GDBMS를 위한  
 이관 도구 구조 설계

V0.2



**개정 이력**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **버전** | **일자** | **내용** | **작성자** | **승인자** | **비고** |
| V0.1 | 2021-10-25 | 최초 문서 작성 | 마인기 |  |  |
| V0.2 | 2021-10-29 | 설계 요구 사항 상위 페이지로 이동 | 마인기 |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

목차

[1 개요 3](#_Toc86420185)

[2 설계 요구 사항 4](#_Toc86420186)

[3 구조 및 세부내용 7](#_Toc86420187)

[3-1. 구조 7](#_Toc86420188)

[3-2. 관련 내용 8](#_Toc86420189)

[3-2-1. Gremlin (Project TinkerPop – by Apache) 8](#_Toc86420190)

[3-2-2. CUBRID Migration Toolkit (큐브리드) 9](#_Toc86420191)

# **1** 개요

본 문서는 IITP-차세대 DBMS 과제 중 TurboGraph++ 를 위한 이관 도구를 개발하기 위한 기초 구조설계 문서이다.

이관 도구는

* JAVA를 개발되며 Eclipse RCP(Rich Client Platform) 기반인CUBRID Migration Toolkit을 기반으로 개발되어, 다양한 플랫폼에서 사용 가능토록 한다.
* CUBRID Migration Toolkit을 이용하여 CUBRID의 데이터를 가져온 후 TurboGraph++로 데이터 이관할 수 있는 프로그램이다.
* Open Source 응용 도구 및 Library를 사용한다.

본 도구는 CUBRID에서 TurboGraph++ 로 이관 기능을 주 목적으로 한다.

IITP-차세대 GDBMS인 TurboGraph++와 연결하여 결과를 산출하는 것이 최종 목표이며,

TurboGraph++에 인터페이스가 완료되기 전까지는 Apache TinkerPop 기반에 Graph

Database(Janus Graph)를 이용하여 개발한다.

# 2 **설계 요구 사항**

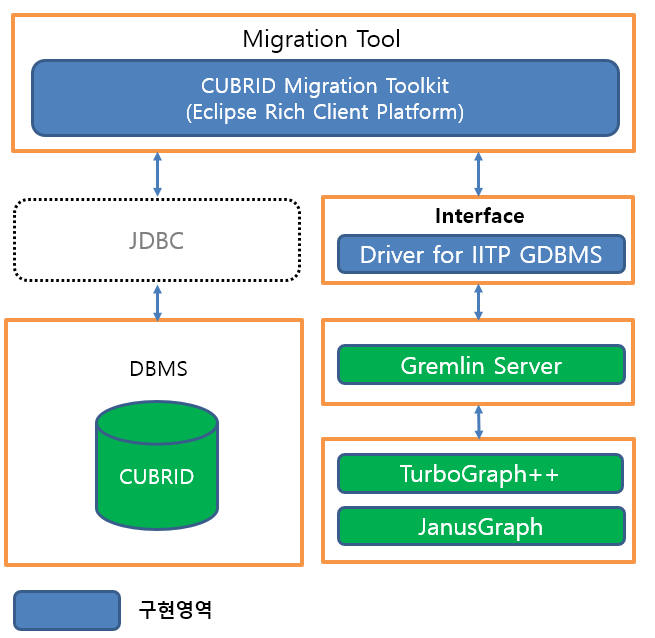
구조 설계에 앞서 요구 사항을 추출하여 이를 기반으로 요구사항 명세서를 작성하고,

이를 토대로 분석, 설계 후 개발한다.

|  |  |  |
| --- | --- | --- |
| **요구사항 ID** | **요구사항명** | **요구사항 내용** |
| RG-TRAN-01 | 메타데이터 획득 | RDB의 meta data를 가져온다 |
| RG-TRAN-02 | 메타데이터를 모델로 변환 | 가져온 Meta data 기반으로 RDB to GDB 변환(매핑) 기능을 제공한다 |
| RG-TRAN-03 | 모델 변환 시각화 | Mapping 된 모델의 시각화 정보를 제공한다 |
| RG-TRAN-04 | Vertex 속성 설정 | 시각화 된 그래프에서 Vertex 속성 설정/변경할 수 있다 |
| RG-TRAN-05 | Vertex 속성 Skip 설정 | 이관 시 불필요한 Vertex를 Skip 설정할 수 있다 |
| RG-TRAN-06 | Edge 속성 설정 | 시각화 된 그래프에서 Edge 속성 설정/변경할 수 있다 |
| RG-TRAN-07 | Edge 속성 Skip 설정 | 이관 시 불필요한 Edge를 Skip 설정할 수 있다 |
| RG-TRAN-08 | Edge 연결 변경 설정 | 시각화 된 그래프에서 Edge 연결을 변경할 수 있다 |
| RG-TRAN-09 | Edge 연결 변경 설정 | FK가 없는 경우 relationship 자동으로 찾아 추천하는 방법 |
| RG-MIGR-01 | Export Data | Meta data에 따라서 RDB Data를 export 한다 |
| RG-MIGR-02 | Import Data | RDB data를 GDB에 online으로 import 한다 |
| RG-MIGR-03 | Export file for GDB | RDB data를 GDB 도구가 import 할 수 있는 형태의 File로 생성하는 기능을 제공한다 |
| RG-MIGR-04 | 증분 이관 | RDB의 증분 데이터 이관 기능을 지원한다. - CUBRID에서 특정 시점 이후 변경된 데이터만 증분 이관 |
| **요구사항 ID** | **요구사항명** | **요구사항 내용** |
| GR-TRAN-01 | GDB 질의 입력 | GDB 질의를 입력 받는다 (사용자 직접 입력 or 파일) |
| GR-TRAN-02 | 질의 결과 변환 | 질의 결과를 기반으로 데이터(GDB to RDB) 변환(매핑) 기능을 제공한다 |
| GR-MIGR-01 | Import Data | 맵핑된 데이터를 JDBC Driver를 이용하여 RDB에 import 한다 |
| GR-MIGR-02 | Export file for RDB | GDB data를 CUBRID가 import 할 수 있는 형태의 File로 생성하는 기능을 제공한다 |
| CM-CONN-01 | RDB 연결 설정 : 추가 기능 | Online RDB 연결 정보를 추가하는 기능을 제공한다 |
| CM-CONN-02 | RDB 연결 설정 : 삭제 기능 | Online RDB 연결 정보를 삭제하는 기능을 제공한다 |
| CM-CONN-03 | RDB 연결 설정 : 수정 기능 | Online RDB 연결 정보를 변경하는 기능을 제공한다 |
| CM-CONN-04 | GDB 연결 설정 : 추가 기능 | Online GDB 연결 정보를 추가하는 기능을 제공한다 |
| CM-CONN-05 | GDB 연결 설정 : 삭제 기능 | Online GDB 연결 정보를 삭제하는 기능을 제공한다 |
| CM-CONN-06 | GDB 연결 설정 : 수정 기능 | Online GDB 연결 정보를 변경하는 기능을 제공한다 |
| CM-CONN-07 | 연결 테스트 | Online migration시에 연결 설정을 이용하여 연결 테스트 기능을 제공한다 |
| CM-MIGR-01 | multi thread support | Migration thread의 수를 조절할 수 있다 |
| CM-MIGR-02 | Export CMT Script | CLI 로 마이그레이션 수행할 수 있도록 CMT Migration Script(XML) 파일을 Export 하는 기능 제공 (GR, RG 용) |
| CM-IF-01 | CLI 제공 | CLI를 이용한 Migration 기능을 제공한다 |
| CM-IF-02 | GUI 제공 | GUI를 이용한 Migration 기능을 제공한다 |
| CM-IF-03 | GDB interface | GDB Online migration을 위해서 Gremlin Java Client Driver를 사용한다 |
| CM-IF-04 | RDB interface | Online migration을 위해서 JDBC Driver를 사용한다 |
| **요구사항 ID** | **요구사항명** | **요구사항 내용** |
| CM-RSLT-01 | 이관 summary | 이관 결과의 summary를 제공한다 |
| CM-RSLT-02 | 에러 로그 기능 | 오류 로그를 제공한다 |

# **3 구조 및 세부내용**

## 3-1. 구조



[그림 3-1] 구조도

본 과제에서 Migration Tool 은 Eclipse Rich Client Platform으로 개발된 CUBRID Migration Toolkit(이하 CMT)을 사용하며 기존에 마이그레이션 단계, DBMS에 연결할 수 있도록 데이터 소스를 확장하는 형태로 개발 진행하도록 한다. CMT는 JDBC Driver를 사용하여 데이터를 가져오고, 가져온 데이터를 GDBMS에 이관할 수 있는 형태로 변환 및 이관한다. 기본적인 연결, 데이터 통신 등을 할 수 있는 Interface부분은 Driver와 같은 형태로 별도 분리하여 다양한 DBMS와 통신할 수 있도록 구조화한다. 여기에서 분리된 Driver는 Gremlin 언어를 이용하며 Gremlin Server를 통해 GDBMS와 통신을 하게 되며, Gremlin Server를 지원하는 GDBMS는 모두 사용이 가능하다.

(추후 해당 부분은 변경될 수 있다.)

개요에서 언급한 바와 같이 TurboGraph++에 interface가 완료되기 전까지 Gremlin Server를 지원하는 Janus Graph를 사용하여 초기 개발될 예정이다.

## 3-2. 관련 내용

### 3-2-1. Gremlin (Project TinkerPop – by Apache)

Apache TinkerPop은 Open Source 그래프 Framework로 그래프 데이터 베이스 및 그래프 분석 시스템에 대한 그래프 Computing 기능을 제공한다. 언어로는 그래프 순회(Traversal) 언어인 Gremlin을 사용한다.

Gremlin으로 작성된 쿼리는 Vertex 또는 Edge를 선택한 뒤, 메서드를 여러 개 이어 붙이면서 데이터를 조작하는 방식으로 쓰인다. 아래는 그 예시이다.

[예시]

g.V().hasLabel('movie').values('year').min()

==>1919

Apache TinkerPop은 Gremlin과 통신을 위한 Gremlin-Server를 제공한다.

본 과제에서는 Gremlin-Server를 Interface로 사용할 예정이다.

### 3-2-2. CUBRID Migration Toolkit (큐브리드)

CUBRID Migration Toolkit은 GUI 환경 기반에서 스키마 객체 및 데이터를 마이그레이션 하는 Open Source인 도구로서, Oracle 및 MySQL DB를 CUBRID DB로 마이그레이션하는 기능을 제공한다. 원본 데이터베이스는 Oracle 10g 이상, MySQL 5 이상 버전을 지원하며, 타겟 데이터베이스는 CUBRID 8.2.1 이상 버전을 지원한다.

텍스트이(가) 표시된 사진

자동 생성된 설명

[그림3-2-2 ] CUBRID Migration Toolkit 화면