ICEPP QC and HPC Research Activity

Ryu Sawada / Masahiko Saito

QC → Ryu Sawada

HPC → Masahiko Saito

3rd Nov, 2023

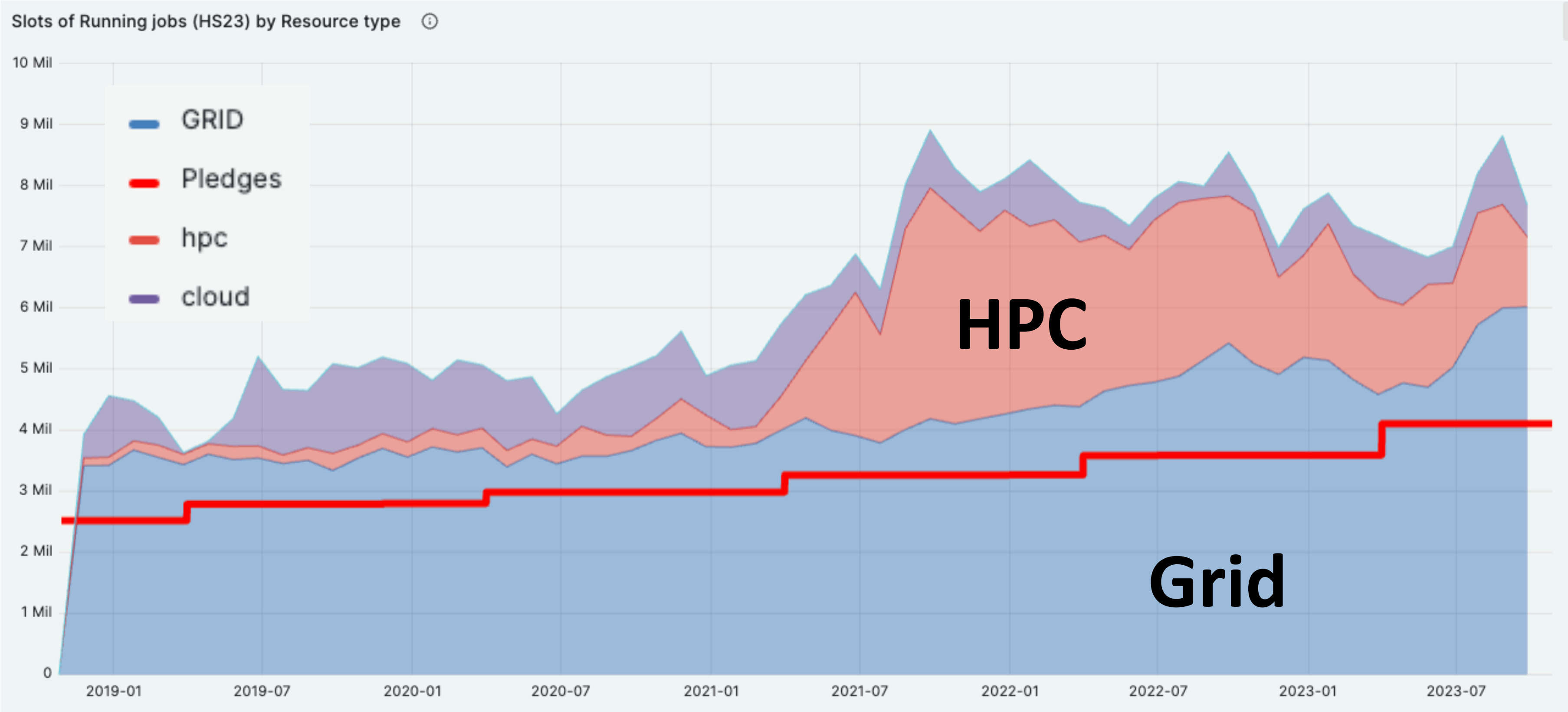
The 7th Asian Tier Center Forum

ICEPP
The University of Tokyo

Thank:
Y. Iiyama : Computer cluster for QC simulation
S. Chen : QC hardware researches

HPC

HPC in ATLAS experiments



May 2021

- EuroHPC Vega (#166 in Top500) is in production for ATLAS form May 2021.
- ATLAS uses a lot of HPC resources. The use of HPC is very promising.

HPC top500

Top500 (June 2023)

	System	Cores (k)	Rmax (PFlops)	Rpeak (PFlops)	Power (MW)
1	**Frontier**, DOE/SC/Oak Ridge National Laboratory, United States	8,700	1,194	1,680	22.7
2	**Supercomputer Fugaku,** RIKEN Center for Computational Science, Japan	7,631	442	537	29.9
3	**LUMI,** EuroHPC/CSC, Finland	2,220	309	429	6.02
4	**Leonardo**, EuroHPC/CINECA, Italy	1,825	239	304	7.40
5	**Summit,** DOE/SC/Oak Ridge National Laboratory, United States	2,415	149	201	10.1
6	**Sierra,** DOE/NNSA/LLNL, United States	1,572	95	126	7.44
7	**Sunway TaihuLight,** National Supercomputing Center in Wuxi, China	10,650	93	125	15.4
8	**Perlmutter,** DOE/SC/LBNL/NERSC, United States	762	71	94	2.59
9	**Selene**, NVIDIA Corporation, United States	556	63	79	2.65
10	**Tianhe-2A,** National Super Computer Center in Guangzhou, China	4,982	61	101	18.5

- “Fugaku” is in second place.
- The total number of CPU cores in the WLCG is ~1M cores. If HPC can be used, it will be a very promising computing resource.

HPC top500 (in Japan)

Top500 (June 2023)

	System
2	**Supercomputer Fugaku**, RIKEN Center for Computational Science
24	**ABCI 2.0**, National Institute of Advanced Industrial Science and Technology (AIST)
25	**Wisteria/BDEC-01 (Odyssey)**, Information Technology Center, The University of Tokyo
41	**TOKI-SORA**, Japan Aerospace eXploration Agency
50	**???**, Japan Meteorological Agency
63	**Earth Simulator -SX-Aurora TSUBASA,** Japan Agency for Marine-Earth Science and Technology
80	**TSUBAME3.0**, GSIC Center, Tokyo Institute of Technology
84	**Plasma Simulator**, National Institute for Fusion Science (NIFS)
97	**Flow**, Information Technology Center, Nagoya University
	…
136	**Wisteria/BDEC-01 (Aquarius)**, Information Technology Center, The University of Tokyo
140	**Oakbridge-CX**, Information Technology Center, The University of Tokyo

- There are several high-performance HPCs in Japan.
- Information Technology Center of the University of Tokyo manages some of them.
  - We have advanced R&D running grid jobs on the ITC HPCs.

History of HPC utilization in ICEPP

- We started R&D on ITC/UTokyo HPC from 2019 using **Reedbush** system (2016-2020)
- From 2020, we moved to the next generation system: **Oakbridge-CX** (2019-2023/09)

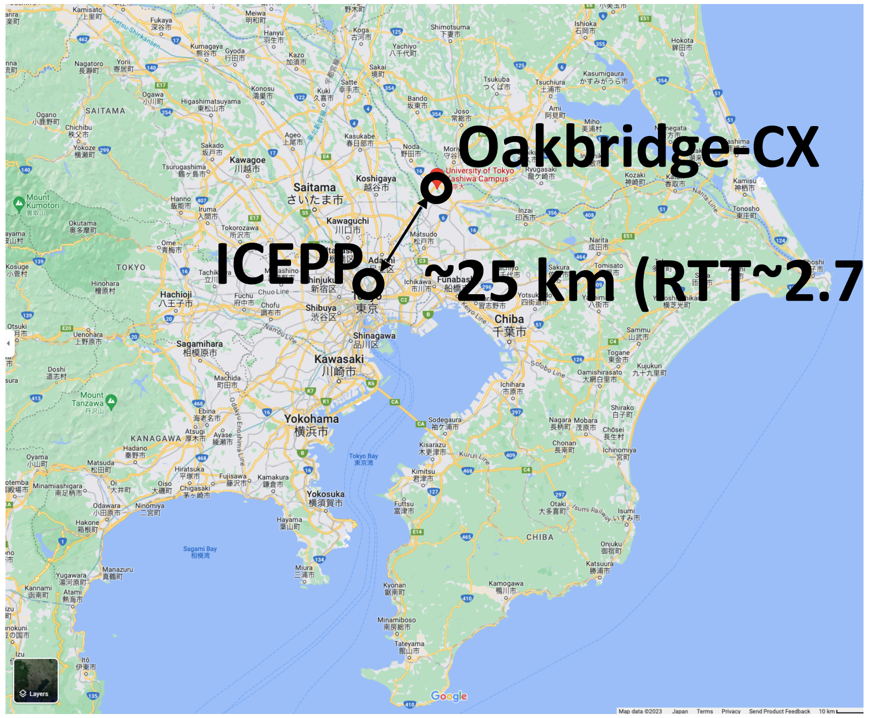
  → We report a **summary** of the integration of HPCs into the Tier2 grid.

- The next generation system is **Wisteria/BDEC-01** (2021-)

  → We report an **overview** of the system and the **difficulties** in using it.

	System	Cores (k)	Rmax (PFlops)	Rpeak (PFlops)	Power (kW)	Year
25	**Wisteria/BDEC-01 (Odyssey)**, Information Technology Center, The University of Tokyo	369	22.1	26.0	1,468	2021-
136	**Wisteria/BDEC-01 (Aquarius)**, Information Technology Center, The University of Tokyo	42	4.4	5.8	184	2021-
140	**Oakbridge-CX**, Information Technology Center, The University of Tokyo	77	4.3	6.6	845	2019-2023
—	TOKYO Tier2	11	1.2	-	120	2022-

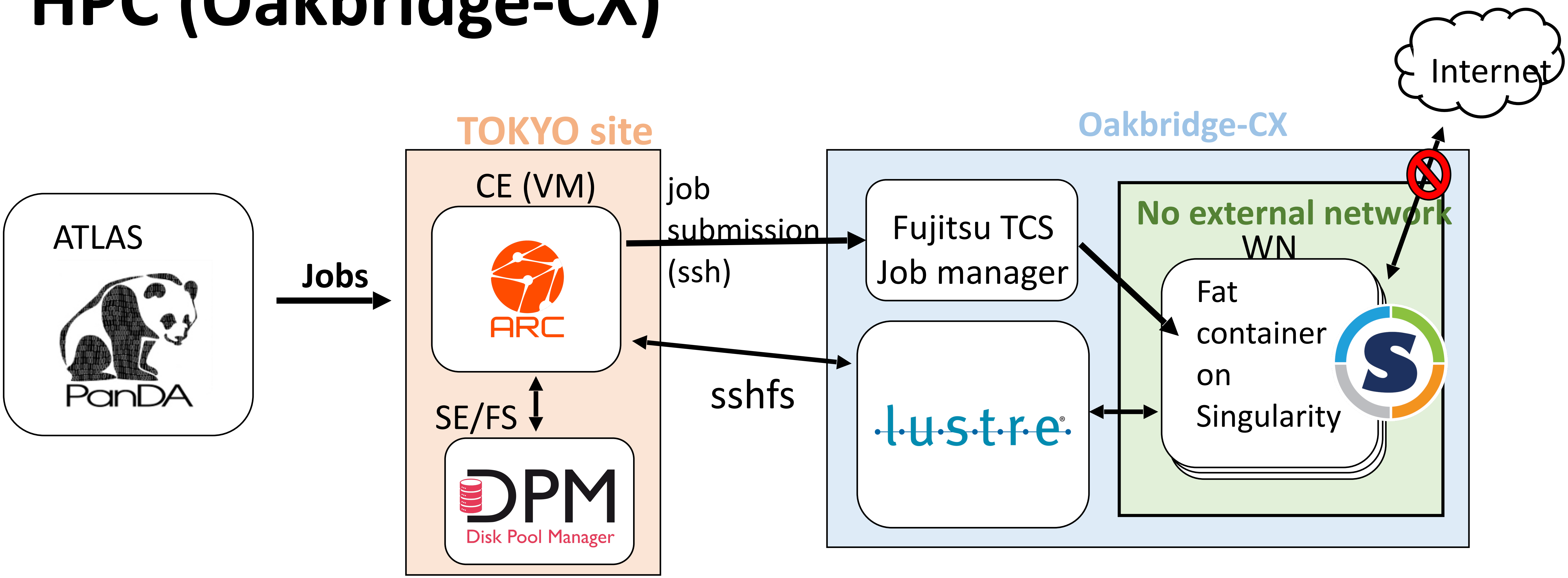
Oakbridge-CX



- Compute nodes (only CPU, no GPU)
  - 1368 compute node, 6.61 PFlops
  - 56 cores / node, 1148 HS06 / nodes
- File system
  - Lustre, 12.4 PB
- Batch system
  - FUJITSU Software Technical Computing Suite (TCS)
- Network connectivity
  - ssh to login nodes, where we can submit jobs and read/write to shared FS.
  - No connections to computing nodes.

  → Grid jobs cannot access storage element, external DB, etc.
- No root privilege → We cannot use CVMFS

HPC (Oakbridge-CX)



- Singularity container image is used.
  - contains all necessary files
  - processes simulation jobs only
- Input/output files are transferred by ARC.

- All necessary files on cvmfs are pre-downloaded to the shared FS on HPC, which can be accessed by compute nodes.
- No negligible overhead.

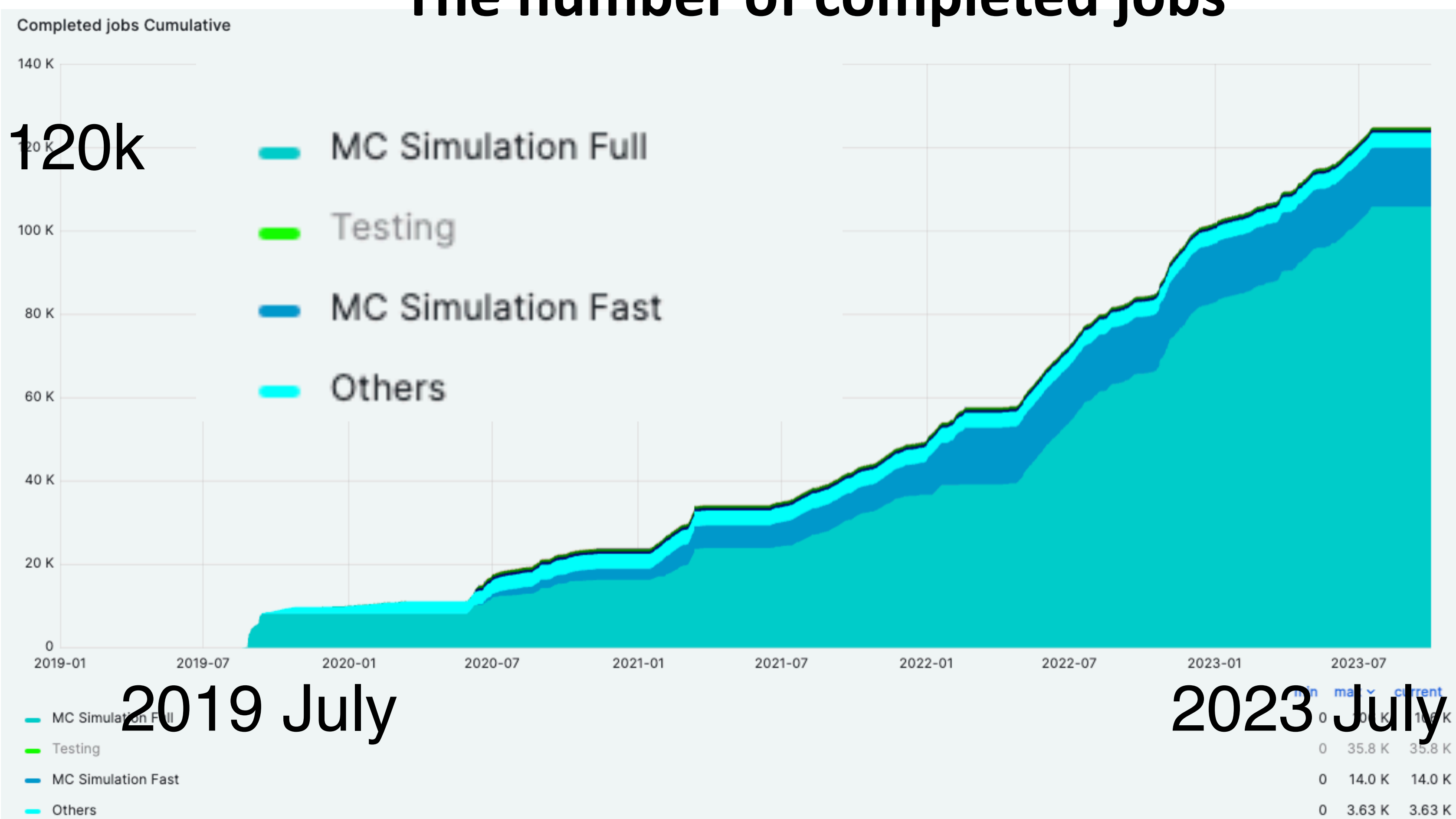
- Before using Singularity container, we used parrot_run + cvmfs_preload.

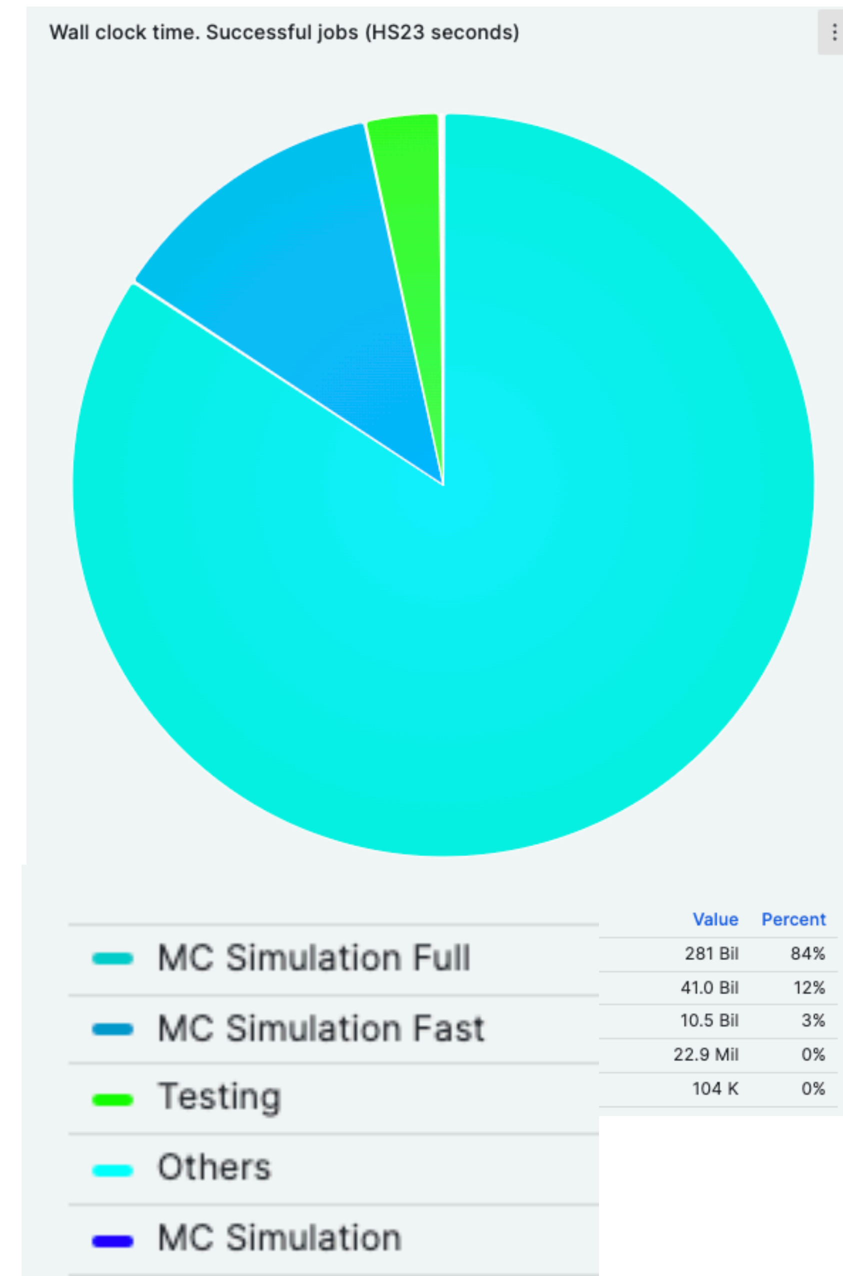
Jobs accounting (History)

[Grafana]

Wall clock time (successful jobs) (HS23 sec)

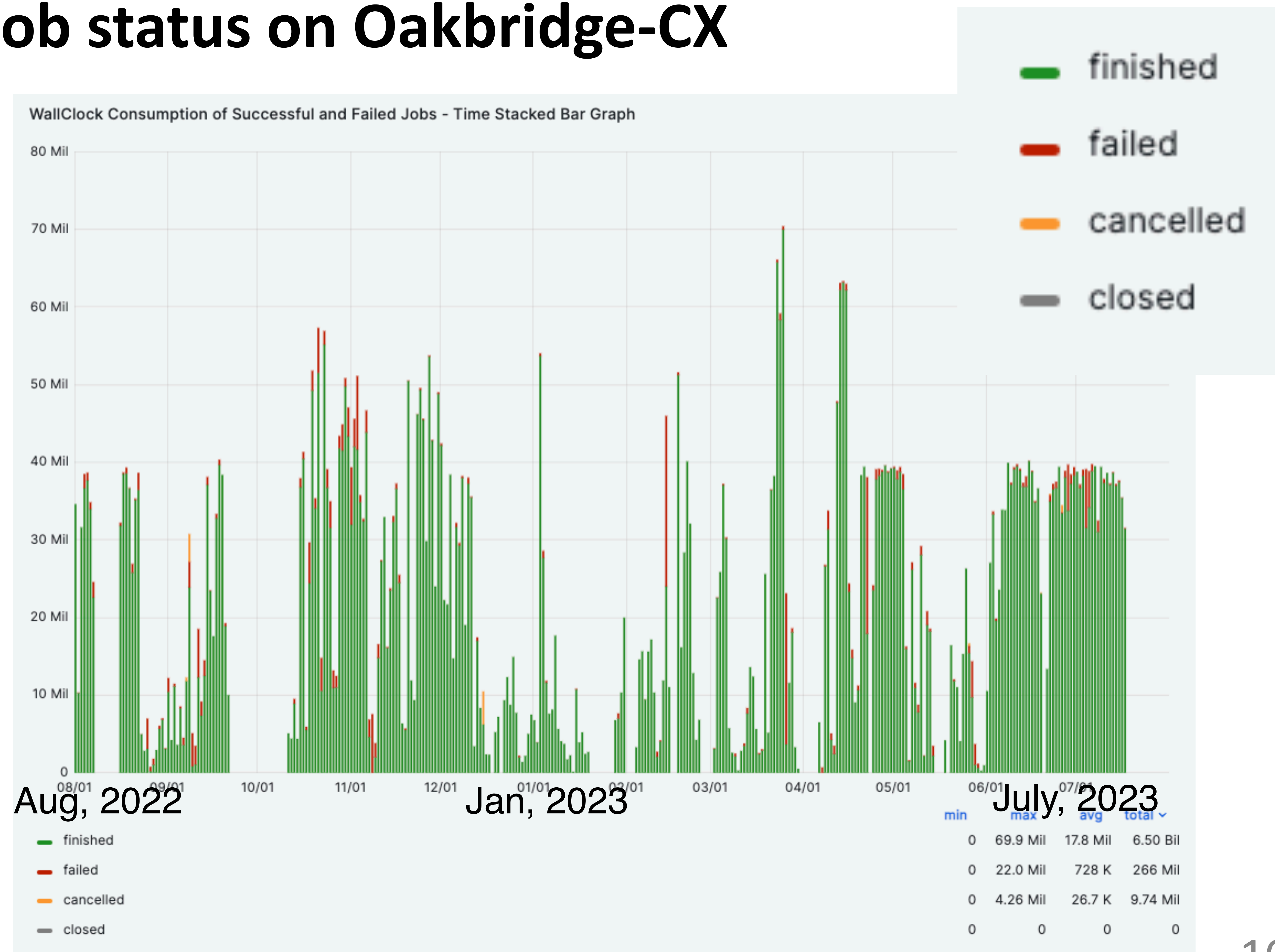
The number of completed jobs



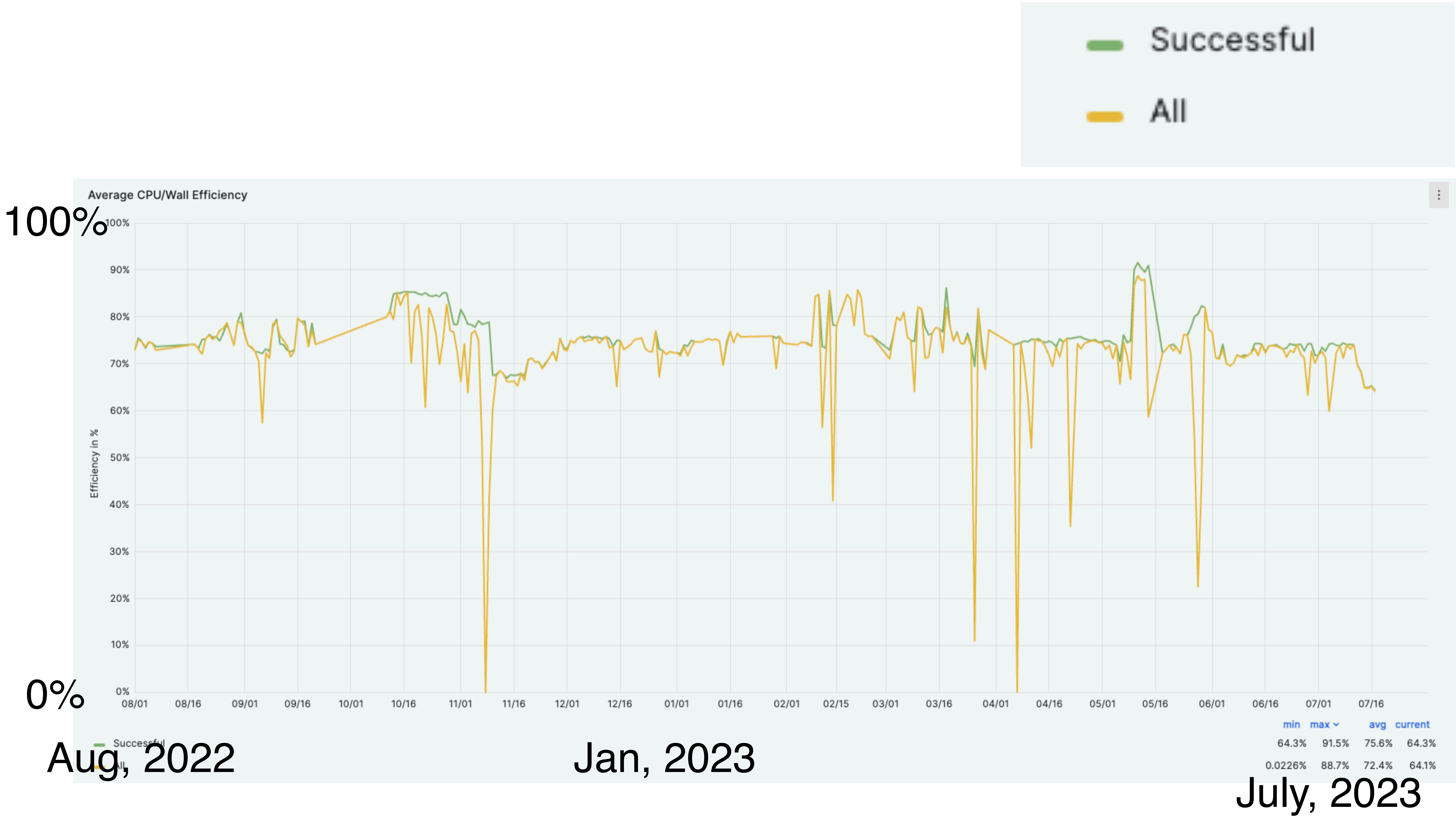


- 120,000 jobs processed
- 330 G HS06 seconds → ~15 days of current Tokyo Tier2 full power

Job status on Oakbridge-CX



CPU/Wall efficiency on Oakbridge-CX



Next HPC systems of ITC/UTokyo: Wisteria/BDEC-01

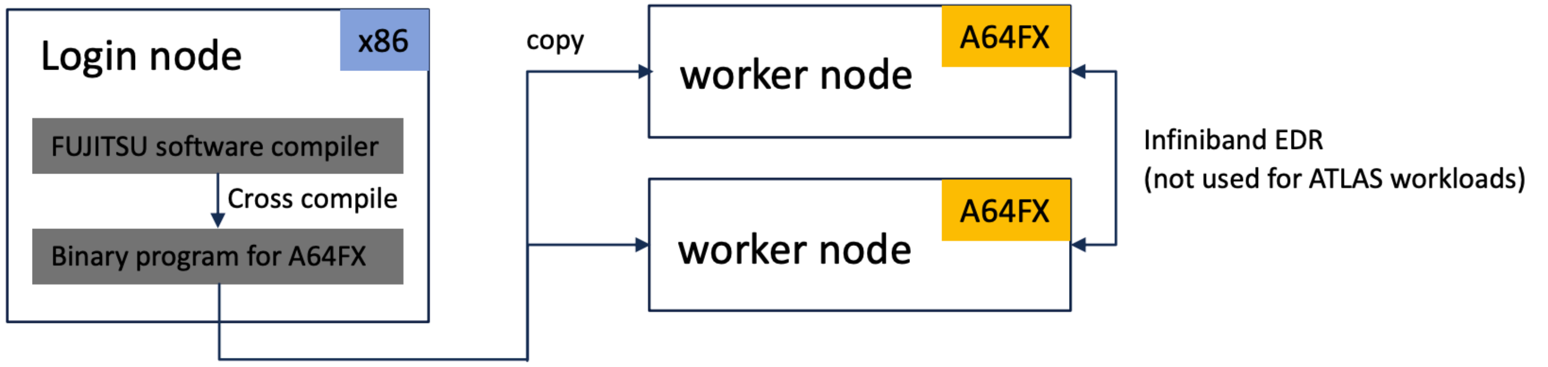
- Consists of two systems
  - Simulation node cluster (25.9 PFlops) → **ARM** CPU
    - FUJITSU Processor A64FX used at Fugaku (the top HPC in Japan)
  - Data/training node cluster (7.2 PFlops) → **GPU**
    - Nvidia A100 x8

  → Suitable for large-scale parallel computing and machine learning

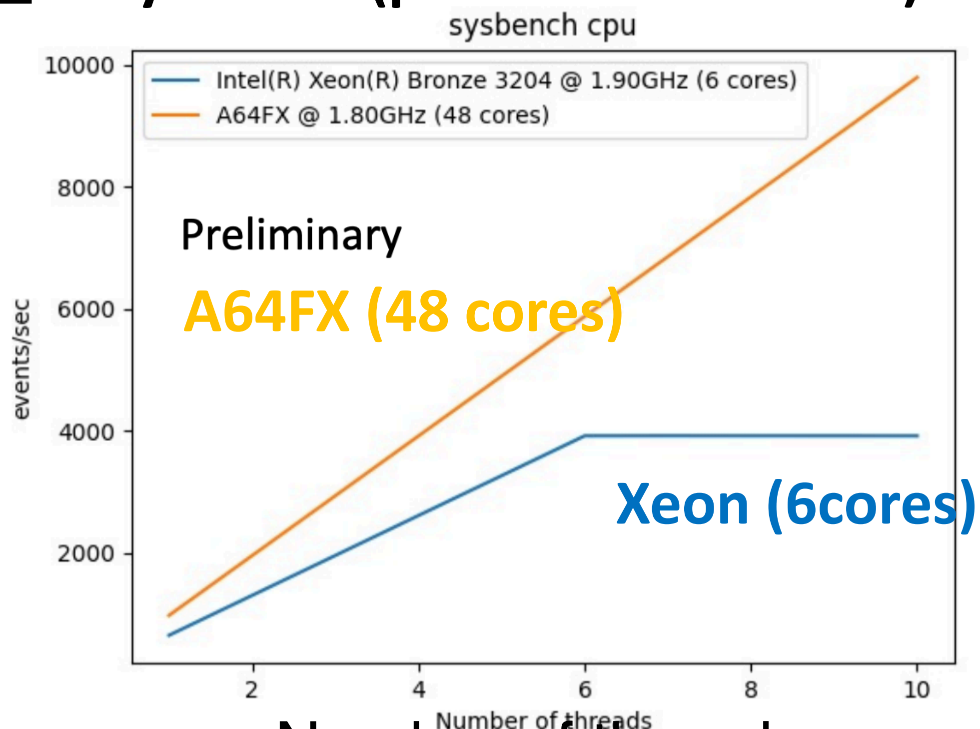
- This HPC cannot use HEP-standard processing unit, such as Intel x86_64 CPU.
  - A lot of R&D is needed to use them with high efficiency.
  - ARM is already supported in ATLAS → Benchmarking (see next page)
  - GPU as a production job is not yet supported in ATLAS.

Performance of A64FX: basic benchmark

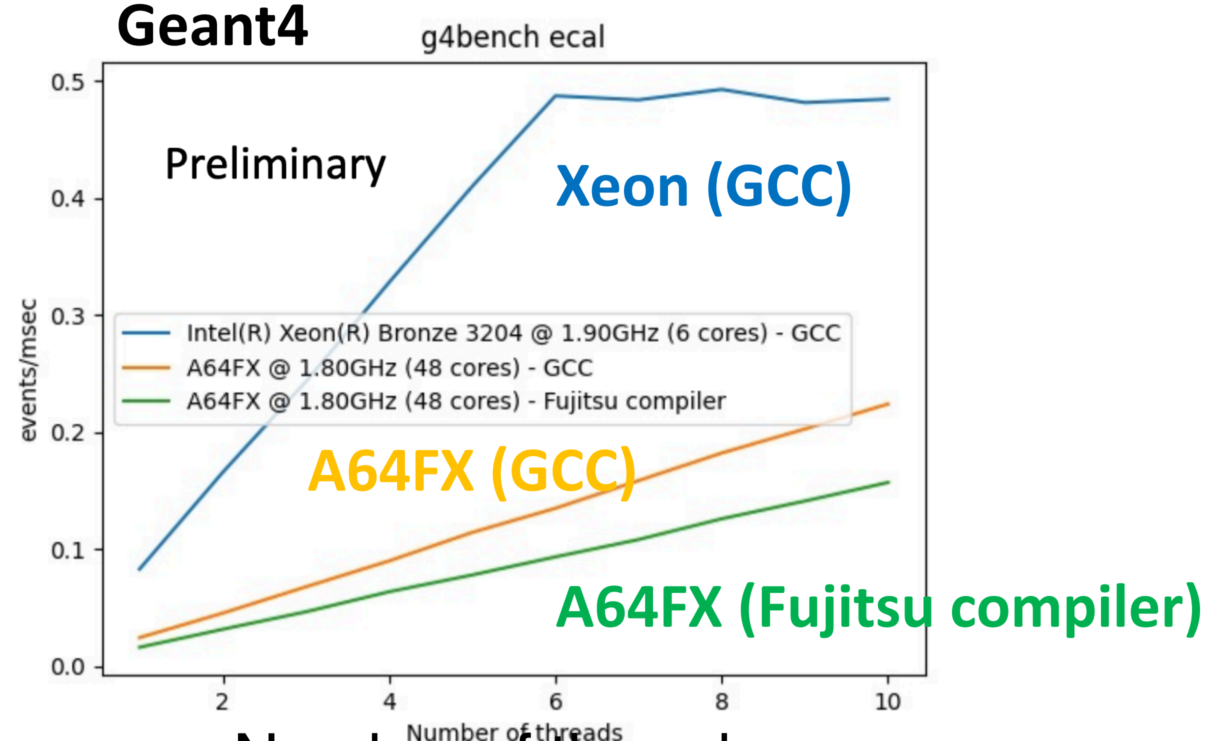
Testbed: PRIMEHPC FX700
(A64FX x 2)



Throughput: **SysBench (prime number calc.)**



Geant4

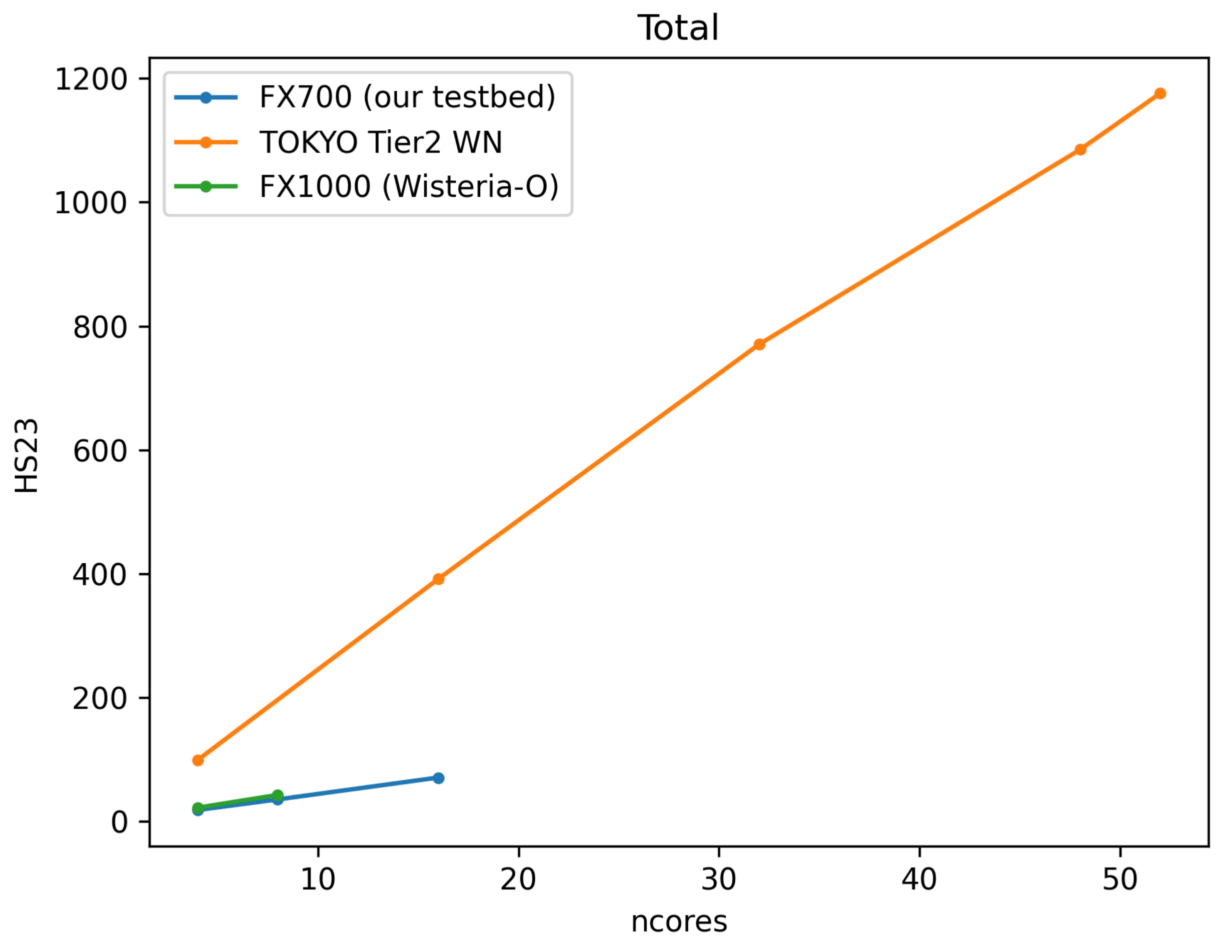


Number of threads

Number of threads

Need to code optimization for Geant4 jobs with A64FX

Performance of A64FX: HEPSCORE23



(Throughput, using 4 threads)

Machine	FX700 (ICEPP testbed)	FX1000 (Wisteria)	TOKYO Tier2 WN
ATLAS Gen	0.62	0.75	3.9
ATLAS Reco	0.17	0.20	0.84
CMS Gen-Sim	0.038	0.046	0.24
CMS Reco	0.086	0.10	0.44
LHCb	30	37	180
Belle2	0.25	0.30	1.6
Alice	0.022	0.026	0.083
HEPSCORE23	18	22	98

- The A64FX has a small amount of memory (32GiB) compared to the number of cores (48).
  - Multithreading code is required to use the memory efficiently.
- To maximise the performance of the A64FX, we may need to fully utilize SVE.

Quantum Computer

System One installed in Kawasaki Japan

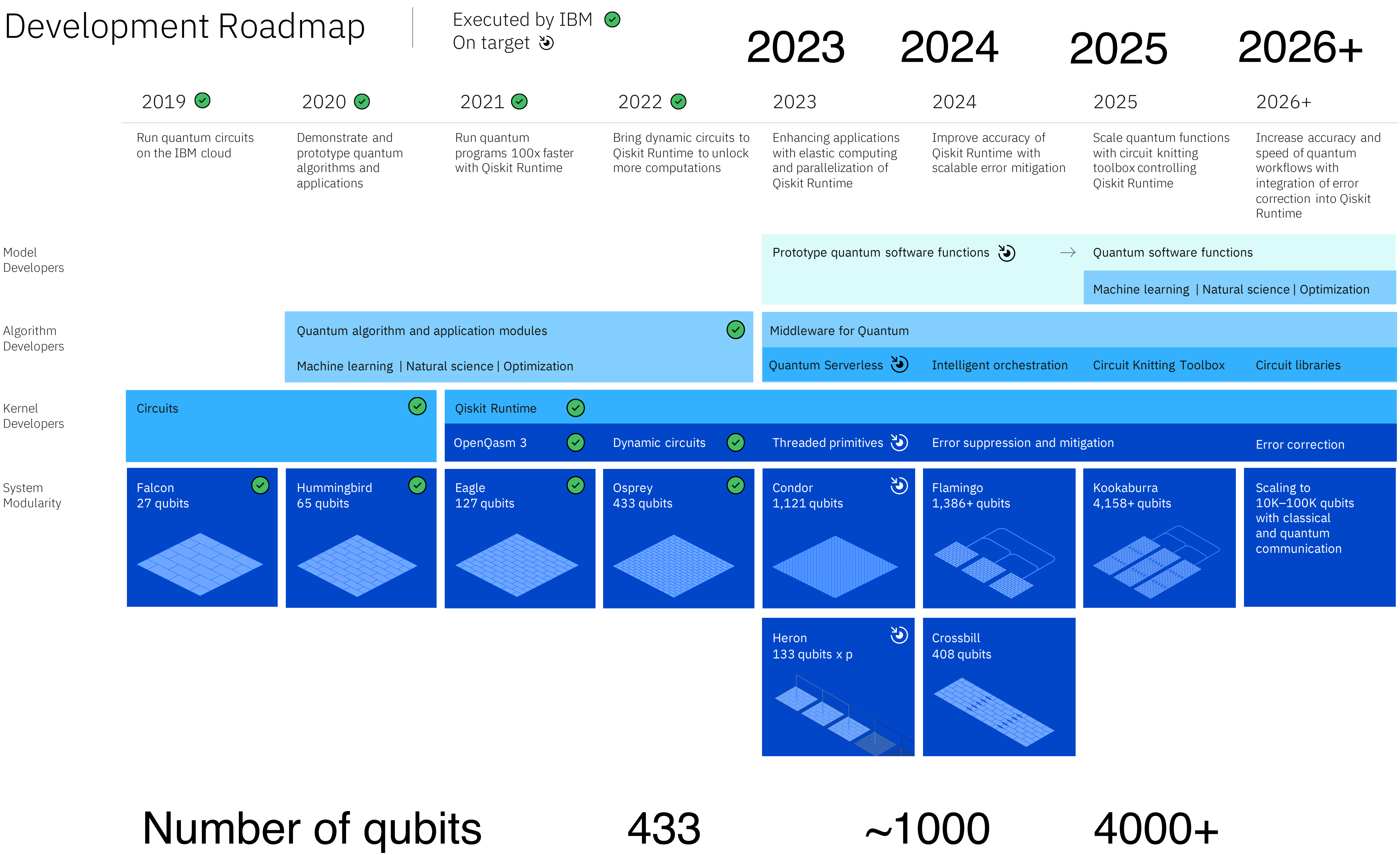
Quantum computer test bed,
in Quantum Hardware Test Center
of UTokyo.

Slide from Alberto Di Meglio at al, CERN QTI 2020

HEP laboratories, e.g. Fermilab, TRIUMF, DESY and ICEPP, have started researches in QC around the beginning of 2020'th

Classical computer v.s. QC

- 1 qubit = two floating numbers.
- The number of quantum states increases **exponentially** as a function of the number of qubits.
  - For simulating a **27 qubit system** using a classical computer, we need **1 GB of memory**.
  - For a **37 qubit system**, **O(1 TB)** is needed.
  - And for **around 49 qubit** or more, **simulation is getting impossible even by HPC**.
  - The number of quantum states that are possible with only **256 qubits exceeds the number of atoms in the solar system**


Development Roadmap
Executed by IBM
On target
2023
2024
2025
2026+
2019
2020
2021
2022
2023
2024
2025
2026+
Run quantum circuits on the IBM cloud
Demonstrate and prototype quantum algorithms and applications
Run quantum programs 100x faster with Qiskit Runtime
Bring dynamic circuits to Qiskit Runtime to unlock more computations
Enhancing applications with elastic computing and parallelization of Qiskit Runtime
Improve accuracy of Qiskit Runtime with scalable error mitigation
Scale quantum functions with circuit knitting toolbox controlling Qiskit Runtime
Increase accuracy and speed of quantum workflows with integration of error correction into Qiskit Runtime
Model Developers
Prototype quantum software functions
Quantum software functions
Machine learning | Natural science | Optimization
Algorithm Developers
Quantum algorithm and application modules
Machine learning | Natural science | Optimization
Middleware for Quantum
Quantum Serverless
Intelligent orchestration
Circuit Knitting Toolbox
Circuit libraries
Kernel Developers
Circuits
Qiskit Runtime
OpenQasm 3
Dynamic circuits
Threaded primitives
Error suppression and mitigation
Error correction
System Modularity
Falcon 27 qubits
Hummingbird 65 qubits
Eagle 127 qubits
Osprey 433 qubits
Condor 1,121 qubits
Flamingo 1,386+ qubits
Kookaburra 4,158+ qubits
Scaling to 10K–100K qubits with classical and quantum communication
Heron 133 qubits x p
Crossbill 408 qubits
Number of qubits
433
~1000
4000+

QC researches in ICEPP, U-Tokyo

including “wish to do” items

HEP application
- Event classification
- Particle tracking

Q-circuit optimisation

Particle dynamics simulation

Quantum ML

Quantum-system Simulation

QC-data learning

QC circuit approximation

MC sampler

ICEPP computer cluster for QC and ML

Y. Iiyama

Cluster shared by QC and machine learning researchers

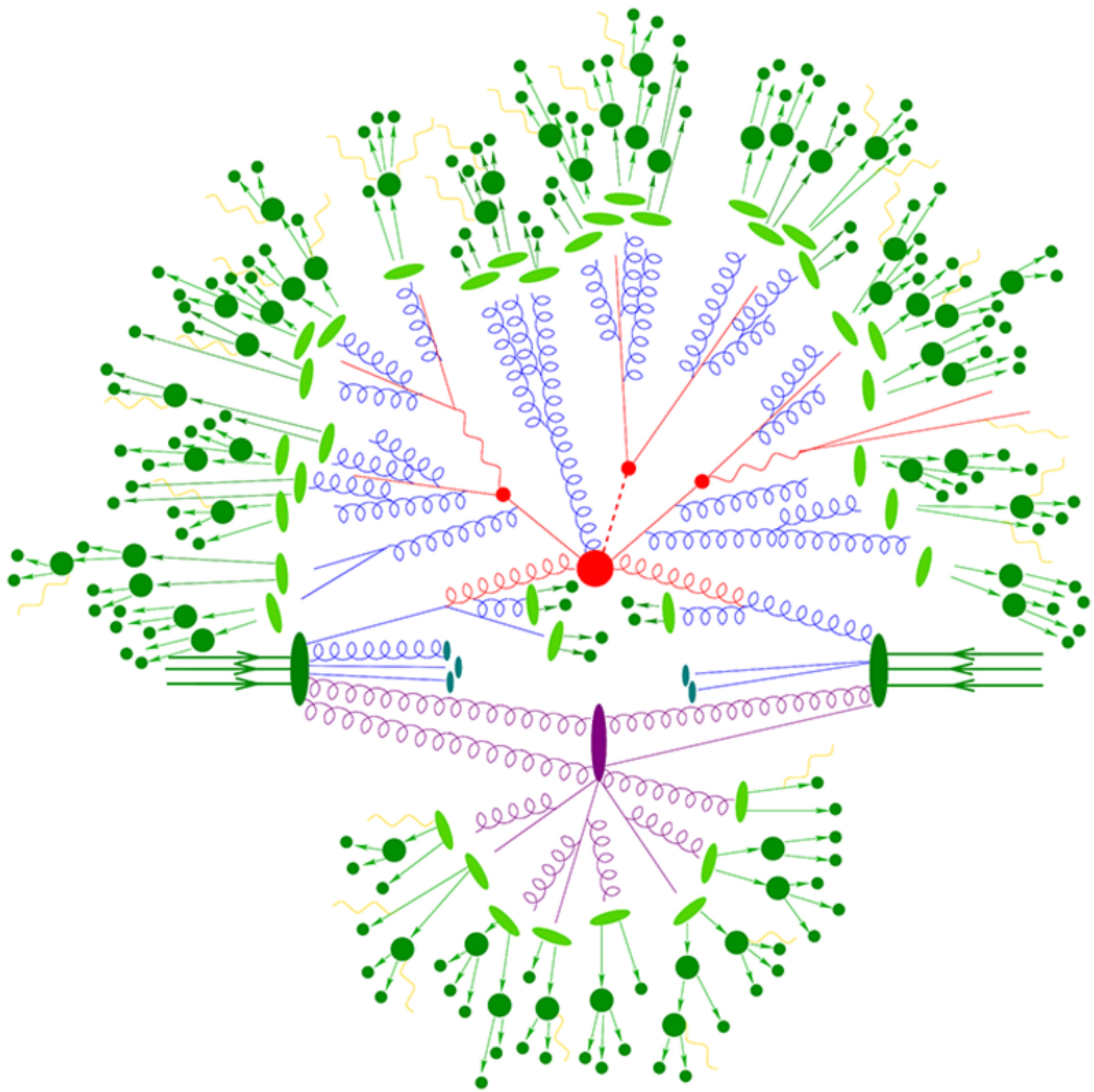
Resource:

- Main investment in GPU
  1 DGX A100, 1 custom node with 10 A100s, 3 various GPU nodes
- Storage 320TB (mostly for ML workloads)
- 2TB & 1.5TB RAM on the two A100 machines

QC usage:

- Qiskit and qulacs heavily used
- Qiskit and related libraries packaged into singularity containers and delivered to users over NFS
  - Spares installation troubles & improves research reproducibility
- GPU utilized extensively in pulse-level simulation of qudits using qutip and JAX

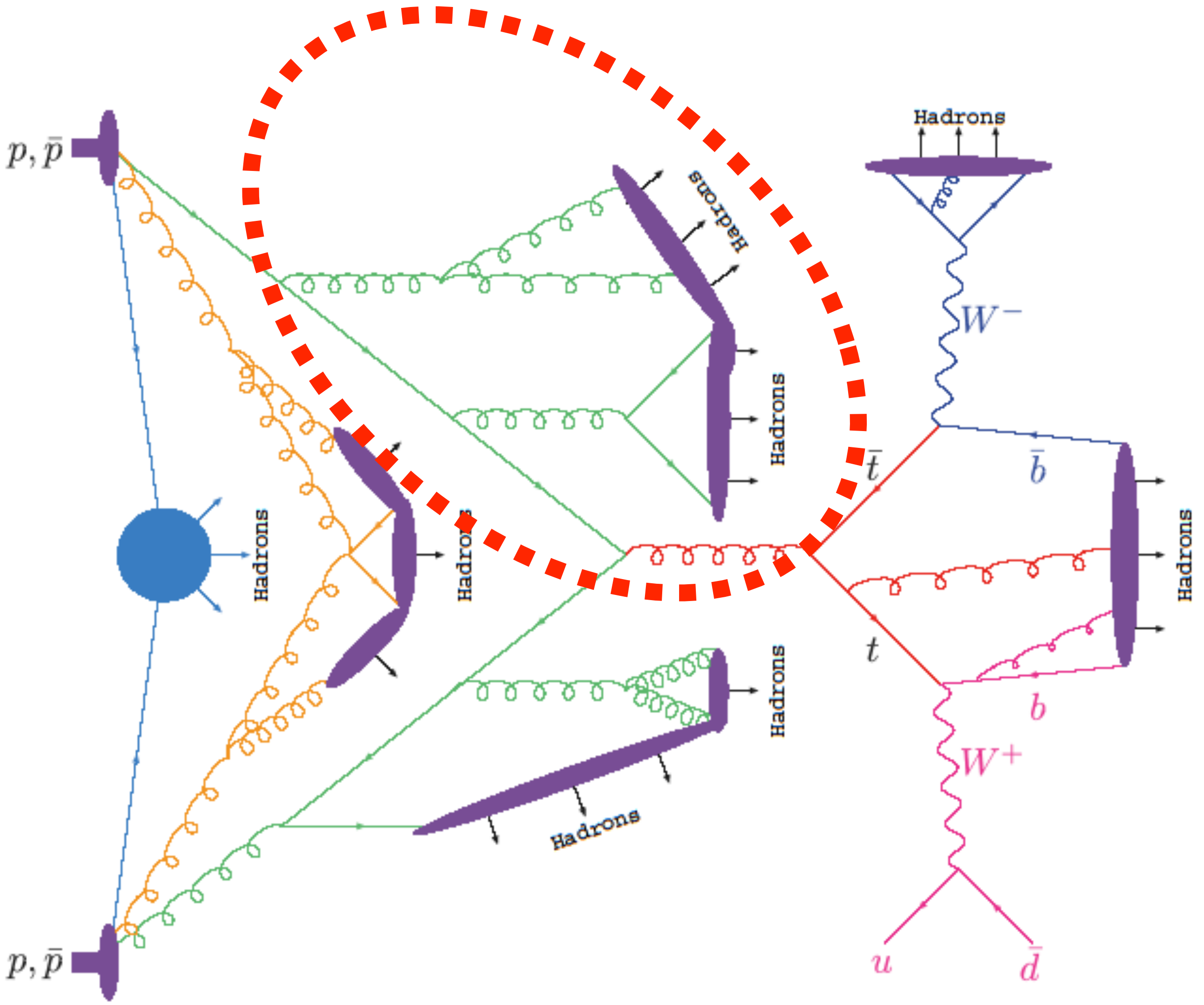
Example : Parton shower simulation



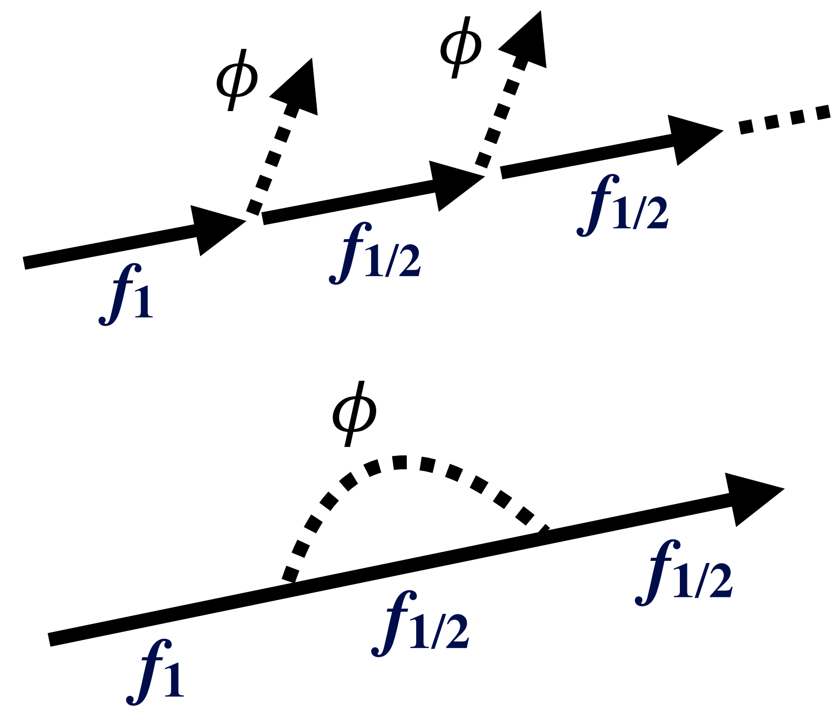
Parton shower simulation

Berkley group

1904.03196



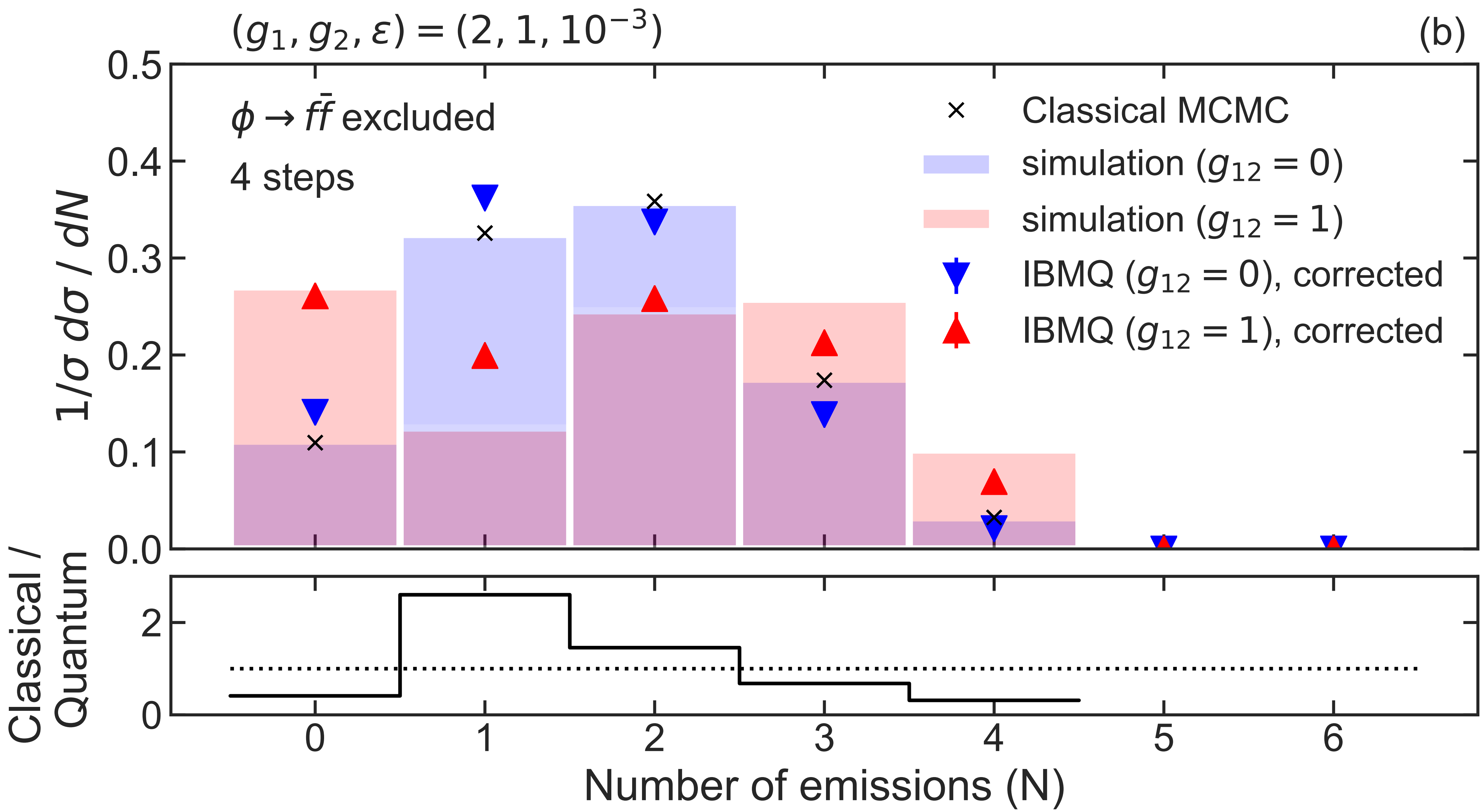
Simplify



- One boson species
- Two fermion flavours
  - Many combinations of flavours of intermediate fermions in a decay chain.

Parton shower simulation : Result

Berkley group



The effect of interference is observed
in difference between blue and red histograms.

R.Sawada

Quantum Dynamics Simulation

$$i\frac{d}{dt}\,\psi(t)\rangle = H\,\psi(t)\rangle$$
$$\Rightarrow\ \psi(t)\rangle = \lim_{\substack{N\to\infty \\ N\Delta t = t}} \prod_{k=1}^{N} e^{-iH(k\Delta t)\Delta t}\,\psi(0)\rangle$$

- Initial state of a many-body system: $\psi(0)\rangle$ → Initial QC state
- Time evolution in unit time: $e^{-iH(k\Delta t)\Delta t}$ → gate operation

$\psi(t)\rangle$ can be obtained approximately

Note: The same calculation can be done classically.

- Initial state → State vector
- Time evolution → Tensor calculation

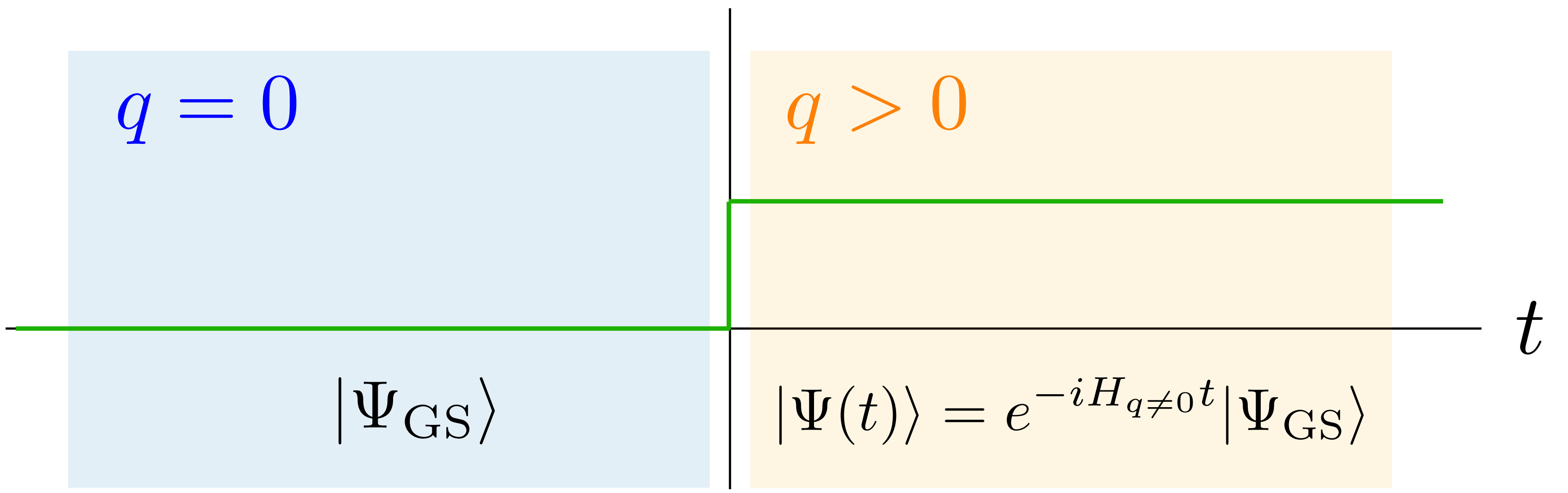
But the size of the vector and tensor can be too large to practically calcuate: 2^{50} complex (128byte) numbers → 2^{54} bytes for 50 qubit system

1dimention, QED Simulation

Variational method

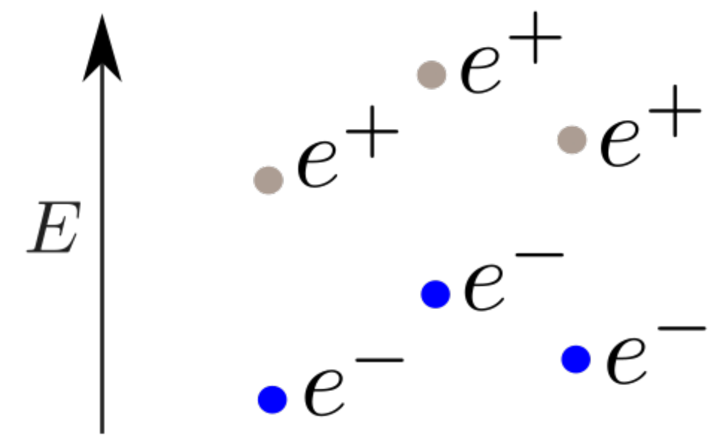
L. Nagano et al, arXiv:2302.10933

q : external electric field



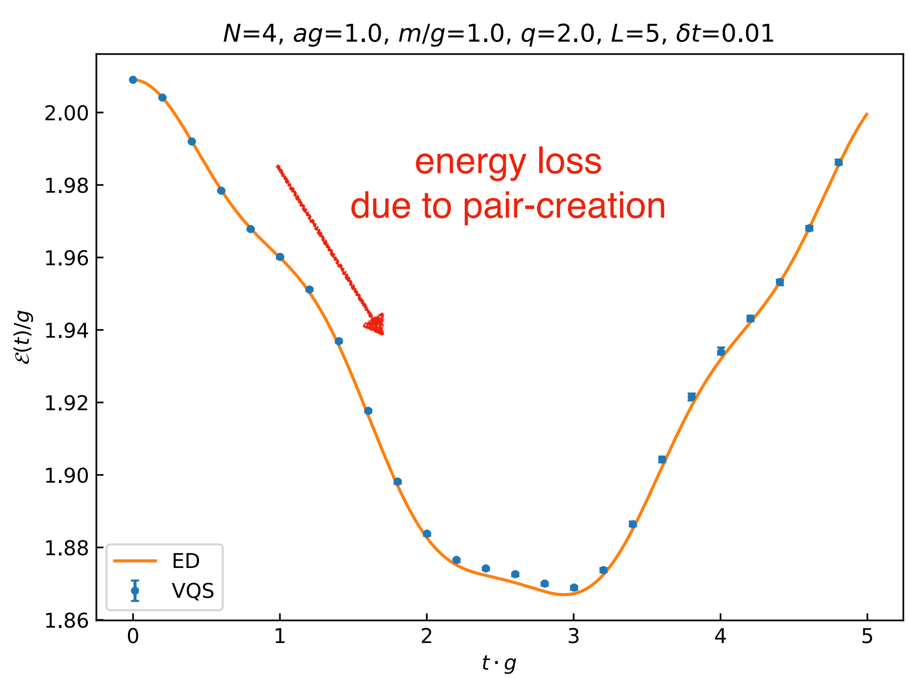
Ground state without external field

time evolution with external field
→pair-creation?

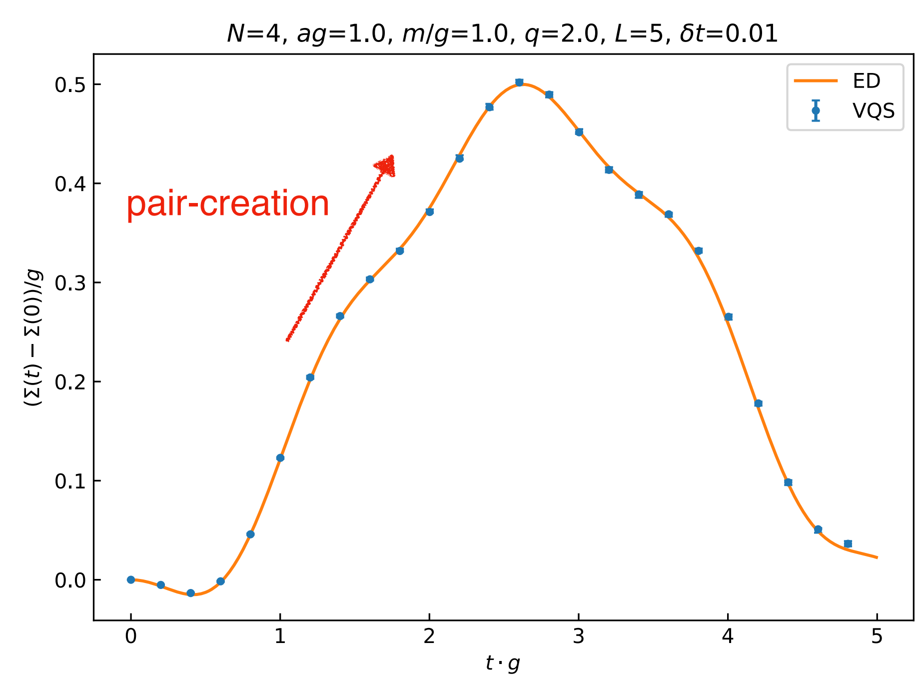


<u>Schwinger effect</u>

electric field

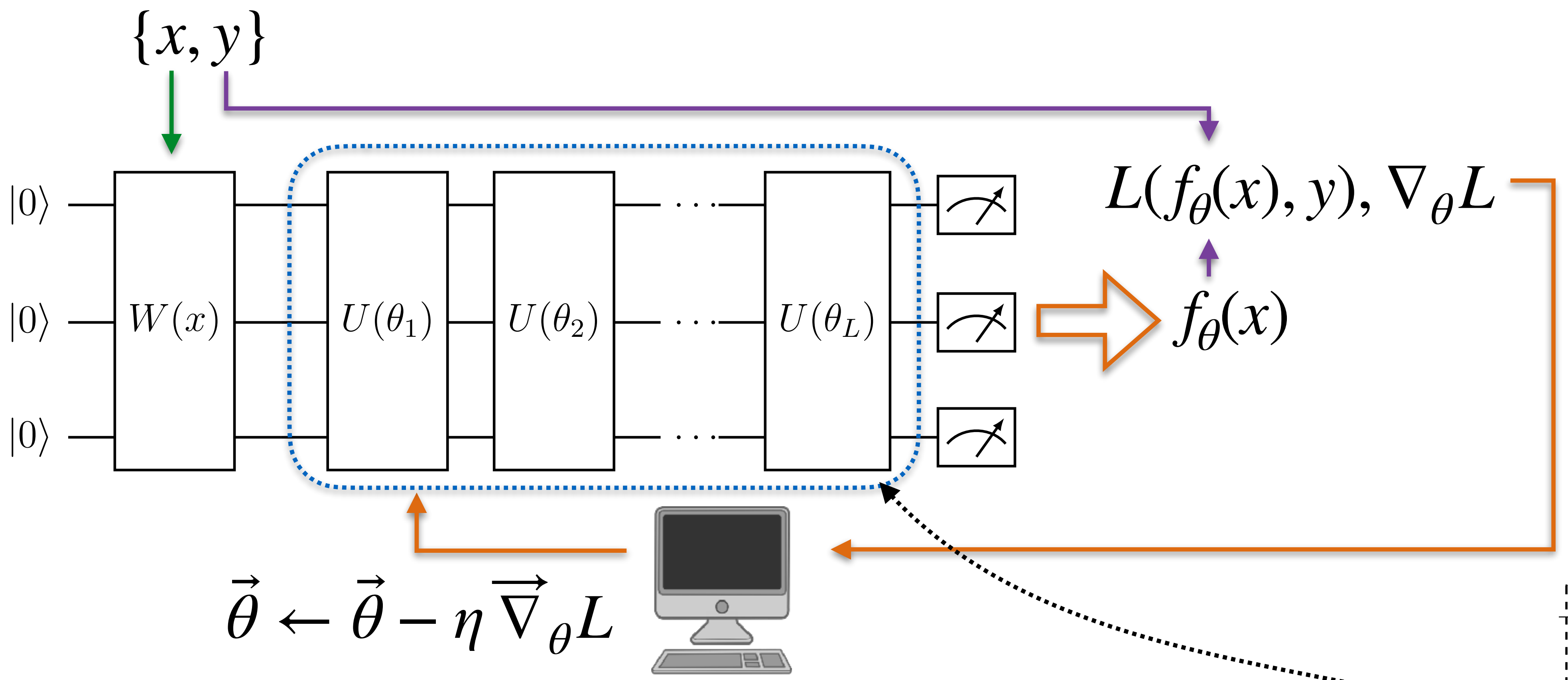


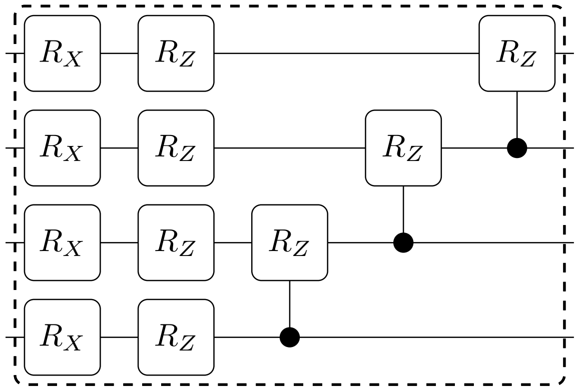
chiral condensation



19

Quantum neural network





Example U

- Data are encoded using a circuit (W)
- Network is formed by a quantum circuit
    - It consists of several “Parametric” gates (U)

      → Such gates contain classical parameters, e.g. rotation angle, that are optimised by training
- Loss function is defined by the circuit output (f) and the label of data.
- Loss function and parameter update are done classically.

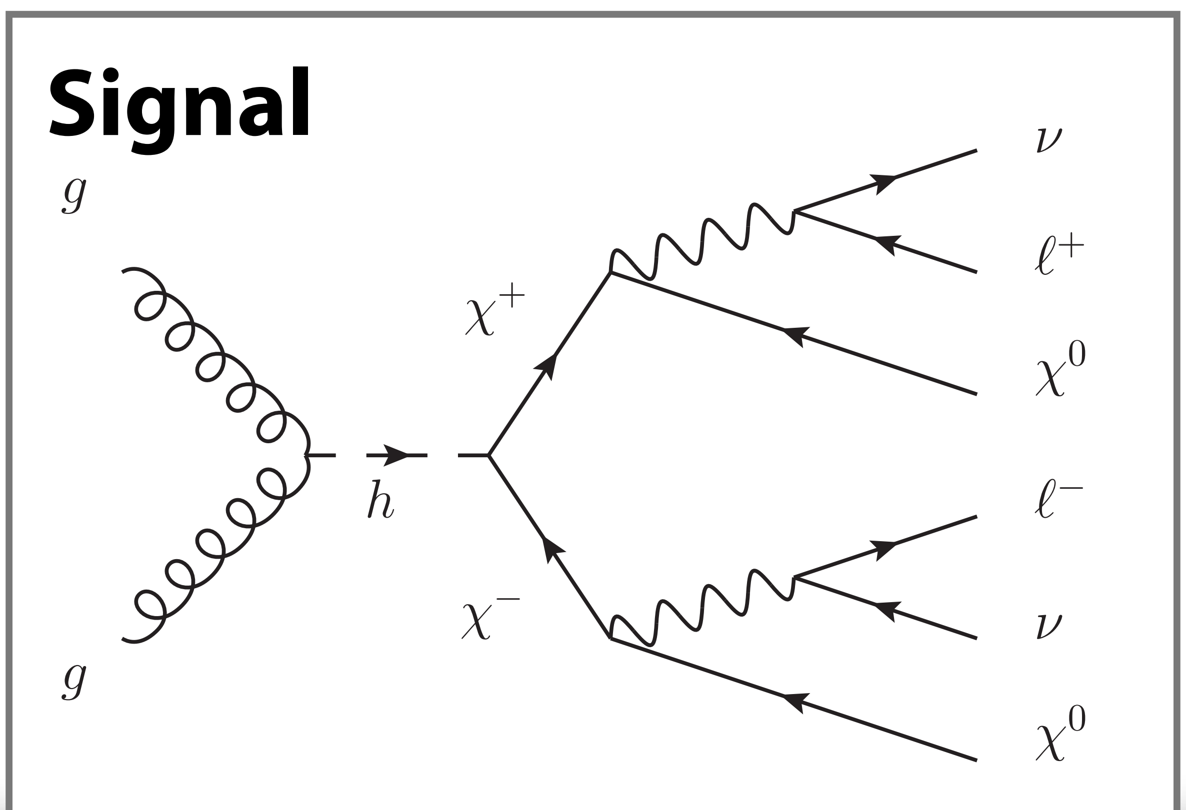
→ Quantum circuit can be simple and short = Suitable to the near-future quantum devices.

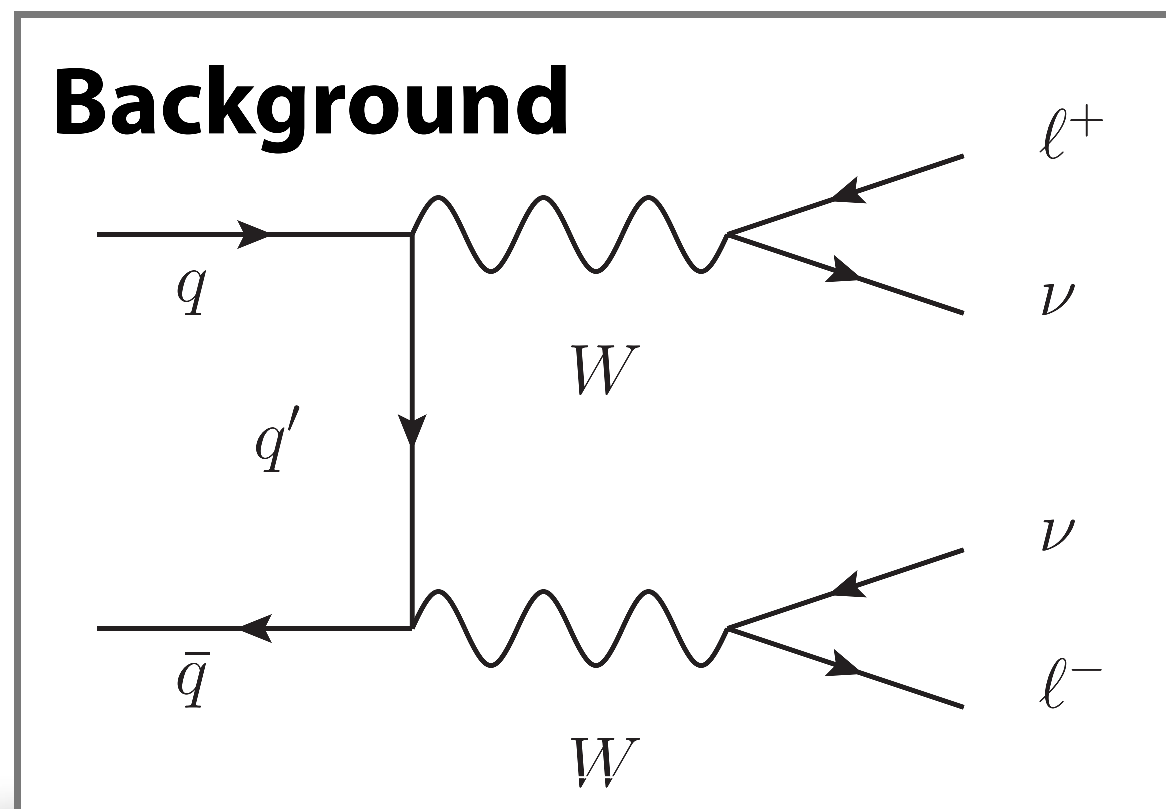
Original Article | Open Access | Published: 03 January 2021

Event Classification with Quantum Machine Learning in High-Energy Physics

Koji Terashi, Michiru Kaneda, Tomoe Kishimoto, Masahiko Saito, Ryu Sawada & Junichi Tanaka

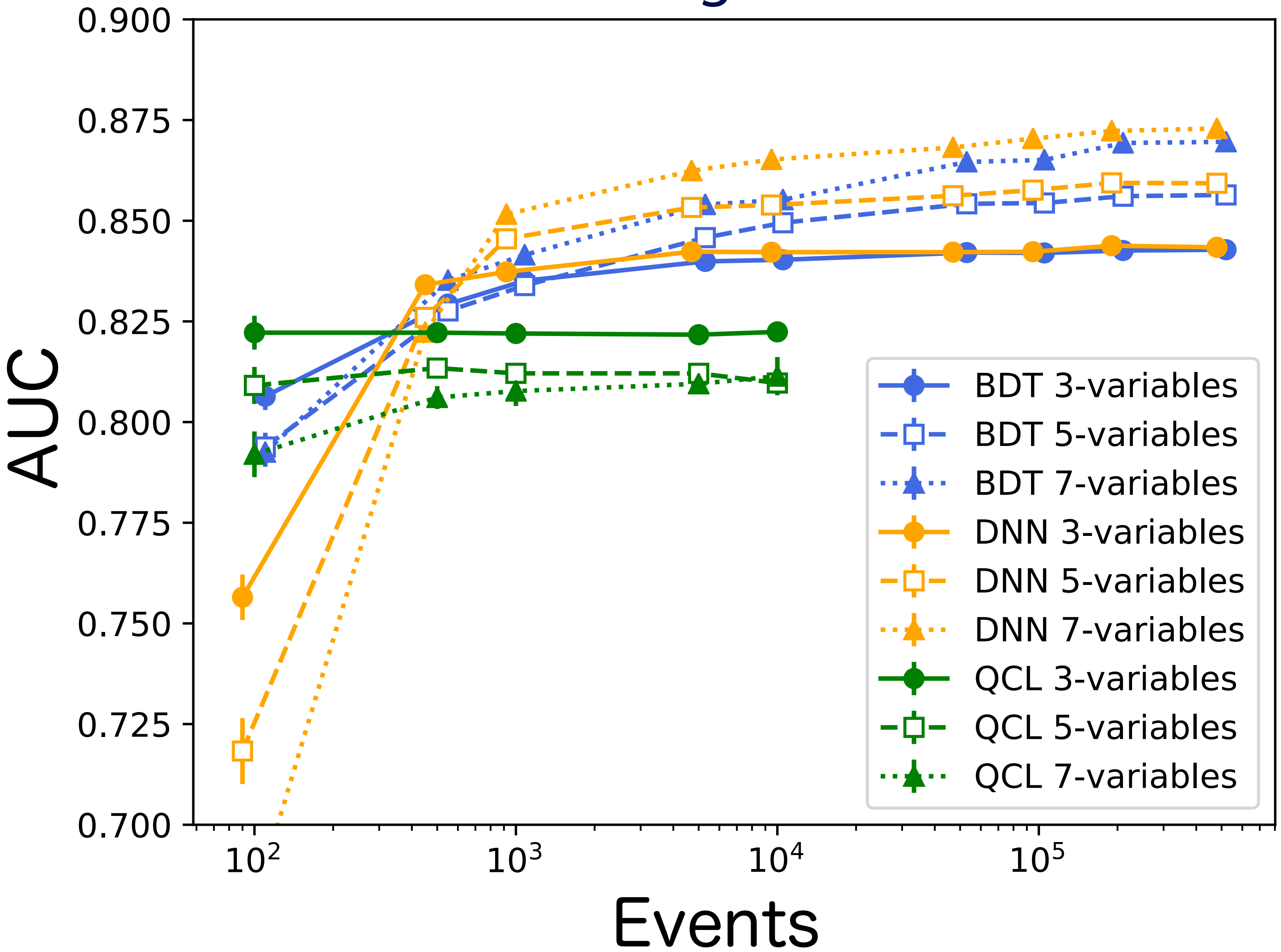
Computing and Software for Big Science **5**, Article number: 2 (2021) | Cite this article





SUSY Classification

QCL = Quantum algorithm with simulator



Compared with BDT and DNN :

- BDT and DNN models optimized at each training set to avoid over-training
- Classical algorithms out-perform at large training set

➡ Performance of quantum algorithm comparable to BDT/DNN at small training set with small # of variables

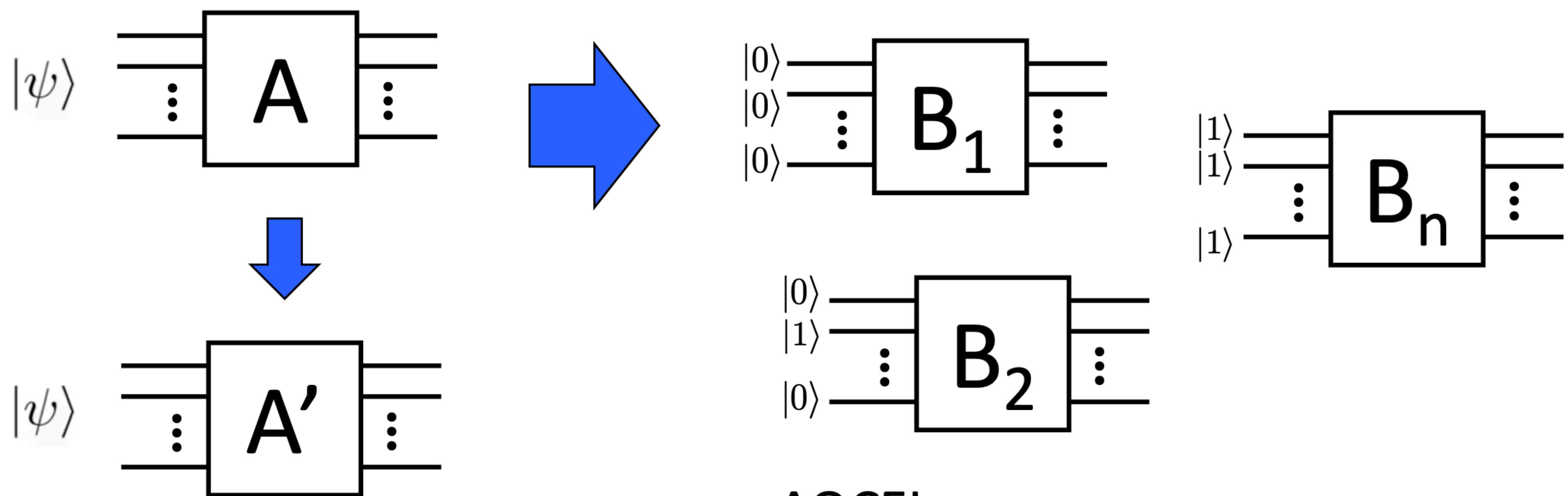
Quantum circuit optimization : AQCEL

W. Jang et al.
Quantum 6, 798 (2022).
Github: UTokyo-ICEPP/aqcel

In physics simulation, many events are generated using a single program with a fixed initial state.

Circuits can be shorter, namely the number of gate operations can be less, by optimising it depending on the initial state.

qc A for any initial states



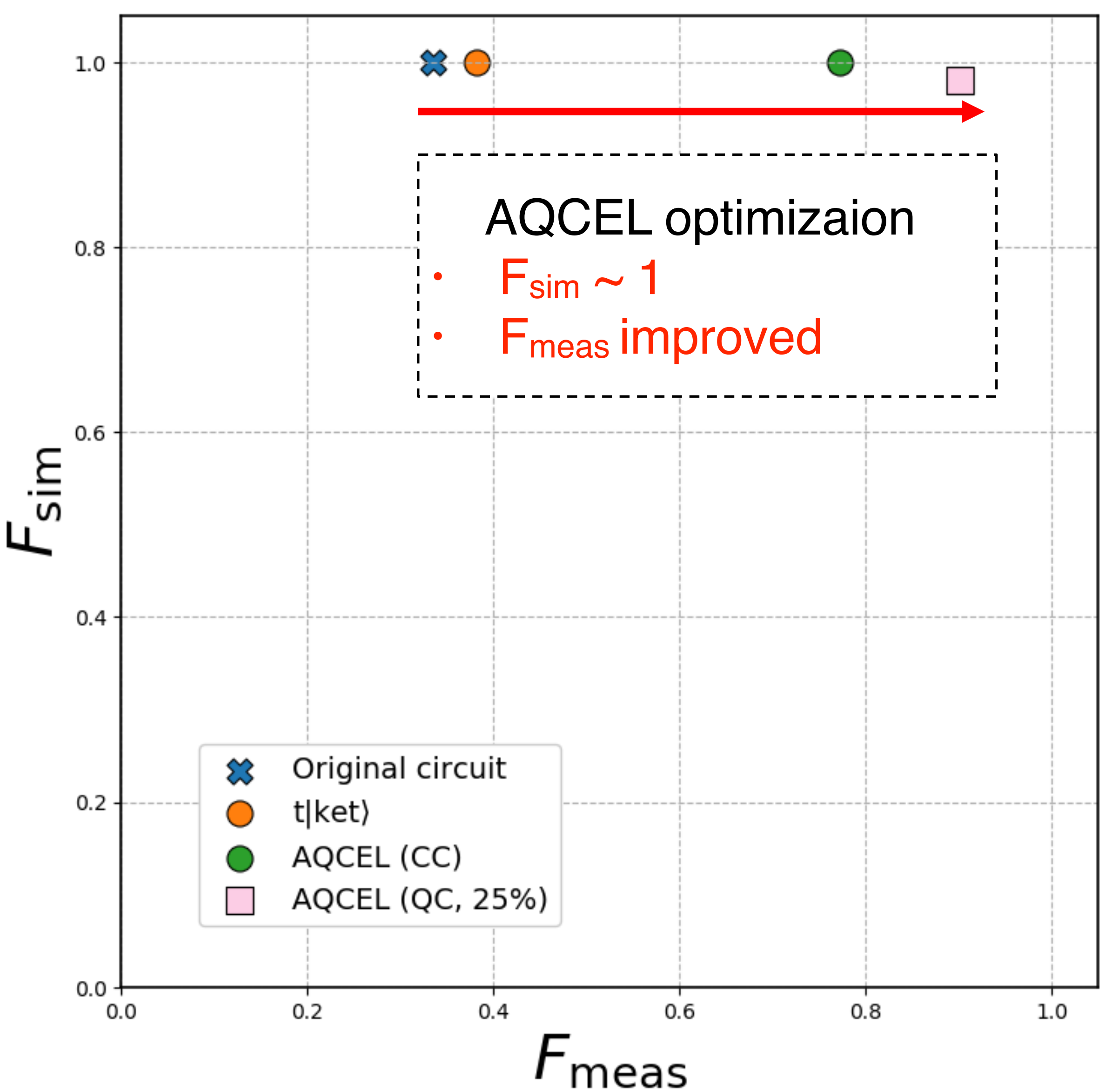
General qc optimizer
(qiskit, tket, etc...)
<u>Preserve circuit equivalence</u>

<u>AQCEL (Advancing Quantum Circuit by ICEPP and LBNL)</u>
Optimize qc depending on initial states.
Circuit equivalence is not always preserved.
Strong reduction can be applied.

AQCEL result for parton shower simulation

Number of gates

	Original	tket	AQCEL(CC)	AQCEL(QC,25%)
CNOT	527	616 (117%)	178 (34%)	64 (12%)
U1, U2, U3	362	331 (91%)	102 (28%)	24 (6.7%)
Total	889	947 (107%)	280 (31%)	88 (9.9%)

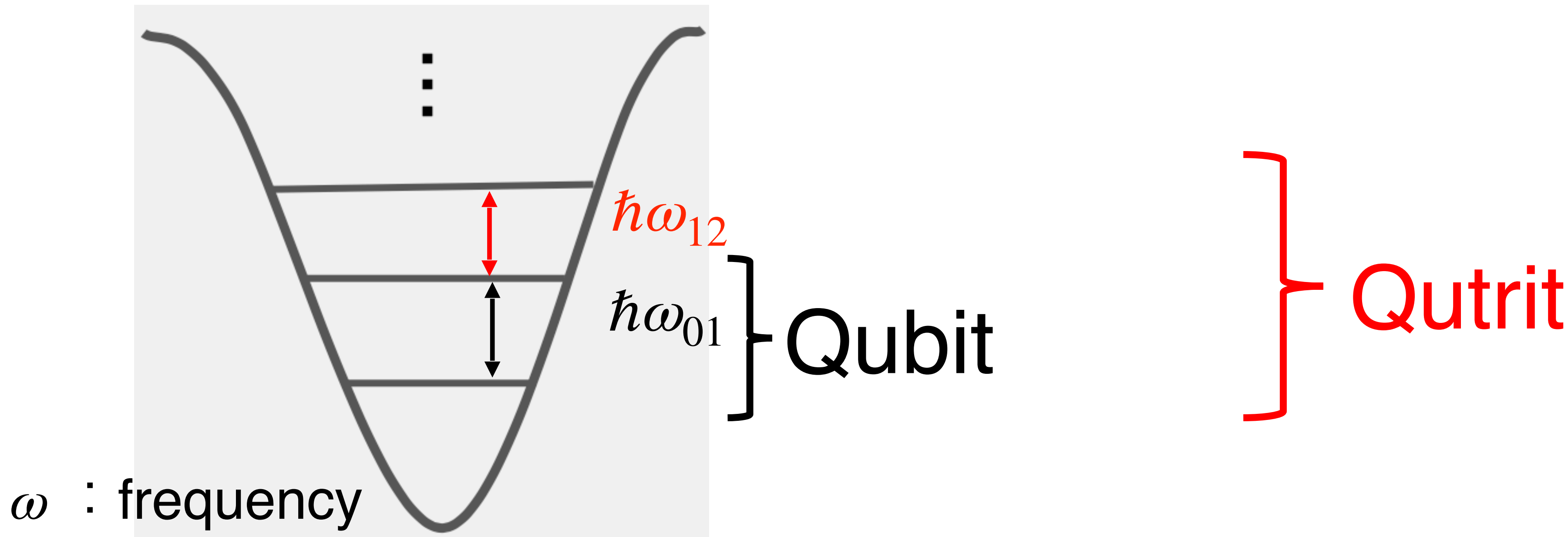


F_{sim} (Fidelity in case of no QC noise) is not decreased by due to the approximation in AQCEL.

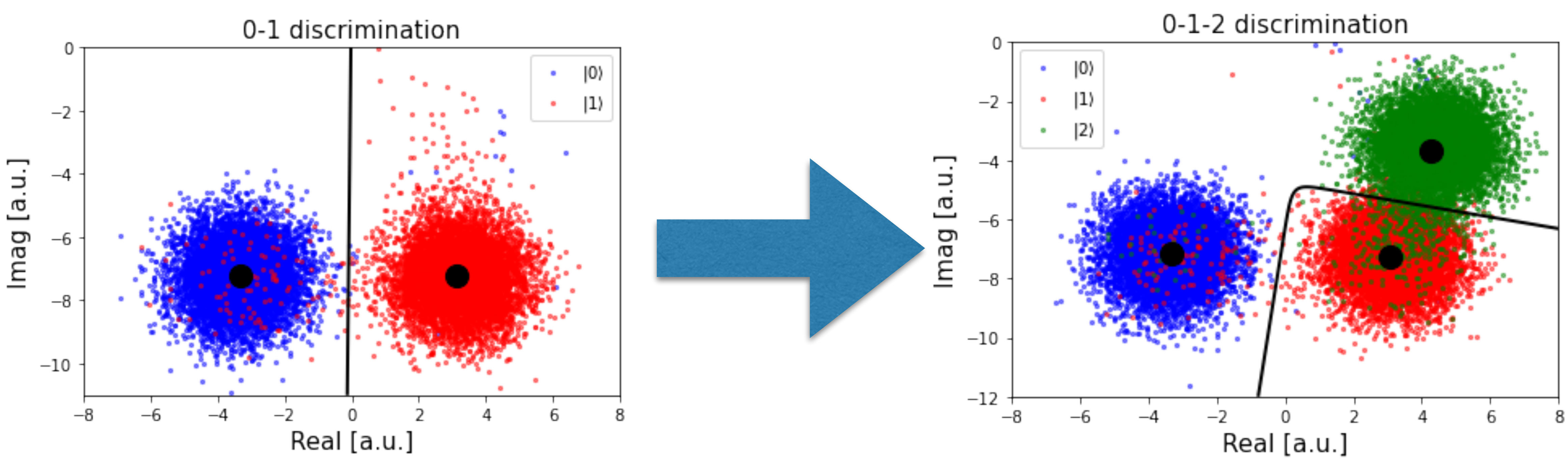
F_{meas} is much improved, ~0.4 → 0.9, by the fewer operations, optimised by AQCEL.

Qutrit : Quantum trit

Energy level of transmon



By programming a customised pule, we can use
$1\rangle \rightarrow \;\; 2\rangle$ や $\;\; 2\rangle \rightarrow \;\; 1\rangle$ transition in IBM Quantum devices.

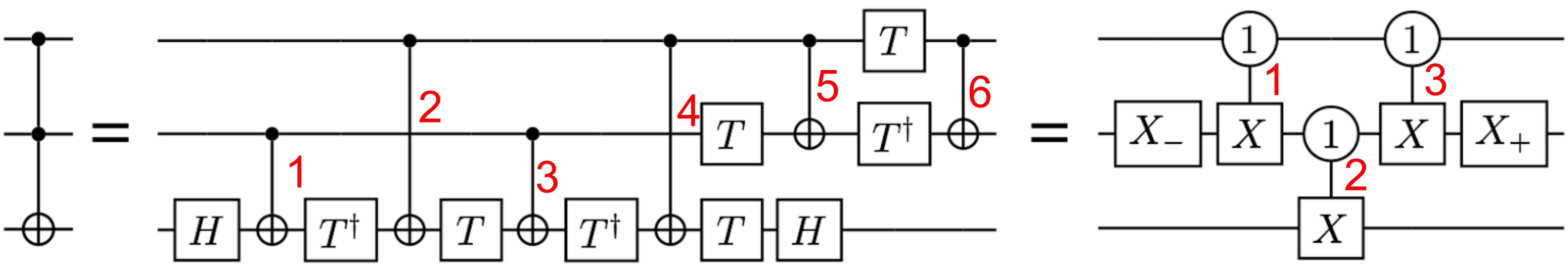


R.Sawada

Multi controlled bit (Toffoli) gate using Qutrit

Toffoli with Qubits

Toffili with Qutrit



※Qutrit is used intermediately

$|0\rangle, |1\rangle$: Qubit

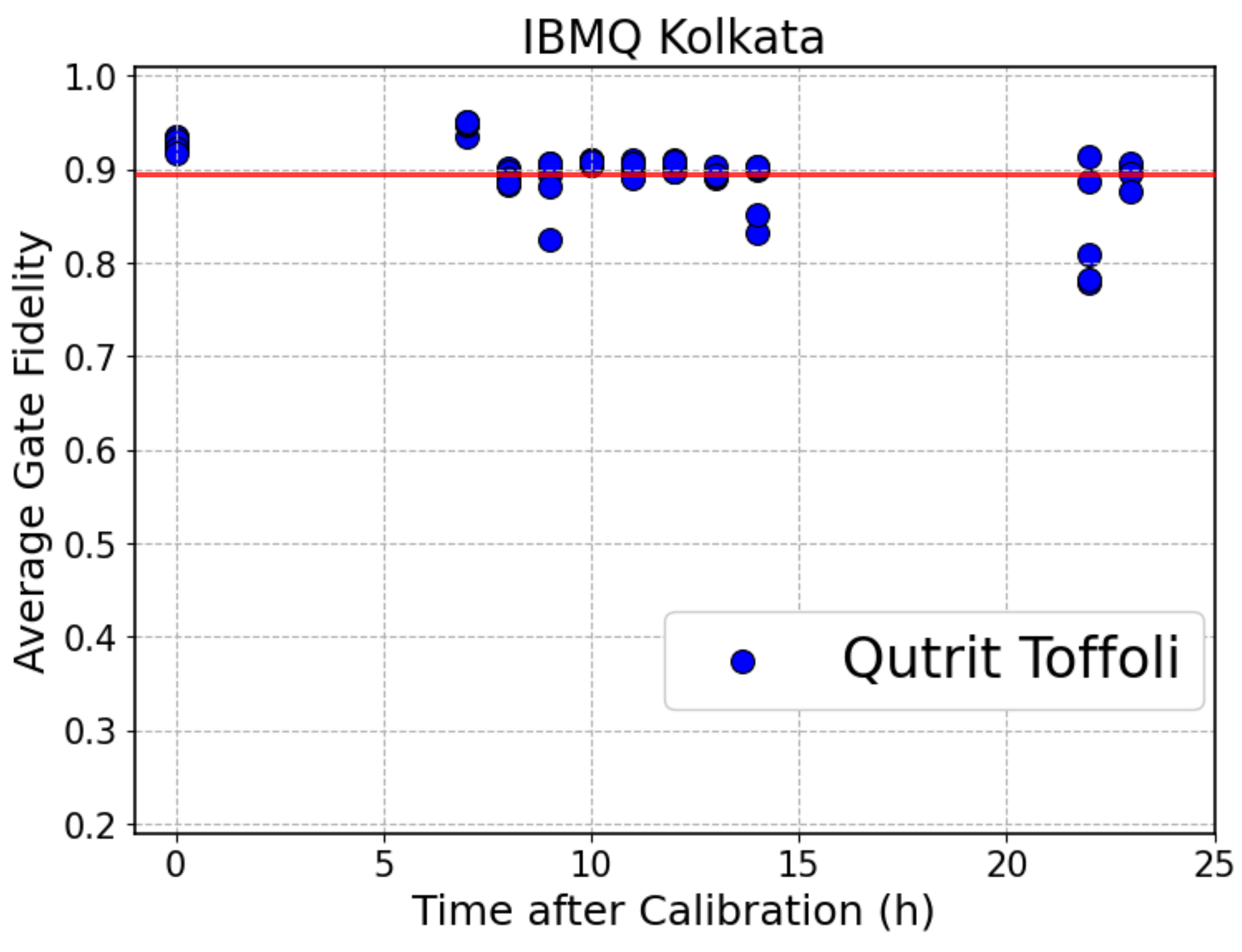
$|0\rangle, |1\rangle, |2\rangle$: Qutrit

Fewer Controlled not gates

Qutrit : gate fidelity

Gate fidelity of Qutrit toffoli gate is measured on ibmq_kolkata

Qutrit Toffoli, averae fidelity over 1 day



- Gate time
  - Qutrit Toffoli : 2.5 μs
  - Qubit Toffili : 3.1 μs
- Fidelity after calibration
  - 0.928±0.007 (1 hour later)
  - 0.896±0.036 (1day later)

Qutrit Toffili fidelity is 5—7% higher than Qubit Toffoli.

Ising Machines : Annealer

	Vender	Product	Number of bits
Quantum	D-Wave Systems	D-Wave Advantage	5760
Quantum-inspired (classical)	Hitachi	CMOS Annealing	147k (ASIC), 256k (GPU)
	Fujitsu	Digital Annealer	8192
	Toshiba	SQBM+	10M
	Fixstars Amplify	Fixstars Amplify AE	131k (Full connect), < 4.3B (Partial connect)
	NTT	Coherent Ising machine	100k

Tracking using Annealing

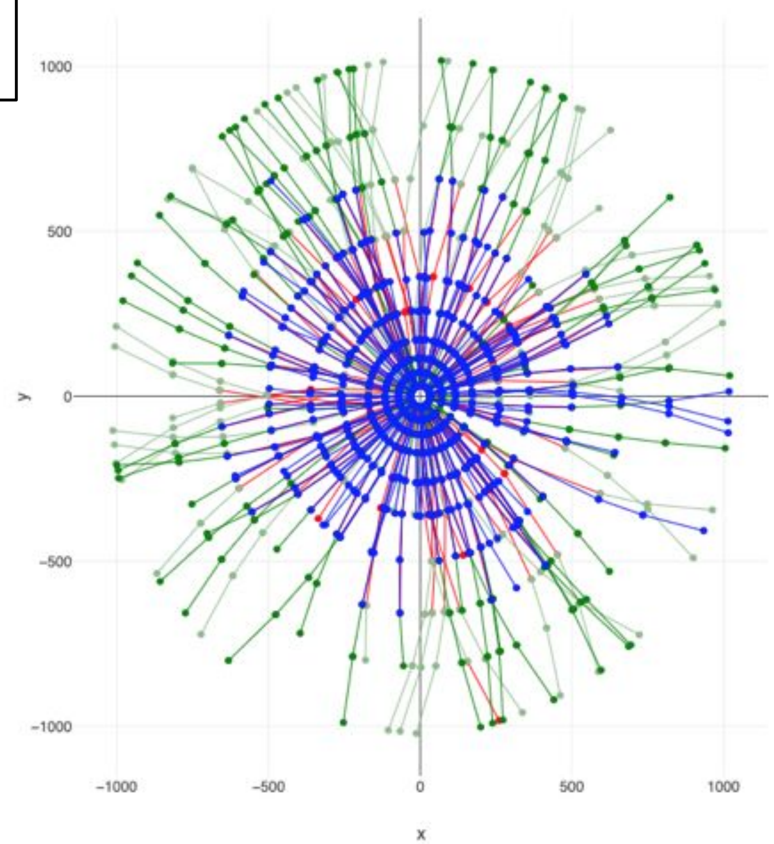
Results

Reconstructed high pT tracks
Reconstructed low pT tracks
Not reconstructed tracks
Fake tracks

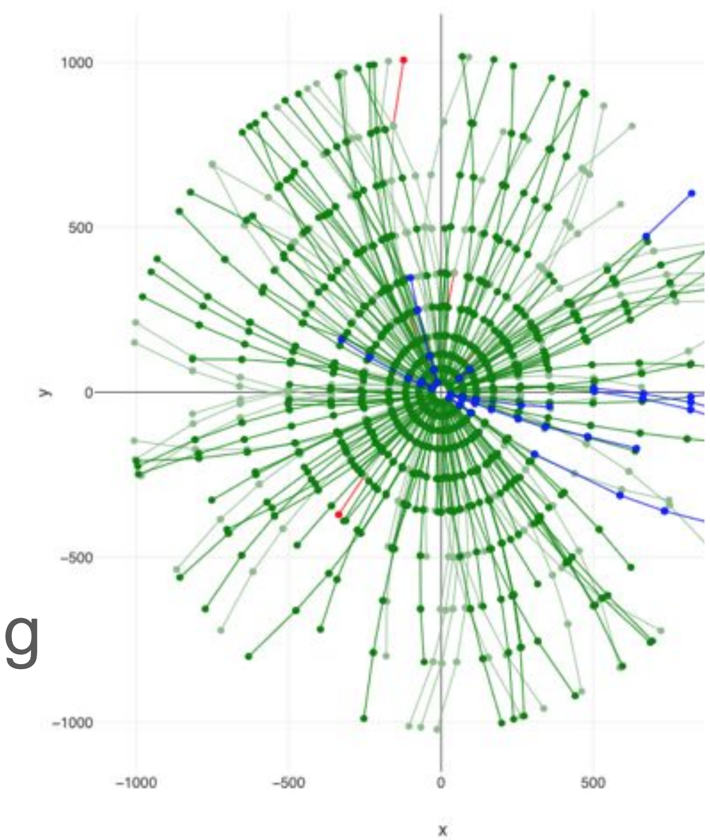
1600 particles (20% of HL-LHC)
- 11000 hits

Input

Doublet selection



Annealing



390000 Doublets	
Purity	0.22 %
Efficiency	99.5 %

2445 Doublets

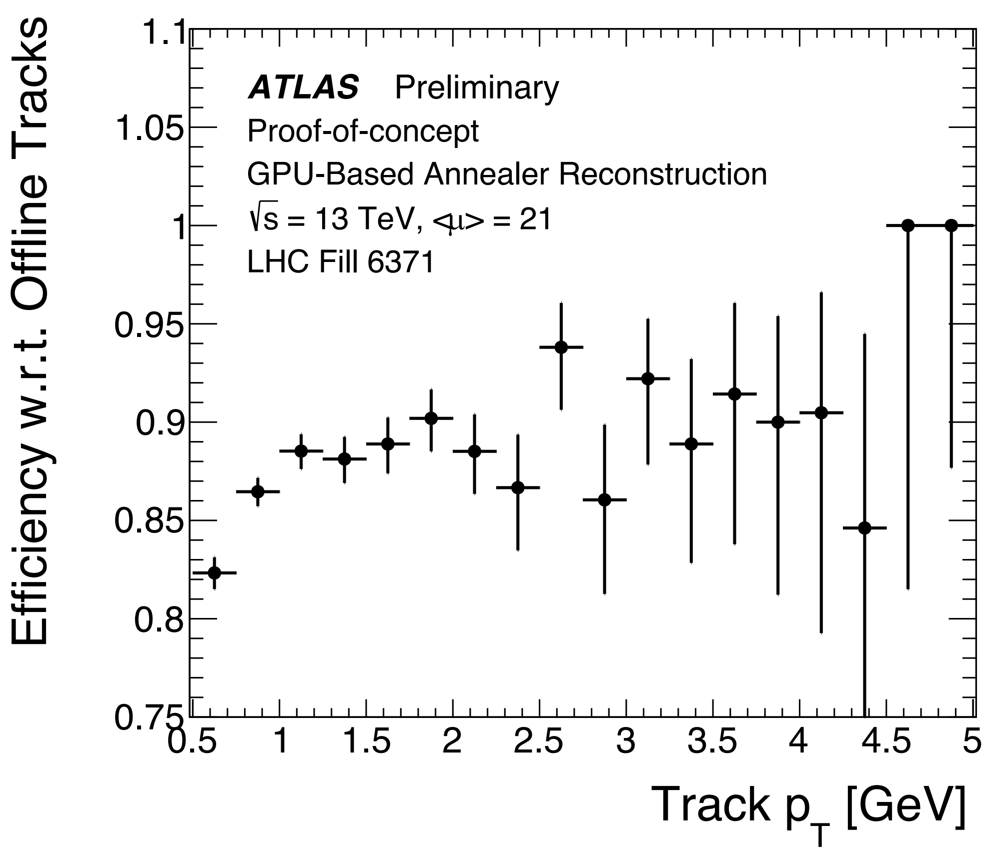
1424 Doublets	
Purity	98.5 %
Efficiency	96.4 %

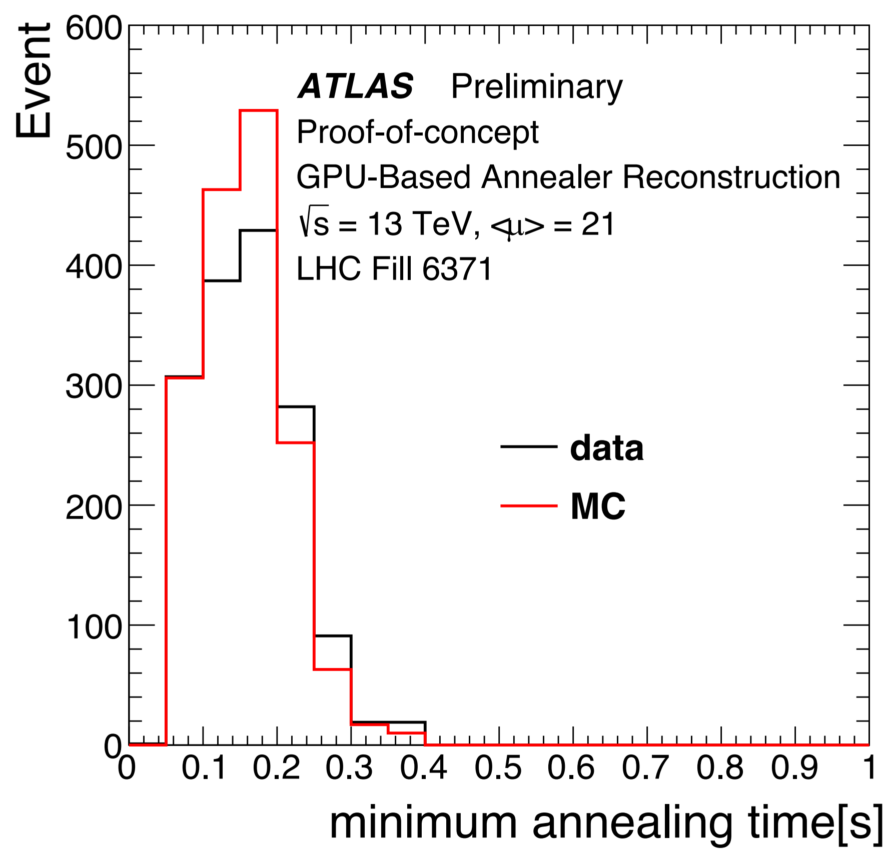
Developed by Berkley group

Annealing tracking with ATLAS Data

Waseda university group

- minBias trigger
- Relative efficiency to offline tracking
- ~90% efficiency for pT > 1 GeV
- Annealing time is similar with read data and MC simulation





Average pre-processing time for data is ~0.6 sec. (single core, 11th Gen Intel(R) Core(TM) i9-11900K @ 3.50GHz)

Hardware development

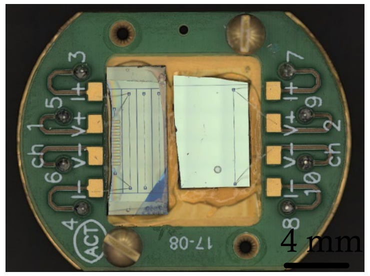
S. Chen

Qubit as a sensor : Direct detection of light dark matter (Axion, darkphoton)

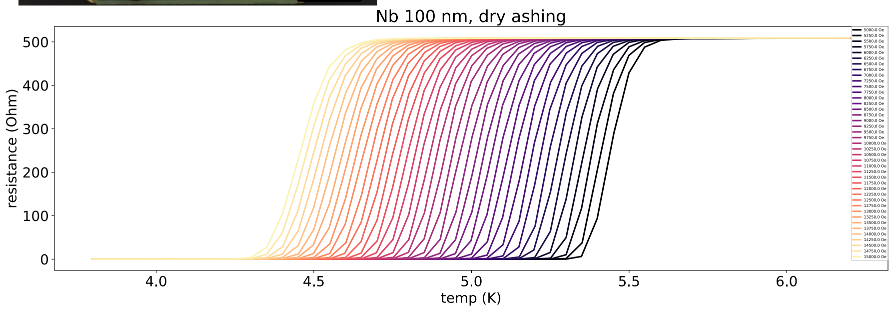
✔ Full-stack development capability established in the first year

○ Pursuing higher quality in the fabrication & new types of quantum devices

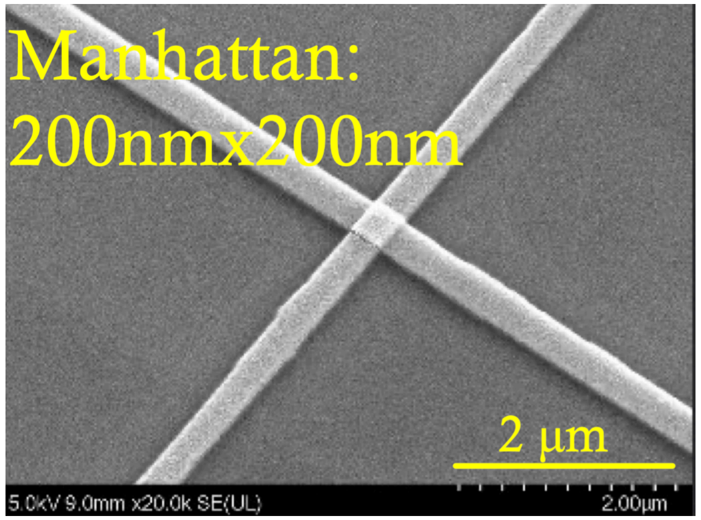
Thin film characterization



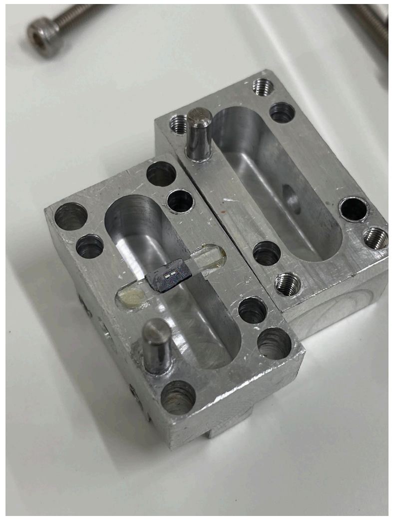
PPMS / XRD for
Al, Nb, TiN, Ta films etc.

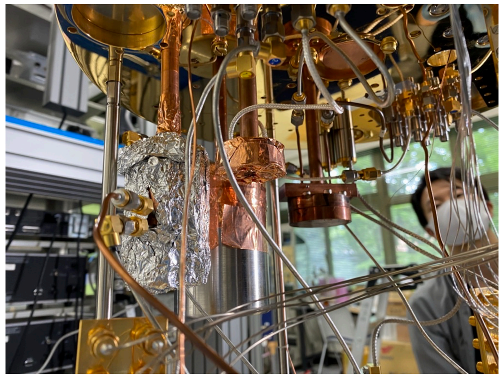


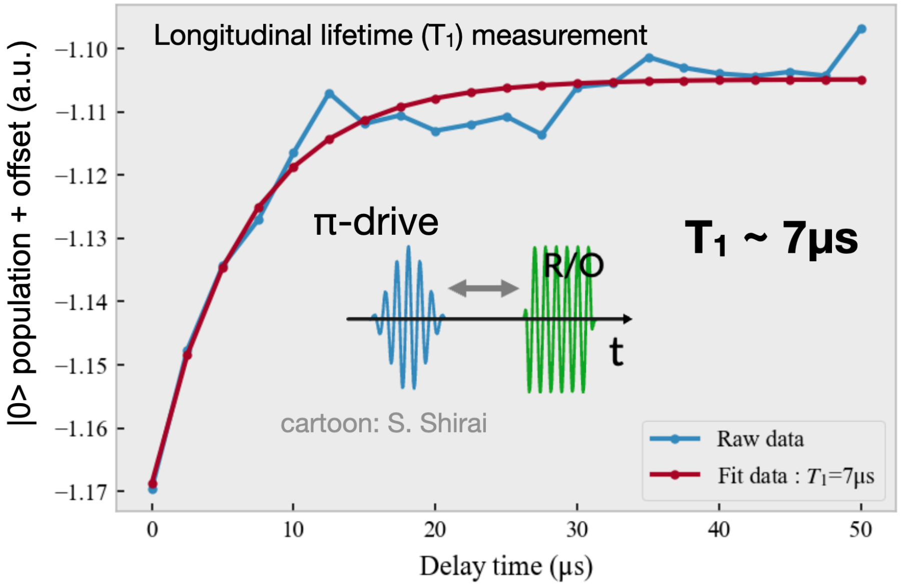
Qubit fabrication



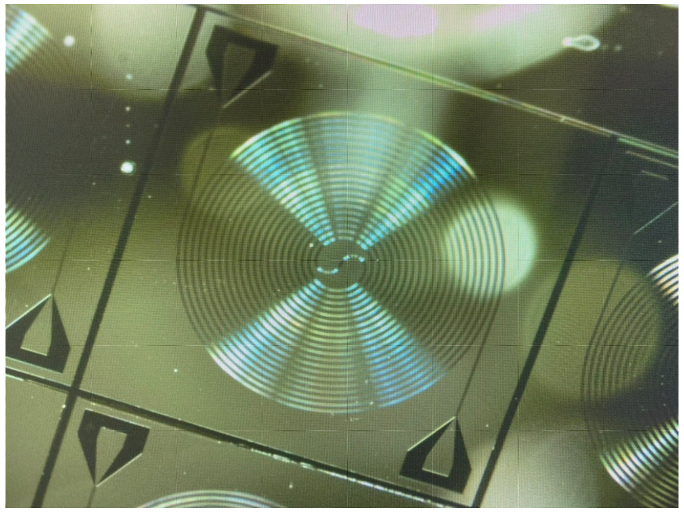
Packaging & Measurement







R&D of new exotic quantum devices



R.Sawada

Conclusion

- Not only traditional computing services, ICEPP is carrying out researches on new technologies
  - HPC, Cloud (not in this talk), ML (not including this talk), QC.
- HPCs in Japan may potentially increase available resource for HEP, but more researches are needed to use them with full computing power.
- QC may outperform classical computer in future. We are try to use QC for different types of problems and seeing how well it works.

Backup

Quantum circuit optimization : AQCEL

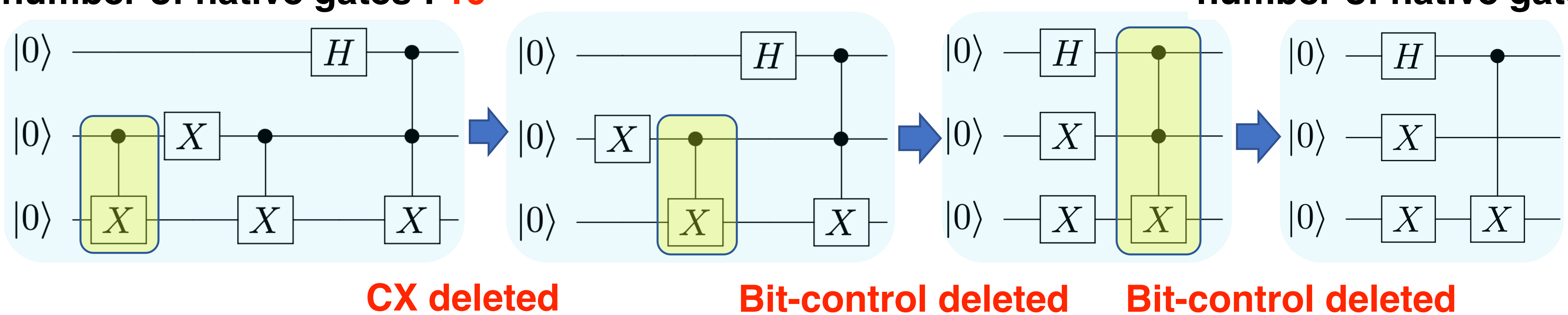
W. Jang et al.
Quantum 6, 798 (2022).
Github: UTokyo-ICEPP/aqcel

In physics simulation, many events are generated using a single program with a fixed initial state.

Circuits can be shorter, namely the number of gate operations can be less, by optimising it depending on the initial state.

Example:

number of native gates : 19 **number of native gates : 4**



	1st CX	2nd CX	CCX
Quantum state	$\|000>$	$\|010>$	$\frac{1}{\sqrt{2}}\|011> + \frac{1}{\sqrt{2}}\|111>$
Control bit states	'0'	'1'	'01', '11'
Deletion	CX	Bit-control	Bit-control