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)

旧ヤフーの場合

LINEヤフー株式会社 Site Operation本部

太田 健也

LINEヤフー

自己紹介

太田 健也

- 2021 年旧ヤフーに新卒入社
  配属以来データセンタ内の L2/L3 スイッチの構築・運用を担当
- 2022 年 10 月から DNS の運用も担当

これまでの変化

JANOG52 発表の振り返り

リゾルバ台数削減

- サーバー台あたりのキャパシティプランニング再検討

デプロイコストの低減

- 構成管理ツールによる config 自動配布
- resolv.conf 自動配布

資料: https://www.janog.gr.jp/meeting/janog52/dnscost/

これまでの振り返り

サーバの規模感

東日本/西日本それぞれ

- 物理サーバ数: 25,000+
- VM 数: 70,000+
- リゾルバ数: 100+

※主要拠点のみ

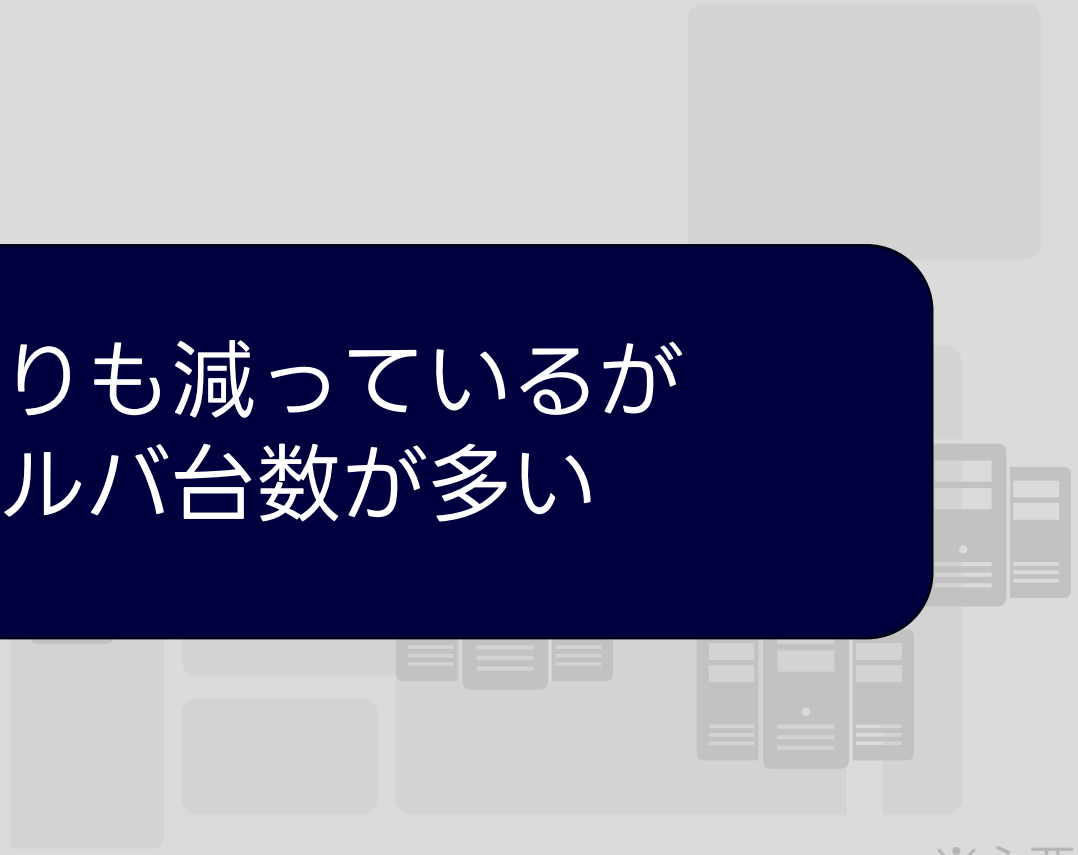
これまでの振り返り

サーバの規模感

東日本/西日本それぞれ

- 物理サーバ数: 25,000+
- VM 数
- リゾ

ピーク時よりも減っているが
依然リゾルバ台数が多い



※主要拠点のみ

なぜリゾルバ台数が多いのか

JANOG52 の議論より

クエリログの取得がボトルネックとなっている可能性

ログ取得の目的	取得内容	ログの量
セキュリティ監査の観点 • DC 内のサーバが どのような名前解決 クエリを出したか	• 時刻 • source IP address • qname, qtype など	• リゾルバ 1 台あたり 平均約 2 MBytes/s

パフォーマンスの改善手法

オプションとしてどのようなものがあったか

Linux kernel parameter や DNS 実装の設定調整

tmpfs の使用

dnstap の導入

パフォーマンスの改善手法

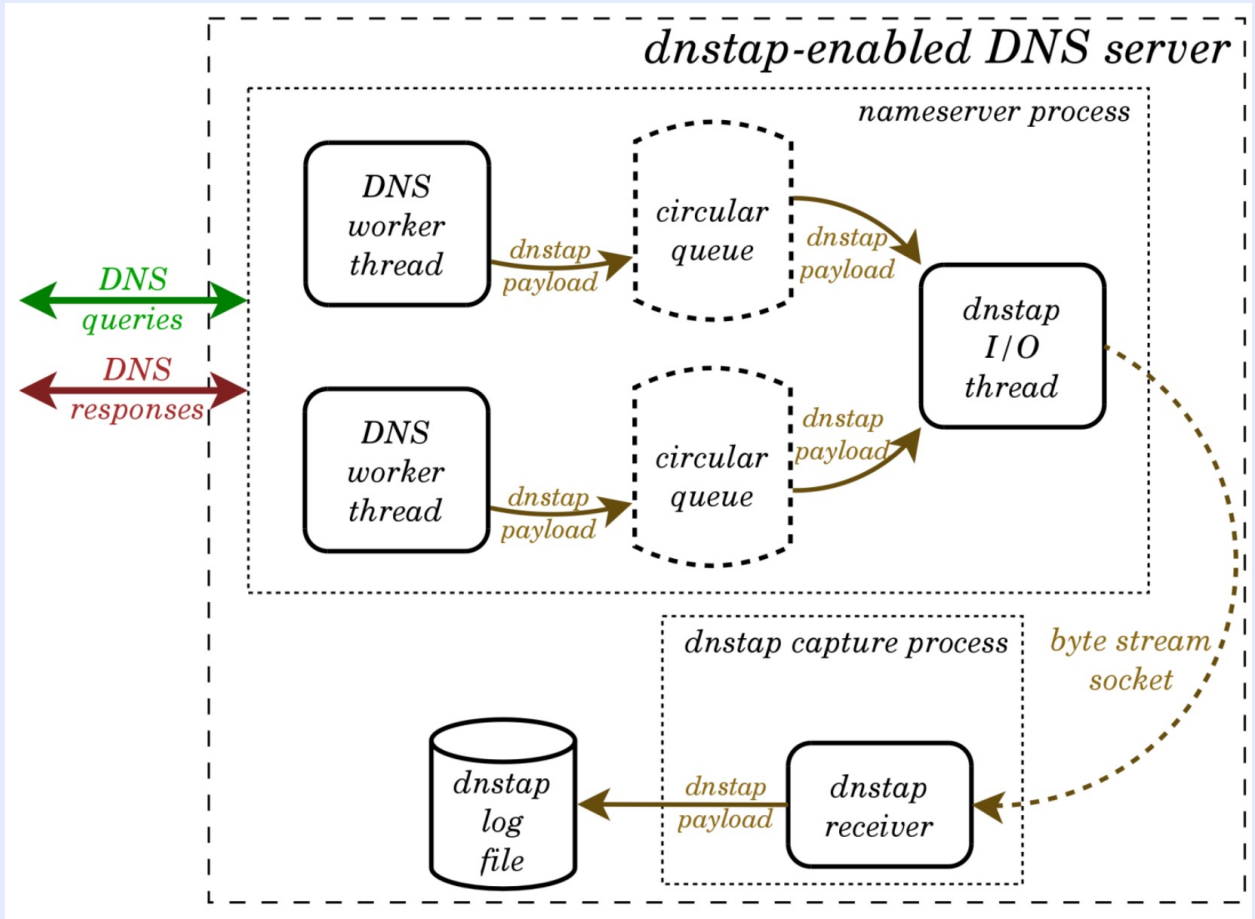
オプションとしてどのようなものがあったか

設定調整・tmpfs は優位な差はみられず

dnstap は disk I/O 削減効果が見込まれたため
そちらを検討

dnstap の導入

dnstap とは



出典: https://dnstap.info/slides/dnstap_vldss2014.pdf

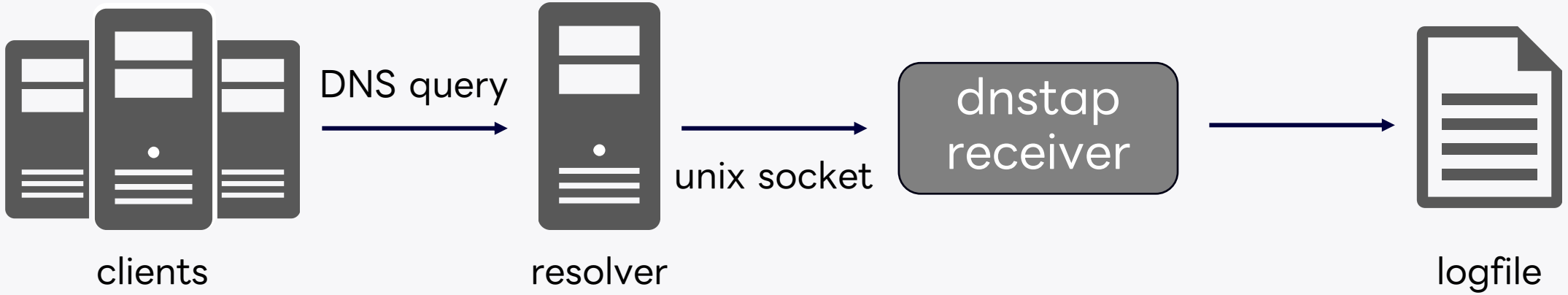
DNS ソフトウェアにおけるクエリログ出力フォーマットの一つ

- 対応している DNS 実装では同じ構造化された出力が得られる
- DNS worker thread と dnstap I/O thread が分離しているためログ出力のパフォーマンス改善が期待できる

dnstap 構成

概要

DNS 実装側のクエリログ出力機能をオフにし、dnstap 経由でログ出力



dnstap 構成

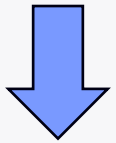
receiver 実装の検討

出力されるログの形式は receiver により異なる

- 柔軟性の観点から go-dnscollector を採用

go-dnscollector config

```
global:
  text-format: "timestamp-rfc3339ns identity queryip queryport family protocol
qname qtype"
```



クエリログ

```
2023-12-19T12:29:01.276695051Z dnstap_id 192.0.2.1 12345 IPv4 UDP example.jp A
```

https://github.com/dmachard/go-dnscollector/blob/main/docs/configuration.md#custom-text-format

dnstap 構成

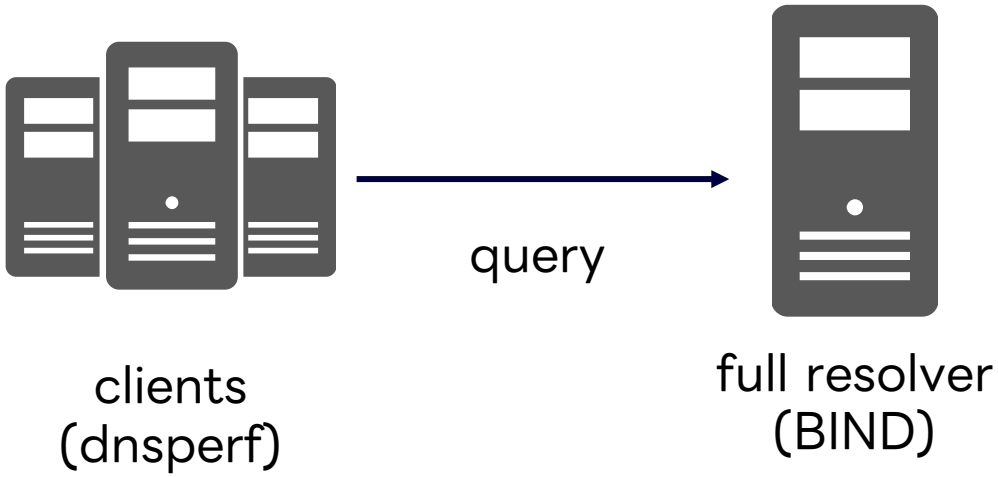
go-dnscollector サンプルコンフィグ

```
multiplexer:
  collectors:
    - name: tap
      dnstap:
        sock-path: "/path/to/unix.sock"
  loggers:
    - name: file
      logfile:
        text-format: "timestamp-rfc3339ns
identity queryip queryport qname qtype"
        file-path: "/path/to/query.log"
        max-size: 600
        max-files: 6
        mode: text
  routes:
    - from:
        - tap
      to:
        - file
```

- yaml 形式で記述
- multiplexer で複数の collectors, loggers, routes を定義可能
  - collector: ログ収集方法を設定
  - loggers: 記録方法やメトリクスの出力を設定
  - routes: collector と logger の対応を設定

ベンチマーク

計測環境と条件



クライアント・リゾルバともに同一性能のサーバを使用し、以下の条件で比較

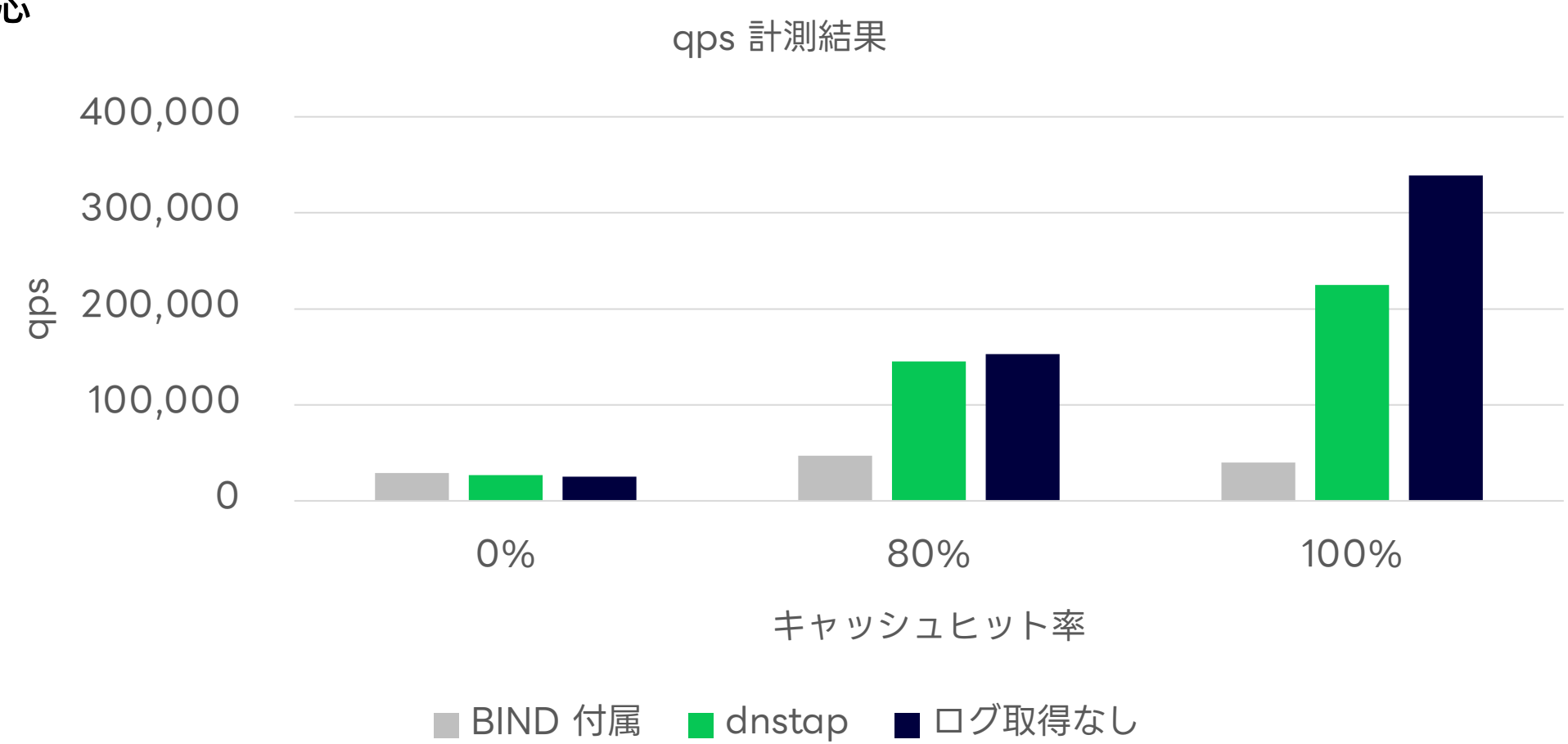
項目	スペック
CPU	Intel Xeon E5-2630L v3 1.80GHz x 2CPU (16C32T)
Memory	128GB
DNS software	BIND 9.16
benchmark	dnsperf 2.11.2

項目	条件
クエリログ取得方法	1: 取得なし　2: BIND 標準 3: go-dnscollector
キャッシュヒット率	1: 0%　2: 80%　3: 100%

ベンチマーク

計測結果

BIND 標準と比較して qps 性能が 3〜4 倍程度向上していることを確認



削減目標

ベンチマーク結果からキャパシティプランニング

考慮点

以下の点から各クラスタ最低台数を設定

- 障害発生時の縮退率
- 大規模災害時の BCP

性能

利用状況からキャッシュヒット率 80% で試算

- qname のほとんどが社内向けのドメイン名

削減目標

ベンチマーク結果からキャパシティプランニング

考慮点	性能
以下の点 を設定 • 障害発 • 大規模災害時の BCP	ヒット率 社内向け のドメイン名

これらを考慮しても半数以上
削減可能であることが判明

これからの課題

新しく標準化されるレコードや機能への対応

- 需要や利便性・必要性のほか、導入による影響・運用コストなど様々な観点で検討した上で対応する必要がある
- 例: HTTPS レコード、XFR over TLS など

まとめ

これまでの変化	現在の取り組み	これからの課題
以下を中心に取り組み • リゾルバ台数削減 • デプロイコスト低減	dnstap 導入によるさらなるリゾルバ台数削減 • クエリログ取得がボトルネック • dnstap により qps 性能が大幅改善	• 新しく標準化されるレコードや機能への対応