

상용 DBMS 제품의 데이터 복제 기술을 이용한 대용량 데이터 이동 시스템

양진영,장혜정,송은정,김영곤,윤명상
KT IT본부 통합솔루션부
(jyyang^o,hjjang,lusia,ygkim85,msyoon)@kt.co.kr

Very Large Data Moving System Using Commercial Copy Technology

JinYoung Yang^o HyeJung Jang, EunJung Song, YoungGon Kim, MyeongSang Yoon
Integrated Solution Dept. of Information Technology Group, Korea Telecom

요약

분산 DB 환경에서 데이터 접근의 병목현상을 방지하기 위하여 테이블 단위의 중복을 허용할 때 데이터 일치성을 유지함으로써 업무처리의 정확성이 요구된다. 분산 DB 환경에서 중복 테이블들은 일정 시간 간격으로 마스터 테이블과의 데이터 일치성을 유지하기 위한 작업은 대상 서버와 해당 테이블 수의 증가에 따라 업무의 복잡성이 가중된다. 본 논문에서는 중복 테이블간의 데이터 일치성을 수행하는 작업을 데이터 이동이라고 하는 일련의 동작으로 모델링하고, 다수 서버 환경에서 다량의 중복 테이블간의 데이터 일치성을 요구하는 업무 환경에 쉽게 적용할 수 있는 데이터 이동 시스템을 소개하고자 한다.

1. 서론

대용량 데이터베이스를 갖는 분산 업무 환경에서는 데이터 접근의 병목현상을 방지하기 위하여 테이블 단위의 중복을 허용하고 있다.[1] 분산 DB 환경에서 중복 테이블들은 일정 시간 간격으로 마스터 테이블과의 데이터 일치성을 유지하여 업무 처리의 정확성을 보장하여야 한다. 중복 테이블간의 일치성 보장을 위하여 일정 시간 동안 발생한 마스터 테이블의 변경 내역을 해당 서버에 이동하고 중복 테이블에 반영하는 일련의 작업을 수행하여야 한다. 이러한 데이터 일치성 작업은 중복을 허용하는 서버와 테이블 수에 기하 급수적으로 증가하는 업무의 복잡성을 가지고 있다. 이러한 문제를 해결하고자 중복 테이블간의 데이터 일치성을 수행하는 작업을 데이터 이동이라고 하는 일련의 동작으로 모델링하고, 다수 서버 환경에서 다량의 중복 테이블간의 데이터 일치성을 요구하는 업무 환경에 쉽게 적용할 수 있는 데이터 이동 시스템을 구축하였다.

대용량 데이터 이동시스템에서는 이동 대상의 업무 특성에 따라서 이동 방법 및 이동 시기를 정의하였다. 이동 방법은 본 시스템의 성능분석을 통하여 데이터 변동율에 따른 소요시간을 측정하여 측정 결과에 따라 이동방식을 선택하여 성능향상을 도모하였으며, 이동절차의 복잡성에 따른 사용자 오류를 방지하고 운영자의 작업 실수를 최소화하기 위해 이동작업을 routine화하여 시스템에 반영하였다. 대용량 데이터 이동시스템은 작업모니터링, 작업프

로세스 관리/ 이동대상 관리, 이동실행, 통계관리등 4개의 서브시스템으로 구성되어 있으며, 데이터 이동의 실제 실행 프로세스는 시스템 사용의 효율화를 고려하여 각 서버로 분산하였다.[2] 이 시스템은 KT의 통합 빌링 시스템인 ICIS(Integrated Customer Information System)에 적용하여 운용업무의 자동화와 안정성을 향상하였다.[3]

본 논문의 구성은 다음과 같다. 제2장에서는 대용량 데이터를 이동시키는 시스템에 대한 요구사항과 시스템 구성에 대해서 살펴본 후 제3,4장에서는 시스템 구축에 필요한 스키마 설계와 구축시 고려할 사항에 대한 테스트 등을 소개한다. 마지막 5장에서는 결론을 맺고 향후 과제를 제시한다.

2. 대용량 데이터 이동 시스템 구성요소

대용량 데이터 이동시스템은 이동 대상과 해당 시스템이 증가할수록 이동되는 데이터량은 기하급수적으로 늘어나기 때문에 분산 시스템간의 데이터 일치성을 보장하면서 이동절차의 복잡성에 따른 사용자 오류를 방지할 수 있도록 운영절차가 routine화 되어 있어 이동이 원활하여야 한다. 본시스템은 그림1과 같이 작업 모니터링, 작업 프로세스 관리/이동 대상관리, 이동 실행,통계 관리등 4가지 기능으로 나누어 대용량 데이터를 이동시 운영자의 업무의 효율성을 높이고 작업자의 실수를 최소화 할 수 있게 구성하였다.[2]

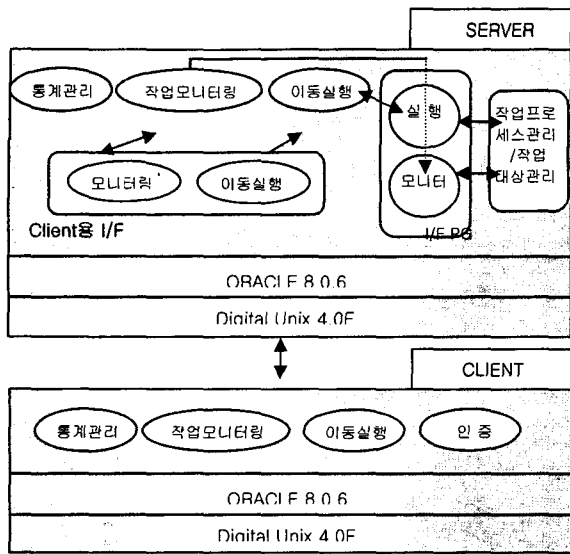


그림1S/W 구성도

2.1 작업 모니터링

작업 모니터링에서는 대상 서버의 상태, 선행 작업 및 이동 작업의 진행 과정 등의 자료를 통합 관리하여, 데이터 이동시스템을 운용하기 위한 정보를 작업자에게 제공하여 준다.

작업자는 데이터 이동 시기가 도래하면 모니터링 대상의 정보를 제공 받아 이동 작업 수행 여부를 결정하고, 만약 서버의 부하 또는 특정 작업의 미완료, 시스템/DB의 작업으로 인해 작업 수행이 불가능할 경우는 작업 보류를 등록하여 해당 작업의 수행을 보류한다.

진행 중인 이동 작업의 수행과정을 주시하면서, 오류 발생한 작업에 대해서는 오류 내역을 조회하면서 오류 조치를 하여야 한다.

2.2 작업 프로세스 관리/ 이동대상 관리

작업 프로세스 관리에서는 이동 작업 단계들을 총괄 제어하여 단계별 작업의 수행, 오류 작업의 재수행, 작업의 보류 및 예외작업을 scheduling하면서 해당 작업을 invoke 시켜주는 기능을 제공한다. 작업자는 금일 작업의 내역을 조회하여, 작업의 수행 여부를 결정하고 작업내역을 보면서 수행 예정인 작업을 보류 작업으로 전환하였을 경우에는 보류 작업에 대한 내역을 인지하여 향후 보류 작업을 수행할 수 있도록 한다. 작업 수행상의 오류내역은 원인을 분석한 후 테이블단위의 사전/이동/사후 등의 단위 작업으로 재수행하며, 재수행이 불가능할 경우에는 예외작업으로 전환하여 향후 수행한다.

이동대상 관리에서는 이동 유형에 따른 단계를 생성/수정/삭제 가능토록 단계작업 정보를 관리하며, 단계 작업에서 정의한 테이블들 단위로 병렬 수행가능토록 프로그램을 생성 관리하는 역할을 수행한다. 작업자는 각 테이블을 이동방법에 따라 SNAPSHOT REFRESH(SR), SNAPSHOT CREATE(SC), SQL COPY (C), FTP(F) 로 정

의하고, 각 테이블의 이동경로를 지정한다. 이동경로(서버)는 서버 A -> 서버 B, 서버 B -> 서버 C, 서버 C -> 서버 A, 서버 B -> 서버 A이며 이동주기는 일, 수시(특정일자), 월 단위 기준으로 등록을 한다. 이동시점은 각 서버별 일 단위 작업을 기준으로 서버별 일 작업과 관계없이 일정한 시간, 서버 A의 특정 작업 A'의 사전작업 시작 전, 서버 A의 A' 사전작업 완료 후, 서버 A의 특정작업 B'작업 완료 후, 서버 B의 C' 작업 완료 후, 서버 C의 D' 작업 후로 나누어져 있으며 테이블 단위의 작업은 사전/이동/사후 세부 작업으로 분할된다. 여기서, 사전 작업은 이동을 위한 사전 환경 구축, 이동은 실제 이동 수행을 위한 작업이며 사후작업은 AP가 사용가능토록 사후 환경 구축까지를 포함하고 있다.

2.3 이동실행

이동 수행에서는 단계별 프로그램을 자동 생성 하고 버전을 관리하는 기능을 제공하며 진행 과정 및 수행 결과를 생성 저장 관리한다.

이동을 수행하기 위한 테이블 단위의 사전/이동/사후 작업을 나타내는 단위 프로그램 및 From , To 해당 서버에서 관련 프로그램이 수행 가능토록 프로그램을 자동 생성하여 프로그램의 수행순위 및 테이블 단위로 병렬 수행 가능한 단계 프로그램을 생성/삭제 한다.

진행 과정에 대한 정보와 수행상의 오류 내역을 생성하여 작업자가 오류발생 작업을 재수행 할 수 있도록 한다. 작업의 모든 수행 결과를 LOG 정보에 저장하며 이 정보는 이력 관리 되어진다.

2.4 통계관리

데이터이동 작업에 대한 결과(로그) 정보로서 분석(통계)자료를 생성, 생성되어진 통계자료를 조회하는 기능이다. 통계정보는 일별 이동대상 테이블 목록 및 건수, Size, Extent 개수,정상여부, 이동량,소요시간, 정상 여부, 실 테이블의 건수와 이동시 변동된 건수 간의 일치성 검증 기능을 제공한다. 추가로 작업시 발생된 오류내역을 조회하는 기능을 제공하며 이를 화면 및 프린터에 출력하여 조회가 가능하게 한다.

3. 시스템 설계

데이터의 일치성을 위한 대용량 데이터 이동시스템은 환경 및 시스템 효율성을 고려하여 아래 3 가지 의 프로세스 구조로 설계하였다. 이에 따른 데이터 이동 메커니즘은 표 1 과 같다. [2],[4]

- 표준시간은 마스터 서버의 정상적인 운용상황을 적용하고 데이터 이동에 영향을 미치는 요인 분석을 통하여 단계별로 이동한다.
- 데이터 배치에 있어서 모든 데이터는 물리적으로 동일한 곳에 배치하며 작업 결과 및 이력데이터는 주로 master 가 아닌 slave 서버에서 입력하도록 한다.
- 프로세스 배치에 있어서는 작업제어 및 통계 관리 등의 전체 작업관리 프로세스는 slave 서버에 배치하여 데이터 이동 관리를 일원화하고 실제 실행 프로세스는 해당 시스템에 배치하여 시스템 사용의 효율화를 고려하였다.

No	Client	송신서버	수신서버
1	송신서버 상태 요구		
2		Master에게 상태 보고	
3	송신 상태 검증		
4	수신 이동PGM invoke		
5			이동 PGM 수행
6			원시통계자료 생성
7			Master에게 결과 통보
8	이동수행 결과 파악		

표1 이동 메커니즘

4. 시스템 구현 및 테스트

본 논문에서 제안한 대용량 데이터 이동시스템의 환경으로 하드웨어는 UNIX시스템 4.0f와 상용 DBMS는 ORACLE 8.0.6을 이용하였다.

시스템의 구현에 있어서 성능보장을 위하여 MASTER 노드의 table별 mlog 최적화(Initial 또는 Extent size 및 highwater 초기화 시기 결정 포함) 및 상대 서버의 데이터 이동 데이터 및 index용 table space 최적화를 고려하였다. 데이터 이동시스템에 있어서의 ORACLE DBMS의 snapshot 기술 제약사항(snapshot create의 경우 log를 생성하지 않으며 수행속도가 빠른 반면, create 중에는 수신측에서 해당테이블에 대한 작업이 없어야 함)을 고려하여 구현하였다.[4]

□ 프로그램에 있어서는 대상테이블 단위로 이동 PGM을 아래의 사전/이동/사후의 SubPGM으로 분할하고, 통계생성과 작업의 오류에 따른 재수행은 SubPGM 단위로 처리하도록 하였으며 그 내용은 표2와 같다.

구분	Snapshot Refresh	Snapshot Create	SQL Copy	FTP
사전	검증/생성 (Mlog, Basetable)	검증/생성 (Mlog, Basetable)	Drop Basetable /Create Basetable	Delete/Overwrite
이동	Fast Refresh	Complete Refresh	Select/Insert	Ftp
사후			Indexing	

표2 PGM 유형

본 시스템의 성능분석을 위해 100만 건의 데이터를 가진 레코드 길이가 251byte인 총, 251Mbyte 크기의 테이블에 대해 데이터 변동율에 따른 소요시간을 측정하였다.[7]

그림2는 snapshot 작업 유형별로 데이터 변동율에 따른 소요시간을 측정해본 결과, 데이터 변동율이 14%를 조금 넘어서는 순간부터 snapshot create 작업보다 snapshot refresh(insert)의 시간이 더 많이 소요됨을 보여준다. 데이터의 변동율이 18%에 육박함에 따라 snapshot create 작업보다 snapshot refresh(update)나 snapshot refresh(delete)의 시간이 더 오래 걸리는 것을 역시 보여주고 있다.[5]

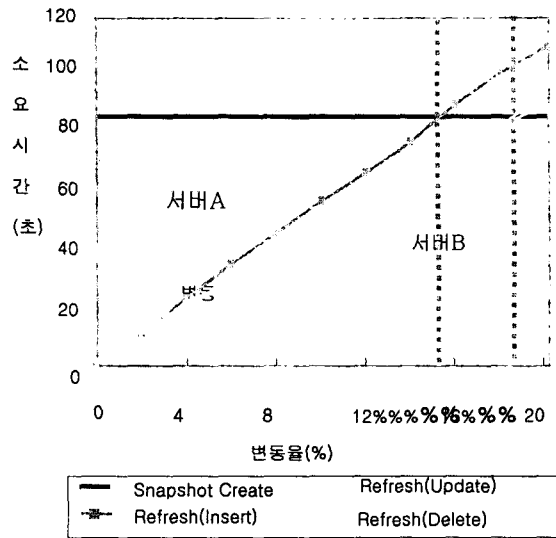


그림2 데이터 성능 분석

본 시스템에서는 Oracle에서 제공하는 snapshot 종류 중 Fast-Refresh 기법을 이용하여 데이터 이동시 초기 전체를 복제하고, 향후 변동 데이터에 대해서는 변동데이터만 복제하는 방식을 이용하였다.

5. 결론 및 향후 과제

현재 분산 환경에서의 정보 공유 방법은 다양하다. 본 시스템에서는 이동시기가 다양한 대상에 대해 성능을 분석테스트를 통하여 데이터량이 많은 초기 구축시에는 전체를 복제하고 향후 변동 데이터에 대해서는 변동분만 복제하는 기법을 이용한 대용량 데이터를 이동하는 시스템을 구축하였다. 이 시스템은 특정 시스템의 요구로 인하여 단일 방향의 데이터 이동만 구축되어 있으나, 향후 이동 시기 및 이동 대상의 다양화에 따른 양방향 이동 시스템으로의 확장이 필요하다.

6. 참고문헌

[1] Michael Stonebraker, "a wide-area distributed database system", VLDB Journal, 1996
 [2] KT, "시스템 구조", KT 원부이동시스템 설계단계 산출물, pp. 3, 2000.10
 [3] Kim EuiKyoung, "Speaker Invitation to Beijing Oracle World - KT- ICIS DB 42 TB", Beijing Oracle World, pp. 20, 2002.6
 [4] KT, "이동흐름", KT 원부이동시스템 설계단계 산출물, pp. 1, 2000.10
 [5] Oracle, "Snapshot 기술서", KT 원부이동시스템 산출물, pp. 2, 2000.10
 [6] Jack G. Conrad, "Database Selection Using Actual Physical and Acquired Logical Collection Resources in a Massive Domain-specific Operational Environment", VLDB 2002 HongKong
 [7] KT, "원부이동 상세설계 방안 수립", KT 원부이동시스템 산출물, pp. 4, 2000.10