

Microsoft Azure for Researcher

日本マイクロソフト株式会社
パブリックセクター統括本部
テクノロジーソリューションセールス本部

平塚 建一郎



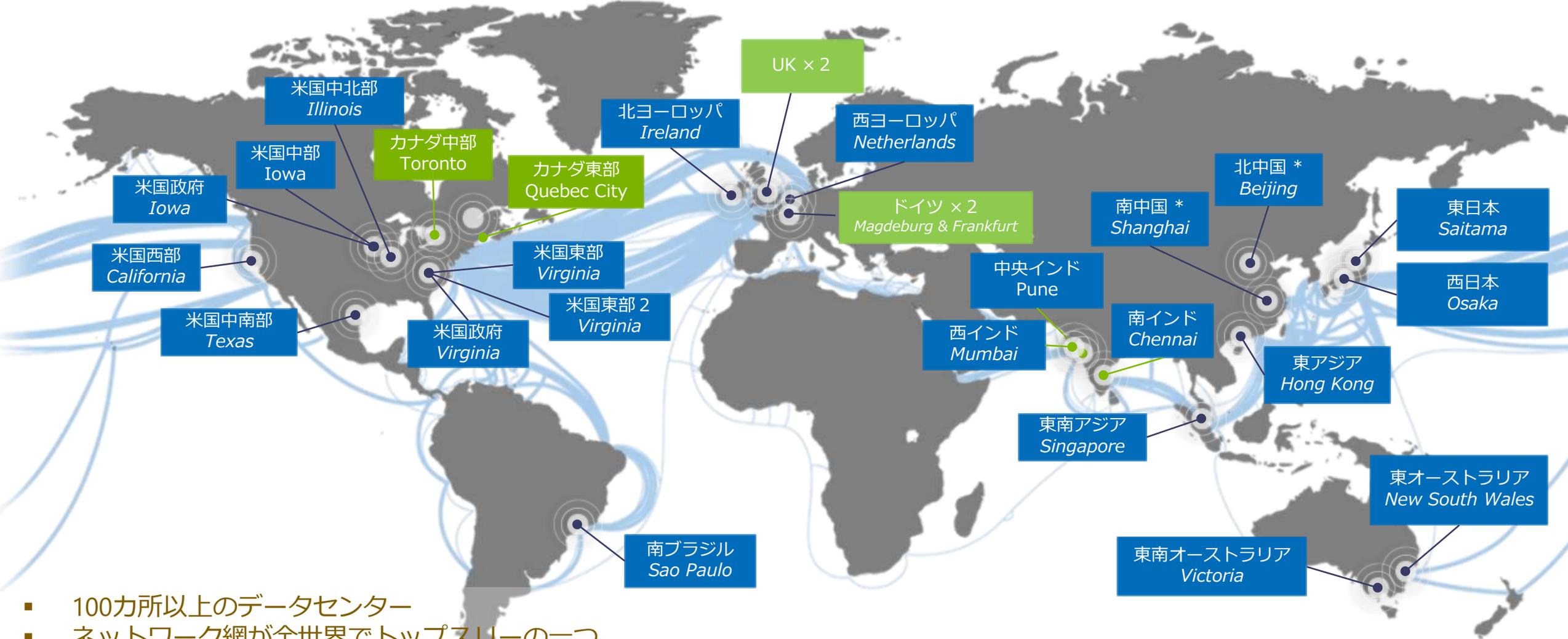
Agenda

- 1 Researcher 向けの Microsoft Azure の機能
- 2 ハイブリット HPC クラスタ環境を提供する HPC Pack
- 3 Linux も利用可能な Microsoft Azure
- 4 HPC 向けインスタンス 1 ～ InfiniBand ～
- 5 HPC 向けインスタンス 2 ～ GPU ～

Microsoft Azure とは

世界最大のインフラストラクチャー

22の地域でサービス中、28の地域まで拡大予定



- 100カ所以上のデータセンター
- ネットワーク網が全世界でトップスリーの一つ
- AWS の2倍、Google 6倍の地域サポート
- G Series – 最大 VM 提供開始 – 32 コア, 448GB RAM, SSD...

■ 稼働中 ■ 構築中 * Operated by 21Vianet

Azure Cloud Platform 全体観

Platform Services

Security & Management

- Portal
- Active Directory
- Multi-Factor Authentication
- Automation
- Key Vault
- Store / Marketplace
- VM Image Gallery & VM Depot

Compute

- Cloud Services
- Service Fabric
- Batch
- Remote App

Web and Mobile

- Web Apps
- API Apps
- API Management
- Mobile Apps
- Logic Apps
- Notification Hubs

Developer Services

- Visual Studio
- Azure SDK
- Team Project
- Application Insights

Hybrid Operations

- Azure AD Connect Health
- AD Privileged Identity Management
- Backup
- Operational Insights
- Import/Export
- Site Recovery
- StorSimple

Integration

- Storage Queues
- Biztalk Services
- Hybrid Connections
- Service Bus

Analytics & IoT

- HDInsight
- Machine Learning
- Data Factory
- Event Hubs
- Stream Analytics
- Mobile Engagement

Data

- SQL Database
- SQL Data Warehouse
- Redis Cache
- Search
- DocumentDB
- Tables

Media & CDN

- Media Services
- Content Delivery Network (CDN)

Infrastructure Services

Compute

- Virtual Machines
- Containers

Storage

- BLOB Storage
- Azure Files
- Premium Storage

Networking

- Virtual Network
- Load Balancer
- DNS
- Express Route
- Traffic Manager
- VPN Gateway
- Application Gateway

Datacenter Infrastructure
(28 Regions, 22 Online)



Researcher 向け Microsoft Azure の機能



Virtual Machine

- ✓ Windows, Linux が利用できます。
- ✓ InfiniBandやGPUを搭載したインスタンスも利用できます。



HDInsight

- ✓ Apache Hadoop の PaaS環境です。
- ✓ Spark、HBase、Storm などが利用できます。



Machine Learning

- ✓ 機械学習を簡単に始められるサービスです。
- ✓ 自由度が高く、「Python」、「R」などを使って新しいアルゴリズムを開発することも可能です。



Event Hub

- ✓ 数百万台の IoT デバイスからデータを収集できるサービスです。
- ✓ 収集したデータをリアルタイムで解析する Azure Stream Analytics なども用意しています。

Windows HPC Server (HPC Pack) とその歴史

ユーザーフレンドリーであるのみならず大規模クラスタにも対応

HPC Pack (旧称: Compute Cluster Pack)

- オンプレミス・クラウドを統合管理できるジョブスケジューラー
- MPICH2 ベースの MPI ライブラリ (MS-MPI)
- 使いやすい GUI 管理ツール
- 効率的なコマンドライン管理ツール
- Excel 高速化機能
- 最新版は HPC Pack 2012 R2 Update 2 (2015 年 7 月リリース)



2006年 Compute Cluster Pack (HPC v1)

- 三菱UFJ証券様のクラスタがTop 500にランクイン。(1760コア、6.52TFlops)
<http://www.top500.org/system/174885>

2008年 HPC Pack 2008 (HPC v2)

- 上海スーパーコンピューティングセンターのDawning 5000AがTop500で11位にランクイン。(30,720コア、180.6TFlops)
<http://www.top500.org/system/176118>

2010年 HPC Pack 2008 R2 (HPC v3)

- 東工大のTSUBAME 2.0で初のペタフロップス越え。1.13PFlops. Top500の5位相当の記録。

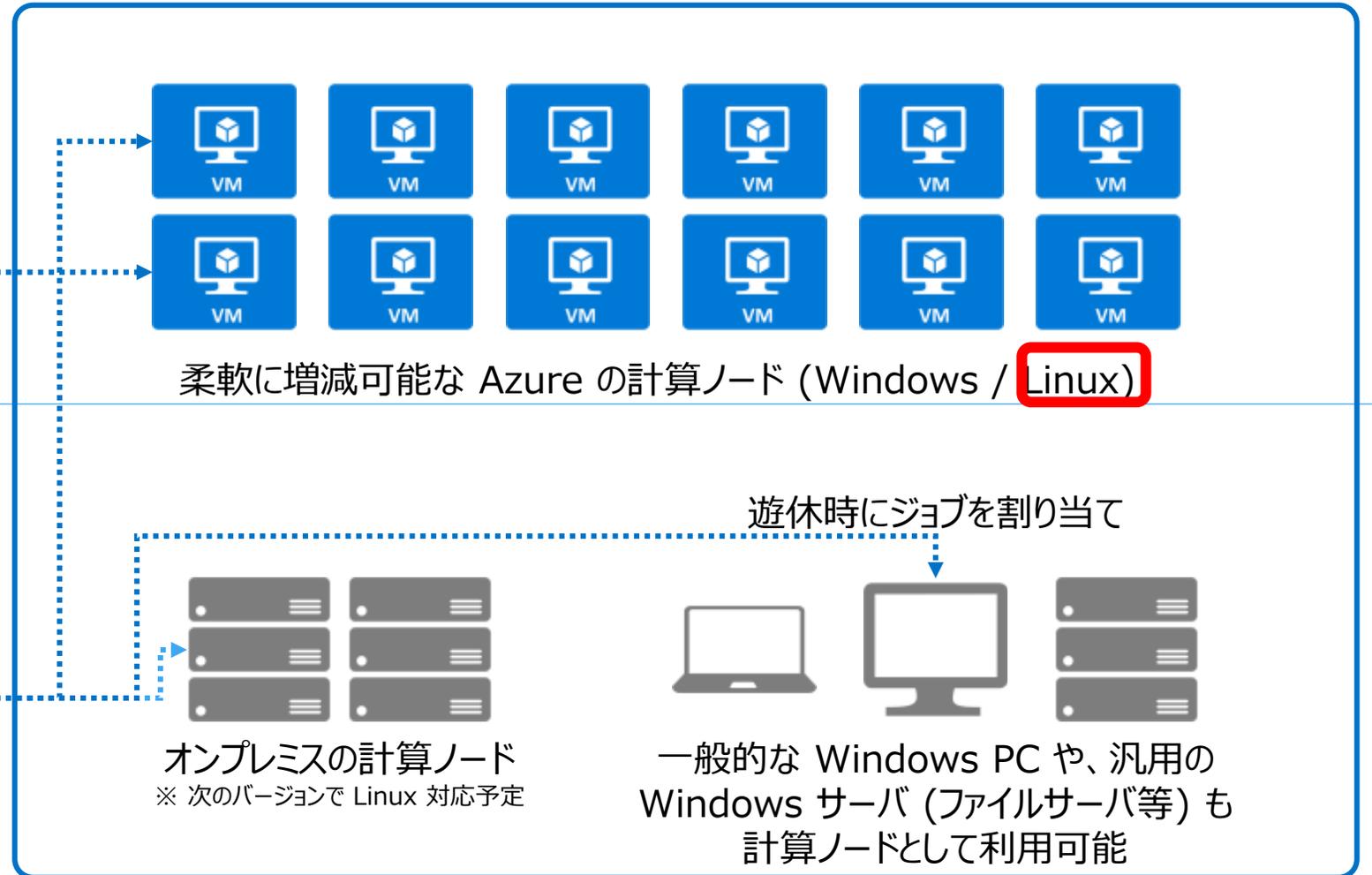
2012年 HPC Pack 2012 (HPC v4)

- HPC用AzureインスタンスでTop500にランクイン (8064コア、151.3 TFlops)
<http://www.top500.org/system/177982>

オンプレミス + クラウドの統合クラスタ

PC, サーバー, クラウド. 様々なコンピューターを計算ノードに

- 社内とクラウドの計算ノードを「一つのクラスタとして」統合管理可能
- クラウドへのノード追加・削除は、数百ノードレベルでも10分程度で完了
- スケジュールに従って自動的にノードを追加・削除することも可能



クラウド (Azure)

社内



管理ツール

ジョブ投入

ジョブの
割り当て

ユーザー端末

ヘッドノード
(Windows Server)

オンプレミスの計算ノード
※ 次のバージョンで Linux 対応予定

遊休時にジョブを割り当て

一般的な Windows PC や、汎用の Windows サーバ (ファイルサーバ等) も計算ノードとして利用可能

Linux も利用可能な Microsoft Azure



高性能インスタンス (A8 - A11)

InfiniBand 搭載

高速 CPU, 大容量メモリ, 高速インターコネクト

サイズ	コア数	メモリ容量	プロセッサ	ネットワーク 1	ネットワーク 2
A8	8	56 GB	Xeon E5-2670 2.6 GHz	10 Gbps イーサネット	QDR InfiniBand (w/ RDMA)
A9	16	112 GB			無し
A10	8	56 GB			
A11	16	112 GB			

A8,A9 = MPI による並列計算に最適化

- TOP500 にランクインしました (2012年11月)
 - 151.3 TFLOPS (効率 90.2%) で 165 位
 - A9を504ノード, 8064 コアで実施
<http://www.top500.org/system/177982>
- MSMPI と Intel MPI に対応。

(A8,A9) - InfiniBand = A10,A11

- 「A8,A9 の速い CPU や大きなメモリは必要だが、MPI は使わないので InfiniBand 無しモデルを」という要望を受けて追加。
- InfiniBand の有無以外は A8, A9 と同一。
<https://msdn.microsoft.com/library/azure/dn689095.as>

検証事例: Particleworks on Azure

実施内容

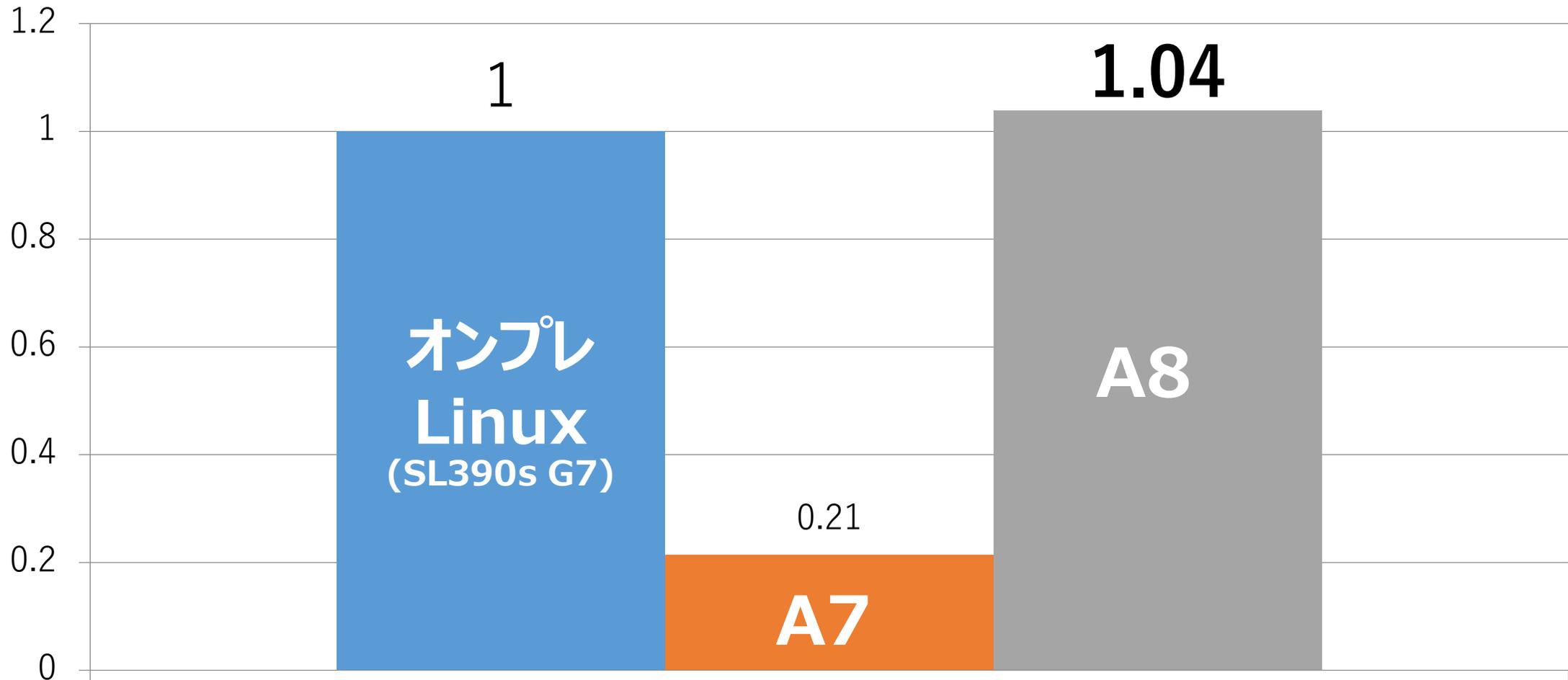
- 4000万粒子規模の解析
- 内容: 自動車の水はね



オンプレミスのLinuxマシンとの比較を実施

- オンプレミスの Linux クラスタと、AzureのA7,A8,A9インスタンスで同じ解析を実施し、実行時間を比較。
 - 機種: ProLiant SL 390s G7 x 4 ノード (計48コア)
 - CPU : Intel Xeon X5675 3.06GHz 6 cores x2
 - RAM : 4GBx12 = 48 GB
 - QDR InfiniBand 40Gbpsx2

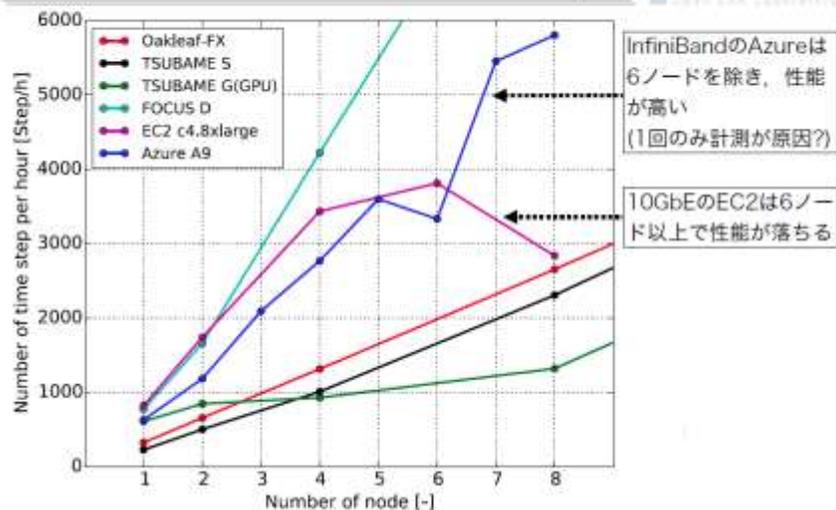
オンプレミスの物理マシンと遜色ない性能



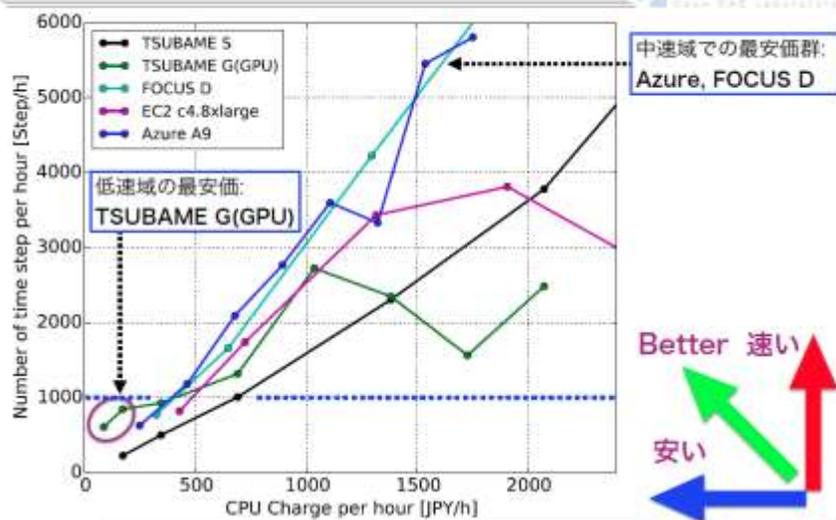
オンプレミスの InfiniBand 付き物理マシンの性能を 1 とした場合の比較
A8/A9 は物理マシンと同等、A7 との比較では 5 倍の性能を記録
(流体計算アプリケーションでの性能検証結果)

OpenFOAM 性能・費用ベンチマークテスト

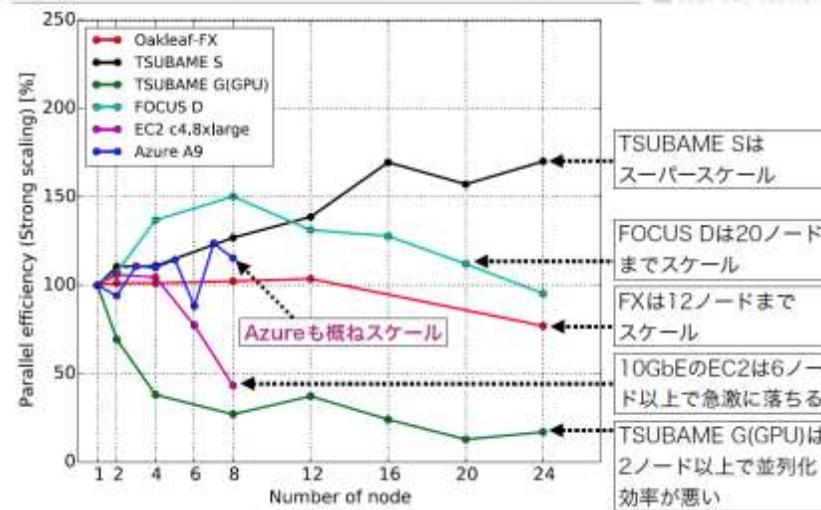
1時間あたりの時間ステップ数(8ノード以下)



1時間の課金と時間ステップ数(成果非公開型)



並列化効率(Strong scaling)

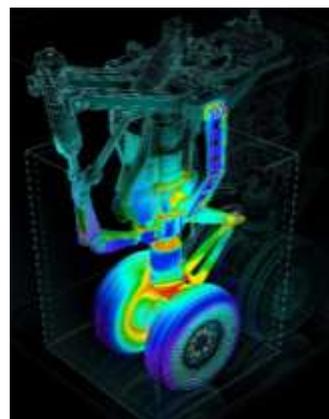


・ 1時間で実行可能な時間ステップ数と課金額の関係から、時間ステップ数の領域毎に安価なシステムが分かれる結果となった。

- ✓ 成果非公開型(TSUBAME, FOCUS, EC2, Azure)
 - ✓ 高速域(6,000時間ステップ/h~): FOCUS D
 - ✓ 中速域(1,000~6,000時間ステップ/h): Azure, FOCUS D
 - ✓ 低速域(~1,000時間ステップ/h): TSUBAME G(GPU)
- ✓ 成果公開型(Oakleaf-FX, TSUBAME)
 - ✓ 高速域(6,000時間ステップ/h~): TSUBAME S
 - ✓ 中速域(1,000~6,000時間ステップ/h): Oakleaf-FX
 - ✓ 低速域(~1,000時間ステップ/h): TSUBAME G(GPU)

NVIDIA Tesla M60 あるいは K80 を備える「Nシリーズ」

サイズ	コア数	メモリ容量 (GB)	SSD (TB)	InfiniBand with RDMA	GPU
N1	6	64	~0.5	-	M60 x 0.5
N10	6	64	~0.5	-	K80 x 0.5
N11	12	128	~1.0	-	K80 x 1
N12	24	256	~2.0	-	K80 x 2
N21	24	256	~2.0	有り	K80 x 2



GPU インスタンスの想定用途

- ゲームなどのクラウドレンダリング
- 動画のエンコーディング
- デスクトップアプリでの OpenGL / DirectX サポート
- GPGPU (CUDA / OpenCL)



