

新 데이터구조 갖춘 통합 DB 구축 필요

운영데이터를 통합하려는 것은 전사적으로 데이터공유를 가능하게 함으로써, 기존 운영데이터의 문제점을 해결하고 업무처리를 효과적으로 지원하기 위함이다. 이를 위해서는 새로운 데이터구조를 갖춘 통합 운영 데이터베이스의 구축이 필요하다.

이용우/ 투이컨설팅 수석 컨설턴트

기업에서 사용중인 기존의 운영시스템은 주로 일상업무처리를 지원할 목적으로 일정기간에 걸쳐 단위업무중심으로 개발되었다. 과거에는 업무가 단순하고 독립성이 컸기 때문에 이와 같은 단위업무중심의 시스템구축이 가능하였다. 아울러 데이터베이스는 운영시스템에 종속되어, 단순히 시스템의 실행결과로 발생하는 데이터의 축적기능만을 담당하였다.

그러나 최근에 업무가 전문화되고 복잡해짐에 따라, 업무간에 유기적인 연계 및 분석을 통한 의사결정 지원이 업무처리에 매우 중요한 이슈로 부각되고 있다. 업무연계와 의사결정지원은 전사적으로 데이터가 공유될 때 가능하게 된다.

데이터 공유란 사용자가 업무수행에 필요한 각종 데이터를 필요시점에 유효한 형식으로 활용하는 것으로, 이를 위해서는 공유데이터가 기업 전체의 관점에서 통합되어야 한다. 은행의 경우, 고객의 기여도를 결정하려면 현재의 거래상태를 포함하여 과거의 대출실적, 예금실적, 카드이용실적 등을 종합적으로 평가하여야 하는데, 이를 위해서는 고객중심으로 통합된 거래실적이 필요하다.

기존 운영데이터의 문제점

기존의 운영데이터베이스(또는 파일)는 운영시스템의 애플리케이션 실행에 필요한 데이터를 관리하기 위하여 운영시스템과 함께 단계적으로 구축되었다. 그리고, 독립적으로 구축된 단위데이터베이스간에 통합은 필요할 때마다 데이터 인터페이스를 별도로 작성하여 연계시키는 방식을 취했다.

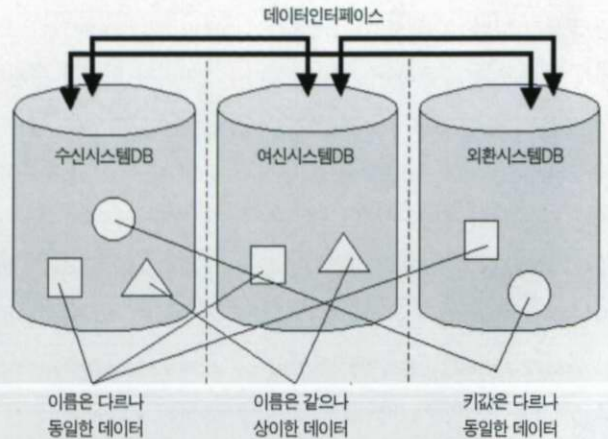
그러나, 처음부터 전사적인 통합데이터체계를 갖지

연·재·순·서

- 1회 : 전사 통합 데이터아키텍처 수립방안
- 2회 : 운영시스템 데이터베이스 구축 방법(이번호)
- 3회 : 정보시스템 데이터웨어하우스 구축방법
- 4회 : 메타모델을 이용한 데이터전송시스템 구축방법

않고, 단위시스템 관점에서 구축된 운영데이터베이스가 비계획적으로 확장됨으로써, 기업전체의 데이터 통합성과 일관성이 저하되는 현상이 발생하였다. 이러한 현상은 단위시스템중심의 운영데이터베이스가 갖는 문제점 때문인데, 그 중 대부분의 기업에서 공통적으로 나타나는 문제점은 다음과 같다.

첫째, 전사 데이터 표준의 부재다. 전사 데이터 표준은 데이터의 설명, 이름, 형식 등에 대한 단일 표준이다. 전사 데이터표준이 없는 경우, 단위시스템별로 데이터 설명, 이름, 형식 등이 상이하게 정의되어, 전사적으로 데이터를 공유하기



(그림 1) 단위시스템중심의 데이터베이스

운영데이터통합 방향

가 매우 어려워진다. 예를 들면, 은행에서 특정 고객의 거래 실적을 집계하는 경우, 수신시스템과 여신시스템의 "고객번호"라는 데이터형식이 상이하면, 다른 고객으로 인식하기 때문에, 결국 특정 고객에 대해 완전한 거래실적을 집계할 수 없게 된다. 데이터이름이 다른 경우도 마찬가지로, 데이터관리자가 '고객'이란 이름을 갖는 모든 데이터를 검색할 때, 데이터이름이 수신시스템에서는 '예금자'로 여신시스템에서는 '고객'으로 되어 있다면, 수신시스템에는 '고객'에 관한 데이터가 없는 것으로 인식하게 될 것이다. 반대로 양 시스템에 모두 '고객'으로 되어 있으나, 수신시스템에서는 '고객'이 실제로 잠재고객을 의미하고 있다면, 데이터 관리자는 결국 잘못된 검색결과를 얻게 된다.

둘째, 동일데이터의 중복이다. 시스템별로 필요한 데이터를 정의하여 사용하기 때문에 전사관점에서는 동일한 데이터(레코드나 필드)가 여러 시스템에 중복되어 관리되는 경우가 발생한다. 이러한 중복데이터는 시스템별로 획득하거나 갱신 시기가 틀리므로, 일정기간이 경과하면 데이터의 불일치가 발생하여, 업무수행에 차질을 초래하게 된다.

예를 들면, 은행에서 고객의 현재거주지가 여신시스템과 수신시스템에 중복되어 있는 경우, 홍길동이란 고객의 현재 거주지가 여신시스템에는 "부산"으로, 수신시스템에는 "서울"로 되어 있다면, 담당자는 어느 것이 정확한 데이터인지를 판단하기 어려울 뿐만 아니라, 판단에 필요한 관련데이터 조회나 고객문의 등의 부가적인 업무에 많은 시간을 할애하여야 한다.

셋째, 데이터 인터페이스의 비효율성이다. 데이터인터페이스는 각 시스템별 데이터베이스간에 중복된 데이터를 일치시키는 기능을 수행하는데, 대부분이 프로그램에 의한 인터페이스방식을 사용하고 있다. 이러한 프로그램방식의 데이터 인터페이스는 완벽한 데이터 정합성 보장에 한계가 있고, 인터페이스체계가 복잡해지는 단점이 있다.

데이터 정합성이 완전하지 못한 경우, 데이터내용 확인 및 보완에 따른 업무지연이나 업무정도의 저하가 발생하여 사용자가 데이터를 신뢰하지 못하게 된다. 또한 데이터 인터페이스용 프로그램의 개발과 운영 그리고 인터페이스데이터에 대한 정기적인 정합성 확인에 소요되는 전산자원도 기업에서는 무시할 수 없는 부분이다.

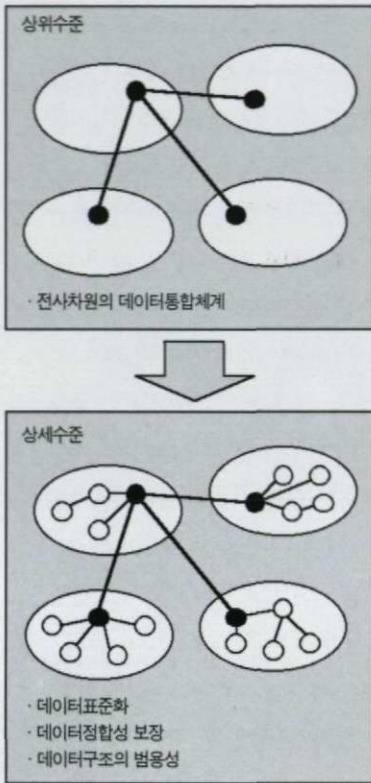
기존의 운영데이터도 단위데이터베이스 관점에서는 통합되어 있다고 말할 수 있다. 그럼에도 불구하고 운영데이터를 통합하려는 것은 전사적으로 데이터공유를 가능하게 함으로써, 기존 운영데이터의 문제점을 해결하고, 업무처리를 효과적으로 지원하기 위함이다. 이를 위해서는 새로운 데이터구조를 갖는 통합운영데이터베이스의 구축이 필요하다.

통합운영데이터베이스가 기존 운영데이터베이스와 다른 데이터구조가 필요한 이유는, 기존 운영데이터베이스가 특정 시스템이나 사용자에 의해 제한적으로 공유되는데 비해, 통합운영데이터베이스는 기업내의 모든 시스템이나 사용자에 의해 공유되기 때문이다. 따라서, 통합운영데이터베이스가 전사적으로 공유될 수 있도록 하기 위해서는 다음의 기본방향에 따라 구축하는 것이 바람직하다.

첫째는 전사적인 통합데이터체계의 수립이다. 통합데이터체계는 기업에서 중요한 주제영역에 대해 핵심데이터와 이들 간에 관계를 정의하고 있는 전사차원의 아키텍처모델로서, 데이터통합의 기본 틀을 제공한다. 통합운영데이터베이스 구축 시 상위수준의 통합데이터체계가 있는 경우, 상세수준의 데이터통합을 체계적이고 계획적으로 진행할 수 있다. 즉, 상위수준의 통합데이터체계를 먼저 정의하고, 이를 바탕으로 상세수준의 데이터를 통합해 나감으로써, 항상 전사차원에서 데이터의 통합성과 통합구조의 일관성을 갖는 통합운영데이터베이스 구축이 가능한 것이다.

둘째는 데이터 표준화이다. 전사데이터공유는 모든 사용자가 데이터를 동일하게 이해하고 사용할 수 있을 때 가능하다. 이를 위해서는 시스템별로 상이한 데이터의 설명, 이름, 형식 등을 전사관점에서 일원화할 수 있도록 표준화하여야 한다. 데이터 표준화는 상세데이터요건이 정의되는 데이터 모델링 단계에서 실시하며, 전사데이터아키텍처 수립단계에서 작성한 대표 도메인 표준안을 적용하여 표준화의 일관성을 유지한다. 데이터 표준화는 업무적인 관점에서 수행되는 것이 중요하며, 정의된 표준안은 데이터 디셔너리 시스템을 통해 관리하여야 한다.

셋째는 전사차원에서의 데이터 정합성 보장이다. 기존 운영데이터 문제점은 해당 시스템 내에서는 데이터의 정합성이 보장되지만, 기업 전체적으로는 정합성을 보장하지 못한다는 점



(그림 2) 운영데이터 통합방향

이다. 따라서, 새로운 통합운영데이터베이스는 전사 데이터 표준을 토대로, 데이터 간에 존재하는 업무 활용규칙을 정확하게 분석하여 이를 데이터베이스 설계에 반영함으로써, 데이터베이스 차원에서 업무 처리시 발생하는 빈번한 데이터의 생성, 갱신, 삭제에 대해 데이터 정합성을 보장할 수 있어야 한다.

넷째는 업무 프로세스에 대하여 독립적

인 데이터 구조이다. 스템의 재개발이나 대폭수정이 불가피하며, 일반적으로 업무 프로세스개선과 함께 추진될 때 진정한 효과를 기대할 수 있다. 후자는 기존의 운영데이터베이스와 애플리케이션 시스템은 그대로 사용할 수 있는 반면에 별도의 통합데이터베이스와 데이터전송시스템의 구축이 필요하다.

상기 두가지 방법 중 어느 것을 선택하는가는 각 기업의 사정이나 특성에 따라 결정하여야 하나, 기본적으로는 데이터 공유의 기대효과 대비 비용측면이 우선적으로 고려되어야 한다. 일반적으로 데이터 공유를 위해 막대한 비용을 투자하여 개발된 기존의 운영시스템을 전면 재개발하는 것은 경제성측면에서 바람직하지 못하다. 모든 운영데이터를 빠짐없이 통합하기 위해서는 운영데이터베이스를 전면 재구축하는 것이 이상적인 방법이나, 애플리케이션 시스템의 재개발이 현실적으로 매우 어렵다는 문제가 있다. 기존의 애플리케이션 시스템은 기업의 중요한 일상업무처리를 지원하고 있기 때문에 재개발에 따르는 위험이 존재하고, 방대한 규모와 복잡한 구조로 인해 많은 기간과 비용이 소요된다.

따라서, 이에 대한 현실적인 대안으로 많은 기업에서 채택하고 있는 것이 후자의 방법이다. 즉, 일상업무처리하는 기존의 운영시스템을 통하여 계속 지원하면서, 운영데이터저장소(ODS : Operational Data Store)라고 하는 별도의 통합데이터베이스를 구축하여, 복수 업무간 연계처리나 분석을 통한 의사결정지원, 보고서작성 등과 같은 데이터공유 기반의 업무를 지원하는 것이다.

운영데이터저장소특성

운영데이터저장소는 기존의 운영시스템으로부터 데이터를 공급 받아 통합된 구조로 관리하는 데이터베이스를 말한다. 기업이 운영데이터저장소를 구축하는 목적은 운영업무처리를 지원하는데 있으며, 이때 운영데이터저장소는 기존 운영시스템에 이질적으로 정의되어 있는 운영데이터를 전사관점으로 통합함으로써, 데이터공유를 가능하게 한다.

운영데이터저장소는 데이터공유를 기반으로 두가지 역할을 수행한다. 하나는 운영시스템의 온라인 트랜잭션 처리를 지원하는 것이고, 다른 하나는 업무운영상의 중요한 의사결정을 지원하는 것이다. 온라인 트랜잭션은 빠른 응답시간과 소량의 데이터처리가 중요한 반면에, 의사결정지원은 응답시

인 데이터 구조이다.

기존의 운영데이터베이스는 단위업무시스템에 종속적인 데이터 구조를 갖는다. 이 때문에 데이터구조가 해당 업무프로세스 수행에는 최적화되어 있으나, 다른 업무프로세스에 대해서는 부적합하거나 비효율적일 수가 있다. 업무의 데이터 공유도를 향상시키려면 관련된 모든 업무프로세스가 데이터를 활용할 수 있는 범용적인 데이터 구조가 필요하다. 데이터구조가 범용성을 갖기 위해서는 특정 업무프로세스 중심이 아니라 기업 전체가 관심을 갖는 공통주제를 중심으로 데이터구조를 설계하여야 한다. 모든 기업의 관심대상이 되는 주제는 고객, 상품, 주문 등이 있다.

운영데이터 통합방법의 선택

운영데이터를 통합하는 방법은 크게 두가지가 있다. 하나는 기존의 운영데이터베이스를 통합구조로 전면 재구축하는 방법이고, 다른 하나는 기존의 운영데이터베이스는 그대로 두고, 별도의 통합데이터베이스를 구축하여 운영데이터를 통합하는 방법이다.

전자는 데이터베이스 재구축에 따른 기존 애플리케이션 시

간 보다는 다량의 데이터에 대한 효율적인 검색과 분석이 상대적으로 더욱 중요하다. 이러한 두가지 역할을 모두 수행하기 위해서, 운영데이터저장소는 다음과 같은 특성을 갖는다.

첫째, 주제중심으로 데이터를 관리한다. 모든 시스템에 대하여 전사적인 관점의 데이터를 제공할 수 있도록, 특정 시스템이나 업무기능관점으로 구성되어 있는 기존 운영데이터를 고객, 제품 등과 같은 전사공통주제 중심으로 재구성한다.

둘째, 통합된 데이터구조를 갖는다. 통합데이터를 요구하는 애플리케이션이나 전사범위의 업무기능을 지원하기 위해서는, 복수의 운영시스템에서 이질적으로 관리되고 있는 운영데이터를 전사관점에서 통합하여 데이터의 정합성과 일관성을 보장할 수 있어야 한다.

셋째, 현재시점의 데이터를 관리한다. 운영데이터저장소의 기본역할은 운영업무를 지원하는 것이기 때문에, 기존 운영시스템에 있는 현재시점의 데이터를 관리한다. 단, 여기서 현재시점이란 각 업무요건에 따라 다르게 해석될 수 있다. 만약에 업무요건이 3개월간의 계정데이터를 현재데이터단위로 본다면, 운영데이터저장소는 3개월치의 계정데이터를 현재시점의 데이터로 관리하게 된다. 즉, 업무요건에 따라 3~6개월정도의 이력데이터를 포함할 수 있다.

넷째, 데이터 갱신이 발생한다. 운영데이터저장소는 현재시점의 데이터를 관리하기 때문에, 업무요건별로 정의된 현재시점에 해당하는 기간에 따라 데이터를 갱신하게 된다. 이는 곧 데이터의 내용이 시간이 경과함에 따라 변경됨을 의미하며, 따라서 동일한 질의를 실행하더라도 질의시점에 따라 다른 결과를 얻을 수 있다.

마지막으로 상세수준의 데이터를 요구한다. 운영데이터저장소는 기본적으로 운영업무처리를 지원한다. 데이터의 상세수준은 운영데이터저장소를 활용하는 업무요건에 따라 달라지게 된다. 온라인 트랜잭션 처리와 같은 업무는 기존 운영데이터와 동일한 수준의 상세데이터가 필요할 것이며, 의사결정을 위한 분석이나 보고서작성 등의 업무는 일정수준의 요약데이터를 필요로 할 것이다.

운영데이터저장소를 통해 완전한 데이터공유를 달성하기 위해서는 통합데이터베이스인 운영데이터저장소 외에, 소스 데이터인 운영시스템의 데이터베이스로부터 운영데이터저장소로 데이터를 공급해주는 데이터전송시스템과 데이터전송

시스템에게 데이터전송경로와 전송방법을 알려주는 메타 데이터베이스 그리고 운영데이터저장소를 활용하는 애플리케이션 시스템으로 구성되는 체계를 갖추어야 한다.

운영데이터저장소 사용자유형

운영데이터저장소는 기업의 정보체계 상에 있는 데이터베이스 - 운영계 데이터베이스, 운영데이터 저장소, 데이터 웨어하우스, 데이터 마트, 메타데이터베이스 - 중에서 가장 복잡한 구조를 가진다. 따라서, 성공적인 운영데이터저장소 구축을 위해서는 최적의 데이터베이스 설계가 무엇보다 중요하다. 운영데이터 저장소 설계의 기본개념을 이해하기 위해서는 운영데이터저장소의 사용자에 대한 이해가 선행되어야 한다.

운영데이터 저장소의 사용자유형에는 업무담당자와 분석가의 두가지 유형이 있다. 업무담당자는 매일 동일한 업무를 반복적으로 처리하는 사람을 말한다. 따라서, 업무담당자는 업무처리시점에서 무슨 데이터를 검색해야 하는지를 알기 때문에, 항상 원하는 데이터를 얻는다. 업무담당자는 매 업무처리 때마다 소량의 데이터를 검색한다. 이러한 업무담당자에게는 정형화된 데이터구조와 업무절차가 요구된다.

분석가는 비정기적으로 분석업무를 수행하는 사람을 말한다. 분석가는 자신이 원하는 데이터를 분석시작시점에서는 알지 못하고, 분석수행과정에서 스스로 발견하게 된다. 분석가는 데이터간에 관련성이나 패턴을 발견하기 위해 대량의 데이터를 검색하며, 때에 따라서는 아무것도 발견하지 못할 수도 있다. 따라서, 분석가는 완전히 비정형적인 작업방식을 갖는다.

운영데이터저장소 모델링

운영데이터 저장소의 데이터 모델링은 전사차원의 정보요구를 반영하고 있는 아키텍처 모델을 출발점으로 한다. 우선 아키텍처모델을 바탕으로 업무관점에서 데이터요건과 업무활용규칙을 분석하여 논리모델을 작성한다. 논리모델은 분석결과를 엔티티 타입과 이들간에 관계로서 표현할 수 있도록 정규화 기법을 적용하여 작성하며, 일반적으로 3차수준까지 정규화를 실시한다.

논리모델이 작성되면 성능향상 관점에서 논리모델을 물리

(표) 논리모델과 물리모델 대응

논리모델	물리모델
엔티타입	테이블, 파일
엔티티	행, 레코드
속성	컬럼, 필드
식별자	키
관계	외부키

모델로 변형시킨다. 변형된 물리모델은 실제 데이터베이스로 구축되는 모델로서, 데이터의 규모, 데이터 활용유형 등

에 대해 최적의 성능을 발휘할 수 있도록 데이터구조를 조율한다. 물리모델에 적합한 데이터구조에는 역정규화구조(Denormalized Structure)와 성형구조(Star Schema)가 있다. 역정규화구조는 복수개의 정규화된 테이블 중에서 공통적인 키와 용도를 갖는 테이블이 결합된 데이터구조로서, 포괄적으로는 정규화 구조에 속한다. 이러한 역정규화를 통해 일대일 관계를 갖는 테이블에 대해, 테이블 참조시 액세스성능을 향상시킬 수 있다. 그러나, 역정규화에도 불구하고 여전히 조인(Join)되는 테이블 수가 많거나, 테이블에 포함되는 데이터 규모가 큰 경우에는 만족할 만한 수준의 성능향상을 기대하기 어렵다.

성형구조는 데이터의 규모나 데이터의 활용패턴이 설계의 중요한 요소가 되는 경우에 적합한 데이터구조다. 성형구조는 사실(Fact)테이블과 차원(Dimension)테이블 그리고 이

들간에 관계로 이루어져 있다. 사실테이블은 사용자의 액세스 대상이 되는 테이블로서 데이터와 차원테이블로부터 반입된 외부키(Foreign Key)로 구성된다. 차원테이블은 사용자가 사실테이블을 다양한 관점으로 액세스할 수 있도록 지원하며, 외부키에 의해 사실테이블과 연계된다.

사실테이블은 데이터를 미리 조인된(Pre-joined) 형태로 관리함으로써 효율적인 데이터엑세스를 지원한다. 데이터를 미리 조인된 형태로 관리하기 위해서는 테이블설계시점에서 이미 사용자의 데이터 액세스 및 활용패턴을 정의할 수 있어야 한다. 운영데이터저장소의 상반된 두가지 역할 - 일상업무처리지원과 의사결정지원 - 을 감안할 때, 정규화 구조나 성형구조 중 어느 것도 단일구조로서 최적화된 데이터구조를 제공하지 못한다. 결국 두가지 역할을 모두 수행하기 위해서는, 정규화 구조와 성형구조가 혼합된 복합구조(Hybrid Structure)가 불가피하다. 운영데이터저장소에서 정규화 구조와 성형구조의 적용범위는 사용자유형과 사용자의 업무성격에 따라 달라진다.

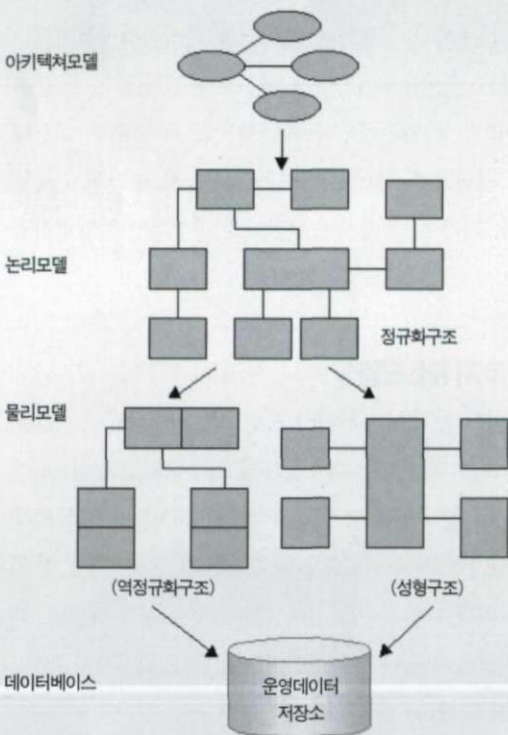
운영데이터저장소가 업무담당자에 의해 의사결정지원 목적으로만 사용된다면, 사전에 데이터 액세스패턴을 정의할 수 있는 성형구조가 적합하다. 그러나, 업무담당자의 온라인 갱신업무가 추가되거나 분석가 사용자유형이 존재한다면 이 부분에 대해 정규화 구조가 필요하다. 그리고 사용자가 분석가 밖에 없는 경우에는 전체데이터구조가 비정형적 질의에 보다 효과적인 정규화 구조를 가질 수 있다.

운영데이터저장소 데이터 획득

운영데이터저장소의 데이터획득 절차는 크게 운영시스템의 데이터베이스로부터 필요한 데이터를 추출하고, 추출된 데이터를 운영데이터저장소의 통합구조에 맞게 재구성하여, 운영데이터저장소로 전송 및 적재하는 세가지의 단계로 구분될 수 있다.

최초의 데이터추출은 복수의 소스데이터로부터 필요한 데이터를 얻는 과정인데, 이를 위해서는 데이터모델링단계에서 운영데이터저장소의 모든 데이터항목에 대해 정확한 데이터를 공급해주는 소스데이터가 결정되어야 한다.

데이터의 재구성은 복수의 소스데이터로부터 추출된 데이터를 운영데이터저장소 구조로 재편성하는 과정인데, 이때



(그림 3) 운영데이터저장소 모델링

시스템기준에 맞게 정의되어 있는 소스데이터의 형식이나 데이터값을 운영데이터저장소의 표준화된 데이터구조에 적합하도록 변형시키게 된다. 따라서 데이터 모델링 단계에서 소스데이터가 결정되면, 소스데이터와 운영데이터저장소 데이터와의 대응관계 및 소스데이터의 변형규칙을 분석하고, 이를 메타데이터베이스에 관리한다.

추출된 데이터의 재구성이 완료되면, 이를 운영데이터저장소가 있는 목표환경으로 전송한 후, 해당 테이블에 적재한다.

운영데이터저장소는 이러한 데이터획득 절차를 통하여 데이터를 공급받고, 운영시스템의 데이터베이스와 데이터를 일치시키게 된다. 데이터획득 절차는 ETL(Extraction Transformation Load)기술을 통해 구현되는데, ETL은 메타데이터베이스(Meta Data Base)에 정의된 데이터의 대응관계와 변형규칙을 근간으로 데이터의 추출, 변형, 적재를 실행한다.

운영데이터저장소 유형

운영데이터저장소의 모든 데이터를 소스데이터와 실시간으로 일치시키기 위해서 갱신할 필요는 없다. 데이터일치는 업무적인 관점에서 데이터의 유효성이 상실되지 않는 주기로 실행되어야 한다.

"Building The Operational Data Store"의 공동저자인 W.H Imnon과 Claudia Imhoff는 운영데이터저장소에 대한 데이터 갱신빈도를 기준으로 운영데이터저장소의 유형을 다음의 세가지로 정의하고 있다.

첫째가 클래스I 운영데이터저장소(Class I ODS)유형인데, 데이터갱신이 운영시스템과 거의 동시에 발생한다. 클래스I 운영데이터저장소는 실시간에 가까운 2~3초 주기로 갱신되어야 하기 때문에 소스데이터로부터 데이터 추출후 변형단계

에 많은 시간을 할애할 수 없다.

따라서, 클래스I 운영데이터저장소는 소스데이터에 대한 변형이 거의 발생하지 않고, 빠른 성능이 요구되는 트랜잭션 처리중심 환경에

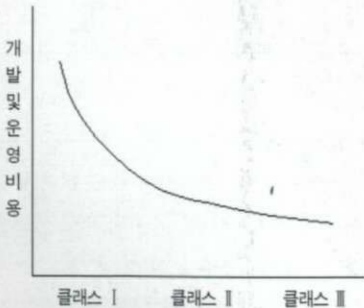
적합하다. 클래스I 운영데이터저장소는 실시간 갱신에 필요한 기술적 요건 때문에 구축에 많은 비용이 소요된다.

둘째는 클래스II 운영데이터저장소(Class II ODS)유형으로, 운영시스템과 1~2시간주기의 갱신빈도를 갖는다. 따라서, 클래스I 운영데이터저장소에 비해 보다 많은 데이터변형 작업을 처리할 수 있다. 클래스II 운영데이터저장소는 운영시스템의 데이터를 로그(Log)와 같은 중간파일에 일단 저장한 후 전송하는 지연처리방식을 사용한다.

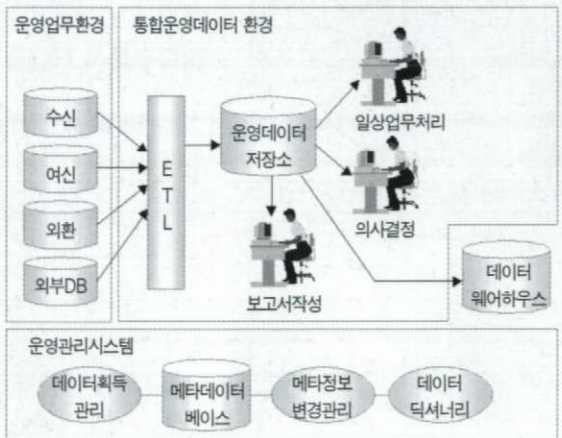
셋째는 클래스III 운영데이터저장소(Class III ODS)유형으로, 하루 1회 배치(Batch)처리에 의해 갱신된다. 이와 같은 여유있는 갱신빈도 때문에, 다수의 소스데이터로부터 추출된 데이터를 대상으로 인덱스 재구성, 오류정정, 데이터값 계산 등과 같은 다량의 변형 작업수행이 가능하다.

운영데이터저장소 운영

운영데이터저장소는 구축이 완료되면, 데이터의 초기적재(Initial Load)를 거쳐 운영단계로 접어 든다. 운영단계에서는 통합된 운영데이터를 바탕으로 업무담당자와 분석가의 업무를 지원하고, ETL시스템을 통해 소스데이터로부터 데이터를 공급받는다. 아울러 데이터웨어하우스가 구축되면, 운영데이터의 일부를 데이터웨어하우스로 공급하는 데이터소스가 되기도 한다. 이상의 운영데이터저장소와 관련된 제반활동은 메타데이터베이스를 기반으로 하는 운영관리시스템에 의해 지원 및 모니터링되며, 모니터링된 결과를 토대로 운영데이터저장소에 대한 보완 및 개선을 수행한다.



(그림 4) ODS 유형 구축비용



(그림 5) 운영데이터저장소 운영환경