



## 「富岳」 Begins! —「富岳」から始まる未来世界—

### 第1部 シミュレーション・ファーストによる「富岳」でのSociety5.0の推進

国立研究開発法人理化学研究所 計算科学研究センター長

まつおか さとし  
松岡 聡



私からは講演の前半で前提知識として「富岳」がどういうものであるか紹介し、また、その未来について解説する。

「富岳」は国家プロジェクトとして設計、開発、構築、運用が行われ、1000億円以上の国費が掛かっている。スーパーコンピューターを民間だけではなく国家プロジェクトでつくる理由について述べる。

我が国と熾烈な開発競争を繰り広げている米・欧・中の投与している国家資金は5000億円や1兆円である。民間だけでは賅えない最先端のシステムを設計することにより、例えばアポロ計画や核融合など、ハイリスクな開発をしてICT技術を進める。現に「富岳」においても理化学研究所と富士通がメインで開発をしたが、富士通が自らのイニシアティブで「富岳」をつくるのは不可能だったと思われる。「つくてなんぼ」と、まずICTの技術をつくてどうなるかということだった。

情報通信技術はそれだけでは駄目で、「つくてなんぼ」、つまり世の中に対してインパクトを与える必要がある。スーパーコンピューターには二重の投資効果がある。

一つは国家所有の情報インフラ資産としての位置付け。

オフザシェルフのチップを使わずに自らが設計開発することにより、3倍近い投資効果を得ることができた。

さらに、コロナのほか様々な研究開発や、今後のIT産業へのインパクトという点で、「富岳」の利活用の効果、ROI (Return on Investment: 投資利益率) は60倍ぐらいある (Hyperion Research試算)。

Society 5.0は、SDGsに対する目標をDXで達成していくことが課題だが、その中心的な課題として高いROIで活躍するために「富岳」は最初の設計開発の段階から、今までの我が国のスーパーコンピューターと違って、単に性能だけではなく非常に広い利活用、ユーザーベース、他のITのインフラストラクチャーとの連携や波及効果を狙って研究開発が行われた。

「富岳」という名前は、頂と裾野が両立している富士山の形状が、われわれの理想像を具現化するものであったためつけられたもので、この広さ、高さを達成するのが大きな研究開発目標となった (図1)。しかし、汎用性と速さを非常に高いレベルで両立させるのは半導体業界にとって大変に難しいことで、まさにムーンショットと呼ばれたアポロ



■ 図1.



計画に匹敵するようなプロジェクトであった。NASAを中心としてアメリカの航空宇宙産業が集結したアポロ計画が、10年間で月に行ったのと同じように、「富岳」も理化学研究所、富士通を中心に数千名が様々な形で関わり、非常に高い目標を達成した。そのために国家はリスクマネーを投与したわけだ。

出来上がった最新プロセッサ「A64FX」は、アメリカの汎用のXeonなどのサーバ向けCPUに対し、少なくともハイパフォーマンス・コンピューティング（HPC）のワークロードに対し3倍の性能を発揮するチップである。

かつArm CPUなので、組み込みには年間10億か200億のチップが使われるが、64ビットのプロセッサと完全な互換性を持っており、独自のアーキテクチャでありながら様々なArmの規格にすべて準拠している。

「富岳」はこのCPUが約16万個集まり、GoogleやAmazonに匹敵するような巨大なデータセンターに集約され、1台のマシンを成している。16万ノードの「富岳」の計算能力を試算すると、最新のスマホで2000万台、iDCのサーバで30万台程度に相当する。これは日本でのスマホ、サーバの年間売り上げ数と偶然にも一致する。「富岳」が2~3台で日本のICTの計算需要がすべてカバーできる計算能力を持っているのだ。

「富岳」は日本の最先端の技術を、世界をパートナーにして結集することによりつくられた世界最高峰のCPUであり、ネットワークの部分においても世界最高の帯域を実現している。世界に先駆けて7nmルールを採用し、かつHBM（High Bandwidth Memory：広帯域メモリ）という毎秒1TBの帯域を持つメモリを初めて汎用CPUとして実現したのが、このチップである。ほかのアメリカ・中国・ヨーロッパの汎用CPUにもいまだに実現されていないHBM接続をいち早く実現して、高性能、省電力、Armに対する互換性を達成した。

計算能力が注目されがちなスーパーコンピューターは単にCPUを並べただけでは駄目で、ネットワークと両立していないと成立しないので、最先端のネットワーク技術が必要になる。データセンターでは400G~Tbps級のファイバが何本か出ているのが通常だが、「富岳」の場合はチップの中にネットワークが内蔵されている。かつ、サーバCPUとして初めて、ネットワークだけでなくスイッチを内蔵している。「富岳」は16万個のネットワークスイッチを内蔵していることにより各チップから外部スイッチを必要とせず400Gbpsの帯域が実現され、400Gbpsのネットワークで16万個がす

べて接続されている。

3分の2が銅線で、3分の1がAOC（Active Optical Cable）で接続されているが、16万個あるので、AOCのファイバの数は10万本、総延長はだいたい2,000kmになる。チップからネットワークに流せるデータ量は、6PB/sのインジェクションバンド幅を持つ。GAFAMの数あるデータセンターの内部トラフィックを全部合算した10倍の容量がある、これが「富岳」のネットワークの実態である。

さらに、システムとしてはDisaggregated（非集約型）なアーキテクチャを採用している。CPUとメモリは接合しているが、ロジカルにはネットワークを通じて、ほかのCPUの介在なくすべてのメモリとアクセスすることが可能である。残念ながらロードストアではなく、技術的にNTTのIOWNのように遠距離でアクセスすることはできないが、少なくともマシン内ではすべてのCPUから5PBのメモリに自由にアクセスできる。ほかのCPUにこのような機能はないので、Disaggregatedアーキテクチャという点でも最先端の技術を実現している。

「富岳」の速度の秘密は、全部で48コア、実際は52コアあることだ。Many-CoreのArm CPUであると同時に、内部の作りが普通のCPUコアとして機能するだけでなく、それぞれのコアがGPUのようなストリーミングアーキテクチャとして動作する。別途GPUがあるのではなく、あくまでもそれぞれのコア及びメモリシステムが、GPUのようなストリーミング動作が可能になっている。昔のスパコンで言えば、ベクトルアーキテクチャを実現しているということになる。汎用のCPUのプログラミングモデルを維持しながら、ベクトルスパコンないしはGPUの性能を得るのが「富岳」のCPUの作りなのだ。

今のデータセンターに欠かせないパワーマネジメントについても、Googleなどの世界のiDCが10年間で5倍の電力効率を達成したと言われていたのと比べると、「富岳」においては10年間で20数倍の電力効率を達成している。この20数倍は、パワーマネジメントで様々な機能が入っていることによるところが大きい。

さらに、普通のクラウドとデータセンターと違い、「富岳」は場合によっては16万ノードで1個のプログラムを動かすため、「富岳」には信頼性を確保するために様々な機構がある。その一つはメインフレームグレードのRAS（Reliability, Availability, Serviceability）機能が入っていることである。メインフレームメーカーとして長年活躍してきた富士通が、その技術を総動員して、「富岳」のチップにも実現している。



インテルなどもその辺をだいぶキャッチアップしてきたが、IBM、富士通などのメインフレームメーカーとはまだ差がある。インテルと比べると数倍の信頼性を確保して、これにより16万ノードがきちんと動くという技術を達成している。

ネットワークはTofu-Interconnectで、各チップ当たり約400Gbpsのインジェクションを実現する。チップの中で、スイッチを含め全部で6次元のトラスを実現している。10ポートのスイッチが入っていて、RDMAコントローラーとスイッチで25mm<sup>2</sup>の面積を占めている。

電力はチップ当たり8~9W、このチップがフルに稼働するとメモリを含め200W近くあるので、ネットワークにかかる消費電力は5%程度である。マシンは最大20~30MWなので、5%といっても数百kWになるが、6PB/sのインジェクションバンド幅を持つネットワークとしては意外とローパワーと言える。このネットワークはI/Oバスを経由せず、各CPUに直接つながっている。CPUのオンチップネットワークに直接結合していて、メモリトランスファーやキャッシュに対するインジェクションなどがすべてできるようになっているものが16万個集まっており、ネットワークのポート数としてはだいたい160万ポートになる。

これをDisaggregateにした一つの理由は、この手のスーパーコンピューターではレイテンシを非常に小さくしなければいけないため、チップ間のレイテンシは500ナノ秒程度になっている。例えば100Gbで10~100マイクロ秒なので、それと比べると200倍の速さである。カタログ値だけではなく実測のInjection rateも38.1GB/s、93%の効率を得ることができている。こういうことで、非常に大規模な並列計算が可能になっている。これが384ノードごとにラックに収められている。普通のiDCのラックは40ノードなので、

その約8倍の密度で、当然ながら水冷にしてある。100kW程度のラックが全部で400ラックほどあるのが「富岳」だ。その総合性能は、FLOPSで考えると1ベタ近辺だが、単に演算速度ではなく、メモリ速度が163PB、ネットワークが6PB/sなど、すべての指標において非常に大きな数値を提供していることが重要である。しかしながら、Arm CPUなので、FORTRANや数値ライブラリなどの従来型のスパコンのソフトウェアスタックだけではなく、ディープラーニングも含め現代的なクラウドのソフトウェアスタックも直接サポートしている。それらが融合して、現代のクラウドやビッグデータ、AIと従来型のスーパーコンピューターの用法をミックスして使えるのが「富岳」の特徴となっている。

毎年発表されるスーパーコンピューターのランキングにおいて、この指標が4連続4冠、しかも、シミュレーションだけではなく、AI、ビッグデータなどの指標において、すべて2位に対してかなりの差をつけて1位を獲得した。加えて今回、新しい深層学習のベンチマークのML Perf HPCができ、5冠になった。従来型のシミュレーションや密行列系、疎行列系のシミュレーション、さらにCNNのブロック処理、ビッグデータ、グラフ処理のものだけではなく、もう少しアプリケーション寄りになっているML Perf HPCが策定されたのである。富士通と理化学研究所でチャレンジして、CosmoFlowというシミュレーションデータからの宇宙論パラメータの予測で1位を取ることができた(図2)。

「富岳」は「つくってなんぼ」だけではなく、「つかってなんぼ」のところも重要視され、アプリケーションファーストで設計されたマシンである。1位を取るとは、あくまでも副次的な話で、本当の設計目標となったのは、図3にあるサステナビリティに関する多くのアプリケーションで、「京」

## 2021年11月17日、SC2021会場(セントルイス)での表彰式

- コロナの影響で、実際の表彰式での富岳の表彰は初めて



■ 図2.



■図3.

に対し数10倍から100倍以上の性能向上を果たすことだった。スポンサーやCSTI（総合科学技術・イノベーション会議）にしてみれば、こちらが評価指標になっている。健康長寿社会、医療・創薬、地震・津波などの防災ないしは気象・カーボンニュートラル性に対する気候モデル、エネルギーの生成や変換効率、蓄積に関するアプリケーション。さらに、産業競争力という点では新しいデバイスや、非常に大きい構造物に対する様々なアプリケーション。もちろん基礎科学に関する目標もあり、平均的には「京」に対し70倍、一部のアプリケーションでは100倍を超える性能向上を10年間で果たした。

さらに、急速な対応が必要とされるコロナウイルスのまん延に対し、プログラムを文科省と一緒に立ち上げて、様々なコロナウイルスのアプリケーション、プログラムの応募に応じた。分子ないしは量子科学レベルで、コロナウイルスの変異があったときに、どのぐらい結合力が高まるか。特にクーロン力とファンデルワールス力の合力がどうなっていくかを検証するようなプロジェクトから、それがどのように創薬に影響していくか。ミクロスコピックなものから、だんだんマクロスコピックなもの、遺伝子レベル、飛沫、さらには社会現象、例えばパンデミックにおけるロックダウンの経済的ダメージのシミュレーションなど、様々なことを行ってきた。総計算容量は、我が国の「富岳」以外のスーパーコンピュータを複数台合わせて丸ごと1年間占有したとしても足りないほどの計算量を、「富岳」で投じてきた。

「富岳」は、ゴードン・ベル賞に2つノミネーションされた。個人よりも成果そのものに対するこの賞を受賞することはアプリケーションチームやマシンにとって最高の栄誉で、非常に権威のある賞である。トップスパコンが受賞してきた

この賞を、坪倉氏をリーダーとした研究チームによる論文で、6つのファイナリストの中から受賞することができた。

今後、「富岳」でSociety 5.0を実現していく。今の感染症に関する様々な仕事も、社会に対しICTがデジタルツインをつくり、サイバー空間の結果を実際の世界にもたらす。逆に実際の空間からセンシングして、それをサイバー空間に持っていく。このような場面で「富岳」が中心的役割を果たす。

「富岳」が得意とする物理空間からサイバー空間でデジタルツインをつくることばかりではなく、さらにICTの部分もデジタルツインをつくることで、スマートシティなどの実際の検証をしたい。スマートシティを実現しようとするときに、それを構成する建物、交通、災害などのシミュレーションをするだけではなく、ICTの部分も同時にシミュレーションして、それを結合する。例えば10万個Arduinoセンサがあり、それがシミュレーションから受け取る。Arduinoセンサはバーチャルマシンとして「富岳」上に実現されていて、バーチャル5Gが動いていて、バーチャルクラウドに SHIPPING され、そこでAIが動いてactuationしたものが、逆にシミュレーションにデータ動向で反映される。このようなインフラストラクチャーをつくり、まさに世界をリードしていく。

最後に、今後デジタルツインをつくっていくことに関して、第一原理のシミュレーション、AI、ビッグデータを技術的に融合させて仮想世界をつくり、それをactuationするプラットフォームになっていくことが、まさに「富岳」に課せられた使命である。「富岳」は結構簡単に使えるので、興味があれば、またそういうことをやりたいというチームがいれば、コンタクトいただきたい。