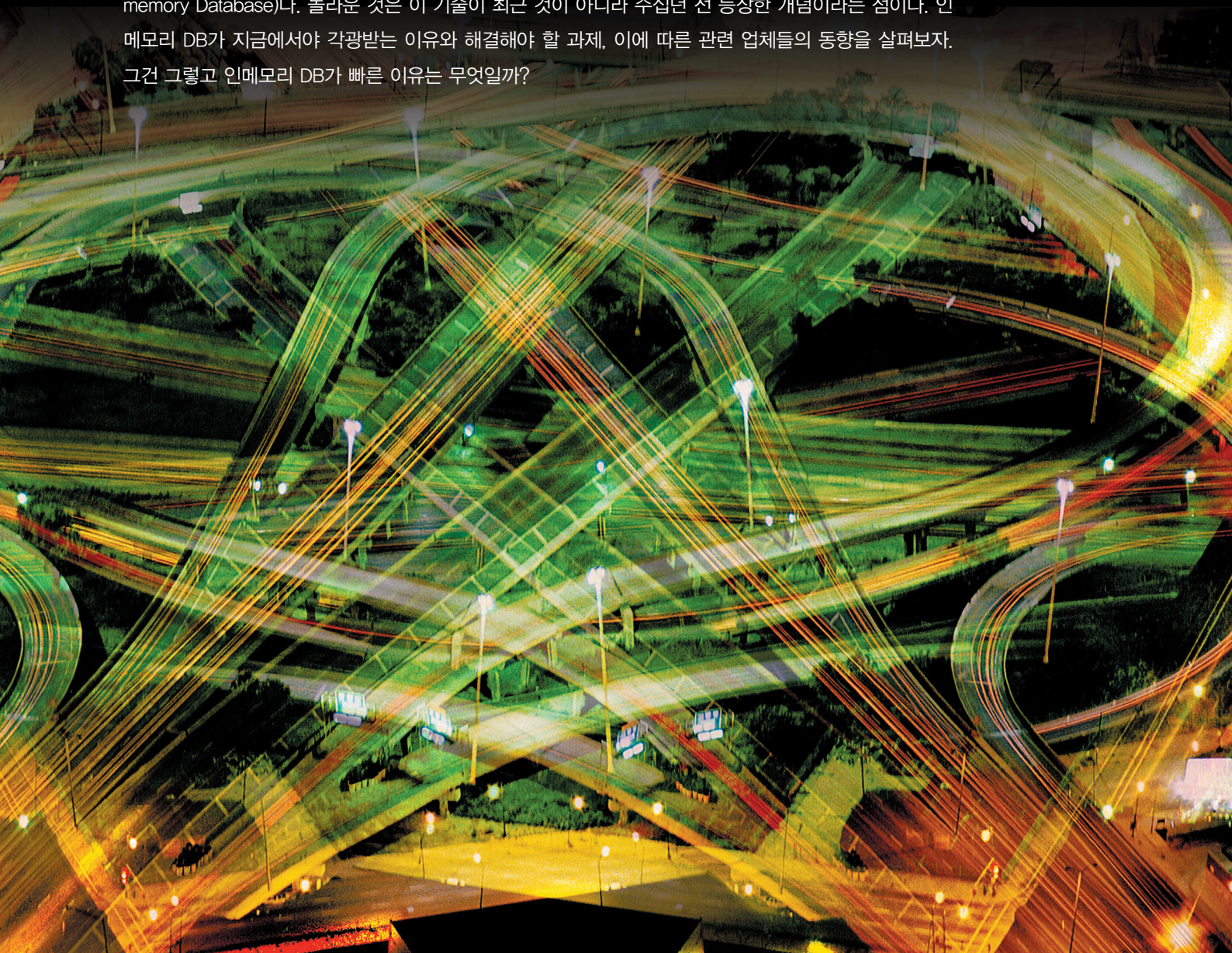


IDG Tech Report

“디스크에서 메모리로” 떠오르는 인메모리 DBMS의 이해

HDD 하드드라이브로 되어 있는 자신의 PC 혹은 노트북을 SSD로 바꿔 본 이들은 부팅 속도의 빠름에 대해 놀라움을 금치 못한다. 처음 SSD로 부팅했을 때 불과 몇 초만에 윈도우 부팅 화면을 뛰어넘어 바탕화면에 도달하는 경험을 했을 것이다. 그런데, 이렇게 빠른 SSD보다 1,000배나 빠른 것이 바로 인메모리 DB(In-memory Database)다. 놀라운 것은 이 기술이 최근 것이 아니라 수십년 전 등장한 개념이라는 점이다. 인메모리 DB가 지금에서야 각광받는 이유와 해결해야 할 과제, 이에 따른 관련 업체들의 동향을 살펴보자. 그건 그렇고 인메모리 DB가 빠른 이유는 무엇일까?



무단 전재 재배포 금지

본 PDF 문서는 IDG Korea의 프리미엄 회원에게 제공하는 문서로, 저작권법의 보호를 받습니다.
IDG Korea의 허락 없이 PDF 문서를 온라인 사이트 등에 무단 게재, 전재하거나 유포할 수 없습니다.

“디스크에서 메모리로” 떠오르는 인메모리 DBMS의 이해

이대영 기자 | ITWorld

2010년, 인메모리 DB인 HANA에 대한 대대적인 마케팅에 들어간 SAP를 필두로, 지난해에는 IBM과 EMC가, 올해는 시장 점유율 2위인 마이크로소프트와 함께 지난 30년 간 DBMS의 강자로 군림해 온 오라클마저 인메모리 기술을 자사의 DBMS 제품에 탑재하기에 이르렀다.

특히 2년전, 인메모리 기술에 대해 비웃기까지 했던 오라클 래리 엘리슨 회장이 지난해 자사의 인메모리 기술 탑재 발표 때, “디스크 드라이브는 이제 구식이 되고 있다. 앞으로 모든 데이터는 메모리에 있게 될 것이다”고 말한 것은 인메모리 기술에 대한 위상의 변화를 단적으로 보여주는 일례다.

인메모리 DB의 정의

인메모리 DB(In-memory Database)는 전통적인 디스크에 저장하던 DB 대신 메인 메모리(Random Access Memory, RAM)에 데이터를 저장하는 DBMS다. 디스크 기반의 DB보다 훨씬 빠른 응답속도를 자랑하는 인메모리 DB는 사실 최근에 나온 개념이 아니다. 인메모리 기술의 역사는 메인프레임 시절까지 거슬러 올라가며, 상용 인메모리 DB 제품은 10년전부터 존재해왔다.

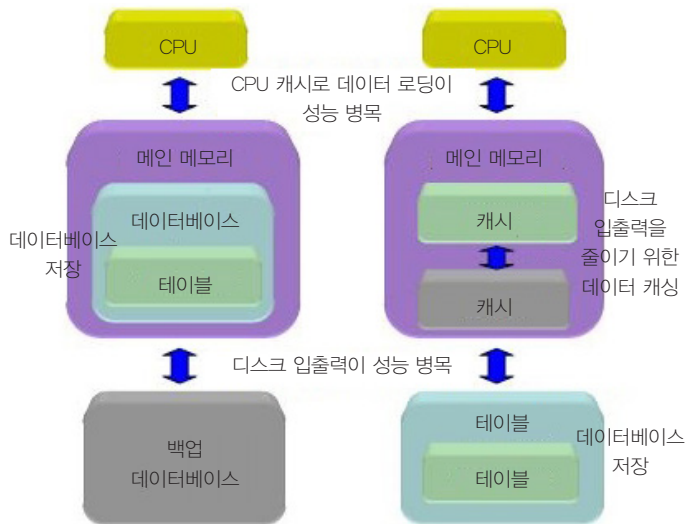
먼저 짚고 넘어가야 할 점은 인메모리 DBMS는 하드웨어인 RAM이 아니라 오라클의 DBMS나 IBM의 DB2, 마이크로소프트의 MS SQL과 같은 순수한 소프트웨어다. 한 마디로 인메모리 DBMS는 기존 디스크 기반의 DBMS가 제공하는 데이터 관리에 대한 모든 기능을 메모리를 기반으로 제공하는 시스템이다. 디스크 기반의 DBMS와는 아키텍처와 성능이 다를뿐, 기능적인 측면에서는 동일하다는 것이다.

디스크 기반 DBMS와 인메모리 DBMS와의 차이

두 DBMS를 구분 짓는 가장 큰 특징은 디스크 기반의 RDBMS의 저장 장소가 디스크인데 비해 인메모리 DBMS의 저장 장소는 메모리라는 점이다. 디스크 기반 DBMS의 기본적인 개념은 모든 데이터를 디스크에 저장, 관리한다는 것이다. 스토리지로서의 디스크는 메모리에 비해 저장된 데이터의 지속성이 우수한데 비해 상대적으로 접근 속도가 느리다.

그래서 디스크 기반의 DBMS는 메모리에다가 디스크에 저장된 데이터 가운

그림 1 | 인메모리 DB와 디스크 DB의 개념적 차이



데 자주 사용하는 데이터를 캐시(Cache)라는 임시 저장소를 두고 디스크 입출력 횟수를 줄이는 데 활용해 데이터 접근 속도를 높였다.

이 때, CPU는 캐시(Cache)에서 캐시 데이터를 관리하는 캐시 관리자를 통해 데이터에 접근하게 되는데, 한정된 캐시를 효율적으로 활용하기 위해 다양한 교체(replacement) 알고리즘을 구현함에 따라 데이터 접근과 질의 처리에 있어 알고리즘의 복잡도를 증가시키는 한계를 갖는다.

또한 캐싱(Caching)은 데이터를 읽어 들이는 데는 확실히 성능을 향상시키지만, 데이터를 업데이트하는 경우, 결국 버퍼를 지

나 디스크에 저장해야 하기 때문에 성능 향상에 도움이 되지 않는다.

그림 1의 경우 디스크 기반의 데이터베이스 구조가 상당히 간단하게 보이지만, 애플리케이션을 실행했을 때, 데이터의 흐름은 애플리케이션과 데이터베이스, 관리시스템 간의 상당히 복잡한 여러 개의 프로세스를 거치게 되면서 오히려 캐시가 성능을 저하시킬 수도 있다.

하지만, 인메모리 DBMS에서 메모리는 원본 데이터가 저장되는 곳이다.

인메모리 DB는 캐시라는 공간 자체를 데이터베이스로 삼았다. 메모리에 데이터베이스를 구축하고 디스크는 백업 데이터베이스 용도로 관리하는 구조다. 테이블과 인덱스를 포함한 데이터는 메모리에서 직접 저장, 관리한다.

이 때 메모리에 있는 데이터들은 디스크에 저장된 데이터에 대한 임시 저장소 역할이 아닌 주 저장소로서의 역할을 한다. 따라서 인메모리 DB는 데이터의 디스크 접근을 없애거나 최소화해 메모리와 디스크 간 병목 시간을 줄였기 때문에 빠른 것이다.

또한 디스크 기반 DBMS는 기본적으로 디스크에 모든 데이터가 있다고 가정하기 때문에, 데이터 접근 방식으로 RID(Record Identifier)를 사용한다. 따라서 메모리 내에서 실제로 해당 데이터에 접근할 때도 메모리의 물리적 주소로 매핑하는 주소변환이 필요하다. 이런 메모리 주소와 데이터베이스 주소 간 매핑의 단위 처리 시간은 아주 짧지만 고속의 데이터 처리시에는 무시할 수 없는 성능 병목 요소가 된다.

인메모리 DBMS는 데이터가 디스크에 존재한다는 가정이 사라져 메모리 내의 데이터베이스에 직접 접근할 수 있어 데이터 아키텍처면에서 효율적인 질의 처리를 할 수 있어 높은 트랜잭션 처리율을 보인다.

그러나 인메모리 DB는 상대적으로 비싼 메모리 가격과 저장 용량 한계 등 여러가지 문제점으로 인해 주류 시장에 편입되지 못하고, 통신, 금융, 국방, 공공,

제조, 유통, 임베디드 등에서 즉각적인 응답 시간과 높은 트랜잭션 처리를 요구하는 특수 영역에 국한되어 활용되고 있었다.

인메모리 DB의 분류

최근 인메모리 DBMS는 처리하는 데이터 종류에 따라 OLTP 중심, OLAP 중심, 그리고 OLTP/OLAP 겸용 인메모리 DB 등 세 가지로 분류된다.

고성능 트랜잭션 처리에 주로 이용되던 OLTP(Online Transaction Processing) 중심의 인메모리 DB 기술은 사용자 수의 폭발적인 증가와 데이터 양의 증가에 따른 대응 방안으로 분산 인메모리 기술로 발전하고 있다.

또한 몇년 전부터 애플리케이션 분야에서 분석 데이터의 좀더 빠른 처리를 위한 OLAP(Online Analytical Processing)용 인메모리 DB 기술이 등장했는데,

표 | 인메모리 DBMS의 분류

구분	인메모리 DBMS
OLTP 중심	IBM 솔리드 DB, 오라클 타임스텐, 애로스파이크, 카우치베이스, 볼트DB, MS SQL서버 인메모리 DB, 알티베이스, 선재소프트, 리얼타임테크
OLAP 중심	EXASOL, 코그니시오, IBM DB2 10.5
OLTP/OLAP 겸용	SAP HANA, 오라클 데이터베이스12c 인메모리, EMC 피보탈

이 기술이야말로 인메모리 DB를 DB 시장의 주인공으로 부상시켰다. 이와 함께 SAP HANA와 같은 양 분야에 모두 쓰이는 기술이 있다.

인메모리 DB가 부상하는 이유

최근 인메모리 DB가 주류 DBMS 시장에서 부상하게 된 것은 무엇보다도 컴퓨팅 환경 자체의 변화, 즉 컴퓨팅 성능의 진화 속도를 하드 디스크와 기존 DBMS 아키텍처가 따라가지 못했다는 점이 가장 큰 원인이다.

지난 20여 년간, CPU의 처리 성능은 약 150배, 메모리는 250배, 네트워크는 100배가 향상됐음에 비해 하드 디스크는 25배에 불과했다. 전체 컴퓨팅 환경이 크게 진화되어 연산 속도나 데이터 전송 속도는 엄청나게 빨라졌지만 정작 데이터를 끌어오고 저장하는 디스크와 이를 관리하는 DBMS가 발목을 잡고 있었던 것이다.

‘함대의 속도는 가장 느린 군함의 속도’라는 격언처럼 지난 20년간 전체 컴퓨팅 환경은 25배정도 향상된 셈이다. 이에 따라 가장 느린 군함의 속도를 높이기 위한 기술이 주목받는 건 당연한 이치다.

또한 ‘이것이 왜 지금이냐’는 물음에는 64비트 컴퓨팅이 보편화된 점을 들 수 있다.

메모리의 종류

메모리는 내부에 저장되는 데이터가 유지되는 시간에 따라 휘발성 메모리와 비휘발성 메모리로 나뉜다.

휘발성 메모리는 메모리칩이 전원을 유지하고 있는 동안에만 데이터를 갖고 있을 수 있는 반면, 비휘발성 메모리는 전원이 꺼져있는 동안에도 데이터를 저장할 수 있다. 따라서 비휘발성 메모리를 사용할 경우 전원이 꺼졌다가 다시 켜져도 이전에 사용하던 데이터를 그대로 사용할 수 있다.

비휘발성 메모리의 대표적인 것이 낸드 플래시 메모리다. 일반적으로 많은 사람들이 사용하고 있는 USB 저장장치, SD 카드, MMC, CFC(Compact Flash Card), 그리고 SSD 등에서 낸드 플래시 메모리가 사용된다.

휘발성 메모리의 대표격인 DRAM(Dynamic Random Access Memory)은 디스크 스토리지인 HDD보다 7만 배 빠른 응답 속도와 142배의 읽기 속도를 보이며, 플래시 메모리인 낸드 플래시보다는 1,000배 이상 빠른 응답 속도를 갖고 있다.

32비트 컴퓨팅 환경에서는 메모리에 4GB 이상의 데이터를 저장할 수 없었지만 64비트 컴퓨팅에서는 이론상으로 16EB(엑사바이트), 즉 1,677만 7,216TB를 메모리에 저장할 수 있게 됐다. 이로써 사실상 데이터나 애플리케이션 크기에 상관없이 메모리 상에서 모두 처리, 저장할 수 있다.

이와 함께 메모리 가격이 지속적으로 하락한 점도 무시 못할 요인 가운데 하나다. DRAM 가격은 매 18개월마다 약 30% 정도씩 떨어지고 있다. 가트너에 따르면, 2011년 GB당 DRAM이 약 10달러, 낸드 플래시는 1달러가 필요했지만, 2015년 DRAM 2달러, 낸드 플래시는 0.25달러가 될 것이라고 예측했다.

비용을 무시하고 오직 트랜잭션 성능과 속도를 중시한 특수 분야에만 활용되던 것이 메모리 가격 하락으로 인해 디스크와의 가격 대비 효과를 따지는 일반 DBMS 시장에까지 내려오게 된 것이다.

인메모리 DB에 대한 관심의 증가는 그간 인메모리 DB 시장이 기능과 성능만을 생각하고 약점에 대해서는 조금씩 해결해 나가던 얼리어답터 계층에서 가격 대비 효과를 따지면서 안정성과 범용성을 추구하는 초기 다수 계층에 다다르고 있다는 신호이기도 하다.

빅데이터와 빠른 분석 처리 요구가 새로운 국면 전환

빅데이터의 등장에 따른 데이터량의 증가, 빠른 분석 처리에 대한 시장의 요구는 인메모리 DB에 있어서도 새로운 국면을 맞이하게 됐다. 앞서 설명한 것처럼 늘어나는 사용자 수와 데이터량 증가에 대응하기 위해 OLTP 중심의 인메모리는 분산 인메모리 기술로 진화하고 있다.

분산 인메모리 기술의 핵심은 인메모리 DBMS가 선형적 확장성을 지원할 수 있도록 ▲애플리케이션 특성을 반영한 데이터 모델 ▲대규모 데이터 분할 ▲분산 질의 처리 ▲장애 대처 등에 대한 구현 방법론이다. 이에 대해 인메모리 DB 업체들은 각 방법론마다 각자 독자적인 해법을 들고 나오으로써 차별화를 꾀하고 있다.

빠른 데이터베이스 분석을 위한 OLAP 중심의 인메모리 DB는 일반적으로 트랜잭션 처리를 위한 OLTP 중심의 그것보다 더 많은 메인 메모리가 필요하다. 그래서 가격이 비싼 메모리에 대용량의 데이터를 저장하기에는 비용 측면에서 무리가 있었다.

무엇보다도 기존 데이터 분석은 데이터웨어하우스에 기반한 과거 상황을 파악하기 위한 분석이었기 때문에 굳이 응답 속도가 빠른 인메모리를 통해 지원할 필요성이 없었다.

하지만 DRAM의 가격이 하락하면서 비용 부담이 사라졌을 뿐만 아니라 데이터베이스 분석을 통해 현재 일어나고 있는 상황을 분석, 대응하거나 앞으로 일어날 상황에 대해 예측, 대비하기 위한 용도로 빠른 분석이 각광받으면서 상황이 달라진 것이다. 즉, 빠른 분석을 위한, OLAP을 지원하기 위한 인메모리 기술에 눈을 돌린 것이다.

인메모리 DBMS가 OLAP 분석을 효과적으로 지원하기 위한 기술은 Δ 데이터 저장 아키텍처에서부터 Δ 분석을 위해 접근하는 데이터 양을 줄이기 위한 방법 Δ 기존 분석툴과의 원활한 연동을 위한 MDX(Multi-Dimensional Expression) 지원 방법 Δ 데이터 중복 저장 등에 대한 방법론이 그 핵심이다.

SAP HANA의 경우 데이터 접근 패턴에 대해 행기반의 저장 방식과 열기반의 저장방식을 위한 저장소를 모두 제공할 뿐만 아니라 데이터가 저장된 방법을 바꿀 수 있는 명령어를 제공함으로써 OLTP와 OLAP을 동시에 지원하고 있다.

인메모리 DBMS의 효과

인메모리 DBMS는 디스크 DBMS 대비 수십 배 빠른 처리 성능 구현이 가능

인메모리 컴퓨팅의 이해

인메모리 컴퓨팅(In-Memory Computing, IMC)이란 애플리케이션 운영을 위한 데이터를 하드 디스크가 아닌 메인 메모리에 올려 서비스를 수행하는 컴퓨팅 기술을 의미한다. 인메모리 컴퓨팅 기술에는 메모리 하드웨어, 인메모리 데이터 관리, 인메모리 메시징, 인메모리 애플리케이션 플랫폼 등으로 나눌 수 있다.

가트너는 기업들이 전략적으로 대응해야 하는 10대 기술 트렌드로 2012년부터 2013년 2년 연속으로 인메모리 컴퓨팅을 선정했다. 사실 이보다 더 오래 전부터 가트너는 인메모리 컴퓨팅 기술에 주목해왔다. 2010년 플래시 메모리(Flash Memory), 2011년 스토리지 클래스 메모리(Storage Class Memory) 기술을 선정한 바 있다.

인메모리 컴퓨팅이 부각된 이유는 빅데이터와 깊은

관련이 있다. 데이터 용량이 빠르게 증가하고, 데이터 생성 속도가 빨라지고, 빠른 처리, 분석을 요구하는 추세가 인메모리 컴퓨팅에 대한 관심을 불러일으킨 것이다.

인메모리 컴퓨팅은 역시 대대적인 기업 환경의 변화를 갖고 올 수 있다. 긴 시간이 걸리는 특정 유형의 일괄 처리 작업 실행을 몇 분 또는 심지어 몇 초 수준으로 단축시킬 수 있기 때문에 클라우드 서비스의 형태로 내외부 사용자를 대상으로 하는 실시간 또는 거의 실시간 서비스로서 제공할 수 있다.

또한 수십ms 내에 수백 만 건의 이벤트를 스캐닝해 새로운 기회와 위협을 나타내는 상관 관계와 패턴을 “즉각적으로” 감지할 수 있다. 동일한 데이터 집합에 대

해 트랜잭션(transaction) 및 분석 애플리케이션을 동시에 실행할 수 있기 때문에 경영 혁신을 위한 새로운 가능성의 세계를 열게 될 것이다.

가트너 보고서에 따르면, 중견 기업과 대기업에 빠른 속도로 침투해 2015년까지 이들 기업의 35%가 IMC를 도입할 것이라고 전망했다. 또한 2021년에는 빅데이터 시장에서 인메모리 컴퓨팅 기술을 약 65%가 채택할 것으로 예상했다.

그림 2 | 인메모리 컴퓨팅 기술 분류



가트너, 2012

하다는 장점을 제공한다.

인메모리 DBMS는 네트워크의 발전과 온라인의 대중화와 맞물리면서 빠른 데이터 처리를 필요로 하는 업무들을 중심으로 빠르게 적용되면서, 대량으로 발생하는 데이터의 고속 처리가 필요한 증권업체의 시세 조회 및 체결, 통신사의 실시간 빌링, 제조 분야의 공정 자동화 업무 등의 핵심 솔루션으로 정착되기 시작했다.

선재소프트 김기완 대표는 “네트워크 발전과 인터넷의 대중화가 비즈니스 패러다임을 온라인으로 이동시켰다면, 인메모리 DBMS는 새로운 패러다임 정착의 일등 공신이라 할 수 있다”고 말했다. 김기완 대표는 “객장이 아닌 컴퓨터 상에서 HTS(Home Trading System)로 주식을 거래하고, 유무선 전화의 음성, 메시지, 데이터 사용량과 과금(Billing)도 실시간으로 확인하며, 온라인 수강 신청이나 온라인 복권 구매 등 현재 익숙한 일상들은 모두 인메모리 DBMS 기술 위에 구현된 서비스들”이라고 설명했다.

애드저글러, “데이터 분석, 빠르다. 정신 못차리게 빠르다”

SaaS 광고지원 플랫폼을 운영하는 애드저글러(AdJuggler)에 물어보라. 이 기업의 광고제공 사업은 항상 빠른 속도로 진행되는데, 실시간 입찰이 출현하면서 속도가 한 차원 업그레이드됐다.

실시간 입찰에서 퍼블리셔(Publisher)는 입찰을 요청하는 온라인 거래소에 광고를 전송한다. 사용자가 특정 웹 페이지에 도착하면 광고주들은 입찰을 열고 입찰이 가장 높은 광고가 해당 페이지에 위치하게 된다. 디지털 광고 판매는 빠르게 이뤄진다. 애드저글러의 기술담당 부사장 벤 린드퀴스트에 따르면 실시간 입찰에 참여하는 구매자는 주어진 광고에 대해 입찰할 수 있는 시간이 1만분의 1초 밖에 되지 않는다.

이런 빠른 속도를 위해 애드저글러는 인메모리 DBMS인 테라코타(Terra-cotta)의 빅메모리(BigMemory)를 도입했다.

테라코타의 제품관리 부사장 마이크 앨런은 경험법칙상 메모리가 디스크보다 1,000배나 빠르다고 말했다. 애드저글러의 경우, 입찰 창구가 매우 협소하고 거래량은 많기 때문에 이런 속도 차이가 인메모리를 선택한 결정적 계기가 되었다.

초당 최소 100만 건의 거래 지원

광고를 웹 페이지 상에 초당 2만 건의 거래 클릭이 이뤄지는 애드저글러의 현재 플랫폼에는 MySQL 데이터베이스가 포함되어 있다. 이 데이터베이스는 다양한 웹 사이트의 광고 표시 등 고객의 광고에 대한 환경설정 데이터가 저장돼 있다.

애드저글러의 기술담당 부사장 벤 린드퀴스트는 이 모든 환경설정 데이터가 디스크 기반의 MySQL 데이터 저장소에서 테라코타의 인메모리 DB로 이전할 것이라고 전했다. 또한 애드저글러는 수 테라바이트에 달하는 익명 청중(Audi-

ence) 데이터를 추가할 계획이다.

린드퀴스트는 새로운 플랫폼의 데이터 처리량이 초당 최소 100만 건의 거래를 지원할 수 있을 것이라고 밝혔다. 또한 실시간 의사결정에 관계된 인메모리 DB에 대한 가능성을 확신했다.

예전에는 데이터베이스가 지금 필요로 하는 수준의 성능을 내도록 하기 위해서 상당한 조율이 필요했다. 메모리를 설정하고 RAM에서 데이터 캐시(Cache)를 별도로 취급해 성능을 향상시켜야 했다.

캐시를 사용하는 것이 디스크로 되돌아가는 것보다 빠르긴 하지만 캐시는 일반적으로 데이터베이스 저장된 데이터의 일부만을 나타낼 수 있다. 린드퀴스트는 “성능은 적절한 시기에 적절한 데이터를 확보하는 기술에 달려 있다”고 말했다.

USPS, 인메모리로 부정 사용을 더 확실하게 적발

한편, USPS(United States Postal Service) 또한 부정사용을 적발하고 메일 전송을 향상시키기 위해 인메모리 DBMS를 구축했다. USPS는 SGI(Silicon Graphics International)의 하드웨어와 오라클의 타임스텐 인메모리 데이터베이스 소프트웨어를 사용하고 있다. USPS는 데이터 부하의 병렬화를 향상시키기 위해 단일 이미지보다는 다수의 타임스텐 인스턴스를 사용하고 있다.

USPS의 데이터관리 서비스 관리자 댄 휴스턴 주니어는 “인메모리 DB의 주된 장점은 매우 신속하고 실시간에 가까운 답변을 제공하면서 상당한 양의 데이터를 확인할 수 있다는 점”이라고 밝혔다.

인메모리 기술은 또 USPS가 부정사용을 적발하고 우편물에 적절한 요금 부과됐는지를 신속하게 파악할 수 있게 해준다. 또한 인메모리 데이터베이스를 통해 동적 경로 설정법, 익일 배송, 예측 전송 등의 작업을 활용할 수 있도록 했다. 특히 USPS는 과거 몇 시간 또는 며칠이 소요되었던 일들을 실시간으로 처리할 수 있게 됐다.

인메모리 DBMS가 넘어야 할 문제와 과제

인메모리 DB는 조직들이 신속하게 많은 데이터를 처리해야 할 때, 신속하게 결과물을 생성할 수 있는 장점을 갖고 있다. 하지만 문제점이 아주 없는 것은 아니다. 기술의 특성에 대한 오해가 이런 문제점 가운데 하나다.

업계 전문가들은 인메모리 배치를 위해서는 단순히 데이터를 메모리로 옮기는 작업 이상의 것이 필요하다고 설명했다. 먼저 반드시 메모리를 활용하기 위한 인메모리 DBMS를 설계해야 한다.

SAP HANA 플랫폼 담당 부사장 크리스 헬렌벡은 “단순히 메모리에 데이터를 옮기는 것이 아니다. 처음부터 데이터베이스 전체를 다시 작성해 디스크 대신에 메모리를 주된 저장 매체로 사용하도록 하는 것이다”고 말했다.

또 다른 문제점도 있다. 인메모리 DB의 속도 때문에 프로세서에 대한 부담이

가중된다. 결과적으로, 기업들은 반드시 데이터에 접근하는 코드를 병렬화하고 클러스터 전반에 걸쳐 로드 밸런싱(Load Balancing)을 적용해야 한다. 린드퀴스트는 “로드 밸런싱은 인메모리 데이터베이스를 활용하는 능력에 있어서 중요한 부분을 차지하고 있다”고 말했다.

실제로 애드저글러는 상용 하드웨어와 자체 개발한 소프트웨어를 사용해 풀(Pull) 기반의 로드밸런싱 시스템을 개발했다. 애드저글러의 거래처리 엔진의 각 인스턴스는 로드밸런싱 구성요소로부터 작업을 가져오고 작업을 완료한 후, 다시 더 많은 작업을 가져온다. 시스템은 추가적인 용량이 필요한 경우에 더 많은 인스턴스를 불러온다.

또한 인메모리 DBMS를 사용하는 기업들은 반드시 데이터베이스 색인을 조심해야 한다. 전통적인 데이터베이스를 사용하는 기업들은 상당한 디스크 공간을 색인에 할당할 수 있다. 하지만 인메모리 DB는 더 높은 정확성이 요구된다.

린드퀴스트는 “인메모리 DB를 사용하는 경우에는 성능을 위해 색인을 작성해야 한다. RAM이 더욱 비싸고 한정적이기 때문에 더욱 정확해야 한다”고 말했다.

30여년 전 기술, 본연의 문제 해결이 경쟁력

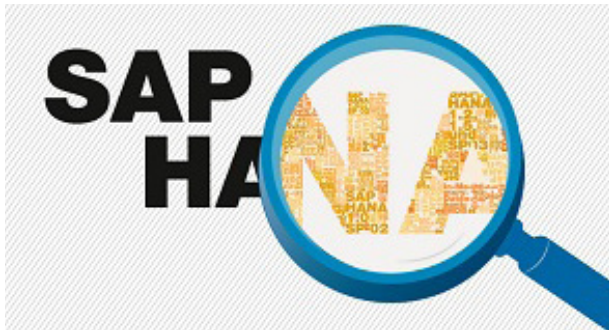
인메모리 도입 시 또 다른 문제점은 RAM의 휘발성에 의해 야기된다. 시스템에 문제가 발생하는 경우, 데이터를 반드시 다시 불러와야 한다. 이 때문에 시간이 소요될 수 있다.

USPS는 인메모리 데이터베이스를 사용하는데 있어 데이터 보호가 가장 큰 문제점 가운데 하나라고 지적했다. USPS는 현재 인메모리로 신속히 작업을 처리하고 나서 관련된 결과를 RDBMS에 다시 전송하고 있다. 또한 인메모리를 통한 처리의 체크포인트 파일을 유지해 문제가 발생하는 경우에 제한적인 복구를 수행하고 있다. 하지만 인메모리 시스템을 복구하는 작업은 쉽지 않다. 디스크 기반의 DB에서 16TB를 복구하기 위해서는 상당한 시간이 소요되기 때문이다.

김기완 대표는 “인메모리 DBMS는 하드웨어 영향을 많이 받는 시스템으로, 하드웨어의 기술적 발전이 긍정적인 요인으로 작용할 수 있다. 최근 호재로 작용할 수 있는 RAM 가격의 급락, 테라급 메모리의 대중화 등이 이어지고 있지만, 인메모리 시장 확산에 청신호로 작용하고 있다는 징후는 어디에서도 찾아볼 수 없다”고 지적했다.

김기완 대표는 “이는 인메모리 DBMS가 하드웨어의 발전 속도를 모두 수용할 수 없는 아키텍처라는 점에 기인한다. 이미 30여 년 전 태동된 기술이다보니, 하드웨어 성능 및 CPU 코어수 증가에도 고객들이 기대하는 수준의 성능을 구현하지 못하고 있는 상태”라고 설명했다.

김기완 대표는 “하드웨어의 발전을 수용할 수 있도록 아키텍처를 개선해 선형적인 성능을 보장하고, 메모리의 휘발성을 커버할 수 있는 데이터 안정성을 보장하는 등 기술적인 개선책을 마련하는데 집중하는 것이 본연의 제품 경쟁력을



살리는 길”이라고 말했다.

SAP, HANA 통해 실시간 기업 제시

2010년 12월, SAP의 인메모리 DB인 HANA의 등장은 인메모리 DB시장뿐만 아니라 전체 DBMS 시장을 뒤흔드는 일대 사건이었다. 2011년 6월 30일, SAP 인메모리 DB인 SAP HANA 어플라이언스 소프트웨어가 전세계 시장에 출시하면서 인

메모리 컴퓨팅 시대가 본격적으로 막이 올랐다. 2010년 11월부터 특정 고객을 대상으로 시범 서비스를 보일 때부터 상당히 주목받았던 HANA는 점차적으로 SAP 애플리케이션과의 결합이 확장, 지원되면서 엄청난 파장을 일으키고 있다. HANA의 강점은 트랜잭션 데이터 처리와 데이터 분석을 실시간으로 구현해 비즈니스 현안에 즉각 대응할 수 있도록 도와준다는 것이다.

SAP는 HANA를 통해 실시간 비즈니스의 가치, 지금껏 잊고 지냈던 RTE(Real-Time Enterprise)에 대한 가치를 다시금 들어올렸다. 30여년간 기업들은 SAP의 기업용 애플리케이션을 사용하는 고객일지라도 DBMS는 오라클이나 IBM, 마이크로소프트 제품을 사용해야 했던 상황이었다. 그런데 SAP는 데이터의 빠른 분석뿐만 아니라 기간제 시스템에서의 빠른 처리를 통해 실시간 기업으로의 길을 제시하면서 DBMS 시장에 돌을 던졌다.

이미 SAP HANA 고객사는 전세계 3,000개에 달하고 있으며 아태 지역에서는 600여 개의 고객사를 확보했다. 국내 시장에서도 SAP HANA를 사용하는 고객이 50여 개사를 넘어서고 있다.

IBM, OLAP용 인메모리 기술 적용

2013년 5월, IBM은 ‘블루 액셀러레이션(BLU Acceleration)’ 기술이 적용된 ‘DB2 10.5’를 출시해 8배에서 최대 25배까지 빨라진 분석 및 리포팅 기능을 제안했다. IBM의 블루 액셀러레이션 기술이란 데이터를 컬럼 기반으로 저장해 인메모리에서 동적으로 처리할 수 있도록 하는, 빅데이터의 분석 질의에 적합한 OLAP용 인메모리 DBMS 기술이다.

EMC, 통합 분석 플랫폼으로 트랜잭션과 데이터 실시간 분석

EMC는 2013년 11월, 피보탈 HD 1.1(Pivotal HD 1.1)과 피보탈 켄파이어 XD(Pivotal Gemfire XD)를 결합한 통합 분석 플랫폼을 출시하고 인메모리 DB 시장에 첫발을 내딛었다.

이 통합 분석 플랫폼은 ‘피보탈 켄파이어 XD’가 ‘피보탈 HD’의 SQL 인터페이스를 제공하는 ‘호크(HAWQ)’ 쿼리 엔진 기술과 결합되면서 OLTP(Online Transaction Processing)와 OLAP(Online Analytical Processing)를 단일 플랫폼에서 동시 지원해 기업 내 발생하는 모든 트랜잭션과 데이터를 실시간 취

합 분석하고 빠른 의사결정을 내릴 수 있도록 지원한다는 것이 특징이다.

마이크로소프트, OLTP용 기술 기본 기능으로 탑재

2014년 4월, 마이크로소프트 또한 5년을 들여 만들어낸 OLTP용 인메모리 기술을 SQL서버2014에 기본 기능으로 탑재했다. 이를 통해 이전 버전 대비 평균 30배 향상된 성능과 100배 빨라진 조회속도, 5배 확장성, 90% 디스크 공간 절감 효과를 제시했다. 이를 통해 마이크로소프트는 기존 SQL 서버 제품을 이용해 왔던 고객들이 추가 비용없이 인메모리 기능을 이용할 수 있도록 해 기존 고객의 이탈을 막겠다는 전략이다.



오라클 마저도 인메모리 DB

2014년 6월, 지난 30년간 DBMS의 강자로 군림해오던 오라클마저도 오라클데이터베이스12c 인메모리 옵션을 출시했다. 오라클의 12c 인메모리 옵션은 인메모리 기술과 함께 디스크 기반의 DBMS를 혼합 사용해 SAP HANA 등 경쟁업체의 공세를 방어한다는 전략이다. 오라클 12c 인메모리 옵션은 OLTP나 데이터 분석용으로 활용할 수 있다.

국산 인메모리 DB 업체 동향

이와 함께 그동안 틈새 시장에서 활약을 보이던 알티베이스, 리얼타임테크, 선재소프트와 같은 국산 인메모리 DBMS 업체들에게도 확대되는 시장 상황에 상당히 발빠르게 움직이고 있다.

특히 인메모리 DBMS가 주류 DB 시장에서 글로벌 업체들과 경쟁하게 되면서, 상당한 구도의 변화가 있을 것으로 예상된다.

이미 400개 이상의 대기업 고객을 확보, 2,000개 이상의 구축 사례를 갖고 있는 알티베이스는 2012년 4월 세계에서 가장 빠른 인메모리 DB라는 알티베이스 HDB 제타(zeta)를 출시하면서 글로벌 기업과의 정면 승부를 시작했다. 알티베이스의 HDB 제품은 OLTP 외에도 OLAP 워크로드 영역을 겨냥한 솔루션으로 최신 버전인 HDB V6 R2는 액티브 클러스터 기능을 강화했으며 대용량 데이터를 신속하게 처리할 수 있는 병렬(Parallel) 기능도 추가됐다.

리얼타임테크는 인메모리 기반의 공간정보 DBMS 제품인 카이로스 스페이셜(KAIROS Spatial)은 다양한 산업 분야의 융복합 서비스 시장을 공략해 나갈 계획이다. 이 제품은 복잡한 공간데이터 처리 속도가 빠를 뿐 아니라 3D 공간 데이터 저장과 처리가 가능해 다양한 3D 기반 서비스에도 적용될 수 있다.

3여년에 걸쳐 개발, 출시한 선DB(SUNDB)를 통해 한국거래소 차세대시스템인 엑스추어플러스, 또 다른 시스템인 시장감시시스템에 적용한 선재소프트는

이 제품이 하드웨어 증가에 따른 선형적인 성능, 효율적인 자원 활용, 개선된 데이터 안정성 등을 모두 충족시킨다고 밝혔다.

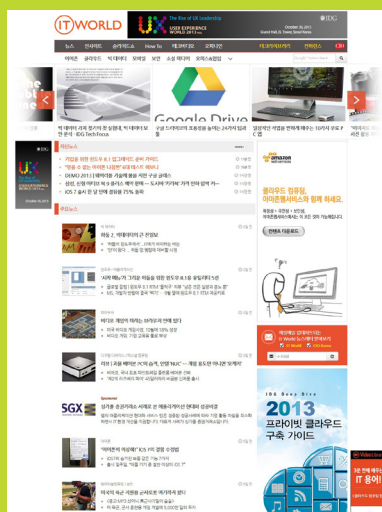
또한 다양한 방식의 백업, 복구 기능을 통해 데이터의 완벽한 정합성 보장하고 있으며, 메모리가 가지는 휘발성으로 인한 데이터 유실에 대한 우려를 최소화한다.

인메모리 DB 시장, 케즘을 극복하라

이렇게 주류 시장의 풍운아로 등장한 인메모리 DB가 시장의 왕좌에 앉기까지는 많은 시간과 노력이 필요할 것으로 보인다. 모든 DBMS 업체들이 인메모리 기술을 채택, 제품을 출시하고, 많은 기업들이 인메모리 DB의 효과를 보고 있지만, 하루아침에 급성장할 것이라고 보는 전문가는 그리 많지 않다. SAP의 HANA가 예상보다 주춤하는 상황에 다른 DBMS 업체들은 자사의 고객 이탈 방지용 제품인 경우가 많아 대대적인 홍보나 마케팅을 하지 않고 있다.

주류 DBMS 시장이 인메모리 DB를 받아들이는 데는 아직 시간이 필요하며, 인메모리 DB 시장 자체는 본격적으로 성장하기 직전의 케즘(Chasm) 기간에 있는 상황이다. 케즘을 극복하기 위해서는 많은 분야에서의 성공 사례와 함께 제품의 안정화, 가격의 합리성 등이 뒷받침돼야 할 것으로 보인다. 

테크놀로지 및 비즈니스 의사 결정을
위한 최적의 미디어 파트너



기업 IT 책임자를 위한
글로벌 IT 트렌드와 깊이 있는 정보

ITWorld의 주 독자층인 기업 IT 책임자들이 원하는 정보는 보다 효과적으로 IT 환경을 구축하고 IT 서비스를 제공하여 기업의 비즈니스 경쟁력을 높일 수 있는 실질적인 정보입니다.

ITWorld는 단편적인 뉴스를 전달하는 데 그치지 않고 업계 전문가들의 분석과 실제 사용자들의 평가를 기반으로 한 깊이 있는 정보를 전달하는 데 주력하고 있습니다. 이를 위해 다양한 설문조사와 사례 분석을 진행하고 있으며, 실무에 활용할 수 있고 자료로서의 가치가 있는 내용과 형식을 지향하고 있습니다.

특히 IDG의 글로벌 네트워크를 통해 확보된 방대한 정보와 전세계 IT 리더들의 경험 및 의견을 통해 글로벌 IT의 표준 패러다임을 제시하고자 합니다.