

LG유플러스 네트워크 데이터 분석/공유 플랫폼으로의 확장 분석의 편의성과 보안성 모두 확보



LG유플러스 네트워크(NW) 빅데이터 엔지니어링팀은 네트워크 장비로부터 받아들이는 수많은 데이터를 효율적으로 활용하기 위해 빠른 분석 성능과 보안을 지원하는 스노우플레이크를 도입, 네트워크의 디지털전환을 꾀하고 있다.

LG유플러스는 모바일, 홈서비스, 인터넷과 IPTV, 통신서비스 등을 제공하는 통신사다. LG유플러스 NW 빅데이터 엔지니어링팀은 각종 통신 서비스 운영상에서 발생하는 빅데이터를 수집하고 저장해 여러 서비스가 문제 없이 제공될 수 있도록 감시하고 분석하는 시스템을 개발, 운영한다. 데이터로 네트워크 품질을 최적화하고 인공지능/머신러닝(AI/ML) 기반의 데이터 활용이 가능한 기능도 만든다. 네트워크 데이터 분석을 통해 인사이트를 발굴하고, AI/ML을 활용해 서비스 품질을 감시하고 네트워크 품질을 최적화하는 네트워크 디지털전환(DX)을 실현시키는 팀이다.

스노우플레이크 도입 배경

• **Pain Point 1** : NW 빅데이터 엔지니어링팀이 운영하던 기존의 빅데이터 플랫폼은 온프레미스 기반으로 구축돼 있었다. 전국의 네트워크 장비로부터 수십만 개 이상의 파일을 수집해 특정 주기에 따라 수천 개 이상의 테이블에 즉시 로딩해서 데이터를 저장했다. 하지만 이 데이터는 실제 활용 빈도가 높지 않은 것들이었다. 그렇다고 평상시에 데이터를 수집하지 않거나 데이터 보관 기간을 줄일 수도 없는 상태였다. 갑작스러운 장애나 분석 이슈가 발생했을 경우 사용자가 즉시 자유롭게 분석을 해야 하는 상황 탓이다.

반대로, 리소스 경합은 굉장히 많았다. 사용자가 데이터를 직접 조회하도록 오픈해 활용하고 있지만 사용자별 접근 대상을 관리하거나 컨트롤 하지 못하고 있었고 데이터 적재와 여러가지 변환 작업이나 배치 작업 등이 갑작스런 사용자 조회와 겹치게 되면 하드웨어 리소스 경합으로 인해 데이터 누락이 발생하는 경우도 있었다. 낮은 플랫폼 버전 탓에 하드웨어와 소프트웨어에 투자가 필요했지만 용량과 리소스 대비 활용도가 많이 떨어져 투자를 하기 어려웠다.

• **Pain Point 2** : 유무선 통합 실시간 감시용 빅데이터 플랫폼도 개선이 필요했다. 통신사 데이터는 철저한 보안관리가 매우 중요하다. NW 빅데이터 엔지니어링팀에서 관리하는 빅데이터 시스템은 외부와 통신할 수 있는 엔드포인트를 차단한 폐쇄망으로 구성돼 있었고, 실시간 네트워크 서비스 품질 감시라는 주 목적에 보안적으로 영향을 줄 수 있는 모든 기능은 철저하게 차단해 분리 운영 중이었다. 무궁무진한 활용도를 가진 엄청난 양의 데이터를 보유하고 있음에도 보안관리가 더 중요했기 때문에 데이터를 공유할 수 있는 아키텍처로 구성하지 않았다. 다양한 인사이트 분석이 가능한 데이터가 이미 있었음에도 불구하고 이를 활용하고자 하는 부서나 사용자 요구에 응하지 못했다. 플랫폼의 주 목적인 실시간 감시 기능에 영향이 없어야 했기 때문이다.

“네트워크 빅데이터 아키텍처 진화를 위해 스노우플레이크를 선택했고, 이로써 더 많은 사람들에게 더 다양한 워크로드로 쉽게 사용될 수 있는 플랫폼의 확장이 가능해졌다.

- 장진수 LG유플러스 NW 빅데이터 엔지니어링팀장

시스템 상의 페인포인트와 함께 빅데이터 플랫폼을 운영하기 위한 전문 인력 확보의 문제도 있었다. 지속적인 하드웨어와 소프트웨어 투자, 버전 업그레이드 작업으로 인한 운영 문제가 점점 증가했다. 한 단계 더 나아가 다양한 데이터 활용 요구에 대응하기 위한 시스템 최적화, NW 빅데이터 플랫폼의 아키텍처 진화 필요성이 늘어났다. 여기에 빠져서는 안 될 DX 가속화를 위한 AI/ML 활용 등 플랫폼 진화와 사용자 요구에 대응하기 위한 수요가 지속적으로 발생했다. 이에 대응하기 위해서는 빅데이터 플랫폼에 전환이 필요했고, 이러한 배경에 의해 스노우플레이크를 도입하게 됐다.

스노우플레이크 도입시 검토한 점

LG유플러스는 NW 부문에 ▲ 사용자 중심의 데이터 ▲ 빠른 변화에 대응하는 애자일 ▲ AI 기반 자동화 ▲ 통합 및 클라우드화 ▲ 보안 강화 ▲ ESG라는 6개의 키워드를 빠르게 달성할 수 있는 협력관계를 원했다. 그래서 이런 변화를 가속화할 수 있는 최적의 솔루션으로 스노우플레이크를 선택했다. 빅데이터 플랫폼 운영의 관리 편의성과 보안 취약점이 없어야 한다는 것 역시 스노우플레이크를 도입한 선택의 기준이 됐다.

퍼블릭 클라우드 빅데이터 플랫폼을 운영하기 위해서 다양한 전문 기술을 단순화해, 구축과 운영에 편의성을 제공해 주는지를 고려하여 검토해야 한다. 아울러 기존 아마존웹서비스(AWS) 에코시스템들과 기술적으로 어떻게 통합되고 함께 운영 관리할 수 있을지에 대한 고민도 새 솔루션을 검토할 때 LG유플러스가 고려한 부분이다.

가장 중요하게 본 것 중 하나가 개인정보 보안에 취약점이 없는 관리 방안을 제공하는 분석플랫폼이어야 한다는 것이다. 데이터 활용 요구 급증에 대응하기 위해서 데이터 공유가 편리하면서도, 이를 개인화해 통제하고 보안관리할 수 있는 표준화된 관리 룰과 기능이 존재해야 했다. LG유플러스는 이러한 고려사항을 기반으로 SaaS 기반 빅데이터 솔루션을 검토했고, 개념검증(PoC) 단계를 통해 2개 솔루션사의 기능과 성능을 검증했으며, 최종적으로 스노우플레이크 솔루션을 선택하게 됐다.

성과

LG유플러스는 스노우플레이크가 상당히 빠른 속도로 데이터 로딩을 지원해 조직이 애초에 기대한 결과를 가져오는 것을 확인할 수 있었다. 아울러 구축기간 역시 짧다는 것도 강점이었다. PoC 수행을 위한 환경 구축과 마이그레이션을 위한 데이터 준비부터 PoC 종료까지 든 시간은 단 2주일에 불과했다. 장진수 LG유플러스 NW 빅데이터 엔지니어링팀장은 “스노우플레이크가 시스템 구축과 데이터 활용에 얼마나 편리한지를 증명할 수 있는 사례”라고 말했다.

LG유플러스는 스노우플레이크를 최종 선정 후, 대용량 데이터 활용을 위한 기존 온프레미스 플랫폼의 마이그레이션이라는 목적과 데이터 공유 효율성 향상을 위한 새로운 플랫폼 구축이라는 목적에 맞게 퍼블릭 클라우드 기반의 아키텍처를 설계했다. 기존에 보유하고 있던 두 개의 데이터레이크 중 데이터 마이그레이션 대상이었던 저사양 대용량 하둡 시스템을 스노우플레이크로 이관했다.

LG유플러스는 현재 스노우플레이크의 사용량과 시간 단위 요금 부과라는 기본 개념을 활용, 최적의 시스템을 운영하기 위해 실제 필요한 데이터만 저장해 목적에 맞게 관리하고 있다. 이렇게 저장된 데이터는 필요한 개인 또는 팀 단위에게 철저한 보안관리를 통해 쉽고 빠르게 공유할 수 있게 됐다. 데이터와 관련한 요구가 들어올 경우, 간단한 절차만으로 데이터를 빠르게 제공하고 수용할 수 있다.

향후 계획

LG유플러스는 그룹 내 데이터 공유에서 더 나아가 스노우플레이크의 마켓플레이스를 활용해 스노우플레이크의 ML 모델 개발과 모델 옹스 기능을 활용해 더 큰 분석 워크로드 이용 확대를 기대하고 있다.

스노우플레이크에 대하여

스노우플레이크는 데이터 클라우드를 통해 데이터를 가치있게 활용할 수 있도록 합니다. 데이터 사일로 없이 통합된 데이터를 활용할 수 있고, 데이터를 안전하게 공유하며, AI/ML 및 다양한 분석 워크로드를 지원합니다. 2023년 7월 30일 기준, Forbes Global 2000 기업 중 639개 기업을 비롯한 다양한 산업 분야의 8,537개사의 고객이 스노우플레이크 데이터 클라우드를 사용하여 비즈니스를 강화하고 있습니다. 자세한 내용은 www.snowflake.com/ko에서 확인할 수 있습니다.