2B-05

データ価値共創プラットフォームにおいて保護されたデータを安全に利活用できるデータ利用環境

坂本 久† 庄司 諒† 大泉 幸太† 真壁 祐† 川北 智弥†

NEC ソリューションイノベータ株式会社†

1. データ価値共創プラットフォームの概要

単一の企業や大学等に閉じず，複数の組織で連携してデータの利活用が求められる中，筆者らは，データを他の組織に安全に共有し効率よく活用するためのデータ価値共創プラットフォーム（以下，本プラットフォーム）を提案した（図 1）．本稿では，保護されたデータを利活用するための環境であるセキュアコンテナの仕組みや効果について論説する．

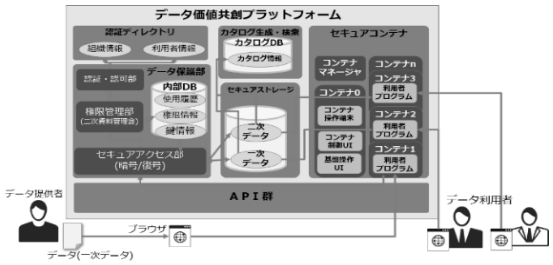

図 1 データ価値共創プラットフォームの構成

2. 保護データを利活用する環境の要件

本プラットフォームでは，データを DRM（デジタル権限管理）により暗号化し，データの所有者に対する所有権限，及び閲覧可能者に対する閲覧権限を設定する事で，権限の無い利用者にデータが漏洩しない様にデータを保護する．こうして保護したデータを利活用する環境は，次の２つの要件が求められる．

１）データ利用時の認証による保護

現在のコンピュータ環境では，概ねログイン認証を必要とし，企業や学校でも社員 ID や学籍 ID 等でのログインが求められる．プログラムで保護データを利用する際に，ログイン認証で用いた認証情報をデータ利用時も用いる事でデータ利用者とプログラム実行者の同一性や，シングルサインオンによる利便性を確保する．

２）DRM に対応した汎用的なデータ利用環境

データの利用環境は，DRM により保護されたデータを利用するクライアントソフトウェアといえる．従来の DRM はデータを処理するためのクライアントソフトウェアは DRM に対応した特別なアプリケーションを使う必要があった．本プラットフォームにおけるデータ利用環境は，データを利用するプログラムに暗復号や鍵交換，認証処理等 DRM 関連の処理を追加せず，それら処理を可能な限り環境側に隠蔽し，利用者に意識させずデータを利活用できる様にする（図 2）．

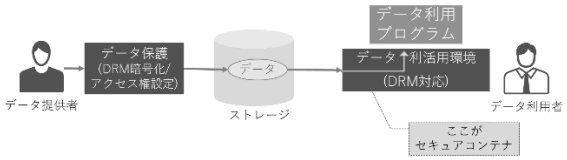

図 2 DRM に対応した汎用的なデータ利用環境

3. セキュアコンテナ

本稿では，これらの要件を満たすデータ利活用環境をセキュアコンテナと呼ぶ（図３）．

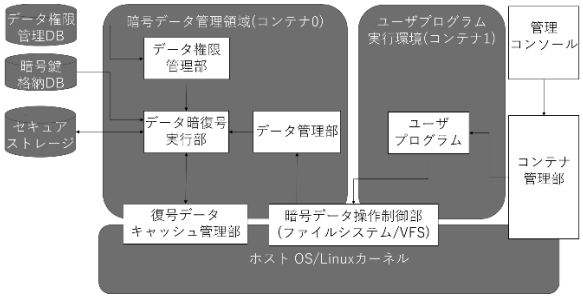

図 3 セキュアコンテナの構成

3-1. セキュアコンテナの構成

セキュアコンテナは，データを暗復号する処理を実行するコンテナ０と，データの利用者が使用するコンテナ１を装備する．これらコンテナはコンテナ管理部がその生成，起動を管理する．コンテナ管理部は kubernetes 等のコンテナ管理システムにより実現する．コンテナ０では，暗号データ操作制御部が，FUSE（Filesystem in UserSpacE）を用いて，保護データを格納するセキュアストレージをファイルシステムとしてマウントする．データ管理部はセキュアストレージのマウントや，プログラムが呼び出すファイル操作システムコールへの介入を制御する．データ暗復号実行部は，システムコールに介入し

The Data Utilization Environment which can leverages protected data securely on Data Value Co-Creation Platform

† Hisashi Sakamoto, Ryou Shoji, Kota Oizumi, Tasuku Makabe, Tomoya Kawakita, NEC Solution Innovators, Ltd.

セキュアストレージから読み込まれたデータを復号する．データ権限管理部はデータ利用者がデータをアクセスする事ができるかを判定し，可能な場合データを復号するための鍵を取得する．復号データキャッシュ管理部はアクセス速度向上の為，復号データをキャッシュする．データ利用者は管理コンソールを通して，コンテナの生成やプログラムの実行を操作する．

3-2. セキュアコンテナの動作

１）コンテナ管理部はコンテナ０を起動し，データ管理部はセキュアストレージをファイルシステムとしてマウントする．コンテナにログインした利用者が利用可能な全てのデータを復号して復号データキャッシュ管理部に登録する．また，一時領域もマウントし初期化する．

２）コンテナ管理部がコンテナ１を起動し，プログラムが実行され，セキュアストレージ内に存在するファイルをアクセスする．

３）データ暗復号実行部が，呼び出されたシステムコールに介入する．

４）データ暗復号実行部が，コンテナの利用者権限を基に，対象データに対して読み込み権限があるかをデータ権限管理部に確認する．権限が確認できた場合，対象のデータを復号データキャッシュ管理部から取得する．

５）読み込まれ復号されたデータはプログラムに返却され処理される．

６）データ管理部はプログラム終了時に．一時領域を全て消去し，コンテナ終了時にセキュアストレージのファイルシステムをアンマウントする．この際に，コンテナ１上でユーザプログラムが生成したファイルは暗号化されセキュアストレージに保存される．

4. 評価

前述した内容に基づきセキュアコンテナを試作し評価した．本フラットフォームはデータの暗復号を伴う為，平文アクセスよりオーバーヘッドが掛かるので，暗号化データのアクセス性能評価を実施した．また，保護データを利活用する環境の要件に対する評価として，外部（権限の無い利用者）からの復号データアクセス評価とプログラムの汎用性評価を実施した．

１）暗号化データのアクセス性能評価

暗号化データのアクセス性能評価を表１，表 2 の内容で評価した．項目１と２は合計サイズが同一でファイル数が異なる場合の評価であり，項目３は比較的大きなファイルの評価である．いずれの場合もアクセス速度向上に対するキャッシュの有効性を確認した．

２）外部からの復号データアクセス遮断

外部からコンテナ１内復号データに対するアクセスを表３の項目で評価した．項目１，２についてはアクセスを遮断できる事を確認した．

表 1 復号データへのアクセス評価結果

項目番号	測定パターン	キャッシュ機能	キャッシュコピー(ms)	ファイル転送(ms)	復号(ms)	読込合計(ms)	キャッシュ有無による読込合計時間差
1	ファイル 4KB×1000個	あり	172	-	-	172	9.55倍
		なし	-	957	686	1,643	
2	ファイル 1MB×40個	あり	145	-	-	145	23.72倍
		なし	-	1,350	2,090	3,440	
3	ファイル 100MB×10個	あり	7,069	-	-	7,069	11.76倍
		なし	-	25,813	57,319	83,132	

表 2 評価環境

インスタンスタイプ	AWS EC2 t3.medium
CPU	Intel(R) Xeon(R) Platinum 8259CL 2.50GHz x 2core
メモリ	4GB
ネットワーク帯域	5Gbps

表 3 復号データへのアクセス評価結果

項目番号	評価内容	アクセス結果
1	コンテナ内データ(復号化データ)に対するホスト外からのアクセス	不可（正常）
2	他利用者のコンテナからのアクセス	不可（正常）
3	自分のコンテナで、他のコンテナからのアクセス	可（正常）

３）プログラムの汎用性評価

セキュアコンテナ上のデータを OS 標準コマンドで利用した．標準コマンドは認証処理や復号処理は含まない汎用的なプログラムであるが，問題なく暗号化データの読み込みに成功し，データ利用プログラムの汎用性を確認した．

5. まとめと今後の課題

DRM で保護されたデータを安全に利用するデータ利活用環境であるセキュアコンテナを試作し，その有効性について検証し一定の効果を確認した．しかし，本稿執筆時点で実施した評価項目については基本的なものである為，今後も継続して評価を実施する．また，市場に存在するデータやユースケースを用いて実用的な検証を実施予定である．

6. 参考文献

[1] 坂本久他，組織が保有する情報を他組織に安全に共有し，複数組織でデータの価値を向上させる「データ価値共創プラットフォーム，情報処理学会研究報告情報基礎とアクセス技術，Vol.2021-IFAT-142，No.10，pp.1-8 (2021).

[2] Kubernetes, https://kubernetes.io/ja/ (参照 2022/1/6)

[3] FUSE, https://github.com/libfuse/libfuse (参照 2022/1/6)

[4] IPA, アプリケーションコンテナセキュリティガイド, https://www.ipa.go.jp/files/000085279.pdf (参照 2022/1/6)