

第 XXIV 部

IRC の運用状況とデータ解析

第 24 部

IRC の運用状況とデータ解析

第 1 章 はじめに

IRC WG は Internet Relay Chat (以下 IRC) に関する研究と、そのためのテストベッドとしての IRC サービス網を運用している。

IRC は文字ベースのチャットシステムとして開発された。IRC システムにはユーザとのインタフェースであるクライアントと、メッセージの転送を受け持つサーバが存在する。IRC ネットワークはサーバを中心とするスター型のネットワークであり、複数のサーバを接続することでサーバ群を形成する。この、複数のクライアントと複数のサーバから構成される IRC ネットワークは IP ネットワーク上に構築されたオーバーレイネットワークであり、IRC はその上で汎用的なユニキャストおよびマルチキャスト通信を提供するインフラストラクチャである。

IRC WG 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開放している。現在 IRC WG が WIDE インターネット上で運用するサーバは以下のものがある。

- irc.tokyo.wide.ad.jp
- irc.nara.wide.ad.jp
- irc.fujisawa.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からの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。これらのサーバは国際的な IRC サービス網である IRCnet の一部として日本国内のユーザからの接続を受け入れるものとして機能している。

本報告書では、まず第 2 章において、過去 5 年間のデータを元に IRCnet 利用状況から見たインターネットの利用分析を報告する。第 3 章では、IRC サーバプログラムの新バージョンである ircd 2.11.0 の新機能についてまとめた。これらの新機能は DoS 攻撃への耐性強化および DoS 攻撃そのものの意味を減少

させるものである。IRC WG の運用するサーバ群は 2005 年 2 月上旬にこの新バージョンへの移行を完了する予定である。

第 2 章 IRCnet 利用状況から見たインターネット利用分析 (2004 年)

概要

WIDE Project IRC WG では、IRC (Internet Relay Chat) 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、特に 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 利用状況を通して見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。

2.1 クライアント同時接続数の持つ意味

IRC は、ユーザが IRC クライアントを用いて IRC サービス網上の IRC サーバに TCP にて接続することで、ユーザ間のテキストベースのリアルタイムコミュニケーションを実現する。したがって、ユーザがメッセージを送信あるいは受信したい間、IRC クライアントはずっと IRC サーバに接続した形で利用される。つまり、ある瞬間に IRC を利用しているユーザ数は、サーバへのクライアントの同時接続数とみなすことができる。

このクライアント同時接続数は時間帯によって大きく変動するが、その変動はほぼ一定のパターンを示すことが観測されており、各ユーザ個別の利用状況を越えたマクロ的な動きが認められる。そのため、このパターンやその変化を追跡することで、ユーザ全体としての利用動向の時間的変化などを掴むことができる。

インターネットの利用動向の時間的変化を示すものとしては、一般には、インターネット自体のトラ

フィック変化、たとえば、IX におけるトラフィックの推移や時間帯別変動分析が挙げられ、これはインターネットの全体状況のパターンを示すものとしては最もふさわしいと思われる。一方、ここで扱っているクライアント同時接続数から見た利用動向は、次のような特徴を持つことから、別の観点による意味を持つ。すなわち、クライアント同時接続数はトラフィック傾向とは異なり、その時に利用しているユーザが1人増えると1つ増えるという対応がある点と、パーソナルユースという面が大きという点から、インターネットの中でも特に個人ユーザの利用動向を示す指標に近いと推定される。

このような点に着目して、IRC WG においては、クライアント同時接続数のデータを記録保存してきた。この蓄積されたデータを分析することで得られた興味深い結果を次節以降で報告する。

2.2.1 日の時間帯別利用動向の5年間における推移

ここでは、国内 IRCnet における1日の時間帯別利用動向の変化パターンが、2000年から2004年にかけての5年間によってどのように推移したかを示す。図 2.1 の横軸は1日の時刻、縦軸は各時刻におけるクライアント同時接続数を示す。この図の特徴的な性質をわかりやすく示すため、横軸の1日の時

刻については昼の12時から翌日の昼の12時までという形をとっている。

2000年のグラフでは、朝からお昼頃にかけての利用が少ない時間帯の同時接続数と、真夜中のピークの同時接続数には5倍という大きな開きが見られるとともに、23時頃の極端な立ち上がり、朝8時までのなだらかな減少および朝8時における小さな立ち下がりが見られる。これは23時から朝8時まで定額利用ができる NTT のテレホーダイの影響であり、非常に多くのユーザが23時のテレホーダイ開始をもってインターネットの利用を行っていたことが顕著に見てとれる。

この2000年という年は、東京めたりっく通信によって1999年12月にADSL接続サービスが開始されたばかりの状況で、2000年12月にはNTTによってもADSL接続サービスが開始されたが、両サービスによるADSL利用者は2000年末の時点でまだ1万人に届いておらず、ほとんどのインターネット利用者はダイヤルアップ接続という状況であった。そのため、特にIRCのようなリアルタイムコミュニケーションによってある程度の時間をインターネット接続するという当時としてはヘビーな使用法では、テレホーダイの影響を顕著に受ける結果となったと推測される。

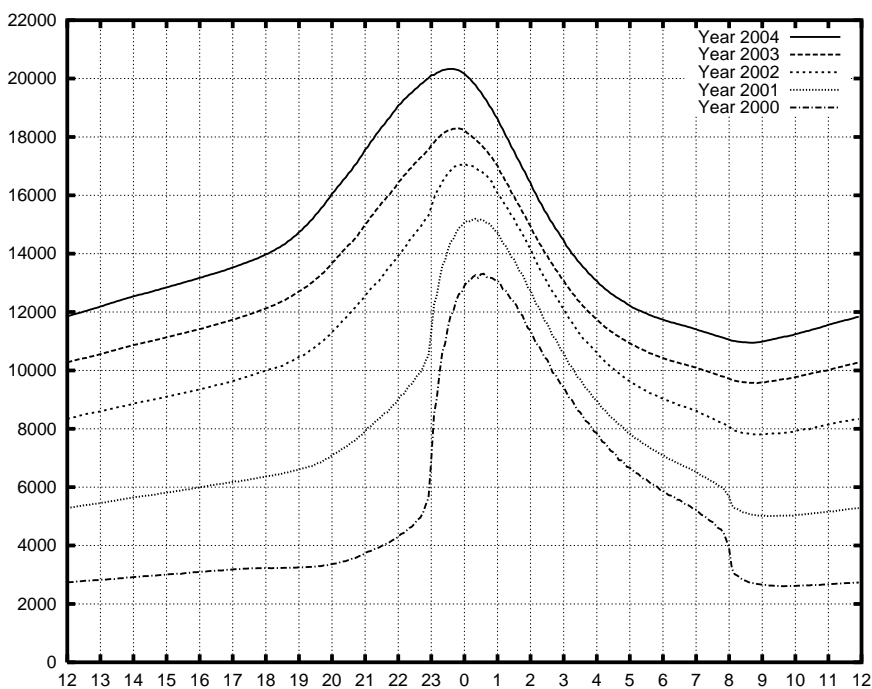


図 2.1. 国内 IRCnet におけるクライアント同時接続数の年別の1日推移

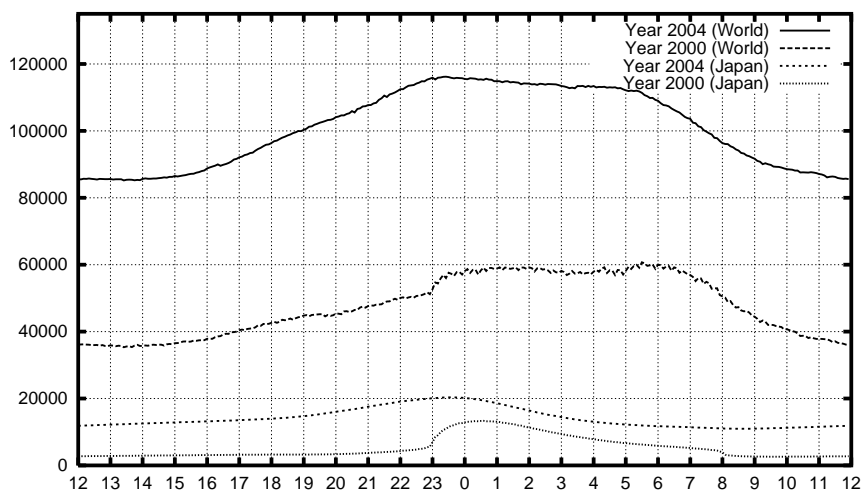


図 2.2. IRCnet におけるクライアント同時接続数の年別の 1 日推移

翌 2001 年、ADSL の普及などによって定額な常時接続環境が徐々に広まり出すと、テレホーダイによるインターネット利用の影響がかなり緩和されており、グラフが示すように、夕方以降ならかに同時接続数の上昇が見られるようになってきているが、23 時の急激な立ち上がりなども依然として見られる。

2002 年以降は、テレホーダイによる 23 時での激変といった特徴はまったく見られなくなっている。また、日中から日が変わるまでの時間帯における同時接続数の大きな上昇に比べて、日が変わってから朝までの深夜帯での上昇は小さくなっており、テレホーダイによって歪んでいた利用状況が、常時接続環境の普及で正常化されていった様子が見てとれる。ピーク時と閑散時の同時接続数の差が、2000 年には 5 倍もあったものが、2004 年には 2 倍程度にまで縮

まっており、ピーク時も日が変わった後から日が変わる前へと移動した。

図 2.2 は、図 2.1 と同様のグラフについて、国内 IRCnet および IRCnet 全体それぞれのものを示したものである。日本国内におけるパターンは 2000 年と 2004 年で大きく変化しているが、IRCnet 全体としてはほぼ同じパターンを示している。

2.3.1 日の時間帯別利用動向の曜日による違い

図 2.3 は、2000 年における図 2.1 と同様のグラフを曜日別で示したものである。また、図 2.4 は 2004 年における同じものを示したものである。これら両方の図において、月曜午後から金曜午前までの 4 日間のパターンはほぼ同一であったため、それを Weekday としてまとめて示し、残りを、金曜午後～土曜午前、

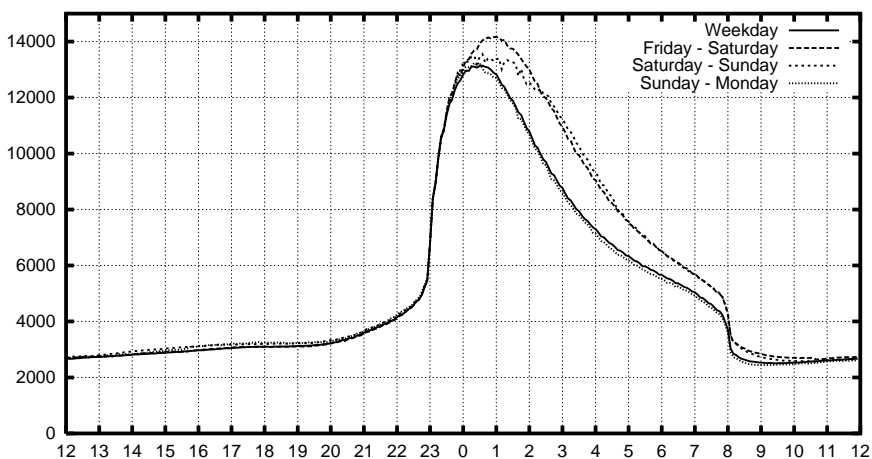


図 2.3. 国内 IRCnet におけるクライアント同時接続数の曜日別の 1 日推移 (2000 年)

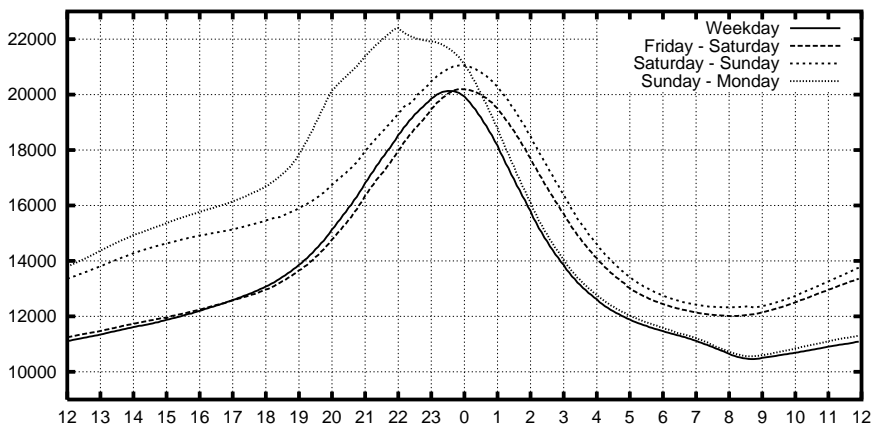


図 2.4. 国内 IRCnet におけるクライアント同時接続数の曜日別の 1 日推移 (2004 年)

土曜午後～日曜午前、日曜午後～月曜午前として、合計 4 つのグラフをそれぞれの図において表示している。

2000 年の状況を示す図 2.3 においては、テレホーダイの時間の深夜から早朝にかけてのみ、曜日的特徴が現れている。すなわち、金曜深夜(土曜未明)と土曜深夜(日曜未明)のその時間帯のみ、ほかの曜日と比べて同時接続数の上昇が見られる。週末といえども日中については利用がほとんど増えることはなく、定額で使えるテレホーダイ時間による制約を大きく受けていたことがわかる。

一方、2004 年の状況を示す図 2.4 においては、曜日における利用状況の差が全時間帯に渡って大きく出ている。まず、最も大きな違いとしては、日中においても、平日と比べて土曜と日曜は大きく増加していることがわかる。この平日との差は、金曜深夜(土曜未明)から土曜の日中、土曜深夜(日曜未明)、日曜の日中および日が変わるまでとほぼ切れ目なく続いている。

もう少し細かく見ていくと、さらに曜日による特徴的な動きが見てとれる。注目すべき点は、金曜の夜はほかの平日と比較して、わずかながら同時接続数が減っているという点である。それでも日が越えるあたりからは平日と比べて大きく増加していることと、世間で一般的に見られる行動から推測すると、特に金曜の夜は飲みに出かけるなどして家でインターネット利用開始が遅くなっている人たちが、ある一定数いることによる影響がグラフに出ていると考えられる。

また、増加している土日の間で比較すると、すべ

ての時間帯において日曜の方が同時接続数が大きくなっている。特に、夕方から夜にかけて日曜は圧倒的に増えており、これも金曜や土曜と比べて、日曜のほうがこの夜の時間に家にいる人が多いためではないかと推測される。さらに、この同時接続数のピーク時間についても、平日は 23 時半、金曜と土曜が 0 時であるのに対して、日曜は 22 時と早くなっている状況がグラフよりわかる。

2.4 クライアント同時接続数の推移

図 2.5 は、国内 IRCnet での、1997 年 7 月から 2004 年 12 月までの各日におけるクライアント同時接続数の最大値と平均値を示している。国内 IRCnet のサーバは 1989 年より存在していたが、1997 年 6 月以前の各日のデータは記録できていないため、各日の記録がある以降のものが図となっている。ただしそれ以前についての同時接続数は多くて 1000 ちょっとであったため、図示されているものから大きな変化はない。

下のグラフとなる同時接続数の 1 日の平均値については、ほぼ一貫して増加が見られる。一方、上のグラフとなる同時接続数の 1 日の最大値については、2000 年くらいまでは大きな傾きで増加しており、平均値との 3 倍近くの差が出ていた。これは前節で示したような、テレホーダイ効果による利用時間集中の影響と思われる。それ以降は、徐々に平均値との差を縮めつつあり、現在の差は 1.5 倍程度となっている。

次に、図 2.6 は、国内 IRCnet におけるユーザが各サーバのどこへ接続したかの変遷を示している。横

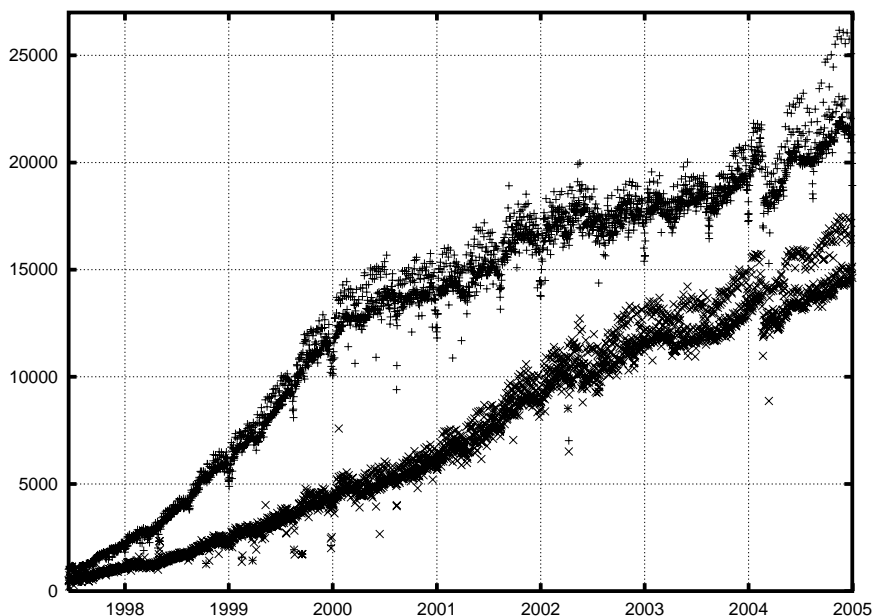


図 2.5. 国内 IRCnet におけるクライアント同時接続数の最大と平均の推移

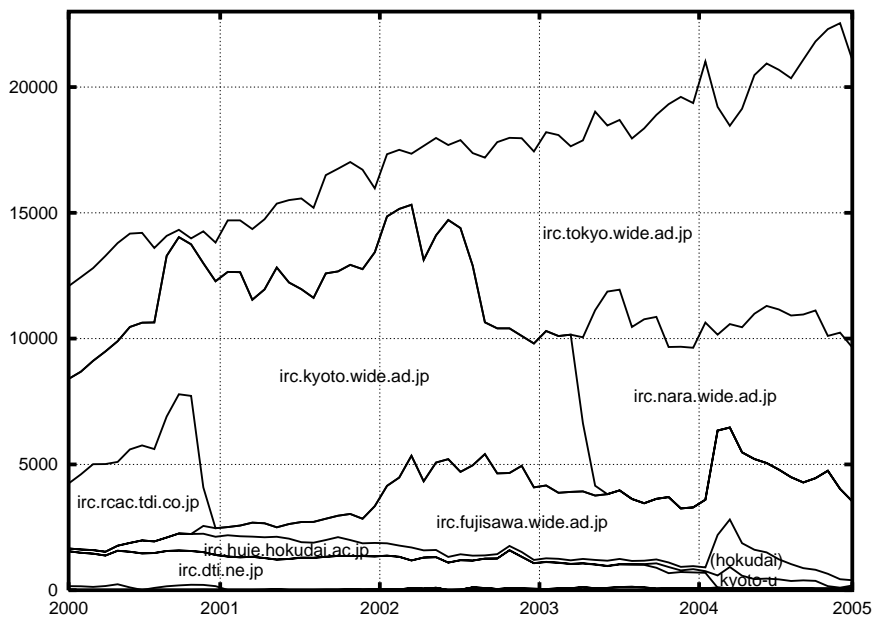


図 2.6. 国内 IRCnet におけるクライアント同時接続数のサーバ別推移

軸は2000年1月から2004年12月までの5年間、縦軸は一番上が各時期における国内 IRCnet 全体でのクライアント同時接続数の最大値を示しており、各サーバ個別におけるクライアント同時接続数の推移がわかりやすいように、その中で層をなす形にて表示している。

各サーバ間でのクライアント同時接続数の変化移動が激しく見える理由は2つある。1つは、サーバの

廃止にともなうもので、たとえば、irc.rcac.tdi.co.jp が廃止となった時には、そのユーザが意識せずとも irc.kyoto.wide.ad.jp にて吸収して引き継ぐ形で移行運用を行ったため、図のようなグラフとなっている。また、irc.nara.wide.ad.jp を新設して irc.kyoto.wide.ad.jp を廃止した時にも、ユーザが意識せずとも移行する運用処置をとったため、図のような形のグラフになっている。

もう 1 つの理由は、サーバのメンテナンスあるいは停電によるダウンやリスタート、および、ネットワーク不調などによるサーバへの不到達トラブルなどである。これらはすべて、一時的にせよサーバとクライアントの間の TCP 接続が切断される。このような場合、IRC クライアントはサーバへの再接続を試みるが、IRC の性質上、同じ IRC サービス網につながっているどのサーバに接続しても同じサービスを受けられるため、同じサーバへの再接続がすぐにできなかった場合、別のサーバへの接続を試みる。そのような再接続の設定や別のサーバの選択などはユーザにより異なるが、多くのユーザが別のサーバ利用へとまとまって移る結果となる。

2.5 まとめ

インターネットの利用状況の一面を分析するため、IRCnet におけるサーバへのクライアント同時接続数について、その推移や 1 日の時間帯別による変化の動向についての統計処理と図示を行った。その結果、ここ 5 年間でインターネット接続環境の変化について、テレホーダイ利用から常時接続環境への移行が顕著に示されるとともに、曜日別によって特徴ある利用のされ方が行われていることなどが判明した。

第 3 章 ircd 2.11.0 の新機能

3.1 ircd 2.11.0 の特徴

本節では ircd 2.11.0 で新たに導入された新機能のうち、特筆すべきものについて説明する。

- Unique ID (UID)

今まで広く使用されていた irc サーバでは、ユーザの識別に nickname が使用されてきた。しかし、ircd 2.11.0 では、Unique ID が実装され、サーバ内部での識別に Unique ID が使用される。Unique ID は、ISO3166 countrycode、Server Number、Serial Number の 9 文字の組み合わせから成り立ち、重複のない ID が生成される。なお、ISO3166 countrycode、Server Number をセットで ServerID (SID) と呼ぶ。

- Classless Inter-Domain Routing (CIDR) のサポート

bans、ban exceptions、invite lists に対して CIDR がサポートされた。

例) mode <channel> +I *!*@192.168.0.0/28

- reop hints

サーバ reop チャンネルフラグ (+r) は、! から始まる safe channel で使用することができる。そのサーバ reop チャンネルフラグが立っているチャンネルで、オペレータ不在 (opless) になってしまった際、一定の法則に基づき自動でサーバからチャンネルオペレータ権限が配布される。これをサーバ reop と呼ぶ。この機能を拡張し、特定のユーザに対し、チャンネルオペレータ権限を配布する機能が搭載された。この reop hints に関しては、safe channel 以外、すなわち # から始まるチャンネルでも使用することが可能となる。なお、この機能についても CIDR がサポートされる。

例) mode <channel> +R *!*@*.wide.ad.jp

3.2 問題回避に使用される Unique ID

Unique ID は、nickname が重複発生した際に代替 nickname として使用される。本来、IRC では nickname の重複は許されないのにも関わらず nickname の重複が発生した場合は、公平に両方のユーザがサーバから切断 (nick collision) される。しかし、この仕様を悪用するユーザが少なからず存在するため、これを回避するために実装された。このようなことは、サーバ間接続が分断 (split) した後、サーバ間接続が復帰したときに稀に発生することがある。

また、IRC では、nickname の最初には数字が使用できないという制限がある。しかし、Unique ID は ISO3166 Countrycode から始まるため、数字で始まる。これは例外的な扱いになり、ほかのユーザが詐称できないことを示し、第三者が Unique ID を nickname として名乗ることができない。そのため、Unique ID を nickname として使用することにより、server split を悪用した nick collision を防ぐことが望める。

3.3 +beI の拡張

今まで +beI を設定する際、'nick!user@host' という形で記述していた。そのため、irc サーバ単位で bans、ban exceptions、invite lists は記述できなかったが、この制限を改善し、サーバ単位での記述を可能

にした。具体的には irc サーバに設定される SID を利用して記述する。また、サーバ内部では nickname と UID が関連付けられており、UID 単位で記述することも可能である。

例) mode +b 000A*!*@*

参考:

+b (ban)チャンネル入室禁止ユーザ

+e (ban exception)チャンネル入室禁止の例外

+I (invite list)チャンネルへ入室可能ユーザ

3.4 ircd 2.11.0 に期待できること

ircd 2.10 までにおいては、nick collision を発生させるために DoS 攻撃を行う有害なユーザが存在したが、新たに Unique ID が導入されることにより、DoS 攻撃を行い server split を発生させても nick collision を発生させることができなくなるため、DoS 攻撃を減らすことが可能になると予想される。

