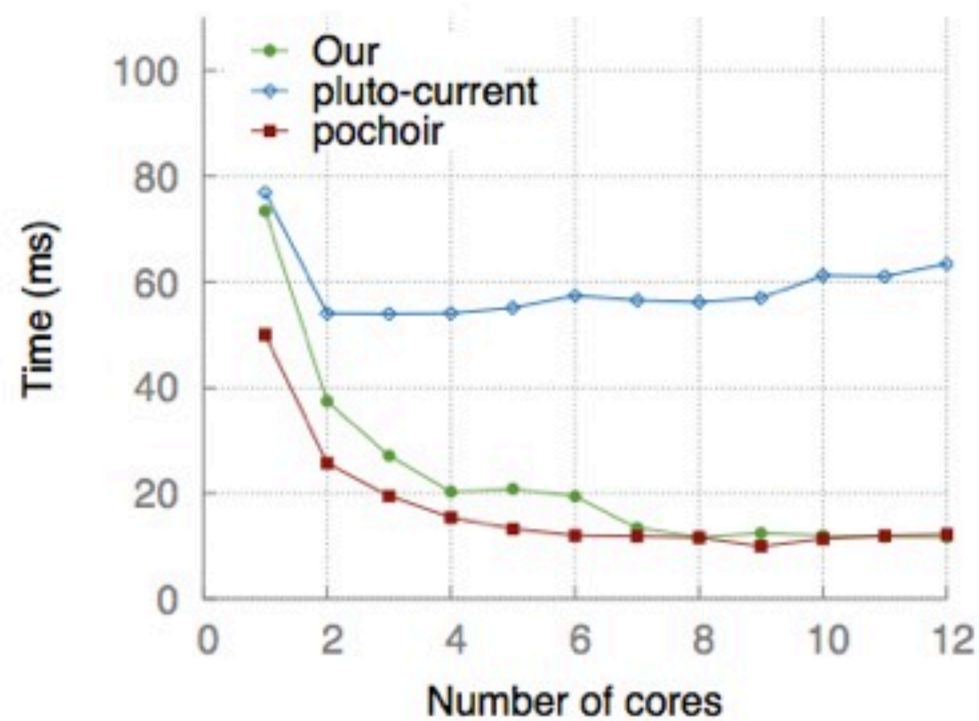


(a) Heat 2D (Intel)



(b) Heat 2D (AMD Opteron)





Thanks!