

비즈니스의 정확도와 컴플라이언스 강화를 위한 데이터 품질 개선

IBM DB2 Time Travel Query를 이용한 이력 데이터 관리



Matthias Nicola
IBM Silicon Valley Laboratory
mnicola@us.ibm.com

시간이란?

알버트 아인슈타인은 "시간이 존재할 수 있는 유일한 이유는 모든 것들이 한꺼번에 일어나는 것이 아니기 때문"이라고 말했습니다. 결국 시간은 이벤트를 차례로 배열하고, 이벤트 간 빈도를 측정하고, 변화 속도를 수량화할 수 있게 해주는 측정 체계를 말하는 것입니다. 실제로, 시간이 중요한 이유는 시간이 바로 변화의 경험을 나타내기 때문입니다. 시간이란 개념이 없으면, 변화도, 움직임도, 비즈니스 수치의 증가와 감소도 존재하지 않을 것입니다. 조직은 변화를 파악하고 측정하기 위한 중요한 목적을 달성하기 위해 시간을 추적하고 파악해야 합니다.

본 백서는 시간을 기반으로 하는 이력 데이터 관리에 대해 설명합니다. 여기서 기존 방식의 문제점을 알아보고 IBM® DB2® 데이터베이스 소프트웨어가 제공하는 이력 기능의 이점에 대해 설명합니다. 이러한 기능을 통틀어 DB2 Time Travel Query라고 하며, 이 기능은 이력 쿼리, 이력 업데이트 및 이력 제한 등을 포함합니다.

시간 여행에 대한 요구

시간 여행이야말로 항상 현재에 국한된 인간의 한계를 넘어 과거를 돌아보거나 미래를 예측하기 위한 인류의 꿈이었습니다. 그럼에도 과거와 미래는 인간의 모든 노력에 대해 가장 중요한 요소입니다. IBM DB2가 제공하는 "데이터 타임머신"은 조직이 비즈니스를 수행할 때 하나 이상의 시간 축을 자유자재로 이동할 수 있게 해줍니다.

이러한 시간 여행을 가능하게 하는 조직의 2가지 주요 요건은 다음과 같습니다.

- GRC(Governance, risk and compliance) 이니셔티브 및 이와 관련해 감사 요구사항을 충족하기 위한 요구
- 의사결정 능력을 개선하기 위해 과거와 미래의 실행일 등 비즈니스 정보를 이용하기 위한 조직의 시간 차원 관리



GRC 및 감사 요구사항 충족

많은 조직들은 컴플라이언스 및 감사 규정에 따라 데이터 변경 이력을 기록하여 과거로 돌아가 원하는 시점에 존재하는 데이터를 식별할 수 있어야 합니다. 가장 많이 적용되는 분야는 금융권으로, 이 분야에서 기업은 컴플라이언스의 확장과 감사 요구사항에 따라 달라질 수 있으며 다음과 같은 질문에 답변할 수 있어야 합니다.

- 2개월 전 주어진 투자에 대한 가격과 리스크를 평가할 당시 어떤 데이터를 사용했는가?
- 해당 과거 상태를 근거로 데이터를 이용해 리스크 평가를 재실행할 수 있는가?
- 과거에 수집된 데이터를 모두 추적하여 수정 유무에 상관없이 과거 데이터를 조회할 수 있는가?
- 조직 내 모든 부서가 동일한 시간대에 같은 데이터를 사용하고 있다고 입증할 수 있는가?

의료 등 타 업계에서도 이와 유사한 질문을 합니다. 어떤 의료진은 6개월 전 진료 방식을 바꾸고 새로운 투약을 권할 당시 환자의 병력에 대해 어떻게 이해하고 있었는지 물을 수도 있습니다.

데이터의 정확성과 완전한 이력이 갖추어지지 않으면 심각한 결과를 초래할 수 있습니다. 불완전한 데이터 이력은 금융 거래 시 투자 손실 등의 비용 오류 또는 심각한 오진으로 이어질 수 있습니다. 여기에 컴플라이언스 위반으로 인한 벌금 부과는 물론, 심각한 법적 문제로 번져 명예 훼손으로 이어지기도 합니다.

DB2는 조직의 컴플라이언스 요구를 해소하기 위해 다음과 같은 기능을 제공합니다.

- 시스템 시한으로 관리되는 시스템 기반의 시간 스탬프를 이용해 모든 시스템 데이터의 변경 이력을 투명하게 유지
- 애플리케이션 또는 최종 사용자가 피코 초 이내에, 즉, 신속하게 과거의 원하는 시점을 선택하여 데이터에 액세스할 수 있는 간단한 쿼리 작성 가능
- ISO/IEC 9075:2011 (SQL:2011) 표준을 지원하여 이력 쿼리의 복잡성 감소

비즈니스 시한 관리

많은 조직들은 비즈니스의 시행 일자 또는 유효 기간 등 자사의 비즈니스에 필요한 시간 차원도 관리해야 합니다. 유효 시간으로 알려진 이 비즈니스 시간은 시스템 기반 데이터의 변경 이력을 나타내는 시스템 시간과 달리 비즈니스 정보의 유효성이 실제로 적용되는 시한을 나타냅니다.

비즈니스 시간을 예로 들자면 다음과 같습니다.

- 어떤 대출에 대해 특정 이자가 유효한 시한
- 어떤 고객에 대해 특정 주소가 유효한 시간
- 어떤 제품이 판매되는 기간
- 계약서 또는 보험증서에 업데이트된 조항 및 조건의 효력일

이러한 비즈니스 기간은 과거, 현재 및 미래 언제든지 적용할 수 있어야 합니다. 그렇기 때문에, 조직은 과거, 현재 및 미래 중 언제든지 비즈니스 정보를 삽입, 수정 또는 쿼리할 수 있어야 합니다.

이에 대한 예로, 대출 이자가 다음 달 1일에 1% 인상되고, 1년 간 유효한 경우를 들 수 있습니다. 이것은 미래의 데이터를 업데이트한 경우입니다. 또는, 어느 고객이 가입한 보험회사에 2주 전 자신의 주소가 변경되었음을 통보하면 자동차 보험료에 프리미엄 비용이 추가될 수 있습니다. 이 고객이 가진 기존 주소의 효력일은 2주 전 끝나고 새 주소의 효력이 그 주부터 시작되어야 합니다. 이것은 과거의 데이터를 업데이트한 경우입니다.

이러한 맥락에서 요구되는 또 다른 공통 사항은 시간 고유의 특성과 같은 이력 제한을 강화하는 것입니다. 예를 들어, 은행은 불확실성을 막기 위해 어떤 시점에서나 오로지 하나의 이자율만 유효하도록 대출에 적용시켜야 합니다. 이처럼 예약 시스템은 2명의 고객이 겹치는 시한에 같은 호텔 객실을 이용하도록 예약해서는 안 됩니다.

“DB2 10은 혁신적인 소프트웨어로 데이터베이스 관리를 급격히 간소화할 수 있게 해줍니다. 테스트를 통해 몇 가지 신기능을 사용한 결과 고객에게 제공되는 서비스가 급격하게 늘어날 것으로 예측할 수 있었습니다. 이러한 고급 기능으로 포함되는 Time Travel Query는 고객이 특정 시한에 데이터를 저장하고 가져올 수 있게 도와주고, 장기간에 걸쳐 이자율의 변동이 심한 보험 증서, 신용카드 및 대출 업무를 처리하는 고객에게 매우 유용합니다.”

- Jithen Bondada, Miracle Software Systems, Inc의 애플리케이션 개발/B2B 담당 부사장.

DB2는 기업이 다음과 같은 오퍼링을 통해 이러한 요구사항을 해소할 수 있게 해줍니다.

- 과거, 현재 및 미래 중 언제든지 유효일을 비즈니스 시한처럼 관리하여 삽입, 수정 및 쿼리할 수 있는 포괄적 능력 제공
- 이력 제한의 선택적 강화
- SQL: 2011 표준 지원을 통해 데이터베이스 내 비즈니스 시간 관리의 간소화

양이측 데이터 관리의 중요성

조직은 종종 비즈니스 시한을 관리하고 모든 변경 및 수정 사항을 모두 갖추고 있어야 합니다. 그렇기 때문에 데이터는 비즈니스 시간과 시스템 시간 등 2개의 시간 차원으로 관리해야 합니다. 이런 탄탄한 결합 방식을 양이측 데이터 관리라고 합니다.

예를 들어, 양이측 데이터 관리만이 "1월 1일을 기준으로 연말까지 예상된 대출액은 얼마였는가?"라는 비즈니스 질문에 답변할 수 있습니다. 이 질문에서 "1월 1일"은 1월 데이터베이스에 기록된 최신 데이터를 감안했기 때문에 시스템 시간을 나타냅니다. 이와 반대로, "연말"은 여기서 미래의 유효한 날짜를 지칭하므로 비즈니스 시간을 나타냅니다.

양이측 데이터 관리는 알고 있는 사항과 알고 있는 시점에 대한 포괄적 정보를 제공합니다. 또한, 과거, 현재 및 미래에 수정된 비즈니스 데이터를 추적할 수도 있습니다. 그 결과, 양측 데이터 관리는 조직이 비즈니스 리스크를 줄이고, 컴플라이언스를 통해 목표를 달성하고 의사결정 능력을 강화할 수 있게 도와줍니다.

DB2는 양이측 데이터를 관리하기 위해 시스템 시간과 비즈니스 시간의 기능을 통합합니다. 또한, SQL: 2011 표준을 지원하기 때문에 시간 차원과 함께 데이터 관리를 간소화할 수 있게 도와줍니다.

시간 기반 처리 로직 구현에 치중하는 기존 방식의 단점

조직은 그 동안 내부 개발 솔루션을 구현하여 이력 데이터 관리의 요건을 해소해 왔습니다. 이런 솔루션은 일반적으로 애플리케이션 코드 또는 맞춤형 저장 프로시저 및 데이터베이스 트리거 단계 등으로 시간 기반 데이터 처리 로직을 구현하는 수작업 코딩을 거쳐야 했습니다. 그러나, 이 방식은 다음과 같은 근본적인 단점을 가지고 있습니다.

- 시간 기반 처리 로직을 구현하는 것은 매우 복잡합니다.
- 애플리케이션 개발과 유지보수에 드는 비용이 높습니다.
- 애플리케이션 성능이 많이 떨어지는 편입니다.
- 맞춤형 코드는 이력 작업의 표준화를 저해함으로써, 상이한 애플리케이션 또는 조직 내 부서마다 시간 기반 처리 로직을 서로 다르게 구현하는 경향이 있습니다.

이러한 과제는 하나의 애플리케이션 또는 여러 애플리케이션에 걸쳐 이력 데이터의 일관성을 유지할 수 없게 만듭니다. 그 결과, 조직은 기껏해야 데이터 품질을 조율하고 리스크에 노출되지 않는 수준에 그치고 맙니다.

다음에 나타난 제품 가격은 시간 기반 로직 구현에 따른 복잡성을 예시한 것입니다. 그림 1의 데이터 행은 제품 ID가 123이고 가격이 \$19.95로 효력 시작일은 2012년 1월 1일이고, 효력 종료일은 2012년 6월 1일입니다. 효력 시작일은 가격을 적용할 수 있는 첫째 날을 말하고, 효력 종료일은 가격을 더 이상 적용할 수 없게 된 첫째 날을 말합니다.

제품 ID	가격	효력 시작일	효력 종료일
123	\$19.95	2012-01-01	2012-06-01

그림 1: 제품 가격의 데이터 행 샘플 예

3월 달 제품 가격에 할인을 적용하여 \$15.00에 판매할 경우, 해당 데이터는 그림 2와 같이 수정해야 합니다.

제품 ID	가격	효력 시작일	효력 종료일
123	\$19.95	2012-01-01	2012-03-01
123	\$15.00	2012-03-01	2012-04-01
123	\$19.95	2012-04-01	2012-06-01

그림 2: 제품 가격의 데이터 수정 예

그림 2의 첫째 행은 정가가 \$19.95로, 1월과 2월에 걸쳐 해당 가격을 적용할 수 있지만, 3월 1일부터는 효력이 종료됩니다. 둘째 행에 나타난 \$15.00은 3월에 적용되는 새 가격이고, 셋째 행은 4월과 5월에 걸쳐 원래 가격으로 다시 적용된다는 의미입니다.

“DB2가 Bitemp 2.0과 같은 2세대 양이측 데이터 관리의 선두가 될 수 있는 이유는 표준화된 SQL 구문을 이력 작업에 이용하고 데이터베이스 엔진을 심층적으로 통합하기 때문입니다!”

- Craig Baumunk, BitemporalData.com 원장

그림 1의 데이터를 그림 2의 데이터로 변환하는 것을 실용 연산(nontrivial operation)이라고 합니다. 이러한 복잡성은 기존 날짜(2012-01-01, 2012-06-01)와 새로 제공된 날짜(2012-03-01, 2012-04-01)가 연관된 방식에 따라 기존 데이터를 수정하고, 새 행을 작성해야 하기 때문에 비롯됩니다.

제품 123의 가격이 3월 15일부터 4월 15일까지 \$10.00로 변경되면 그림 2는 어떻게 바뀔까요? 이 새 기간은 그림 2에 있는 2개의 기존 기간과 겹치기 때문에 필수 변환을 더 복잡하게 만듭니다. 애플리케이션은 기존 행을 모두 검토하여 어떤 행이 3월 15일부터 4월 15일까지 일어나는 가격 변경에 영향을 받는지 결정합니다. 그런 다음, 해당 행을 업데이트하는 방법과 삽입할 새 행을 결정합니다. 이와 동시에, 애플리케이션은 제품 123에 대한 상이한 가격들이 기간에 겹치지 않도록 해야 합니다. 이는 또 다른 복잡한 작업입니다.

DB2를 통한 이력 데이터 관리의 이점

DB2를 이용하면 데이터 수정 등의 시간 작업을 애플리케이션 로직으로 코딩할 때 더 이상 수작업을 거치지 않아도 됩니다. 해당 애플리케이션은 표준 SQL을 이용해 원하는 변경사항을 서술형으로 표현할 수 있고, DB2가 필요한 데이터 수정을 결정하고 수행할 수 있게 해줍니다. 그림 3은 2개의 간단한 SQL 문으로 어떻게 제품 가격 예에서 논의했던 가격 변경을 수행할 수 있는지 보여줍니다.

```
UPDATE product
FOR PORTION OF BUSINESS_TIME FROM '2012-03-01' TO '2012-04-01'
SET price = 15.00
WHERE productID = 123;
```

```
UPDATE product
FOR PORTION OF BUSINESS_TIME FROM '2012-03-15' TO '2012-04-15'
SET price = 10.00
WHERE productID = 123;
```

그림 3. IBM DB2는 간단한 SQL 문을 이용해 복잡한 이력 데이터 변환을 수행합니다

FOR PORTION OF BUSINESS_TIME 절은 그림 3에 있는 것처럼 SQL: 2011 표준으로 구성된 여러 언어 중 하나로 이력 데이터 관리를 간소화합니다. 이 SQL 표준을 지원하는 DB2는 다음과 같은 이점을 제공합니다.

- 이력 작업 시 복잡하고 순차적인 애플리케이션 코딩 방식이 아닌 짧고 간단하고 서술적인 코딩 이용

- 개발자 생산성을 높이고 애플리케이션 개발 및 유지보수의 비용 감소
- 애플리케이션에 걸쳐 시간 처리 로직을 표준화함으로써 일관성을 높이고 IT 투자 보호 지원
- 전사적 데이터 품질 향상 및 비즈니스 리스크 감소
- 이력 작업을 DB2 엔진에 심층적으로 통합함으로써 성능 향상

IBM 연구에 따르면, 제품 가격의 예에 소개된 것처럼 비즈니스 이력 작업을 수행하는 데 필요한 3개의 상이한 옵션을 비교한 결과는 다음과 같습니다.

1. 이력 작업에 맞는 SQL의 표준화를 이용한 DB2 기반 옵션
2. 일련의 저장 프로시저로 구현된 내부 개발 옵션
3. Java로 구현된 내부 개발 옵션

3개의 옵션 모두가 똑 같은 이력 작업을 구현했다더라도, DB2가 제공하는 비용 절감은 코딩과 테스트 노력 측면에서 매우 놀라웠습니다. DB2 기반 방식은 양 쪽의 내부 구현과 비교할 때 코딩 요구사항을 90%나 줄여주었습니다(그림 4 참조). 주요 비즈니스 시간 로직을 SQL 보관 절차 또는 Java로 구현할 경우 코드 라인이 각각 16배, 45배가 필요했는데 이것은 새 DB2 기능을 이용하는 간이 SQL 문과 맞먹는 분량입니다.

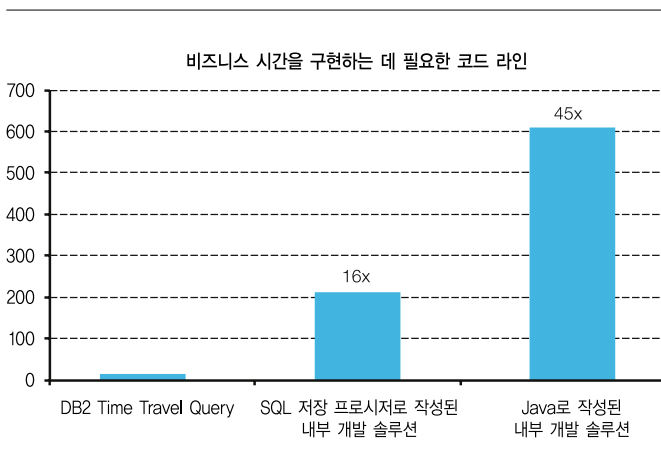


그림 4: 코드 라인은 애플리케이션 개발 및 유지보수 비용에 필요한 기준을 제시합니다

뿐만 아니라, IBM 테스터가 필수 SQL 문을 개발하고 테스트하는 데 1시간도 걸리지 않습니다. 이와 반대로, 내부 개발 방식 중 하나로 코딩하고 테스트하는 데는 4 ~ 5주가 걸렸습니다. 코드의 길이와 복잡성은 개발비 외에도 애플리케이션 유지보수 비용에 막대한 영향을 미칩니다.

결론: DB2 타임머신을 시작하기 위한 준비

IBM DB2는 애플리케이션이 같은 데이터에 액세스할 수 있게 해주는 단일 타임머신과 같은 기능을 발휘합니다. 이력 데이터 처리는 DB2 엔진과 심층적으로 통합되어 성능 향상은 물론, 타사 데이터베이스 기능 및 유틸리티와 양이측 데이터 관리의 통합을 간소화시켜 줍니다.

또한, IBM 연구에 따르면 DB2는 복잡한 시간 작업에 대해 SQL:2011 코딩 시간을 단축하고 효율성 향상을 지원함으로써, 개발자의 생산성을 높이고 시간 기반 애플리케이션을 구축하고 유지보수하는 비용을 최대 45배나 줄여줍니다.

요약하면, DB2는 전사적으로 데이터의 일관성과 품질을 개선하고, 비용 효율적인 방식을 제공하여 감사 및 컴플라이언스 문제를 해결해 줍니다. 그 결과, 조직은 컴플라이언스 위반으로 인한 리스크를 줄이고 비즈니스의 정확도를 높일 수 있습니다.

추가 정보

IBM DB2의 이력 기능에 대한 자세한 내용은 다음 리소스를 참조하십시오.

- 백서 다운로드 및 전문 보기: "A Matter of Time: Temporal Data Management in DB2" (ibm.com/developerworks/data/library/techarticle/dm-1204db2temporaldata/index.html)
- DB2 이력 토론 포럼에 방문하여 궁금한 사항에 대해 질문하십시오. ibm.com/developerworks/forums/forum.jspa?forumID=2316
- DB2 이력 관련 모범 사례 사이트: ibm.com/developerworks/data/bestpractices/db2luw/#bestpractices



© Copyright IBM Corporation 2012

(135-270) 서울시 강남구 도곡동 467-12
군인공제회관빌딩

한국아이비엠주식회사
고객만족센터

TEL: (02)3781-7114
www.ibm.com/kr

2012년 7월

Printed in Korea
All Rights Reserved

IBM, IBM 로고, ibm.com 및 DB2는 전 세계에 등록되어 있는 International Business Machines Corp.의 상표입니다. 기타 회사, 제품 및 서비스 이름은 타사의 상표 또는 서비스표입니다. 현재 IBM 상표 목록은 웹사이트 "저작권 및 상표 정보" (ibm.com/legal/copytrade.shtml)에 있습니다.

Java 및 모든 Java 기반 상표 및 로고는 Oracle 및/또는 해당 자회사의 상표 또는 등록상표입니다.

이 문서는 처음 발행될 당시의 날짜를 기준으로 업데이트되었으며 IBM은 언제든지 문서 내용을 변경할 수 있습니다. 일부 오퍼링은 IBM 매장이 있는 국가에서도 제공되지 않습니다.

여기서 소개된 성능 데이터와 고객 사례는 오로지 예시를 목적으로 작성된 것입니다. 실제 성능 결과는 특정 구성 및 운영 환경에 따라 달라질 수 있습니다. IBM 제품 및 프로그램과 함께 사용되는 기타 제품 또는 프로그램을 평가 및 검증하는 것은 사용자의 책임입니다. 이 문서의 정보는 상품성에 대한 보증, 특정 목적의 적합성 여부 및 저작권을 침해하지 않는다는 보증 또는 조건을 포함해 명시적 또는 암묵적 보증 없이 "있는 그대로" 제공됩니다. IBM 제품은 제공된 약정에 명시된 조항 및 조건에 따라 보증됩니다.

고객은 관련 법령과 규정을 반드시 지켜야 할 책임이 있습니다. IBM은 고객이 법령 또는 규정을 준수한다고 해서 당사의 서비스 또는 제품이 보증하는 법적 상담을 제공하거나 보증을 대신하지 않습니다. IBM의 향후 방향 및 의도와 관련된 내용은 통보 없이 변경되거나 취소될 수 있으며, 오로지 목표와 목적만을 나타냅니다.

¹ IBM 내부 테스트.



재활용 하십시오.