

IDG Tech Insight

“데이터베이스 주류가 바뀐다”

오픈소스 DBMS 선택 가이드

오늘날 기업이 선택할 수 있는 DBMS는 상용 제품부터 오픈소스까지 매우 다양하다. 수백 종의 제품이 시장에서 치열하게 경쟁하고 있다. 그러나 지난 10여 년간 DBMS 시장의 뚜렷한 변화를 꼽는다면 단연 오픈소스의 부상이다. 최근에는 높은 신뢰성과 성능이 필요한 업무에도 오픈소스 DBMS가 속속 도입되고 있다. DBMS 시장의 변화를 살펴보고 성능과 안정성 등에서 좋은 평가를 받는 대표적인 오픈소스 DBMS 포스트그레SQL의 장단점을 분석한다. 기업이 오픈소스 DBMS를 선택할 때 가장 중요하게 확인해야 할 사항을 알아보고, 유통, 공공, 금융 등 다양한 업종의 도입 사례를 통해 실무자가 바로 참고할 수 있는 실질적인 비교 기준을 제시한다.

- > “느리지만 확실하게” 데이터베이스는 세대교체 중
- > “아는 만큼 더 잘 쓴다” 포스트그레SQL의 특징과 장점
- > 오픈소스 DBMS 선택 시 핵심 체크리스트 6가지



**무단 전재
재배포 금지**

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

“느리지만 확실하게” 데이터베이스는 세대교체 중

Matt Asay | InfoWorld

한번 손에 익은 DBMS(Database Management System)는 여간해서 바꾸기 어렵다. 오라클이 그렇다. 매년 수십억 달러 매출을 올리고 있지만, 프로그래밍 커뮤니티 스택 오버플로우(Stack Overflow)가 개발자 7만 2,517명을 설문한 결과, '가장 꺼리는(most dreaded)' DBMS 중 하나로 오라클이 꼽혔다.

여전히 많이 사용하지만 가장 꺼리는 DBMS라는 모순적 상황이다. 하지만 이를 꼭 나쁘게 볼 필요는 없다. 지난 수년간 레디스(Redis)는 '개발자가 가장 쓰기 좋은 DBMS'로 꼽혔고, 포스트그레SQL과 몽고 DB가 뒤를 이었다. 이는 지난 2017년 순위와 대체로 비슷하다. 굳이 변화를 꼽자면 SQL 서버의 순위가 하락했고 구글의 파이어베이스는 반대로 올랐다.

스택 오버플로우 자료를 보면, 개발자의 DBMS 선호도는 변화가 빠르지 않다는 것을 알 수 있다. 순위가 급변하는 웹 프레임워크와 비교된다. 가트너의 애널리스트 머브 아드리안이 레거시 DBMS를 유지하는 가장 큰 힘을 '관성'이라고 한 것도 같은 맥락이다.

실제로 새로운 DBMS를 도입하는 데는 매우 오랜 시간이 걸린다. 그러나 한때 사랑 받았던 DBMS를 폐기하는 데는 더 오랜 시간이 걸린다. 심지어 해당 DBMS를 운영하던 개발자가 이직해도 기업은 기존 DBMS를 바꾸지 않곤 한다. 하지만 반대로 생각하면 이런 질문도 가능하다. 현재 개발자가 쓰기 좋아하는 DBMS가 앞으로 10년간 점차 기업 전반으로 확산해 자리 잡지 않을까?

느리지만 확실한 변화

오늘날처럼 DBMS가 다양한 시절도 찾기 힘들다. DB-엔진(DB-Engines)에 따르면 현재 373가지 DBMS가 존재한다. 그중 인기를 누리는 것에는 오라클, 마이SQL, 마

DBMS 인기 톱 10 목록 변화

2012년 DBMS 인기 톱 10

- 01 오라클
- 02 마이크로소프트 SQL 서버
- 03 마이SQL
- 04 마이크로소프트 액세스
- 05 IBM DB2
- 06 포스트그레SQL
- 07 몽고 DB
- 08 SQ라이트
- 09 카산드라
- 10 멤캐시드

2021년 DBMS 인기 톱 10

- 01 오라클
- 02 마이SQL
- 03 마이크로소프트 SQL 서버
- 04 포스트그레SQL
- 05 몽고 DB
- 06 레디스
- 07 IBM DB2
- 08 일래스틱서치
- 09 SQ라이트
- 10 마이크로소프트 액세스

이크로소프트 SQL 서버처럼 비교적 역사가 오래된 제품이 있고, 몽고 DB, 레디스처럼 상대적으로 새로운 DBMS도 있다. 포스트그레SQL은 오래전 출시됐지만 최근 들어 인기가 높아진 독특한 경우다.

DBMS의 인기가 높다는 것은 어떤 의미일까? DB-엔진은 2012년부터 검색 횟수, 커뮤니티에서의 언급 정도, 관련 구인 수요, 링크드인 정보 등을 종합해 DBMS 인기 순위를 매겼다. 오라클, SQL 서버, IBM DB2 등이 몽고 DB 같은 신규 DBMS에 비해 인기가 줄긴 했지만, 여전히 광범위하게 쓰이며 높은 인기를 누리고 있다.

2012년과 2021년 DBMS 인기 톱 10 목록을 비교하면 거의 같다는 것을 알 수 있다. 2012년 이후 대략 9년이 지났지만, 아파치 카산드라가 11위로 밀려난 것을 제외하면 목록에 올라온 DBMS의 이름은 별로 변화가 없다. 하지만 개별 DBMS의 순위가 오랜 기간 어떻게 바뀌는지를 보면 변화의 움직임이 감지된다. DBMS의 순위는 10여 년을 주기로 서서히 변화하는 걸로 보이는데, 오라클, DB2, SQL 서버 같은 레거시 DBMS에서 포스트그레SQL, 몽고 DB, 레디스로 전환하는 경향이 뚜렷하다. 실제로 그렇다. 주변의 개발자에게 물어보면 이런 변화를 당장 체감할 수 있다. 지금부터 이 부분을 더 자세히 알아보자.

‘쓰고 싶은’ DB 순위가 곧 미래의 트렌드

스택 오버플로우 설문은 개발자 7만 명 이상을 조사했으므로 오차는 매우 적을 것이다. 쓰기 좋은 DBMS와 사용을 꺼리는 DBMS에 관한 조사 결과가 <그림 1>이다.

<그림 1>을 보면 흥미로운 점을 발견할 수 있다. 일단 개발자가 쓰기 좋아하는 DBMS와 지난 9년간 DB-엔진에서 인기가 많았던 DBMS가 다르지 않다. 이 부분은 어찌면 당연하다. 대신 더 주목해야 할 것은 개발자가 꺼리는 DBMS 순위다. 지난 9년간 DB-엔진의 조사에서 순위가 떨어진 DBMS와 같다는 것을 알 수 있다.

또 하나 살펴봐야 할 수치가 스택 오버플로우가 조사한 ‘아직 써 본 적 없지만 앞으로 쓰고 싶은’ DBMS 순위다. DBMS 시장의 미래를 엿보는 힌트가 될 수 있다. 스택 오버플로우가 이 설문을 시작한 2017년 같은 질문에 대한 답변은 몽고 DB, 포스트그레SQL, 레디스였다. 2021년 순위는 <그림 2>와 같다.

한때 오라클은 많은 사랑을 받는 DBMS였다. 오라클과 IBM, 마이크로소프트 같은 업체는 관계형 DBMS 사업을 크게 확대했다. 많은 기업이 오라클과 협업하고 싶어 하지 않을지언정 오라클의 DBMS를 수십 년 동안 사용했다. 그러나 이제 개발자는 점차 관계형 DBMS의 오픈소스 대안, 즉 포스트그레SQL 등 다른 선택지로 눈을 돌리고 있다. 몽고 DB나 레디스를 사용해 속도와 확장성을 비약적으로 개선하기도 한

그림 1 | 개발자가 쓰기 좋은 DBMS와 사용을 꺼리는 DBMS

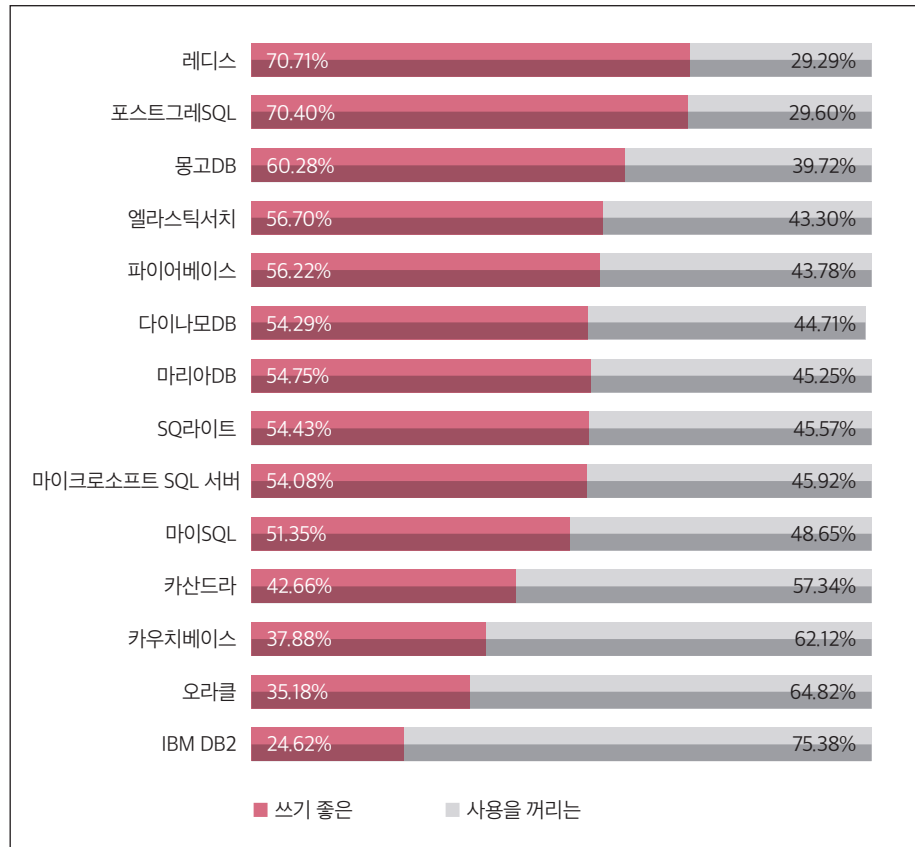
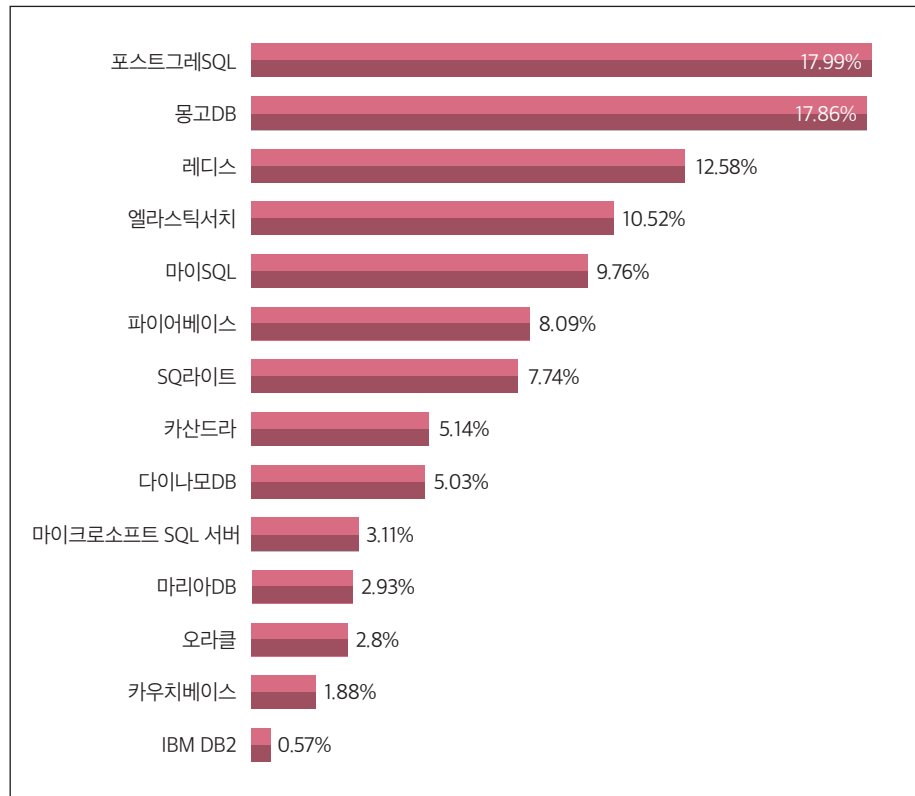


그림 2 | 아직 써 본 적 없지만 앞으로 쓰고 싶은 DBMS



다. 개발자는 기업의 혁신을 모색하는 과정에서 여러 대안을 비교하고 결국 새로운 DBMS에 정착한다.

여기서 말하는 '새로운'이란 '최근'을 의미하는 게 아니다. 포스트그레SQL은 1986년부터 이런저런 형태로 존재했다. 몽고 DB는 2007년부터, 레디스는 2009년부터 있었다. 카산드라와 마찬가지로 이 DBMS는 개발자 사이에서 인기 있고 안전한 선택지로 자리매김하는 데 10년 이상 걸렸다. 그리고 기업의 기본 DBMS로 입지를 굳히는 데도 앞으로 상당한 시간이 걸릴 것이다.

분명 DBMS 순위 변화는 느리다. 기업 내 IT 부서가 기존 워크로드를 새로운 DBMS로 재구성하는 것을 위험하다고 여기는 것이 한 요인이다. 그래서 신규 DBMS는 보통 새로운 영역에서 먼저 채택되는 경향이 있다. 그러나 이런 변화가 느리다고 해도 천천히 그리고 확실하게 나타나는 것은 분명하다. 즉, 현재 가장 쓰고 싶은 DBMS가 기업 내에 느리지만 확고하게 자리매김할 것임을 의미한다. 지금은 가장 인기 있는 DBMS라고 해도 20년쯤 후에는 외면받는 대상이 될 수 있다.

“아는 만큼 더 잘 쓴다” 포스트그레SQL의 특징과 장점

Ivan Panchenko | InfoWorld

DBMS(Database Management System) 시장은 오픈소스와 상용 제품으로 나뉘어 있고 진영마다 다양한 선택지가 있다. 나온 지 30년이 된 포스트그레SQL은 커뮤니티 주도의 오픈소스 프로젝트로, 여전히 높은 인기를 누리며 전 세계 기업 업무 현장에서 사용되고 있다.

예를 들어 러시아 최대 검색 엔진 얀덱스(Yandex)는 포스트그레SQL에 페타바이트 규모의 데이터를 저장하고, 하루 1억 5,000만 건 이상의 이메일을 처리한다. 오래전부터 포스트그레SQL을 사용해 온 깃랩은 초당 18만 1,000건의 트랜잭션을 처리하는 대규모 클러스터를 구축했다. 글로벌 가구 기업 이케아는 총소유비용(TCO) 절감을 위해 포스트그레SQL로 전환해 테라바이트 규모의 데이터를 운영하고 있다.

포스트그레SQL의 장점 8가지

포스트그레SQL은 다양한 애플리케이션에 적합한 여러 특징을 갖고 있다.

- **코드 품질** : 포스트그레SQL의 모든 코드는 여러 전문가의 검토를 거친다. 전체 개발 프로세스가 커뮤니티를 중심으로 진행되므로 버그 보고, 수정, 검증이 매우 신속하게 이뤄진다.
- **확장성** : 포스트그레SQL은 거의 모든 기업 활용 사례를 포괄하는 확장 기능을 갖춘 매우 유연한 솔루션이다. 특정 데이터 유형이나 확장된 로깅 등 세밀한 기능이 필요하다면, 직접 만들거나 서드파티 포스트그레SQL 솔루션 업체를 통해 조달할 수 있다.
- **SQL과 NoSQL** : 포스트그레SQL은 트랜잭션 또는 통계 데이터 행에 대해 전통적인 SQL RDBMS로 사용하는 것은 물론, JSON 문서를 저장하는 NoSQL 솔루션

으로도 쓸 수 있다. 이런 유연성은 비용을 낮추고 보안을 개선하는 데 도움이 된다. 다양한 DBMS 솔루션을 설정, 관리, 보호, 업데이트하는 데 필요한 전문가를 추가로 확보하지 않아도 된다.

- **공간 데이터** : 많은 기업이 공간 데이터를 사용하는 애플리케이션에 PostgreSQL을 쓴다. PostgreSQL에 공간 데이터를 처리하는 유용한 기능이 많기 때문이다. 예를 들어 PostgreSQL은 기하학적 객체를 다루는 데이터 유형을 지원하고, PostGIS(PostGIS)를 이용해 공간 DBMS를 쉽고 빠르게 만든다. PostgreSQL이 운송, 유통 기업에서 인기가 높은 것도 이 때문이다.
- **데이터 가용성과 탄력성** : 공공기관, 금융 서비스 기업, 의료기관과 같은 미션 크리티컬 환경에서는 기술 지원이 포함된 유료 PostgreSQL을 사용할 수 있다. 이 버전은 추가적인 고가용성과 탄력성, 보안 기능을 제공한다.

PostgreSQL은 다른 오픈소스 DBMS와 비교해도 여러 강점이 있다.

- **단일 솔루션 업체에 의해 좌우되지 않음** : PostgreSQL은 특정 업체의 솔루션이 아니다. 보통 단일 업체 솔루션의 운명은 사용자 커뮤니티가 아닌 해당 업체의 결정에 따라 좌우된다. 버그 수정이나 다음 릴리스에서 무엇을 수정할지도 전적으로 해당 업체의 결정에 달려 있다. 사용자 커뮤니티가 업체에 압력을 행사할 수 있고 현명한 업체라면 사용자 커뮤니티의 의견을 경청하겠지만, 사용자의 요구가 무시될 가능성은 항상 열려 있다.
- **업체 종속 없음** : 오픈소스 DBMS가 특정 업체의 소유가 되면 해당 업체가 사용자를 종속시키려 하는 경우가 종종 있다. DBMS 외에 클라우드 인프라와 같은 보안 제품군을 끼워 파는 방식이 대표적이다. 이런 제품은 DBMS를 최대한 활용하는데 도움이 될 수 있지만, 기업이 다른 솔루션으로 전환하기 어렵게 만드는 족쇄가 된다.
- **대중성** : PostgreSQL은 방대한 사용자층을 가진 매우 인기 있는 솔루션이다. 만약 기업이 사용자 기반이 비교적 작은 오픈소스 DBMS로 표준화하면, 서드파티 솔루션 설치, 오류 해결 또는 컨설팅 서비스를 받을 때 어려움을 겪을 수 있다. 더구나 직원을 채용하든 서비스 계약을 맺든 비교적 드문 전문 기술에는 보통 더 높은 가격표가 붙기 마련이다. PostgreSQL은 이런 문제에서 상대적으로 자유롭다.

포스트그레SQL 도입 시 고려해야 할 6가지

포스트그레SQL은 여러 장점을 갖고 있지만, 도입 시 꼭 고려해야 할 사항도 있다.

- **비용** : 포스트그레SQL은 오픈소스 솔루션이므로 100% 무료라고 생각하기 쉽다. 그러나 다운로드와 설치, 사용은 무료지만 마이그레이션, 기술 지원, 유지보수에는 거의 항상 비용이 따르고, 이런 비용을 과소평가해서는 안 된다. 포스트그레SQL 컨설턴트나 전문 업체와 미리 비용에 대해 상의하는 것이 좋다.
- **시간** : 포스트그레SQL 개발 또는 마이그레이션 프로젝트가 예상보다 오랜 시간이 걸릴 수 있다. 예기치 못한 문제에는 부가적인 조사가 필요한 경우가 많고, 이에 따라 전체 일정이 지연되기 때문이다. 따라서 처음부터 프로젝트 일정에 추가적인 연구 조사 기간을 넣어두는 것이 좋다.
- **한 업체에 의해 좌우되지 않음** : 이는 장점이자 단점이기도 하다. 포스트그레SQL로 전환하면 같은 기능을 제공하는 여러 업체와 솔루션 중 선택의 기로에 놓인다. 예를 들어 포스트그레SQL 백업 솔루션은 5개 업체의 제품이 있다. 즉, 기업에 가장 잘 맞는 제품을 선택하려면 시간을 투자해 각 제품의 장단점을 파악해야 한다. 포스트그레SQL을 막 시작하는 기업이라면 컨설턴트의 조언을 받는 것이 좋다.
- **온프레미스 vs. 클라우드** : 포스트그레SQL을 클라우드에서 사용하는 것과 온프레미스에서 쓰는 것은 다르다. 클라우드에서는 많은 포스트그레SQL 확장 기능을 사용할 수 없고, 클라우드 업체가 일부 설정을 막아 포스트그레SQL 구성과 기능이 제한되는 경우도 있다. 클라우드 리소스 사용에 따른 비용도 상당히 커질 수 있다. 실제로 많은 기업이 아마존 RDS(Relational Database Service)를 도입했다가 비용을 줄이고 확장성을 확보하기 위해 온프레미스 포스트그레SQL로 돌아갔다.
- **배포 규모** : 포스트그레SQL로의 전환은 DBMS 크기가 작은 소규모 기업에는 부담이 적지만, 중견 기업과 대기업은 전문 업체의 마이그레이션 서비스 지원이 필요할 가능성이 크다.
- **내부 전문 인력** : 모든 DBMS를 유지보수하는 데는 전문 인력이 필요하고 포스트그레SQL도 마찬가지다. 필요한 인력을 새로 채용하는 방법과 기존 팀이 포스트그레SQL을 배우도록 한 다음 때에 따라 서드파티 업체나 컨설턴트의 도움을 받는 방법을 모두 검토해야 한다. 신규 채용한다면 새 팀이 개발 또는 마이그레이션

을 시작하기 전에 애플리케이션의 비즈니스 요구사항을 완벽히 이해할 수 있도록 프로젝트 초기에 합류시키는 것이 좋다. 그다음 가능한 과제를 예상해 단계별 마이그레이션 계획을 수립하면 된다.

내부 포스트그레SQL팀을 구성한다고 해도, 오랜 기간 전문성을 보유한 포스트그레SQL 컨설턴트나 전문 업체의 도움을 받아 개발 또는 마이그레이션 계획을 검토할 필요가 있다. 여러 업체 솔루션을 평가해 적절한 것을 선택하는 데도 도움이 된다. 이렇게 하면 전체 프로젝트 과정에서 많은 시간과 노력, 비용을 줄일 수 있다.

포스트그레SQL은 인기 있고 잘 지원되는 DBMS이며, 가장 복잡하고 미션 크리티컬한 환경의 요구사항도 충족하도록 계속해서 발전하고 있다. 점점 더 많은 기업이 주 DBMS로 포스트그레SQL을 선택하는 이유다. 적절한 수준의 전문성을 확보해 포스트그레SQL 프로젝트를 수행하면 기업의 미래 요구사항까지 충족할 강력한 데이터 인프라를 구축할 수 있다.

ITWORLD

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 전 세계 IT 리더들의 경험 및 의견을 통해 글로벌 IT의 표준 패러다임을 제시하고자 합니다.

오픈소스 DBMS 선택 시 핵심 체크리스트 6가지

이용우 | 인젠트 데이터사업본부 컨설팅팀 수석

오늘날 기술 혁신에서 가장 주목해야 할 것은 그 내용이 아니다. 신기술이 만들어내는 구체적인 모습도 중요하지만, 오히려 그 변화 속도에 적응하고 점점 더 빨라지는 혁신의 흐름을 읽는 것이 더 중요하다. 실제로 지난 몇 년 동안 빅데이터, 인공지능, 머신러닝 등 새로운 유형의 비즈니스가 잇달아 등장하면서, 생성되는 데이터 유형이나 활용법, 활용 분야 역시 빠르게 세분화됐다. 이런 요구사항에 맞춰 DBMS (Database Management System)도 신기술을 적용하면서 진화했고, 종류 또한 크게 늘었다.

현재 기업이 고를 수 있는 DBMS의 폭은 매우 넓다. 먼저, 데이터 간의 관계에서 출발하는 전통적인 RDBMS, 데이터를 JSON 형식으로 쉽게 저장, 추출하고, 자유롭게 처리할 수 있는 도큐먼트 스토어가 있다. 대량으로 적재된 데이터에서 특정 데이터만 매우 빠르게 골라내는 데 특화된 키값 스토어가 있고, SNS에 적합한 그래프 DBMS, 타임 시리즈 DBMS 등 다양하다. 실제 기업 현장에는 각종 업무가 복잡하게 얽혀 있어 여러 가지 DBMS를 혼용하기도 한다.

이런 DBMS의 발전을 주도하는 것이 바로 오픈소스다. 포스트그레SQL, 몽고DB, 카우치베이스, 레디스, 다이نام오DB, 카산드라, H베이스 등 다양한 유형의 오픈소스 DBMS가 있는데, 상용 DBMS와 달리 새로운 기술을 빠르게 추가하면서 혁신을 주도하고 있다. 그렇다면 이렇게 많은 오픈소스 DBMS 중 우리 기업에 꼭 맞는 것을 어떻게 골라야 할까? 오픈소스 DBMS는 상용 DBMS와 비교해 어떤 장단점이 있을까?

오픈소스 DBMS 선택 기준 6가지

오픈소스 DBMS 선택 시 중요하게 고려해야 할 사항은 크게 6가지다. 안정성, 호환성, 확장성, 편의성, 비용 그리고 기술 지원이다. 하나씩 자세히 알아보자.

DBMS는 기업의 핵심 자산인 데이터를 저장하고 관리하므로, 다른 요인에 영향을 받지 않고 안정적으로 운영되어야 한다. DBMS의 안정성에는 여러 가지가 포함된다. 예기치 않은 시스템 장애나 서비스 중단 등 코어 엔진이 자체적으로 보증해야 할 부분이 있고, 가용성 측면에서 지원해야 할 부분도 있다. 가용성의 경우 싱글 노드가 아닌 다중 노드를 동시에 운영하는 클러스터 구성을 통해 높일 수 있다. 이때 기업은 클러스터 간 데이터 동기화와 정합성 보장 방안을 살펴야 한다. 클러스터를 구성하는 각 노드로 트랜잭션을 적절히 나누는지, 일부 노드에 장애가 발생했을 때 데이터 유실 없이 얼마나 빠르게 복구해 정상적인 서비스를 제공할 수 있는지 등도 검토해야 한다.

둘째는 호환성이다. 기업이 오픈소스 DBMS를 도입하는 이유는 다양하다. 기존 상용 DBMS의 고비용 문제를 해결하기 위해서든, 특정 DBMS의 종속성에서 벗어나거나 빠르게 신기술을 적용하고 클라우드에 대응하기 위해서일 수 있다. 이런 경우 시스템을 처음부터 새로 구축하기도 하지만, 기존에 운영 중인 시스템에서 전환하는 사례가 매우 많다. 필자의 경험으로는 60% 이상이 전환 프로젝트였다. 이때 이기종 DBMS 간에 데이터 이관, 애플리케이션 전환 문제 등에 직면하게 되는데, 이를 해결하려면 호환성 보장이 필수적이다.

구체적으로는 마이그레이션 관점에서 서로 다른 기종 간 다양한 데이터 유형을 수용할 수 있어야 하고 애플리케이션 전환이 원활하게 이뤄져야 하는데, 여기서 중요한 것이 SQL 표준이다. 일부 상용 제품은 해당 제품에서만 동작하는 명령어 또는 함수가 있어서 다른 제품으로 바꾸기 힘든 종속이 발생한다. 이런 경우 신규로 개발해야 하는 범위가 늘어나고 애플리케이션 전환 공수도 증가한다. 따라서 되도록 SQL 표준 구문을 활용해 전환할 수 있어야 하고, 초기 개발 시 SQL 표준으로 작성하는 것이 애플리케이션 간에 호환성을 보장하고 전환하는 데 유리하다. 이를 위해서는 DBMS 자체적으로 SQL 표준 기능과 구문을 풍부하게 지원하는지 확인해야 한다.

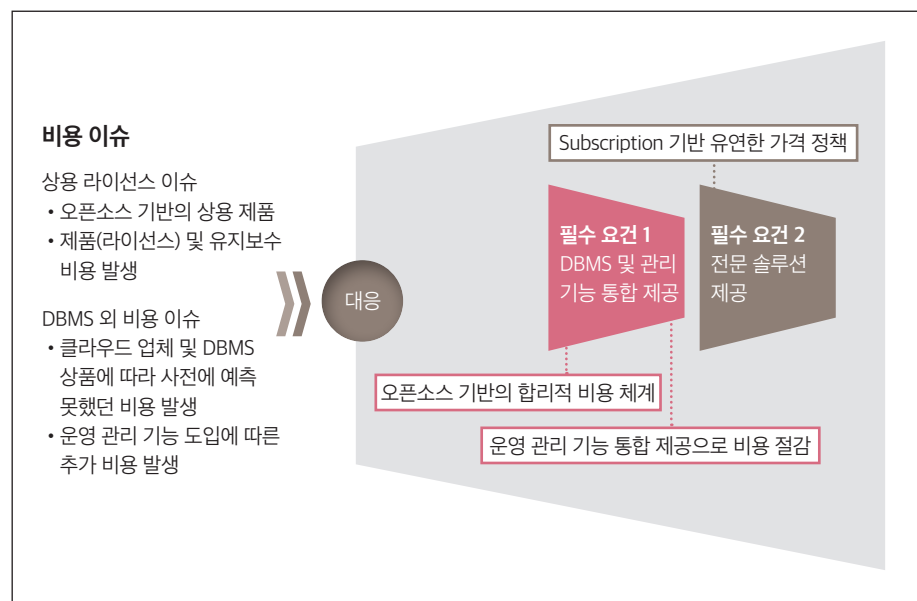
확장성은 시스템 확장과 기능의 확장, 2가지로 검토해야 한다. 최근에는 온프레미스에서 클라우드로 전환이 활발하게 이뤄지고 있다. 클라우드를 통해 확장성의 이점을 누릴 수 있기 때문이다. 운영 시스템을 스케일 아웃 방식으로 확장하는 이유는 여러 가지다. 시스템 가용 자원이 포화상태이거나 확장을 통해 부하를 분산하거나 복제본을 구성해 장애를 대비하는 목적 등이다. 이런 요건에 따라, 얼마나 빠르고 손쉽게 확장할 수 있는지, 데이터 유실을 원천적으로 차단할 수 있는지, 장애 시 신속한 복구가 가능한 유연한 아키텍처를 제공하는지 등을 확인해야 한다.

기능적 확장도 살펴보자. 일반적으로 DBMS 코어 엔진에서 제공하는 기능은 정해져 있다. 기업이 DBMS에서 제공하지 않는 특정 기능을 필요로 할 때 업체에서 이를 개발해 주지 않으면 활용하기 어렵다. 이때 DBMS 코어 엔진에 플러그인 형태로 원하는 기능을 추가해 활용할 수 있다면 상당한 장점이 될 수 있는데, 이것이 바로 기능적 확장이다. DBMS 관련 기능은 계속해서 진화하고 기업의 요구사항도 다양해지고 있다. 기업이 요구하는 기능을 업체가 제공할 계획이 없다면 비즈니스까지 제약받을 수 있다. 따라서 원하는 시기에 새로운 기능을 자유롭게 추가해 활용할 수 있는지 확인해야 한다.

그다음으로 고민해야 할 것이 편의성이다. 보통 DBMS를 도입하면 모니터링, 데이터 암호화, 백업/복구 등 다양한 서드파티 기능이 필요하다. 이런 관리 기능을 개별적으로 도입하는 대신 DBMS 자체적으로 활용, 관리하면 효율성이나 편의성, 총소유비용(TCO)에서 상당한 이점을 누릴 수 있다. 전문화된 관리 기능을 저렴하게, 통합적으로 이용할 수 있는 DBMS를 선택하면 편의성, 효율성, 비용을 줄이는 최선의 선택이 될 수 있다.

비용은 가장 민감한 부분 중 하나다. 불과 몇 년 전만 해도 오픈소스라고 하면 일단 무료라고 생각했지만, 지금은 기술료, 서비스 비용을 합당하게 지불해야 한다는 인식이 확산하고 있다. 오픈소스 소프트웨어는 대부분 연간 구독 서비스를 제공한다. 이를 5년 TCO로 환산하면 상용 제품과 비교해 현저하게 낮다.

DBMS 도입 시 비용 이슈



한 가지 더 고려해야 하는 사항은 클라우드 네이티브 DBMS다. 클라우드 DBMS는 애초에 상용 DBMS는 물론 온프레미스와 비교해 낮은 비용을 강점으로 내세웠는데, 최근에는 이렇게 마케팅하는 업체가 거의 없다. 클라우드 DBMS에 부가 솔루션을 추가하고 스토리지나 네트워크 등 리소스까지 고려하면 오히려 비용이 온프레미스보다 늘어나는 경우도 많기 때문이다. 실제로 클라우드가 너무 비싸서 온프레미스로 회귀하는 기업도 많다. 따라서 온프레미스건 클라우드건 비용 절감이 가능한 상품 구성 또는 적절한 서비스 정책을 제공하는지 살펴봐야 한다.

오픈소스 DBMS를 도입하기 전에 검토해야 할 마지막 요건은 기술 지원이다. 기업이 가장 큰 어려움을 호소하는 부분이기도 하다. 아무래도 생소한 솔루션일 경우 제품에 대한 학습이 선행되어야 하는데, 기존 인력을 활용하는 데 어려움이 있다. 오픈소스 자체만으로는 기술 보증을 하지 않으므로 도입하려는 DBMS에 대해 전문 지식을 가지고 기술 서비스를 제공할 수 있는 파트너가 꼭 필요하다.

특히 오픈소스 DBMS를 개발해 공급하면서 기술 서비스를 직접 제공하는 조직을 갖춘 전문 업체와 솔루션을 선택하는 것이 좋다. 그래야 개별 기업 고객에 대한 이력 관리와 담당자 지정을 통한 책임 관리가 가능하기 때문에 이슈 대응이 빠르고, 안정성 확보나 성능 개선, 튜닝 같은 지원에서도 원하는 수준의 전문 서비스를 받을 수 있다. 서비스하는 제품에 대해 얼마나 전문성을 가졌는지, 전담 지원 조직을 갖추고 있는지 확인해야 한다.

신개념 통합 데이터 플랫폼 '엑스퍼DB'

인젠트는 기업이 오픈소스 DBMS의 장점을 충분히 누릴 수 있도록 앞서 설명한 6가지 선택 기준을 만족하는 솔루션 '엑스퍼DB(eXperDB)'를 공급하고 있다. 인젠트가 포스트그레SQL을 기반으로 개발한 신개념 통합 데이터 플랫폼이다.

엑스퍼DB를 살펴보기 전에 빠뜨릴 수 없는 것이 포스트그레SQL이다. 포스트그레SQL은 30년 이상 개선을 거듭한 오픈소스 DBMS다. 전신인 인그레스(Ingres) 프로젝트까지 포함하면 전체 개발 기간은 더 늘어난다. 오랜 기간 다듬어진 만큼 안정적으로 작동한다. 포스트그레SQL은 가장 진보적인 DBMS로도 꼽힌다. 기능이 매우 풍부하고 사용자의 새로운 요구사항을 적극 수용해 계속해서 진화하고 있다. AWS 오로라(Aurora), 사이베이스(Sybase), SQL 서버, 그린플럼 등이 포스트그레SQL의 영향을 받았다.

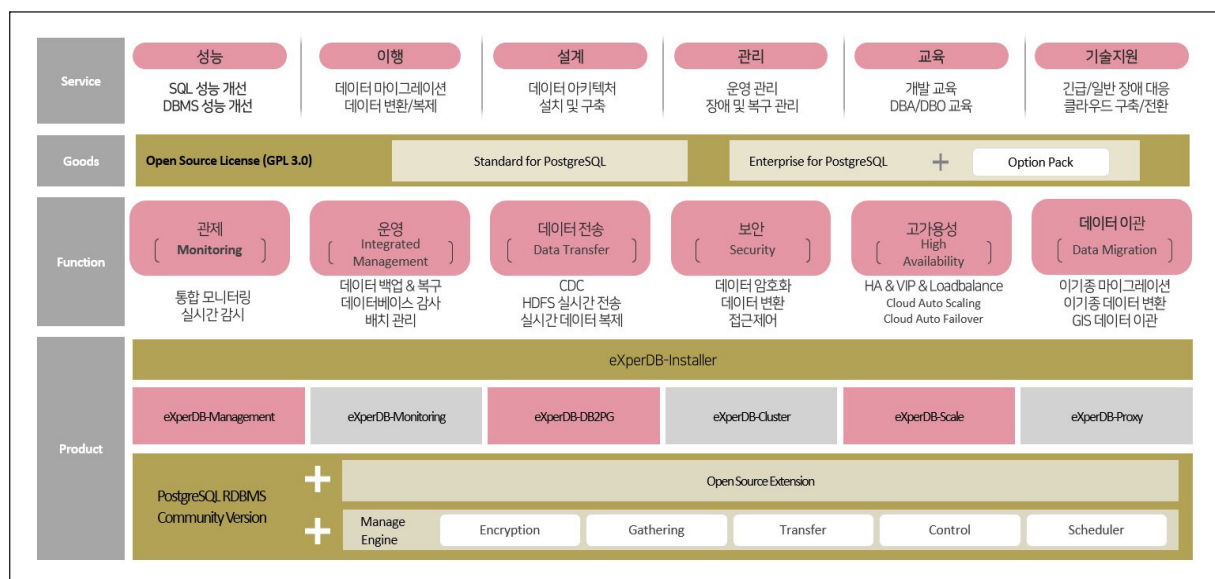
포스트그레SQL이 오랜 기간 꾸준히 발전할 수 있었던 이유 중 하나는 완전한 오픈 소스이기 때문이다. 소스코드의 수정과 배포는 물론 상업적 이용에도 아무런 제약이 없는 라이선스다. 그러나 더 중요한 것은 포스트그레SQL이 매우 활발하게 활동하는 오픈소스 프로젝트라는 사실이다.

오픈소스는 그 특성상 얼마나 많은 기여자가 활동하는지, 커뮤니티의 방향성은 어떤지 등이 매우 중요하다. 기여자가 적으면 프로젝트 자체가 중단돼 더는 코드 개선이 이루어지지 않고, 특정 기업에 인수된 후 공개 저장소가 갑자기 비공개로 바뀌는 경우도 있다. 포스트그레SQL은 이런 위험 없이 지속 가능한 오픈소스 프로젝트로 계속 혁신하고 있다.

포스트그레SQL은 그 역사만큼 장점이 많다. 현존하는 DBMS 중 가장 많은 SQL 기능을 지원해 추가 애플리케이션 개발을 최소화할 수 있고, 최다 표준 구문과 표준 함수를 지원해 다른 DBMS나 애플리케이션과 호환성이 매우 높다. 이 밖에 다양한 데이터 타입을 지원하고 대량 데이터 운영에 최적화해 단순한 OLTP(Online Transaction Processing)뿐만 아니라 복잡한 트랜잭션 처리에도 강점이 있다.

엑스퍼DB는 이런 포스트그레SQL을 코어 DBMS로 소스 수정 없이 그대로 도입해 포스트그레SQL의 장점을 온전히 활용한다. 여기에 순수 오픈소스 DBMS에는 없는 여러 가지 운영, 관리 기능을 추가로 제공하고 다수의 전문 엔지니어를 통해 조직적이고 차별화된 기술 서비스를 지원한다.

엑스퍼DB 플랫폼 아키텍처



구체적으로 보면, 엑스퍼DB에는 포스트그레SQL DBMS 외에 백업/복구, 데이터 암호화, 시스템 모니터링, 통합 관리, 로드 밸런싱, CDC(Change Data Capture), 마이그레이션, 고가용성 구성 등 인젠트가 직접 개발한 기능이 통합돼 있다. 일반적으로 DBMS는 각종 운영 관리 솔루션과 병행 운영하는데, 상용 DBMS는 높은 라이선스비에 서드파티 솔루션 비용을 따로 부담, 관리해야 한다. 반면 엑스퍼DB는 이를 단일 솔루션으로 제공해 기존 비효율을 대부분 해소한다.

엑스퍼DB는 성능과 기능, 클라우드 확장성, 비용 등의 관점에서 기존 국내외 상용 DBMS 대비 장점이 있다. 먼저 성능을 보면, 엑스퍼DB는 코트라, 태광그룹, 국가정보자원관리원 프로젝트 등의 OLAP(Online Analytical Processing), OLTP 성능 비교 개념검증(PoC)에서 오라클 대비 최소 30%, 최대 200% 더 뛰어난 수치를 기록했다.

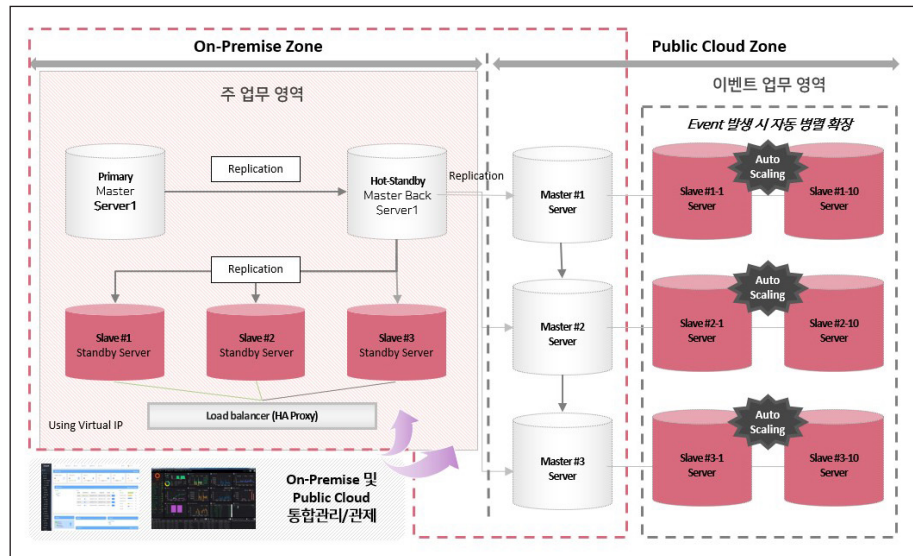
기능에서는 쿼리, 인덱싱 등 표준 SQL을 기준으로, 엑스퍼DB가 90%를 지원해 71%에 그친 오라클을 앞선다. 클라우드 확장성에서도 엑스퍼DB는 무제한 자동 확장을 지원하는 반면 오라클과 티베로 등의 상용 DBMS는 온프레미스에 최적화돼 자동 확장이 어렵다. 마지막으로 비용 측면에서, DBMS와 암호화, 모니터링, 관리 등을 포함한 5년 TCO를 비교했을 때, 엑스퍼DB는 오라클 대비 80% 이상, 티베로와 비교해도 40% 더 저렴하다.

이런 엑스퍼DB의 장점은 다른 오픈소스 DBMS와 비교해도 마찬가지다. 마이SQL, 큐브리드 등 국내외 오픈소스 DBMS와 기본 기능, 플랫폼 기능 등을 비교하면, 엑스퍼DB는 가장 많은 SQL 표준을 지원하고 프로그래밍 언어와 병렬 처리 등에 강점이 있다. DBMS 구축 시 필수적인 모니터링과 암호화, 데이터 변환, 고가용성 등도 다른 오픈소스 DBMS는 지원하지 않거나 별도로 구매해야 하지만 엑스퍼DB는 자체적으로 제공한다.

유통에서 공공, 금융으로 전방위 확산하는 엑스퍼DB

엑스퍼DB는 이런 비교 우위 덕분에 이미 다양한 업종의 여러 기업에 도입돼 현장 실무에 쓰이고 있다. 먼저, SSG닷컴은 이커머스 업계의 대표적인 엑스퍼DB 성공 사례다. 본래 SSG닷컴은 주문, 배송 등 대고객 서비스 주요 업무를 오라클 4노드 RAC(Real Application Cluster)로 구성해 사용했다. 사업이 점차 성장하면서 고객 증가와 서비스 확대에 시스템의 부하가 가중돼 노드를 확장해야 했는데, 기존 상태로는 구조적으로 노드 추가가 어렵고 특정 분석/집계 쿼리를 실행할 수 없는 등 어

SSG닷컴 하이브리드 클라우드 아키텍처



려움이 많았다.

SSG닷컴의 선택은 엑스퍼DB였다. 주요 시스템 전체를 엑스퍼DB 5노드로 전환한 이후에는 마케팅 전략을 원하는 대로 구상해 실행할 수 있었다. 온라인 '쓱데이' 이벤트가 대표적이다. 쓱데이 이벤트의 특성상 특정 시간에 몰리는 대규모 트랜잭션을 감당해야 하는데, 엑스퍼DB의 병렬 자동 확장 기술이 결정적인 역할을 했다. 대량의 트랜잭션이 급격하게 유입될 때 최대 30대까지 노드를 한 번에 확장해 트랜잭션을 분산한다. 실제로 쓱데이 기간 일일 트랜잭션이 800GB에서 1.2TB까지 늘었지만 전혀 문제가 없었다.

신용평가 업체 이크레더블 사례는 엑스퍼DB와 다른 오픈소스 DBMS와의 차별점을 잘 보여준다. 오라클을 사용하던 이 기업은 애초에 엑스퍼DB가 아닌 다른 오픈소스 DBMS로 전환할 계획이었다. 그러나 프로젝트를 진행하며 성능 이슈가 있었고, 전환 후 데이터 정합성이 보장되지 않거나 일부 데이터가 전환되지 않는 문제가 잇달아 불거졌다. 여전히 오라클 기술에 종속되는 문제도 있었다. 결국 새롭게 발견된 문제를 해결하고 모니터링, 백업 등 플랫폼 기능까지 활용하기 위해 엑스퍼DB로 전환했다.

엑스퍼DB가 특히 주목받는 것은 이크레더블 외에 금융권으로 계속 확산하고 있기 때문이다. S 은행 사례가 대표적이다. 이 은행은 최근 비금융권 비즈니스로 사업 영역을 확장했는데, 그 첫발이 배달 플랫폼 사업이다. 새로운 전략 사업을 시작하는 것

이므로, 안정적인 플랫폼을 구축하기 위해 개념증명(PoC), 성능비교(BMT)를 철저히 진행했다.

이 배달 플랫폼 사업의 업무 영역은 크게 결제/정산과 배달 등 2가지다. 결제/정산용 DBMS는 엑스퍼DB로 확정됐고, 배달 앱 업무는 AWS의 RDS를 활용할 예정이었다. 그러나 프로젝트를 수행하던 중에 RDS가 금융기관의 까다로운 보안 요구사항을 충족하지 못하는 것이 확인됐다. 결국 AWS IaaS에 엑스퍼DB를 설치하고 확장 기능을 추가하는 형태로 보안 요구사항을 맞출 수 있었다. 결제/정산 업무는 물론 배달 앱 업무까지 모두 엑스퍼DB로 구축했다.

기존 질서를 뒤흔들 새로운 무기가 필요하다면

최근 들어 성능과 기능, 구축사례에 있어 오픈소스 DBMS의 약진이 두드러진다. 새로운 기술과 신제품, 대규모 구축 사례가 쏟아지고 있다. 하지만 DBMS 장기 트렌드를 보면, 오픈소스로의 전환은 이미 오랜 기간 진행됐고 이제는 거스를 수 없는 대체로 굳어지고 있다고 보는 것이 더 정확하다. DB-엔진이 지난 9년간 오픈소스 DBMS와 상용 DBMS의 인기도 변화 추이를 조사한 결과를 보면, 2013년 65%에 육박했던 상용 DBMS의 인기가 계속 하락해 2021년에는 오픈소스 DBMS에 역전당했다. 뒤바뀐 두 DBMS 간 인기도 격차는 앞으로 더 벌어질 것이다.

이처럼 오픈소스 DBMS의 인기가 꾸준히 상승한 것은 상용 DBMS가 시장의 변화에 대응하지 못하고 안주하는 사이 오픈소스 DBMS가 기업과 사용자가 원하는 새로운 혁신을 빠르게 제품에 반영했기 때문이다. 실제로 급성장하는 기업, 혁신적인 비즈니스 모델 뒤에는 이를 안정적으로 뒷받침하는 오픈소스 DBMS가 있다. 새로운 비즈니스에 꼭 필요하지만 기존 상용 DBMS가 지원하지 못하는 기능과 성능을 오픈소스 DBMS를 통해 구현하고, 이 오픈소스 DBMS를 기반으로 시장의 변화를 이끌고 있다.

더 중요한 것은 이런 혁신의 흐름이 끊기지 않고 계속된다는 사실이다. 엑스퍼DB 역시 2022년에 DML(Data Manipulation Language) 자동 분기, 쿠버네티스 배포 고도화, DBaaS 등의 개발을 진행 중이다. 앞으로 새로운 파괴적 비즈니스로 시장을 혁신하고자 하는 기업에 강력한 무기가 될 것이다. 이는 곧 기존 시장 질서를 뒤집기 위한 새로운 도약이 필요하다면 오픈소스 DBMS가 그 출발점이 될 수 있음을 의미한다. 바로 지금, 아니 앞으로 점점 더 오픈소스 DBMS에 주목해야 하는 이유다.