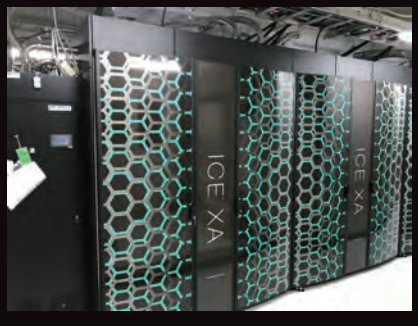
東京工業大学　学術国際情報センター

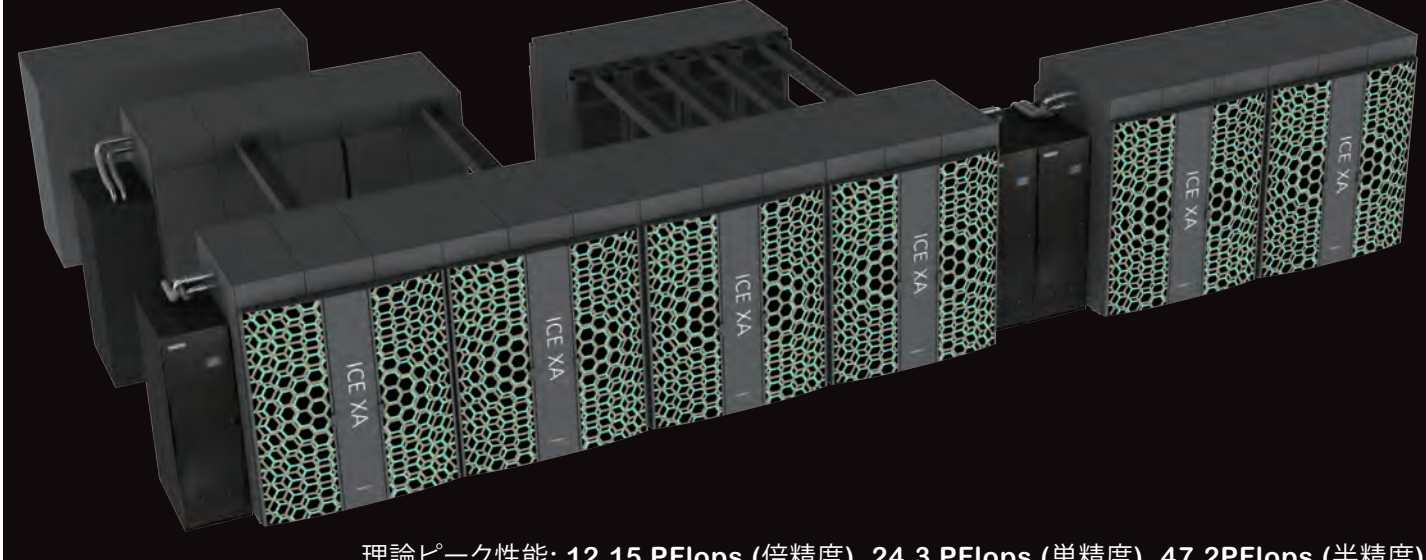
HARDWARE SOFTWARE SPECIFICATIONS



TSUBAME3.0

ハードウェアおよびソフトウェアの仕様

- **GPU**搭載の高性能計算ノード
- 高速なネットワークインターコネクト
- 高性能・高信頼ストレージシステム
- 高い電力効率を実現する冷却システム
- システムおよびアプリケーションのソフトウェア



理論ピーク性能: **12.15 PFlops** (倍精度), **24.3 PFlops** (単精度), **47.2PFlops** (半精度)
フルバイセクションバンド幅のファットツリー型Omni-Pathネットワーク
容量**15.9PB**, アクセス速度**150GB/s**の**Lustre**ファイルシステム高速ストレージ
冷却塔を利用する温水冷却

GPU搭載高性能計算ノード

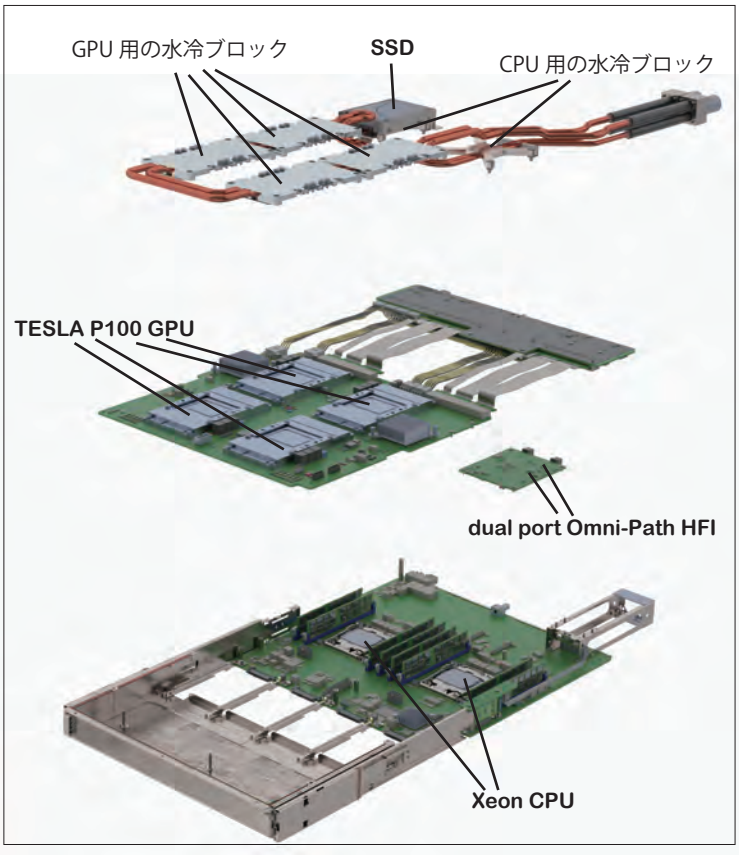
TSUBAME3.0システムは**540**台の計算ノードを備え, 総計12.15ペタフロップスの演算性能を供給する. 各計算ノードは2基のCPUと4基のGPUを極めてコンパクトに設計のブレードに搭載する. さらに4本のOmni-Pathインターフェイスや高速・大容量のSSDなどを搭載し, ビッグデータや人工知能などの分野の計算にも対応する.

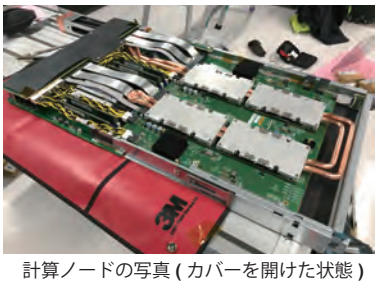
HPE SGI ICE-XA (SGI 8600) IP139-SXM2 540 ノード

CPU: **Intel Xeon E5-2680 V4 (Broadwell-EP, 2.4GHz) ×2** ソケット
ソケットあたり **14** コア, ノードあたり合計 **28** コア.
GPU: **NVIDIA TESLA P100 for NVlink-Optimized servers ×4** 基 .
メモリ : **256GB (DDR4-2400 32GB** モジュール **×8** 本 **)**
SSD: **Intel DC P3500 2TB (NVMe, PCI-E 3.0 x4, R2700/W1800)**
ネットワーク : **Intel Omni-Path Architecture HFI (100Gbps) ×4** 本

SGI ICE XA のキャビネット
(フロントドア・パネルを外した状態 **)**

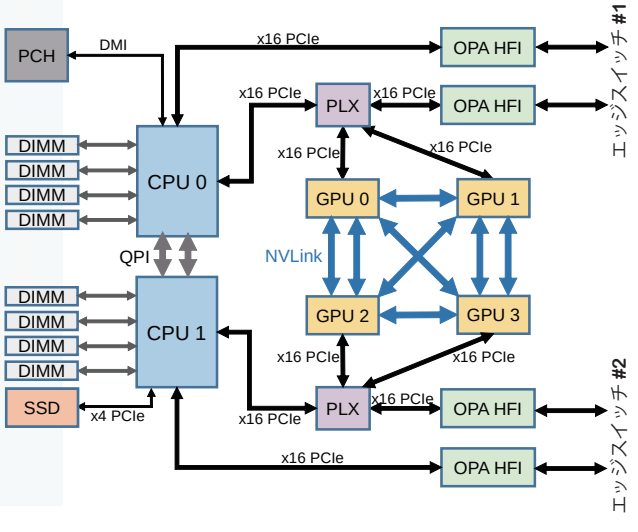
各 E-Rack には4個のシャーシが格納され , 各シャーシには 9 台の計算ノードが収容される .





計算ノードの写真 **(** カバーを開けた状態 **)**

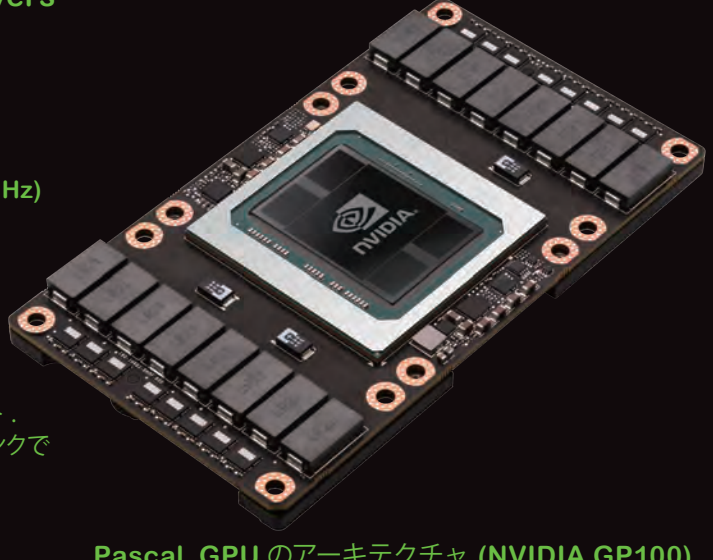
ブロックダイアグラム



Tesla P100 for NVlink-Optimized servers

理論ピーク性能 :
5.3 TFLOPS (倍精度 **)**
10.6 TFLOPS (単精度 **)**
21.2 TFLOPS (半精度 **)**
クロック周波数 : **1328 MHz (** ブースト時は **1480MHz)**
CUDA コアの数 : **3,584**
Streaming Multiprocessor 数 : **56**
オンボードメモリ : **16GB HBM2**
メモリバンド幅 : **720GB / sec**
消費電力 (TDP) : **300W**

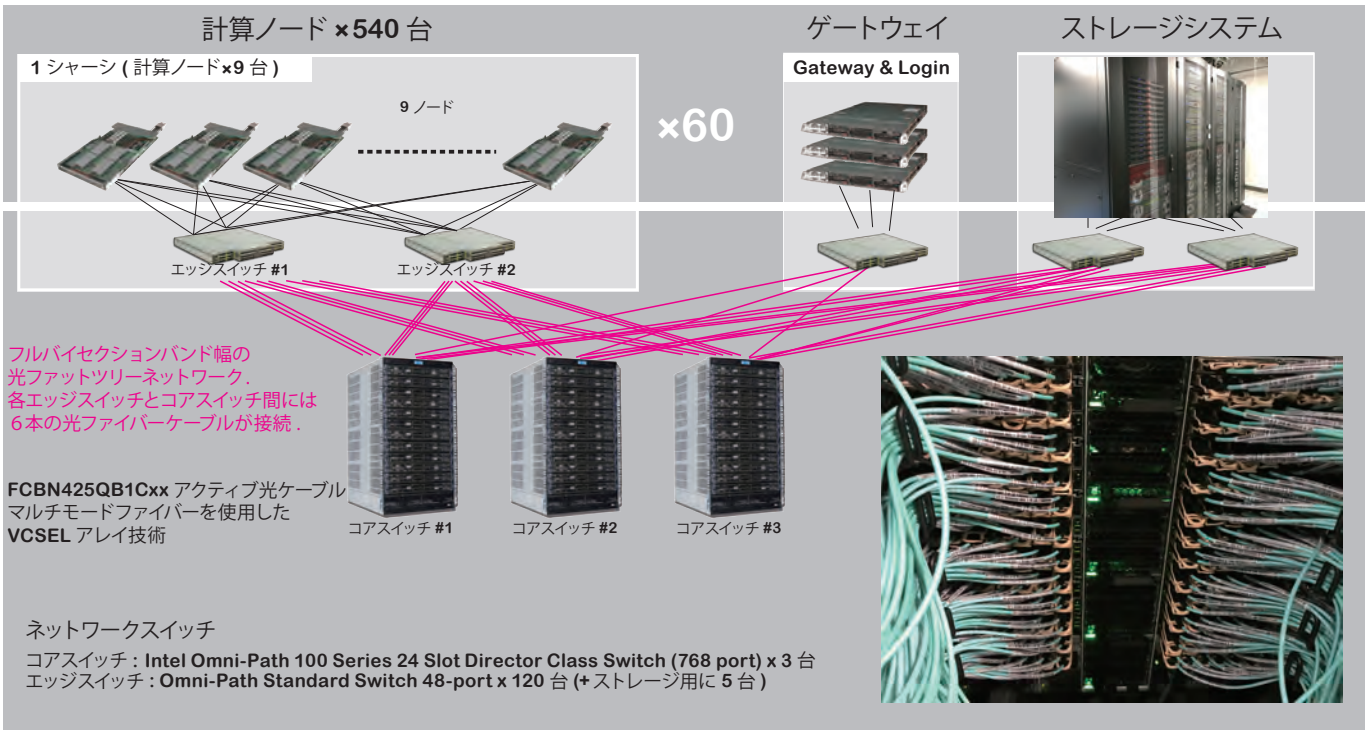
NVLink は GPU 間を直接接続するインターコネクト .
1 リンクで各方向 20GB/s のバンド幅を持ち , 4 リンクで合計 160GB/s のデータ転送が可能 . PCI-Express インターフェイスも備え , 同時にデータ転送可能 .



Pascal GPU のアーキテクチャ **(NVIDIA GP100)**

高速インターコネクト

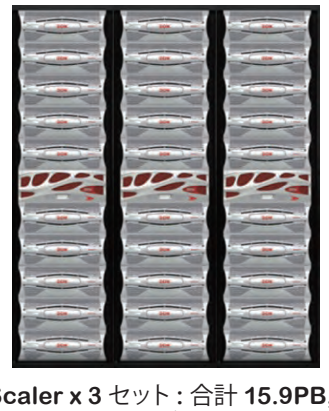
TSUBAME3.0の計算ノード群はOmni-Pathアーキテクチャを採用するフルバイセクションバンド幅432Tbpsのファットツリーネットワークで相互接続されている. 計算ノード間の通信遅延もマイクロ秒レベルと低く, 大規模アプリケーションのスケーラビリティを高めているだけでなくまたストレージシステムへの高速アクセスや高い信頼性にも寄与している.



ネットワークスイッチ

コアスイッチ : **Intel Omni-Path 100 Series 24 Slot Director Class Switch (768 port) x 3** 台
エッジスイッチ : **Omni-Path Standard Switch 48-port x 120** 台 **(+** ストレージ用に **5** 台 **)**

Lustre



DDN EXAScaler x 3 セット : 合計 **15.9PB, 150GB/s**
CIFS gateway server x 2 台

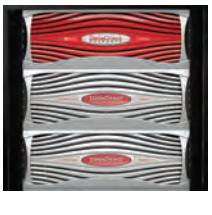
各 **EXAScaler** の構成は以下の通り
SFA14KXE x 1 台 **+ SS8462 x 10** 台
EF4024(MDS) x 2 台 **+ ED4024(MDT) x 2** 台
10TB 7.2Krpm NL-SAS HDD x 700 (スペア **20** を含む **)**
実効容量 **5.3PB (** 物理容量 **7.0PB)**

ホームストレージ

DDN GridDirector : 容量 45TB

SFA7700X x 1 台 **+ SS8460 x 2** 台
300GB 2.5" 10Krpm SAS HDD x 226
(データ : **200,** メタデータ : **20,** スペア : **6)**
NAS Gateway(NFS) x 4 台

学内キャンパスストレージ



DDN GridDirector : 容量 36TB

SFA7700X x 1 台 **+ SS8460 x 2** 台
300GB 2.5" 10Krpm SAS HDD x 186
(データ : **160,** メタデータ : **20,** スペア : **6)**
NAS Gateway(CIFS) x 4 台

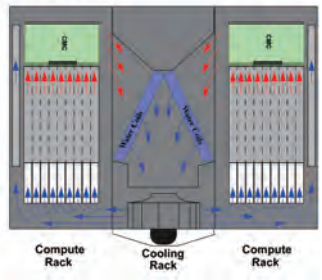
高性能・高信頼ストレージシステム

TSUBAME3.0は多種多様なストレージを提供する. 各計算ノードには合計1.08PBのSSDをスクラッチ領域として搭載. 全計算ノードから参照可能なストレージとして15.9PBの大容量・高速Lustreファイルシステムと45TBのホームストレージを備え, さらに36TBの学内キャンパスストレージを備える.

高い電力効率を実現する冷却システム

温水冷却

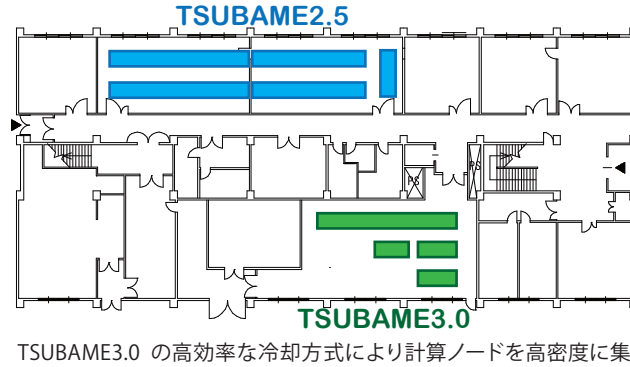
TSUBAME 本体に加えてその冷却システムも電気を消費する．これを最小化するために TSUBAME3.0 では低消費電力な蒸散冷却塔を屋上に配備している．冷却塔が供給する冷却水の温度は夏場では最高 32℃になる推定で，**TSUBAME3.0** の計算ノードはこの温度の水で動作するように設計・検証されている．



CPU と **GPU** は直接水冷となっており，その他の部品はラックが内蔵する熱交換器（図中の Water Coils）で生成される空気を用いる間接水冷を採用することによって低コスト，高メンテナンス性，および高電力効率を実現している．ストレージやネットワーク・I/O 機器の収容されるラックはラックの背面にリアドアを配備して機器の廃熱を水で冷却してから室内に戻している．

屋上に設置された冷却塔

省スペース



TSUBAME3.0 の高効率な冷却方式により計算ノードを高密度に集約することが可能となる．TSUBAME3.0 の設置面積は TSUBAME2.0/2.5 の半分以下．

システムソフトウェア

ジョブスケジューラはユーザのジョブに計算ノードを割り当てる機能を持つ．ユーザのジョブは確保された計算ノード内の全ての計算資源（CPU コア，GPU，メモリなど）を使用しない場合が多い．**UNIVA Grid Engine** は **cgroup** と **Docker** の両方に対応し，複数のジョブによる計算ノードの共有によって計算資源の有効利用を実現している．

OS	SUSE Linux Enterprise Server 12 SP2
Batch System	UNIVA Grid Engine

商用ソフトウェア

(* GPU完全対応または部分対応)

コンパイラ・デバッガ

Intel Compiler (C/C++/Fortran)
PGI Compiler*
(C/C++/Fortran, OpenACC, CUDA Fortran)
Arm Forge*

アプリケーション

ANSYS Workbench*, Mechanical*
ANSYS CFD, Fluent*, HFSS*
COMSOL Multiphysics
LS-DYNA
Gaussian*, Gauss View
MATLAB*
AVS/Express, AVS/Express PCE
Schrödinger Small-Molecule Drug Discovery Suite*
ABAQUS*, ABAQUS CAE
MSC Nastran*, Patran, Marc*
CST STUDIO SUITE* (MW-Studio*)
AMBER*
Materials Studio, Discovery Studio
Mathematica*
Maple*

Yellow: The license for all users　White: The license for Tokyo Tech users

発行：　東京工業大学　学術国際情報センター
〒152-8550 東京都目黒区大岡山2-12-1　　電話：03-5734-2087 FAX：03-5734-3198 E-mail：tsubame@gsic.titech.ac.jp

https://www.gsic.titech.ac.jp/

201906