

# 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

○山中郷史, 田中伸広, 磯貝瑞希, 小澤武揚, 亀谷和久, 卷内慎一郎, 藤川真記子, 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では、現在運用中の多波長データ解析システムに対する共同利用ユーザーからの要望に対応するため、メールによるサポートを行なっている。メールにて送られてきたシステムへの要望に関する情報を集約し、今後の運用改善と効率化を目的としたメールの管理・共有のためのソフトウェア開発を行なった。本ソフトウェアは PostgreSQL、PHP 等から構築され、現在はベータ版の開発が完了し試験運用を開始している。本ソフトウェアは 2018 年 3 月から稼働開始予定の次期多波長データ解析システムと同時に本格運用を開始する予定である。

## 1. 背景

### 1.1 多波長データ解析システムとは

多波長データ解析システム（以下、解析システム）とは、国立天文台天文データセンターが運用しているシステムであり、あらゆる波長で観測された観測データの解析を目的として構築された計算機群である。大学共同利用の計算機システムであり、『国内外の天文学関連分野における大学院生以上の研究者』であれば誰でも使用することができる。

### 1.2 Consult メール（問い合わせメール）について

解析システムに関する質問・要望等に対応するために設けている問い合わせ用メールアドレスが“consult\_アット\_マーク\_ana.nao.ac.jp”である（\_アット\_マーク\_は@で置き換え）。アドレスに consult が含まれていることから、我々のグループ（以下、運用グループ）では Consult メールと呼んでいる。このメールアドレスは以下の問い合わせの窓口となっている。

- ・解析システムを利用するための新規アカウントの申請
- ・既存アカウントの更新
- ・解析システム専用 VPN の利用申請
- ・アカウントのグループ化申請
- ・システムに対する要望（ソフトウェアのインストール、不具合対応など）
- ・システムに関する質問

（・天文データセンターで管理している共同利用室の設備に関する質問）

Consult メールが運用グループに配送されるまでを模式的に表したものが次頁の図 1 である。最初にユーザーから送られてきたメールは、運用グループで管理しているメールサーバーへと配送される。こ

ここでメール 1 通 1 通に **Consult ID** と呼ばれる固有の数字が件名に付与される。その後、ID が付与されたメールは運用グループ全体に一斉に配送される。問い合わせの種類によって運用グループ内の担当者は決まっており、メール受信後は担当者が各々対応を行う。この際 **CC** に上記 **consult** のアドレスを加えて返信することで、メールの返信に対しても固有の **Consult ID** が付与され、かつ運用グループ全体にもメールが共有されるようになっている。

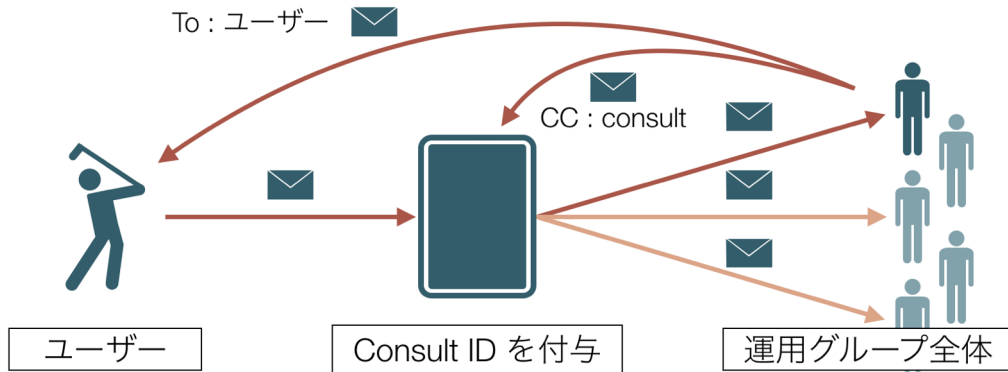


図 1. メール配送の模式図

しかし、この配送モデルにおいては『あとから対応したメールを参照することが手間である（個々人が自分の PC 内を検索する必要がある）』や『ユーザーからの要望・質問が既に解決しているのかが不明瞭である（もちろん運用グループで相互に確認をしており、アカウント関係の申請は漏れがないよう別の方法で毎日チェックはされている）』などいくつかの問題点があった。

### 1.3 本システムの目的

1.2 で述べたような問題点を可能な限り減らし、今後の運用業務改善と効率化を図ることは非常に重要である。実際、1ヶ月の間に送られてくる **Consult** メール総数は 200 通に上るため、決して少なくない時間がユーザー対応業務に割かれている（ただし、実際の対応件数は 200 よりもっと少ない数字である。理由については 3.1 で述べる。）。そこで以下の 3 つを目的としたソフトウェア開発を行った。

- ① **Consult** メールを対応後も簡単に参照できるようにデータベースを用いて適切に管理する
- ② データベースに集約したメールの統計情報を計算し、日々の業務を数値化する
- ③ 運用グループ内で簡単に共有できるようにインターフェースを整備する

## 2. システム開発

### 2.1 システム概略

1.3 で述べた目的のために、右記図 2 で示すような大きく分けて 3 つに分類されるソフトウェアの開発を行った。詳細については 2.2 以降から順番に述べていく。

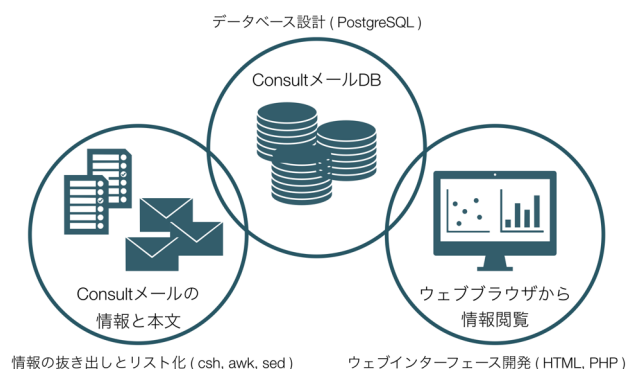


図 2. システムの概略図

## 2.2 データベース設計

データベースの構築には PostgreSQL を使用した。カラムには Consult ID、メール受信日、件名、送信者、宛先、などの一般的な情報の他に“Parent Consult ID”という情報を用意した。これは最初に送信されたメールを“親”と定義し、後に続く返信を“子”と定義した際の“親”の Consult ID を入れるカラムである。この工夫を施すことにより『メールのスレッド表示』『質問・要望の解決までに要した時間の計算』『相談件数（≠メール総数）の計算』などの点で便利になると考えて取り入れた。

## 2.3 Consult メール自動データベース化

データベースでメールの情報を取り込むためには、受信した Consult メールから必要な情報のみ取り出して、かつ PostgreSQL で読み取れる形式でリスト化する必要がある。このために、`csch`・`awk`・`sed` に SQL スクリプトを組み合わせたシェルスクリプトを作成した。作成したシェルスクリプトは `cron` により毎日自動で実行&ログを管理者へ送るよう設定を施した。

## 2.4 メールデータベース共有のためのウェブインターフェース

最後に、作成したデータベースの情報をウェブブラウザから閲覧できるよう、インターフェースの開発を行なった。HTML/CSS でウェブページを作成し、PHP を HTML に埋め込むことで PostgreSQL と連携させた。また、Python を用いた CGI スクリプトを用意して、最新の情報を参照してその場で集約情報をグラフ化するページの作成まで行った。最終的に、メールの到着順表示・スレッド表示・メールの基本情報を用いた検索・本文のキーワード検索・情報集約、に関するページを作成した。以下の図 3 ではその内の 2 つの例を示す。図 3 左のスレッド表示のページの一例を見ると、親 Consult ID が正しく設定され、スレッド表示と解決日の決定が適切に行われていることがわかる。

ID	親ID	受信日	件名	From	To	...	解決日	更新日	
17003	17003	2017-10-02 20:53:15	【作業依頼】開発系サーバへのCASA 5.1.0のインストール	伏せています	consult@ana.nao.ac.jp		2017-10-10 17:45:00	2017-12-13 15:52:22	詳細
17006	17003	2017-10-03 15:43:01	Re: 【作業依頼】開発系サーバへのCASA 5.1.0のインストール	consult <consult@ana.nao.ac.jp>	伏せています		2017-10-10 17:45:00	2017-12-13 15:52:22	詳細
17031	17003	2017-10-05 19:57:21	Re: 【作業依頼】開発系サーバへのCASA 5.1.0のインストール	伏せています	consult <consult@ana.nao.ac.jp>		2017-10-10 17:45:00	2017-12-13 15:52:22	詳細
17038	17003	2017-10-06 11:56:53	Re: 【作業依頼】開発系サーバへのCASA 5.1.0のインストール	consult <consult@ana.nao.ac.jp>	伏せています	...	2017-10-10 17:45:00	2017-12-13 15:52:22	詳細
17039	17003	2017-10-06 12:20:35	Re: 【作業依頼】開発系サーバへのCASA 5.1.0のインストール	伏せています	consult <consult@ana.nao.ac.jp>		2017-10-10 17:45:00	2017-12-13 15:52:22	詳細
17053	17003	2017-10-10 10:16:20	Re: 【作業依頼】開発系サーバへのCASA 5.1.0のインストール	consult <consult@ana.nao.ac.jp>	伏せています		2017-10-10 17:45:00	2017-12-13 15:52:22	詳細
17060	17003	2017-10-10 12:21:39	Re: 【作業依頼】開発系サーバへのCASA 5.1.0のインストール	伏せています	consult <consult@ana.nao.ac.jp>		2017-10-10 17:45:00	2017-12-13 15:52:22	詳細

### Filtering Mail Database

メールデータベースの検索を行う。\*指定なしの項目にチェックが入っている場合にはそちらが優先される。

Reset Submit

unique\_id  
指定なし / [ ] - [ ]

consult\_id  
指定なし / [ ] - [ ]

parent\_consult\_id  
指定なし / [ ] - [ ]

mail\_date  
指定なし / 2017 [ ] [ ] [ ] - 2017 [ ] [ ] [ ]

mail\_subject  
指定なし / [ ]

図 3. メールのスレッド表示ページから抜粋した一例（左）とメールの検索のページ（右）

## 3. ベータ版システムの運用結果

今回、Consult メールをデータベース化したことで得られた集約情報の一部が次頁の図 4 である。図 4 左は 2017 年 4 月から 10 月までに届いた Consult メールの内、何通が親メール（相談件数）であったかをヒストグラムで表している。緑が親メールの数を表しており、青+緑がメールの総数である。注意

点として、アカウント申請などの自動応答メールは1度に3通 Consult メールとして配送される（それぞれ配送先が異なっているが、運用グループには全て配送されている）。このため『メール総数=親メールとその返信』という関係性になっていない。注意すべき点はあるものの、この図を見ると4月が最も相談件数が多く、5月以降は減少し一定値へと漸近していくことがわかる。

次に図4右は Consult メールをカテゴリー分けした場合のヒストグラムになっている。カテゴリー分けの際には件名を利用して自動で分類し、分類できたメールのみ総数を表示している。赤は新規申請を表し、今回の結果からは分散が大きく明確なことは言えないが、1年を通して概ね一定であるように見える。緑はアカウント更新を表しており（正式な更新受付期間は2-3月なのでここでは期間外の更新申請を意味している）、こちらは4月をピークに以降減少している。青とシアンはそれぞれVPN申請とグループ化申請を表し、こちらは毎月一定数の申請があることがわかる。

これら傾向が正しいのかを確認するためにも、過去の Consult メールを取り込み、今後もメールを絶えず取り込んでいくことでより正確な傾向を掴むことができると考えている。

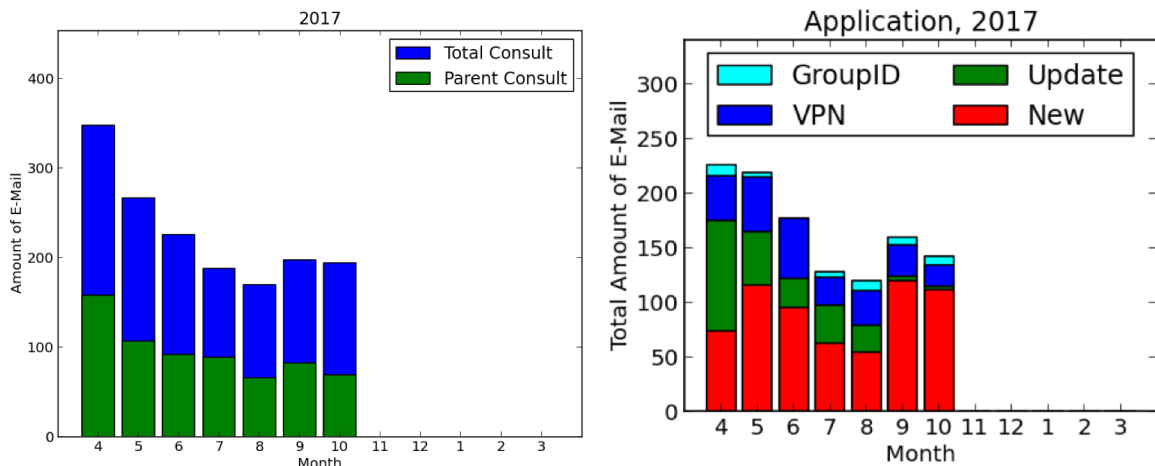


図4. メール総数のうち親メールの占める数（左）とカテゴリーごとの Consult メール総数（右）

#### 4. 今後の展望

現在運用中の解析システムは2018年3月1日から新システムへと移行される。この新多波長データ解析システムは、現行のシステムと比べて“総合的な計算機性能が4倍”、“ユーザーの利用可能ディスク容量が2倍（2018年3月1日時点の数値であり、最終的には3倍にまで増強される）”を達成する見込みである。今回開発した Consult メールデータベースは、新システムの稼働と同時に本格運用を予定している。また、新解析システムからはユーザー管理システムも刷新される予定である。新しいユーザー管理システムと連携することで、ユーザー単位で情報をまとめることができると考えている。すなわち、分野・役職ごとに Consult メールをまとめることで、例えば、大学院生やポストドクからの質問を集約し数値化できるはずである。これは今後の天文データセンター主催の講習会の内容に反映できるのではないかと期待しており、さらに User-Friendly な運用を進めていければと考えている。