

趣旨説明

大阪大学サイバーメディアセンター
応用情報システム研究部門 伊達 進

SQUID
Supercomputer for
Quest to
Unsolved
Interdisciplinary
Datascience

利用説明会の位置付け

大阪大学、あるいは、それ以外の計算機センターでスーパーコンピュータを使ったことのある方を対象とした説明会

※ **SQUIDに興味・関心のある方はどなたでも参加いただけます。**

※ ※ **本日と同じ内容の利用説明会を4月26日に開催しています。**

※ ※ ※ **スーパーコンピュータを未経験な方向けの説明会については2021年5月25日に別途開催予定です。**

利用説明会での達成目標

- SQUIDの概要・スケジュール
 - SQUIDの概要を理解する。
 - SQUIDの運転計画を理解する。
- SQUIDの利用方法
 - ログイン、ジョブ投入・実行、ジョブ確認の基本を理解する。
(OCTOPUS/SX-ACEとSQUIDの共通点、相違点を理解する。)
- SQUID利用申請
 - SQUIDを利用するために必要となる申請方法、課金の考え方を理解する。

本日のプログラム

13:00-13:20 趣旨説明



サイバーメディアセンター応用情報システム研究部門

伊達 進

13:20-14:50 SQUIDの利用方法

日本電気株式会社

14:50-15:30 利用申請方法

情報推進部 情報基盤課

技術職員

※ 適宜各時間帯で質問を受け付ける予定です。

クラウド連動型HPC・HPDA用スーパーコンピュータシステムSQUID (Supercomputer for Quest to Unsolved Interdisciplinary Datascience)

- わが国の学術・産業を支える研究者による未解決のデータサイエンス問題への探求を支援するため、高性能計算および高性能データ分析分野の多様な計算ニーズを収容可能なクラウド連動型スーパーコンピュータ

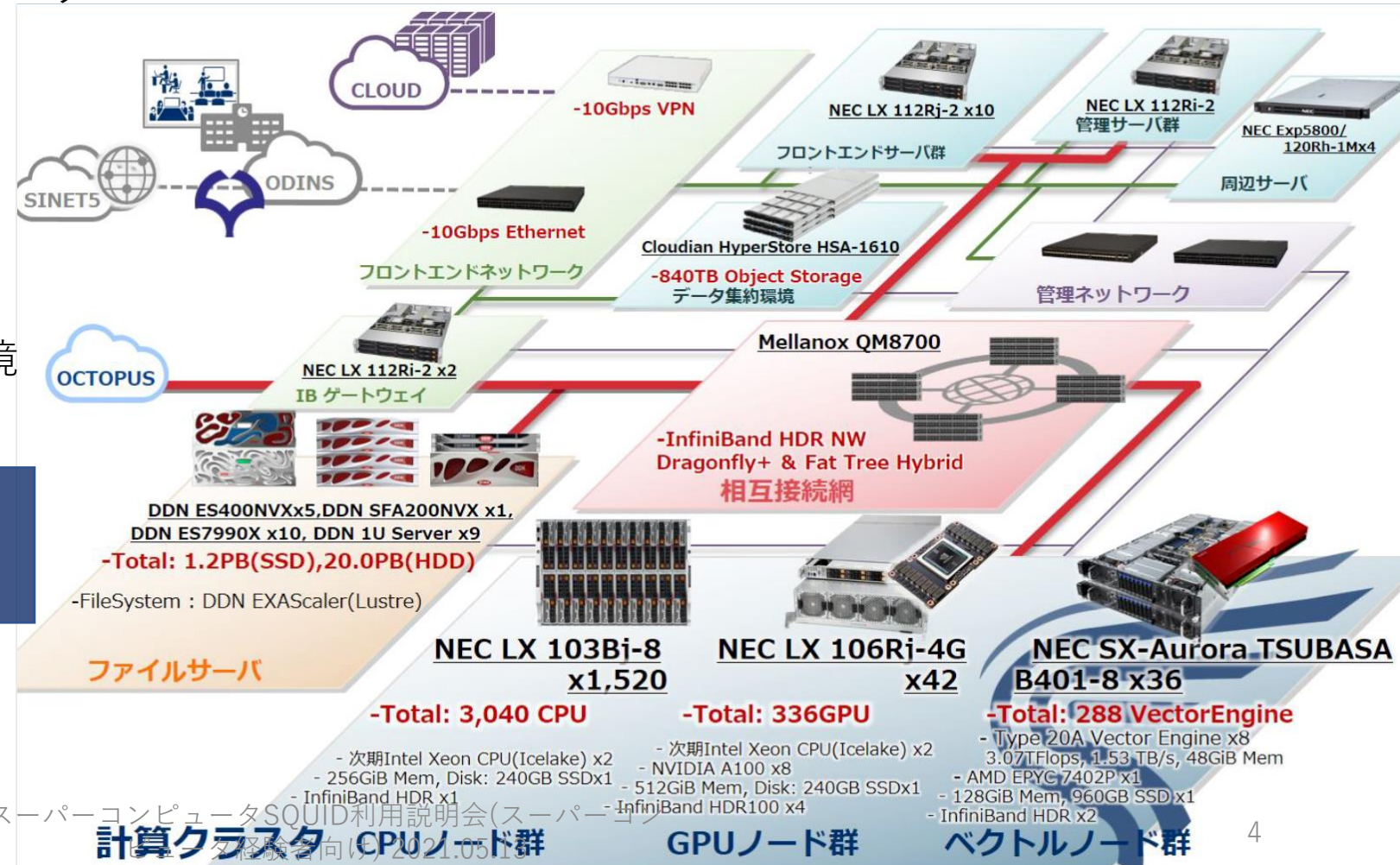
特徴

- HPC・HPDA融合
- テイラーメイド環境
- データ集約基盤ONION
- セキュアコンピューティング環境
- クラウドバースティング環境

本日は、これらの詳細については説明せず、基本的な利用方法にとどめます。
別途予定いたします。

稼働

- 2021年5月 ~



クラウド連動型HPC・HPDA用スーパーコンピュータシステムSQUID (Supercomputer for Quest to Unsolved Interdisciplinary Datascience)



- 16.591 PetaFlopsの総理論演算性能を予定

総演算性能	16.591 PFLOPS	
ノード構成	汎用CPUノード群 1,520 ノード(8.871 PFLOPS)	プロセッサ：Intel Xeon Platinum 8368 (Icelake / 2.40 GHz 38コア) 2基 主記憶容量：256GB
	GPUノード群 42 ノード(6.797 PFLOPS)	プロセッサ：Intel Xeon Platinum 8368 (Icelake / 2.40 GHz 38コア) 2基 主記憶容量：512GB GPU：NVIDIA A100 8基
	ベクトルノード群 36 ノード(0.922 PFLOPS)	プロセッサ：AMD EPYC 7402P (2.8 GHz 24コア) 1基 主記憶容量：128GB Vector Engine：NEC SX-Aurora TSUBASA Type20A 8基
ストレージ	DDN EXAScaler (Lustre)	HDD：20.0 PB NVMe：1.2 PB
ノード間接続	Mellanox InfiniBand HDR (200 Gbps)	



5.836 Tflops /node
(2.918 Tflops/CPU)
0.83 kW/node

161.836 Tflops /node
(2.918 Tflops/CPU,
19.5Tflops/GPU)
5.4 kW/node

26.71 Tflops /node
(2.150 Tflops/CPU,
3.07 Tflops/VEC)
3.3 kW/node



1.996 Tflops /node
(0.998 Tflops/CPU)
0.44 kW/node

5.836 Tflops /node
(2.918 Tflops/CPU,
5.3Tflops/GPU)
1.89 kW/node

276 Gflops /node
0.56 kW/node

ノード性能の大幅UP 😊
消費電力UP ☹️

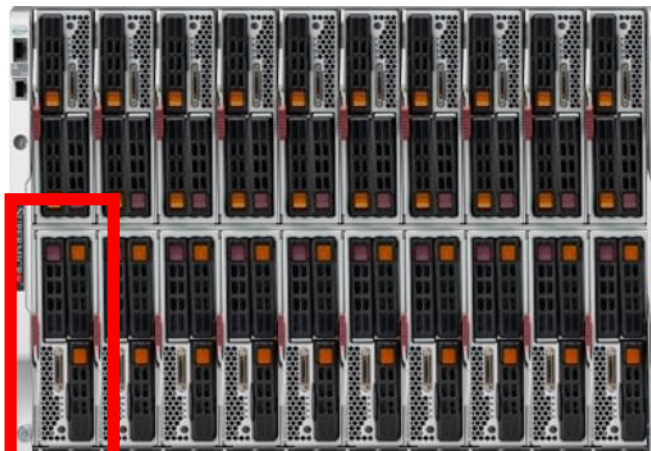


SQUID-CPU

- 1,520ノード (NEC LX 103Bj-8)

NEC LX 103Bj-8

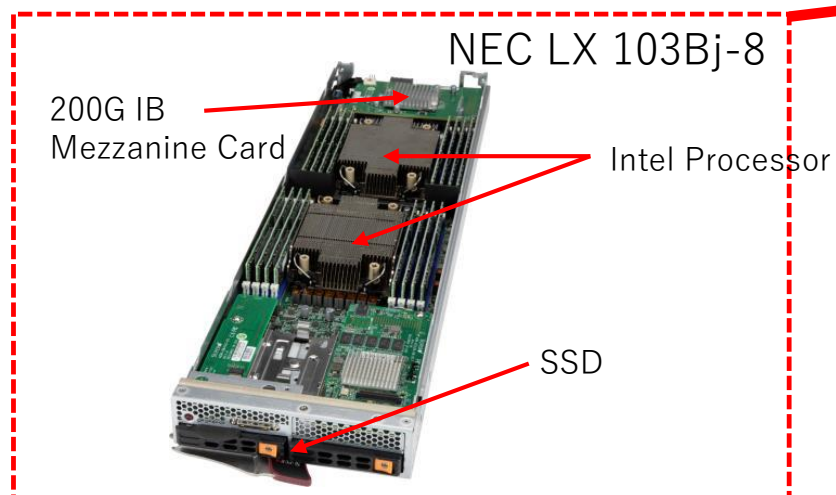
プロセッサ	Intel Xeon Platinum 8368 (Icelake) x 2
メモリ構成	256GiB (16GB DDR4-3200 DIMM x 16)
ストレージ	240GB SSD
NW I/F	InfiniBand HDR x 1, 25G/10GbE x2, BMC
OS	Cent OS



8U/シャシー

4シャシー/ラック

水でCPUを冷却し、安定的な性能を供給します。



SQUIDでは:

- 1シャシーあたり19ノードを搭載
- 1ラック あたり 4シャシー
- SQUID-CPUに20ラック
- ➡ **$19 \text{ ノード/シャシー} \times 4 \times 20 = 1520 \text{ ノード}$**
- 1ノードあたり2CPU
- ➡ **$1520 \times 2 = 3040 \text{ CPU}$**

SQUID-GPU

- 42ノード (NEC LX 106Rj-4G)

NEC LX 106Rj-4G

プロセッサ	Intel Xeon Platinum 8368 (Icelake) x 2
メモリ構成	512GiB (32GB DDR4-3200 DIMM x 16)
ストレージ	240GB SSD
NW I/F	InfiniBand HDR100 x 4, 25G/10GbE x2, BMC
GPU	NVIDIA A100 (SXM4) x 8
OS	CentOS



水でCPU、GPUを冷却し、安定的な性能を供給します。



SQUIDでは:

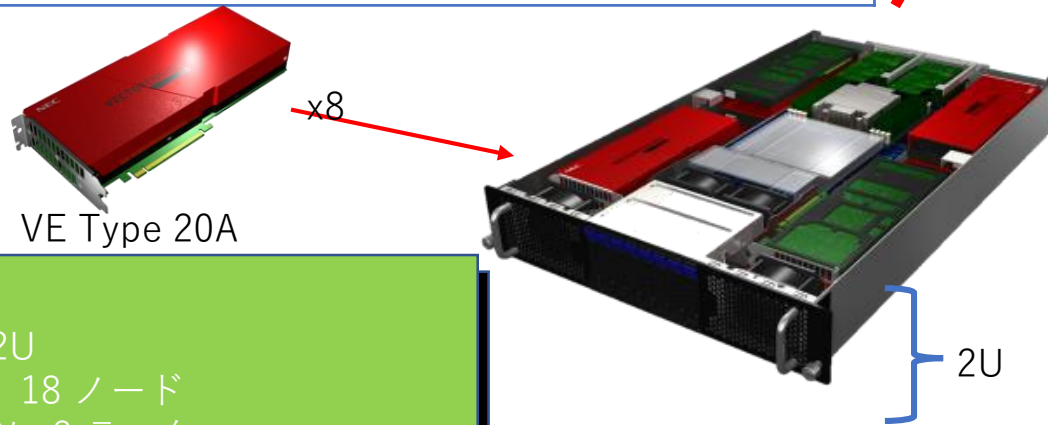
- 1ラック あたり 6ノード
- SQUID-CPUに 7 ラック
- ➡ **$6 \times 7 = 42$ ノード**
- 1ノードあたり2CPU, 8GPU
- ➡ CPU: **$42 \times 2 = 84$ CPU**
- ➡ GPU: **$8 \times 42 = 336$ GPU**

SQUID-VECTOR

• 36ノード (NEC SX-Aurora TSUBASA B401-8)

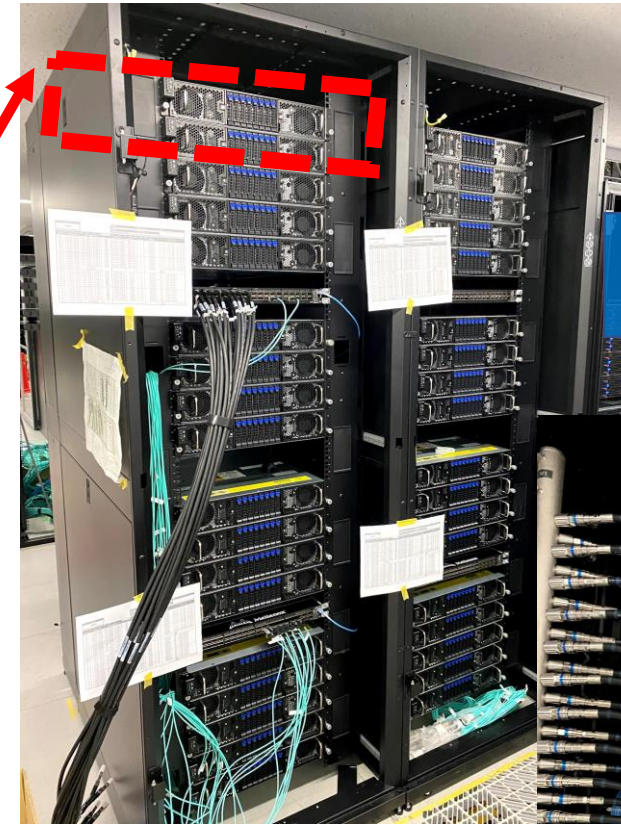
NEC SX-Aurora TSUBASA B401-8

プロセッサ	AMD EPYC 7402 (2.8Ghz/24c) x 1
メモリ構成	128GiB (16GB DDR4-3200 DIMM x 8)
ストレージ	960 GB SSD
NW I/F	InfiniBand HDR x 2
ベクトルエンジン	VE Type 20A x 8
OS	CentOS

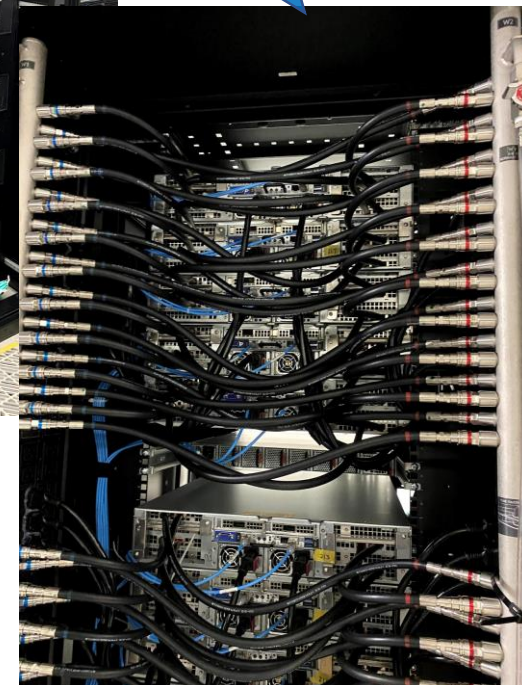


SQUIDでは:

- 1ノードあたり 2U
- 1ラックあたり 18ノード
- SQUID-VECTORに 2ラック
- ➡ **$18 \times 2 = 36$ ノード**
- 1ノードあたり 8 VE を搭載
- ➡ **$8 \times 36 = 288$ VE**

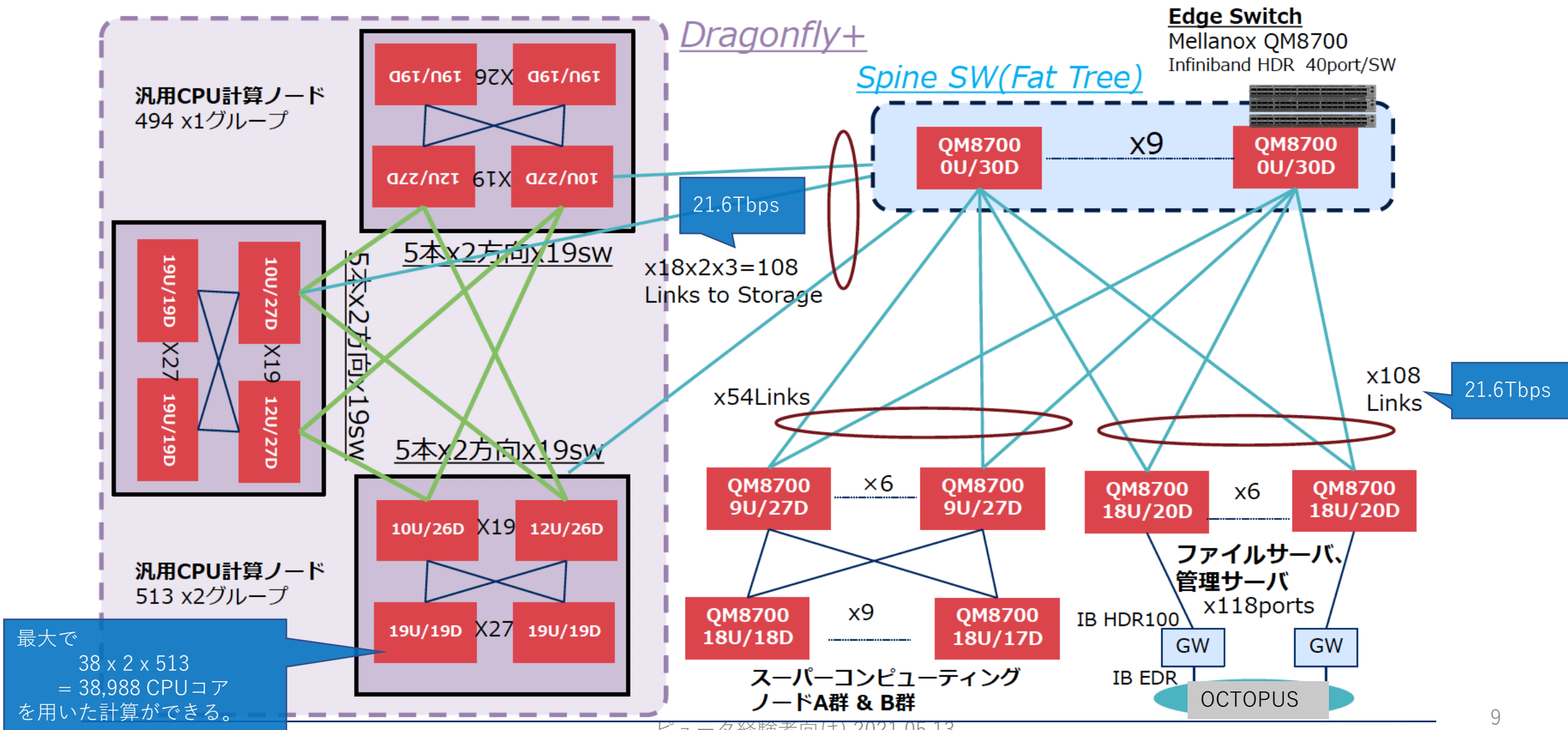


水でCPU、VECを冷却し、安定的な性能を供給します。



SQUID-INTERCONNECT

- Dragonfly+ と Fat tree の併用:



データ集約基盤ONION

- 特徴5. テイラーメイド型計算
 - Osaka university Next-generation Infrastructure for Open research and open Innovation
 - 世界最高水準の基礎的、基盤的研究や学際融合研究が生み出す多様な知の創出と深化に寄与すべく、**総合大学である大阪大学内で創出された「利用可能な超大量データを将来に渡る持続可能性を保持しつつ責任をもって活用」可能にする**とともに、**新たな社会的価値の創出を目指した「産学共創」「国際共同研究」のための学内外でのデータ利活用を支援する**データ集約基盤
 - スーパーコンピュータSQUIDの調達に合わせて試験的に導入**（うまく機能しない、予算獲得できない場合はやめる）

クラウドノード群

汎用計算ノード群

SCノード群 (Type A: GPU)

SCノード (type B: VE)

総計 300+ TFlops
4th S2DHシンポジウム - 2021.03.05

セキュアフロントエンド

セキュアステージング機能

特徴2. クラウド連動・連携機能

高負荷状態時にオフロード
新製品の早期実践配備

民間クラウド(IaaS)事業者

ODINS網を活用した
学内部局からのデータ集約

ODINS

SINET5

オブジェクトストレージ

データ集約ストレージ 500TB

クラウドストレージを活用した
海外研究拠点とのデータ共有

特徴4. データ集約環境

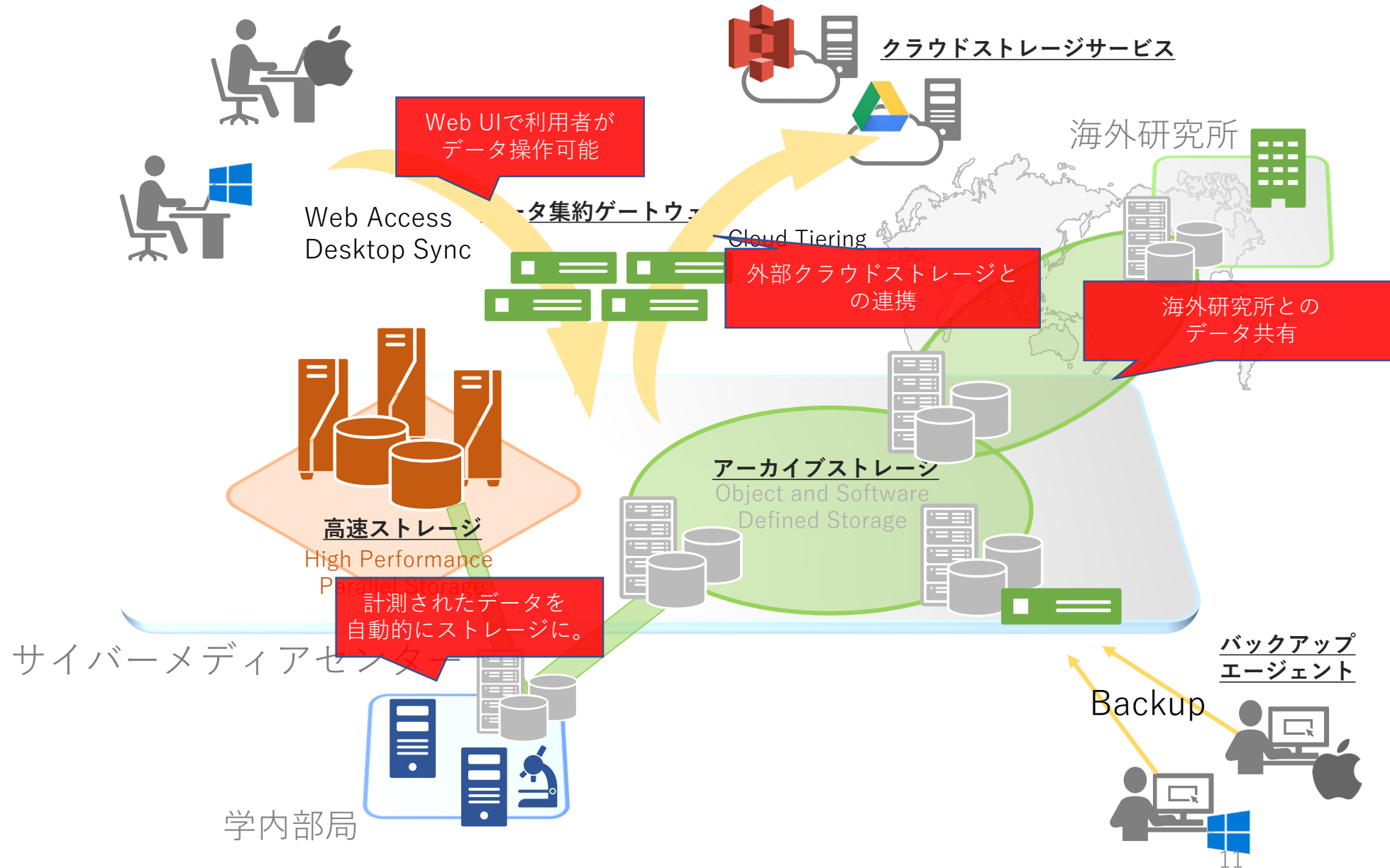
民間クラウド(ストレージ)事業者

データ集約基盤 ONION

並列ファイルシステム

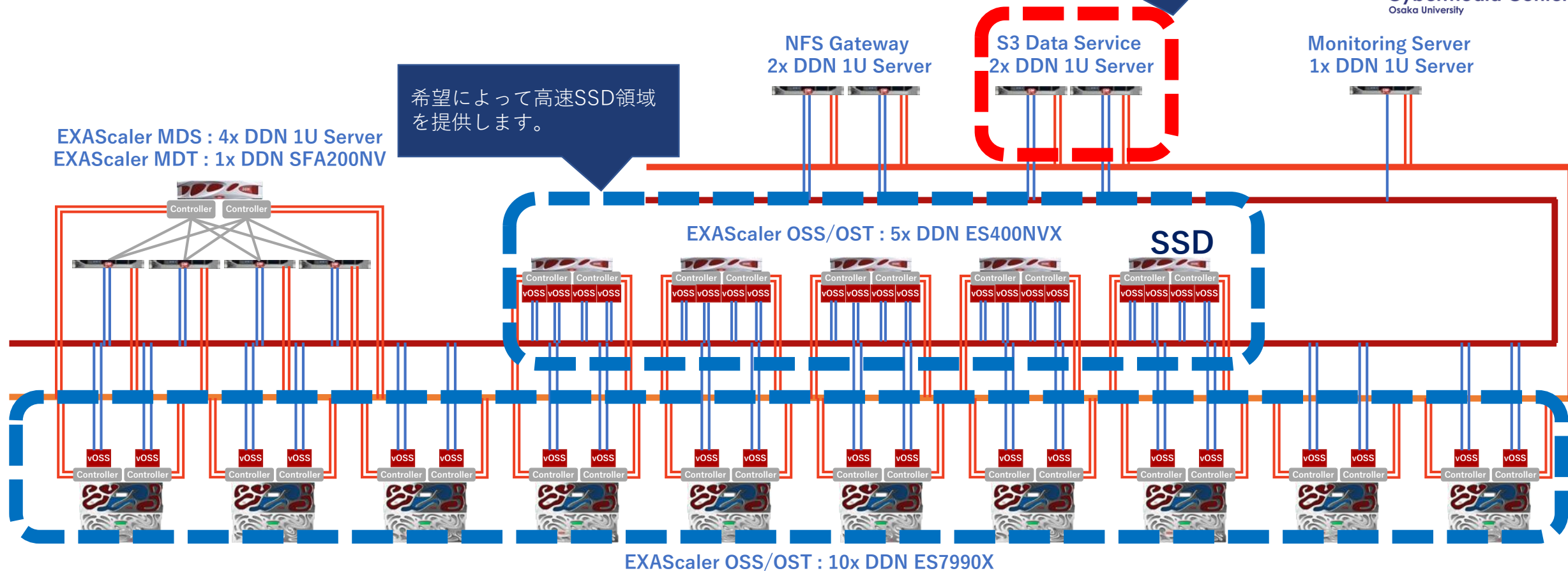
大容量高速ストレージ
20+1 PB (HDD+SSD)

データ集約基盤 ONION 概要



ONION構成Filesystem (DDN ExaScaler)

S3を用いて、S3対応ストレージとのデータ移動ができるようになります。

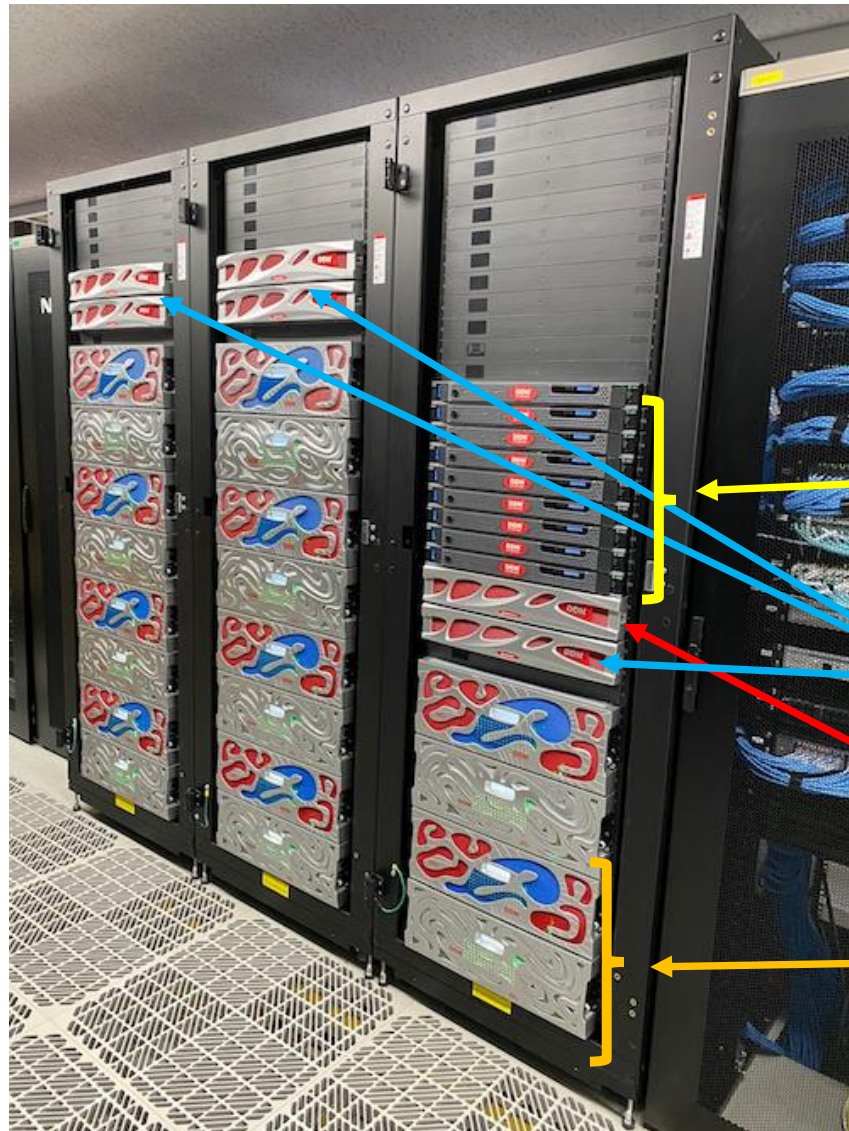


実効容量(HDD)	20.00PB
実効容量(NVMe)	1.20PB
最大合計inode数	約80億個
最大想定実効スループット(HDD)	160GB/s以上
最大想定実効スループット(NVMe)	Write : 160GB/s以上, Read : 180GB/s以上



ONION構成Filesystem (DDN ExaScaler) 外観

- 20PB(HDD)+1PB(SSD)PBを3ラック
- 1700本のHDD(16TB)
- 105本のSSD(15.36TB)



モニタリングサーバ (1個)
NFSサーバ (2個)
S3DSサーバ (2個)
MDSサーバ (4個)

SSD用ES400NVX (5個)

MDT用SFA200NV(1個)

HDD用ES7990X (10個)
HDD用SS9012 (10個)



2021年3月1日頃

スーパーコンピュータSQUID利用説明会(スーパーコンピュータ経験者向け) 2021.05.13

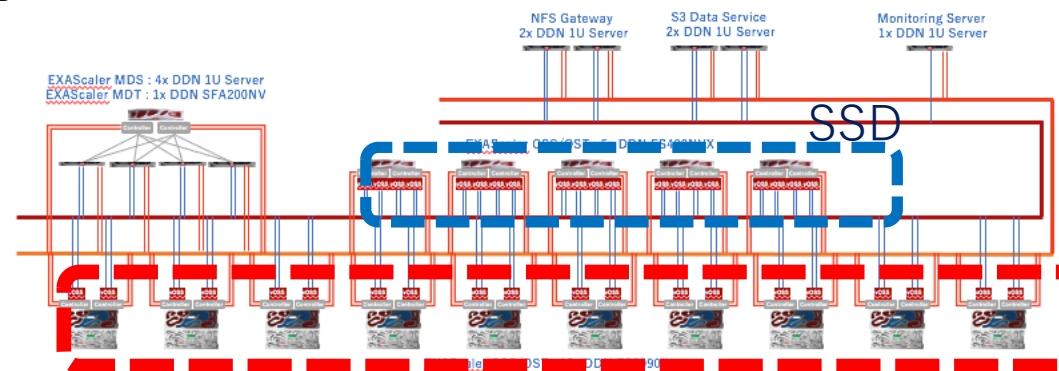
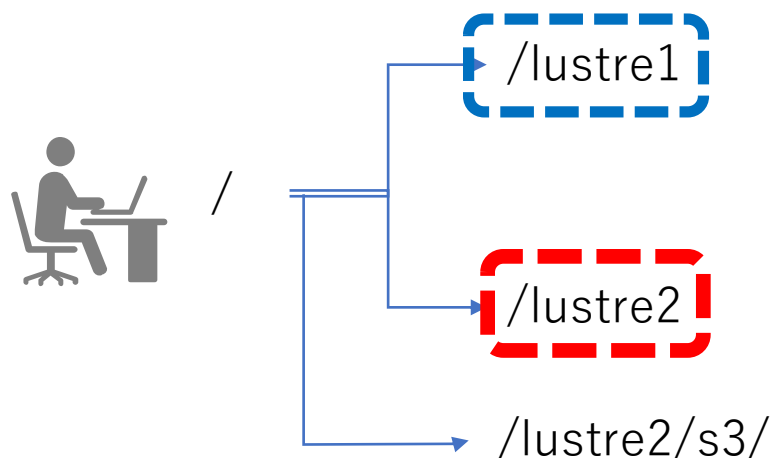
2021年3月3日頃

ONION構成Filesystem

• Lustreファイルシステム構成

領域	ファイルシステム名	MDS	MDT数	ファイル数	DNE	OSS数	OST数	容量	機器
SSD	lustre1	#1-2	1	約20億	No	40	40	約1.2PB	ES400NVX
HDD	lustre2	#3-4	3->2	約60億	Yes	80	80	約20PB	ES7990X

- SSD, HDD毎にファイルシステムを構成
- HDDは5TBをデフォルトで利用者申請グループ単位で提供
- SSDを希望される利用者には1TB単位で提供
- ディレクトリの違いでSSD/HDDを利用者は区別



ONION構成ObjectStorage (Clouidian HyperStore)

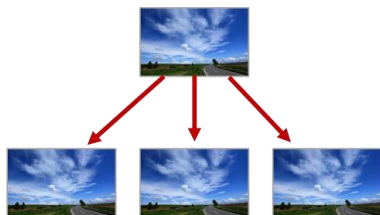
• ONIONを支える基盤技術

並列ファイルシステムに利用者が慣れ親しんだPOSIXのデータアクセス方法で、外部のクラウドストレージ等のS3対応ストレージのデータを扱える仕組みを補強

- S3対応型オブジェクトストレージ
- 高い拡張性
- 単一障害点のない完全分散処理
- 高い保護強度
 - イレジャーコーディング(DC内)
 - レプリケーション(DC内, DC間)

レプリカ（複製）方式

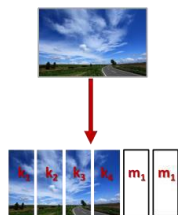
比較的小さいデータやアクセス性を確保する場合



3 個のレプリカ = 3 倍の格納領域

イレジャー・コーディング（EC）方式

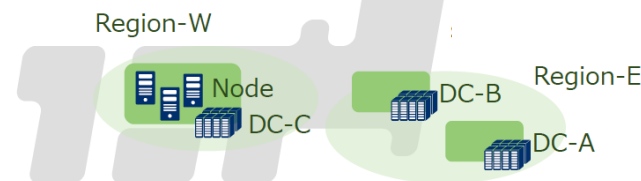
大きなデータや格納効率を確保する場合



4 + 2 コーディング = 6/4または
1.5倍の格納領域

HyperStoreの概念

- Node : HyperStoreサーバ
- Data Center: 複数のサーバの集合
- Region: 複数のDCの集合



現在のIT環境



S3 APIに対応する
各種アプリケーション

SMB/NFS/iSCSI
→ S3 API 変換

- キャッシュ
- ファイル共有同期
- ファイルバックアップ
- 仮想サーババックアップ
- 圧縮、重複排除
- ファイル管理
- ファイルアーカイブ
- 暗号化
- ファイル仮想化
- ファイル階層化

S3 API

高信頼性

複数サーバ、データセンターに
データを自動的に複製、分散配置

高拡張性

ノードを追加するだけでクラスタ全体容量を拡張

高可用性

一部のサーバ障害時でもシステム無停止



CLOUDIAN HYPERSTORE



クラウド階層化

Amazon S3/Glacier、
他のHyperStoreシステムに
データ階層化

ONION ゲートウェイ : NextCloud

- ONIONを支える基盤技術

ONION構成ファイルシステム(DDN)、ONION構成オブジェクトストレージ(Cloudian)にデータをウェブブラウザを通じてアップロード/ダウンロード可能な機能を補強。外部ストレージ連携機能により、クラウドストレージにも対応。

NextCloudを通じて、
オンプレミスストレージと
クラウドストレージを利用可能



1. エンドユーザ操作で連携



2. エンドユーザ操作で連携



3. エンドユーザ操作で連携



SQUIDの運転計画

- 無料お試し利用

5/6 10:00am: OCTOPUS継続利用者向けSQUIDお試し利用開始 (7末)

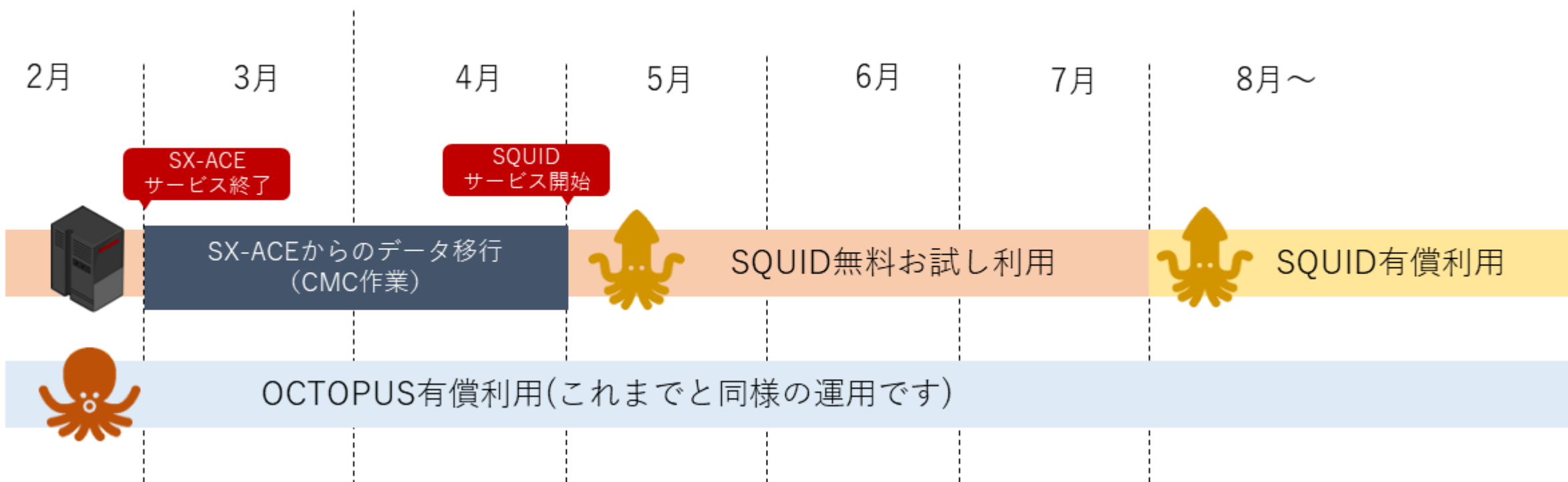
5 中以降 : 新規利用者向けお試し利用開始

- 正式サービス運用 (有料)

5/6 HPCIおよびJHPCN向け正式運用

8月: 全ての利用者を正式サービス運用

無料お試し利用期間中は無料でご利用いただけますが、センター側でも試行運転、実験を行いますので、全ての計算機資源が利用できない場合があります。



スパコン構築・運用の立場から

- クラウド連動型スーパーコンピュータSQUID
 - わが国の学術・産業を支える研究者による未解決のデータサイエンス問題への探求を支援するため、高性能計算および高性能データ分析分野の多様な計算ニーズを収容可能に設計・構築
- 大阪大学サイバーメディアセンターが”今”導入できる最高のスーパーコンピュータSQUIDをご利用ください。
- そして、創出された成果は些細なことでも構いませんので、是非フィードバックください。（皆様にお使いいただく、スーパーコンピュータをよりよくすることにつながります）
- 情報提供・共有は引き続き重要課題と認識しています。