

ペタフロップス級ハイブリッド型スーパーコンピュータ OCTOPUS: Osaka university Cybermedia center Over-Petascale Universal Supercomputer ~サイバーメディアセンターのスーパーコンピューティング 事業の再生と躍進にむけて~

伊達 進

大阪大学 サイバーメディアセンター

1. はじめに

2017年12月、大阪大学サイバーメディアセンターは、スカラ型スーパーコンピュータ OCTOPUS (Osaka university Cybermedia center Over-Petascale Universal Supercomputer)の運用を開始した。このOCTOPUSは、総理論演算性能 1.46 ペタフロップス (Peta Flops: 浮動小数点演算を1秒間に 1460 兆回実行可能) を有する本センター初のペタフロップス級性能を有するスーパーコンピュータである。

このOCTOPUSの実現には、本センターでスーパーコンピュータシステムの運用に携わる教員や技術職員、仕様策定委員らによる2年以上の議論・検討、提案候補ベンダとの打ち合わせ・情報交換がなされており、関係者らにとってはようやく実現にこぎつけた待望のシステムとなった。本稿の筆者はOCTOPUSシステム導入に際して、仕様策定委員長を務めるなど中心的な役割と責任を担う立場でシステム導入に携わった。自画自賛のようでみっともないが、OCTOPUSの調達に限られた予算内で当初想定していた性能目標も達成でき、また、高性能に安定した素晴らしいシステムとして完成できたと考えている。

本稿では、サイバーメディアセンターの事業記録、また、将来の後任者への引き継ぎ・情報共有という視点も含めて、OCTOPUS導入の背景、OCTOPUS調達概要を記したのち、OCTOPUSの構成・特徴について解説したい。なお、OCTOPUSの構成・特徴についてのみ関心がある方は、本稿の1~3節を読み飛ばし、4節から読んでいただくことをお勧めしたい。

2. OCTOPUS 導入の背景・経緯

OCTOPUS導入以前のサイバーメディアセンターにおいては、ベクトル型スーパーコンピュータとし

てSX-ACE、スカラ型スーパーコンピュータとしてHCC (汎用コンクラスタ: Hanyou Computer Cluster) とVCC (大規模可視化対応PCクラスタシステム: PC Cluster for Large-scale Visualization) がサービス提供されていた。しかし、この当時、スカラ型計算サービスとして提供していたHCC、VCCは、スカラ型スーパーコンピュータの利用を切望する利用者にとって切実なる問題を抱えており、スカラ型スーパーコンピュータの拡充が極めて深刻な課題となっていた。

以下、HCC、およびVCCの利用者視点での問題を整理し、HCCの後継として導入されたスーパーコンピュータOCTOPUSがどのような背景から誕生したのかについて記す。

2.1 HCC に対する利用者の失望

HCCは調達「汎用コンピュータシステム」に基づいて2012年に導入されたHPC用途向けPCクラスタシステムである。当該システムは、NEC製計算機Express5800/53Xh 575 台上で稼働させた仮想Linuxをクラスタシステムとして提供するものであった。Express5800/53Xh自体はIntel Xeon E3-1225v2 (4プロセッサコア)、8GBあるいは16GBの主記憶容量を備えており、導入当時は高性能なハードウェア性能を有していたといえる。また、それらの計算機を接続するネットワーク (相互結合網) についてもInfiniBand QDR (Quad Data Rate) で接続されており、最先端のネットワーク性能であったといえる。

しかし、上述の調達「汎用コンピュータシステム」は、サイバーメディアセンターの財政的な事情 (筆者はスーパーコンピューティング事業に携わっていない時期であったので、詳細な事情はわからないのであるが。) もあったのか、サイバーメディアセンタ

一が全学支援業務として行う、教育、CALL、高性能計算、図書館のすべての要望を取り込んだシステムとして一括調達・導入された。言い換えれば、高性能計算サービスを提供する HCC は、「汎用コンピュータシステム」を構成する”一部のシステム”あるいは”一機能”として導入されたことになる。さらに、その”一部のシステム”あるいは”一機能”は、物理的にシステム資源を切り分けるのではなく、当該調達当時に期待の高かった”仮想計算機技術”によりシステム資源を切り分けられる手法が選択されていた。より具体的・技術的には、「汎用コンピュータシステム」の高性能計算に係る機能は、計算機 Express5800/53Xh 上でオペレーティングシステムとして Windows を稼働させ、仮想化技術 Parallels を導入した上で、当該計算機のもつ 4 コアのうち 2 コアを使う Cent OS 6.1 を稼働させ、それらをクラスタ化することで実現されていた。一方、他の 2 コアは教育、CALL 用途に利用される仕組みとなっていた。

仮想計算機技術は、IaaS (Infrastructure As A Service) 型クラウドを実現する用途にも多く利用されており、今日では個別化した計算環境提供のために高性能計算用途での利用実績も報告がある。しかし、HCC 調達・導入当初においては、仮想クラスタで高性能計算環境を提供するというコンセプトが新しすぎたのか、仮想化技術 Parallels から InfiniBand QDR へのパススルー技術利用時の不安定な挙動、MPI (Message Passing Interface) プログラムの不安定な動作などの数多くの技術問題が発生していた。また、高性能計算用途として、性能面での課題があった仮想 Linux クラスタを提供していたことに対する本センターへの不信感も多く寄せられていた。さらに、「汎用コンピュータシステム」を一括に導入した本センター事業間でも当該システムの資源競合に対する問題(例えば、HCC での計算負荷が高まると、教育系システムの挙動(マウスの動きが遅くなる)にも影響がある等の報告)が発生しており、HCC の安定的かつ継続的な運転が困難な状況が発生してしまっている現状があった。

このように、HCC は安定的に高性能を提供するという、スーパーコンピュータとしての基本的な要件

を満たすことができない状況になってしまっており、スカラ型スーパーコンピュータへ期待する利用者の皆様方を失望させる結果となってしまうていた。

2.2 VCC の規模を超える資源要求

スカラ型スーパーコンピュータである VCC は、2013 年の補正予算「高性能汎用計算機高度利用事業「京」を中核とする HPCI の産業利用支援・裾野拡大のための設備拡充」の一環として導入した計算クラスタシステムである。

VCC は、Intel Xeon E5-2670v2 プロセッサ 2 基、64GB の主記憶が搭載された計算ノード (66 ノード)、Intel Xeon E5-2690v4 プロセッサ 2 基、64 GB の主記憶が搭載された計算ノード (3 ノード) を InfiniBand FDR (Fourteen Data Rate) によるインターコネクで接続したクラスタシステムとなっている。また、当該システムは、59 基の NVIDIA 製 GPU Tesla K20 がリソースプール化されている構成となっており、システムハードウェア仮想化技術 ExpEther [1] により、任意の枚数の GPU を計算ノードに比較的高い自由度をもって割り付けることのできる再構成可能システムとしての特徴を有している。そのため、例えば、1 計算ノードに対し GPU 4 枚、1 計算ノードに対し GPU 2 枚といった構成を、利用者の計算要求に応じた設定が可能である。

本システムの利用者の多くは、本システムの性能に不満があるわけではない一方、(1) 高い並列度数でジョブを実行しようとする際の待ち時間の長さ、(2) 本システムで投入できる並列度数のジョブの大きさ、に対して数多くの不満の声を上げていた。これらは、本システムを構成する計算ノード数が 69 と小さいために、本センターのシステム管理上、ノード数の少ない本システムが並列度数の高いジョブをもつ利用者によって占有されることを防止する観点から、ジョブスケジューラのキュー構成の最大並列度数を小さくしていること、本センターの採用するジョブスケジューラの特性上並列度の小さいジョブが並列度の高いジョブに先行して実行される傾向にあることに起因している問題である。

また、当該システムは、上述したように、該当年度でのみ有効である補正予算での導入であるため、通常、本センターが利用者に計算機利用負担金として課す電気代相当分の費用に加え、導入業者による保守運用費用が利用負担金として課されている。そのため、利用者からは、(3) 利用負担金が高い、という不満の声が寄せられていた。

上述のように、OCTOPUS 導入以前においては、本センターではスカラ型スーパーコンピュータとして提供を行っている HCC、VCC に問題・課題を抱えており、本センターにおける高性能計算機事業の品質向上にはスカラ型スーパーコンピュータの拡充が極めて深刻な課題となっていた。このような背景のもと、「汎用コンピュータ」の契約更改をターゲットとして、新スーパーコンピュータシステム OCTOPUS の導入に向けた検討・議論が 2015 年頃より開始することとなった。

3. OCTOPUS 調達の基本方針

HCC の後継となる OCTOPUS の導入は、前調達「汎用コンピュータシステム」の経験を踏まえ、2 節で記載した問題・課題を解決し、高性能計算用途のみに特化した調達とし、大阪大学だけではなく全国の研究者にご利用いただけるようお願いを込め、調達名を「全国共同利用大規模並列計算システム (Large-scale Parallel Computing System for Nation-wide Shared Use)」と設定し、下記のスケジュールで調達を実施した。

2016 年 2 月 17 日 導入説明会
2016 年 7 月 28 日 仕様書案説明会
2017 年 2 月 6 日 入札官報公告
2017 年 2 月 13 日 入札説明会
2017 年 4 月 3 日 入札期限
2017 年 5 月 31 日 開札
2017 年 12 月 1 日 稼働

本節では、上記 OCTOPUS 調達を通じて検討・重視した基本方針について記す。

3.1 基本方針

前節で記したように、本センターの有するスカラ型スーパーコンピュータは、システムの安定性、提供可能な計算資源量で課題があった。また、利用負担金についての要望も多く寄せられていた。そうした背景から、本システムで設定したシステム調達に係る基本方針として、HCC での安定的な稼働・運転による高性能計算サービスを提供できなかった反省を踏まえ、

- (1) 安定稼働・運転できることを最重要方針と定め、
 - (2) 総理論演算性能 1 ペタフロップス以上であること
- を目指し、本センターの脆弱なスカラ型計算資源の増強・拡充をねらった。

3.2 調達概要

OCTOPUS 調達に際しては、利用者の要望・ニーズを把握し、仕様書への反映を行うため、利用者を対象としたアンケートを実施した。より多く回答数を得られるよう、できるだけ簡易な質問項目に限定し、本センターの Web ページで実施した。

アンケートの質問内容としては、希望するスカラ型計算資源に対して、a. 並列利用できる最大ノード数、b. ノードあたりのコア数、c. ノード(コア)あたりのメモリ、d. 利用できるアクセラレータ、のうち、優先する利用基準について、希望するアクセラレータについて、希望するアプリケーションについてを中心として設定した。

その結果、コア数よりもノード数を優先し、メモリーコア型プロセッサおよび GPU については「希望する」「希望しない」がおおよそ半数程度となる結果となった。アプリケーションについては、特定のアプリケーションに対する希望が集中するのではなく、様々なアプリケーションに対する希望があったことがわかった。アンケートのより詳細については、参考文献 [2] を参照されたい。

3.1 節に記した基本方針と利用者ニーズに基づき、新システム OCTOPUS に向けた調達では、調達予算と性能のトレードオフを考慮しつつ、

- * 利用者の多様な計算ニーズを収容可能とするため、単一構成とするのではなく、複数種の計算ノード群からなるハイブリッド型スーパーコンピュータとする
- * 最大利用できるプロセッサコア数よりも最大利用できるノード数を優先する
- * 導入するハードウェア資源であるプロセッサ、アクセラレータが継続的に安定した高性能を提供できるよう冷却方式は、空冷よりも液冷を優先する
- * アプリケーションはオープンソース系ソフトウェアを拡充する
- * ストレージは 3 ペタバイト (PB) 以上を確保する
- * 計算クラスタを構成するノードおよびファイルサーバを 100 Gbps 以上のネットワークで接続すること。任意のノード間でフルバイセクション帯域幅を有する構成、トポロジであることが望ましい

を調達の中核とする仕様策定を行なった。

なお、調達は総合評価方式を採用しており、プロセッサ性能のよいプロセッサを搭載したノード数を多く含むシステム提案、冷却方式として水冷方式を採用したシステム提案、計算ノード間の接続がフルバイセクション帯域幅を有する構成・トポロジであるシステム提案、大容量主記憶搭載ノードに搭載される主記憶容量が大容量であるシステム提案がより高得点となるよう設定した。その結果、本センターの利用できる予算内で、提案候補ベンダ間の競争も働き、本センターの利用者の望む、ノード数を多く搭載し、多様な計算ニーズを収容可能な総理論演算性能 1.46 ペタフロップスのハイブリッド型スーパーコンピュータの調達が実現できたと考えている。

4. OCTOPUS

4.1 OCTOPUS 名称の由来

OCTOPUS は、1 節に示した通り、Osaka university Cybermedia cenTer Over-Petascale Universal Supercomputer を略したシステム名称である。本システム名称の背景には、本システムの調達に際して、

1 ペタフロップス以上の性能を有することを目指したこと (over-petascale)、および、幅広い利用者の方に利用されるようにとの思い (universal) が込められている。

また、本センターのこれまでのスーパーコンピュータシステムの名称は、お世辞にも愛着があり覚えやすいネーミングとは言えなかった。OCTOPUS の名称は、そのような観点も考慮されており、大阪から連想されやすいたこ焼き、たこ焼きといえバコという関連性も採用の一因である。

本稿の読者の皆様方には、本稿を通じて、少なくとも「大阪大学サイバーメディアセンターには OCTOPUS というスーパーコンピュータがある」とだけでも理解いただければありがたい。すでにご利用いただいている方には、皆様方の研究成果をまとめられた論文の謝辞欄に OCTOPUS を利用した旨を記載いただき、皆様方の研究成果とともに OCTOPUS のプレゼンス向上に貢献いただければ幸いである。

なお、トリビア的になるが、本システム名称については、3 節で記した、利用者に向けたアンケートの時点で codename: octopus と称して一部公開を行っていた。読者のなかには早くから次期システム名称が OCTOPUS であることに気づかれた方がおられたかもしれない。

4.2 OCTOPUS 計算ノード概要



図 1 OCTOPUS 外観

さて、ようやく本特集の目玉部分である OCTOPUS について説明する。図 1 は本センター IT コア棟サーバ室に設置された OCTOPUS の外観である。上図が IT コア棟 3 階見学室側の外観 (表面)、下図が見学室側から遠い側 (裏面) の外観である。表面には、OCTOPUS の導入・運用開始を記念し、2017 年 8-9 月に実施した OCTOPUS ラックデザインコンテスト [3] の最優秀受賞作品 (図 2) である宮本要子氏のデザインが描画されている。なお、ラックデザインは、「大阪」、「皆に愛される」、「最先端の科学研究」、「スーパーコンピュータ」のキーワードより一般公募を行っている。



図 2 OCTOPUS ラックデザイン

OCTOPUS は、IT コア棟内に区画分けされている 1 区画内の 11 ラック (分電盤ラックを除く) に搭載されている。

表 1 OCTOPUS ノード構成

ノード構成	汎用 CPU ノード 236 ノード(471.24 TFlops)	プロセッサ：Intel Xeon Gold 6126 (Skylake / 2.6 GHz 12 コア) 2 基 主記憶容量：192 GB
	GPU ノード 37 ノード(858.28 TFlops)	プロセッサ：Intel Xeon Gold 6126 (Skylake / 2.6 GHz 12 コア) 2 基 GPU：NVIDIA Tesla P100 (NV-Link) 4 基 主記憶容量：192 GB
	Xeon Phi ノード 44 ノード(117.14 TFlops)	プロセッサ：Intel Xeon Phi 7210 (Knights Landing / 1.3 GHz 64 コア) 1 基 主記憶容量：192 GB
	大容量主記憶搭載ノード 2 ノード (16.38 TFlops)	プロセッサ：Intel Xeon Platinum 8153 (Skylake / 2.0 GHz 16 コア) 8 基 主記憶容量：6 TB
相互結合網	InfiniBand EDR (100Gbps)	
ストレージ	DDN EXAScaler (Lustre/3.1PB)	

図 1 の写真上部にはガラス扉が見えるが、このガラス扉で区切られた空間 (キャッピング空間) を共有するラックに OCTOPUS を構成する機器群は分散搭載されている。具体的には、表面側に空冷 4 ラック、裏面側に水冷 7 ラックの構成である。このキャッピング空間により、それら搭載機器類の発熱を閉じ込め、サーバ室内の冷えた空気と機器類によって温められた空気が混ざらないようにすることで、IT コア棟内の空調効率を高め、低消費電力で高性能を提供する工夫がなされている。

次に、OCTOPUS のシステム構成について解説する。表 1 にノード構成について示す。OCTOPUS は汎用 CPU ノード群 (236 ノード)、GPU ノード群 (37 ノード)、Xeon Phi ノード (44 ノード)、大容量主記憶搭載ノード (2 ノード) の総計 319 ノードから構成される。また、利用者のプログラム開発環境として、フロントエンドサーバ群を 4 基備えている。以下、各ノード群について概説する。

4.2.1 汎用 CPU ノード



図3 汎用 CPU ノード (LX 2U Twin2 Server 406Rh-2) の外観

汎用 CPU ノードには、2017 年後半にリリースされたばかりの最新プロセッサ Intel Xeon Gold 6126 (Skylake, 2.6 GHz, 12 コア) を 2 基、192 GB の主記憶が搭載される。このプロセッサは 1 基あたり 0.9984 テラフロップス (TFlops: 1 TFlops は 1 秒間に 1 兆回の浮動小数点演算を行うことができることを表す) の理論演算性能を供給する。そのため、ノード単体で 1.9968 TFlops の理論演算性能を有しており、OCTOPUS の汎用 CPU ノード群全体で 471.24 TFlops の理論演算性能、45.312 TB の主記憶容量を提供可能である。

この Skylake 世代のプロセッサの最大の特徴は、AVX-512 命令セットがサポートされている点にある [4]。これにより、SIMD (Single Instruction Multiple Data) ベクトル幅が 512 ビットに拡張されており、以前の命令セットでは不可能あるいは困難であったベクトル化が可能となり、並列処理の向上が可能となっている。

なお、CPU ノードは、NEC LX 2U Twin2 Server 406Rh-2 という型番のサーバとして供給され、2 U の筐体に 4 ノードが搭載される構成となっている (図 3)。

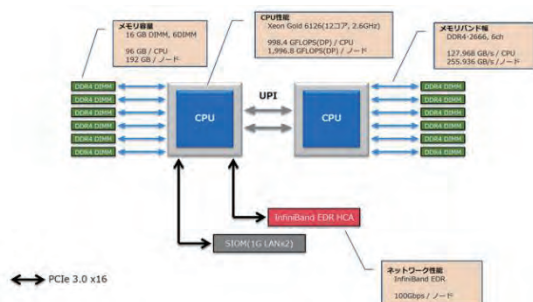


図4 汎用 CPU ノードのブロック図

図 4 に汎用 CPU ノードのブロック図を示す。上述したように、汎用 CPU ノードでは、2 基の Skylake 世代のプロセッサが搭載されている。これらの 2 基のプロセッサソケット間は 2 本の UPI (Ultra Path Interconnect) で接続されており、10.4 GT/s (Gigatransfers per second) \times 2 での通信速度で CPU 間のデータ交換が可能である。また、各 CPU は 6 チャンネルメモリコントローラを通じて 6 枚の DDR4 DIMM-2666 (16GB) と接続されている。そのため、各 CPU は主記憶装置に対して、127.968 GB/s (2666 M \times 8B \times 6 チャンネル) でアクセスが可能であり、ノード全体で 255.936 GB/s のメモリバンド幅を提供する。さらに、CPU ノードは InfiniBand EDR HCA (Host Channel Adapter) を備えるが、InfiniBand EDR HCA は CPU ノード内の 1 CPU に PCI Express 3.0 (Gen3) x16 で接続されており、100 Gbps のデータ送受信が問題なく行える構成となっている。

4.2.2 GPU ノード

図 5 に示す GPU ノードは、CPU ノードと同一のプロセッサ、主記憶を搭載し、アクセラレータとして 4 基の NVIDIA 製 GPU Tesla P100 [5] を備えている。4 基の Tesla P100 は NVIDIA の高速インターコネクト技術 NVLink によって相互接続されており、GPU 間で最大 80 GB/sec の通信を可能にしている。この NVLink に対応した Tesla P100 は、16 GB の HBM2 (High Bandwidth Memory 2) メモリを備えており、理論演算性能 5.3 TFlops (倍精度) を提供する。また、Tesla P100 では、732 GB/s のメモリバンド幅での高速アクセスが可能となっている。



図5 GPU ノード (LX 1U 4 GPU Server 102Rh-1G) の外観

このように、GPU ノード単体では、上述の汎用 CPU ノードの演算性能に加えて、4 基の Tesla P100 の演算性能が利用可能であるため、23.196 TFlops の理論演算性能を有しており、GPU ノード群全体では 858.28 TFlops の理論演算性能、7.252 TB の主記憶容量、2.368 TB の GPU メモリを提供可能である。なお、この GPU ノードは、NEC LX 1U 4GPU Server 406Rh-2 という型番のサーバとして供給され、1U の筐体に 1 ノードが搭載される構成となっている (図 5)。図中やや中央部の 4 つの箱状の部分が Tesla P100 が実装されている部分である。

図 6 に GPU ノードのブロック図を示す。CPU、主記憶、InfiniBand EDR HCA を接続する構成は、上述した CPU ノードと同じ (CPU 間の UPI 接続を除く) であるので、ここでは説明を割愛する。GPU ノードの最大の特徴は、名称が示す通り、4 基の GPU を搭載していることである。前述したが 4 基の GPU 間は、図に示すように、各 GPU から 4 本の NVLink で相互に接続されており、GPU 間では最大 80 GB/sec の通信が可能となっている。また、各 GPU は PCIe ブリッジを経て PCI Express 3.0 (Gen3) x16 で接続されており、CPU-GPU 間の通信は 16 GB/sec で実現される。

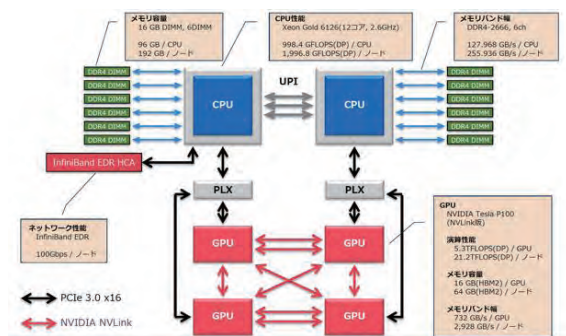


図 6 GPU ノードのブロック図

4.2.3 Xeon Phi ノード

Xeon Phi ノードは、その名前が示す通り、Intel 製 Xeon Phi プロセッサを搭載した計算ノードである。Xeon Phi は、本調達以前の段階では、MIC (Many Integrated Core) アーキテクチャベースのコプロセッサ (co-processor: Knights Corner) としてリリースされていたが、OCTOPUS 調達時には、プロセッサ

製品も利用可能となっていた。OCTOPUS では、利便性の観点から、メニーコア型プロセッサとして Xeon Phi を選択している。Xeon Phi ノードでは、Intel Xeon Phi 7210 (Knights Landing, 1.3Ghz, 64 コア) を 1 基、192 GB の主記憶を搭載している。この Xeon Phi 7210 [6] は、単体で理論演算性能 2.662 TFlops の性能を有しており、Xeon Phi ノード全体では 117.14 TFlops、8.448 TB の主記憶容量が利用可能である。



図 7 Xeon Phi ノード (Express5800/HR10c-M) の外観

さらに、この Xeon Phi 7210 プロセッサには MCDRAM (Multi-Channel Dynamic Random Access Memory) (オンパッケージ HBM メモリ) を 16 GB 搭載しており、フラットモードで利用することによりさらに大規模なメモリ空間を利用可能である。本稿執筆時には、OCTOPUS の MCDRAM 設定はフラットモードとして運用している。この MCDRAM 設定はユーザの要望に応じて、他のキャッシュモード、ハイブリッドモードに適宜変更することも可能であるので、リクエストのある読者の方がおられたらお知らせいただければ幸いです。なお、Xeon Phi ノードでは、Express5800/HR10c-M という型番のサーバとして供給され、2U の筐体に 4 ノードが搭載される高密度構成となっている (図 7)。

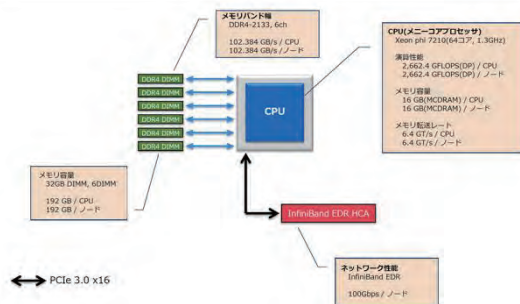


図 8 Xeon Phi ノードのブロック図

図 8 に Xeon Phi ノードのブロック図を示す。Xeon Phi ノードでは、Xeon Phi 7210 プロセッサは 6 チャンネルメモリコントローラを通じて 6 枚の DDR4 DIMM-2133 (64GB) と接続されている。そのため、各 CPU は主記憶装置に対して、102.384 GB/s (2133 M × 8B × 6 チャンネル) でアクセス、すなわち、102.384 GB/s のメモリバンド幅を提供する。CPU、GPU ノードと同様に、Xeon Phi ノードについても、InfiniBand EDR HCA を装備しており、プロセッサとは PCI Express 3.0 (Gen3) x16 で接続されており、ノード内外の通信にストレスなく通信が可能な設計となっている。

4.2.4 大容量主記憶搭載ノード



図 9 大容量主記憶搭載ノード (LX116Rh-7) の外観

大容量主記憶搭載ノードは、その名前の通り、1 ノードあたり 6TB の大容量主記憶を搭載した計算ノードである。本ノードには、Skylake 世代の最新プロセッサ Intel Xeon Platinum 8153 (Skylake, 2.0GHz, 16 コア) が 8 基搭載されており、最大 128 コアでの並列処理が可能となっている。大容量主記憶搭載ノード単体では、8.192 TFlops の理論演算性能を提供可能である。OCTOPUS では、この大容量主記憶搭載ノードを 2 基備えている。なお、大容量主記憶搭載ノードは、7 U の筐体の NEC LX116Rh-7 という型番のサーバとして供給されている (図 9)。

図 10 に大容量主記憶搭載ノードのブロック図を示す。大容量主記憶搭載ノードに搭載されているプロセッサ Intel Xeon Platinum 8153 は、CPU ノード、GPU ノードに搭載されるプロセッサよりも上

位クラスに位置づけられ、8 ソケット構成を可能としている。大容量主記憶搭載ノードでは、このプロセッサを 8 基備え、各 CPU が 3 本の UPI で相互に接続されている。また、各 CPU は 6 チャンネルメモリコントローラを通じて 12 枚の DDR4 DIMM-2666 (64GB) と接続される。そのため、各 CPU は主記憶装置に対して、127.968 GB/s (2666 M × 8B × 6 チャンネル) でアクセスが可能であり、ノード全体で 1023.744 GB/s のメモリバンド幅を提供する。また、ノードあたりの総メモリ容量は 6.144TB となる。さらに、図には記載していないが、汎用 CPU ノード、GPU ノード、Xeon Phi ノードと同様に、InfiniBand EDR HCA を装備しており、プロセッサとは PCI Express 3.0 (Gen3) x16 で接続されており、ノード内外の通信にストレスなく通信が可能な設計となっている。

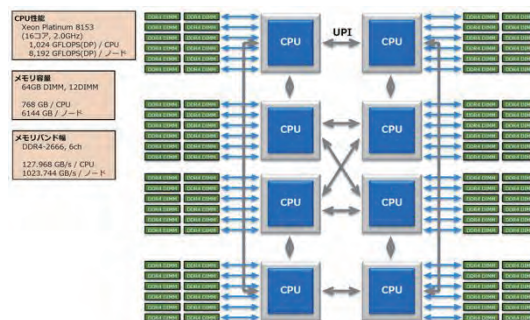


図 10 大容量主記憶搭載ノードのブロック図

4.2.5 フロントエンドノード



図 11 フロントエンドノード(NEC LX 4U-GPU Server 108Th-4G)の外観

OCTOPUS では、利用者の皆様方がプログラム開発を行う環境としてフロントエンドノードを 3 基準備している。図 11 にフロントエンドノードの外観を示す。このフロントノードでは、OCTOPUS の

計算ノードでのプログラム開発が容易に行えるように、最もノード数の多い汎用 CPU ノードと同一のプロセッサである Intel Xeon Gold 6126 を 2 基、主記憶を 192 GB 備えている。

図 12 にフロントエンドノードのブロック図を示す。フロントエンドノードの内部構成は、汎用 CPU ノードのそれと類似するが、CPU 間の接続が 3 本の UPI でなされている点が異なる。また、フロントエンドノードには、NVIDIA Quadro P5000 を搭載している点も異なる。このフロントエンドノードは、計算ノード群（汎用 CPU ノード、GPU ノード、Xeon Phi ノード、大容量主記憶搭載ノード）との通信用に InfiniBand EDR HCA を備えている点はいずれの計算ノードとも同一の点であるが、10GBase-SR を 2 ポート備えている。この 10GBase-SR の 2 ポートは、大阪大学の学内ネットワークである ODINS に接続されており、20 Gbps での高速通信を可能としている。なお、このフロントエンドサーバには、NEC LX 4U GPU Server 108Th-4G という型番が付与されている。

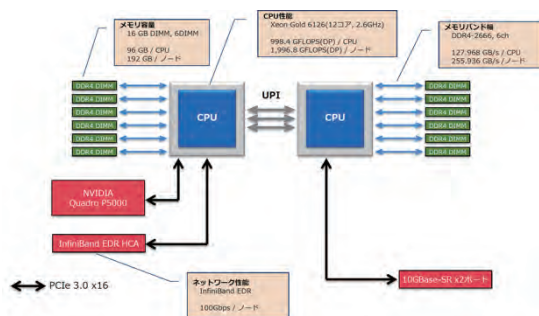


図 12 フロントエンドノードのブロック図

4.3 大容量ストレージ概要

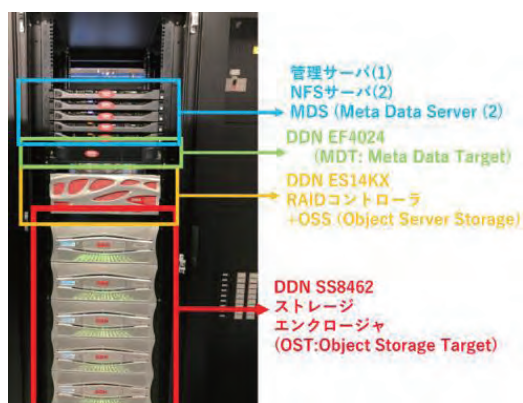


図 13 大容量ストレージの外観

次に OCTOPUS の大容量ストレージについて概説する。図 13 に OCTOPUS の大容量ストレージ外観を示す。

OCTOPUS の大容量ストレージは、3 PB を並列ファイルシステム Lustre ベースのファイルシステムである DDN (DataDirect Networks) 製 ExaScaler [7] を基軸として提供する。OCTOPUS では、わずか 1 ラックに 3 PB を高密度に収容している。

OCTOPUS の大容量ストレージは、ExaScaler ファイルサービス (Lustre ファイルサービス) 実現のためのサーバ群 (Object Storage Server (OSS)、Object Storage Target (OST)、Meta Data Server (MDS)、Meta Data Target (MDT))、ストレージ管理サーバ、および、NFS サーバより構成されている。これらは図 14 に示す通り、相互結合網である InfiniBand EDR ネットワークに接続されており、OCTOPUS への高速なファイルサービスを提供している。

Lustre ファイルシステムは、実際のデータであるファイルと、ファイルに付属するデータ (メタデータ) を分離して管理する点に特徴があり、次の通り動作する。いまクライアントがファイルにアクセスしようとする場合を考える。この時、クライアントはアクセス対象とするファイルがどこに格納されているかといったメタデータが格納された MDT を管理する MDS に問い合わせる。問い合わせを受けた MDS は、アクセス対象のファイルに関するメタデータをクライアントに返す。クライアントは、そのメタデータに基づき、実際のデータが含まれる OST を管理する OSS にアクセスする。

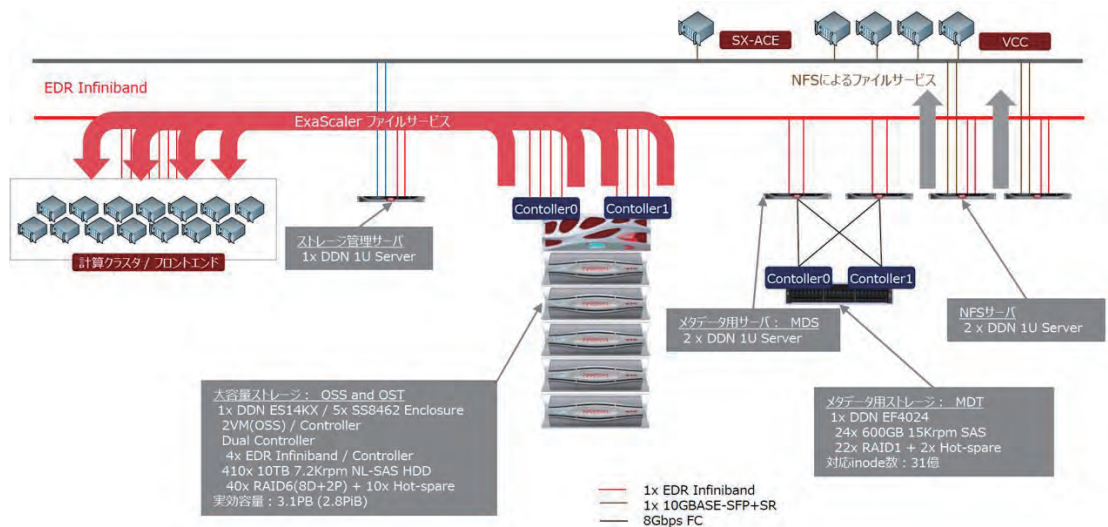


図 14 OCTOPUS ファイルシステムの概要

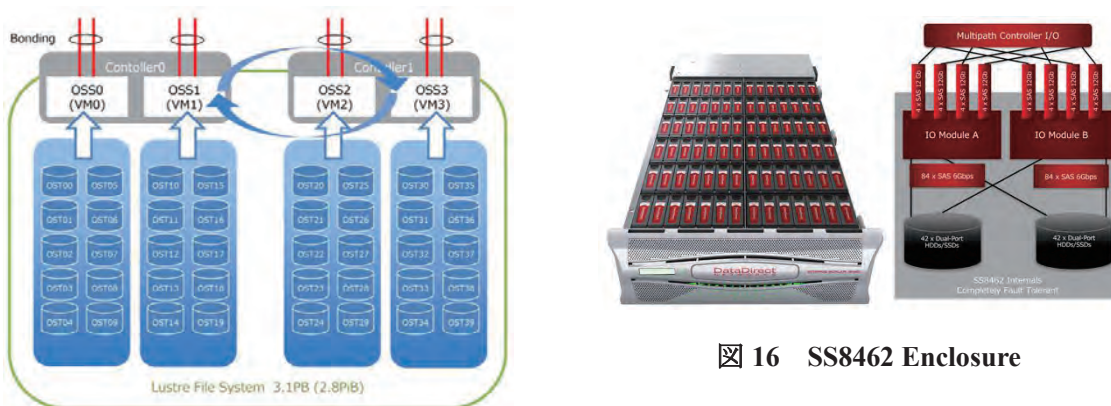


図 15 ES14KX

OCTOPUS の大容量ストレージでは、Lustre ファイルサービスの OSS 機能を DDN 製 ES14KX で提供する。ES14KX では、2 基の RAID コントローラそれぞれに Active/Active の冗長構成設定をされた 2 VM (virtual machine: 仮想計算機) を配備する。この ES14KX は 8 本の InfiniBand EDR コネクションで相互結合網に接続されており、各 VM が 2 本ずつをボンディングし、耐故障性、負荷分散性能を高めたうえで、Lustre ファイルサービスを提供する構成となっている。(図 15)。

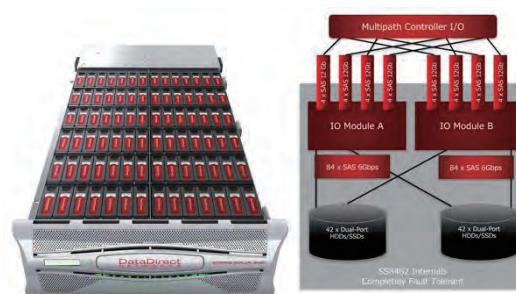


図 16 SS8462 Enclosure

また、OST としては、5 基の DDN SS8462 Enclosure を設置している。この Enclosure は 4U の筐体内に 84 個のドライブを収容可能な高密度ストレージエンクロージャ (図 16) であり、SSD、SAS、SATA の各種ドライブを混在可能である。また、このエンクロージャでは、すべての構成要素が冗長化されており、高い耐障害性を有する構造となっている。OCTOPUS では、10 TB の NL-SAS 7/2krpm HDD 410 個 (82 個 × 5) を 5 基の SS8462 Enclosure 内に格納している。OCTOPUS の Lustre ファイルサービスでは、これら 5 基の

SS8462 エンクロージャー内に格納されている 10 本のディスクを用いて RAID6 (8D+2P) で構成し、各々 RAID 6 ボリュームを LUN (Logical Unit Number) として管理する(図 17)。この筐体間にまたがった RAID6 構成により、OCTOPUS の Lustre ファイルサービスでは、1 エンクロージャーがアクセス不能になったとしてもデータが保護される耐障害性を持たせている。

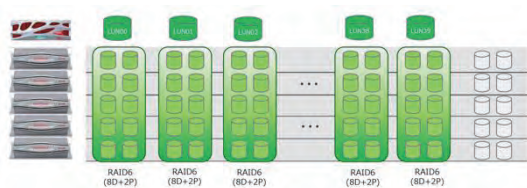


図 17 ストレージ構成

Lustre ファイルサービスを実現する上での重要な役割を担う MDS については、プロセッサとして Intel Xeon E5-2640v4 (2.4GHz, 10 コア) を 2 基、主記憶として 128 GB を備えた 1 U サーバを 2 ノード設置している。各 MDS は、InfiniBand EDR 2 本をボンディングすることにより相互結合網に接続され、Active/Standby の冗長構成により動作する構成となっている。また、実際のメタデータを格納する MDT には DDN EF4024 が採用されており、MDS とは 8 Gbps FC にて冗長構成を持たせて接続される構成となっている。このように、OCTOPUS では、メタデータアクセスについても高い耐障害性を提供し、ユーザのファイルアクセスへの信頼性を高めている。

さらに、補足的ではあるが、OCTOPUS の大容量ストレージでは、近い将来に Lustre ファイルサービスを本センターの他のスーパーコンピュータから利用できるよう、2 基の

NFS サーバを導入している。これらの NFS サーバもまた相互結合網、センター内のスーパーコンピュータネットワークに 2 本ずつのネットワークコネクションをもたせた冗長構成としている。

最後に、ストレージ管理サーバであるが、この管理サーバ上にはファイルシステムのモニタリング機能を搭載しており、日々の運用状況の監視に利用されている。

4.3 相互結合網 (インターコネクト)

次に、OCTOPUS の相互結合網 (インターコネクト) について説明する。本節以前に何度か記載したように、OCTOPUS では、InfiniBand EDR を相互結合網 (インターコネクト) として採用している。OCTOPUS を構成する計算ノード群 (汎用 CPU ノード群、GPU ノード群、Xeon Phi ノード群、大容量主記憶搭載ノード群)、ストレージ、フロントエンドサーバ (計算ノードから合計 319 本、ストレージより合計 18 本、フロントエンドノードより 3 本の合計 340 本) を収容するために、OCTOPUS では Mellanox 製 CS7500 InfiniBand Director Switch [8] を導入している (図 18)。

通常、高性能計算クラスタを構成する計算ノード間をインターコネクト接続する際、その計算ノード間の通信が他の計算ノード間の通信と干渉しないよう、また、その通信も低遅延で行えるよう、トポロジ構成、帯域等を熟慮する必要がある。また、インターコネクトを構成するスイッチ台数の増加は調達コストを増加させ、一方で計算ノード数を減少させてしまう。そのため、予算的な側面からの考慮も必要となる。

事実、OCTOPUS 調達においても、このインターコネクトにかかる部分の記載は、導入

されるノード数が仕様書記載の段階では(仕様書ではノード数増加に応じた加点設定としていたため)未定であったこと、また、同様に 100 Gbps の通信性能を有するインターコネクト技術 Intel Omni-Path の提案も可能としたこと (InfiniBand EDR と Omni-Path で利用できるスイッチのポート数が違う：それぞれ 36 ポートと 48 ポート) などもあり、困難を極めた部分であった。調達当初はこれら 36 ポートあるいは 48 ポートのスイッチを用いたファットツリーあるいはハイパーキューブなどのトポロジを想定し、任意のノード間でフルバイセクション帯域幅を有し、低遅延での通信が可能となるインターコネクト構成を想定していたが、OCTOPUS の調達では、非常に高価であるため想定範囲外であった Mellanox 製 CS7500 InfiniBand Director Switch が導入されることとなった。



図 18 CS7500 InfiniBand Director Switch

この CS7500 は、28 U の筐体に 648 EDR ポートを備え、最大 130 Tb/s のノン・ブロッ

キング帯域幅を有し、そのポート間は 0.5 μ s (マイクロ秒) 未満での通信が可能である” モンスタースイッチ “である。現状では、OCTOPUS の CS7500 は、OCTOPUS を構成する計算ノード、ストレージ、フロントエンドノード全てを直取 (この状態では、最大約 70 Tb/s のノン・ブロッキング帯域幅を供給) しており、上述のノード同士が互いに 1 ホップでの低遅延通信を行うことが可能となっている。

4.4 水冷方式

次に、OCTOPUS の調達導入に際して、こだわった部分でもあるサーバの冷却方式について記載する。

3 節で記したが、OCTOPUS 調達の基本方針として、「安定稼働・運転できること」を定め、具体的にはプロセッサ、アクセラレータに液冷を優先することとした。この背景としては、2013 年に竣工した IT コア棟の水冷設備の活用という視点もあるが、プロセッサやアクセラレータの温度上昇による動作周波数低下による性能低下を回避した安定稼働を行いたいという願いがあった。さらに、仕様書策定時に、ターゲットとなりうるプロセッサやアクセラレータの消費電力がそれぞれ 130 W、300 W 前後であろう、という情報を得ていたことは、IT コア棟の空調だけではなく、プロセッサおよびアクセラレータの直接液冷にこだわる主な要因であった。

このこだわりの結果、OCTOPUS の調達では、汎用 CPU ノード、GPU ノード、Xeon Phi ノードの計算ノード群については、空気よりも比熱の高い水を用いて、これらのノード群に搭載されたプロセッサ、アクセラレータで発生した熱を取り去る直接水冷方式を行う

Asetek 社の Asetek Rack CDU Direct-to-Chip (D2C) [9][10]を採用することとなった。以下、この水冷の仕組みについて概説する。

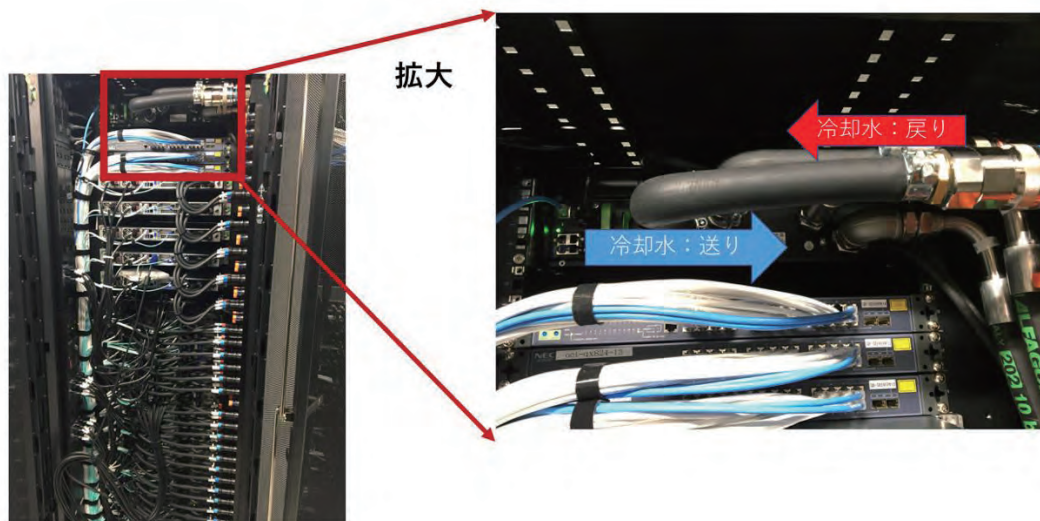


図 19 OCTOPUS 水冷ラック背面

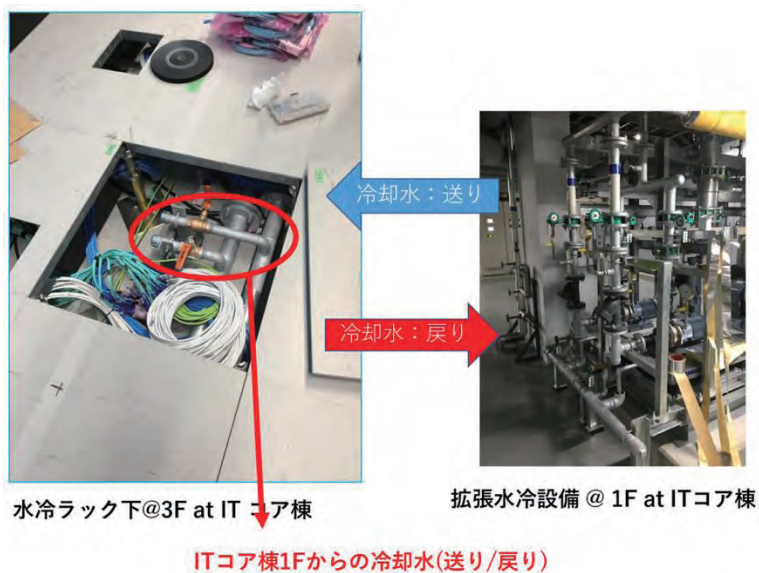


図 20 IT コア棟水冷設備から OCTOPUS への冷却水

図 19 に OCTOPUS の水冷ラック背面を示す。左図を見ると、ラック内のサーバ筐体に対して黒いホースが何本も挿入されていることがわかるが、このホースを使ってサーバ筐体内のプロセッサ、アクセラレータに冷却水を送り、プロセッサ、アクセラレータの熱で温まった水を回収する仕組みとなっている。この冷却水の送り出しと回収をラック単位で行うために、冷却水循環装置 (Coolant Distribution Unit: CDU) がラック上部に配備されており、IT コア棟水冷設備との間で冷却水と暖められた水を交換している (図 19 右)。

なお、プロセッサ、アクセラレータの直接冷却は、図 21、図 22、図 23 に示すコールドプレートと呼ばれるものをプロセッサ、アクセラレータ表面に装着することで行なっている。



図 21 CPU 用コールドプレート

図 20 は IT コア棟水冷設備と水冷ラックへの循環の仕組みを示したものである。左図は設置工事中の水冷ラックの床面を撮影したものである。左図の赤丸で囲んだ部分は、右図の水冷設備との間で冷却水および暖められた水を循環させるための配管である。この配管を各ラックの CDU と接続することで、IT コア棟 1 階冷却設備－IT コア棟 3 階のラック CDU－サーバの冷却水循環を行う。これにより、プロセッサ、アクセラレータで発生する熱を定期的に取り去り、プロセッサ、アクセラレータが動作周波数を低下させることなく高性能を継続的に供給できるように実現している。



図 22 GPU 用コールドプレート



図 23 KNL 用コールドプレート

5. 利用のために

本節では、これまでに説明した OCTOPUS を実際に利用者の皆様にご利用いただくための参考情報を記す。ここまで読んでいただけた読者の皆様方には、ぜひ OCTOPUS のご利用をご検討いただければ幸いです。なお、ここに記載する内容は、本センターの大規模計算機システム (スーパーコンピューティング) 事業の Web ページ [11] にも記載しているので、詳細については当該 Web ページを参照いただければ幸いです。

5.1 利用負担金の考え方

多くの利用者の方の関心が高いのは、計算機利用に伴う負担金であろう。事実、現利用者の方からも負担金についての多くの質問・相談・要望が寄せられている。とりわけ、多く寄せられる要望としては、利用負担金がもう少し安くないか?、あるいは、無料にならないのか? というものである。

スーパーコンピュータ上で大規模なシミュレーションやデータ解析を行う研究者にとっては、どうしても利用負担金が大きくなってしまふ。このことは

容易に想像できる。スーパーコンピューティング事業を担当する者としても、研究者としても、上記の利用負担金に対する要望は痛いほどよくわかる。しかし、言い訳ではないが、サイバーメディアセンターの利用負担金は、計算機利用に相当する消費電力量に基づいており、できるかぎり低額な設定としている。

スーパーコンピュータの保守・運用には、スーパーコンピュータシステムをただ設置し「はい、どうぞ」と置いておくだけではだめで、ハードウェア保守、ソフトウェア整備、テクニカルサポート等の費用も発生する。サイバーメディアセンターのスーパーコンピュータの調達においては、スーパーコンピュータを構成するハードウェア機器だけでなく、ハードウェア保守、整備・サポートに関する人件費等の運用・保守にかかる費用の多くを含めて調達を行っている。そのため、サイバーメディアセンターのスーパーコンピュータの利用負担金は、利用者の計算機利用に伴う消費電力量相当のみを負担いただく構造となっている。言い換えれば、スーパーコンピュータを運転するのに必要最低限となる費用だけを利用者に負担いただいている。この利用負担金をゼロにするためには、あらかじめ調達に消費電力を見込んで含めておく、あるいは、どこからかそのための予算をとってくるしかないのが現状である。

ここで、前者でいいのではないかとと思われる読者の方もおられるかもしれない。その場合、消費電力分相当の費用をスーパーコンピュータシステム調達より取り置いておかなければならなくなり、調達できるスーパーコンピュータの規模は小さくならざるを得ない。場合によっては、スーパーコンピュータではなく、ただのコンピュータになってしまい、利用者の計算要望やニーズを充足することはさらに難しくなってしまう可能性が高い。今日科学研究に伴うデータ量はますます大きくなり、計算要求・ニーズもますます大きくなりつつある。結果として、より大規模なシミュレーション、データ分析が可能なスーパーコンピュータの必要性・重要性が高まりつつある。このことを鑑みれば、計算機を利用した利用者自身が消費電力相当分を利用負担金とし

て負担いただく受益者負担が最善ではないだろうか。

電気、水道等の生活ライフラインのように、科学研究のインフラであるスーパーコンピュータの運転にもどうしても費用が発生する。この費用については誰かが払わなければならない構造である。この構造をご理解いただき、OCTOPUS をご利用いただければ幸いである。

5.2 OCTOPUS 利用負担金

さて、前置きが長くなってしまったが、OCTOPUS の利用負担金について説明する。表 2 に設計・策定した負担金制度の概要を示す。OCTOPUS の利用負担金では、OCTOPUS ポイントを導入している。OCTOPUS は、上述してきたように、プロセッサ、アクセラレータの構成が異なる複数のノード群より構成されるハイブリッド型スーパーコンピュータである。そのため、それぞれのノード群ごとに消費電力が異なるため、原則に従えば異なる利用負担金単価となるのが適切であるが、利用者の皆様方にノード毎に契約いただく方式をとることはかなり不便である。そこで、OCTOPUS では、OCTOPUS ポイントを採用することで、利用者は事前にノード種別を確定しておく必要がなく、臨機応変に複数のノードを使い分けることができるよう利用負担金を設定した。

表 2 OCTOPUS 利用負担金

(A) 占有	
基本負担額	占有ノード数
191,000 円/年	汎用 CPU ノード群 1 ノード
793,000 円/年	GPU ノード群 1 ノード
154,000 円/年	Xeon Phi ノード群 1 ノード

(B) 共有		
コース	基本負担額	OCTOPUS ポイント
	10 万円	1,000 ポイント
	50 万円	5,250 ポイント
	100 万円	11,000 ポイント
	300 万円	34,500 ポイント
	500 万円	60,000 ポイント

(C) ディスク容量追加		
基本負担額	提供単位	
3,000 円/年	1TB	

設計した利用負担金制度は、ノードを一利用者がある一定期間を予約して占有利用する占有、および、複数の利用者が時間分割で利用する共有の種別がある。占有利用は、サイバーメディアセンターの計算拠点としての一角を担う革新的ハイパフォーマンス・コンピューティング・インフラストラクチャ (HPCI) [12] や、サイバーメディアセンターが参画する学際大規模情報基盤共同利用・共同研究拠点 (JHPCN) [13] で採択された課題の利用を想定しており、多くの場合は共有利用となる。

表 3 OCTOPUS ノード種別消費係数

ノード群	消費係数	季節係数
汎用 CPU ノード群	0.0520	平成 30 年度は 通年 1 で稼動 (季節変動無し)
GPU ノード群	0.2173	
メニーコアノード群	0.0418	
大容量主記憶搭載ノ ード群	0.3703	

表 3 にノード種別の消費係数を示す。表は 1 ノード時間あたりの消費係数を示しており、例えば、汎用 CPU ノードの場合、1 ノードを 1 時間利用した場合、0.0520 OCTOPUS ポイントが消費される。言い換えれば、1000 ポイント (10 万円税別)で、

CPU ノードの 1 ノードを 19230 時間利用できることがわかる。表中季節係数とあるのは、平成 31 年度以降の運用時に前年度の利用状況に基づいて、季節ごとに消費係数を変更することで利用者の計算計画を制御しようとする係数である。一般的に、年度末には、ジョブが多く投入され混雑する傾向にある一方、年度初めには比較的ジョブが走りやすい状況がある。これを緩和するために、例えば、年度初めころには消費係数を低くし、利用者の利用を年度初めに移動させることをねらう。初年度となる平成 30 年度は、季節係数を年間通じて 1 として運用を行う。平成 31 年度以降については、OCTOPUS の平成 30 年度の利用率を鑑みつつ、サイバーメディアセンターに設置されている高性能計算機システム委員会での議論をへて、季節係数を決定していく計画としている。

5.3 公募型利用制度

それでもやはり利用負担金では十分な計算資源量が確保できないという方がおられるかもしれない。その場合は、サイバーメディアセンターの公募型利用制度を通じた計算機利用も可能である。

本センターでは、本センターの大規模計算機システムを活用する研究開発の育成・高度化支援の観点から、本センターが参画する「ネットワーク型」学際大規模情報基盤共同利用・共同研究拠点 (JHPCN) [12] や革新的ハイパフォーマンス・コンピューティング・インフラ (HPCI) [13] の目的を踏まえつつ、今後の発展が見込まれる萌芽的な研究課題や本センターのスーパーコンピュータを最大限活用することで成果が見込まれる研究課題を公募する制度を 2016 年度より開始している。本制度には、それぞれ JHPCN および HPCI での課題として採択されることを目的とし、(1) 若手・女性研究者支援萌芽枠、(2) 大規模 HPC 支援枠を設定している。本センターが設置する高性能計算機委員会において厳密な審査の上採択された研究課題については、本センターが研究課題推進のための利用負担金の全額を負担する。また、本セン

ターの施設(会議室)もまた利用可能となる。ただし、その一方で、本センターが開催する公募型利用制度成果報告会での発表、成果報告、発表論文への謝辞記載等の義務も存在するので、その点にご留意いただきたい。

本制度は、例年、11月半ば頃から12月半ば頃までの期間、翌年度の公募型利用制度への課題提案・応募を受け付けている。この応募期間は、科学研究費の応募を終えた後に設定しており、研究者にとって応募しやすい(科研を提出した後に引き続き申請書を記載しないとイケないが。)よう考慮して設定してある。本公募型利用制度の制度設計を行い、運営を行う筆者の立場からは、上記の公募利用制度整備趣旨をご理解頂ける方は、本センタースーパーコンピューティング事業 Web ページ内の公募型利用制度募集要項に従い、是非 OCTOPUS を利用する課題の提案・応募を積極的に検討いただければと思う。

5.4 HPCI & JHPCN

また、上述してきたように、サイバーメディアセンターは、学際大規模情報基盤共同利用・共同研究拠点 (JHPCN) [12] や革新的ハイパフォーマンス・コンピューティング・インフラ (HPCI) [13] に参画している。HPCI では、国立研究開発法人理化学研究所のスーパーコンピュータを中核とし、国内の大学や研究機関の計算機システムやストレージを高速ネットワークで結ぶことにより、全国の HPC リソースを全国の幅広い HPC ユーザー層が効率よく利用できる科学技術計算環境を実現することを目的とする。また、JHPCN においては、北海道大学、東北大学、東京大学、東京工業大学、名古屋大学、京都大学、大阪大学、九州大学にそれぞれ附置するスーパーコンピュータを持つ 8 つの施設を構成拠点とし、超大規模計算機と超大容量のストレージおよび超大容量ネットワークなどの情報基盤を用いて、学際的な共同利用・共同研究を実施することにより、我が国の学術・研究基盤のさらなる高度化と恒常的な発展を目的としている。これらの枠組みを利用し

て、サイバーメディアセンターのスーパーコンピュータを利用することが可能である。

HPCI および JHPCN においても、前節 5.3 で記した公募型利用制度と同様に、課題申請書を記載の上、課題提案・応募を行う必要がある。応募された課題は、それぞれで設置された課題審査委員会で採否が決定される。採択された課題は、利用負担金を研究者が負担することなく利用することが可能となる。例年、JHPCN では12月半ば頃より1月初旬、HPCI では3月頃から5月初旬にかけて課題募集が行われている。本センターのスーパーコンピューティング事業からも Web ページやメーリングリストを通じて、これらの課題募集についてはアナウンスを行っている。HPCI および JHPCN に対して、OCTOPUS を含む本センターのスーパーコンピュータを利用する研究課題を是非提案・応募して頂ければ幸いである。

6. おわりに

本稿では、2017年12月よりサービスを開始したハイブリッド型スーパーコンピュータ OCTOPUS について、その導入の背景、調達概要を記したのち、システム構成・特徴について解説した。OCTOPUS は、総理論演算性能 1.46 ペタフロップス (Peta Flops: 浮動小数点演算を1秒間に 1460 兆回実行可能) を有する本センター初のペタフロップス級性能を有するスーパーコンピュータとなった。冒頭にも記したが、本システムの導入に際しては、センターでスーパーコンピュータシステムの運用に携わる教員や技術職員、仕様策定委員らによる 2 年以上の議論・検討、提案候補ベンダとの打ち合わせ・情報交換がなされており、関係者らにとってはようやく実現にこぎつけた待望のシステムである。OCTOPUS に少しでも興味を持っていただけた方がおられたら、是非ご利用いただければ幸いである。皆様方のご支援がサイバーメディアセンター スーパーコンピューティング事業の再生・躍進につながります。是非とも、率直なご意見、ご要望等のフィードバックを引き続きお寄せ頂きますようお願い致します。

参考文献

- (1) ExpEther Consortium,
<http://www.expether.org/jindex.html>.
- (2) 伊達進, 木越信一郎, “全国共同利用大規模並列計算システム調達の背景”, 大学 ICT 推進協議会 2017 年度年次大会, 2017.
- (3) OCTOPUS ラックデザインコンテスト,
<http://www.hpc.cmc.osaka-u.ac.jp/octopus-rack-design/>.
- (4) インテル アドバンスド・ベクトル・エクステンション 512 (Intel AVX-512),
<https://www.intel.co.jp/content/www/jp/ja/architecture-and-technology/avx-512-overview.html>.
- (5) NVIDIA Tesla P100, <http://www.nvidia.co.jp/object/tesla-p100-jp.html>.
- (6) インテル Xeon Phi プロセッサ7210,
<https://www.intel.co.jp/content/www/jp/ja/products/processors/xeon-phi/xeon-phi-processors/7210.html>.
- (7) DDN EXAScaler,
https://ddn.co.jp/products/storage_software/exascal.html.
- (8) Mellanox CS7500 Series,
http://jp.mellanox.com/page/products_dyn?product_family=191&mtag=cs7500.
- (9) ASETTEK, RackCDU D2C (Direct-to-Chip),
<https://www.asetek.com/data-center/oem-data-center-coolers/rackcdu-d2c/>.
- (10) Cybermedia Center, Osaka University Adopts Asetek Liquid Cooling for New Computer System “OCTOPUS”,
<https://www.asetek.com/pressroom/news/2017/cybermedia-center-osaka-university-adopts-asetek-liquid-cooling-for-new-computer-system-octopus/>.
- (11) 大阪大学サイバーメディアセンター大規模計算機システム事業, <http://www.hpc.cmc.osaka-u.ac.jp/>.
- (12) 革新的ハイパフォーマンス・コンピューティング・インフラストラクチャ (High Performance Computing Infrastructure), <http://www.hpci-office.jp/>.
- (13) 学際大規模情報基盤共同利用・共同研究拠点 (Joint Usage/Research Center for Interdisciplinary Large-scale Information Infrastructure), <https://jhpcn-kyoten.itc.u-tokyo.ac.jp/ja/>.