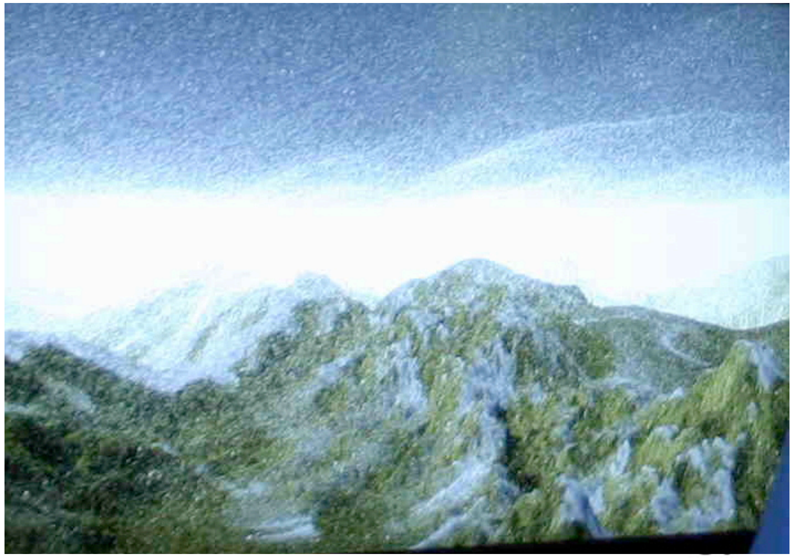


Real-Time Parallel Computing Using GPUs

Anne C. Elster & her students
Dept. of Computer & Info. Science
NTNU (Norwegian Univ. of Sci.&Tech.)

Wroclaw Cathedral

Taken by Elster
on Sep. 14, 2008

NTNU
Norwegian University of
Science and Technology

PARA 2008 & 2010

Para 2008 Proceedings hopefully in by end of year

Most extended abstracts now on web

PARA 2010 will be held June 6-9
in REYKJAVIK, ICELAND

NTNU
Norwegian University of Science and Technology

ParCo-related Activities & Events:

EU COST Action IC0805: Open European Network for High Performance Computing on Complex Environments (2009-2014)

Meeting in Lisboa, mid-October. See: www.complexhpc.org

Talk to Elster re. WG on Numerical Algorithms

NTNU
Norwegian University of Science and Technology

HPC History: Personal perspective

- 1980's: Concurrent and Parallel Pascal
- 1986: Intel iPSC Hypercube
  - CMI (Bergen) and Cornell (Cray at NTNU)
- 1987: Cluster of 4 IBM 3090s
- 1988-91: Intel hypercubes
  - Some on BBN
- 1991-94: KSR (MPI1 & 2)
- 1995 -2005: SGI systems (some IBM SP)
- 2001-current: Clusters
- 2006:
  - IBM Supercomputer @ NTNU (Njord, 7+ TFLOPS, proprietory switch)
  - GPU programming (Cg)
- 2008:
  - Quadcore Supercomputer at UiTø (Stallo)
  - HPC-LAB at IDI/NTNU opens with
    - several NVIDIA donation
    - Several quad-core machines (1-2 donated by Schlumberger)
- 2009: More NVIDIA donations:
  - NVIDIA Tesla s1070 and
  - two Quadro FX 5800 cards (Jan ´09)

NTNU
Norwegian University of Science and Technology

The Wal-Mart Effect
(PARA02)

- Wal-Mart – bigger than Sears, K-mart and JC Penney's combined
  ➔ predicted to influence $40 billion of IT investments (MIT Review)
  ➔ has much more impact than Microsoft and Cisco could ever hope for…

  – Not driven my latest technology, but by business model
    – bad news for HPC?

  – Game market --> HPC market ➔ Future high-performance chips and systems --> NVIDIA Tesla!

NTNU
Norwegian University of Science and Technology

”COT”-based SUPERCOMPUTER HARDWARE TRENDS:

- Intel iPSC (mid-1980’s)
  - The first iPSC had no separate communication processor ...
  - Specialized OS
  - 2-128 nodes

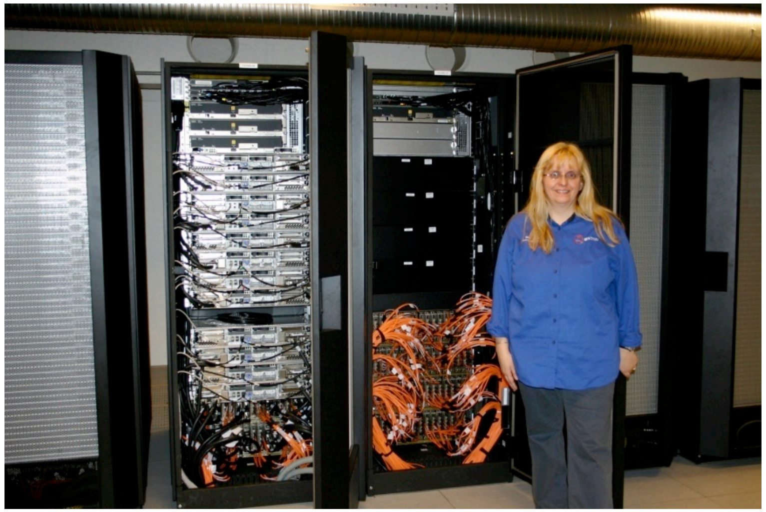
- Today’s PC clusters
  - Fast Ethernet or better (more expensive interconnect)
  - Linux OS
  - 32-bit cheapest, but many 64-bit cluster vendors ☺
  - Top500 supercomputers

Today´s GPU farms entering Top500 list..

Clustis 1 ca. 2003

SGI Griddur/Embla

SGI Altix



HPC Hardware Trends at NTNU/IDI

NTNU
Norwegian University of Science and Technology

NTNU´s Supercomputer ("Njord") from IBM

- Also runs Norway´s operational weather forcasts (met.no)

- NOK 30 millions for system
- NOK 20 million for infrastructure
    - incl. new machine room, back-up power generator & batteries ++

Jørn Amundsen, Roar Skålin (it-sjef met.no) og Bjørn Lindi ved Njord

Digging of hole for 2000 hp (horsepower) back-up power generator

- **Comparisons:**
    - **Our 1986-Cray cost NOK 130 millon in today´s currency!**
    - **NVIDA s1070 Tesla w/ 960 cores costs ca NOK 100 000! (4TFlops)**

NTNU
Norwegian University of Science and Technology

HPC Hardware Trends at IDI

NVIDA Tesla card

Unpacking NVIDA s1070 and Quadro FX 5800 cards

NTNU
Norwegian University of Science and Technology

GPUs: Graphical Processor Units

HISTORY:

- Late 70's/ Early´80's: Grafic drawing calculations on CPUs
- Xerox Alto computer: first special *bit block transfer* instruct
- Comodore Amiga: first mass-market video accelerator able to draw fills shapes & animations in HW. Graphics sub-system w/ several chips, incl. Dedicated to *bit blk xfer*
- Early 90's: 2D accelleration
- **Ca. 1995: VIDEO GAMES! -->** 3D GPUs

NTNU
Norwegian University of Science and Technology

GPU History continued:

- 1995-1998:
  - 3D rasterization
    (converting simple 3D geometric primitives
    (e.g. lines, triangles, rectangles) to 2D screen pixels)
  - *Texture mapping*
    (mapping 2D texture image to planar 3D surface)

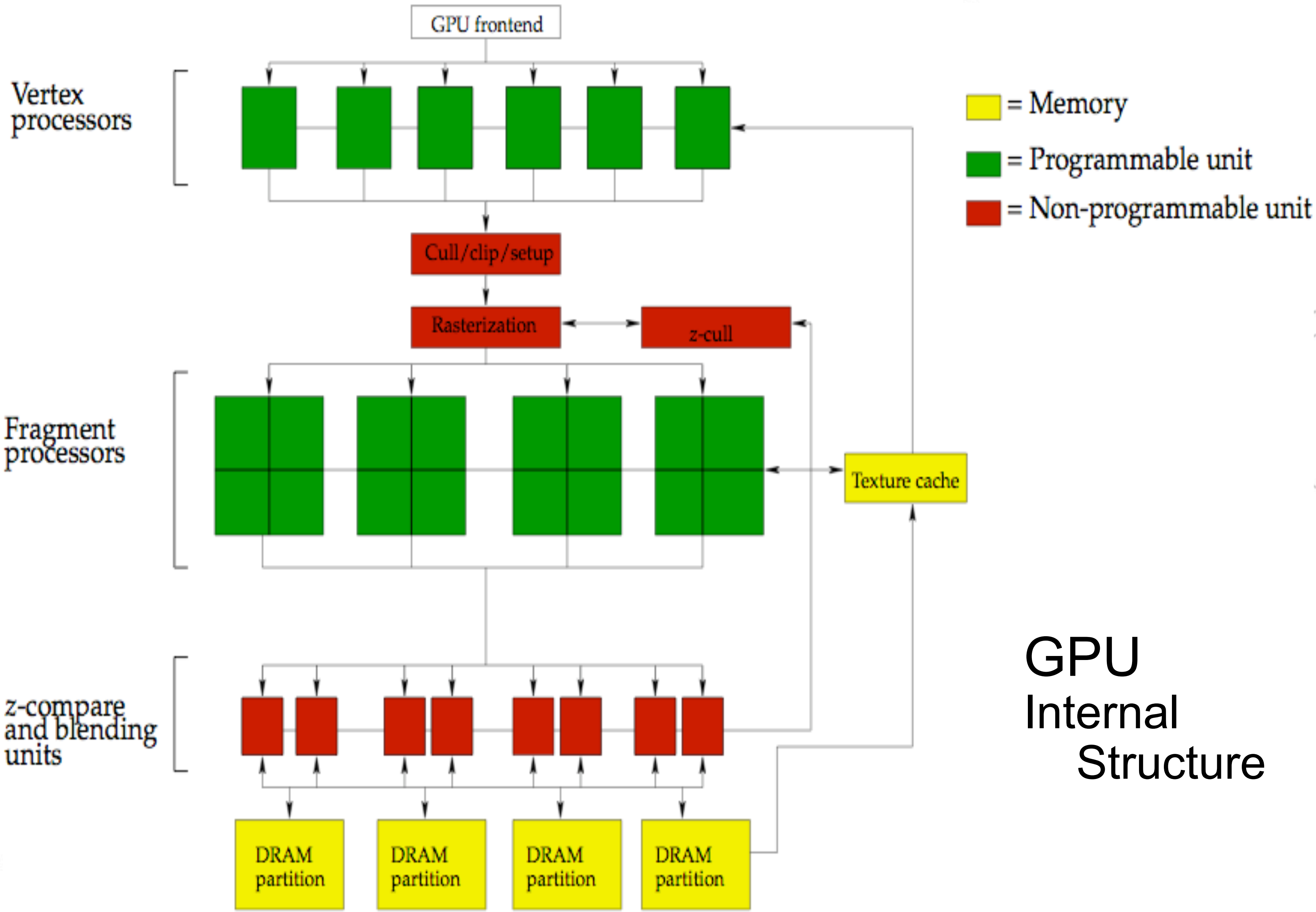
- 1999-2000: 3D translation, rotation & scaling

- Towards 2000: GPUs more configurable

- 2001-2007: **programmable**
  (ability to change individual pixels)

- 2008 and beyond: **more programmable**
  (NVIDIA CUDA, OpenCL ...)

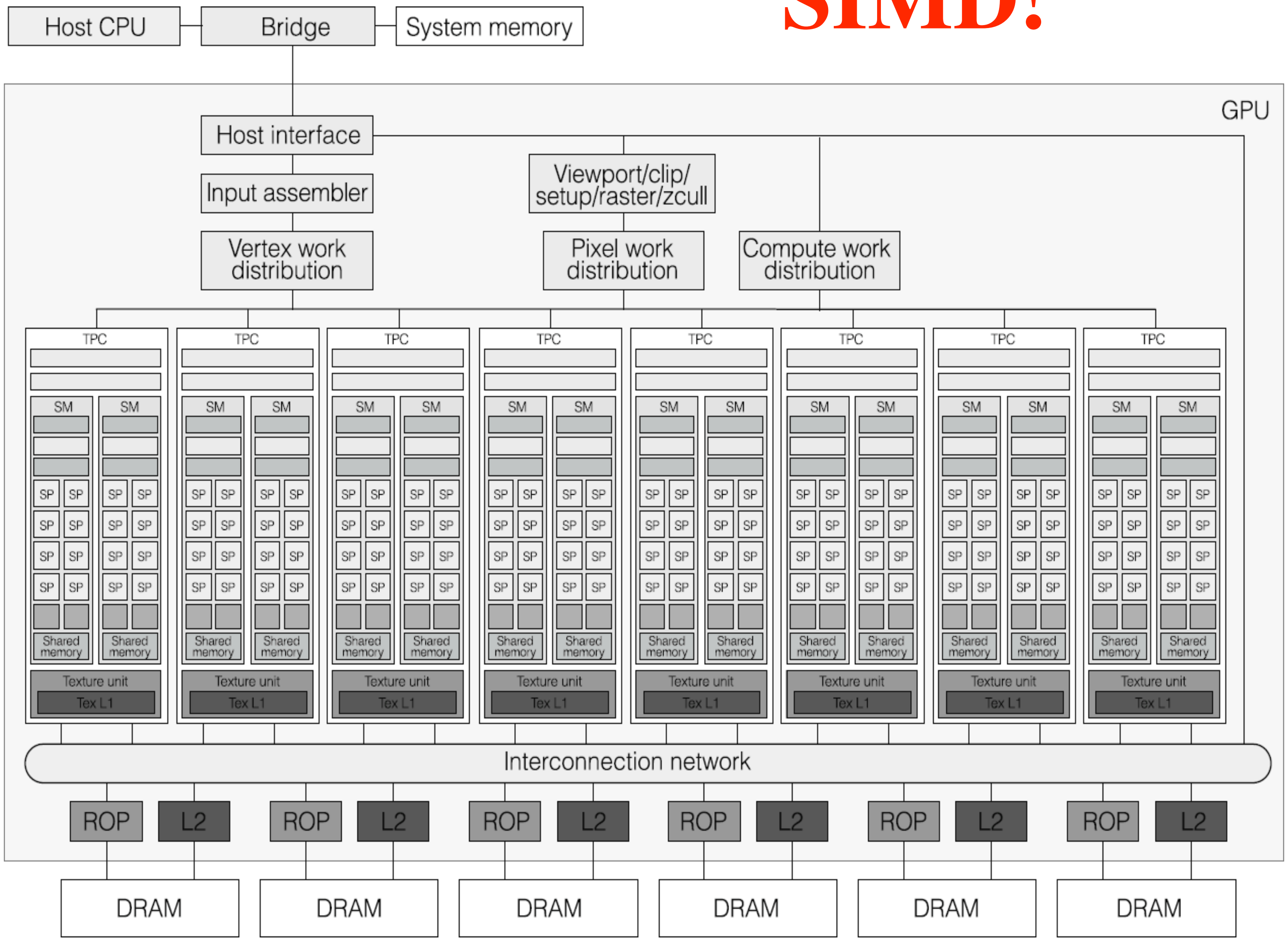
NTNU
Norwegian University of Science and Technology



GPU Internal Structure

The Nvidia Tesla Architecture

SIMD!



General programming on GPUs

- Rendering = executing
- GPU textures = CPU arrays
- Fragment shader programs = inner loops
- Rendering to texture memory = feedback
- Vertex coordinates = computational range
- Texture coordinates = Computational domain

- Now have NVIDIA´s CUDA library! (BLAS & FFT)

Limitations

- **Branching usually not a good idea**
- Random memory access problematic
- GPU cache is different from CPU cache
  - Optimized for 2D locality

Floating point precision

NTNU
Norwegian University of Science and Technology

Check out: http://gpgpu.org/developer/ppam2009

Session 1: GPU Basics
・Introduction (Strzodka) /\(PDF)
・Why GPUs? (Strzodka) /\(PDF)
・Prog. Environments and Ready-to-use Libraries (Goedeke) /\(PDF)
・GPU Architecture (Strzodka) /\(PDF)

Session 2: Introduction to OpenCL
・Introduction to OpenCL (Behr) /\(PDF)
・Hands-on Examples

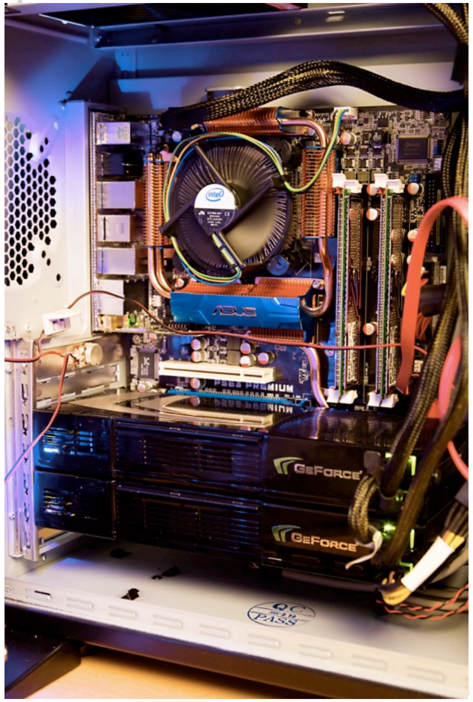
Session 3: OpenCL Basics
・OpenCL API (Behr) /\(PDF)
・OpenCL Language (Behr) /\(PDF)
・Hands-on Examples

Session 4: Scientific Computing on GPUs
・Aspects of Scientific Computing on GPUs (Strzodka)/\(PDF)
・Case Study: GPU Cluster Computing for FEM (Goedeke)/\(PDF)

Session 5: Advanced OpenCL
・OpenCL Architecture and Optimization on AMD GPUs (Behr)/\(PDF)
・Hands-on Examples
・AMD OpenCL GPU Demo

NTNU
Norwegian University of Science and Technology

HPC-Lab at IDI



NTNU
Norwegian University of
Science and Technology

OpenCL

Olav Fagerlund

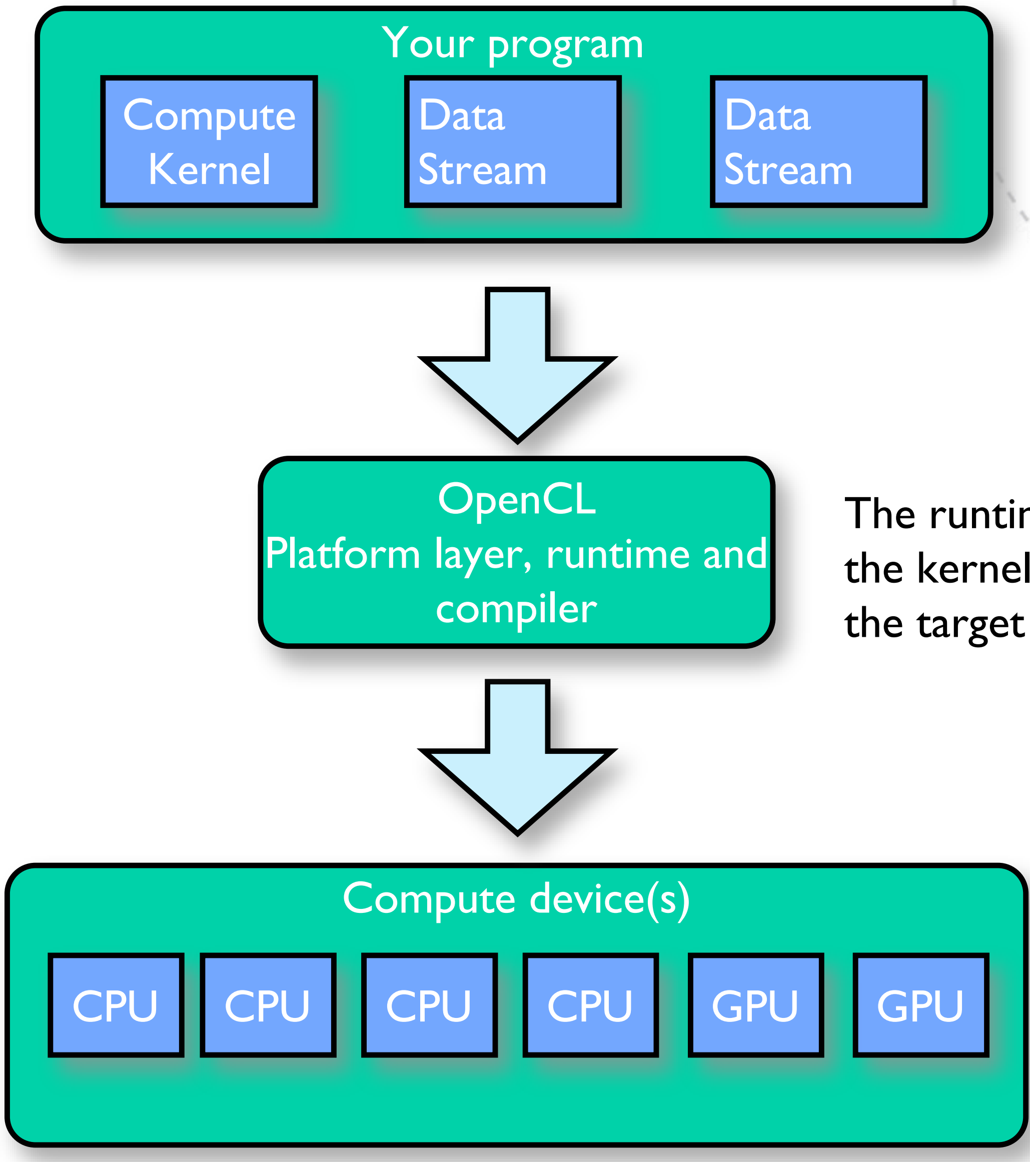
Goals:

- leverage all computational resources in the system.
- unleash computational power to everyone;
- to be integrated with ordinary commercial applications. Example: next gen. games.

- For any data-parallel algorithm.
- Uses GPU, CPU, or both as compute device. Or any combination of GPUs and CPUs. Designed for heterogeneous parallel data computation.
- Simple and clean API.
- A Khronos group standard. Royalty free.
- Future: Support for multiple devices on multiple platforms. Implementations on the way from AMD, NVIDIA and Apple.
- Platform + device independence in focus. Must be in order to become successful.

NTNU
Norwegian University of Science and Technology

OpenGL Overview



The runtime compiles the kernel, optimizes for the target device(s).

NTNU
Norwegian University of Science and Technology

Recent Activities & Events:

IEEE International Parallel & Distributed Processing Symposium

MTAAP'09 (Friday, May 29)

Workshop on Multithreaded Architectures and Applications

Jan Christian Meyer
PhD Student

Two presentations from HPC-Lab:

- Jan Chr. Meyer and Anne C. Elster: Super-Fast Adaptable Bit-Reversal on Multithreaded Architectures

- Daniele Spampinato and Anne C. Elster: Simplex-Based Linear Optimization on Multithreaded Architectures (Linear Programming on GPU!)

Daniele Spampinato
Future PhD Student?

NTNU
Norwegian University of Science and Technology

EuroGPU´09 Organizers

Anne C. Elster, NTNU, Norway

and

Stephane Requena, Genci, Paris, France
with

Guillaume Colin de Verdière, CEA

NTNU
Norwegian University of Science and Technology

EuroGPU at ParCo 2009:

TUESDAY Sept. 1, 2009 Euro GPU 2009 – DAY 1

Time	*Title/Speaker*
11:30-12:00	**Intro. – GPU Computing**, *Anne C. Elster, Norwegian University of Science and Technology (NTNU), Trondheim, Norwa y*
12:00-12:30	**Throughput Computing on Future GPUs** *Rune Johan HOVLAND and Anne C. ELSTER, NTNU, Norway*
12:30-14:00	LUNCH
14:00-14:30	I1: **OpenCL a new standard for GPU programming** *Francois BODIN -- Caps Entreprise, Rennes, France*
14:30-15:00	I2: **Heterogeneous Multicore Parallel Programmi ng** *Stéphane BIHAN -- Caps Entreprise, Rennes, France*
15:00-15:30	I3: **Cosmological reionisation powered by multi-GPUs** *Dominique AUBERT[a], Romain TEYSSIER[b]* *[a] Université de Strasbourg, France [b]CEA, France*
15:30-16:00	Coffee/Tea Break & Exhibitions
16:00-16:30	I3: **Efficient use of hybrid computing clusters for nanosciences** *Lugi GENOVESE[a], Matthieu OSPICI[b], Jean Francois MÉHAUT[c], Thierry DEUTSCH[d]* *[a] ESFR, Grenoble, [b] BULL, UJF/LIG, CEA, Grenoble* *[c] UJF/INRIA, Grenoble, [d] CEA, Grenoble, France*
16:30-17:00	I4: **Accelerating depth imaging seismic application on GPUs, status and perspectives,** *Henri CALANDRA, TOTAL, Pa u*
17:00-17:30	I5: **Debugging for GPUs with DDT** *David LECOMBER Allinea Ltd, Bristol, UK*
19:00-20:00	Reception: City Hall

NTNU
Norwegian University of Science and Technology

WEDNESDAY Sept. 2, 2009 Euro GPU 2009 – DAY 2

Time	*Title/Speaker*
10:00-10:30	**Porus Rock Simulations and Lattice Boltzmann on GPUs** *Erik Ola AKSNES and Anne C. ELSTER, NTNU, Norway*
10:30-11:00	**An efficient multi-algorithms sparse linear solver for GPUs** *Stéphane VIALLE; Thomas JOST and Sylvain CONSTASSOT-VIVIER, Supélec Campus de Metz, France*
11:00-11:30	Coffee/Tea Break & Exhibitions
11:30-12:00	**Abstraction of Programming Models Across Multi-Core and GPGPU Architectures,** *Ian GRIMSTEAD and David R. WALKER, Cardiff University, UK*
12:00-12:30	**Modeling Communication on Modern GPU Systems**, Anne C. ELSTER, Thorvald NATVIG, and Daniele G. SPAMPINATO, NTNU, Norway
12:30-14:00	LUNCH
14:00-15:00	PANEL DISCUSSION – ParCo 2009
15:00-15:30	Coffee/Tea Break & Exhibitions
15:30-16:30	**PANEL DISCUSSION on GPU Computing** Including informal presentation by **Rune J ensen NTNU/CERN on Compiler Issues and Challenges** **PANEL:** - **Anne C. Elster (Organizer)** – GPU scientific computing/academia - **G u i l lame Coline Vedriére, CEA, France** – GPU Applications - **T i m Lanfear, Nvidia, UK** – GPU HW vendor - **Stéphane Vialle, Supelec Metz, France** – GPU financial computing/ academia
	Excursion & Conference Dinner

NTNU
Norwegian University of Science and Technology

Rune Jensen: Optimizing BLAS

- Beat ATLAS, an auto-tunable BLAS library routine
- ATLAS
  - Self-tuning at install
  - Compiles lots of code versions
  - Compares their speed
  - Find patterns
  - Makes the best code for your CPU/memory/mainboard
  - Cores handwritten in assembly

  - Now moving to multi-core. How about GPUs...

NTNU
Norwegian University of Science and Technology

Courses Taught by Dr. Elster:

.

TDT4200 Parallel Computing
(Parallel programming with MPI, threads and NVIDIA CUDA)

TDT24 Parallell environments & Numerical Computing

- **2-day IBM CellBE Course (Fall 2007)**
- **GPU & Thread programming**

TDT 4205 Compilers

DTD 8117 Grid and Heterogeneous Computing

NTNU
Norwegian University of Science and Technology

HPC -LAB Sponsors/Collaborators

- NVIDIA

- CERN (birthplace of Internet; EU´s largest GRID project;
  Norwegian CTO (Sverre Jarp) :-)
  - 7 Master students from my group have ahd summer jobs there
  - 4 took their Master´s thesis there
  - 1 still a staff member there, 2 will join this summer

- GE Healthcare
  Our GPU wavelet algorithm now in their high-end cardiac ultrasound scanner!

- SCALI (Commercial MPI implementer)

- Schlumberger (formerly Voxel Vision)

- StatoilHydro

- several other departments at NTNU
  including Petroleum, Physics and Chemistry

NTNU
Norwegian University of Science and Technology

The IDI HPC-Lab focuses on research related to novel GPU and multi-core architectures

- GPGPU for HPC
- Parallel and Distributed Algorithms
- Performance Evaluation and Benchmarking
- Parallelization of Seismic and Image Related Applications on GPUs and Multi-Cores
- Adaptive and Auto-Tuneable Algorithms and Implementations

NTNU

Collaborators / Supporters:

ARM, CERN, NVIDIA, StatoilHydro, Schlumberger, GE-Healthcare, and others

NTNU
Norwegian University of Science and Technology

HPC-Lab

Lab Director

Dr. Anne C. Elster

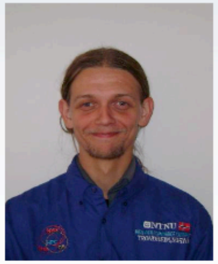
Master Students

Robin Eidissen

Rune E. Jensen

PhD Students

Thorvald Natvig

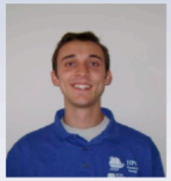


Jan Christian Meyer

Olav Fagerlund

Åsmund Eldhuset

Daniel Haugen



Daniele G. Spampinato

Eirik O. Aksnes

Henrik Hesland

Åsmund Herikstad

Owe Johansen

Rune Hovland

Safrudin Mahic Hon. Mbr.

HPC Research Group - Spring 2008

GROUP MEMBERS:

- Assoc. Prof. Anne C. Elster
- Adjunct Assoc. Prof-Jørn Amundsen
- Henrik Nagel (PostDoc/now at NTNU HPC-Ctr)

- 2 PhD students
  - Jan Christian Meyer
  - Thorvald Natvig

- Recent Master students (grad. dates)
  - Øystein Borgen(June´06) - Schlumberger
  - Ingar Saltvik (June´06) - Fast
  - Nils Magnus Larsgård (Aug´06) - IBM
  - Erik Axel Nielsen (May´07) - consultant
  - Idar Borlaug (June´07) - StatoilHydro cons
  - Knut Imar Hagen (June´07) - --" --
  - Leif Christian Larsen (June´07) - Roxar
  - Jerôme Dubois (Feb ´08) - back in France
  - Andreas Bach (su ´08) - Uninett
  - Atle Rudshaug (su´08) - Numerical Rocks
  - Robin Eidissen (Feb. ´09) - IDI
  - Rune Jensen (May´09) - CERN

NTNU
Norwegian University of Science and Technology

High-Performance Computing Group

Alumni

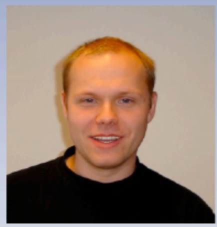
HPC Research Group

NTNU Norwegian University of Science and Technology

Master Student Alumni

Andreas Bach (2008)

Atle Rudshaug (2008)



Idar Borlaug (2007)

Thibault Collet (2007)

Knut Imar Hagen (2007)

Nils Magnus Larsgård (2007)

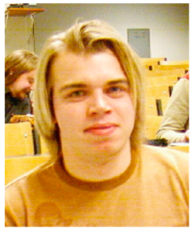
Erik Axel Nielsen (2007)

Christian Larsen (2007)

Øystein Borg (2006)

Ingar Saltvik (2006)


Håvard Bjerke (2005)


Andreas Braathen (2005)


Rune Johan Andresen (2005)


Snorre Boasson (2004)


Glenn Hisdal (2004)


Einar Råberg Rosenvinge (2004)


Tor Arvid Lund (2004)


Frode Nilsen (2004)


Morten Rodal (2004)


Torbjørn Vik (2003)


Åsmund Østvold (2003)


Robin Holtet (2003)

Others

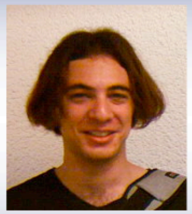
Post Doc


Henrik R. Nagel (2005-2007)

MS Project


Jostein Tveit (2004)

Summer Project


Paul Sack (2002)

NTNU
Norwegian University of Science and Technology

HPC-Lab Fall´09

Dr. Anne C. Elster
Lab Director

Dr. Jørn Amundsen
Adjuct Assoc Prof

TBA
Contact
me!

Post Doc
TBA

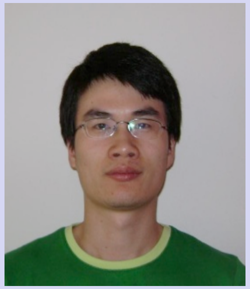
Jan Christian
Meyer (PhD stud)

Thorvald Natvig
(PhD stud.)

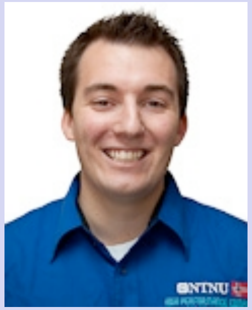
Master Students – Fall 2009

Ahmed Aqrawi
Assist. TDT 4200

Aleksander
Gjermundsen



Gaojie He



Øystein Krog

Holger Ludvigsen
Assist TDT 4200

Affiliates /Visitors – Fall 2009

Runar Refsnæs
(Math)

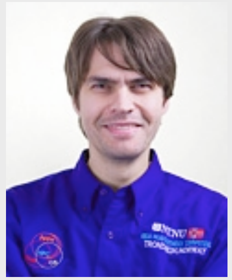
Gagandeep
Singh (Math)

Roald Fernandez
(Cybernetics)

Peter Sveistrup
(Cybernetics)

+ 3 visualization students
+ 1-2 || arch/ multicore students

Eirik O. Aksnes
(tentative PhD



Rune E. Jensen
(tentaive) PhD

Daniele G.
Spampinato
(tentative PhD

HPC-LAB Spring/Summer 2009 Master projects

- 1 worked on real-time snow simulations
- 1 works on “beating Atlas”
- 2 work with IO Center on flow through porous media on GPU
- 1 work with Schlumberger on line finding algorithms on GPU
- 1 works on GPU-CPU system configurations
- 1 work with Statoil-Hydro on check-pt restart of large applications
- 1 on sound processing on GPU
- 1 on LP on GPU
- 1 on special transforms for GPU

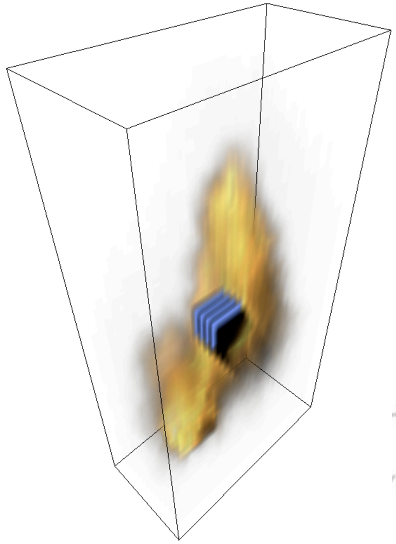
NTNU
Norwegian University of Science and Technology

Smoke & Snow
Particle Simulations

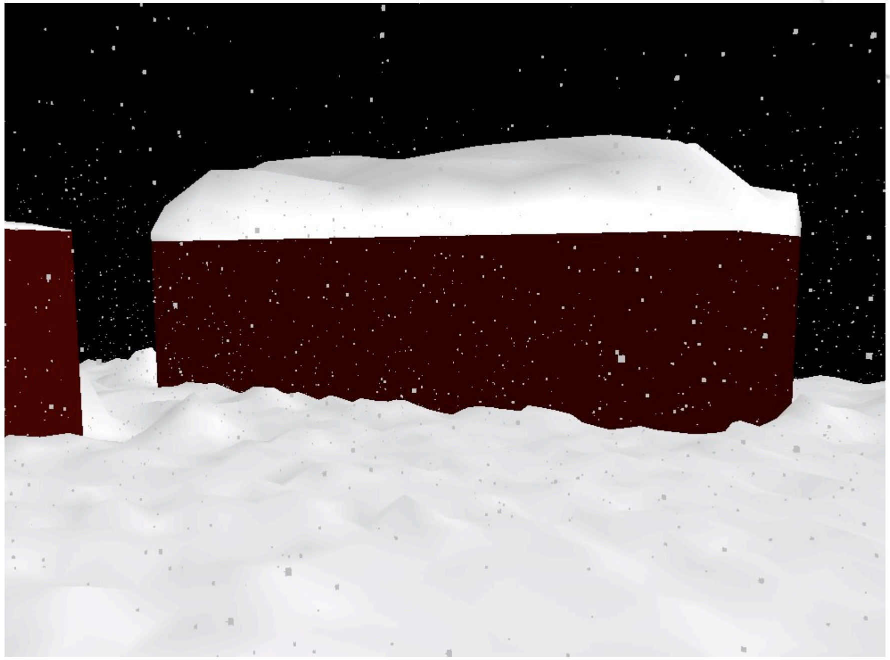
Started out as Real-time smoke simualtion on dual core lap top (Torbjørn Vik, 2003)

Crude snow simulation on Multicore (Ingar Saltvik, 2007)

Snow simulation simulating several million snow flakes as Particles, widn field interactions++ Using compute-power of the GPU (Robin Eidissen, 2009)

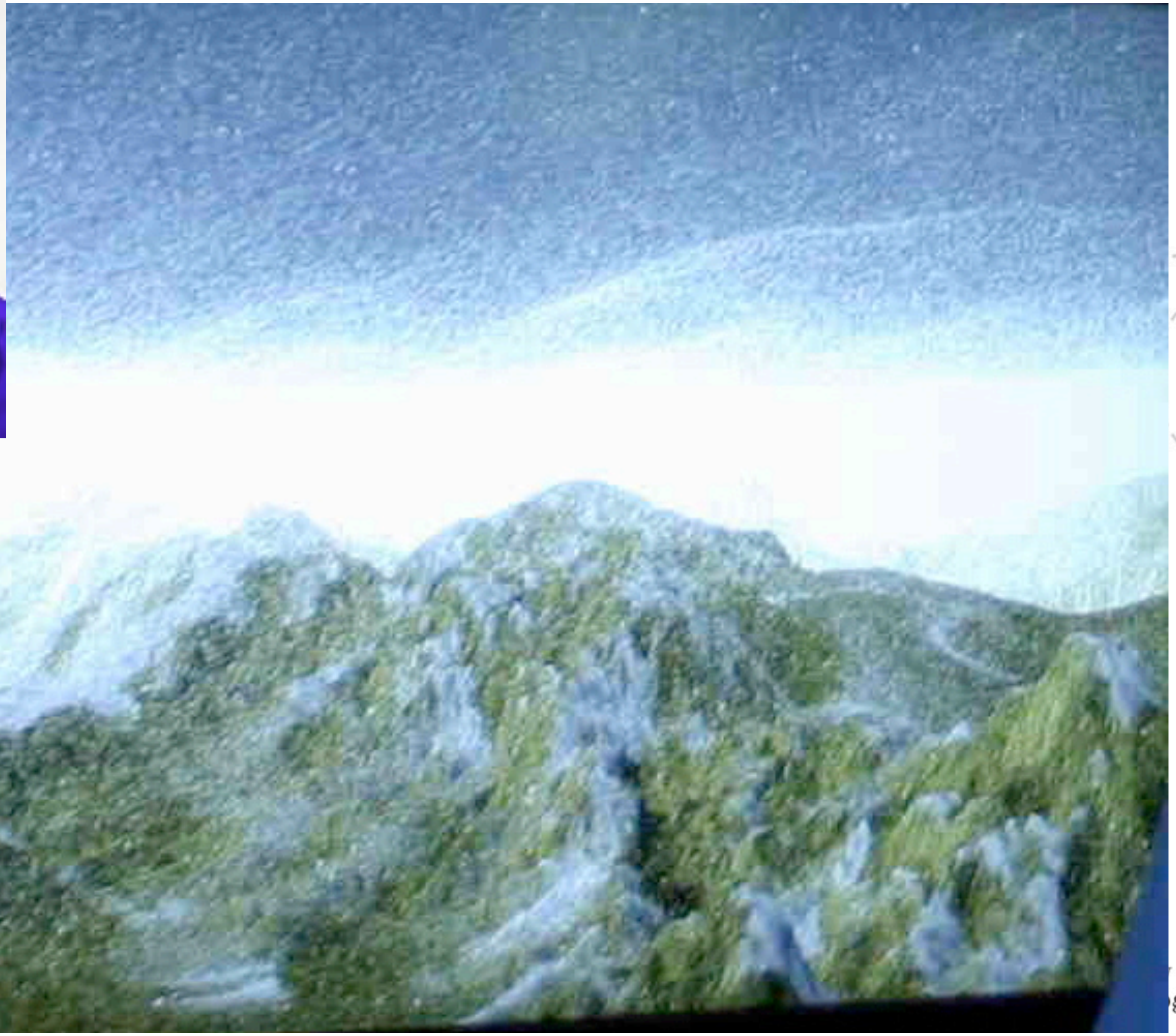


Paralellized
Snow &
Smoke
Simulations



SNOW SIMULATION DEMO!

Robin Eidissen
(Teaching Assitant)

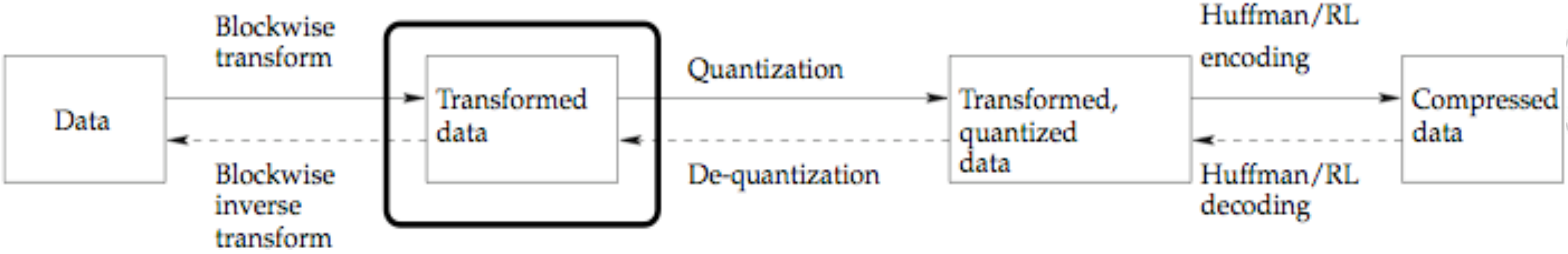



Nara In
Snow simulation
Using NVIDIA CUDA

Åsmund Eldhuset: Line finding algorithms (in collaborations with Schlumberger)

NTNU
Norwegian University of Science and Technology

DCT Compression



NTNU
Norwegian University of Science and Technology

See also Daniel Haugen´s NOTUR 2009 Poster:

“Strategies for Handling Large Amounts of Data from Storage to GPUs”

(also in collaborations with Tore Fevang, Schlumberger)

Daniel Haugen

NTNU
Norwegian University of Science and Technology

Some of the many other applications

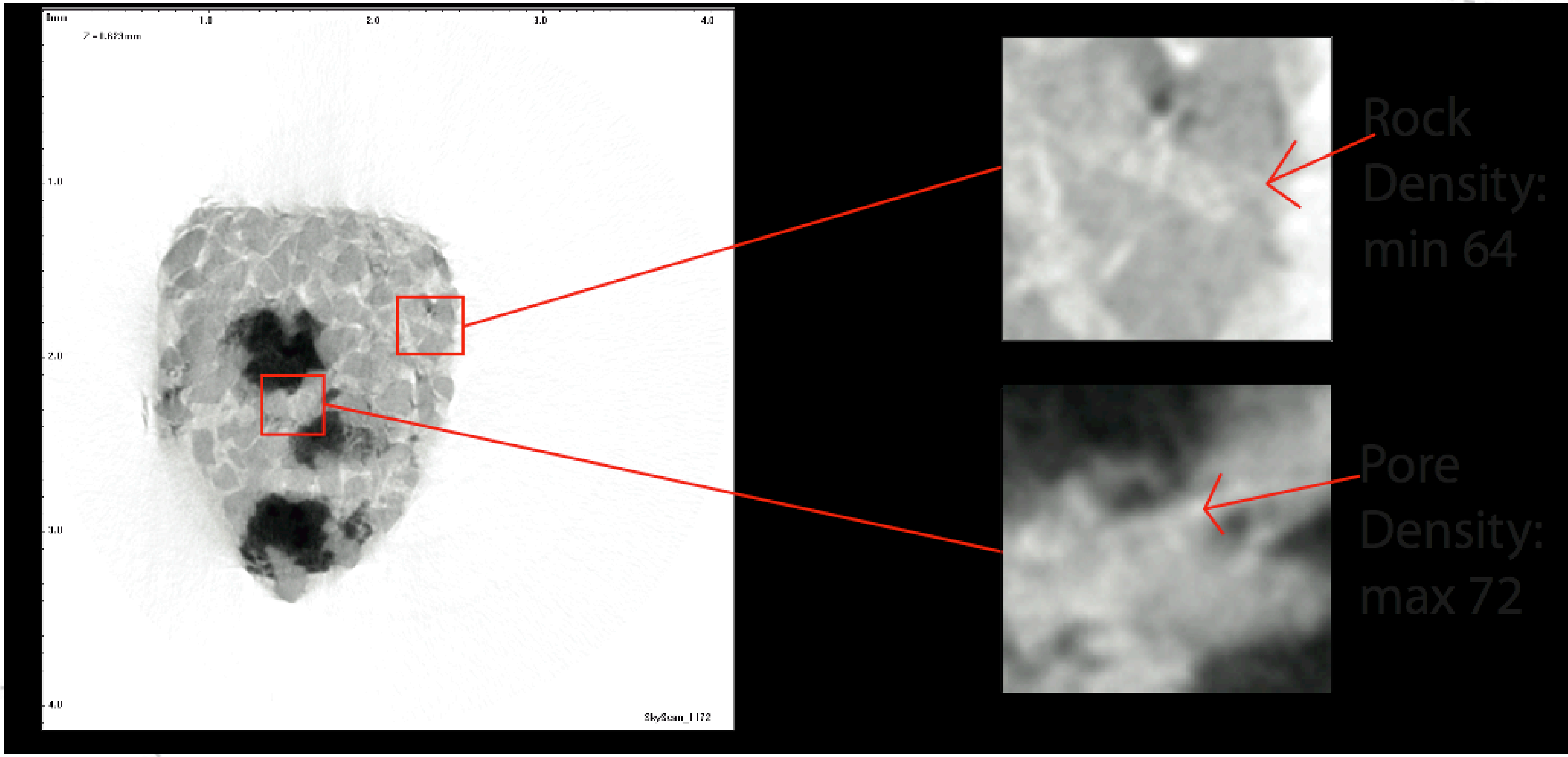
- Medicine:
  - ultrasound imaging
  - imaging of vessels before surgery
- Chemistry
  - molycular analysis and simulation
  - CFD
- Physics
  - particle simulations

- Marine technology
- Mathematical methods
- Computer algorithms, benchmarking ...
- ...

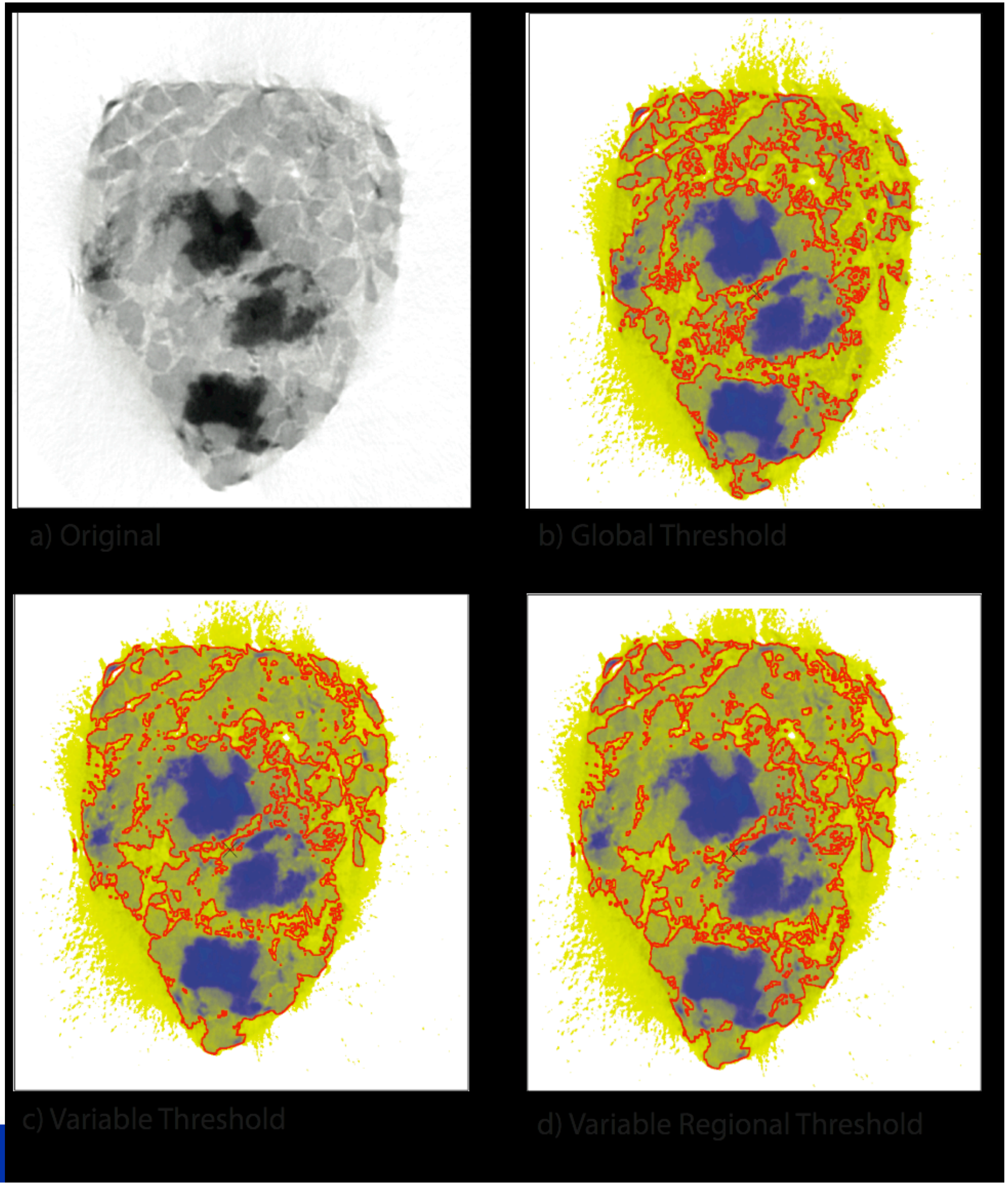
NTNU
Norwegian University of Science and Technology

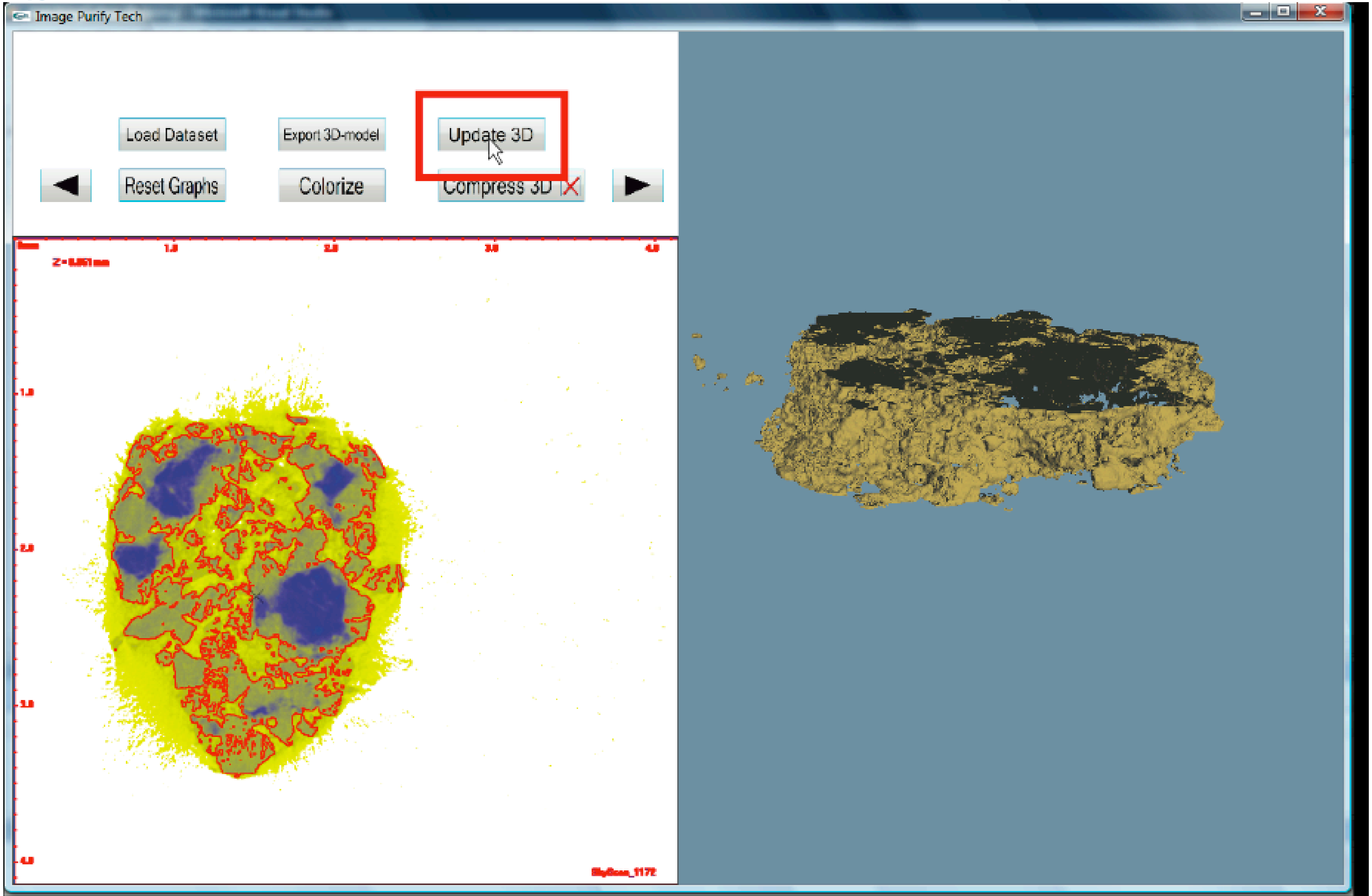
Real-time Image enhancement for Porous Rocks (w/ Henrik Hesland)

- In collaboration with Dept. Of Petroleum Engineering



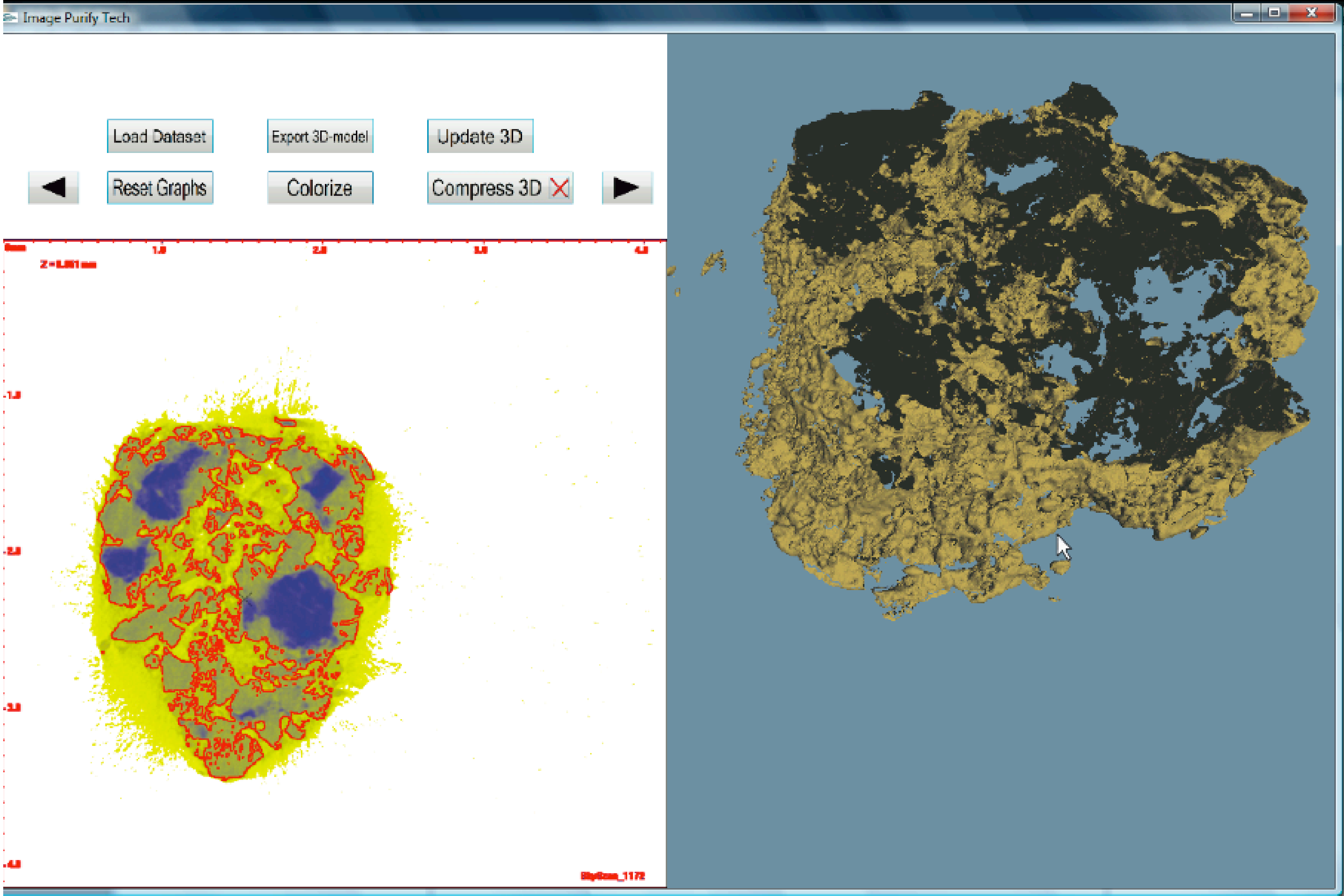
Global vs. Variable vs. Variable regional thresholds





a) <Update 3D> pressed, and volume loaded.

a) <Update 3D> pressed, and volume loaded.



b) Volume rotated.

Simulations of Fluid Flow through Porous Rocks using GPUs

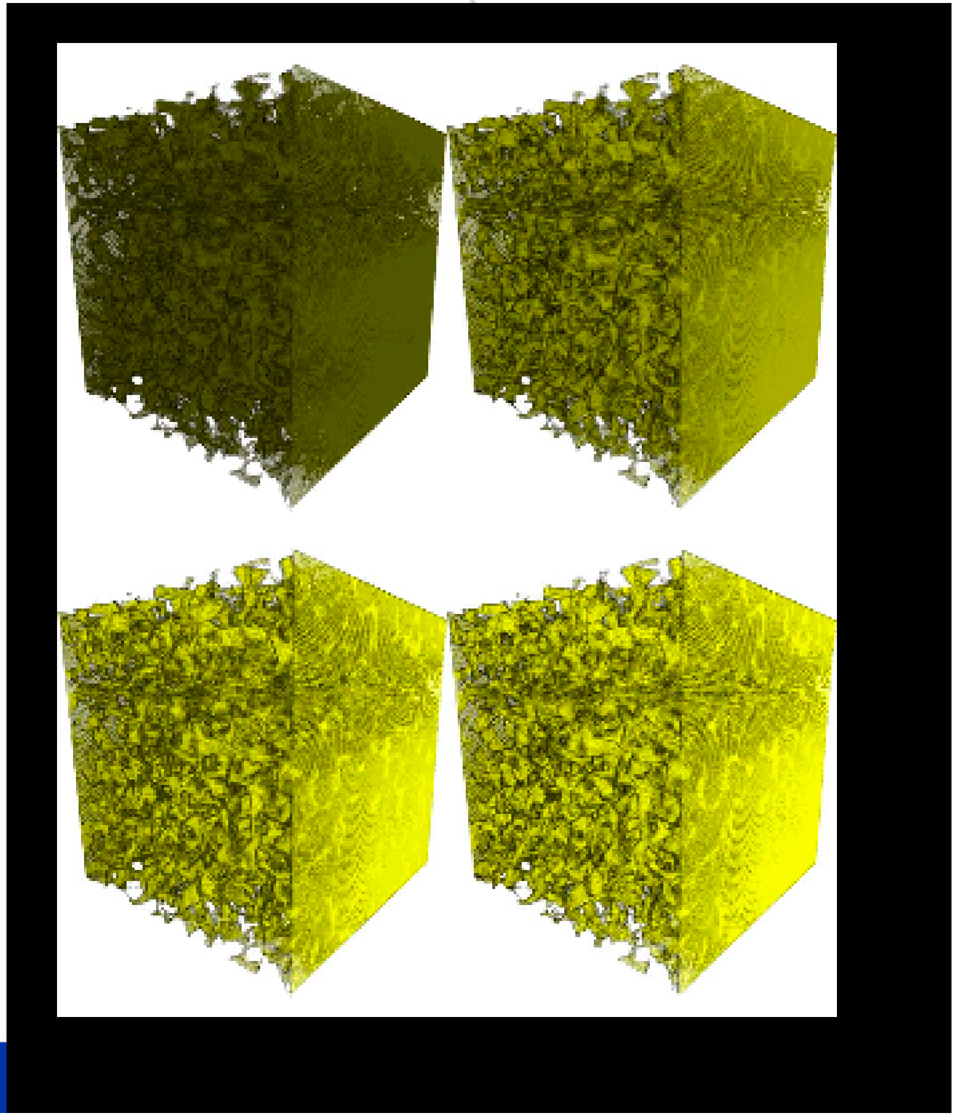
Eirik Ola Aksnes

(supervisor: Anne C. Elster)

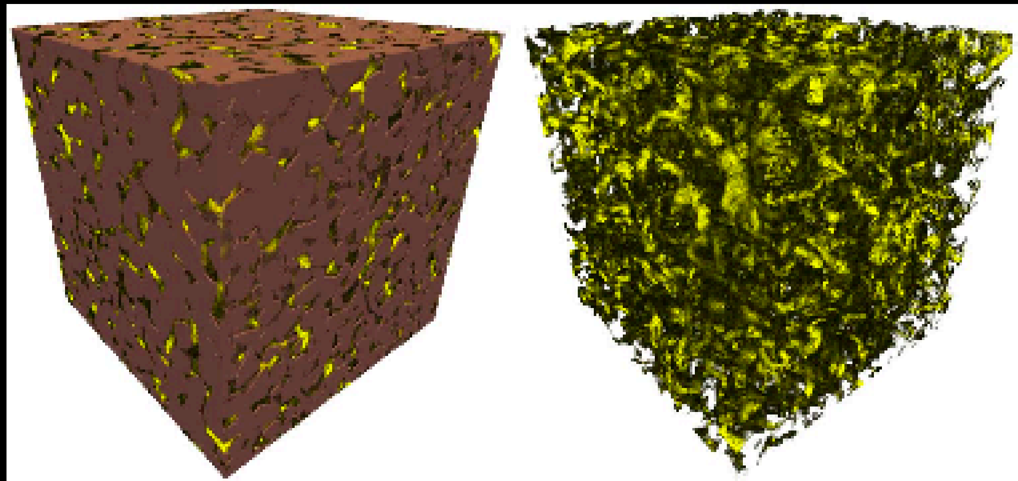
In collaboration with :

- Numerical Rocks &
- NTNU Chemistry Dept.

Use Lattice Bolzmann Method



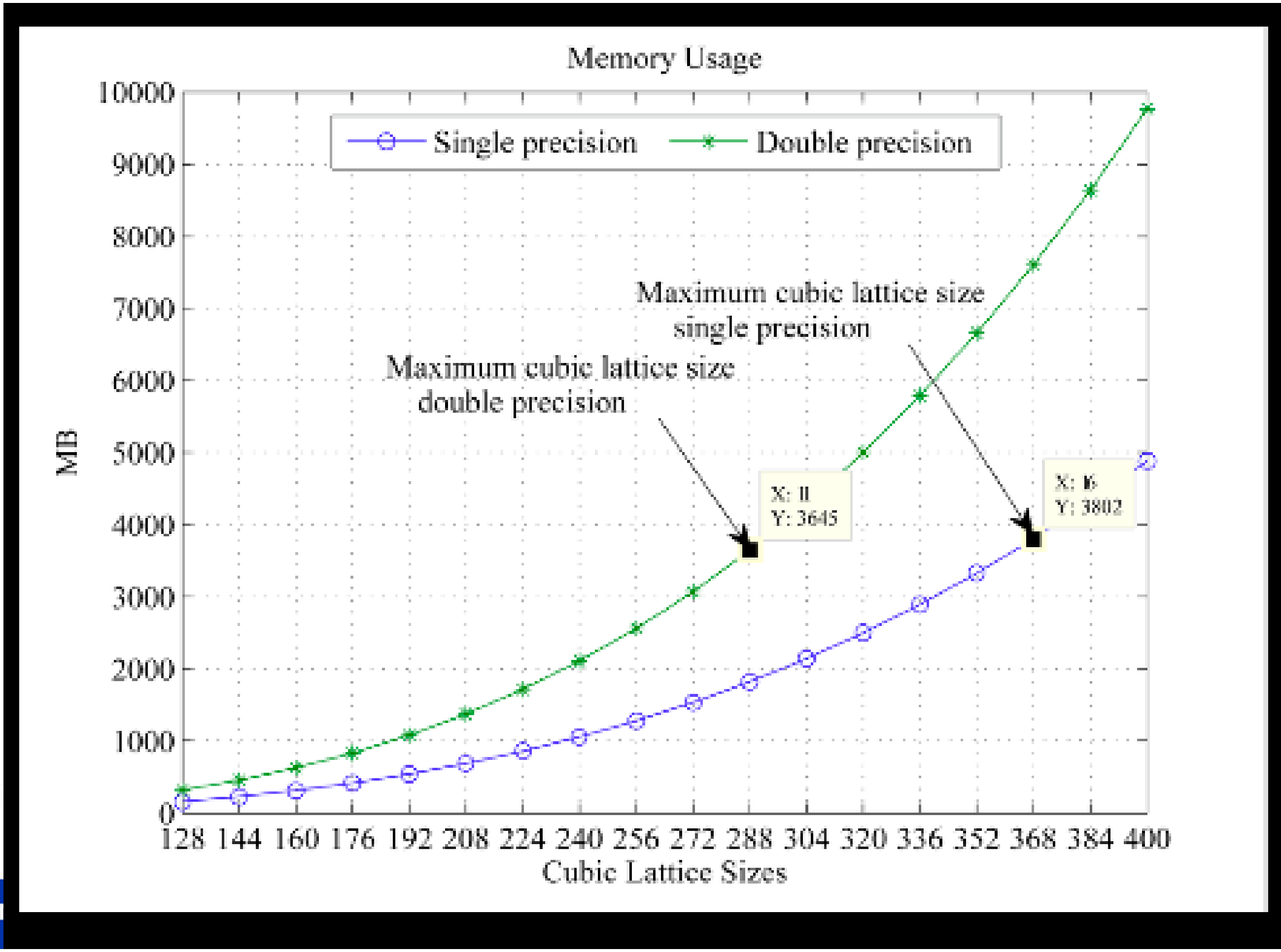
Benchmarks: Fontainebleau



Implementation	Average MLUPS	Maximum MLUPS	Total Time	Number Of Iterations	Permeability Obtained
CPU 32	1.03	1.04	2152 s	445	1247.80 mD
GPU 32	58.81	59.15	38.0 s	445	1247.81 mD
CPU 64	0.94	0.94	2375.4 s	445	1247.80 mD

LBM Global memory usage

Our GPU implementation support lattice sizes up to 368^3, which fit into the 4 GB mem. of the NVIDIA Quadro FX 5800 card.



Summary of LBM/Porous Simulation:

- To support large lattices and to get high performance: Swap-instead-of-copy-approach.
- The configuration of grids and thread blocks of kernels were properly configured.
- Register and shared memory usage of the kernels were minimized. Structure-of-arrays / coalescing.
- To get matching result from CPU and GPU with single floating-point precision: Round-off errors in the simulation model were reduced.

NTNU
Norwegian University of Science and Technology

"Large Seismic Application on GPU"

Owe Johansen

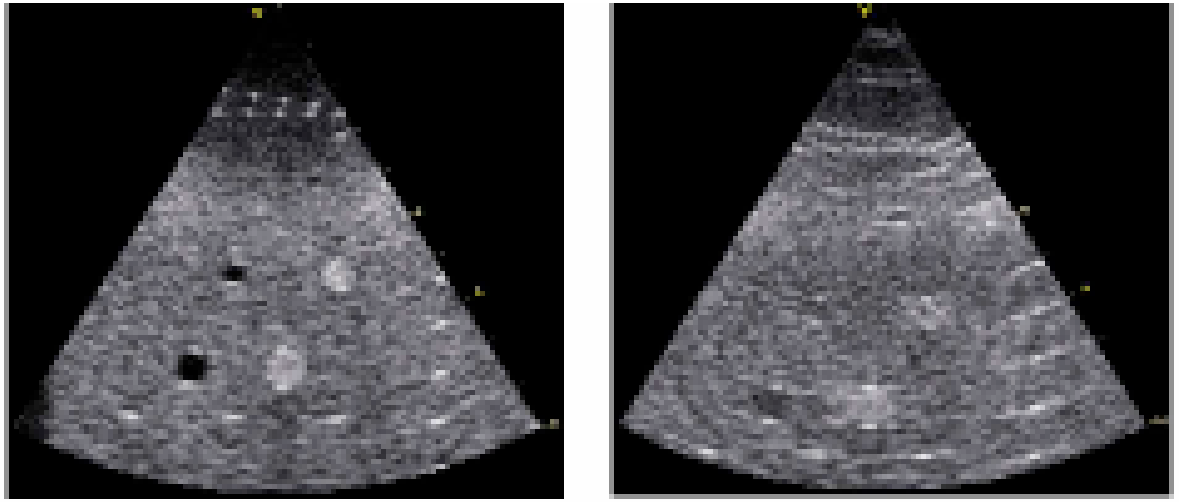
-- collaboration with StatoilHydro
(to be presented at a later conference)

NTNU
Norwegian University of Science and Technology

"Parallel Techniques for Estimation and Correction of Aberration on Medical Ultrasound Images"

Åsmund Herikstad

-- Collaboration w/ NTNU Med. Tech.
(NOTUR 2009 Poster)



Left: Unaberrated image, right: Aberrated Image

NTNU
Norwegian University of Science and Technology

HPC-Lab supports research and studies of novel GPU and multi-core architectures

- Parallel and Distributed Algorithms
- **Performance Evaluation and Benchmarking**
- Parallelization of Seismic and Image Related Applications on GPUs and Multi-Cores
- Adaptive and Auto-Tuneable Algorithms and Implementations

Project/Master Thesis Topics

- GPGPU for HPC
- Compiler Techniques for Parallel Linear Algebra
- Memory Latency Impact GPU-CPU Configurations
- Other projects in parallel and distributed computing: **suggest something that interests you!**

Collaborators / Supporters

ARM, CERN, NVIDIA, StatoilHydro, Schlumberger, GE-Healthcare, and others

"Communication Challenges on Multi-GPU Systems" (NOTUR 2009 Poster)

Daniele G. Spampinato

NTNU
Norwegian University of Science and Technology

"Modelling Overlapping Communication and Computation"

(NOTUR 2009 Poster)

Jan Christian Meyer
PhD Student

NTNU
Norwegian University of
Science and Technology

Zotac Nvidia ION

Specifications:

Nvidia ION GPU,

Graphics Engine Clockl: 450MHz

Shader Click: 1100MHz

16 stream processors

MS DirectX 10

Intel Atom (integrated) up to 533 MHz FSB

DDR2 667/800

Up to 4GB RAM

1 mini PCIExpress

“High Data Volumes and Streaming on Future GPU Systems”

Rune Hovland
Master of Tech at
NTNU June 2008.
Now consultant and
Sirius IT,
Oslo, Norway

-- Collaboration w/ Dr. Magnus Lie Hetland (IDI)
And Dr. Øystein Thorbjørnsen (Fast/MicroSoft)

On-board GPUs

NTNU
Norwegian University of
Science and Technology

Throughput Computing

- Information Retrieval (IR), including searches, focuses on throughput and large data volumes.
- Goal for IR:process large amounts of data efficiently, something current GPUs are not good at.
- Here are some modelling and suggestions on how to fix this

Application Related Topics

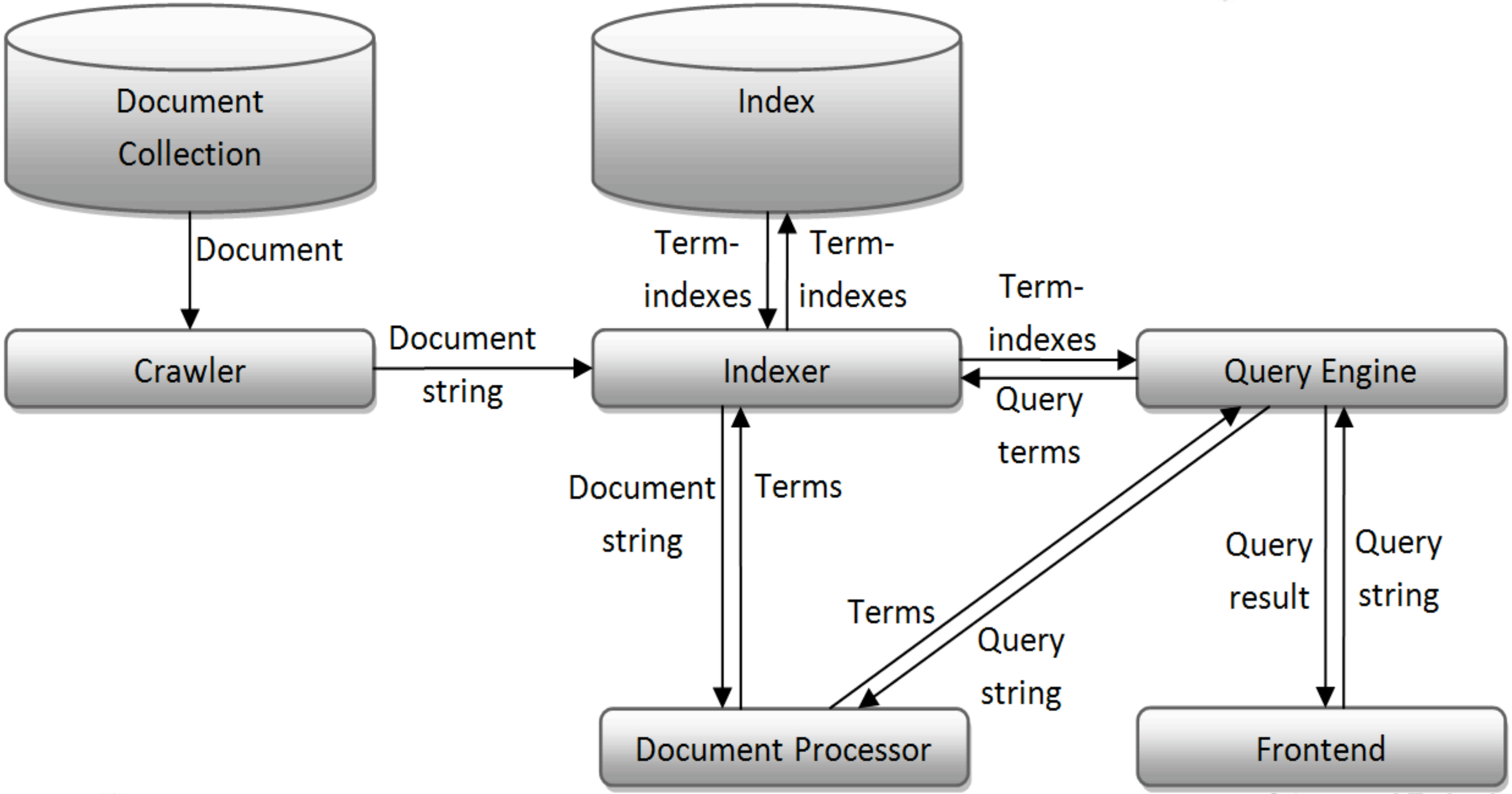
- Search engines
  - Architecture
  - Inverted index (datastructure optimized for term look-ups e.g. Loooking for freqency of term
    - Document-level inverted index
    - Word-level
    - Block-level)

Compression (application dependent)

- variable byte encoding
- Golomb encoding (bit-wise)
- Pfor Delta (avoiding branches and optimizes pipeline

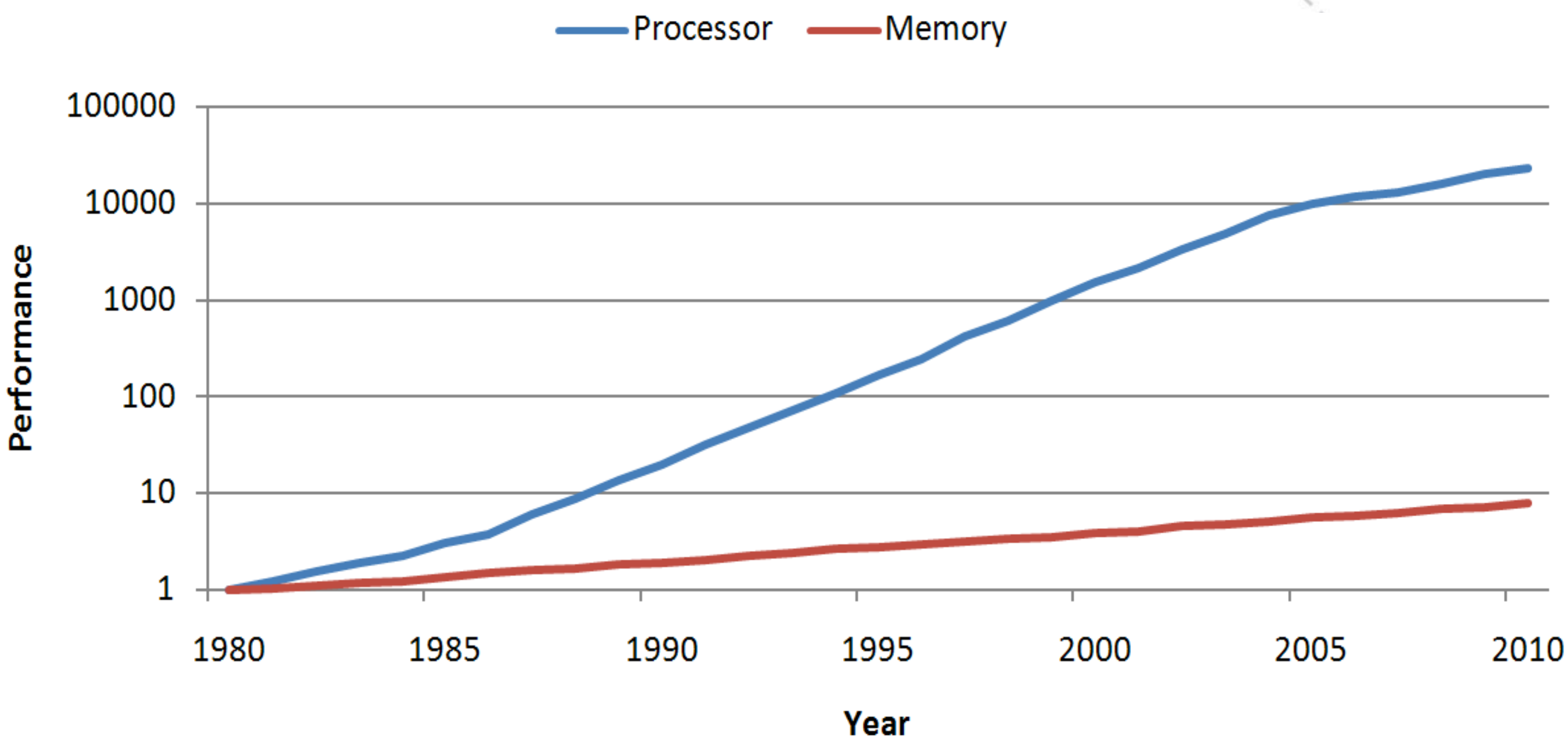
NTNU
Norwegian University of Science and Technology

Search Engine Architecture



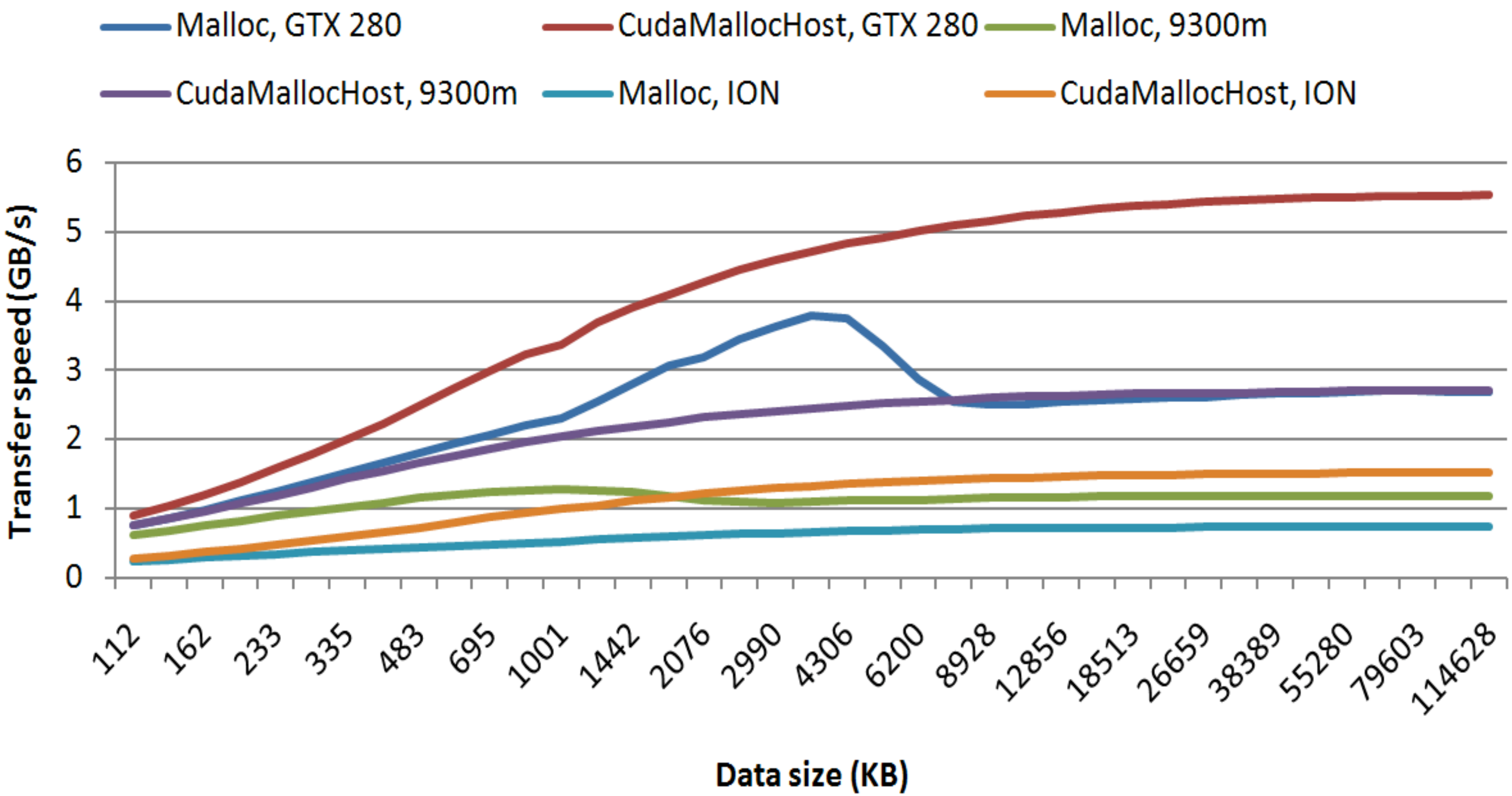
of
Science and Technology

Memory Gap
(Patterson & Hennesey)

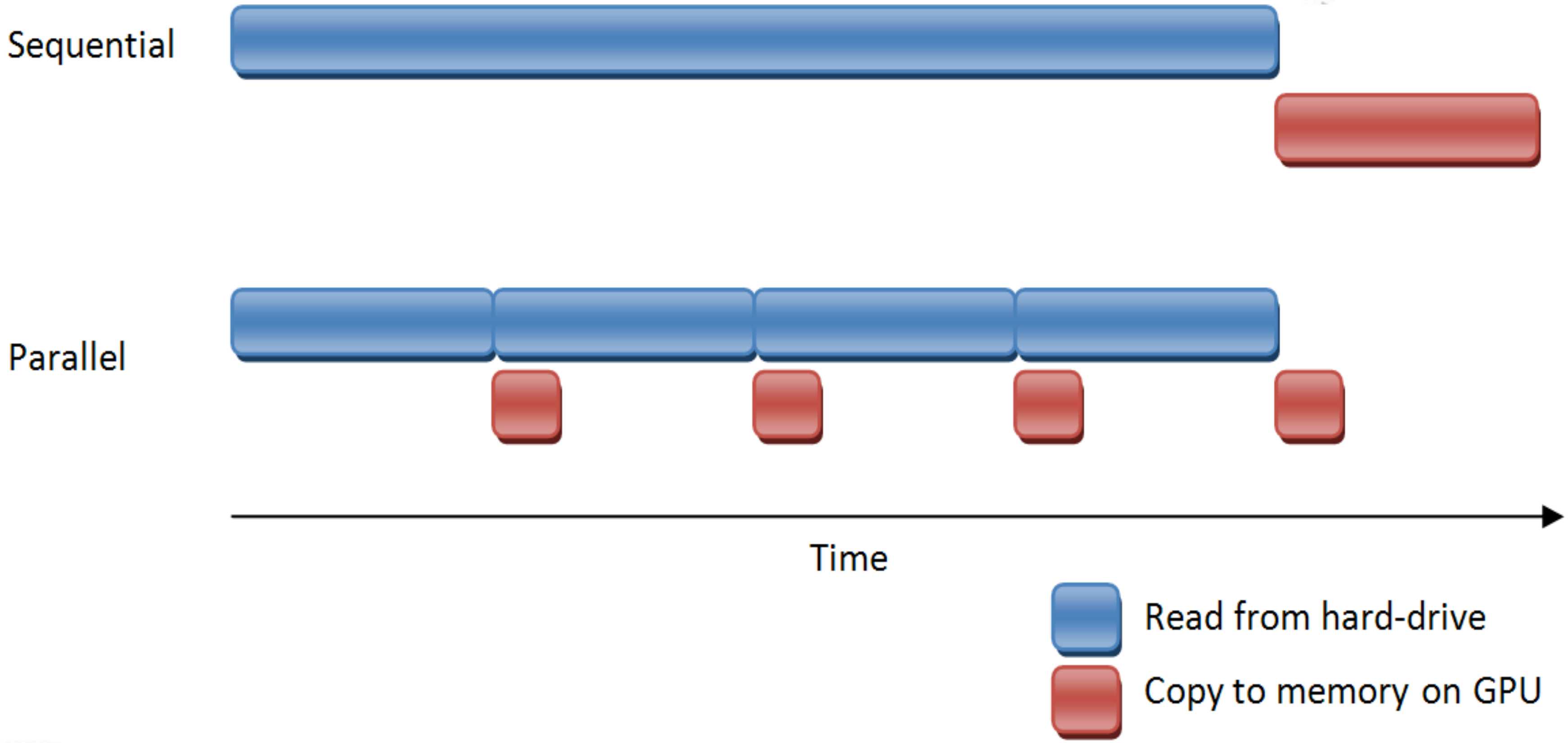


NTNU
Norwegian University of Science and Technology

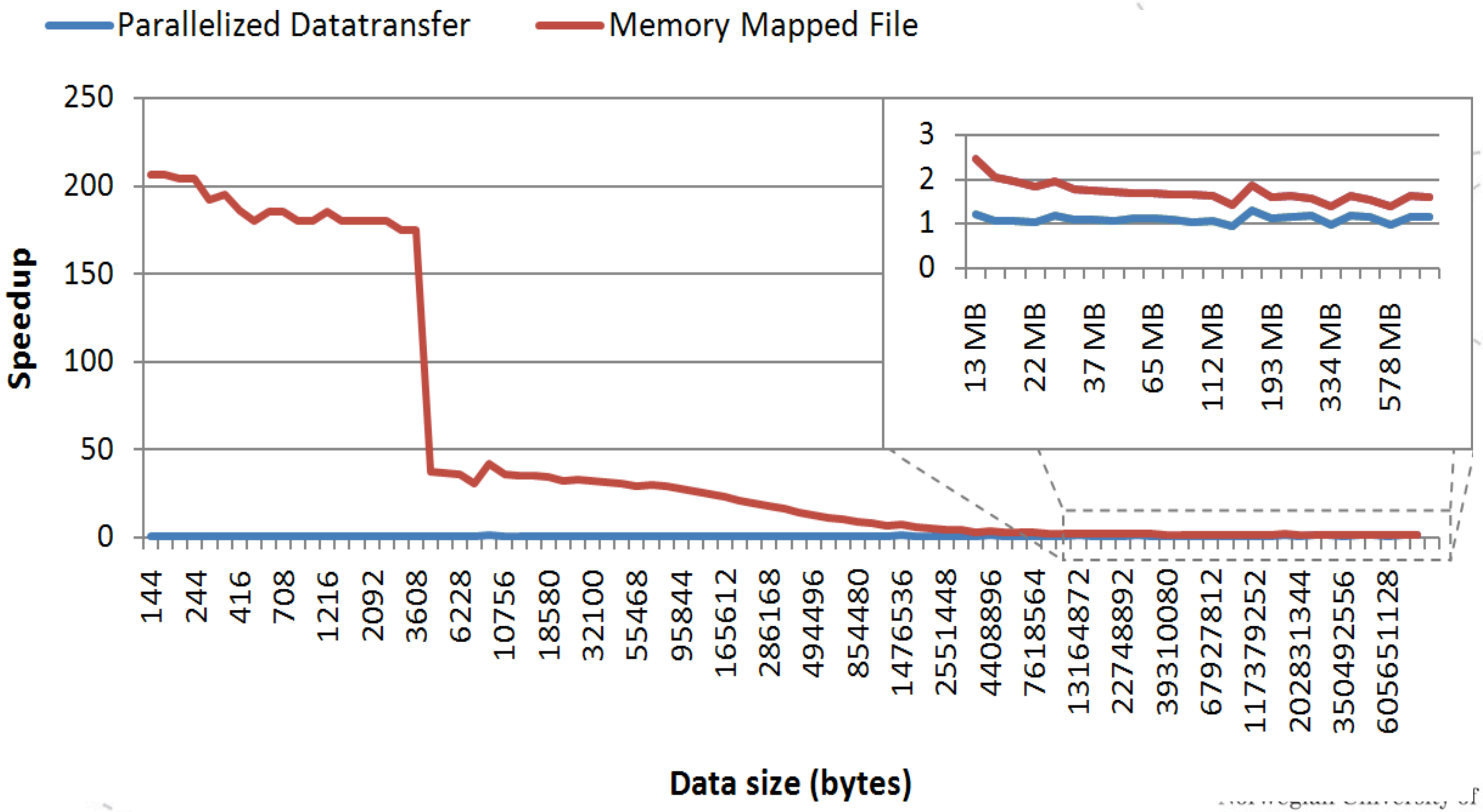
ION vs 9300m vs GTX 280



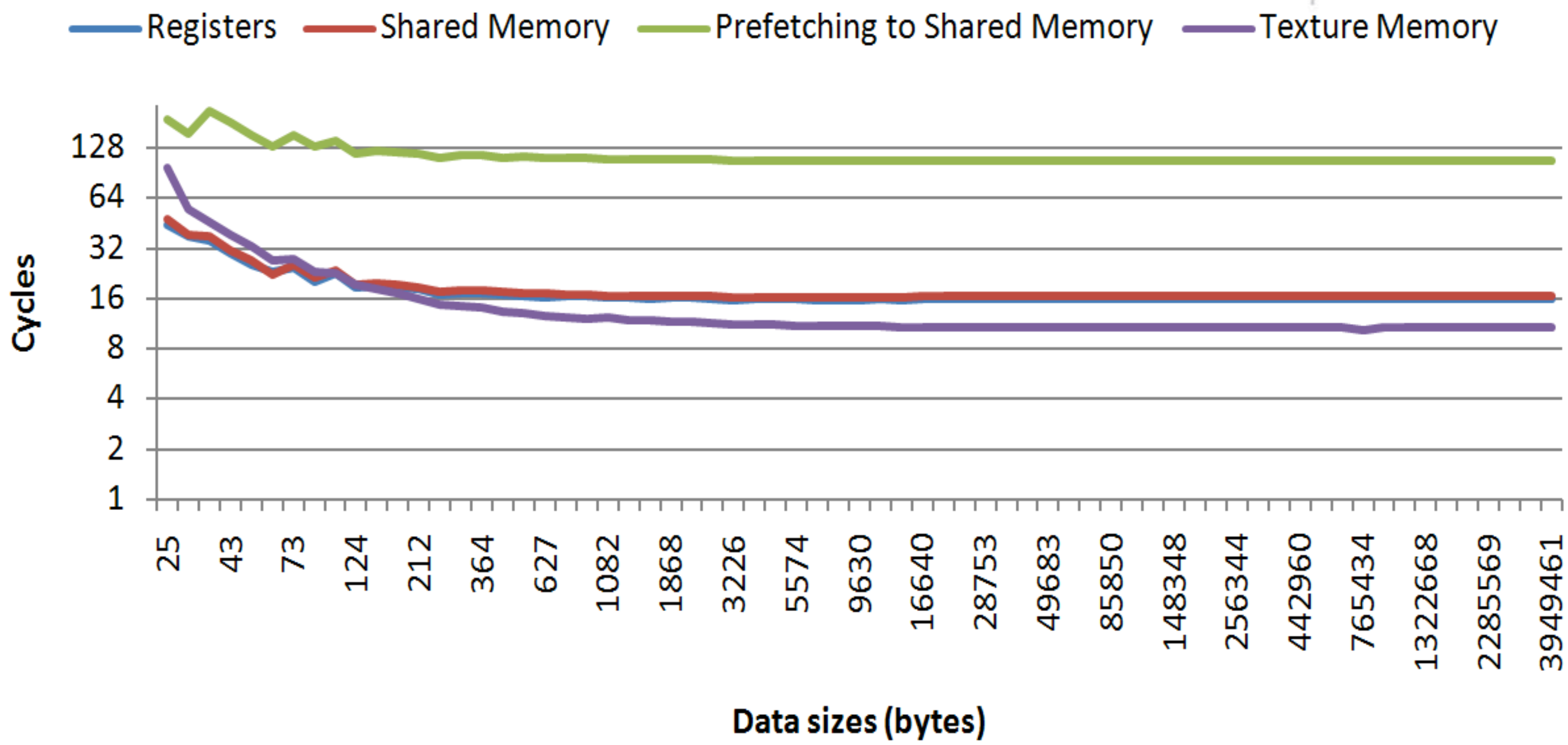
Hide datacopy between host & GPU while accessing disk:

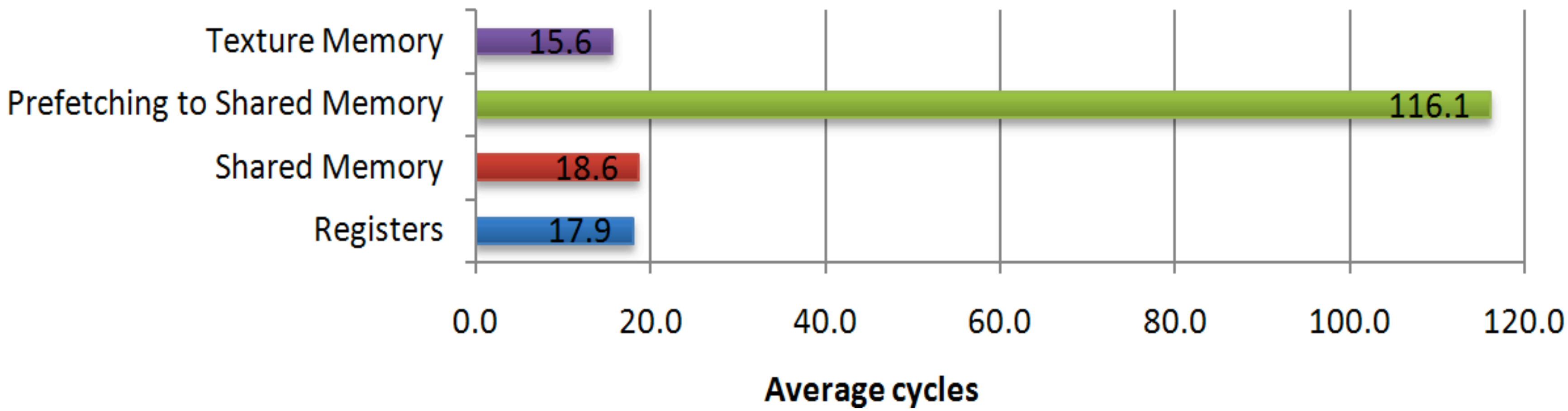


But, memory mapped files best ...



Data access figures





Expanding Memory hierarchy

- Nvida Tesla architecture
- Allows GPU to have:
  - Its own dedicated memory
  - Use part of host memory

Any data used by GPU must be copied to device

Any results copied back (or displayed)

Would like: **Direct access to host memory!**

-> Change memory hierarchy

NTNU
Norwegian University of Science and Technology

Alternative: Combine Dedicated & Shared Memory

- Allow GPU to use both dedicated and local memory, but divided into two levels in memory hierarchy
- Would require CPU can "DMA" to host memory used as device memory
- All data GPU will use, stored here

NTNU
Norwegian University of Science and Technology

Zero-copy in CUDA

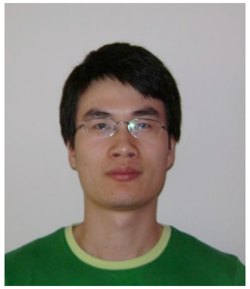
- To be introduced in CUDA 2.2
- Alllows user to page-locked memory from GPU
- Only paged-locked memory is limiting since a scarce resource
- User may still need to extra copy ops if there is extensive memoery usage on host (CPU)

NTNU
Norwegian University of Science and Technology

Fall 2009 Master Projects

Ahmed Aquari
Schlumberger project on GPU

Aleksander Gjermundsen
LBM/solvers - Snow Simulation



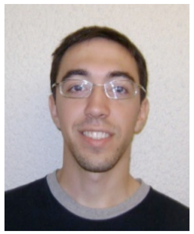
Gaojie He
GPU-enhanced parallel games

Øystein Krogh
Snow Simulations - terrain interactions

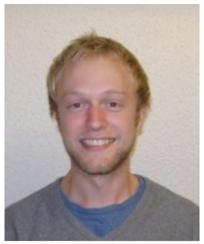
Holger Ludvigsen
GPU Ray tracing using OptiX

Runar Refsnæs (Math)

Gagandeep Singh (Math)



Roald Fernandez (Cybernetics)



Peter Sveistrup (Cybernetics)

+ 3 visualization students
+ 1-2 || arch/ multicore students

NTNU
Norwegian University of Science and Technology

Dzienkuje! /Thank you!

(jen-koo-ye)

Contact Info:

elster@idi.ntnu.no

http://www.idi.ntnu.no/~elster/hpc-lab

GPU-related Activities & Events:

cost

EU COST Action IC0805: Open European Network for High Performance Computing on Complex Environments (2009-2014)

ParCo2009

Lyon, France, Sep 1-4, 2009

MS on GPU Computing /EuroGPU2009

NTNU
Norwegian University of Science and Technology

HPC-LAB at SC´08

NTNU
Norwegian University of Science and Technology

NTNU Gløshaugen

(formerly Norwegian Institute of Technology)

NTNU
Norwegian University of Science and Technology

"Dynamic Optimization of MPI Communication"

(NOTUR 2009 Poster)

Thorvald Natvig
PhD Student

NTNU
Norwegian University of
Science and Technology

HPC Group contin.

At SC´07 in Reno, NV, USA

NTNU
Norwegian University of Science and Technology

GPU Apps
(from NVIDIA's web pages)

CUDA applications actively in use today by these researchers and organizations include:

NTNU
Norwegian University of Science and Technology

Oil and gas

- **Acceleware:** Kirchoff Time Migration library
- **ffA:** 3D Seismic processing software
- **Headwave:** Prestack data processing
- **Mercury Computer systems:** 3D data visualization
- **SeismicCity:** 3D seismic imaging for prestack depth migration
- **SMT:** Kingdom – Seismic Processing

Computational Chemistry and Molecular Dynamics:

- GROMACS molecular dynamics
- HOOMD molecular dynamics
- NAMD molecular dynamics
- VMD visualization of molecular dynamics

NTNU
Norwegian University of Science and Technology

Bio-Informatics and Life Sciences:

- GPU HMMER: CUDA version of HMMER
- LISSOM: Human neocortex modeling
- MUMmerGPU: High-throughput DNA sequencing

NTNU
Norwegian University of
Science and Technology

Financial Computing and Options Pricing:

- Aqumin: 3D Visualization of market data
- Exegy: Risk Analysis
- Hanweck: options pricing
- SciComp: derivatives pricing

NTNU
Norwegian University of Science and Technology

Mathematical Computing

- Jacket CUDA plugin for MATLAB from Accelereyes
- LabVIEW from National Instruments

NTNU
Norwegian University of
Science and Technology

GeoSciences:

- Tsunami simulation – Tokyo Institute of Technology
- Weather Research and Forecast (WRF) model
- Geographical Information Systems - Manifold

NTNU
Norwegian University of
Science and Technology

Medical Imaging, CT, MRI:

- AxeRecon CT reconstruction library from Acceleware
- SnapCT tomographic reconstruction software from Digisens

NTNU
Norwegian University of Science and Technology

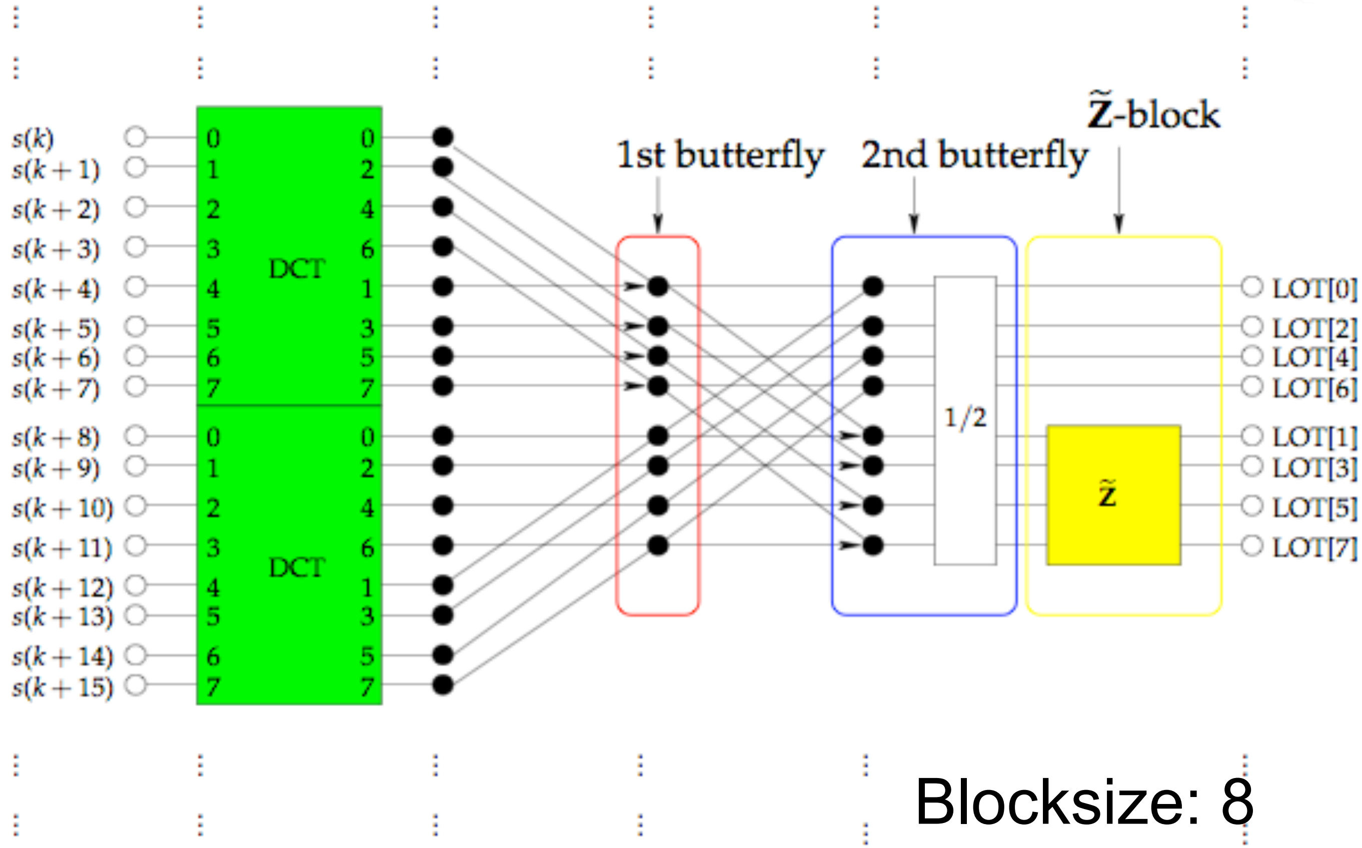
Electrodynamics and Electromagnetics

- CST Microware Studio
- FDTD solver from Acceleware

Electronic Design Automation

- ADS SPICE simulator from Agilent EESof
- OmegaSim GX SPICE simulator from Nascentric
- Sentaraus TCAD from Synopsys

LOT : Lapped Orthogonal Transform (H.S: Malvar 1980s, Textbook: 1992)



Blocksize: 8

NTNU
Norwegian University of Science and Technology

Why LOTs?

- Attempts to solve blocking problems w/ DCTs

-->Better objectively & subecttively compression

NTNU
Norwegian University of
Science and Technology

Feig-Linzer DCT *vs* Fast LOT

Algorithm	Multiplies	Adds/ Subtractions	Multiply-adds	FLOPs
FL mult-add DCT	64	256	160	640
Fast LOT w/ FIL mult-add DCT	288	512	256	**1312**

But LOT very vectorizable!

NTNU
Norwegian University of Science and Technology

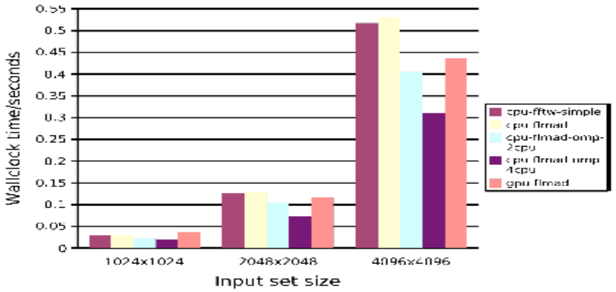
Compression of Image Data on Clusters using GPUs and Quad-Core CPUs

Leif Christian Larsen, IDI-NTNU leifchl@idi.ntnu.no
Anne Cathrine Elster, IDI-NTNU (*main supervisor*) elster@idi.ntnu.no
Tore Fevang, Schlumberger Ltd (*co-supervisor*) tfevang@slb.com

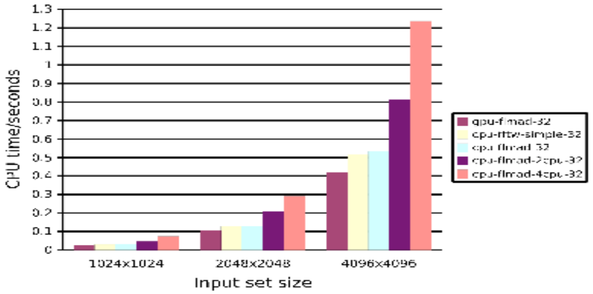
• In client-server visualization applications using clusters for computations, image compression is used to reduce communication time between cluster nodes and from cluster nodes to the client, thereby improving application responsiveness

• Transform coding methods are employed for compressing image data. Transform computation may be offloaded to the graphics processing unit (GPU), which is well suited for such computations. Offloading computations to the GPU releases CPU time for other tasks, and in certain cases offers significantly improved performance over using multi-core CPUs.

Discrete Cosine Transform – Wallclock time



Discrete Cosine Transform – CPU time



• GPU transform implementations compute hundreds of transform values simultaneously by using SIMD vector instructions and multiple GPU vertex and pixel processors in parallel.

• The GPU is particularly well suited for transforms involving heavy computations and large amounts of data, since data upload/download/initialization time becomes less significant as arithmetic complexity increases

Acknowledgement. This project is done in collaboration with Schlumberger Limited, Trondheim, which provides hardware equipment and other resources.

NTNU
Innovation and Creativity

NTNU
Norwegian University of Science and Technology

NTNU Collaborations with CERN:

CERN openlab
for DataGrid applications

- Two Norwegian students in summer 2003:
  1 Master student from Elster's NTNU group (Hisdal)
  + 1 from Ousada's Univ. of Oslo's Physics group
  Hisdal went on to do his MS thesis at CERN
  (with CERN staff and Elster)

- **Six MS** students from NTNU in summer 2004-2005
  Three of these went on to do their MS Theses at
  CERN (with CERN staff and Elster)

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of Science and Technology

Dr. Anne C. Elster
Lab Director

Thorvald Natvig
PhD Student

Jan Christian Meyer
PhD Student

Master Students

Robin Eidissen
(Teaching Assitant)

Rune E. Jensen

Olav Fagerlund

Eirik O. Aksnes

Daniel Haugen

Henrik Hesland

Åsmund Herikstad

Owe Johansen

Rune Hovland

Daniele G. Spampinato