Vol.2016-IOT-35 No.16
Vol.2016-SPT-20 No.16
2016/9/24

公開用NTPサービスの運用と課題

藤村 丞[1,a)]　谷崎 文義[2,b)]

概要：福岡大学では日本初の公開用 NTP サーバの運用を 1993 年 10 月から開始し、22 年が経過した。この間にトラフィック量が増え続けていることはもちろんであるが、当時に様々な問題点も生じてきた。本発表では公開用 NTP サービスについての現状分析と課題の取り組み状況について述べる。

The operation and challenges for the future of the public NTP Servers

FUJIMURA Sho[1,a)]　TANIZAKI Fuminori[2,b)]

1. はじめに

福岡大学は福岡県福岡市に所在地を置き、9 学部 31 学科、10 研究科 33 専攻、学生数約 21,000 名（学部生+大学院生）、大学病院 2 病院、附属高校 2 校、付属中学校 1 校を有する私立の総合大学である。

この福岡大学では、1993 年（平成 5 年）10 月に日本で初めて全世界に向けて NTP サービスを開始し、22 年が経過した現在も変わることなく、同サービスの提供を行っている。この 22 年の間にトラフィック量が増え続けていることはもちろんであるが、この公開用 NTP サービスが引き金となったネットワーク障害が発生するなどの様々な問題点も生じてきた。

本稿では、この公開用 NTP サービスについて、障害の事例とその対処法、現状分析と課題の取り組み状況について述べていく。

2. サービス開始の動機とその概要

1993 年頃の福岡大学では、大型計算機を学内の各所と同軸ケーブルで結んだ中央集中型（スター型）の構成を取っていた。この頃の計算機は、起動後に正確な時刻を手入力にて設定する必要があった。あるときこの設定を間違えて入力してしまったことがあり、時刻設定を自動化する仕組みを整えればこの間違いを防ぐことができると考え、自動的に時刻同期をする仕組みを整えることにした。当時の郵政省の標準電波では実用に耐えうる精度が出なかったため、GPS 受信機から時刻情報を取り出し、それをコンピュータにて自動化する仕組みを整えた。また、この頃日本国内ではまだ一般向けの NTP サービスが提供されていなかったので、1993 年 10 月から一般公開し本サービスを運用している。現在運用中の公開用 NTP サーバは、以下の 2 台である。

- 133.100.9.2（clock.nc.fukuoka-u.ac.jp）
- 133.100.11.8 （clock.tl.fukuoka-u.ac.jp）

これらの FQDN と IP アドレスは、公開用 NTP サービス運用開始時から現在も変更していない。また、公開用 NTP サーバは、2004 年（平成 16 年）9 月より ntp.org[*1]の Public Time Server Lists [*2]に登録している。

3. 2015 年 8 月までのネットワーク構成

福岡大学では、2010 年（平成 22 年）8 月より運用してきた教育研究システム[*3]FUTURE[*4]を、2015 年（平成 27 年）

[1] 福岡大学
8-19-1, Nanakuma, Jyonan-ku, Fukuoka 814–0080, Japan
[2] 西日本電信電話株式会社
3-2-28, Hakataeki-higashi, Hakata-ku, Fukuoka 812–0013, Japan
[a)] fujimura@fukuoka-u.ac.jp
[b)] fuminori.tanizaki@west.ntt.co.jp

[*1] http://www.ntp.org/
[*2] http://support.ntp.org/bin/view/Servers/StratumOneTimeServers
[*3] 医療系・事務系・図書館以外のシステム、学内ネットワーク・PC 教室など
[*4] Fukuoka University Telecommunication Utilities for Research and Education

8 月に第 5 世代目の FUTURE5（FUTURE version 5）として更改した。この更新時期と同時に、公開用 NTP サーバの構成を変更したが、その変更前の構成を図 1 に示す。この図からもわかるように公開用 NTP サーバは、2015 年

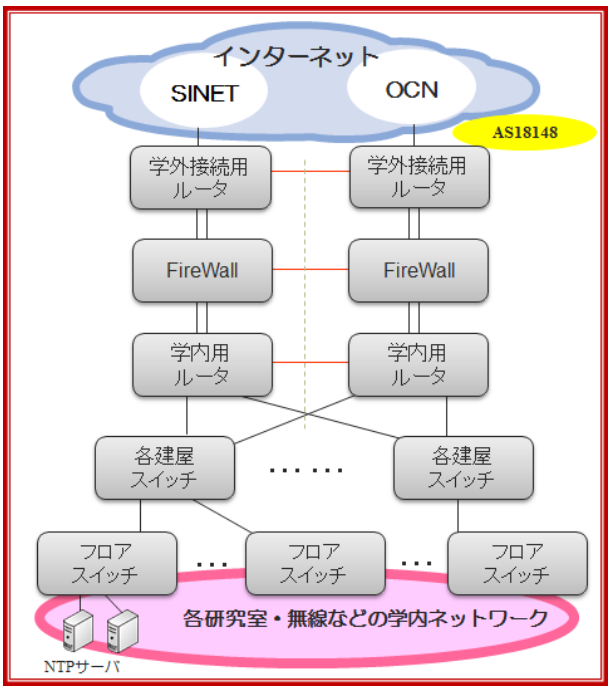

図 1　2015 年 8 月までのネットワーク構成

8 月までは学内ネットワークの末端（一研究室）にて運用してきた。このため、周りのユーザとネットワーク帯域を共有していたことや、徐々にアクセス数が増加して毎時 0 分になると極端にアクセスが集中したことなどからネットワークへの負荷が大きくなっていった。このため、2005 年（平成 17 年）1 月 20 日に負荷分散のお願いを、当時の管理者より”2 ちゃんねる”の”時刻合わせ総合スレッド”を通じて行った経緯がある。この時のアクセス数は、秒間約 900 件であった。なお、2016 年（平成 28 年）7 月現在のネットワーク構成については、第 6.1 節にて述べる。

4.　障害とその対処

先に述べたように 2015 年 8 月までは図 1 の構成にて、公開用 NTP サーバを運用してきた。この間、この NTP サーバが引き金となった障害が数回発生し、学内ネットワークが停止する事象が発生した。本節では、その幾つかの障害事例とその対処法を述べていく。

4.1　計画停電による障害

2012 年（平成 24 年）12 月 23 日（日・祝）に、公開用 NTP サーバーに起因したネットワーク障害が起こった。

同日、公開用 NTP サーバを運用している建物で計画停電があり、以前から行っているとおり、その前日計画停電に備えてサーバをシャットダウンした。だが、翌日の 23 日になると図 1 の FireWall が停止ししてしまった。原因は、NTP クライアントからの大量のリクエストパケットにより、FireWall の有効セッション数を超えてしまったために、FireWall がそのトラフィックを処理することができなくなったからである。本学の学外接続は図 1 からわかるように、SINET と OCN に接続している。大量の NTP リクエストパケットは、OCN 側から流入していることが調査により判明（約 230Mbps）したため、OCN 接続ルータに ACL を設定して、公開用 NTP サーバ宛てのトラフィックを一時的に破棄することにした。こうすることで、FireWall を復旧させることができた。

翌日公開用 NTP サーバの起動を確認後、その ACL を削除した。だが、再度すぐに FireWall が停止した。原因は、大量の NTP リクエストパケットであった。このため、OCN 側の接続ルータに QoS を設定し、10kbps から徐々にこの帯域を広げていくことで、約 2 時間ほどで FireWall を復旧することができた。この時の QoS の設定値は 8Mbps であり、OCN 側からの NTP トラフィックの流入は、この 8Mbps に少し満たないくらいであった。よって、正常なサービス提供ができているものと判断し、この QoS の設定はそのまま残すことにした。

4.2　大量トラフィックによる障害

2014 年（平成 26 年）2 月 14 日（金）に、再び学内ネットワークが停止した。

原因は、SINET 側から大量の NTP のトラフィックにより FireWall の有効セッション数を超えてしまったためにトラフィックを処理することができなくなったからである。この時、OCN 用接続ルータには QoS として 8Mbps を設定していたが、SINET 接続用ルータには QoS を設定していなかった。このため、SINET 接続用ルータにも OCN 接続用ルータと同様に QoS として 8Mbps の設定を行った。これにより、多少遅延はあるもののインターネット接続は行えていたので、しばらく様子を見ることにした。

だが翌日の 15 日には、OCN 接続用ルータがダウンした。一旦再起動してみたが、起動はするものの再びダウンしてしまった。そこで、QoS 設定の 8Mbps を 8kbps に変更するなどを試みたが、この日は OCN 接続用ルータを復旧させることができなかった。この OCN 接続用ルータがダウンした原因は、QoS 設定である。この設定を入れると、大量の NTP のトラフィックを QoS 処理するため CPU 使用率が 100%となり、機能しなくなるのである。よって、この日は OCN 接続用ルータの電源を切った状態で暫定対応することにした。この時、SINET 接続用ルータには 8Mbps の QoS が入っていたが、遅延はあるものの動作はしておりインターネット接続は行えていた状態であった。SINET 接続用ルータが動作していたのは、OCN 接続用ルータに比べて上位機種であったため QoS を CPU 処理できていたか

らである。16 日および 17 日も引き続き様々な対処を行ってみたが、OCN 接続用ルータを復旧させることはできなかった。

18 日になって、一つの解決策を見いだすことができた。OCN と SINET 接続用ルータの QoS 処理は CPU で行っていたため、大量のトラフィックを処理することはできない。だが OCN 接続用ルータの前段に、QoS の処理を CPU ではなくハード処理できる QoS 用スイッチを図 2 のように設置し、このスイッチに QoS 処理をさせて OCN 接続ルータの処理を軽減させることにした。この QoS 用スイッチ

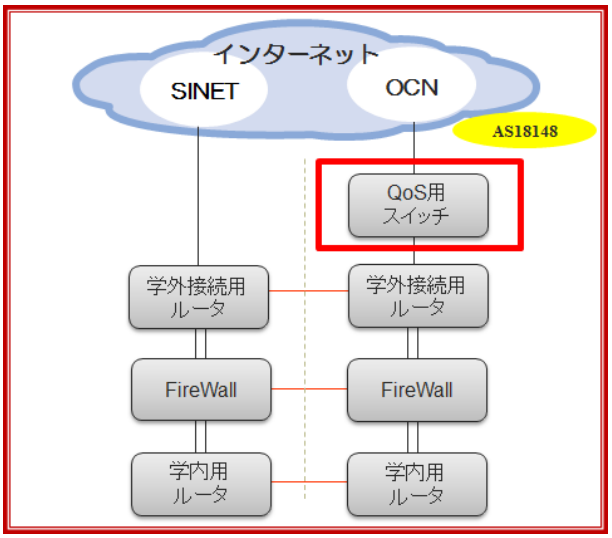

図 2　QoS 用スイッチを導入したネットワーク構成

に QoS の設定を行いその値を徐々に広げていき、SINET 用接続ルータ、OCN 用接続ルータともに正常復旧することができた。よって、この QoS 用 L2 スイッチは正常運用のためには必要不可欠と判断し、このまま設置することにした。また最終的には、SINET 用接続ルータと QoS 用スイッチに QoS 設定として、8Mbps の設定をして運用することにした。

NTP のトラフィックについては今回のように応答を返すことができない場合、大量のリトライが発生することが判明した。また後の調査で、SINET 側からの最大流入は、約 900Mbps であることが判明し、クライアントの数が大幅に増えてきていることも判明した。

なお、根本の原因であるなぜ SINET 側から大量の NTP トラフィックが来たのか、また SINET 接続用ルータに QoS を設定した次の日なぜ OCN 側から大量の NTP トラフィックが来たのか、これらの原因は不明なままである。

5.　2015 年 8 月以降のネットワーク構成

第 4 章での事象をはじめとして、公開用 NTP サーバが原因による学内ネットワーク停止が数回起こっていた。このため、運用を一研究室から総合情報処理センターへと移管した。また先にも述べたように、福岡大学では教育研究システム FUTURE5（The fifth generation FUTURE）を 2015 年 8 月に更改した。これに合わせて公開用 NTP サーバのネットワーク構成も変更し、その構成が図 3 である。ネットワーク構成上末端に設置・運用していた公開用 NTP

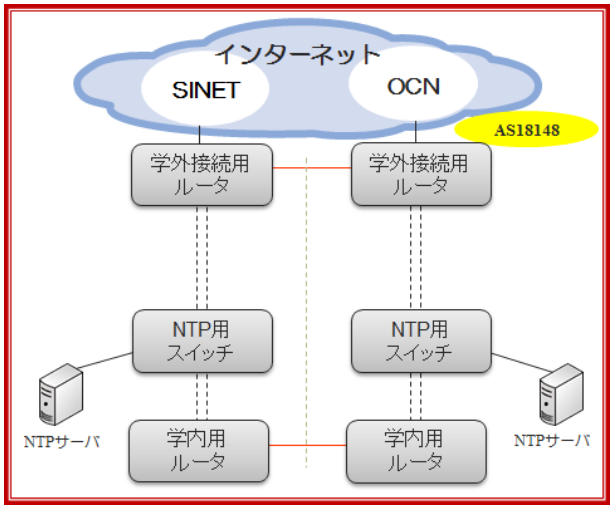

図 3　2015 年 8 月以降のネットワーク構成

サーバを、可能な限りネットワーク的に見て上流に移動した。これにより、耐障害性向上とサービスの安定化を目指すこととした。なおネットワーク構成を変更したことに加えて NTP サーバも更新し、リクエスト応答能力も向上したため、前構成での 8Mbps の QoS 設定は行なわず、基本的に全リクエストに対して応答を返すことができる構成となった。

6.　現状の構成と分析

6.1　公開用 NTP サーバ構成

2016 年 7 月現在の公開用 NTP サーバの構成は、図 4 の通りである。公開用 NTP サーバを学内ネットワークの上流に設置し、教育研究システム（FUTURE5）とは完全に独立させ、可能な限り学内ネットワークへの影響を少なくした。また可用性対策を施して安定したサービス提供を行うとともに、サービス停止時には NTP クライアントから大量の NTP リクエストパケットが発生してしまうため、これを引き起こさないことを一番の目標としてこのような構成をとった。

公開用 NTP サーバは Stratum2 として運用しており、Stratum1 は学内に設置している NTP サーバ（Stratum0 として GPS と CDMA、学内専用）から正確な時刻を取得している。

6.2　トラフィック量

2016 年 7 月現在でのトラフィック量は、図 5 のとおりである。使用帯域としては約 120Mbps、リクエスト数でみる

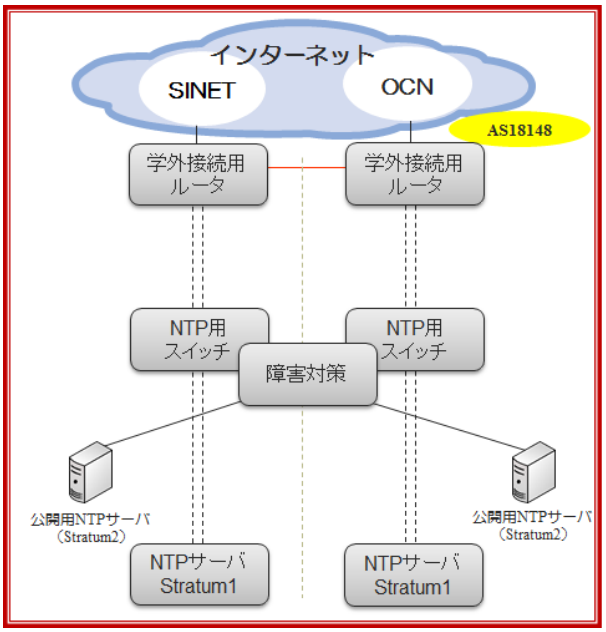

図 4　NTP サーバ構成

と約 150,000 リクエスト/秒にものぼる。2015 年 8 月までの構成では QoS に 8Mbps を設定していたため実際のトラフィック量を確認することができなかったが、図 4 のように構成変更を行い QoS 制限を解除したため、現状のトラフィックを確認することが可能となった。

6.3　サーバチューニング

このように大量の NTP トラフィックを処理しなければならないため、NTP サーバの OS には独自のチューニングを実施した。なお、サーバ OS は現時点で CentOS、カーネルは 2.6 系を使用している。

まずはハイパースレッディングである。ハイパースレッディングは音声や動画のフォーマット変換などを行う場合には処理の高速化などの効果が見込めるが、今回のような大量の NTP トラフィックを処理する場合には意味のない処理である。また、当該の NTP サーバのコア数は 8 であったが、後述のキュー処理で用いられるネットワークカード側のキューも 8 であり、1 コアに 1 キューを割り当てる構成で十分であることから、ハイパースレッデイングは無効にしている。

ネットワークカードについては、受信キューと送信キューの処理を複数の CPU コアに分散させて処理している。受信に関しては Receive Packet Steering（RPS）と Receive Flow Steering（RFS）を、送信に関しては Transmit Packet Steering (XPS) を設定し、8 つの CPU コアに送受信処理の分散を行っている。

C-ステートについては、この機能を BIOS にて OFF にしている。電源管理機能には ACPI（Advanced Configuration and Power Interface）が用いられているが、Intel の Nehalem 以降の CPU ではさらに細かい電源管理を行う機能（C-ステート）が実装された。この C-ステートは、各コアを独立してスリープさせる機能である。この機能を使用した場合、CPU に対して高頻度でハードウェア割り込みをかけるため CPU の負荷があがってしまう。図 6 は、C-ステートの使用を ON にした場合と OFF にした場合の比較で、CPU の各コアの負荷を積み上げたグラフである。ON にした場合には図の左半分であるが、OFF にした場合の右半分と比べて明らかに CPU の負荷が違う。この結果からもわかるように、サーバなど常時処理を行うような場合はコア毎の負荷を観察し必要に応じて C-ステートを OFF にしておくべきである。

6.4　トラフィック量増加に対する分析

公開用 NTP サーバへのトラフィック量は図 5 の通りであるが、もはやここまでのリクエスト数になると、ntp.org のパブリックタイムサーバリストや Web で調べた情報から手入力して、公開用 NTP サーバを使用しているとは考えにくい。よって、アクセス傾向を分析してみた。

図 7 は、ntopng[*5]というソフトウェアを用いて分析した結果で、リクエスト元 IP に対する AS（Autonomous System）ごとの結果である。また、リクエスト元 IP に対する国別の分析は、図 8 である。ただしこれら結果は、複数の NTP サーバのうちその 1 台のトラフィックを分析した結果である。とはいえ、総リクエスト約 150,000 リクエスト/秒からみると、結果と実際はほぼ一致していると考えられる。

このアクセス傾向をみると、全世界の様々な国がアクセス元になっている。ネットワーク的に遠いところ（国）からでは遅延が大きいため、正確な時刻を取得することはできていないと推測される。またリクエスト数については、近年の増加傾向から見るとユーザー側が手動で設定してるのではなく、なんらかの機器の初期設定などに福岡大学の NTP サーバーが使用されていることが推測される。

7.　結論

このように本学の公開用 NTP サーバについてはアクセス数が今でも増加しており、2016 年 7 月現在で約 150,000 リクエスト/秒までになっている。また、アクセス元については様々な国に広がっている。

本学にとってこれ以上のアクセス増（帯域占有）は、他の学内サービスや機器・回線の費用、人的負担などの面から好ましくない。また NTP サーバが停止すると、全世界の NTP クライアントから大量のリクエストパケットが送られ、本学のサービスに何らかの支障をきたすことは、過去の事例からも明らかである。よって、現時点で最大限に優先する事項は、NTP リクエストに対して「正しく返答し

[*5] http://www.ntop.org/

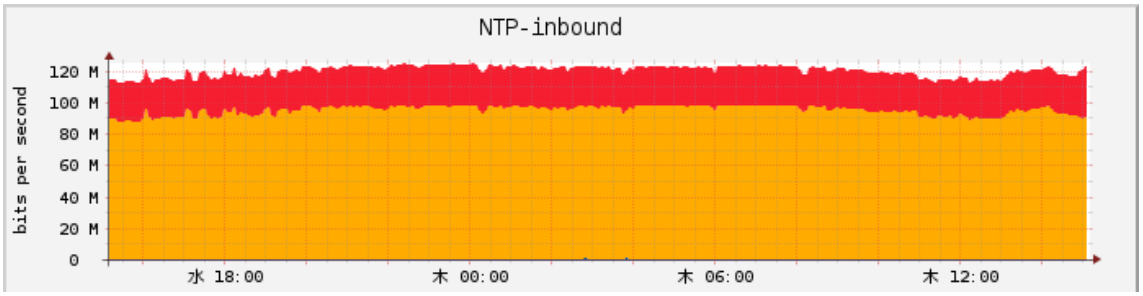

図 5　NTP トラフィック

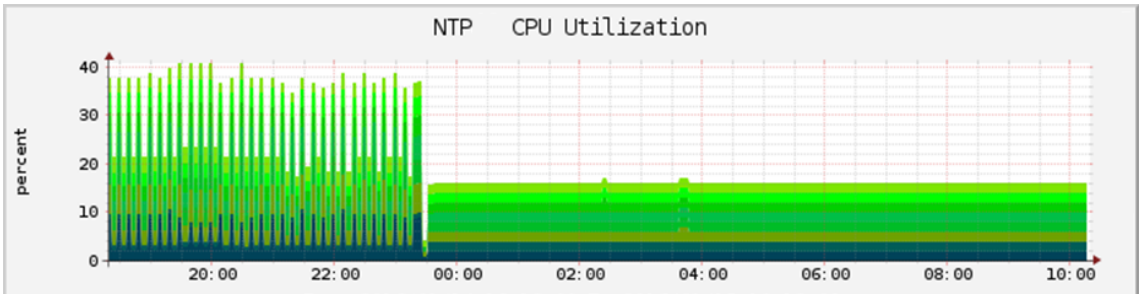

図 6　C-ステート機能の比較

Autonomous Systems

AS number	Hosts	Alerts	Name
56042	6,318	0	China Mobile communications corporation
28573	5,954	0	S.A.
4134	5,168	0	Chinanet
8048	2,917	0	Servicios, Venezuela
4837	2,877	0	CNCGROUP China169 Backbone
9808	2,086	0	Guangdong Mobile Communication Co.Ltd.
27699	1,169	0	TELEFÃNICA BRASIL S.A
18881	1,317	0	Global Village Telecom
10481	1,044	0	Prima S.A.
6830	858	0	Liberty Global Operations B.V.
11830	725	0	Instituto Costarricense de Electricidad y Telecom.
3320	685	0	Deutsche Telekom AG
7303	667	0	Telecom Argentina S.A.
1267	628	0	Wind Telecomunicazioni SpA
8151	590	0	Uninet S.A. de C.V.
22927	525	0	Telefonica de Argentina
3269	522	0	Telecom Italia S.p.a.
8167	521	0	Brasil Telecom S/A - Filial Distrito Federal
7738	513	0	Telemar Norte Leste S.A.
7018	473	0	AT&T Services, Inc.

図 7　アクセス分析（AS 別）

Hosts by Country

Name	Hosts
CN	4,995
BR	2,239
US	657
VE	674
IT	383
ES	365
AR	371
IN	374
DE	244
RU	165
PL	202
CR	236
ID	116
NL	105
VN	99
MX	107
GB	88
PT	69
AE	75
HK	100

図 8　アクセス分析（国別）

続ける」ことである。

現在のところトラフィック量を減らす有効な対策を見いだすことができていない。今後はこのトラフィックの詳細な分析方法を考えてそれを行い、発生原因の追求とトラフィックを減らす活動を行っていきたい。

参考文献

[1] 鶴岡 知昭「楽しかりし年月」、Column 情報の糧、福岡大学総合情報処理センター Web ページ、2008 年 10 月 https://www.ipc.fukuoka-u.ac.jp/column/y2008/m10/