The USQCD Collaboration

Overview

USQCD Collaboration

http://www.usqcd.org

Paul Mackenzie

Fermilab

mackenzie@fnal.gov

USQCD/FastMATH/SUPER SciDAC-3 workshop
Livermore
November 10-11, 2011

The USQCD Collaboration

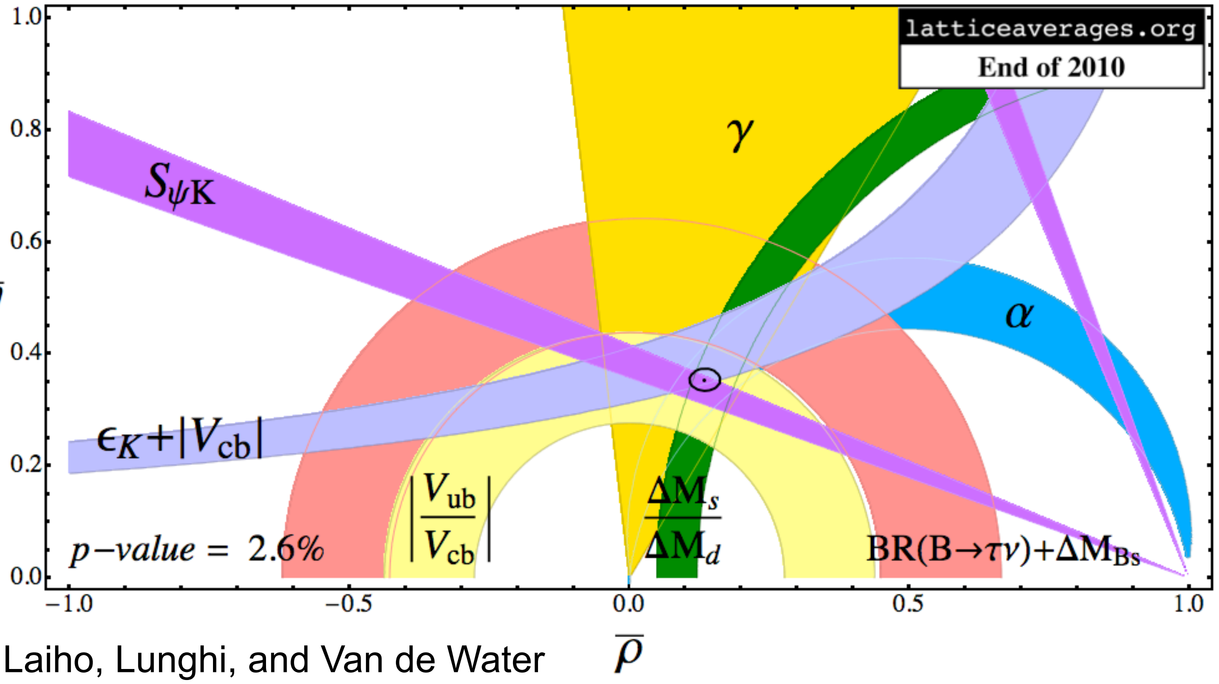
- Organizes computing hardware and software infrastructure for lattice gauge theory in the US.
- Represents almost all of the lattice gauge theorists in the US; ~ 170 people.
- Physics calculations are done by smaller component collaborations within USQCD:
  - Fermilab, HotQCD, HPQCD, JLab, LHPC, LSD, MILC, NPLQCD, RBC, ...

Major areas of physics research

Lattice QCD calculations are essential to accomplishing the physics goals of high energy and nuclear physics.

- Standard model (Fermilab, SLAC, KEK, LHC-b)
    - Hadron spectrum; determining the parameters of the standard model: the CKM matrix, the quark masses, and the strong coupling constant, ..., and searching for inconsistencies.
- Beyond the standard model (LHC)
    - Search for new particles and forces not yet discovered, ...
- Quark-gluon plasma in heavy ion collisions (RHIC)
    - De-confinement temperature; QCD plasma equation of state; transport coefficients (viscosity, ...)
- Hadronic and nuclear structure and interactions (JLab)
    - Resonance and exotics spectra, scattering lengths, and phases shifts; hadronic structure, ...

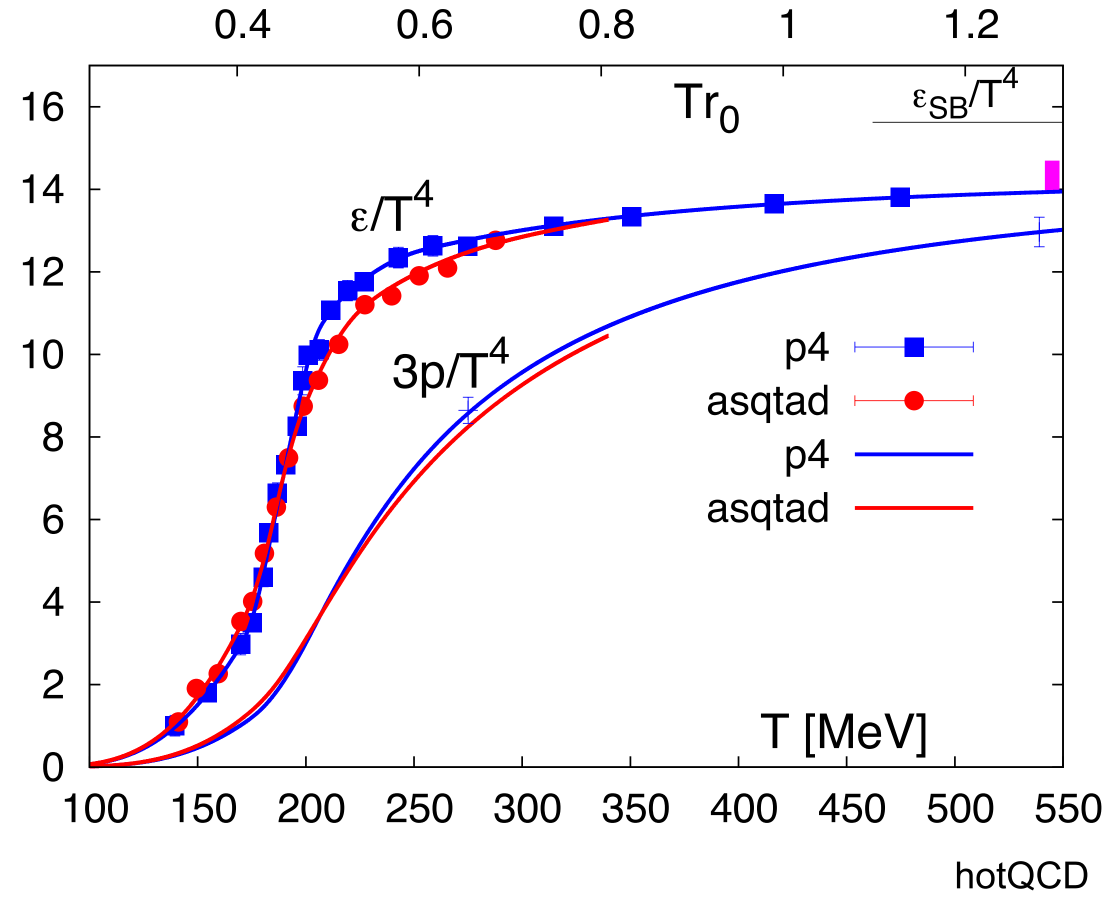
Science drivers: Fermilab and the LHC (HEP)



Laiho, Lunghi, and Van de Water

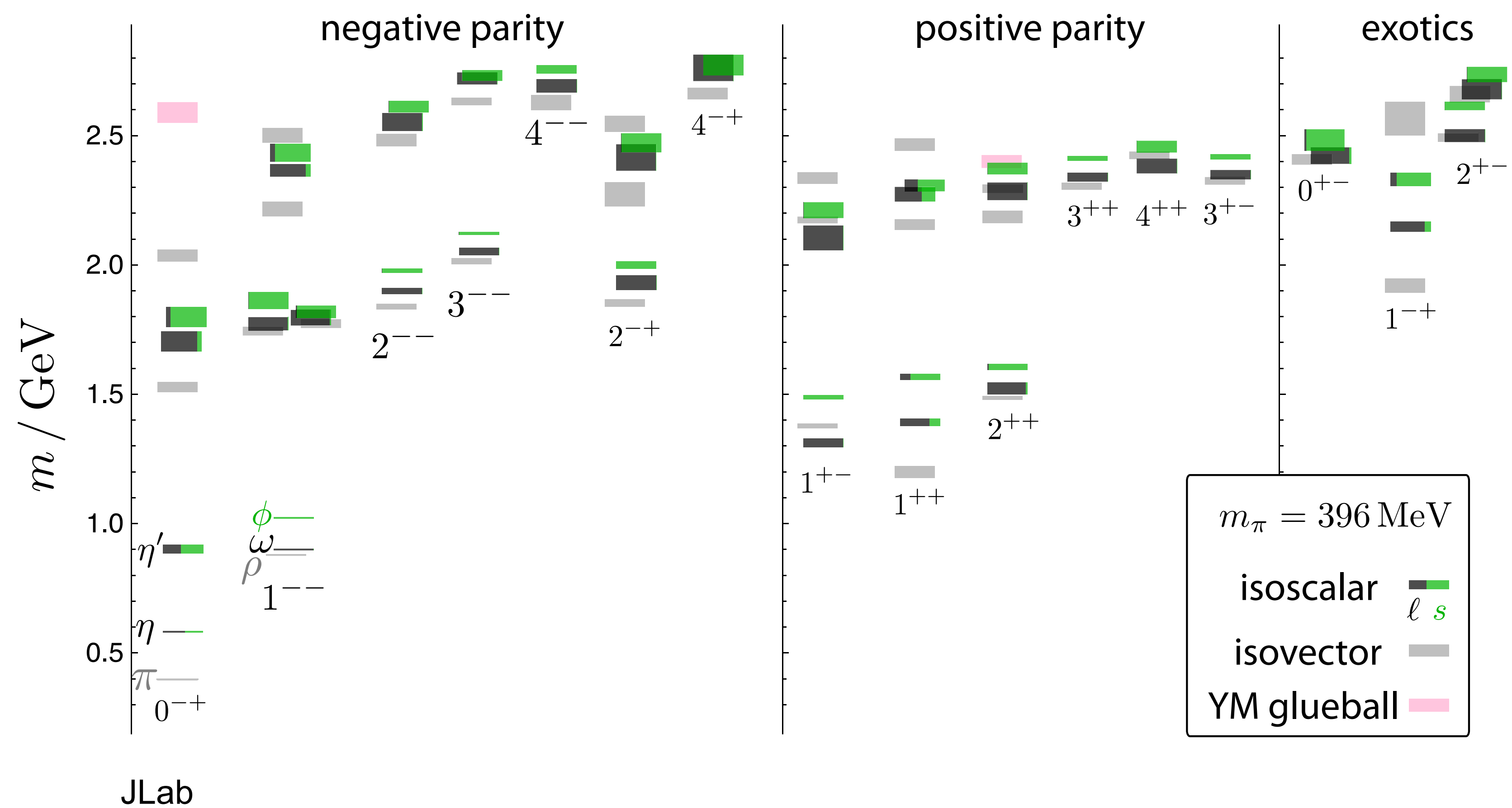
Fermilab director Pier Oddone: Lattice QCD calculations will make the data we obtain from quark factories (both electron-positron colliders as well as the Tevatron and LHC) far more useful in determining the fundamental parameters of the standard model and revealing any model inconsistencies indicative of new physics. For example, the existence of good lattice calculations allowed Fermilab's discovery of *BsBs*-bar mixing to make an important bound on the CP violating elements of the CKM matrix. Much more accurate calculations of this and other quantities are now needed to make full use of the data from the Fermilab's program

Science drivers: Brookhaven National Lab (NP)



BNL Physics Department hear Tom Ludlam: The lattice QCD calculations performed at BNL have direct relevance for the experimental program at RHIC, where an accurate determination of the equation of state of dense QCD matter with lattice gauge calculations is of central importance to the understanding of hydrodynamic properties from experimental data. In addition, we are counting on the USQCD research program to provide guidance in the search for a QCD critical point in heavy ion collisions, and an understanding of the properties of strongly interacting matter near this landmark point on the QCD phase diagram.

Science drivers: Jefferson Lab (NP)



The structure and interactions of hadrons (nucleons and mesons).

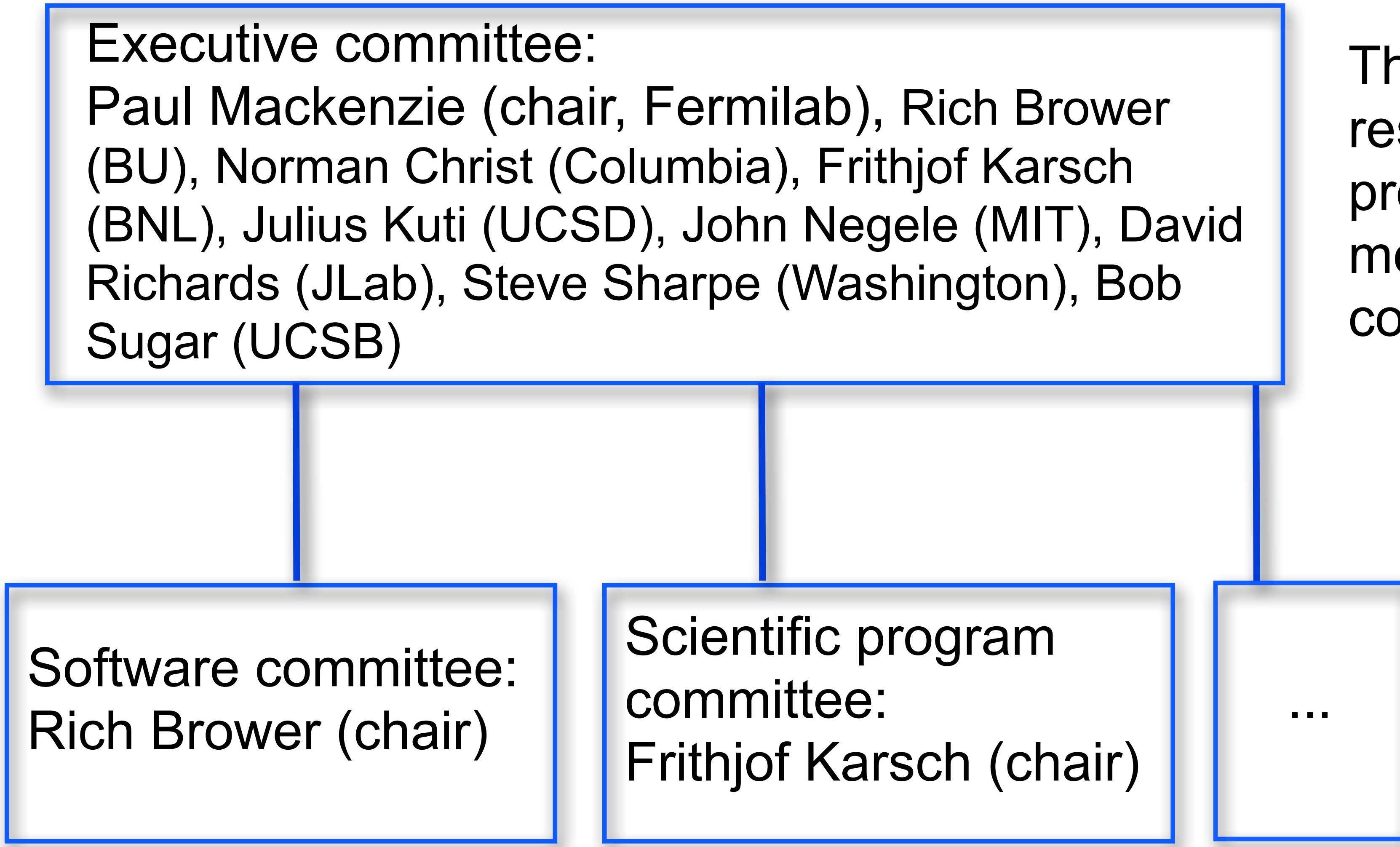
JLab director Hugh Montgomery: The national efforts of the USQCD collaboration are key to the success of the lattice program at Jefferson Lab... A continued strong national program will ensure both the algorithmic developments, and the software infrastructure, to further exploit both frontier leadership-class and special-purpose computers, and thus provide the calculations that will capitalize on the DOE investment in the Jefferson Lab experimental program.

USQCD organization

On behalf of the US lattice community, USQCD oversees:

- SciDAC software grant (OHEP, ONP, OASCR).
  - ~$2.3 M/year since 2001.
  - Creates community libraries, optimized production programs, research on new approaches (GPUs are hot now), ...
- Community INCITE grants on ASCR Leadership Computing Facilities for capability computing.
  - Last year, were allocated 67 M core-hours and used 187 M core-hours at ALCF, used 40 M core-hours at OLCF.
- Design and deployment at national labs of cost-efficient capacity hardware funded by LQCD-ext (OHEP and ONP).
  - ~ $4 M/year.
  - Infiniband clusters, now adding GPUs.

USQCD Executive Committee



The Executive Committee is responsible for writing USQCD's proposals and for appointing the members of the other committees.

USQCD Software Committee

Software Committee:
Richard Brower (chair, BU), Carleton DeTar (Utah), Robert Edwards (JLab), Rob Fowler (UNC), Donald Holmgren (Fermilab), Robert Mawhinney (Columbia), Pavlos Vranas (LLNL), Chip Watson, (JLab).

- Organizes software work done under our SciDAC grant.
- Weekly conferences calls with 12-20 people, 40 people on mailing list.
- SciDAC grant pays for less than half of our software work.
    - \$2.3 M/year, ~12 FTEs.
    - Much of the work of the software program is done by people on their regular salaries working to accomplish the goals of their physics collaborations.

USQCD timeline

USQCD Executive Committee formed.

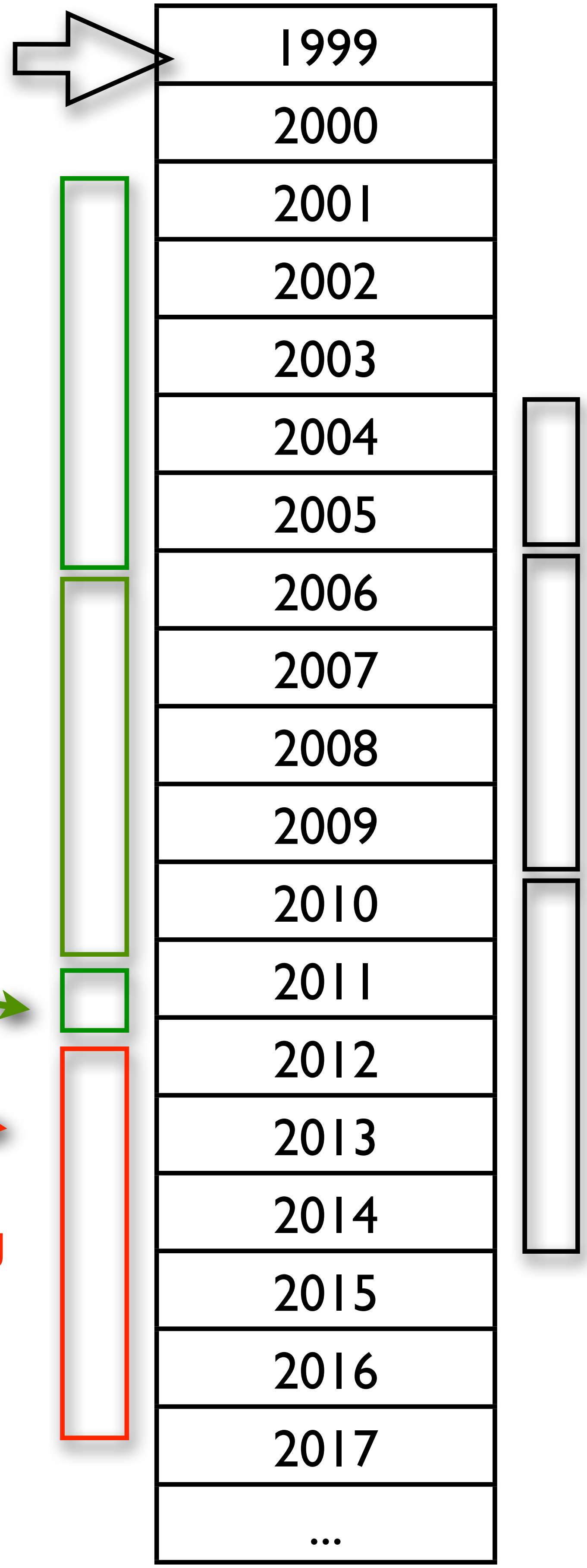
Software grants

First two five-year SciDAC grants for lattice computing R&D.

SciDAC extension

SciDAC-3?

Absolutely essential for making effective use of Leadership Computing Facilities and our dedicated hardware, and for accomplishing our physics objectives.



Hardware grants

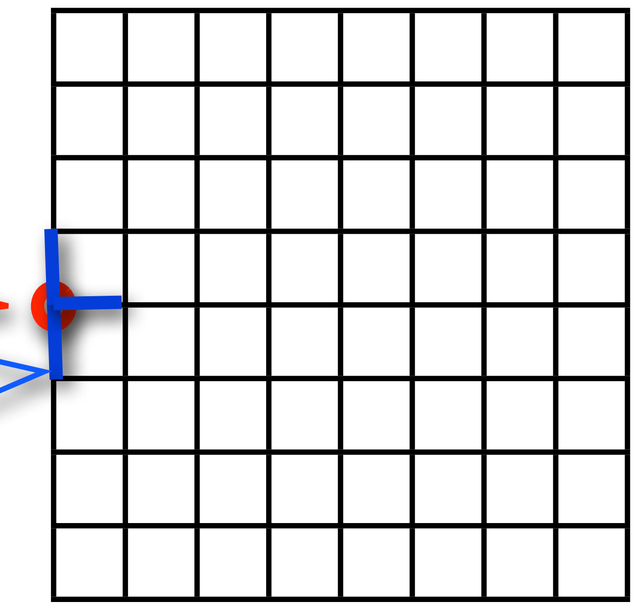
Construction of the purpose-built QCDOC.

Funding from HEP and NP for hardware through LQCD and LQCD-ext projects.

The lattice QCD computing problem

Lattice QCD approximates the continuum theory by defining the fields on a four dimensional space-time lattice.

Quarks (*complex three vectors*) are defined on the sites of the lattice, and the gauge field gluons (*complex 3x3 matrices*) on the links.

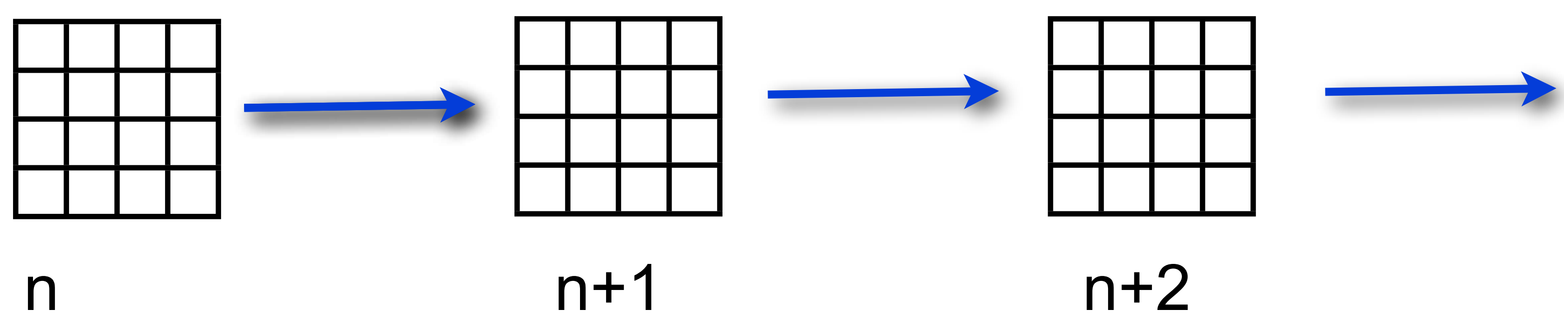


Monte Carlo methods are used to generate a representative ensemble of gauge fields. Relaxation methods for sparse matrices are used to calculate the propagation of quarks through the gauge field.

Continuum quantum field theory is obtained in the zero lattice spacing limit. This limit is computationally very expensive.

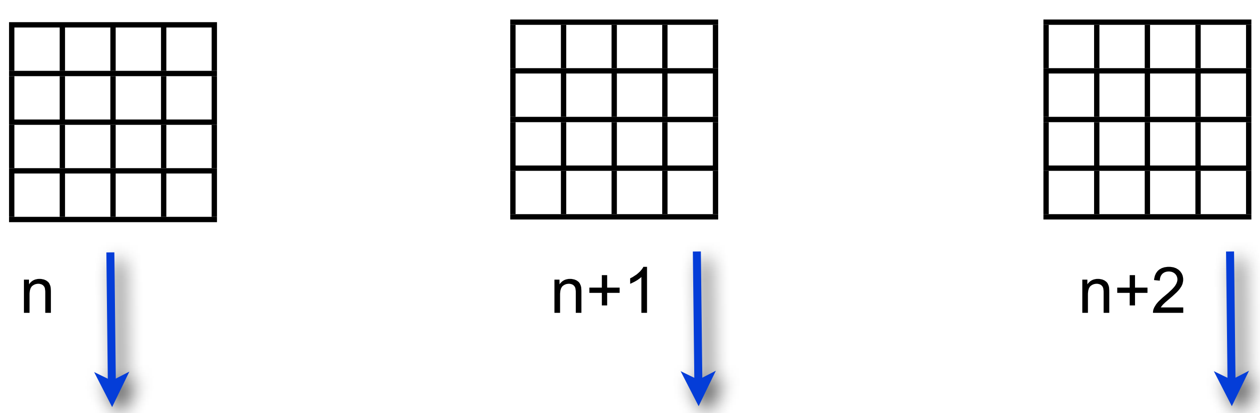
Algorithms and methods

An ensemble of gauge configurations is created with Monte Carlo methods with symplectic Hamiltonian integration. A Markov chain of configurations is made, each one from the previous.



Kostas Orginos's talk

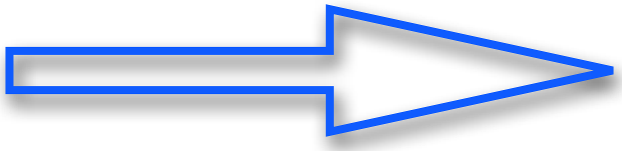
Once created, each configuration can be analyzed in parallel.



The main numerical component of both jobs is solving a sparse matrix equation $Ax=b$, with, for example, the bicongradstab algorithm.

Rich Brower's talk.

Two main components of a typical lattice calculation



TB file sizes

Generate $O(1{,}000)$ gauge configurations on a leadership facility or supercomputer center. Tens of millions of BG/P core-hours.

Transfer to labs for analysis on clusters. Larger CPU requirements.

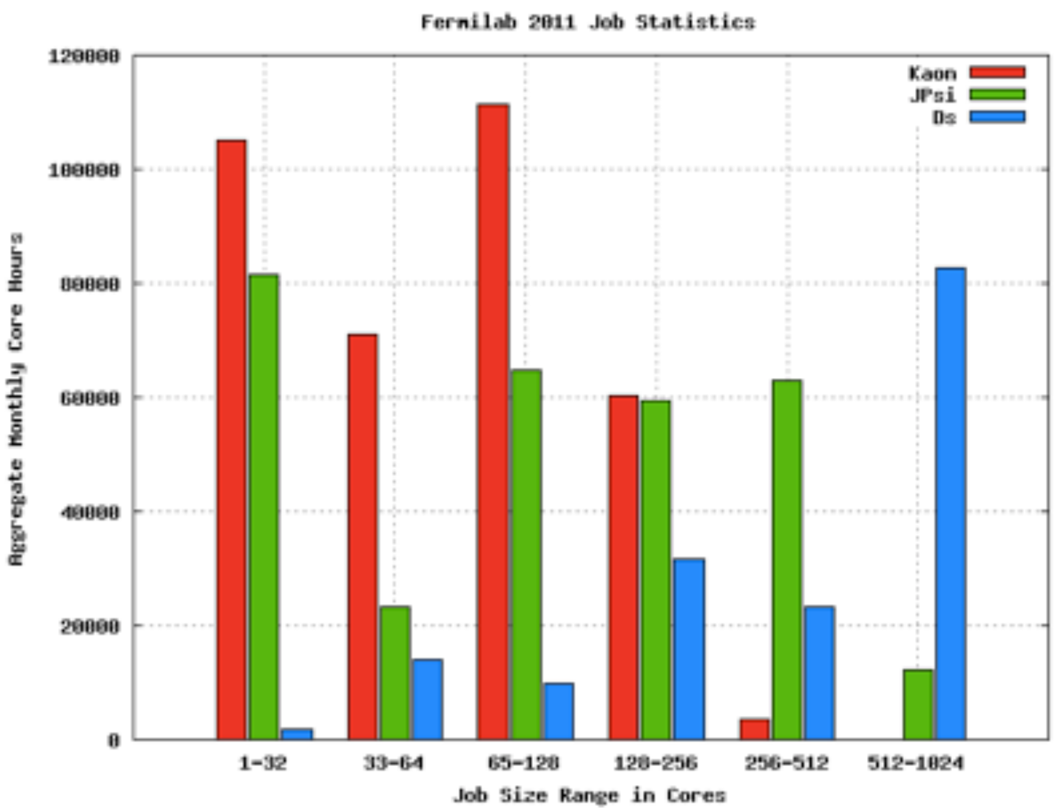
Gauge configuration generation:
a single highly optimized program,
very long single tasks,
moderate I/O and data storage.

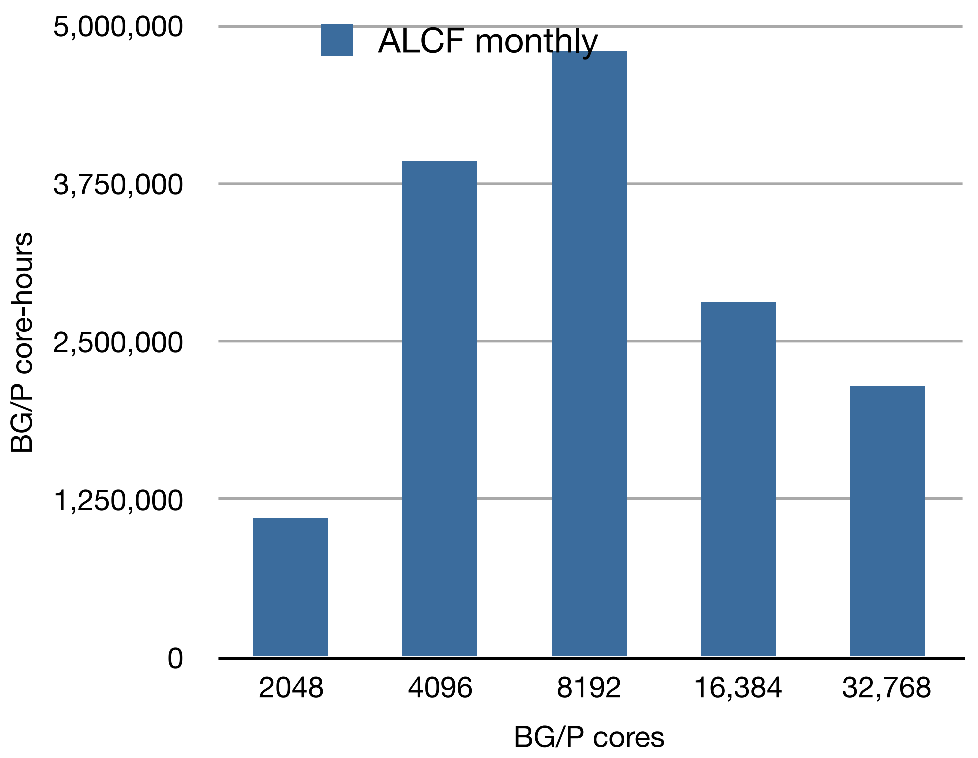
Hadron analysis.
Large, heterogeneous analysis code base,
10,000s of small, highly parallel tasks,
heavy I/O and data storage.

Two comparably sized jobs with quite different hardware requirements.

Current hardware resources

- Last year, used 187 M core-hours at ALCF, 40 M core-hours at OLCF.
  - Expect about the same this year.
- The SPC is allocating on USQCD's dedicated hardware
  - 262.3 M Jpsi-core hours on clusters at JLAB and FNAL. (Jpsi core~2 BG/P cores.)
  - 4.2 M GPU-hours on GPU clusters at JLAB and FNAL.
- (Large resources at NERSC and the Teragrid are also used for lattice QCD, managed by individual member collaborations, not USQCD.)

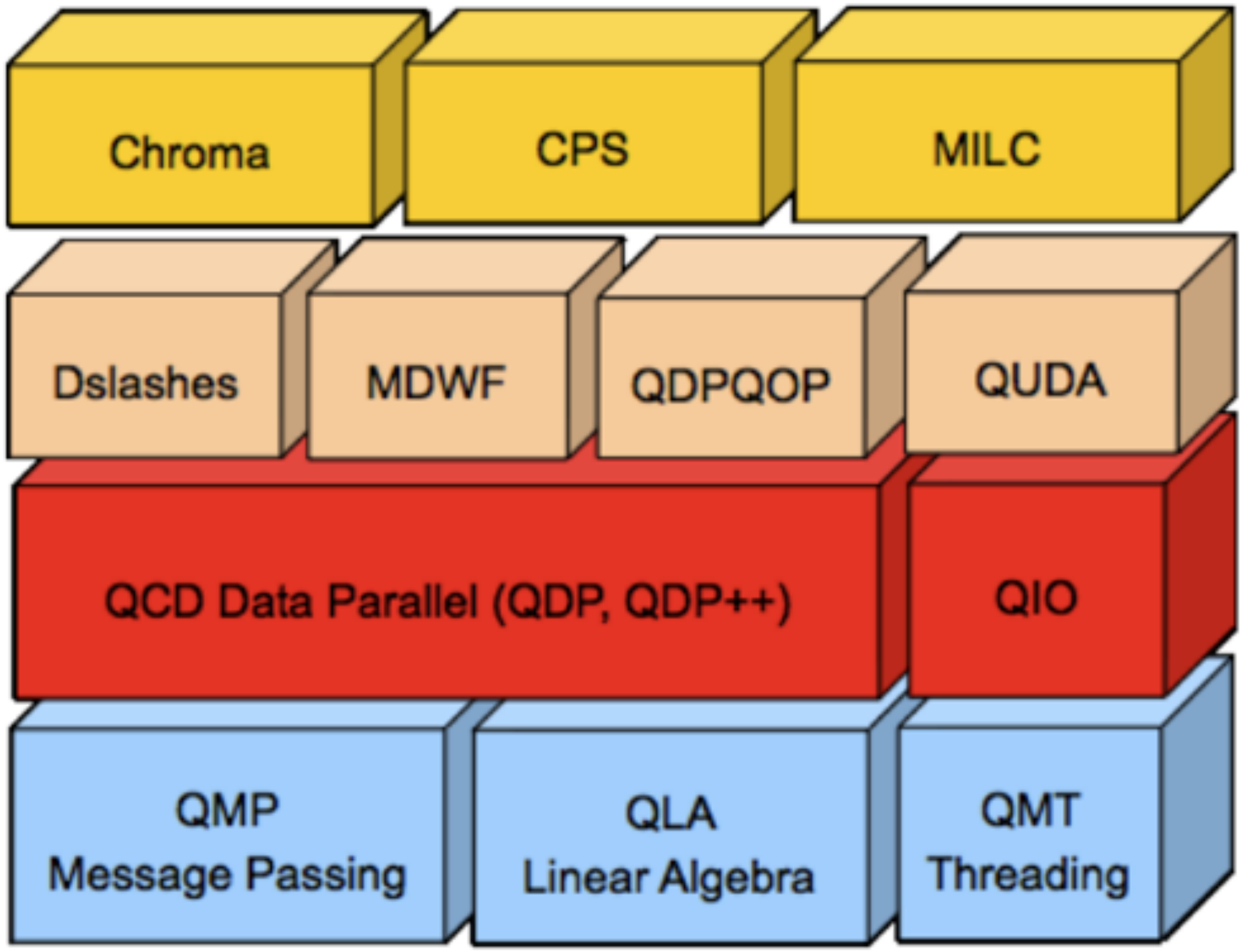




SciDAC software program

- Organized by the the USQCD Software Committee.
- Essential to our program
  - for using hardware resources efficiently, both our INCITE resources and our LQCD-ext hardware,
  - for integrating new methodological developments,
  - for accomplishing our physics goals.
- Includes community libraries for QCD programming, called the QCD API, optimized high-level QCD codes and software packages, porting to new platforms, work with SciDAC centers and institutes and with computer scientists.

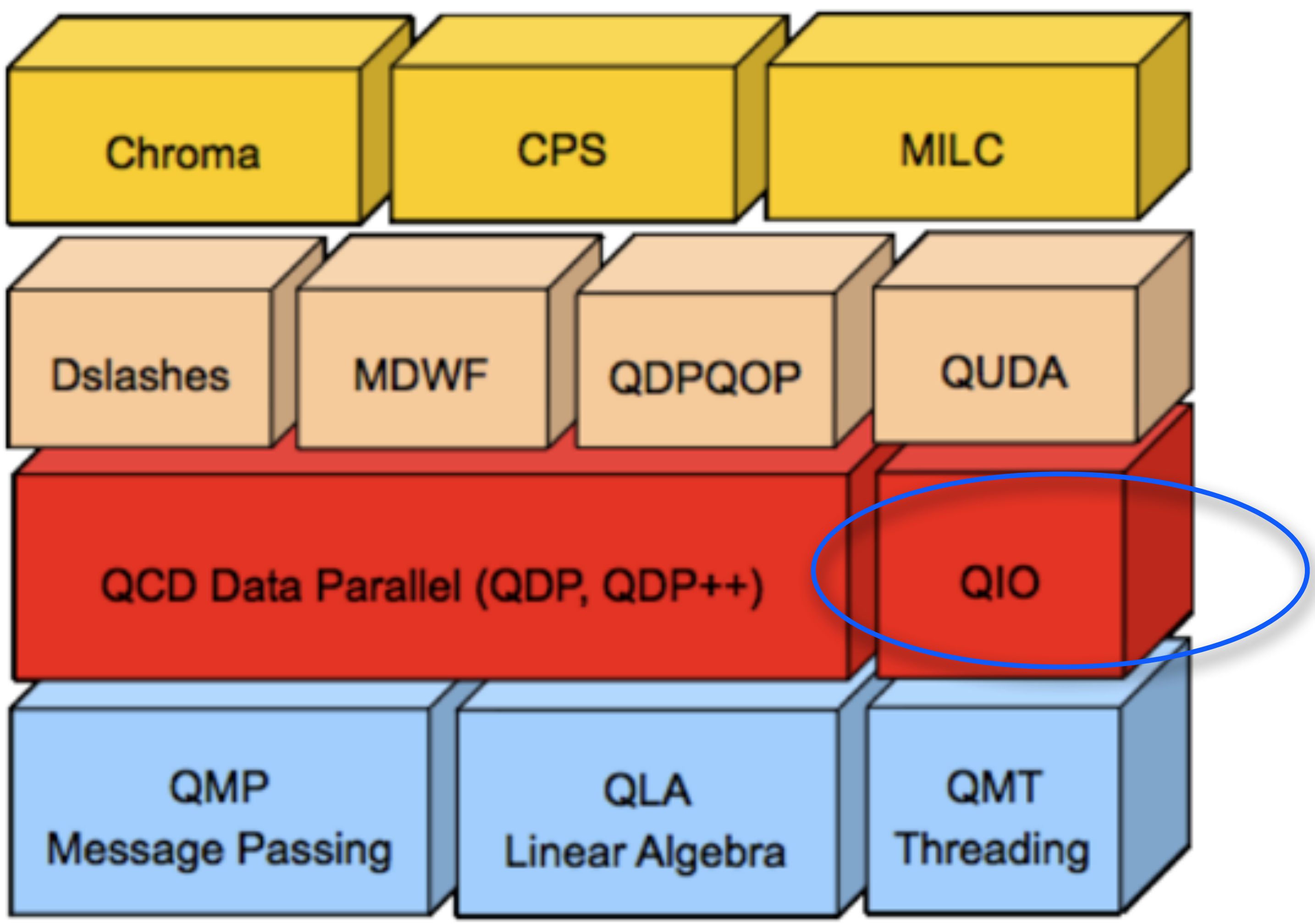
The QCD API



Basics created in SciDAC-1.

Rich Brower's talk.

The QCD API

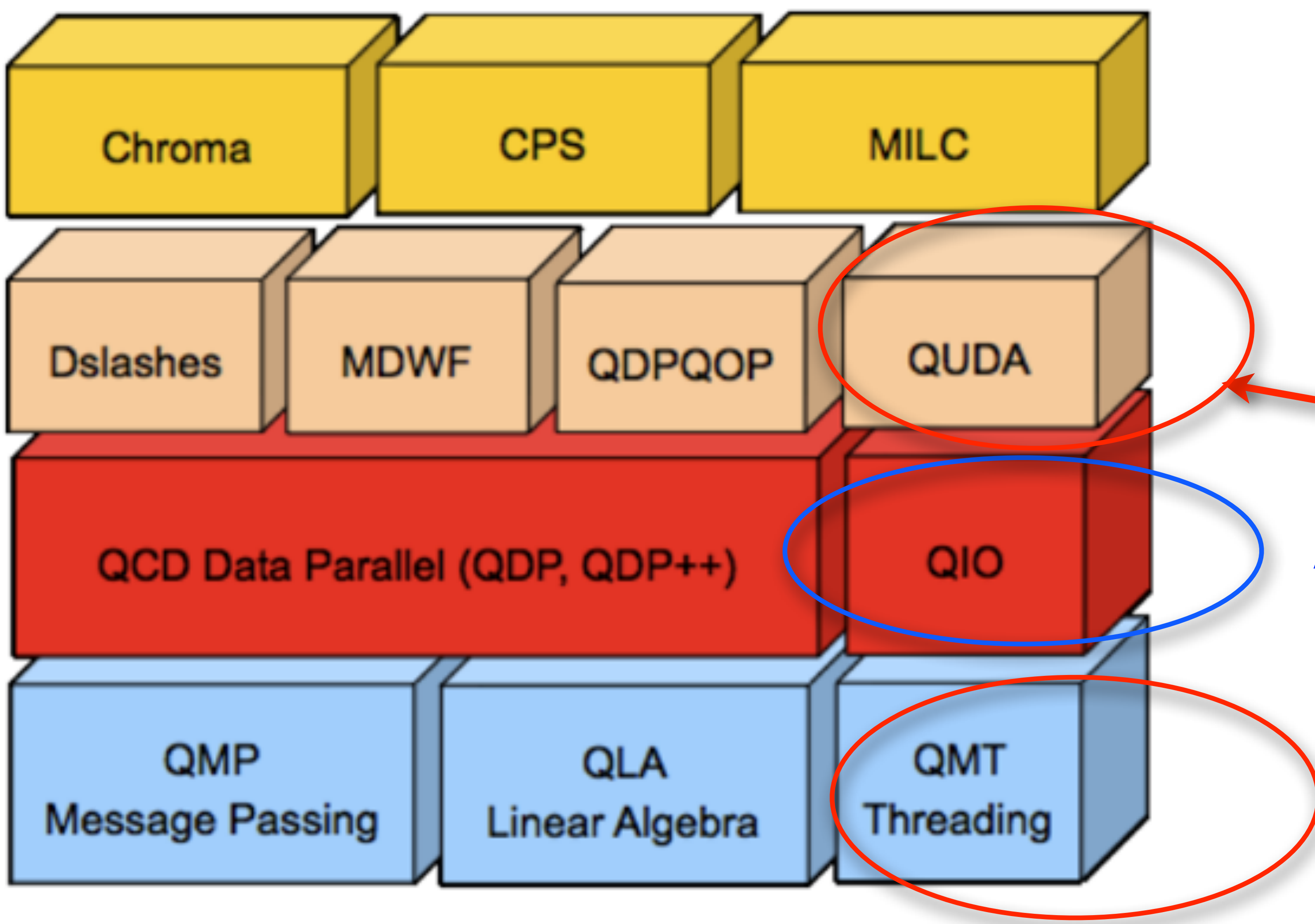


Basics created in SciDAC-1.

Added in SciDAC-2.

Rich Brower's talk.

The QCD API



Basics created in SciDAC-1.

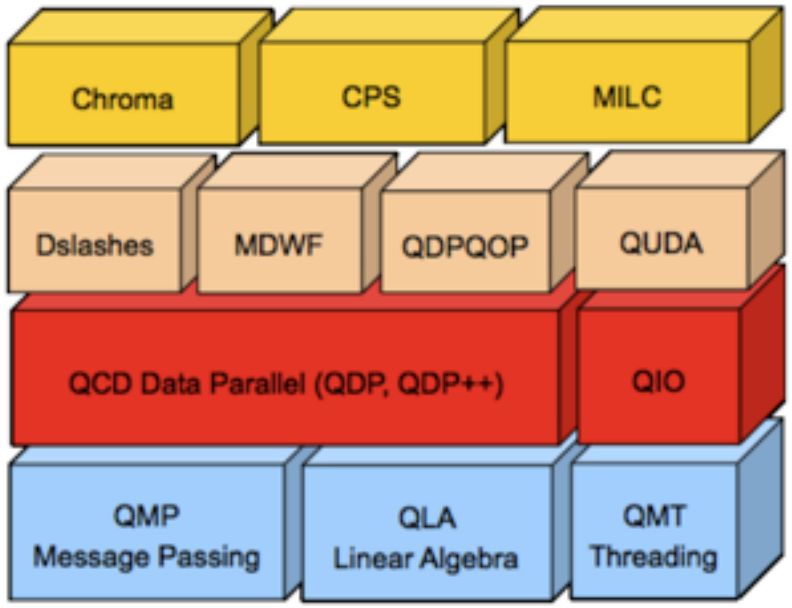
Added in SciDAC-2.

Active areas of development now.

Rich Brower's talk.

Community software

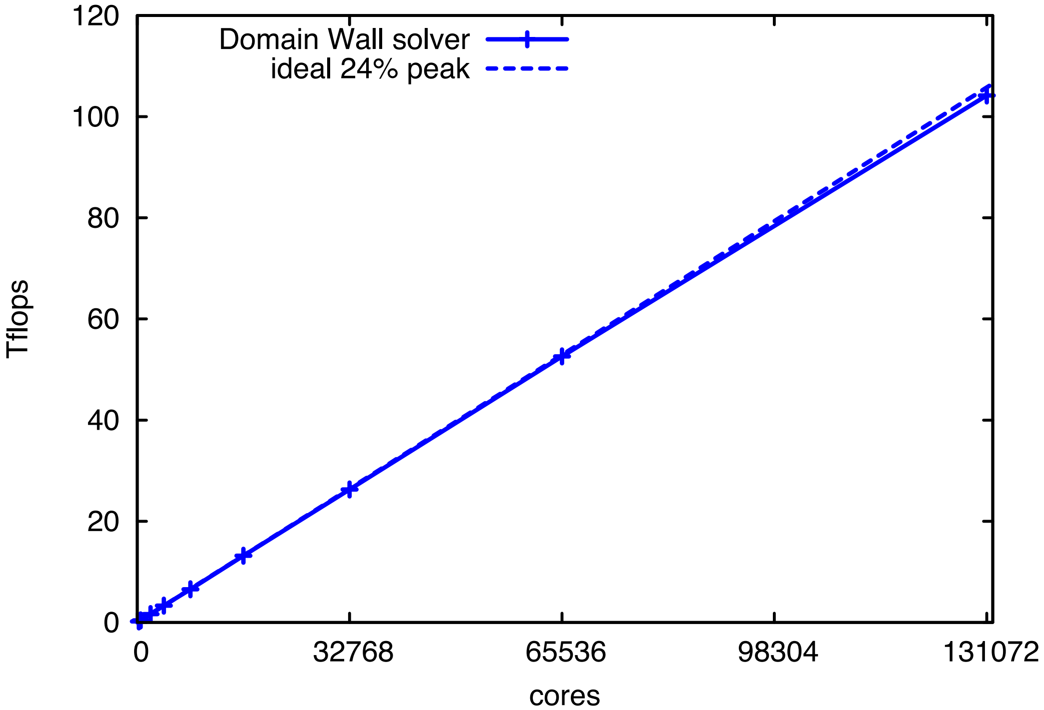
E.g., Chroma and QDP++.



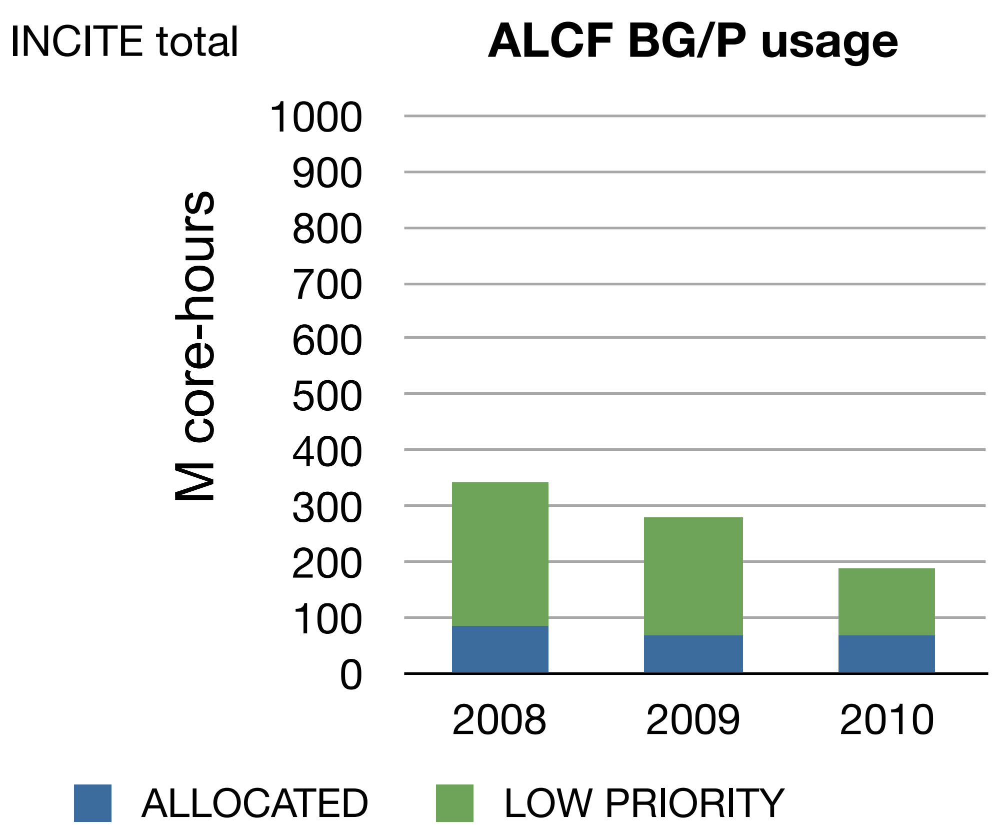
- Chroma was designed from bottom up in the USQCD era along with the QDP++ version of the QCD API.
- Level 2 (QDP++) Data parallel abstraction.
  - Hides architectural implementation and many optimizations.
  - Supports expressions & communications – close to pure math.
  - Eases rapid prototyping. Lowers entrance barrier for newcomers.
  - Use of expression templates in QDP++ hides loops over lattice site and internal space indices. Designed using modern software engineering techniques (design patterns, nightly test builds and regressions).
- Wide variety of highly optimized code available for various platforms.
- Nucl. Phys. Proc. Suppl. 140 (2005) 832, 174 citations.

Balint Joo's talk on an approach to updating QDP++ and Chroma for the many-core, heterogeous era.

The USQCD SciDAC program has enabled us to make optimal use of the hardware resources available.



In 2008, Chulwoo Jung, James Osborn, and Andrew Pochinsky had highly optimized QCD software for the BG/P ready to go when it became available at the ALCF. Chulwoo's codes were able to identify a hardware error in the machine when it was delivered.



When Blue Gene/P use began at the ALCF in 2008, USQCD was one of only a few projects that had highly optimized software in place, and was the only one of these with a large-scale, three-year program mapped out. As a result, we were able to make full use of the machine when it was relatively empty.

Coming hardware challenges

System attributes	2010	"2015"		"2018"	
System peak	2 Peta	200 Petaflop/sec		1 Exaflop/sec	
Power	6 MW	15 MW		20 MW	
System memory	0.3 PB	5 PB		32-64 PB	
Node performance	125 GF	0.5 TF	7 TF	1 TF	10 TF
Node memory BW	25 GB/s	0.1 TB/sec	1 TB/sec	0.4 TB/sec	4 TB/sec
Node concurrency	12	O(100)	O(1,000)	O(1,000)	O(10,000)
System size (nodes)	18,700	50,000	5,000	1,000,000	100,000
Total Node Interconnect BW	1.5 GB/s	20 GB/sec		200 GB/sec	
MTTI	days	O(1day)		O(1 day)	

from Dongarra and Beckman, via Thakur

Two paths to the exascale are being thought out. (We are actively involved with both.)

Fewer cores/node (like BG/Q)

More cores/node (like GPUs)

Memory per chip will grow by 100x.
Each core will communicate quickly only with nearby memory on chip.
~Six levels of (user-controlled?) cache.

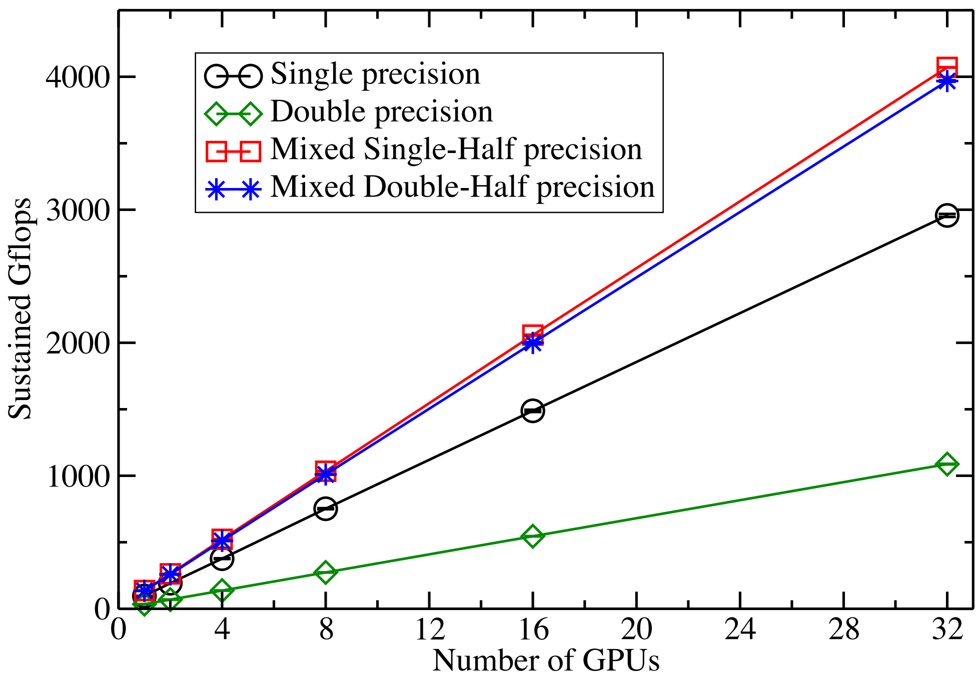
The BG/Q

- The Columbia lattice QCD group participated in the design of the BG/Q. They designed and implemented:
  - The interface between the processor core and the level-2 cache, and
  - The look-ahead algorithms used to prefetch data from level-2 cache and main memory, anticipating misses in the level-1 cache.
- Lattice QCD code of Columbia's UK collaborator Peter Boyle was first production code to run on prototype BG/Q hardware:
  - Was used to optimize the design, identify bugs in early hardware.
- The cache management problem on current computers will become nightmarish on exascale computers. (6 levels of cache?)
  - These BG/Q prefetching methods may serve as an approach to the exascale cache problem. (Pete Beckman.)

Chulwoo Jung's talk

GPUs

- USQCD members Ron Babich and Mike Clark have been hired by NVidia. One goal: get USQCD production code ready for Titan.
- Up to 10X peak price/performance vs. clusters for parts of code resident in GPU. But ... very low memory/core, bandwidth/core
    - E.g., C2050 has 6 MB (!)/core, 320 MB/sec/core.
    - Problem will grow worse on *all* computers throughout this decade. Lessons learned now are important.
- Optimization must reduce data movement, floating point not as important.
    - Easily reconstruct 8 or 12 of 18 SU(3) matrix components. Transfer only half.
    - Calculate desired double precision solution of $Ax=b$ in single or half precision, use double precision residual of result $r=b-Ax$ as a source to "polish" the result to double precision.



Ron Babich's talk

Coming software challenges

- ~50-100 person-years have been invested in our current set of community libraries and production codes for QCD.
  - SU(3) gauge theories with several quarks.
  - Codes and algorithms are highly polished.
- As the search for new forces and particle at the LHC comes to the forefront of the HEP program, lattice investigations of new strongly coupled theories has heated.
  - 100s of possible combinations of forces and fermions.
  - Increased need to get “pretty good” code for new theories up and running rapidly.
  - This need has long existed in development of new algorithms.

Several talks this afternoon on approaches to more rapid methodological development.