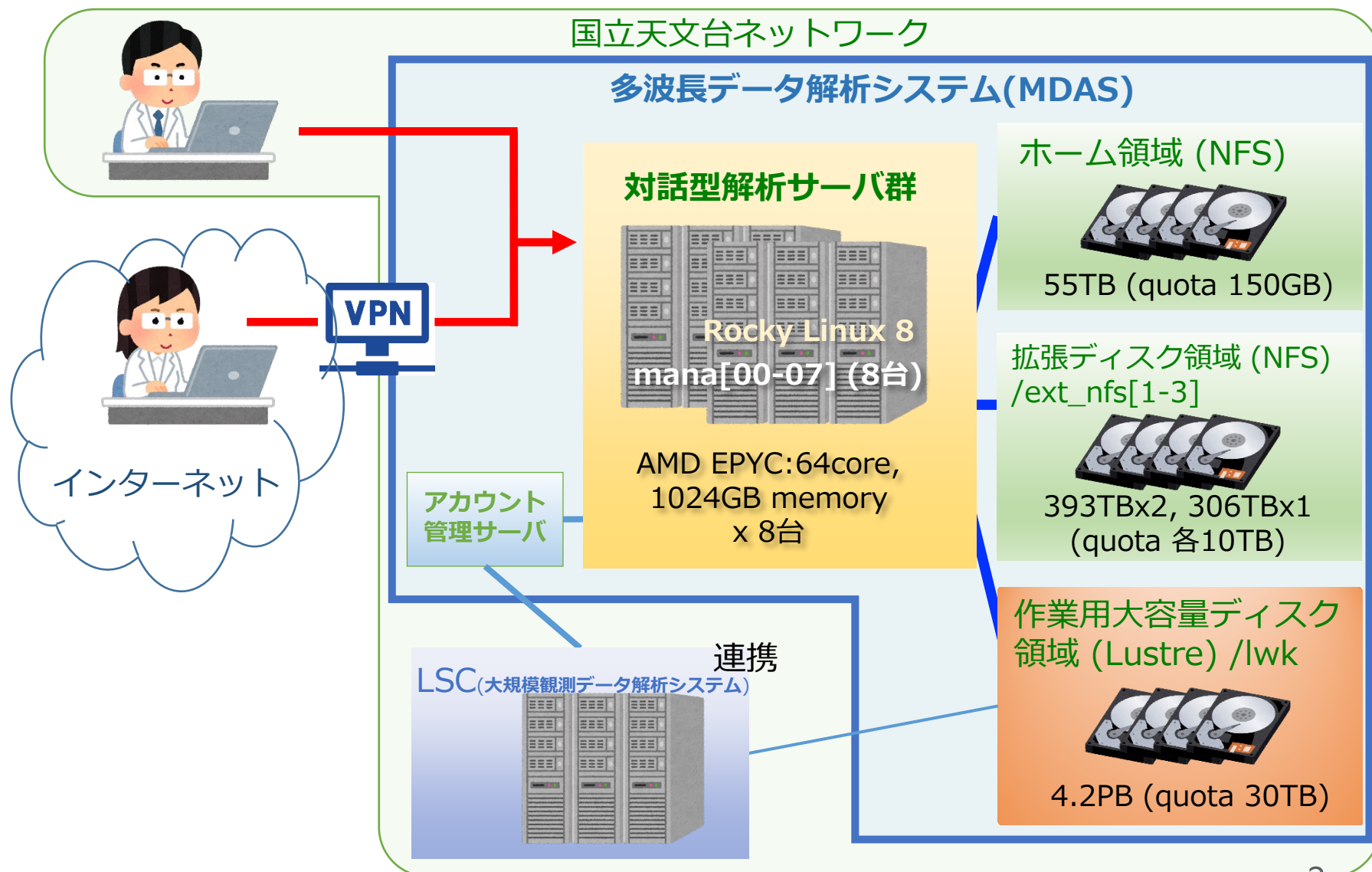


多波長データ解析システム

The Multi-wavelength Data Analysis System (MDAS)

巻内慎一郎、共同利用運用グループ (ADC)

システムの構成と接続



MDAS 解析サーバ スペック一覧

	対話型解析サーバ
ホスト名	mana[00-07].ana.nao.ac.jp (8台)
使用機器	Gigabyte R182 base CMS Custom Server
OS	Rocky Linux 8 (RHEL 互換)
CPU	AMD EPYC 7543 (32 コア/2.8GHZ/MAX3.7GHz) ×2 (合計 64 core)
メモリ	1024GB (DDR4/3200Mhz ECC対応 64GB ×16)

システム全体

★ 総CPUコア数: **512**

★ 総メモリ量: **8TB**

★ 総ディスク容量: **5.2PB** (ホーム 55TB, NFS 1PB, Lustre 4.2PB)

☆現時点では全領域データ保存期間の設定無し(定期削除無し)

システムリプレイス(2024)方針

- 前回ユーザズミーティング(2021)の参加者およびその他の利用者からのご意見・ご要望の内容を検討してシステム規模や構成を決定
- システム全体のCPUコア数は同程度を維持
 - システム合計512コア、解析サーバ1台の搭載コア数を増やして(16→64)、サーバ数を減少(32台→8台)
- 解析サーバ1台のメモリ容量は増強
 - 1台当たりの搭載メモリ 192 or 256GB →1024GB
- ストレージ容量は大幅に増強
 - ユーザが扱う観測データの巨大化に伴い、ストレージ容量不足が常態化していたため、容量増大の要望が多かった
 - サーバごとのローカルディスクは廃止し、共通の大容量作業領域(Lustre FS)に集約

リプレイス後のシステム構成

- レンタル機器とリース機器によるハイブリッド構成
- アカウント管理サーバ、VPN接続装置、検証用解析サーバ1台は引き続き**レンタル機器**で構成する
 - 基本的な管理は**レンタル業者**が行う
- MDAS の主要部分(解析サーバ、大容量ストレージ)は**リース機器**により構成する
 - 天文台側によるシステム構築と管理・運用
- MDAS のバッチ型解析サーバは廃止
 - 代替として LSC の一般利用キューが使用可能

システムリプレイス経過

- 2024年3月 リース機器の借入開始
 - リース契約期間 2024/3/1 - 2029/2/28
- システム構築作業 2024年3月～6月
- ユーザデータ移行期間 2024/05/09 – 6月末
 - 旧システムの全ユーザデータ(home, 各サーバのローカル領域, 作業用NFS領域)について、ユーザ自身によるNFS領域 (/ext_nfs[1-3]), 新作業用領域(/lwk)への移行
- システム停止(現新切り替え作業) 7/1 – 7/3
- 新システム稼働開始 2024/7/4
- 解析サーバやVPNサーバのホスト名はすべて変更

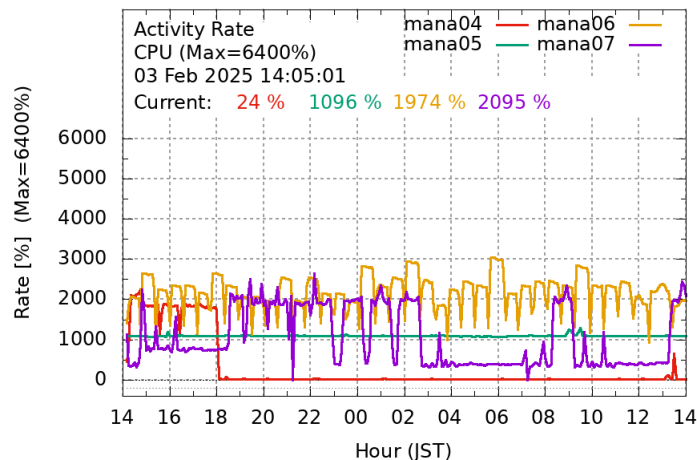
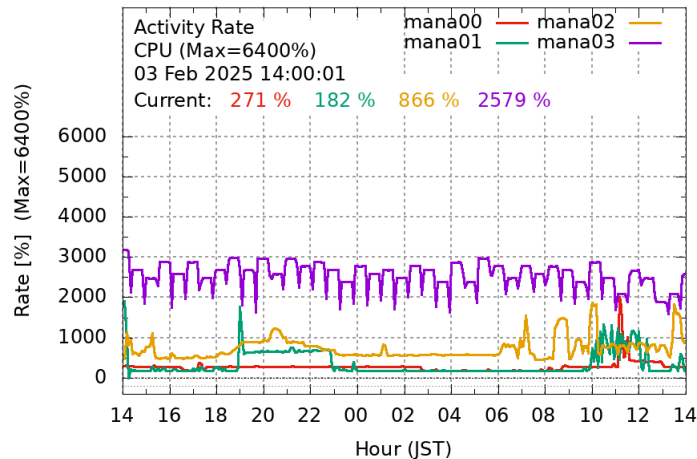
運用の現状

- 新システム稼働後、大きなトラブルはなく、概ね安定的に動作している
- 計算機室内でのシステムのラック移設作業のため一時停止を実施(2024年9月25-27日)
- 今後の定期保守は、前システムで実施していたのと同様に3ヶ月に一度の実施を予定
- ソフトウェア整備については主要なものはインストール済。前システムで利用があった一部のソフトで未インストールあり(対応継続中)
- ソフトウェアに関してはユーザからの要望等も踏まえて適宜対応している
 - 前システムには無かった新たなソフトインストールも実施
 - ユーザ自身でユーザ領域にインストールしたソフトの利用も増えている

現在の利用状況

2025年2月3日の例

解析サーバのCPU稼働率



	CPU	Memory	ログイン ユーザ数	ロードアベ レンジ値
	(Max=6400%)	(Max=100%)		(5min)
Hostname	CPU(=usr+nice)	Memory(user)	N_Login	Load avg.
mana00	271.4%	0.7%	7	5.66
mana01	182.4%	0.0%	19	7.39
mana02	865.9%	0.5%	14	23.66
mana03	2578.6%	7.9%	4	22.78
mana04	25.0%	0.6%	19	0.60
mana05	1100.2%	0.0%	8	14.41
mana06	1997.4%	9.5%	25	45.88
mana07	2313.6%	7.1%	12	59.67

- 各サーバが定常的に利用されている
- 長時間の処理プロセスを流すユーザが増えている
- 時々、非常にヘビーな処理を行うユーザも見られるが、ガイドライン(次ページ)に沿った利用をお願いしている

リソース使用ガイドライン

- 最近では多くのユーザが、あるいは、利用ソフトウェアが、マルチコアを使用した並列処理を要求する
- マシンのリソースを使い切ると、処理のレスポンス低下や他のユーザの作業を妨げる原因となるため対策が必要
- 前システムに比べると、サーバ1台のコア数増加により1ユーザがサーバを完全に占拠するような可能性は低くなったが、依然起こりうる
 - ➡ 一人のユーザがリソースを占有しないようにガイドラインを策定
 - ガイドラインを超過して使用中のユーザには自動メールで通知

新システムのガイドライン(ユーザ 1 人による同時利用の最大目安)

- **mana00**: 対話的使用優先サーバ。最大4コア
- **mana[01-07]**: 最大16コア
- MDAS 全体での1ユーザの同時使用は、合計で最大80コアまでとする

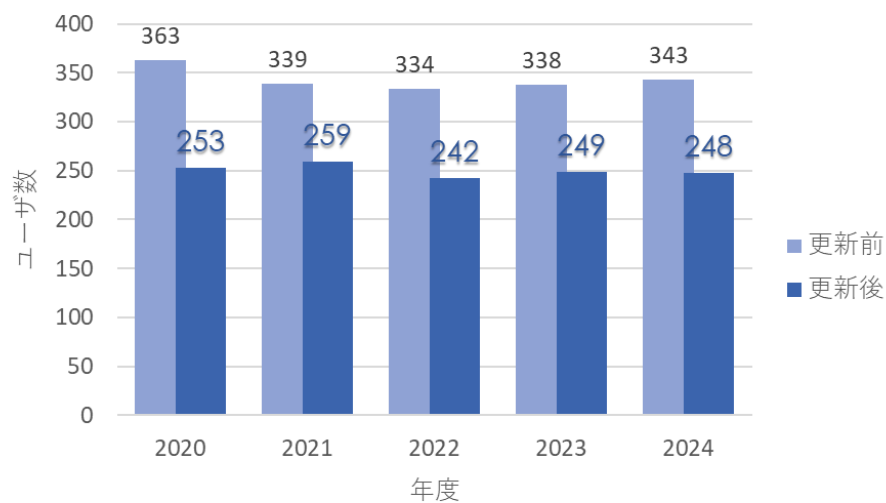
現在の運用上の問題点

高ロードアベレージ問題

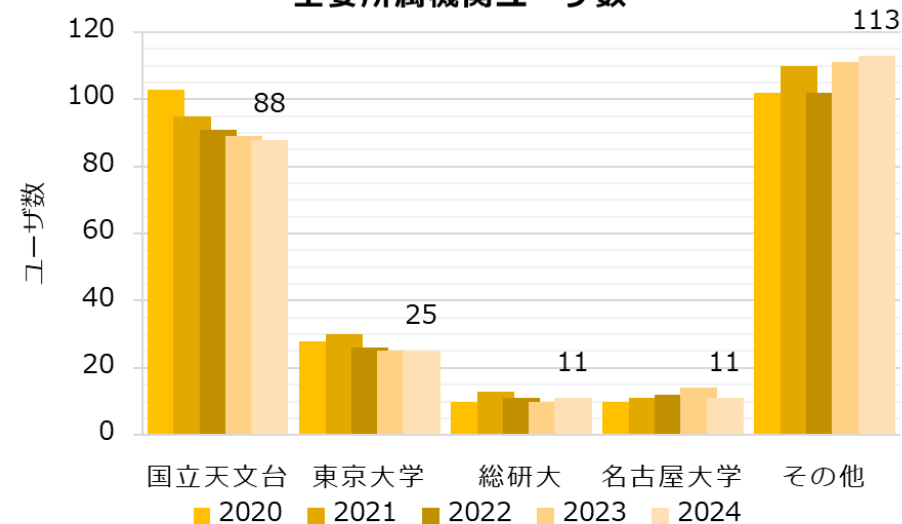
- サーバ搭載コア数64に対して二桁から三桁(～数万)以上大きなロードアベレージ(LA)値が観測されている
- サーバの動作が非常に重くなる。ただし、時間の経過により自然復旧する(完全フリーズによるマシンの再起動などは必要なし)
- (主に)CASA のプロセスが引き起こしている模様
- 問題発生の原因(メカニズム)は未解明。現在のところ再現性も不明
- 調査のため、原因プロセスのユーザへ個別にコンタクト予定です。ご協力よろしくお願いします

ユーザ利用動向

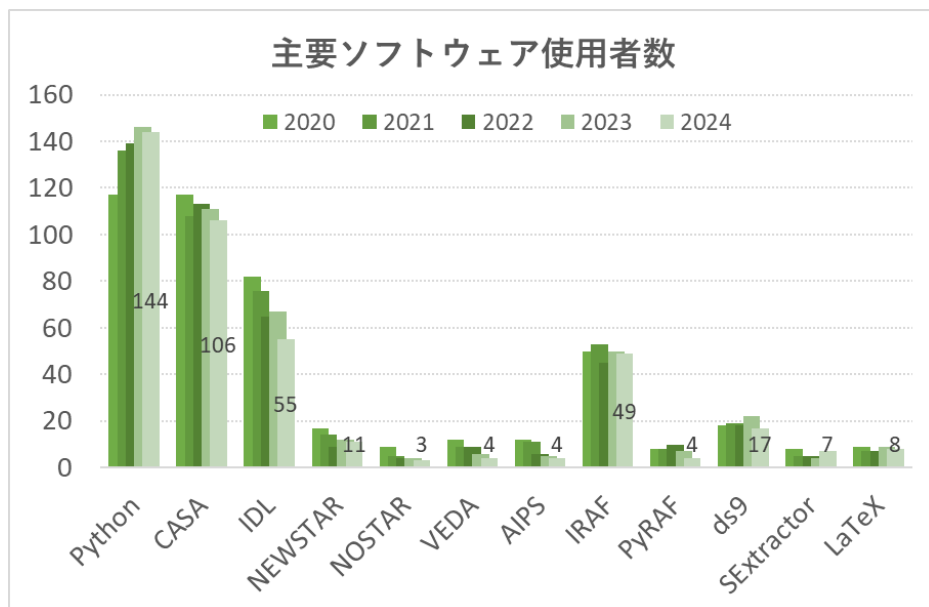
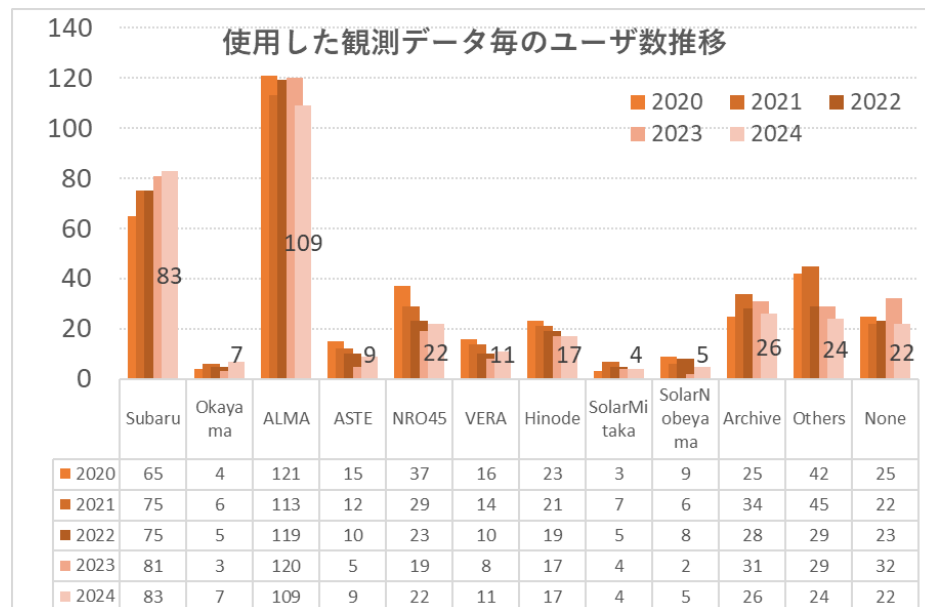
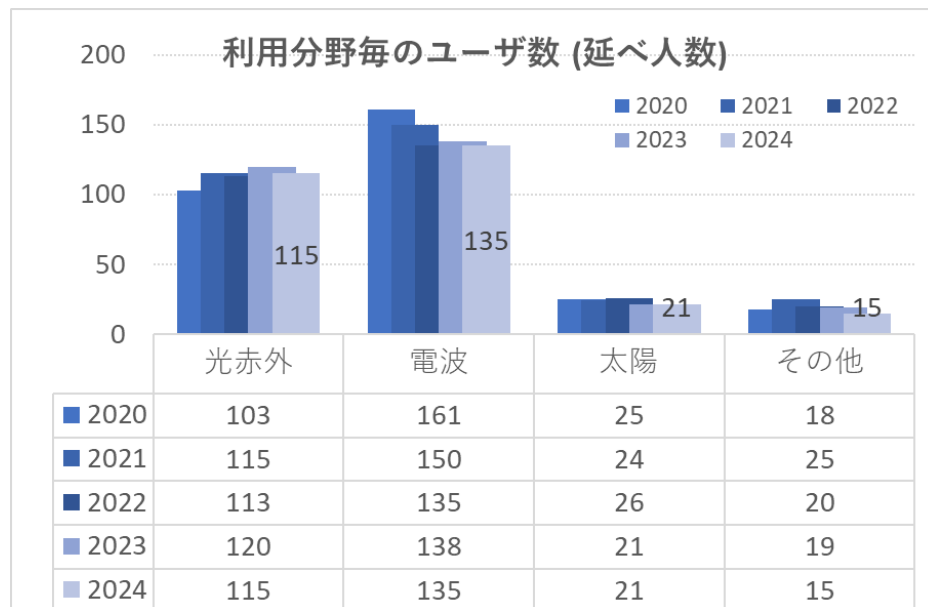
アカウント更新前後の有効アカウント数



主要所属機関ユーザ数



- 最近のユーザ数はほぼ一定
- アカウント更新後で250名程度、年間で100名ほどが増加してアカウント更新直前は350名程度になる
- 国立天文台のユーザが多い一方、ユーザーの約2/3は大学や他研究機関の研究者



- ユーザの分野別の分布(重複ユーザあり)では、電波がもっとも多く、光赤外も近い数
- MDASで使用した観測データは ALMA が多く、CASA がよく使われている
- すばるデータも多く、若干増加傾向

今後の方針と方向性

次期リプレースに向けての検討

- 現行システムの OS には Rocky Linux 8 が採用されている
 - Rocky Linux 8 のメンテナンスサポート期限は**2029年5月31日**
 - リース機器借入契約期間も**2029年2月**で終了
- 2029年以降のMDASをどうするか？
- レンタル？ リース？ 全面的な買取機？
- 予算的には厳しい状況が続くことが予想される。
検討にあたってはユーザからの要望や意見が重要