

Magic Mirror Mailing: 個人情報データベースを利用する柔軟なメール配送システム

高畑 理 † 藤沼健太郎 ‡ 石橋 玲 § 遠山元道 †

† 慶應義塾大学大学院 理工学部 情報工学科 4年

‡ 株式会社イイガ § 有限会社さくらソフト

† 慶應義塾大学 理工学部 情報工学科/JST さきがけ研究 21

E-mail: † takahata@mmm-keio.net, ‡ fuji@iiga.co.jp § ishibashi@sakurasoft.co.jp, † toyama@mmm-keio.net

近年幅広く E メールが利用されるようになってきている。その1つの使われ方としてメーリングリストがある。同時に多数の人とコミュニケーションを取るためには有効であるが、現在リストごとに管理が行われており、統一されたフレームワークは存在しない。今回、慶應義塾大学理工学部同窓会のシステムを構築するにあたって、3万人の卒業生が所属することが想定するシステムを設計し、実装した。本稿では、そのシステムの仕組みについて述べる。

キーワード: E メール、コミュニケーション、非同期システム

Magic Mirror Mailing: A Flexible Mail Delivery System using a Personal Information Database

Osamu Takahata † Motomichi Toyama ‡ Kentarou Fujinuma § Rei
Ishibashi †

†

Faculty of Science and Technology, Keio University.

‡ Department of Information and Computer Science, Faculty of Science and Technology,
Keio University. PRESTO, JST

‡ Iiga Inc. ‡ Sakurasoft Inc.

E-mail: † takahata@db.ics.keio.ac.jp, ‡ fuji@iiga.co.jp § ishibashi@sakurasoft.co.jp, † toyama@ics.keio.ac.jp

E-mail has come to use broadly in recent years. There is a mailing list as the one usage. Although it is effective in order to take communication with many men simultaneously, management is performed for every present list and the unified framework does not exist. The system which that 30,000 graduates belong assumes was designed and mounted in building the system of the Keio University department-of-science-and-engineering class reunion this time. This paper describes the structure of the system.

keyword : electronic mail , group communication , interence system

1 はじめに

近年企業、大学問わず EMAIL の利用が必須となってきたり、仕事、プライベートに関わらず様々な利用シーンで使われている。そのメールを利用したコミュニケーションの1つとして、メーリングリストが昔から存在した。しかし、既存のメーリングリストでは配送リストごとの管理であり、全てのメーリングリストに対しての統一された管理プラットフォームは現在でも存在していない。規模が大きくなった場合メーリングリストを維持管理することが極めて困難となる。

今回、慶應義塾大学理工学部の同窓会の事業として、生涯に渡るコミュニケーションを行うためのインフラを作るにあたって、卒業生3万人に対してのメールサービスを提供することとした。そこで、卒業生名簿を元にしたデータベースの個人情報をもとに、メーリングリストをルールベースを元に自動的に生成し、メール転送するシステム Magic Mirror Mailing (MMM) を開発した。このシステムについて報告する。

なお本システムは5月よりサービスを開始している。現在、卒業生の1割以上に当たる3700人程度がサービスを利用している。

2 システムの概要

2.1 サービスの概要

本システムには大きく分けて2つのサービスが存在する。

2.1.1 Personal Mirror Service

Personal Mirror は一般のメール転送サービスと同様で、xxx@mmm-keio.net へのメールを、登録された個人のアドレス xxx@hoge.net に転送する機能を持つ。

2.1.2 Group Mirror Service

Group Mirror は理工学部同窓会の名簿データを元に、上記の転送先メールアドレスを結合し、学科別、学年別、研究室別などのメーリングリストを自動的にメンテし、配送を行う機能を持つ。

2.2 データモデルの概要

名簿データをマスターとしてRDBに格納し、メール転送に必要なデータをLDAPサーバに置く。RDBからLDAPへのデータのロードはRDBのデータを元にルールベースに基づいて行われる。

メールルーティングはLDAPにある情報を元にアドレス解決が行われる。

エンドユーザはアドレスの登録・変更はWeb経由で行うことができる。

同窓生にクレジットカード状のカードでパスワードを配布し、そこにかかっている会員コードとアクセスキーを元に個人の認証を行う。

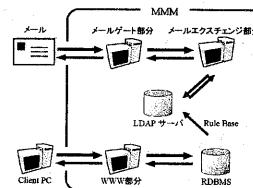


図1: 本システムの概要

2.3 メールルーティングの概要

メールルーティングは3つのパートに別れている。その3つとは、メールゲートウェイ部分、メールエクスチェンジ部分、ディレクトリーサーバ部分である。

ルーティングに関するシーケンス図を図2に示す。

- ・メールを受け取ったメールゲートウェイ部分はメールエクスチェンジに転送をする。
- ・メールエクスチェンジ部分は、LDAPに問い合わせをかけて、アドレスの変換を行う
- ・アドレスの変換を行ったメールをメールゲート

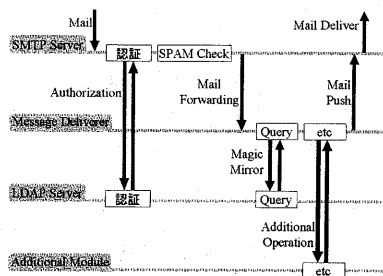


図 2: シーケンス図

ウェイ部分に戻す

- ・メールゲートウェイ部分がインターネットに対して配送する

2.4 システムソリューション

本システムは以下の組み合わせで実装を行っている。

- ・ RedHat6.2
- ・ Postgres7.1
- ・ OpenLDAP
- ・ Apache + mod_ssl
- ・ Tomcat
- ・ JDK1.3
- ・ qmail

3 データサーバ

本システムではRDBMS(Postgres)とLDAP(OpenLDAP)を採用した。その特徴をここで示す。

3.1 ディレクトリーサーバ

ディレクトリーサーバは転送のためのアドレス情報を保持する。今回利用した理由としては、3点ある。第1に検索のみであればRDBMSよりも非常に高速である。Join操作のような集合演算はできないが、サブツリーの抽出操作はとて高速であり、MMMのような「大量のデータの高速検索」という用途に向く。

第2にエンティティ(テーブルにおけるレコードに相当)の構造を柔軟に決定できる。エンティティごとに備えるべき属性を変更したりすることができる。エンティティをオブジェクトに例えると、様々なタイプ(型)を付与することができるということ。またRDBMSでは正規化違反になるような繰り返しデータも許される。

第3にMMMでメーリングリスト以外のサービス(グループウェア、掲示板、本人認証等)を提供するときに、LDAPを使うと認証・プロフィール情報を格納するときに便利である。

以上のような理由によってディレクトリーサーバを利用した。

3.2 RDBとLDAPの使い分け

本システムではRDBとLDAPの2つのデータサーバを使用した。用途に応じて2つのデータベースを使い分けた。

RDBMS サーバー: マスター情報を保存

LDAP サーバー: 転送先情報の保存

これには、それぞれの特性を生かし、アクセスパターンの違いに応じて使い分けをすることによって高い性能を引き出している。トランザクションをサポートしているRDBMSでは追加・更新作業を得意としており、マスターデータの保持には非常に優れている。

一方、LDAPはトランザクションをサポートしていないため、追加・更新作業を得意としていないが、単純な検索においては非常に高速である。MMMではLDAPをRDBMSのスナップショットデータという位置付けで検索専用のものとして利用している。

4 データモデル

MMMにおけるデータモデルを説明する。MMMでは、いくつかのエンティティがある。

4.1 People

People は個人に関する情報が入り、主に Personal Mirror で使われる転送先の情報、MMM でのアカウント名、所属した研究室の情報などがこのエントリーにある。

4.2 List

List にはメーリングリストに関するエントリーが入る。このエントリーの子供には、People ツリー配下にある個人のエントリーへの alias があり、最終的に検索した結果、転送先アドレスの集合としてクエリーがかけられる状態となっている。

4.3 Lab

Lab には研究室に関するエントリーが入る。研究室の所属には複雑な問題があり、複合的な制約条件がかかる場合がある。例えば、

- 1998 年度においては、x 研究室は y 学科に属していた

- 1998 年度においては、z 学科には a, b, c の研究室が属している

- 2000 年度に x 研究室は z 学科に所属が変更された

その結果、z 学科以下の研究室は a, b, c, x になった

- 2000 年に、1998 年度の z 学科全員というメーリングリストを作成した

- a, b, c 研究室は上記メーリングリストに含まれるが、x 研究室は上記メーリングリストには含まれない

という風な結果を保証する。

このようなケースでは個別のメンテナンスが必要である。その部分は RDBMS → LDAP のためのルールベースにて 1 つずつ表現することにした。

実際の操作としては、エンドユーザは Web を通じて自分と関係のある研究室を選ぶが、その情報は一度 RDBMS の LabHistory に入る。その後、ルールベースを通じて、Lab ツリー以下のエントリーとして展開される形でエントリーは作られる。

5 メールルーティング

この章ではメールルーティング部分のそれぞれの詳細について説明する。

5.1 メールゲートウェイ

この部分は SMTP に関して外部のインターネットと接点となる。

この部分は、世間一般で利用されている sendmail、qmail などを利用できるが、MMM では、sendmail を採用した。

これにより、SMTP に関しての DoS 攻撃や SPAM メールなどのセキュリティ対策をある程度のレベルで保持することが可能である。

メールを受け取った時点で、メールに対する認証チェックを行い、宛先が正しいことが確認されたメールをメールエクスチェンジ部分に転送する。また後述のメールエクスチェンジ部分にてアドレスが展開され、再度転送されてきたメールを外部のインターネットに対して配送する。

5.2 メールエクスチェンジ

メールエクスチェンジ部分は、受け取ったメールアドレスをディレクトリーサーバに問合せをかけ、メールアドレスの変換を行う。そして、それをメールゲートウェイに戻し、最終的にユーザにメールを届ける。

メールを受け取った時点で、サブドメイン名より、配送ルールを決定する。サブドメイン名には、

なし Personal Mirror 用

mech,ics など 学科

lab 研究室

all 全員

などがある。例えば、toyama@lab.mmm-keio.net というメールが送られてきた場合は、研究室宛のメールということを判断する。

次に SMTP Envelope から Receipt を抽出する。

Personal Mirror では

sample@mmm-keio.net → sample

といった具合である。その Receipt に対して LDAP サーバに問い合わせを行う。

```
cn=sample, ou=People, dn=mmm-keio, dn=net
```

のようなクエリーに対し、以下のような応答が帰ってくる。

```
dn: cn=sample, ou=People, dc=mmm-keio, dc=net
mail: sample@keio.ac.jp
```

この mail アドレスを実際の配送先として使い、メールの Receptient のヘッダーを書きかえ、その後メールゲートウェイ部分に送り返す。

6 ルールベース

RDB から LDAP へのデータ移行の処理を行うためのルールベースについて述べる。現在は、RDB と LDAP のそれぞれのマッピング情報を一般的に表現して展開している訳でなく、エンティティごとにそれぞれ独自にマッピング情報を書いている。

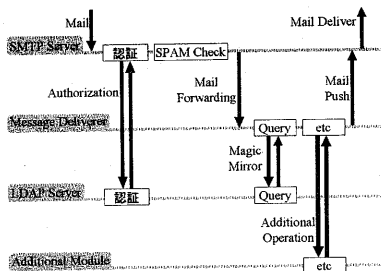


図 3: ルールベース クラス図

ルールベースが起動するタイミングは現在 2 つある。1 つが Web の Interface より情報が登録・変更された場合であり、もう 1 つがバッチ処理として RDBMS より LDAP のエントリーを再生成する場合である。前者の場合、登録・変更されたデータが RDBMS にはいり、同時に LDAP にも書き出される。これにより、ほぼリアルタイムに RDBMS と

LDAP の同期を取ることが可能である。後者の場合、RDBMS のデータを 1 件ずつ読み、各エンティティに付属するクラスによって、LDAP にデータを移行していく。これによって、何らかの理由で RDBMS と LDAP の間でデータの矛盾が生じたとしても RDBMS のデータを正として、再度 LDAP のデータを作り直すことによってデータの整合性を保つことが出来る。

7 評価・検討

7.1 性能評価

LDAP に対する問い合わせは、4 万件のテーブルへのクエリーが 4ms で完了した。同様に RDBMS で join が必要な問い合わせにも、10ms で応答することが分かった。これにより、理論上 1 台で毎秒 100 通、1 日で 864 万通のメールを転送可能が可能であり、非常に高速な配送エンジンであることが証明された。

7.2 実サイトでの性能評価

MMM はサイト開始直後に 1000 人前後の登録があった。その後、3000 人の登録を越した時点で、登録されたメンバーに対して一斉配送を行ったところ、配送に 2 時間程度かかった。これは上記の性能とは大きな差があるが、それは以下のボトルネックとなる点が考えられるためである。

- ・DNS による MX レコードの検索が時間がかかっている
- ・sendmail の配送手順が順次配送のため

特に後者の sendmail によるボトルネックは非常に大きいことが分かった。そのため、サイトではメールゲートウェイ部分に qmail もしくは postfix を利用する方向に変更する予定である。

7.3 今後の目標

現在ルールベース部分に関してアドホックな対応を取っているため、同様のシステムを構築する際も毎回実装のレベルで作直さないといけない。このルールベースの部分に関して非手続的な処理として表現をできるようにし、他のシステムへの展開を

容易にしたい。

参考文献

- [1] H. Lachman, Internet-Draft LDAP Schema for Intranet Mail Routing
- [2] Timothy A. Howes, Ph.D., Mark C. Smith, and Gordon S. Good, UNDERSTANDING AND DEVELOPING LDAP DIRECTORY SERVICES
- [3] Tim Howes, Mark Smith, Programming Directory-Enabled Applications With Lightweight Directory Access Protocol (MacMillan Technology Series)
- [4] Sendmail with LDAP
<http://www.stanford.edu/~bbense/Inst.html>
- [5] OpenLDAP
<http://www.openldap.org>
- [6] W3C World Wide Web Consortium
<http://www.w3.org>
- [7] 萩野浩明, 門林理恵子, 塚本昌彦, 西尾章治郎: 推論システムを利用したメッセージ管理システムにおける分類・検索機構, 1996年度人工知能学会全国大会(第10回) 論文集, pp. 557-560 (June 1996).