

JP DNSの裏側

JP DNSの更新間隔短縮について

民田雅人 <minmin@jprs.co.jp>

株式会社日本レジストリサービス

2006年7月14日

JANOG18@パナソニックセンター東京

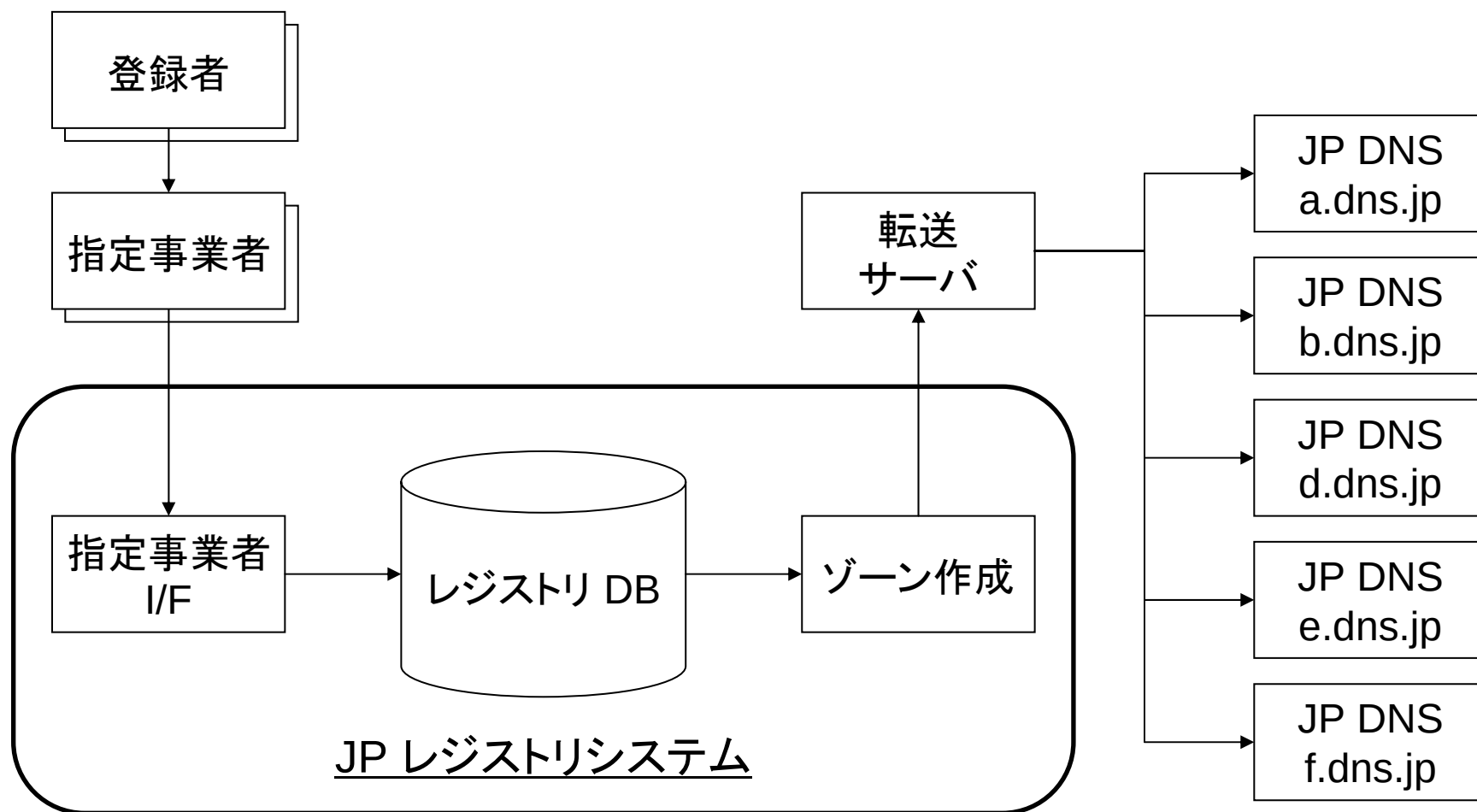
JP DNSの更新間隔短縮

- 2006年4月3日(月)
JP DNSの更新間隔が、
1日1回から15分毎に1回への大幅短縮
- 大幅なサービス性の向上
 - 「申し訳ありません。ネームサーバの設定間違えました。まる1日アクセスできなくなりそうです」といったことが無くなった(はず?)
- さて、舞台裏では何が?

更新間隔短縮の背景

- ドメイン名利用形態の変化
 - 会社、組織に加え、個人への普及
- 指定事業者の要求サービスレベルの変化
 - 即時性を必要とするサービスの提供
 - Webで登録⇒即時利用開始
- ドメイン利用ユーザの利便性
 - データ変更の即時反映
 - 登録ドメイン名の即時利用開始

JPレジストリとJP DNSの構成



旧環境のJP DNSの更新

- 1日1回属性型と汎用型を別々にゾーン作成
 - 01:00 属性型・地域型のデータ抽出
 - 03:00 汎用型のデータ抽出
 - 05:00 ゾーンファイルの作成
 - 05:15 頃 JP DNS にデータ転送開始
- 毎朝5時30分頃、JP DNSのゾーン更新完了
- トータルで4時間30分以上
 - 1日1回の更新を前提にしたシステム構成

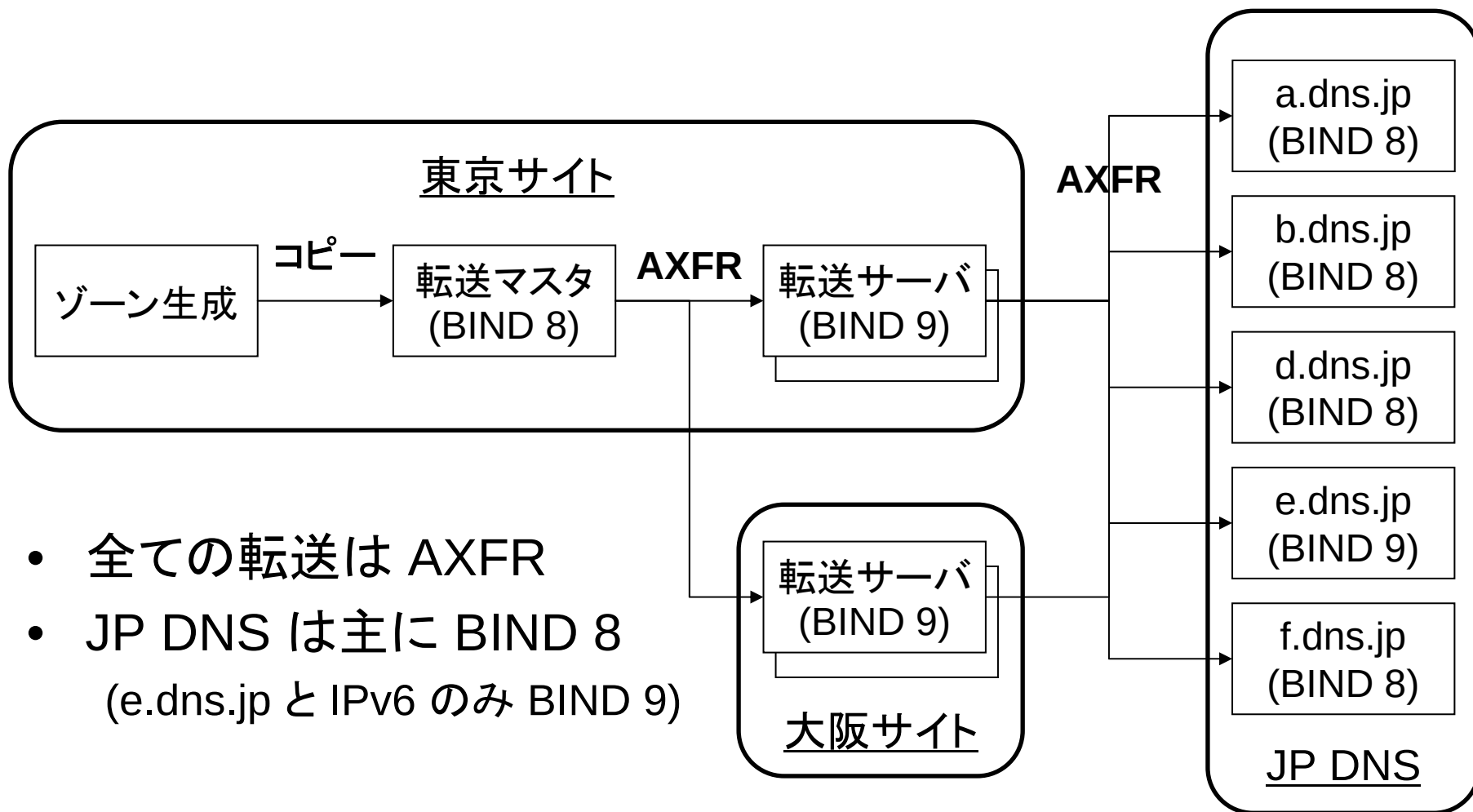
更新間隔短縮に向けた 2つのパートの分析

- JPレジストリ側の処理時間 約4時間
 - 空き時間(cronの処理間隔等)が結構あり、単純に縮めると1時間半程度 ⇒ まだ不十分
- DNSサーバ側の処理時間 約30分
 - ほとんどがBINDのリロードとゾーン転送
⇒ 単純な高速化はあまり望めない
- ゾーン転送の全転送(AXFR)による更新を
差分転送(IXFR)による更新へ

IXFRのための差分情報

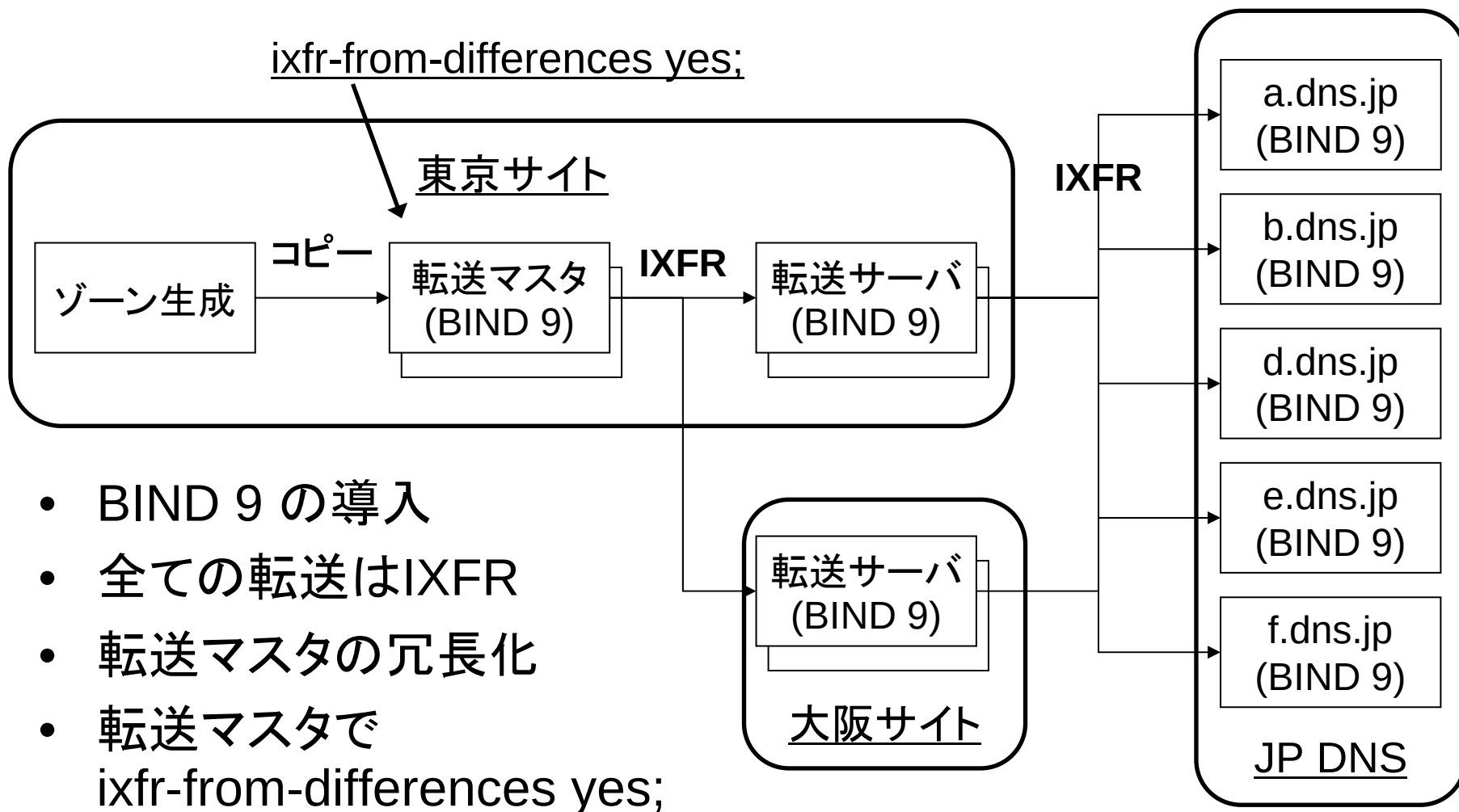
- DB側で生成しDNSサーバへ送る
 - レジストリシステムに大きな修正が必要になり新たなトラブルの可能性が否めない
 - 実施までの期間が長期化する可能性
- ixfr-from-differences(BIND 9.3.0～)で生成
 - ゾーンファイルのリロード時に差分情報を生成
 - 既存システムとの親和性がよい
 - サービス決断から実施までの期間の短縮化

JP DNSのデータ更新 (旧)



- 全ての転送は AXFR
- JP DNS は主に BIND 8
(e.dns.jp と IPv6 のみ BIND 9)

JP DNSのデータ更新 (新)



JPレジストリ側の高速化

- 徹底した処理の見直し
 - 各処理の高速化と無駄な待ち時間の削減
 - 逐次処理のうち並列可能な部分を並列化
- 日本語JPナビ用のRRの削減
 - 旧: 4～12RR ⇒ 新: 2RR
 - <http://日本語JPナビ.jp/> 参照
- サーバ機材の一斉更新による処理能力向上
 - 実は、これが一番効果が大いい😊
 - 機材の更新は各DNSサーバでも行われた

初めてのIXFRは不安がいっぱい

- 実験環境でのIXFRの基本テスト
 - BIND 9のIXFRの挙動と性能の確認
 - ジャーナルファイルの生成、更新時の挙動
- 2ndary組織と実環境に近い形でテスト
 - IXFRでの高頻度の更新
 - 継続的なIXFR後のデータ整合性の確認
 - AXFRとの併用運用技術の確立
 - ⇒万が一のときにAXFR

ゾーンデータ更新スケジュール

- 15分毎にプログラムを起動
 - 1回あたりの作業時間が15分を超える場合、次のデータは作成しない方針
- AM3:00 – 5:00 は15分毎の更新を停止
 - メンテナンス用の時間枠の確保
 - AXFRを使いデータの整合性確認を実施
 - 明示的にAXFRを実施して差分情報をリセット

BIND 8 vs BIND 9

- JP DNSが、co.jpとne.jpのゾーンを保持する場合に、www.example.co.jpのAを問合せる

```
$ORIGIN co.jp.
```

```
example.co.jp.      NS ns.example.ne.jp.
```

```
$ORIGIN ne.jp.
```

```
example.ne.jp.      NS ns.example.ne.jp.
```

```
ns.example.ne.jp.  A  10.10.10.10
```

- BIND 8はグルー有り、BIND 9はグルー無し
 - 仕様の差異であり、不具合ではない
 - NSDはBIND 9と同様・tinydnsはBIND 8と同様

多段グルーレス問題(JPRS用語)

- BIND 8.2.7(BIND 4を含む)までの不具合
 - BIND 8.2.7までのキャッシュサーバでは、あるドメイン名の検索中に、グルーが2回連続して得られない場合、当該ドメインの検索が不能となる
- JPドメイン名のゾーン構成
 - 汎用JP、属性型、地域型でドメイン名の種類毎にゾーンを構成し、全部で64ゾーン
 - BIND 9に単純に移行すると、古いBINDで検索不能ドメインが4万以上発生する

差分更新とゾーンマージ

- 安定したIXFRにはBIND 9が必須
- 多段グルーレス問題の回避も必須
 - 登録者への呼掛けだけでは対応不可
 - BIND 9の改造 → ×
 - 応答するNSに細工して現象を回避 → ×
- JPドメイン名のゾーン構成を
 - JPのみの1つのゾーンに集約 → ○
 - ゾーンマージ(これもJPRS用語?)

ゾーンマージの問題と検証

- ゾーンマージにより起こる変化
 - Authority Section の情報が変わる
 - 個別のゾーンの NS が存在しなくなる
 - 移行時にキャッシュと異なるデータになる
 - 移行時にゾーンマージと旧ゾーンが併設される
- キャッシュサーバの挙動を理論的に検証
- テスト環境を構築し、実データによる検証
 - BIND 8 と BIND 9 の挙動確認 ⇒ OK 😊

JP DNSの変更点(その他)

JANOG17「.JP Update」参照

- JPゾーンのSOAのシリアル値を
UNIX epochからの秒数へ変更
 - 以前はYYYYMMDDNNの10桁
 - "date -r <SOAシリアル>" で見やすくなります
- JPゾーンのSOAのminimum値を900秒へ
 - ネガティブキャッシュの値(以前は86400秒=1日)
 - 15分毎の更新で、最悪は30分の遅延になる

BIND 9化 + ゾーンマージ

- 2006年3月22日(水) 実施
 - 各2ndaryを1時間ずつずらして
 - 冗長化構成を利用し、断を極力避ける
 - トラフィックの監視をしながら
- 移行は問題なく完了
 - 利用者からの問合せなし
 - トラフィックの異常なし
- JP DNS史上最大の変更は無事成功！

4月3日以降のJPドメイン名運用

- ネームサーバの登録がすぐに確認できる
- 情報の変更がすぐに確認できる
 - 誤った情報もすぐに修正できる
- 実際の変更にはキャッシュが関係する
 - ネガティブキャッシュの900秒の影響
 - JPの各RRのTTL値は86400秒(=1日)
- 更新所用時間は概ね3分以内
- 運用上大きなトラブルは皆無

BIND 9での差分更新運用TIPS

- デフォルトでIXFRを行う
 - 通常、差分情報が無いためAXFRになる
- ixfr-from-differencesは**マスタ側でのみyes**
 - スレーブ側でyesにすると、差分情報を作ろうとして、転送がAXFRになる

Q and A

