JÜLICH
APPLICATION LAB
NVIDIA

JÜLICH
FORSCHUNGSZENTRUM

NVIDIA Application Lab at Jülich

1st Dresden CCoE GPU Programming Workshop Workshop | 27. May 2013 | Dirk Pleiter (JSC)

Mitglied der Helmholtz-Gemeinschaft

JÜLICH
FORSCHUNGSZENTRUM

NVIDIA Application Lab at Jülich

Collaboration between JSC and NVIDIA since July 2012

- Enable scientific applications for GPU-based architectures
- Provide support for their optimization
- Investigate performance and scaling

Work focus

- Application requirements analysis
- Kepler and CUDA feature analysis
- Parallelization on many GPUs
- Collaboration with performance tools developers
- Training

Small team JSC+NVIDIA: Andrew Adinetz, DP + Jiri Kraus

Mitglied der Helmholtz-Gemeinschaft

Features of New Kepler Architecture

Hardware capabilities comparison

	M2090	K20x
Core clock speed	1.3	0.73
Number of SMs	16	14
Cores per SM	32	192
Peak DP flops [TFlop/s]	0.665	1.31
Memory bus width	384	384
Data rate [Gbit/s]	3.7	5.2
Memory bandwidth [GByte/s]	178	250
Number or registers	32,768	65,536
L1 cache per SM [kBytes]	64	64
L2 cache [kBytes]	768	1536

Mitglied der Helmholtz-Gemeinschaft

Features of New Kepler Architecture (cont.)

L2 Atomics

- Naïve locking mechanisms conflicts with warp divergence
- Increased number of faster atomic processors

Dynamic parallelism

- Ability to launch new work from GPU

HyperQ

- Allow multiple CPU tasks to use same device

Mitglied der Helmholtz-Gemeinschaft

Multi-GPU Parallelisation

Today: CUDA-aware MPI

- Data transport device ↔ host managed by MPI library
  - Data continues to be routed through host
- Implementations: MVAPICH, OpenMPI, ...

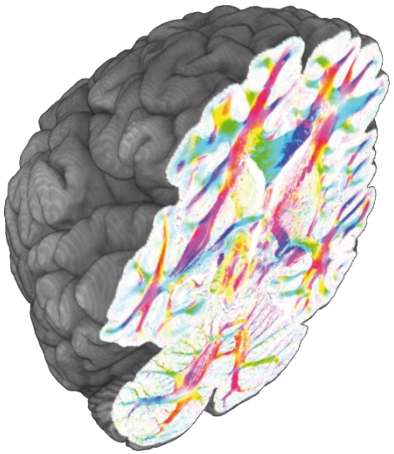
Outlook: GPUDirect RMDA

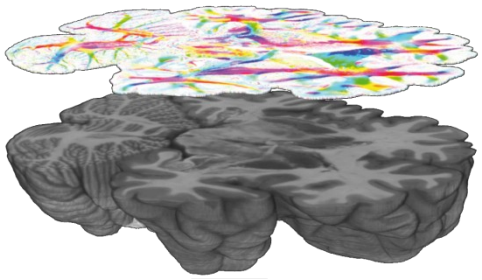
- Direct data transport device ↔ IB HCA
- Early evaluation: small message MPI latency reduces from 21 to 6.5 µs [DK Panda, 2013]
- Limitations:
  - Communication in dual-socket Xeon system
  - Lack of fine-grained communication models

Application: JuBrain

Application developed at the Institute of Neuroscience and Medicine (INM-1) at Forschungszentrum Jülich

Katrin Amunts, Markus Axer, Marcel Huysegoms





Research goal

Accurate, highly detailed computer model of the human brain

Mitglied der Helmholtz-Gemeinschaft

JuBrain Challenges

Complexity of the brain

- Billions of nerve cells
- 10^4 times larger number of interconnects

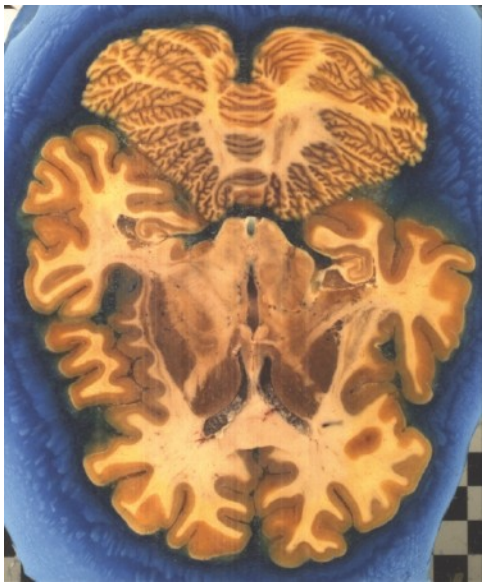
Building an atlas requires thousands of brain sections

- High-resolution histological data
- Need re-alignment of 2D images into a coherent 3D model

☞ **Major computational problem**

Mitglied der Helmholtz-Gemeinschaft

Brain Section Images

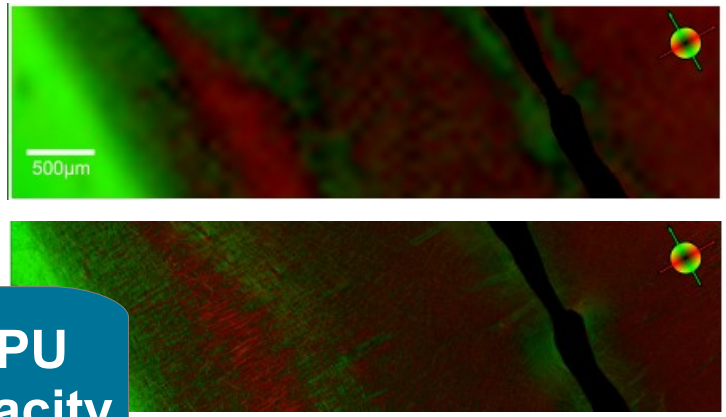


Block-face pictures

- Created while cutting brain in sections

Histological images

- Polarized light images
- Low resolution vs. high resolution
  - *100 µm → 3 µm pixel size*
  - *30 MBytes → 40 Gbytes data*

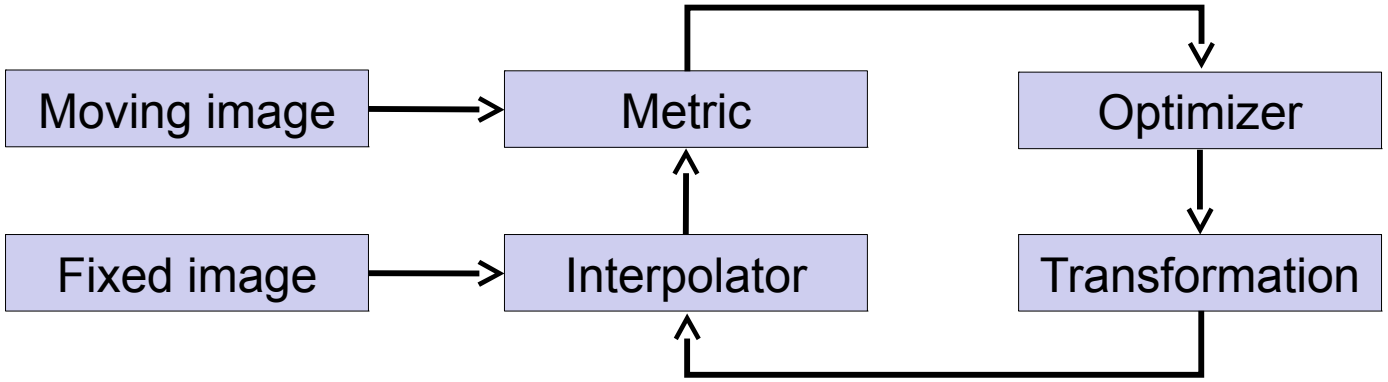


Computational task: Image registration

Exceeds GPU memory capacity

Mitglied der Helmholtz-Gemeinschaft

Image Registration



Registration algorithms

Rigid registration → 3 parameters
Affine registration → 6 parameters
Elastic registration → O(100) parameters

Mitglied der Helmholtz-Gemeinschaft

Metric computation

Metric = mutual information

- Metric based on Shannon entropy $H = \sum_i p_i \log \frac{1}{p_i}$
- Need to compute probability of certain pixel values to occur
- Maximize mutual information ↔ minimize joint entropy

Determination of joint histogram → joint distribution

- Pixels can be processed in parallel
- But: Update of common data structure ☞ atomic operations
- Significant performance improvements observed on Kepler

Parallelization challenges

- Irregular communication
- Load imbalance

Mitglied der Helmholtz-Gemeinschaft

Application: Lattice Boltzmann Method

Application developed at U Rome Tore Vergata/INFN, U Ferrara/INFN, TU Eindhoven

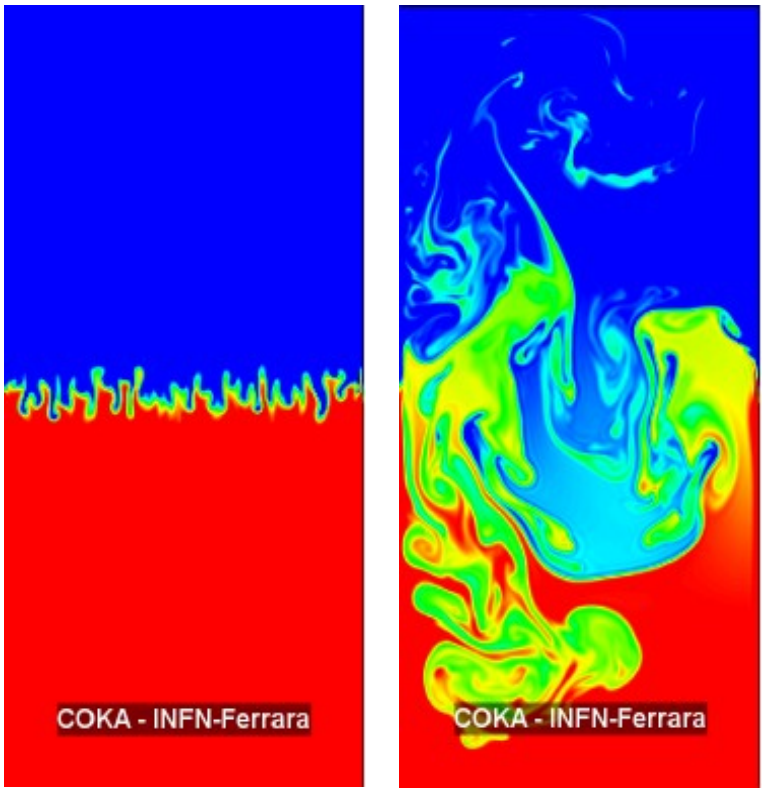
F. Schifano (U Ferrara) et al.

UNIFE

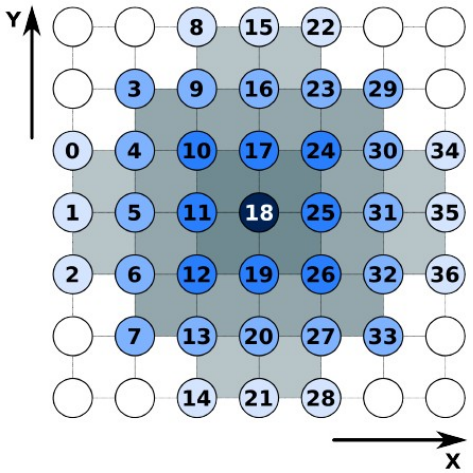
Research goal

Simulation of Rayleigh-Taylor instability

- Need sufficiently large systems
- Requires
  - Double precision computation
  - Many GPUs



Lattice Boltzmann Method



Model ansatz

- Reproduce dynamics of fluid by simulating virtual particles which collide and propagate
- Evolution equation:

$$f_l(\boldsymbol{x} + \boldsymbol{c}_l \Delta t, t + \Delta t) - f_l(\boldsymbol{x}, t) = -\frac{\Delta t}{\tau}\left(f_l(\boldsymbol{x}, t) - f_l^{(eq)}\right)$$

- D2Q37 model → 37 populations

Kernels

- Collide = evolve populations
- Propagate = collect populations

Mitglied der Helmholtz-Gemeinschaft

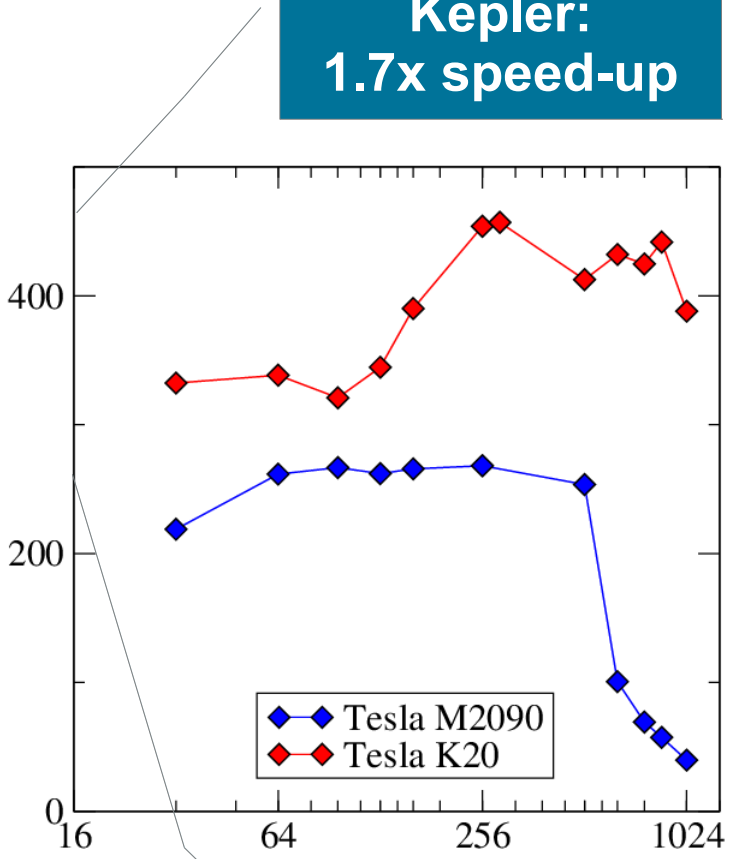
Collide Kernel: Fermi vs. Kepler

Kernel characteristics

- About 7,600 DP Flops per site, load/store 37 DP values
- Arithmetic operations dominate
  → Relevant metric = performance [GFlop/s]

Performance analysis observation

- Significant increase of L1 cache misses: 17% → 67%
- SM performance increased, but L1 cache capacity remained unchanged



Mitglied der Helmholtz-Gemeinschaft

Collide Kernel Optimization

Root cause for cache misses

L1 use due to memory indirection

```
for (i = 0; i < NPOP-1; i++) {
  popTemp = localPop[i];
  u = u + param_cx[i] * popTemp;
  v = v + param_cy[i] * popTemp;
}
```

Problem mitigation by simple code change

Enforce loop unrolling to eliminate indirect memory accesses

```
#pragma unroll
for (i = 0; i < NPOP-1; i++) {
  popTemp = localPop[i];
  u = u + param_cx[i] * popTemp;
  v = v + param_cy[i] * popTemp;
}
```

Mitglied der Helmholtz-Gemeinschaft

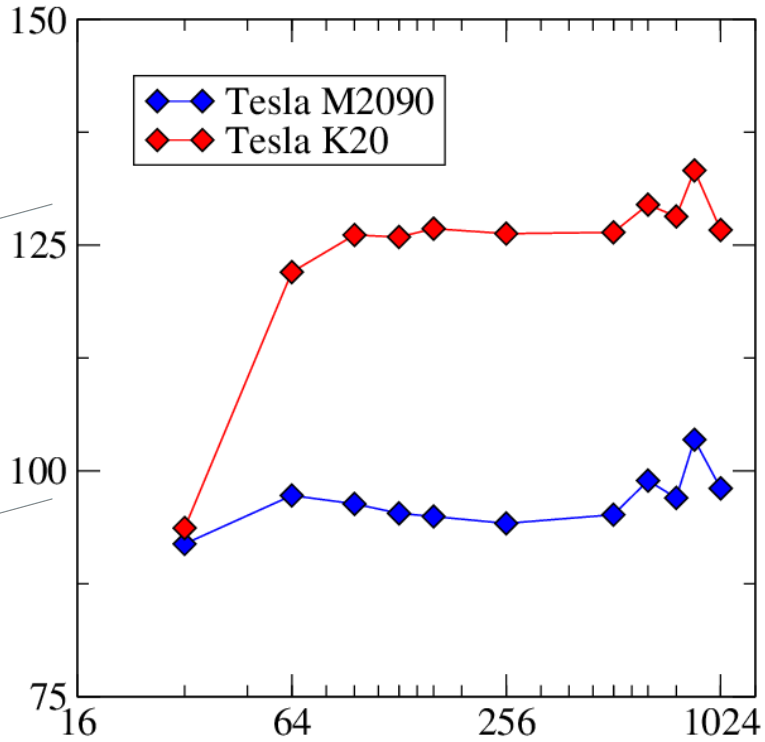
Propagate Kernel

Kernel characteristics

- Memory bandwidth limited
  → Relevant metric = BW [Gbyte/s]

Kepler:
1.4x speed-up

Fermi:
O(50%) of peak

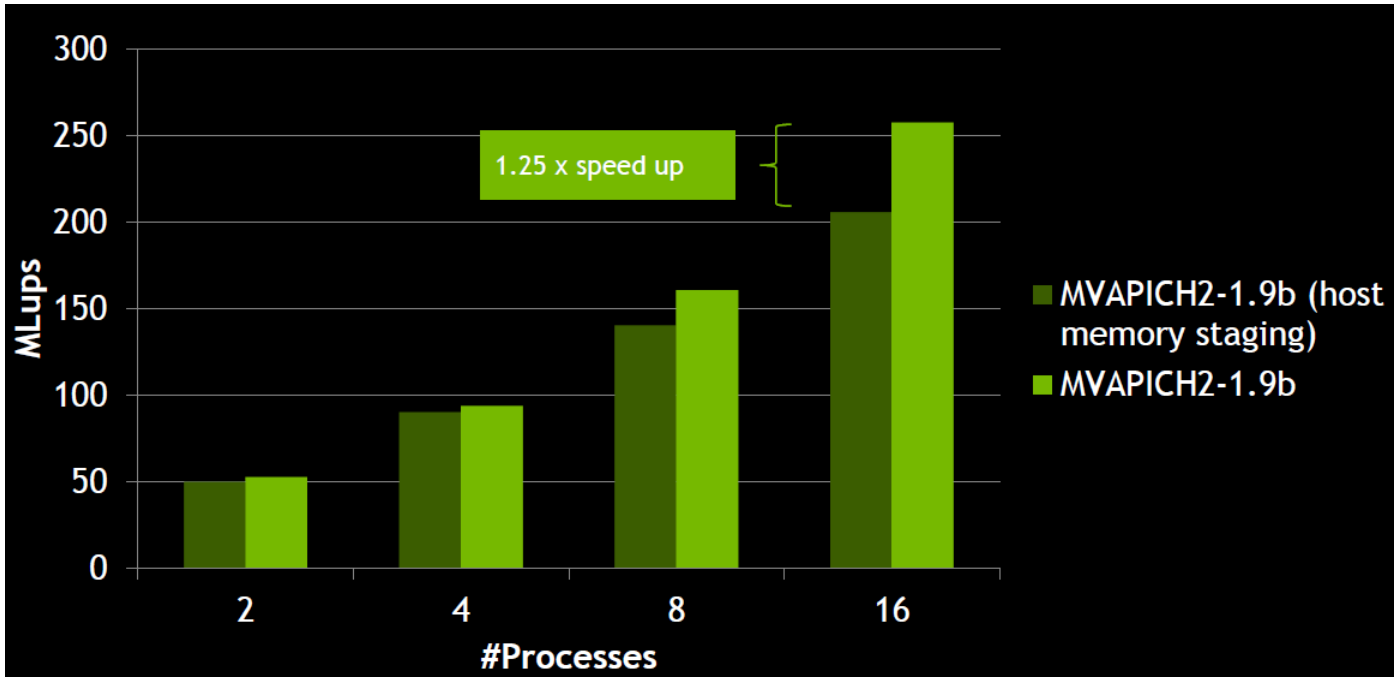


Multi-GPU Parallelization

1-dimensional parallelization

- Benefits from CUDA-aware MPI
- Reasonable parallel efficiency up to 16 GPUs

[Jiri Kraus, 2013]



Application: MAFIA

Jan Meinke,
Andrew Adinetz

Scientific goal

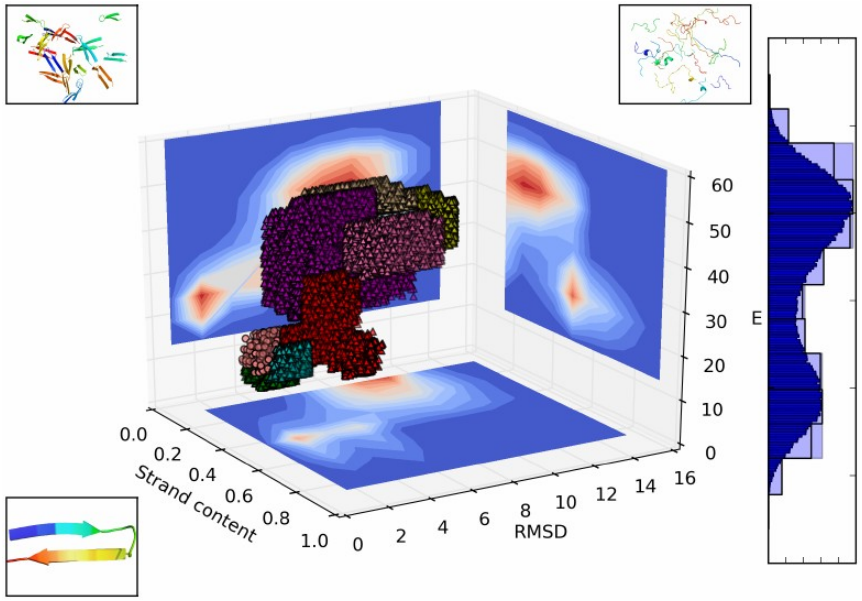
- Identify clusters in data from Monte Carlo simulations of protein folding

MAFIA

- Sub-space density clustering technique → data mining tool
- Adaptive grids

Lab focus

- GPU implementation
- Algorithmic optimisation
- Performance analysis



Mitglied der Helmholtz-Gemeinschaft

GPUMAFIA

Input data scaling

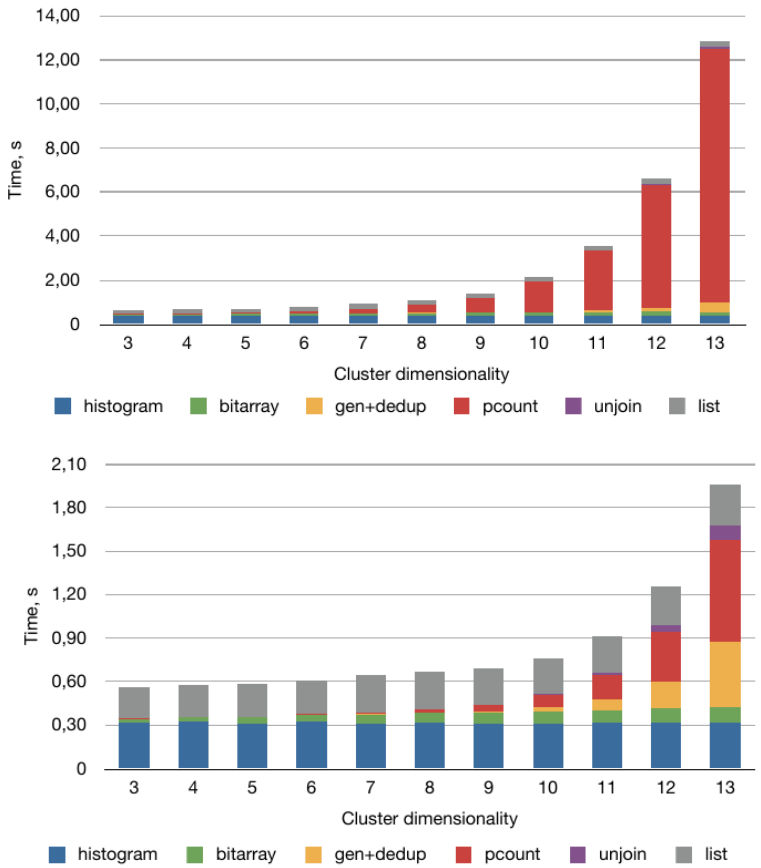
- Operation count for point-counting kernel scales ~k 2^k

Optimization

- Atomics
- Bit arrays → high data re-use

Performance comparison 16 CPU cores vs. K20x

- Speed-up allows for interactive analysis of large data sets



Mitglied der Helmholtz-Gemeinschaft

Application: Pulsar Search

Application developed at
MPI for Radio-Astronomy Bonn

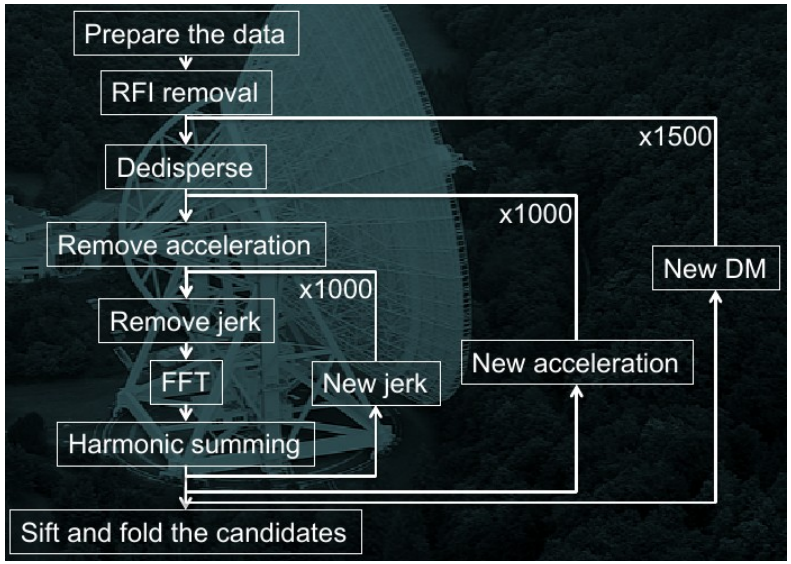
David Champion et al.

Max-Planck-Institut für Radioastronomie

Scientific goal

Discover pulsars from radio-astronomy data

- Large amount of data to be processed
- Compute intensive pipeline executed many times



Kernels

- FFT ☞ efficient on GPU
- Harmonic summing

Application: HEP Trigger

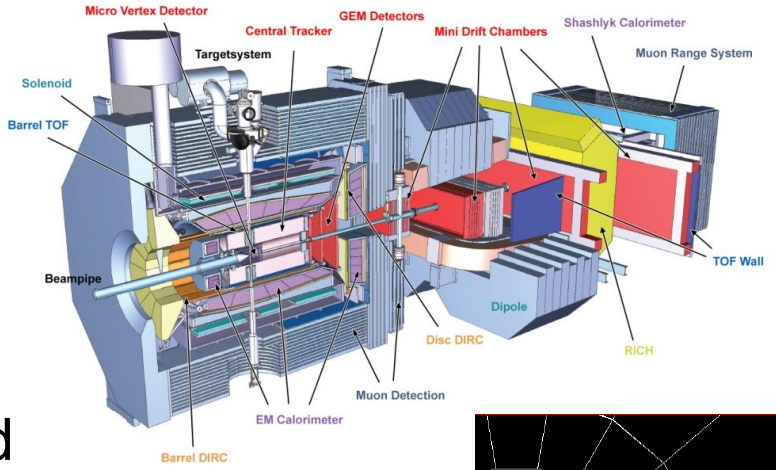
Application developed at FZJ (IKP-1)

Andreas Herten et al.

Scientific goal

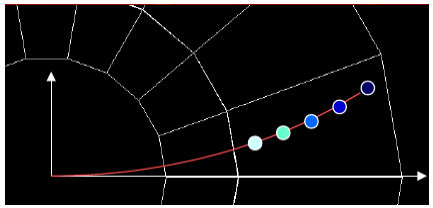
Reconstruction of particle tracks in future high-energy physics experiment

- Target: PANDA experiment at GSI
- Hough transformation or other algorithms



Advantages of GPU

- Easier to program compared to, e.g., FPGAs
- Latencies more predictable than for CPUs



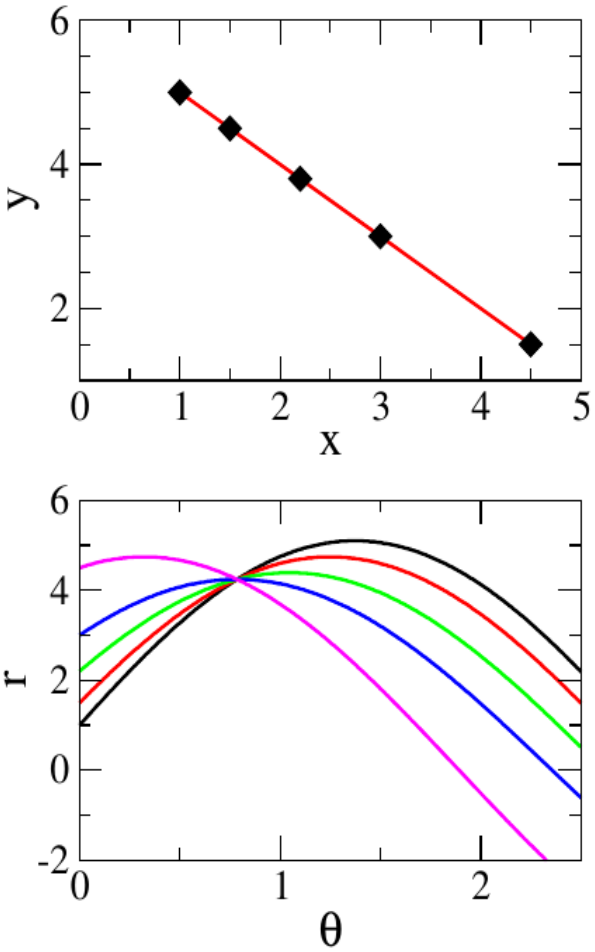
Mitglied der Helmholtz-Gemeinschaft

Hough Transformation

Technique to identify lines

$$y = a + b\,x$$
$$r(\theta) = x \cos\theta + y \sin\theta$$

- Requires construction of histogram to determine intersection of curves in Hough space ☞ **atomic operations**
- Use conformal mapping to identify circles



Adaptive Hough transformation

- Start with coarse grid and dynamically refine grid ☞ **dynamic parallelism**
- Significant performance improvements for very fine grids

Mitglied der Helmholtz-Gemeinschaft

Summary and Conclusions

- **NVIDIA Lab: new form of collaboration**
  - Investigation of application requirements
  - Analysis of hardware/software usability
- **Challenges to exploit increasing performance**
  - Multi-GPU parallelization
  - Use of new hardware features
- **Work driven by applications:**
  - Life sciences: JuBrain, MAFIA
  - Physics: Lattice Boltzmann, HEP trigger, pulsar search
- **Exploited new Kepler features**
  - Atomics
  - Dynamic parallelism

Mitglied der Helmholtz-Gemeinschaft