

HPC usage in geo-resources exploration and modeling

Oscar Peredo

Associate Researcher

ALGES Lab, Advanced Mining Technology Center, U. de Chile



ALGES

ADVANCED LABORATORY FOR
GEOSTATISTICAL SUPERCOMPUTING

amtc

ADVANCED MINING TECHNOLOGY CENTER

Outline

Motivation

ALGES Lab

HPC usage in exploration geophysics

3D Electromagnetic Inversion with Fortran90/MPI/OpenMP

HPC usage in geostatistical modeling

Variogram acceleration using C#/Threads

Optimization of k-NN kernel using CUDA

Comments

Outline

Motivation

ALGES Lab

HPC usage in exploration geophysics

3D Electromagnetic Inversion with Fortran90/MPI/OpenMP

HPC usage in geostatistical modeling

Variogram acceleration using C#/Threads

Optimization of k-NN kernel using CUDA

Comments

Motivation

What kind of problems can be solved/accelerated using HPC in geo-resources exploration and modeling?

What are we doing today at ALGES/AMTC?

Outline

Motivation

ALGES Lab

HPC usage in exploration geophysics

3D Electromagnetic Inversion with Fortran90/MPI/OpenMP

HPC usage in geostatistical modeling

Variogram acceleration using C#/Threads

Optimization of k-NN kernel using CUDA

Comments

Our goal is to develop tools, software and models for the characterization of geological phenomena, specifically regarding geological modelling and resource assessment.

Fundamental research areas:

- Multivariate geostatistics
- Multiple-point geostatistics
- Image and texture analysis
- Heuristics for combinatorial optimization
- Data mining tools
- **HPC/Supercomputing**

People:

- 2 full time Faculty researchers
- 9 full time researchers
- 6 part time researchers
- 2 Postdoc
- 11 MSc students
- 2 PhD students
- 3 undergrad students

Achievements:

- Publications: 20 ISI / 35 Non-ISI
- Grants: US\$ 1.826.000
- Graduates: 26
- Software Licence Agreements: 2

Industrial/Scientific sponsors/partners:

- BHP Billiton
- CODELCO
- El Tesoro
- Yamana Gold
- **Nvidia (1st chilean CUDA research center)**
- **Barcelona Supercomputing Center (Spain)**
- University of New South Wales (Australia)
- Facultad de Ciencias Físicas y Matemáticas (Chile)

Outline

Motivation

ALGES Lab

HPC usage in exploration geophysics

3D Electromagnetic Inversion with Fortran90/MPI/OpenMP

HPC usage in geostatistical modeling

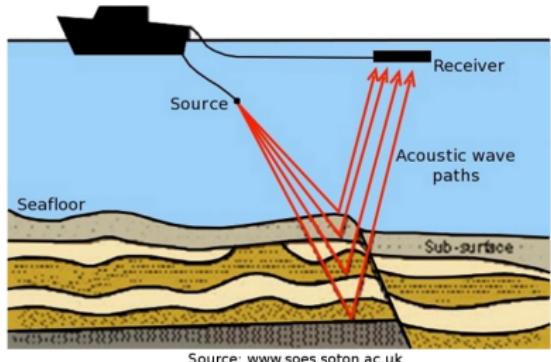
Variogram acceleration using C#/Threads

Optimization of k-NN kernel using CUDA

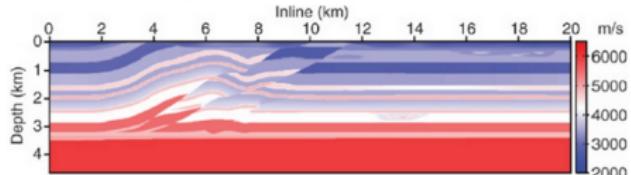
Comments

Exploration Geophysics

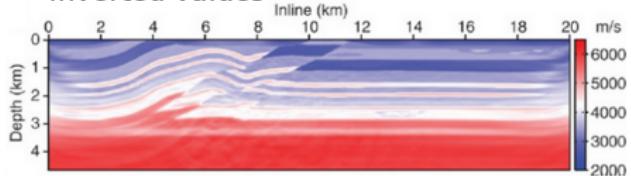
Seismic methods



True model

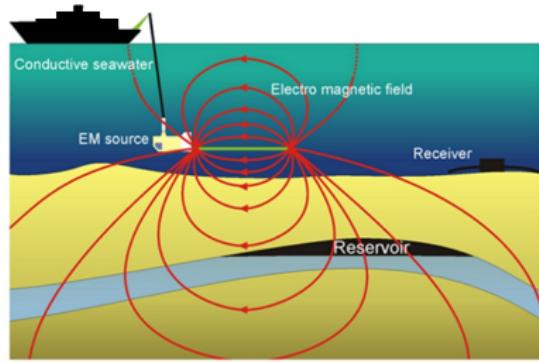


Inverted values



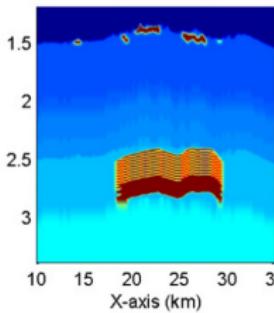
Source: Ben-Hadj-Ali, Operto, Virieux (2011)

Electromagnetic methods



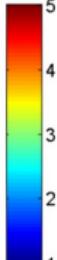
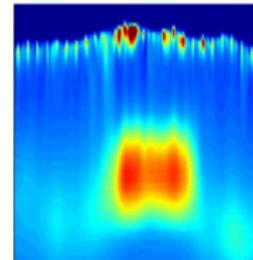
Source: noc.ac.uk

True model



Source: Plessix, Mulder (2008)

Inverted values



Outline

Motivation

ALGES Lab

HPC usage in exploration geophysics

3D Electromagnetic Inversion with Fortran90/MPI/OpenMP

HPC usage in geostatistical modeling

Variogram acceleration using C#/Threads

Optimization of k-NN kernel using CUDA

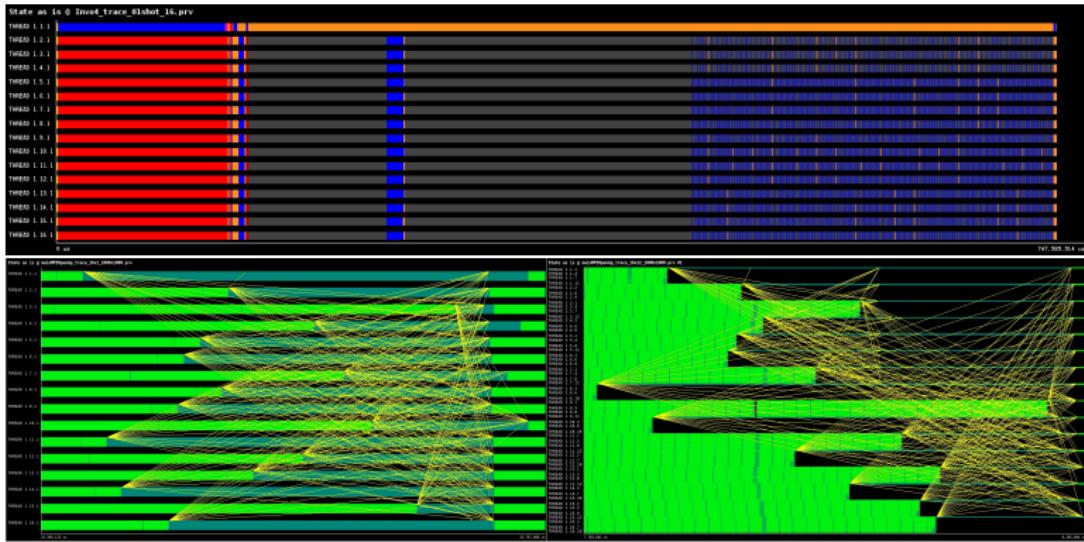
Comments

3D Electromagnetic Inversion

- Task:
 - Implementation of reduced gradient calculation (discrete adjoint), to be used in a non-linear optimization framework (inverse problems). Applied case: 3D Electromagnetism.
- Environment:
 - Software: Alya
 - Parallel Multi-physics simulator (PDE solver, Finite Elements)
 - Developed by Barcelona Supercomputing Center
 - Fortran 90, MPI+OpenMP, +750000 lines of code
 - User interface: configuration files, command-line
 - Hardware: Marenostrum III
 - +40000 CPUs (3028 compute nodes, Intel SandyBridge-EP E52670 2 CPUS 8-cores each at 2.6 GHz)
 - Operating system: SUSE Linux
- **Parallel patterns (dwarfs, Tech-Rep. UCB/EECS-2006-183):**
 - **Unstructured grid** (non-uniform 3D tetrahedral mesh with local refinements)
 - **Sparse linear algebra** (parallel sparse matrix-vector operations, iterative methods)

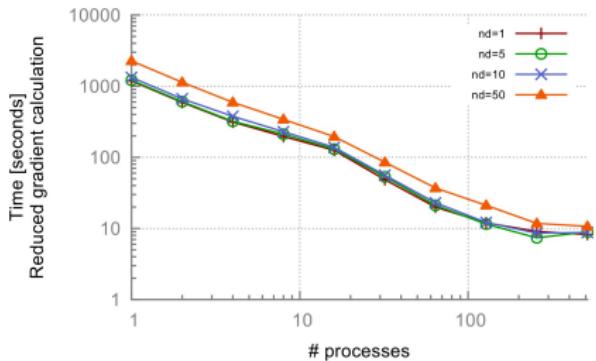
Profiling/Debugging

- Profiling: Extrae/Paraver, Vampir, Oprofile, gprof, kcachegrind, ...
- Debugging: Valgrind, gdb/idb, ...
- Trace views:

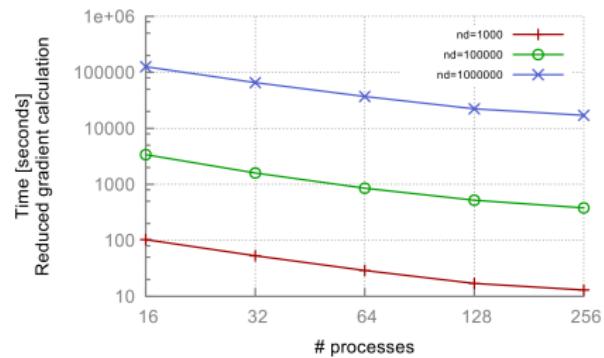


Speedup

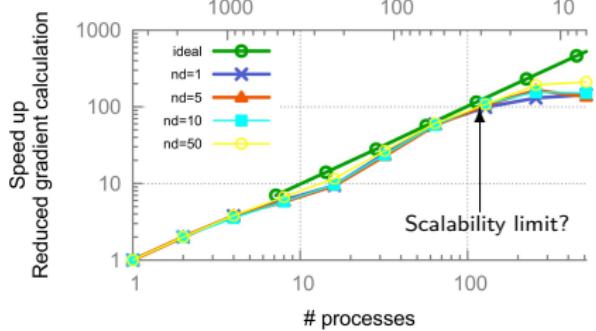
2D mesh, 3.7M tetra elems



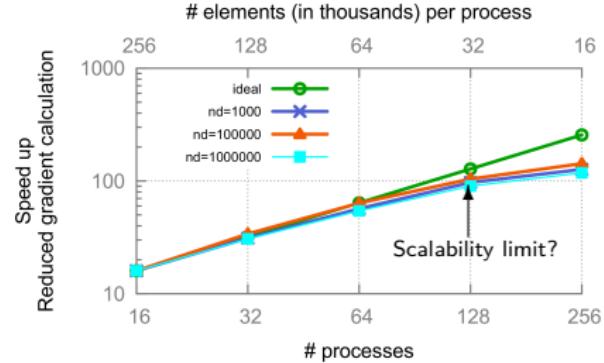
3D mesh, 4.7M tetra elems



elements (in thousands) per process



elements (in thousands) per process



Outline

Motivation

ALGES Lab

HPC usage in exploration geophysics

3D Electromagnetic Inversion with Fortran90/MPI/OpenMP

HPC usage in geostatistical modeling

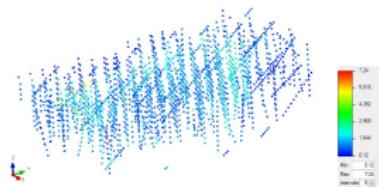
Variogram acceleration using C#/Threads

Optimization of k-NN kernel using CUDA

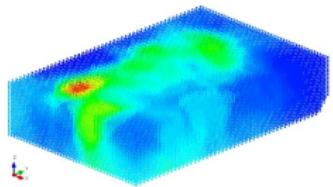
Comments

Geostatistical modeling

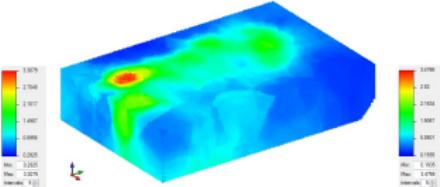
Kriging estimation



2380 data points



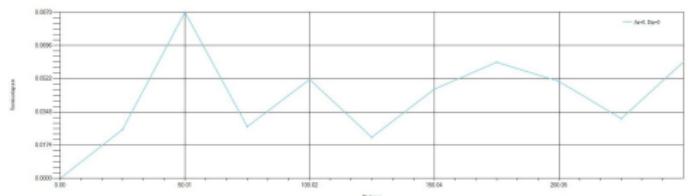
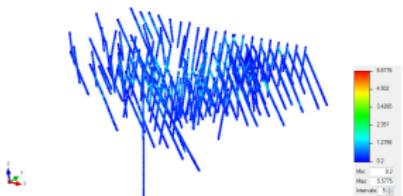
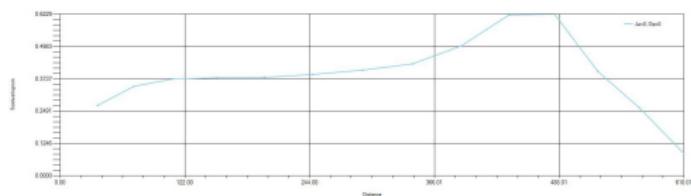
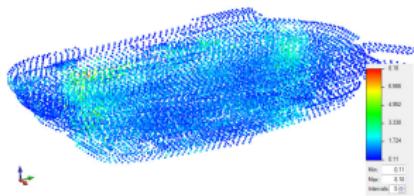
grid= $50 \times 50 \times 26$



grid= $100 \times 100 \times 52$

Source: generated with Umine software.

Variogram calculation



Source: generated with Umine software.

Outline

Motivation

ALGES Lab

HPC usage in exploration geophysics

3D Electromagnetic Inversion with Fortran90/MPI/OpenMP

HPC usage in geostatistical modeling

Variogram acceleration using C#/Threads

Optimization of k-NN kernel using CUDA

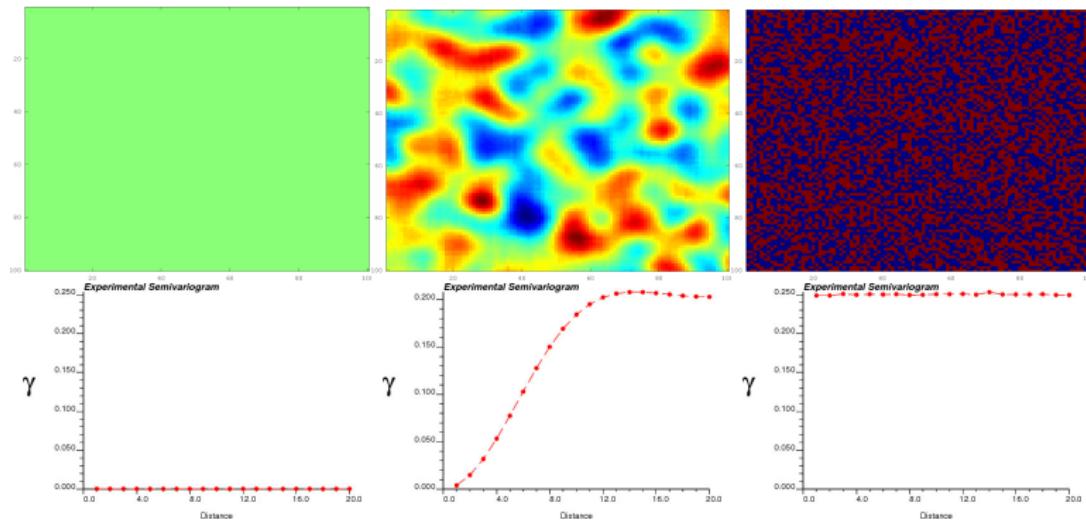
Comments

What is the variogram?

Definition:

It is a tool that quantifies spatial correlation. The variogram for lag distance \mathbf{h} (vector) is defined as the average squared difference of values separated approximately by $|\mathbf{h}|$: $2\gamma(\mathbf{h}) = \frac{1}{N(\mathbf{h})} \sum_{N(\mathbf{h})} (z(\mathbf{u}) - z(\mathbf{u} + \mathbf{h}))^2$

Examples:



Variogram acceleration

- Task:
 - Acceleration of the variogram calculation in 3D datasets (code optimization and multi-threading implementation, no OpenMP pragmas available)
- Environment:
 - Software: Umine
 - Univariate geological resources evaluation and modeling tool (estimation, simulation, variographic analysis, ...)
 - Developed by ALGES Lab, Universidad de Chile (licensed software)
 - C#, no optimized, sequential, +100000 lines of code
 - User interface: GUI
 - Hardware: any Desktop computer
 - 1~8 CPUs
 - Operating system: Windows (XP, Vista, 7, 8)
- **Parallel patterns (dwarfs):**
 - **N-body** (sum reduction without synchronization, upper-triangular loop partitioning using Windows threads)
 - **Unstructured grid** (non-uniform data point locations)

Profiling/Debugging

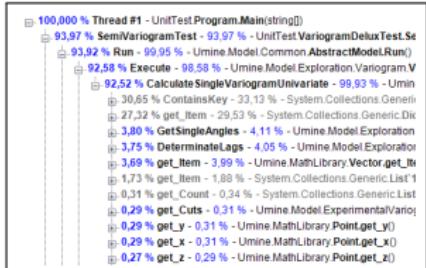
- Profiling: ANTS Performance Profiler, SlimTune, Visual Studio Profiler, ...

6,929,874	0,000	6,713
6,929,874	0,000	6,714
6,929,874	0,000	4,739
6,929,874	0,000	4,726
6,929,874	0,000	4,725
6,929,874	0,000	0,219
6,929,874	0,000	0,219
6,929,874	0,000	0,219
6,929,874	0,000	0,328
6,929,874	0,000	0,164
.....

```
    }
    // Definition of the lag corresponding to the current pair
    //
    point2 = points[j];
    pairJ = values[j];

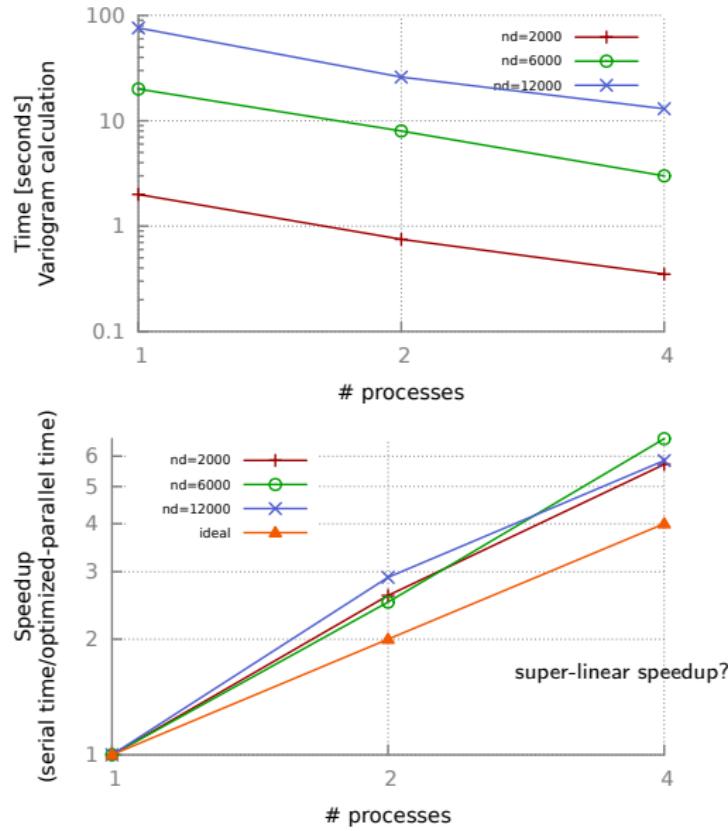
    dx = point2.x - point1.x;
    dy = point2.y - point1.y;
    dz = point2.z - point1.z;
    dxs = dx * dx;
    dys = dy * dy;
    dzs = dz * dz;
    hs = dxs + dys + dzs;

    if (hs < 0.0) hs = 0.0;
```



- Debugging: Visual Studio, code instrumentation, licensing overhead issues (check if your licensing software is lowering your performance being too strict)...

Speedup



Outline

Motivation

ALGES Lab

HPC usage in exploration geophysics

3D Electromagnetic Inversion with Fortran90/MPI/OpenMP

HPC usage in geostatistical modeling

Variogram acceleration using C#/Threads

Optimization of k-NN kernel using CUDA

Comments

Optimization of k-NN kernel

- Task:
 - Code optimization of existing CUDA kernel that calculates k nearest neighbours. Possible techniques: inlining, memory coalescing, loop transformation, loop unrolling, register blocking,...
- Environment:
 - Software: kNN.cu
 - k nearest neighbours GPU calculation routine.
 - Developed by ALGES Lab, Universidad de Chile
 - CUDA, no optimized, 200 lines of code
 - User interface: command-line
 - Hardware: Desktop + Nvidia GPU Quadro NVS 200
 - 4 CPUs + 8 SM (1024 threads per block)
 - Operating system: openSUSE 12.2
- **Parallel patterns (dwarfs):**
 - No parallel patterns, just code optimization

Profiling/Debugging

- Profiling: nvprof, NVIDIA Visual Profiler, ...
- Debugging: CUDA-gdb, code instrumentation, ...
- Auto-tuning: Flamingo (mistymountain.co.uk/flamingo) (in progress)

Speedup

Test: we measured 1000 queries centered in random points, searching k nearest neighbours over 64000 structured nodes ($40 \times 40 \times 40$ grid):

Version	Time [sec]	Speedup
CPU sequential code (-O3)	13.180	1.000
GPU no optimized	2.173	6.035
GPU manually optimized	1.039	12.685
GPU automatically optimized	?	?

This results are still *slow*, but we are working in the acceleration (binning or local searches, other low-level optimizations).

Outline

Motivation

ALGES Lab

HPC usage in exploration geophysics

3D Electromagnetic Inversion with Fortran90/MPI/OpenMP

HPC usage in geostatistical modeling

Variogram acceleration using C#/Threads

Optimization of k-NN kernel using CUDA

Comments

Comments

- We have shown three usages of HPC in exploration geophysics and geostatistical modeling, however, there are MANY other usages, this is just a glimpse.
- In ALGES/AMTC we are trying to push forward the usage of HPC in the Mining landscape, working together with national/international companies and research centers.
- However, we are short on hardware resources, due to an increasing usage of compute power and new applications....
- New collaborations?



Thanks for your attention!
Contact (author): operedo [at] alges.cl
<http://www.alges.cl> — <http://www.amtc.cl>