Lattice QCD simulations using new high-performance computing systems

Mario Schröck
for the Lattice QCD group @ RM3

Roma, 05/05/2014

Introducing myself

- 2004-2010: Diploma studies of Physics at Universität Tübingen (Germany)

- 2009/2010: Diploma thesis at Institut für Theoretische Physik, Universität Tübingen with Prof. H. Reinhardt:

  "*Coulomb Gauge Quark Propagator from Lattice Quantum Chromodynamics*"

- 2010-2013: Ph.D. thesis at Universität Graz (Austria), with Prof. C.B. Lang:

  "Dynamical Chiral Symmetry Breaking and Confinement: Its Interrelation and Effects on the Hadron Mass Spectrum"

- *since 01/2014: postdoc INFN SUMA @ RM3*

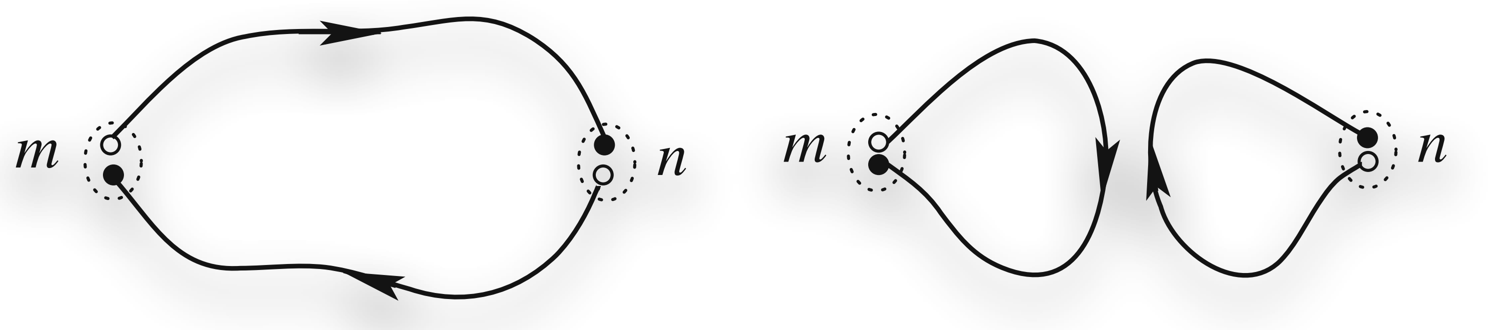
INFN SUMA grant

- The aim is the optimization and the porting of codes for lattice QCD simulations on both existing and new architectures for high-performance computing included in the INFN program SUMA for massive supercomputing

- http://web2.infn.it/SUMA

- Currently: adopting GPUs for large matrix inversions in lattice QCD

Motivation



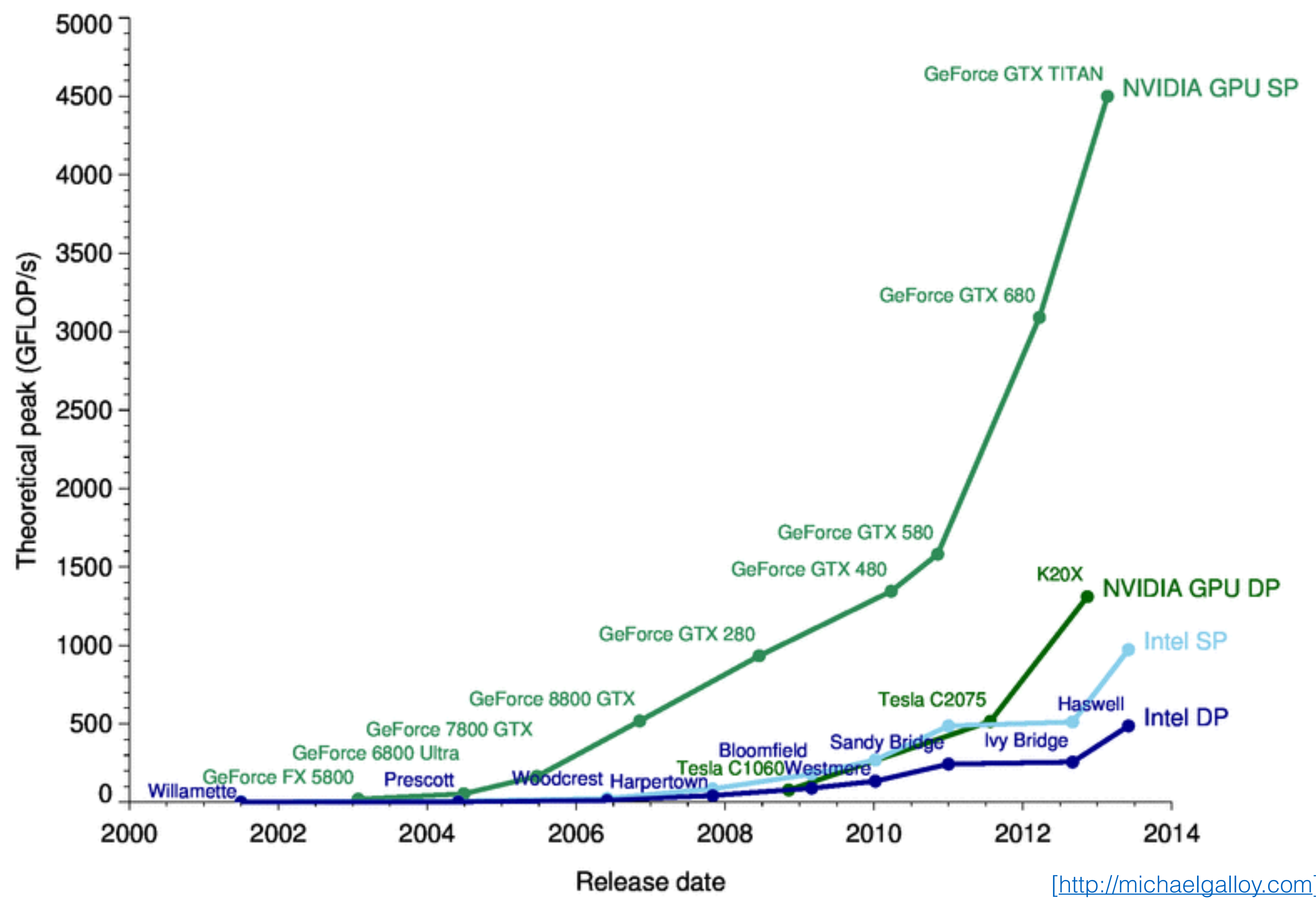
- inclusion of disconnected quark loops in lattice QCD requires *ab initio* the complete inversion of a rank ≳ one million matrix

- clever algorithms lower this to 100-1000 solutions per gaugefield configuration

- costs of the inversions still highly dominate the post gaugefield generation analysis

⇒ adopt modern hardware to accelerate the inversions!

Motivation



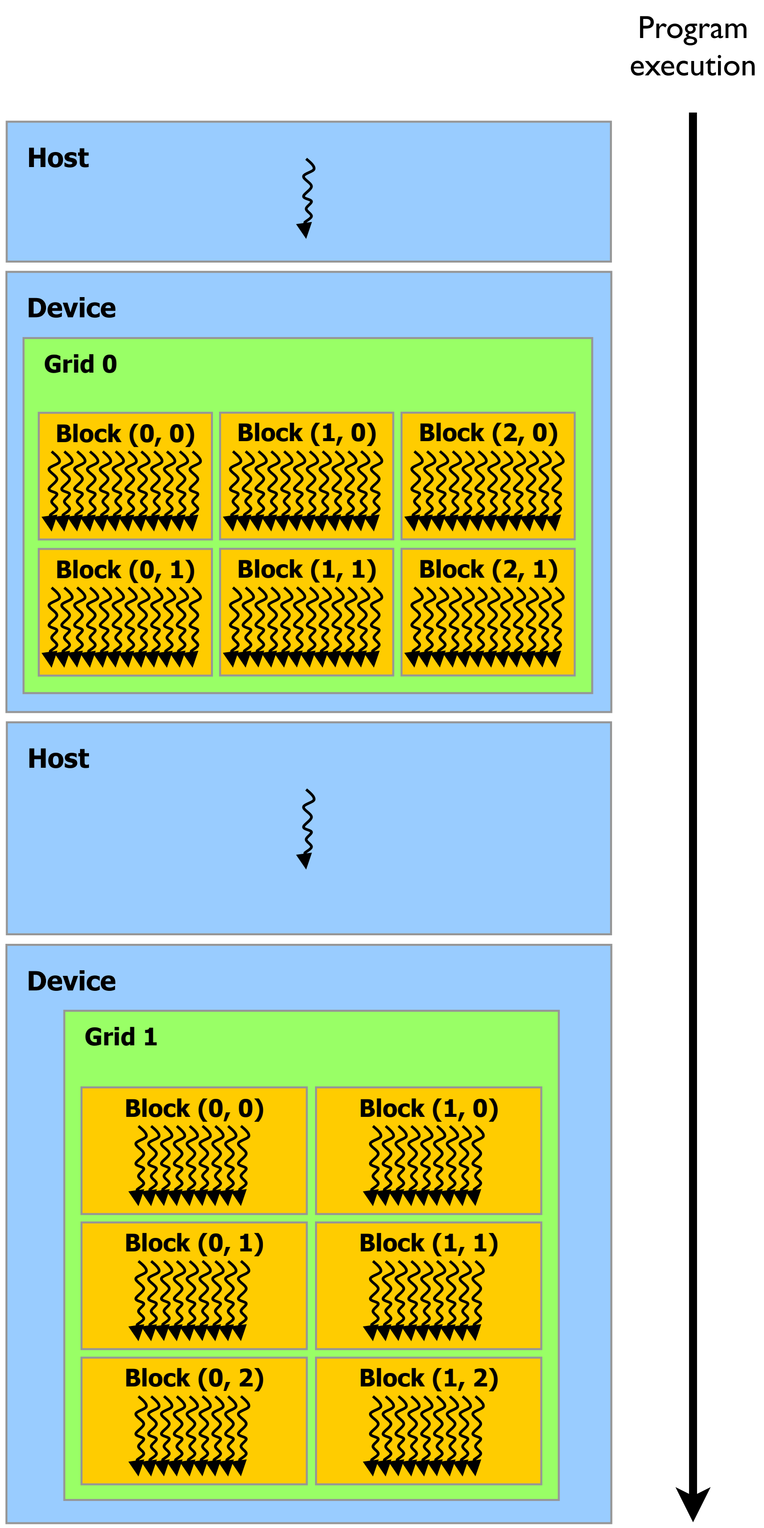
[http://michaelgalloy.com]

The NVIDIA CUDA Programming Model

- Program is executed on host system (CPU)

- host calls kernels that run on the device (GPU)

- each kernel starts many threads that perform the same work on different data, e.g., one thread per lattice site



Code optimizations

- GPU kernels in lattice QCD are bandwidth bound: transfer only 12 (or 8) parameters of each SU(3) matrix and recalculate the matrix when needed (virtually for free!)
- maximize memory throughput: reorder gaugefield/ spinorfields to allow coalesced memory accesses
- combine low precision arithmetic in the conjugate gradient with high precision reliable updates

QUDA

- “QCD on CUDA” – http://lattice.github.com/quda

- Effort started at Boston University in 2008, now in wide use as the GPU backend for Chroma, MILC, and various other codes.

- Various solvers for several discretizations, including multi-GPU support and domain-decomposed (Schwarz) preconditioners.

- *Developers*: Ron Babich (NVIDIA), Mike Clark (NVIDIA), Kip Barros (LANL), Rich Brower (Boston University), Justin Foley (University of Utah), Joel Giedt (Rensselaer Polytechnic Institute), Steve Gottlieb (Indiana University), Bálint Joó (JLab), Claudio Rebbi (Boston University), Guochun Shi (NCSA -> Google), Alexei Strelchenko (FermiLab), Frank Winter (JLab)

Eurotech Eurora @ Cineca

Model: Eurora prototype

Architecture: Linux Infiniband Cluster

Processors Type:

- Intel Xeon (Eight-Core SandyBridge) E5-2658 2.10 GHz (Compute)
- Intel Xeon (Eight-Core SandyBridge) E5-2687W 3.10 GHz (Compute)
- Intel Xeon (Esa-Core Westmere) E5645 2.4 GHz (Login)

Number of nodes: 64 Compute + 1 Login

Number of cores: 1024 (compute) + 12 (login)

Number of accelerators: 128 nVIDIA Tesla K20 (Kepler)

RAM: 1.1 TB (16 GB/Compute node + 32GB/Fat node)

OS: RedHat CentOS release 6.3, 64 bit

CINECA

Performance comparison

Fermi: IBM Blue Gene/Q

- 10.240 nodes PowerA2 sockets @1.6GHz, 16 cores each
- #15 of (Nov.'13) TOP500 SUPERCOMPUTER SITES

Eurora: prototype

- 64 nodes, two Intel Xeon Sandybridge CPUs and two NVIDIA K20s
- #4 of Little List (Nov.'13) THE GREEN 500

Performance comparison

32^3 x 64 lattice: twisted mass inverter in double precision

Fermi

- nissa code (DP/SP)
- 128 nodes (2048 cores)
- 0.009898 sec./iter.

Eurora

- QUDA (DP/HP, 12 recon.)
- 1 node (2 GPUs)
- 0.017016 sec./iter.

Performance comparison

32^3 x 64 lattice: twisted mass inverter in double precision

Fermi

- nissa code (DP/SP)
- 128 nodes (2048 cores)
- 0.009898 sec./iter.

Eurora

- QUDA (DP/HP, 12 recon.)
- 1 node (2 GPUs)
- 0.017016 sec./iter.

ratio = 0.58

Performance comparison

32^3 x 64 lattice: twisted mass inverter in double precision

Fermi

- nissa code (DP/SP)
- 128 nodes (2048 cores)
- 0.009898 sec./iter.

Eurora

- QUDA (DP/HP, 12 recon.)
- 1 node (2 GPUs)
- 0.017016 sec./iter.

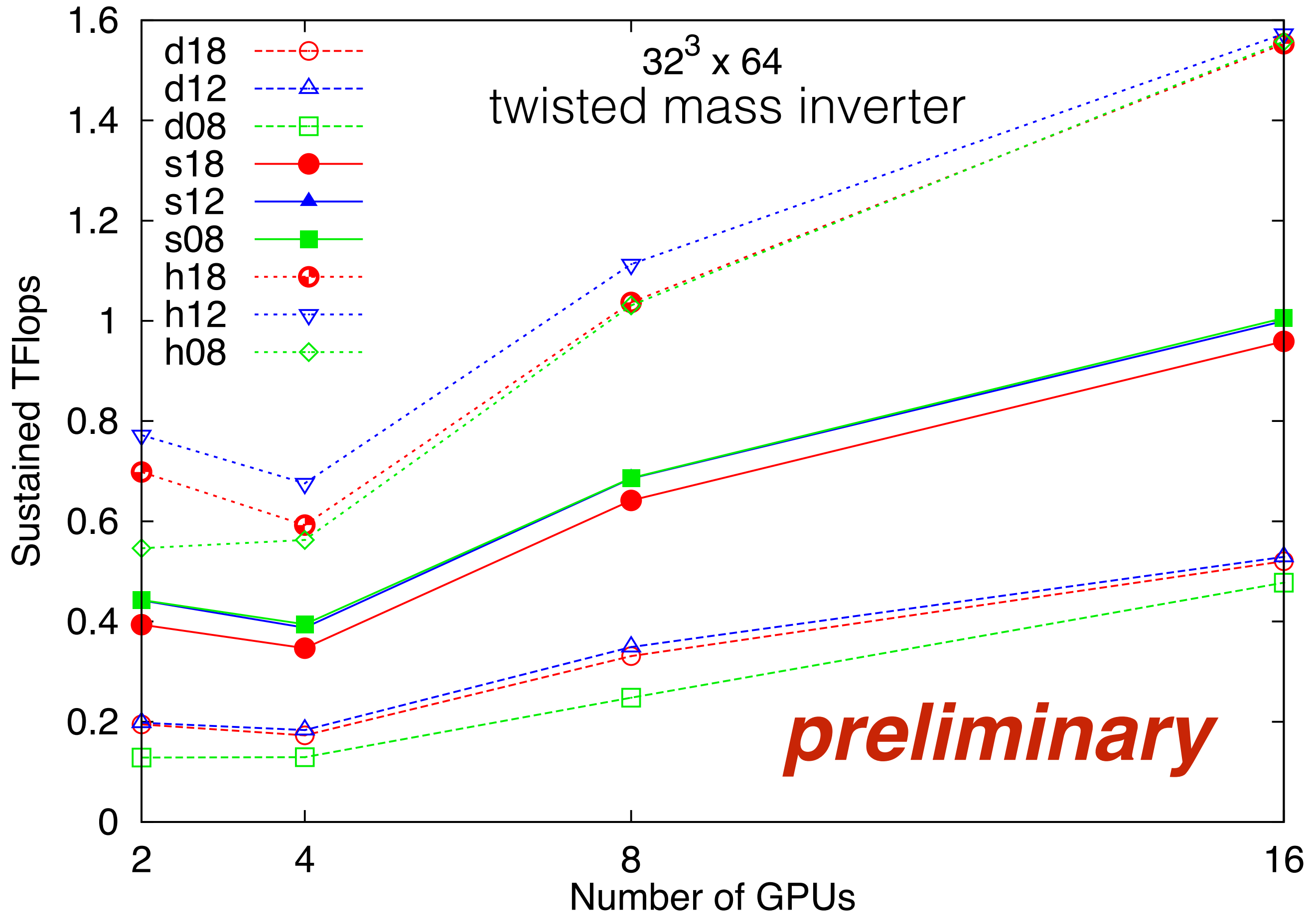
ratio = 0.58

75 Fermi nodes ~ 1 Eurora node

596 Fermi cores ~ 1 GPU

Multi-GPU on Eurora

- QUDA shown to scale up to O(100) GPUs
- in practice highly dependent on the lattice size
- hide inter GPU communication behind calculations in the inner part of the domain



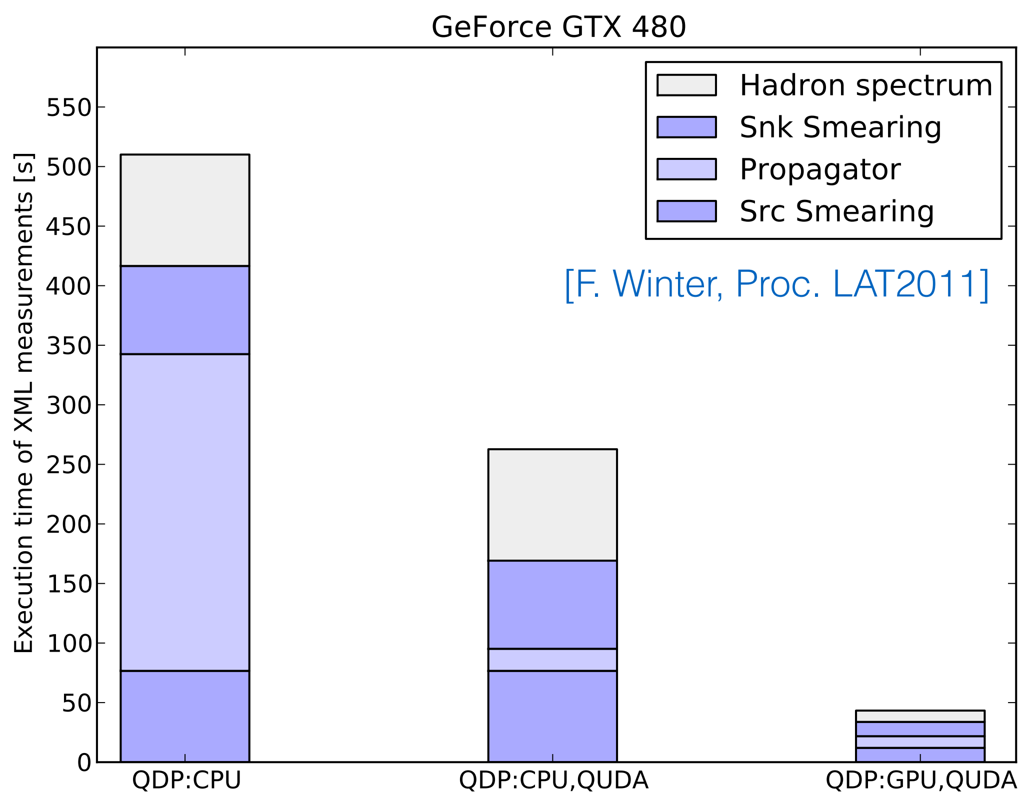
Status of the project

- some final steps in the integration of the QUDA inverter into our code have to be completed
- then we will start to use the Eurora cluster for production, in particular to study
  - isospin breaking effects in QCD+QED
  - nucleon sigma term

Outlook

- Amdahl's law states that a program fraction *P* subject to acceleration with the according program part sped up by a factor *S* gains a total speedup factor of

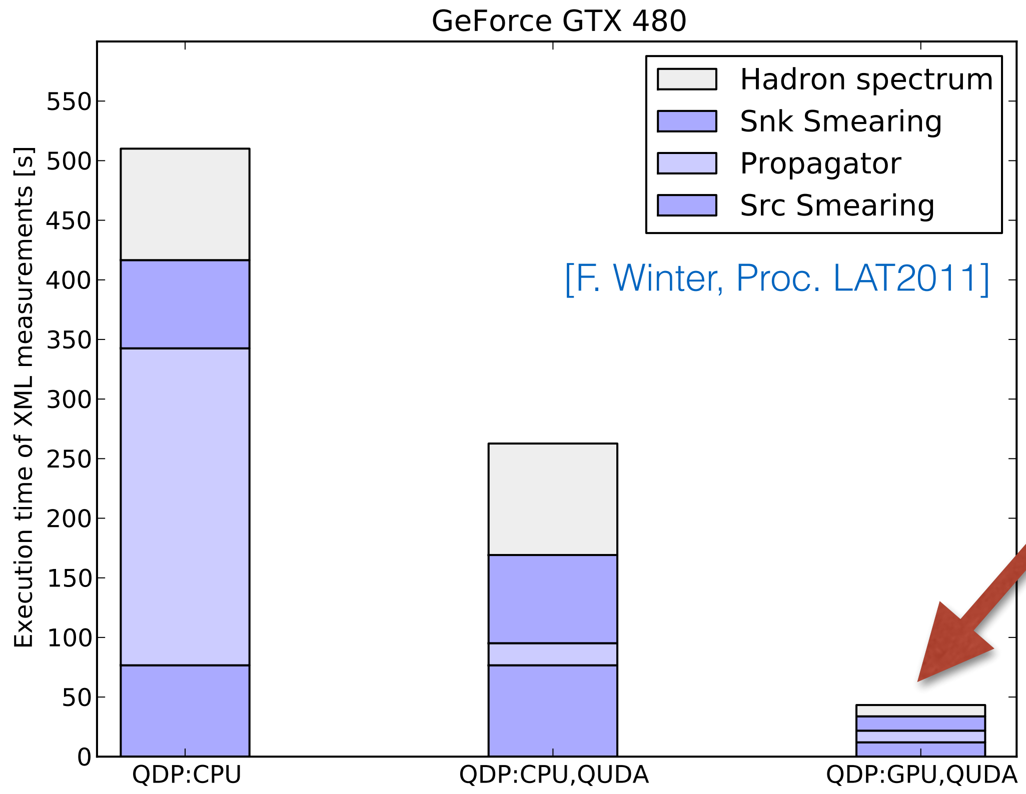
$$S_{\text{total}} = \frac{1}{(1-P)+\frac{P}{S}}$$



Outlook

- Amdahl's law states that a program fraction *P* subject to acceleration with the according program part sped up by a factor *S* gains a total speedup factor of

$$S_{\text{total}} = \frac{1}{(1-P)+\frac{P}{S}}$$



perform smearing and contraction on the GPU as well