JANOG54 Day1

2024年7月3日(水) 15:45～16:30

マンガ海賊版サイト動向2024 -対策状況アップデート-

マンガ海賊版サイトの大量停止についての技術的考察

もっと素敵な伝え方を。

株式会社Ｊストリーム
エンジニアリング推進室　高見澤信弘

自己紹介

▶名前：高見澤信弘

▶出身地：山形県天童市

▶所属：株式会社Ｊストリーム (AS24253)

- 新卒でＪストリーム入社
- エンジニアリング推進室＆プロダクト企画部（アーキテクト）

▶お仕事

- 動画配信のためのネットワーク企画、構築
- CDN(Content Delivery Network)の企画、構築

▶好きなもの

- ロードバランサー　→　家にBIG-IP
- 19インチラックあるよ

もっと素敵な伝え方を。

Jstream

- 活動
  - IPoE協議会 IPv6地理情報共有推進委員会 幹事
  - 海賊版対策実務者意見交換会 海賊版対策技術検証チーム(WG) メンバー

J-Streamについて

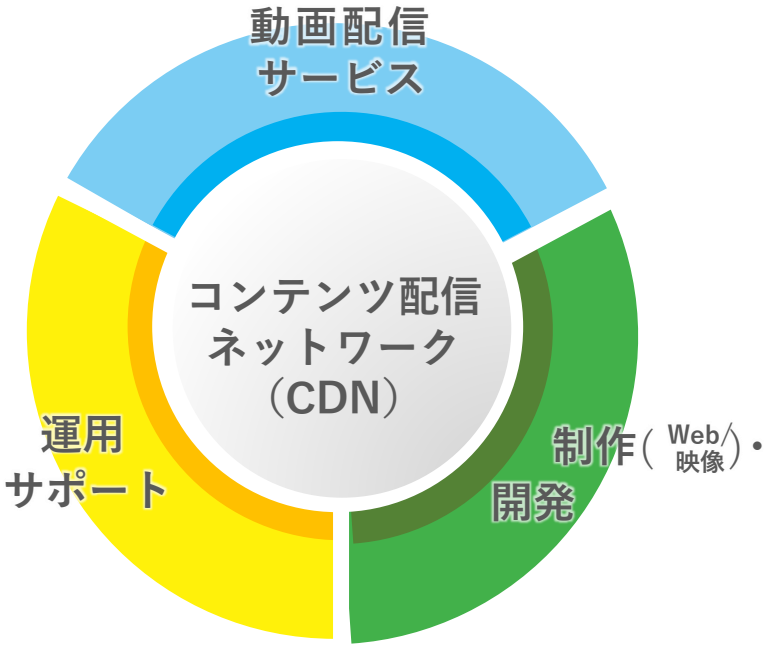
もっと素敵な伝え方を。

Jstream

www.stream.co.jp

株式会社Ｊストリーム

Ｊストリームは1997年の設立以来、動画配信を主軸として事業展開を続けております。
自社で保有･運営する独自のコンテンツ配信ネットワーク（CDN）を活用した動画配信に加え、
これまで積み上げてきたノウハウを活かした動画の企画・制作・運用から、Webサイト制作、
システム開発、動画広告による収益化支援まで、総合的なサービスとソリューションを提供し、
企業のマーケティングやコンテンツビジネスなどを支援しております。



設立　　1997年5月
証券コード 4308（東京証券取引所 グロース市場）
本社：105-0014 東京都港区芝二丁目5-6 芝256スクエアビル6階
西日本オフィス：530-0003 大阪府大阪市北区堂島2-4-27 JRE堂島タワー5階
資本金 21億8,237万円 （2023年12月末現在）

本日の目次

- マンガ海賊版サイトの大量停止についての紹介
  - 時系列での状況の整理
  - 調査の経緯の振り返り
- 技術的な裏付けを踏まえた考察
  - DNS（レジストラ）の対応
  - CDN事業者の対応
- その後の展開

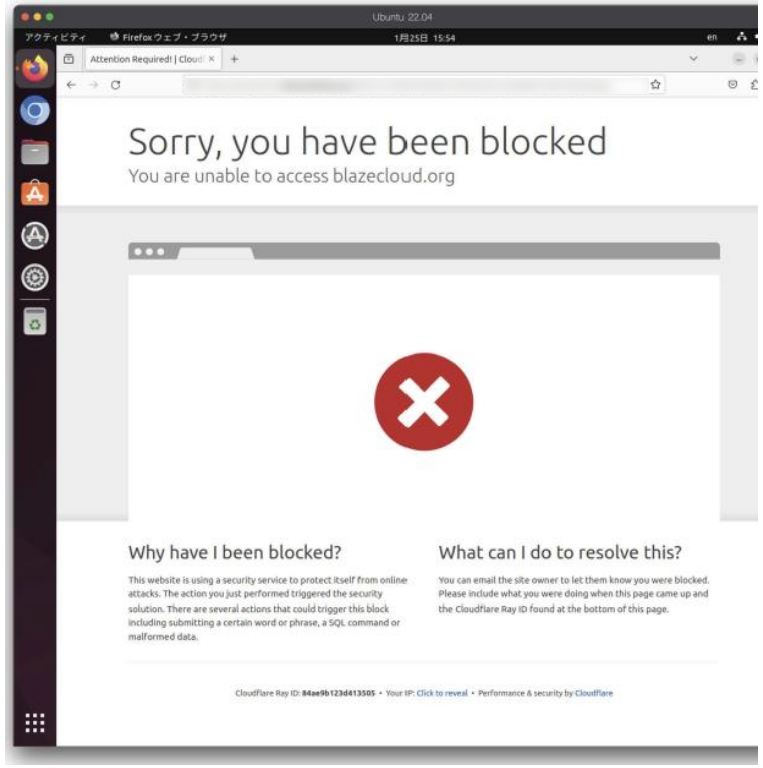
マンガ海賊版サイトの大量停止についての時系列

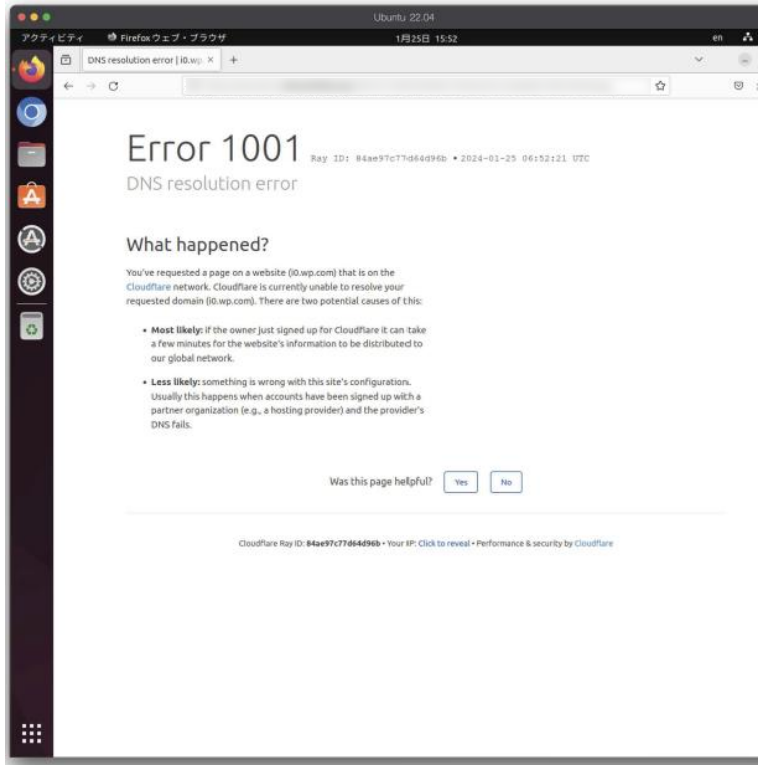
- ▶2024/1/15(月)に、複数の海賊版サイトがアクセスできない状態であることを発見
  - 20件以上の海賊版サイトに対して継続的にアクセスできない状態となっていた
  - その中にはアクセス数トップ10以内のサイトも複数含まれていた
  - 技術検証チームとしても調査を開始
- ▶1月下旬：調査により原因を2パターンに絞って調査継続
  - 特定のレジストラがホスティングしているドメイン名について設定変更を行った
  - 特定のCDN事業者が配信設定の変更を行った
- ▶2月中旬
  - 複数のレジストラによる設定変更で、海賊版サイトの名前解決が正常に行えない状態になっていることを確認
  - CDNについては、キャッシュ設定の変更により、海賊版サイトが正常に動作しない状態になっていることを確認
  - しかし、CDN事業者が設定変更を行ったのか、サイト運営者が間違って設定したのかの確証が取れなかった
- ▶3月上旬
  - CDNについても事業者側の設定変更により、海賊版サイトが正常に動作しない状態になっていることを再現

意図的にサイトを停止させる対策が取られた可能性が高いと結論付けた

サイトの状況

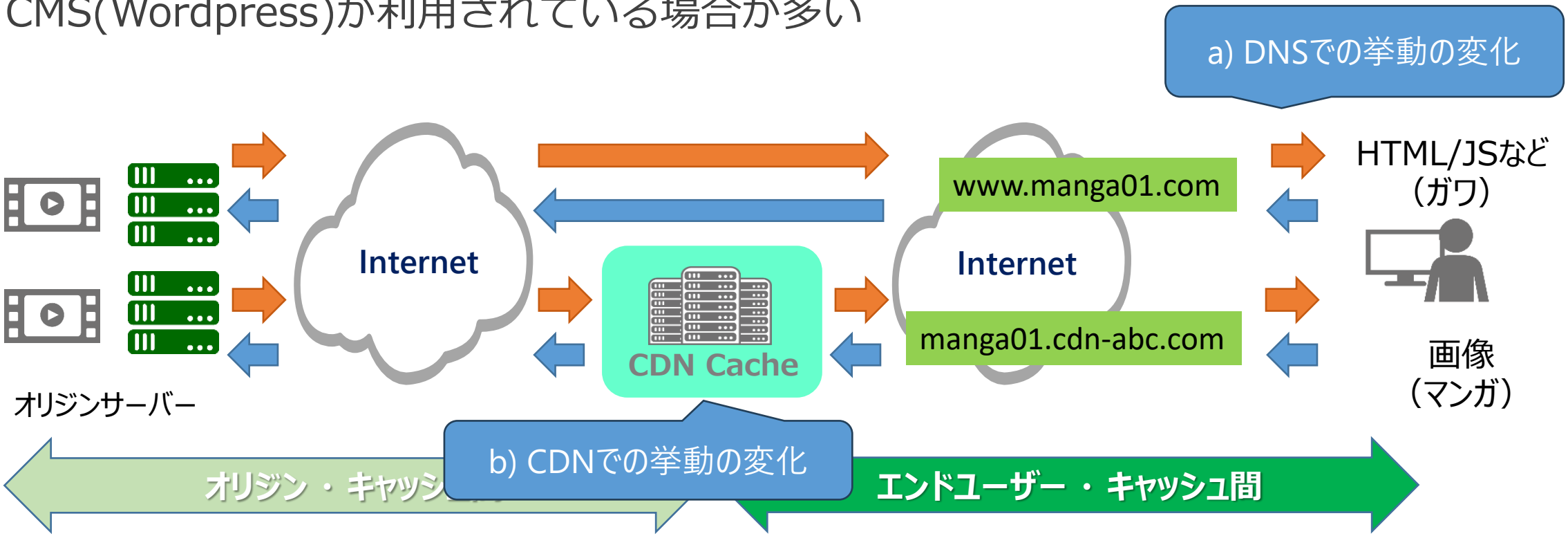
- サイトによりエラー内容が複数パターン存在する状態
  - 名前解決のエラー
  - タイムアウトエラー　など





マンガ海賊版サイトの仕組みについて

- Webページ（ガワ）と画像配信は別FQDN
  - Webページはオリジンサーバーを直接見せている（場合が多い）
  - 画像配信についてはCDNを利用
- オリジンサーバーはWebページも画像共通のサーバーが利用されている
  - CMS(Wordpress)が利用されている場合が多い



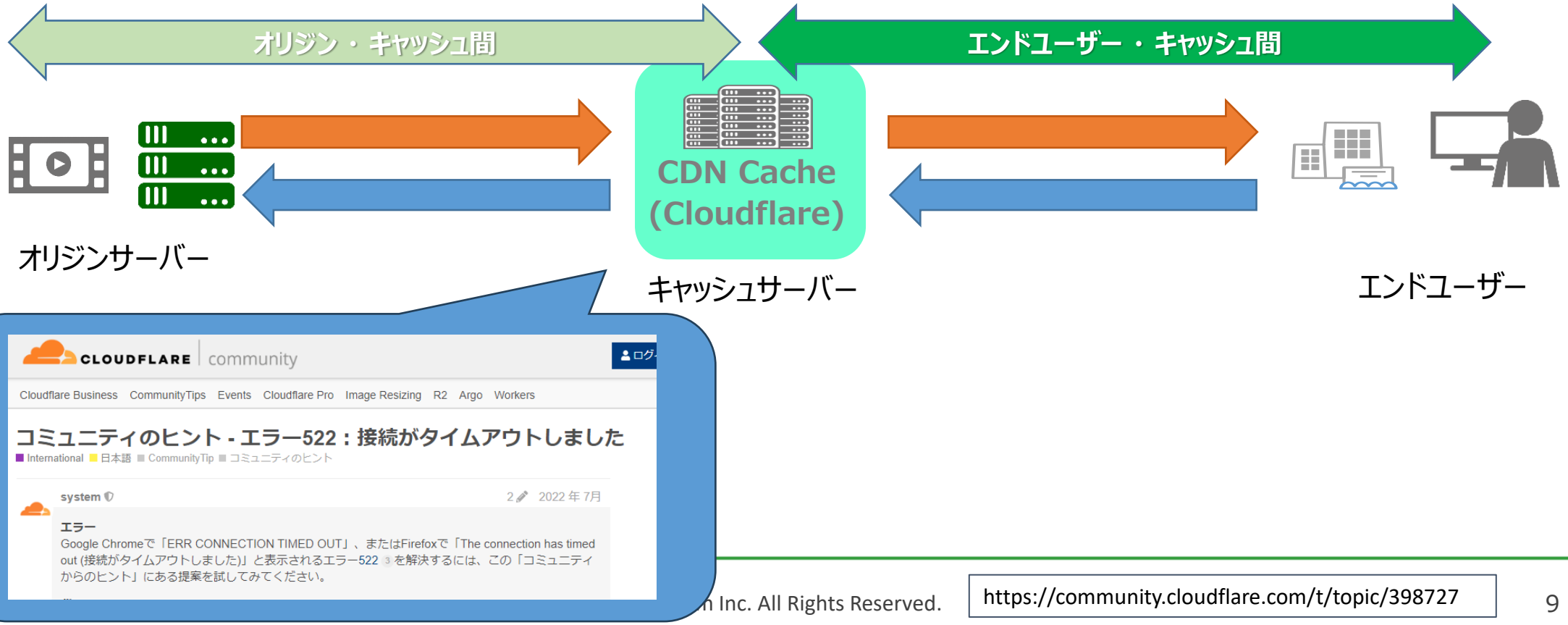
a) DNSでの挙動の変化

- マンガ海賊版サイトで利用されていた一部のRRを回答しなくなった
  - ドメインのリセラーであり権威サーバを提供する njalla サスペンドさせたように見える
  - 回答しなくなったRR(リソースレコード)はWebページで利用されていたFQDNなど
  - SOAレコードのシリアル番号が複数のサイトで同じ値となっていた

No.	ドメイン名	NS	SOA (2024/1/25 11:00 JST頃に確認)
1	[redacted]	njalla	1-ceci.njalla.do. you.can-get-no.info. **2024012119** 21600 7200 1814400 3600
2	[redacted]	njalla	1-ceci.njalla.do. you.can-get-no.info. **2024012119** 21600 7200 1814400 3600
3	[redacted]	njalla	1-ceci.njalla.do. you.can-get-no.info. **2024012119** 21600 7200 1814400 3600
4	[redacted]	njalla	1-ceci.njalla.do. you.can-get-no.info. **2024012119** 21600 7200 1814400 3600
5	[redacted]	njalla	1-ceci.njalla.do. you.can-get-no.info. **2024012119** 21600 7200 1814400 3600
6	[redacted]	njalla	1-ceci.njalla.do. you.can-get-no.info. **2024012119** 21600 7200 1814400 3600
7	[redacted]	njalla	1-ceci.njalla.do. you.can-get-no.info. **2024012119** 21600 7200 1814400 3600

b) CDNでの挙動の変化

- HTTPステータスコードとしては522が返ってきている
  - オリジンサーバーへのTCPセッションが確立しなかった場合のステータスコード・タイムアウトは15秒程度
- コンテンツをキャッシュしなくなったことですべてのリクエストがオリジンサーバーに到達、オリジンサーバーが返答できない状態が発生



https://community.cloudflare.com/t/topic/398727

b) CDNでの挙動の変化

- HTTPステータスコードとしては522が返ってきている
- HTTPヘッダー(キャッシュコントロールヘッダー)の比較
  - キャッシュファイルの扱いをブラウザやキャッシュサーバーに対して指示するヘッダー
- 停止中のサイトのコンテンツはキャッシュしない設定となっている
  - 配信中のサイトのmax-ageは 31536000s (= 365日)キャッシュする
  - max-age=0：キャッシュ保持時間を”0”にする＝キャッシュしない

■配信中のサイトの場合

```
< HTTP/2 200
< expires: Tue, 18 Mar 2025 02:09:27 GMT
< cache-control: public, max-age=31536000
< last-modified: Mon, 18 Mar 2024 02:09:27 GMT
< cf-cache-status: HIT
< age: 546113
```

■停止中のサイトの場合

```
< HTTP/2 522
< cache-control: private, max-age=0, no-store, no-cache, must-revalidate, post-check=0, pre-check=0
< expires: Thu, 01 Jan 1970 00:00:01 GMT
```

配信停止しているサイトの画像蔵置先CDNでは、ファイルをキャッシュをせずオリジンサーバーからコンテンツを中継するだけの設定になっている模様
※リファラー制限などの設定は残っており、設定が消えたわけではない

まとめ

- マンガ海賊版サイトの大量停止について
  - 時系列の調査状況の紹介
- 技術的な裏付け（検証）の中身を紹介
  - 複数のサイトでのSOAレコードのシリアル番号の一致
  - キャッシュコントロールヘッダの事業者側での変更

- 意図的にサイトを停止させる対策が取られた可能性が高い
- 改めて、CDNがないと大量の配信は難しいことが証明されてた

- その後の展開
  - 一部のサイトはCDNを変更するなどして再開したが、多くのサイトは閉鎖されたまま
  - CDNを変えて再開するより、新しいドメインで新たに立てる場合が多いと想像される
  - 量産型サイトの脅威へ、、、

CFにおけるキャッシュコントロールヘッダの制御

参考

- 利用者側が意図的に設定した可能性の確認
  - Web管理画面の設定だけで、今回の事象を再現できるか検証した
  - 今回の事象を再現することはできなかった(No.3,4)
- なんらかの意図をもって設定していると考えるのが妥当

No	設定内容	キャッシュコントロールヘッダー	備考
1	初期状態での設定	**cache-control: max-age=14400**	4時間キャッシュする
2	配信中のサイトでの設定（画像）	**public, max-age=31536000**	365日キャッシュする
3	キャッシュの適格性 キャッシュをバイパスする	-	キャッシュコントロールヘッダーが付与されなくなった
4	ブラウザTTL キャッシュをバイパスする	-	キャッシュコントロールヘッダーが付与されなくなった
5	今回の事象	**private, max-age=0, no-store, no-cache, must-revalidate, post-check=0, pre-check=0**	