

#2

하이브리드 HPC 플랫폼을 통한 R&D의 혁신 : 아렌티어 하이브리드

글. 서진우 클루닉스 대표이사

왜 하이브리드 고성능 컴퓨팅(HPC) 플랫폼인가

최근 디지털 전환(DT)과 초격차 R&D가 이슈화되면서 고성능 컴퓨팅(이하 HPC)에 대한 요구가 점차 증가하는 추세다. 2000년대 초반만 하더라도 HPC를 필요로 하는 그룹 내 연구원들은 팀 혹은 부서 단위로 각각의 요구 방식에 맞춰 HPC를 직접 구매하고 이용하는 형태를 갖춰왔다. 그러다 보니 소속 연구원의 HPC 운영 역량이 R&D 역량과 맞닿아 있는 상황이 연출되었다. 즉, HPC 역량이 부족한 조직에서는 R&D 역량을 강화할 수 있는 힘조차 잃게 되었다. 팀 혹은 부서 단위로 운영되던 HPC는 중복 투자를 야기하여 투자 효율성 면에서도 아쉬운 점이 많았다.

그러다 2010년, 국내 제조 산업의 큰 발전으로 R&D 연구원이 급증하게 되면서 HPC 수요가 급증하게 되자 대규모 연구소에서는 팀 단위로 HPC를 사용하기보다는 중앙에 통합하여 소속 연구원들이 좀 더 효과적이고 효율적으로 HPC 환경을 서비스 받을 수 있길 원했다. 즉, 연구원들은 오직 연구에 몰입하고, HPC 운영과 관리는 전문기업을 통해 HPC 환경을 서비스 받고자 했던 것이다. 그러나 지속적인 HPC 수요 증가에 따라 매번 복잡한 검증 절차와 도입 절차를 반복하게 되는 상황에 이르렀다. 또한 늘어나는 HPC 인프라를 관리하고 점검하는 에너지와 비용도 만만치 않았다.

이로써 시장은 자연스럽게 클라우드로 눈을 돌리게 된 것이다. 필요할 때, 필요한 만큼, 즉각 이용할 수 있는 클라우드를 통해 도입 절차를 축소하고, 관리 비용을 절감하고자 하는 욕구가 치달아 오르는 것이다. 하지만 클라우드로의 마이그레이션은 예

정된 몇 가지 이슈가 발생하게 된다.

첫째는 특수한 국내 R&D 상황과 맞지 않다는 것이다. 퍼블릭 클라우드는 표준화된 서비스 체계로 특수한 연구 환경에 맞춤형 솔루션을 제공하는 데에는 기술적 한계가 있다.

두 번째는 국내 제조 산업은 국가 산업으로 보안 이슈가 크다는 것이다. 그 때문에 대부분의 연구기관에서는 폐쇄망을 이용하고 있고, 이러한 상황에서 기업의 보안 정책에 맞춰 클라우드로 HPC를 이용한다는 것은 한계에 부딪힐 수밖에 없다.

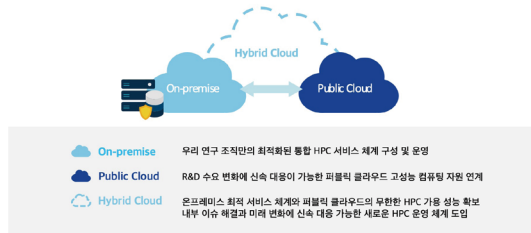
마지막으로 비용 이슈이다. 물론 소규모 HPC 작업을 위해 클라우드를 이용하는 데에는 비용적 이점이 있는 것이 분명하다. 하지만 대규모 연구소에서 오직 클라우드 기반의 HPC를 이용한다는 것은 실제 온프레미스를 새롭게 구축하고 운영하는 것보다 비용면에서 비효율적이라는 통계 보고가 충분하다.

이러한 까닭에, 초기에 클라우드로 전환했던 대부분의 기업들이 최근 다시 온프레미스로 회귀하고 있는 것이다. 그럼에도 불

그림 1. 고성능 컴퓨팅(HPC)은 다양한 산업 분야에서 핵심적인 전략 자산이 되었다.

예시 화면			
용도	CAD/설계(그래픽)용 <small>*온프레미드 HPC</small>	해석/시뮬레이션(계산)용 <small>*퍼블릭 클라우드 HPC</small>	딥러닝/AI/빅데이터용 <small>*온프레미드 HPC</small>
목적	컴퓨터로 제품 형상 CAD, 모델링, 해석결과 가시화	모델의 타당성률 수치 연산으로 예측, 검증	데이터 기계 학습으로 개발기간 단축, 정확도 향상
분야	반도체 전기/전자	자동차 조선/항공	AI/딥러닝 빅데이터
	에너지/화학	바이오	국방/안보 공공
	건설/건설	금융/보안	위험관리

그림 2. 변화에 신속 대응 가능한 하이브리드 HPC 플랫폼이 요구되고 있다.



구하고, 오직 온프레미스로 HPC를 운영하던 시대는 다시는 오지 않을 것이다. 왜냐하면 그 어느 때보다도 변화가 빠르기 때문이다. 예상치 못한 상황의 HPC 수요는 매번 발생할 것이고, 초격차 R&D를 향한 욕구는 끊임없이 진화할 것이기 때문이다.

그럼, 우리는 이러한 상황에서 어떤 선택을 할 수 있는 것일까? 급격한 변화 속에서 신속하게 대응 가능한 새로운 HPC 체계가 과연 있을까?

클루닉스는 새로운 HPC 체계로 하이브리드 HPC 플랫폼을 제안한다. 하이브리드 HPC 플랫폼은 내부에 구축되어 있는 최적의 연구 개발 환경은 그대로 유지하고, 상황에 따라 클라우드 기반의 HPC 자원을 이용하는 것이다.

하이브리드 고성능 컴퓨팅(HPC) 플랫폼의 개념

우선 하이브리드 HPC 플랫폼의 정확한 개념부터 살펴보자. 하이브리드 HPC 플랫폼은 우선 3단계로 나누어 그 개념을 설명할 수 있다. 첫 번째 단계는 인프라 통합이다. 인프라 통합이란 온프레미스 내 개별적으로 이용되던 모든 HPC 환경들을 중앙에 통합하고, 그 상태에서 퍼블릭 클라우드와 연계하는 것까지를 의미한다.

두 번째 단계는 통합된 인프라 위에 계산 처리를 위한 시뮬레이션, 그래픽 처리를 위한 엔지니어링 디자인, AI를 위한 딥러닝 플랫폼 등 다목적 다기능의 HPC 환경을 구성하는 것이다. 그리고 마지막 단계는 웹 기반의 단일화된 서비스 환경에서 온프레미스 자원과 클라우드 자원을 자유롭게 이용할 수 있도록 다양한 기능을 제공하는 것이다.

이러한 단계를 거쳐 구성된 하이브리드 HPC 플랫폼을 통해 사용자는 막힘없이 더욱 쉽고 빠르게 HPC를 이용할 수 있고 운영자는 기존의 운영 방식보다 더 효율적이고 안전하게 HPC 환경을 제공할 수 있다. 더불어 IT 투자 혹은 기획자의 경우, 기업

그림 3. 하이브리드 HPC 플랫폼은 인프라 통합, 플랫폼 구성, 서비스 환경 제공의 총 3단계를 거친다.



의 전략에 맞춰 필요한 HPC 환경을 즉각 투입하여 기업의 경쟁력을 확보할 수 있다.

‘아렌티어 하이브리드’의 6가지 이점

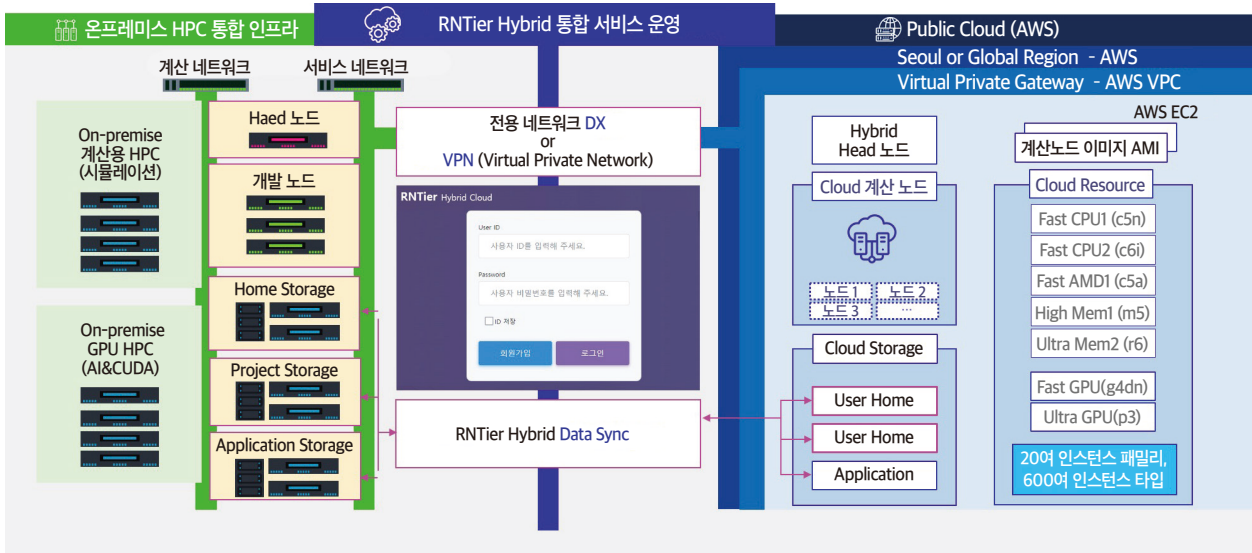
그럼 구체적으로 클루닉스가 제안하는 하이브리드 HPC 플랫폼 ‘아렌티어 하이브리드’의 이점에 대해 알아보자. 우선 ‘아렌티어 하이브리드’의 가장 큰 이점은 R&D 작업 성능에 최적화되어 있다는 것이다. 온프레미스는 물론 클라우드 자원 선정을 위해 PoC, BMT 등 엄격한 검증 절차를 거친다. 이를 통해 응용 S/W별 최적화된 인프라를 제공하여 부족함 없이 일관성 있는 HPC 자원을 이용할 수 있다.

다음으로 탁월한 확장성이 있다. 이는 클라우드가 지니고 있는 기본 속성과도 맞닿아 있다. 더욱이 ‘아렌티어 하이브리드’의 경우 온프레미스에서의 작업을 기본으로 하되, 자원이 부족한 경우 동일한 웹 환경에서 클라우드 자원을 즉각 이용할 수 있으며, 작업 대기 시간이 길어질 경우에도 클라우드로 자동 연결해 준다.

안전한 보안 체계도 ‘아렌티어 하이브리드’의 이점이다. 대기업 연구소의 경우 대부분 중앙 통합 방식의 HPC 환경을 운영하고 있지만, 여전히 많은 연구소에서는 부서 혹은 팀 단위의 HPC 환경을 구축하기도 한다. 그러다 보니 데이터 유출, 해킹 등의 보안 이슈에 노출되는 경우가 허다하다. 이런 경우 ‘아렌티어 하이브리드’ 도입을 통해 안전한 보안 관리가 가능하다. 또한, 온프레미스와 클라우드 간의 전용 네트워크를 통해 기업의 보안 정책을 준수하면서 클라우드로 접근할 수 있는 정책을 마련할 수 있게 된다.

클라우드의 가장 큰 이슈가 되고 있는 비용 문제도 ‘아렌티어 하이브리드’를 통해 해결이 가능하다. ‘아렌티어 하이브리드’에서는 비용을 최소화하기 위해 다양한 통제 기능을 제공한다. 그

그림 4. 아렌티어 하이브리드 인프라 구성도



룹별, 사용자별, 프로젝트별 비용을 별도 관리할 수 있고, 프로젝트가 종료된 경우에는 자원이 자동 회수되어 불필요한 비용 발생을 막는다. 또 실제 비용은 하나의 계정으로 통합 처리되어 내부 비용 정책에 맞춰 클라우드 자원을 운용할 수 있다.

‘아렌티어 하이브리드’를 통해 고객의 요구에 부합하는 최적화된 업무 환경을 구성할 수 있다. 즉, 기본 온프레미스 내 시스템과 클라우드가 연동되어 기존 운영 환경을 그대로 유지하여 업무를 효율적으로 이행할 수 있는 것이다. 그 밖에 쉽고 간편

한 서비스 환경을 제공하여 누구나 더욱 편리하게 HPC 업무를 수행할 수 있다는 것도 큰 이점이다. 한마디로 리눅스, HPC, 도커, 클라우드 등의 전문 지식 없이도 클릭만으로 HPC를 이용하여 업무를 가속할 수 있는 것이다.

각 기업의 최적화가 우선, ‘아렌티어 하이브리드’의 솔루션 구조

이쯤 되면 ‘아렌티어 하이브리드’가 어떻게 안정성을 확보하

그림 5. 아렌티어 하이브리드 동작 개념도

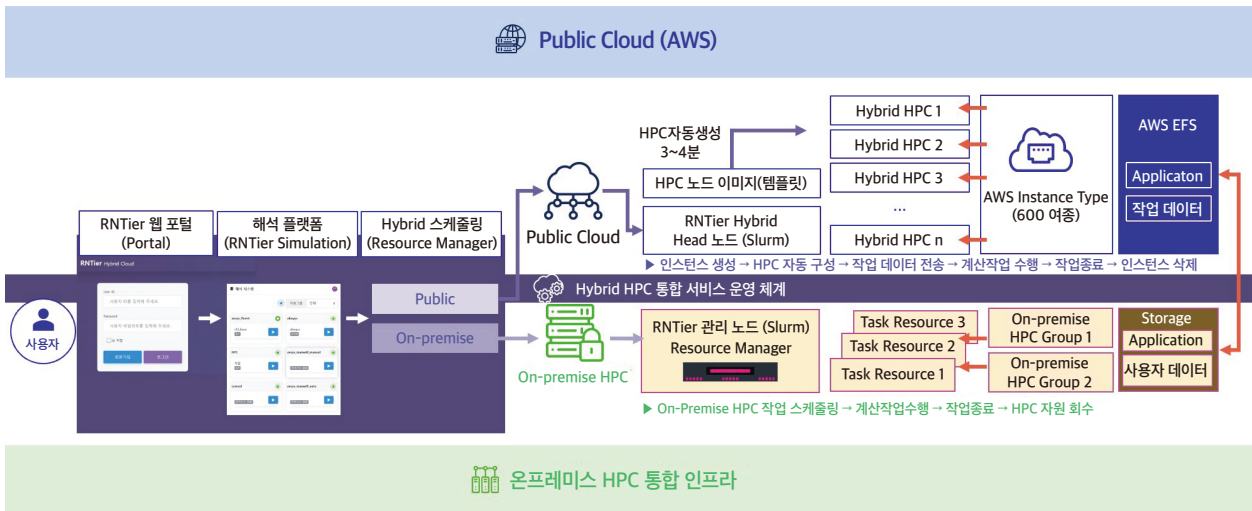
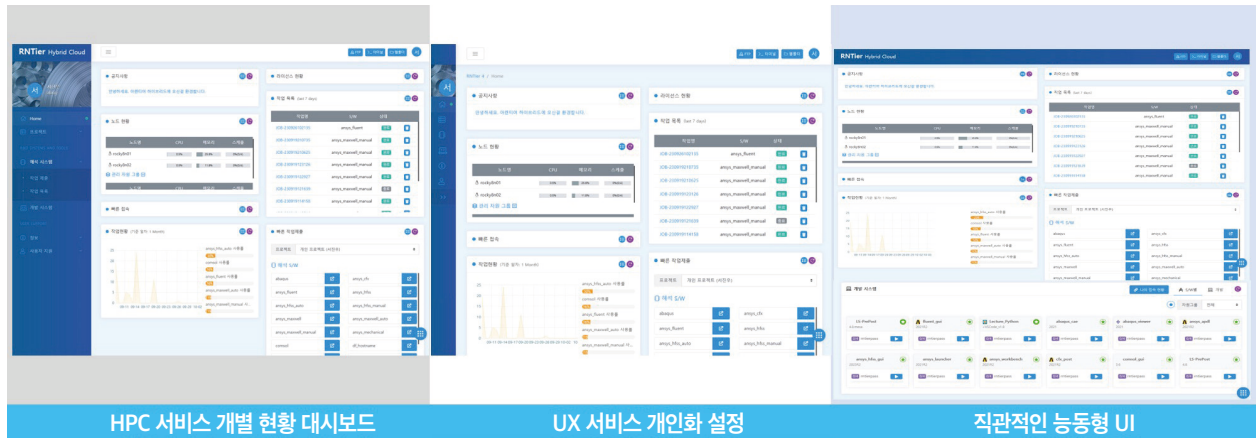


그림 6. 아렌티어 하이브리드 웹 UI 화면 예시



고 비용을 최적화할 수 있는지 가장 궁금할 것이다. 이 부분은 '아렌티어 하이브리드' 구조도를 통해 설명할 수 있다.

'아렌티어 하이브리드'의 기본 구조는 온프레미스 부분과 클라우드 부분으로 나뉜다. 우선 온프레미스 부분은 다양한 형태의 계산용 HPC, 엔지니어링 디자인용 HPC, 멀티 GPU, 스토리지 등을 연구 개발 환경에 맞춰 중앙에 통합한다. 여기에 HPC 서비스 환경 역시 고객 요구에 맞춰 최적화시킨다. 다음으로 클라우드 부분은 고객의 업무에 맞춘 최적의 인스턴스 자원을 추가 구성한다. 실제 AWS의 수천 종의 인스턴스 중 R&D 업무에 적합한 인스턴스 그룹이 선별된 것이다. 여기에 HPC 구성을 위한 기본 운영 체제 등을 이미지로 제공한다. 사용자는 클라우드 HPC 자원이 필요할 때, 이미지를 통해서 5분 이내에 필요한 HPC 환경을 제공받을 수 있다.

가장 중요한 부분은 온프레미스에 있는 데이터를 어떻게 클라우드 쪽에 동기화할 것인가이다. '아렌티어 하이브리드'는 온프레미스에 있는 특정 볼륨이나 데이터를 작업이 진행될 때 순간적으로 동기화시키고 작업이 끝나면 작업 결과를 다시 온프레미스로 동기화하는 데이터 싱크를 이용하고 있다.

마지막으로 네트워크에 대한 언급이 필요할 것이다. 왜냐하면 대부분 기업에서는 내부의 사설 네트워크를 통한 접근만 허용하기 때문이다. 이에 클루닉스는 실제 물리적인 전용선을 제안하거나 VPN(가상 사설 네트워크)를 이용하여 내부 정책에 맞춰 보안을 최적화하고 있다.

내부 자원의 최적화와 외부 자원의 확장성을 통해 R&D 경쟁력을 확보해야

최근 응용 S/W의 유연한 자원 옵션 덕분에 기업은 HPC 기능과 사양을 더욱 확장해 나갈 것으로 보인다. 게다가 과학 및 공학의 발전으로 시뮬레이션 및 모델 설계가 더 복잡해지고 있다. 또한 데이터의 확산으로 대규모 데이터 세트에서 의미 있는 통찰력을 분석하고 추출하기 위해서도 HPC는 전략 자산으로써 그 가치와 중요도가 더욱 높아지고 있다.

이러한 상황에서 기업은 '아렌티어 하이브리드'를 통해 일상적인 워크로드에 맞게 온프레미스 HPC 시스템 크기를 조정하고 시간에 민감하고 긴급한 워크로드를 위해서 피크 기간 동안 클라우드로 버스트할 수 있다. 즉, 내부 자원의 최적화는 물론 외부 클라우드 자원을 활용함으로써 새로운 R&D 경쟁력을 확보할 수 있을 것이다. 단순히 오늘의 필요성을 넘어 내일의 가능성까지 고려한다면, '아렌티어 하이브리드'는 가장 좋은 대안이 될 것이다.

...	저자소개	↗
서진우 클루닉스 대표이사는 2002년 클루닉스에 입사하여 국내의 취약한 연구개발 업무 환경을 개선하고자 국내 최초로 HPC 통합 솔루션 '아렌티어'를 개발하였다. 이후 삼성전자, LG디스플레이, 포스코, 서울대학교, 카이스트, 석유공사, 국방과학연구소 등 국내 다양한 연구기관에 아렌티어를 도입시킴으로써 국내 R&D 경쟁력을 크게 개선하였다. 이후 2021년에는 클루닉스 대표이사로 취임하여 클라우드 기반의 HPC 플랫폼 '아렌티어 클라우드'와 온프레미스와 클라우드를 연동하는 하이브리드 HPC 플랫폼 '아렌티어 하이브리드' 개발을 주도지휘하고 있다.		