

Tech Info(vol.4)

Enterprise Replication 고성능 정보공유 및 분산솔루션

지난 10년 동안, 기업 전산 환경은 메인 프레임의 역할을 바꾸어 왔습니다. 기업체 정보는 이제 더 이상 하나로 집중되어 있지 않습니다. 이러한 경향은 경쟁적이고 비용 절감 효과가 있는 조직적인 힘에 의하여 이루어지며, 이러한 힘은 회사들이 자사의 IT 자원의 극대화를 요구합니다. 이러한 분산화 현상의 결과 중의 하나는 IS 관리자들도 기업 내에 퍼져 있는 모든 정보를 관리하는 분산된 데이터베이스를 다루고 있다는 것입니다.

데이터의 복제(Replication)는 정보를 효과적으로 분산시키고 공유시키는 주요 기술입니다. 그러나, 복제는 데이터를 이곳에서 저곳으로 옮기는 것 이외의 의미가 있습니다. 그것은 기술적인 요구 사항이 매우 많은 작업입니다. 그리고, 복제 시스템의 성공 여부는 기업과 응용 프로그램이 필요로 하는 모든 범위의 사항을 다룰 수 있는 기술에 의존하고 있습니다.

복제 시스템은 원래의 데이터베이스에게 무거운 짐을 지우지 않고, 응용 프로그램을 변화시키지 않으면서 높은 성능을 제공해야 합니다. 이들은 시스템의 성능을 최대로 만들어야 하며, 데이터의 지속적인 전달을 보장해야 합니다. 마지막으로, 데이터베이스 운영자들은 기업 전산화 자원이 가장 효과적으로 사용될 수 있도록, 복제 시스템의 모든 구성 요소를 설정하고 관리할 수 있어야 합니다.

이 글에서는 이러한 요구 사항들과 기업들이 이 중요한 기술을 성공적으로 사용하도록 하는 Informix의 지원에 대하여 다루겠습니다. 이제 Informix 기술 개요를 설명한 후, 복제에 대한 중요한 개념과 고려 사항을 설명하겠습니다. 그 다음 Informix의 복제 솔루션인 Enterprise Replication(ER)을 보다 상세히 설명하겠습니다.

Enterprise Replication: 개요

Informix의 진보된 복제 솔루션인 Enterprise Replication은 기업 정보의 분산과 공유의 기초를 제공합니다. Enterprise Replication은 작업 그룹으로부터 고급의 MPP와 클러스터까지의 데이터 복제를 지원합니다.

Enterprise Replication의 수많은 주요 특성들은 기업체와 응용 프로그램에서의 넓은 영역의 요구를 충족시키는 뚜렷한 장점을 갖고 있습니다.

- **높은 성능** Informix의 첨단 서버 기술은 우수한 성능과 확장성 및 대규모의 용량 그리고 기업 기반의 응용 프로그램들을 제공합니다. Informix는 대부분의 전산 환경에서의 요구를 충족시키는, 확장성 있는 병렬처리 서버 구조의 모든 이점을 살릴 수 있는 최초의 RDBMS 공급업체로서 이러한 일을 이루어 왔습니다.
- **높은 이용성** Informix의 복제 제품은 다양한 선택 사양을 통한 높은 이용도를 보증합니다. 서버에서 하나 이상의 사이트로 복제할 데이터의 비동기 전송이 가능합니다. 데이터의 전송이 비동기적이기 때문에 네트워크와 대상 노드의 작동 중단이 허용되어, 모든 단일 지점 오류를 해소하여 줍니다. 전체 ER 목록의 변화를 포함한 변동 사항은 자원이 복원될 때 자동적으로 원격 사이트로 전달될 것입니다.
- **트랜잭션 수준의 무결성** 무결성은 원본 데이터베이스의 변화가 대상 데이터베이스로 즉각적으로 전달됨으로서 유지됩니다. 트랜잭션들은 그들의 배열 순서를 유지하면서 복제됩니다. ER은 데이터의 수정 작업이 충돌할 경우, 이를 감지하고 해결하는 많은 방법을 제공합니다.
- **유연한 구조** ER은 기업체에 특수한 모델이나 방법론을 강요하지 않는 복제 구조를 제공하며, 조직체들이 기업 환경이나 응용 프로그램의 필요에 따른 그들만의 복제 환경을 정의할 수 있게끔 허용합니다. ER의 구현은 응용 프로그램에 대하여 투명하기 때문에 코드의 변화를 피할 수 있습니다. ER은 Master/Slave, 데이터 흐름 수정, 모든 데이터에 대한 수정 등의 다양한 용도의 복제 기능을 제공합니다.
- **관리성** 시스템 관리 체제와 부속된 그래픽 기능의 도구를 이용하여 환경 설정 작업과 모니터 작업을 할 수 있습니다. ER의 세부 사항에 대해서 다루기 전에 기업들이 복제 기술을 도입하는데 고려해야 할 사항과 핵심이 되는 개념을 다룰 것입니다.

Data Replication이란 무엇인가?

복제(Replication)는 여러 개의 사본 데이터를 하나 이상의 사이트에 만들고 다시 복제하는 작업이라고 간단하게 정의할 수 있습니다. 데이터베이스의 복제는 기업체로 하여금 사용자들이 원하는 시간과 장소에서 현재의 데이터를 다룰 수 있도록 해주기 때문에 중요합니다. 데이터의 복제는 중앙에 집중된 자원이 과부하를

초래하는 경우의 성능 향상, 데이터의 이용도 향상, 의사 결정 지원을 용이하게 해주는 데이터 저장소에 대한 지원을 포함하는 넓은 범위의 이점을 갖고 있습니다.

복제에는 동기적(synchronous), 비동기적(asynchronous)의 두 가지 형태가 있습니다. 각각은 서로 다른 목적에서의 기능과 장점이 있습니다. 동기적 복제에서는 복제된 데이터는 원본 데이터가 수정되는 즉시 수정됩니다. 데이터가 동기적으로 복제되는 동안에는 데이터의 무결성은 2단계 완료(two-phase commit) 프로토콜에 의해서 보호됩니다. 각각의 서버는 데이터를 수정한 후에 확인 메시지를 원본 데이터베이스에게 보내고, 트랜잭션을 끝내라는 메시지를 기다리게 됩니다. 동기적 데이터 복제(최종 트랜잭션 등과 같은)가 즉각적인 데이터 동기화를 필요로 하는 프로그램에 좋은 기술인 반면, 이용도나 성능의 저하를 가져올 수도 있습니다.

비동기적인 복제는 원본 데이터베이스가 수정된 이후에 복제 대상 데이터베이스가 수정되는 방식입니다. 수정하는데 걸리는 지연 시간은 환경에 따라, 수 초에서 수 시간까지가 될 수 있습니다. 그러나, 데이터들은 모든 장소에서 항상 같은 값을 유지합니다. 특정 장소의 고장이나, 액세스 불능시, 비동기적 전송에서는 작업을 계속하기 위한 로컬 프로세싱 요구를 수행합니다. 예를 들면, 주문 입력 시스템에서 심지어 선적 시스템이 연결되지 않을 경우에도 주문을 입력할 수 있습니다. 고장난 사이트가 다시 온라인으로 연결되면, 시스템은 입력된 주문 데이터를 전송하는 작업을 할 것입니다.

Data Replication에서의 핵심 고려 사항

분산 처리 환경에서 데이터의 복제가 중요한 역할을 하기 때문에, 기업들은 그 용도를 이해하고 이 기술을 가장 잘 적용할 방법을 강구할 필요가 있습니다. 근본적인 문제는 데이터의 무결성과 이용도간의 상관 관계입니다. 이미 설명했듯이, 동기적 기술 방식은 이용도와 응답 시간을 희생함으로써 데이터의 무결성을 확립시킵니다. 복제 기술의 다른 쪽인 비동기적 기술은 데이터의 무결성을 보장하고 수정 작업에서의 충돌을 방지하는 설계를 필요로 하지만 이용도와 응답 시간을 극대화시킵니다.

궁극적으로 기업과 응용 프로그램의 요구 사항은 적절한 분산 방법을 찾는 것이고, Data Replication 시스템은 이러한 요구 사항을 지원하는 것입니다. 다음으로 우리는 비동기적 데이터의 또 다른 근본적인 문제인 소유권에 대해서 다루겠습니다.

데이터 소유 관계(Data Ownership)

소유 관계의 개념은 어떤 사이트가 데이터를 수정할 수 있는 권한을 갖느냐는 것입니다. 소유 관계의 종류는 가능한 충돌 감지와 해결의 필요성을 결정합니다. 예를 들어, Master/Slave 소유 관계에서는 하나의 사이트만이 데이터를 수정할 수 있을 것입니다. 오직 하나의 사이트만이 데이터를 수정할 수 있기 때문에 이러한 방법은 충돌의 감지와 해결을 필요로 하지 않습니다. 나중에 다른 다른 소유 관계 모델들은 여러 장소에서의 데이터 수정을 허용하기 때문에, 충돌의 감지와 해결을 요구합니다.

Master/Slave 소유 관계

Master/Slave 소유 관계는 데이터 분배 방식, 데이터 통합 방식, 작업 파티션 방식 등의 많은 경우에 적용될 수 있습니다.

- 데이터 분배 방식(Data Dissemination)

데이터 분배 방식에서는 중앙에서 데이터가 수정된 후 각 지역의 읽기 전용 사이트에 복제되는 환경을 생각할 수 있습니다. 이러한 방법은 데이터를 각 지점에 분배할 필요가 있는 회사들에게 유용합니다. 예를 들면, 대규모 서점들은 야간에 책들의 수정된 가격 목록을 보낸 필요가 있을 것입니다. 이 데이터의 일관성을 보장하기 위해서는 본점에서는 읽고 쓰기가 가능한데 반하여, 각 지점들은 정보를 읽을 수만 있어야 합니다.

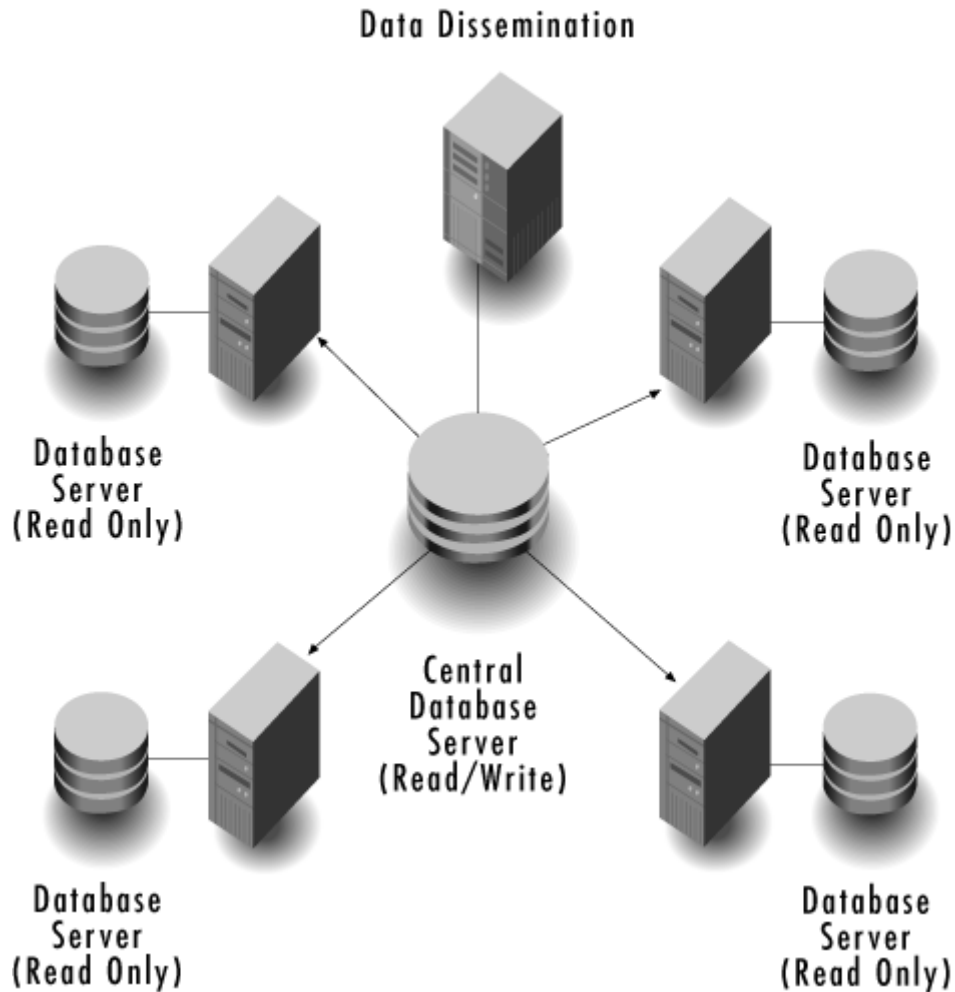


그림 1. 데이터 분배의 경우 데이터는 중앙에서 수정되며, 여러 개의 읽기 전용 사이트로 분배

- 데이터 통합 방식(Data Consolidation)

데이터 통합에서는 데이터가 각 지역에서 수정된 후 중앙에 있는 데이터베이스 서버의 읽기 전용 저장소에 모아지는 환경을 생각할 수 있습니다. 이러한 방법은 지역 수준의 데이터의 소유 관계와 자율권을 보장합니다. 이러한 예로는 하루 동안의 각 지점 영업 사항 정보를 모으는 유통 상점이 해당됩니다. 각 점포는 일별로 이 영업 정보를 본점으로 보낼 것입니다. 일과가 끝나면 이러한 정보는 본점으로 전송되고, 중앙의 데이터 저장소에 저장될 것이며 관리부에서는 이것을 사용하여 각 점포의 사업에 대한 분석을 할 것입니다.

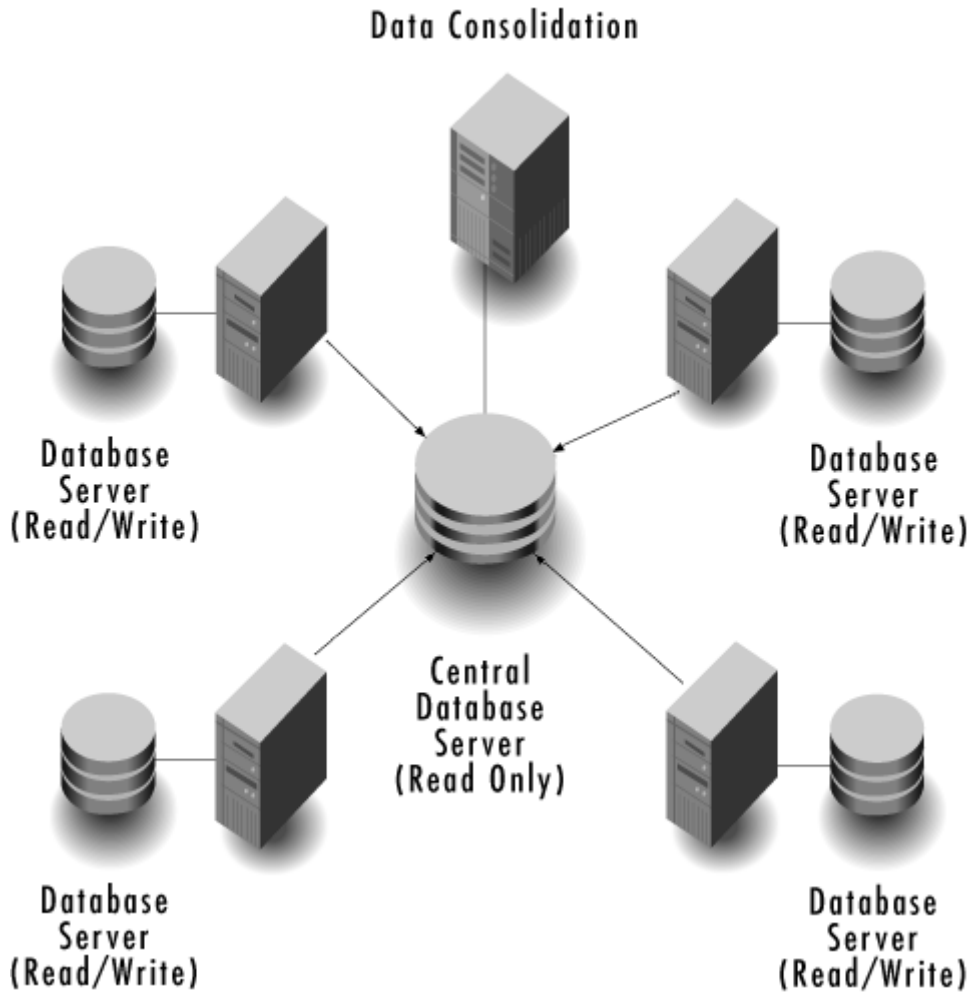


그림 2. 데이터 통합의 경우 여러 사이트에서 수정된 데이터들이
중앙의 읽기 전용 사이트에 복제

- 작업 파티션 방식(Workload Partitioning)

작업 파티션 방식에서는 데이터베이스 운영자(DBA)들이 테이블 파티션 수준의 데이터 소유 관계를 정할 수 있도록 합니다. 그 실례가 아래에 설명되어 있습니다. 데이터 복제 스키마는 사원 테이블의 파티션 스키마에 맞게 정해져 있습니다. 아시아 태평양 지역의 컴퓨터들이 자신의 파티션에 대한 소유권을 갖고 있으며, 그래서 개별적으로 사원 레코드들의 수정, 추가 및 삭제 작업을 할 수 있습니다. 그 다음 이러한 변동 사항이 미주와 유럽 지역으로 전달됩니다. 아시아, 태평양 지역에서는 큰 파티션들을 지역적으로 질의하거나 읽을

수 있으나 그 파티션들을 수정할 수 없습니다. 이러한 사항은 다른 지역에서도 마찬가지로 적용됩니다.

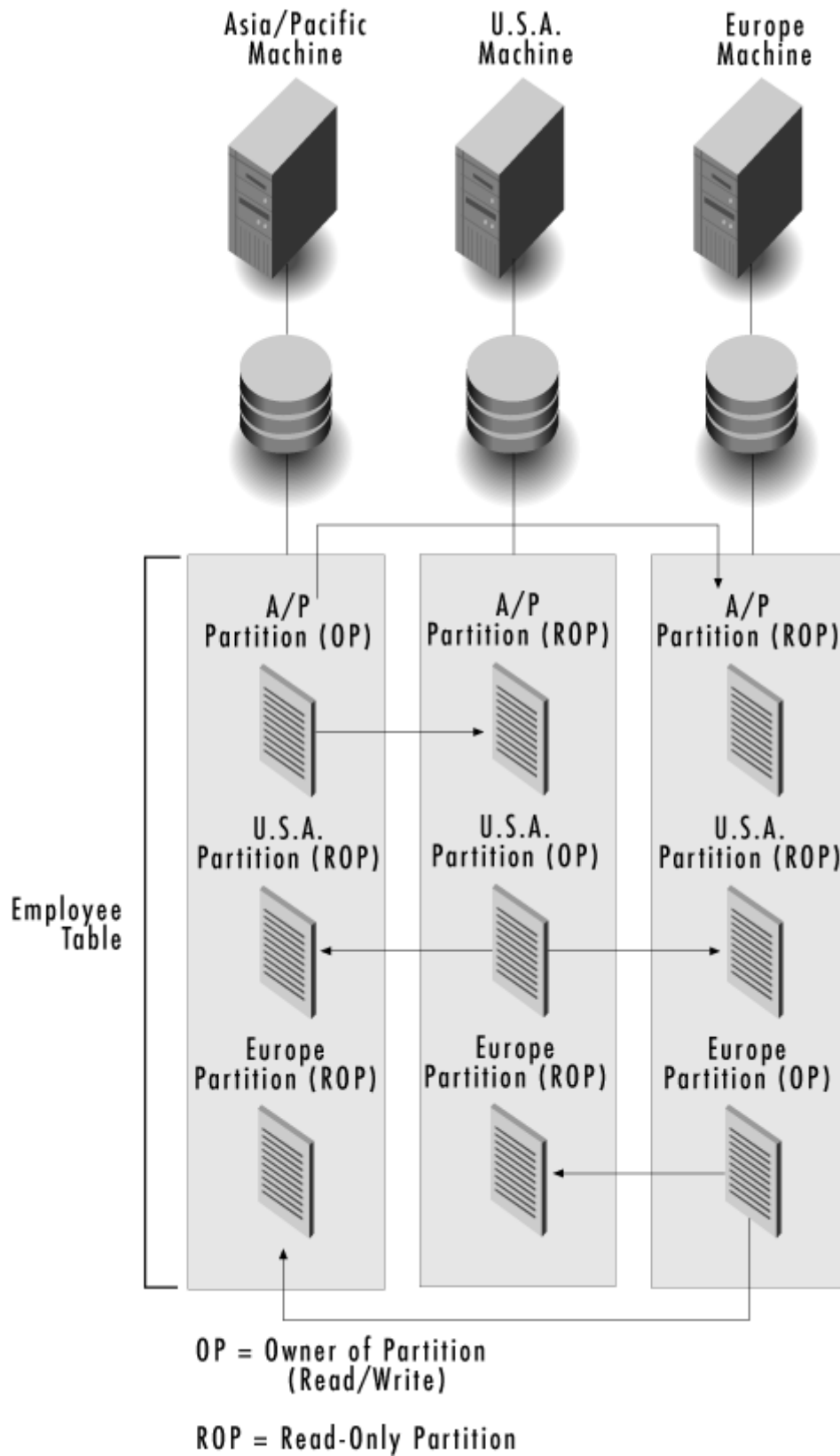


그림 3. 작업 파티션 방식에서는 각 지역의 사용자들이
테이블 전체를 볼 수 있을 뿐 아니라 수정 가능

Workflow 소유 관계

Master/Slave 소유 관계와 같이 이 모델은 수정 작업이 충돌 현상을 막을 수 있으며, 동시에 보다 동적인 수정 기술을 제공합니다. Workflow 소유 관계는 복제된 데이터를 변경하는 권한을 한 사이트에서 다른 사이트로 옮겨 주는 것을 허용합니다. 그러나, 한 번에 한 사이트만이 데이터를 수정할 수 있습니다. 각 사이트는 이전 사이트의 데이터에 의존적이며 기업 업무에 따라 레코드를 수정할 것입니다. 각 단계가 끝나면 이 정보는 수정되어 다음 사이트로 복제됩니다.

전형적인 workflow 소유 관계의 실례는 주문 처리 시스템입니다. 주문은 처음에 주문 접수 부서에서 접수됩니다. 그 다음 신용 조회와 송장 작성을 하는 회계부로 보내지고, 다음으로 자재 관리부의 승인을 얻어, 마지막으로 포장과 배달 업무를 맡은 배달부로 보내집니다.

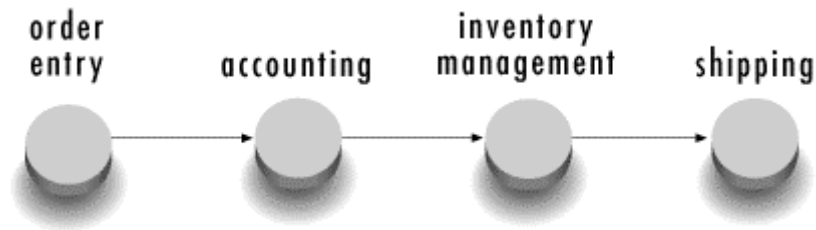


그림 4. 데이터를 변경할 권한이 부서에서 부서로 옮겨지는
전형적인 workflow 소유 관계의 실례

수정 권한 개방(peer-to-peer)

회사들은 여러 사이트에서 동시에 실시간으로 수정이 이루어지도록 요구할 수 있습니다. 복제된 데이터가 읽기 전용인 Master/Slave 소유 관계와는 달리 수정 권한 개방 복제 방식은 여러 사이트가 데이터를 수정할 수 있는 동등한 권한을 갖는 peer-to-peer 환경을 만듭니다. 이것은 심지어 분산 처리 환경의 다른 시스템이나 네트워크를 사용할 수 없는 경우에도 지역 사용자들이 자치적인 권한을 갖는 것을 허용합니다. 수정 작업이 충돌하는 것을 막기 위해서는 넓은 범위의 자동 충돌 감지와 해결 루틴을 갖고 있는 데이터 복제 구조가 이러한 환경을 지원하여야 합니다.

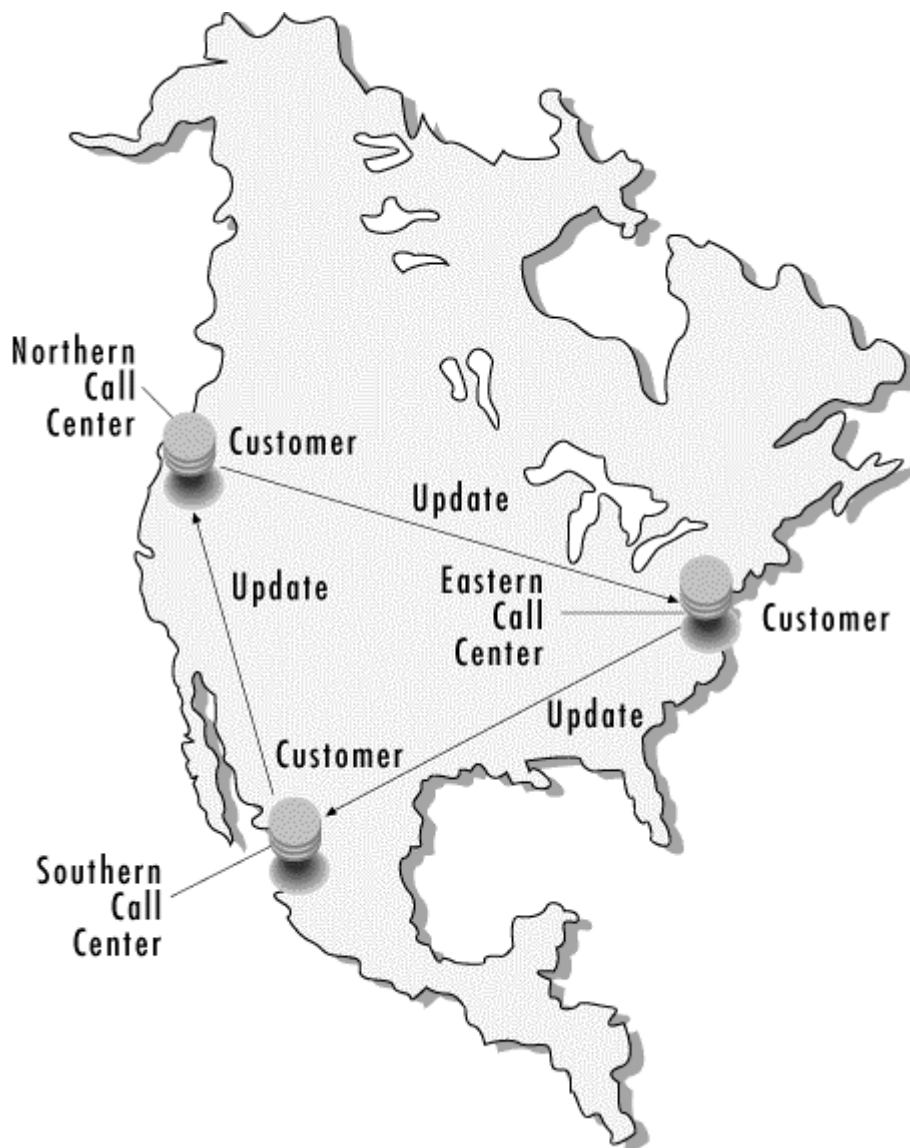


그림 5. 다른 사이트들이 읽기 전용 사본을 갖고 있는 동안 데이터가

수정 권한 개방 방식의 실례는 위의 그림에 잘 나타나 있습니다. 이 예에서는, 800 개의 수신자 부담 전화 배달 주문을 받는 우편 배달 회사가 있습니다. 통화 내용은 하루의 시간과 통화 분산에 의해 세 곳의 전화 센터에 연결될 것입니다. 각 전화 센터는 레코드들을 수정하는 능력을 필요로 하며 주문이 처리되는 동시에 그 정보를 다른 전화 센터에도 복제합니다.

트리거 기반 방식 대 로그 기반 방식

가장 빈번하게 사용되는 복제 변경 수용 알고리즘은 트리거 방식과 로그 방식입니다. 트리거 방식에서의 데이터 변경은 데이터베이스 내에 트리거를 형성합니다. 이것이 원본 데이터베이스 내의 복제 관련 코드를 활성화시키고, 다음으로 이 코드가 복제 프로세스를 실행하게 합니다. 이 프로세스는 사용자 트랜잭션과 같은 동일한 작업 내에서 발생하며, 그러므로 소스 데이터베이스의 트랜잭션 처리와 직접적으로 충돌합니다.

트리거 방식의 또 다른 핵심 문제는 트리거 기반의 시스템이 어떻게 일반적인 데이터베이스 트리거로부터 발전할 수 있는지는 것입니다. 데이터베이스 트리거는 처음에는 데이터베이스 업체에 의하여 소개되었으며, 기업 데이터의 무결성을 보호하고, 기업 규칙을 강화하기 위한 것이었습니다. 이러한 업체들에게는 트리거 기반의 복제 방법은 논리적이었으며, 그들 제품의 올바른 확장법이었습니다. 또한, 대부분의 DBMS 회사들이 내부 로그 형식을 공개하지 않았기 때문에 트리거 기반의 복제 방법은 또한 하청업자들이 이질적인 복제 방법을 구현할 수 있는 유연성을 제공하였습니다. 그러나, 트리거 기반의 복제 방법은 트랜잭션의 무결성을 보장하기 위해서 업체들이 추가적인 기능을 제공하여야 된다는 단점을 갖고 있습니다. 이러한 무결성 문제는 해결되었지만, 심각한 성능과 관리 부담을 가질 수 있습니다.

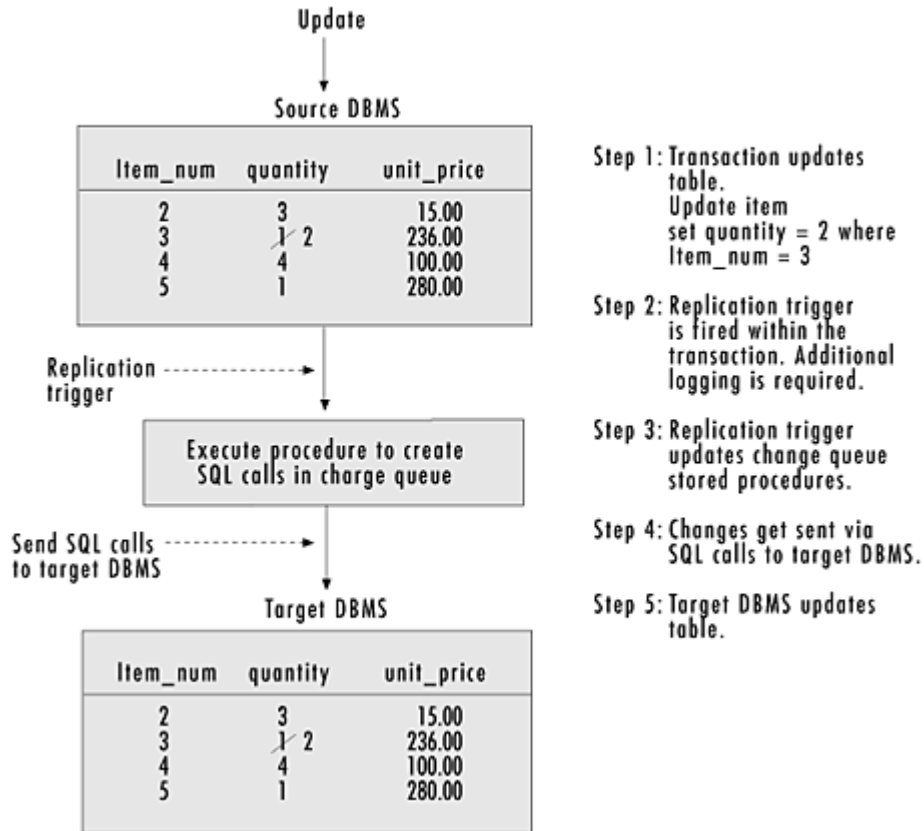


그림 6. 원래의 트랜잭션과 충돌하는 트리거 기반의 데이터 복제 방법

반면, 로그 기반의 시스템은 보통의 데이터베이스 로깅 프로세스의 일부로 작동할 수 있으며, 그래서 시스템에 최소의 부담만을 유발합니다. 이것의 개념은 트랜잭션 로그가 원본 데이터의 변화를 감지하는 최고의 데이터라는 것입니다. 원본 데이터베이스의 변경 사항은 원본 데이터베이스 시스템의 정상적인 연산들을 방해하지 않으면서 감지되고 전달될 수 있습니다.

그러나, 모든 로그 기반 복제 시스템이 다 같은 것은 아닙니다. 중요한 차이점은 로그 인식 알고리즘이 얼마나 데이터베이스 시스템과 잘 결합되어 있는냐는 것입니다. 예를 들어, 외부의 서버를 사용하는 로그 기반 방식의 제품들과 로그 기반 방식의 이점을 상쇄시키는 프로세스들이 추가적인 데이터 전송 부하로 인하여 발생할 수 있습니다. 로그 기반 방식의 성공 여부를 결정하는 또 다른 특성은 로그의 모든 엔트리가

트랜잭션 감지 프로세스의 일부로서 읽혀졌는가 입니다.

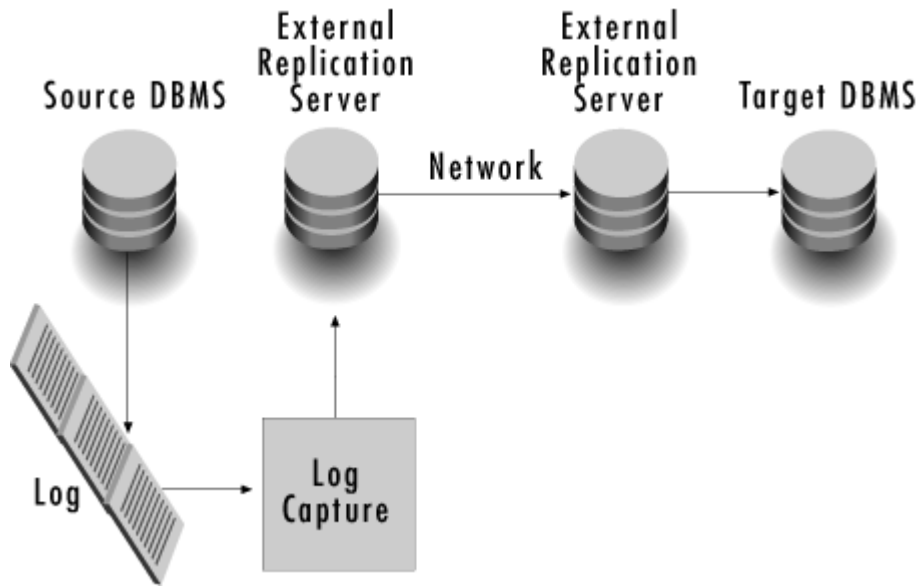


그림 7. 로그 기반 방식에서 추가적인 데이터 전송을 일으키는 외부의 데이터 복제 서버

Informix 는 Enterprise Replication 이전의 트리거 기반 구조와 낮은 품질의 로그 기반 구현 방식의 취약점을 분석하였습니다. 그 결과로, ER 은 효율적인 로그 기반 방식과 분배 메커니즘을 데이터베이스 시스템의 일부로 발전시켰습니다.

Replication Integrated with DBMS:
reduces overhead by minimizing data transfer.

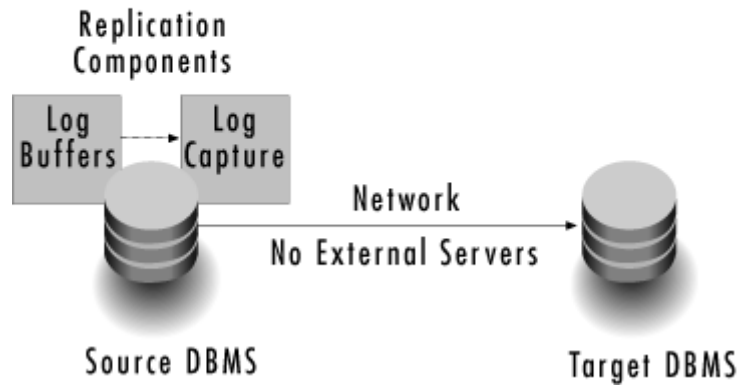


그림 8. DBMS 와 밀접하게 결합된 데이터 복제 방식.

외부 서버가 없으며, 로그 감지할 때의 적은 부하

Informix Data Replication의 구조와 장점

Informix Enterprise Replication은 산업 시장에서의 가장 진보된 Data Replication 솔루션이라는 것입니다. 첨단 동적 확장 구조(DSA)로 이루어진 Enterprise Replication(ER)은 다른 Data Replication 제품들의 한계를 극복한 최고 수준의 기술을 보여 줍니다. ER은 MPP 클러스터에 대한 작업 그룹을 지원합니다. Informix는 ER을 모든 범위의 산업 영역을 다룰 수 있도록 설계하였습니다.

- 높은 성능
- 높은 가용성
- 일관적이고 신뢰성 있는 트랜잭션 무결성
- 유연한 구조
- 관리 용이성
- 이종 데이터(heterogeneous) 지원

높은 성능

Informix의 Enterprise Replication은 기업들의 Data Replication 시스템에 대한 많은 요구를 수용하는데 적합한 고성능이며, 확장성 있는 플랫폼을 제공합니다. 일찍이 Informix는 트리거 기반의 복제 방식의 한계성을 인식하였고, 그래서 매우 효율적인 로그 기반의 트랜잭션 감지 방식과 병렬적인 데이터 배분을 구현하였습니다. Informix는 외부 서버를 사용하는 다른 로그 기반의 시스템이 겪는 성능의 저하나 관리의 어려움이 없이 이러한 사항을 구현할 수 있었습니다.

덧붙여 말하면, DSA는 SMP와 약하게 결합된, MPP 구조의 병렬처리 성능을 사용함으로써 모든 작업을 병렬처리로 구현하였습니다. 그 결과, 모든 복제 작업이 병렬로 이루어지고, 시스템의 성능은 크게 향상되며, 병목 현상을 최소화하여 줍니다.

높은 가용성

이용성과 고장에 대한 내구성을 최대화하기 위해서 Informix Enterprise Replication은 주 데이터베이스를 다른 2차 서버에 복제하도록 해 줍니다. 이러한 옵션은 주 서버가 작동하지 않을 경우에 대비한 예비 서버를 만드는데 유용합니다.

Informix의 enterprise replication은 하나 이상의 2차 사이트에 대한 비동기적 데이터 복제를 통하여 주 시스템을 고장으로부터 보호해 줍니다. 전체 ER 목록의 변경을 포함한 중심 사이트에서의 모든 수정사항이 자동적으로 부속 사이트로 전달됩니다. 수정 사항의 전송은 즉시 이루어지거나, DBA가 수정에 대한 시간 간격을 정해 놓았을 경우 시간이 걸릴 수도 있습니다. 수정 작업은 트랜잭션이 실행된 이후나, 사용자 정의한 경우에는 이벤트에 의해 이루어질 수도 있습니다.

ER은 데이터를 지역적 기반으로 저장하고, 이 데이터를 독립된 각각의 트랜잭션의 형태로 원격지의 서버로 전송하는, 신뢰성 있는 메시지 전송 메커니즘을 갖고 있습니다. 서버나 네트워크에 이상이 생길 경우, 각 지역의 서버는 해당 지역의 사용자들에게 서비스를 계속할 수 있습니다. 고장났던 서

거나 네트워크가 다시 정상적으로 작동하게 되면, 원본 데이터베이스의 모든 변경 사항은 원격지의 데이터베이스에게로 전달되게 됩니다.

일관적이고 신뢰성 있는 트랜잭션 무결성

Enterprise Replication은 분산된 데이터 환경에서, 여러 개의 지속적인 데이터들의 사본들을 다른 여러 사이트에 유지시킵니다. 이러한 일은 트랜잭션이 생성한 데이터베이스 변동사항이 원래의 서버에서 작업이 행하여진 직후 대상 서버로 전달되기 때문에 이루어집니다. 덧붙여서, 저장되는 트랜잭션들은 트랜잭션의 일관성과 순서를 유지합니다. 이것은, 동일한 레코드가 여러 번 수정되는 경우, 이 변화는 중앙의 데이터베이스에서와 동일한 순서로 일정하게 원격지의 데이터베이스로 전달됩니다.

데이터의 수정시 충돌이 발생할 경우, ER은 예비된 자동 충돌 감지 및 해결 루틴이나, 이를 대신하는 사용자 정의 해결 방식을 제공합니다.

관리 용이성

Informix는 데이터 복제 기술과 실용성의 양쪽에 많은 양의 투자를 하여 왔습니다. Informix는 하나의 통합된 데이터 복제 관리 환경을 제공합니다. DBA들은 어떠한 콘솔을 통해서도 그래픽이 지원되는 환경에서, 데이터 복제 시스템의 환경을 설정하고 감시할 수 있습니다. 예를 들면, DBA들은 네트워크의 모든 서버들을 아이콘으로 표현된 그래픽 화면에서 볼 수 있습니다.

관리 도구들을 사용하여 DBA들은 복제 방법의 종류를 선택할 수 있으며, 대상 서버의 결정과 검색/기록 권한의 결정, 복제 빈도를 결정할 수 있습니다. 덧붙여서, 다른 데이터 복제 제품들과는 달리 DBA 복제 환경 설정을 동적으로 변경할 수 있습니다. 이것은 DBA들이 전체 시스템을 끄지 않고 온라인으로 새로운 서버를 추가하고, 환경 설정 작업을 할 수 있도록 하여 줍니다.

다른 핵심적인 특성은 ER의 환경 설정 정보에서 찾을 수 있는 전체적 복제 목록입니다. 여기에는

복제에 대한 설정이나, ER 서버, 서브스크라이버 목록, 기본 질의 스톱시홀드, 충돌 감지와 해결 규칙 그리고 관련된 내장 프로시저 정보가 들어 있습니다. DBA에게 전체적 복제 목록의 주된 장점은 이것이 복제 설정의 모든 변경 사항을 네트워크상에 있는 다른 서버로 자동적으로 전달한다는 것입니다. 이것은 DBA가 원격지에 있는 서버로 데이터를 일일이 복사해야 하는 부담을 덜어 주며, 전체 목록에서의 데이터의 불일치를 최소화하여 줍니다.

감시 성능은 각 목적지의 복제 작업 큐의 크기를 보여 주는 능력과 원격지 서버의 사용 가능성을 보여 주는 기능, 데이터 복제가 일어나는 동안의 즉각적인 충돌 감지를 포함합니다.

그리고, 결론적으로 Informix는 Informix Enterprise Command Center(IECC)와 밀접하게 결합된 관리 도구들을 제공합니다. Informix의 객체 지향적인 운영 환경은 데이터베이스의 관리를 위하여 특별히 설계되었으며, 데이터 복제에 대한 DBA들의 골칫거리를 해결하여 줍니다.

이질적인 데이터에 대한 지원

Informix는 Praxis International의 OmniReplicator를 사용하여 Informix 이외의 제품으로부터 Informix OnLine Dynamic Server로의 양방향의 이종간 데이터의 복제를 지원합니다. SQLServer, Oracle7, DB2를 포함한 많은 Informix 이외의 제품들에 대한 Data Replication가 지원됩니다.

Informix Enterprise Replication의 기술적 개요

우리는 이제 다음의 내용에 초점을 맞추어 Enterprise Replication에 대하여 보다 상세히 다루고자 합니다.

- 환경 설정과 제어
- 트랜잭션 감지
- 트랜잭션 질의와 전달
- 대상 서버로의 트랜잭션 삽입

- 충돌 감지와 해결

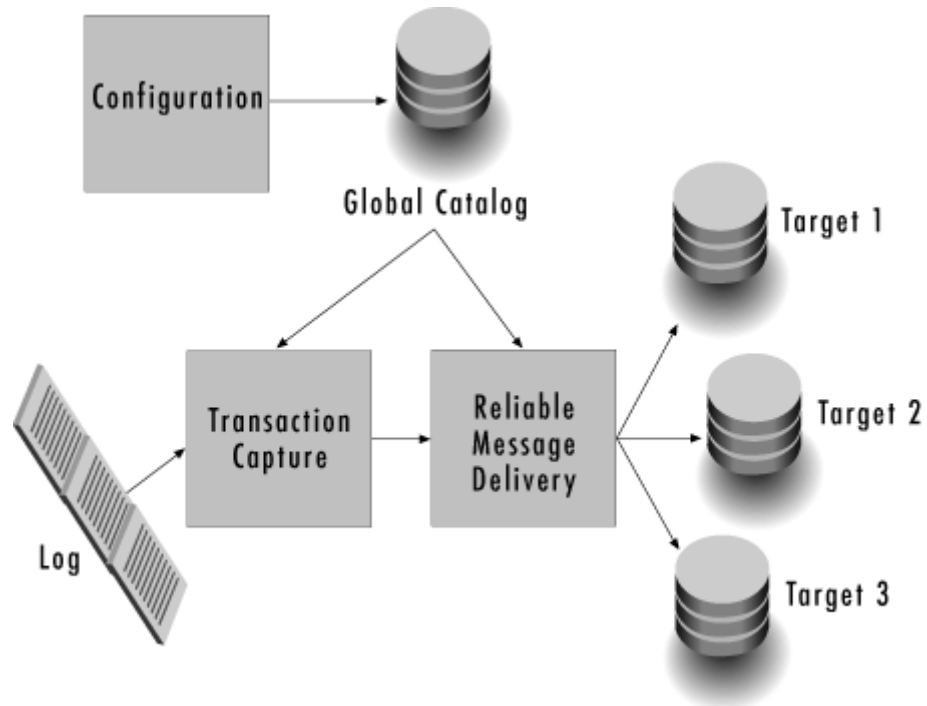


그림 9. Informix Enterprise Replication 의 구성 요소

환경 설정과 제어

ER은 지속적인 트랜잭션의 감지와 전달을 보장하는 데이터 복제 환경 설정에 대한 유연한 방법을 제공합니다. 데이터 복제는 테이블 수준에서 정의될 수 있으며, 파티션이나 뷰와 같은 테이블의 하부 요소들을 사용하여 정의될 수도 있습니다. 데이터 복제 그룹의 특성은 복제에 대한 설정 사항도 같이 묶여서 복제되게끔 합니다. 테이블 수준의 모든 데이터 복제 명령어가 그룹 수준에 적용되기 때문에, 데이터 복제 그룹들은 운영을 단순화시키며 다루기가 용이합니다. 이러한 특성은 ER로 하여금, 서로 다른 복제 그룹의 병렬 트랜잭션들을 적용함으로써, 대상 데이터베이스에 대한 처리 속도를 향상시키게 합니다.

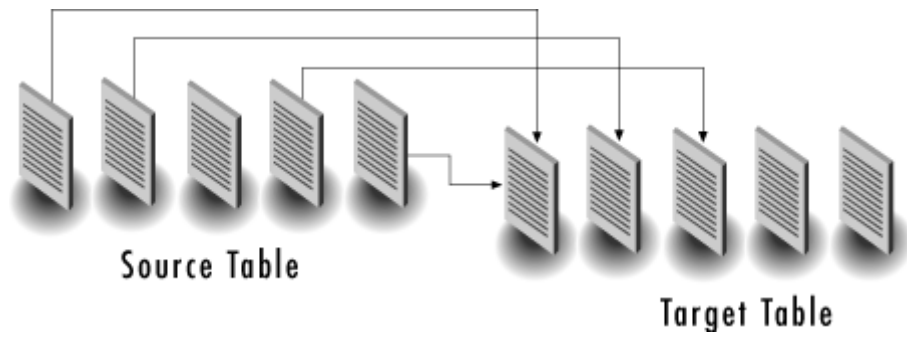


그림 10. 데이터 복제는 테이블의 하부 요소들을 포함하도록 정의될 수 있습니다.

트랜잭션 감지

복제된 트랜잭션을 감지하기 위해서 ER은 트랜잭션 로그 사후 처리를 사용합니다. 사후 처리는 트랜잭션 작업에 대한 충돌을 최소화함으로써 특별한 성능상의 이점을 제공합니다. 먼저, ER은 복제되기로 정해져 있는 레코드들에 대한 로그를 전체 로그와 대조하여 읽습니다. 그 다음, 해당 레코드들은 그들에 대한 복제 설정이 이루어질 것인가의 여부를 결정하기 위하여 필터 연산합니다. 해당되는 레코드들은 트랜잭션 그룹으로 모이게 됩니다. 높은 성능을 보장하기 위하여 ER은 레코드들에 대한 필터 연산에 병렬 처리를 이용합니다. 또한, 이 작업은 데이터베이스 엔진과 결합되어 있기 때문에, 어떠한 추가적인 I/O 연산도 이루어지지 않습니다.

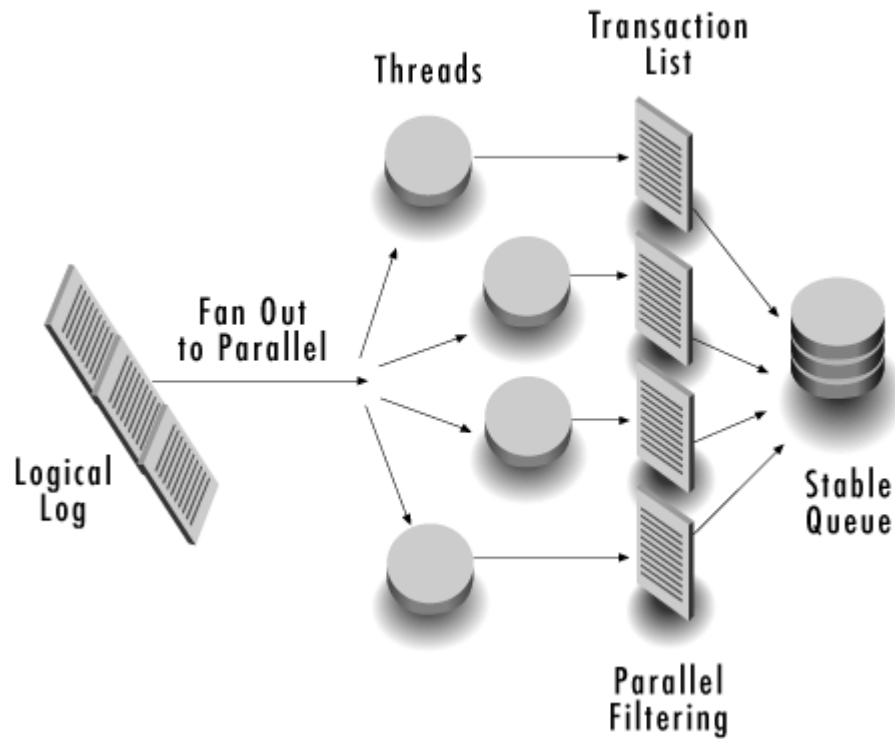


그림 11. Enterprise Replication 은 병렬 처리를 사용하여
레코드들에 대한 필터 연산을 합니다.

트랜잭션 질의와 전달

복제된 트랜잭션이 실제로 전달되면, 이 트랜잭션은 안정된 전달 큐 테이블에 놓여지게 됩니다. 취소된 트랜잭션은 여기서 제외됩니다. 네트워크에 이상이 있을 경우에는 이 안정된 큐는 트랜잭션을 네트워크가 정상화될 때까지 보유하고 있습니다. 무결성을 보장하기 위하여, 전달되는 동안 전달 순서가 유지됩니다. 전달의 신뢰성을 확보하기 위하여 대상 사이트에서 전달 사항을 받아들이면 원래의 서버로 확인 신호가 전달됩니다.

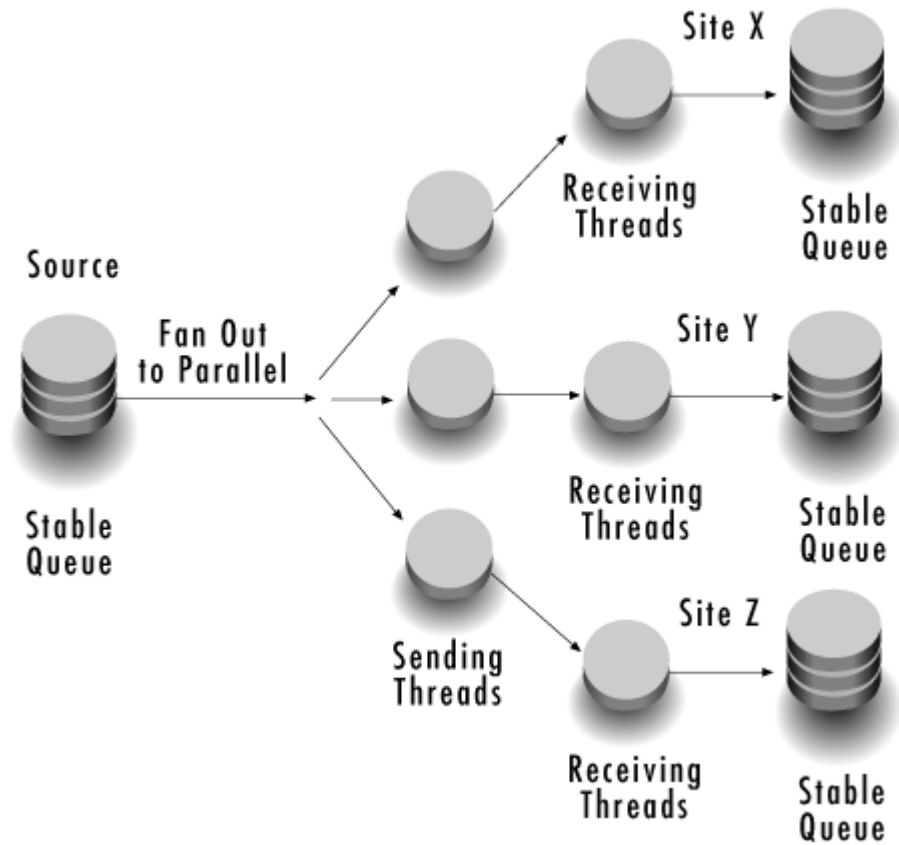


그림 12. 질의와 전달 트랜잭션

대상 서버로의 트랜잭션 삽입

트랜잭션 그룹은 데이터베이스 트랜잭션에 따라 안정된 큐로부터 이동하여 해당 테이블에 삽입됩니다. 성능을 향상시키기 위하여, 이 연산에서는 삽입 작업에 대한 병렬 처리 옵션을 지원합니다. 복제된 레코드가 대상 테이블 내의 레코드와 동일한 기본 키를 갖고 있을 경우에는 어떠한 레코드가 우선인가를 결정하기 위하여 미리 정해져 있는 충돌 감지 및 해결 규칙이 적용됩니다. 모든 필수적인 환경 설정 데이터는 전체 목록에 저장되어 있습니다.

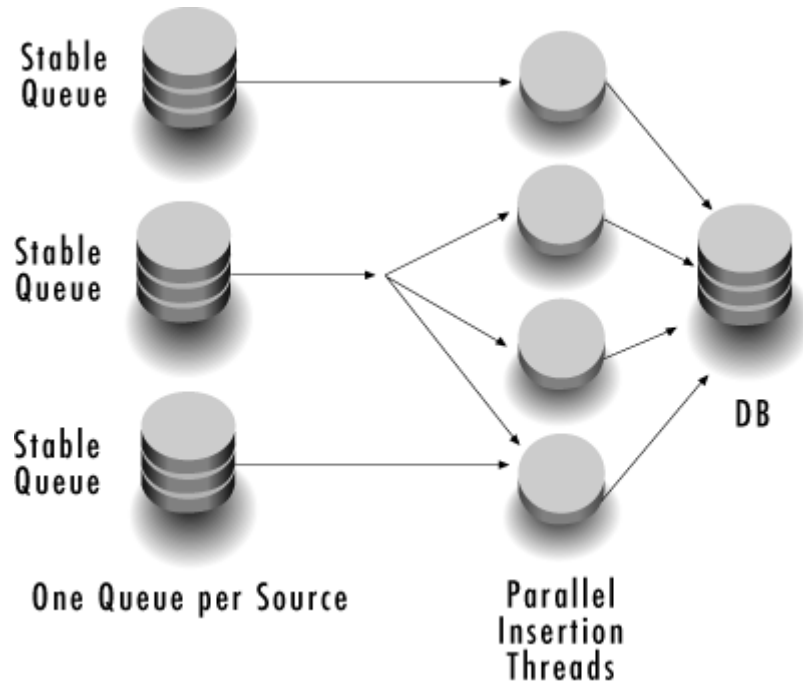


그림 13. 데이터베이스 트랜잭션에 따라 안정된 큐로부터 이동하여 해당 테이블에 삽입되는 트랜잭션 그룹들

충돌의 감지와 해결

Enterprise Replication은 데이터의 일관성을 복구하는 프로그램별 루틴을 호출하여 주는 데이터의 완벽한 충돌 감지와 해결 기능을 합니다. 수정 작업시 충돌이 생겼을 경우, 레코드들간의 우선 순위를 결정하기 위하여 다음의 루틴들이 테이블 사이에 적용됩니다.

- 가장 최근의 타임스탬프 가장 최근의 타임스탬프를 가진 레코드에게 우선권을 줍니다.
- 사용자 정의 방식 특정 프로그램에 대해서는 DBA가 개발한 우선 순위 결정 프로시저가 있다면 이 프로시저가 우선 순위를 결정합니다. 예를 들어, DBA는 특정 서버가 높은 우선 순위를 갖도록 설정할 수 있습니다.

- 충돌 무시 Master/Slave 환경에서 충돌 감지와 해결 기능을 정지시키는데 사용합니다.

Informix를 선택해야 하는 이유

현재, 점점 분산되어가는 데이터 시장에서 분산된 데이터베이스 기능은 사치가 아니라 필수 사항이 되어가고 있습니다. 조직들이 보다 축소되고 분산되는 기업 환경에서 데이터 복제는 분산된 데이터베이스 시스템을 실현하는데 매우 중요한 역할을 하고 있습니다. 복제 작업을 고려하고 있는 곳에서는 그들의 필요를 충족시키기 위하여 이 기술을 어떻게 가장 적절하게 이용할 것인가를 계획할 필요가 있습니다. 복제는 향상된 성능의 전산 데이터 처리, 데이터 이용성의 향상, 편리성, 의사 결정 지원 기능을 갖고 있는 데이터 저장소의 제공과 같은 이점을 포함하고 있습니다.

Informix Enterprise Replication은 높은 성능과 유용성, 일관된 데이터의 분배, 기업 환경에서의 정보의 분배와 공유에 대한 요구를 만족시키기 위한 필수적 요소의 운영성을 갖고 있습니다.