

第 XXIII 部

IRC の運用状況とデータ解析

第 23 部

IRC の運用状況とデータ解析

第 1 章 はじめに

IRC ワーキンググループは Internet Relay Chat (以下 IRC) に関する研究と、そのテストベッドとしての IRC サービス網の運用を通じて、大規模サービスの安定運用に関する研究を行っている。また、WIDE プロジェクト内のほかのワーキンググループと連携した研究活動も行っている。

IRC とは、人と人とのコミュニケーション手段の 1 つである文字ベースによるリアルタイムチャットを、インターネットのサービスとして提供するシステムである。

IRC のシステムは、メッセージ転送を受け持つサーバと、ユーザとのインタフェイスであるクライアントにより構成される。IRC のサービスは、これらをサーバを中心にスター型に接続した IRC ネットワーク上でメッセージを順次パケットリレーのごとく転送することにより提供される。この IRC ネットワークは広域分散技術に基づいて設計され、世界規模で運用されている。

IRC ワーキンググループでは複数の IRC サーバを運用・接続し、研究に利用するとともに、国内の一般ユーザにも開放している。現在 IRC ワーキンググループが WIDE ネットワーク上で運用するサーバは以下の 4 台である。

- irc.tokyo.wide.ad.jp
- irc.fujisawa.wide.ad.jp
- irc.nara.wide.ad.jp
- irc6.nara.wide.ad.jp

このうち、irc6.nara.wide.ad.jp は IPv6 によるクライアント接続を受け付けており、ほかのサーバは IPv4 による接続を担当する。

以下、各章で、2005 年の IRC ワーキンググループの研究活動の報告と、WIDE で運用している IRC サーバ群の利用状況について報告する。

第 2 章 2005 年の IRC ワーキンググループの活動

本年の IRC ワーキンググループとしての活動として、まず運用上の作業として、「IRC サーバのバージョンアップ作業」を行った。次に研究課題として「SCTP を利用したサーバ間接続の冗長化に関する研究」といった課題で調査と研究を行った。以下、これらを順に述べる。

2.1 IRC サーバ群のバージョンアップ

IRC サーバは全世界の多数のサーバを相互に接続して運用しているため、その安定した運用のためには全体のバージョンが適切にそろえられている必要がある。2004 年の年末にバージョン 2.11 が正式にリリースされ、その後 2 ヶ月以内に全世界の IRC サーバをバージョンアップすることが求められた。

しかしながら、日本語独自の事情により国内のサーバに対してこのバージョンをそのまま利用することはできない。そこで、IRC サーバの実装である“ircd”の Ver2.11 へのバージョンアップに対応した「日本語対応パッチの作成」がワーキンググループ内で行われ、この成果により 2 月 14 日に国内 IRC サーバの一斉バージョンアップが行われた。(irc.tokyo.wide.ad.jp はそのバージョンアップに時間がかかってしまい、バージョンアップ期限間際である 2 月 28 日の午前中に作業が完了した。)

2.1.1 Ver. 2.11 以前の ircd における問題点

IRC サーバとして広く利用されている ircd では、チャンネル名やユーザ名(nick)としてアルファベットが利用された場合に、大文字と小文字とを区別することなく、内部的には同一のものとして扱うように実装されている。また、ircd は日本語のチャンネル名を扱う際には、内部では JIS コードを用いてチャンネル名を管理している。

ところが、JIS コードで利用するバイト列は

US-ASCII の符号化表からみると大文字・小文字の相当の符号化領域を利用するため、現在の ircd の実装方法では IRC のチャンネル名などとして日本語を扱う場合に問題となる。

すなわち、大文字と小文字が同一視されることにより、たとえば「#もも」というチャンネルと「#ぢぢ」というチャンネルとが同一のチャンネルとして扱われることになり、先に作成されたチャンネルのみが利用できる（「も」は JIS コードで 0x2462、「ぢ」は 0x2442）。

また、IRC はフィンランドを中心として開発されてきたことから、これまでの Ver. 2.10 系のサーバにおいてはフィンランド語の文字において大文字と小文字の関係にある“{”と“[”をも同一視している。

この影響により、「#日本語」というチャンネルと「#日万語」というチャンネルは同一のものとして扱われてしまう（「本」は JIS コードで 0x4b5c、「万」は JIS コードで 0x4b7c）。

2.1.2 Ver. 2.11.0 での変更点と日本語利用の問題

ircd の比較的大きなバージョンアップの機会である Ver. 2.10 系から Ver. 2.11 への移行に際して、これまでの Ver. 2.10 系における大文字・小文字の同一視の扱いから、フィンランド独自の問題であった部分が適用されなくなることが取り決められた。

このフィンランドにとっては大きな問題となる仕様変更を全世界のサーバにおいて一斉に行うことは困難であるため、その経過措置として、Ver. 2.11.0 においてはチャンネル名に含まれる“{”の文字は、プロトコルの処理過程において自動的に“[”に置き換えられるという処理を採用した上で、全サーバがこのバージョンに移行するのを待つ。その後 Ver. 2.11.1 にバージョンアップする際に、これらの文字を区別して扱うように変更するという、2 段階の移行手順が採用されることとなった（ほかに、“}”と“]”、“{”と“[”、“^”と“~”もあわせた計 4 組が同様の扱いを受けていた）。

2.1.3 日本語対応パッチの必要性と IRC ワーキンググループの活動

前項で述べた「IRC プロトコルにおける文字の扱い」と「日本語の文字コードの特徴」の不整合という日本独自の問題に対処するために、これまで IRC ワーキンググループでは日本語対応のためのパッチ

を作成してきた。

このパッチは、日本国内にあるサーバにおけるチャンネル名の文字の取り扱いを変更し、大文字と小文字とをサーバ内では区別して取り扱いつつ、海外サーバにおいて同一視されることになる 2 つのチャンネルが同時に生成されないようにするためのものである。

このような措置により海外サーバと接続する際のプロトコルを変更することなく、日本語チャンネル名に利用可能な文字数を飛躍的に増やすことができ、利用者の利便性を向上させてきた。

2.1.4 日本語対応パッチの Ver. 2.11 対応の必要性

上述のようなアルファベットの文字と小文字を同一視する問題に加えて、Ver. 2.11.0 では“{”を自動的に“[”に置き換える」という処理を行うことになった。この取り扱いは、海外のユーザにとっては「小文字が自動的に大文字に変更される」ことを意味するため、とくに大きな影響を与えることにはならないが、国内の利用者（日本語チャンネル名の利用者）にとっては大きな影響を与えることとなる。

たとえば「#日本語」というチャンネルに入ろうとした場合にはこの処理の影響により自動的に「#日万語」に入ろうとしたものとして取り扱われる（「本」は JIS コードで 0x4b5c、「万」は JIS コードで 0x4b7c）。

これは利用者に大きな影響を与えると思われるため、Ver. 2.11 系への移行にあたっては日本独自の対策が求められた。そこで、IRC ワーキンググループではこの問題に対処するためのパッチを作成することとした。

また、WIDE の IRC サーバが接続されている世界的な IRC ネットワークである IRCNET では Ver. 2.11 のリリースとともに、できるだけ速やかにすべてのサーバで移行することが取り決められたため、WIDE のサーバでも早急に上述の日本独自の問題に対策をした上で移行作業を行うことが求められた。

2.1.5 Ver. 2.11 系における日本語対応パッチの詳細

日本語対応パッチの作成にあたっては、Ver. 2.11.0 における経過措置が盛り込まれることとなった経緯をふまえて、日本語を利用しているユーザへの影響を最小限に留め、かつ海外のサーバへの影響を与えないことを前提とした。

まず、Ver. 2.11.0 相当の変換は日本語の利用にあたっては影響が非常に大きいことを考え、この変換

をそのまま導入することは見送った。その上で海外に先んじて Ver. 2.11.1 相当のプロトコルを利用するようにしつつ、海外の Ver. 2.11.0 のサーバとの間での整合性を持たせるために、Ver. 2.11.1 相当のプロトコルを利用しているチャンネル名は海外サーバに伝達させないようにすることにした。

すなわち、Ver. 2.11.0 対応のパッチにおいては、チャンネル名に対して次のような規則で変換を行うことになる。

- “A-Z” と “a-z” とを同一視して扱う。その結果同一のものとなる複数のチャンネル名は、最初に作成されたもののみを有効なものとして扱う。
- チャンネル名の日本語部分で “{” のような変換を必要とする文字が含まれている場合には、そのチャンネルは国内のサーバの間でのみ伝播するようにする。
- チャンネル名の日本語部分以外にある “{” のような変換を必要とする文字は、海外サーバと同様に置き換える。

IRC ワーキンググループではこの規則に基いたパッチを作成し公開した。次いでこのパッチを利用して、国内のサーバ群の Ver. 2.11.0 への移行を急ぐこととした。

ところで、上記 3 つの処理のうち、3 つ目の処理を行わないようにすることで、Ver. 2.11.1 相当のプロトコルと互換性が保てることとなる。そのため、海外サーバのすべてが Ver. 2.11 系に移行し、Ver. 2.11.1 がリリースされた段階で、上記 1 つ目と 2 つ目のみの処理を行うパッチを公開し、Ver. 2.11.1 への移行を行う予定としている。

2.2 SCTP を利用したサーバ間接続の冗長化に関する研究

IRC ネットワークを安定して運用するために、IRC サーバ間の接続を安定して維持することは重要な課題である。これまでこのサーバ間接続には TCP を利用していたが、TCP は両端で 1 つずつの IP アドレスのみを利用するため、特定のネットワークの障害によりサーバ間接続が切断されるといった問題点があった。この問題に対しては IRC ワーキンググループでは両端のサーバに対して複数のネットワークにつながるアドレスを割り当て、それらのアドレスを使いわけることにより冗長化を実現してきたが、これらの設定を手作業にて行うことは管理コストの増

大や設定ミスの原因ともなっていた。

また、IRC はそのアプリケーションとしての性質から、サーバとクライアントの接続は長時間にわたり持続されることが期待される。しかし、一般にクライアントが起動されている端末が移動し、複数のネットワークを渡り歩いた際には、端末の IP アドレスが次々と変化する。そのため、クライアントと IRC サーバとの接続を持続することは困難であった。

ところで、最近 SCTP (Stream Control Transmission Protocol) と呼ばれる新しいトランスポート層のプロトコルが注目を集めている。SCTP は通信を行う両端に複数の IP アドレスが割り当てられている際に、それらを適切に切り替えて通信を行うことで、信頼性の高い接続を提供できる。このような SCTP の特徴を効果的に活かすことで、上述した 2 つの問題のいずれにも適切な解決策が見つかることが期待される。

そこで、2005 年度後半および今後の活動において、WIDE プロジェクト内において SCTP に関する議論と開発・研究を行っている SCTP ワーキンググループと連携して研究を行うこととした。今年度においては、まず、SCTP の仕様策定状況や SCTP を利用したアプリケーションの開発技法についての調査を行い、ついで SCTP を IRC 網で利用する方向性についての議論を行った。

今後はこの結果を活かした研究や実験を行う。すなわち、作成されたアプリケーションの試験や実証実験を通じて SCTP の可能性を探り、安定したネットワーク網の構築と運用に貢献していきたい。

第 3 章 IRC サーバの利用状況

IRC ワーキンググループでは、IRC に関する研究のテストベッドとして、国際的な IRC サービス網である IRCNET の一部を運用しており、それを構成する各 IRC サーバにおいて、クライアントの接続状況などを記録している。ここでは、それらのうち、とくに IRC サーバへのクライアント同時接続数からみた利用状況の分析を行い、その IRC 利用状況を通してみたインターネット利用状況とその変化について、得られた分析結果の報告を行う。

3.1 IRC の利用状況と分析

IRC ワーキンググループでは 2005 年度に

- irc.tokyo.wide.ad.jp
- irc.fujisawa.wide.ad.jp
- irc.nara.wide.ad.jp
- irc6.nara.wide.ad.jp

の 4 台の IRC サーバを運用した。このほかにも、irc.media.kyoto-u.ac.jp、irc.huie.hokudai.ac.jp があるが、とくにここでは取り上げない。

3.2 サーバごとの年間クライアント接続数の変化

今年度は、irc.nara.wide.ad.jp において、3 分ごとに集計した、国内の各サーバの収容ユーザ数をもとに解析した。各サーバにおける 1 日ごとの最大接続数をグラフ化したのが、図 3.1 である。

これが、各サーバごとの年間クライアント数の変化となる。図 3.1 によると、週末になると、接続数が大きく増えるが、年間を通してユーザ数の変化はあまりみられない。なお、2 月中旬に irc2.11.0+jp8 へのアップグレードを行い、3 月に irc6.nara.wide.ad.jp および irc.nara.wide.ad.jp のサーバのラック移動などにより、立て続けにシャットダウンがあったため、irc.nara.wide.ad.jp の利用者が激減した。しかし、ほかのサーバの接続数は増えていないため、全体的にみて、利用者減があったようにグラ

フから読み取ることができる。また、8 月および 10 月頃に、irc.nara.wide.ad.jp の利用者が急に減り、irc.tokyo.wide.ad.jp の利用者が増えたのは、計画停電により、すべてのサービスが停止したためにほかのサーバに利用者が流れたものと考えられる。また 12 月中旬には全体的に利用者が減っているが、原因として大雪にともない、大規模停電が発生したため利用者が減ったものと思われる。irc.fujisawa.wide.ad.jp に関しては年間を通して大きな変化はみられないが、法定停電後、ユーザ数が減っている。

3.3 サーバごとの週平均のクライアント接続数の変化

年間を通して、週の各時刻ごとの平均をとり、グラフ化したのが図 3.2 である。図 3.2 によると、昨年度と同じく、サーバごとの変化のパターンに違いはほとんどない。また、曜日ごとの変化は少ないが、例年と同じく、金曜日・土曜日の夜に平日の夜よりもクライアント数が増えている。また、常時接続環境が一般的になりつつあり、国内での平日の日中のクライアント数はほとんど変化していない。

3.4 サーバごとの時間ごとクライアント接続数の変化

年間を通して、1 日の各時刻ごとの平均をとり、グラフ化したのが図 3.3 である。図 3.3 によると、すべてのサーバで同じ変化を示している。午前中から昼間にかけて接続ユーザ数が少なく、夕方から夜に

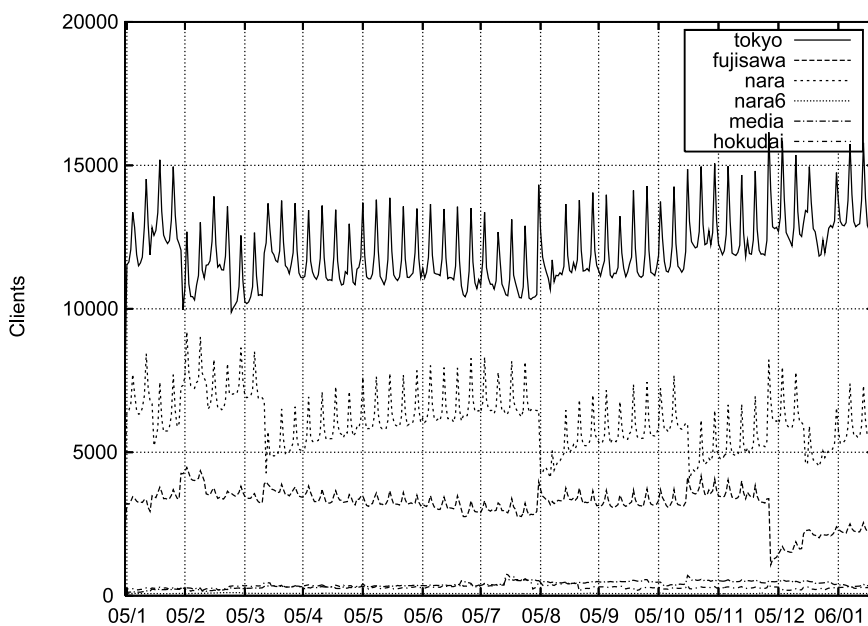


図 3.1. 年間クライアント接続数の変化

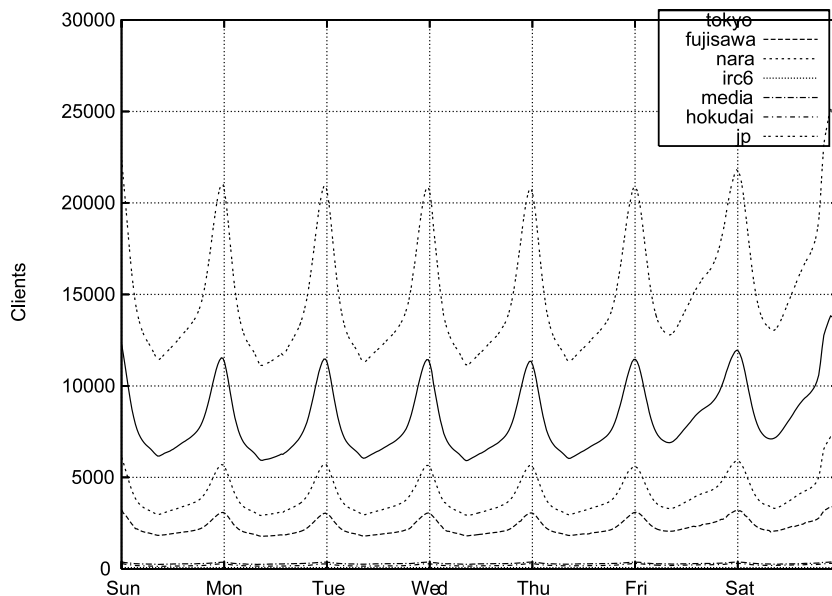


図 3.2. 週平均のクライアント接続数の変化

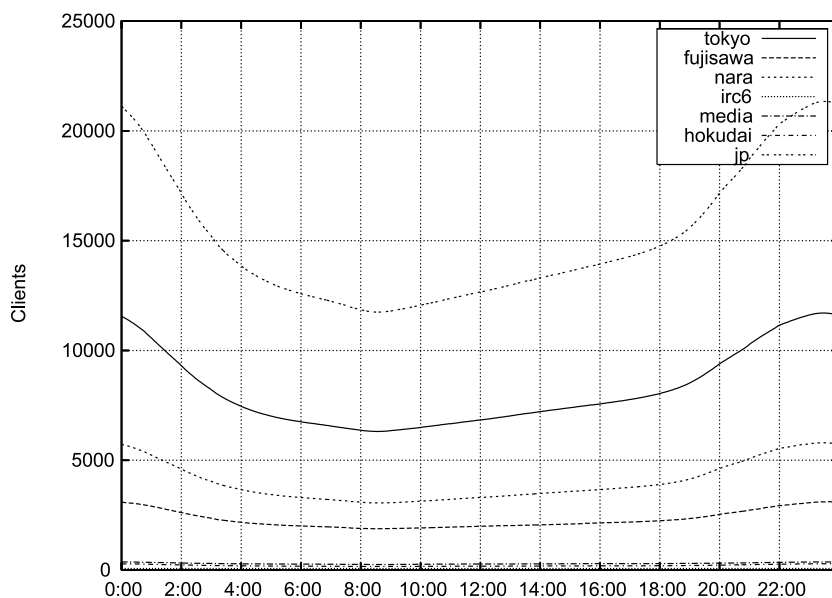


図 3.3. 時間ごとのクライアント接続数の変化

かけてユーザ数が増加し、23 時過ぎから徐々にユーザ数が低下している。常時接続環境がさらに普及したこともあり、昼夜の格差が毎年減ってきてはいるが、夕方帰宅した後、接続するという傾向があるため、夕方の接続数が増加している。

第 4 章 まとめ

本年度の IRC ワーキンググループの活動は、その活動の趣意書の変更に象徴的に現れているとおり、「運用を中心とした活動から研究を中心とした活動への転換」、そして「IRC に関する活動から、そのほか

の分野との連携を模索する活動への転換」を目指したものであった。とくに SCTP ワーキンググループとの共同研究は来年も大きなウェイトを占めてくることと思われる。

今後は、IRC というインターネットを利用した特徴的なアプリケーションの運用技術を核に、他方面の研究・運用活動との連携を模索していくことになるとと思われる。