気象庁XML等のオープンデータにおける電子署名とタイムスタンプの活用アイデア

2014年2月4日
クラウド・テクノロジー活用部会
宮地（miyachi@langedge.jp）
有限会社ラング・エッジ

オープンデータとセキュリティ技術

そもそも以下の疑問が…

疑問１:オープンデータに使えるセキュリティは？
疑問２:オープンデータにセキュリティは必要？

本日は気象庁XMLを例に上記の２疑問について考察してみます。

その上で利用可能と考えられるセキュリティ技術として電子署名とタイムスタンプの利用についてのアイデアを説明します。

まず情報セキュリティ技術を分類

大分類	小分類	機能	対象
暗号応用	暗号化	情報秘匿	データ
	秘密分散	機密データ保存	
電子署名	電子署名（PKI方式）	改竄防止、発行者確認	
	タイムスタンプ	改竄防止、時刻保証	
	非PKI電子署名	改竄防止、証拠保存	
電子透かし	追跡・秘匿通信	追跡（不正コピー対策）	
電子認証	認証	本人性確認	通信運用
	認可	権限（アクセス権等）付与	
対策	ウイルス・侵入	ハッキング対策	操作

秘密にする意味が無いので除外

対象が限定なので除外

オープンデータだと使えそうなのは「電子署名」。

データの１次取得と2次取得の違い

直接取得の場合

気象庁のURLを指定

直接データを取得

間違い無く気象庁のデータだ！

2次取得の場合

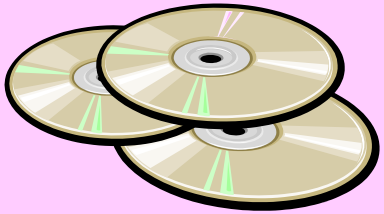
取得

蓄積中継サーバ

流通

取得

改竄ができるぞ



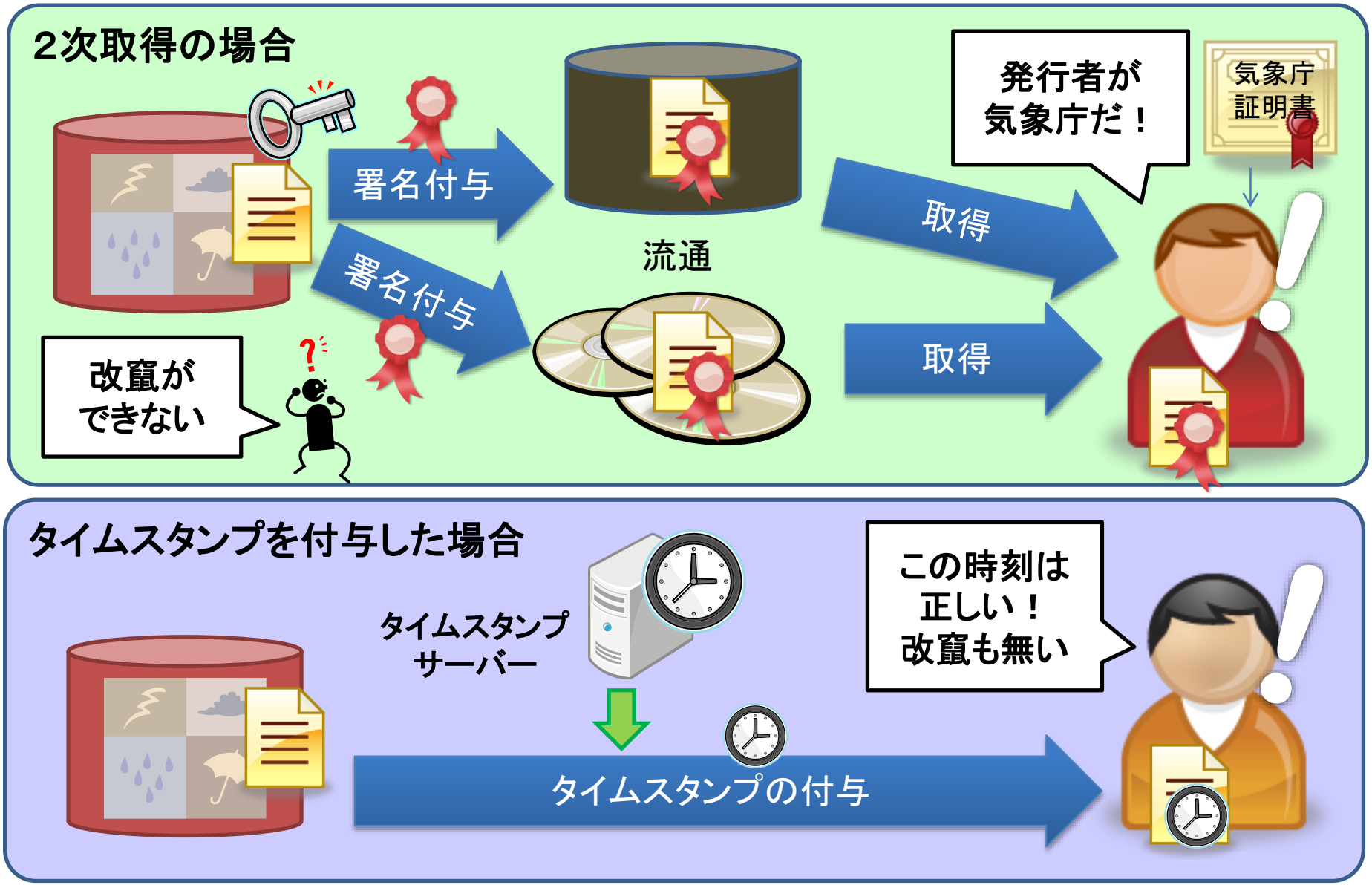
データファイル

第三者URLで取得

本当に気象庁のデータ？

ファイル取得

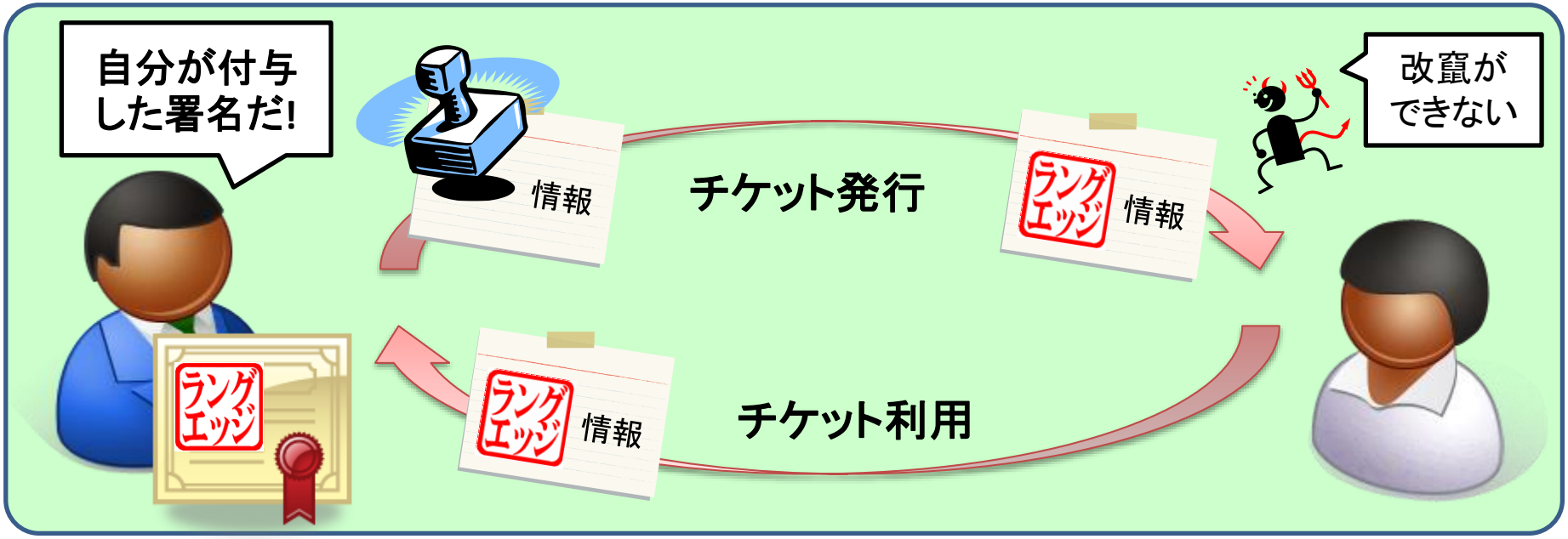
電子署名により発行者を確認できる



電子署名は意外に使われている

認証トークン（電子チケット）

- 自分で署名したトークンやチケットを配布して利用時にそれらの署名を確認して判断。



クライアント認証

- サーバが認めた認証局が発行した証明書を持つクライアントだけが接続可能。

コード署名

- 決められた証明書で署名されたアプリだけが実行可能。または発行元を証明書で確認。

ではタイムスタンプは？

法的対応や知財保護には使われている。
海外（中国やタイ等）でも普及しそうな情報あり。

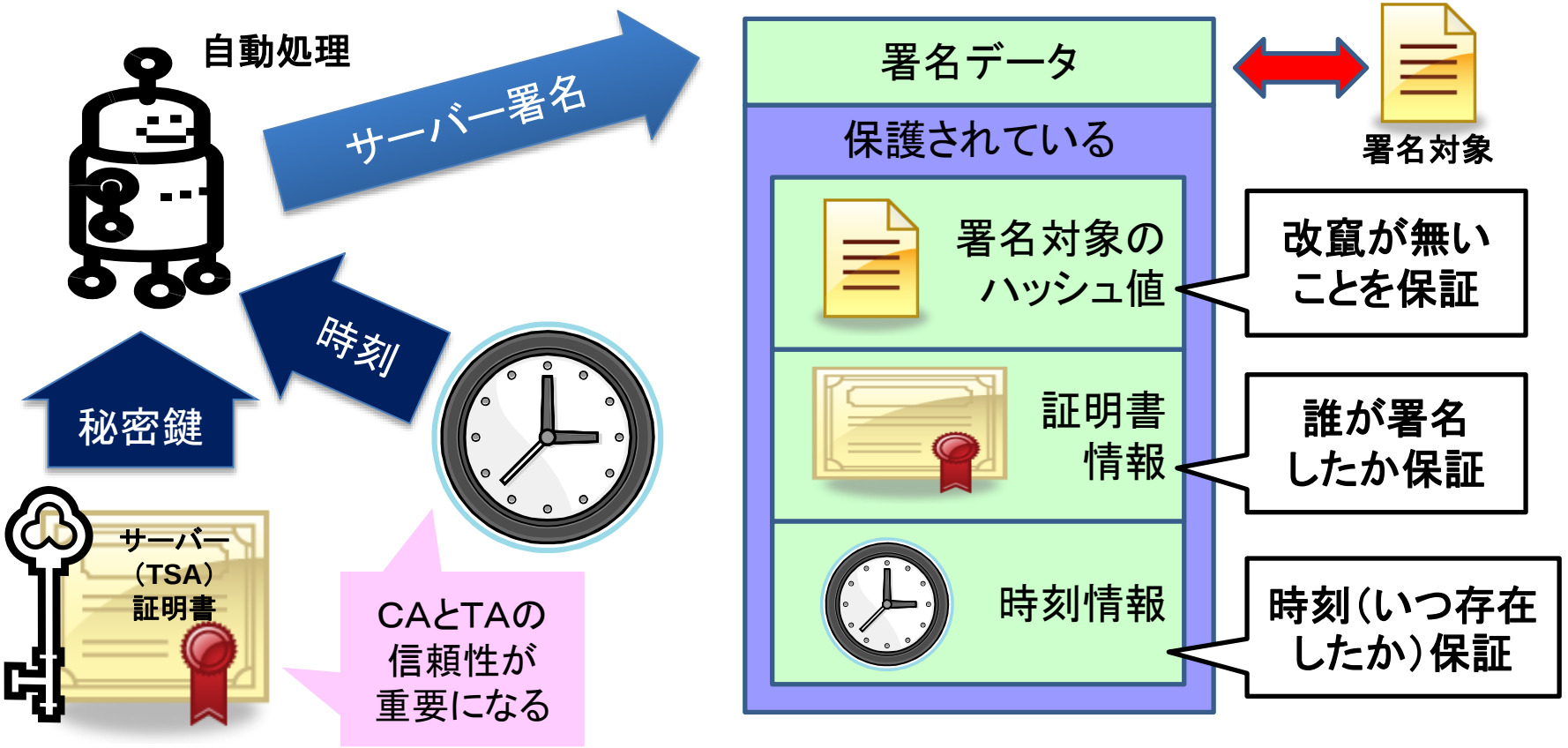
でもオープンデータ･クラウドでは使われていない。

サーバ署名なので一般的なRESTfulでは無いが
マッシュアップ用のAPIとしての利用が可能。

認定タイムスタンプ局では無くもっと手軽に使える
タイムスタンプサービスがあればもっと使える？

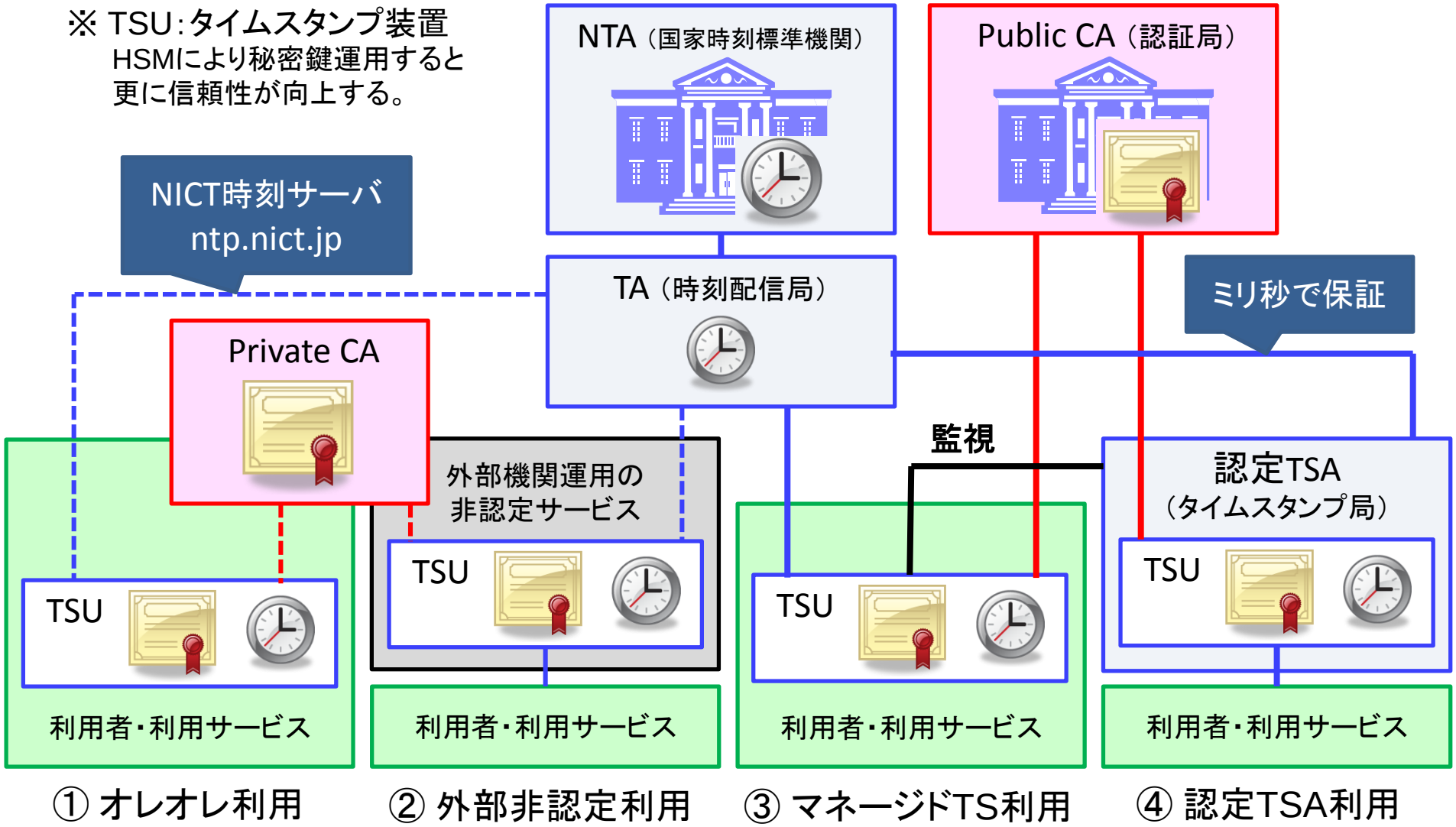
タイムスタンプ（RFC3161）って何？

タイムスタンプはPKI方式のサーバ署名の一種。
署名時に「時刻」も埋め込むので時刻を保証可能。
サーバにHTTP経由で要求するので自動化可能。



タイムスタンプ運用パターン

※ TSU：タイムスタンプ装置
HSMにより秘密鍵運用すると
更に信頼性が向上する。



FreeTSAプロジェクト（お試しに！）

OpenSSLコマンドによる簡単タイムスタンプ（RFC3161）サービス

- タイムスタンプって面白そうだから使ってみたい！
- タイムスタンプ クライアントの試験で使えるサーバは？

と言うような要望に応える為にOpenSSL 1.0.0のコマンドを利用して簡易タイムスタンプサービス（サーバ）の構築手順をまとめた。ラング・エッジのサーバで稼働させて手軽に使えるタイムスタンプサービスも提供中。プロジェクトとしては以下の2つ。なお作るには別途TSA証明書は用意が必要。

作る：FreeTSA Project：10分でできるタイムスタンプ局
使う：FreeTSA Service：自由に使えるタイムスタンプ局

http://www.langedge.jp/tsa

- 参考 LangEdge Weblog: フリータイムスタンプ局（FreeTSA）のすゝめ
  http://www.langedge.jp/blog/index.php?itemid=665

タイムスタンプを使ってみよう！

How to generate and send an RFC3161 timestamp with OpenSSL and curl.

http://unmitigatedrisk.com/?p=395

OpenSSL 1.0.0 の ts オプションを使って、タイムスタンプクライアントとしてタイムスタンプサービスを利用する方法（タイムスタンプ取得と検証）が書いてある。

HTTP通信はcurlコマンドを利用。curlはhttpsも対応可能。

FreeTSAのサーバと組み合せればタイムスタンプクライアントの試験も出来ます！

効率良い署名・タイムスタンプの付与

オープンデータやビッグデータは対象が多い。
大量の対象に効率良く署名/タイムスタンプを付与。
工夫が必要、例えば…

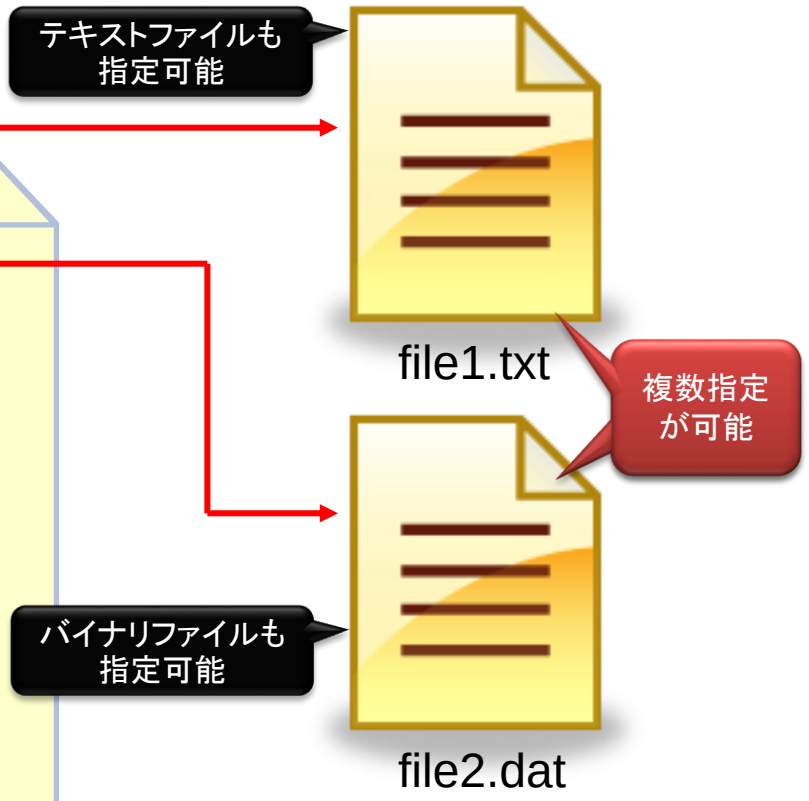
1. 署名数とタイムスタンプ数を減らす。
2. フリーのタイムスタンプをうまく利用する。

既に色々と仕様が策定・検討されています。
また実績もあります。

複数対象に1つのXML署名付与1

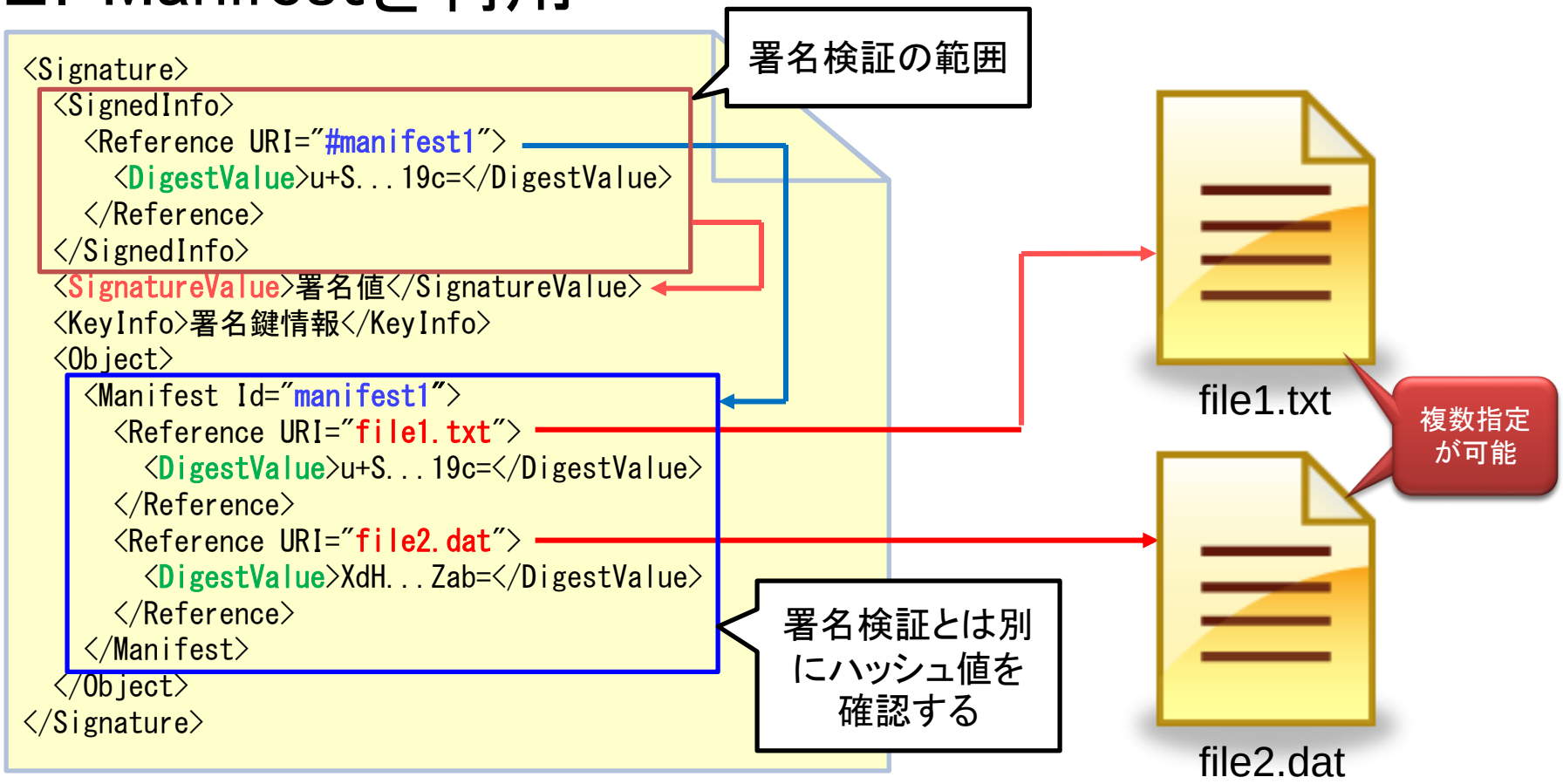
1．複数のReferenceを利用

```
<Signature>
  <SignedInfo>
    <Reference URI="file1.txt">
      <DigestValue>u+S...19c=</DigestValue>
    </Reference>
    <Reference URI="file2.dat">
      <DigestValue>XdH...Zab=</DigestValue>
    </Reference>
    <Reference URI="#SignedProp">
      <DigestValue>C6Q...OQ2=</DigestValue>
    </Reference>
  </SignedInfo>
  <SignatureValue>署名値</SignatureValue>
  <KeyInfo>署名鍵情報</KeyInfo>
  <Object>
    <QualifyingProperties>
      <SignedProperties Id="SignedProp">
        署名対象データ（長期署名）
      </SignedProperties>
      <UnsignedProperties>
        非署名対象データ（タイムスタンプ等）
      </UnsignedProperties>
    </QualifyingProperties>
  </Object>
</Signature>
```



複数対象に1つのXML署名付与2

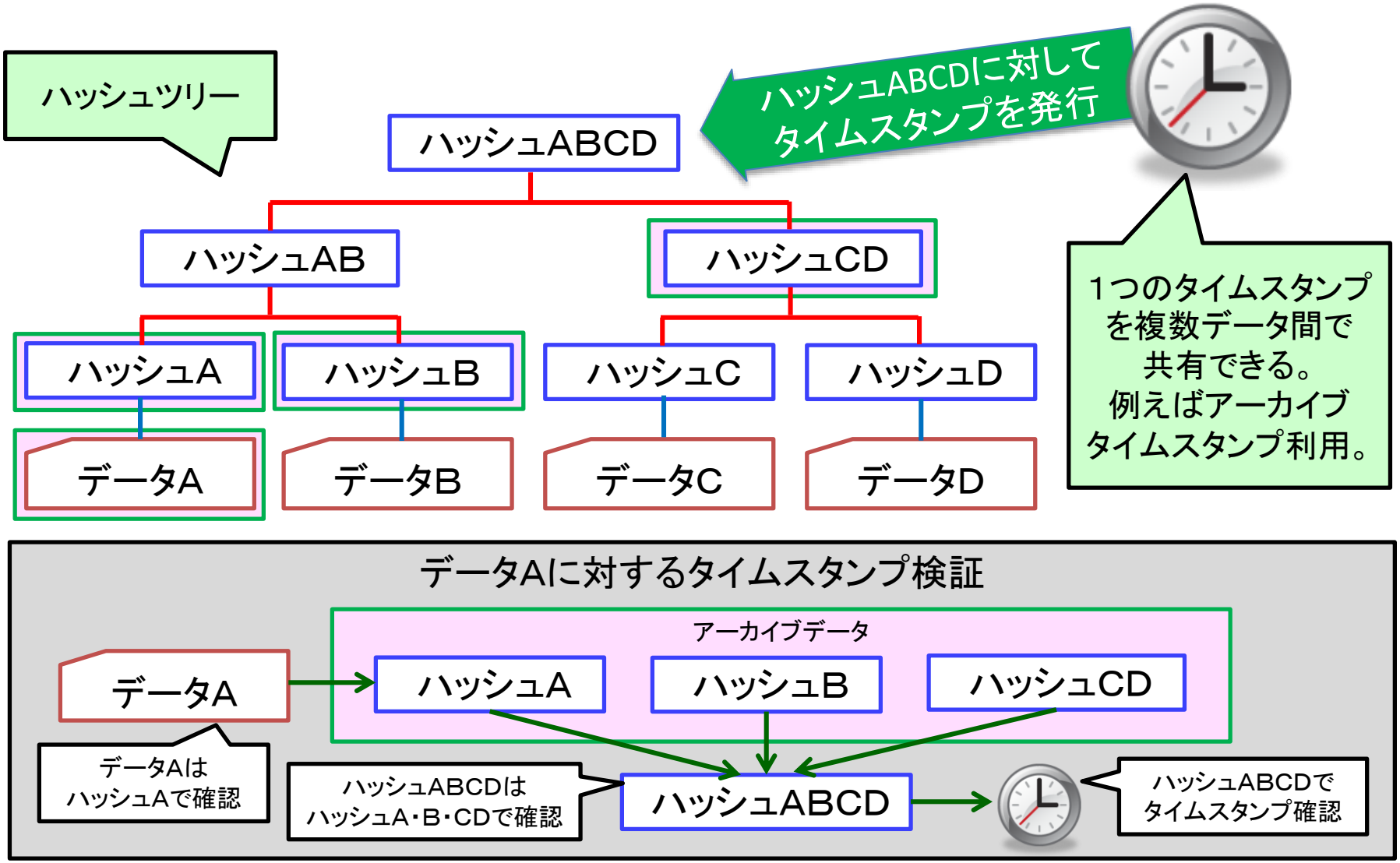
2. Manifestを利用



- XmlDsigのManifest要素は署名検証の範囲外
  - この為に署名検証が容易かつ短時間で終わらせることが可能。
  - Manifest から Reference されている署名対象のどれが不正か明確で他の署名対象に影響がないと言うメリットがある。

複数対象に1つのタイムスタンプ付与

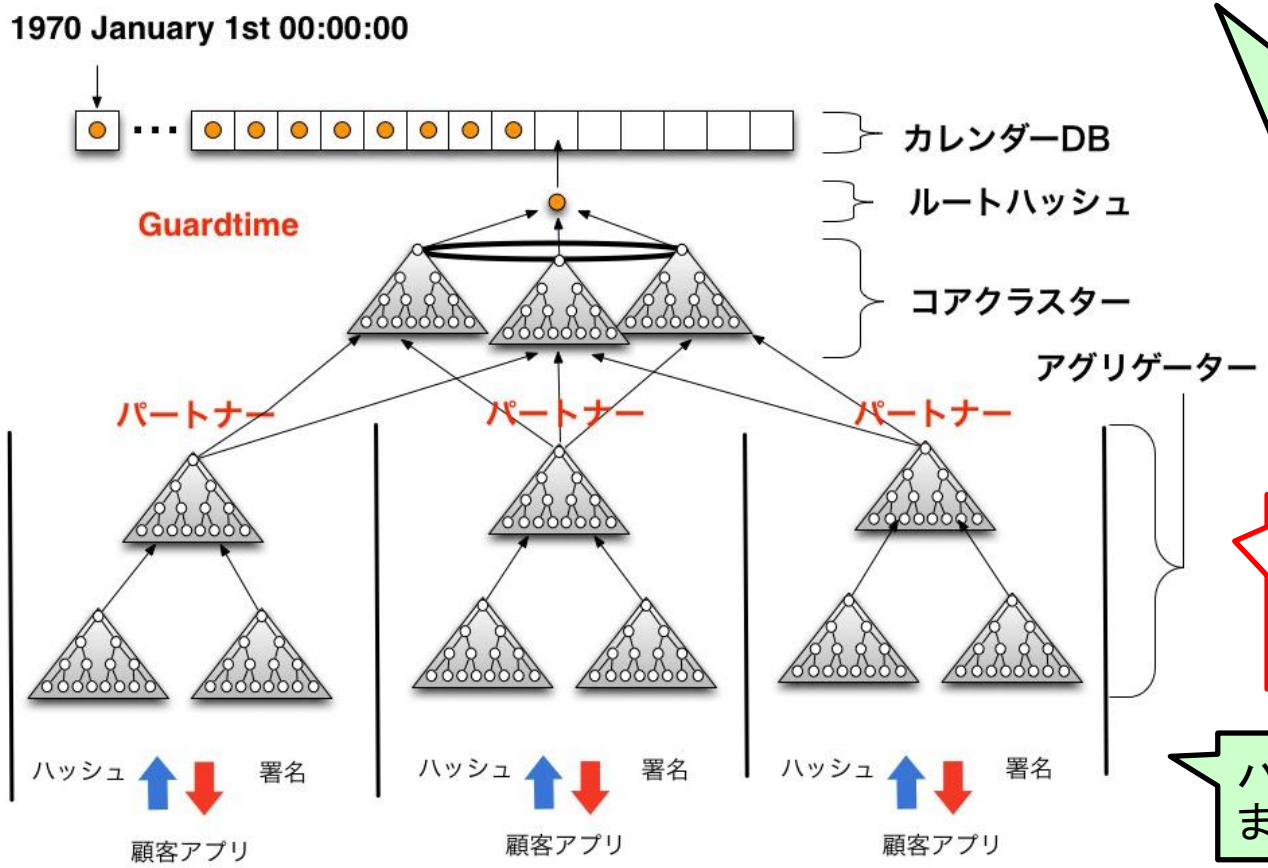
1．ハッシュツリーの利用（XMLERS/RFC6284）



複数対象に1つのタイムスタンプ付与

2. ガードタイム方式（非PKI署名）の利用

公表コード：カレンダーを結合して毎月１個のハッシュ値を生成し、新聞に掲載。
検証カレンダー：複数のデータから生成したハッシュ値を結合し世界で毎秒１個のハッシュ値を生成しネット上に公開。
検証：オリジナルデータ・取得したキーレス署名・カレンダーを使って計算したハッシュ値を新聞に公表されたハッシュ値と照合し、データの非改ざん性を検証。



非PKIの理由は、証明書等の信頼チェーンでは無く、新聞に公開されたハッシュ値を利用して信頼性を確保する為。単独で検証が可能。

全て数学的な仕組みの為にハッシュアルゴリズムが危殆化するまでは有効である。

ガードタイム社はシンガポールに本社があり欧米で使われている。標準化が望まれる。

ハッシュツリーを構築するところまではERSとほぼ同じ。

http://www.guardtime.com/

ハイブリッド方式のタイムスタンプ付与

独自運用と商用運用を組み合わせて利用。

既に医療分野にてCAdES方式だが実績がある運用形態。

独自運用と言っても運用規定を作りきちんとした運用は必要。

独自運用のタイムスタンプサーバ

商用運用のタイムスタンプサーバ

データA → ES（電子署名） → ES－T（署名タイムスタンプ）

データB → ES（電子署名） → ES－T（署名タイムスタンプ）

データC → ES（電子署名） → ES－T（署名タイムスタンプ）

ES－T → ES－A（アーカイブタイムスタンプ）

アーカイブタイムスタンプにERSを利用すると更に有効。

参考：http://www.jipdec.or.jp/esac/promotion/h24pdf/24-3.pdf

まとめ

疑問1：オープンデータに使えるセキュリティは？

電子署名やタイムスタンプの技術はオリジナリティを維持する為に使えるのではないだろうか。
暗号化はオープンなデータには適用する意味がない。

疑問2：オープンデータにセキュリティは必要？

直接データを取得する1次取得では意味がないかもしれないが、データが流通する場合は意味がある。
データのオリジナリティや発行元を確認・保証できる。

- タイムスタンプはマッシュアップに利用できる。
- 電子署名・タイムスタンプ技術は進歩している。
- 適用するデータに最適な仕様を検討できる。