

솔루션 브리프

정밀 의학을 위한 유전체 분석 가속화

퓨어스토리지로 구현하는 차세대 시퀀싱

게놈 데이터는 현대 의료에서 질병을 예방하고 환자에게 맞춤형 치료 계획을 구축하는 데 중요한 도구입니다. 차세대 시퀀싱(next-generation sequencing; NGS) 등 새로운 분석 기술의 출현으로, 정밀 의학 연구원들은 방대한 양의 게놈 데이터를 확보하고 짧은 시간 내에 치료법을 정의할 수 있게 되었습니다. 그러나 이는 IT 인프라가 지속적으로 유지되는 경우에만 가능한 일입니다. 데이터 증가에 보조를 맞추고 동시 시퀀싱 워크플로우를 지원하기 위해, 제약 및 생명공학 연구 기관들은 IT 인프라를 업그레이드하여 속도, 규모 및 민첩성을 비약적으로 향상시키고 있습니다.

현대적인 시퀀싱과 분석의 핵심

게놈 시퀀싱 워크플로우는 염기쌍 시퀀싱 리드를 생성하는 것으로 시작됩니다. 단일 시퀀서는 하루에 20~48개의 시퀀스 리드를 생성하며, 이는 테라바이트 규모의 데이터를 생성합니다. 대부분의 생명과학 조직들은 다양한 시퀀싱 장비를 사용하기 때문에 리드는 페타바이트 수준으로 빠르게 축적되며, 이러한 리드를 전체 게놈으로 조립해 생물학적 데이터 마이닝, 시각화 및 해석 작업을 수행합니다.

시퀀스를 조립, 마이닝 및 해석(2차 및 3차 분석)하려면, 로우 레이턴시 IOPS와 높은 처리량을 제공하는 스토리지 아키텍처에 구축된 고성능 컴퓨팅 환경이 필요합니다. 특히, 2차 분석에는 메타데이터에 대한 액세스와 동시성 요구사항을 지원하는 고성능 스토리지가 필요합니다. 시퀀스 리드 수가 증가함에 따라, 시퀀싱 실행과 과학적 분석이 중단되지 않도록 무중단 용량 확장이 필요합니다.

퓨어스토리지는 이러한 요구사항을 잘 이해하고 있으며, 비교할 수 없는 간단함과 안정성으로 속도, 확장성 및 신속한 구매(구매, 임대 또는 서비스형 제공)라는 고유한 요구사항을 충족하는 다양한 솔루션을 제공하고 있습니다.



분석 가속화

최대 수억 개의 파일에서
엔드 투 엔드 시퀀싱 성능을 개선하고
데이터 레이턴시를 낮춥니다.
게놈 시퀀싱 인덱스 로드 소요되는
시간을 3분의 1로 단축합니다.



다운타임 및 중단 최소화

스냅샷, 무중단 확장, 예측적 지원 및
랜섬웨어 보호를 통해 다운타임과
중단을 최소화합니다.



효율성 향상

유연한 소비 모델로 필요한 용량에
대해서만 비용을 지불하고 필요에
따라 신속하게 확장할 수 있어 예산을
효율적으로 관리할 수 있습니다.

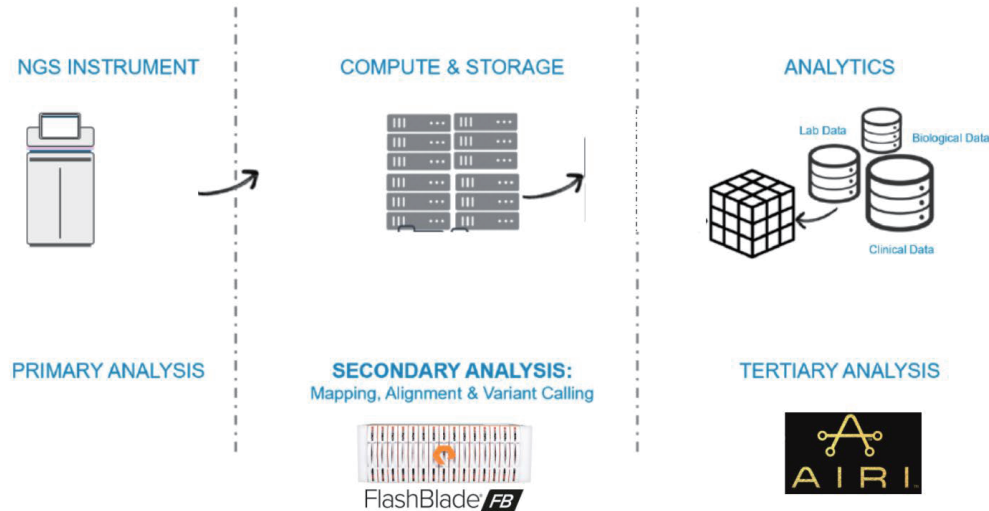


그림 1. 차세대 시퀀싱 워크플로우

유전체학을 위한 현대적인 데이터 경험(Modern Data Experience™)

퓨어스토리지는 현대적인 데이터 경험이 연구원들의 역량을 강화하는 가장 좋은 방법이라고 생각합니다. 그래서 고급 올플래시 솔루션 포트폴리오에 구현된 4가지 개념을 기반으로 유전체학을 위한 현대적인 데이터 경험을 구축했습니다.

- 빠른 속도:** 새로운 생체의학적 발견을 위한 게놈 시퀀싱 및 분석 워크플로우에는 레이턴시가 낮은 고성능 스토리지가 필요합니다. 퓨어스토리지는 이러한 요구를 염두에 두고 플래시블레이드(FlashBlade[™]) 어레이를 설계했습니다. 성능과 용량을 독립적으로 확장할 수 있는 맞춤형 엔지니어링 플래시 아키텍처를 기반으로 하는 플래시블레이드는 페타바이트급 데이터 처리를 지원하며 수억 개의 파일에 로우 레이턴시 액세스를 제공합니다. 게놈 정렬과 변이 검출이 포함된 2차 및 3차 분석 프로세스의 경우, 플래시블레이드 아키텍처는 프로세스를 3~4배 이상 가속할 수 있습니다.
- 간단함이 곧 스마트함:** 동시성이 증가하고 DNA 및 RNA 시퀀싱 데이터 양이 기하급수적으로 증가하면, 쉽게 확장할 수 있는 아키텍처를 보유하는 것이 중요합니다. 조직은 연구원의 액세스를 지연시키거나 인적 오류로 인해 다운타임이 야기되는 복잡한 설정 프로세스를 제거해야 합니다. 이를 돕기 위해 플래시블레이드는 내부 네트워크 설정을 자동화하고 로드 밸런싱을 동적으로 관리하여 지속적으로 성능을 향상시킵니다. 퓨어스토리지는 지원을 간소화하기 위해 AI 기반의 선제적 지원은 물론 스냅샷, 재해 복구, 랜섬웨어 보호 등 업계 최고의 데이터 보호 기능을 제공합니다. 이러한 기능들이 결합되어, 중단이 발생하기 전에 문제를 식별하고 심각한 영향을 주는 중단으로부터 비즈니스를 보호해줍니다.
- 모든 곳에 존재하는 클라우드:** 생명과학에 사용되는 애플리케이션과 프로세스는 끊임없이 진화하며, IT 인프라도 더불어 진화해야 합니다. 이를 위해, 퓨어스토리지는 클라우드 같은 경제성으로 컨테이너화된 워크로드를 지원하는 솔루션을 설계했습니다. 서비스형 퓨어(Pure-as-a-service)는 예산을 최적화하고 온디맨드 방식으로 확장할 수 있는 진정한 소비 기반 모델을 통해 퓨어스토리지 제품을 사용할 수 있도록 지원합니다. MSP(매니지드 서비스 공급업체)와 코로케이션 제공업체를 통한 컨테이너 및 오케스트레이션, 하이브리드 클라우드 구성을 지원하기 때문에, IT 담당자들은 워크플로우 설계와 리소스 프로비저닝, 애플리케이션의 위치를 자유롭게 변경할 수 있는 민첩성을 확보할 수 있습니다.

솔루션 브리프

혁신을 위한 구축: 퓨어스토리지는 고객 관계 중심의 접근 방식을 통해 지속적으로 업계를 선도해왔습니다. 퓨어스토리지에서 제품을 구매하는 순간 구축되는 고객 관계는 컨트롤러 업그레이드, 어레이 소프트웨어, 플래시 스토리지 업그레이드까지 포함하기 때문에, 숨겨진 비용이 두려워 업그레이드를 하지 못하는 경우는 발생하지 않습니다. 이러한 관계를 통해, 계층 데이터를 보호하고 시퀀싱 워크플로우를 지속적으로 실행하는 데 필요한 최신 툴을 확보할 수 있습니다. 또한, 입증된 무중단 업그레이드를 통해 시퀀싱 실행이나 시간에 민감한 3차 분석을 방해하지 않고 필요할 때마다 업데이트할 수 있습니다.

AI 기반 3차 분석

세계 최초의 풀스택 AI 인프라인 에이리(AIRI®)는 인사이트를 얻기까지 걸리는 시간을 단축함으로써 과학자들이 계층 변이에 대한 3차 분석을 실행하여 추가적인 생물학 및 임상 정보를 기반으로 계층 변이의 중요성을 이해할 수 있도록 지원합니다. 퓨어스토리지와 NVIDIA가 공동으로 설계한 에이리는 전체 GPU 워크플로우를 가속화해주는 최신 플랫폼으로, 수집에서 추론까지 데이터 파이프라인의 모든 측면을 가속화하도록 구축되었으며, 즉시 사용이 가능한 통합 솔루션입니다.

참고자료

- [현대적인 데이터 보호 솔루션\(영문자료\)](#)에 대해 더 알아보세요.
- [에이리](#), [퓨어1](#), [서비스형 퓨어](#), [퓨어스토리지 프로페셔널 서비스](#)에 대해 자세히 알아보세요.

고객 사례

- [UC 버클리\(영문자료\)](#) 연구원들이 어떻게 DNA 시퀀싱을 가속화하고, 가치 있는 새로운 인사이트를 생성하며, 생명이 달린 중요한 질문에 대한 답을 보다 빠르게 얻고 있는지 알아보세요.
- 바르셀로나의 [농업유전체학연구센터\(영문자료\)](#)는 HPC 환경을 업데이트하여 속도, 확장성 및 안정성을 개선했습니다.
- 캐나다 해밀턴에 위치한 [맥마스터 대학교\(영문자료\)](#)의 과학자들은 전염병 및 코로나19 연구를 가속화하고 있습니다.

purestorage.com/kr

02-6001-3330

