JP DNS ソフトウェアダイバーシティ 導入その後＋α

2025年11月26日 Internet Week 2025

DNSOPS.JP BoF

株式会社日本レジストリサービス(JPRS)

池田和樹

自己紹介

名前

- 池田　和樹

所属

- 株式会社日本レジストリサービス(JPRS)

出身

- 和歌山県

経歴

- 2018年 JPRS新卒入社 システム部に配属
- 2021年～2024年大阪へ、2025年から東京へ戻ってきました

業務内容

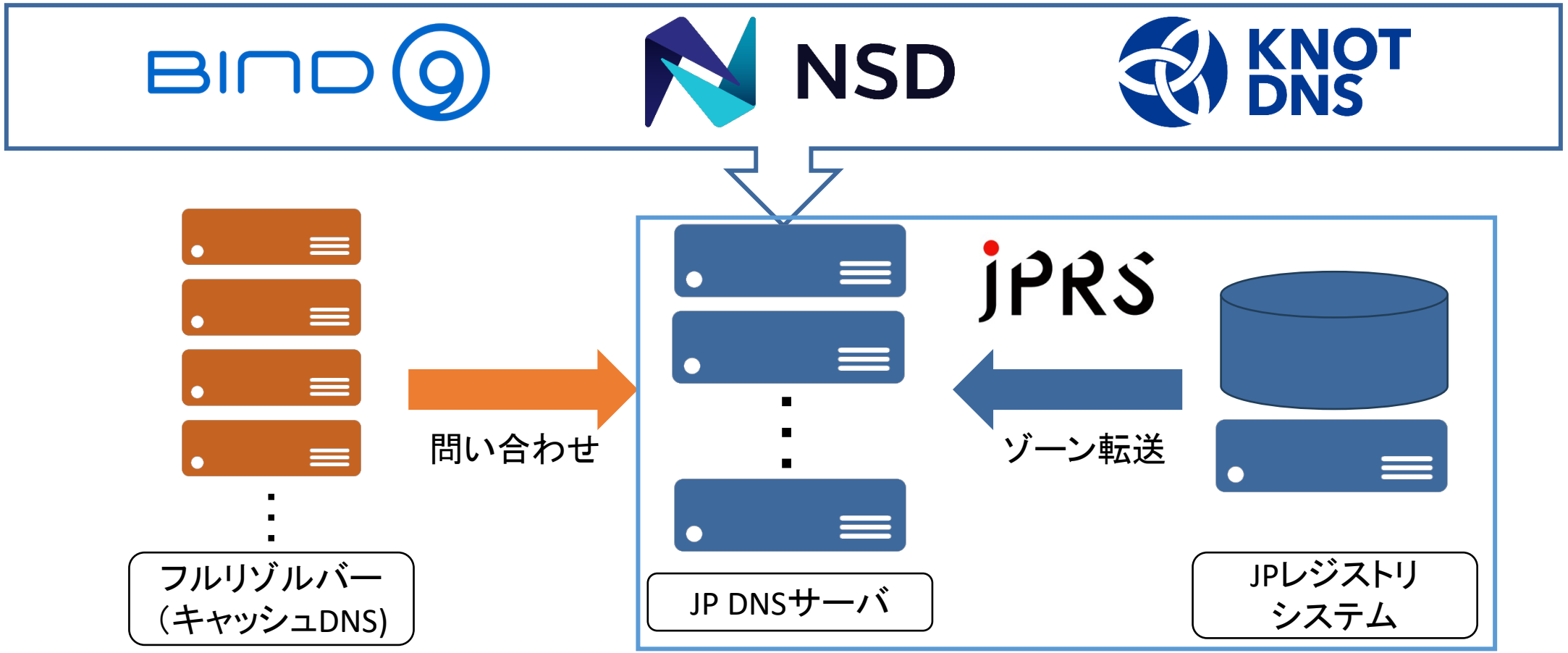
- JP DNS/JP レジストリシステム等のサーバ・ネットワークの管理・運用
- 対外系はDNSOPS.JP 事務局

目次

- （振り返り）DNSソフトウェアダイバーシティ確保
- （振り返り）導入時の評価結果の一覧
- DNSソフトウェアダイバーシティ導入の効果
- DNSソフトウェアダイバーシティ導入時のトラブルや課題例
- BIND 9.20について
- BIND 9.20を含めたDNSソフトウェア性能評価
- さいごに

（振り返り）DNSソフトウェアダイバーシティの確保

■ 2024年、安定運用を目的としてJP DNSのソフトウェアを複数種類に
  - ソフトウェアに脆弱性が発見された際のリスク分散



（振り返り）導入時の評価結果の一覧

#	評価項目	確認方法	評価ポイント	NSDの結果	KnotDNSの結果
1	ソフトウェア機能評価	JP DNSに求められる機能を有しているか確認する	社内で定めている機能評価項目の一覧表を用いて確認し、問題点がないこと	○	○ ゾーン転送用の鍵名がログに出力されなかったため、開発元に依頼し修正してもらった
2	ゾーン転送評価	JPゾーンをロードしたBIND 9, NSD, Knot DNSを準備し、ゾーン転送時間を測定する	社内で定めているゾーン転送のサービスレベル目標を満たすこと	○	○
3	クエリ応答内容評価	BIND 9, NSD, Knot DNSの応答内容を比較・評価する	JPドメイン名の名前解決の支障とならない応答を返すこと	○	○
4	クエリ応答性能評価	性能評価ツールで測定※	BIND 9と同等以上の応答性能を有すること	○ BIND 9の約4倍の性能を有することを確認	○ BIND 9の約4倍の性能を有することを確認
5	運用手順への影響	定型業務の手順やドキュメントを準備し、実際にオペレーションを実施する	ソフトウェアのバージョンアップ手順等を準備し、問題なく動作すること	○	○

※ DNS Summer Day 2024　DNSソフトウェアのパフォーマンステストをしてみた
（ https://www.dnsops.jp/event/20240621/20240621_abe.pdf ）

導入から1年が経ちました

JPRS DNS関連技術情報ページ(https://jprs.jp/tech/)

セキュリティ情報

- 2025-10-27 Unboundの脆弱性情報が公開されました(CVE-2025-11411)New!
- 2025-10-27 PowerDNS Recursorの脆弱性情報が公開されました(CVE-2025-59023、CVE-2025-59024)New!
- 2025-10-23 (緊急)BIND 9.xの脆弱性(DNSキャッシュポイズニングの危険性)について (CVE-2025-40778) - バージョンアップを強く推奨 -New!
- 2025-10-23 (緊急)BIND 9.xの脆弱性(DNSキャッシュポイズニングの成功確率向上)について (CVE-2025-40780) - バージョンアップを強く推奨 -New!
- 2025-10-23 (緊急)BIND 9.xの脆弱性(過剰なCPU負荷の誘発)について(CVE-2025-8677) - バージョンアップを強く推奨 -New!
- 2025-07-24 PowerDNS Recursorの脆弱性情報が公開されました(CVE-2025-30192)
- 2025-07-23 Knot Resolverの脆弱性情報が公開されました
- 2025-07-18 Unboundの脆弱性情報が公開されました(CVE-2025-5994)
- 2025-07-17 BIND 9.20.xの脆弱性(DNSサービスの停止)について(CVE-2025-40777) - バージョンアップを強く推奨 -
- 2025-05-22 (緊急)BIND 9.20.xの脆弱性(DNSサービスの停止)について(CVE-2025-40775) - フルリゾルバー(キャッシュDNSサーバー)／権威DNSサーバーの双方が対象、バージョンアップを強く推奨 -
- 2025-04-28 Knot Resolverの脆弱性情報が公開されました
- 2025-04-10 PowerDNS Recursorの脆弱性情報が公開されました(CVE-2025-30195)
- 2025-03-17 Windowsドメインネームサービスの脆弱性情報が公開されました(CVE-2025-24064)
- 2025-01-30 (緊急)BIND 9.xの脆弱性(過剰なCPU負荷の誘発)について(CVE-2024-11187) - バージョンアップを強く推奨 -
- 2025-01-30 (緊急)BIND 9.xの脆弱性(パフォーマンスの低下)について(CVE-2024-12705) - バージョンアップを強く推奨 -
- 2025-01-21 サービス終了後に残っているDNS設定を利用したサブドメインの乗っ取りについて
- 2025-01-21 終わったWebサイトのDNS設定、そのままになっていませんか？(パンフレット[PDF])
- 2024-11-15 Windows DNSの脆弱性情報が公開されました(CVE-2024-43450)

われわれの安眠は？

改めて導入の効果はどうよ（その1）

■ 稼働実績とダイバーシティの実効性

- 結論：1年間安定して稼働し、耐障害性は向上
  - ➔権威DNSサーバーはそもそも分散配置されていることも有効に機能
  - ➔脆弱性による不安感はかなり低減
  - ➔Knot DNSで非計画停止が1件（2024年 DNSOPS.JP BoFでも触れた不具合、後述）
- BIND 9で脆弱性修正リリースが何件かあり
  - ➔対象がフルリゾルバー、9.20系のものが多かった
- サーバのリソース消費だとNSDとKnot DNSではNSDの方がより省エネ
  - ➔RRLをかけていることもあり、そもそもサーバが火を噴くことは…

導入直後に踏んだKnot DNSの不具合

■ まだ監視設定そのものをチューニングしていた最中、それは発生した

全台Core dump 吐いとる！

導入直後に踏んだKnot DNSの不具合

■ 兎にも角にもサポート(CZ.NIC)に問い合わせ
- CZ.NICの有償サポートは事前に契約済み
- サポートでも再現できず、被疑箇所にデバッグパッチを当てて様子見

➔不具合再現せず、3カ月ほどが経過

■ 3か月後、やっとデバッグログが出力された
- 手元の環境でもログ出力のタイミングのゾーンデータ・クエリで再現

➔原因はNSEC3周りの動作不具合と判明、パケットを投げつけるとコロリ
- サポートでも再現、無事修正される

➔v3.4.3/v.3.3.10にて修正

改めて導入の効果はどうよ（その2）

■ 運用負荷や効率化

- 複数ソフトウェアのサポート契約のコスト・リリース情報調査の負荷が増加
  - ➔2.5倍程、コストや工数が増えた
- ソフトウェア以外は基本的に共通手順、手順書作成の負荷は1.5倍ぐらい
  - ➔バージョンアップ、設定変更など…
  - ➔可能な限りAnsibleで構成差分を吸収
    - ➢ログの出力先
    - ➢設定ファイルのテンプレート化
- 監視負荷は大きく変わり無し
  - ➔外形監視はこれまで通り（SOA Check, 応答時間など）
  - ➔ソフトウェア毎の監視はPrometheusで吸収

運用も落ち着いてはきましたが、導入前後でいくつかの課題があり

導入前・導入後の課題例

#	ソフトウェア	課題	対応
1	共通	BIND 9 viewに依存した実装の代替	IPアドレスとプロセスを分離する形の構成に
2	NSD	公式、あるいはコミュニティでメンテナンスされているExporterがない	4.12.0からNSD自身でmetricsを出力できるようになった
3	共通	クエリログの代替	dnstap経由で統一したフォーマットでロギング
4	共通	クエリログサイズの肥大化	ログ圧縮方式の変更

実装・運用上の課題もいくつか発生しましたが、とりあえずは運用に載せられた（かなと）

課題の詳細#1

BIND 9 view機能

```
view view_name [ class ] {
    match-clients
{ address_match_list } ;
    match-destinations
{ address_match_list } ;
    match-recursive-only <boolean> ;
  [ view_option ; ... ]
  [ zone_statement ; ... ]
} ;
```

送信元IPアドレス、宛先IPアドレス等で、
split DNSを実装できる機能

- NSDでは類似機能は無いため、ロードバランサーを実装し、前段で振り分け先をコントロールする構成にするか、他の何らかの方法で実装する必要があった
- 色々検討した結果、サーバへ複数IPアドレスを設定し、各ソフトウェアがListenするIPアドレスを別にした状態で起動させることに
  ➡ 力技ではあるが、シンプルな設定で実装することとした

課題の詳細#2

動作状況を確認出来るnsd-controlコマンド

```
$ nsd-control -c config stats_noreset
server0.queries=10
num.queries=10
～
num.type.A=5
num.type.NS=5
～
zone.primary=0
zone.secondary=1
```

読み込んでいるゾーン数、クエリタイプ等、
一通りのステータスが確認可能

metricsを出力する機能が内蔵されていないので、
個人開発のnsd-controlをラッパーしたexporterを利用

NSD 4.12.0

```
nsd_queries_total{server="0"} 10
～
nsd_queries_by_type_total{type="A"} 5
nsd_queries_by_type_total{type="NS"} 5
～
nsd_zones_primary 0
nsd_zones_secondary 1
```

NSD自身がmetricsを出力

課題の詳細#3,4

■ dnstapとDNS-collectorにより新クエリログ基盤を構築
- 検討段階であった2024年上期にJANOGでdnstapに関するプログラム※にて紹介
- 複数のlogger機能があり、ログ解析基盤との連携機能も考え採用

➔インプット：BIND 9, Knot DNS, NSD

➔アウトプット：テキストデータ（JSON形式）のクエリログ

■ 現在も安定運用しているが、データが詳細化されたことによりファイルサイズが肥大化
- BIND 9のクエリログと比較して、5倍程のサイズに

➔データ圧縮方式をZstandardに変更（圧縮後は約1/20のサイズ）

➔今後ログデータ用DBへの移行を検討中

※変化するDNS運用とこれからの課題について(DNS設計/運用者の目線から) LINEヤフー株式会社 太田さんパート
https://www.janog.gr.jp/meeting/janog53/dnsops/

安定した運用を進める中でも、次の世代(BIND 9.20)はやってくるもので…

BIND 9の今

■ 2026年Q2にBIND 9.18がEOL予定※

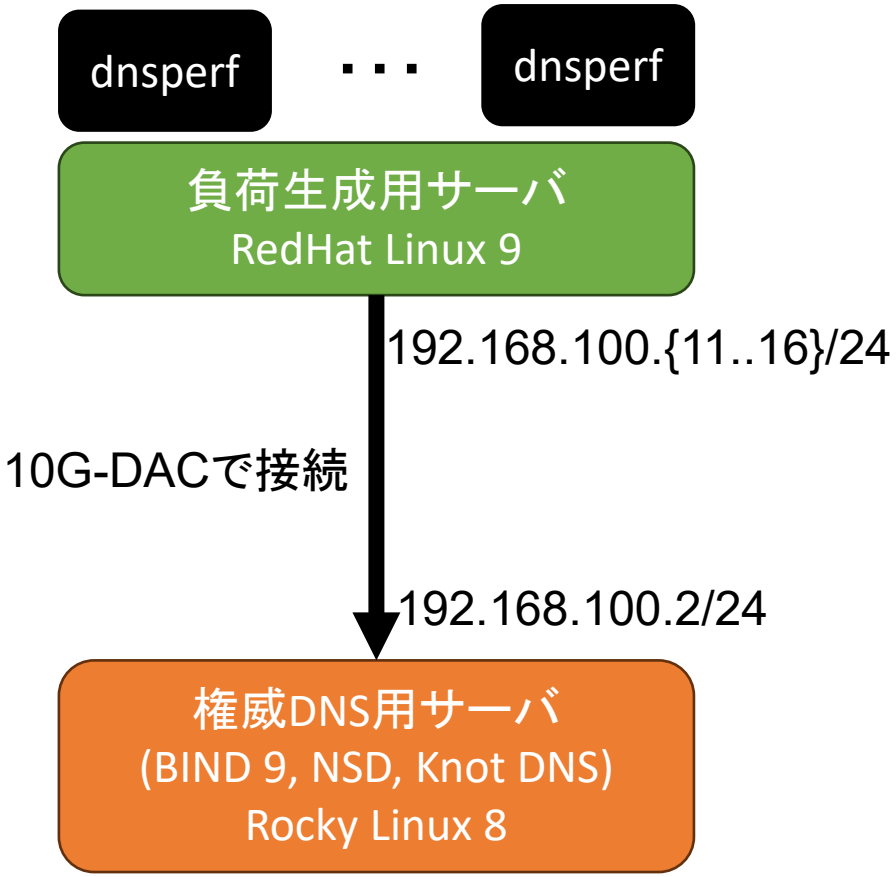
- BIND 9.20のESV宣言はまだではあるが、評価も進めないとなあ…
- BIND 9.20でピックアップされている特徴
  - libuvへの完全移行
  - 専用スレッドプールの利用による長時間タスク(ゾーン転送、DNSSEC検証時)のレイテンシ改善
  - バックエンドDBを赤黒木からQP trieに置き換え。Userspace RCUが必須となり、メモリ解放機構をQSBR(Quiescent State Based Reclamation)に置き換え
  - 圧縮アルゴリズムの改善

■ どれぐらいのパフォーマンスが出るか確かめよう

- ついでに、NSD, Knot DNSの最新版も
- 宿題となっていた実環境のクエリを模して実施

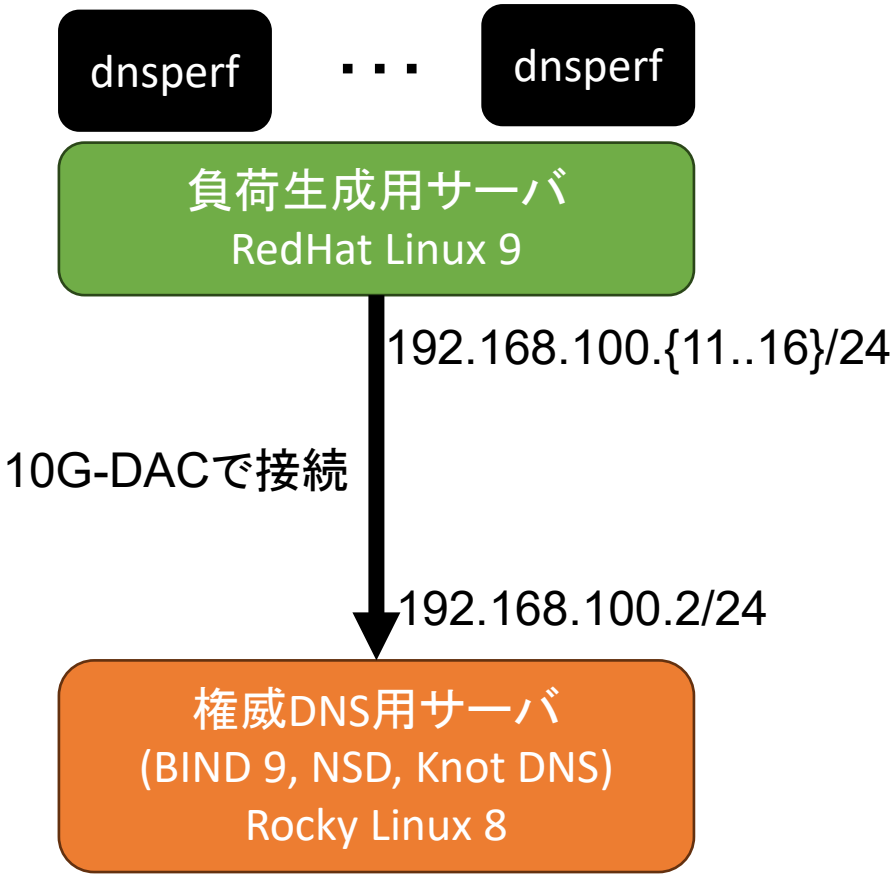
※ ISC's Software Support Policy and Version Numbering (https://kb.isc.org/docs/aa-00896)

測定環境（その1）



- 機材情報
  - CPU: Intel Gold 5315Y@ 3.20GHz
  - メモリ: 64GB
  - NIC: 10Gbps(Broadcom BCM57412)
- ソフトウェア系
  - BIND 9.18.28, 9.20.13
  - NSD 4.13.0
  - Knot DNS 3.5.1
  - dnsperf 2.12.0
  - DNS-collector 1.4.0
- 権威DNS側の情報
  - ゾーン数 1
  - DNSSEC署名済み
  - RR数 500万弱
  - tcp-client系 8192
  - スレッド数 デフォルト&CPUに合わせた16
  - reuseport on (NSDはデフォルトoff)

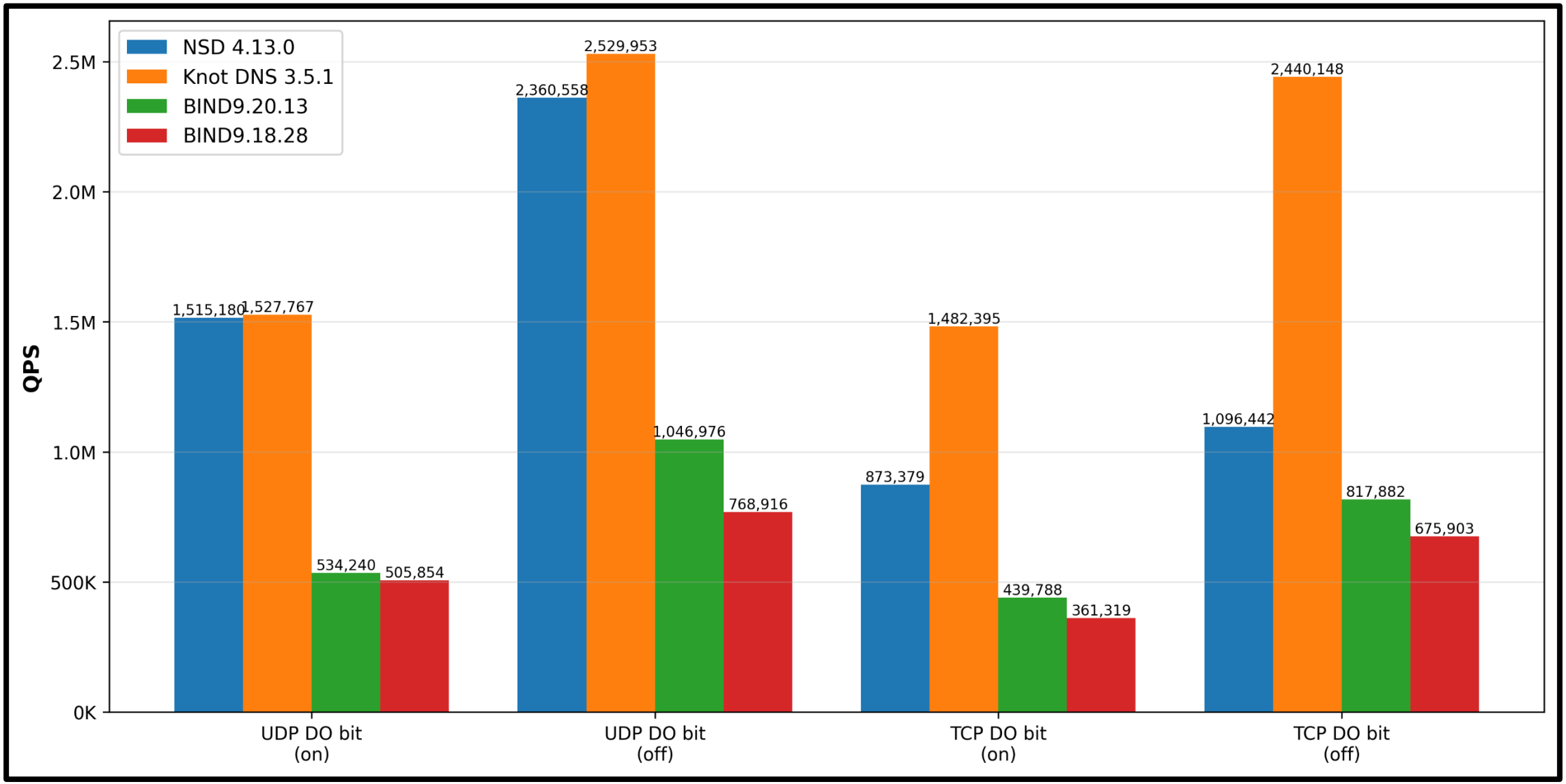
測定環境（その2）



- 負荷生成情報
  - クエリデータは実環境のデータをベースに生成
    - NOERROR　58%
    - NXDOMAIN　38%
    - REFUSED　4%
  - 6プロセスから並列実行
  - DO(DNSSEC OK) bitはonとoff両方
  - UDP(TCP fallback無し)とTCPで計測

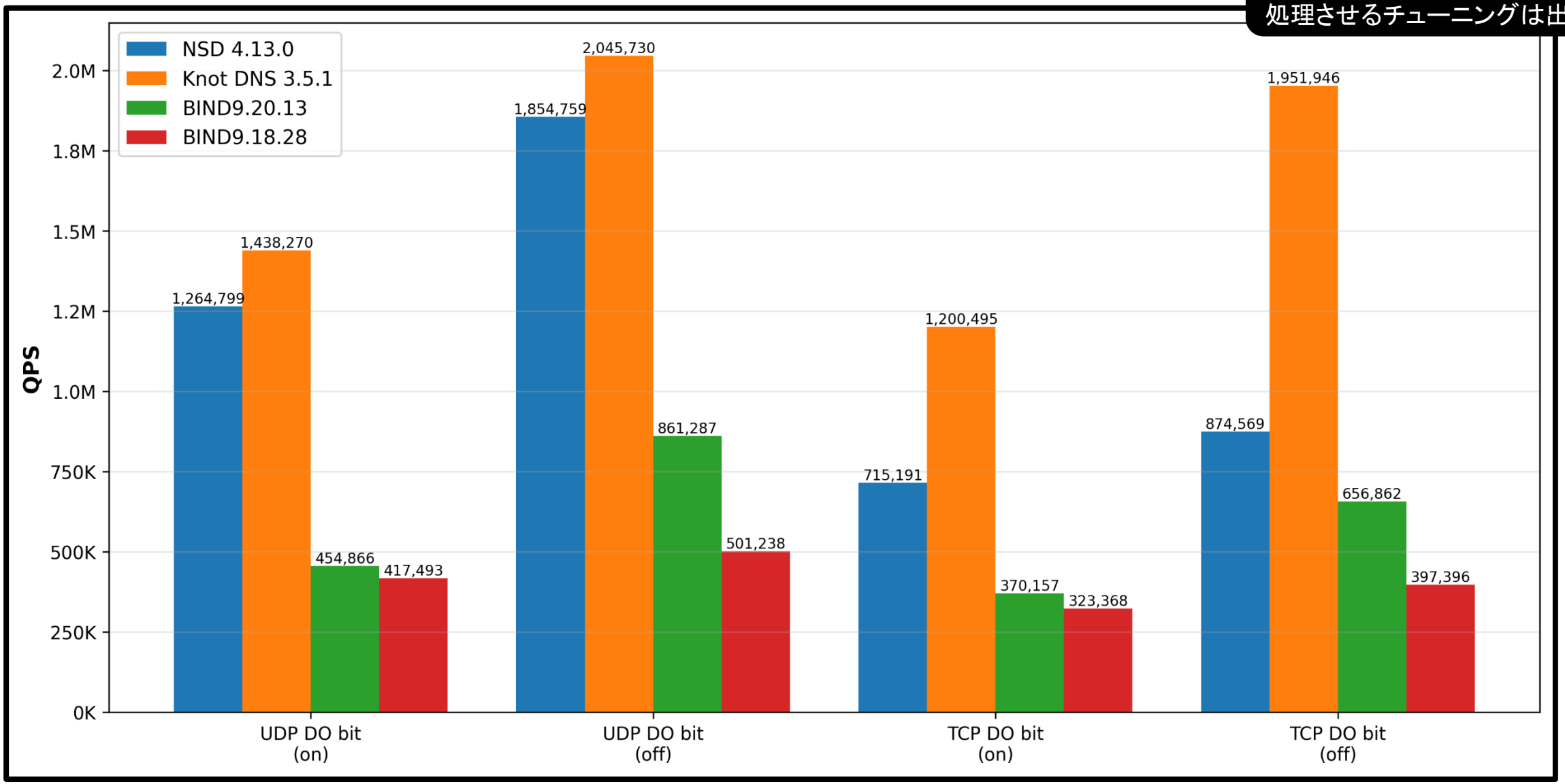
- パラメータ系
  - チューニングの余地しかないのでKnot DNSのドキュメントを参考に
    - https://www.knot-dns.cz/docs/latest/html/configuration.html#performance-tuning
  - ソケットバッファ　1048576
  - backlog　40000

測定結果（ロギング無効）



測定結果（dnstapのログ出力を有効）

（注）バッファ溢れあり
JSONデータを1秒間で100万行
処理させるチューニングは出来ず



考察とこれからの展望

■ BIND 9.20について、9.18と比較するとパフォーマンス向上
- 権威DNSサーバーとしての性能
- QP trie・RCUにより性能向上したか

■ 言わずもがな、NSD/Knot DNSのパフォーマンスはBIND 9よりかなり高い
- DO bit onのパターンでNIC帯域の限界まで、BIND 9側は結果からも分かるように余裕あり
- CPU使用率では16スレッド全て100%に張り付くテストも
  ➔CPUを強化、ネットワークを25Gまたはそれ以上の環境を用意できれば…
- 前回のテストではパフォーマンスが出なかったKnot DNS(dnstap有効)は改善？

性能測定についてコメント・ご意見あれば是非お願いします

さいごに

■ ソフトウェアダイバーシティを確保したことで耐障害性は向上

- お金・運用に関わる部分の負担は増えるので、トレードオフとなる部分の整理が付くかどうか
- 安定を一番に考えたJP DNSの運用では導入に動けた
- 運用改善はこれからも続いていく

  ➔取り組み・技術要素等、この場でお話いただけるものがあればコメント頂けると嬉しいです

■ BIND 9.18を利用されている方はEOL情報にご注意ください

- 9.20は権威DNSサーバーとしての性能向上は確認出来た

  ➔評価についてもコメント・知見を共有いただけると嬉しいです
- JPRSとしても注視し、引き続き情報展開を進めていきます