
# Pacific Teck
# 過去 8年半と未来

設立20周年記念PCクラスタシンポジウム

December 2021

Howard Weiss ワイスハワード

PacificTeck
HPC and Machine Learning Experts

# Agenda

- 会社概要と自己紹介
- 過去の８年半の実績
- 将来に向けての課題

Managing Director

# Howard Weiss

ハワード ワイス

- アメリカ ミシガン州出身
- 1990年日本移住
- 英語・日本語
- Phoenix Technologies / Sales Director, Japan
- BakBone Software / VP APAC
- Cofio Software (acquired by HDS) / co-founder BOD member
- Voltaire / Mellanox / VP APAC
- DataDirect Networks / VP APAC
- Incorporated Pacific Teck Limited in July 2013

PacificTeck
HPC and Machine Learning Experts

# 世界中の最先端技術製品にフォーカス

- 日本を拠点にインドを含むアジア太平洋エリア(APAC)に製品を提供
- 英語 / 中国語 / 日本語のグロバールな言語での支援が可能
- ハードウェアには依存しないソフトウェアソリューションを提供
- APAC最大のスパコンでの採用実績多数
- ストレージ/コンテナー/ジョブ管理のエキスパート

# Pacific Teck in the TOP500! Nov 2021 list


1st Riken/Fugaku (SingularityPRO, iRODS, ARM Forge)

16th AIST/ABCI (Grid Engine, SingularityEnterprise, BeeGFS)

54th NCHC/Taiwania 2 (Grid Engine, SingularityPRO)

59th TiTech/TSUBAME3.0 (Grid Engine, BeeGFS)

79th Osaka U/Squid (SingularityPRO)

98th Nagoya U/Flow (BeeGFS, NVMesh)

110th Tokyo U/Oakbridge-CX (BeeGFS)

433th Tohoku U/AFI-NITY (Grid Engine)

# 取扱製品カテゴリー

- ジョブ管理システム
  - ALTAIR
- HPC仮想コンテナーシステム
  - PRO
  - ENTERPRISE
- ストレージソフトウェア
  - 並列ファイルシステム
    - BeeGFS
    - BeeOND
  - NVMe高速ストレージ
    - NVMesh
  - S3 オブジェクトストレージ
    - DATACORE
- S3クラウドストレージサービス
  - RSTOR
- ティアリングスソフト・データ移動のツール
  - iRODS
- クラスターマネジメントシステム
  - Bright Computing
- プログラム開発者用ツール
  - 

  - arm FORGE

# 過去８年で達成できたこと

- GPU-NUMA bus aware ノード分割 ***job scheduling***
- Promoting the ***containerization*** of applications to improve portabilty
- Usage of a compute node as a ***high performance storage scratch*** system
- Creation of ***high speed storage systems*** from commodity hardware

# Altair (Univa) Grid Engine

## CPU・GPUの構成を考慮した最適なジョブ実行が可能


@

RIKEN

産総研

東京工業大学
Tokyo Institute of Technology

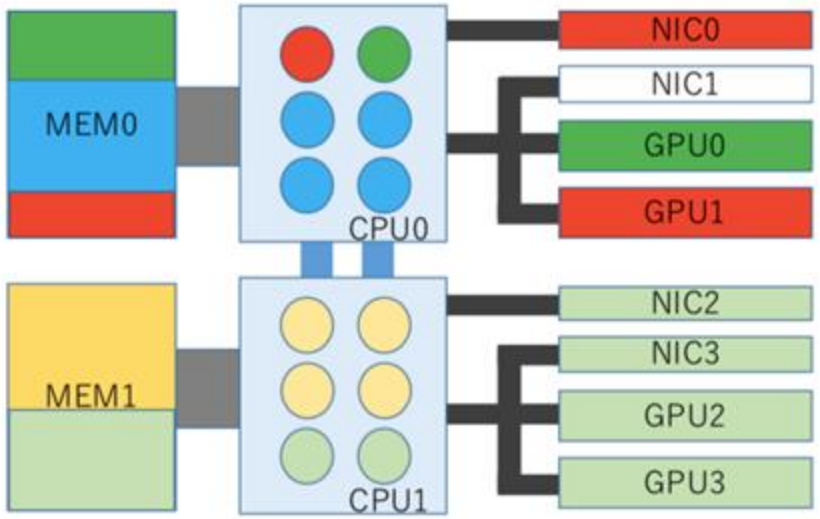
NARO 農研機構

IFS 東北大学 流体科学研究所
Institute of Fluid Science, Tohoku University

# Altair (Univa) Grid Engine - TSUBAME3.0

東京工業大学
Tokyo Institute of Technology

## TSUBAME3.0 Container-Based Fine-grained Spatial Resource Allocations of Fat Nodes



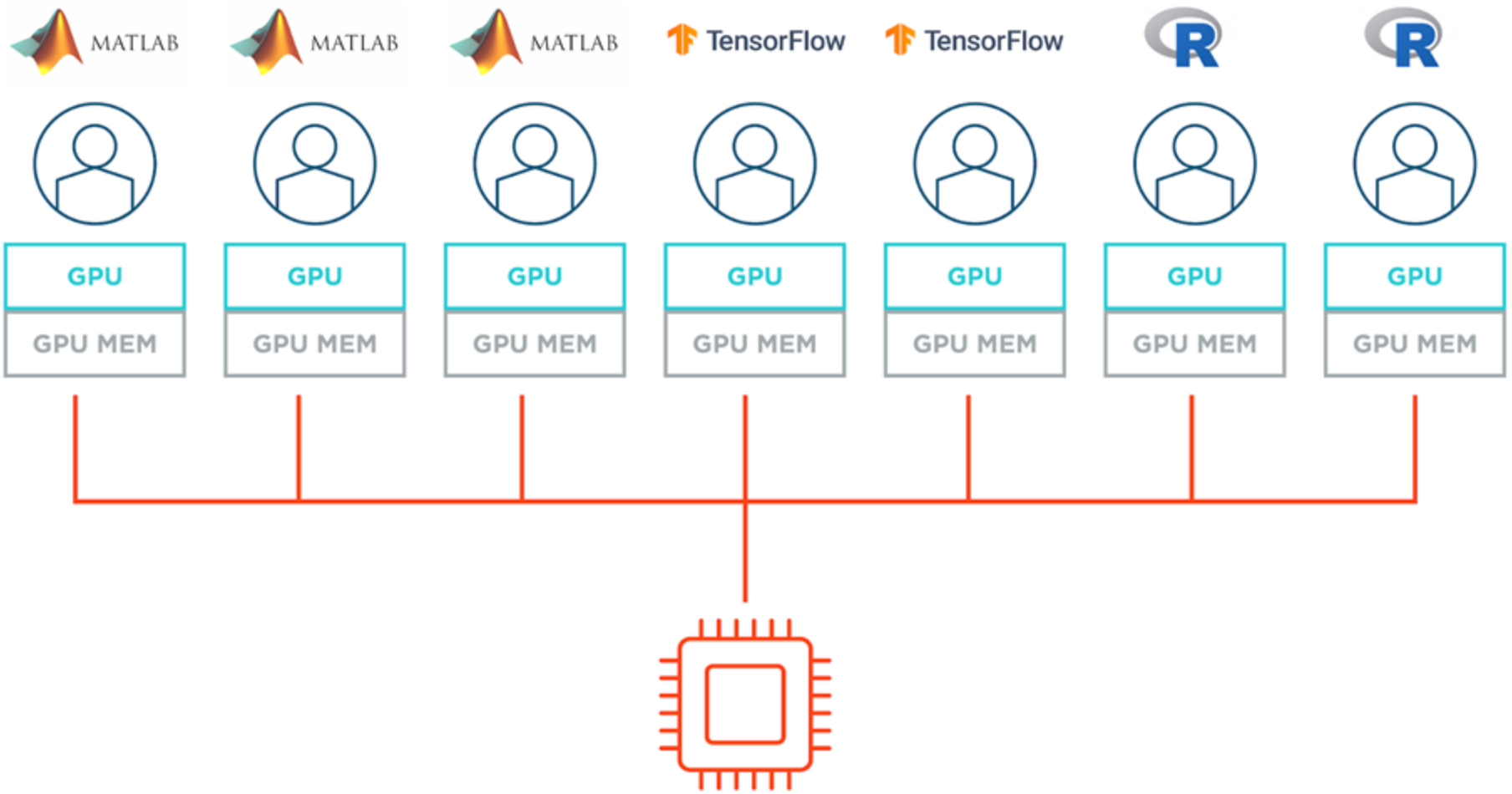

Resource Isolation via UGE Containers (future Docker etc.)

| Job | Allocated Resource |
|---|---|
| 1 | CPU 2Cores, NIC0, GPU1, 32GB Mem |
| 2 | CPU 8 Cores, 64GB Mem |
| 3 | CPU 4 Cores, GPU0、16GB Mem |
| 4 | CPU 8 Cores, 64GB Mem |
| 5 | CPU 4 Cores, NIC2&3, GPU2&3, 48G Mem |

Container configuration and deployment tied to Univa Grid Engine

# Multi-Instance GPU on the NVIDIA A100

## Realize the full potential of the A100 GPU



# 過去８年で達成できたこと

- GPU-NUMA bus aware ***job scheduling***
- Promoting the ***containerization*** of applications to improve portabilty
- Usage of a compute node as a ***high performance storage scratch*** system
- Creation of ***high speed storage systems*** from commodity hardware

# SingularityPRO・Singularity Enterprise Sample Endusers




## Singularity Enterpriseの特徴

- PGP公開鍵を追加
- コンテナライブラリを提供（オンプレ版・Cloud版）
- Remote Build Services を提供（オンプレ版・Cloud版）


## SingularityPROの特徴

- Sylabs社により技術確認されたSingularityPROのバイナリ版を提供
- 同じバージョンのロングランサポート
- お客様のご要望を次期バージョンに反映

# 過去８年で達成できたこと

- GPU-NUMA bus aware ***job scheduling***
- Promoting the ***containerization*** of applications to improve portabilty
- Usage of a compute node as a ***high performance storage scratch*** system
- Creation of ***high speed storage systems*** from commodity hardware

# BeeOND Parallel File System
## Sample End users

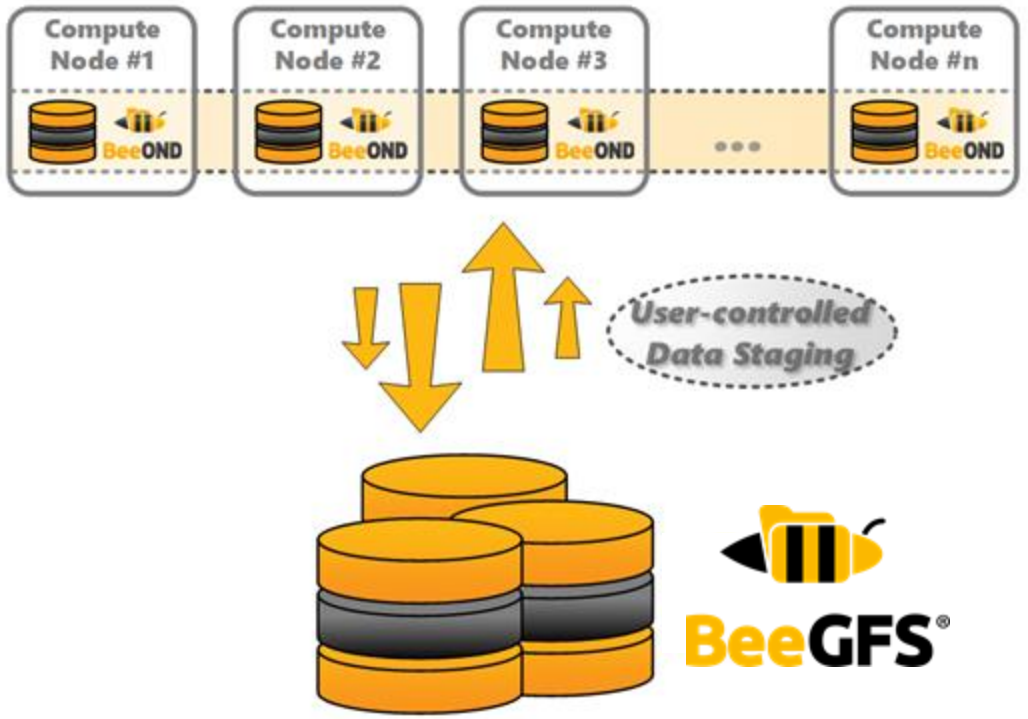




# BeeOND - BeeGFS On Demand

- パラレルファイルシステムをその場で作成
- 簡単な起動/停止コマンド
- Jobスクリプトに記述しジョブスケジューラによる

作成・削除の制御が可能（ジョブスケジューラ: Altair Grid Engine）




**<< これから流行 >>**

**計算ノードの中のファイルシステムは**
**パーマネントユースとして使用**

東京工業大学
Tokyo Institute of Technology

ABCI
AI Bridging Cloud Infrastructure

東京大学情報基盤センター
INFORMATION TECHNOLOGY CENTER, THE UNIVERSITY OF TOKYO

# 過去８年で達成できたこと

- GPU-NUMA bus aware ***job scheduling***
- Promoting the ***containerization*** of applications to improve portabilty
- Usage of a compute node as a ***high performance storage scratch*** system
- Creation of ***high speed storage systems*** from commodity hardware

# CSIRO Storage Architecture
# 2.048 PB usable capacity all NVMe

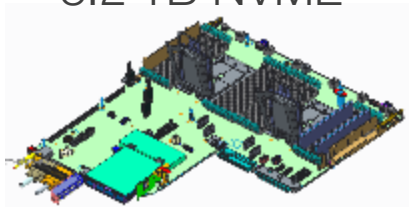
Metadata: Dell PowerEdge R740XD

x 4

Storage: Dell PowerEdge R740XD

x 32

3.2 TB NVME



x 24 per server

# 将来に向けての課題

- How to support storage, containers, debugging tools on upcoming architectures such as ARM and accelerators
- Moving of the parallel permanent file system to the compute node
- Seemless operation of on premise and cloud resources with job schedulers, containers, etc.
- How to move / archive data on an S3 compliant cloud
- Various use cases for S3
- Application specific accelarator eco system support

# Altair Breeze™ の概要

- Altair Breeze™ （以後、Breezeと略します。）は、Mistralより詳細に、全てのファイルとプロセスについてのI/Oを監視する事が出来ます。
  I/Oエキスパートユーザー、管理者に大変役に立つツールです。
  - 内製アプリの最適化のために利用したり、レガシーアプリケーションのマイグレーションのために利用出来ます。



Breezeは、I/Oエキスパートの方向けの製品です。



ALTAIR

# 将来に向けての課題

- How to support storage, containers, debugging tools on upcoming architectures such as ARM and accelerators
- Moving of the parallel permanent file system to the compute node
- Seemless operation of on premise and cloud resources with job schedulers, containers, etc.
- How to move / archive data on an S3 compliant cloud
- Various use cases for S3
- Application specific accelarator eco system support

# コンピュートノード上にストレージを実現する




- タイプIIサブシステム100ノードはBeeGFSを使用
- 50ノードがNVMeshを使用
- 50ノードでRAID 1とイレイシャーコーディング使用パーマネントファイルシステム
- 2020年7月1日に運用開始
- 永続ファイルシステムとして使用可能

# 将来に向けての課題

- How to support storage, containers, debugging tools on upcoming architectures such as ARM and accelerators
- Moving of the parallel permanent file system to the compute node
- Seemless operation of on premise and cloud resources with job schedulers, containers, etc.
- How to move / archive data on an S3 compliant cloud
- Various use cases for S3
- Application specific accelarator eco system support

# クラウドジョブ管理・クラウドバースティング


NEW

Altair Control
navops

- Migrates workloads to the cloud
- Helps organizations control spending

*"We gained practically infinite capacity with Univa's hybrid cloud solution in a very cost-effective manner"*
Enterprise Computing Director, *HPC* Mellanox

Sample clients

Mellanox TECHNOLOGIES

Wharton University of Pennsylvania

# 将来に向けての課題

- How to support storage, containers, debugging tools on upcoming architectures such as ARM and accelerators
- Moving of the parallel permanent file system to the compute node
- Seemless operation of on premise and cloud resources with job schedulers, containers, etc.
- How to move / archive data on an S3 compliant cloud
- Various use cases for S3
- Application specific accelarator eco system support

# 将来に向けての課題

- How to support storage, containers, debugging tools on upcoming architectures such as ARM and accelerators
- Moving of the parallel permanent file system to the compute node
- Seemless operation of on premise and cloud resources with job schedulers, containers, etc.
- How to move / archive data on an S3 compliant cloud
- Various use cases for S3
- Application specific accelarator, eco system support


# Thank you

sales@pacificteck.com