

Expert Insights

IBM Institute for
Business Value

하이브리드 멀티클라우드로 데이터 패브릭 결합

연결된 데이터로 더 큰 가치 실현

IBM

이 분야의 전문가



Sandipan Sarkar 박사

IBM 석학 엔지니어, Global CTO
Data, Hybrid Cloud Transformation
Service Line, IBM Global Business
Services
sandipan.sarkar@in.ibm.com
linkedin.com/in/sandipansarkar

Sandipan은 대규모 글로벌 기업들이 하이브리드 멀티클라우드를 활용하는 데이터 기반 조직이 되도록 돕고 있습니다. 그는 인공지능, IoT(Internet of Things), 블록체인과 같은 기하급수 기술 분야에서 세계적으로 존경받는 전문가입니다. 그는 여러 학술지와 비즈니스 잡지에 정기적으로 기고하며, 여러 산업 포럼과 학술 기관에서 강연하고 있습니다.



Varun Bijlani

글로벌 경영 파트너, Hybrid Cloud
Transformation Service,
IBM Global Business Services
varun.bijlani@uk.ibm.com
linkedin.com/in/varunbijlani

Varun은 IBM의 글로벌 하이브리드 클라우드 혁신 서비스 비즈니스를 이끌며 고객들이 클라우드 전략과 아키텍처를 설계한 후 마이그레이션, 현대화, 새로운 클라우드 네이티브 기능을 통해 이를 실행하도록 돕고 있습니다.

그는 분야별 지식과 글로벌 프로그램 관리 전문성, 그리고 컨설팅 및 산업 분야에서 전략적 운영 리더십을 함께 발휘하며 26년 넘게 경력을 쌓았습니다. 그는 영국 UWE에서 비즈니스 리더십 PGD를, 인도 Mumbai University에서 전자공학 학사를 취득했습니다.



Richard Warrick

글로벌 연구 책임자,
Cloud Computing
IBM Institute for Business Value
rwarric@us.ibm.com
linkedin.com/in/richardwarrick

Richard는 기업들이 더욱 완전한 디지털 기업으로 진화함에 따라 새로운 업무 수행 방식의 일부로서 클라우드를 바라보는 IBM Institute for Business Value(IBM)의 클라우드 연구 의제를 지휘하고 있습니다. Richard의 컨설팅 경력은 20년이 넘으며 이 기간 동안 그는 글로벌 1000대 기업인 고객들이 비즈니스와 기술의 교차로를 통과할 수 있도록 지원했습니다. 그는 주로 비즈니스 + IT 전략, 운영 모델 설계 및 구현, 대규모 혁신 프로그램과 관련된 업무를 수행했습니다. 그는 George Mason University에서 조직 학습 전공 석사 학위를 취득했습니다.

기여자

Subhendu Dey

이그제큐티브 아키텍트/어소시에이트
파트너, Data & AI 부문 클라우드 자문,
IBM

Sujay Nandi

이그제큐티브 IT 아키텍트, Connected
Solutions & IoT, Cloud Advise, IBM

핵심 사항

분명한 이점

데이터 패브릭 기술은 복잡해 보일 수 있지만 그 이점은 이해하기 쉽습니다. 애플리케이션이 하이브리드 멀티클라우드 환경의 어느 곳에서 실행되든 적합한 애플리케이션에 최적의 데이터를 적시에 활용할 수 있다는 이점 말입니다.

경계가 중요합니다.

데이터 패브릭은 분산된 하이브리드 클라우드 환경에서 지능적 데이터를 활용하는 데 중요한 3가지 경계를 관리합니다. 그 3가지 경계는 데이터 플랫폼이 공유하는 경계, 클라우드와 클라우드 제공업체가 공유하는 경계, 그리고 엔터프라이즈의 트랜잭션 및 분석 데이터 운영 커뮤니티가 공유하는 경계입니다.

데이터의 모든 기능 구현

인공 지능(AI), 머신러닝, IoT(Internet of Things) 및 엣지 컴퓨팅에 투자하고 있는 기업은 데이터 패브릭을 사용하여 트랜잭션 측면과 분석 측면의 데이터 기능을 결합할 수 있습니다.

가치 실현 경로 개척

데이터 패브릭 구현을 통한 비즈니스 가치 실현 여부는 해당 기업이 직면하게 될 비기술적 장애물을 얼마나 효과적으로 제거하느냐에 따라 크게 좌우됩니다.

데이터 패브릭은 데이터 엔지니어에게만 중요한 것이 아닙니다.

기술 기반의 비즈니스 혁신에 관여하는 모든 사람은 데이터 패브릭의 개념을 이해해야 합니다. 대규모 기업들이 끊임없는 경쟁의 압박과 현재의 팬데믹과 같은 예상치 못한 사건에 대응하여 계속 진화하고 있는 가운데 데이터 패브릭은 많은 기업이 조직이 미래에 반복할 수 있는 “유전자 전달”에 도움이 될 것입니다. 현재, 기업들은 확장된 데이터 및 데이터 분석 기능에 계속 많이 투자하고 있으며, 하이브리드 멀티클라우드 아키텍처에도 많은 투자를 하고 있습니다. 소프트웨어는 *지능적 데이터와 하이브리드 클라우드*로 구성된 고단백 식단으로 “세상을 먹여치우고 있습니다”.

IBM Institute for Business Value의 연구에 따르면 데이터에서 더 큰 가치를 실현하면 건전한 디지털 전략을 추진하는 데 도움이 되는 것으로 나타났습니다. 우수 기업들은 데이터 전문가 조직(CoE), 데이터 사이언티스트, 데이터 분석 툴에 엔터프라이즈 규모의 투자를 하고 있습니다.¹ 대부분의 대규모 기업들은 현재 고객용 애플리케이션, 즉 디지털 제품에 고객 경험을 향상하고 고객 여정을 지원하며 새로운 서비스를 가능하게 하는 정보를 활용하기 위한 전략을 갖고 있습니다. 이러한 전략은 B2C뿐만 아니라 B2B 비즈니스 모델에서도 효과를 발휘하고 있으며, 새롭게 떠오르고 있는 플랫폼 비즈니스 모델의 중심적 요소입니다.

기술에 규모를 더하고 사람을 더하면 *복잡해집니다*. 우리는 한 번에 복잡한 시스템의 한 부분에만 집중할 수 있도록 경계를 그려 복잡성을 관리합니다. 이 방법은 어느 정도까지는 효과적이지만, 우리는 그려진 *경계들* 사이에 여백을 두는 일의 중요성을 간과하는 경우가 많습니다. 데이터 패브릭은 애플리케이션, 데이터, 클라우드, 그리고 이들을 설계하고 만든 사람들을 구분하는 경계를 관리할 수 있는 새로운 방법을 제공하므로 데이터 패브릭을 이해하는 것이 중요합니다. 이러한 경계를 어떻게 관리하느냐에 따라 성공이 결정됩니다.

정의: 데이터 패브릭

데이터 패브릭은 특정 기술 하나가 아니라 개념을 의미합니다. 데이터 패브릭을 아키텍처 및 통합된 데이터 서비스 세트를 포함하는 환경이라고 생각하셔도 됩니다. 이러한 서비스들은 함께 조직의 자체 네트워크, 즉 온프레미스와 다수의 클라우드 환경 전반에서 일관적인 데이터 기반 기능을 지원합니다.

“패브릭”이라는 용어는 오해를 불러일으키는 동시에 유용한 정보를 전달할 수 있습니다. 데이터 패브릭은 그 자체로 하나의 실체가 아니지만, 그 구성 요소들은 직물이 짜여진 방식을 닮았다고 생각해도 됩니다. 즉, 전체를 연결하고 함께 붙어 있도록 유지합니다.

이 문서에서는 이러한 경계들 중 3개에 대해 살펴볼 것입니다. 첫 번째는 데이터 플랫폼 사이의 경계입니다. 두 번째는 클라우드와 클라우드 제공업체 사이의 경계입니다. 세 번째는 엔터프라이즈의 *트랜잭션 및 분석* 데이터 운영 커뮤니티 사이의 경계입니다. 실무자들에게는 이 세 번째 유형의 경계가 첫 번째와 두 번째 유형보다 낫설 것입니다. 세 번째 경계는 첫 번째와 두 번째보다 더 중요하지는 않을지라도 최소한 동등하게 중요합니다.

또 하나의 복잡한 기술 그 이상인 데이터 패브릭

디지털 비즈니스 혁신의 많은 요소들과 마찬가지로 데이터 패브릭 기술과 아키텍처는 중요하지만 꽤 복잡합니다. Google에서 “데이터 패브릭 아키텍처”를 검색해 보면 그 복잡성이 이해되실 것입니다(“정의: 데이터 패브릭” 참조). 이러한 복잡성과 하이브리드 클라우드 채택, AI, 엣지 컴퓨팅 및 IoT가 주도하는 빠른 변화가 결합되면 데이터 사이언티스트, 데이터 엔지니어, AI 모델 수립자 등 관련 종사자가 아닌 사람들은 데이터 패브릭을 이해할 수 없는 대상으로 느끼게 될 수 있습니다. 즉, 대부분의 사람들은 데이터 패브릭을 이해하기 어려울 것입니다.

하지만 두려워하지 마십시오. 이 문서를 읽기 위해서는 데이터 패브릭이 항상 부딪히는 문제에 대한 새로운 해결책이라는 점만 이해하면 됩니다. 데이터는 큰 가치를 제공하고 대규모 기업은 엄청난 양의 데이터를 갖고 있지만, 이러한 데이터는 불완전하게 관리됩니다. 데이터는 데이터센터, 데이터 웨어하우스, 데이터 레이크와 같은 크고 작은 용기에 담겨 엔터프라이즈 전체에 분산되어 있습니다. 데이터는 여러 가지 방법을 통해 다양한 형식으로 저장되고 검색됩니다. 데이터는 이리저리 이동하며, 때로는 이러한 이동이 느리고 어렵습니다. 데이터는 관리하기 어렵고 저장하는 데 비용이 많이 들며 항상 신뢰할 수 있는 것은 아닙니다. 그리고 데이터는 제한된 조건에서 매개체를 통해서만 액세스할 수 있습니다.

이제 클라우드 컴퓨팅을 추가해 보십시오. 클라우드 컴퓨팅은 데이터를 클라우드로 이동하여 저장 비용을 낮추지만, 데이터가 더 많은 사일로, 엔드포인트 그리고 다양한 클라우드 서비스 제공업체가 만든 “폐쇄형 공간(walled garden)”에서 증식할 수 있도록 새로운 장소를 제공하여 이러한 혼란을 가중시킵니다. 클라우드는 더 많은 데이터 사일로를 생성할 수 있는 새로운 온-클라우드 애플리케이션의 형태로 혁신을 장려합니다.

데이터가 새로운 석유라면, 대부분의 대규모 기업은 방금 터진 유정입니다. 데이터가 사방에 흘러 넘치지만 현재의 원유와 같은 상태에서는 어떤 데이터도 별로 가치가 없습니다. 데이터 패브릭은 이 문제를 해결하려고 합니다.

데이터가 새로운 석유라면, 대부분의 대규모 기업은 방금 터진 유정입니다. 데이터 패브릭은 이 문제를 해결하려고 합니다.

데이터 패브릭: 데이터베이스 수준을 벗어난 이동

데이터 패브릭의 기본 개념을 이해하려면 데이터베이스, 데이터 레이크, 데이터 웨어하우스 또는 데이터 마트를 고정된 데이터 저장소로 생각하는 일을 멈추어야 합니다. 대신, 적시에 적합한 위치와 적합한 애플리케이션에 최적의 데이터가 제공되도록 데이터가 광범위한 “수도관” 네트워크를 통해 흐르고 있다고 생각해야 합니다. 데이터 패브릭은 사용자의 관점에서 잘 작동하는 시스템입니다.

데이터베이스 수준을 벗어난 이동은 “데이터 중력”에 대처하는 한 방법입니다. 데이터 중력은 한곳에 데이터가 많이 모일수록 이동하기 어려워지는 현상을 말합니다. 데이터 패브릭은 데이터의 이동을 더 용이하게 할 수 있습니다. 또는 이동할 수 없는 데이터, 예를 들면 온프레미스에 머물러야 하는 데이터에 대한 액세스를 더 용이하게 할 수 있습니다.

데이터 패브릭과 인드라의 그물

수천 년간 전해져온 동양의 가르침에서 등장하는 인드라의 그물 (Indra's Net)은 데이터 패브릭에 대한 적절한 은유를 제공합니다.² 이 가르침에 따르면 인드라는 신들의 왕이고 인드라의 그물은 격자 모양을 구성하는 줄이 교차하는 부분마다 귀중한 보석이 박혀 있는 거

대한 우주 같은 망입니다. 이 그물은 무한한 수의 보석을 꿰고 있으며, 모든 보석의 각 면은 다른 모든 보석의 각 면을 반사합니다. 모든 것은 다른 것과 연결되어 있어 원인과 결과의 영향이 이 무한한 그물 전체에 물결처럼 퍼져나갑니다(그림 1 “인드라의 그물: 데이터 패브릭에 대한 은유” 참조).

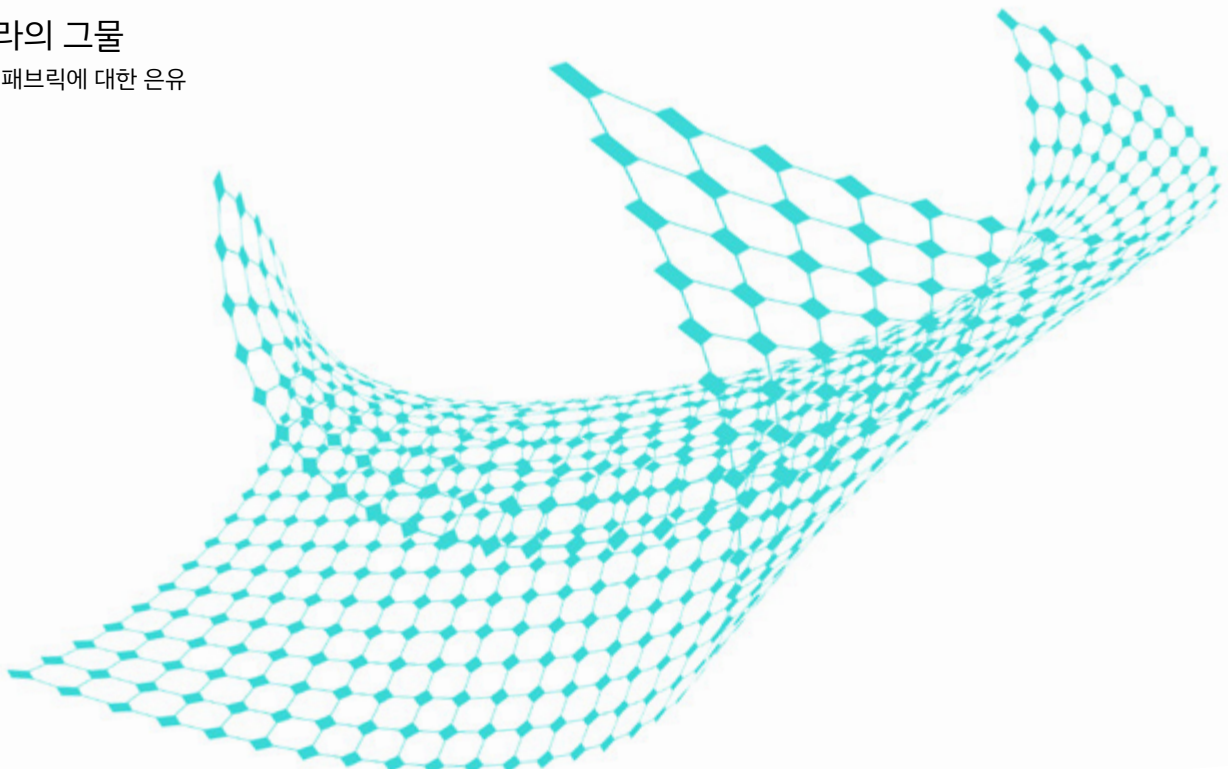
이 은유에서 인드라의 그물에 있는 보석은 엔터프라이즈가 보유한 수천 개의 애플리케이션과 고객, 비즈니스 파트너, 업계 생태계 참여자 및 기타 관계자가 운영하는 애플리케이션을 각각 나타냅니다. 각 보석은 다른 모든 보석과 연결되어 있습니다. 각 보석은 별개이지만 그물, 즉 *패브릭*으로 연결되어 있습니다. 이 패브릭은 컴퓨팅 성능, 스토리지, 연결성, 데이터 등 각 보석이 필요로 하는 모든 것을 제공합니다. 각 보석이 그물의 어느 위치에 있든 그 보석은 가장 가까이에 위치한 최적의 *데이터* 소스로부터 필요한 데이터를 받을 수 있습니다.

데이터 패브릭과 데이터 패브릭이 해결하는 문제를 알아보았으니, 이제 기업들이 데이터 저장소 전반에서 직면하게 되는 세 가지 경계와 인드라의 그물과 유사한 접근법이 어떻게 도움이 되는지를 살펴보겠습니다.

그림 1

인드라의 그물

데이터 패브릭에 대한 은유



고객을 통합적 시각에서 파악하는 데 필요한 데이터를 단일 데이터 플랫폼에서 찾기가 어렵습니다.

데이터 플랫폼 간의 경계

데이터 실무자들이 흔히 하는 말은 “두 번째 데이터베이스를 추가하면 통합 문제가 발생합니다”입니다. 빅데이터가 IT 분야의 으뜸가는 화젯거리가 된 후 몇 년간 데이터의 활용도를 높여 수익을 낼 수 있는 기회가 크게 증가했지만, 규모에 맞게 적합한 애플리케이션에 최적의 데이터를 제공하는 과제 역시 같은 속도로 증가했습니다. 클라우드 이전에도 기업들은 다양한 데이터베이스에 위치한 데이터를 통합하는 기술 솔루션을 의미하는 *데이터 플랫폼*을 구축하고 있었습니다. 데이터 플랫폼은 서비스형으로 작동하도록 설계되었습니다. 보호벽 안에서, 데이터가 필요한 사람들은 데이터에 액세스하거나 데이터를 사용자, 애플리케이션 또는 기타 기술에 제공할 수 있습니다.

각 데이터 플랫폼의 경계는 일반적으로 저장되는 데이터의 유형 또는 데이터가 사용되는 방식에 의해 결정됩니다. 대규모 기업의 경우 HR 데이터 플랫폼 또는 공급망 데이터 플랫폼 또는 특정 사업부를 위한 고객 데이터 플랫폼이 필요할 수 있습니다.

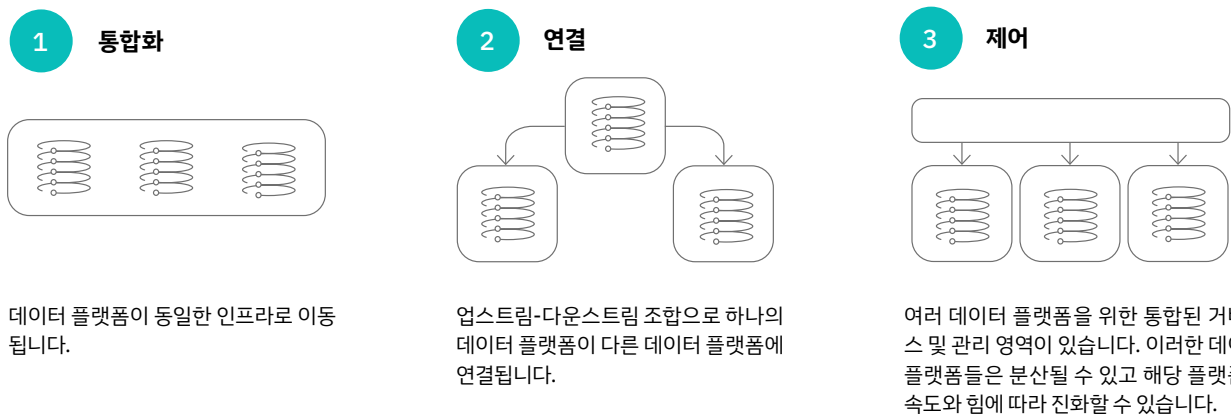
이러한 경계를 두는 것이 합리적이었으며 이러한 경계는 가용 데이터에서 더 큰 가치를 얻기 위한 편리한 방법을 제공했습니다. 그러나, 요즘에는 여러 가지 변화가 발생하여 상황이 달라졌습니다. 첫째, 기업들이 새로운 비즈니스 모델을 활용하고 평생 지속되는 고객과의 상호작용을 단일 시각에서 파악하게 되면서 사업부별로 고객 데이터를 관리하여 발생하는 사일로를 변호하기가 어려워졌습니다. 둘째, 공급망이 점점 더 디지털화되면서 고객 가치 사슬을 통합적 시각에서 파악하고 관리하는 데 필요한 데이터를 단일 데이터 플랫폼에서 찾기가 어려워졌습니다. 셋째, 데이터 분석 기능이 향상되어 데이터 플랫폼 설계자가 생각할 수 없었던 연결 관계와 인사이트를 여러 데이터 플랫폼 사이에서 찾아낼 수 있는 가능성이 생겼습니다.

데이터 플랫폼 간의 경계를 관리하는 방법은 3가지입니다. 바로 통합, 연결, 제어입니다(그림 2 “경계 전략: 데이터 플랫폼 경계를 관리하기 위한 3가지 방법” 참조). 이 방법들 중 어느 것이든 데이터 패브릭의 일부가 될 수 있지만, 분산된 하이브리드 클라우드 환경에 가장 적합한 방법은 세 번째 옵션입니다. 그러나, 각 방법 모두에게 중요한 것은 플랫폼의 데이터를 해당 데이터를 필요로 하는 애플리케이션과 느슨하게 결합하는 것입니다. 이러한 느슨한 결합이라는 개념이 데이터 패브릭 아키텍처 전반에서 중요합니다(“교차 판매 고객을 위한 가치 제안” 참조).

그림 2

경계 전략

3가지 데이터 플랫폼 경계 관리 방법



클라우드 및 클라우드 제공업체 간의 경계

대규모 기업들은 일반적으로 최대 수십 개의 클라우드와 다양한 클라우드 서비스 제공업체를 사용하는 하이브리드 클라우드 아키텍처를 채택했습니다. 이러한 이질성은 장점이 될 수 있지만, 데이터 공유를 반드시 용이하게 하는 것은 아닙니다. 별개의 클라우드와 클라우드 제공업체를 사용하면 위에서 설명한 데이터 플랫폼의 수직적 사일로가 또 다른 형태로 발생할 수 있습니다.

일부 아키텍처 설계 관련 결정이 그렇듯이 데이터 패브릭은 이러한 경계를 관리하는 데 핵심적인 역할을 수행합니다. 두 경우 모두에서 중요한 것은 함께 엮어야 할 여러 클라우드 위에 수평적 관리 계층을 적용하는 것입니다(“데이터 수신 및 송신 문제 해결” 참조). 하이브리드 멀티클라우드 환경에서 데이터 패브릭으로 이러한 경계를 관리하는 경우, 하나의 클라우드 제공업체가 “우두머리”가 되어야 합니다. 이 우두머리 제공업체는 개방형 아키텍처를 사용해야 합니다. 또한 데이터 패브릭을 적용하면 관리 대상인 클라우드 중 광고된 것만큼 개방적이지 않았던 클라우드가 드러날 수 있습니다.

트랜잭션 및 분석 데이터 운영 커뮤니티 간의 경계

우선, 몇 가지 기본적인 사항을 짚고 넘어가겠습니다. 대규모 기업은 많은 *트랜잭션 데이터*를 생성합니다. 트랜잭션 데이터는 고객과의 상호작용 동안 생성되거나 교환되는 데이터입니다. 현재, 트랜잭션 데이터를 효과적으로 사용하고 관리해야 한다는 과제는 사람들이 잘 이해하고 있지만, 아직 완전히 해결되지 않았습니다. 예를 들면, 대부분의 기업들이 다수의 채널, 애플리케이션, 제품 및 서비스에서 생성되는 트랜잭션 데이터를 사용하고 교환하기 때문에 고객을 하나의 시각에서 파악하는 데 어려움을 겪고 있습니다.

적극적으로 관리하지 않으면 트랜잭션 데이터는 엔터프라이즈 전반에서 사일로를 유발하여 비즈니스 성과를 향상하는 데 중요한 데이터 공유를 방해하는 경향이 있습니다. 적극적 관리가 얼마나 중요할까요? Amazon의 Jeff Bezos는 트랜잭션 데이터를 자신만 알 수 있도록 보관하는 사람은 누구든지 해고할 것이라고 경고한 것으로 유명합니다.³

두 번째 데이터 유형인 *분석 데이터*는 다릅니다. 이 데이터 유형은 숨겨진 패턴을 드러내고 예측과 의사 결정을 수행하는 데 사용되는 다양한 대량의 데이터 세트의 구성입니다. 빅데이터 물결은 이러한 분석 데이터의 활용을 장려했습니다. 현재, 분석 데이터는 AI, 머신러닝, IoT, 엣지 컴퓨팅, 5G와 같은 소위 말하는 기하급수 기술의 토대입니다.

분석 데이터의 두 가지 특징이 데이터 패브릭에 중요합니다. 첫째, 분석 데이터가 사용되는 많은 경우, 대량의 데이터를 고속으로 수집해야 합니다. 둘째, 분석 데이터에는 데이터 셀에 입력할 수 없는 *비정형* 데이터가 포함될 수 있습니다. 영상, 이미지, 오디오, 텍스트, 소셜 미

인사이트: 교차 판매 고객을 위한 가치 제안

데이터 플랫폼 간의 경계를 역동적으로 조정하기 위한 “제어” 패턴은 별개의 고객 오퍼링을 중심으로 데이터 플랫폼을 발전시킨 조직들에서 흔히 볼 수 있습니다. 예를 들면, 모기지 상품을 위한 데이터 플랫폼과 신용 카드 고객을 위한 데이터 플랫폼을 보유한 대규모 은행이 있을 수 있습니다. 이 은행이 모기지를 신용 카드 고객에게 교차 판매하거나 신용 카드를 모기지 고객에게 교차 판매할 경우 비즈니스 타당성이 우수합니다. 이는 대규모 기업이 이미 보유한 데이터에서 수익을 창출할 수 있는 전형적인 사례입니다.

이 예에서는 플랫폼들을 통합할 실질적 필요성이 없습니다. 그리고 사업부들은 이러한 통합에 동의하지 않을 것입니다. 플랫폼들을 단순히 연결하기만 하면 장기적으로 확장성이 크게 향상되지 않습니다. 제어 패턴을 따르면 각 플랫폼을 계속 해당 사업부가 담당하되, 이 대규모 기업이 결합된 플랫폼들이 발전되는 방식을 관리하고 자금을 지원하는 데 일정한 역할을 수행하도록 할 수 있습니다.

인사이트: 데이터 수신 및 송신 문제 해결

클라우드 간 경계는 매우 가시적인 방식으로 그 모습을 드러냅니다. 즉, 클라우드 서비스 제공업체 비용이 높아집니다. 클라우드 서비스 제공업체는 고객의 데이터가 클라우드에 들어오면 그곳에 머문다는 것을 전제로 비용을 책정합니다. 클라우드에 데이터를 추가하면 비용이 올라갈 수 있습니다. 그러나, 제공업체의 클라우드에서 데이터를 빼내어 다른 클라우드에 이동해도 비용이 많이 발생할 수 있습니다.

이전에 살펴본 교차 판매 은행의 예를 사용하여, 이 은행이 모기지 애플리케이션을 클라우드 서비스 제공업체 A에 호스팅하고 신용 카드 애플리케이션을 클라우드 서비스 제공업체 B에 호스팅했다고 가정해 봅시다. 일상적인 운영 환경에서 실시간과 같은 속도로 한 클라우드에서 다른 클라우드로 직접 데이터를 이동하는 일은 어렵고 비용이 많이 들 것입니다. 데이터 패브릭을 사용할 경우, 데이터는 클라우드 A에서 클라우드 B뿐만 아니라 모든 클라우드의 애플리케이션이 요청할 경우 액세스 가능한 데이터 서비스로 기능하는 데이터 플랫폼으로 이동할 것입니다.

데이터 패브릭은 API를 활용하여 트랜잭션 애플리케이션 또는 비즈니스 워크플로우 안에 직접 이식된 서비스형인 사이트(Insights-as-a-Service)를 제공할 수 있습니다.

디어, 센서 데이터는 모두 데이터이지만 스프레드시트에 입력할 수 있는 데이터는 아닙니다.

현대적인 분석 기술은 실시간 또는 거의 실시간 데이터를 기반으로 실행해야 합니다. 이러한 속도를 달성하려면 분석 기술은 분석을 위해 일괄 처리와 비용이 많이 드는 데이터의 이동을 요구하는 기존의 데이터 웨어하우스나 데이터 레이크에 의존해서는 안 됩니다. 인사이트를 얻게 될 즈음 이미 그러한 인사이트는 구식이 되어버릴 수 있습니다.

데이터 패브릭은 데이터가 이동 중일 때에도 실행되는 이벤트 스트리밍, 애플리케이션 통합, 실시간 분석과 같은 기능을 제공하여 이러한 과제를 해결할 수 있습니다. 또한, 데이터 패브릭은 분산된 데이터 소스로 분석을 수행할 수 있도록 데이터 가상화 기능을 제공합니다. 데이터 소스가 여러 클라우드에 분산되어 데이터를 클라우드 및 클라우드 제공업체 안팎으로 이동하는 일이 많은 비용을 초래할 경우 이 기능은 특히 중요합니다.

데이터 패브릭은 API를 활용하여 트랜잭션 애플리케이션 또는 비즈니스 워크플로우 안에 직접 이식된 서비스형 인사이트(Insights-as-a-Service)를 제공할 수 있습니다. 이는 데이터 사용자가 활용 가능

한 인사이트를 제공하기 위해 데이터 대시보드에 대한 비용이 많이 드는 수동 해석을 기다릴 필요가 없음을 의미합니다. 경우에 따라, 인간이 개입하지 않고도 이러한 인사이트를 토대로 시스템 내에서 바로 자동화된 조치가 시작될 수도 있습니다.

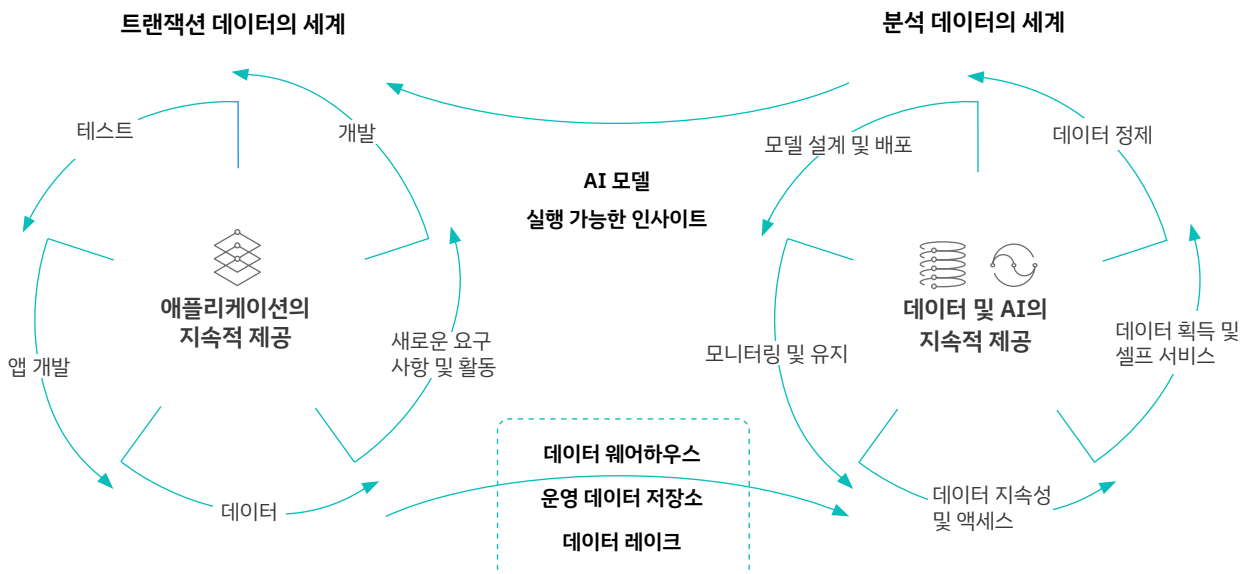
커뮤니티 경계 관리를 통해 운영 모델과 관련된 기회 창출

트랜잭션 데이터의 세계와 분석 데이터의 세계를 합치면 엔터프라이즈의 운영 모델에 중대한 영향을 줄 수 있습니다.⁴ 데이터 패브릭은 트랜잭션 데이터의 세계와 분석 데이터의 세계를 기술적으로 더 가깝게 만들어 애플리케이션 개발 팀과 데이터 및 분석 팀이 더 가까워질 수 있는 새로운 기회를 열어줍니다. 디지털 제품 개발 및 관리 주기는 이제 데이터로부터 인사이트를 생성하고 AI를 위한 모델을 적용하는 활동의 주기와 결합되어 가치를 추가하는 새로운 워크플로우를 형성할 수 있습니다. 이 두 커뮤니티의 팀들은 애플리케이션 사용자와 고객에게 풍부한 데이터 기반의 가치를 제공한다는 공통된 목표에 따라 일할 수 있습니다(그림 3 “득이 되는 연대: 트랜잭션 데이터 및 분석 데이터 커뮤니티 간의 경계 관리” 참조).

그림 3

득이 되는 연대

트랜잭션 및 분석 데이터 커뮤니티 간의 경계 관리



애플리케이션 제공 및 데이터 분석 커뮤니티 사이에 고객-공급자 관계 형성

데이터 패브릭이 하이브리드 멀티클라우드 IT 환경 전반의 데이터를 연결하고 이러한 데이터가 디지털 제품에 의해 소비되도록 할 때 마법과 같은 일이 일어납니다. 데이터 패브릭은 또한 디지털 제품이 생성한 데이터가 데이터 CoE에 의해 소비 가능하도록 합니다. 데이터 패브릭은 중앙집중화된 아키텍처가 필요하지 않고, 분산된 이질적 하이브리드 멀티클라우드 환경과 잘 작동하기 때문에, 데이터 패브릭이 없다면 차단되었을 혁신을 가능하게 합니다. 데이터 패브릭이 대규모 기업에 일반적으로 존재하는 기술적, 조직적 장벽을 깨부수면 좋은 일이 일어날 수 있습니다.

이와 같은 장점은 모든 플랫폼에서 데이터에 대한 통합된 거버넌스와 분석을 가능하게 할 수 있으므로 더 많은 종류의 사일로로 해소할 수 있습니다(“공급자 및 고객의 관계 형성” 참조). 통합된 거버넌스는 공통된 데이터 거버넌스 프로세스를 통해 다양한 분석 이니셔티브가 함께 묶일 수 있도록 지원하면서 사업부가 개별 플랫폼으로 앞으로 나갈 수 있도록 지원합니다. 예를 들어, 데이터 패브릭을 사용할 경우, 엔터프라이즈 수준의 데이터 카탈로그와 모델 레지스트리가 엔터프라이즈의 신뢰할 수 있는 데이터와 분석 자산의 가시성과 가용성을 향상할 수 있습니다.

데이터 패브릭은 이해하기 어려운 기술일 수 있지만 그 유용성은 분명합니다. 데이터 패브릭은 개별적 데이터를 엔터프라이즈와 엔터프라이즈가 사용하는 여러 클라우드 환경 전반에서 연결할 수 있습니다. 가장 중요한 이점은 아마도 데이터 패브릭이 엔터프라이즈의 애플리케이션 제공 팀과 데이터 분석 CoE를 연결해 줄 수 있다는 점일 것입니다. 데이터 패브릭, 하이브리드 클라우드, 애플리케이션 개발, 데이터 분석이 함께 활용될 수 있는 영역을 열심히 찾아볼 것을 귀사에게 권장합니다. 이러한 영역은 잠재적 비즈니스 가치가 가득한 영역입니다.

인사이트: 공급자 및 고객의 관계 형성

은행의 예를 세 번째로 사용해 보겠습니다. 이 은행의 디지털 모기지 상품과 디지털 신용 카드 상품은 모두 트랜잭션 데이터를 생성합니다. 고객 집단 간의 교차 판매를 뒷받침하는 비즈니스 타당성은 모기지 애플리케이션이 생성한 데이터의 전부는 아닐지라도 일부가 신용 카드 비즈니스에 유용할 것이라는 생각, 또는 그 반대의 경우도 유용할 것이라는 생각을 기반으로 합니다. 이 두 가지 유형의 애플리케이션을 사용하는 상품 담당자는 나머지 애플리케이션의 데이터에 대한 분석을 수행해야 합니다. 예측 분석 또는 워크플로우 자동화와 같은 고차원적인 기능을 일부 활용해야 할 수도 있습니다. 이를 위해 이들에게는 데이터 분석 커뮤니티가 필요합니다.

반대로, 데이터 분석 커뮤니티에는 고급 데이터 모델을 테스트 및 구축하고 머신 러닝 및 AI 상품을 훈련시킬 때 활용할 수 있는 대량의 데이터가 안정적으로 공급되어야 합니다. 이러한 관점에서 보면 이 두 커뮤니티는 서로에게 자연히 공급자이자 고객이 됩니다. 그러나 현실에서는 이 두 커뮤니티는 거의 상호작용을 하지 않습니다. 그 결과 두 커뮤니티는 좋은 성과를 거두지 못합니다.

트랜잭션 커뮤니티와 분석 커뮤니티 사이의 장벽을 부수는 일은 운영 모델과 관련된 문제입니다. 이 문제를 해결하려면 사고의 전환이 필요합니다. “트랜잭션 애플리케이션 실행”과 “데이터 과학 활용”을 별도의 영역으로 간주하는 대신 이 두 가지 활동 모두 고객에게 가치를 제공하는 데 필요한 동일한 수평적 업무 흐름에 기여한다고 생각해야 합니다.

실행 지침

비즈니스 가치 실현 경로 개척

클라우드 채택, 데이터 패브릭, 지능적 데이터에 관한 논의는 점점 더 상호 관련성이 커지고 있으며 이러한 관련성은 커져야 합니다. 그러나 *관리 방법과 사고 방식의 변화*에 대해 이와 유사하게 통합된 대화를 하지 않는 경우가 많습니다. 이러한 변화는 단지 데이터 패브릭 아키텍처 기술을 구현하는 데 머물지 않고 비즈니스 가치를 실현하는 방식으로 이러한 기술을 구현하는 데 필요합니다.

기술 구현을 넘어 가치를 실현하는 3가지 방법은 다음과 같습니다.

1. 데이터 패브릭 이니셔티브를 프로젝트가 아닌 제품으로 간주하고 자금 지원

비즈니스 기술 이니셔티브는 착수 후 사전 설정된 요구 사항을 충족하고 책임을 유지 보수 및 지속 업무를 담당하는 IT의 부분에 넘기면서 사라지는 프로젝트로 자금을 지원받는 경향이 있습니다. 데이터 패브릭과 같은 이니셔티브에는 이러한 접근 방식이 효과가 없을 것입니다.

데이터 패브릭 이니셔티브는 제품으로서 자금을 지원하고 관리하는 것이 훨씬 더 낫습니다. 즉, 제품 관리 라이프사이클에 따라 자금 지원이 결정되어야 합니다. 기술을 구현하고 비즈니스 성과가 실현되기 전에 승리를 선언하는 대신, 이러한 제품 접근 방식은 제대로 작동하는 데이터 패브릭 버전을 내놓는 일은 시작에 불과한 것으로 여깁니다.⁵ 수년의 기간에 걸쳐 자금 지원과 인력 배정이 지속적으로 이루어져야 할 것입니다.

대부분의 대규모 기업에서는 디지털 이전, 그리고 클라우드 이전의 자금 지원 프로세스와 정책이 여전히 시행되고 있기 때문에 이러한 접근 방식을 취하는 일은 어려울 수 있습니다. 이러한 프로세스와 정책은 회복력이 뛰어나며 잘 보호됩니다. 기존의 자금 지원 체계를 직접 현대화하려고 하기보다 현재의 정책을 일회성으로 면제할 것을 요청하는 것이 더 나은 조직도 있습니다.

2. 민첩한 구현 원칙 활용

수치로 데이터 패브릭을 규정하는 키트 같은 것은 없습니다. 각 기업은 자사에 효과적인 데이터 패브릭을 구현하는 데 필요한 것을 알아내야 합니다. 데이터 패브릭을 위한 업무가 클라우드 채택 업무 또는 확장된 AI 구현 업무와 동시에 진행될 경우 특히 그러합니다. 시스템의 한 부분을 변경하면 다른 부분이 변경될 것이며 이러한 변화는 예상할 수 없는 방식으로 일어나는 경우가 많습니다.

작게 구축하고, 작게 테스트하고, 작게 배우는 등 민첩한 원칙을 활용하면 데이터 패브릭을 구축할 때 피할 수 없는 복잡성에 대처하는 데 도움이 됩니다. 협력을 통해 더욱 완성도 높고 확장된 버전의 데이터 패브릭을 연속적으로 구축, 통합, 운영하는 여러 분야를 아우르는 소규모의 팀을 활용할 때 최고의 결과를 얻을 수 있습니다. 클라우드 이니셔티브와 AI 이니셔티브가 동시에 진행되는 경우 팀의 노력을 통합하는 데 도움을 주는 이니셔티브 간 제품 지원 사무소를 만들어 팀을 집단적으로 지원하는 것이 가장 좋습니다.

많은 대규모 기업은 상명하달식 명령 및 통제 관리 원칙에 익숙해져 있으므로 민첩한 구현 원칙을 활용하는 일은 어려울 수 있습니다. 이니셔티브를 시작할 때는 모두가 효과적인 데이터 패브릭을 구현하는 방법을 잘 모르기 때문에 고정된 계획, 일정, 중간 목표를 수립하는 것은 좋은 생각이 아닙니다. 최적의 인재, 최적의 기술 그리고 레퍼런스 아키텍처와 같은 유용한 아티팩트가 있는 경우에도, 이러한 이니셔티브가 해당 비즈니스 환경의 해당 조직의 사람들에 의해 전에는 추진된 적이 없다는 사실은 변하지 않습니다.

3. 서비스 방식으로 데이터 패브릭 실행

이상적으로는, 데이터 패브릭은 사용자, 예를 들면 개발 팀이 *어떻게 무엇*이 수행되고 있는지 알지 못하는 상태에서 백그라운드에서 실행되어야 합니다. 개발 팀은 풀타임 데이터 패브릭 전문가를 고용할 필요가 없어야 합니다. 데이터 패브릭은 사용자가 고객의 요청에 대응하는 과정에서 때로는 데이터 패브릭을 소비한다는 사실을 인식하지 못한 채 소비하는 서비스로 제공되는 것이 가장 좋습니다. 서비스 기반 비즈니스 기술 모델에서 데이터 패브릭은 운영 모델의 다른 곳에 있는 다양한 “고객”이 사용할 수 있는 지속적인 기반 서비스 세트로서 완벽한 기능을 수행합니다.

데이터 패브릭을 고객이 소비하는 서비스로 제공할 경우 얻을 수 있는 이점 한 가지는 데이터 패브릭의 범위와 규모를 자동으로 적절하게 유지할 수 있다는 점입니다. 수요가 증가하면 데이터 패브릭 서비스도 확장되고 고객 수요가 줄어들거나 변경되면 데이터 패브릭 서비스도 축소되거나 변경됩니다.

일반적인 IT 조직의 체계는 수직적 기능 사일로로 기반으로 하고 서비스 기반 운영 모델은 궁극적으로 고객 중심적인 방식으로 디지털 제품을 제공할 수 있도록 더욱 수평적인 구조로의 변화를 요구하므로 서비스 방식으로 데이터 패브릭을 실행하는 일은 어려울 수 있습니다.

Expert Insights 소개

Expert Insights는 뉴스로서 가치가 있는 비즈니스 및 관련 기술 주제에 대한 사고 리더의 의견을 전달합니다. 이 문서의 내용은 전 세계의 선도적인 분야별 전문가와의 대화를 기반으로 합니다. 자세한 내용은 IBM Institute for Business Value(iibv@us.ibm.com)에 문의해 주십시오.

참고사항 및 출처

- 1 “From Data Science to Data Diplomacy.” (데이터 과학에서 데이터 외교로) IBM, 2019, www.ibm.com/thought-leadership/institute-business-value/c-suite-study/cio?mhsrc=ibmsearch_a&mhq=data+diplomacy. https://www.ibm.com/thought-leadership/institute-business-value/c-suite-study/cio?mhsrc=ibmsearch_a&mhq=data%20diplomacy, Giordano, Tony. “A Blueprint for Data in a Multicloud World.”(멀티클라우드 세상에서 데이터를 위한 블루프린트) IBM, 2019, www.ibm.com/thought-leadership/institute-business-value/report/multicloud-data-strategy
- 2 O'Brien, Barbara. “Indra’s Net: Buddhist Metaphor for the Interpenetration of All Things.”(인드라의 그물: 모든 것의 상호침투성에 대한 불교적 은유) Learn Religions, www.learnreligions.com/indras-jewel-net-449827
- 3 Settle, Mark. “CIOs Can Weaponize Their API Architectures Through Better Governance.”(CIO는 더 나은 거버넌스를 통해 API 아키텍처를 무기로 활용할 수 있음) Forbes, Forbes Magazine, 2020년 6월 1일, www.forbes.com/sites/marksettle/2020/05/04/cios-weaponize-apis-through-better-governance/?sh=27972aee6076
- 4 “How to Move Beyond a Monolithic Data Lake to a Distributed Data Mesh.”(모놀리식 데이터 레이크에서 분산된 데이터 메시로 이동하는 방법) Martinfowler.com, martinfowler.com/articles/data-monolith-to-mesh.html, “Data Mesh Principles and Logical Architecture.”(데이터 메시 원칙 및 논리적 아키텍처) Martinfowler.com, martinfowler.com/articles/data-mesh-principles.html
- 5 Kersten, Mik. “Project to Product: How Value Stream Networks Will Transform It and Business.”(프로젝트에서 제품으로: 가치 창출원 네트워크가 IT와 비즈니스를 혁신하는 방식) IT Revolution PR, 2018. <https://itrevolution.com/book/project-to-product>

© Copyright IBM Corporation 2021

(07326) 서울특별시 영등포구 국제금융로 10
서울국제금융센터(3IFC)
Tel. 02-3781-5114
Produced in the United States of America
2021년 4월

IBM, IBM 로고, ibm.com은 전세계 여러 국가에 등록된 International Business Machines Corp.의 상표입니다. 기타 제품 및 서비스 이름은 IBM 또는 타사의 상표일 수 있습니다. 현재 IBM 상표 목록은 웹 “저작권 및 상표 정보”(ibm.com/legal/copytrade.shtml)에 있습니다.

이 문서는 최초 발행일을 기준으로 하며, 통지 없이 언제든지 변경될 수 있습니다. IBM이 영업하는 모든 국가에서 모든 오퍼링이 제공되는 것은 아닙니다.

이 문서의 정보는 상품성, 특정 목적에의 적합성에 대한 보증 및 타인의 권리 침해에 대한 보증이나 조건을 포함하여(단, 이에 한하지 않음) 명시적이든 묵시적이든 일체의 보증 없이 “현상태대로” 제공됩니다. IBM 제품은 제품이 제공되는 계약의 조건에 따라 보증됩니다.

이 보고서는 일반적인 안내 목적으로 작성되었으며, 상세한 연구 또는 전문가적 판단을 대체할 수 없습니다. IBM은 이 문서에 의존하여 조직 또는 개인이 겪은 어떠한 손실에 대해서도 책임을 지지 않습니다.

이 보고서에서 사용된 데이터는 제3자 출처를 통해 확보되었을 수 있으며, IBM은 이러한 데이터를 독립적으로 검증, 확인 또는 감사하지 않습니다. 이러한 데이터를 사용한 결과는 “현상태대로” 제공되며 IBM은 명시적이든 묵시적이든 어떠한 진술이나 보증도 하지 않습니다.

