Vol.2012-EVA-39 No.5
Vol.2012-IOT-19 No.5
2012/9/28

Gmail が大学メールサーバへ与える負荷状況の分析

笠原義晃[†]　伊東栄典[†]　堀良彰[††]　藤村直美[†††]

九州大学では，従来から大学ドメインのメールサーバを学内に構築し，構成員へメールサービスを提供してきた．2012 年 1 月頃より，学内の情報サービスに対し利用者認証機能等を提供する全学認証サーバの負荷の高さが問題となり，その原因の一つが学生向けメールサーバであることが明らかになった．詳細な分析の結果，Google 社の Gmail から本学のメールサーバへ持続的なアクセスがあり，中でも既に卒業などで消滅したアカウントへのアクセスが多数あることが分かった．本稿では，本学の学生メールサーバのアクセスログ解析に基づいて，Gmail がメールサーバへ与える負荷状況の分析とその理由について述べ，対応策について検討する．

Google makes a chronic big load to university mail server

YOSHIAKI KASAHARA[†]　EISUKE ITO[†]
YOSHIAKI HORI[††]　NAOMI FUJIMURA[†††]

Traditionally, Kyushu University has been providing email service using its own domain name for staff members and students of the university. Around January 2012, we noticed that the high load of the university authentication server, and we realized that one of causes was the access from the mail server for students (called Student Primary Mail Service). Detailed analysis showed that there was chronic big load produced by Google's Gmail, especially toward nonexistent accounts removed due to graduation. In this paper, we explain the current situation and reasons of the big load induced by Gmail and its possible countermeasures based on the analysis of access logs for Student Primary Mail Service.

1. はじめに

情報通信サービスは，大学における教育研究活動でも欠かせないものとなっている．様々なサービスの中でも，電子メールはインターネットが普及する前から現在まで基本的かつ重要なサービスとして使われている．大学が提供する電子メール環境は設備の視点から見ると二つに分類でき，一つは大学ドメインのメールサーバを独立メールサーバとして設置する方法，もう一つは SaaS（Software as a Service）型のメールサービスである．SaaS 型のメールサービスは 2007 年頃から普及し始め，Google 社の Gmail や Yahoo!メール，Windows Live@edu などがある．

大学以外が提供するメールサービスの利用は多く，現在大学構成員のほぼ全員が，大学ドメインではないメールアドレスを保有していると推測される．その代表は携帯電話に付随するメールアドレスである．次に無料で提供されているメールサービスであり，Google 社の Gmail, Yahoo!メール，Microsoft 社の Outlook（旧 Hotmail）などがある．他にも ISP が提供するメールアドレスの利用者もいる．

筆者らが所属する九州大学では，大学ドメインのメールサーバを学内に構築し，構成員へ提供する方法を継続してきた．全学生向けの電子メールサービスを 1995 年から現在まで継続して提供しており[1][2]，2009 年 7 月からは全職員向けに基本的なメール環境を提供するサービスも開始している[3][4].

我々は本学の全学認証サーバの負荷を分析してきた[5]．認証サーバの負荷としては，無線 LAN の接続時と，電子メール利用時の認証処理が膨大で，その二つのサービスで全体の 95%以上の認証処理を占めていることが分かった．学生メールサーバから来る認証処理について詳細に分析した結果，Google 社の Gmail から本学のメールサーバへアクセスする処理数が膨大であることが分かった．実際，本学の学生メールサーバに来る POP アクセスのうち，55%が Gmail からのものである．しかも，Gmail から来るアクセスのうち，65%が卒業などでサーバ側に存在しないアカウントへのアクセスである．

本稿では，Gmail が九州大学の学生メールサーバへ与える負荷状況の分析について述べる．本稿の構成は以下のとおりである．第 2 節で本学のメール環境を説明する．第 3 節で認証サーバおよび学生メールサーバのログから分析した，メールサーバの負荷状況を述べ，Gmail が大学の学生メールへ負荷を与える理由を述べる．第 4 節で，Gmail が全世界に大学メールサーバに与える影響についての定性的分析と，Gmail の影響に対する対応策を検討する．最後に第 5 節でまとめと今後の課題を述べる．

2. 九州大学の全学メールサーバ

まず初めに，我々が所属する九州大学情報統括本部が提供しているメールサーバ[1][2][3][4]について説明する．

2.1 学内構成員数

大学の主たる構成員は学生と教職員である．学生は，学

† 九州大学　情報基盤研究開発センター
Research Institute for Information Technology, Kyushu University
†† 九州大学　システム情報科学研究院
Department of ISEE, Kyushu University
††† 九州大学　芸術工学研究院
Department of Design, Kyushu University

Vol.2012-EVA-39 No.5
Vol.2012-IOT-19 No.5
2012/9/28

部学生と大学院生からなる．正規の学生（正課生）以外に，研究生や科目等履修生などの非正課生も在籍している．学生と教職員以外にも，特別研究員や，外部組織で雇用されている派遣社員，訪問研究者など，様々な身分の職員が在職している．九州大学における 2012 年 3 月現在の発行 ID 概数を表 1 に示す．表 1 に示す ID の数が，メールサーバの利用者数である．

表 1　九州大学の ID 数（2012 年 3 月現在）
Table 1 The Number of IDs in Kyushu University (Mar 2012)

種類	ID 総数（概数）
正課生（学部生と大学院生）	19,000
非正課生	500
職員	9,000
派遣等	800
合計	29,300

構成員の年間の入れ替わり数は，学生が約 6,000 人である．内訳はおよそ学部学生が 3,000 人，大学院生が 3,000 人である．そのため，学生用の学生基本メールでは，毎年 6,000 個のアカウントが作成・削除されている．

一方，職員の入れ替わり数は毎年約 1,000 人である．入れ替わりは 3 月と 4 月に集中しており，約 1,000 人が 3 月に移動または退職し，4 月に 1,000 人が追加される．そのため，職員用のメールシステムでは，1,000 個のアカウントが作成・削除されている．

2.1 職員用メールサーバの構成

職員用メールサーバのシステム構成を図 1 に示す．

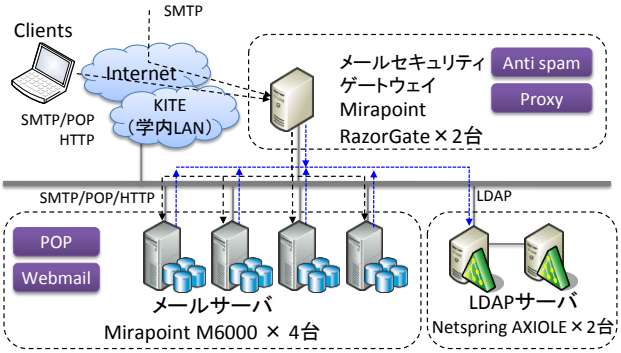

図 1 職員用メールシステムの構成
Figure 1 The outline of the Primary Mail System in Kyushu University

職員用メールサービスは，大学の活動に公的に従事する全員へ提供している．メールサービスを提供する中心となる機器は Mirapoint 社の製品で構成した．認証用の LDAP サーバは，当初は ID 統合管理システムが提供する LDAP サーバを直接参照していたが，後述する有料サービスクラスの導入に際して Mirapoint 専用のスキーマを利用する必要があったため，別途専用の LDAP サーバを導入し，ID 統合管理システムから利用者データを投入してもらう形に構成を変更している．

サービス概要は以下のとおりである．

- 1 人あたりの容量は 300MB（開始当初は 100MB）
- 受信から 60 日で削除（開始当初は 30 日）
- POP3 と Web メールを提供 (SSL を使用)
- 職員でなくなったら三ヶ月後にアカウント削除
- 受信可能なメールの大きさは 20MB 未満

また，保存容量を拡大（10GB）し，保存期間を制限しないサービスクラスを有料で提供している．このサービスでは上記に加えて IMAP を利用できる．

2.2 学生基本メールの構成

九州大学では，従来から教育情報システムの一部として，学生のメールサービスを提供してきた．現在の学生基本メールは，2011 年 6 月に開始したサービスであり，学内の全学生に対して提供している．図 2 に全体のシステム構成を示す．

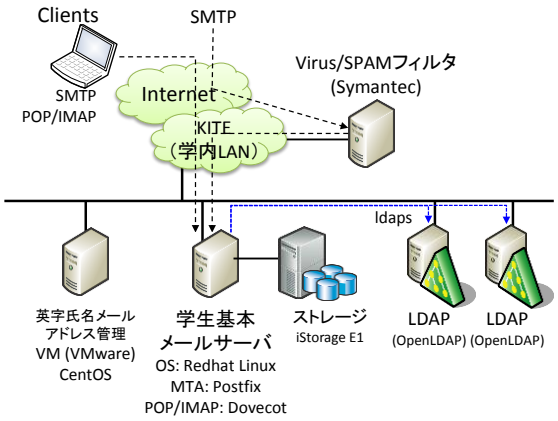

図 2 学生基本メールシステムの構成
Figure 2 The outline of the Student Primary Mail System

メール送信時には，迷惑メール送信を防ぐため，SMTP-AUTH によるユーザ認証を行なっている．メール受信には POP3 および IMAP4 を受け付けており，メーラで着信メールを受け取る際に利用者認証が行われる．これらの認証は，ID 統合管理システムが提供する全学共通の LDAP サーバで行われる．利用者が直接サーバにログインすることはない．

学生基本メールは，汎用のサーバ機に Red Hat Enterprise Linux と Postfix，Dovecot を導入した OSS (Open Source Software) システムとして構成した．職員用と異なり，学生向けには高い可用性よりもメールスプール容量と使い勝手の良さを優先すべきであると判断したためである．また人数の多い学生に大きめのスプール容量を割り当てるため，ストレージに費用をかけている．

Vol.2012-EVA-39 No.5
Vol.2012-IOT-19 No.5
2012/9/28

3. Gmail が大学メールサーバへ与える負荷状況

この節では，Gmail が本学のメールサーバに負荷を与えていることが発見された経緯と，その状況について述べる．

3.1 経緯

2012 年 1 月頃より，ID 統合管理システムが提供する LDAP サーバを利用する各種情報サービスにおいて，認証処理の不具合が発生してサービスが利用できなくなるという問題が認識され始めた．調査の結果，LDAP サーバに認証要求が集中することにより過負荷になり，認証処理が失敗していることがわかった．

この LDAP サーバを最も利用しているサービスは，802.1x 認証を実施している学内無線 LAN サービスと，学生基本メールの 2 つであった（職員向けサービスは既に専用の LDAP サーバに移行済）．特に，学生基本メールは LDAP の bind 処理を用いてユーザ認証していたため，LDAP サーバに与える負荷が高いと予想された．メールサーバのログや設定を詳しく調査したところ，Google の所有するアドレスブロックからの POP による認証要求が非常に多いことが判明した．なお LDAP サーバについても，導入当初メール等の高負荷なサービスでの利用は想定されていなかったためチューニング設定が不十分であることがわかったが，この件については本稿では扱わないこととする．

3.2 アクセス状況分析

学生基本メールの POP/IMAP での利用状況を確認するため，メールソフトウェアである Dovecot と，認証バックエンドとして利用している PAM システムのログを 2012 年 6 月の一ヶ月分について調査し，ログイン試行回数とその対象となったアカウント数，接続元ネットワークについて分析した．また 7 月下旬時点での有効アカウント総数と，メール転送設定数を調査した．調査結果を表 2 に示す．

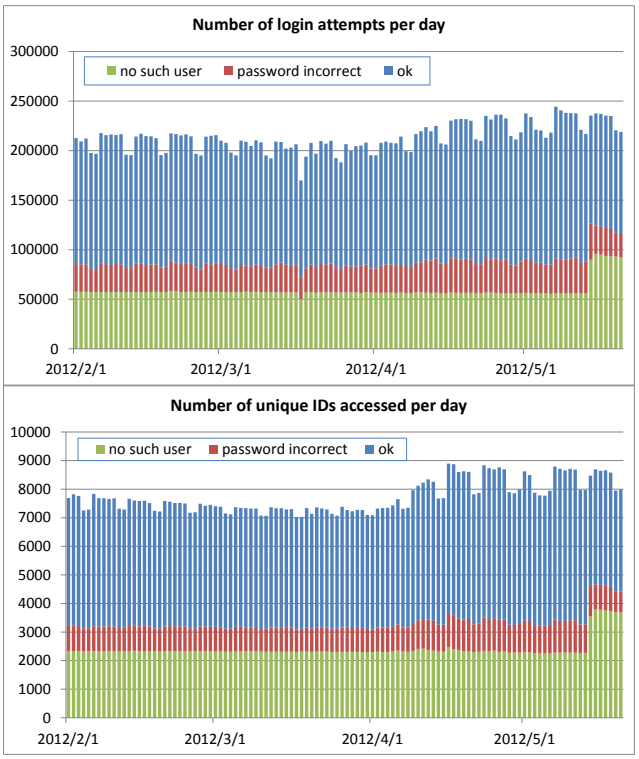

図 3 日毎の POP/IMAP アクセス数と対象アカウント数
Figure 3 The number of POP/IMAP accesses and unique IDs per day

表 2 から，有効アカウント数の約 3 分の 1 の利用者が POP や IMAP でメールサーバにアクセスしており，うち 4 割の利用者が Google の所有するネットワークから，すなわち Gmail を利用しているであろうことがわかった．また卒業するなどして無効になったアカウントへの POP/IMAP アクセスはほぼ 9 割が Google からのアクセスであった．アクセス数で見ると，無効アカウントへの POP/IMAP 接続が全体に占める割合は約 40%あり，そのうち 95%は Google からのアクセスであった．もし Google から無効アカウントへの

表 2 学生基本メールのアクセス状況
Table 2 Access statistics of Student Primary Mail

	POP/IMAP 平均アクセス／日		アカウント数	
総数（アカウントは有効なもの）	234,843	(100.0%)	19,996	(100.0%)
別アドレスへ転送	-	-	3,513	(17.5%)
正常利用数	113,491	(48.3%)	6,435	(32.2%)
Google から	33,835	(14.4%)	2,732	(13.7%)
学内から	19,081	(8.1%)	3,432	(17.2%)
パスワード違いによる認証エラー	29,824	(12.7%)	1,761	(8.8%)
Google から	12,422	(5.3%)	542	(2.7%)
学内から	6,483	(2.8%)	689	(3.4%)
無効アカウントへの接続試行	90,773	(38.7%)	3,989	-
Google から	86,495	(36.8%)	3,567	-
学内から	1,266	(0.5%)	190	-
その他のエラー（ID なし等）	755	(0.3%)	-	-

アクセスが消滅すれば，それだけで POP/IMAP のログイン認証処理が約 4 割減ることになる．

次に, 2011 年 2 月から 2012 年 5 月にかけての POP/IMAP でのログイン試行数とその対象アカウント数を日別に時系列で表したグラフを図 3 に示す．このグラフから，無効アカウントへのアクセス数とその対象アカウント数にはほとんど変化がないことと，5 月中旬に卒業生等が確定しアカウントが削除された時点から，アクセスされる無効アカウント数が 1,200 程度増加していることがわかる．これらのことから，卒業生の多くは卒業後アカウントが消滅しても特に設定を変更せず, また Google 側も認証エラーを放置していることが読み取れる．

3.3 Google の Mail Fetcher サービスについて

Google からの POP アクセスは，Gmail の提供する Mail Fetcher サービスによるものであると考えられる．これは，他のメールサーバに蓄積されているメールを POP で取得し Gmail のメールボックスに格納する機能である．これにより，複数のメールアカウントに分散したメールを Gmail で一括して購読・管理できる．

九州大学の学生基本メールに対して Mail Fetcher でのアクセスが多い主な理由としては，2009 年度の教育システム更新の際にそれまで提供されていたウェブメールが廃止されたことから，メールへの簡便なアクセス手段として他のメールサービスへの転送方法を案内していたこと，特に Gmail については転送と Mail Fetcher の両方の方法を案内していたこと，また一部の学部において情報教育の一環として Mail Fetcher 利用の実習を行なっていたことが挙げられる．

Gmail に他のメールサービスのメールを集約する方法としては，元のサーバから転送機能で Gmail のアドレスに再配送するという手段もある．しかし，一般には転送よりも Mail Fetcher の利用が推奨されている．その主な理由は，一旦あるサーバで受信したメールを他のサーバに転送すると，迷惑メール対策が誤作動するという問題があるためである．ある利用者が単純にすべての着信メールを転送する設定をすると，迷惑メールも転送されることになる．多くの利用者が同様のことを行っていると, (判定方法にも依存するが) 転送先のサーバで転送元が迷惑メールの発信元と認識され，ブラックリストに登録されたり，ドメインの評価が下がったりするという問題が発生する場合がある．学生向けに一斉同報メールを送信した場合に，転送先のサーバから見ると多数の利用者に同じ内容のメールを送信しているように見えるため，迷惑メールを送信するサーバとして判定される危険性もある．ブラックリストやドメインの評価はサーバ間で共有されている事が多く，あるサーバで評価が下がると他の組織にもメールが届かなくなる可能性がある．また，単純なメール転送と SPF（Sender Policy Framework）による発信者認証は相性が悪いという問題も知られている [6].

転送ユーザが少ない場合はあまり問題が顕在化することはないが，Gmail のように利用者が多いと問題が発生しやすくなる．また以前の学生向けメールサービスには迷惑メールフィルタがなかったため，迷惑メールが転送先に流入しやすく，誤動作の可能性が高いという事情もあった．

しかし，多数の学生が Mail Fetcher を設定し，そのままその対象アカウントが消滅した場合にどうなるかということについては，少なくとも本学においてこれまで問題は認識されておらず，全く検討もされていなかった．

3.4 Gmail の POP アクセスの状況

Gmail の Mail Fetcher が対象のメールサーバにアクセスする頻度の詳細は公開されておらず，Gmail のヘルプには「前回試みたアカウントごとのメールの取得結果に応じた頻度で，新着メールをチェック」すると記述されているのみである．そこで，2012 年 6 月 1 日～7 日の一週間分のアクセス記録について，Google から正常にアクセスされているアカウントと存在しないアカウント全てについて，一日あたり平均アクセス回数を集計し，アカウント数に対する累積密度関数のグラフを作成した．

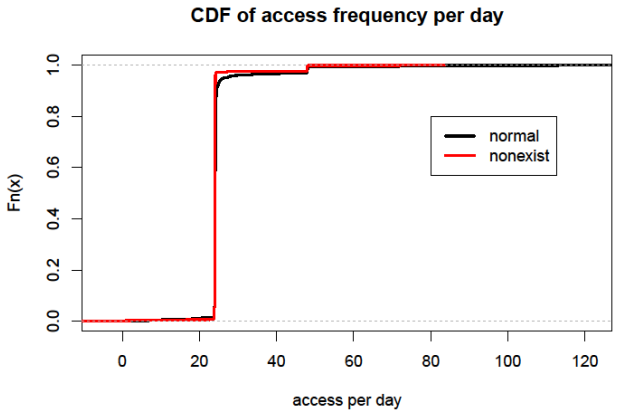

図 4 Google からの 1 日平均アクセス頻度
Figure 4 CDF of access frequency per day from Google

図 4 から，アクセス頻度の傾向は Gmail がアカウントにアクセスできているかいないかにかかわらずほぼ同じであり，9 割以上のアカウントに対して 1 時間に 1 回（1 日 24 回）のアクセスであること，48 回・72 回等 24 の倍数に回数が集中していることから 1 時間あたりの回数で頻度を調整しているらしいことがわかった．アカウントが存在しないにもかかわらずアクセス頻度が高いアカウントがある理由は，対象アカウントが存在していた時のメール到着履歴をそのまま利用し続けているためと考えられる．

4. Gmail の影響と対策

Gmail が大学メールサーバに与える影響についての定性的分析と，Gmail の影響に対する対応策を検討する．

4.1 影響

2.1 節と 3.2 節より，九州大学においては，毎年約 6,000 アカウントが入れ替わっていることと，その約 1/5 である 1,200 アカウントが消滅後も Gmail からアクセスされ続けていることがわかった. 現在の学生基本メールは 2011 年度に運用を開始したものだが，サーバの FQDN 等は入れ替え前のシステムを踏襲しており，2009 年度から年度が変わるごとに同程度ずつ増加していると考えることができる.

文部科学省の学校基本調査（平成 22 年度）によると，日本全国の国立・公立・私立を合わせた大学学生数は 288 万 7 千人である[7]. これら学生全てに大学のメールサービスが提供されていると仮定し，毎年九州大学と同程度，すなわち 1/3 が入れ替わるとすると，96 万アカウントが入れ替わることになる. 九州大学においては，このうちの 2 割が Gmail の Mail Fetcher を利用していたが, 少なく見積もって 1 割としても毎年約 10 万の無効アカウントへのアクセスが増加することになる. これに対し 2012 年現在 Gmail ユーザは約 3.5 億人であり[8], その負荷を処理する Gmail のサーバ群にとって毎年 10 万の無効アカウントに対する POP アクセスは無視できるようなものであろう.

しかし，この状況を放置すると日本のみならず世界中のメールサービスに無駄な負荷を与え続けることになる. 大学などの教育機関は企業よりアカウントの入れ替わりが激しいと考えられるため，消滅アカウントへのアクセスが残存し続けると悪影響が大きい. 在学生向けサービスは負荷の増減があまりないことを想定して設計する場合も多く，毎年一定の割合でアクセスが増加すると本来の機能に支障をきたす可能性が高い.

4.2 対策

Gmail からのアクセス増加に対する対応策としては，以下のような方法が考えられる.

1) Google からの POP アクセスを遮断する
2) Mail Fetcher の利用者に設定変更を呼びかける
3) サーバ側で処理を変更し，負荷を軽減する
4) Google が仕様を変更する

4.2.1 Google からの POP アクセスを遮断する

これは Gmail からの負荷に対しては最も即時的かつ効果的な対策であろう. 同時に利用者の利便性を大きく損なう対策でもあるため，最後の手段と考えられる. 負荷の高さに対し他に打つ手が無い場合にはやむを得ない対応であり，実際に北海道教育大学においてこの対策を実施したという事例があった[9].

4.2.2 Mail Fetcher の利用者に設定変更を呼びかける

これは, Mail Fetcher の設定が負荷の原因であることから，それを取り除くという対策である. 利用者が大学に在籍しており，Mail Fetcher の設定を間違っている場合には，アクセスされているアカウントに対し連絡することで一定の効果が期待できる. しかし，削減したい負荷のほとんどは卒業等で消滅したアカウントへのアクセスが原因であり，その利用者への連絡は困難である. 卒業しているため学内の連絡網は利用できない. また Mail Fetcher でアクセスされる側からは，Mail Fetcher を有効にしている Google アカウントの情報を取得できないため，Gmail 側に通知を送ることも困難である.

一時的に無効アカウントを復活させ, Mail Fetcher 設定を外す方法を指示するメッセージをスプールに入れておき，これを取得させる，という手段も考えられる. この場合 Mail Fetcher に設定されているパスワードを抽出して POP アクセスできるようにする等の手当が必要であり，単純ではない. また設定はしたがその後 Gmail を利用していないような利用者には効果がない.

4.2.3 サーバ側で処理を変更し負荷を軽減する

これはサーバ側で認証処理等の負荷を軽くすることで問題を回避する対策である. 具体的には，例えば Dovecot であれば認証情報をキャッシュする機能があるため，これを有効にする，といった方法が考えられる. しかしキャッシュにはパスワード変更時等にキャッシュの不整合による認証エラーの問題があるため, 慎重に検討する必要がある. 負荷分散装置で Google からのアクセスのみを別のサーバにルーティングするといった方法も考えられる. しかし，Google 側で何らかの対処がなければ無効アカウントへのアクセスは毎年増加することになるため，これらの対策は一時的には効果があっても本質的な解決にならない. 削除したアカウントへのアクセスに耐えるためにコストをかけ続けるのは難しい.

4.2.4 Google が仕様を変更する

問題の本質は, Mail Fetcher が無効アカウントへのアクセスを永続的に試行し続ける点であり, Mail Fetcher のソフトウェアを改変することで解決可能である. 具体的には，連続的なエラーへの一般的な対策として, バックオフ処理（連続するエラーでは試行間隔を延ばす）とタイムアウト処理（一定期間試行してエラーが解消しなければ処理を停止する）を実装すればよい. タイムアウトにより処理を停止した際には, Mail Fetcher を利用している Gmail アカウントに対しメールとして Mail Fetcher 設定を無効にした旨を通知すればよい. Gmail でメールを読んでいれば通知に気づいて対処するだろうし，もともと Gmail を利用していなければそのまま無効でも問題がない.

この対策の問題は, このような仕様変更を Google に要望する窓口がはっきりしていない点である. Google への機能要望は Google Product Forums に投稿することになっている. 記事を検索したところ 2010 年 8 月に同様の要望が書き込まれているが，フォローは特になく，問題は改善されないままとなっている.

Vol.2012-EVA-39 No.5
Vol.2012-IOT-19 No.5
2012/9/28

5. おわりに

本稿では，Gmail が九州大学の学生メールサーバへ与える負荷状況の分析結果とその対策について述べた．認証サーバおよび学生メールサーバのログを分析し，メールサーバの負荷状況を述べた．その結果，Gmail が大学の学生メールへ高い負荷を与えていることが分かった．Gmail が大学メールサーバに与える影響についての定性的分析と，Gmail の影響に対する対応策を検討した．

九州大学では，現在までのところ 4.2.3 節に示したサーバ側での負荷軽減対応を可能な範囲で実施し，様子を見ている状況である．しかし，サーバ側での対応には限界があることから，何らかの方法で Google に働きかけ，Mail Fetcher の仕様を変更してもらう必要がある．

参考文献

1) 藤村直美，戸川忠嗣，笠原義晃，伊東栄典：姓名をベースにしたアドレスによる学生基本メールの運用について，情処研報 2011-IOT-14(10)，pp.1-6（2011）．
2) Naomi Fujimura, Tadatsugu Togawa, Yoshiaki Kasahara, Eisuke Ito: Introduction and Experience with the Primary Mail Service based on their Names for Students, Proc. of ACM SIGUCCS2012, (to appear)(2012).
3) 伊東栄典，笠原義晃，藤村直美：九州大学における職員向け電子メールサービスの現状，平成 21 年度情報教育研究集会 D3-4（2009）．
4) 伊東栄典，笠原義晃，藤村直美：九州大学全学基本メールの機能改善と有料サービスクラスの開始，情報教育研究集会 2011，B2-4（2010）．
5) 伊東栄典，笠原義晃，藤村直美：全学認証サーバの負荷状況と負荷分散，情処研報 Vol.2012 CSEC-57/IOT-17 No.11，pp.51-56（2012）．
6) 迷惑メール対策委員会：SPF と転送の相性問題に対する解決案，（財）インターネット協会（オンライン），入手先〈http://salt.iajapan.org/wpmu/anti_spam/admin/operation/suggestion/spf-sugg_a02/〉（参照 2012-07-30）．
7) 文部科学省：学校基本調査-平成 22 年度（確定値）結果の概要，文部科学省（オンライン），入手先〈http://www.mext.go.jp/b_menu/toukei/chousa01/kihon/kekka/k_detail/__icsFiles/afieldfile/2010/12/21/1300352_2.pdf〉（参照 2012-07-30）．
8) Google：2012 Update from the CEO - Investor Relations，Google（オンライン），入手先〈http://investor.google.com/corporate/2012/ceo-letter.html〉（参照 2012-07-30）．
9) 北海道教育大学函館校情報システム，入手先〈http://system.hak.hokkyodai.ac.jp/〉（参照 2012-07-30）．