

# AI橋渡しクラウド（ABCI）と ビッグデータ転送

産業技術総合研究所 高野了成

2019年10月4日 ADVNET2019@東大

# 世界最大級・超省電力・オープンAIインフラストラクチャ



- 経産省「人工知能に関するグローバル研究拠点整備事業」(H28二次補正)の一環として整備
- 我が国における産学官共同によるAI研究開発を加速するオープンイノベーションプラットフォーム
- 高い計算能力を活用した人工知能技術の研究開発・実証、社会実装の推進、AI分野の最重要課題への挑戦が目的

## 理論ピーク性能

550 PFLOPS (半精度)

37 PFLOPS (倍精度)

## 実効性能

TOP500 (HPL): 19.88 PFLOPS

HPCG : 508.85 TFLOPS

GREEN500 : 14.423 GFLOPS/W

使用電力 (最大) : 2.3 MW

年間平均PUE : 1.1以下 (推定値)



2018年8月1日運用開始

# 世界トップクラスの実運用される省電カクラウド型計算システム

Top500 (June 2019) Linpack性能

Rank	System	Rmax
1	Summit	148.6PF
2	Sierra	94.6PF
3	Sunway TaihuLight	93.0PF
4	Tianhe-2A	61.4PF
5	Frontera	23.5PF
6	Piz Daint	21.2PF
7	Trinity	20.2PF
8	ABCI	19.9PF
9	SuperMUC-NG	19.5PF
10	Lassen	18.2PF

HPCG (June 2019) 共役勾配法性能

Rank	System	HPCG
1	Summit	2.9PF
2	Sierra	1.8PF
3	K computer	603TF
4	Trinity	546TF
5	ABCI	509TF
6	Piz Daint	497TF
7	Sunway TaihuLight	481TF
8	Nurion	391TF
9	Oakforest-PACS	385TF
10	Cori	355TF

Green500 (June 2019) 省エネ性能

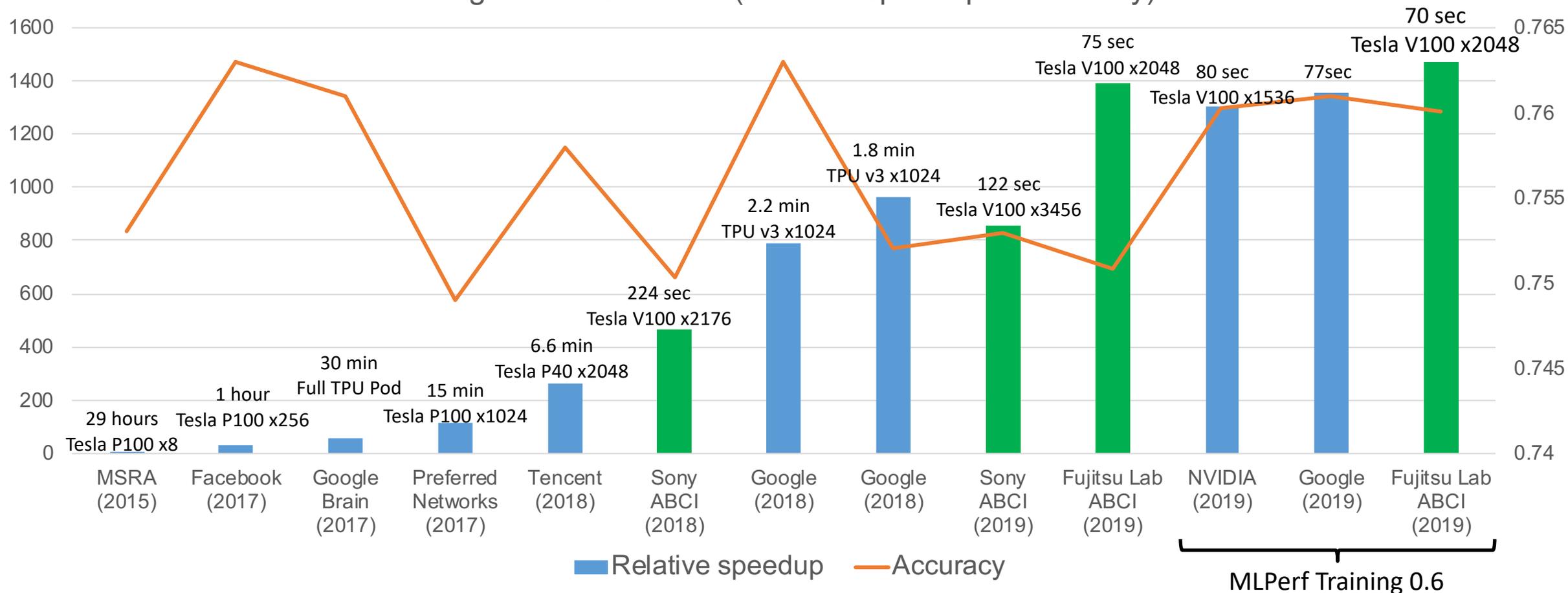
Rank	System	GF/W
1	DGX SaturnV Volta	15.1
2	Summit	14.7
3	ABCI	14.4
4	MareNostrum P9 CTE	14.1
5	TSUBAME3.0	13.7
6	PANGEA III	13.1
7	Sierra	12.7
8	PreE	11.4
9	Taiwania 2	11.3
10	Huawei G5500	10.8



3指標でトップ10に入るのは現在ABCI、米国DOEのSummit、Sierraのみ  
 実性能と省エネ性能を世界トップレベルで両立  
 公的かつオープンで、AI技術開発に特化した計算インフラとして世界初の先進的システム

# ディープラーニングの学習速度の世界記録を大幅に更新

ImageNet / ResNet-50 (Relative speedup & Accuracy)



今年7月、富士通研究所がABCI上で記録した70秒が現在の世界記録  
 ABCIにより、Google、Tencentなど先行する技術開発への対抗が可能に

# ABCI のハードウェア構成

## 高性能計算システム

550 PFlops (FP16), 37.2 PFlops (FP64)  
476 TiB メモリ, 1.74 PB NVMe SSD



### 計算ノード (GPU搭載) x 1,088台

- GPU** NVIDIA Tesla V100 SXM2 x 4 (計 4,352基)
- CPU** Intel Xeon Gold 6148 x 2 (計 2,176基)
- Memory** 384GiB
- Local Storage** Intel SSD DC P4600 (NVMe) 1.6TB x 1
- Interconnect** InfiniBand EDR x 2

### マルチプラットフォームノード (GPU非搭載) x 10台

- Intel Xeon Gold 6132 (2.6GHz/14cores) x 2
- 768GiB Memory, 3.8TB NVMe SSD

インタラクティブノード x 4

管理サーバ・ゲートウェイノード x 15

## 計算ネットワーク (Infiniband EDR)

- Mellanox CS7500 x 2
- Mellanox SB7890 x 229

## サービスネットワーク (10GbE)

## 大容量ストレージシステム

22 PB GPFS + 1 PB フラッシュストレージ +  
17 PB クラウドストレージ



### DDN SFA14K (w/ SS8462 Enclosure x 10) x 3

- 12TB 7.2Krpm NL-SAS HDD x 2400
- 3.84TB SAS SSD x 216
- NSD Server x 12

### DDN S

- 7.68TB
- MDS S

### HPE A

- 12TB
- 3.2TB

### ■ポイント1: 世界最大級の最新GPUファーム

全部で4352基。誰でも2048基までの最新GPUを使って大規模・高速なディープラーニングを実行可能。

### ■ポイント2: スパコン由来のネットワーク技術の活用

すべてのGPUとCPUが2.5μ秒以下の低遅延、200Gbpsの広帯域で通信可能。  
(商用クラウドでは数10倍の遅延、1/2以下の帯域)

FortiAnalyzer 400E x1

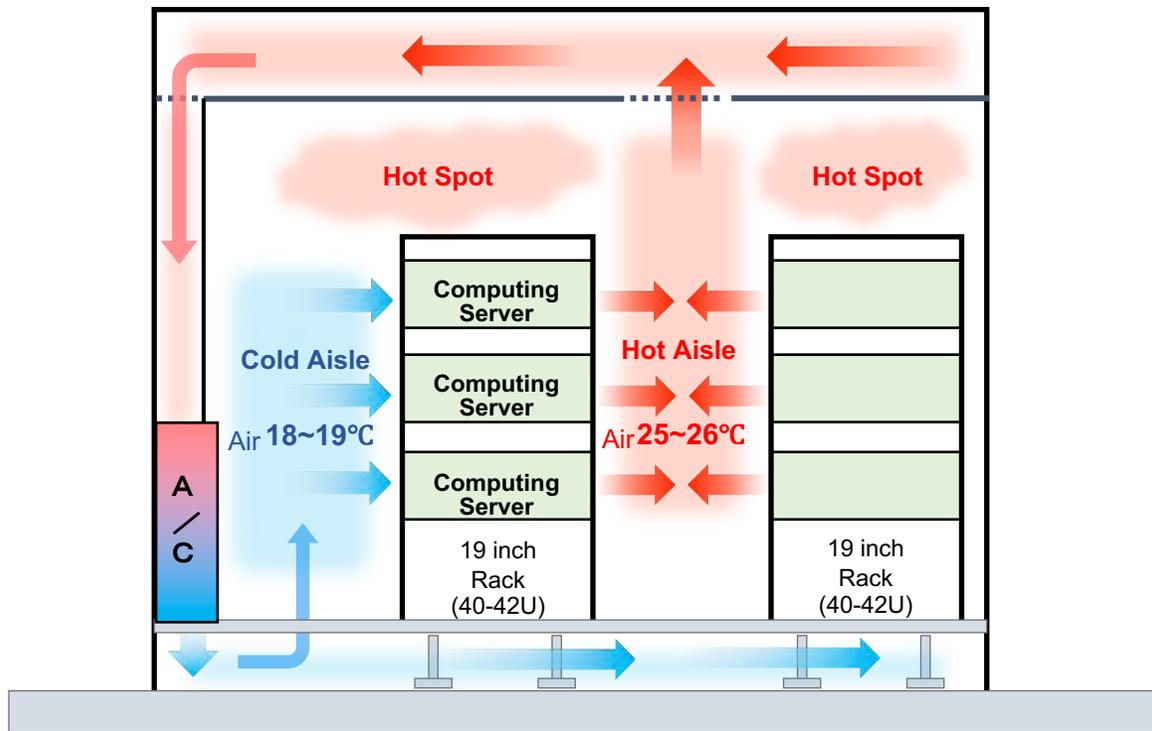
SINET5

# ハイブリッド冷却システム

## ■ 従来型データセンターの冷却システム

「床吹き出し方式」による全体空調のモデル  
 近年は、冷却装置をホットアイルの上部に設置するなどの方式で、冷熱混合を低減させることで冷却の効率を上げている

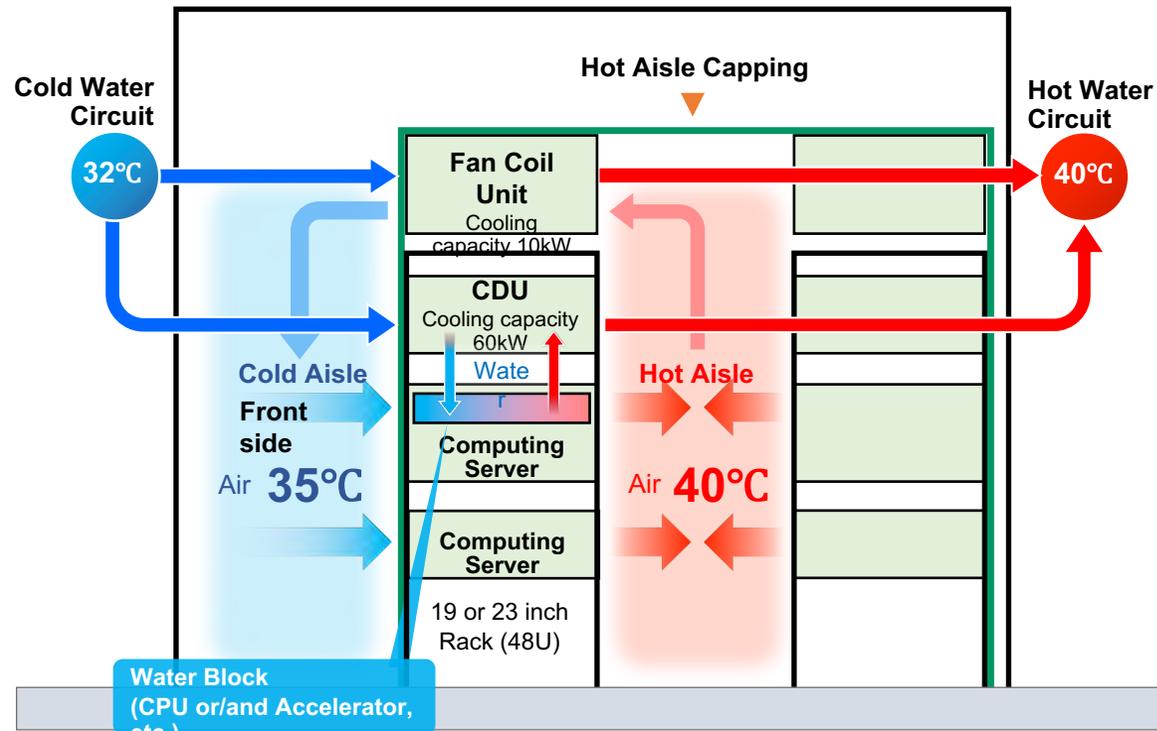
年間平均PUE：1.7~2.0



## ■ A B C I の冷却システム

「水冷+空冷のハイブリッド冷却方式」  
 外気を使ったフリークーリングにより「32°Cの冷却水」を作り、その水でラック内冷却水（2次冷却水）と、ホットアイル内の空気を冷却する

年間平均PUE：1.1 以下



※PUE (Power Usage Effectiveness) とは、右の計算式で定義されるデータセンターにおける電力効率を示す指標で、1.0に近いほど高効率。

$$PUE = \frac{\text{データセンター全体の消費電力}}{\text{IT機器による消費電力}}$$

# ABCIを用いたビッグデータ転送事例

- 大規模動画画像解析
- Global Research Platform
- データ活用社会創成プラットフォーム
- ABCIクラウドストレージ

# ABCIを活用した大規模動画画像解析

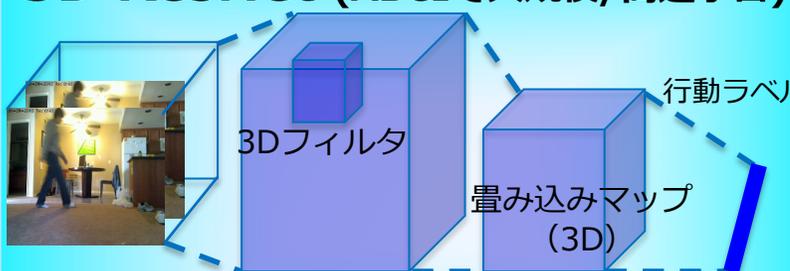
## ABCI

AI Bridging Cloud Infrastructure

<https://abci.ai/ja/>

産総研開発、世界トップレベルの  
計算処理能力/データ処理能力

### 3D ResNet (ABCIで大規模/高速学習)



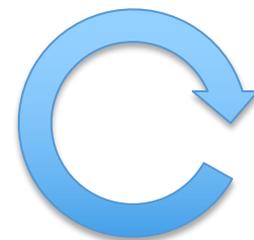
行動ラベル

畳み込みマップ (3D)

産総研開発 動画画像解析 深層学習  
構造、世界的に用いられている

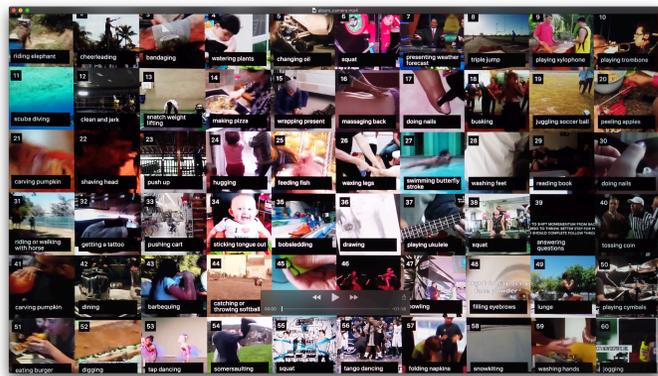
大規模AI基盤であるABCIを活用し大量の動画画像を同時にリアルタイムで解析する技術を開発

ABCIを使って動画画像を事前に学習

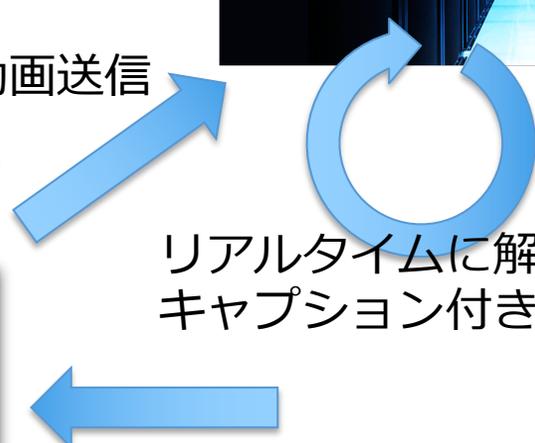


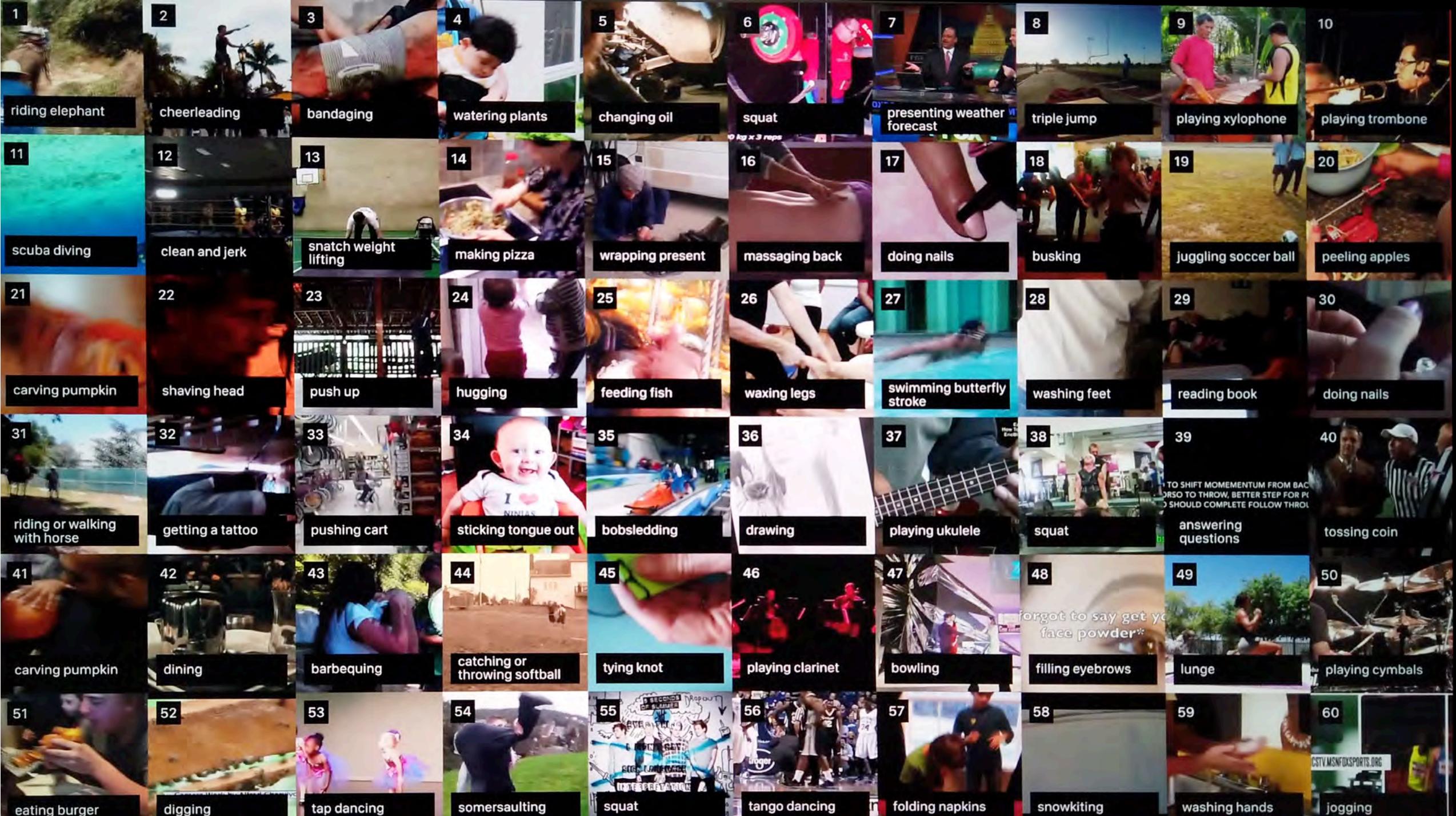
動画送信

東京デモ会場とABCIをSINET5で接続  
動画画像をリアルタイム解析



リアルタイムに解析して、  
キャプション付き動画生成





1 riding elephant

2 cheerleading

3 bandaging

4 watering plants

5 changing oil

6 squat

7 presenting weather forecast

8 triple jump

9 playing xylophone

10 playing trombone

11 scuba diving

12 clean and jerk

13 snatch weight lifting

14 making pizza

15 wrapping present

16 massaging back

17 doing nails

18 busking

19 juggling soccer ball

20 peeling apples

21 carving pumpkin

22 shaving head

23 push up

24 hugging

25 feeding fish

26 waxing legs

27 swimming butterfly stroke

28 washing feet

29 reading book

30 doing nails

31 riding or walking with horse

32 getting a tattoo

33 pushing cart

34 sticking tongue out

35 bobsledding

36 drawing

37 playing ukulele

38 squat

39 answering questions

40 tossing coin

41 carving pumpkin

42 dining

43 barbequing

44 catching or throwing softball

45 tying knot

46 playing clarinet

47 bowling

48 filling eyebrows

49 lunge

50 playing cymbals

51 eating burger

52 digging

53 tap dancing

54 somersaulting

55 squat

56 tango dancing

57 folding napkins

58 snowkiting

59 washing hands

60 jogging

# Global Research Platform実現に向けた動向

- 2010年頃、米国ESnetが大量の科学実験データを転送するためのネットワーク Science DMZ を提案
- 2015年、カリフォルニア州のScience DMZをつないだPacific Research Platform (PRP)プロジェクト開始
  - FIONA：データ転送に最適化したコスト効率性の高い計算機
  - Nautilus：Kubernetesベースの広域クラスタシステム
  - ユーザ事例：IceCube
- 2018年、National Research Platform、Asia Pacific Research Platformなどが登場
- 2019年9月、1st Global Research Platform workshop@UCSD

# AIST's Machine Learning Supercomputer Has Been Connected to PRP/CHASE-CI at 6.4 Gbps!

Larry Smarr@UCSD  
GRP Workshop

SDSC CTO Phil Papadopoulos and Satoshi Sekiguchi sign 5-Year MOU on behalf of UC San Diego and AIST.



January 10, 2018

"The two sides will collaborate on cyberinfrastructure projects, notably PRP/CHASE-CI and AIST's AI Bridging Cloud Infrastructure (ABCI)."

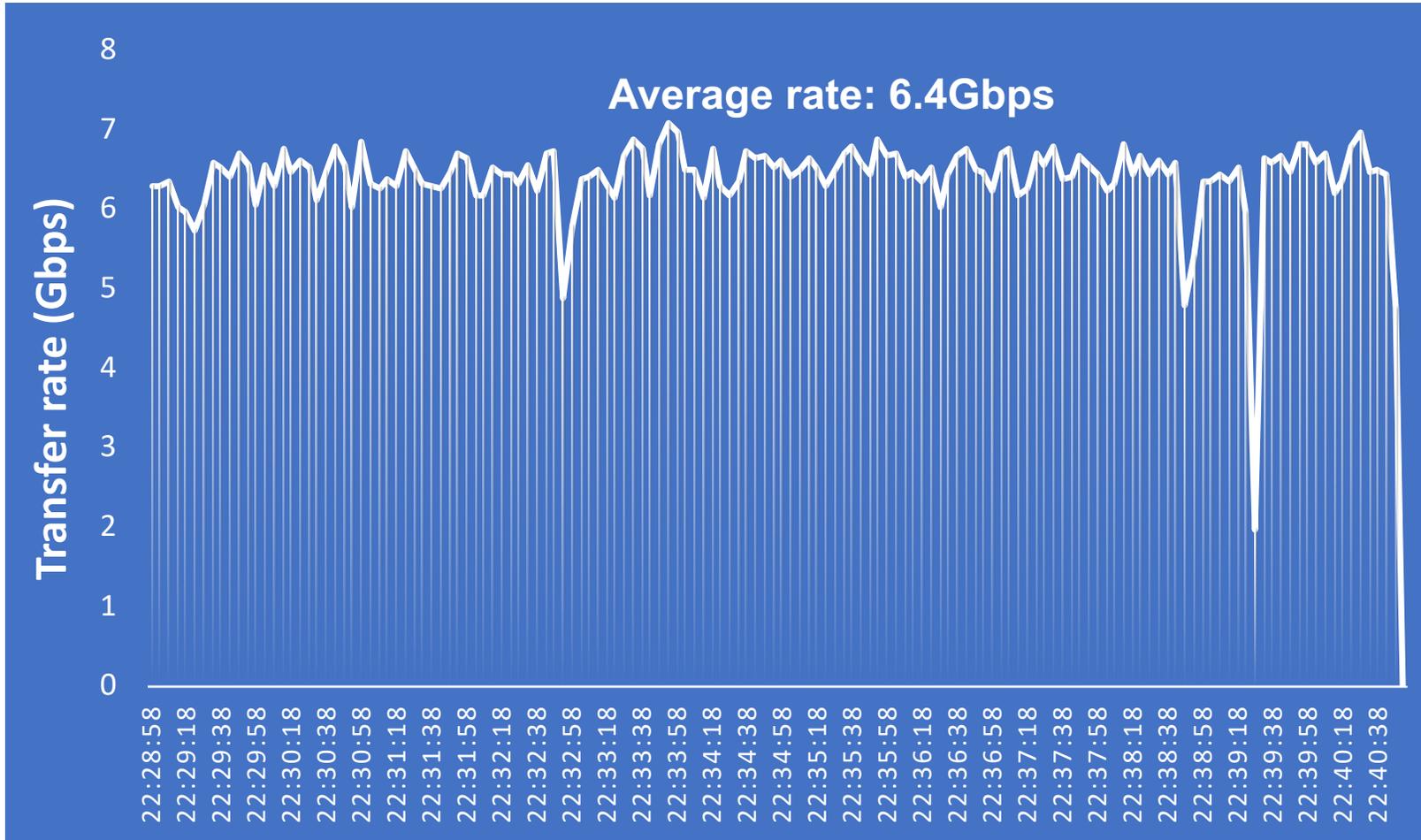


Source: Ryousei Takano, AIST

- **ABCI Has 4352 NVIDIA Tesla V100 GPUs**
  - Peak Performance: 550 PFlops (FP16)
  - Linpack: 20 PFlops (#8 in TOP500)
- **ABCI-PRP: Grand Challenge Project**
  - Gerald Pao, Salk Institute
  - George Sugihara, SIO
  - Creation of Neuromorphic Deep-Learning Architectures by Large-Scale Dynamic Modeling of Transparent Fish Brains on ABCI
  - Disk-to-Disk data transfer from UCSD to ABCI
    - FDT version 0.26.1
    - 574 GB Transferred in 12 Minutes
    - 6.4 Gbs



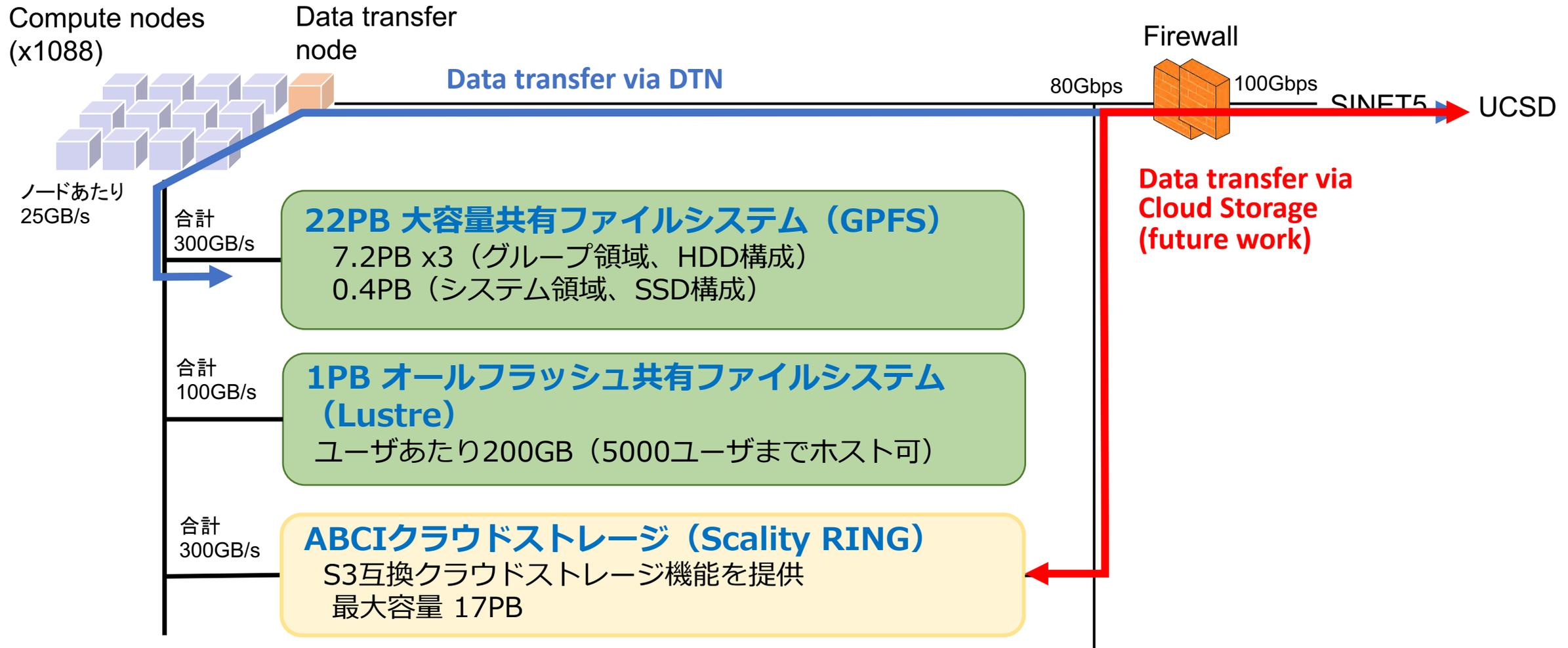
# Disk-to-Disk data transfer from UCSD to ABCI



- FDT version 0.26.1
- Total data size: 574GB
- File size: 193MB
- Number of files: 2977
- Elapsed time: 12m10s
- Average rate: 6.4Gbps

\*) Zebrafish brain image raw data

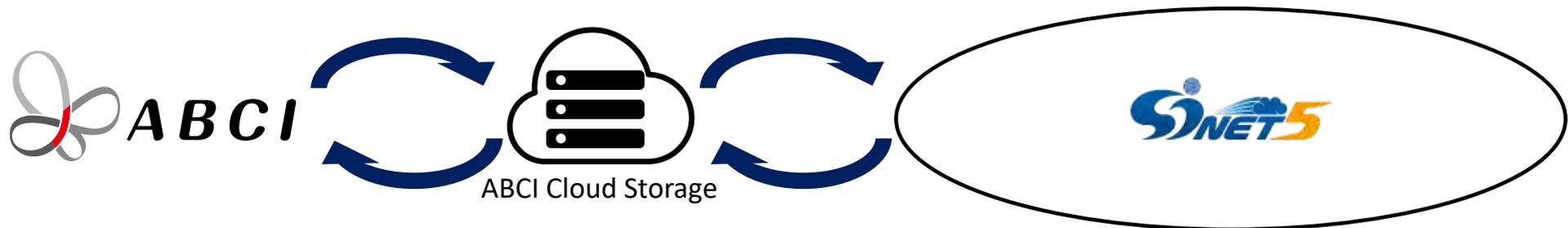
# ABCIストレージ構成とData Transfer Node



# ABCIクラウドストレージの主な特徴

## ■ SINET5に直結した「データハーバー」

- SINET5に直結した各機関から、高速・安全にデータを収集・蓄積
- ABCI上で生成された、高性能な汎用学習モデル等の共有・配布



## ■ Amazon S3互換API

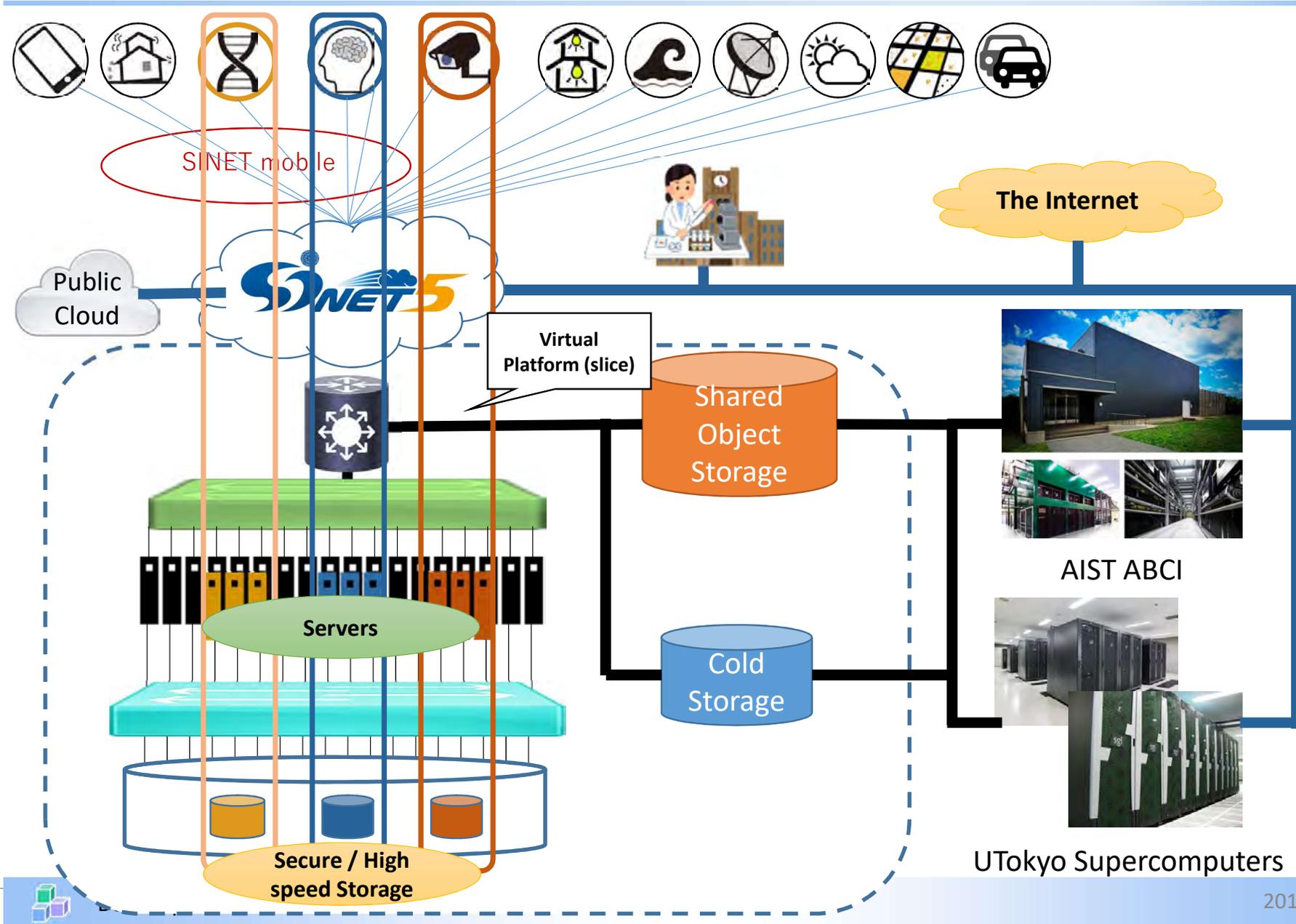
- クラウドストレージのデファクトAPI、そのソフトウェアエコシステムの活用
  - Oracle Cloud, IBM Cloud、国内でもIIJ、さくらインターネットのサービスなどで採用
  - S3向けの様々なソフトウェアやサービスをそのままABCIにポート可能
- 共有ファイルシステムと異なり、バケット単位での細粒度のアクセス制御を実現

## ■ エンタープライズクラスの通信・データ暗号化

- ABCIクラウドストレージへのアクセスは、ABCI内部・外部問わず、SSLにより暗号化
- S3準拠のサーバサイド暗号化（SSE）に対応、保管時にデータ暗号化、読出し時にデータ復号が可能

# データ活用社会創成プラットフォーム

1. Society 5.0における「ビッグデータ等の新たな技術をあらゆる産業や社会生活に取り入れてイノベーションを創出」するための「基盤」
  - ステークホルダー間のマッチングハブ機能を提供
  - ターゲットはBig ScienceではなくEveryday Application
  - 地方の中小企業や自治体のデータ利活用の実証実験の場
2. 自ら設備を持たなくとも誰もがデータ利活用
  - SINET5、モバイルSINET網、東大スパコン、ABCIを活用
3. 応用ごとに必要な機能を用途に応じて利用。オンデマンドで短時間に構築・拡張・融合可能
  - ユーザごとに隔離されたスライスを提供



# まとめに代えて

- スパコンがスタンドアローンだったのは今は昔
- 大容量ネットワークの整備は必須に
- ネットワークの安定運用に日々携わってくださる皆様に感謝！

ご清聴ありがとうございました。

詳しくはWebで！  
<https://abci.ai/>

