The co-evolution of HPC computing and LQCD

ACAT 2024

March 14, 2024

Stony Brook University

Norman H. Christ

Columbia University

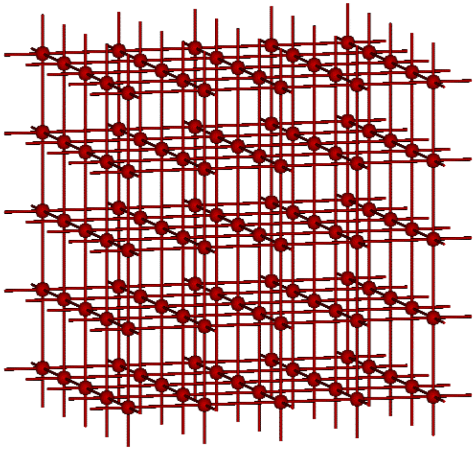
RBC and UKQCD Collaborations

Outline

- Lattice QCD (2024)
- QCD → Technology
  - Lattice QCD in 1980
  - Commercial computers in 1980
  - Purpose built hardware 1980-2005
  - Blue Gene codesign 2005-2012
- Technology → QCD:
  - All errors controlled, <1% accuracy
  - QCD → QCD + QED

Lattice QCD

Lattice QCD



- Introduce a space-time lattice.
- Evaluate the Euclidean Feynman path integral
  - Study $e^{-H_{QCD}\,t}$
  - Foundational non-perturbative formulation of QCD!
  - Permits numerical evaluation

$$\sum_n \langle n|e^{-H(T-t)}\mathcal{O}e^{-Ht}|n\rangle = \int d[U_\mu(n)]e^{-\mathcal{A}[U]}\det(D+m)\mathcal{O}[U](t)$$

- Evaluate using Monte Carlo importance sampling with a hybrid of molecular dynamics & Langevin evolution. (HMC)

Lattice QCD

$$\sum_n \langle n| e^{-H(T-t)} \mathcal{O} e^{-Ht} |n\rangle = \int d[U_\mu(n)] e^{-\mathcal{A}[U]} \det(D+m) \mathcal{O}[U](t)$$

- Very large computational challenge:
  - For a 96^3 x192 lattice: Integrate over five billion variables
  - Integrand contains the determinant of (100 Billion) x (100 Billion) matrix

Frontier – ORNL

- Fast code running on 2048 nodes of Frontier sustains 16 Petaflops [10^{15} (adds + mults)/sec]
- Soon to be overtaken by Intel Aurora machine at Argonne

The RBC & UKQCD collaborations

University of Bern & Lund
Dan Hoying

BNL and BNL/RBRC
Peter Boyle (Edinburgh)
Taku Izubuchi
Yong-Chull Jang
Chulwoo Jung
Christopher Kelly
Meifeng Lin
Nobuyuki Matsumoto
Shigemi Ohta (KEK)
Amarjit Soni
Raza Sufian
Tianle Wang

CERN
Andreas Jüttner (Southampton)
Tobias Tsang

Columbia University
Norman Christ
Sarah Fields
Ceran Hu
Yikai Huo
Joseph Karpie (JLab)
Erik Lundstrum
Bob Mawhinney
Bigeng Wang (Kentucky)

University of Connecticut
Tom Blum
Luchang Jin (RBRC)
Douglas Stewart
Joshua Swaim
Masaaki Tomii

Edinburgh University
Matteo Di Carlo
Luigi Del Debbio
Felix Erben
Vera Gülpers
Maxwell T. Hansen
Tim Harris
Ryan Hill
Raoul Hodgson
Nelson Lachini
Zi Yan Li
Michael Marshall
Fionn Ó hÓgáin
Antonin Portelli
James Richings
Azusa Yamaguchi
Andrew Z.N. Yong

Liverpool Hope/Uni. of Liverpool
Nicolas Garron

LLNL
Aaron Meyer

University of Milano Bicocca
Mattia Bruno

Nara Women's University
Hiroshi Ohki

Peking University
Xu Feng

University of Regensburg
Davide Giusti
Andreas Hackl
Daniel Knüttel
Christoph Lehner
Sebastian Spiegel

RIKEN CCS
Yasumichi Aoki

University of Siegen
Matthew Black
Anastasia Boushmelev
Oliver Witzel

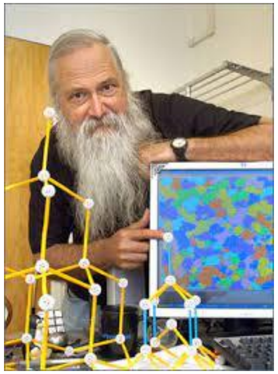
University of Southampton
Alessandro Barone
Bipasha Chakraborty
Ahmed Elgaziari
Jonathan Flynn
Nikolai Husung
Joe McKeon
Rajnandini Mukherjee
Callum Radley-Scott
Chris Sachrajda

Stony Brook University
Fangcheng He
Sergey Syritsyn (RBRC)

Lattice QCD

1980

Lattice QCD – 1980



- Invented six years earlier by Wilson
- Strong coupling expansion explained quark confinement
- Mike Creutz extracted the perturbative QCD beta function from the scale dependence of lattice QCD string tension.

1980: Promise of Lattice QCD demands enhanced computer resources

- Commercial computers were IBM or Control Data Corporation main frames or a Cray supercomputer.
- Increasing component integration was creating single-board computers.
- Hobbyist PC’s beginning to appear with integrated microprocessors.
- Parallelism was not being exploited but was natural for QCD – an opening to far exceed commercial computers.
- Integrated circuits + parallelism = 100x

1980-1985 many HE Theory groups began parallel computer construction for QCD

- Caltech: Cosmic Cube, Seitz and Fox
- Columbia: Terrano and N.C.
- Tsukuba: PACs, Hoshino, Iwasaki, Ukawa
- Edinburgh: Transputers, Bowler, Kenway, Wallace
- Rome: APE, Cabibbo, Parisi, Marinari, Trippicione
- Femilab: ACP-MAPs, Mackenzie, Eichten, Hockney, Fischler
- IBM: GF11, Beetem, Denneau, Weingarten

Columbia
QCD Machines

Overview

- 1980-1983: matrix multiplier 1 Mflops
- 2D mesh machines
  - 1983 – 1985: 16 nodes, 256 Mflops
  - 1985 – 1987: 64 nodes, 1 Gflops
  - 1987 – 1989: 256 nodes, 16 Gflops

Design
Manufacture

- 1989-1992: Thinking Machines/MIT (failed)

Codesign

- 4 & 6 D mesh machines
  - 1992 – 1998: QCDSP, 400 + 600 Gflops
  - 1998 – 2005: QCDOC, 10+10+10 Tflops

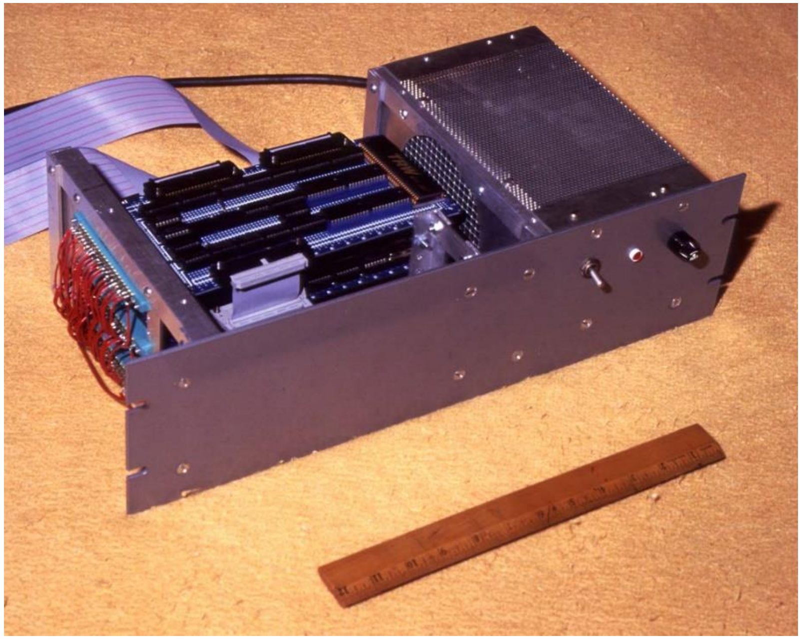
Design
Manufacture

- 2002 – 2012 Blue Gene series, IBM
- 2016 – 2019 CSA, Intel (failed)
- 2016 – Aurora, Intel (some QCD input)

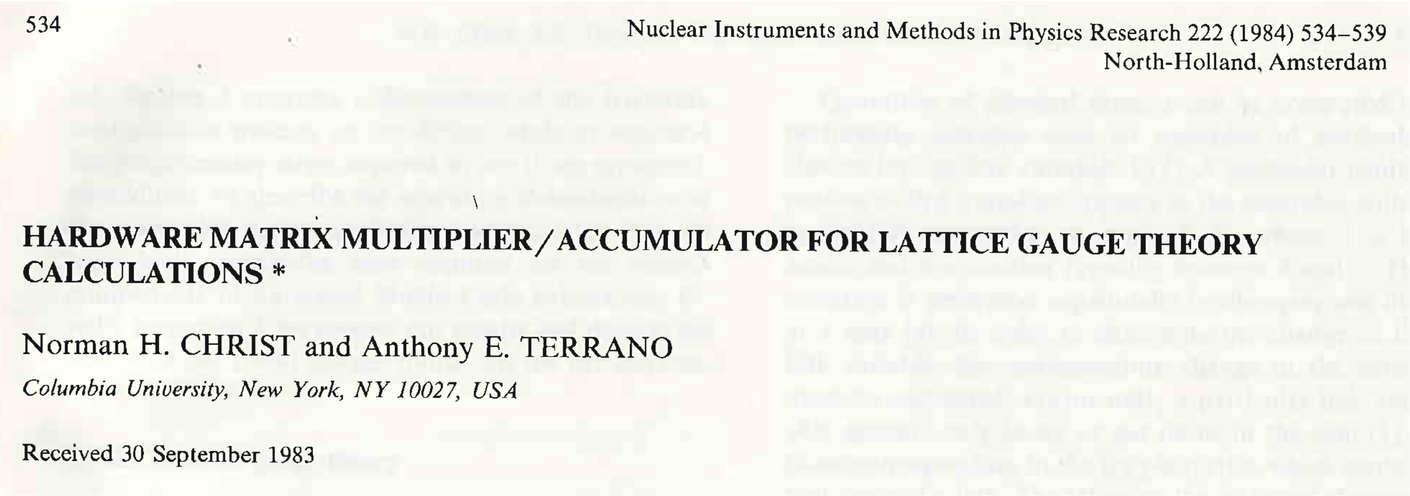
Codesign

1980-1983 Columbia Matrix Multiplier

(A. Terrano)



- Built as a peripheral for a DEC PDP11/23
- Download SU(3) matrices
- Accumulate the product of three matrices in a ``staple”
- Upload the result
- 16-bit integer multiplier and adder
- Three small wire-wrap boards controlled by preset counters
- 20X speed-up making the PDP11 2x faster than a VAX 780



534

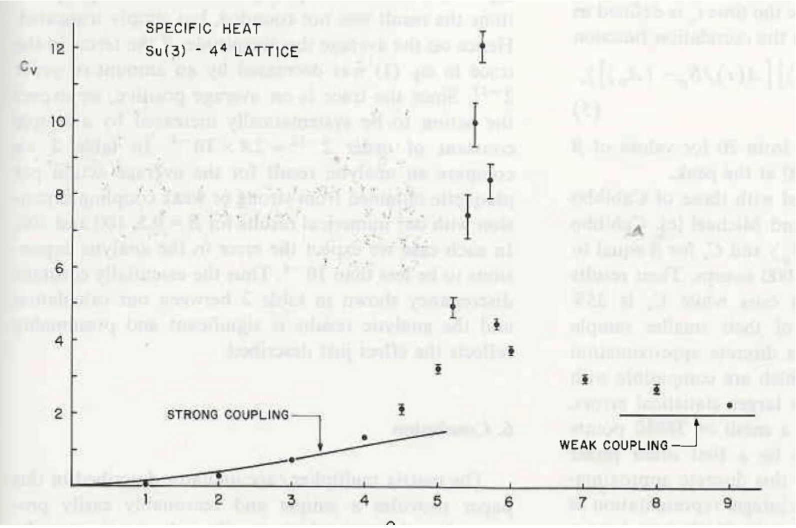
Nuclear Instruments and Methods in Physics Research 222 (1984) 534–539
North-Holland, Amsterdam

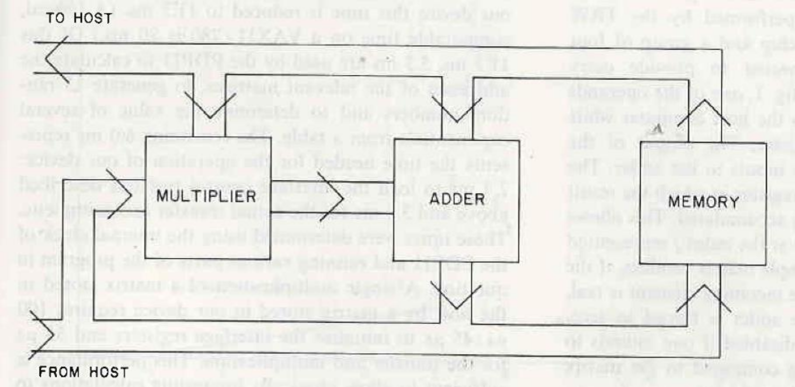
HARDWARE MATRIX MULTIPLIER / ACCUMULATOR FOR LATTICE GAUGE THEORY CALCULATIONS *

Norman H. CHRIST and Anthony E. TERRANO

Columbia University, New York, NY 10027, USA

Received 30 September 1983

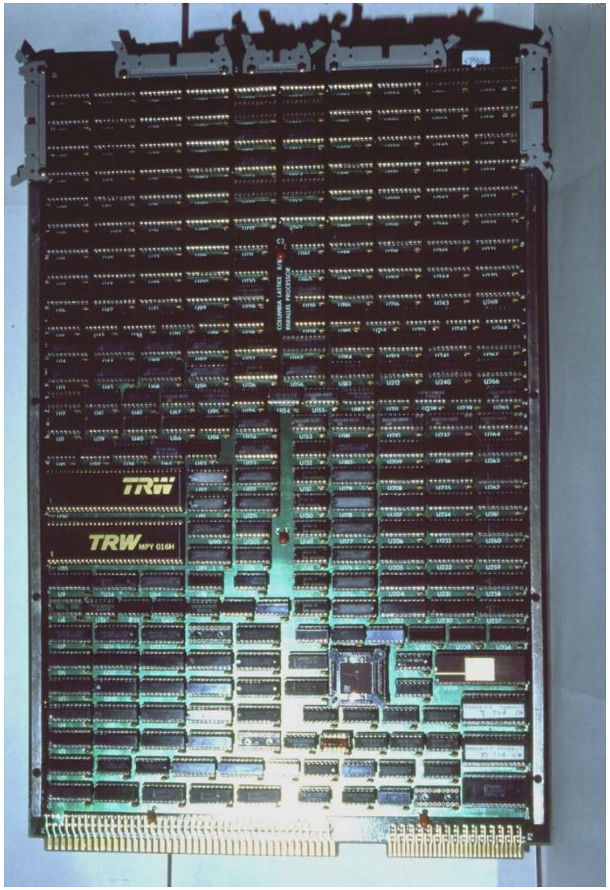




- Fluctuations in the action
- A competitive result in 1983

1983-1985 Columbia 16-node

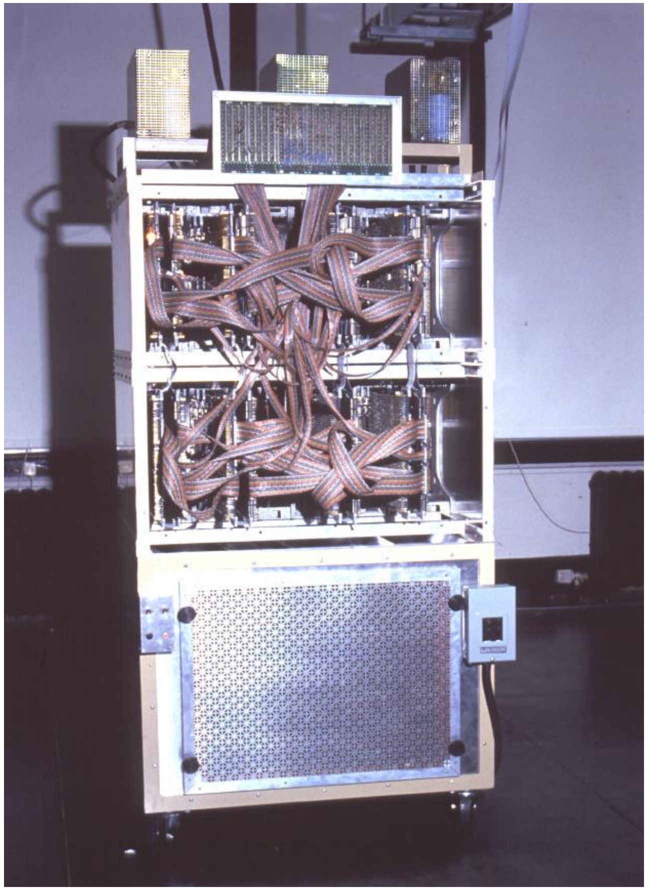
- TRW 22-bit floating point adder
- TRW 16-bit integer multiplier
- Driven by 4K 56-bit microcode words.
- Controlled by an Intel 80286
- 12”x18” wire-wrap board
- Two 64 Kbyte memory banks
- Connectors at top give direct read/write access to +x, and +y neighbors’ memories.



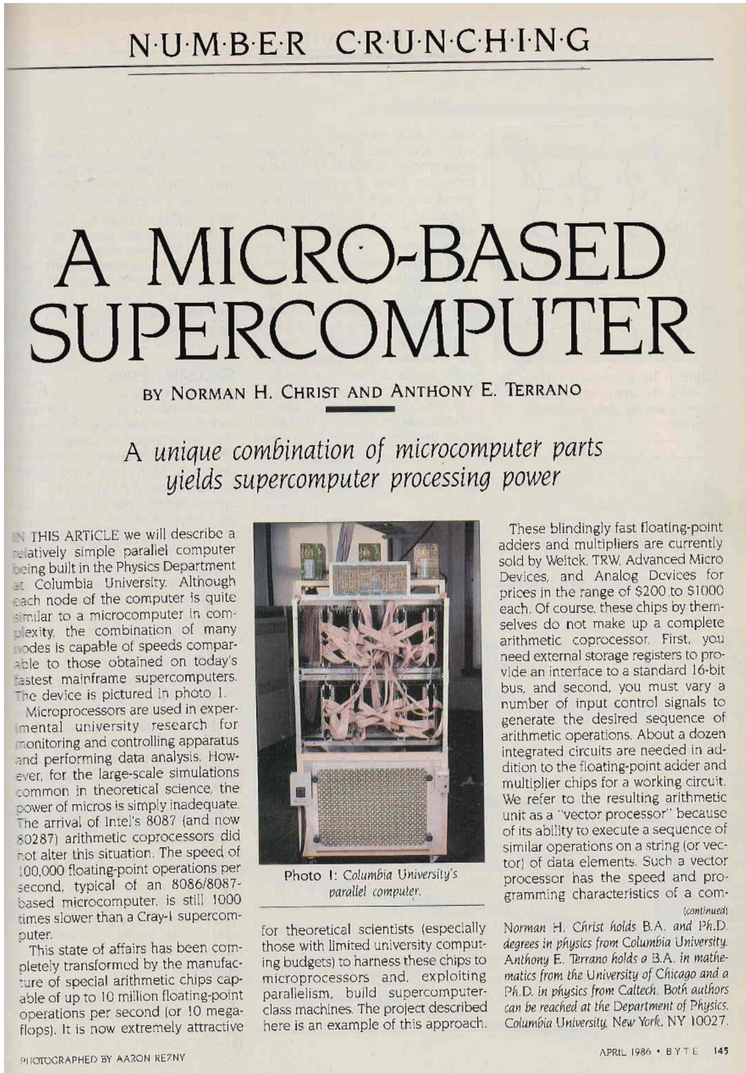
1983-1985 Columbia 16-node

- Central controller distributes clock and daisy chain IO.
- 2D periodic mesh
- Access with equal latency to local and neighboring memory.
- Efficient for 2^4 local volume.
- $150K, 256 Mflops

Machine	**Peak Mflops**	**Link update m sec**
Cray 1	160	80
16 Node	256	65
Cray XMP-4	1000	15



Readers' Favorite
Byte Magazine – 1986



A MICRO-BASED SUPERCOMPUTER

BY NORMAN H. CHRIST AND ANTHONY E. TERRANO

A unique combination of microcomputer parts yields supercomputer processing power

IN THIS ARTICLE we will describe a relatively simple parallel computer being built in the Physics Department at Columbia University. Although each node of the computer is quite similar to a microcomputer in complexity, the combination of many nodes is capable of speeds comparable to those obtained on today's fastest mainframe supercomputers. The device is pictured in photo 1.

Microprocessors are used in experimental university research for monitoring and controlling apparatus and performing data analysis. However, for the large-scale simulations common in theoretical science, the power of micros is simply inadequate. The arrival of Intel's 8087 (and now 80287) arithmetic coprocessors did not alter this situation. The speed of 100,000 floating-point operations per second, typical of an 8086/8087-based microcomputer, is still 1000 times slower than a Cray-1 supercomputer.

This state of affairs has been completely transformed by the manufacture of special arithmetic chips capable of up to 10 million floating-point operations per second (or 10 megaflops). It is now extremely attractive for theoretical scientists (especially those with limited university computing budgets) to harness these chips to microprocessors and, exploiting parallelism, build supercomputer-class machines. The project described here is an example of this approach.

Photo 1: *Columbia University's parallel computer.*

These blindingly fast floating-point adders and multipliers are currently sold by Weitek, TRW, Advanced Micro Devices, and Analog Devices for prices in the range of $200 to $1000 each. Of course, these chips by themselves do not make up a complete arithmetic coprocessor. First, you need external storage registers to provide an interface to a standard 16-bit bus, and second, you must vary a number of input control signals to generate the desired sequence of arithmetic operations. About a dozen integrated circuits are needed in addition to the floating-point adder and multiplier chips for a working circuit. We refer to the resulting arithmetic unit as a "vector processor" because of its ability to execute a sequence of similar operations on a string (or vector) of data elements. Such a vector processor has the speed and programming characteristics of a com-

(continued)

Norman H. Christ holds B.A. and Ph.D. degrees in physics from Columbia University. Anthony E. Terrano holds a B.A. in mathematics from the University of Chicago and a Ph.D. in physics from Caltech. Both authors can be reached at the Department of Physics, Columbia University, New York, NY 10027.

PHOTOGRAPHED BY AARON REZNY

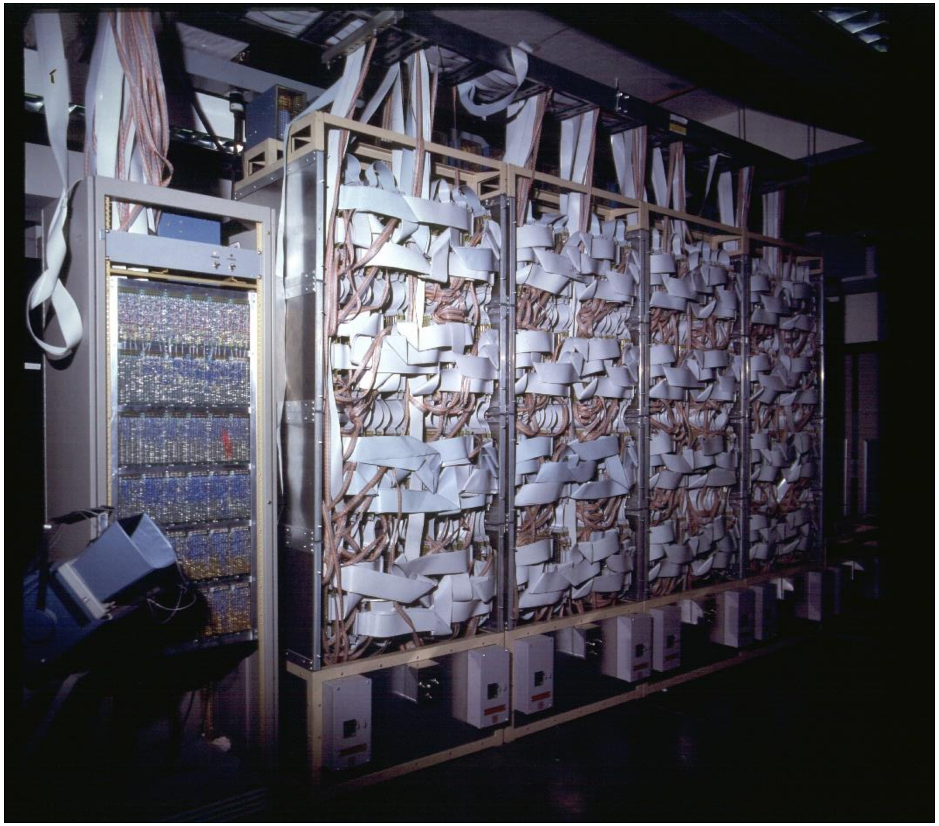
Follow-on 2D machines

American Association for the Advancement of Science

SCIENCE

18 March 1988 $3.00
Vol. 239 ▪ Pages 1349–1464

1987: 64 nodes, 1 Gflops



1989: 256 nodes, 16 Gflops

6.7 Gflops sustained

(later equaled by CM2 and GF11)

1992-1998 QCDSP

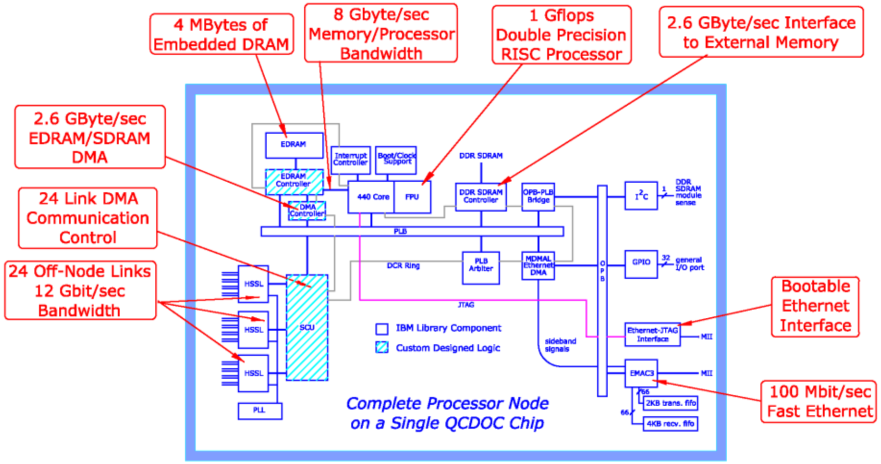
(A. Gara, R. Mawhinney)

- 1000’s of credit-card size nodes
  - TI DSP, custom ASIC, 2MB memory
- 64 nodes/board.
- Fast DMA access to 4D neighbors.
- Efficient for 4^4 local volume.
- 400 Gflops (Columbia)
- 600 Gflops (BNL)
- 1998 Gordon Bell prize for price performance
- Used by larger BNL/Columbia collaboration

1998-2005 QCDOC

(P. Boyle, A. Gara, R. Mawhinney)

- IBM custom ``system on a chip'' ASIC
- 6D mesh, 128 MB memory
- Efficient for 8^4 local volume.
- 2 Mbyte EDRAM
- 8 Watts/node



4 MBytes of Embedded DRAM

8 Gbyte/sec Memory/Processor Bandwidth

1 Gflops Double Precision RISC Processor

2.6 GByte/sec Interface to External Memory

2.6 GByte/sec EDRAM/SDRAM DMA

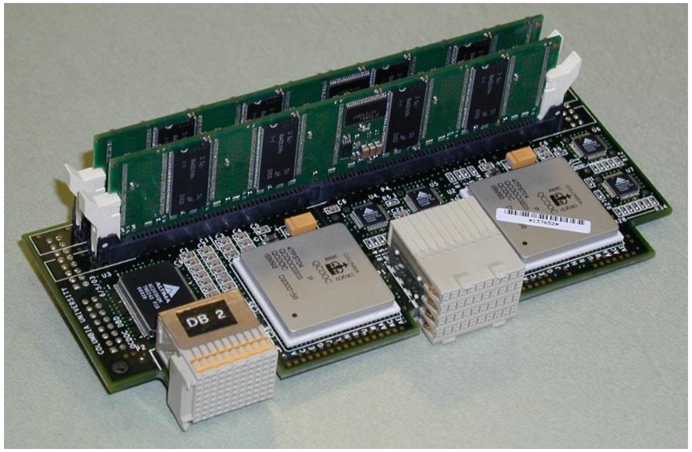
24 Link DMA Communication Control

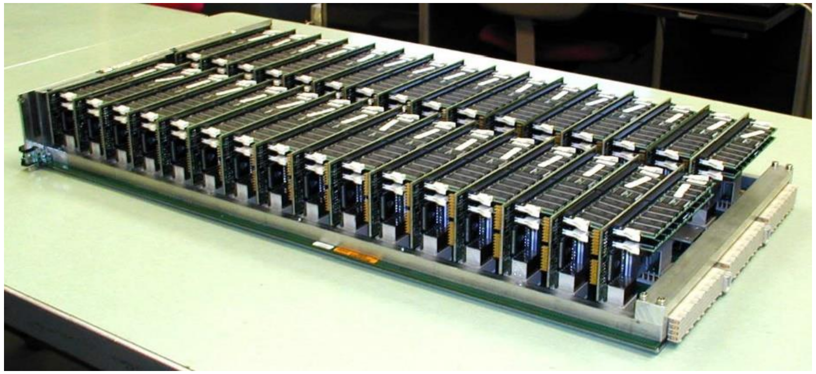
24 Off-Node Links 12 Gbit/sec Bandwidth

Bootable Ethernet Interface

100 Mbit/sec Fast Ethernet

Complete Processor Node on a Single QCDOC Chip





1998-2005 QCDOC

(P. Boyle, A. Gara, R. Mawhinney)

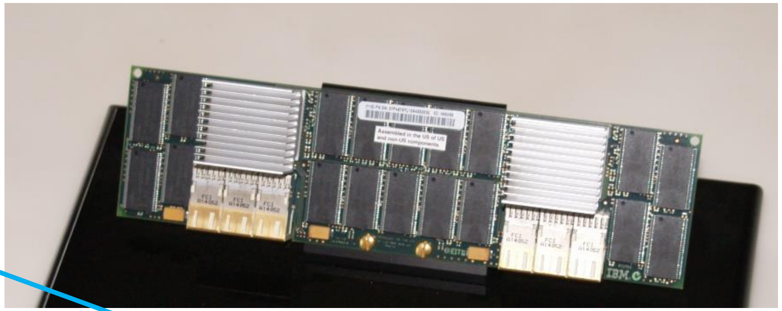
- $2 M NRE, $7 M IBM ASCI Fab, $20M total
- RBRC / BNL / Edinburgh: 10+10+10 Tflops
- First USQCD computer

Technology Transfer: BG/L & /P

(A. Gara)

- IBM Blue Gene: replace QCDOC with IBM product
- Columbia students/postdoc: J. Sexton, D. Chen, P. Vranas
- BG/L top of Top 500, ’04–’07

Two BG/L nodes

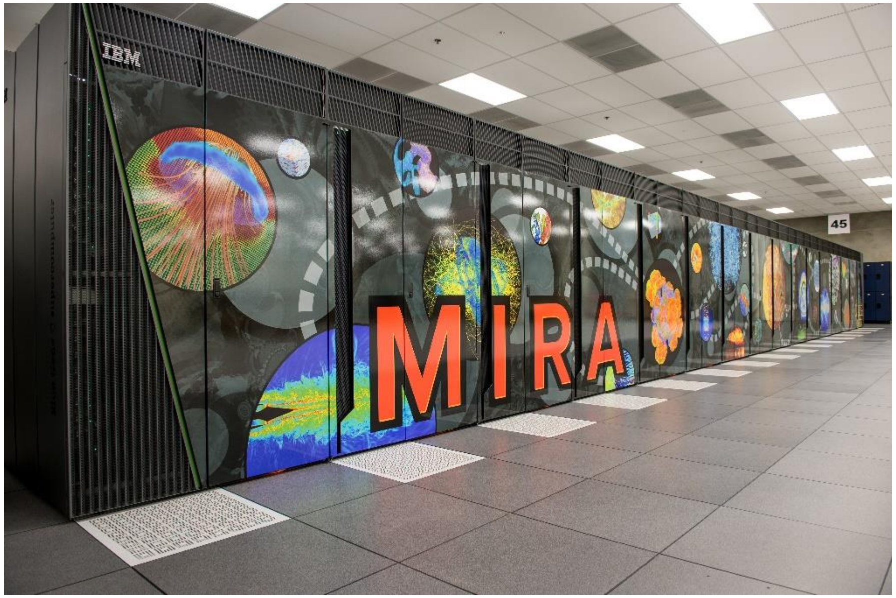


Gara, Palmisano (IBM CEO) & Obama
National Medal of Technology

Codesign: Blue Gene/Q

(A. Gara, P. Boyle)

- Cost of ASIC design exceeded our funding – join IBM Blue Gene/Q project.
- P. Boyle, C. Kim, N.C. designed the L1P, interface between Power CPU and BG/Q the system.
- QCD code was run on ASIC simulator and first hardware
- P. Boyle was a critical member of 4-5 person team making ASIC work
- Decommissioned in March 2020 after billions of core hours for QCD



MIRA, ANL 10 Pflops

Another Chapter: Early GPUs

- Eötvös, Wuppertal:

  Z. Fodor (2006)

Computer Physics Communications

Volume 177, Issue 8, 15 October 2007, Pages 631-639

Lattice QCD as a video game

Győző I. Egri[a], Zoltán Fodor[a b c], Christian Hoelbling[b], Sándor D. Katz[a b], Dániel Nógrádi[b], Kálmán K. Szabó[b]

- Began with gaming cards
- Among the first science applications using GPUs

← Boston University & JLab:

Babich, Brower, Clark,

Edwards, Joo

9g cluster JLab 100 Tflops (2011)

Software Innovation

- Neglected in this presentation.
- Custom OS designed for
  - High parallel performance
  - Fault detection
  - Diagnostic power
- Non-Linux kernel essential:
  - QCSDP
  - QCDOC
  - Blue Gene
- Require talks from Bob Mawhinney and Peter Boyle.

QCD Predictions

→

Discovery…

Physics driven by lattice QCD

- Explore confinement and chiral symmetry breaking
- Nucleon → nuclear structure (EIC)
- Quark and lepton flavor physics:
  - heavy quarks: CKM unitarity
  - light and strange quarks: rare processes
  - leptons: $g_\mu - 2$

Standard Model Tests

(RBC/UKQCD Collaboration)

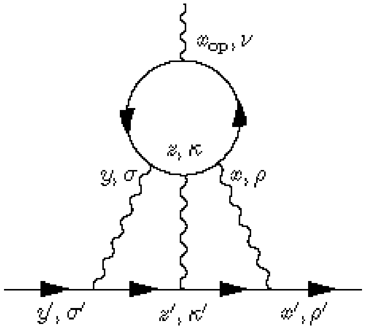
$(g_\mu\text{-}2)_{\text{expt}}$ = 0.00116592089(63)

(Expt) – (Theory) = 0.00000000249(87)

HLbL = 0.00000000079(35)

[Phys.Rev.Lett.124, 132002 (2020)]

Calculation of light-by-light contribution (L. Jin)

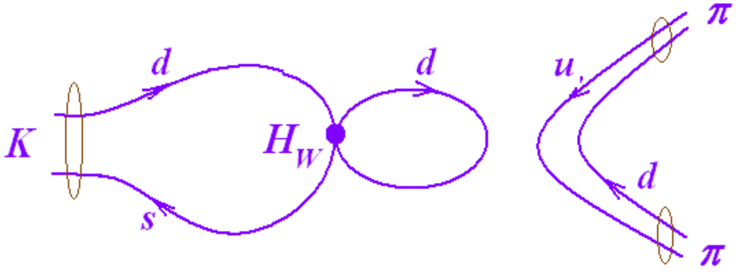


Direct CP violation $K \rightarrow \pi\,\pi$

[Phys.Rev.D 102, 054509 (2020)]

Re $(\varepsilon'/\varepsilon)$ = 21.7(6.9) x10^{-4}

[16.6(2.3) x10^{-4} expt] (C. Kelly, T Wang)



Standard Model Tests

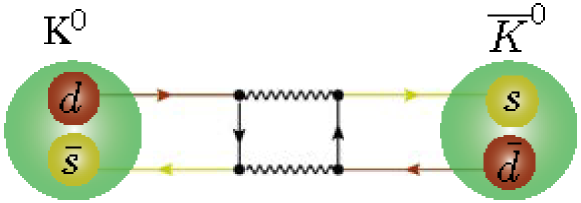
(RBC/UKQCD Collaboration)

$m_{K_L} - m_{K_S}$ = 5.8(2.4) x 10^{-12} MeV [PoS, Lattice 2021]

[3.484(6) x 10^{-12} expt]

- Sensitive to new physics at the 1,000 TeV scale
- More accurate calculation about to start

  (B. Wang, Yikai Huo)



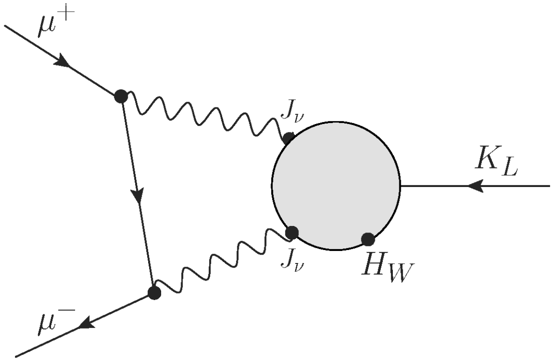
$\pi^0 \rightarrow e^+ e^-$ BR = 6.30(06) x10^{-8}

[6.87(36) x10^{-8} expt]

- First lattice QCD calculation (Y. Zhao)

$K_L \rightarrow \mu^+ \mu^-$ Initially without disconnected graphs

- BSM-sensitive, strangeness-changing neutral current process (EH Chao, C. Hu)



Add QED: QCD → QCD + QED

(L. Jin)

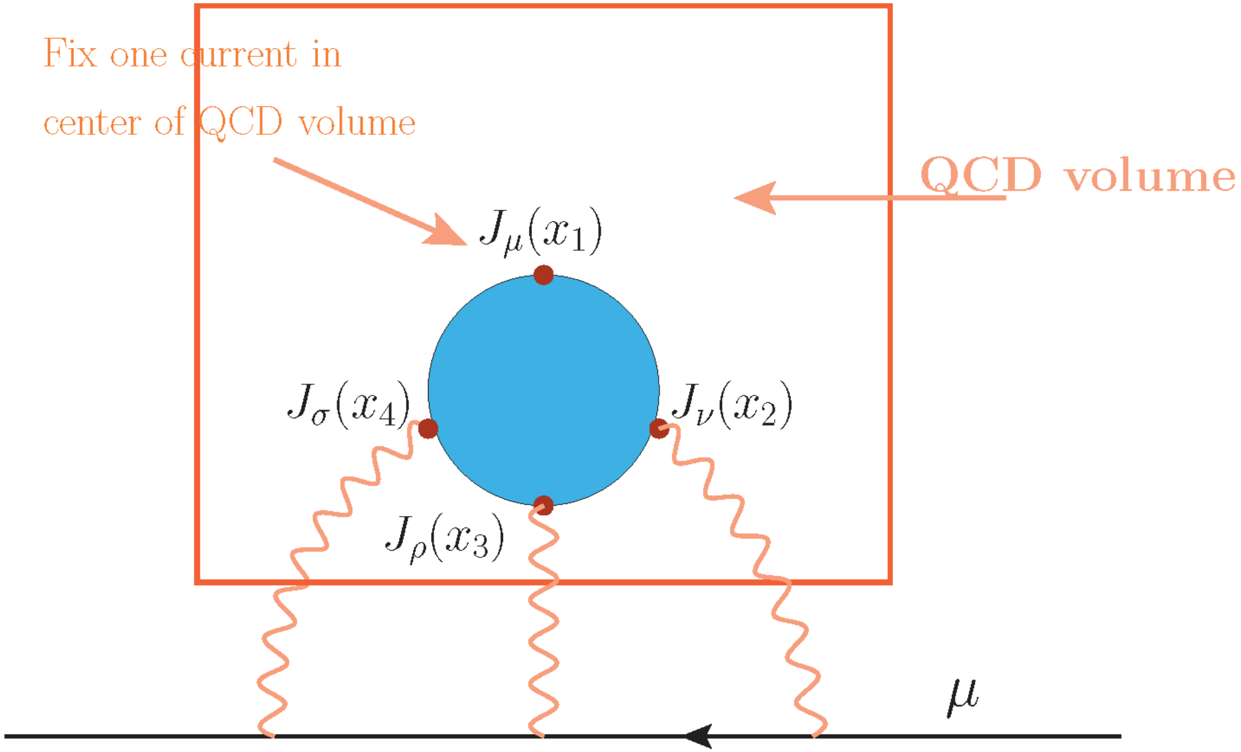
- Lattice formulation for QCD + ~~QED~~ ?

- Lattice QCD + Continuum QED

(Kenneth Wilson)	(Richard Feynman)
Finite volume	Infinite volume

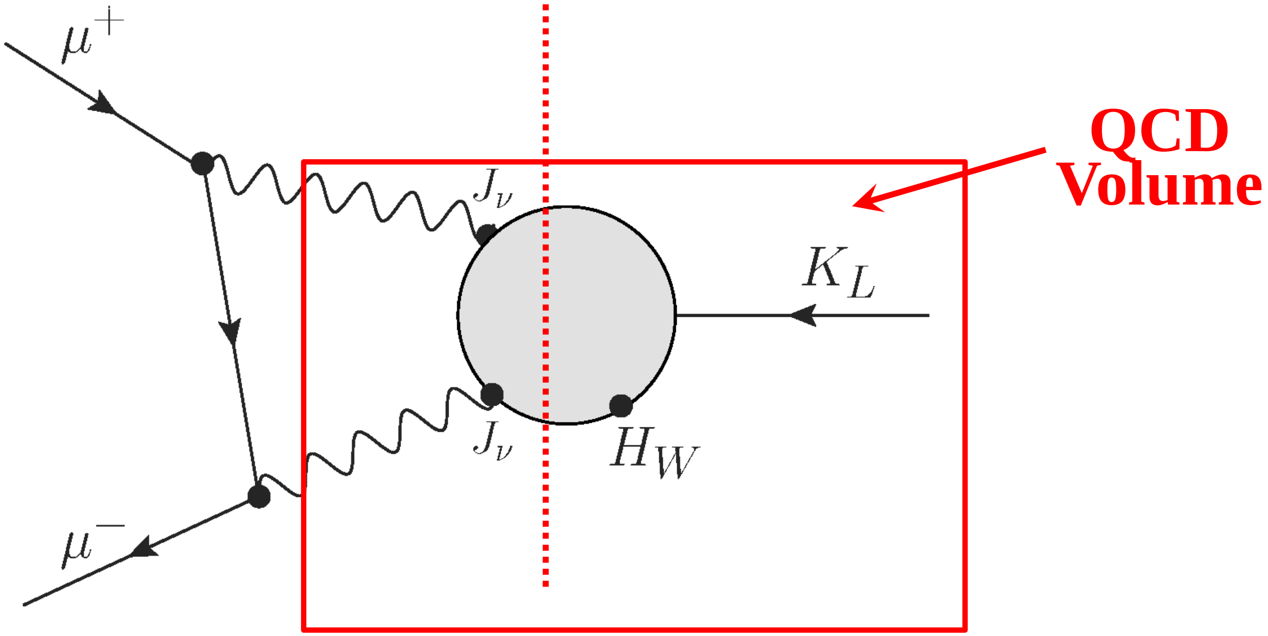
Hadronic light-by-light scattering from lattice QCD

(L. Jin)



$\pi^0 \rightarrow e^+e^-$ & $K_L \rightarrow \mu^+\mu^-$ Decays

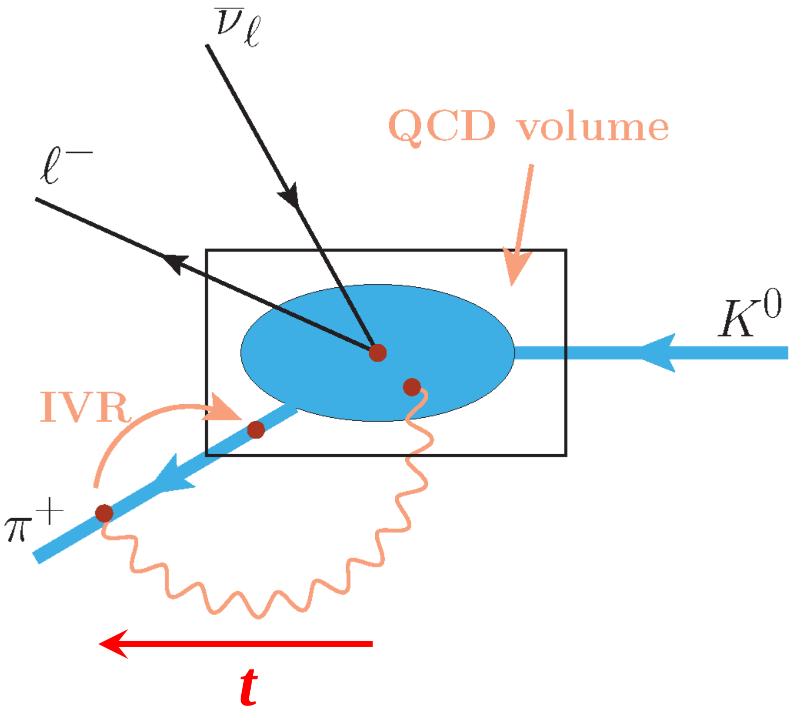
(Y.Zhao, EH Chao, C. Hu)

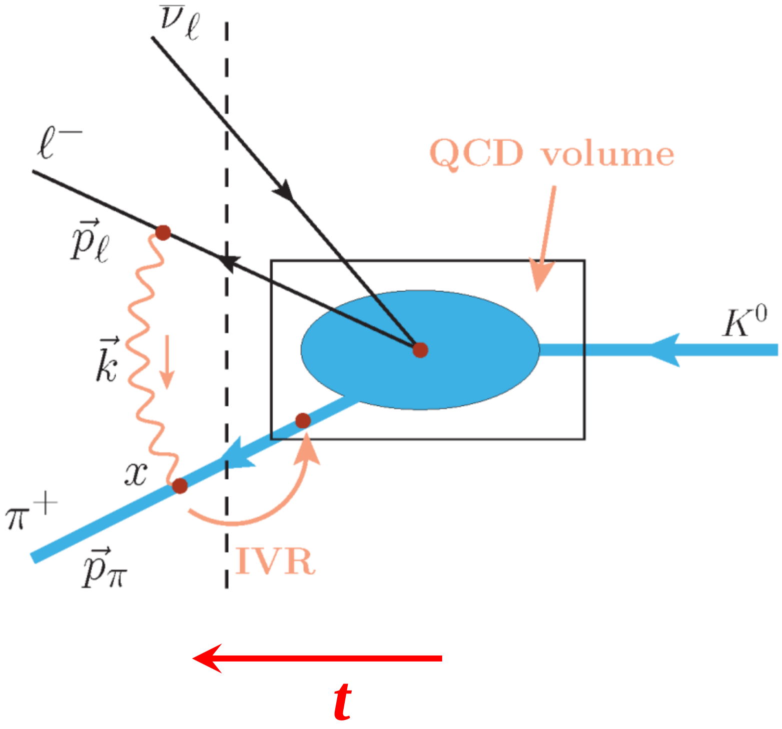


QCD Volume

- Must remove a p intermediate state
- Perform a Wick rotation resulting in complex E&M factor and Euclidean QCD Green’s function

QED Corrections to $K^0 \rightarrow \pi^+ l^- \nu$





- Large time ***t*** is not further suppressed
- Reconstruct full Minkowski amplitude from Euclidian pion-emission Greens function, (IVR: L. Jin and X. Feng)

Conclusion

- Lattice QCD → HPC → Lattice QCD
- Symbiosis of lattice QCD and HPC remains strong [Listen to Peter Boyle this afternoon for the software dimension and g_μ-2.]
- First-principles LQCD calculations profoundly changing search for new physics.
- Precision advances:

10% (2015) → 1% (2020) → 0.1% (2025)

→ 0.01% (20??)

Thank you!

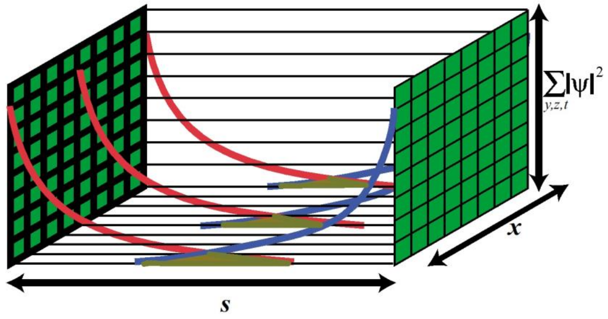
Lessons learned

- Focus on real physics goals: clear metric for what is needed and what is not
- Aim for 10-100x enhancements
  - Hardware projects take longer than imagined.
  - Final performance always compromised.
- Not theoretical physics: team and mentors are critical.
  - Highly-talented collaborators critical
  - Ethical, reliable colleagues essential
  - Powerful supporters necessary.
    - T.D. Lee (Columbia)
    - Nick Samios (BNL)
    - Randy Issacs (IBM)

Columbia
QCD Machines

Elaborate methods required

- Use 5-D, domain wall lattice fermions – physical quarks bound to 4D boundaries



- Use a 96^3 x 192 → 128^3 x 288 → 160^3 x 360 lattice
- Compute 8000 lowest Dirac eigenvectors to speed up Dirac operator inversion.
- Frontier machine at ORNL has 9472 nodes, each with one AMD EPYC CPUs and 4 MI250X GPUs, complex memory and communications hierarchies
- Broad collaboration needed.

Path Integral Formulation

- Asymptotic freedom justifies a lattice action defined at weak coupling.
- Stochastic evaluation exploits the large number of field theory variables while treating them exactly.
- Monte Carlo Markov chain importance sampling is extremely effective: 1% accuracy from 10 samples.
- Euclidean $e^{-H_{QCD}\,t}$ projects onto stable H_{QCD} eigenstates:
  - Correlation lengths give particle masses.
  - Matrix elements of physical operators directly evaluated.
- Clearly the key to solving non-perturbative gauge theory at low-energy