Fugaku Update and its Future Perspectives

Satoshi Matsuoka, Director Riken R-CCS
Hyperion HPC Presentation
26 Aug 2021

What is a ‘Exascale’ Supercomputer?

1. FP64 Performance > 1 Exaflop (EF)

1.1. Achieve Rpeak (FP64) > 1 EF

1.2. Achieve Top500 – Linpack Rmax > 1 EF

- *Fugaku Rmax = 0.442 EF, Rpea = 0.537 EF-> NG*
- *However, very little correlation to real apps, symbolic*

2. Any floating point precision performance > 1 Exaflop

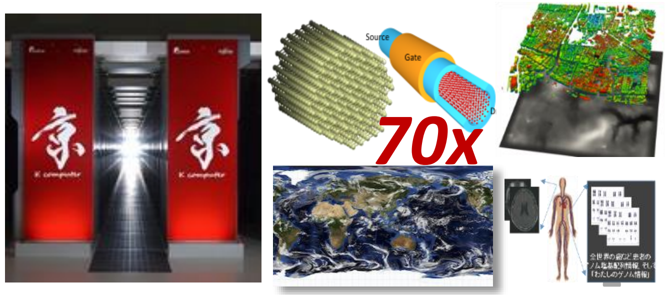
1.1. Peak FP performance > 1 EF

1.2. Measured performance in credible app or benchmark

- *Fugaku FP32, FP16 Peak, HPL-AI (2EF) > 1 Exaflop -> OK!*
- *However, ORNL Summit: FP16 Peak ~= 3 EF, GB2018 App ~= 2EF*

3. Real apps ~= 50~100x 2011~12 10~20PF SCs

- **Fugaku ~70x c.f. K (11PF Rmax) on 9 target apps**
- **“Applications First” -> The most important metric**



Fugaku: Largest & Fastest Supercomputer Ever

‘Applications First’ R&D Challenge--- High Risk “Moonshot” R&D

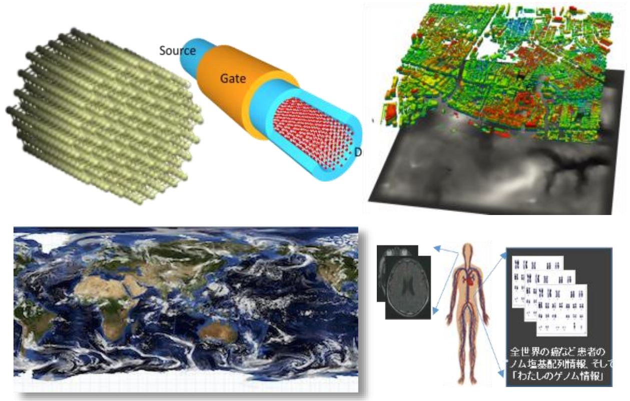
- **A new high performance & low power Arm A64FX CPU co-developed by Riken R-CCS & Fujitsu along with nationwide HPC researchers as a National Flagship 2020 project**

- 3x perf c.f. top CPU in HPC apps
- 3x power efficiency c.f. top CPU
- General purpose Arm CPU, runs same program as Smartphones
- Acceleration features for AI

“Moonshot” R&D Target

- **Fugaku x 2~3 = Entire annual IT in Japan**

	Smartphones		Servers (incl. IDC)		Fugaku		K Computer
Untis	20 million ~annual shipment in Japan	=	300,000 (~annual shipment in Japan	=	1 (160K nodes)		Max 120
Power (W)	10W×2,000万台＝ 200MW	=	600-700W×30万台＝ 200MW (incl cooling)	> >	30MW (very low)		15MW (less than 1/10 efficiency c.f. Fugaku)



- **Developed via extensive co-design**

”Science of Computing"

By Riken & Fujitsu & HPCI Centers, etc., Arm Ecosystem, Reflecting numerous research results

”Science by Computing"

“9 Priority Areas” to develop target applications to tackle important societal problems

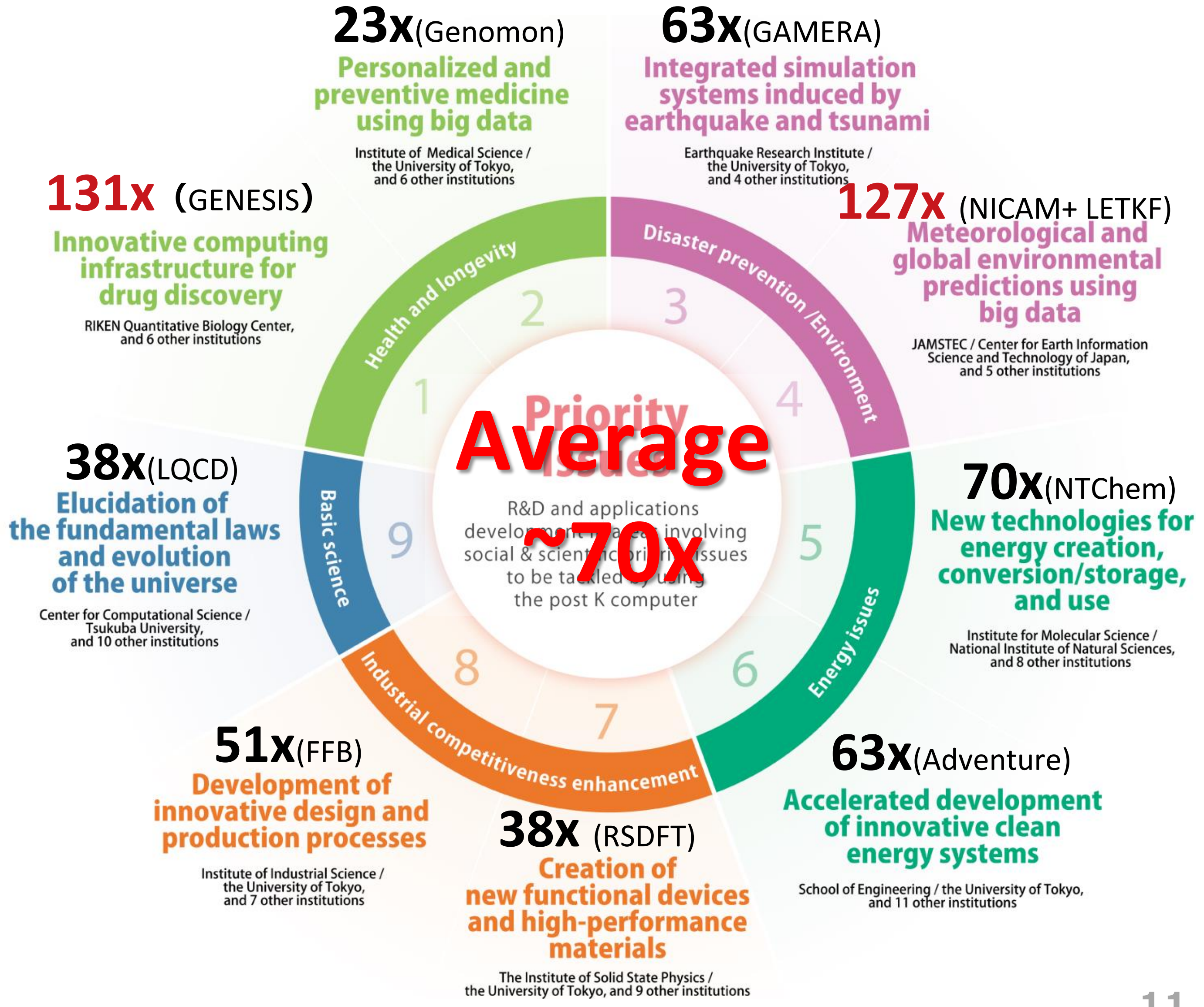
"Applications First" Exascale R&D

Fugaku Target Applications – Priority Research Areas

- **Advanced Applications Co-Design Program to Parallel Fugaku R&D**
- **Select one representative app from 9 priority areas**
  - Health & Medicine
  - Environment & Disaster
  - Energy
  - Materials & Manufacturing
  - Basic Sciences

SDGs Goals

- **Up to 100x speedup c.f. K-Computer => achieved!**

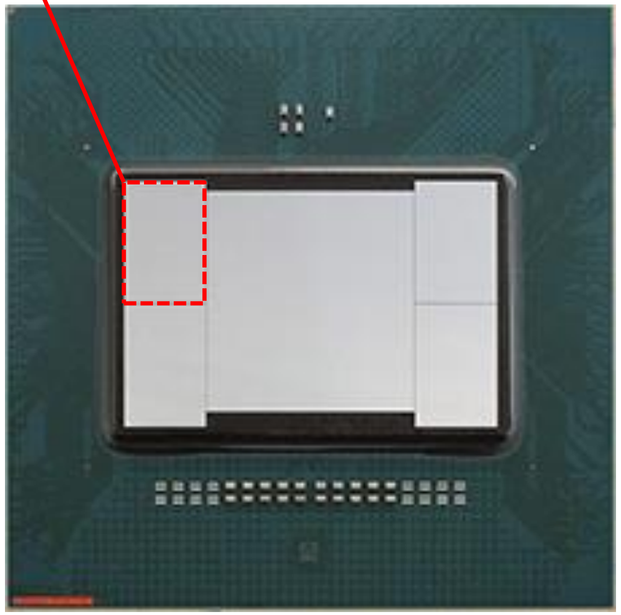


A64FX CPU for supercomputers

FUJITSU

- All-in-one 7nm SoC w/ low power consumption
  - Armv8.2-A, 512-bit SVE (Scalable Vector Extension)
  - Four HBM2, 32 GiB per package
  - Tofu Interconnect D integrated
  - HW inter-core barrier & sector cache
  - 48 compute cores & 4 assistant cores for OS daemon & MPI offload

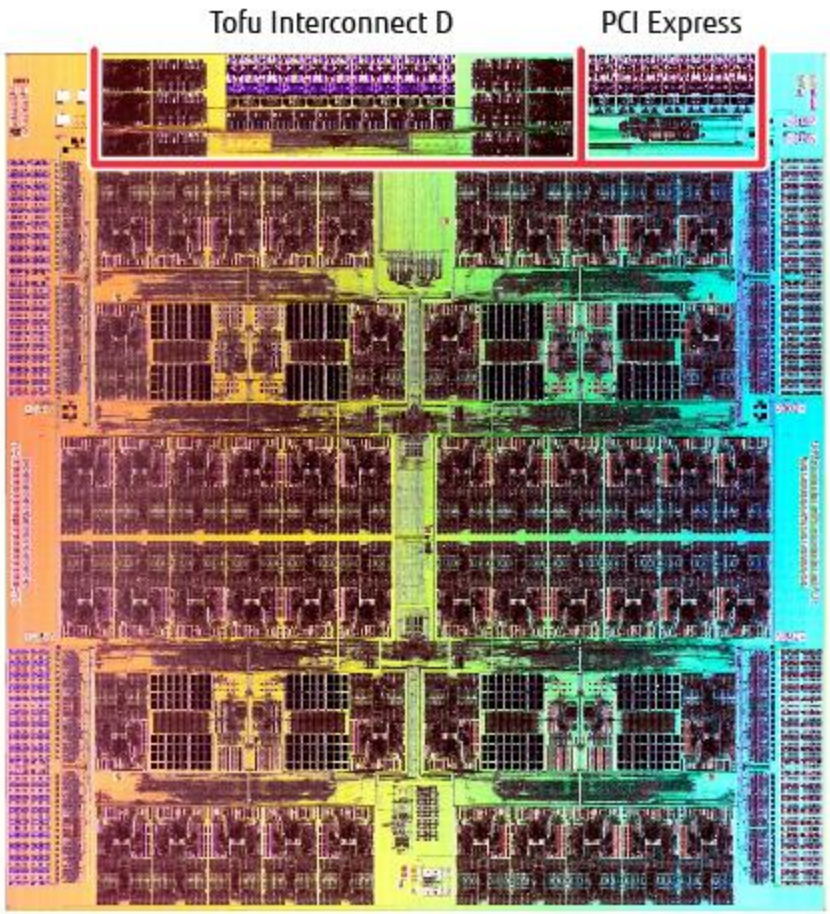
CPU core frequency	1.8	2.0	2.2	GHz
Peak DP perf (FP64)	2.7	3.0	3.3	TFLOPS
Peak SP perf (FP32)	5.5	6.1	6.7	TFLOPS
Peak HP perf (FP16)	11	12	13	TFLOPS
Memory peak bandwidth	1024			GB/s



A64FX w/o LID

Fugaku: World Leading Advanced IT (not just SC)

- **CPU: Highest performing general purpose CPU for high-end computing**
  - First server CPU w/7nm process
  - 3x faster c.f. latest CPUs from US competitors w/SVE & HBM2, etc.
  - 3x power efficient -> GPU-class power efficiency
  - Arm v8.2 ISA compliant (own μ-architecture) => e.g. RHEL works out of the box
- **Network/Interconnect: highest bandwidth & lowest latency (Tofu-D)**
  - 400Gbps-class network/node, 0.5μs latency (c.f. IDC 10~100Gbps, 10~100μs latency)
  - First server CPU w/ on-die NIC & switch => 160K nodes interconnected w/o external switch, 1.6 million switch ports, > 100K AoC cables
  - ~6 PetaByte/s injection bandwidth => 10x aggregate GAFAM IDCs traffic
- **System Architecture => World's first ultra-scale disaggregated architecture**
  - CPU cores (esp. L2 Cache), memory (HBM2) and NIC all connected via on-chip network with multiple DMACs => any memory region in the system of 160K noes accessible by any CPU via RDMA and injected onto on-die L2 cache w/sub-μs latecy



- **Tofu-D logic Embedded into CPU die**
- **25mm² die area (~6% of entire die)**
- **Power: 8~9W (incl. SerDes&AOC, very low power c.f. 100GbE, EDR/HDR IB @ 25-30W)**
  - Constant irrespective of state
  - ~ 4~5 % of entire node
- **Directly connected to on-chip torus network**
  - No I/O bus inbetween e.g. PCI-E
  - Direct DMAC access to L2 cache
- **6-D torus router switch + DMAC**
  - ~160,000 low dimension switch on Fugaku
  - ~1.6 million ports total
- **CPU, Memory, and Tofu-D directly connected to on-chip Xbar & NW => disaggregated architecture**

Fugaku Tofu-D Performance

FUJITSU

- 8B Put transfer between nodes on the same board

	Communication settings	Latency
Tofu1(K)	Descriptor on main memory	1.15 µs
	Direct Descriptor	0.91 µs
TofuD	**To/From far CMGs**	**0.54 µs**
	To/From near CMGs	**0.49 µs**

C.f. 100GbE in IDC
Latency 10~100µs

- Total Injection Bandwidth

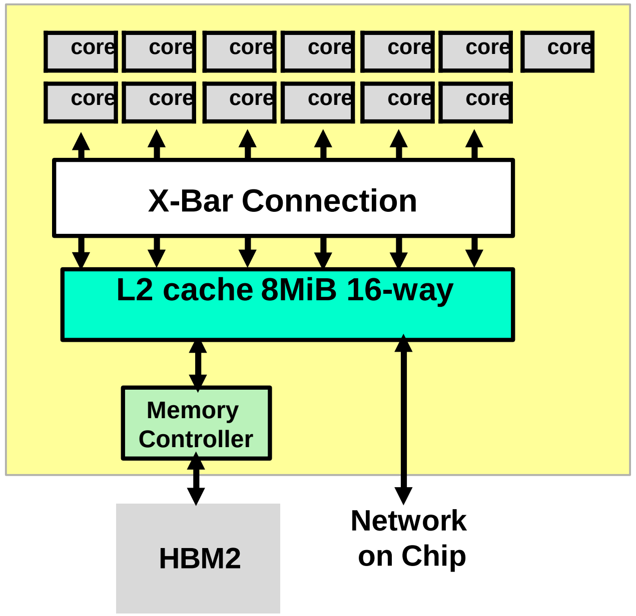
	Injection rate	Efficiency
Tofu1 (K)	15.0 GB/s	77 %
Tofu1 (FX10)	17.6 GB/s	88 %
TofuD	**38.1 GB/s**	**93 %**

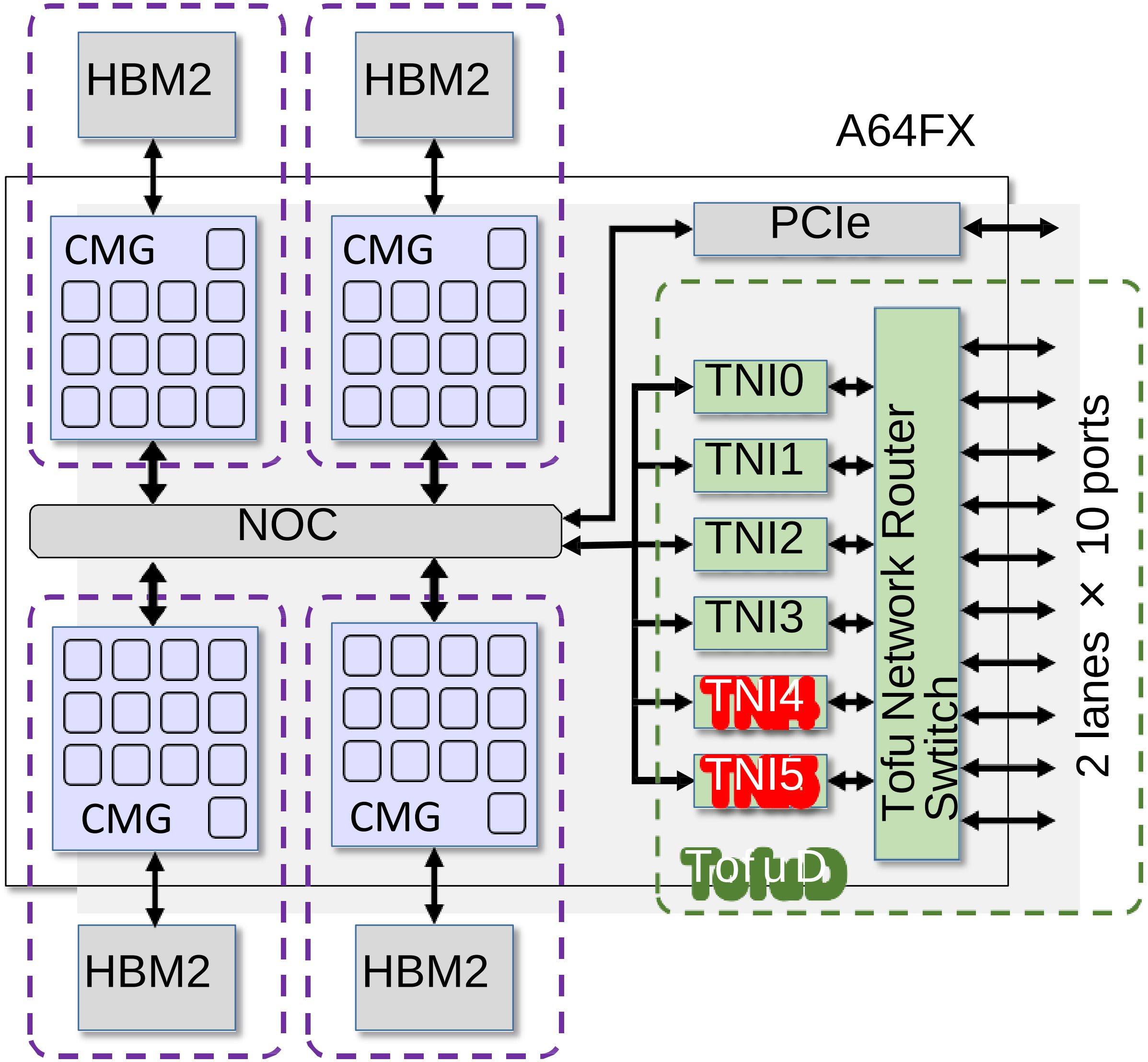
C.f. 100GbE in IDC
Bandwidth ~10GB/s

Disaggregated Architecture of A64FX

- Any CPU can access any memory in system via RDMA (TNI) to its L2
  - Entire 160K Fugaku Nodes
  - Sub microsecond latency
  - NOC + Tofu-D NW Switch on every node (on-die)

CMG Configuration (13 cores + L2 + MC=>HBM2)



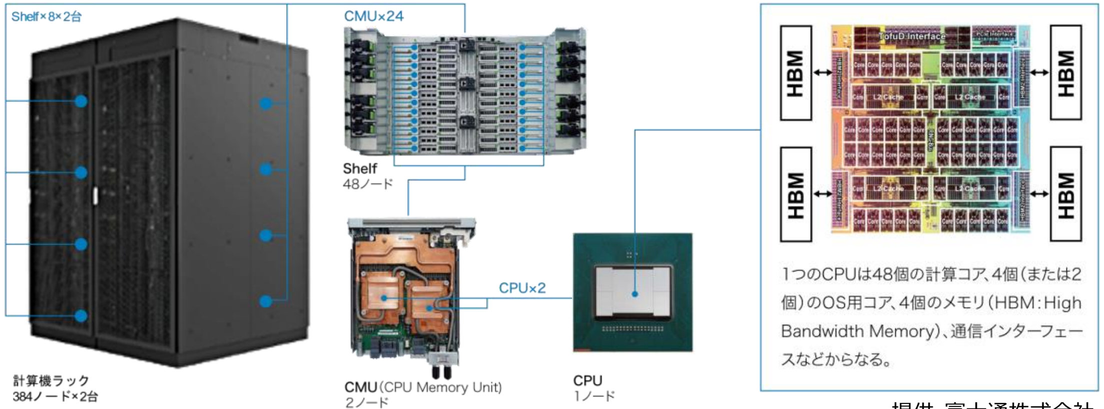


HBM2=>NoC=>TNI=>SW…AoC…SW=>TNI=>NoC=>L2&HBM2

Fugaku Total System Config & Performance

- **Total # Nodes: 158,976 nodes**
  - 384 nodes/rack x 396 (full) racks = 152,064 nodes
  - 192 nodes/rack x 36 (half) racks = 6,912 nodes

  c.f. K Computer 88,128 nodes
- **Theoretical Peak Compute Performances**
  - Normal Mode (CPU Frequency 2GHz)
    - 64 bit Double Precision FP: 488 Petaflops
    - 32 bit Single Precision FP: 977 Petaflops
    - 16 bit Half Precision FP (AI training): 1.95 Exaflops
    - 8 bit Integer (AI Inference): 3.90 Exaops
  - Boost Mode (CPU Frequency 2.2GHz)
    - 64 bit Double Precision FP: 537 Petaflops
    - 32 bit Single Precision FP: 1.07 Exaflops
    - 16 bit Half Precision FP (AI training): 2.15 Exaflops
    - 8 bit Integer (AI Inference): 4.30 Exaops
- **Theoretical Peak Memory Bandwidth: 163 Petabytes/s**



提供: 富士通株式会社

- <u>C.f. K Computer performance comparison (Boost)</u>
  - 64 bit Double Precision FP: 48x
  - 32 bit Single Precision: 95x
  - 16 bit Half Precision (AI training): 190x
    - K Computer Theoretical Peak: 11.28 PF for all precisions
  - 8 bit Integer (AI Inference): > 1,500x
    - K Computer Theoretical Peak: 2.82 Petaops (64 bits)
  - Theoretical Peak Memory Bandwidth: 29x
    - K Computer Theoretical Peak: 5.64 Petabytes/s

Fugaku HPC+Big Data+AI+Cloud 'Converged' Software Stack

Traditional Clouds eg EC2

Fugaku AI (DL4Fugaku)
RIKEN: Chainer, PyTorch, TensorFlow, DNNL…

Live Data Analytics
Apache Flink, Kibana, ….

~3000 Apps sup-ported by Spack

Math Libraries
Fujitsu: BLAS, LAPACK, ScaLAPACK, SSL II
RIKEN: EigenEXA, KMATH_FFT3D, Batched BLAS,,,,

Cloud Software Stack
OpenStack, Kubernetis, NEWT…

Open Source Management Tool
Spack and other DoE ECP Software

Compiler and Script Languages
Fortran, C/C++, OpenMP, Java, python, …
(Multiple Compilers suppoted: Fujitsu, Arm, GNU, LLVM/CLANG, PGI, …)

Batch Job and Management System

Hierarchical File System

ObjectStore
S3 Compatible

Fugaku and future HPCI systems

Tuning and Debugging Tools
Fujitsu: Profiler, Debugger, GUI

Red Hat Enterprise Linux 8 Libraries

High-level Prog. Lang. XMP	Domain Spec. Lang. FDPS	Communication Fujitsu MPI RIKEN MPI	File I/O DTF	Virtualization & Container KVM, Singularity

Process/Thread PIP	Low Level Communication uTofu, LLC	File I/O for Hierarchical Storage Lustre/LLIO

Red Hat Enterprise Linux Kernel+ optional light-weight kernel (McKernel)

Traditional HPC system eg K-computer

Most applications will work with simple recompile from x86/RHEL environment to t he Arm processor.
LLNL Spack automates this.

Standard Software Ecosystem & OSS Contributions

- Arm v8.2 + SVE and other server standards fully compliant
- Standard Linux distributions work out of the box, most Cloud, HPC, BD OSSs as well
- Standardized configurations via frameworks (e.g., OneAPI, Spack), VMs, Containers
- High Performance AI being developed w/OneDNN & others)

openHPC oneAPI python PyTorch TensorFlow

julia QEMU Red Hat Ruby KVM

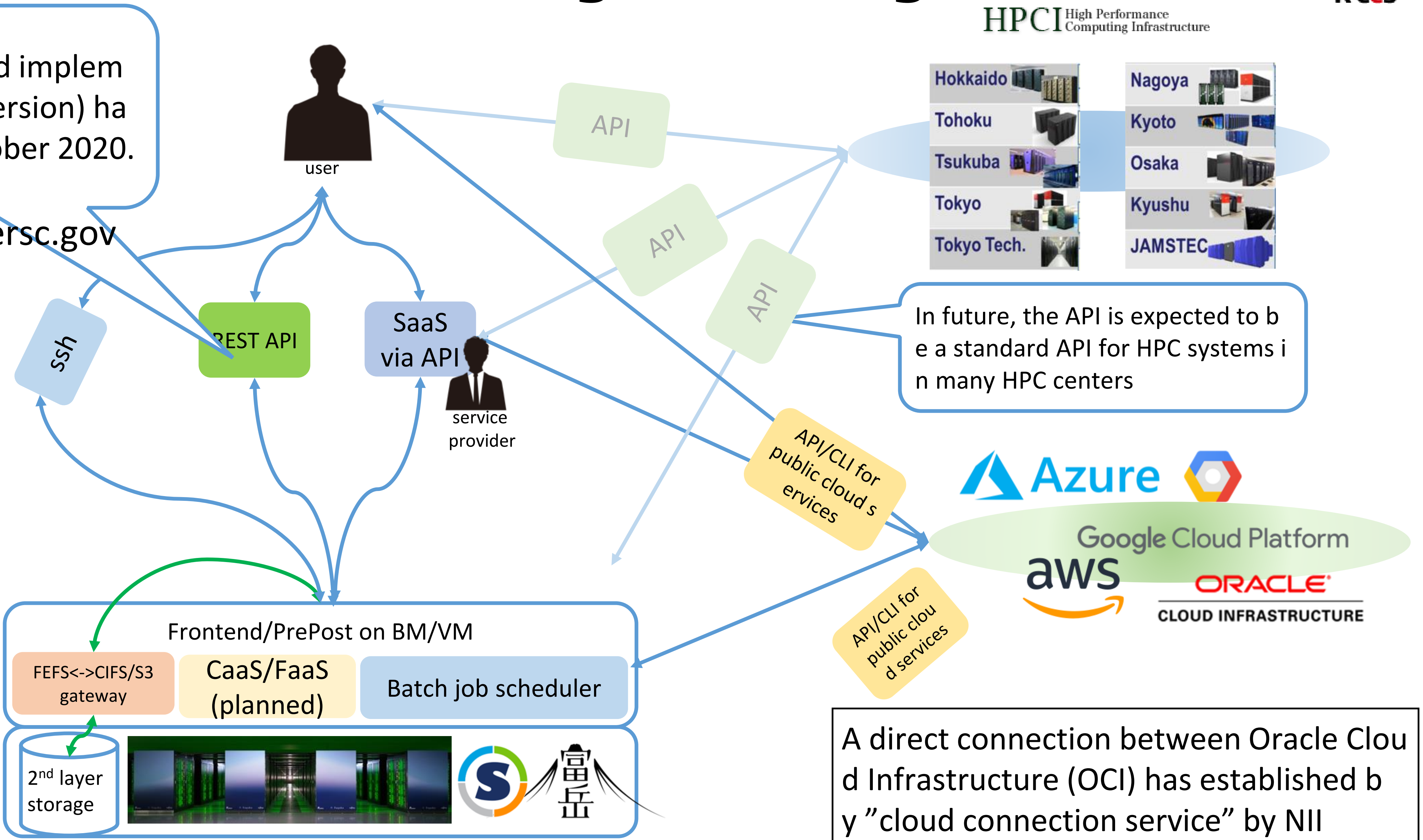
Red Hat Enterprise Linux 8（RHEL8）

Most Software on x86 HPC Clusters & Clouds Simply Work on Fugaku

“Cloud” technologies on Fugaku

NEWT2.0(*) based implementation (alpha version) has released at October 2020.

(*)https://newt.nersc.gov



In future, the API is expected to be a standard API for HPC systems in many HPC centers

A direct connection between Oracle Cloud Infrastructure (OCI) has established by ”cloud connection service” by NII

Collaboration Partners

Recently, **Fujitsu** and **Rescale, Inc.** joined our collaboration.

https://www.riken.jp/pr/news/2021/20210113_1/index.html

Fugaku HPL-AI Results Comparisons (update Jun 2021)

- Compute units utilized (FP16)
  - A64FX: 32-element vector FP16 & FP64 mixed precision
  - GPUs: FP16 Matrix Engine (Tensor Core) & FP64 mixed precision
- FP16 vast difference in efficiency, while FP64 efficiency similar
- See our latest paper "Matrix Engines for High Performance Computing: A Paragon of Performance or Grasping at Straws?" [IEEE IPDPS 2021] https://arxiv.org/abs/2010.14373
- We will also release our code as OSS RSN to become a standard like HPL

	Main Processor	**HPL-AI Measured Performance**	**FP16 Peak Performance (full machine)**	**Efficiency**	**HPL-AI Performance /Chip**	Top500 /Linpack FP64 Measured Performance	FP64 Peak Performance	Efficiency
1. Fugaku	Fujitsu A64FX	**2.00 EF**	**2.14 EF**	**93.2%**	**12.6TF**	442.01 PF	537.21 PF	82.3%
2. Summit	NVIDIA V100	**1.15 EF**	**3.46 EF**	**33.2%**	**42.6TF**	148.60PF	200.79 PF	74.0%
3. Selene	NVIDIA A100	**0.63 EF**	**1.40 EF**	**45.0%**	**140.6TF**	63.46 PF	79.22 PF	80.1%

Note: Selene node count based on prerelease info

Development of DL software stack for Arm SVE

FUJITSU

Framework & oneDNN porting & tuning

Naoki Shinjo, Akira Asato, Atsushi Ike, Koutarou Okazaki, **Yoshihiko Oguchi**, Masahiro Doteguchi, **Jin Takahashi**, Kazutoshi Akao, Masaya Kato, Takashi Sawada, **Naoto Fukumoto**, Kentaro Kawakami, Naoki Sueyasu, Kouji Kurihara, Masafumi Yamazaki, Takumi Honda

Fugaku AI project

Technical support

Fugaku AI project
Signed on Nov. 25, 2019

RIKEN

Tuning for Fugaku

Satoshi Matsuoka, High Performance Artificial Intelligence Systems Research Team Leader

Kento Sato, High Performance Big Data Research Team Leader

Kazuo Minami, Application Tuning Development Unit Leader

Akiyoshi Kuroda, Application Tuning Development Unit

arm

Linaro

Cybozu Labs

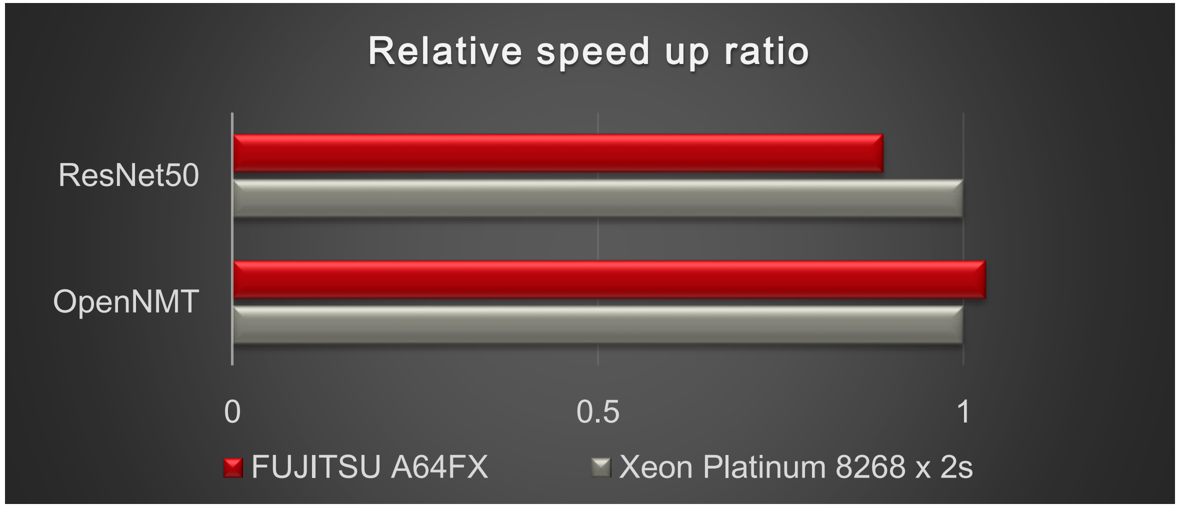
Shigeo Mitsunari

A64FX preliminary results for Deep Learning

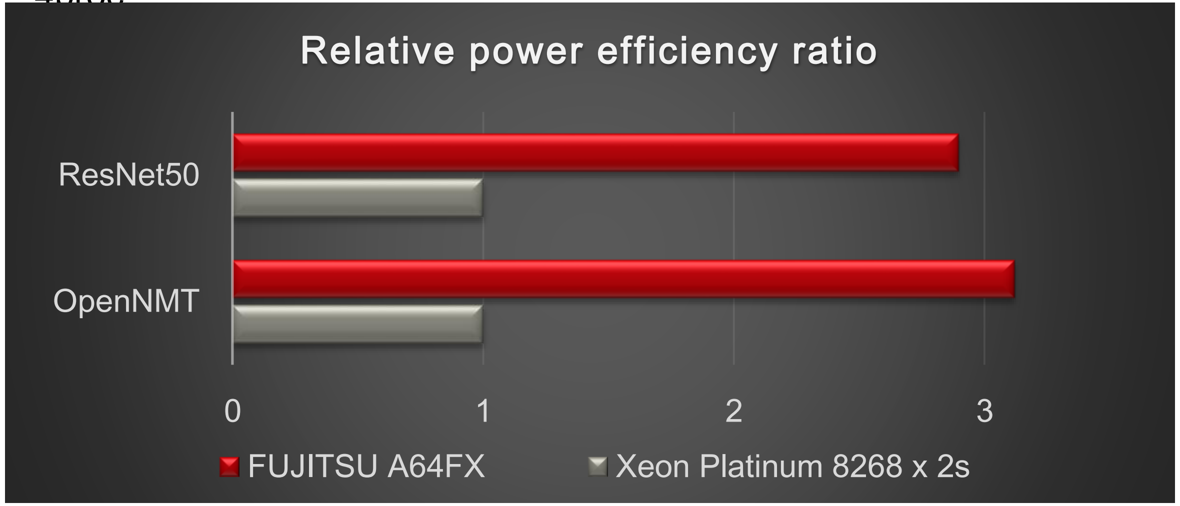
FUJITSU

- Setup
  - Using the same number of CPU cores
    - FX1000 single node (A64FX 2.2 GHz) vs. Xeon Platinum 8268 (24 core, 2.9GHz) x2
  - ResNet50 (image classification)
  - OpenNMT (natural lang. processing)
- Results
  - Performance:
    - Almost the same performance as Xeon
  - Energy efficiency:
    - Up to 2.8x more efficient over Xeon

FX1000



Training using fp32, PyTorch v1.5.0, OneDNN_aarch64, batch size 75 x 4proc.



Training using fp32, PyTorch v1.6.0, OneDNN_aarch64, batch size 3850 x 2proc.

Global | Change

Please enter search word

Services | Products | Solutions | Support | About Fujitsu

Fujitsu, AIST, and RIKEN Achieve Unparalleled Speed on the MLPerf HPC Machine Learning Processing Benchmark Leveraging Leading Japanese Supercomputer Systems

Fujitsu Limited,National Institute of Advanced Industrial Science and Technology,RIKEN

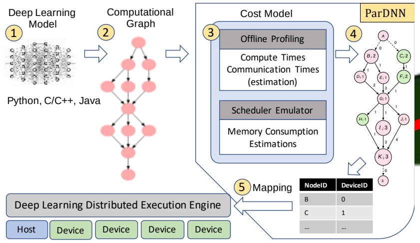
Tokyo, November 19, 2020

Fujitsu, the National Institute of Advanced Industrial Science and Technology (AIST), and RIKEN today announced a performance milestone in supercomputing, achieving the highest performance and claiming the ranking positions on the MLPerf HPC benchmark[(1)]. The MLPerf HPC benchmark measures large-scale machine learning processing on a level requiring supercomputers, and the parties achieved these outcomes leveraging approximately half of the "AI-Bridging Cloud Infrastructure" ("ABCI") supercomputer system, operated by AIST, and about 1/10 of the resources of the supercomputer Fugaku, which is currently under joint development by RIKEN and Fujitsu.

Utilizing about half the computing resources of its system, ABCI achieved processing speeds 20 times faster than other GPU-type systems. That is the highest performance among supercomputers based on GPUs, computing devices specialized in deep learning. Similarly, about 1/10 of Fugaku was utilized to set a record for CPU-type supercomputers consisting of general-purpose computing devices only, achieving a processing speed 14 times faster than that of other CPU-type systems.

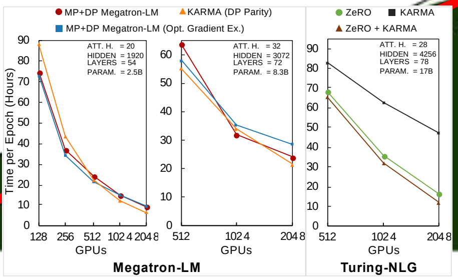
The results were presented as MLPerf HPC v0.7 on November 18th (November 19th Japan Time) at the 2020 International Conference for High Performance Computing, Networking, Storage, and Analysis (SC20) event, which is currently being held online.

Exploring and Merging Different Routes to O(100,000s) Nodes Deep Learning



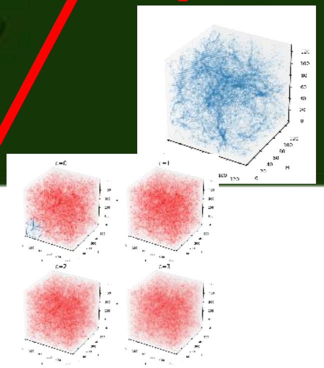
Non-intrusive graph-based partitioning strategy for large DNN models achieving superlinear scaling [1]

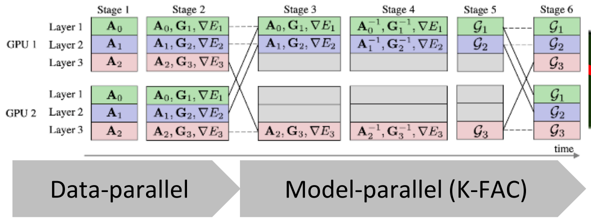
AIST, Koc U.



KARMA: Out-of-core distributed training (pure data-parallel) outperforming SoTA NLP models on 2K GPUs [2]

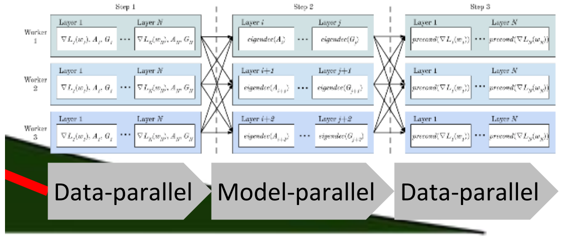
AIST, Matsuoka-lab, RIKEN





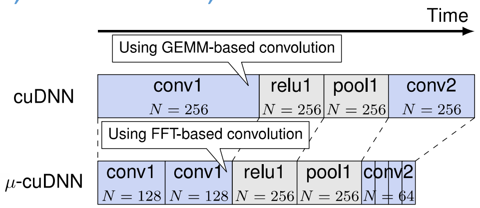
A model-parallel **2nd-order method (K-FAC)** trains ResNet-50 on **1K GPUs** in 10 minutes [4]

TokyoTech, NVIDIA, RIKEN, AIST



Layer-wise distribution and inverse-free design further accelerate K-FAC [5]

UT Austin, UChicago, ANL



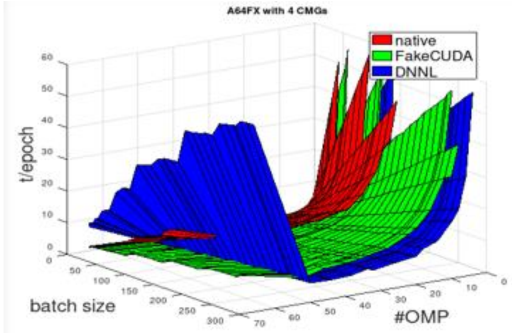
Model-parallelism enables 3D CNN training on **2K GPUs** with 64x larger spatial size and better convergence [3]

Matsuoka-lab, LLNL, LBL, RIKEN

Merging Theory and Practice

Layer-wise loop splitting accelerates CNNs [6]

Matsuoka-lab, ETH Zurich



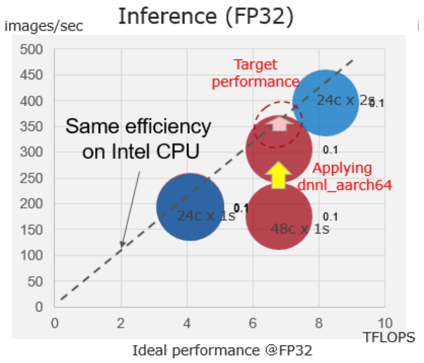
MocCUDA: Porting CUDA-based Deep Neural Network Library to A64FX and (other CPU arch.)

RIKEN, Matsuoka-lab, AIST

Engineering for Performance Foundation

Porting High Performance CPU-based Deep Neural Network Library (DNNL) to A64FX chip

Fujitsu, RIKEN, ARM

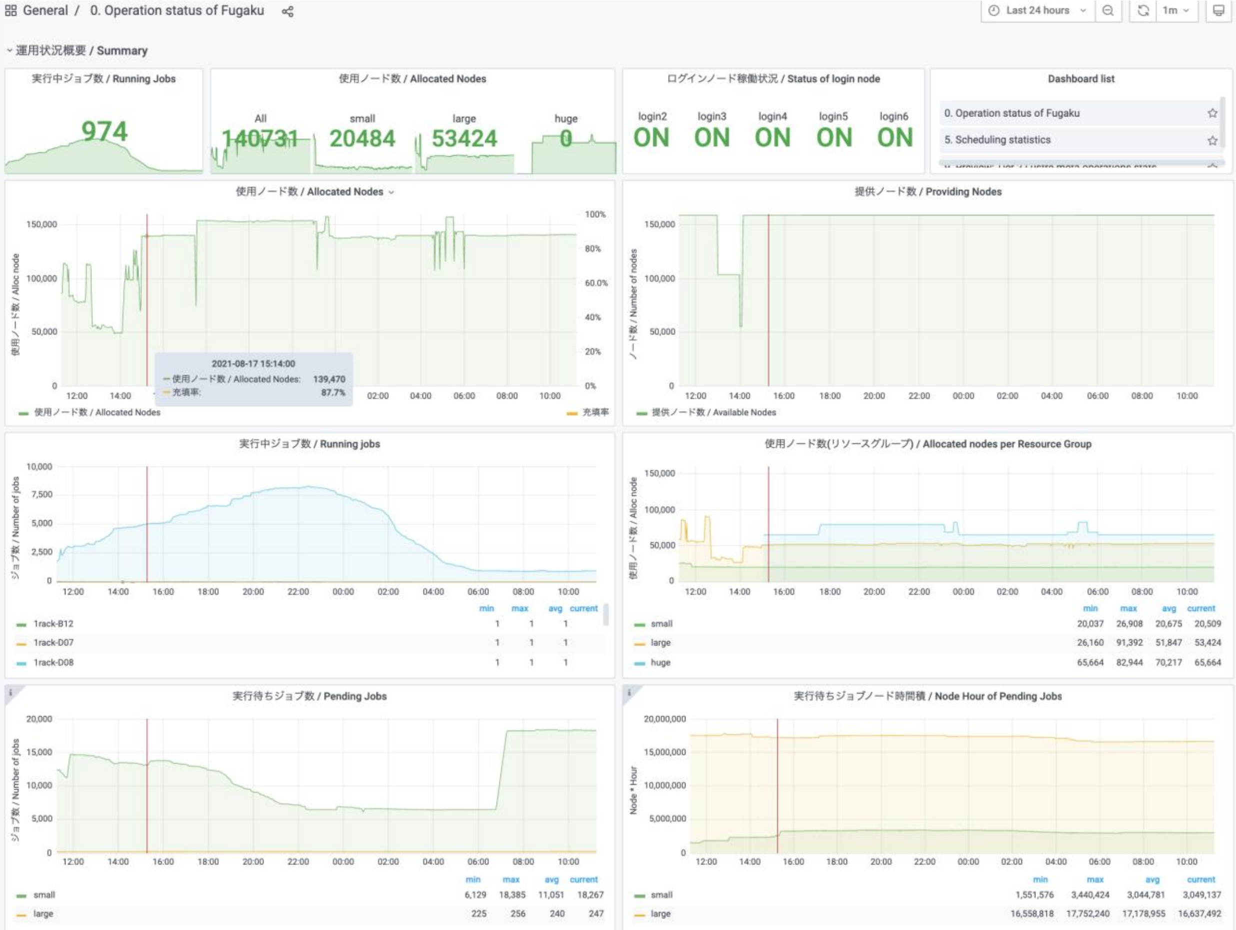


[1] M. Fareed et al., “A Computational-Graph Partitioning Method for Training Memory-Constrained DNNs”, Submitted to PPoPP21
[2] M. Wahib et al., “Scaling Distributed Deep Learning Workloads beyond the Memory Capacity with KARMA”, ACM/IEEE SC20 (Supercomputing 2020)
[3] Y. Oyama et al., “The Case for Strong Scaling in Deep Learning: Training Large 3D CNNs with Hybrid Parallelism,” arXiv e-prints, pp. 1–12, 2020.
[4] K. Osawa, et al., “Large-scale distributed second-order optimization using kronecker-factored approximate curvature for deep convolutional neural networks,” Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 2019-June, pp. 12351–12359, 2019.
[5] J. G. Pauloski, Z. Zhang, L. Huang, W. Xu, and I. T. Foster, “Convolutional Neural Network Training with Distributed K-FAC,” arXiv e-prints, pp. 1-11, 2020.
[6] Y. Oyama et al., “Accelerating Deep Learning Frameworks with Micro-Batches,” Proc. IEEE Int. Conf. Clust. Comput. ICCC, vol. 2018-September, pp. 402–412, 2018.

Visualize all of the operation and service (cont`d)

infrastructure

HPC system



Recent power consumption trend

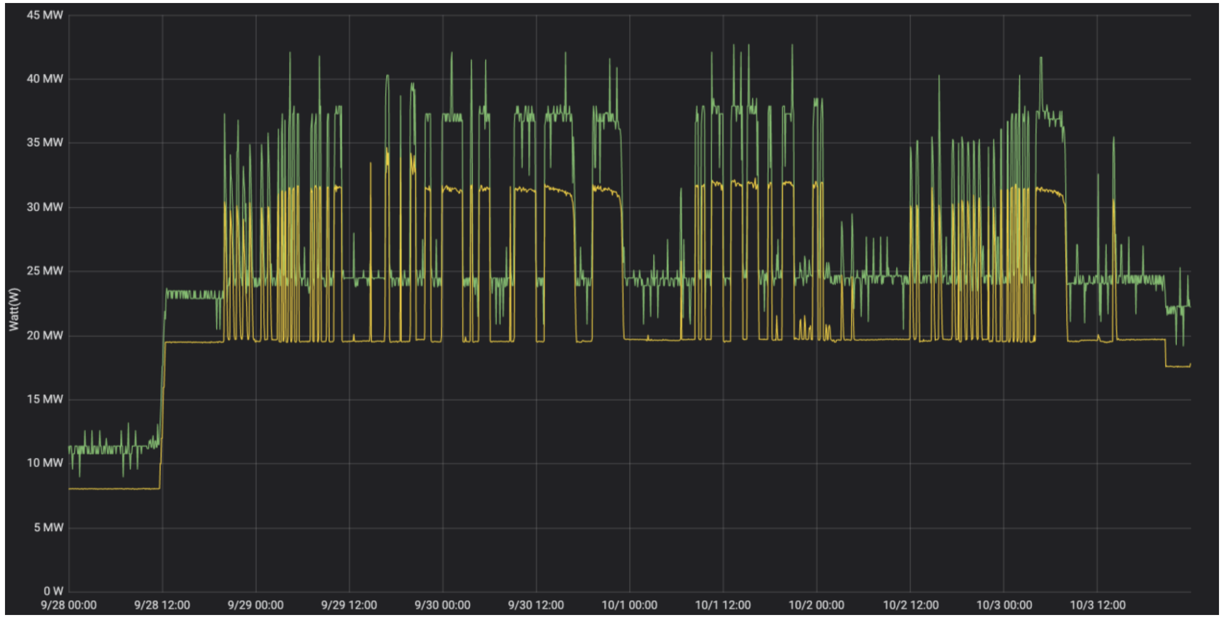
site total

Fugaku

last 30 days power consumption history

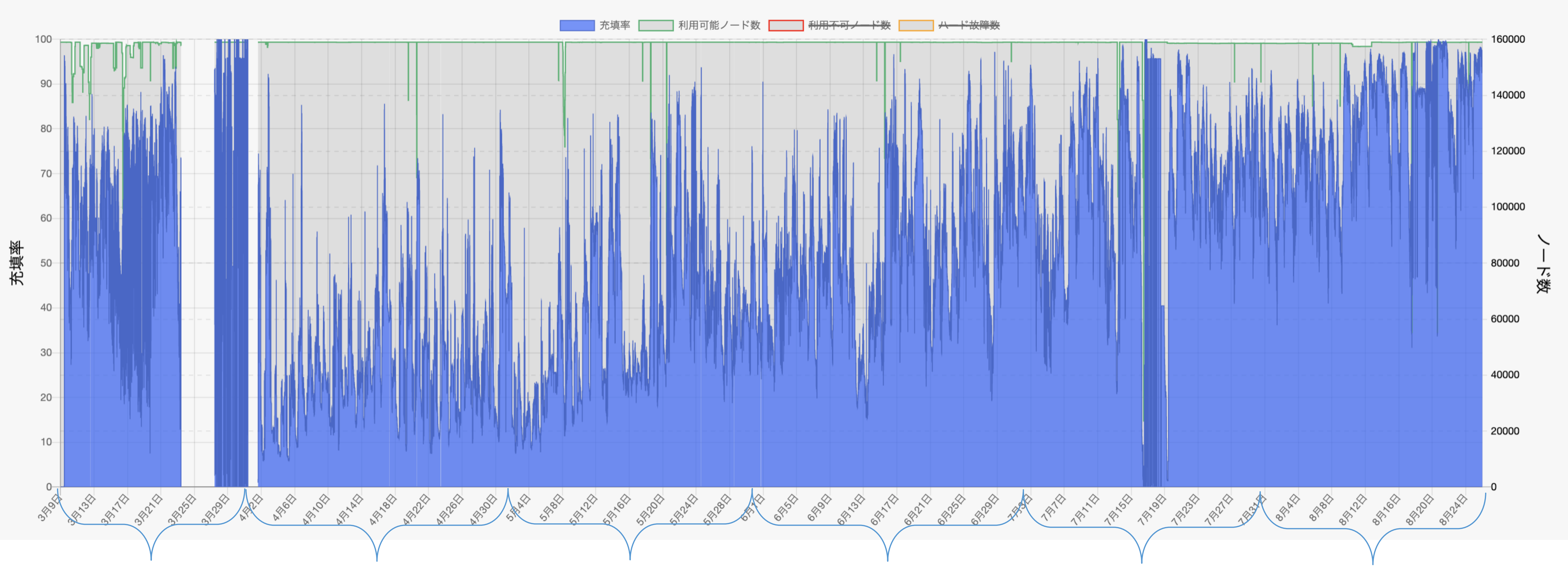
average power consumption
~22-23MW(site total)
~18-19MW(Fugaku)
"DoE Goal: Exascale at 20 MW"

Full node HPCG/HPL measurement



max power consumption (HPCG)
42.70MW(site total)
34.66MW(Fugaku)
power swing ~15MW

Job filling rate (-8/25)



March
64.3%

April
28.3%

May
38.8%

June
52.3%

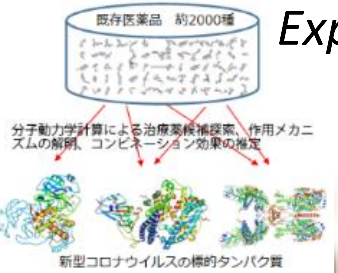
July
71.6%

August
82.2%

MEXT Fugaku Program: Fight Against COVID19

Fugaku resources made available a year ahead of general production (more research topics under international solicitation, also joined US-lead COVID-19 High Performance Computing Consortium)

Medical-Pharma

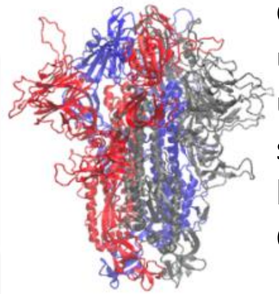


Exploring new drug candidates for COVID-19

Large-scale MD to search & identify therapeutic drug candidates showing high affinity for COVID-19 target proteins from 2000 existing drugs

(Yasushi Okuno, RIKEN / Kyoto University)

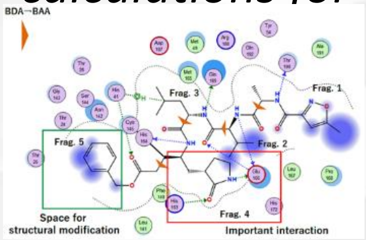
Prediction of conformational dynamics of proteins on the surface of SARS-Cov-2



GENESIS MD to interpolate unknown experimentally undetectable dynamic behavior of spike proteins, whose static behavior has been identified via Cryo-EM

((Yuji Sugita, RIKEN)

Fragment molecular orbital calculations for COVID-19 proteins

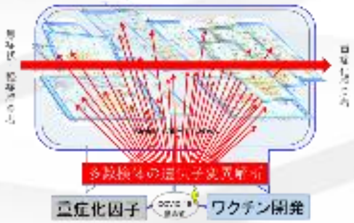


Large-scale, detailed interaction analysis of COVID-19 using Fragment Molecular Orbital (FMO) calculations using ABINIT-MP

(Yuji Mochizuki, Rikkyo University)

Host genetic analysis for severe COVID-19

Whole-genome sequencing of severe cases of COVID-19 and mild or asymptomatic infections, and identify risk-associated genetic variants for severe disease

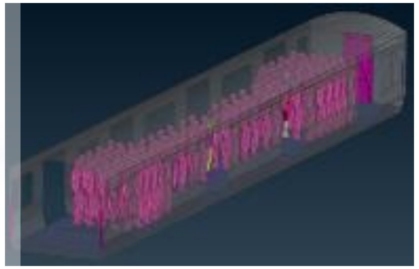


(Satoru Miyano, Tokyo Medical and Dental University)

Societal-Epidemiology

Prediction and Countermeasure for Virus Droplet Infection under the Indoor Environment

Massive parallel simulation of droplet scattering with airflow and hat transfer under indoor environment such as commuter trains, offices, classrooms, and hospital rooms



(Makoto Tsubokura, RIKEN / Kobe University)

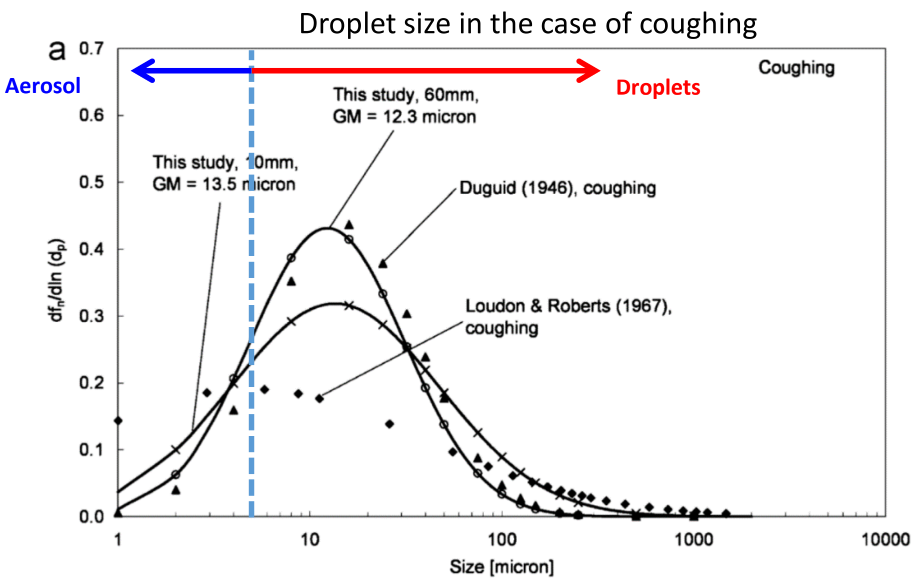
Simulation analysis of pandemic phenomena

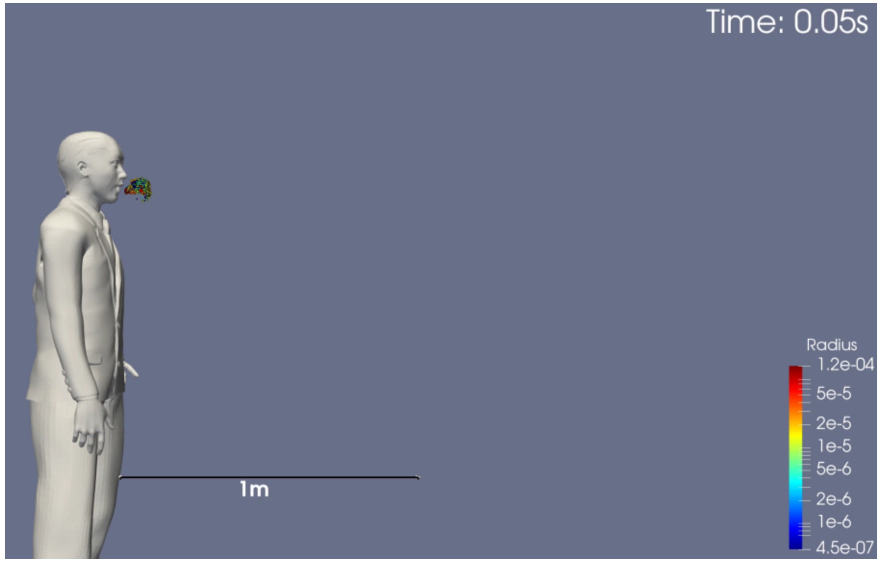
Combining simulations & analytics of disease propagation w/contact tracing apps, economic effects of lockdown, and reflections social media, for effective mitigation policies

(Nobuyasu Ito, RIKEN)

Difficulty in COVID 19 transmission

- **Basically the risk of airborne transmission can be determined by four factors:**
  - Behavior (breathing, speaking, singing…), Staying time, Room volume, Ventilation rate
- **How droplets disperse in the air?**





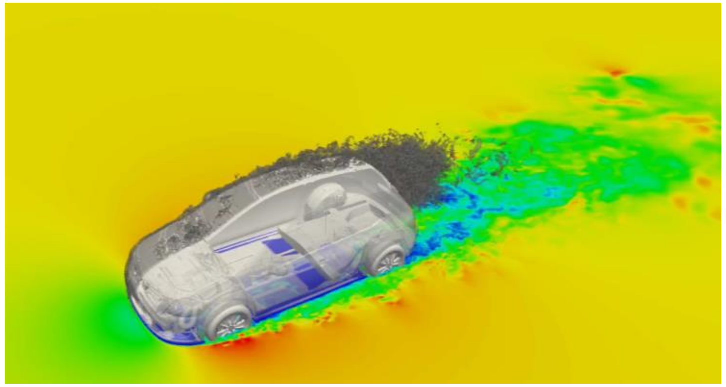
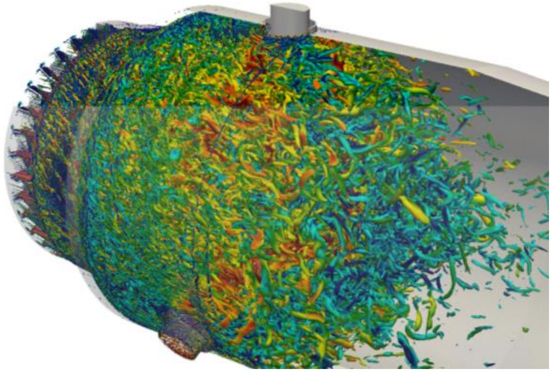
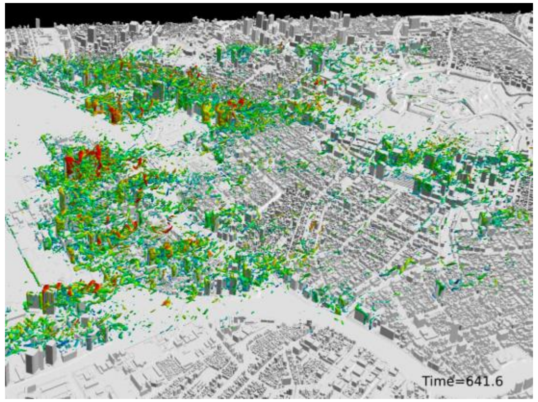
- **COVID 19 does not cause as strong airborne infections as tuberculosis and measles, and thought to be at high risk of inhaling droplets especially smaller than 5 microns at close range to the infected person.**
- **Evaluation based on “instantaneous homogeneous dispersion” does not work!**

C.Y.H. Chao et al./Characterization of expiration air jets and droplet size distributions immediately at the mouth opening/Aerosol Science 40(2009)122-133

Software “CUBE” realizing the Society 5.0

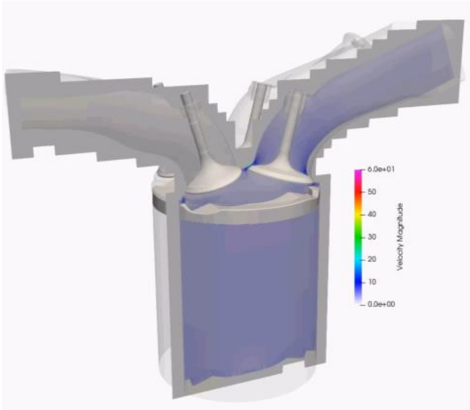
Realizing a huge number of simulations at very high speed including pre- and post-processing

- Having been developing since 2012.
- Many achievements on the supercomputer K, for vehicle aerodynamics, combustion systems, and high-rise buildings.

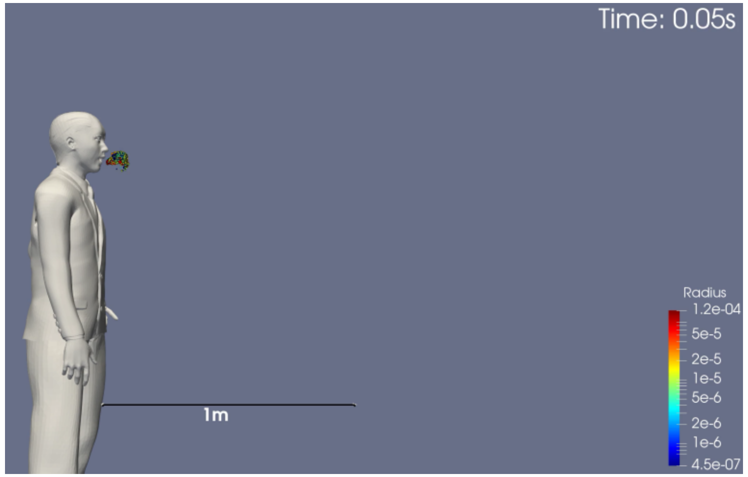




- COVID-19 pandemic early in 2020, when we were tuning “CUBE” on the supercomputer “Fugaku”

IC combustion engine simulation by CUBE on the supercomputer K and fuel spray injection.



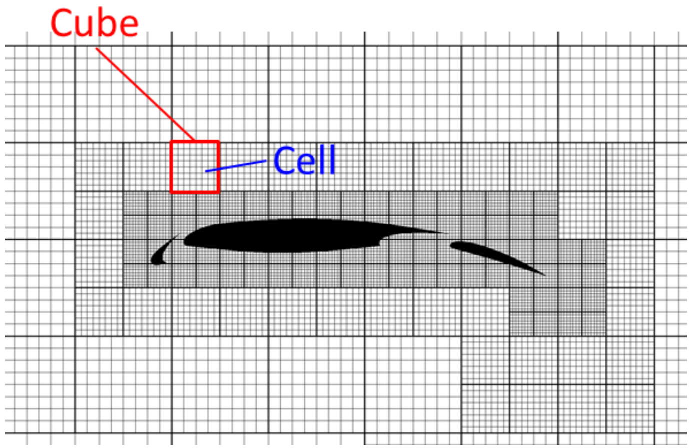

Droplet/aerosol dispersion simulation on the supercomputer “Fugaku”



Complex Unified Simulation Framework: CUBE

Hierarchically structured Finite Volume Method

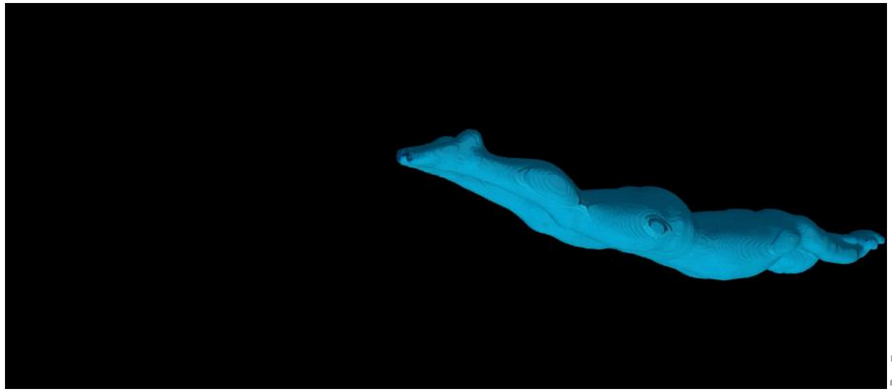
- A solver for coupled phenomena: fluid/structure/acoustics/chemical reaction…
- Building Cube Method for the unified data structure (Nakahashi et al., 2003)
    - Easy tune for both single node and parallel performance



- Immersed Boundary Method (Fadlun et al., 2002)

(1) Dirty CAD treatment (Onishi et al., 2013)
(2) Moving Boundary Method (Bale et al., 2016)
(3) Unified Compressible/Incompressible analysis (Li)
(4) Unified Fluid/Structure analysis (Nishiguchi)

K. Nakahashi, Building-Cube Method for Large-Scale, High Resolution Flow Computations, Am. Inst. Aeronaut. Astronaut. 42nd AIAA (2004) 1–9.



Eulerian Air and Lagrangean Spray Coupling

Flow Solver

Eulerian Mesh

Conservation Equations

Navier-Stokes

$$\frac{\partial U}{\partial t} + \nabla \cdot \mathbf{F} = S, \quad S_{\rho Y_k} = -\frac{1}{\Delta V}\sum_n \frac{dm_{d,k}}{dt}$$

Species transport

time stepping

Lagrangian Particles

Spray equations

Particle tracking

$\overline{W}$ - Average molecular wt of the gas phase
W_V - Molecular wt of water vapor
P_{sat} - Saturated vapor pressure
$Y_{V,s}$ - Vapor surface mass fraction
Y_V - mass fraction of vapor in the far field.
$X_{V,s}$ -Mole fraction of vapor at droplet surface
Sc – Schmidt number
Pr – Prandtl number

Drag Model

$$\frac{d\mathbf{x}_d}{dt} = \mathbf{u}_d, \quad \frac{d\mathbf{u}_d}{dt} = \frac{6}{8}\frac{\rho_g}{d_d \rho_d}|\mathbf{u}-\mathbf{u}_d| (\mathbf{u}-\mathbf{u}_d)C_d$$

$$C_d = \begin{cases} 0.424 & \mathrm{Re}_p > 1000 \\ \frac{24}{\mathrm{Re}_p}\left(1+\frac{1}{6}\mathrm{Re}_p^{2/3}\right) & \mathrm{Re}_p \le 1000 \end{cases}$$

$$Re_p = \frac{\rho_g |\mathrm{u}-\mathrm{u_d}| \mathrm{d_d}}{\mu}$$

Evaporation Model

$$\frac{dT_d}{dt} = \frac{Nu}{3Pr}\left(\frac{c_p}{c_l}\right)\left(\frac{f_2}{\tau_d}\right)(T-T_d)\frac{1}{m_d}\left(\frac{dm_d}{dt}\right)\frac{L_V}{c_{p,d}}$$

$$\dot{m}_d = -\frac{m_d}{\tau_d}\left(\frac{Sh}{3Sc}\right)\ln(1+B_M)$$

$$Nu = 2 + 0.552 Re_s^{1/2} Pr^{1/3}$$

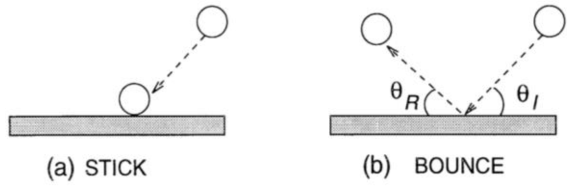
$$Sh = 2 + 0.55 Re_s^{1/2} Sc^{1/3}$$

$$B_M = \frac{Y_{V,s} - Y_V}{1 - Y_{V,s}} \qquad \tau_d = \frac{\rho_d d_d^2}{18\mu}$$

$$Y_{V,s} = \frac{X_{V,s}}{X_{V,s} + (1 - X_{V,s})\overline{W}/W_V} \qquad X_{V,s} = \frac{P_{sat}}{P}$$

Wall Reflection Model

Regime transition state	Critical Weber number
Adhesion (stick/spread) → splash	$\mathrm{We}_c \approx 2630 \cdot \mathrm{La}^{-0.183}$



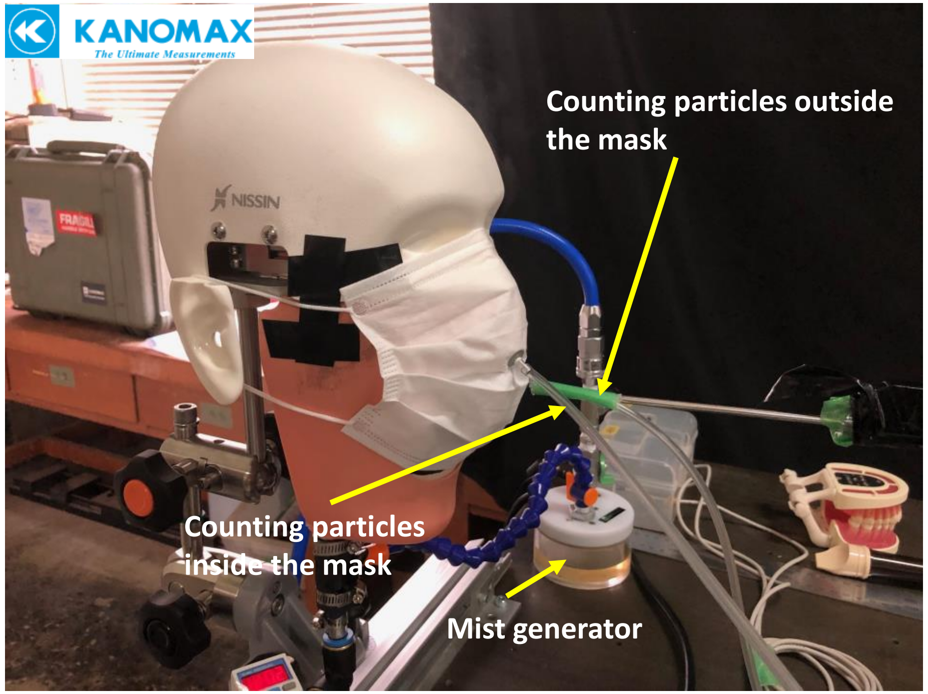
$\mathrm{We} = \rho V_{IN}^2 d_I/\sigma \qquad \mathrm{La} = \rho\sigma d_I/\mu^2$

Face Masks of Different Filter Materials

Experimental setups for the simulation input.(Conducted at Toyohashi Institute of Tech.)

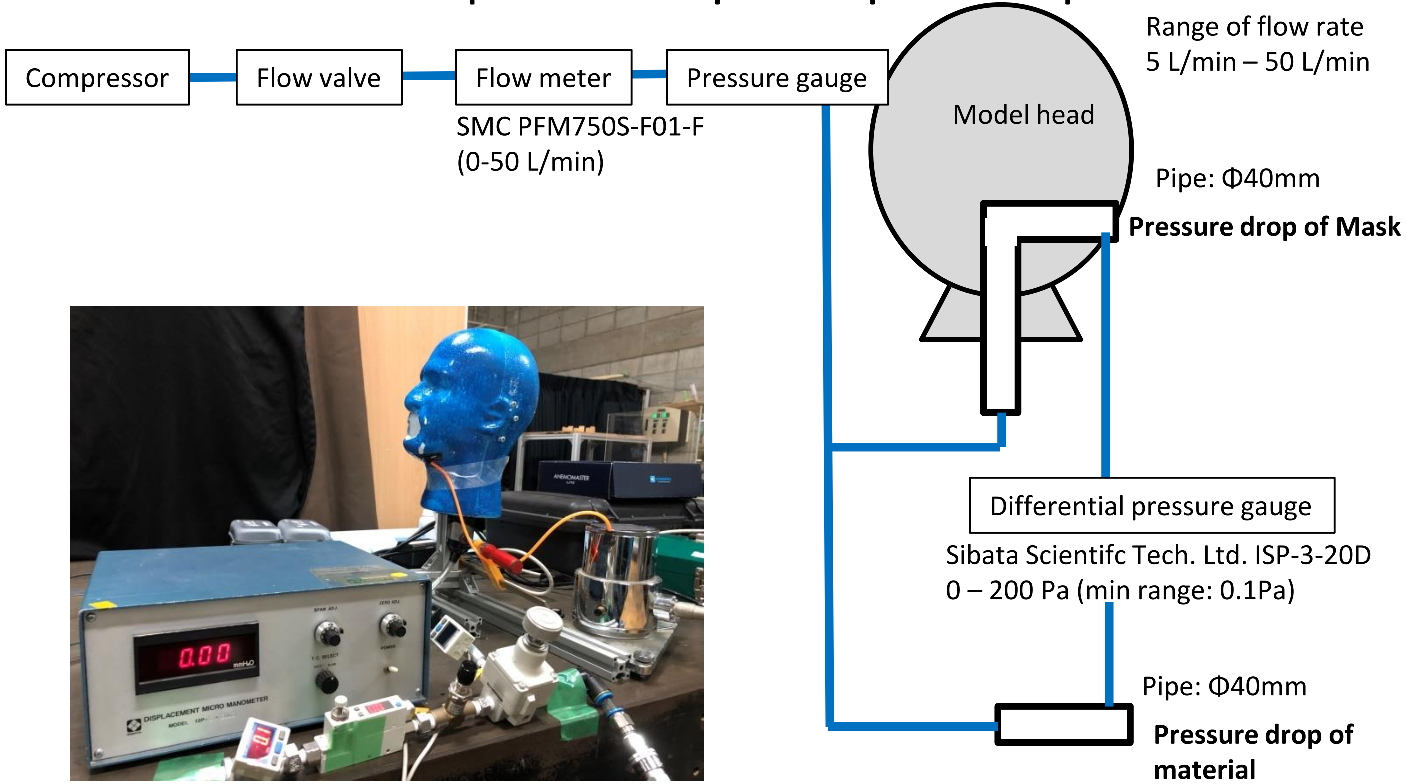
- The filter efficiency and the pressure drop of each material as input, we conduct mask simulation and investigate how the filter materials affect droplet and aerosol spreading.
- Filter efficiency: counting particles before and after each material using particle counter (0.3μm～10μm).
- Pressure drop: measuring pressure before and after each material.

Experimental setups for the filter efficiency

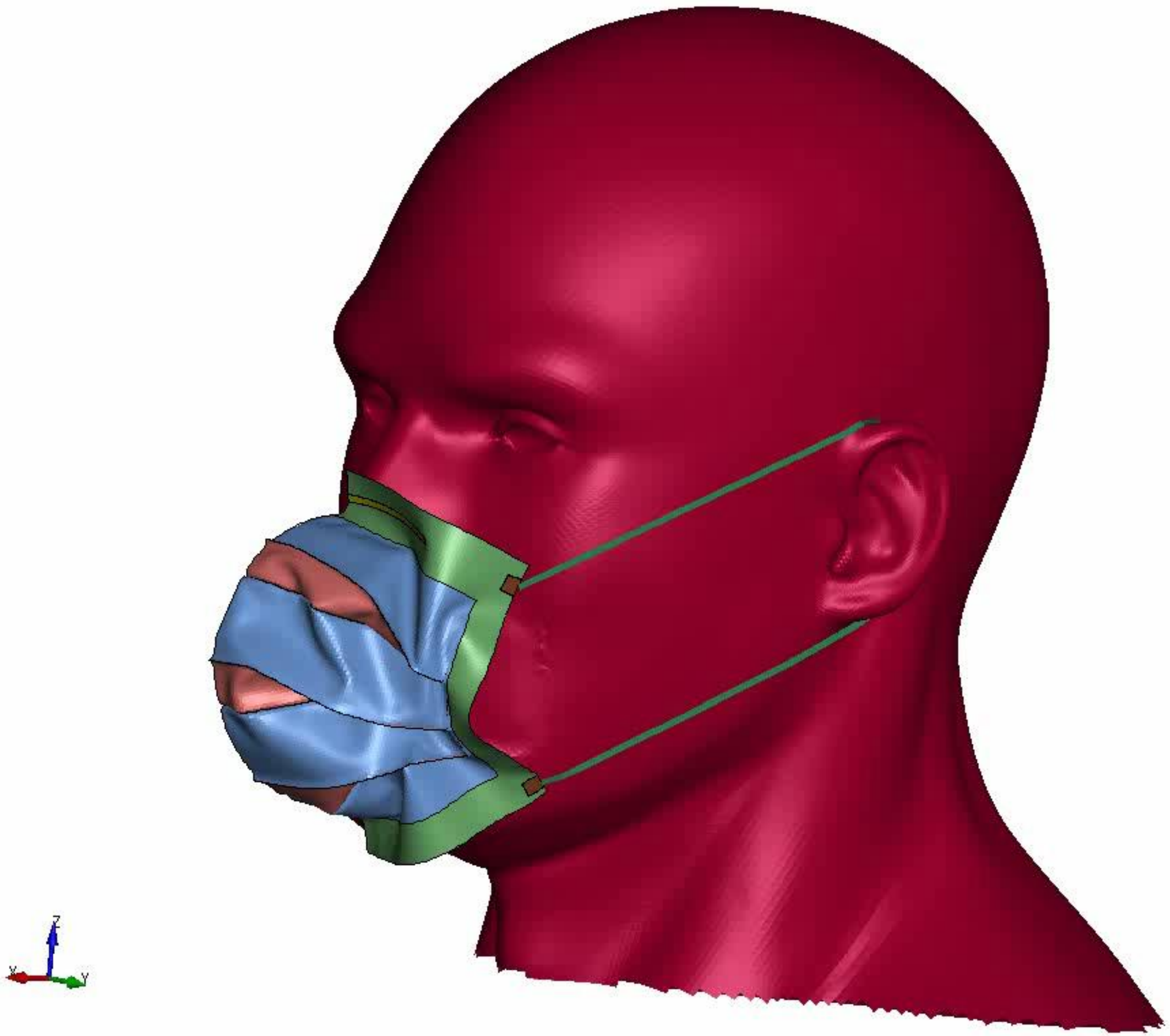


Mask fit tester（Kanomax）Model 3000
Particle counter（Kanomax）Model 3889

Experimental setups for the pressure drop



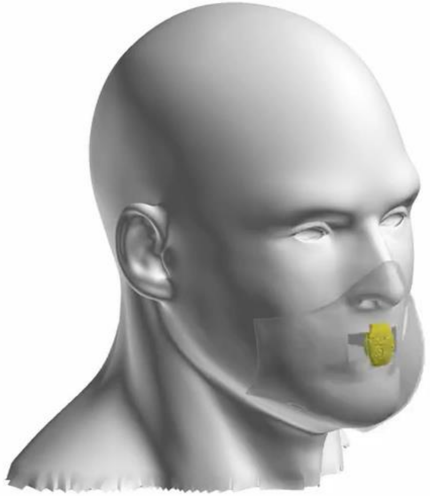
Face Mask's deformation by FEM



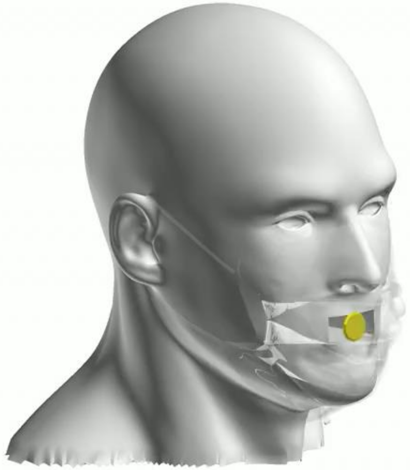
Face Masks of Different Filter Materials

How the filter materials of masks affect droplet/aerosol spreading?

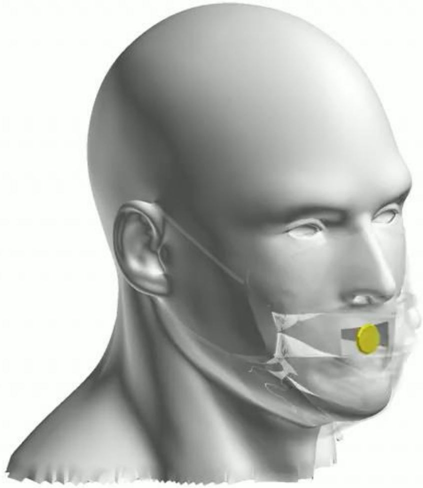
Urethane



Fabric (cotton)

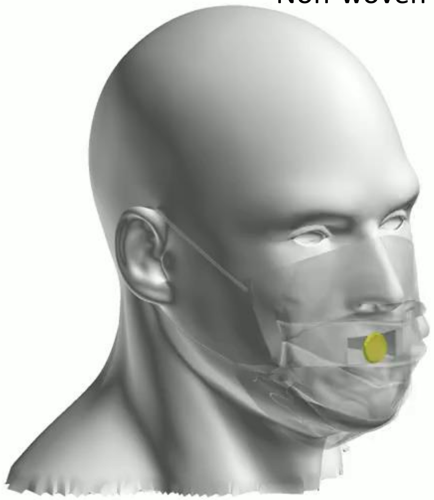


Fabric (polyester)

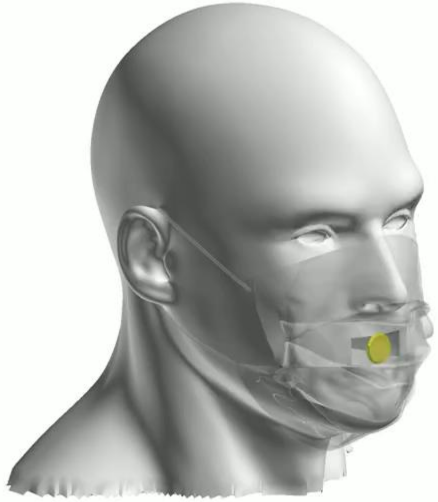


Yellow : leaking from the gap
Red : trapped by the mask
Blue : permeating through the mask

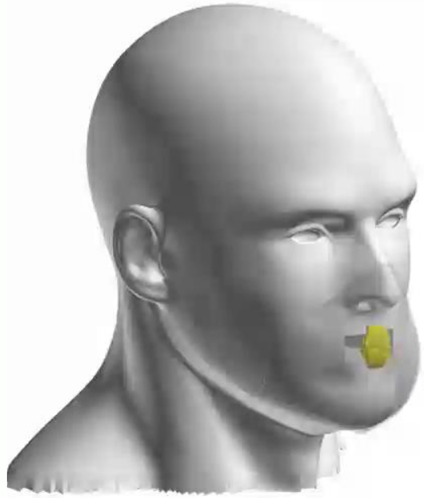
Non-woven ①



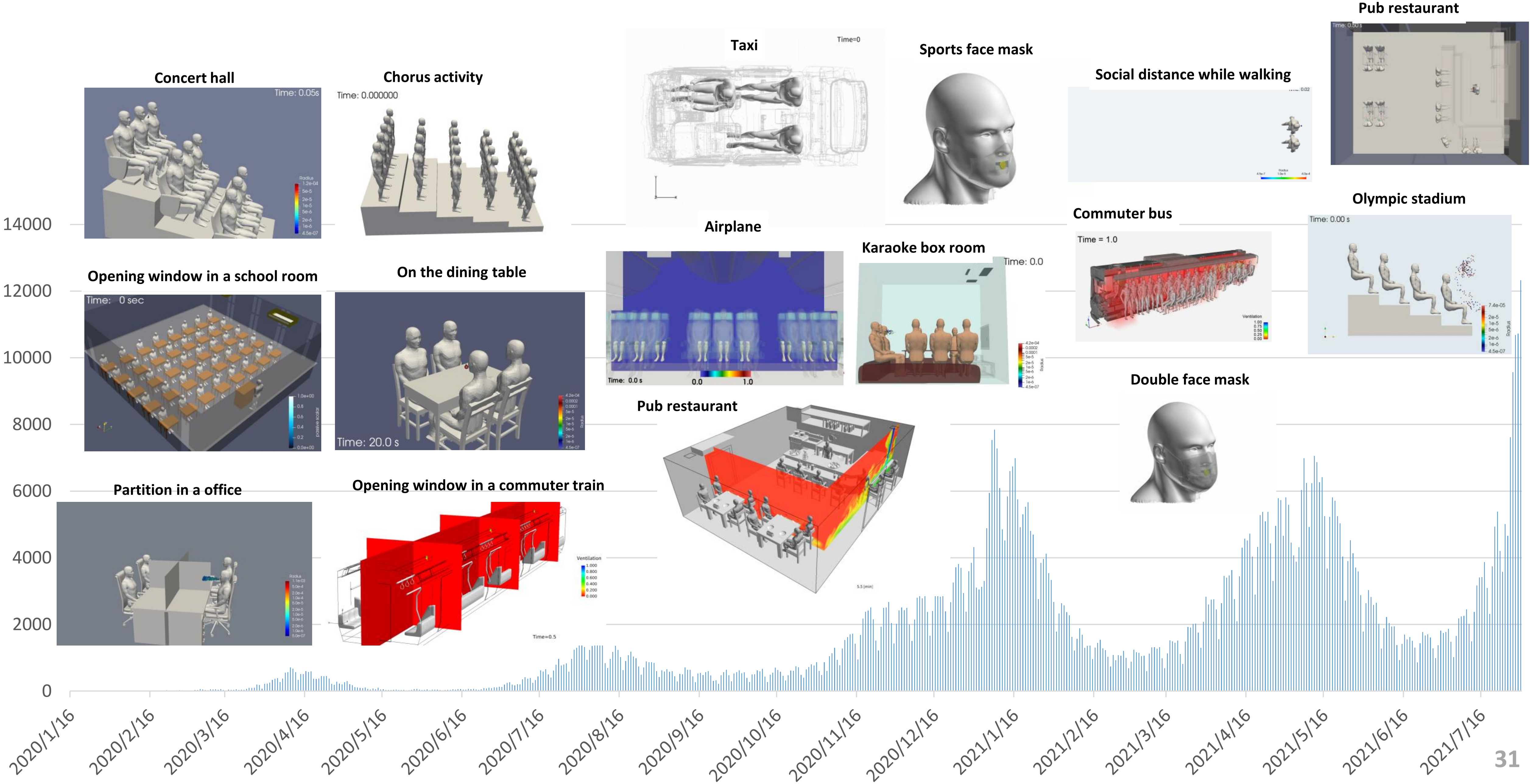
Non-woven ②



N95



Capacity computing made only by “Fugaku”

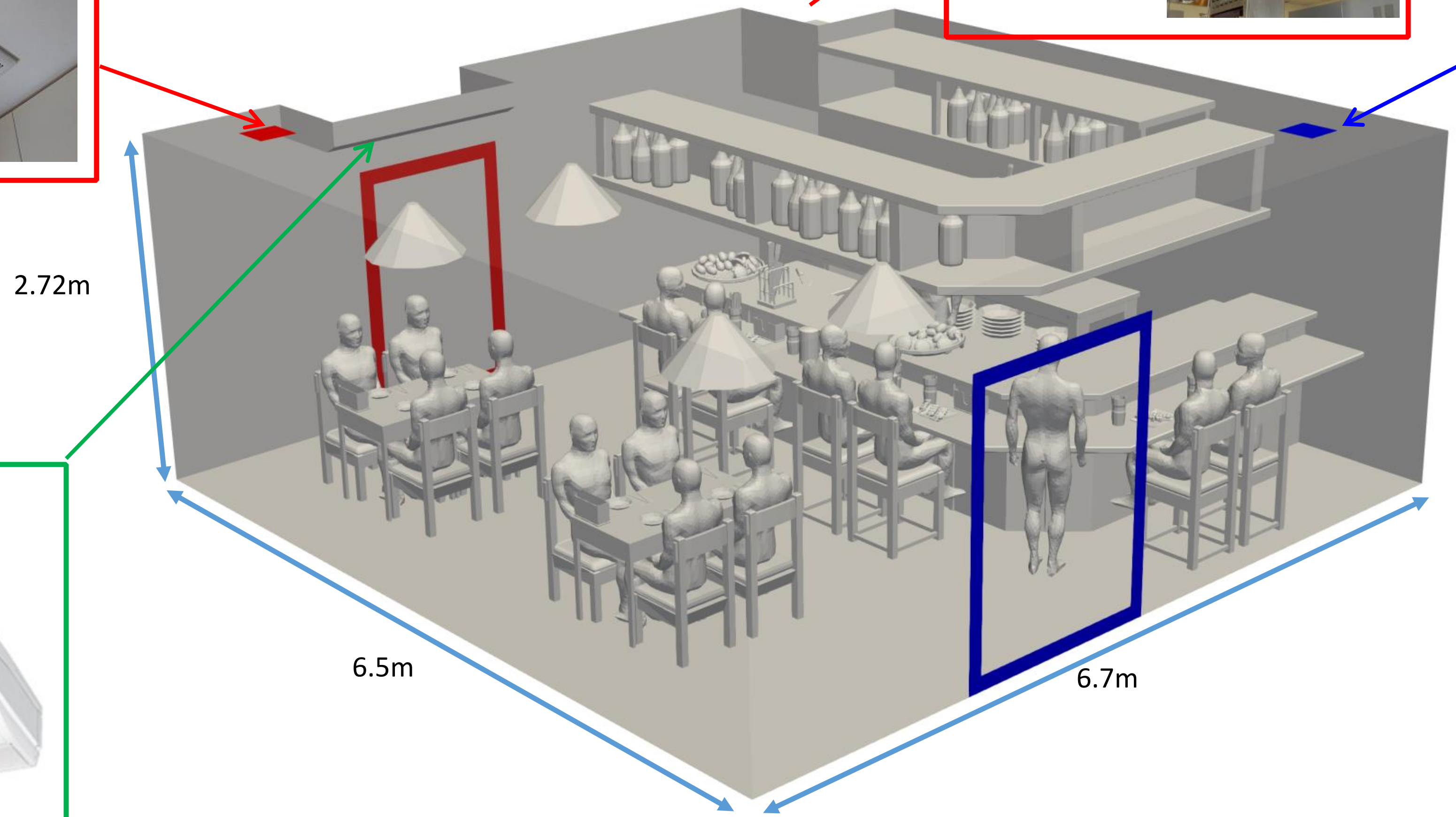


Risks at a izakaya pub.

Exhaust (ventilation)
440m^3/h

Kitchen Duct
1156 m^3/h

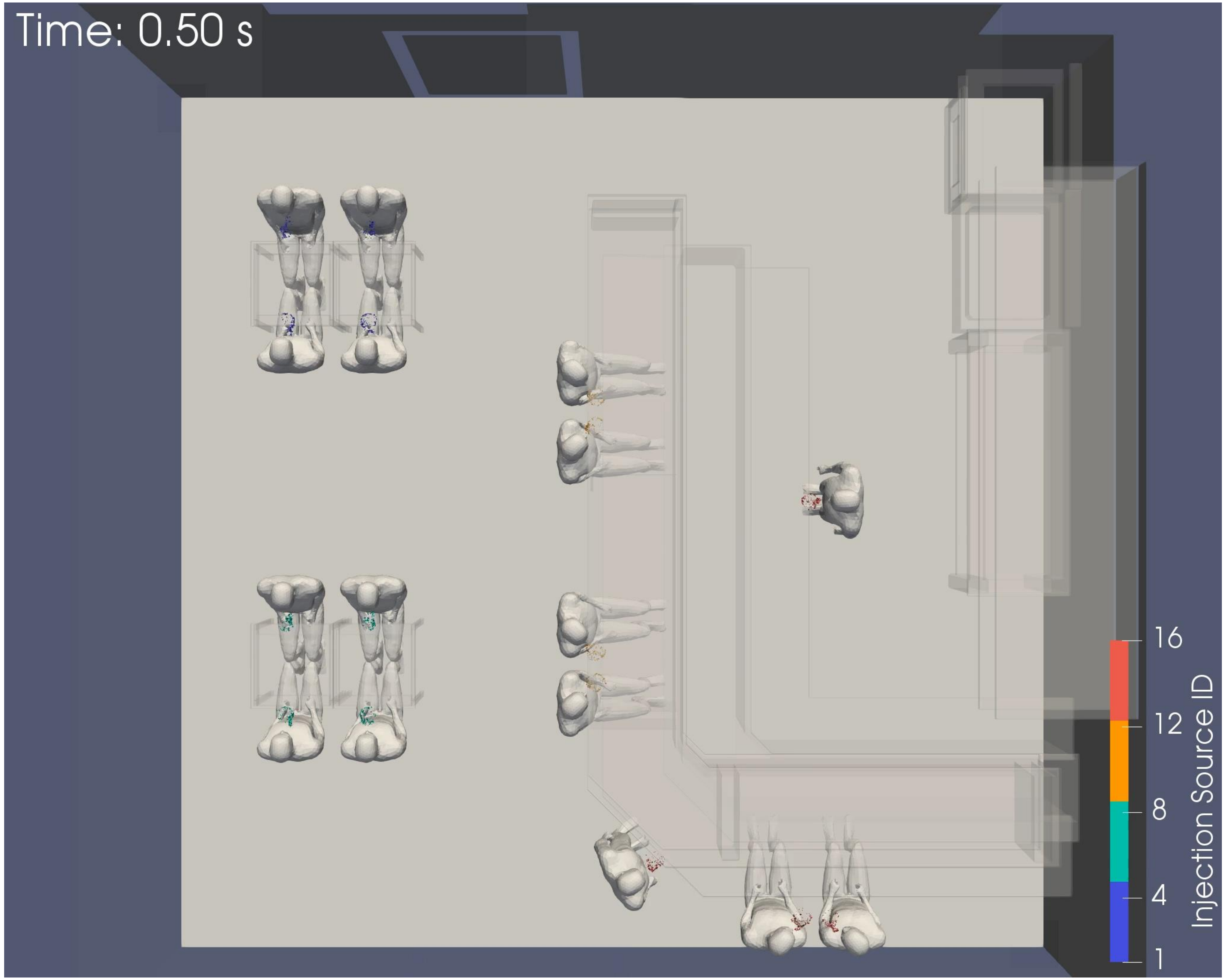
Fresh air supply
540m^3/h
20 deg.



Air Conditioner (no ventilation)

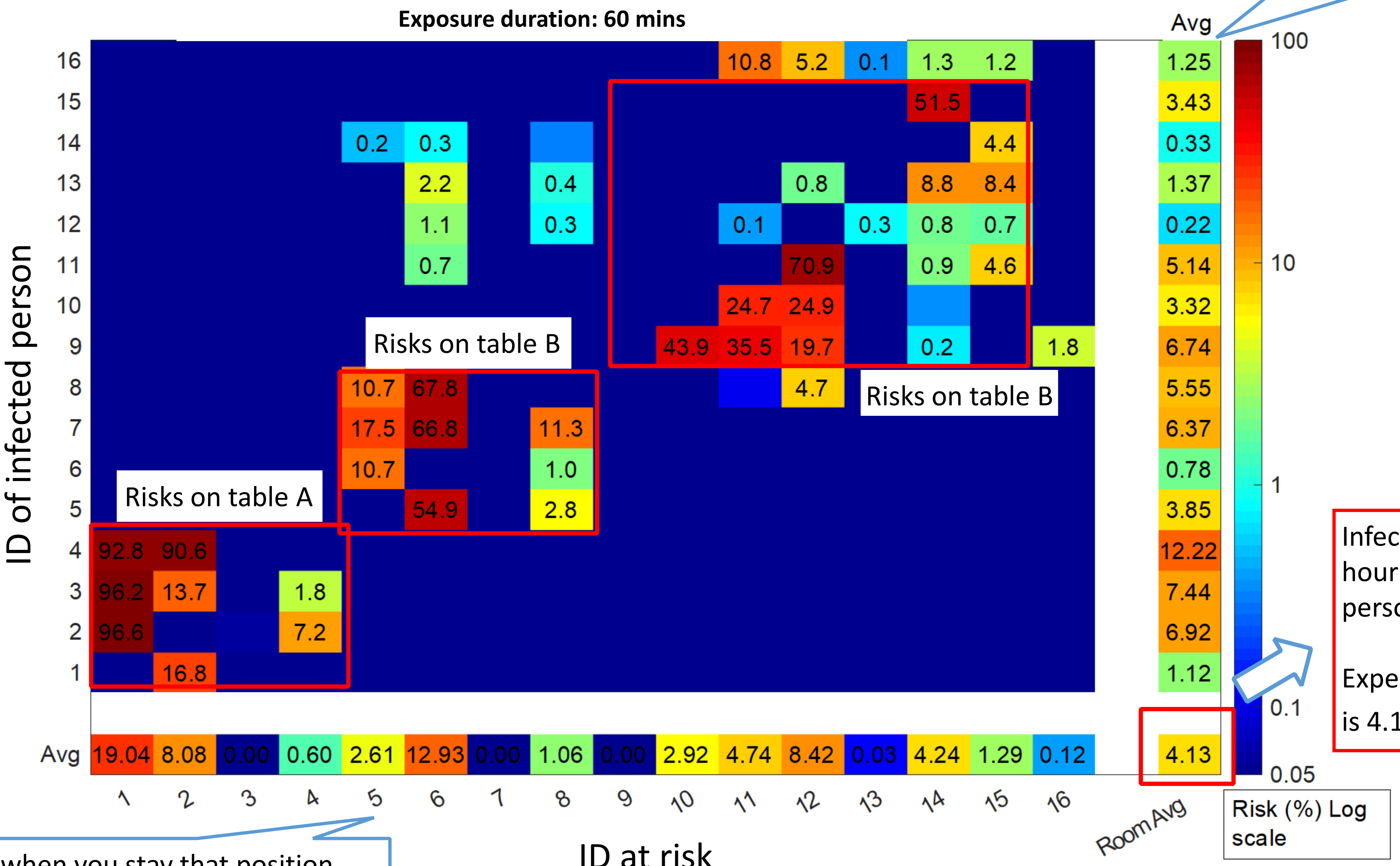
Risks at a Izakaya Pub.

- One person infected.
- Indexing each droplet emitted from each person in the room.
- Counting total droplets reaching from each person to each person for one hour.

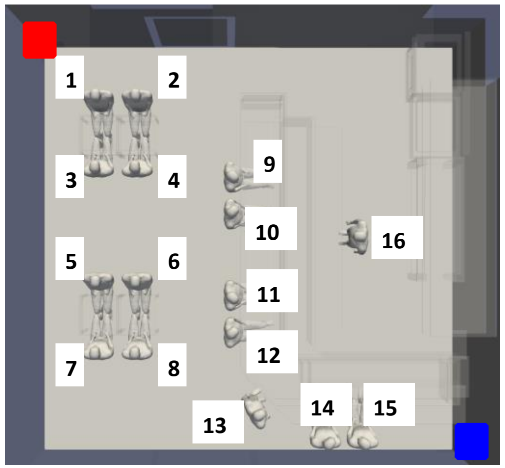


Infection probability for one-hour stay

Risks when an infected person stay that position.



Risk (%) Log scale



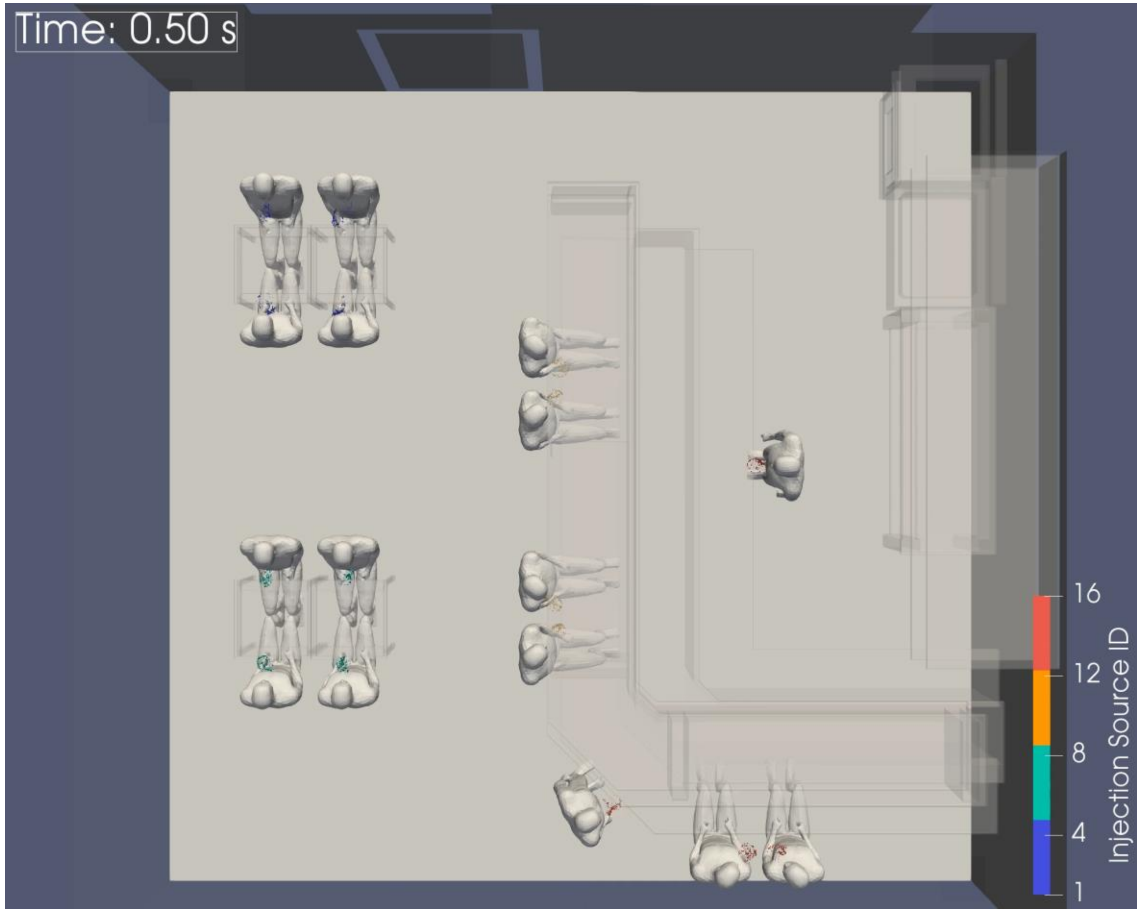
Infection probability when you stay one hour with one unrecognized infected person.

Expected new infected person in this room is 4.13% times 15 = **0.62 person**

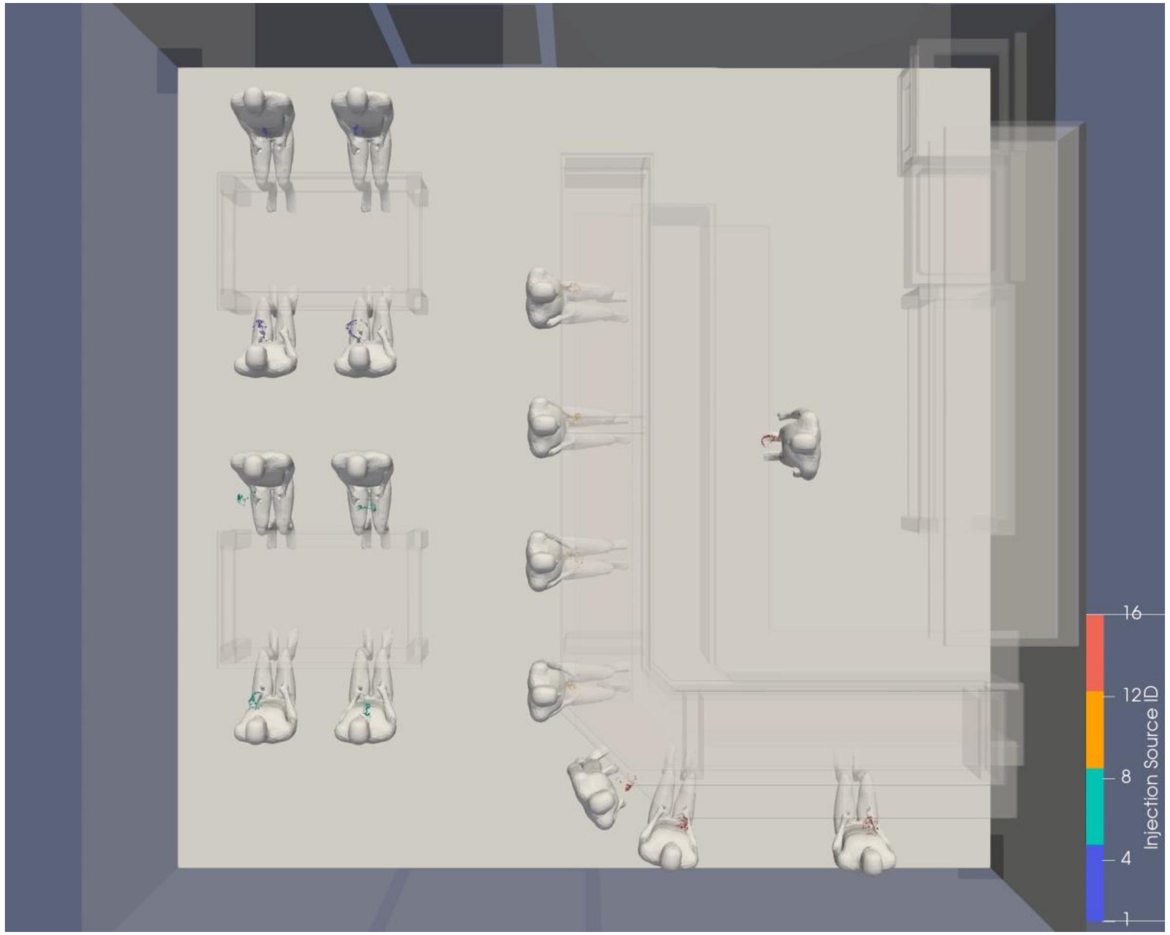
Risks when you stay that position.

Effect of far seating

Close seating

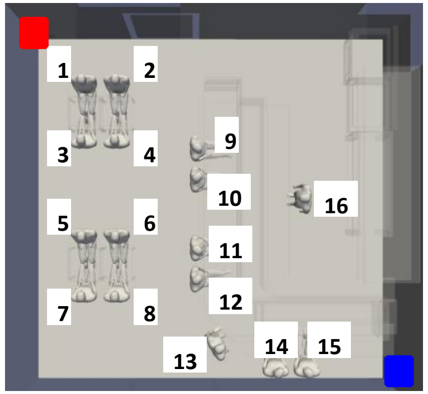


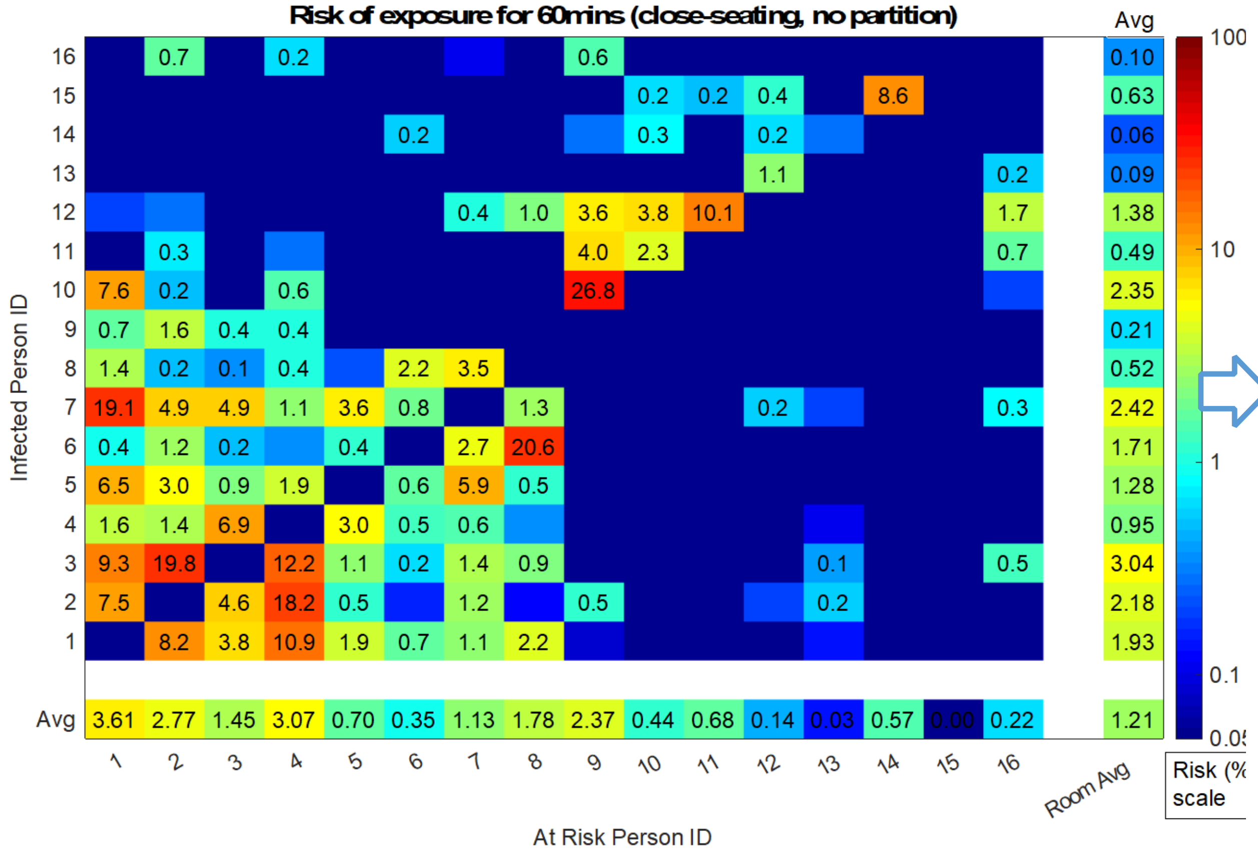
Far seating

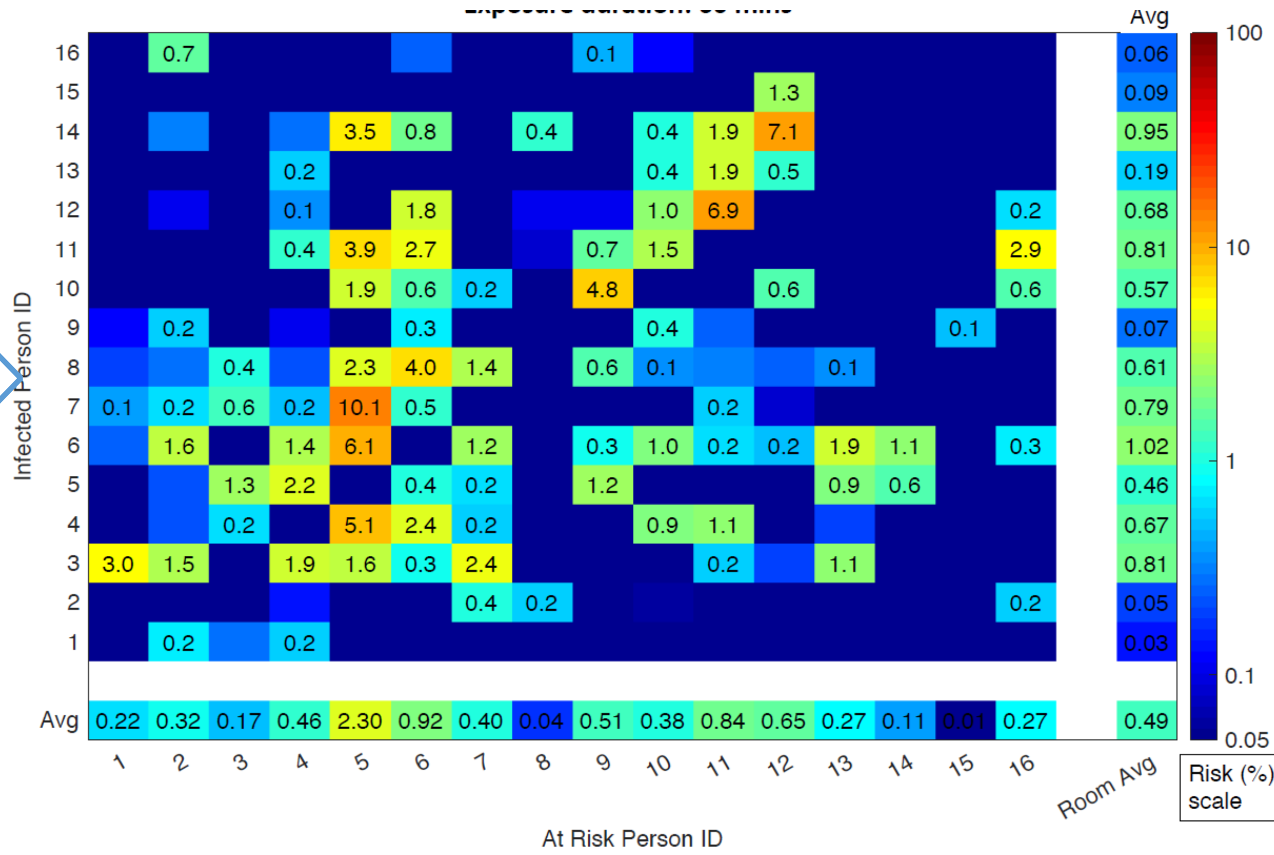


Effect of far seating

- **Expected new infected person reduced from 0.62 to 0.18**





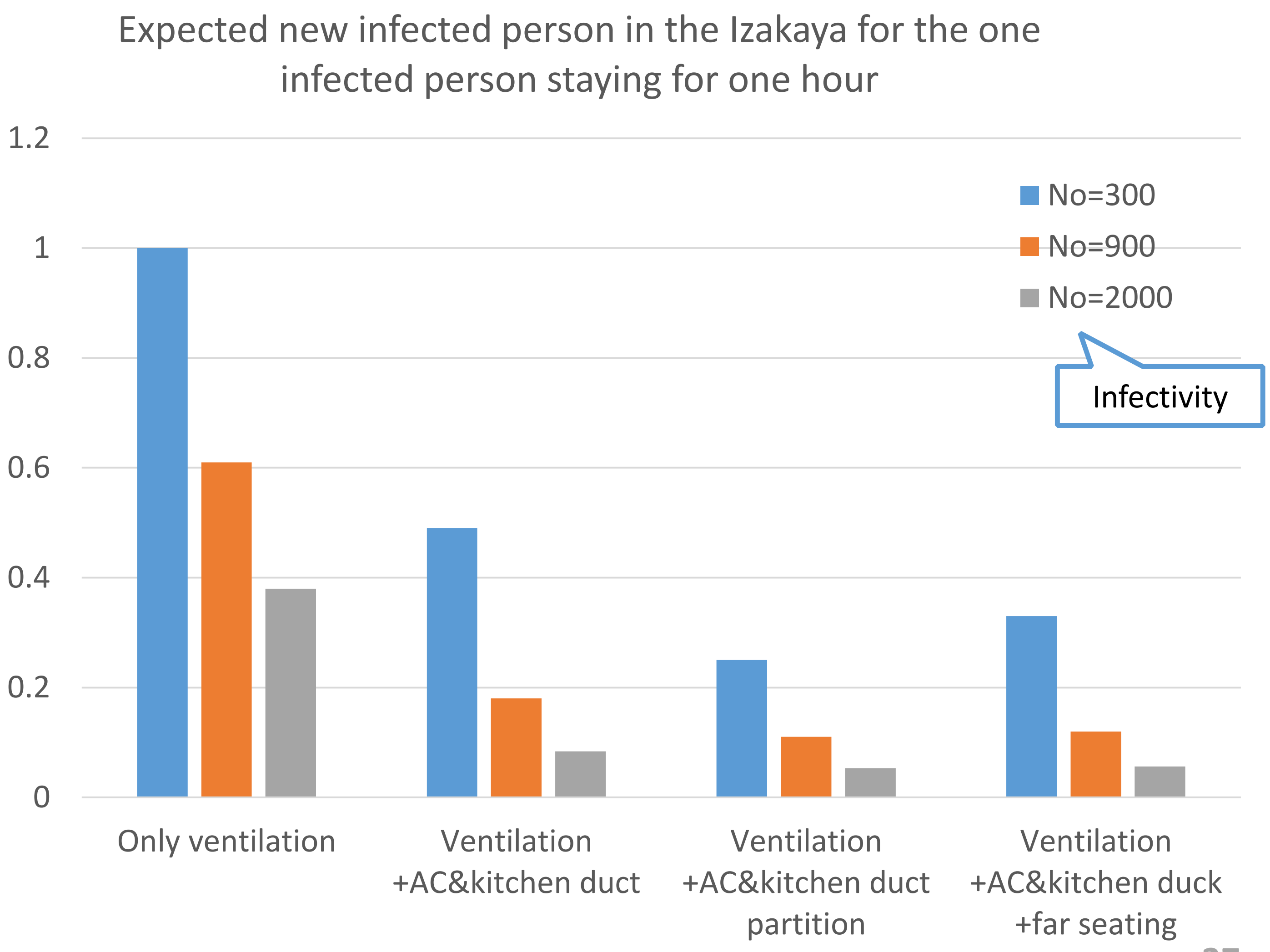


Expected new infected person

- **Effect of countermeasures by expected new infected person**

Small room

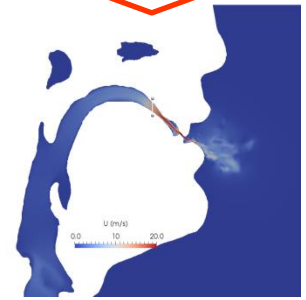
	No=300	No=900	No=2000
Only Ventilation	1.00	0.61	0.38
Ventilation +AC&Kitchen duct	0.49	0.18	0.0844
Ventilation +AC&Kitchen duct +partition	0.25	0.11	0.053
Ventilation +AC&Kitchen duct +far seating	0.33	0.12	0.056

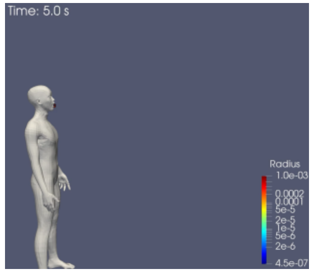


Integrated droplet/aerosol infection risk assessment system

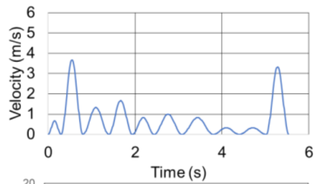
Generation of droplet/aerosol inside human body

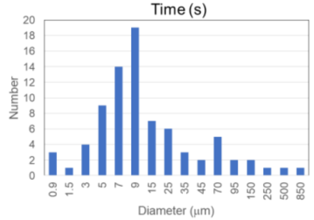
Condition of droplet/aerosol generation (breathing, speaking, coughing, sneezing…)





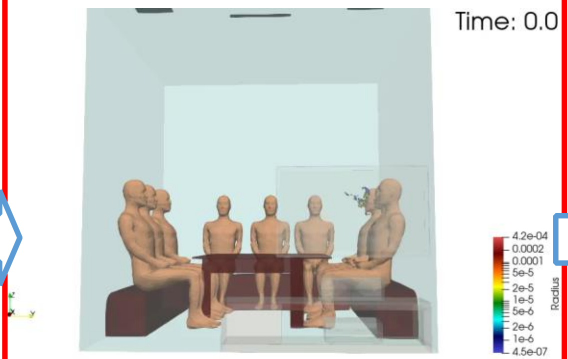
Breath flow rate, droplet size distribution



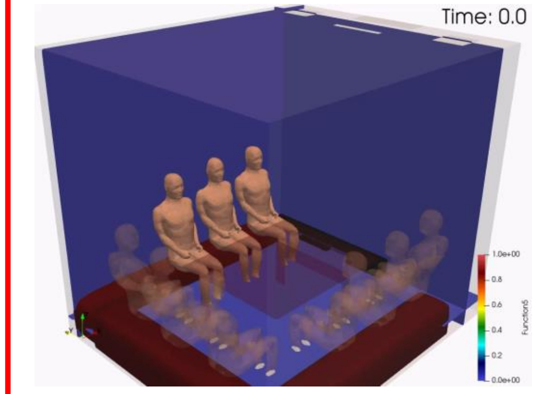


Droplet/aerosol dispersion in indoor environments

Indoor environment and human allocation



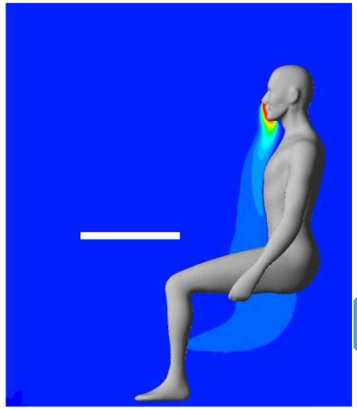
Coupling simulation of droplet/aerosol and indoor flow



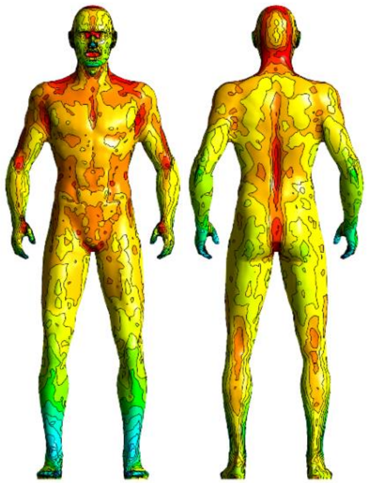
Indoor environment evaluation based on HPC simulation

Numerical human body

Biological information of an at-risk person



Precise reproduction of human breathing

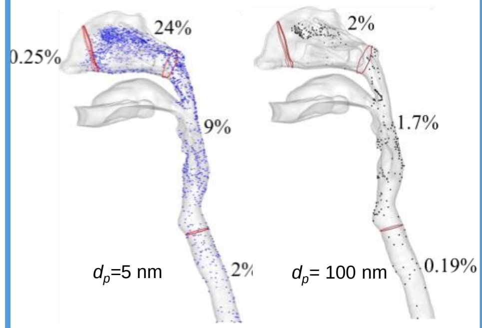


Precise reproduction of body temperature

Numerical respiratory tract

Biological information of an at-risk person

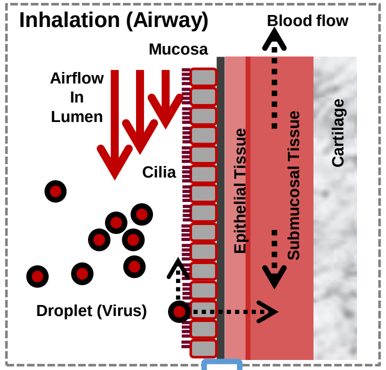
Reproduction of nasal/oral cavity and respiratory tract



Prediction of deposit distribution of droplet/aerosol on the airway surface and its dependence of droplet size.

Infection risk assessment based on the bio-regulation model

Biological information of an at-risk person and target virus



Bioregulation

(Host cells, Pathogen, Adaptive Immune System)

$$\frac{dT_T}{dt} = -\beta_T T_T V - \phi F T_T + \xi R \frac{dR}{dt}$$ (Target Cells)

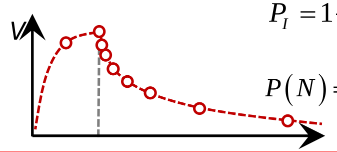
$$\frac{dI}{dt} = \beta_T T_T V - \kappa_F I F - \kappa_E I T_C - \delta_X I$$ (Infected Cells)

$$\frac{dV}{dt} = \beta_E I - \delta_V V - \kappa_V V A$$ (Virus)

$$\frac{dF}{dt} = \beta_F I - \kappa_A F$$ (Interferon)

$$\frac{dT_H}{dt} = \left[\frac{\pi_{H2} D_M}{\pi_{H2} + D_M}\right](1 - T_H / K_H) - \left[\frac{\delta_{H2} D_M}{\delta_{H2} + D_M}\right] T_H$$ (Helper T Cells)

Quantitative evaluation of infection risk



$$P_I = 1 - \exp(-\frac{Iqpt}{Q})$$

$$P(N) = 1 - \exp(-\frac{N}{N_0})$$

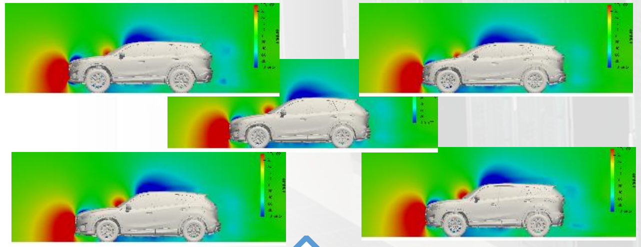
"Smart Design" in the Society 5.0 Era

R-CCS's Trinity Researches for the Advanced Use of HPC

Computational Science | Computer Science | Data Science

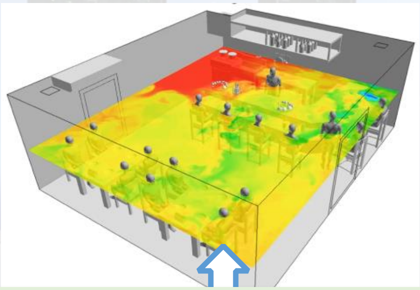
Supporting the software (CUBE, FrontFlow/red) and data science technology (AI, Data assimilation) usage/tuning on the supercomputer "Fugaku"

Sub-Task A (Kobe Univ.)
AI supported Vehicle Aerodynamics Optimization Considering Stylists' Design Space

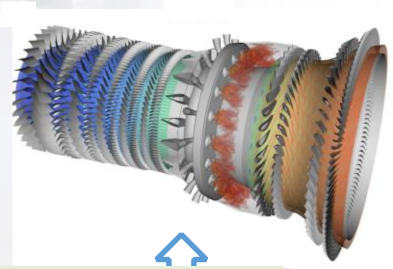


Sub-Task B (Tokyo Tech.)
Performance Design of Transforming Cities under Natural Disturbance

Sub-Task C (Kyushu Univ.)
Indoor-Environment Design Robust for the Infectious Diseases



Sub-Task D (Kyoto Univ.)
Carbon-Free Gas Turbine Engine Design by the Multi-Component Unified Simulation



Toward the social Implementation through the tight collaboration and system development with industries

Academia-Industry-Government Collaboration

Steering members

理化学研究所 RIKEN
R-CCS
神戸大学 KOBE UNIVERSITY
国立大学法人 豊橋技術科学大学
京都工芸繊維大学 KYOTO INSTITUTE OF TECHNOLOGY
大阪大学 OSAKA UNIVERSITY
東京工業大学 Tokyo Institute of Technology
九州大学
鹿島 KAJIMA CORPORATION
DAIKIN
National Univ. of Singapore

Cooperative members

TOYOTA
JAPAN AIRLINES
大王製紙株式会社
ZEN-ON ZEN-ON Music Company LTD Since 1931
TOYOTA CUSTOMIZING & DEVELOPMENT
FUSO
SUNTORY
TOPPAN

Administrative organizations

文部科学省 MINISTRY OF EDUCATION, CULTURE, SPORTS, SCIENCE AND TECHNOLOGY-JAPAN
国土交通省 Ministry of Land, Infrastructure, Transport and Tourism
内閣府 Cabinet Office
Kobe City

Timey Simulations and Media Dissemination

- We have been staging multiple press conferences on the latest research results
- Extremely high interest from the media, with immediate national news coverage
- Most people in Japan have seen the Fugaku COVID19 news, esp. droplet simulation, with high trust in being scientifically grounded
- Visualization extremely effective in raising public understanding & awareness of COVID19 & its mitigation
- Prime Minister Suga holds a press conference 22 Nov., urging everyone to wear masks even during group dining, as "it's effectiveness has been proven by a supercomputer (Fugaku)".

Performance projection of many-core CPU systems based on IRDS roadmap

Predictions based on the IRDS Roadmap(2020 ed.), extrapolation of traditional many core architectures relying merely on advances of semiconductor technologies will achieve only 1.8EFLOPS Peak (3.37x c.f. Fugaku), if a machine with broad applicability will be built

- Methodoogies(CPU part): Assumptions from IRDS Roadmap Systems and Architectures
  - Cores/socket=70 cores
  - SIMD width=2048-bit x 2
  - Clock frequency=3.9GHz
  - Socket TDP = 351W
- System assumptions
  - System Power=30, 40, 50MW
  - PUE=1.1
  - CPU power occupy=60,70,80%

https://sites.google.com/view/ngaci/home


White Paper on Next-Generation Advanced Computing Infrastructure
(次世代先端的計算基盤に関する白書)

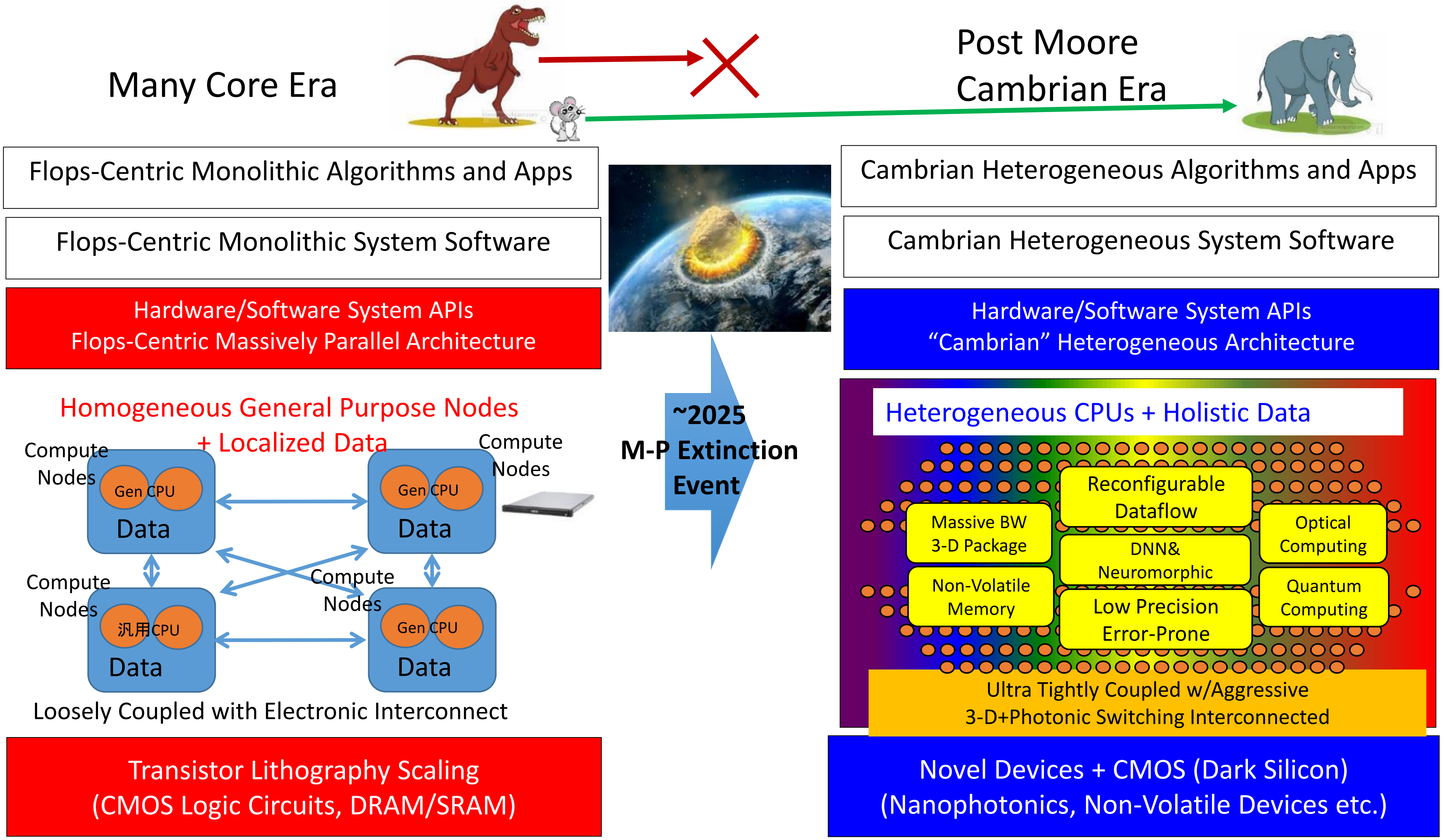
2020年11月

NGACI white paper

From NGACI white paper

	30MW			40MW			50MW		
	60%	70%	80%	60%	70%	80%	60%	70%	80%
Socket Cores	46620	54390	62160	62160	72520	82880	77700	90650	103600
	3.3×10^6	3.8×10^6	4.4×10^6	4.4×10^6	5.1×10^6	5.8×10^6	5.4×10^6	6.3×10^6	7.3×10^6
	815	950	1086	1086	1267	1448	1358	1584	1810
DDR 総 BW (PB/s)	102	120	137	137	160	182	171	200	228
HBM 総 BW (PB/s)	307	358	410	410	478	547	512	598	683
DDR 総容量 (PB)	17	20	23	23	27	31	29	34	39
HBM 総容量 (PB)	4	5	5	5	6	7	7	8	9
Injection BW(Tb/s)	1.6	1.6	1.6	1.6	1.6	1.6	1.6	1.6	1.6
総 I/O 性能 (TB/s)	34	34	34	34	34	34	34	34	34
Storage (EBytes)	3.45	3.45	3.45	3.45	3.45	3.45	3.45	3.45	3.45

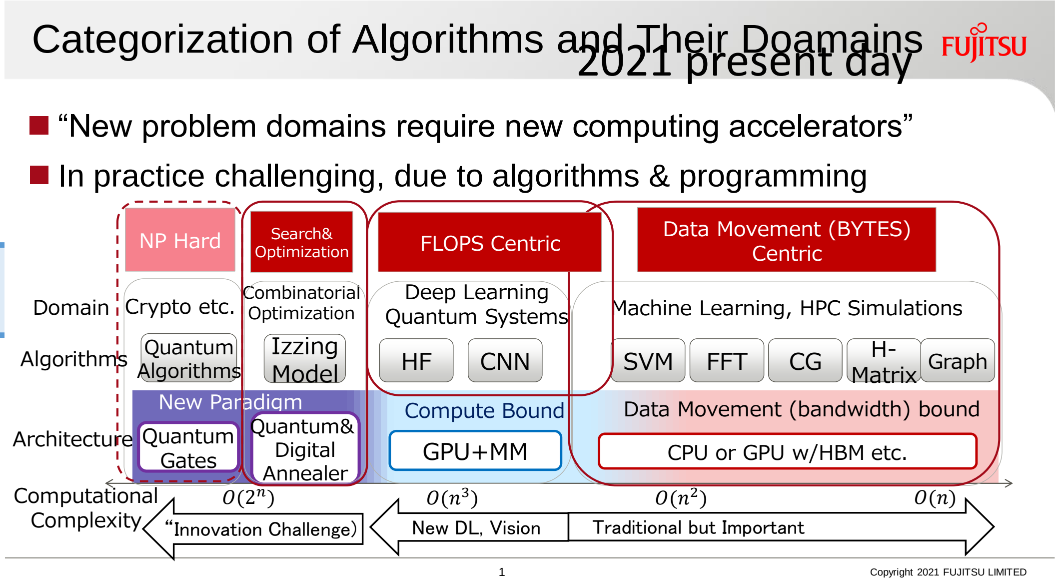
最もアグレッシブなシステム構成（50MW電力バジェット、CPUで80%電力消費）においても1.8EF程度の性能と予測

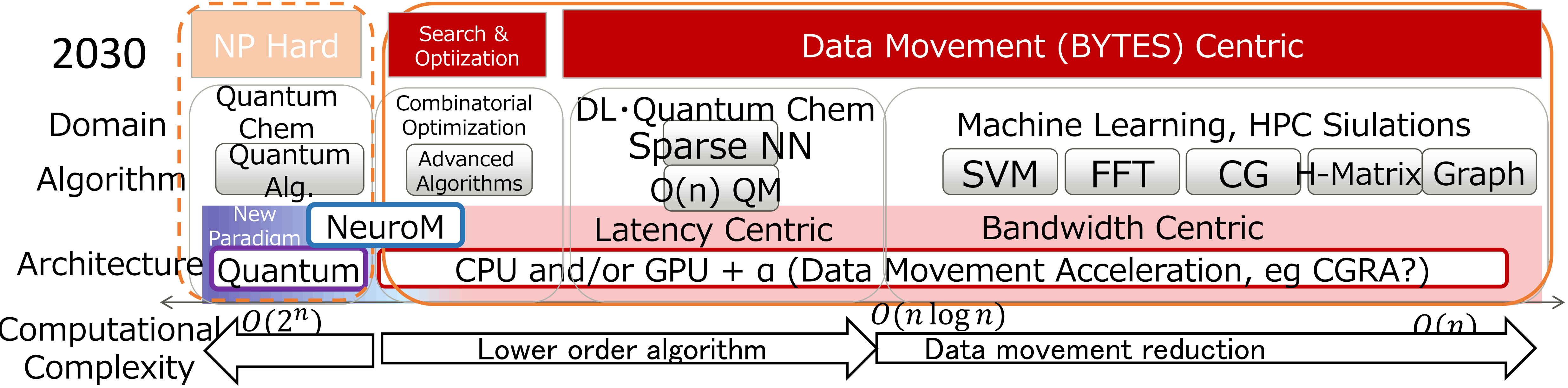

Many Core Era
Post Moore
Cambrian Era
Flops-Centric Monolithic Algorithms and Apps
Flops-Centric Monolithic System Software
Hardware/Software System APIs
Flops-Centric Massively Parallel Architecture
Homogeneous General Purpose Nodes
+ Localized Data
Compute Nodes
Gen CPU
Data
Compute Nodes
Gen CPU
Data
Compute Nodes
汎用CPU
Data
Compute Nodes
Gen CPU
Data
Loosely Coupled with Electronic Interconnect
Transistor Lithography Scaling
(CMOS Logic Circuits, DRAM/SRAM)
~2025
M-P Extinction
Event
Cambrian Heterogeneous Algorithms and Apps
Cambrian Heterogeneous System Software
Hardware/Software System APIs
"Cambrian" Heterogeneous Architecture
Heterogeneous CPUs + Holistic Data
Reconfigurable Dataflow
Massive BW 3-D Package
Optical Computing
DNN& Neuromorphic
Non-Volatile Memory
Quantum Computing
Low Precision Error-Prone
Ultra Tightly Coupled w/Aggressive
3-D+Photonic Switching Interconnected
Novel Devices + CMOS (Dark Silicon)
(Nanophotonics, Non-Volatile Devices etc.)

Post-Moore Algorithmic Development

- **Towards 2030 Post-Moore era**

- End of ALU compute (FLOPS) advance
- Disrupritve reduction in data movement cost with new devices, packaging
- Algorithm advances to reduce the computational order (+ more reliance on data movement)
- Unification of BD/AI/Simulation towards data-centric view





Our Project: Exploring versatile HPC architecture and system software technologies to achieve 100x performance by 2028

Problems to be solved and goals to be achieved

- General-purpose computer architectures that will accelerate a wide range of applications in the post-Moore era have not yet been established.
- What is a feasible approach for versatile HPC systems based on bandwidth improvement?
- **Goal:** to explore architectures that can achieve 100x performance in a wide range of applications around 2028

Approaches and subtasks

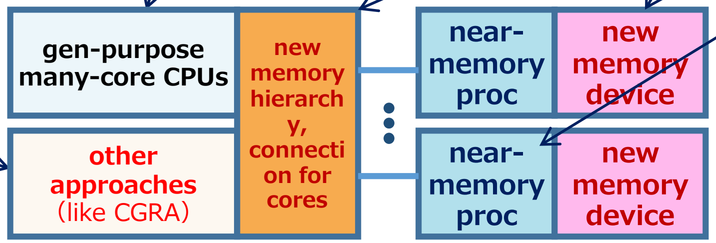
- Exploration of future CPU node architectures and necessary technologies

Subtask1.1 Performance characterization and modeling with benchmarks to identify directions for exploration and improvement (Riken R-CCS)

Subtask2 Exploring innovative memory architectures with ultra-deep and ultra-wide bandwidth (Tokyo Tech.)

Subtask1.2 Exploring a reconfigurable vector data-flow architecture (CGRA) that can exploit increased data transfer capability (Rijen R-CCS)

Subtask3 Exploring near-memory computing for highly effective bandwidth and cooling efficiency for general purpose computing (U-Tokyo)



目標とするノードアーキテクチャの例

Planned

Subtask4 Exploration of node architectures as extension of existing many-core CPUs with non von-Neumann methods (unnamed company)

Plan

2018	2019	2020	2021	2022	2023～

- Explore individual technologies
- Integrate promising technologies for a target node architecture
- Developing stage ...