Lattice QCD and the Standard Model in the Exascale Era

Paul Mackenzie

Fermilab

mackenzie@fnal.gov

For the USQCD Collaboration

http://www.usqcd.org

SciDAC

Scientific Discovery through Advanced Computing

SciDAC 2010
Chattanooga
July 11-16, 2010

The Standard Model of Particle Physics

- Three forces (strong (*quantum chromodynamics or QCD*), weak, and electromagnetic), with coupling strengths:

$$\alpha_{\rm s},\ \alpha_{\rm w},\ \alpha_{\rm em}$$

- Six quark and six lepton masses

$$m_u,\ m_d,\ m_c,\ m_s,\ m_t,\ m_b$$
$$m_e,\ m_\mu,\ m_\tau,\ m_{\nu_1},\ m_{\nu_2},\ m_{\nu_3}$$

- Mixings among the quarks, the Cabibbo-Kobayashi-Maskawa matrix (2008 Nobel Prize), and (as of the last few years) among the leptons:

$$\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \qquad \begin{pmatrix} V_{e\nu_1} & V_{e\nu_2} & V_{e\nu_3} \\ V_{\mu\nu_1} & V_{\mu\nu_2} & V_{\mu\nu_3} \\ V_{\tau\nu_1} & V_{\tau\nu_2} & V_{\tau\nu_3} \end{pmatrix}$$

Where do these parameters come from?
Can we predict them with a more fundamental theory?

The Standard Model of Particle Physics

- Three forces (strong (*quantum chromodynamics or QCD*), weak, and electromagnetic), with coupling strengths:

$$\alpha_{\rm s},\ \alpha_{\rm w},\ \alpha_{\rm em}$$

Domain of lattice QCD

- Six quark and six lepton masses

$$m_u,\ m_d,\ m_c,\ m_s,\ m_t,\ m_b$$

$$m_e,\ m_\mu,\ m_\tau,\ m_{\nu_1},\ m_{\nu_2},\ m_{\nu_3}$$

- Mixings among the quarks, the Cabibbo-Kobayashi-Maskawa matrix (2008 Nobel Prize), and (as of the last few years) among the leptons:

$$\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \qquad \begin{pmatrix} V_{e\nu_1} & V_{e\nu_2} & V_{e\nu_3} \\ V_{\mu\nu_1} & V_{\mu\nu_2} & V_{\mu\nu_3} \\ V_{\tau\nu_1} & V_{\tau\nu_2} & V_{\tau\nu_3} \end{pmatrix}$$

Where do these parameters come from?
Can we predict them with a more fundamental theory?

The Standard Model is *maddeningly* successful. It accounts for every particle physics experiment performed so far, sometimes to great precision (one part in a billion for the electron anomalous magnetic moment).

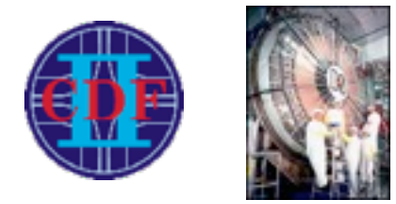
Why maddeningly? It contains obvious gaps and puzzles!

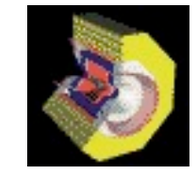
- A mathematically consistent theory cannot be constructed from the currently observed particles.
  - At least one additional, undiscovered particle is required. Is it the ''Higgs" boson, or something more complicated?
- How can gravity be incorporated?
- Why is there more than one generation of quark?
- What is the relation between the three forces?
- ...

The search for a more fundamental theory underlying the Standard Model is the central task of particle physics today.

To to understand what physics lies "Beyond the Standard Model",

- Pin down the parameters of the Standard Model, so that they can be compared with Beyond the Standard Model theories.
  - Heavy flavor factories at SLAC, KEK, and Cornell have poured out huge amounts of high precision data to pin down the CKM matrix elements.
  - The "Intensity frontier".





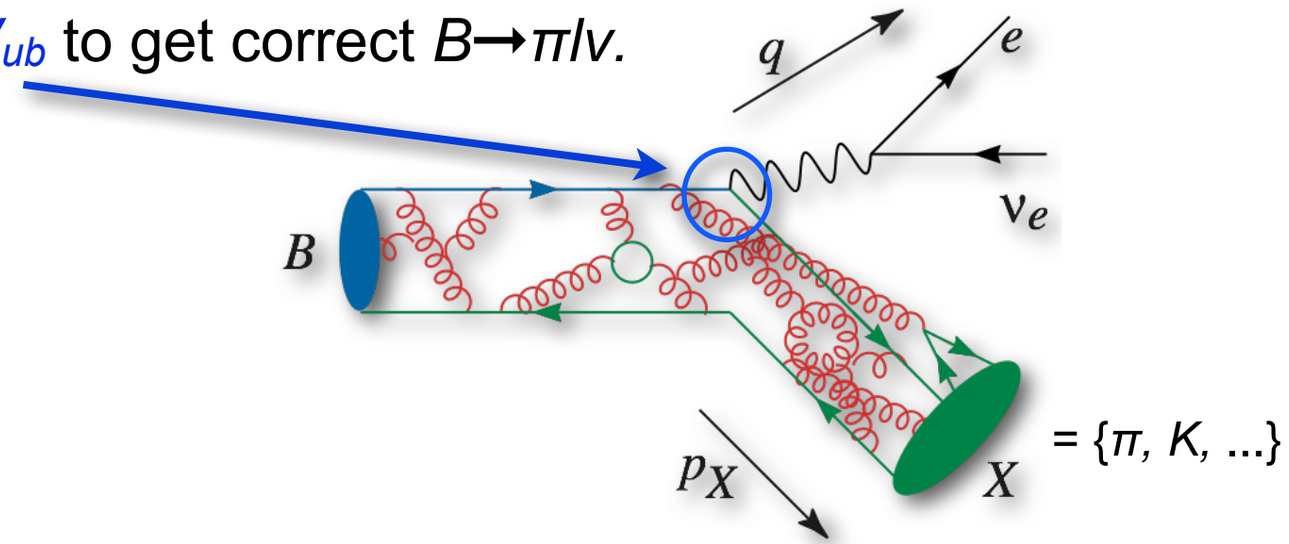
- Search for new particles and forces.
  - Exhibit A: the Higgs. Proton colliders at Fermilab, and now at the LHC at CERN are extending the search for the Higgs to higher and higher energies.
  - The "Energy frontier".

Lattice QCD calculations are essential to this program in two ways.

A, they are required to use QCD to extract quark properties from hadron properties.

Covered in this talk.

Tune V_{ub} to get correct $B \to \pi l \nu$.



B, they are essential to prepare for possible new nonperturbative phenomena in coming experiments.

Covered in Julius Kuti's talk.

- Lattice gauge theory is the first general tool for nonperturbative quantum field theory.
- New BSM interactions are likely to be contain nonperturbative sectors.
  - Of the interactions known to particle physics, only one (quantum electrodynamics) is known to be described by a perturbative theory.

The USQCD Collaboration

To support the necessary lattice QCD calculations required by high energy and nuclear physics experimental programs, the USQCD Collaboration was established in 1999 to organize computing hardware and software infrastructure for US lattice gauge theory.

- Composed of almost all US lattice gauge theorists, about 150 people.
- SciDAC-2 grant, Incite grants at ALCF and Oak Ridge, commodity clusters for capacity computing at Fermilab and JLab.
- It organizes computing infrastructure only.
  - Physics projects are done by much smaller physics collaborations within USQCD in the usual competitive and cantankerous way.

Today's talk

- The Standard Model and lattice QCD
- Current results in high energy physics
- Give an overview of USQCD's activities
- The next ten years of the Standard Model and lattice QCD

Quantum field theory

- Classical field theory (Einstein, 1905 EM, ’15 gravity).
  - Particles move on simple paths through force fields such as the electric and magnetic fields.
  - Dynamics of the EM field can be derived from a simple function of the fields, the “action”: A=1/4 (E^2-B^2).
  - Fails at atomic distances.
- Quantum mechanics (1925, Heisenberg, Schrödinger, ...)
  - At atomic distances, particle paths are spread out in a “wave function”.
  - Wave function can be derived as the sum over all possible classical paths, with the proper mathematical weighting.
- Quantum electrodynamics (QED) (1946-9, Tomanaga, Schwinger, Feynman)
  - At atomic and subatomic distances, EM field evolution is also spread out. Can be derived as sum over all classical field configurations with the weighting exp($i\,A$).

QED and QCD

- QED may be solved as an perturbative expansion in the coupling constant α.
  - This has been carried out to 4th and 5th order for some quantities, yielding predictions of amazing precision.
  - α=1/137.035 999 679(94).
- In QCD, this procedure works for particle collisions at high energies, but the effective coupling "constant" α_s is large in low energy scattering, and the perturbative method fails completely.
  - Nonperturbative methods are needed.

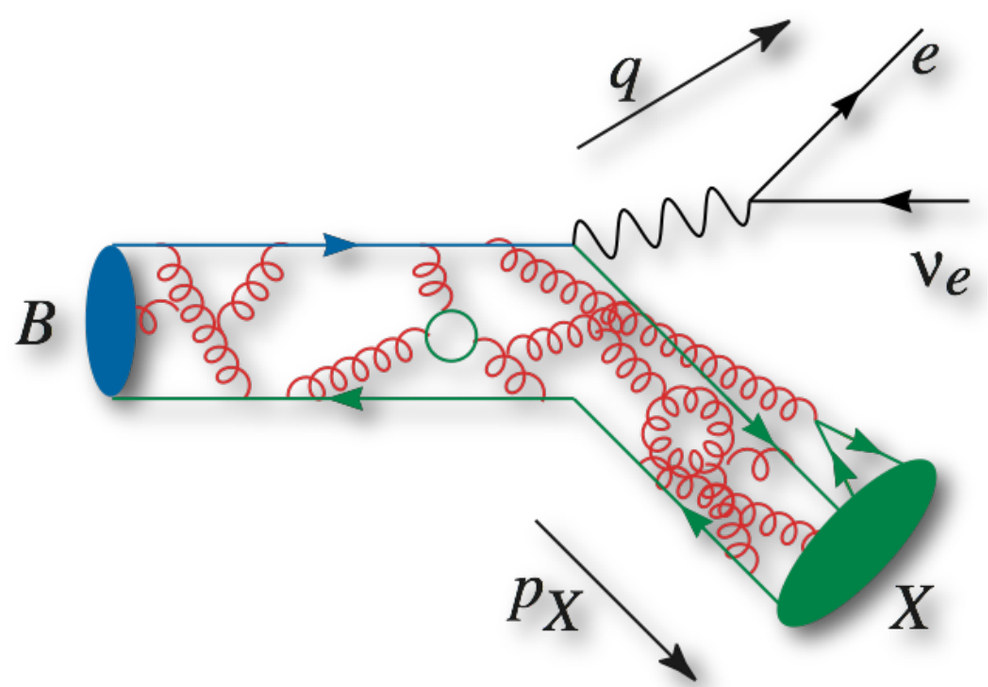
QCD

QCD is the theory of quarks and gluons. Quarks and gluons cannot be directly observed because the forces of QCD are strongly interacting.

Quarks are permanently confined inside hadrons, even though they behave as almost free particles at asymptotically high energies.

"Asymptotic freedom", Gross, Politzer, and Wilczek, Nobel Prize, 2004.

Lattice QCD is used to relate the observed properties of hadrons from the properties of their quark and gluon constituents.

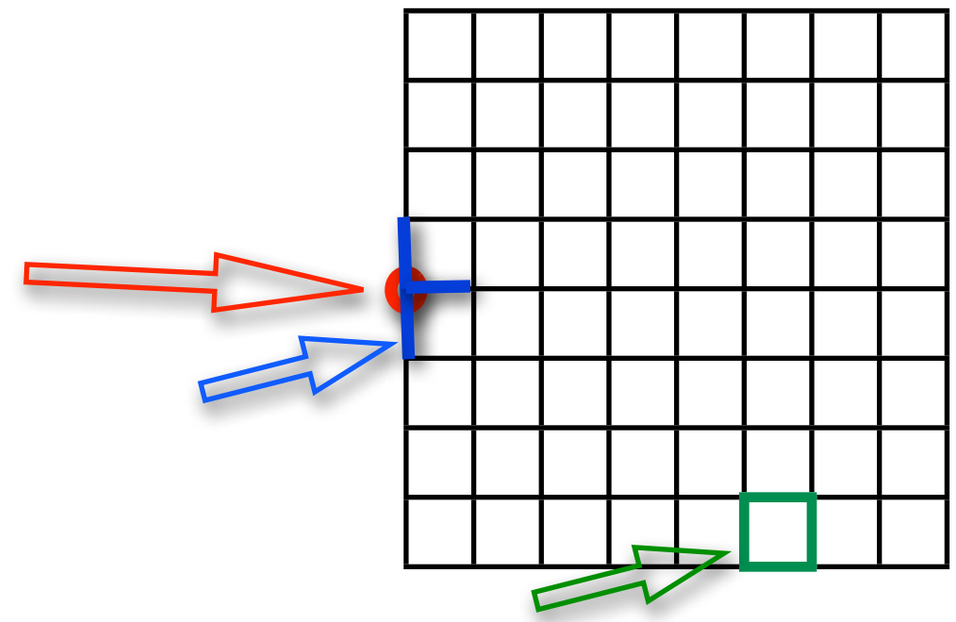


E.g,
$B \to \pi l \nu$
semileptonic decay

Luke

Lattice quantum field theories

Approximate the path integral of quantum field theory by defining the fields on a four dimensional space-time lattice.



Quarks (ψ) are defined on the sites of the lattice, and gluons (U_μ) on the links. Product of U matrices around a square is analogue of classical action.

Calculations are analytically continued to imaginary time $t_E = it$.
Leads to the real weight factor exp(-A). Can be treated as a probability.
Monte Carlo methods are used to generate a representative ensemble of gauge fields.
Relaxation methods 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.

The Dirac, or “Dslash”, operator

The fundamental operation that consumes the bulk of our cycles is the solution of the Dirac equation on the lattice.

The fundamental component of the Dirac operator is the discrete difference approximation to the first derivative of the quark field on the lattice.

$$\partial_\mu \psi(x) \to \Delta_\mu \psi(x) \approx \frac{1}{2a}(\psi(x+\hat{\mu}a) - \psi(x-\hat{\mu}a)) + \mathcal{O}(a^2)$$

Quarks in QCD come in three colors and four spins.
The color covariant Dslash operator of lattice QCD is

$$D_\mu \gamma_\mu \psi(x) \equiv \frac{1}{2}\left(U_\mu(x)\gamma_\mu\psi(x+\hat{\mu}) - U_\mu^\dagger(x-\hat{\mu})\gamma_\mu\psi(x-\hat{\mu})\right)$$

The bulk of the $O(10^{22})$ flops used in current calculations are consumed in multiplying complex 3-vectors by 3x3 complex matrices.

γ operates on spin four-vector.

U operates on color three-vector of the quark.

The computational challenge of lattice QCD

Lattice	Quark			Gauge	ensembles		Analysis propagators, correlators		
spacing *a (fm)*	mass *ml/ms*	Volume (sites)	Configu rations	Core-hours (M)	TB/ ensemble	Files/ ensemble	Core-hours (M)	TB/ ensemble	Files/ ensemble
0.06	0.1	64^3*144	1000	32.36	10.9	1,000	32	696	155,000
	0.15	56^3*144	1000	14.04	7.3	1,000	14	466	“
	0.2	48^3*144	1000	6.74	4.6	1,000	7	294	“
	0.4	48^3*144	1000	3.66	4.6	1,000	4	294	“
0.09	0.1	40^3*96	1000	3.43	1.8	1,000	3	113	155,000
	0.15	28^3*96	1000	0.81	0.6	1,000	0.8	39	“
	0.2	28^3*96	1000	0.62	0.6	1,000	0.6	39	“
	0.4	28^3*96	1000	0.36	0.6	1,000	0.4	39	“
0.12	0.1	24^3*64	1000	0.38	0.3	1,000	0.4	16	155,000
	0.15	20^3*64	1000	0.15	0.1	1,000	0.2	9	“
	0.2	20^3*64	1000	0.12	0.1	1,000	0.1	9	“
	0.4	20^3*64	1000	0.07	0.1	1,000	0.1	9	“

Example gauge ensemble library.

CPU times normalized in BG/P core-hours.

Operationally, lattice QCD computations consist of

1) Sampling a representative set of gauge configurations with Monte Carlo methods,

   E.g., the Metropolis method, the hybrid Monte Carlo algorithm, ...
   Consists of one long Markov chain. A capability task.

2) Calculating the propagation of quarks through the gauge configurations,

   Solve the Dirac equation on each configuration with relaxation methods, e.g., biconjugate gradient algorithm, etc. A capacity task.

3) Constructing hadron correlation functions from the quark propagators (smaller task).

Anatomy of a typical lattice calculation

TB file sizes

Generate gauge configurations on a leadership facility or supercomputer center.
Tens of millions of BG/P core-hours in a single job.

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

A single highly optimized program,
very long single tasks,
moderate I/O and data storage.
Needs high capability computing.

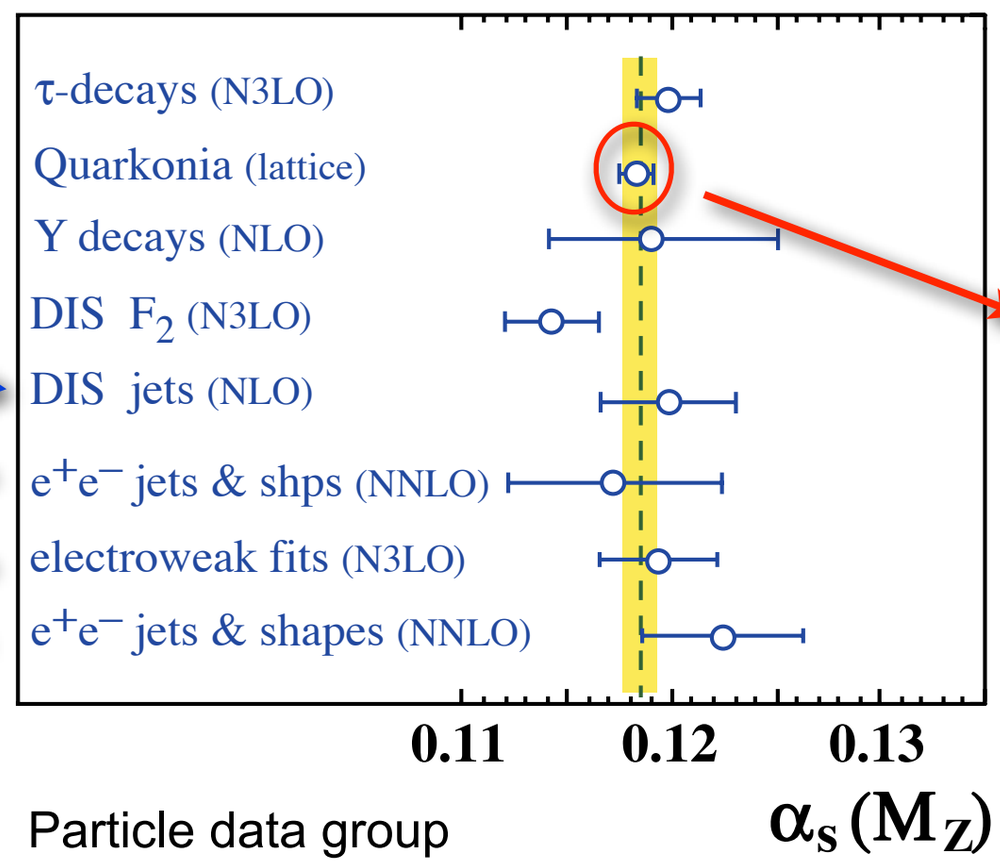
Large, heterogeneous analysis code base,
10,000s of small, highly parallel tasks,
heavy I/O and data storage.
Needs high capacity computing.

Two comparably sized jobs with quite different hardware requirements.

Results of the current generation of lattice QCD calculations: the strong coupling constant α_s.

Parametrizes the strength of the forces between quarks.

α_s can be determined by comparing high energy scattering processes with perturbative series in α_s.



Particle data group

Can also be determined with lattice QCD. Results agree. Lattice is most precise.

Results: quark masses

Light quark masses can only be determined with lattice QCD.

quark	mass (MeV)
charm	1095(11)
strange	85.5(1.4)
down	4.79(16)
up	2.01(14)

HPQCD
MILC

MS bar masses at scale 2 GeV

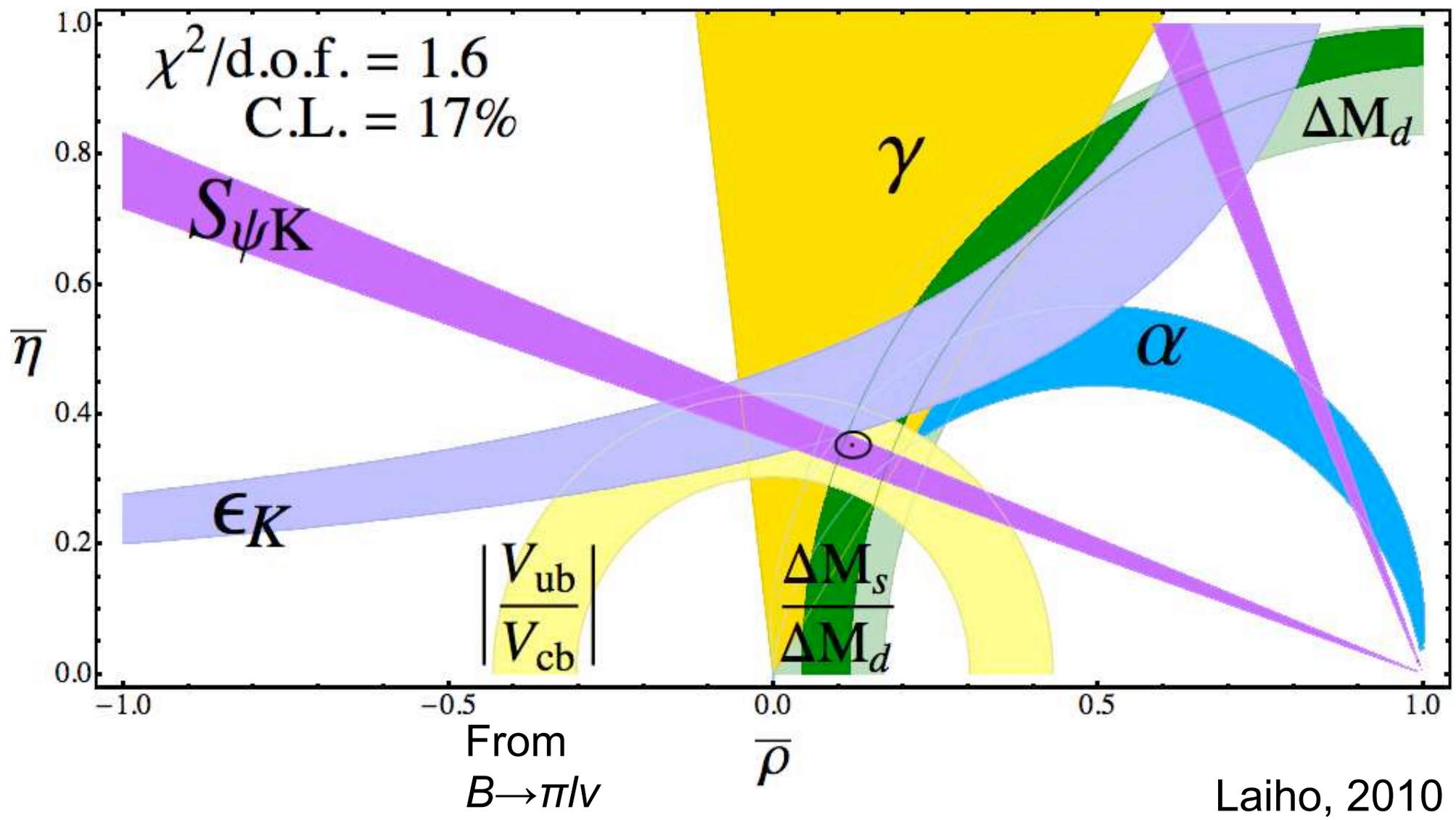
2010

Five years ago: δ m_s ~ 10%.
Now, <2%!
Big progress! (Required both more computing and better methods.)

news.sciencemag.org/sciencenow/2010/04/mass-of-the-common-quark-finally.html

Results: elements of the quark mixing (CKM) matrix

Example: $V_{ub} \propto \overline{\rho} + i\overline{\eta}$



From *B-B* bar mixing

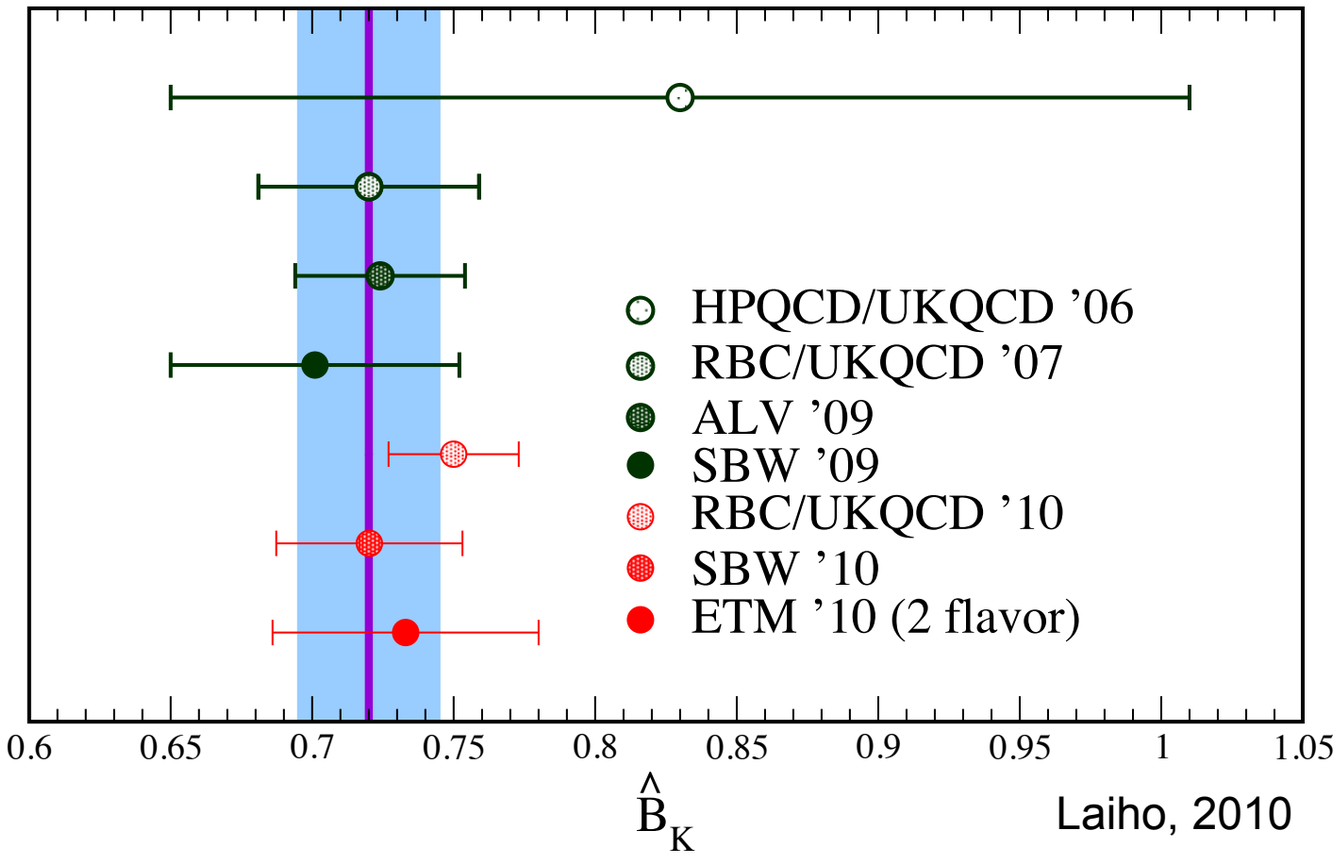
From *K-K* bar mixing

From *B→πlν*

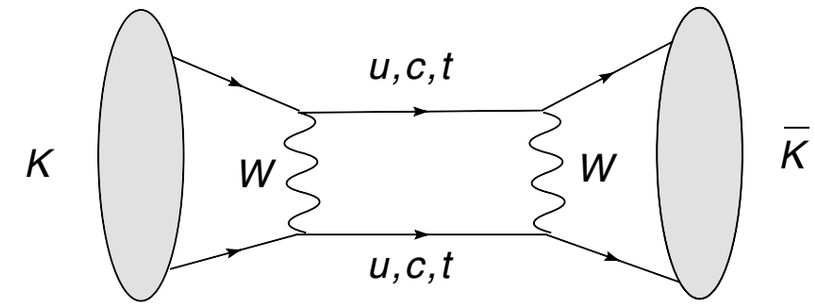
Laiho, 2010

B_K

The limit from *K-K* bar mixing is limiting by the kaon mixing parameter B_K. (Roughly speaking, related to the probability of the quarks in a kaon being on top of each other).



Laiho, 2010

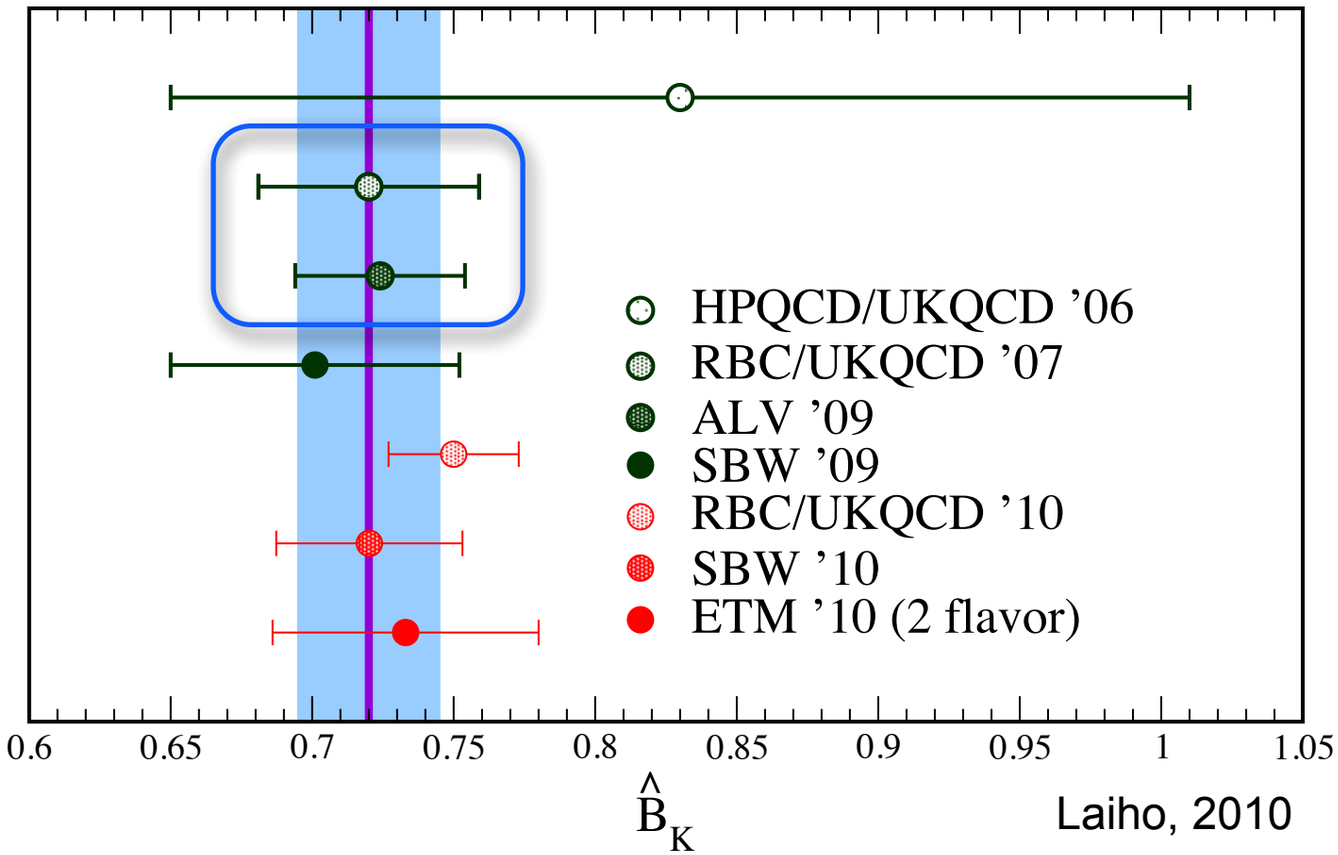


In 2009, the most precise results were from
1) RBC, led by Norman Christ, Ephraim Gildor Professor at Columbia, and from
2) three post-docs: Aubin, Laiho, and Van de Water.

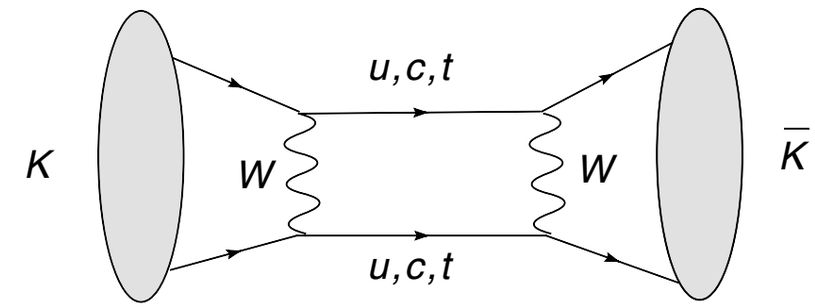
Computing infrastructure developed by USQCD using SciDAC and other resources has leveled the playing field, allowing young people to compete with established groups on an equal footing.

B_K

The limit from *K-K* bar mixing is limiting by the kaon mixing parameter B_K. (Roughly speaking, related to the probability of the quarks in a kaon being on top of each other).



Laiho, 2010



In 2009, the most precise results were from
1) RBC, led by Norman Christ, Ephraim Gildor Professor at Columbia, and from
2) three post-docs: Aubin, Laiho, and Van de Water.

Computing infrastructure developed by USQCD using SciDAC and other resources has leveled the playing field, allowing young people to compete with established groups on an equal footing.

Uncertainty analysis

Complete lattice calculations come with full uncertainty analysis.

Correlation analysis
Chiral perturbation theory
Operator product expansion
Etc. ...

uncertainty	B_K
statistics	1.2%
chiral & continuum extrapolation	1.9%
scale and quark mass uncertainties	0.8%
finite volume errors	0.6%
renormalization factor	3.4%
total	4.2%

Aubin, Laiho, and Van de Water

Each uncertainty is understood theoretically, and theory is compared with numerical data.

The CKM matrix

All of the elements of the CKM quark mixing matrix can be determined from experiment using lattice calculations and meson leptonic decays, semileptonic decays, and mixings.
For several, the lattice calculations are the only determinations.

$$\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ f_\pi & f_K & f_B \\ & K \to \pi l \nu & B \to \pi l \nu \\ V_{cd} & V_{cs} & V_{cb} \\ f_D & f_{D_s} & B \to D^* l \nu \\ D \to \pi l \nu & D \to K l \nu & B \to D l \nu \\ V_{td} & V_{ts} & V_{tb} \\ < B|\overline{B} > & < B_s|\overline{B_s} > & \end{pmatrix}$$

In the Standard Model, this matrix is described by four parameters, and the matrix elements are all related.
Is this true in reality?

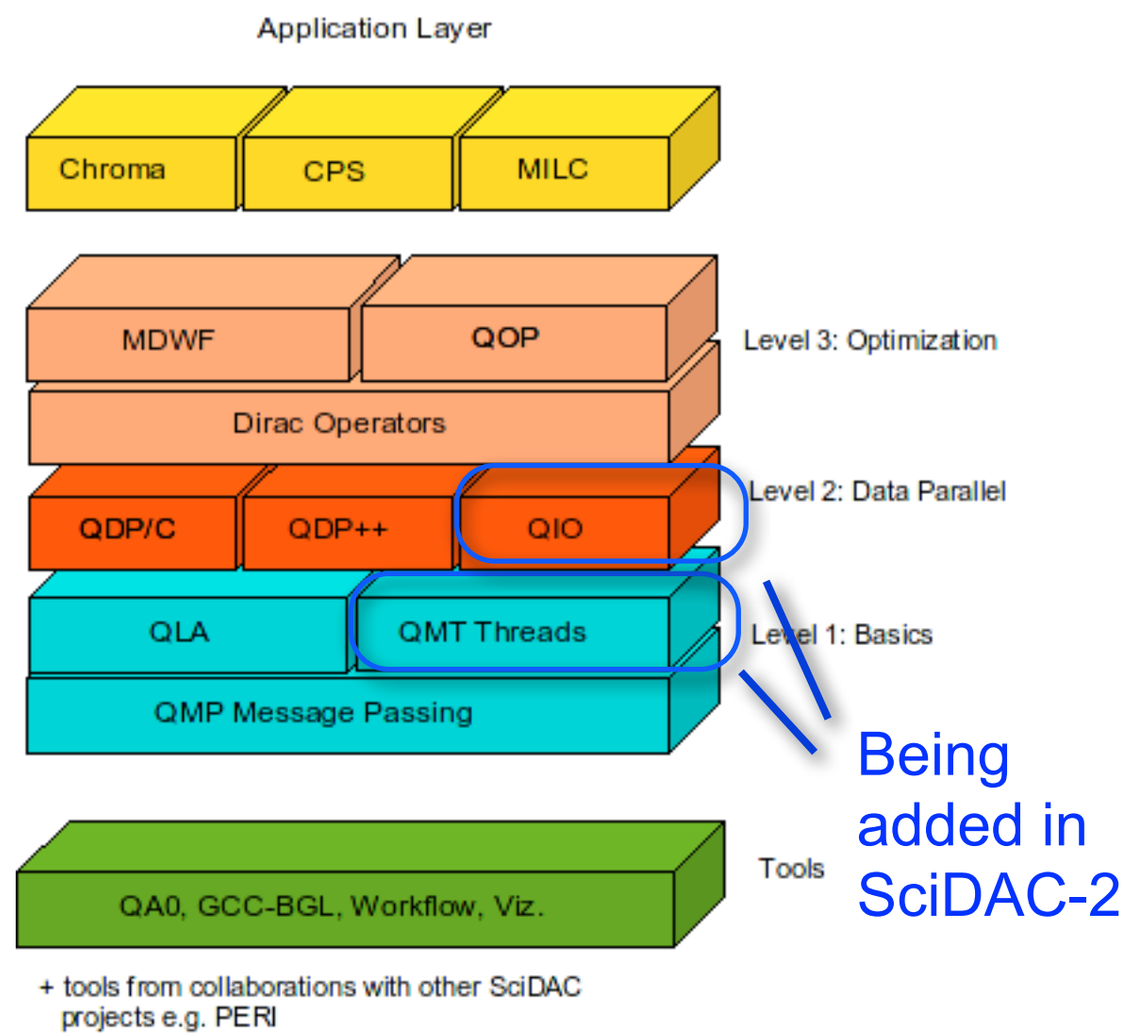
Computational infrastructure
enabling these calculations

Core components of the computing program of US lattice gauge theory:

- National program for community lattice gauge theory software funded by a SciDAC-2 grant.
- Capability computing delivered through an Incite grant to USQCD.
- Capacity computing delivered by Infiniband clusters at Fermilab and JLab funded by OHEP and ONP.
  - The larger of the two hardware needs.

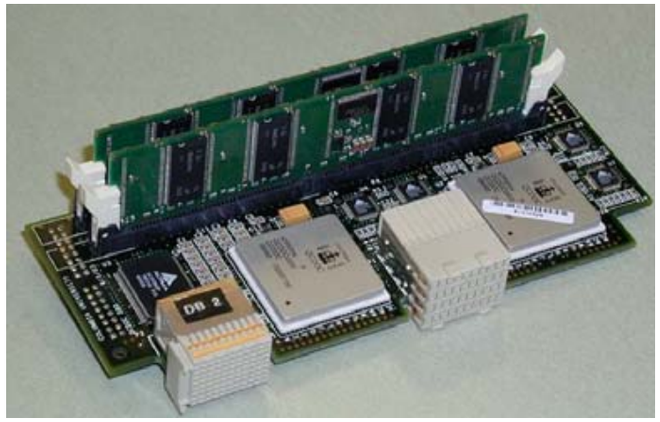
Software

- The QCD API and community libraries
  - Lower entrance barriers to lattice QCD.
  - Enable postdocs to run major projects without being part of major collaborations.



- Porting and optimizations for new platforms
  - In 2008, USQCD was the only project with a multiyear program ready to run on the Argonne BG/P from the start. Used ~1/3 of cycles in 2008, accomplished a three-year program of configuration generation in one year.
  - Now undertaking serious programs getting ready for new platforms:
    - GPUs (now!), Blue Waters (2011?), BG/Q (2012?).

Capability hardware: the QCDSP, QCDOC, the BG/L, BG/P, and beyond



QCDOC daughter card

- QCDSP, 1998 Gordon Bell Prize
  - It and its successor, the QCDOC, designed by Columbia team led by Christ.
  - Used IBM design tools for computer optimized for lattice QCD, contributed design experience to design of the BG/L and BG/P.
- Columbia/Edinburgh/RBC team, working as IBM contractors, have designed the prefetching interface between the processor and L2 cache in a next generation HPC chip targeted to be in products in the 2011/2012 timeframe.
  - QCD code was the first realistic application to run on the full chip simulator.

Capacity hardware: Commodity clusters at Fermilab and JLab

Two large-scale computing jobs in lattice gauge theory:

- Generate a long sequence of gauge configurations
  - Each lattice generated from the previous in a Markov chain. One long computing job; needs capability computing.
- Calculate many quark propagators independently on each gauge configuration.
  - Over half of the cycles, but each job is 10^{-4} the size of the job that generated the gauge configurations; needs high capacity computing.

2010 infiniband cluster at JLab: 544 dual quadcore Intel Westmere nodes.

Capacity hardware: GPUs

GPU enabled nodes now supply a significant and growing part of our capacity resources.

The 2010 JLab cluster contains 115 nodes with 530 GPUs, a mixture of cheaper gaming chips and higher quality chips such as the NVidia Fermi pictured.

Up to 32 GPUs have been made to work together on a single quark propagator. Extending the application of GPUs to larger fractions of our program is now an important focus of the software committee.

Speedups of up to x10 have been achieved for a small fraction of our workload. What fraction of new purchases should be allocated to GPUs?
The part of the workload that has been sped up becomes more valuable; how much bigger a part of our allocated workload should it become?

Coming capability hardware: Blue Waters

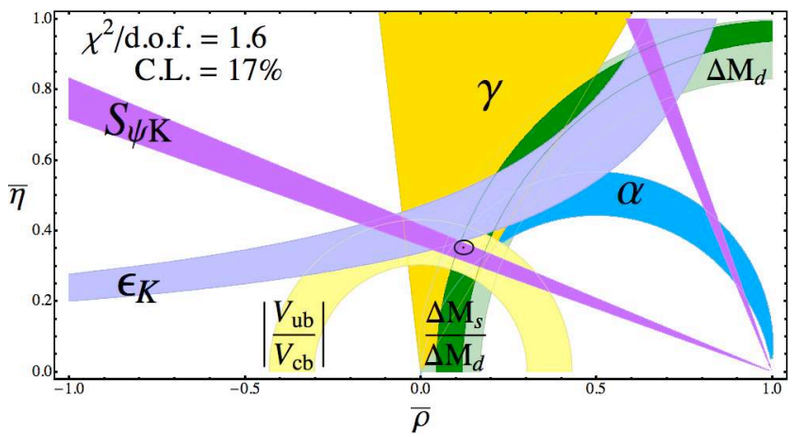
- >25,000 Power7 8-core nodes
- Acceptance test: three codes will achieve > 1 petaflop *delivered*. These include USQCD's MILC code for configuration generation. A MILC expert is now on sabbatical at NCSA to help.

IBM POWER7 processor node

Beyond the petascale: exascale physics challenges

The current physics accomplishments of lattice QCD are making a crucial contributions to the experimental programs of nuclear and particle physics, but we are only scratching the surface of what experiment needs from theory.

- precision



LHC-b aims to reduce the uncertainty in the ratio of *B* mixing and B_s mixing to 0.1%, an order of magnitude better than today. Need the same from lattice!

Precision improves as a slow function of CPU power.

$$\epsilon \propto \frac{1}{CPU^{\frac{1}{2+\cdots}}} \Rightarrow \epsilon \propto \frac{1}{CPU^{\frac{1}{\{6 \text{ or } 7\}}}}$$

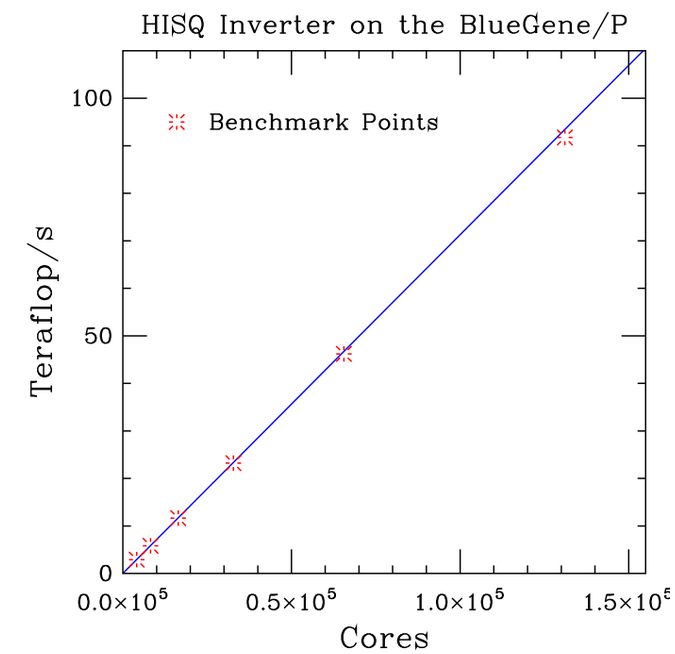
statistics

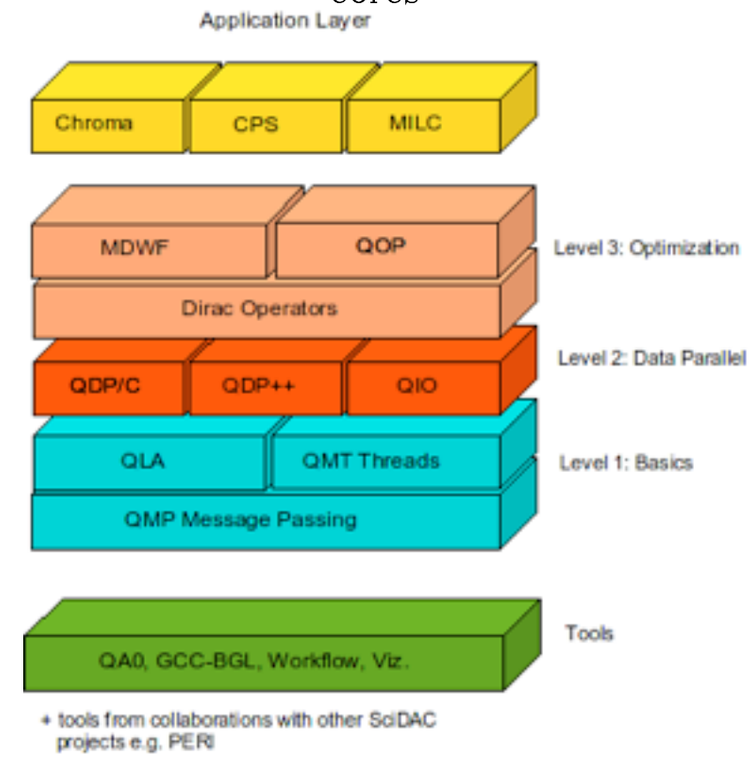
discretization, volume, ...

A factor of 1,000 increase in CPU power will improve the accuracy of current calculations by a naive factor of 3. Need both exascale computers and further improvement in methods to achieve the required factor of 10.

- New quantities
  - Processes with single, hadronically stable mesons are “Golden Quantities” of lattice QCD.
  - Many simple-looking processes are now done poorly or not at all:
    - Hadronic decays: $K \to \pi\pi$, ...
    - Hadron scattering
    - Multi-nucleon systems
    - ...
- Beyond the Standard Model
  - Hundreds of candidates for new physics beyond the Standard Model exist, each as tough to do currently as QCD:
    - E.G., SU(N) gauge theory with any N, any number of quarks,
    - Non-QCD-like theories,
    - ...

Exascale computing challenges





- Beyond one-MPI-process/core
  - weak scaling OK, but we will need strong scaling
  - starting work on threads, has helped a little on the XT5
- Dealing with small mean time to failure and silent errors
  - Most of our cycles are in quark propagator calculation. A relaxation problem; can be made highly fault tolerant.
  - A key component in configuration generation requires detailed balance; not clear how to make compatible with silent errors.

Conclusion

- Lattice QCD simulations have become an essential component of the high energy (and nuclear) physics programs, but we are only scratching the surface of what needs to be done.
- Exascale computing *and* continued dramatic advances in methods are required in order to accomplish the needs that we can clearly see in front of us.
- The work of the QCD SciDAC project has been essential in effectively using current machines such as the BG/P and the XT5, and will continue to be necessary in getting ready for the BG/Q, Blue Waters, and machines further in the future.