

DCS-7280PR3-24-Fを採用してスーパーコンピュータ「富岳」を ダークファイバ経由でSINET6と400Gbpsで接続

Highlights

Challenge

「富岳」へのアクセス高速化

Arista Solutions

DCS-7280PR3-24-F

400G回線、400G-ZR+LineSystemで
SINET6と計1.2Tbpsで接続

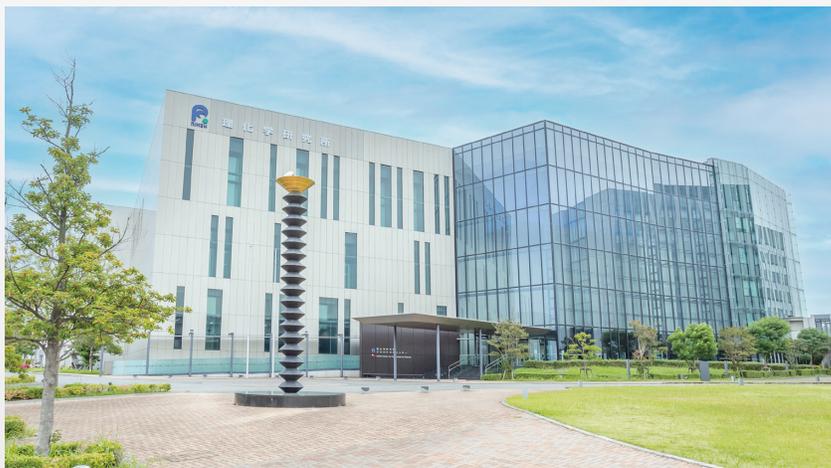
基幹ネットワークを400Gで構成

CloudVision®を使用しているテレメトリー取得

Results

Arista 7280R3シリーズとトランシーバの組み合わせは、従来の専用線接続サービスで利用されていた終端装置の役割も兼用することが可能。機器構成がシンプルになり、コスト削減

国立研究開発法人理化学研究所 計算科学研究センター (R-CCS) は、「社会の革新的発展を目指し、科学や社会が抱える課題の解決に貢献するため、『シミュレーション』『ビッグデータ解析』『AI』を融合した最先端の研究を、ハイパフォーマンス・コンピューティング (HPC) で実践する」ことを使命とする研究機関で、いずれも世界最高性能にランキングされた「京」と「富岳」の2世代のスーパーコンピュータの設置／管理・運用拠点となっている。2012年から「京」の運用が開始され、2019年に共用停止。「富岳」は2020年4月から新型コロナウイルス感染症対策を目的とした優先的な試行的利用を開始し、2020年6月から4期連続でグローバルの4つのスーパーコンピュータランキングで世界第1位となる“四冠”を達成、現在も世界最高性能のスーパーコンピュータとして幅広い分野で活用されている。



理化学研究所
計算科学研究センター
RIKEN Center for Computational Science



Customer Data

国立研究開発法人理化学研究所 計算科学研究センター

所在地

〒650-0047

兵庫県神戸市中央区港島南町7-1-26

URL

<https://www.r-ccs.riken.jp/>

スーパーコンピュータとそれを支えるネットワーク

スーパーコンピュータ「富岳」は、前世代システム「京」の100倍以上のアプリケーション実効性能を達成し、2022年11月時点で、実アプリケーションの実効性能を示す「HPCG」、ビッグデータ解析性能を示す「Graph500」で6期連続の世界第1位、演算速度性能を示す「TOP500」で第2位、AI性能を示す「HPL-AI」で第3位を獲得した世界最高峰のスーパーコンピュータです。「富岳」は、ランキングトップになることが存在意義ということではなく、日本政府が第5期科学技術基本計画で提唱した「Society 5.0」の理念である「サイバー・フィジカル」の実社会への実装検証を担う不可欠のHPCインフラと位置づけられています。

「富岳」は学術情報ネットワーク「SINET6」に接続されており、日本全国の大学や研究機関のみならず、世界中からインターネットを経由して利用可能になっています。「富岳」が本格的な共用を開始した2021年3月の時点では100Gbpsの専用線 1回線でSINET5に接続されていましたが、2022年4月から運用を開始したSINET6では神戸DCおよび大阪DCとの間で400Gbpsの回線を3回線、総帯域幅1.2Tbpsという大容量で接続しています。この接続のために導入されたのが、アリスタネットワークスの“Arista 7280R3シリーズ”スイッチです。

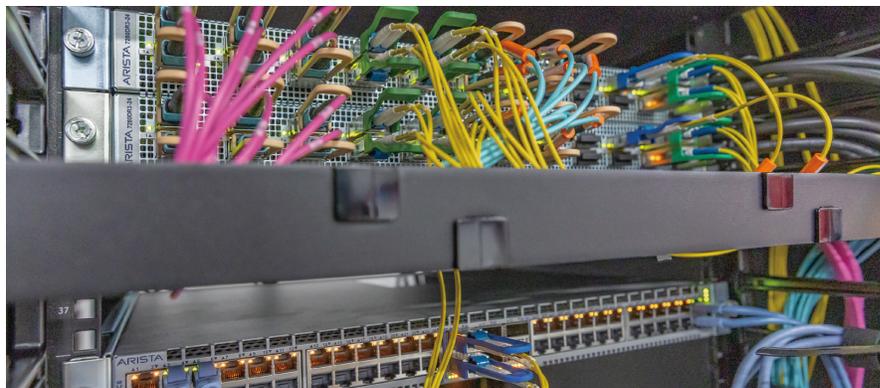


そもそもスーパーコンピュータの接続回線はどのような役割を担っているのでしょうか。スーパーコンピュータの演算性能を活かすために、データアクセスがボトルネックにならないように計算に必要なデータはあらかじめ「富岳」側の高速ストレージにコピーしておき、超高速な内部インターコネクトを介して読み書きしています。SINETとの外部接続回線は、ユーザーのデータを「富岳」のストレージにコピーするために利用される形になります。

一方で、スーパーコンピュータが扱うデータ量は年々増大の一途を辿っているため、「富岳」での計算の準備のために送り込まれるデータ量もどんどん大きくなっていると言います。もちろんアプリケーションやワークロードによって必要とするデータサイズはマチマチなので一概にはいえませんが、少ない場合では数GB程度で済む場合もあれば、大規模な場合はPB級のデータを扱う場合もあるそうです。「富岳」の運用や外部接続ネットワークを担当するR-CCSの運用技術部門 先端運用技術ユニットのユニットリーダーを務める山本 啓二氏は、「5~6年前にSINET5の100Gbps回線に接続したときには

100Gbpsもの帯域を使うとは思っていませんでしたが、数年で100Gbpsをほぼ使い切るようになりました。SINET6に切り替えて400Gbpsに増強した現時点ではまだ400Gbpsも使っていませんが、SINET6の運用期間中に400Gbpsを使い切る状況になるだろうと予想しています」と言います。

大量のデータを超高速で処理することがスーパーコンピュータの役割であり、さらに現在社会活動のあらゆる状況がデジタルデータ化されるようになってきていることから、データ量の爆発的増大は今後もさらに継続していきます。よって、最新の規格にいち早く対応する高速ネットワークスイッチの役割も大きなものとなると言えます。



導入の経緯

400Gbps回線導入に備えたネットワーク機器の選定経緯について、R-CCSの先端運用技術ユニット技師の三浦 信一氏は、「400Gbpsのトラフィックをワイヤレートで処理できるとなると大手ベンダー製品に限られてしまいますが、中でもアリスタネットワークスのコストパフォーマンスが高かったことに加え、Aristaでは光ファイバの多重化伝送に対応するトランシーバが使えるので、長距離を高速に接続したいというニーズにマッチしていたという点も重要でした」と振り返ります。

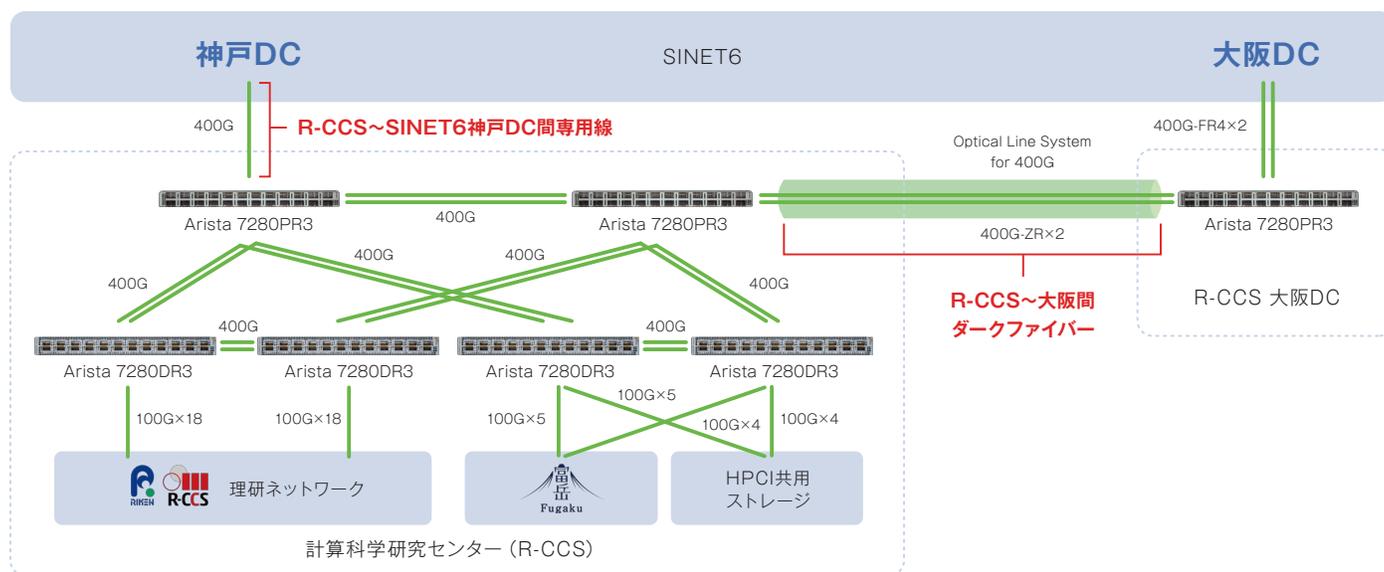
「富岳」のSINET6との接続は、神戸DCと大阪DCの2箇所に対して行なわれています。神戸のR-CCSとSINET6の神戸DCとの間は専用線接続による400Gbpsですが、SINET6の大阪DCとの接続はR-CCSの大阪DCを経由しています。R-CCS大阪DC～SINET6大阪DCの間は400GBASE-FR4を用いた近距離接続となっています。一方でR-CCSと大阪DCとの間は直線距離で30kmほど離れています。この区間にダークファイバを使い、400G-ZRモジュールとLine Systemを活用して、400Gbps回線を2チャンネル(800Gbps)で接続することで、神戸DCの専用線と合わせて、合計で1.2Tbpsの帯域幅を確保することができました。更に、帯域要求の声が上がったとしても、R-CCS～大阪DC間のそれぞれのArista 7280PR3に、400G-ZRモジュールを追加することで、3.2Tbpsまで帯域増強が可能になっています。神戸DCと大阪DCの2箇所に接続しているのは、万一の障害に備えた多重化の意味がありますが、同時に新技術を試す場としても機能していると言います。R-CCSの先端運用技術ユニット 上級テクニカルスタッフの野村 明広氏は、「神戸DC側の400Gbps専用線は、通信事業者の専用線接続サービスを利用したもので、いわば実績のある技術として導入しています。

一方、大阪DC側ではダークファイバを借りて自分たちで接続するという新しい手法にチャレンジしています」と説明しました。自分たちの責任でネットワークを構築し、日々運用していく負担は発生しますが、そうした作業を外部のサービス業者に委託するのと比べると、コスト面で大幅な削減が見込めます。スーパーコンピュータを開発して実際に構築し、運用を続けていくためには多額のコストを要



します。世界最高峰のシステムですのでそれも当然と言えますが、出来る限りコストを下げ、高コストパフォーマンスを維持していく努力は不可欠となります。実はArista 7280R3シリーズとトランシーバの組み合わせなら、従来の専用線接続サービスで利用されていた端末装置の役割も兼用することができるので機器構成がシンプルになり、コストも下がります。専用線接続サービスとダークファイバでは回線使用料も大幅に軽減されることから、「毎年新しいArista 7280PR3スイッチを買えるくらいの額が節約できます」（三浦氏）というほどです。

当然ながら、最先端の規格に基づく高速ネットワークを構築し、日々運用していく技術力があって初めて実現可能な手法ですが、スーパーコンピュータを熟知した専門家であるR-CCSならではの取り組みだと言えます。また、Arista 7280R3シリーズの製品品質や運用管理のしやすさもこうした体制を可能とすることに貢献していると評価されています。



Arista 7280R3シリーズの評価

前述の通り、R-CCSにとってもダークファイバを導入してネットワーク構築を行なうのは初めての挑戦だったことから「何かトラブルが起こるだろう」という想定で構築準備が行なわれたそうですが、実際には「接続して電源を入れたらそのまま普通にリンクアップしてしまい、何も起こらず普通に使えてしまいました」（三浦氏）というほど簡単だったとのことでした。運用管理面では、Arista EOS CloudVisionも高評価を得ています。

「富岳」の共用開始はコロナ禍の時期と重なっており、R-CCSでもリモートワーク/テレワークを中心とした体制になっていましたが、そうした環境でも各スイッチの状態をリモートから一元管理可能なので都合が良かったといえます。

また、内部での運用に切り替えたことから、従来以上に詳細な情報を把握する必要も生じましたが、CloudVisionではログ情報を過去に遡って詳細に参照することも可能なので、仮に何らかの障害が発生した場合にも、そのとき何が起こっていたのかを後から確認できます。実際に、BGPの不適切なメッセージが流れて来たためにルーティングテーブルがおかしくなってしまう、接続トラブルが発生する、といったことも経験したそうですが、その場合にもCloudVisionで状況を詳細に把握できる点が便利だったとの

ことでした。グローバルで見ても400Gbps接続で光多重化を活用している例はまだほとんど出てきておらず、R-CCSのシステムは最新技術をいち早く実運用した最初期の事例となっています。

多重化技術を活用したことにより、今後更なる帯域拡張に関しても、既設のダークファイバをそのまま活用しながら、スイッチのトランシーバーモジュールの変更だけで最新規格に対応したり帯域拡大を図ったりできる余地が確保されている点もポイントとなっています。製品のコストパフォーマンスの高さや運用管理性について高評価を得られていることから、R-CCSのネットワークではアリスタネットワークス製品の導入が拡大しており、結果としてCloudVisionによる一元管理がさらに効果を発揮できる環境になりつつあるともいえます。このほか、実はこのクラスのスイッチでは他社製品ではシャシー型の製品が主流となりますが、アリスタネットワークスではコンパクトなBox型のラインナップが充実しているため、省スペース性に優れている点も高評価でした。

アリスタネットワークスでは、今後も最新技術を高コストパフォーマンスで提供していくことはもちろん、運用管理性や信頼性／安定性の部分でも最高水準の製品を提供していくことで最高レベルのネットワークを運用していく必要のある先端ユーザーの皆さまをご支援していきます。

左から

国立研究開発法人理化学研究所
計算科学研究センター
運用技術部門 先端運用技術ユニット
ユニットリーダー 博士(情報理工学)

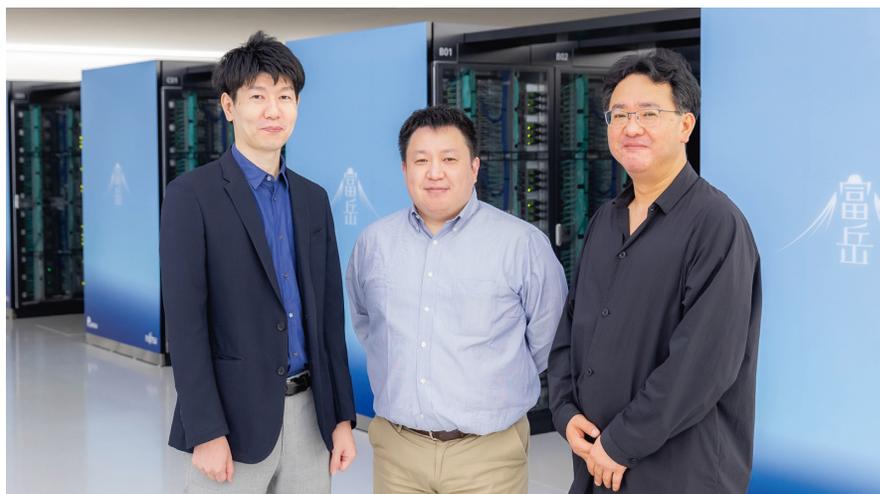
山本 啓二氏

国立研究開発法人理化学研究所
計算科学研究センター
運用技術部門 先端運用技術ユニット
技師 博士(工学)

三浦 信一氏

国立研究開発法人理化学研究所
計算科学研究センター
運用技術部門 先端運用技術ユニット
上級テクニカルスタッフ

野村 明広氏



お問い合わせ先

contact-japan@arista.com

URL

www.arista.com/jp

東京

〒100-0004
東京都千代田区大手町1-7-2
東京サンケイビル 27F

Tokyo

Level27, Tokyo Sankei Building,
1-7-2 Otemachi, Chiyoda-ku,
Tokyo 100-0004, Japan

大阪

〒530-0001
大阪市北区梅田2-2-2
ヒルトンプラザウエスト19F

Osaka

Level19 Hilton Plaza West Office Tower,
2-2-2 Umeda Kita-ku,
Osaka 530-0001, Japan

