

スーパーコンピュータ「富岳」の システム実装技術

江口 進 梅松 三三雄 大嶋 修

あらまし

スーパーコンピュータ「富岳」（以下、「富岳」）は、2020年6月に発表されたスーパーコンピュータのランキングにおいて、計算性能4部門で世界一を獲得した。「富岳」のプロトタイプ機は2019年11月発表のランキングにおいても、消費電力あたり性能で世界一を獲得した。このような高性能および省電力の実現は、新規に開発したCPU性能だけではなく、ラック当たりのCPU搭載数を前モデルであるスーパーコンピュータ「京」の4倍とするなどのシステム実装技術によるところも大きい。

本稿では、「富岳」の高性能・省電力を実現したシステム実装技術について述べる。

1. まえがき

スーパーコンピュータ「富岳」(以下,「富岳」)は,世界一の計算性能と世界最高クラスの消費電力性能を実現している。

身近なコンピュータであるパソコンの場合,多くはCPUを1つのみ搭載しているので,その計算性能はCPU性能で決まる。一方,スーパーコンピュータの場合,超高速な計算能力が求められるため,CPUを大量に搭載している。そのため,計算性能はCPU性能だけでなく,CPU間の連結能力も大きく影響する。したがって,CPUをより近くに実装する高密度実装が必要となる。

また,世界最高レベルの消費電力性能を実現するためには,CPUの省電力だけではなく,システム全体の電力ロスを徹底的に削減する必要がある。

本稿では,「富岳」の筐体の特長を説明し,高性能化および省電力化を支えているシステム実装技術について述べる。

2. 高性能化を支えるシステム実装技術

「富岳」のような多くのCPUをつなぐことで性能を高めるスーパーコンピュータでは,CPU間的高速データ通信を可能とする高密度実装が必要となる。本章では,まず「富岳」の構造を紹介する。更に,高密度実装を実現するための課題,および課題を解決する技術について述べる。

2.1 「富岳」の構造

図-1に,「富岳」の筐体の構造を示す。2個のCPU(それぞれのCPUには4つの積層メモリーを内蔵)を搭載したCPU Memory Unit(以下CMU)が基本単位であり,そのCMUをラックに前面96枚,後面96枚の合計192枚搭載している。CMUには,背面側のラック中央部で接続する高速通信コネクタ,水冷配管と,前面側で接続する高速通信ケーブルが配置されている。

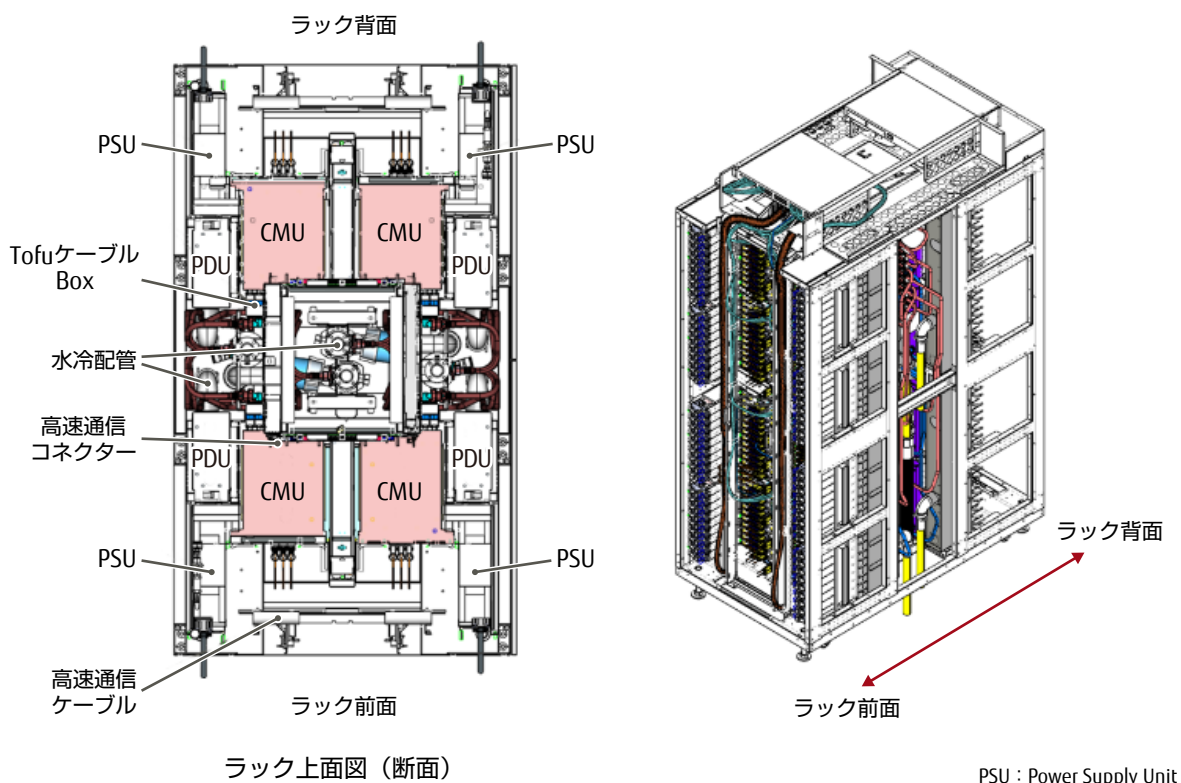


図-1 筐体構成

2.2 高密度化における課題

本節では、「富岳」において高い計算性能を実現する際の、システム実装における二つの課題について述べる。

(1) 少ない流量での効率的な水冷の実現

「富岳」では、スーパーコンピュータ「京」(以下、「京」)の4倍のCPUを搭載するとともに、CPU自体も高性能化し、1つのCPU当たりの発熱量も大きくなっている。その結果、ラック当たりの発熱は「京」の約6倍となり、冷却性能の強化が必要となる。一般に、水冷の流量を増やせば冷却効果は向上するが、配管が太くなるため、本来の目的である高密度実装の妨げになる。そこで、少ない流量での効率的な水冷の実現が課題となる。

(2) 少ない作業スペースでの活性保守

「富岳」では、CMUの保守を行う場合には、システムとしては動作を継続したまま保守を行う活性保守が求められる。しかし、CMUには高速信号、水冷配管、電源などの多くの接続があり、保守時にはそれらを挿抜する必要がある。活性保守を実現するためには、挿抜時の作業スペースを十分確保する必要があるが、これは高密度実装の妨げとなる。そこで、少ない作業スペースでの活性保守が課題となる。

2.3 高密度化を実現する技術

本節では、前節で述べた二つの課題を解決し、高速化を実現する技術を紹介する。

(1) 少ない流量での効率的な水冷の実現

少ない流量で高いCPU冷却性能を実現するために、新しい高効率クーリングユニットの開発、および複数あるCMUの流量差を抑える配管の設計を行った。

・高効率なクーリングユニット

CPUを冷却するクーリングユニットは、一般に一方の入口から入った冷却水をユニット内全体に分散させ、もう一方の出口から排出する方式を取っており、「京」ではこの方式を採用した。

一方、「富岳」の水冷ユニットLCU (Liquid Cooling Unit) では、図-2のように入口から入った冷却水をユニットの中央部に集中させ、両側に折り返して外側を冷やしてから排出する流路構造を採用した。これによって冷却水の流速は上がり、冷却効果を上げている。また、発熱密度の高い中央部から流入しているので、より効率的な冷却を行うことができている。

・流量差を抑える配管設計

「富岳」のラックには192枚のCMUがラックの前後に96枚ずつ搭載されるため、一番上に搭載されるCMUと一番下に搭載されるCMUの高低差は約2mにも及ぶ。そのため、重力による水圧差の影響で、CMUへの流量に差が発生する。この問題を解決するために、各CMUへ水流を分岐させるポイントを1次分岐管と2次分岐管に分け、1次-2次間を接続する配管の長さを調整することで、流量の差を±5%以内に抑えた(図-3)。

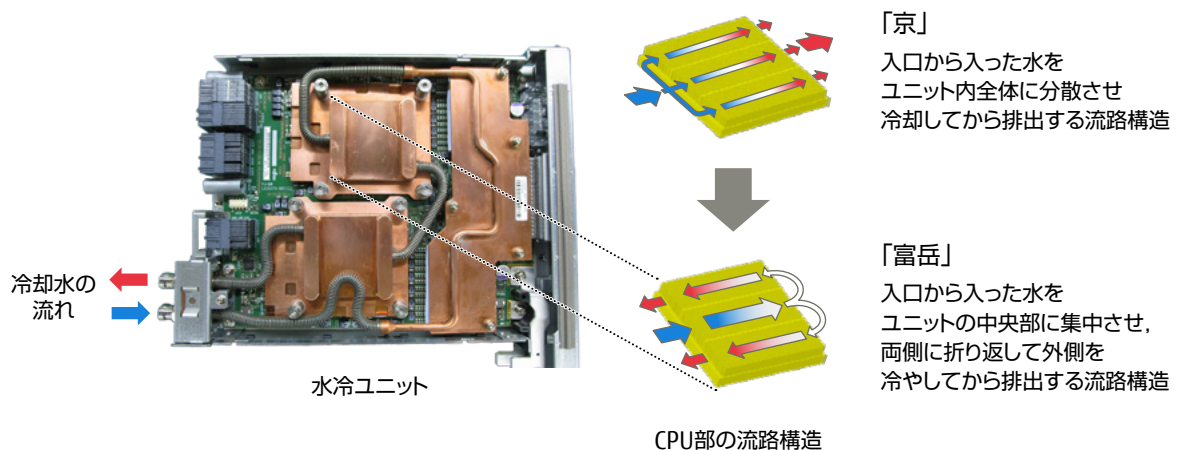


図-2 水冷ユニットの流路構造

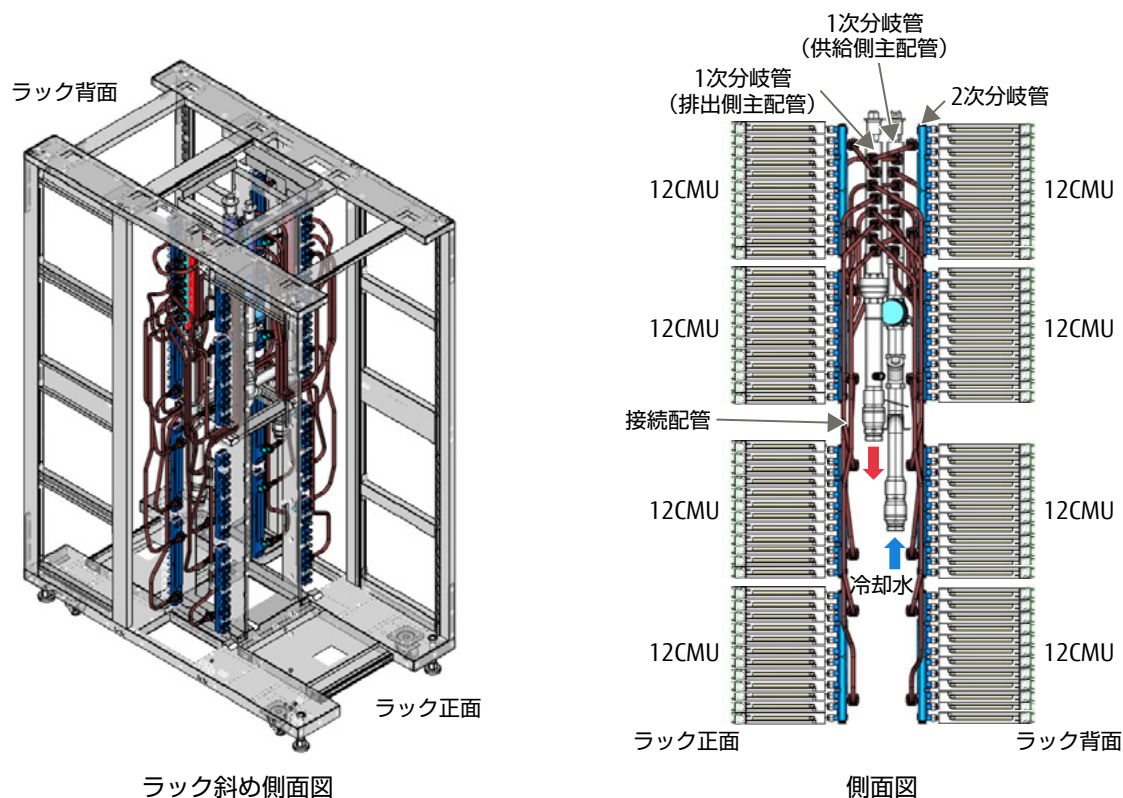


図-3 冷却水を均一に流す水冷システム

(2) 少スペースでの活性保守の実現

少ない作業スペースで高速信号、水冷配管、電源など多くの接続があるCMUの活性保守を実現するために、複数の接続を同時に行う電気-水一括嵌合システムと可動式ケーブルガイドを開発した。

・CMU電気-水一括嵌合システム

「富岳」では、ラックの前後両面にCMUをそれぞれ96枚搭載している。CMUなどを交換しやすくするために、配管、高速通信ケーブル、電源ケーブルを装置の内側に集約し、小型化と可用性を両立させている。「京」と同じ水供給方式を採用すると、配管の実装エリアおよび保守作業エリアが必要となり、高密度実装は困難である。それを解決するためには、電気コネクタおよび水冷コネクタを一括で嵌合することが必要となる。

一括で嵌合するには、水冷コネクタには位置合わせ機能がないため新たに機能追加することが課題となる。従来は手差しでコネクタの角度を合わせていたが、「富岳」では水冷コネクタに独自のフローティング構造を採用している。すなわち、水冷コネ

クターが3軸方向にフローティング（回転および平行移動）できるようにすることで、ユニットの挿入する動作で複数のコネクタ、水冷コネクタがそれぞれのガイド機能により正常位置に嵌合する（図-4）。

・可動式ケーブルガイド

「京」や既存の機種では、CMU交換作業エリアとケーブルを束ねて固定するケーブルフォーミングエリアは個別に必要であった。「富岳」ではそれらを共有するために、可動式ケーブルガイドを開発した。これによって、交換作業エリアとケーブルフォーミングエリアの共有が可能となり、スペース削減による高密度化を実現している（図-5）。ケーブルガイドは通常、CMUの前面に配置して、空冷のBIOU（Boot I/O Unit）の排気に影響を与えないようにしている。一方、CMUの交換作業時には中央に移動させることで、作業エリアが十分に確保でき、短時間でスムーズな交換作業が可能となっている。

以上のシステム実装技術によって、水冷を効率化するとともに、計算に関与しないスペースを可能な限り削減し、高密度にCPUを実装することが可能

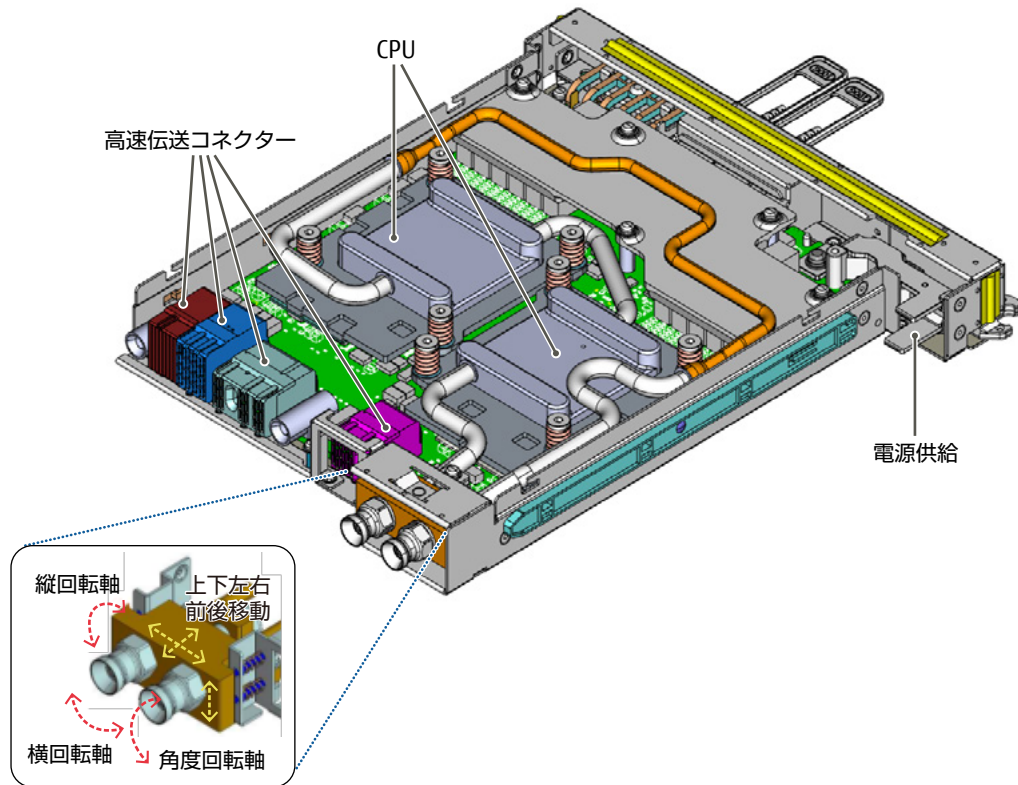


図-4 水冷コネクタのフローティング機構

となった。これによって、CPU間の接続距離を短くでき、世界最高速のCPUの計算能力を引き出すことが可能となった。

3. 省電力化を支えるシステム実装技術

世界最高の消費電力性能を実現するためには、CPUだけでなく、システム全体の徹底的な省電力が必要となる。システム内において、CPU以外で発生する電力ロス、電源ユニット（PSU）と電源回路（DC-DCコンバーター）によるものが大きいことが分かっている。省電力化には、PSUの電力ロスの削減とPSU-CMU間の電力ロスの削減が必要になる。

本章では、電力ロスを減らすために取り組んだ技術開発について紹介する。

3.1 PSUの電力ロスの削減

PSUには、メインとスタンバイの2系統の出力回路がある。従来の省電力対策においては、メイン出力回路を中心とした電力ロスの削減が行われて

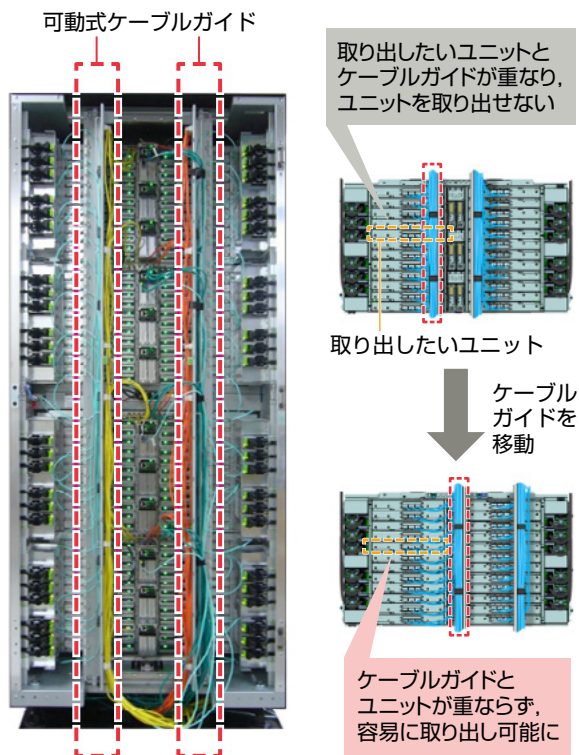


図-5 可動式ケーブルホルダー

いた。しかし「富岳」では、メイン出力回路だけでなく、従来あまり手が付けられていなかったスタンバイ出力回路の電力ロスの削減にも取り組んだ。

スタンバイ出力回路の電力ロスを低減するために、富士通独自のスタンバイ回路切り替え方式を考案した。この方式では、メイン出力回路が有効になっている場合は、スタンバイ出力回路の動作を止め、メイン出力回路からスタンバイ回路にバイパスして出力することとした。これによって、スタンバイ出力回路の電力ロスを全て削減し、電力ロスを大幅に削減できた。

メイン出力回路に対しても、セミブリッジレスPFC (Power Factor Correction)、フェーズシフトフルブリッジ回路など、最新の高効率回路を採用して、電力ロスを低減した。また、新電力デバイスである低損失半導体（スーパージャンクションFET：Field Effect Transistor）の採用によるリカバリー電流などの主要なスイッチング損失の低減、および部品レイアウトの最適化によるパターンロスの低減を図った。

ここまで述べた富士通独自の回路方式や最新高効

率回路の採用によって、PSUの電力ロスを9%から4%と半減させ、電力変換効率を世界最高水準まで向上させることができた。

3.2 PSU-CMU間の電力ロス削減

「京」では、PSUが集中的に配置されていたため（図-6）、PSU-CMU間の経路が長かった。この経路

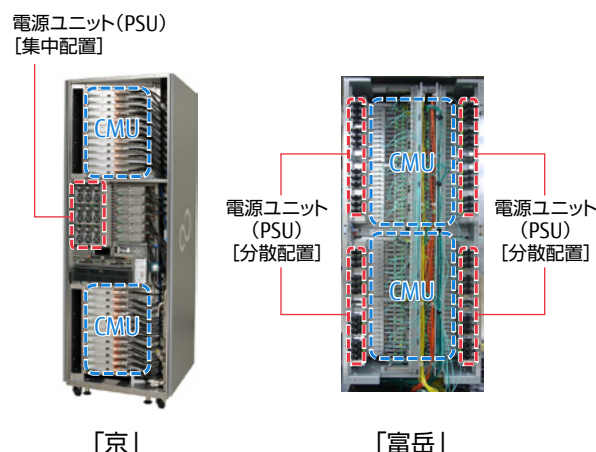


図-6 PSU配置

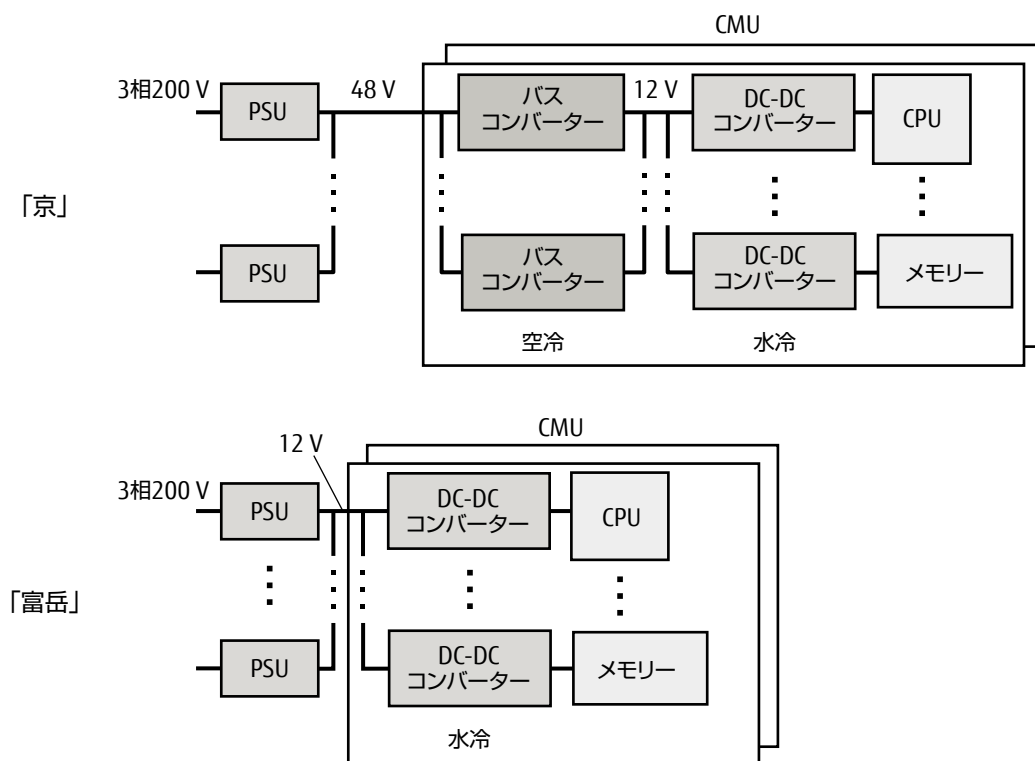


図-7 電源システム構成

を流れる電流による電力ロスや電圧ドロップが大きいため、CMUに供給する電源は48Vという高電圧であった。したがって、12V入力のDC-DCコンバーターへ電源を供給するためには、電圧変換回路であるバスコンバーターを置く必要があり、2段構成となっていた（図-7）。しかし、バスコンバーターは電力ロスが大きいため、PSU-CMU間の電力ロスを減らすとともに、バスコンバーターの削減にも取り組んだ。

図-6に示すように、「富岳」では、PSUをラックの両サイドに分散配置したことで、全てのPSU-CMU間の距離を短くでき、平均で従来の5分の1へ短縮することができた。これによってPSU-CMU間の電力ロスを削減でき、CMUに12Vの電源を供給することが可能となった。また、「京」で必要としていたバスコンバーターを削減し、電力ロスを改善している。更に、DC-DCコンバーターも最新のデバイスを採用することで、電力ロスを減らしている。これらの施策を合計すると、電力ロスを16%から10%に削減している。

以上の技術開発の結果、システム全体の総合電力ロスを従来の24%から14%と大幅に改善することができた。

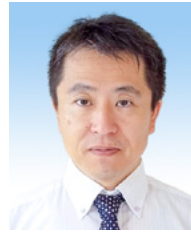
4. むすび

本稿では、スーパーコンピュータ「富岳」の高性能・省電力および優れた交換作業性を高いレベルで両立させたシステム実装技術について、具体例を交えて紹介した。

近年はスーパーコンピュータだけでなく、サーバ装置全体で高性能・省電力のシステム要求が高まっており、今後もその要求にこたえる技術開発を行っていく。

本稿に掲載されている会社名・製品名は、各社所有の商標もしくは登録商標を含みます。

著者紹介



江口 進 (えぐち すずむ)

富士通株式会社
プラットフォーム開発本部
サーバ装置の電源技術開発に従事。



梅松 三三雄 (うめまつ みさお)

富士通株式会社
プラットフォーム開発本部
サーバ装置の実装構造技術開発に従事。



大嶋 修 (おおしま おさむ)

富士通株式会社
プラットフォーム開発本部
サーバ装置の冷却技術開発に従事。

この記事は、富士通の技術情報メディア「富士通テクニカルレビュー」に掲載されたものです。他の記事も是非ご覧ください。

富士通テクニカルレビュー

<https://www.fujitsu.com/jp/technicalreview/>

