■IRS17 ログ

□XXX日問題ふたたび
JPIX 西野大さん

IRS14では、特定のベンダーで830日目に再起動が発生する問題があることを
お知らせしたが、同時期に別のベンダからも似たような障害が報告された。

よく調べると、XXX日問題はポピュラーな問題のようだ。

注意:
ネットワークの事故防止などを目的として、運用者の情報共有が問題
特定のベンダーさんを問題視する意図はない

・497日問題
カーネルのタイマー（tick）が最大値になってゼロに戻る
OSなどのシステムタイマーが
32bit (unsigned)
100Hz (10msec刻み)
いろいろと、いやんなことがおこる

例: デバイスドライバでロックやクラッシュによる再起動
例: プロセス間通信が出来なくなり、不具合

・830日問題
上記の
32bit (unsigned)
60Hz
バージョン

・事例: MS Windows の場合
問題: Windows XP/2000/NT 4.0 で WM_TIMER メッセージがプログラムに
通知されないことがある
Windowの再描画がうまくいかない→再起動しかない
MSサポート文章番号: 322913, 323328, 811756

・事例: Linux の場合
497日問題
kernel 2.4ではカーネルとモジュールに問題が生じる場合がある
kernel 2.6以降は基本的に大丈夫

C: kernel 2.2 でプロセスの情報を正しく取得できなくなる問題が発生した

・事例: BSD系の場合
497日問題はない

2038年問題
BSDに限らず UNIX系の源流を持つOSでは
1970年1月1日0時0分0秒(UTC)からの経過秒のカウンタをkernel内部に
持っている。このタイマーが32bitのsignedだと
2038年1月19日
に桁あふれをおこし、問題を生じることが予想されている。
この問題は Linux にも共通

組み込み系のシステムに取り込まれていることを考えると問題

・事例: SNMPの場合
SNMPの標準MIBにおけるuptime
.iso.org.dod.internet.mgmt.mib-2.system.sysUpTime

ロールオーバーするはずではないが、ロールオーバーする実装
いろいろありそう、大きな問題ではない

・これまでの自分の考え方…
「Linux や Windows だと、しょちゅうパッチをあててリブート
しているので関係ないぜ」
「悪魔崇拝者（BSD教）には、関係ないぜ」
「われらネットワーク屋には、関係ないぜ」

しかし、ネットワーク機材には、ペンギン（Linux）の遺伝子が
はいっている（組み込みシステムのカーネルとして使用している）
ものも多い、…ということを忘れていた

・事例: Force 10 Networksの場合
スイッチの RP の再起動から 830 日が経過すると、IRC（RPモジュール間）と
IPC（プロセス間通信）がタイムアウトを起こす

・対象
ハードウェア: Force 10 Networks 社　E-series
ソフトウェア: 7.7.1.0 以前

・発見日
2007年9月28日

2008年3月に改修版のファームウェアをリリース
ベンダ側で土日をあげて再現試験をやっていただき問題切り分けは早かった

- 参照先
Force 10 Networks
PR #72905

- その他
これとは別に497日問題も存在する。が、FTOS 7.0 から
「心を入れ替えて」別の OS をベースになったので
バージョン 7.0 以上を使っていれば497日問題は存在しない

・事例：Extreme Networks の場合
- 問題：
スイッチを再起動してから 497日が経過すると自動的にリブートする

- ハードウェア
- ソフトウェア
ExtremeOS 11.x　XOS 11.6.2 以前
ExtremeOS 12.x　XOS 12.0.1 以前

- 参照先
Field Notice FN0315
- その他
そういえば Extreme Networks 社製品は XOS になってから
Penguin 教（Linuxベース）になったという情報もあり

・仮想化OSの場合（余談）
VMware ESX server 3.5で（個人的に）遊んでいる。
VMwareのカーネルは、
kernel 2.4
ベースに構成されている。

VMware ESX 3.5 は 2008年1月10日にリリースされている
（まだ、497日経過していない）
リリースノートやKBにも記載はないが、大丈夫だろうか？

- 仮想化ホストが死ぬときは、ゲストOS10台以上を引き連れて集団自殺…

・ベンダーさんに訊いてみたい
　　こういうバグって発見しにくいものですか？
　　こういうバグって取れないものなのですか？

　　kernelの初期値が大きいものをつかってテストする?
　　発見するのは497日待つしかない？

□質疑応答

C: 悪魔教なので497日問題の影響を受けなかった
　2000年問題のときは tick を変えてテストした
　2番目の理由だと品質管理部門の仕事ではない: それより前の段階の問題

C: 組み込み系では、クロックを10倍にするなど、すごいクロックを与えて
　加速試験をする場合もあるらしい

・会場の方に訊いてみたい
　　ユーザはどうしたら良いと思うか？
　　怪しそうな装置は定期的にリブートを行う？

・SNMPで自衛策を講じる？
　　mib-2.system.sysUpTimeを監視する

・まとめ
　　XXX日問題は、いろいろとありそう
　　「ゴキブリ(バグ)は、一匹見つけたら、30匹いると思え」
　　の一般法則によれば、あちこちに地雷が埋まっていそうな気がする

　　＞メーカーさん、ベンダさん
　　＞オペレータのみなさん
　　装置を起動してから497日目と830日にはご注意ください

□質疑応答

C: 冗長系を構築する際には、(両系とも) 同時に電源上げるのが普通
　数時間、時間をおいてから上げるのはどうか。

C: Force10 の E シリーズの古いものを利用しているが、古い装置は直せません
　という回答があった。解決するには買い換えか、497日おきに再起動するか…

C: F10はカードの密度があがっていくので、それに伴いどんどん FTOS が新しく
　なっていく

C: Linux をベースにしている機器には、Load Balancer、Shaper、STB などがあ
　るので確認が必要ではないか

C: 冗長化していた機器で証明書を使っているものがあるが、証明書の有効期限が
　切れて制御できなくなるケースがある

C: ベンダからのリリースノートを見ないとわからないので、Google で調べても
　出てこない。

C: 組み込み系でも、組み込みOSの都合で加速度試験ができない場合もある。
　タイマーをたくさん持っている装置もあるので、どうやって品質検査するのか
は課題。
　各タイマーの持ち方が異なる場合も多いので、
　システム全体ではなく、特定のプロセスだけ落ちるということが発生する
　場合もある。

□ホスティングサービスにおけるARP Spoofing対策
さくらインターネット(株)
大久保修一さん

□質疑応答

Q: ポート単位で L3 のフィルタを書ける L2 スイッチ製品もある。
　自分がソースの IP アドレスしか出さないというやりかたもあるのではないか
。
　IP ヘッダのソース IP アドレスでフィルタする方法。

A: ARP は IP ではない。
　そもそも通信できなければ悪意のあるアタックは防げるのではないか。
　他のホストの ARP の改竄もできてしまう。
C: ウィルスによるインジェクションは防止できても、(通信はできないので)
　障害が発生してしまうのではないか。
C: コンテンツを改竄されるよりはましではないか。
　(このあたり少しdropしてしまった記憶があります..すいません..)

C: 未使用のアドレスを使って SPAM を送るのには有効だとおもう

C: この問題はハウジングやホスティングだけではなく、ケーブルインターネット
　でも起こりうる問題ではないか

Q: Port isolate 機能を利用するのはどうか

C: (ホスト同士が直接通信できないようにするため、) ルータの ICMP redirect
　機能を無効化する必要があるだろう

C: Webサーバとメールサーバを同一セグメントに設置している場合には、相互に
　通信できないと困る

C: 顧客毎に VLAN を分割する方法もある
　US ではサーバー台に対して /29などを割り当ている事業者もある
Q: IPv4アドレスが枯渇しているが、日本ではIPアドレスが取りにくい
A: JPNIC においても、必要性を示すことができれば割り当てが行われる
C: /31の利用も検討対象である

C: IX側ではMACアドレスやARPについて制限をしていない例が多い
　装置を急いで変えることがあるので運用上やりにくい
　学習数を制限する方法については、途中に変なスイッチが入っていると1は難し
い

C: お客さんが間違えてネットマスクを設定している + proxy arp を設定している
場合、
　トラフィックが吸い込まれる
C: わざと arp spoofing している例もある
　　やめたはずのお客さん宛のARPがすごく多いのでうるさい
　　スポンジARPと言っている

C: ping, udp(traceroute)は止められない
　TCP Port 179は止めている

C: 監視装置から昔ピアを張っていたところにpingしているケースがある
　high-arper というか、明らかに大量の arp が発生
　5分に数発が普通だろうが、1分60発のようなケースも。すごい人は
　1分あたり700発とかある

C: port isolation の場合、router に local proxy arp という機能を
　もった場合もある。同じ足から来て同じ足から出る

Q: IPv6 の場合にはどうか
A: NDP Storm が出る

Q: ARP Spoofing 検出システムでポートを抜去する際、抜いたらどうするか?
A: ぬきっぱなし
C: 抜去したあと、Gratuios ARPで戻してあげるとよい

A: arp spoofing には悪い使い方だけではなく、良い使い方の例もある。
検疫ネットワークで、認められていないクライアントに対して 妨害で
ARPを送る手法がある。
C: 監視する方法としては、arpwatch もある。IXでも動いている。

Q: static arp の設定を cfengine とか puppet でばらまくということは
できないか。puppet は SSL の証明書をつかって認証できる。
A: 多数のノードの設定変更が必要な方法は難しい

□経路ハイジャック対策のための Prefix Filter について考える
～これからの IPv4アドレス枯渇時代に向けて～

Verizon Business 伊賀野さん
NTT Communications 浜田さん

このセッションは思考実験。PIアドレスでの接続サービスを提供されている
プロバイダの皆様が今まで継続して取り組んできた内容ではある。

「他人のIPv4アドレスを騙って接続に来るユーザとかって出てこないかな？」

C: 経路ハイジャックといっているが、むしろ「Prefixハイジャック」ではないか

・考えられるケース:
海外のどこかのアドレスを使って日本向けのサービス
台帳管理はされている(IANA/RIR/NIR/LIR)
オペレーションしているのは各ISPの運用

送信フィルタ
接続種別毎に自分で適切なPrefixを広報するしかない
受信フィルタ
上流接続&ペイドピア
Full Route/Partial Routeを全て受けるしかない
フリーピア
ピア先が自分自身とその顧客等の希望するPrefixを広報
顧客接続
顧客が利用予定のPrefix情報を知らせてくる
BGP接続
Static接続

・Prefixジャックを防ぐには?

怪しいPrefixを使いたいといったときにISPは接続を拒否できる？
顧客A社が「B社のPrefixだけど、B社が許可している」といっている
・さらにB社に確認する？
・全てのPrefixのOwnerをチェックして、そこに連絡する？
A社の接続完了後、C社からクレームが来た
顧客A社のフィルタを変更したいが、A社の許可が取れない
A社の確認は必要？

C: ケースバイケース

C: フィルタリングのガイドラインが必要ではないか

C: (このようなケースは) 日本では余りない。アジアではある

C: 正しくやっていればpunching holeである。
日本では「性善説」(ピア先が意図的に不正な経路を広告しないという暗黙の前提がある)。
そもそも他のISPのアドレスを広報する場合があまりない。

C: US では普通に行われる運用である。ぼくが SE ならチェックする。
社内フロー上は RADB を確認し、どういう色のIPアドレスかチェックする。
多国籍企業等で、US から（日本法人は）この IP アドレスを使えと言われて、

punching holeになっているケースもある。

C: 日本だけで PI のフィルタを書きましょうというわけではなく、グローバル顧客接続（水際）で適切なオペレーションが必要だと思う

Q: 問題がよくわからない。接続のときにチェックすれば終わりではないか。

C: IIJ は（ピア向けに）AS_PATH update のメールによる通知をやめた。JPIRRベースでやっている。ntt.net は（顧客向けフィルタを）IRR から自動生成している。

（フィルタの内容については）IRRに登録した人の設定の責任。何を持って正しいというかが課題。

C: BGPのピアってIXの接続があって、顧客接続があって、上流接続がある
ピア：自分と自分のお客さんをながす
上流：フィルタ出来ない
顧客：顧客から来るものはフィルタする
一番面倒なのはIXのピア

Q: 顧客からのフィルタの質が問われてくる

C: 顧客からの経路を受けるかどうかというより、ポリシー問題である。
ガイドラインみたいなものが必要

C: 一社だけで（フィルタを）やっても、汚れてきたものが入ってきたら対応が難しい

C: 顧客を収容する機器では、顧客のプレフィックスだけは受けて、それ以外のプレフィックスはフィルタする。

C: 新規にインストールする場合、オーダーのプロセスに課題がある。インストール後は、お客様からのアップデートの依頼が来るので、ここも対応が必要で、短納期の場合もしばしば見受けられるので、チェックを適正に継続的に実施する難しさがある。

C: プレフィックスジャックというのは悪意をもってやる
自分のネットワークのコネクティビティがあって、ほしいとやっているから、お金をもらいながら悪意を持っている。

ある程度チェックすればなんとかなる。経路ジャックという線を切るとすれば、そういうことをやっていないところが問題。

C: カスタマー接続するときには大原則として「Prefix Filter をしましょう」というのがある。そのクオリティというのを考えましょうかというのが問題提起。

C: チェックするしかない。ユニークネスを持っていることは確認できる

C: IRR は誰がどんなものでも記載できる。

議論しているのは Internet Registry。
自分が登録者ではないIPv4アドレスを持ってきて、使わせてくれといったら、(IPv4アドレスの)オーナーに確認するしかない

Q: punching hole を（顧客に）してくれといわれた（サービスプロバイダの）人はいるか？

A: いる

Q: （その際、IPv4 アドレスの登録者に）確認したか？

A: Yes

C: ISPによっては（punching holeの）数が半端じゃなかったりする。

C: 数が多いからって確認しなくていいわけではない

C: PI はレジストリ上があるからいいんじゃないの

Q: PI を提供している人は？

A: …数名が挙手

Q: PI でチェックしている人

A: …挙手者中ほぼ全員

A: 使っている人のプレフィックスはIRRには載っていなかったが、whoisに載っていたため正しい接続として確認した。

A: /8 か/16 を持っている、とある US の大きな会社が、そのうち /24 が日本ブランチ向けに割り当てたといわれたことがあった。そのときは、法人名が一緒で Japan とついてるからいいと判断。「え、」とは思った。

C: 新しく使うときはわかる。

運用をスタートしてプレフィックスを追加するときはどうしているか。
新規のときと追加のときで、社内のルーティンは違うのではないか。

みなさん事情があって「明日までにお願い」等。
ルーチンワークに紛れ込んでいるとチェックをしにくいのではないか。
トラブルを事前に防ぐ手段を考えたい。

C: それはオペミス、確認不足ではないか。正規の手続きを踏んでいるわけではない。

悪意があるものから防ぐというものと、正規の手続きを踏んでいるものは別。
経路ジャックと書かれていることに違和感を感じる

Q: ドメイン名でも同じでようなケースがあるのではないか。
DNSはISP、xSPで持ってくれという場合、追加・更新とか。

A: 契約者であることが確認できれば、依頼内容をそのまま反映する。

A: お客さんの書いたゾーンがあった場合、そのまま反映される

A: どこに Authority を作るかが結論

C: 与信をあたえる方法と一緒
　　機器発注が来たが、信用できるかどうかと同様
　　信頼できる経路のソースがあれば同じ
実は whois でチェックしているかもしれないが、信頼できないなら別の方法を採る必要がある

C: (APNIC の) リソース証明書は、証明書内に名前や組織名をいれず、プレフィックスのみをいれることを考えているようだ。

C: お客さんが申告した内容をチェックする手順は必要。
ハイジャックが実際に行われた後に、ハイジャックであるといって問題にするには、何らかのチェック方法が必要

Q: クレームが来たらどうする？

A: 土地の場合には登記簿がある
JPNIC のような IRR のサービスは一つの解決
APNIC の方が危機意識は高い。APNIC はリソース証明書を準備している

C: RIR は、リソース証明書+ROA(Route Origination Authorization)で出来るようにする方向に走っている

C: sBGP の署名を考えているのではないか
IRRだと自動化するとそれを信用するしかない

C: 登録を Validation していい
DoS を受けたときに Blackhole して欲しいというリクエストがあった場合、(リクエストの内容を) Validate しないと吸い込まれたりする。ただし、経路が来ているかどうかの確認なので楽ではある。

C: ガイドラインは必要だという人が多ければ、こういう問題を未然に防げるなら防ぐというアクティビティを起こさないか

C: グローバルに share していく必要があるのではないか

C: 結構そういう事例は多いのか

C: ガイドラインだから守る必要はない。ルールではない。

C: 海外にも出さないと意味はないと思う。

C: ルールのようになることに違和感を感じる。
コンセンサスは必要だが、どれだけ問題が起きているのか
問題が多発していていてそれだけ対処する必要があるのか。

事前対処か、事後対処か。

ソーシャルアタックや BGP オペレータを抱き込めば、(不正な経路の広報も)出来るのではないかガイドラインをつくるだけで防ぐのは難しい。

お客様の経路を流すかどうかは ISP の自由であるが、そういう変な経路ばかり広報してくるような ISP があれば、信頼を落とすことになる。そうなれば、

ピアを落されてしまうこともあるだろう。従ってコミュニティで対策していれば、それなりのチェックをするのではないだろうか。

C: やっていることは同じ
プレフィックスフィルターを書きましょうというのは、自由。
だれをガイドするのか

C: SFBay Packet Radio という組織が持っていた /16 のアドレスブロックが、元の組織との関係性が不明な、似たような名前の会社に使われていた。
乗っ取りが発覚しにくいよう、こうした名前にしたのかは不明であるが、whois の名前だけで判断することに課題がある。

参考：http://blog.washingtonpost.com/securityfix/2008/04/a_case_of_network_identity_the_1.html

C: この問題は与信の問題である。
100%防げるかというとNO、スクリーニングするならNO、more better

C: 今必要ないものはいま作る必要があるのか。
起きるかどうかはわからない、危険性がわからない状態でいまやらなければ
ならないのかが疑問

C: 危険性は高いのではないかと思っているが、必要に応じて実施すべき。ただ、
質の問題なので、品質を高める事をすれば効果があると思われる。

C: 100%スクリーニング出来ているかとかというとそうでない。
その文章があるとうれしいことが見えてこない。
その文章がないと社内に周知できない、というような状況であれば価値はある

特にそういうものがないなら、Prefix Filterをします、というのが、潜在的に

問題を発生させる可能性があるためゴールが見えない。
USにガイドラインはあるのか。本当に必要があるのか。

C: 他のISPでは受けるが、うちでは受けない理由が必要
約款みたいなので、あやしいのは受けないとか。

C: なるべくインターネットの自由を下げたくない
自由度を下げて運用したくない

C: 技術屋さんと営業でせめぎあってない？
このお客さんをやるなら技術的に検証ができていないことをやらないと、
他のプロバイダに行ってしまう。

C: お客さんから来るアドレスでフィルタしても、他のISPから流入するのをフィ
ルタするのは大変。フィルタするには理由が必要。

C: 内容として whois と IRR の登録内容を調査することが書かれているだけでは
、
ガイドラインとしては嬉しくない気がする

C: せっかく4人集まったのだから、一度具体的なものを書いてみるとよいのでは。

C: ハイジャックされるのは使われていないPI、使っているPI？
使っているならすぐわかる？
使われていないのなら whois と IRR でチェックする必要がある
帝国バンクのデータを参照する等

C: そこに証明書がついてくると登記と証明書がマッチングできない
PI は whoisに載るまでの壁は低い

C: こないだ（JPNICが未使用の）PIのリストを公開した。リストの中のIPアドレ
スには、海外で使われてしまっている例もある。

C: 具体的に誰が、どこに言いたいガイドラインなのか明確にしてほしい。
経路ハイジャックというのはテクニカルにセキュリティホールを
踏んでくるイメージがある。

今回の問題提起は手続きを踏んでやってくるのは別に考える

経路を乗っ取られるにはいくつかのパターンがあり、それを整理した上で
フォーカスを定義すべき

C: キックオフのミーティングをやりたいと思っている
参加したい方は手を挙げてください

□JANOG 23
2009年1月22日(木)、23日(金)
高知県立県民文化ホール内グリーンホール

□IRS
日時：11月ぐらい
場所：提供いただける方大募集！！