

# スーパーコンピュータ「不老」の運用状況と 次期システムへのファイル移行について

岩瀬 雄祐<sup>1)</sup>, 宇田川 暢<sup>1)</sup>, 林 秀和<sup>1)</sup>,  
山田 一成<sup>1)</sup>, 田島 嘉則<sup>1)</sup>, 毛利 晃大<sup>1)</sup>,  
椋木 大地<sup>2)</sup>, 星野 哲也<sup>2)</sup>, 片桐 孝洋<sup>2)</sup>

1) 名古屋大学 全学技術センター

2) 名古屋大学 情報基盤センター

iwase@nagoya-u.jp, udagawa@icts.nagoya-u.ac.jp, hd-hayashi@cc.nagoya-u.ac.jp,  
{yamada, ytajima}@itc.nagoya-u.ac.jp, a49982a@cc.nagoya-u.ac.jp,  
{mukunoki, hoshino, katagiri}@cc.nagoya-u.ac.jp

## Operations of Supercomputer "Flow" and File Migration to The Next System

Yusuke Iwase<sup>1)</sup>, Mitsuru Udagawa<sup>1)</sup>, Hidekazu Hayashi<sup>1)</sup>,  
Kazunari Yamada<sup>1)</sup>, Yoshinori Tajima<sup>1)</sup>, Akihiro Mouri<sup>1)</sup>,  
Daichi Mukunoki<sup>2)</sup>, Tetsuya Hoshino<sup>2)</sup>, Takahiro Katagiri<sup>2)</sup>

1) Technical Center, Nagoya University

2) Information Technology Center, Nagoya University

### 概要

名古屋大学情報基盤センターでは、2020年7月より、スーパーコンピュータ「不老」の運用を開始した。本稿では、今年度で運用を終了する現システムの運用状況、コールドストレージの性能検証、ならびに次期システムへのファイル移行について報告する。

## 1 はじめに

名古屋大学情報基盤センターは、文部科学省認定の共同利用・共同研究拠点として、研究、教育、社会貢献を目的にスーパーコンピュータシステムなどの計算資源を提供している。

本センターのシステムはスーパーコンピュータを取り巻く環境に対応するため数年おきに更新しており、2020年7月にスーパーコンピュータ「不老」の運用を開始した[1]。スーパーコンピュータ「不老」は4つのサブシステムと大容量のホットストレージ群およびコールドストレージシステム、可視化システムなどが高速ネットワークによって接続された複合的なシステムである。

現システムは、導入当初は2025年3月末の運用停止を予定していたが、1年の延長運用を行っており、2026年3月に運用を終了する。次期システムは2026年10月からの運用を予定している。

本稿では、今年度で運用を終了する現システムの運用状況、コールドストレージの性能検証、ならびに次期システムへのファイル移行について報

告する。

## 2 スーパーコンピュータ「不老」

スーパーコンピュータ「不老」は4つのサブシステムと大容量のホットストレージ群およびコールドストレージシステム、可視化システムなどが高速ネットワークによって接続された複合的なシステムである（図1）。

各サブシステムの特徴は以下の通りである。

- Type I サブシステムは超並列大規模計算用であり、理化学研究所のスーパーコンピュータ「富岳」と同じ計算ノードを搭載した FUJITSU PRIMEHPC FX1000 により構成される。
- Type II サブシステムは機械学習や AI などの研究分野向けであり、1ノードにつき NVIDIA Tesla V100 を 4 台搭載した FUJITSU PRIMERGY CX2570M5 により構成される。
- Type III サブシステムは、大容量メモリを使用する可視化処理用のプリポストサーバであり、

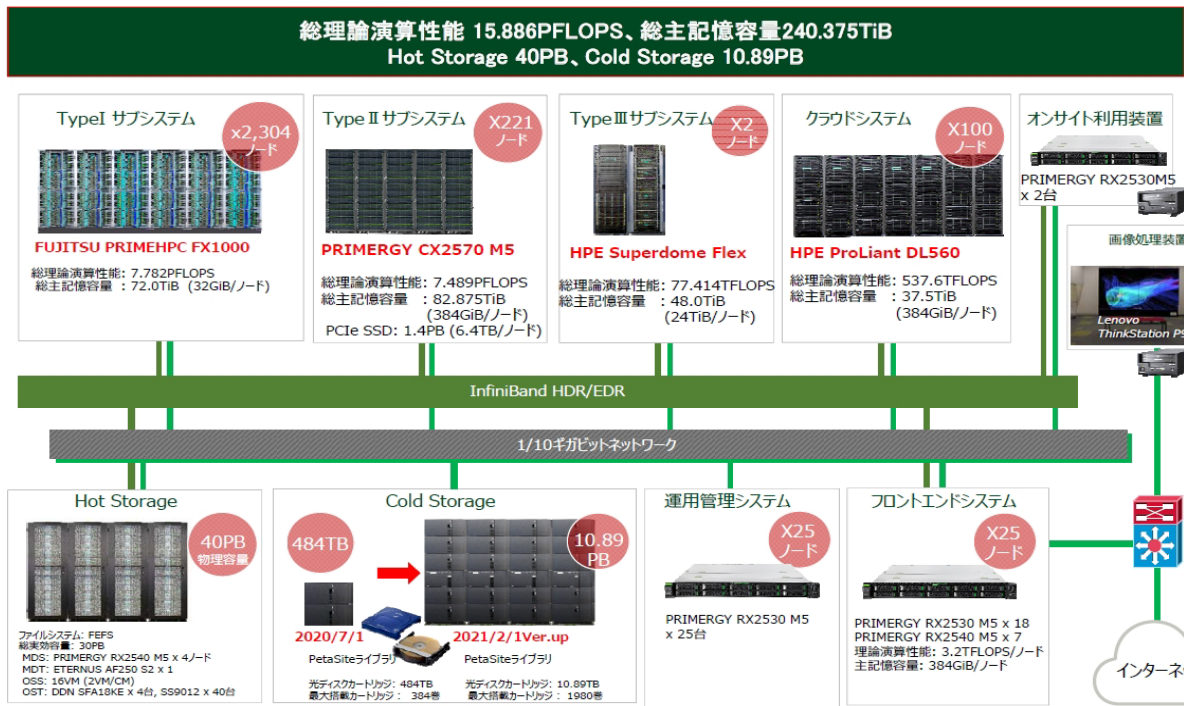
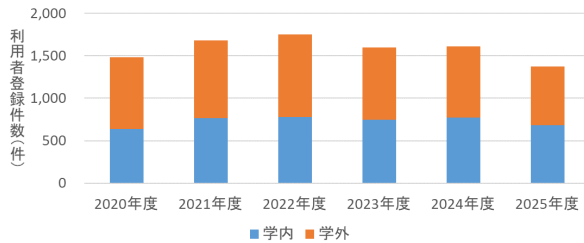


図1 スーパーコンピュータ「不老」



2025年度は7月までの値。

登録件数 (件)	学内	学外	合計
2020年度	635	849	1,484
2021年度	769	914	1,683
2022年度	776	978	1,754
2023年度	744	852	1,596
2024年度	773	839	1,612
2025年度	686	689	1,375

図2 利用者登録件数の推移

2 ノードで総メモリ容量が 48TiB の HPE Superdome Flex により構成される。

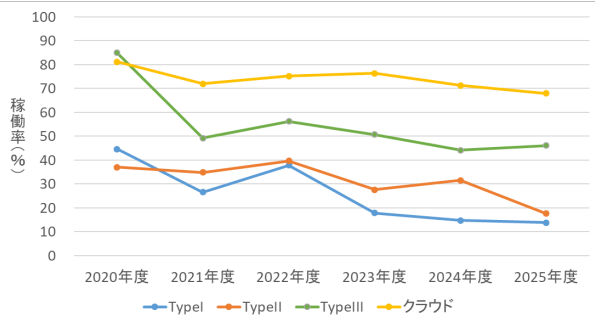
- クラウドシステムは、前システムの FUJITSU PRIMERGY CX400 の後継として HPE ProLiant DL560 を導入している。バッチ処理以外に、Web システムから時刻を指定しての利用にも対応している。

### 3 現システムの運用状況

現システムについて、年度毎の利用者登録件数の推移を図2に示す。利用者は学内と学外でおおよそ半分に分かれ、学外利用者が多い傾向にあり、共同利用・共同研究拠点として役割を果たしている。年度末における利用件数の増加は、2022年度の1,754件(学内776件、学外978件)をピークに落ち着いている。2025年7月現在の利用件数は

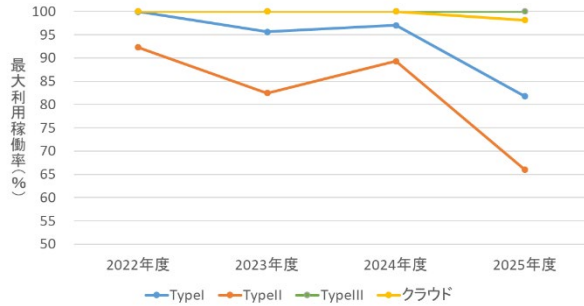
1,375件(学内686件、学外689件)となり、例年と比較して若干少ないが、例年に準ずる利用者登録件数が見込まれる。

現システムの各サブシステムについて、年度毎の稼働率の推移を図3に示す。電力料金の高騰により、2022年度から縮退運転を実施しており(表1)、同図は縮退運転ノードを除いた値となっている。富岳と同型のTypeIサブシステムの稼働率は、富岳の共用が開始された2021年3月を含む初年度が44%と高いが、富岳の次世代機の方針[2]が示された現在は低く、富岳の練習機としての役割を終えつつある。GPUを搭載したTypeIIサブシステムの稼働率は、初年度が37%と高くなかった一方、アーキテクチャが古くなった現在も大きく下がらず、GPUサーバとしての需要があるとみられる。可視化用のTypeIIIサブシステムの稼働率は、初年



2020年7月～2025年7月における年度毎の平均値。

図3 稼働率の推移



2022年4月～2025年7月における年度毎の最大値。

図4 最大稼働率の推移

表1 現システムの縮退運転

	期間	縮退運転
TypeI	2022年6月～2023年3月	2ラック停止
	2023年4月～2023年10月(10/20から稼働)	2ラック停止
	2024年4月～2024年11月(11/12から稼働)	2ラック停止
	2025年4月～2025年7月現在	2ラック停止
TypeII	2023年4月～2023年11月(12/01から稼働)	40ノード停止
	2024年4月～2024年10月(10/18から稼働)	38ノード停止
	2025年4月～2025年7月現在	40ノード停止

度が84%と高く、それ以降、50%程度の推移し、少ないノード数が有効活用されている。x86型のクラウドシステムの稼働率は、全期間を通じて70～80%程度と高い値を推移しており、高いニーズがある。

現システムの各サブシステムについて、年度毎の最大稼働率の推移を図4に示す。先の稼働率と同様に、同図は縮退運転ノードを除いた値となっている。TypeIサブシステムは、稼働率が低い一方、最大稼働率が高く、全ノード稼働による大規模な計算が多い。TypeIIサブシステムは、稼働率が低く、最大稼働率も低く、全ノードを利用しない小規模な計算が多い。TypeIIIサブシステムとクラウドサブシステムの稼働率と最大稼働率の両方が高く、全ノードを効率良く利用できている。

現在のスーパーコンピュータ「不老」は、2020年7月から運用開始して6年目に突入しているが、運用最終年度となっても、利用者登録件数が減らず、稼働率が低くなく、有効利用されている。現システムにおいて、富岳と同じアーキテクチャの採用、GPUサーバの採用は先見性があり、現システムの活用と延命に寄与したと考えられる。ただし、x86アーキテクチャは依然としてニーズがあり、クラウドシステム相当のシステムの拡充が必要とみられる。

## 4 コールドストレージシステム

### 4.1 コールドストレージシステムの概要

コールドストレージシステムは、国内のスーパーコンピュータで初となる業務用次世代光ディスク

クを使用したソニー社製のオプティカルディスク・アーカイブ PetaSite 拡張型ライブラリシステム（以下、ODA ライブラリ装置）<sup>1</sup>を採用した大容量・低コストのコールドストレージであり、2020年7月、総物理容量 484 TB にて運用を開始した。その後、2021年2月、総物理容量 6 PB、最大搭載容量 10.89PB に増強して運用している[3]。

ODA ライブラリ装置は、ODA カートリッジ（5.5TB/巻）を格納するキャビネット、ドライブ、キャビネット内の ODA カートリッジを識別してドライブにセットするロボット機構、ODA カートリッジを ODA ライブラリ装置に収容および搬出する機構で構成される。表 2 に、ODA ライブラリ装置の諸元を示す。

ODA ライブラリ装置は、製造元から告知があり、2025年3月末をもって販売終了となった[4]。

#### 4.2 コールドストレージシステムの仕様

不老のコールドストレージシステムは 4 台の ODA ライブラリ装置で構成され、また、ODA ライブラリ装置はそれぞれ 5 つのドライブを備えている。そのため、最大で 20 巻の ODA カートリッジを同時にマウント可能となっている。ユーザはマウントされた ODA カートリッジに対して、ファイルシステムを介してアクセスが可能である。その際に各カートリッジに対するユーザのアクセス制限を目的として、ODA カートリッジが持つシリアル番号に対して各ユーザのアクセス権を設定することが可能となっている。本アクセス制限は ODA ライブラリ内でのみ有効となる。

光ディスクは図 5 左側の ODA カートリッジで提供され、温湿度変化、水濡れ、紫外線、電磁パルスなど外的影響に強いなど、100 年間保存可能な耐久性を備えている。また、ODA カートリッジ内部に複数枚の光ディスクを備えており、不老で提供している ODA カートリッジでは WORM (Write Once Read Many) の光ディスクを 11 枚内蔵することで、ODA カートリッジ当たり 5.5TB の容量となっている。

また、高密度記録を実現するために光ディスクの物理構成はディスクの両面を使用し、片面当たり 3 層の記録層を持っており、ディスクの案内溝と溝間の両方に対する記録を可能とするランド&

表 2 ODA ライブラリ装置の諸元

項目	仕様
機種名	PetaSite 拡張型ライブラリシステム
総物理容量	2.72PB
総スロット数	495
総ドライブ数	5



図 5 ODA カートリッジと単体ドライブ

グループ記録方式を採用している<sup>2</sup>。

#### 4.3 コールドストレージシステムの性能

ODA カートリッジの特徴的な物理構成で性能を推測することが困難なため、シーケンシャル性能を測定したところ、図 6 のようになった。このベンチマークでは 1GB のファイルをコピーし続けることで測定しており、書き込み性能は最大でも 1.2Gbps 以下となっているが、これはファイルサイズに起因していると推測する。グラフより、書き込み性能は周期的であること、470GB 程度を 1 周期として、各周期において変化が 3 回生じていることがわかる。このことから記録層の 3 層構造と光ディスク交換の様子が見て取れる。ディスク交換のタイミングでは前後と比較して 120 秒以上処理が遅くなっていることから、この時間がドライブ内で光ディスクの交換に必要な時間と推測する。両面記録やランドおよびグループへの記録についてはドライブ側で隠蔽されているためかグラフから読み取ることが出来なかった。

また、ファイルサイズによるアクセス性能の変化の測定を行った結果を図 7 として示す。このテストでは各ディスクの 1 層目のみを利用するようにディスクの先頭から 100GiB 以内の領域で調整しながら、ファイルサイズを変更し測定を行った。図より、ファイルサイズが大きいほどスループットが向上していることがわかる。特に、ファイルサイズが 4096MiB 以上である場合には 1.4Gbps を示しており、これは先述したカタログスペックに

3Gbps、書き込みで最大 1.5Gbps とされている。

<sup>1</sup> <https://www.sony.jp/oda/>

<sup>2</sup> カタログスペックでは ODA ライブラリ装置に内蔵されたドライブの性能は読み取りで最大

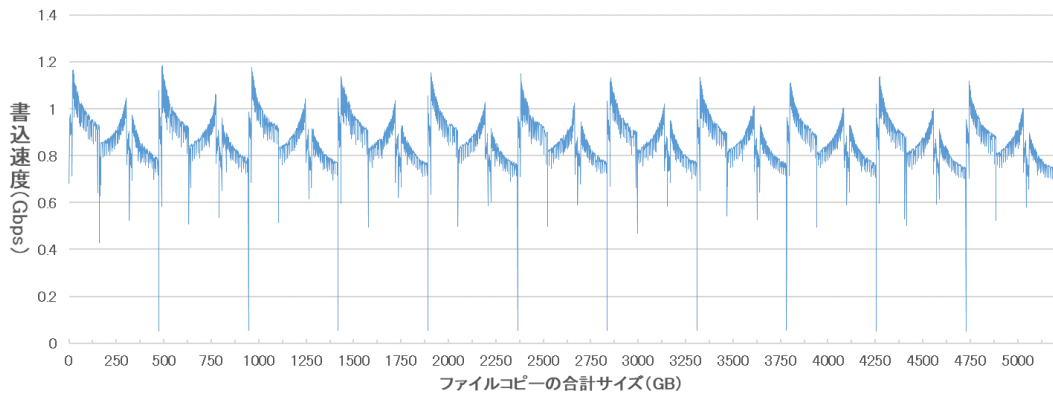


図6 ODA カートリッジのシーケンシャル性能

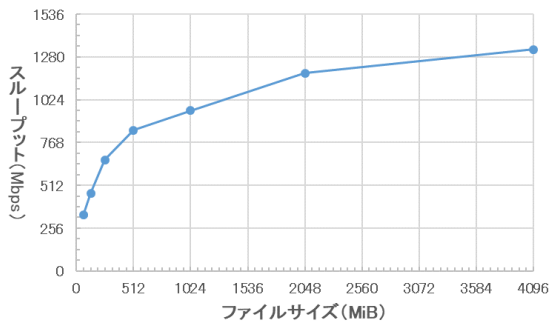


図7 ファイルサイズによる書き込み性能の変化

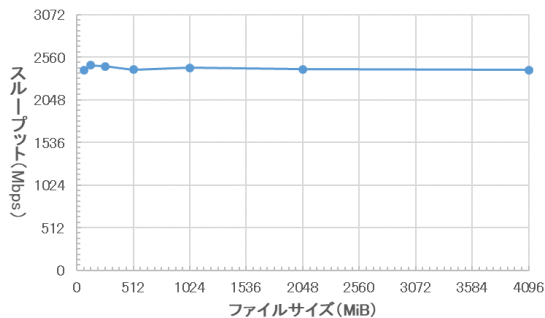


図8 ファイルサイズによる読み込み性能の変化

近い性能である。一方、ファイルサイズが小さいほどスループットが低下していることが分かる。これは、ディスクへの書きこみにあたり、データと同時にメタデータの記録も行われていること、ファイルサイズが小さいほどメタデータの記録頻度が向上し、時間を要していることが推測される。以上のことから、書き込むファイルサイズは大きいほど、特に4096MiB以上であるほど高いパフォーマンスを維持できると分析可能である。

ファイルサイズによる読み込み性能の変化の測定を行った結果を図8として示す。このテスト

では書き込み性能テストで書き込まれたファイルを読み込むことで測定をおこなった。同図より、ファイルサイズ間でスループットに変化がないことがわかる。

#### 4.4 コールドストレージ運用にあたる反省点

コールドストレージを導入するにあたり、データ科学や機械学習といった新しい利用スタイルにより、膨大なデータを収集し低コストで長期保存する需要を想定していた。しかし導入時に引き合いのあった大口利用者を含めた少数の大口利用にとどまり、小口の利用者が想定に対して増えなかった。

また、ベンダの都合により製品のロードマップがキャンセルされた。ODA ライブラリの場合は、製品が市場投入された2013年以降、長期的な製品ロードマップが示されていたものの、ビジネス上の判断により2025年3月末でODA ライブラリ本体の販売終了となった。製品サポートについては2032年3月末まで提供される見通しである。LTOのようなオープン規格で複数社から製品化されているものと異なり、独自規格の場合は提供しているベンダの都合により製品ロードマップが大きく見直され、場合によっては製品自体の提供およびサポートが突然終了してしまうリスクがあることを身をもって知る事となった。

## 5 次期システムへのファイル移行

### 5.1 ファイル利用状況と対応案

先述した通り、スーパーコンピュータ「不老」に保存されているファイルは、次期システムへ移行する計画である。ファイル移行を計画するにあたり、前提条件は以下の通りである。

ファイル利用状況 (3.1PB, 有効なアカウント数, 1,420 アカウント, 7/30 現在)

	ファイルサイズ	アカウント数	使用容量	
①	0~1GB まで	932	268TB	412TB
②	1GB 超~1TB まで	344		
③	1TB 超~5TB まで	75		
④	5TB 超~10TB まで	22	144TB	1,227TB
⑤	10TB 超~50TB まで	33	678TB	
⑥	50TB 超~100TB まで	8	549TB	
⑦	100TB 超~	6※	1,449TB	

※ファイルサイズ内訳 484TB, 421TB, 170TB, 141TB, 130TB, 102TB

- ①~④ 移行期間 (12 月中旬から 3 月) に利用者自身が最大 1~2 週間程度で移行可能.  
 ⑤~⑦ 転送に時間がかかる恐れがあり, なんらかの事前対応が必要と思われる.

図 9 ファイル利用状況と対応案

- ・ 次期システムの仕様書では「1PB 分の移行を支援すること」とあるが, ”支援”の内容が明確ではない.
- ・ そのため, 落札ベンダがホットストレージを用意し, 現行ベンダからのデータ吸い上げを行わない前提で検討.
- ・ スーパーコンピュータ「不老」は 80Gbps (40Gbps×2) で NICE (名古屋大学キャンパスネットワーク) と接続されている. この回線を利用して利用者はデータの引き上げを行う.
- ・ 特にデータ量が多い利用者については, 利用者支援室に設置されている, ホットストレージに対して直接データ読み込みを可能とするオンサイト利用装置を使ってデータ引き上げも可能.

ファイル利用状況と対応案について図 9 に示す. ファイル利用状況として, 全体では 3.1PB 程度となる. そのうち 10TB 超は 50 人程度, 10TB 以下は 1400 人程度となる. 10TB 移行の場合, 1 人あたり 1 週間で可能と推測される (学内の場合). 10TB 以下の利用者には 12 月頃のアナウンスにより各利用者自身で対応できると考えた. 10TB 超えの 50 人は早期のアナウンスが必要と考える.

学内の研究室のパソコン (1Gbps で NICE に接続) を用いて, スパコンに接続しファイル転送を行った場合, 容量 10TB で約 7 日程度を見込む. 理論値は約 28 時間だが, 帯域を占有出来ないことや, ファイルサイズによる効率悪化を考慮すると, 100TB あたり約 2 ヶ月, 400TB あたり半年を要する計算となる.

対応案として以下を考えた.

- ・ ①~④ 利用者自身で移行してもらう

- ・ ⑤~⑦ 利用者 (計 47 アカウント) へ早期より個別連絡し, 移行対象データの整理, 不要データの削除を促す.

## 5.2 コールドストレージの利用

余剰の ODA カートリッジを利用し, コールドストレージにより 1.5PiB 分のデータ移行を支援するシミュレーションを行った. ODA ライブラリ装置を利用してバックアップを行うまでをスコープとし, バックアップ作業はセンター側で実施する場合と利用者側で実施する両方のシナリオを検討した.

図 7 で示した通り, コールドストレージはファイルサイズにより大きくアクセス性能が変化するため, 効率よくバックアップを行う場合は事前に tar などでファイルをアーカイブしておくことが望ましい. センター側でバックアップ作業を行う場合は, 作業を自動化した際には最速 1 週間程度となる. これを実現するために, 作業開始前の指定した期日までに利用者側でバックアップ対象のファイルを tar でアーカイブしておくよう依頼しておけばよい. 一方, 利用者側でバックアップ作業を行う場合は, 20 台の内蔵ドライブのスケジューリングを考える必要があり, 利用者側で高効率なアクセス性能となるよう期待通りにアーカイブ作業を行っているかが問題となる. 理想的な条件でも 5TiB 分のバックアップにおよそ 8 時間必要であり, 平均ファイルサイズによって容易に数倍に膨れ上がってしまう. 利用者が割り当てられたスケジュールどおりに ODA カートリッジをマウントしてバックアップを行い, アンマウントまで完了して次の利用者が作業できる状態になるかも重要となる. このように不確定要素が多いため, バ

ックアップ作業を利用者に任せることは現実的でない。

シミュレーションのスコープ外としたが実際には次期システムへのデータの復元も行う必要があり、現時点では現契約期間終了後の ODA ライブラリの扱いも未定である。センターが所有する単体ドライブは今後も利用可能だが、300 巻の ODA カートリッジを数時間ごとに手動で入れ替えながらデータをリストアすることは作業コストだけでなく、1 日あたりドライブ 1 台で 2 巻程度、2 台のドライブを利用しても 1 日 4 巻しかリストアできないと考えた場合、完了するまでに最悪 75 日程度必要となることから実現性は厳しい。最終的に ODA ライブラリのレンタルアップ品買い上げの検討や、次期システムの落札ベンダとの調整を踏まえて決定する必要がある。

## 6 おわりに

本稿では、今年度で運用を終了する現システムの運用状況、コールドストレージの性能検証、ならびに次期システムへのファイル移行について報告した。スーパーコンピュータ「不老」は、2020 年 7 月から運用開始して 6 年目に突入しているが、富岳と同じアーキテクチャの採用、GPU サーバの採用によって、運用期間を延ばすことができた。コールドストレージは、膨大なデータを収集し低コストで長期保存することに優れているが、製造元から販売終了の告知があり、対応に苦慮した。次期システムへのデータファイル移行について、ファイルサイズ毎にクラス分けすることで見通しが立った一方、大量データを保有するユーザの対応が課題として残り、コールドストレージの再利用案が急浮上していることに戦々恐々としている。

次期システムは 2026 年 10 月からの運用を予定しているが、スーパーコンピュータ本体の性能や利用者の利便性だけでなく、運用業務の遂行が効率的になるように努めたい。

## 参考文献

- [1] スーパーコンピュータ「不老」,  
<https://icts.nagoya-u.ac.jp/ja/sc/>, 2025 年.
- [2] スーパーコンピュータ「富岳」の次世代となる新たなフラッグシップシステムの開発・整備を開始,  
[https://www.riken.jp/pr/news/2025/20250122\\_1/index.html](https://www.riken.jp/pr/news/2025/20250122_1/index.html), 2025 年.
- [3] 山田 一成, 田島 嘉則, 高橋 一郎, 毛利 晃大, 林 秀和, 片桐 孝洋, 大島 聡史, 永井 亨, スーパーコンピュータ「不老」コールドストレージと民間利用制度について, 大学 ICT 推進協議会 2021 年次大会論文集, pp. 99-105, 2021 年.
- [4] オプティカルディスク・アーカイブ商品 販売終了のお知らせ,  
<https://www.sony.jp/oda/info2/20230131.html>, 2023 年.