Heavy Quarks physics, codes and machines

Yasumichi Aoki : RIKEN Center for Computational Science (R-CCS)

EXALAT kickoff meeting
June 15, 2020

contents

1) HQ application

2) machines

3) codes

this is not quite logical (maybe not pedagogical) talk on the future lattice QCD

sorry about this.

But, mostly (my) mixture of activities related to Lattice QCD towards Fugaku.

thanks to

B meson applications / Estimates

Takashi Kaneko

QWS

Yoshifumi Nakamura, Issaku Kanamori

Fugaku / A64FX etc

Yoshifumi Nakamura, Issaku Kanamori, Keigo Nitadori

knowledge acquired through discussion with

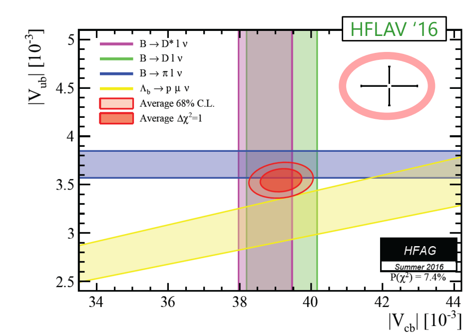
Grid: *Peter Boyle, Guido Cossu,*

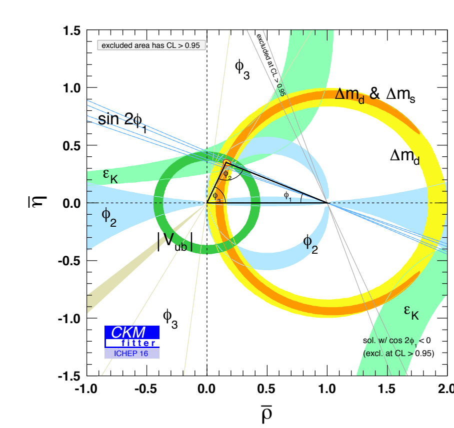
A64FX extension of Grid: *Tilo Wettig, Nils Meyer*

Bridge++: *Issaku Kanamori, Hideo Matsufuru*

HQ applications - in particular B physics

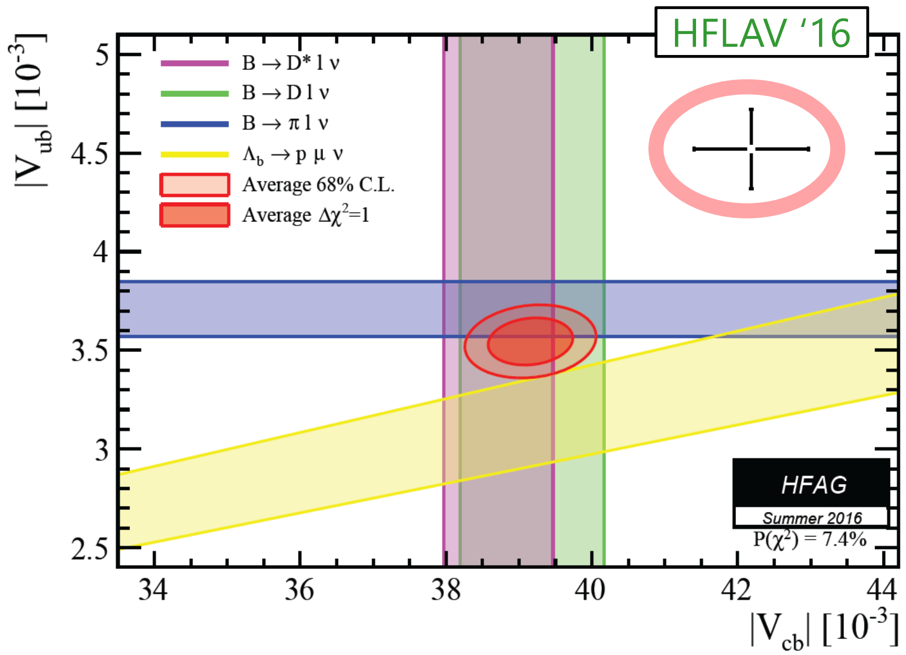
- stringent test of the Cabibbo-Kobayashi-Maskawa(CKM) paradigm
- new physics appear as inconsistency
  - over-constraining the parameters
  - check of unitarity
- Example Hints
  - tension in $R\left(D^{(*)}\right)=\frac{\Gamma\left(B\to D^{(*)}\tau\nu\right)}{\Gamma\left(B\to D^{(*)}\{e,\mu\}\nu\right)}$
  - tension in exclusive <-> inclusive on V_{ub} - V_{cb} plot
  - tension in unitarity triangle
  - tension in rare processes: B→K l l,

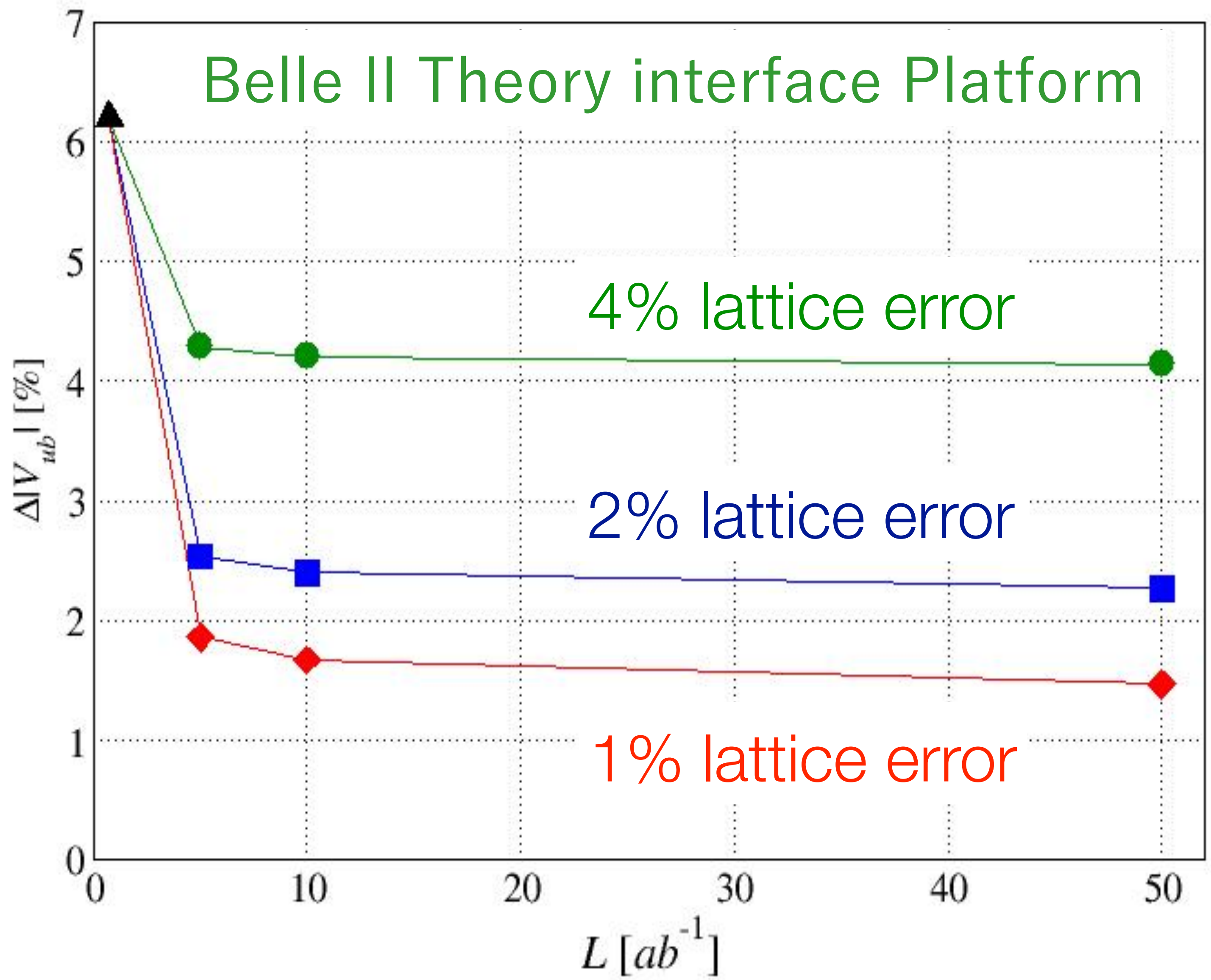




Forecasting experiment, plan for lattice

- Ex: B→πlν @ Belle II /
  - 2020-2022: 10 ab^{-1}
    - 5x Belle
  - -2025: 50 ab^{-1}



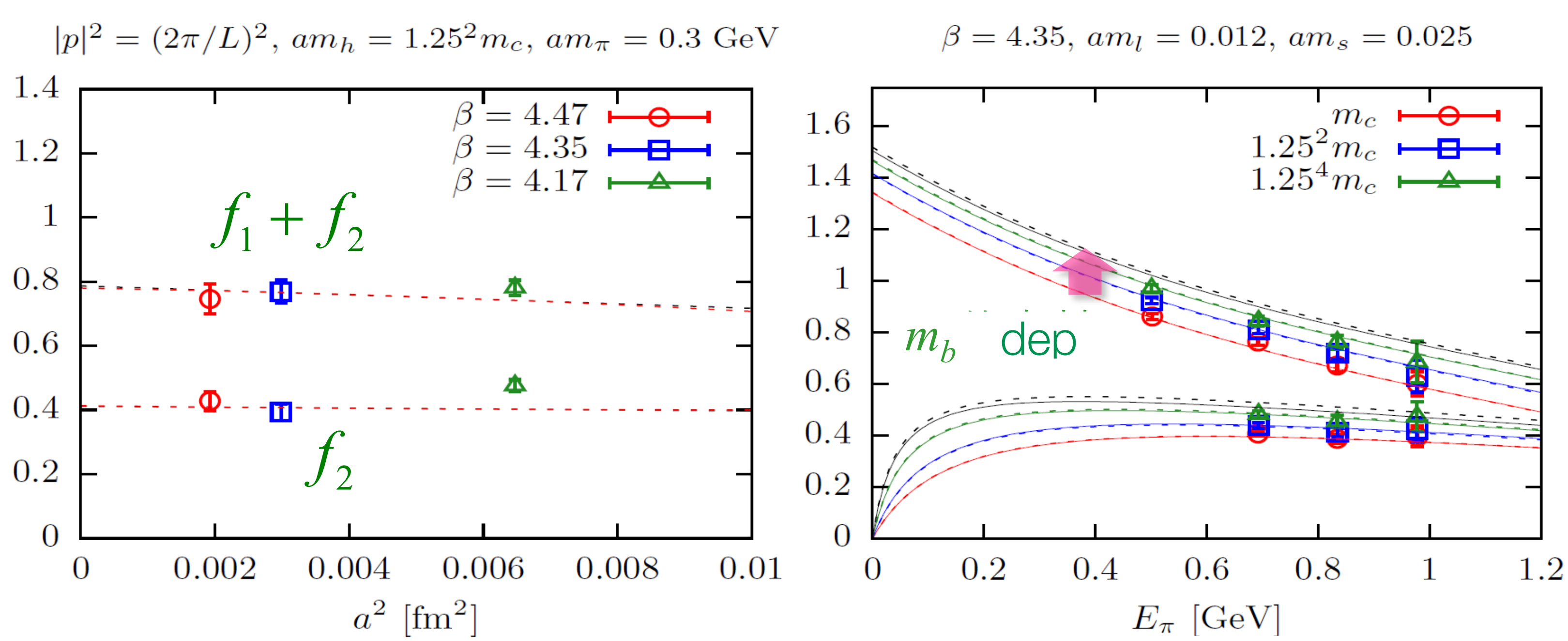


courtesy: Takashi Kaneko

HQ applications on the Lattice

- Lattice difficulties
  - m_b ~ 4 GeV, 1/a ~ 2-3 GeV
  - m_b a > 1 for 1/a < 4 GeV → light quark formulation fails → EFT
  - recent JLQCD comp. $1/a|_{max}$ = 4.6 GeV for Domain wall fermions
    - should work if m_b << 4 & check m_b → m_b^{phys} dependence
      - $m_b = 1.25^4\ m_c^{phys}$ ~ 3 GeV seems OK

B. Colquhoun, J. Koponen @ Lattice'19



HQ applications on the Lattice: Why DWF

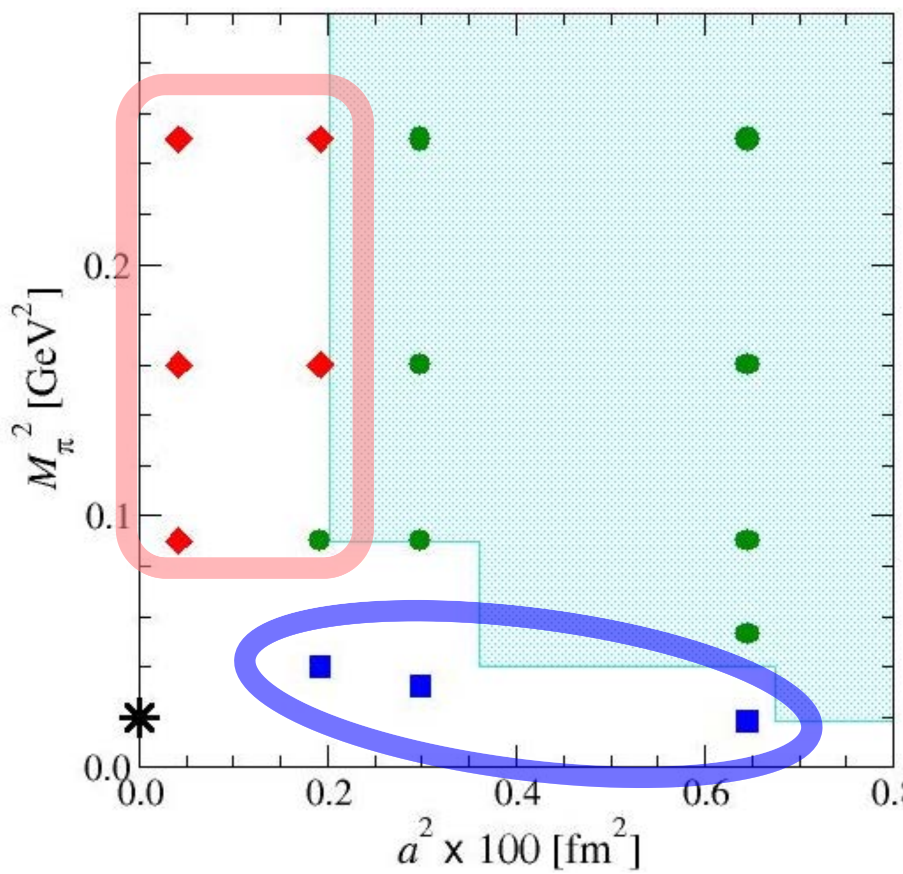
- symmetry of continuum QCD
- good continuum scaling
- most simple operator renormalization over various lattice approaches
  - once configurations were generated, lots of opportunities for applications
- useful for light meson / baryon physics too

HQ applications on the Lattice

- 2 step approach in plan in JLQCD / 成果創出加速プログラム

  Program for Promoting Researches on the Supercomputer Fugaku (Simulation for basic science: from fundamental laws of particles to creation of nuclei)

  - Shoji Hashimoto, Takashi Kaneko et al
  - (2020)-2022: b quark physical point
    - $1/a = 9.6$ GeV > m_b
    - 5.5 PFlops*yrs
    - aiming $\Delta|V_{ub}| < 5\%$
  - later.. u, d quark physical point
    - L~ 6 fm
    - 8.2 PFlops*yrs
    - together with/light quark physics, possibly larger collaboration



HQ applications on the Lattice

- All lattice quarks = light quarks
  - Domain wall fermions (DWF)
  - most of the computational effort goes in the DWF light quark solver
    - predominantly DWF Mult (mostly 4D Wilson matrix multiplication)
    - typical efficiency ~ 4% (KNL/Grid), 20% (Skylake/Grid)
  - assuming ~15% efficiency on Fugaku (to be checked)
  - assuming effort for HMC : measurements = 1 : 2.5
  - assuming existing simulation statistics and a scaling as defined in https://hpci-aplfs.r-ccs.riken.jp/document/roadmap2017/roadmap_170713.pdf (Japanese Computational Science Road Map 2017)

machines: pre Fugaku

Oakforest-PACS

KNL(68 cores) & Omni-Path

memory/node: 96 GB(DDR4) + 16 GB(MCDRAM),

115 GB/s, 490 GB/s

0.04 Byte/Flop, 0.16 Byte/Flop

eff: ~4% (Grid DWF mult)

Skylake

(RIKEN - Hokusai BigWaterFall: network: EDR 12.6GB/s (bidirectional)

Intel Xeon Gold 6148 (2.4GHz) x2 x840

memory/node: 96 GB, 255GB/s, 0.08Byte/Flop

eff: ~20% (Grid DWF mult)

machines

Fugaku https://postk-web.r-ccs.riken.jp/

A64FX

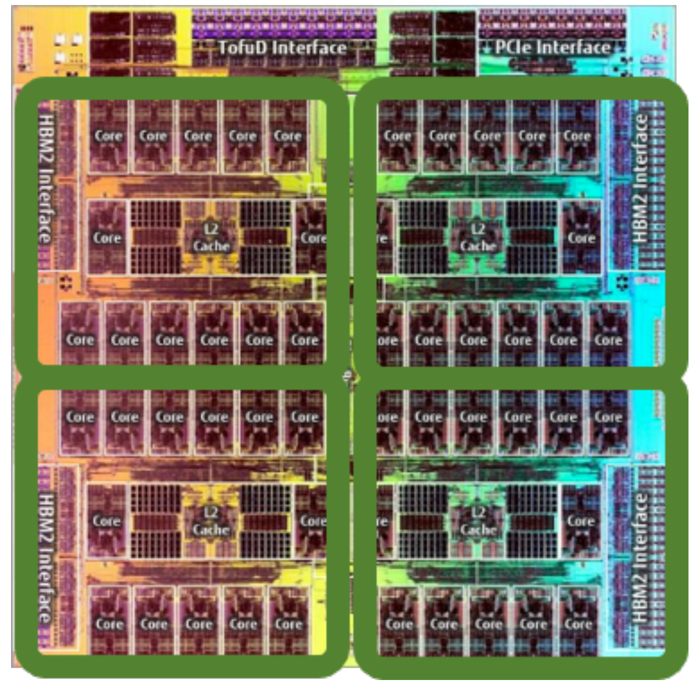
Architecture Information: Download from https://github.com/fujitsu/A64FX

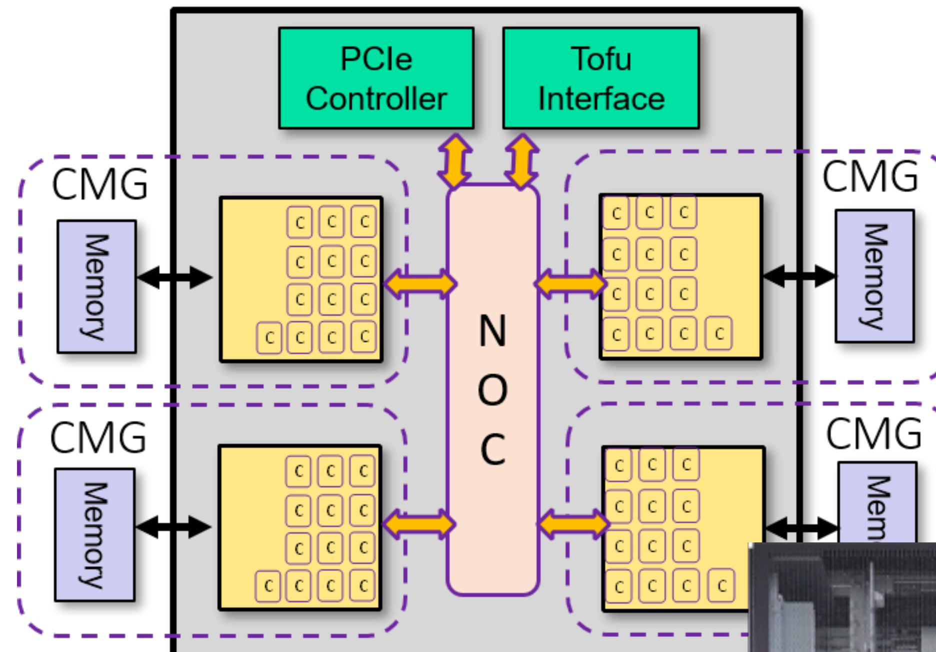
	Description	
Architecture	Armv8.2-A SVE (512 bit SIMD)	
Core	48 cores for compute and 2/4 for OS activities	
	Normal: 2.0 GHz	DP: 3.072 TF, SP: 6.144 TF, HP: 12.288 TF
	Boost: 2.2 GHz	DP: 3.3792 TF, SP: 6.7584 TF, HP: 13.5168 TF
Cache L1	64 KiB, 4 way, 230+ GB/s (load), 115+ GB/s (store)	
Cache L2	CMG(NUMA): 8 MiB, 16 way Node: 3.6+ TB/s Core: 115+ GB/s (load), 57+ GB/s (store)	
Memory	HBM2 32 GiB, 1024 GB/s	~0.3 Byte/Flop
Interconnect	TofuD (28 Gbps x 2 lane x 10port)	6d torus
I/O	PCIe Gen3 x 16 lane	
Technology	7nm FinFET	

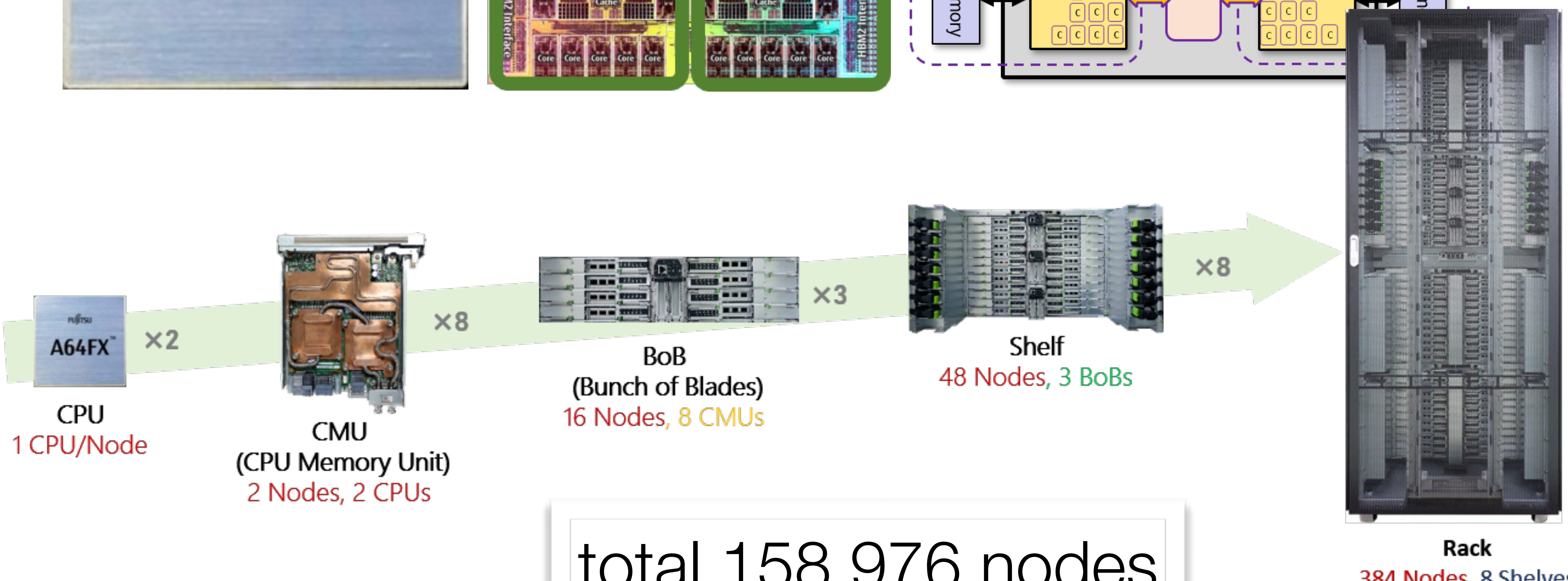
machines

- Fugaku - A64FX

https://postk-web.r-ccs.riken.jp/







total 158,976 nodes

Courtesy of FUJITSU LIMITED

codes

B meson applications

Iroiro++ → Grid

Configurations Generation

Iroiro++ → Grid

(thermodynamics uses Grid)

On fugaku

QWS (QCD wide SIMD library) → to be used in:

- Bridge++
- BQCD https://www.rrz.uni-hamburg.de/services/hpc/bqcd.html
- “Ishikawa code” by Ken-Ichi Ishikawa https://www.ccs.tsukuba.ac.jp/qcd/

Grid

codes - Grid

QCD package from Edinburgh: P. Boyle et al https://github.com/paboyle/Grid

- SIMD efficient
  - 4d hyper cuve → SIMD lane
- efficient on modern Intel type CPUs
- flexible to lattice size variation
- SIMT → GPU's
- may physics application classes

- A64FX extension being developed by Regensburg/RIKEN collaboration
  - N. Meyer's github: https://github.com/nmeyer-ur/Grid

codes - QWS

QCD library for Fugaku by Y.Nakamura, Y.Mukai, KI.Ishikawa, I.Kanamori

(LQCD co-design WG for Fugaku) https://github.com/RIKEN-LQCD/qws

- one lattice dimension (x) → SIMD lane
- tuned on Fugaku: the LQCD target application
  - O(a) improved Wilson fermion w/ $L^4 = 192^4$
  - mixed precision BiCGstab + Lüscher's DD preconditioning (25+ * K)
  - good efficiency on Fugaku
    - local volume is small: 32*6*4*3
      - mostly fit in L2 cache (no scaling study for larger V so far)
    - for SIMD : L_x = 16*n (for double prec.), 32*n (for single prec.)
- open to public from March 2020
- slides by Y.Nakamura:

Performance Targets

- ✓ 100 times faster than K for some applications (tuning included)
- ✓ 30 to 40 MW power consumption

Peak Performance

	PostK	K
Peak DP (double precision)	400+ Pflops (34x +)	11.3 Pflops*
Peak SP (single precision)	800+ Pflops (70x +)	11.3 Pflops
Peak HP (half precision)	1600+ Pflops (141x +)	--
Total memory bandwidth	150+ PB/sec (29x +)	5,184TB/sec

* Reported in TOP500 (including I/O nodes)

Geometric Mean of Performance Speedup of the 9 Target Applications over the K-Computer

37x +

Predicted Performance of 9 Target Applications

As of 2019/05/14

Area	Priority Issue	Performance Speedup over K	Application	Brief description
Health and longevity	1. Innovative computing infrastructure for drug discovery	125x +	GENESIS	MD for proteins
	2. Personalized and preventive medicine using big data	8x +	Genomon	Genome processing (Genome alignment)
Disaster prevention and Environment	3. Integrated simulation systems induced by earthquake and tsunami	45x +	GAMERA	Earthquake simulator (FEM in unstructured & structured grid)
	4. Meteorological and global environmental prediction using big data	120x +	NICAM+ LETKF	Weather prediction system using Big data (structured grid stencil & ensemble Kalman filter)
Energy issue	5. New technologies for energy creation, conversion / storage, and use	40x +	NTChem	Molecular electronic simulation (structure calculation)
	6. Accelerated development of innovative clean energy systems	35x +	Adventure	Computational Mechanics System for Large Scale Analysis and Design (unstructured grid)
Industrial competitiveness enhancement	7. Creation of new functional devices and high-performance materials	30x +	RSDFT	Ab-initio simulation (density functional theory)
	8. Development of innovative design and production processes	25x +	FFB	Large Eddy Simulation (unstructured grid)
Basic science	9. Elucidation of the fundamental laws and evolution of the universe	25x +	LQCD	Lattice QCD simulation (structured grid Monte Carlo)

codes - Bridge ++

A common QCD code set developed in Japan

http://bridge.kek.jp/Lattice-code/

- original paper: S.Ueda, S.Aoki, T.Aoyama, K.Kanaya. H.Matsufuru, "Development of an object oriented lattice QCD code 'Bridge++'", J.Phys.Conf.Ser. 523 (2014) 012046

- Our goal is to develop a code which includes various lattice actions and numerical algorithms, supports wide range of architectures from laptop to supercomputer, has sufficient performance for practical researches, while simultaneously is easy to be handled.

- eg: data array type can be easily changed

- → calling QWS for Fugaku

Code for DWF on Fugaku

- plan
  - QWS → Bridge ++ *mostly for HMC*
    - if no severe bottle necks found, this would be most efficient
    - size limitation : $L_x = 16n$ (double prec):
      - OK with large lattice
      - not good for finite temperature
  - Grid w/ Regensburg/RIKEN extension *measurements / HMC*
    - running version already there
    - tuning for A64FX underway / Regensburg