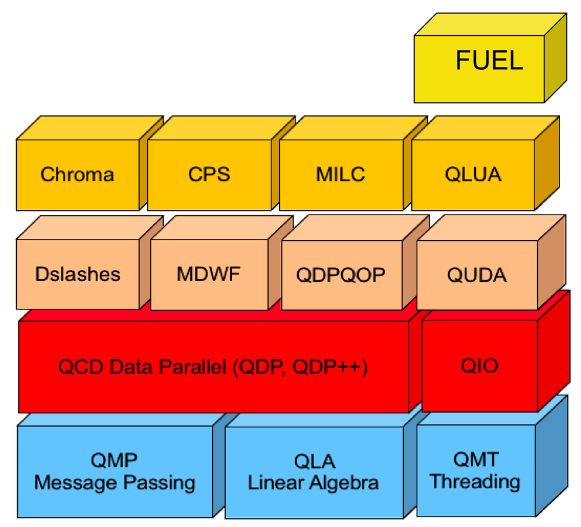
Multi-scale Lattice Field Theory in SciDAC Software

Presented by Rich Brower (software co-ordinator)

QUDA (QCD in CUDA) library started in 2008 with NVIDIA's CUDA implementation by Kip Barros and Mike Clark at Boston University. It has expanded to a broad base of USQCD SciDAC **[1]** software developers and is in wide use as the GPU backend for HEP and NP SciDAC application codes: Chroma, CPS, MILC, etc.



Provides:

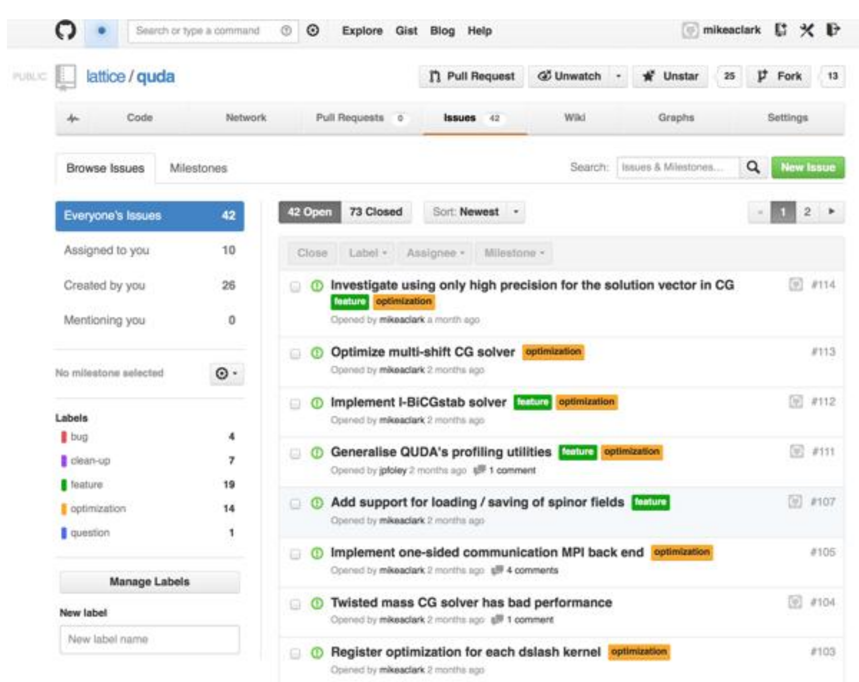
- Various solvers for several discretizations, including multi-GPU support and domain-decomposed (Schwarz) preconditioners
- Additional performance-critical routines needed for gauge-field generation

Maximize performance:

- Exploit physical symmetries
- Mixed-precision methods
- Autotuning for high performance on all CUDA-capable architectures
- Cache blocking

"QCD on CUDA" team – http://lattice.github.com/quda

- Ron Babich (NVIDIA)
- Kip Barros (LANL)
- Rich Brower (Boston University)
- Michael Cheng (Boston University)
- Mike Clark (NVIDIA)
- 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 (Cyprus Institute -> FNAL)
- Hyung-Jin Kim (BNL)
- Frank Winter (UoE -> Jlab)



Data Compression: Local Memory Reduction

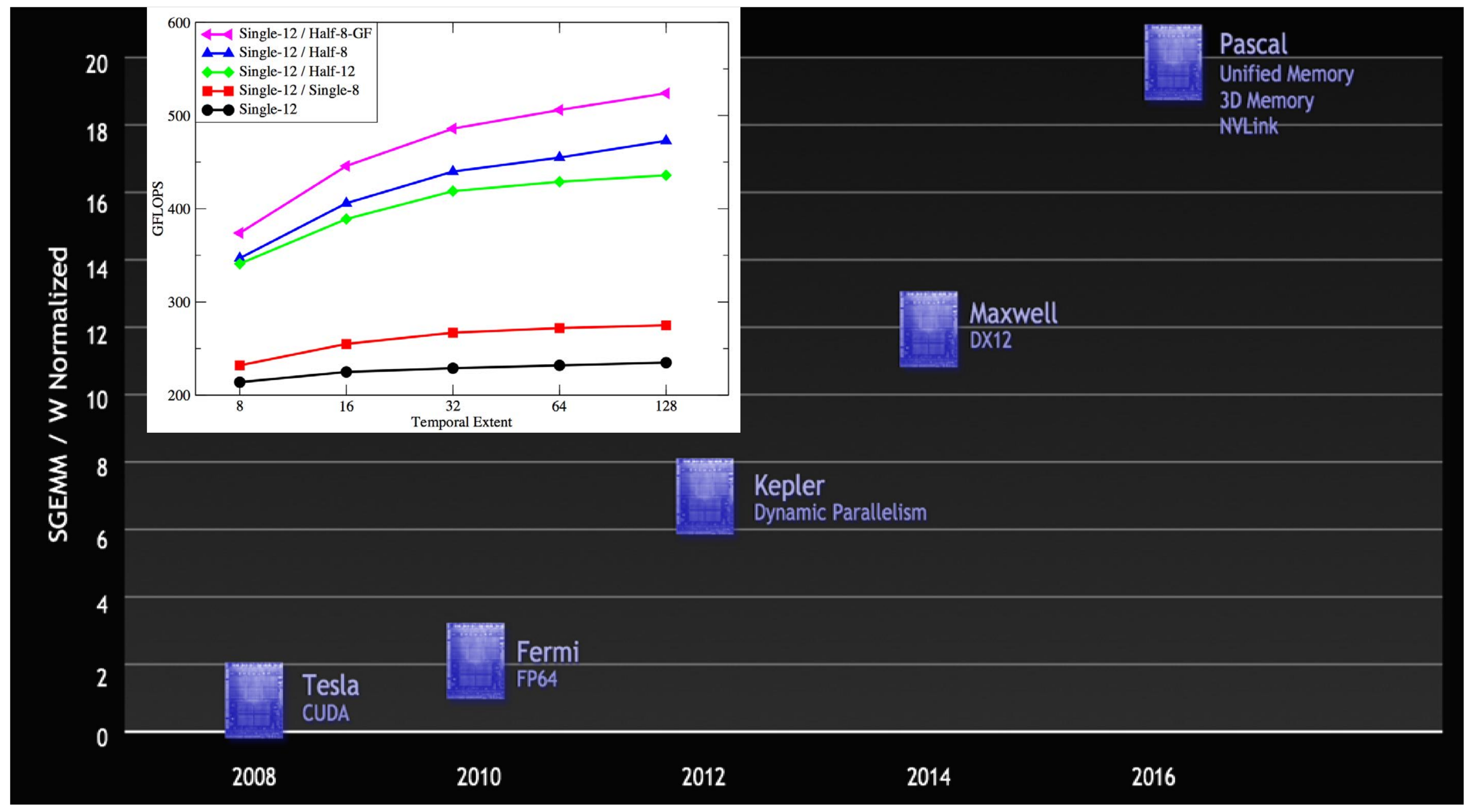
- **SU(3) matrices are all unitary complex matrices with det = 1**
  - **12-number parameterization: reconstruct full matrix on the fly in registers**

$$\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} \longrightarrow \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix} \quad \mathbf{c} = (\mathbf{a} \times \mathbf{b})^* \qquad \text{Group Manifold:} S_3 \times S_5$$

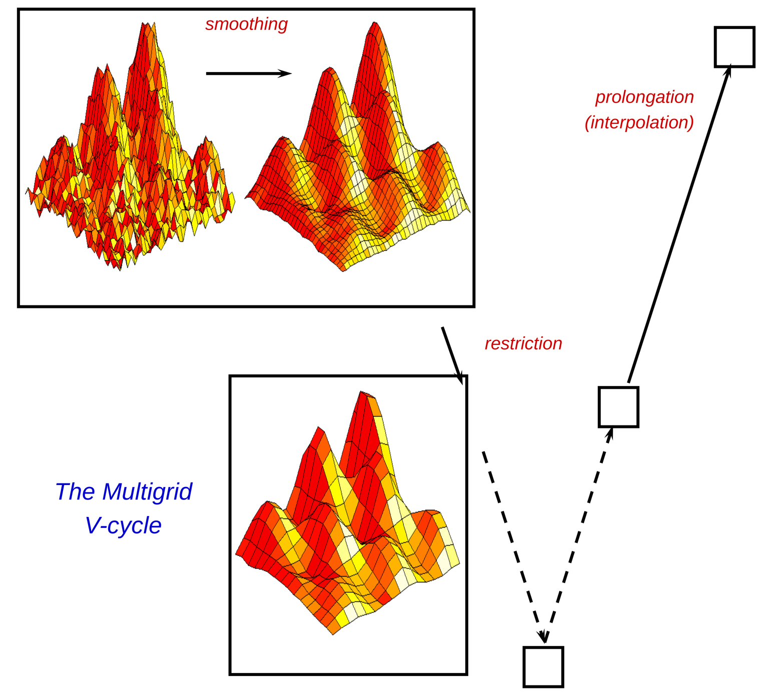
- **Additional 384 flops per site**
- **Also have an 8-number parameterization of SU(3) manifold (requires sin/cos and sqrt)**
- **Impose similarity transforms to increase sparsity**
- **Still memory bound - Can further reduce memory traffic by truncating the precision**
  Use 16-bit fixed-point representation
  No loss in precision with mixed-precision solver
  Almost a free lunch (small increase in iteration count)

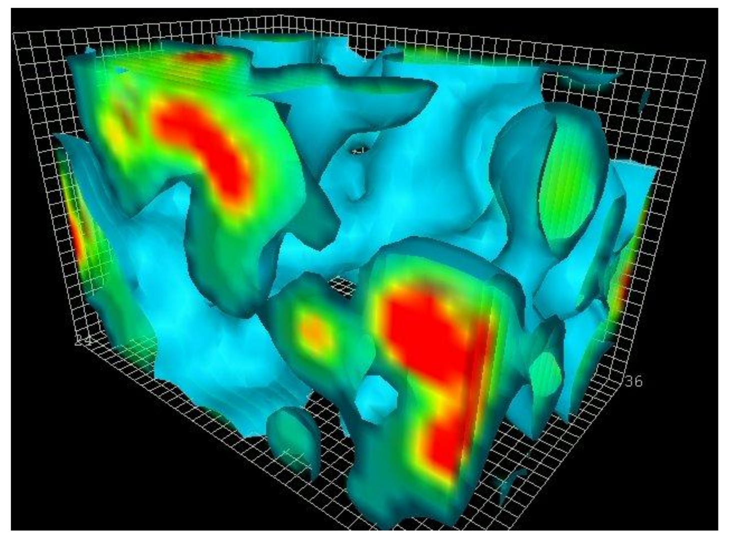
K20X performance
$V = 24^3 \times T$
Wilson-Clover is ±10%
BiCGstab is -10%

Kepler Wilson-Solver Performance



Multi-scale Physics ➔ Multilevel Solvers

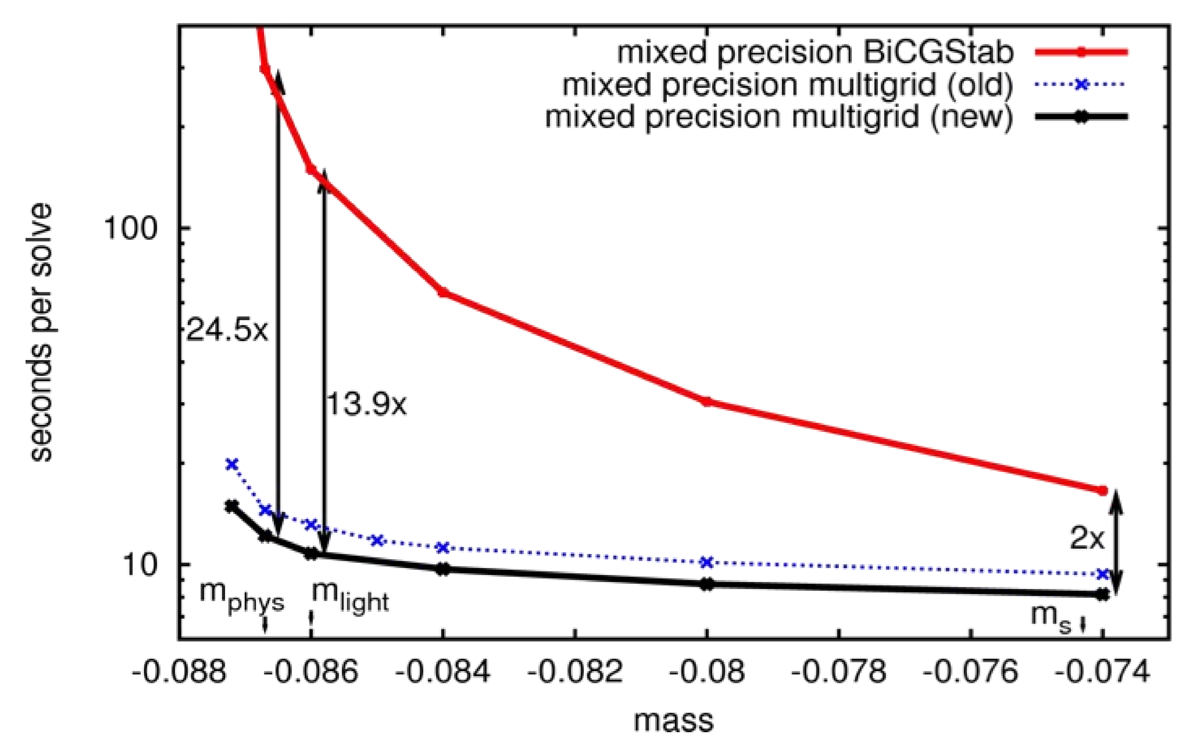




Multiple Distance scales of the "Chromodynamic Plasma" in which quarks propagate.

Adaptive Smooth Aggregation Multigrid

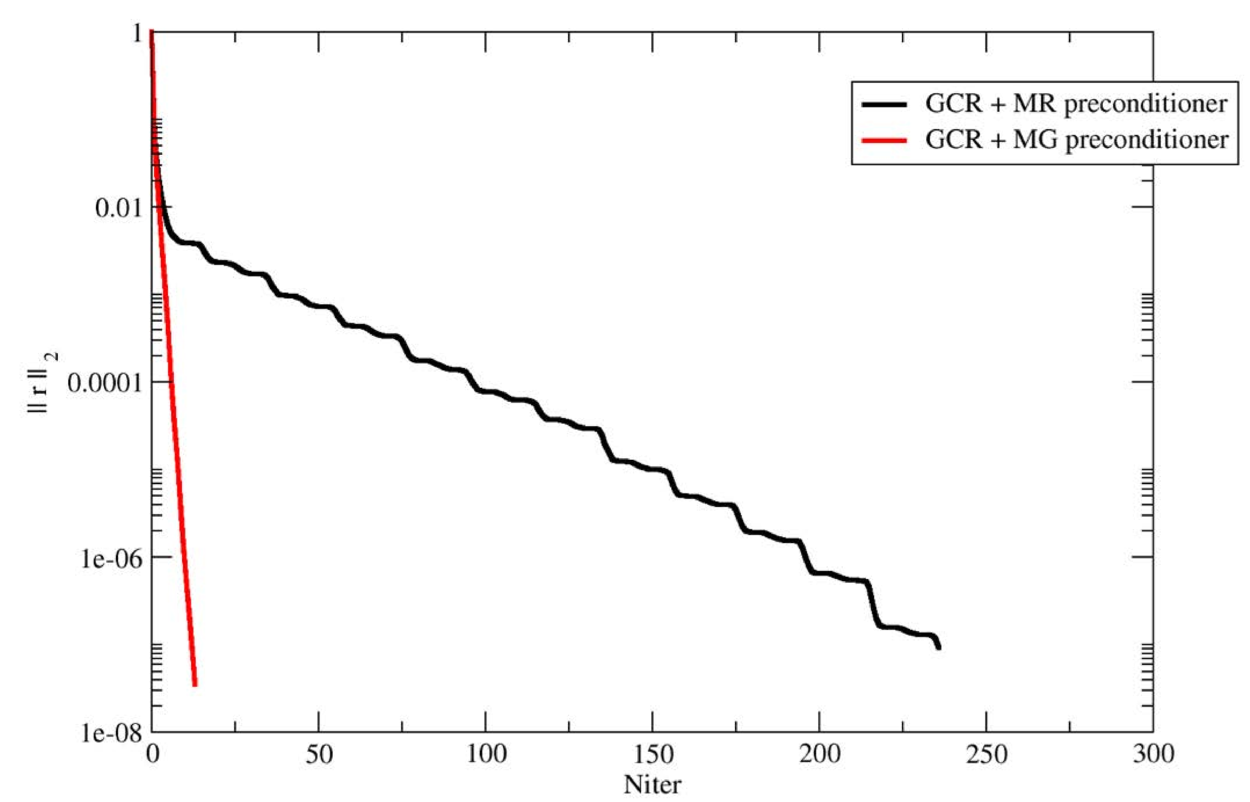
32^3x256 aniso clover on 1024 BG/P cores



20 Years of QCD MULTIGRID
In 2011 **Adaptive SA MG [3]** successfully extended the 1991 **Projective MG [2]** for algorithm to long distances.

Performance on BG/Q [3]

MG on Accelerators



GPU technology + MG ➔ Reduce $ cost by over a factor of 1/100.
(GPU/MG project: Rich Brower, Michael Cheng and Mike Clark report at Lattice 2014)

References

[1] *SciDAC-3 HEP:* ***Searching for Physics Beyond the Standard Model: Strongly-Coupled Field Theories at the Intensity and Energy Frontiers***

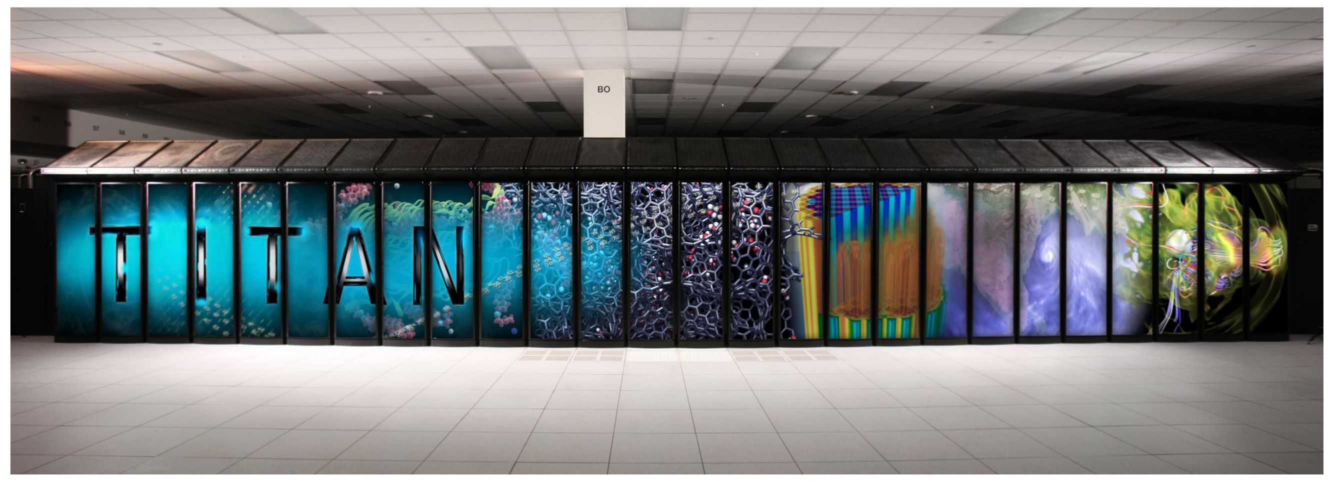
[2] *R. Brower, R. Edwards, C. Rebbi, E. Vicari,* ***Projective Multigrid for Wilson Fermions.*** Nucl.Phys. B366 (1991)

[3] *R. Babich, J. Brannick, R. Brower, M. Clark, T. Manteuffel, S. McCormick, J. Osborn, C. Rebbi, Adaptive* ***Multigrid algorithm for the lattice Wilson-Dirac operator*** *Phys. Rev. Lett. 105, 201602 (2010)*

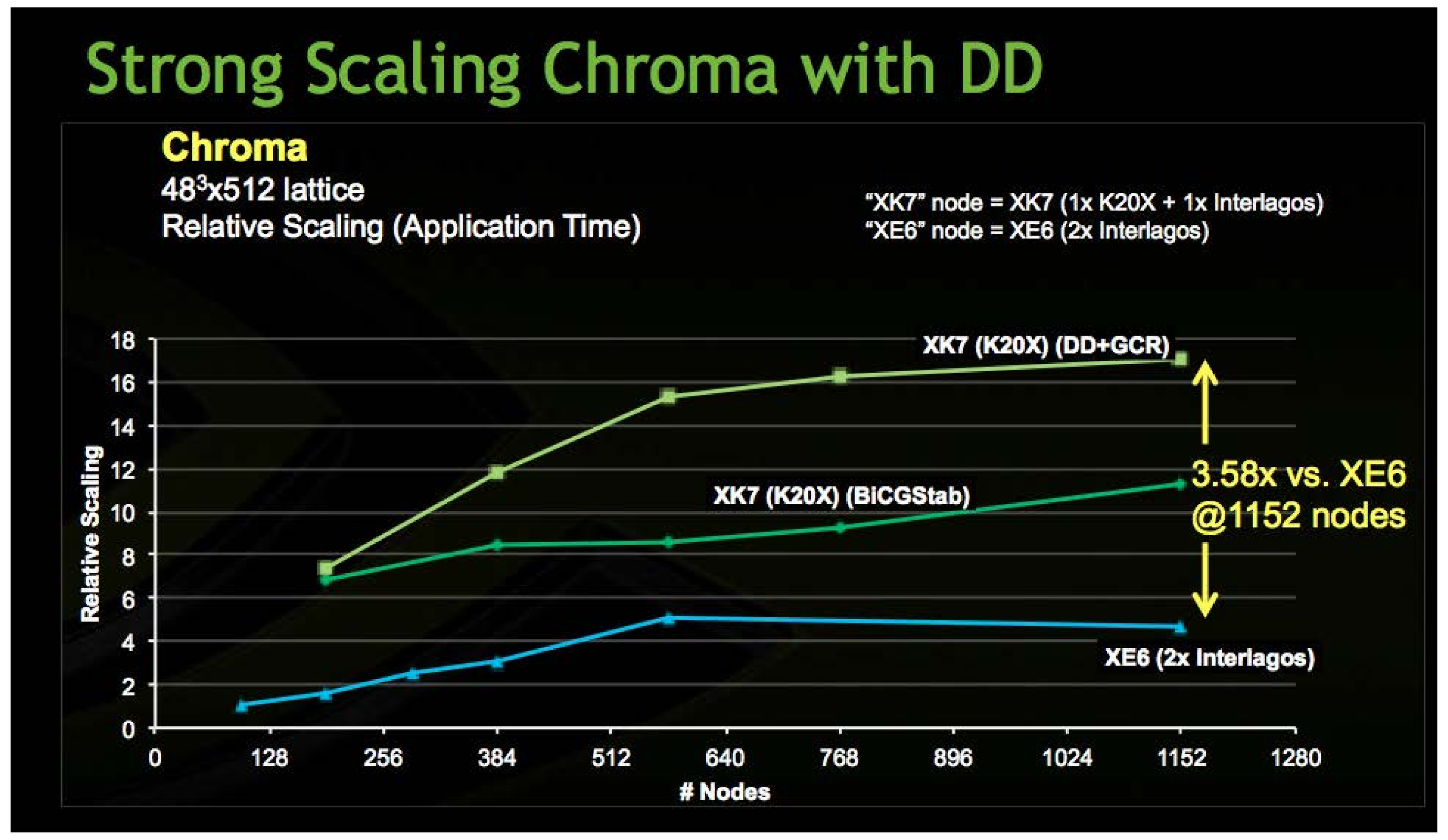
[4] *R. Babich, M. A. Clark, B. Joo, G. Shi, R. C. Brower and S. Gottlieb,* ***Scaling Lattice QCD beyond 100 GPUs*** *Super Computing 2011, arXiv:1109.2935 [hep-lat].*

[5] **Lua: an extensible embedded language http://en.wikipedia.org/wiki/Lua_%28programming_language%29**

Hierarchical Architecture ➔ Multilevel Solvers

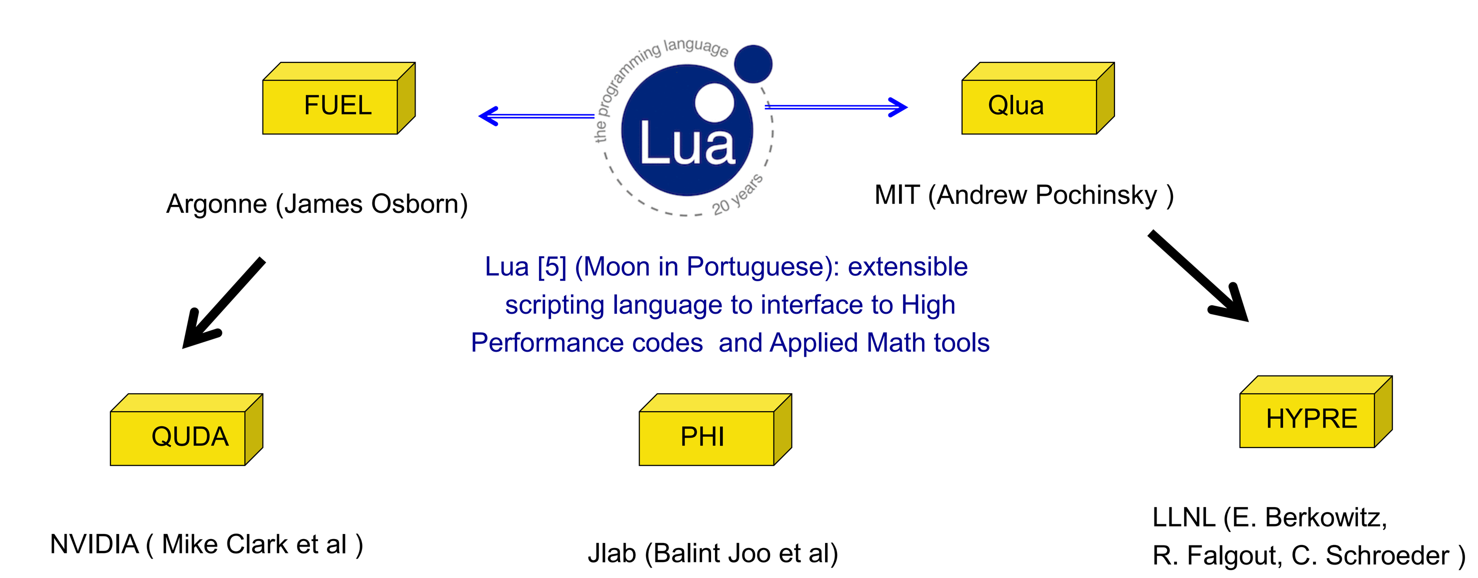


Domain Decomposition: Network Data Reduction



Block Jacobi Domain Decomposition **[4]**

Future: Hierarchical Algorithmic Frameworks (rapid prototypes, code opt and auto-tuning)



Hardware Targets: May you live in interesting times!

Need new research to adapt multi-scale physics to hierarchical (multi-scale) computers. Need compromise and auto-tuning to bring them into a happy marriage!