

kinetica

키네티카(Kinetica)는 GPU 가속 기반 인메모리 데이터베이스 솔루션으로 복잡하고, 양이 많은 스트리밍 데이터 세트를 대상으로 실시간 쿼리를 수행할 수 있습니다. 실시간 처리에 있어 키네티카는 전통적인 하드웨어와 데이터베이스 환경과 비교할 때 1/10 수준의 인프라와 플랫폼 규모로 100배 더 빠른 성능을 끌어낼 수 있습니다. 또한, 키네티카는 완전한 분산 구조 아키텍처로 설계되었으며, 자료 구조 역시 단순하게 가져갑니다. 따라서 업계 표준 하드웨어 장비 상에서 운영할 때 필요한 만큼 정확히 용량과 성능을 예측하여 스케일아웃 확장을 할 수 있습니다. 이 밖에도 키네티카의 UDF(User-defined functions)를 이용하면 분석 플랫폼을 거치지 않고 데이터베이스 상에서 바로 분석을 수행할 수 있습니다. 따라서 GPU 가속을 활용해 AI와 BI 워크로드를 더 빠르게 처리할 수 있습니다. 키네티카는 인터랙티브한 위치 기반 분석을 지원하기 위해 지형 공간(Geospatial) 데이터 처리와 시각화 기능을 탑재하고 있습니다. 키네티카를 사용하는 조직은 고객 경험 개선, 신규 매출 창출, 비용 절감, 업무 효율 개선, 중요 의사 결정 지원을 위한 데이터 분석을 밀리 초 내에 끝낼 수 있습니다. 키네티카는 거의 동시에 데이터 수집, 탐색, 분석, 시각화 작업을 처리합니다.

GPU 가속으로 데이터 처리 병목 현상 해소

GPU는 수천 개의 코어를 내장하고 있어 CPU처럼 명령어를 순차적으로 처리하는 것이 아니라 병렬로 반복해 처리하는 데 잘 맞습니다. 이런 특징으로 GPU는 많은 양의 데이터 세트를 처리하는 머신 러닝, 딥 러닝, NLP, OLAP 같은 컴퓨트 집약적인 워크로드 처리에 이상적입니다.

- NVIDIA의 고성능 GPU는 4,000개의 코어를 내장하고 있습니다. 반면에 전통적인 CPU는 16개에서 최대 32개코어를 지원합니다.
- 전체 데이터 세트를 스캔하거나, 무차별 대입의 경우 병렬 처리가 매우 이상적입니다.
- 고성능 컴퓨팅(HPC) 업계는 대용량 처리라는 도전 과제 해결을 위해 GPU 도입을 확대하는 추세입니다.
- GPU 가속은 범용 하드웨어의 컴퓨트 성능을 극대화합니다.

키네티카는 200개가 넘는 스트리밍 데이터 세트를 통해 테러리스트와 그 밖에 국가 안보를 위협하는 활동을 실시간으로 추적하기 위해 미국 육군과 국가 보안 관련 기관에 의해 구축된 플랫폼입니다.

국가 안보 기관이 처한 상황과 같이 오늘날 많은 기업은 IoT 등 다양한 원천에서 생성되는 데이터가 홍수처럼 넘쳐나는 시기를 맞이하고 있습니다. 이런 이유로 실시간으로 데이터를 스트리밍하고 분석하는데 있어 컴퓨팅 측면의 병목 현상이 발생하고 있습니다.

키네티카는 이미 현장에서 검증된 솔루션이며, NVIDIA와의 긴밀한 파트너십을 통해 하드웨어 기반 가속도 최적화되어 있습니다. 더불어 키네티카는 IBM, HP, Dell, SuperMicro, Cisco 등 유명 업체의 인증을 획득하였습니다.

키네티카를 통한 혁신 사례로 미국 우정국(USPS)의 예가 있습니다. USPS는 우편 수화물의 정시 배달을 위해 실시간 경로 최적화를 합니다. 이를 위해 매우 복잡한 스트리밍 데이터를 수집, 분석, 시각화하는데, 이때 쓰이는 것이 바로 키네티카입니다.

Kinetica의 장점

성능

대규모 스트리밍 데이터를 실시간으로 수집하면서 분석을 동시에 수행하고, 밀리 초 단위로 시각화된 분석 결과까지 제공하는 것, 키네티카는 이를 현실로 만들 수 있습니다. 키네티카의 분산 아키텍처가 제공하는 컴퓨팅 파워는 타의 추종을 불허합니다. 수십억에 달하는 스트리밍 데이터를 분 단위로 수집할 수 있으므로 애플리케이션과 사용자는 가장 최근의 분석 결과를 받아 볼 수 있습니다.

- CPU 기반 인메모리 솔루션과 비교할 때 최대 100배 빠른 쿼리 처리 성능 실현
- 밀리 초 단위로 분석 및 시각화된 결과를 제공
- 인메모리에 100테라바이트 용량의 데이터를 넣어 둘 수 있어 중요 의사결정을 지원하는 분석 작업에 있어 지연을 최소화

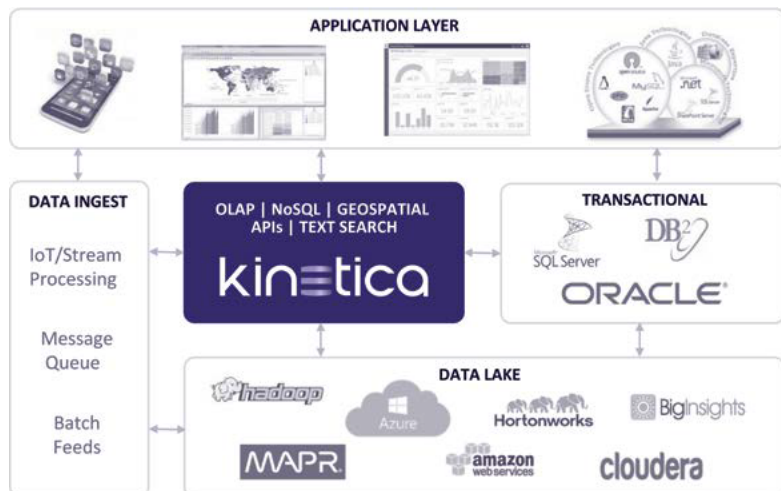
인메모리 데이터베이스 프로세싱을 통한 첨단 분석과 머신 러닝

- UDF(User-defined functions)를 이용해 데이터베이스 내에서 데이터 처리뿐 아니라 컴퓨트 연산까지 수행
- TensorFlow, Caffe, Torch 등의 머신 러닝/AI 라이브러리를 데이터베이스 플랫폼에 직접 올려 실행할 수 있으며, 이를 통해 단일 플랫폼에서 머신 러닝, BI 워크로드를 함께 운영할 수 있음
- TensorFlow를 번들 형태로 제공해 배포가 매우 빠르고 간편함
- Fuzzy Logix가 통합되어 있어 데이터베이스 내에서 직접 분석이 쉬움

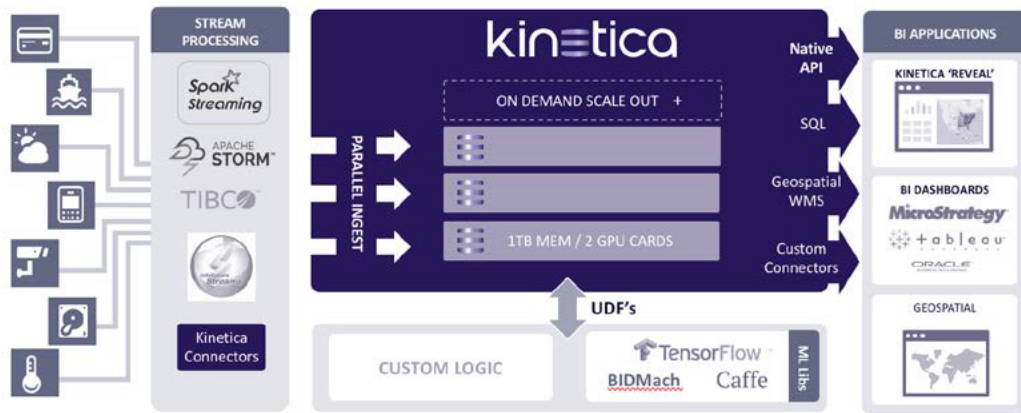
엔터프라이즈 수준의 데이터베이스

키네티카는 단순성, 관리성, 확장성, 가용성, 보안성 측면에서 기업의 요구를 충족하는 엔터프라이즈 수준의 데이터베이스입니다. 키네티카로 기업은 엔터프라이즈 데이터 플랫폼 운영의 복잡성을 낮추는 동시에 비즈니스 가치를 극대화할 수 있습니다.

레퍼런스 아키텍처



고성능의 인데이터베이스 애널리틱스 아키텍처



- **단순성:** 인덱싱과 튜닝 작업이 필요 없습니다. 쿼리를 수행하면 원하는 데이터를 실시간으로 읽어 옵니다. 관계형 데이터베이스와 마찬가지로 현재 사용 중인 데이터베이스, 데이터 레이크, 비즈니스 애플리케이션, BI나 ETL 도구 등과 통합도 간편합니다.
- **관리성:** 노드 추가나 삭제 같은 자원 프로비저닝을 매우 역동적으로 처리할 수 있습니다. 프로세스 관리뿐 아니라 압축, 디렉토리 인코딩 등을 이용한 리소스 최적화까지 첨단화되어 있습니다. 이를 통해 기업은 관리를 간소화하고, 비용을 절감할 수 있습니다.
- **확장성:** 스케일업, 스케일아웃이 모두 쉽습니다. 키네틱카에 쓰기 작업을 한 데이터는 자동으로 클러스터에 분산되어 기록됩니다. 키네틱카는 클러스터 전반에 걸쳐 병렬 프로세싱을 하므로 컴퓨터 자원 집약적인 분석과 머신 러닝 워크로드의 성능을 극대화합니다.
- **고가용성:** 키네틱카는 고가용성 보장을 위해 클러스터 내에서 액티브/액티브 구성을 지원합니다. 또한, 99.999%의 업타임 보장을 위한 데이터 복제 기능도 제공합니다.
- **보안성:** 키네틱카는 LDAP, 액티브 디렉토리를 하며, 이를 통해 역할 기반 접근 제어를 합니다. 또한, 저장이나 전송 중인 데이터를 보호하기 위해 SSL, PLS, AES-256 등의 암호화도 지원합니다.
- **온프레미스와 공용 클라우드에 배포:** 자유로운 배포를 통해 시간을 절약하고 비즈니스 가치를 높입니다.

사용하기 쉬운 API와 간편한 통합

키네틱카는 현재 사용 중인 인프라스트럭처, 비즈니스 애플리케이션, EDW와 하둡 환경에 플러그인하듯이 간편하게 연계할 수 있습니다. 키네틱카는 업계 표준 커넥터를 제공합니다. 이를 이용하면 다음과 같이 기업에서 사용하고 있는 환경에 손쉽게 통합할 수 있습니다.

- **오픈 소스 통합 구성 요소**
Apache NiFi
Apache Spark, Spark Streaming
Apache Storm
Apache Kafka
Apache Hadoop
- **Kinetica의 API를 통한 연계**
REST, Java, Python, C++, Javascript, Node.js
- **Kinetica의 ODBC와 JDBC 드라이버**
업계 표준의 BI, ETL, SQL 도구와 통합
- **SQL-92 쿼리 지원**

완벽한 데이터 시각화와 지형 공간 데이터 처리 기능

키네틱카는 실시간으로 지형 공간(geospatial) 데이터를 처리하고, 이를 시각화할 수 있습니다. 이는 매우 빠르게 움직이는 위치 기반 IoT 데이터 처리에 이상적입니다. 레거시 지형 공간 솔루션과 비교했을 때 키네틱카는 더 빠른 성능을 매우 낮은 비용에 제공합니다.

- Points, Shapes, Tracks, Labels 등 지형 공간 데이터 형식을 지원
- GPU 가속 기반의 분산 렌더링 파이프라인
- 초당 수십억 Points 처리를 토대로 한 시각화
- 각각의 Point를 실시간으로 드릴다운
- 광범위한 지형 공간 분석 지원
- 데이터 변화에 따른 업데이트도 실시간으로 처리
- 키네틱카에서 처리된 지형 공간 정보는 OGC(Open Geospatial Consortium)와 호환되는 WMS(Web Map Services)로 시각화되며, 업계 표준을 따르는 다른 지형 공간 정보 애플리케이션과 호환됩니다.

비용 절감

키네틱카 아키텍처와 GPU가 결합한다는 것이 뜻하는 바는 장치당 4,000개의 코어를 쓸 수 있다는 것입니다. CPU 기반 장치의 경우 보통 8개에서 최대 32개 코어를 쓸 수 있는 것과 비교할 때 매우 차이가 큼니다.

- 1/10 수준으로 하드웨어 비용을 낮출 수 있습니다. 서버 운영을 위한 전력과 냉각 비용은 1/20 수준으로 더 낮습니다.
- 고가의 관계형 데이터베이스로 물리는 작업 부하를 키네틱카로 분산할 수 있습니다. 더 나아가 단일 데이터 플랫폼에 장기 데이터 저장을 위한 하둡 환경까지 통합할 수 있습니다.

Uclick

디지털 트랜스포메이션 사업 본부
이메일: Kinetica@uclick.co.kr
전화: 02-6320-0121

(주)유클릭은 4차 산업 혁명 시대 기업의 트랜스포메이션을 주도하는 전문가 그룹입니다. CA, Oracle, McAfee 등 세계적인 IT 기업들의 최신 기술을 시장에 전달해 왔고, 4차 산업 혁명 시대를 맞아 AI 경제로의 전환을 돕기 위해 NVIDIA Preferred 파트너와 Kinetica 총판으로 국내 주요 금융, 제조, 소매, 공공 고객에게 AI와 첨단 실시간 분석 솔루션을 제공하고 있습니다.

kinetica

Kinetica 및 GPU 가속화 데이터베이스에 대한 더 자세한 정보를 위해서는 kinetica.com를 방문하세요.