

国立天文台天文データセンター共同利用計算機システムの 運用効率改善：ユーザ情報のデータベース化

○亀谷和久、巻内慎一郎、田中伸広、磯貝瑞希、小澤武揚、藤川真記子、市川伸一、高田唯史（国立天文台天文データセンター）

概要

国立天文台天文データセンターの共同利用計算機システムである多波長データ解析システムのユーザ情報のデータベース(DB)化について報告する。本システムは国内外の研究者へデータ解析用の計算機環境を提供している。そのユーザアカウントの新規申込や登録済みの数百名のユーザから届く各種登録情報を一元管理するためにDB化を進めている。これは各種ユーザ情報の統計分析のための基本情報となることも期待されている。

1. 背景と開発目的

国立天文台天文データセンター(以下、ADC)は、共同利用の天文データ解析環境として、「多波長データ解析システム」[1]を国内外の天文学及び周辺分野の研究者、大学院生を対象に提供している。本システムは合計34台の高性能サーバ群と合計2.6PBの大容量ストレージ、そして三鷹キャンパスの共同利用計算機室等に設置された端末群から成る。天文データ解析や科学計算のためのソフトウェアが100種類以上インストールされており、ユーザは端末あるいはリモート計算機からログインして利用することができる。2018年末の時点でのユーザ数(使用できる状態にあるもの)は、336名である。

本システムの利用を希望する者は、まずADCのウェブサイト上に設けられた申請フォームを用いてアカウントIDを申請し、承認されればアカウントが発行される。ユーザは、年に1回のアカウント更新、VPNアカウント申請、グループID申請等の各種申請を同様にウェブサイト上のフォームを用いて行なう。これらの申請情報は、従来はADC共同利用運用グループ内の担当者が管理するファイル等に手動で登録し管理していた。また、申請情報には、ユーザの需要や利用傾向を知る貴重な情報が含まれると考えられるが、一部は十分に活用できていなかった。

そこで、ユーザ情報を一元化して管理を効率化するために、ユーザ情報のDB化を行なった。このユーザ情報DBへの情報登録作業を可能な限り自動化するため、ウェブインターフェース(I/F)を開発した。さらに、ユーザ情報の傾向を定量的に捉えるため、統計情報の自動的な取得及び可視化を行なった。

2. アカウント申請と登録の手順

新規アカウント申請を例として、申請からユーザ情報DBへの登録までの手順を図1に示す。①申請者(ユーザ)はウェブサイト上の申請フォームに登録情報を記入し送信する。これをウェブサーバが受け取ると、サーバ上にテキスト形式の申請内容ファイルが生成される。それと同時に申請者とADC担当者宛てに②申請内容確認メールが送信される。その内容をADC内で③確認・審査し、承認されると④アカウント登録作業が実施される。登録後、今回開発した⑤ユーザ情報集約DBへの登録作業へと進む。

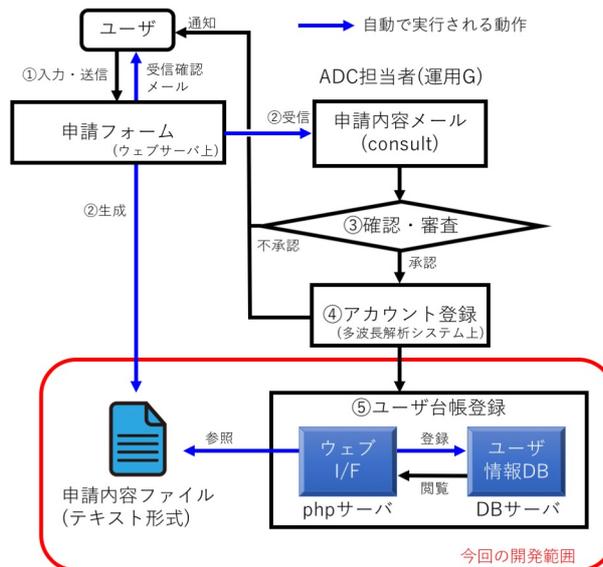


図 1. 申請から DB への情報登録までの手順(新規アカウント申請時)

3. データベースの構成

ユーザ情報 DB の構築には PostgreSQL 10.6 を利用した。ユーザ DB を構成するテーブルは申請フォームの種類毎に作成した(図 2)。最も重要なテーブルはユーザ情報を統括するユーザ台帳である。この他に年度に 1 回実施されるアカウント更新、期間外アカウント更新(ロック解除)、VPN 利用申請、グループ ID 利用申請の各申請フォームの登録情報を格納するテーブルを用意した。ユーザ台帳のカラムは、新規アカウント申請フォームで登録する各項目に加えて、アカウント等の認証情報、VPN 取得状況や使用データ等のシステム利用情報、そしてユーザ情報管理のための新規登録日、最終更新日、更新内容等の付加情報を用意した。ユーザ台帳以外の各テーブルには、各申請フォームの登録内容が全て格納される。ユーザ台帳に存在する項目については、後述のウェブ I/F を介してユーザ台帳にも反映される。

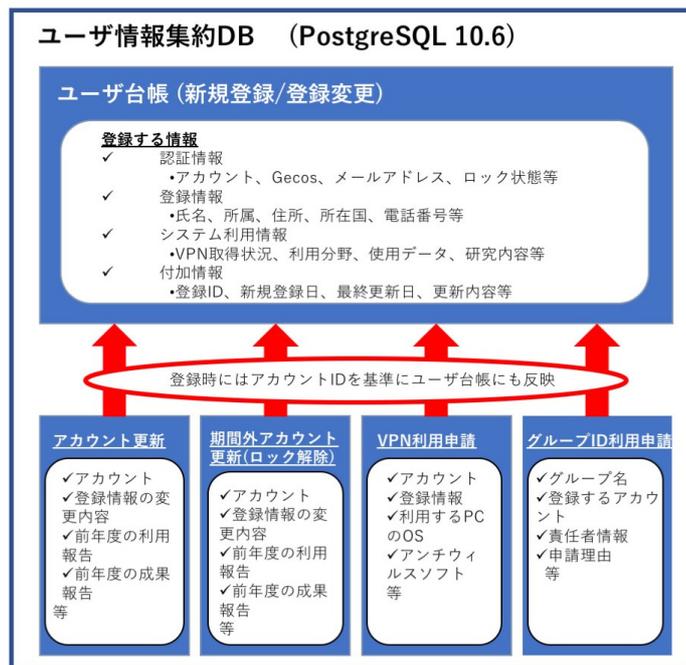


図 2. ユーザ情報集約 DB のテーブル構成

4. ウェブインターフェースの開発

ウェブ I/F の開発には html および php 7.2 を利用した。html でウェブサイトを作成し、php を埋め込むことにより PostgreSQL と連携させた。ウェブ I/F は大別して DB 上の登録情報の閲覧のためのページと登録情報の確認、変更のためのページの 2 種類を作成した(図 3)。閲覧ページでは、任意のカラムを選択して表示できるように工夫した。また、新規情報登録のためのフォームおよび各レコードの内容確認、修正、削除のためのボタンを用意した。例えば、新規登録の際は、未処理の申請内容ファイルリストから該当のものを選択し登録ボタンを押下すると、申請内容が自動的に整形され、確認ページに移行する。その内容を確認、修正した上で登録を行なう。一方、申請情報確認変更ページでは、各レコードに含まれる全てのフィールドの情報を表示し、適宜変更することができる。上記のユーザ情報集約 DB とウェブ I/F により、ユーザ情報の一元管理を実現した。



図 3. 開発したウェブ I/F の一例(左：閲覧ページ、右：登録情報確認変更ページ)

5. 統計情報の自動取得と可視化

ユーザ情報集約 DB に登録された情報から、PostgreSQL の機能を利用して統計情報を自動的に抽出し、ウェブサイト上に phplot 等のグラフ描画ツールを利用して表示する機能を実装した。この目的は、一つは各種報告を作成する際に迅速に現状の統計情報を把握するため、もう一つはシステムの運用・管理および次期システム仕様検討等に資する基本的な利用状況を把握するためである。取得した統計情報の例を図 4 に示す。登録ユーザの内訳(a)では、これまでに多波長データ解析システムのアカウントを取得したユーザのうち、現在使用できる状態にあるものが約 6 割であることがわかる。また、ユーザのデータ別利用者数(b)では、多波長データ解析システムを利用して各観測所・望遠鏡のデータを解析しているユーザ数を示しており、すばる望遠鏡や ALMA のデータを解析するユーザが多いことがわかる。

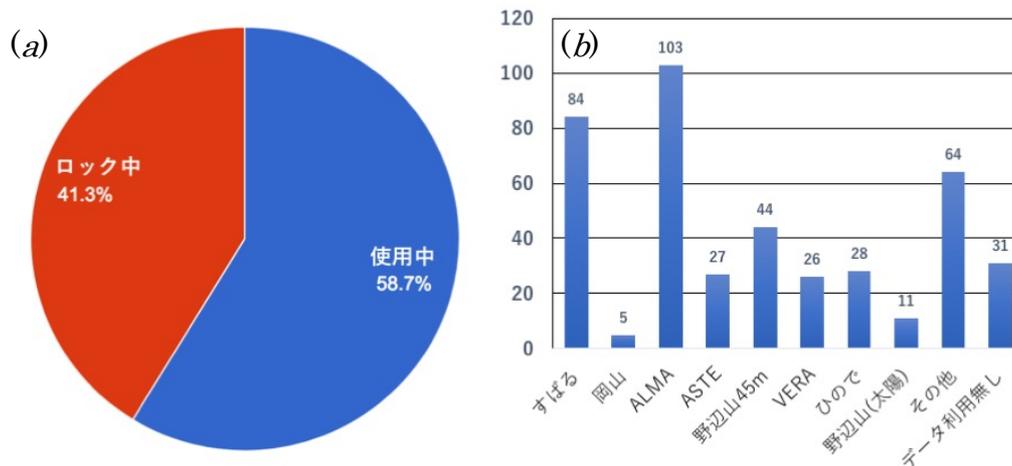


図 4. 取得した統計情報の例。(a)登録ユーザの内訳(ロック状態) (b)ユーザのデータ別利用者数

6. セキュリティ対策

今回開発したユーザ情報集約 DB は、個人情報も含むため、セキュリティには十分な配慮が必要である。セキュリティ対策として、サーバのネットワーク環境の制限、サーバ間の接続の暗号化、認証制限、およびソフトウェアの脆弱性対応等を施している。今後もサーバのプライベートネットワーク内での運用など、順次セキュリティを強化していく予定である。

7. まとめと今後の予定

本稿では、多波長データ解析システムのユーザ情報集約のためのデータベース開発、ウェブ I/F 開発、統計情報の自動取得と可視化、セキュリティ対策について報告した。今後は引き続き主に以下の項目の開発を行ない、このシステムの改善を進める予定である。

- ウェブ I/F の機能改良
- 取得できる統計情報の拡充
- 多波長データ解析システム関連で開発された他データベース[2][3]との連携
- セキュリティ対策の強化

これらにより、ユーザ情報の管理の効率化と利用促進を実現し、システム運用の効率化に貢献する。

参考文献

- [1] 多波長データ解析システム <https://www.adc.nao.ac.jp/MDAS/> (2019年1月31日閲覧)
- [2] 小澤武揚ほか, 「国立天文台天文データセンター共同利用計算機システムの運用効率改善 1: システムログのデータベース化」, 第 37 回天文学に関する技術シンポジウム 2017 集録
- [3] 山中郷史ほか, 「国立天文台天文データセンター共同利用計算機システムの運用効率改善 2: 問い合わせメールの DB 化と共有のためのウェブインターフェースの開発」, 第 37 回天文学に関する技術シンポジウム 2017 集録