

메타데이터 관리시스템의 이관 관리 프로세스 개선을 위한 테이블 형상 비교 알고리즘 설계

양승연*, 박석천**, 문승식***, 이진희***, 인태환***

*가천대학교 모바일소프트웨어학과

**가천대학교 컴퓨터공학과 (교신저자)

***(주)데이터스트림즈 DA본부

e-mail:201240184@gc.gachon.ac.kr

Design of Table Shape Comparison Algorithm for Improvement of Transformation Managing Process in Meta Data Management System

Seungyeon Yang*, Seok-Cheon Park**, Seung-Sic Moon***, Jinhee Lee***, Teahwan In***

*Dept. of Mobile Software, Gachon University

**Dept. of Computer Engineering, Gachon University

***DataStreams DA Headquater

요 약

오늘날 기업의 시스템은 훨씬 복잡해지고 다루는 데이터의 규모는 점점 대형화됨에 따라 환경 변화에 기업이 대응하기가 어려워지고 있다. IT 환경의 변화에 기업이 대응하기 위해서는 메타데이터 관리시스템의 기능을 개선할 필요가 있다. 본 논문에서는 데이터 모델링에 대한 새로운 요구사항을 신속히 반영하기 위해서 모델의 변경 이력을 참조하지 않고 이관해야 할 목표 데이터베이스의 테이블의 형상을 비교하여 이관 관리 프로세스를 향상시킨 알고리즘을 설계하였다. 본 논문서 설계한 알고리즘을 통해 모델의 변경 이력을 참조할 경우 발생하는 데이터베이스 변경의 처리 횟수를 줄임으로써 프로세스 개선의 효과를 기대할 수 있는 알고리즘을 제시하였다.

I. 서론

오늘날 기업에게 데이터란 시간이 지날수록 가치가 상승하는 자산과 같다. 데이터 관리의 목적이 과거 쓸모 있는 데이터를 구별해내는 것으로부터 새로운 가치를 창조하기 위한 정보 자원관리로 변화되었다. 데이터는 정적인 상태가 아니라 천연 자원의 흐름처럼 유동적으로 흘러가고 있다. 즉, 환경이 변화하면 데이터도 변화하게 되는 것이다. 데이터가 변화하게 되면 기업은 이에 신속하게 대응하여야 한다. 그렇지 않으면 데이터를 잘 쌓아 놓기 어려우며, 어느 순간 가치를 잃게 된다. 데이터를 신속하고 정확하게 관리하기 위해서 기업들은 여러 가지 관리시스템을 도입하게 되었다.

IT 환경의 변화에 대해 기업이 대응하기 위해서는 필히 메타데이터 관리시스템의 기능을 개선할 필요가 있다. 새로운 개념의 데이터가 생성되고 형태가 변화하며 언젠가는 시대의 흐름에 따라 사라지게 된다. 기업의 데이터는 계속 변화하기 때문에 데이터를 관리하기 위해서는 메타데이터의 변경 관리가 필요하다. 이러한 요구사항의 증가에 따른 연쇄적인 필요 작업으로 메타데이터 관리시스템은 메타데이터 관리 프로세스를 반복적으로 수행한다. 표준을 변경하여 모델을 만들고 실제 데이터베이스에 반영하는 일련의 과정을 반복하게 된다. 반복하는 횟수는 변경 횟수에 비례하게 되므로 프로세스의 효율적인 운영은 메

타데이터 관리시스템의 성능개선의 필수적인 요구사항이 된다.

메타데이터의 변경을 반영하는 프로세스를 관리하는 것은 이관관리에 해당된다. 모델이 변경 되면 데이터베이스에 반영하기 위해 표준 요소들로 구성된 테이블을 이관 신청하게 되고 업무 관련자의 검토를 거쳐 테이블 이관이 실행된다. 이 과정에서 반복적인 프로세스를 실행해야 하는 비효율적인 측면이 존재한다. 기존의 연구에서는 메타데이터 관리시스템의 데이터 표준관리에 초점을 맞추어 기능개선을 하였지만 현재의 급변하는 환경에서 메타데이터 관리시스템을 적극적으로 활용하기 위해서는 이러한 불필요한 프로세스를 개선하여 데이터를 보다 효율적으로 관리할 필요가 있다.

본 연구에서는 데이터 모델링에 대한 새로운 요구사항을 신속히 반영하기 위해서 모델의 변경 이력을 참조하지 않고 이관해야 할 목표 데이터베이스의 테이블의 형상을 비교하여 프로세스의 반복적인 절차를 개선한다. 특히 모델의 변경 이력을 참조할 경우 발생하는 데이터베이스 변경의 처리 횟수를 줄임으로써 작업량 감소의 효과를 기대할 수 있다. 또 한 테이블 형상 비교를 통해 모델과 동일한 변경과정을 거쳐야만 유지할 수 있었던 모델과의 정합성도 지키며, 계층별 데이터베이스 관리를 통해 보다 안정적인 시스템 운영을 할 수 있다.

II. 관련 연구

본 장에서는 메타데이터 관리시스템의 이관관리 프로세스 개선을 위한 테이블 형상 비교 알고리즘을 설계하기 위해 테이블 형상 비교와 테이블 이관에 대한 정의를 설명한다. 다음으로 메타데이터 관리시스템의 구성과 기능에 대해 알아보고 이관관리의 기능과 기존 이관 관리 프로세스의 문제점에 대해 알아본다.

2.1 테이블 형상 비교 정의

테이블 형상 비교란 두 테이블 간의 형상 즉, 테이블의 형태를 이루는 요소를 비교하는 것이다. 비교 대상이 되는 요소는 테이블, 컬럼, 그리고 인덱스이다. 모델에 생성된 테이블과 운영 데이터베이스 사이에 존재하는 테이블의 형상 차이를 비교하여 새롭게 정의된 내용은 신규, 변경된 내용은 수정, 모델에서 지워진 내용은 삭제로 구분한다. 테이블 형상을 비교할 때 중요한 것은 비교하는 시점의 모델의 변경이 반영된 개발 데이터베이스의 테이블 형상을 추출하는 것이다. 그리고 모델과 개발 데이터베이스는 동일하다고 간주하고 변경 시점의 개발 데이터베이스의 테이블 형상을 추출한다. 모델은 지속적으로 변경되고 있기 때문에 한 시점의 형상을 추출해야 한다.

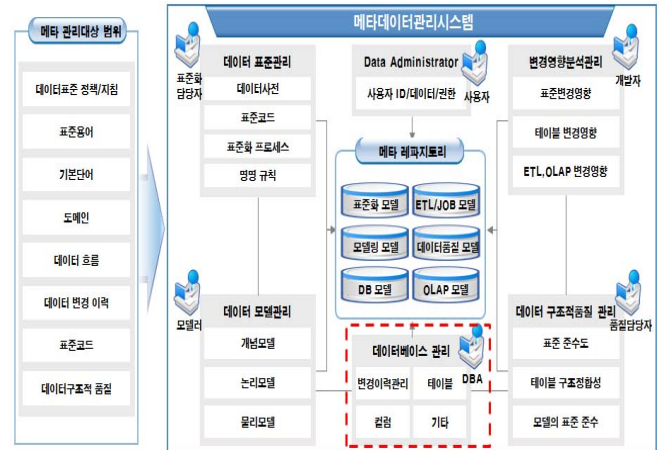
2.2 테이블 이관

테이블 이관이란 다른 데이터베이스 간에 테이블의 형태를 옮기는 것이다. 데이터를 저장하기 위한 테이블의 형태를 모델로 생성하고 실제 데이터가 저장되는 데이터베이스에 모델에서 만든 테이블 양식을 동일하게 생성하는 것이다. 이때 이관은 틀을 옮긴다는 의미로 사용한다. 메타데이터 관리시스템에서 테이블 이관을 실행할 때 데이터베이스를 조작하기 위한 DDL을 생성하였다. 즉, 모델의 변경 횟수와 데이터베이스의 변경 횟수가 동일하다고 생각할 수 있다. 기존의 시스템에서는 모델과 데이터베이스 간의 정합성 확보를 위해 모델의 변경이력을 참조하여 DDL을 생성하였다. 즉, 모델의 변경 횟수와 테이블 이관 횟수가 동일하다고 생각할 수 있다. 이와 같이 테이블 이관 프로세스가 진행될 경우 테이블 이관의 관리적 의미가 국한되며, 단지 모델과 데이터베이스 간의 정합성 확보라는 목표만을 달성하게 된다. 또한 모델과 데이터베이스 간의 물리적인 형태가 동일하지 않을 경우 테이블 이관 시 이를 반영하기 위한 추가적인 노력이 필요하게 된다.

2.3 메타데이터 관리시스템

메타데이터는 일반적으로 데이터관리를 위한 데이터로서 일반적으로 기업에서는 업무 활동에서 다루는 데이터를 정의하고, 분석하기 위해 조직된 정보라고 할 수 있다 [1]. 메타데이터 관리시스템은 이러한 메타데이터를 사용자 측면과 시스템 측면 사이에서 요구사항을 반영하며 데이터의 흐름과 변화를 분석한다[2]. 또한 각 시스템의 메타데이터를 추출하고 가공하여 메타데이터 데이터베이스

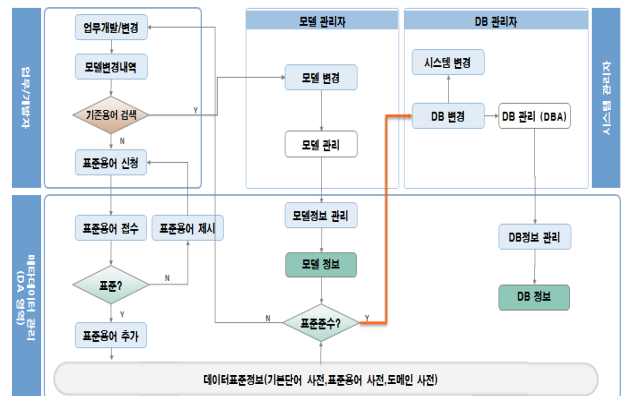
에 저장하고 사용자들이 쉽게 조회할 수 있도록 한다. 향상 데이터의 일관된 관리 및 지속적 품질관리를 목표로 하고 비즈니스 데이터에 대한 이해도를 향상시키고 데이터에 대한 중앙관리를 가능하게 하며 표준화된 용어로 사용자와 인터페이스를 설계함으로써 운용시스템의 활용도를 향상시키는 기능이 있다[3]. [그림 2.1]은 메타데이터 관리시스템의 구성도이다.



(그림 2.1) 메타데이터 관리시스템 구성도

2.4 이관관리 프로세스

메타데이터 관리시스템에서 이관관리는 아래 [그림 2.2]와 같은 프로세스를 거친다.



(그림 2.2) 이관관리 프로세스

업무 담당자 의해 업무개발 또는 기존의 업무의 변경이 발생된다. 수립한 데이터 표준 지침을 검토하여 모델에 반영이 되면 개발 데이터베이스에 반영이 되어 검토하게 된다. 그 다음 업무 개발자가 필요에 의해 테스트 데이터베이스로의 이관 신청을 하면 결재 과정을 거쳐 DBA에게 이관 실행을 지시하게 된다. DBA는 메타데이터 관리시스템에서 제공하는 DDL 생성 기능을 통해 획득한 DDL을 해당 DBMS에 적용시켜 테스트 DB를 변경한다. 변경하는 과정에서 DBA의 수작업에 진행되는 부분은 모델과의 정합성 확보에 위험 요소로 작용되고 있다.

2.5 기존 이관관리 프로세스의 문제점

기존의 이관관리 프로세스는 여러 가지 문제점이 있다. 먼저 현재 기업의 환경적인 측면에서 신규시스템이나 변경사항이 발생하여 모델에 대한 변경작업이 빈번하게 발생할 경우 신속하고 정확하게 대응하기 어렵다. 모델을 기준으로 물리 DB를 관리할 경우 먼저 동일하지 않은 환경에서는 DBA의 수작업에 대한 위험성이 존재하며, 모델의 변경 이력을 관리하더라도 실제 물리 DB와 동일하지 않다면 실제에서는 유용한 이력정보가 될 수 없다. 이러한 상황에서 모델과의 정합성을 최대한 확보하기 위해서는 모델의 변경 작업과 순서와 동일하게 이관을 실행하는 것이다. 이 과정에서 또 하나의 문제가 발생한다. 즉 한건의 변경 작업이 진행될 경우 추가적인 변경작업은 진행될 수 없다. 한 건이 완료되어야만 다음 단계를 진행할 수 있다. 만약 모델의 경우 생성과 수정의 과정을 거쳤다면 이관도 생성과 수정을 처리해야 한다. 만약 이 프로세스를 개선하여 수정된 내용으로 한번만 생성하게 된다면 상당한 업무 개선 효과가 발생할 것이다.

III. 테이블 형상 비교 알고리즘 설계

본 장에서는 테이블 형상 비교 알고리즘을 적용한 테이블 이관 프로세스의 전체적인 설계를 기술하였다. 테이블 형상 비교 알고리즘을 적용한 시나리오를 정의하고 전체적인 데이터의 흐름을 흐름도를 통해 파악하고 본 알고리즘을 상세하게 설계하였다.

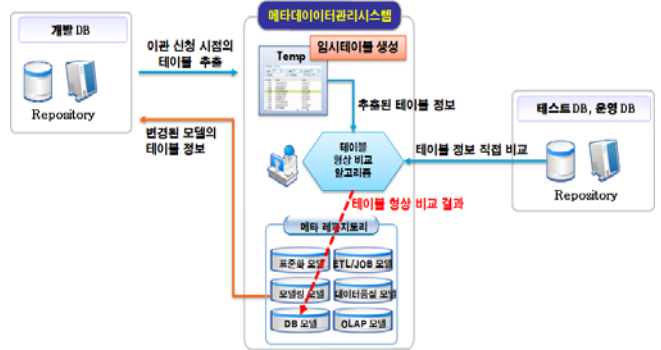
3.1 요구사항

테이블 형상 비교 알고리즘에 대한 요구사항을 도출한다. 테이블 형상 비교하기 위해 먼저 개발 DB로부터 테이블 이관 시점의 테이블을 추출한다. 모델 변경이 바로 적용되는 개발 DB는 계속 새로운 변경이 진행되고 있으므로 테이블 이관 시점의 테이블을 추출하여 고정된 테이블의 형상을 지정할 필요가 있다. 이관할 목표 DB의 테이블과의 형상을 비교한다. 테이블의 구성 요소인 테이블, 컬럼, 인덱스의 신규, 수정, 삭제 사항을 비교하여 해당 사항에 대해 플래그 체크를 하여 사용자에게 보여준다. 이관 신청 시 플래그 체크 된 내용을 참조하여 DDL생성 한다. 테이블 형상 비교한 내용을 바탕으로 DDL문이 생성되면 최종 변경 내용만을 테이블 이관 처리하게 된다. 테이블의 구성 요소의 변경 이력을 기록하여 전후 비교하여 볼 수 있도록 하여야 한다. 테이블 이관을 통해 변경된 내역을 확인 가능하도록 하며, 테이블 이관 실행에 관한 이력도 각 계층별 DB에 따라 기록하여 관리하도록 한다.

3.2 테이블 형상 비교 알고리즘

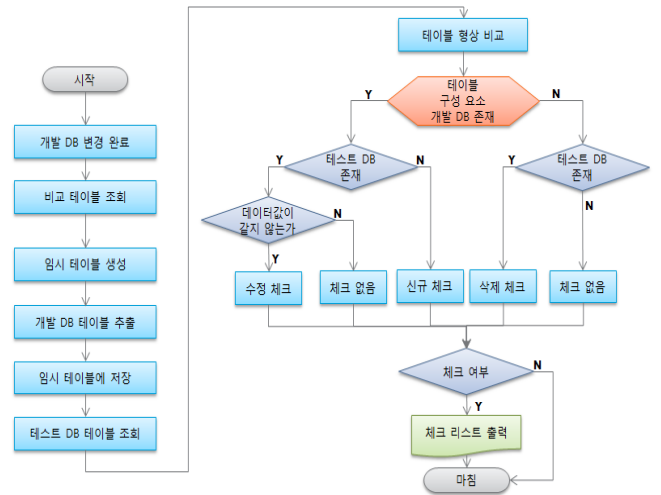
모델의 변경과 동시에 개발 DB에 반영된다는 가정 하에 본 알고리즘을 설명하고자 한다. 테이블 형상 비교 알고리즘의 데이터 흐름도는 [그림 3.1]과 같이 설계하였다.

각 계층의 데이터베이스와 메타데이터 관리시스템의 리포지토리 사이에서 데이터 흐름을 설명하였다.



(그림 3.1) Data Flow Diagram

테이블 형상 비교 알고리즘은 [그림 3.2]와 같이 설계하였다.



(그림 3.2) 테이블 형상 비교 알고리즘

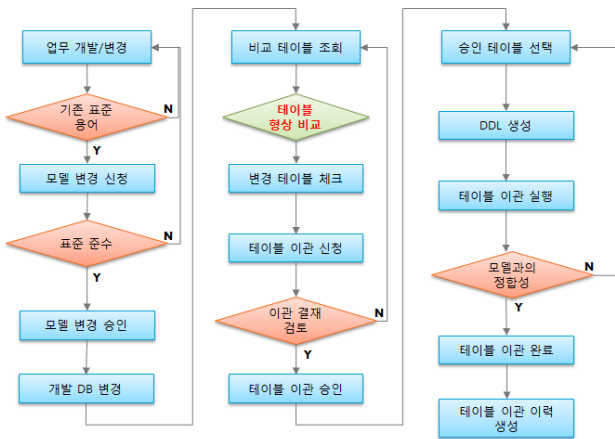
개발 DB 변경이 완료되면 사용자에게 의해 테이블 이관을 하고자 비교 테이블 조회를 한다. 여기서 비교 테이블이란 개발 DB의 테이블과 테스트 DB의 테이블을 말한다. 테이블 조회를 하면 메타데이터 관리시스템 내 DB모델에 임시테이블이 생성이 되며 여기에 개발 DB의 테이블이 추출되어 저장된다. 그 후 테스트 DB의 테이블을 조회한 후 테이블 형상을 비교한다. 테이블 구성 요소인 테이블, 컬럼, 인덱스에 대해 각각 비교한다.

테이블 구성 요소에 대해 먼저 개발 DB에 존재여부를 확인한다. 다음으로 테스트 DB에 존재여부를 확인한다. 개발 DB에 존재하며 테스트 DB에도 존재하는 경우는 데이터 값이 같지 않는지 확인한다. 데이터 값이 같지 않을 경우 '수정'이라고 플래그 체크를 한다. 데이터 값이 같을 경우는 플래그 체크를 하지 않는다. 이 경우는 동일한 데이터 값으로 변경사항에 해당되지 않는다. 개발 DB에 존재하며 테스트 DB에 존재하지 않는 경우는 새롭게 생성된

경우로 '신규'라고 플래그 체크 한다. 마지막으로 개발 DB에 존재하지 않으며 테스트 DB에 존재하는 경우는 해당 테이블 구성 요소가 삭제된 경우이다. '삭제'라는 플래그 체크를 한다. 반대로 개발 DB에 존재하지 않으며 테스트 DB에도 존재하지 않는 경우는 존재하는 테이블 구성 요소가 아니므로 플래그 체크의 사항에 해당되지 않는다. 비교에 의한 플래그 체크 여부가 끝나고 나면 플래그 체크 여부를 확인한다. 플래그가 존재하는 항목만을 모아서 사용자의 조회 화면에 출력하도록 한다.

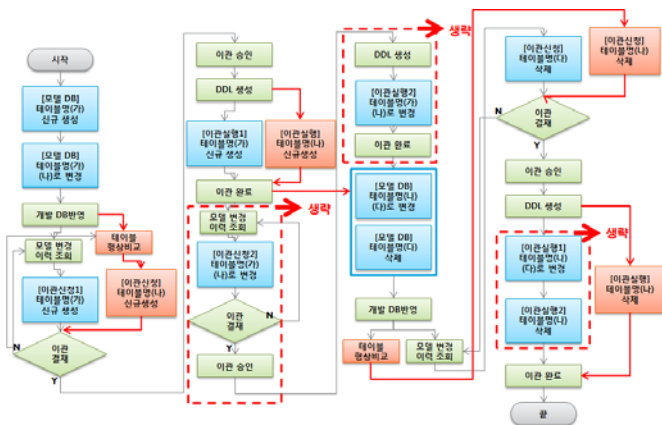
3.3 테이블 이관 프로세스

테이블 형상 비교 알고리즘이 적용된 테이블 이관 프로세스에 대해 정의한다. 테이블 이관 프로세스는 메타데이터 관리시스템의 전체 프로세스에 관련된 사용자와 관련이 되어 있다. 테이블 이관 프로세스는 아래 [그림 3.3]과 같이 설계하였다.



(그림 3.3) 테이블 이관 프로세스

다음 [그림3.4]는 테이블 형상 비교 알고리즘을 적용한 시나리오를 통해서 테이블 이관 프로세스가 개선됨을 알 수 있다. 기존의 프로세스에서는 2회의 프로세스를 거쳐야 하지만 본 알고리즘을 적용하면 1회의 프로세스로 감소하는 것을 확인하였다.



(그림 3.4) 테이블 형상 비교 적용 이관 프로세스 시나리오

4. 결론

IT 환경 변화에 신속하게 대응하기 위해 기업의 시스템의 성능을 개선하는 한 방법으로 메타데이터 관리시스템의 이관관리 프로세스를 개선하는 것을 제안하였다.

본 연구에서 제안하는 테이블 형상 비교 알고리즘은 기존의 모델의 변경 이력을 기준으로 테이블 이관을 실행했던 방식에서 나타나는 반복적인 테이블 이관 처리와 테이블이관 신청 프로세스를 개선하였다. 이관 대상의 데이터베이스의 현재 테이블 형상과 모델로부터 추출한 테이블의 형상을 비교하여 차이에 대해 변경하도록 DDL문을 생성하여 DBA에게 제공함으로써 모델의 변경 횟수에 상관없이 최종 사항에 대해서만 테이블 이관을 함으로써 테이블 이관 처리상의 성능향상과 테이블 이관 신청 회수의 감소에 따른 업무 효율성 제고 효과를 얻을 수 있었다. 또한 모델과 데이터베이스의 정합성 유지기능도 테이블 형상 비교 알고리즘을 통해 편리하게 관리 할 수 있다. 이로 인해 데이터에 대한 신뢰성이 향상 될 수 있다.

테이블 형상 비교 알고리즘을 통해 테이블 변경 전후 비교 기능을 구현하여, 산출물로 활용할 수 있는 가용성이 확보되었으며, 실제 데이터베이스 이관 이력을 기록할 수 있게 되어 이관 이력관리 기능을 제공할 수 있도록 메타데이터 관리시스템의 기능적 확장을 고려할 수 있게 되었다. 마지막으로 계층별 데이터베이스 관리가 가능하여 완성된 모델을 만들고 테스트 데이터베이스에서 성능을 검증하여 실제 운영 데이터베이스에 반영하는 안정적인 데이터베이스 시스템이 될 수 있다.

앞으로의 IT 환경은 더욱 빠르게 변하게 될 것이며, 기업의 시스템은 점점 복잡해지고 대형화될 것이다. 메타데이터 관리시스템의 기능 개선을 위해 본 연구에서 제안하는 알고리즘의 적극적인 활용이 필요하며, 데이터 품질 관리시스템과의 연계를 통해 통합적인 데이터 관리가 이루어져야 앞으로의 환경에 대응해 나갈 수 있을 것이다.

사사의 글

본 연구는 2013년도 지식경제부의 SW전문인력양성사업의 재원으로 정보통신산업진흥원의 고용계약형 SW석사과정 지원사업(HB301-13-1003)으로부터 지원받아 수행되었습니다.

참고문헌

- [1] 조상혁 "대형 금융정보 데이터 모델링을 위한 메타데이터 시스템의 설계 및 구현" 2012 한국컴퓨터정보학회지 17(1) 81-85
- [2] 김광우 "DB 물리모델 적용 시 DDL트리거를 이용한 표준데이터 준수를 상향방법: 데이터품질관리 성숙도 기반" 2012 서강대 석사학위논문
- [3] 정재군 "대형 금융정보 Data 처리를 위한 DB I/O 모듈 생성엔진 개발" 2009 한국 정보과학회 가을 학술발표 논문집 Vol.36, No.2(B) 66-70