

JANOG51 Meeting会場限り

JANOG51 Meeting

ISPにおけるNTPのありかたを考える

KDDI CORPORATION
MATSUMOTO Akira

Jan 27, 2023

■ 自己紹介

● 名前

- 松本 章(まつもと あきら)

● 所属

- KDDI株式会社 プラットフォームエンジニアリング部

● 仕事

- KDDI事業用のNTPのシステム設計構築・運用支援
- KDDIグループ内におけるNTP構築・運用の技術支援
- DNSサーバの(権威・キャッシュDNS)設計構築・運用支援
 - 業務量としてはDNSの方がメイン

- NTP時刻同期システムのプラクティスの議論
 - 提供者間、利用者間でのより良い構成を議論する。
- 大手ISPにおけるNTP構築・運用例の提示
 - KDDIの事例を一例として提供

■ 今回の発表の内容

● KDDIのNTPの構成の概略と紹介

- 全体構成
- Stratum1
- Stratum2

● 当社で抱えるNTPに関する課題

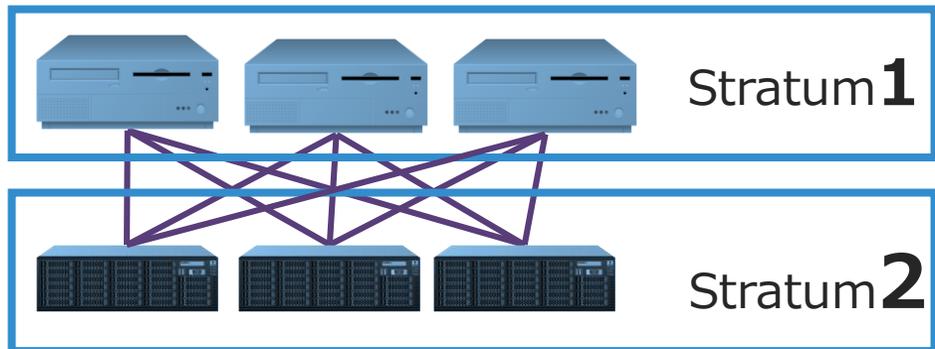
- NTP全体
- Stratum1
- Stratum2

● 考えるNTP構成のありかた

● みなさまへのといかけ

- そもそもISPに(の)NTPサービスって必要ですか？

うるう秒配信系



LIビット配信、うるう秒挿入有

うるう秒非配信系



LIビット非配信、所謂“LeapSmear”実施

※うるう秒非配信システムは社内向けのみとなっております。

■ 構築・運用するNTPサーバ

- Stratum1/Stratum2
 - ・ Stratum2が公開用

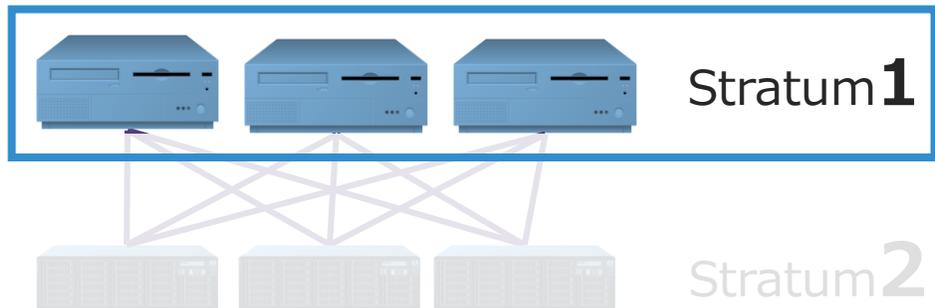
■ 構成概要

- うるう秒の**配信システム**と**非配信システム**の2システムを保持
 - ・ それぞれのシステムは独立
 - ・ Stratum毎に3台ずつ設置
 - ・ 同一Stratum間での同期無し

■ 設置場所

- 関東・関西

うるう秒配信システム



LIビット配信、うるう秒挿入有

うるう秒非配信システム



LIビット非配信、所謂“LeapSmear”実施

■ アプライアンスを使用して運営

● 主時刻源はGNSS

- アンテナを屋上に設置
 - アンテナ系統は2重化

■ 動作モード毎に2系統存在

- うるう秒を配信する系統
- LeapSmearを行う系統
 - 各系統は独立して設置

■ 参照者

- 同一系統同士のStratum2

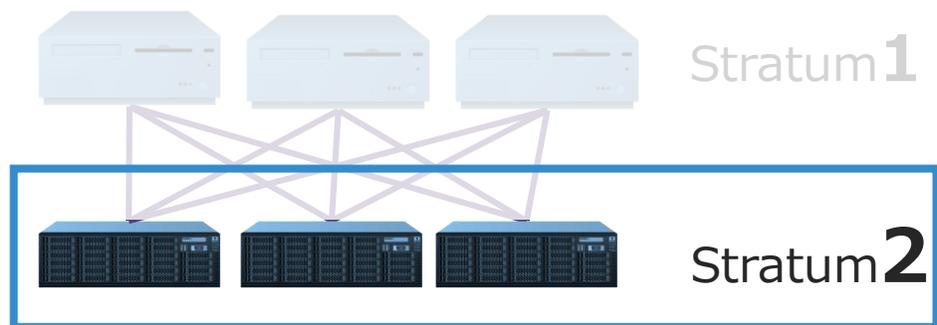
■ 一般的な監視

- 死活監視(Ping)
- ログ文字列監視
 - ログ転送先にてZabbixで実施
- SNMP監視
 - OID取得による監視
 - SNMP Trap監視

■ Stratum1特有の監視

- 現在の同期先
 - GNSSに同期しているか否か
- 現在見えている衛星数
- C/N比
- Ntpdateによる疑似呼監視
 - うるう秒の有無
 - 現在のStratum
 - etc

うるう秒配信システム



LIビット配信、うるう秒挿入有

うるう秒非配信システム



LIビット非配信、所謂“LeapSmear”実施

■ オンプレミスLinuxサーバを使用

- 時刻源は同一系統のStratum1
- ソフトウェアはchronyd
 - ・ コンパイル版

■ 主な参照者

- 法人のお客様
 - ・ うるう秒配信システムのみ
- KDDI事業用システム
 - ・ 両系統とも参照有

■ うるう秒

- 配信も非配信システムも明示的に設定
 - ・ ただし設定は異なる

■ アクセス制限

- nftablesを利用して外部で実施。
- chronycはchronyd側でBINDするアドレスを制限
 - この設定は卒業できない

■ Hardware Timestamp

- 使用する全NICに設定

■ Tracking

- 障害時に備え、tempcompをモニタ項目に追加
 - ・ 記録のみ

■ あとは大体デフォルト値

■ 一般的な監視

- 死活監視(Ping)
- Zabbix Agent監視
 - ログ監視など
- ネットワークトラヒック
- SNMP
 - 主にHW監視

■ Stratum2特有の監視

- Chronyc sources情報監視
 - */+/-の数
 - Reach / Poll / LastRx/ Lastsample
- Chrony tracking情報
 - Stratum
 - LastOffset/root dealy/root dispersion
- ntpdate疑似呼監視
 - 管理サーバからStratum3視点で監視
 - Leap Indicator
 - Precision
 - Root delay

\$ chronyc sourcesの例

MS	Name/IP address	Stratum	Poll	Reach	LastRx	Last sample
Λ+	Stratum1#1-Service IP	1	6	377	22	-79us[-79us] +/- 2767us
Λ*	Stratum1#2-Service IP	1	6	377	3	-12us[-12us] +/- 1556us
Λ-	Stratum1#3-Service IP	1	6	377	29	-72us[-72us] +/- 5836us

会場限り

課題

■ ベストプラクティスが少ない。

● 試験、設定値、動作挙動諸々

- ・ 設置している人は多いはずなのに。

■ 技術者問題

● “ガチでやろうとする”と色々難しい。

- ・ 広範な知識が求められる。
 - 特にStratum1

● 魅力が薄い。

- ・ レガシーサービス
- ・ 参照の引き合いはあるが、人の引き合いは少ない。

■ うるう秒

● 試験されないうるう秒

● Leap Smearという誘惑

■ 利用者の期待が謎

● 大体あってればOK?

■ 時刻はどこから取るのがいいのか？

● NICTじゃダメですか？

■ 構築・運用コストが高額

- 屋上と通信機械室の距離によっては工事費が高額

- 工事費up = 減価償却費up = 固定資産税up

- 売るのは難しい

- 有料でも使います？
 - (むしろいくら出しますか?)

■ 立地問題

- GNSSは高層ビルが必要

- 安定した同期にはほぼ必須
- 自社物件より高いビルが立つリスク
 - やられた

- 屋上も至適な場所があるとは限らない

- 特に古いビル

■ 技術者問題

- 一気通貫で建てられる人がいない。

- サーバ・ネットワーク以外の知識も要求される。
 - 建物、工作物、GNSS、RF、電気等

- レガシー技術

- 魅力を示すのも難しい。

■ 構築や運用のノウハウが少ない

● 理想とされる動作がわからない。

- 例えば…
 - slewrateの理想値は？
 - 上位同期先全喪失のあるべき動作は？

● 適正範囲のノウハウが少ない

- 例えば、Jitter値
 - 少ないほど好ましいのはわかる。
 - でも、どこまで下がってもいいの？

■ ラストオンプレミス？

● 仮想化は現実的なのか？

● 白眼視

- 仮想化しないんですか？
- えっ？今時箱(物理サーバ)必要なんすか？
- 今時オンプレミスとか冗談ですよ？
- 仮想化してないのお宅だけですよ？
 - 全部実際に言われた内容です

■ 技術者問題

● レガシー技術

- 魅力的ではない

会場限り

あり方の提案

■ 自分でStratum1を持つべきか？

● Stratum1を自社で提供しないならNICTを参照すればいいのでは？

- 自社制御下に置くのが必須でなければ、他所のNTPを見る方が断然安価。
- NICTが止まったら？
 - そもそもGNSSも確実ではない。
 - 殊更にNICTへの依存がハイリスクとは言えないのでは？

● 高精度？ 本当？

- そもそも本当に精度が必要ならNTPよりPTPの方がいいのでは？

■ Stratum1時刻源のありかた

● 自社内でStratum1が必要なら下記が理想的？

- マルチGNSS
 - GPS+QZSS or GPS+Galileo等
- 時刻源 + 高精度ホールドオーバークロック
 - (GNSS or PTP) + (光テレホンJJY or 高精度Holdoverクロック)

■ 同期先のありかた

- 少なくとも物理的に異なるマシン3台に同期を。
- 可能なら複数サイトの時刻源を。

■ Symmetric(同じStratum同士の同期)は必要？

- ネットワークが死んだら意味がない。
 - ・ 横にあるなら不要では？

■ ベアメタ/オンプレミスは必要か？

- chronyd使えばVMでも差し支えないのでは？
- PTP+VM Host等も選択肢もあるのでは？

■ サービス提供のありかた

● 可能ならFQDN、少なくとも持ち運べるIPで。

- 一度公開すると公開先の制御は不可能。
- 無断参照されてても無断で切ると叩かれる。
 - 知らんがな…

● 公開しているサーバは全部参照してもらう。

- 可能ならpoolで参照してもらえるようにする。
- どれがプライマリかは明示しない。
 - Active・Standby型の参照もある。
- 全部参照してもらう前提で運用を明示
- “天然由来”のサービスを強調したい。
 - 時刻源は自分たちの制御下に無い(ことが多い)

■ 何故か試験されないうるう秒

● そもそもうるう秒の試験をすべき。

- うるう年の試験をしていなかったらシステム担当者はタコ殴り。
 - しかしうるう秒はなぜかNTP担当が叩かれる
 - 過去12年で3回入っている = 廃止予定の12年後(2035)までに2~3回来る可能性がある。

■ LeapSmearは本来あるべきではない

● LeapSmearが許容できるなら自走すべき

- LeapSmear = 実施中の時刻は不正確。
 - 時刻を許容出来るなら別に自走させとけばいいのでは？

● ワークアラウンドでは？

- 本来は試験して備えるべき。
 - 常習すべきではない。
 - 弊社でも作ったが新規参照希望者がちよくちよくいて(担当としては)困惑。

■ Stratum1はどうあるべきだと思いますか？

● 時刻源

- 精度・耐障害性(セキュリティ)・コストの両立
 - GNSSもJJYも完全無欠ではない

● 台数

- 3台が良いと言われるが本当に必要な台数は？
- 自社で全部賄うのか否か？

■ Stratum2より下のStratumってどうあるべきだと思いますか？

● 時刻源

- 上位Stratumを何台見るのがいいのか？
- Stratum2でもholdoverは必要ではないか？

● サーバ&ソフトウェア

- 物理サーバって必要？
 - CPUをPinningしてchrony使えば仮想環境でもいいのでは？
- Chrony? ntpd?

● Symmetric同期って必要ですか？

■ 監視

● 実はみんな精度なんてたいして気にしてない？

- 何だか知らんが大体同期出来てればヨシ!?

■ うるう秒

● なんでうるうびょうのしけんをしていないんですか？

● 「2035年までに廃止」

- 1回ぐらいは入るかも。

■ そもそもISPに(の)NTPサービスって必要ですか？

● 公開NTPじゃダメなんですか？

- NICTさんの方が優秀だと思いませんか？

■ 免責事項

- **記載事項について正確性や再現性を保証するものではありません。**
 - ・ 構成図はいずれもシンプル化しています。
- **特定の構成、製造元、組織等の利用等を推奨するものではありません。**

■ 著作権表示

- **スライド内の画像の一部にAdobe Stockを使用しています。**
 - ・ 当該画像の著作権は、当該画像の著作者が保有します。下記の画像を使用しています。
 - Ovidiu – stock.adobe.com
 - Old vintage computer vector illustration
 - Vadim
 - The 1u server for mounting into a 19 inch rack with six 2.5 inch hard drives and an optical drive. Black on a white background. Vector illustration.

「つなぐチカラ」を進化させ、
誰もが思いを実現できる社会をつくる。

KDDI VISION 2030

