

# 表形式データをグラフ形式データモデルとして 利用するためのデータ移行方式

古跡 進† 田中 覚† 楓 仁志†

三菱電機株式会社 情報技術総合研究所†

## 1. インフラマネジメントシステム

インフラ設備の老朽化が進む日本では、事故リスクの顕在化や修繕維持費の増大が問題となっている[1]。一方で、近年の製造業では、様々なデータを連携させて活用することで、生産性の向上や新たな付加価値を生み出すデジタルトランスフォーメーション(DX)と呼ばれる取り組みが推進されている[2]。

そこで、本稿ではインフラ設備の保守効率を向上させるため、図1のようなインフラマネジメントシステムを考える。このシステムは、機器メーカーが保有する生産情報と、インフラ事業者が保有する機器の運用情報をデータ共有基盤に集め、分析する。このように、互いの情報を活用することで、機器メーカーにおける生産品質向上や、インフラ事業者における予防保全による事故防止、及びメンテナンス効率の向上などが期待できる。

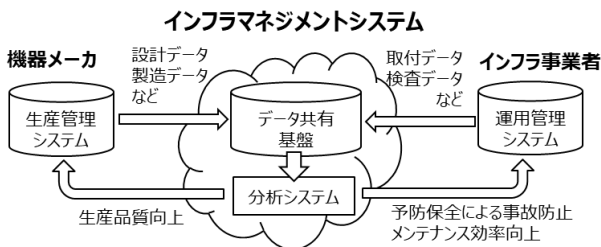


図1 インフラマネジメントシステムの概要

## 2. データ移行の目的と課題

### 2.1. 複数事業者間でのデータ共有方式

本稿では、上記システムを構築するために、各事業者は連携するデータを複製して、データ共有基盤へ移行させると想定する。データ共有基盤ではデータを実体として管理する。そのため、各事業者は初回に初期データを移行し、定期的に増分データを移行する必要がある。

また、データ共有基盤では、各種機器データの時間的、物理的な関連を柔軟に管理するため、機器をノードとし、機器間の関係をリンクとしたグラフ形式のデータモデルを用い、ノードとリンクが分離されたデータとして管理する。これにより、例えば故障情報が報告された部品のリンクを辿ることで、同じ製造ロットの部品を搭載している設備の検索などが容易になる。

### 2.2. データ移行の要件

目的のデータ移行におけるビジネス上の2点の制約について述べる。1点目は連携事業者向けのデータのみを移行することである。2点目は移行元データベース(DB)で使用されるリレーショナルモデルでは、ノードとリンクが同じテーブルで管理されており、ノードとリンクを分離するようなモデル変換が必要なことである。

以上を踏まえ、目的のデータ移行では、移行元テーブルにおいて特定の条件(起点データ)からリンクを辿れるデータを抽出し、モデル変換した上で複製して移行する方式を提案する。これにより、例えば連携事業者に紐づくデータのみを移行することや、新規注文に紐づくデータを抽出することで増分データの移行を実現できる。データ移行の具体例を図2に示す。図では、移行元テーブル(a)の起点テーブルから、リンクを辿れるデータを移行対象としている。この時、モデル変換を行うと移行先テーブル(b)となる。

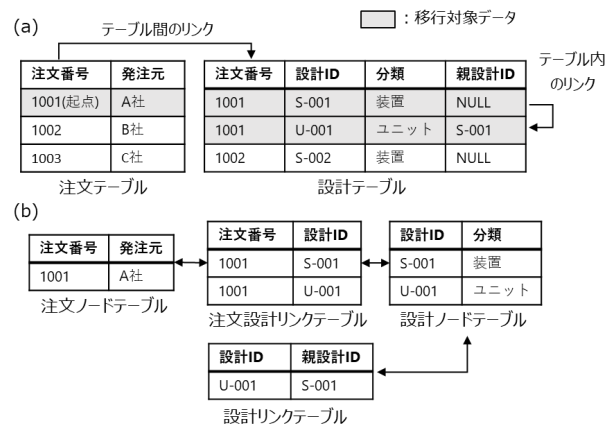


図2 移行元(a)から移行先(b)へのデータ移行

Data migration method for handling relational data as a graph data model

† Susumu Koseki, Satoru Tanaka, Satoshi Kaede  
Information Technology R&D Center, Mitsubishi Electric Corporation

2.3. 開発における課題

企業におけるそれぞれの業務システムは、異なる機能要件に基づいて設計されているため、複数システム間のデータを繋げるためのリンク条件が設定されていないことが多い。例えば、テーブル間での参照制約などが存在せず、データのリンクを辿るような検索を行う場合には、テーブルの結合条件を手動で管理しなければならない。また、本システムでは、BOM(部品表、Bill of Material)と呼ばれるような機器構成に関するデータを主に対象としている。このようなデータは規模の大きい階層構造を持つため、リンクが複雑になりやすい。

目的のデータ移行ではデータのリンクを辿るような検索処理が必要であるが、データの複雑なリンク条件を手動で管理することは、システム改修時などの開発効率の低下に繋がる。

3. 提案するデータ移行方式

3.1. 提案手法の概要

開発効率における課題を解決するため、データのリンク条件などをメタ情報として管理しておき、処理に関する情報と組み合わせて利用することで、データ移行システムを効率的に開発する手法を提案する。

まず、目的のデータ移行を実現するための具体的な処理方式について述べる。目的のデータ移行は、起点データからリンクを辿れるデータを検索して抽出する処理、抽出されたデータをモデル変換する処理、及び変換されたデータを転送する処理に分割できる。本手法では、抽出データを中間結果として中間テーブルに登録しておき、その後の処理で利用する。図3は起点データからリンクを辿れるデータを検索し、中間テーブルに登録する処理を示している。中間テーブルはノードとリンクの情報を含む形式のため、以降の変換処理などが可能である。

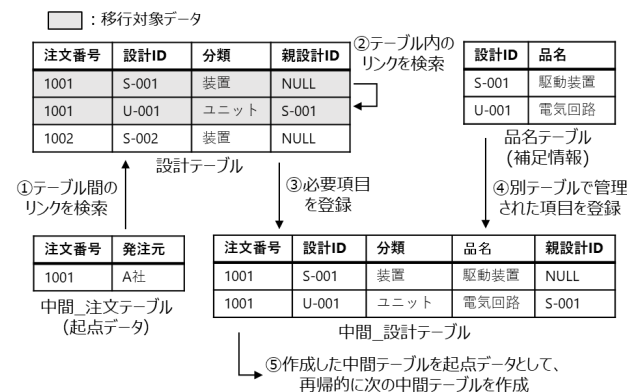


図3 中間テーブルへの抽出データ登録手順

3.2. データ抽出におけるメタ情報の利用

中間テーブルへ抽出データを登録するための処理(図3)は、SQL クエリにより実行することができる。また、各中間テーブルへデータを登録するためのそれぞれの処理は、図中の手順①から④で表される一定の手順で実現できる。そのため、各 SQL クエリは一定形式の構文となり、メタ情報を用いて SQL クエリを生成できる。メタ情報としては、手順①から④に対応する、テーブル間及びテーブル内のリンク情報と、移行するデータ項目などのデータ移行を実行するためのノード情報が必要となる。

3.3. 期待される効果

リンク条件などのメタ情報を一元管理しておき、その情報を基にデータ抽出 SQL クエリを生成することで開発の手間を削減できる。システム改修時などには、人手によるソースコードの修正量を抑えることができ、システムの保守性を向上できると考えられる。

また、提案する SQL クエリ生成機能を ETL ツールと組み合わせることも考えられる。ETL ツールとは、GUI(Graphical User Interface)を用いて、ローコードによるデータ移行処理の開発が可能なツールである。メタ情報を管理する記憶部と、そこから SQL クエリを生成するプログラムを ETL ツールの拡張機能として組み合わせることで、データ移行の開発効率をさらに高めることができると考えられる。

4. まとめ

本稿では、データのリンク条件などをメタ情報として一元管理しておき、その情報から、移行対象データ抽出のための SQL クエリを生成するデータ移行方式を提案した。この移行方式は、移行元テーブルにおいてデータのリンク条件が複雑な場合などに、開発効率の面で有効である。移行先のデータ共有基盤では、グラフ形式のデータモデルとなっているため、機器についてデータのリンクに基づく分析が容易となり、インフラ設備の保守効率化に貢献できると考えられる。今後は、必要なメタ情報について整理し、情報の活用範囲の拡大について検討する。

参考文献

[1] 経済産業省. 製造業の企業変革力を強化するデジタルトランスフォーメーション(DX)の推進. 2020年版ものづくり白書. 2020.  
 [2] 国立研究開発法人 科学技術振興機構. インフラ技術総覧. 2019.