



AI 시대의 핵심 동력 :

고성능·고확장 인프라 최적화 전략

AI 시대, 기업이 '인프라'에 주목하는 이유

AI 기술의 급속한 확산은 기업의 IT 인프라에 대한 접근 방식을 근본적으로 바꾸고 있습니다. 과거에는 IT 인프라를 주로 업무 지원과 비용 최적화 관점에서 구축해왔지만, 이제는 초거대 AI 모델의 학습과 운영을 위해 과거와 비교할 수 없는 수준의 연산 성능과 데이터 처리 능력이 요구되고 있습니다. 실제로 AI 학습에 필요한 연산량은 2012년 이후 평균 3.4개월마다 두 배씩 증가하고 있으며, 이러한 수요는 기업의 인프라 전략을 더욱 고도화하도록 만들고 있습니다.

이에 따라 글로벌 기업들의 AI 인프라 투자는 전례 없는 속도로 증가하고 있습니다. 2024년 상반기, 전 세계 AI 관련 컴퓨트 및 스토리지 인프라 지출은 전년 대비 97% 증가해 474억 달러에 달했고, 서버 투자는 105% 성장하며 전체 지출의 95%를 차지했습니다. 이는 AI를 제대로 활용하기 위해 기업들이 고성능·고확장 인프라에 집중하고 있음을 보여줍니다.

하지만 이러한 투자가 항상 성공으로 이어지는 것은 아닙니다. 많은 기업들이 PoC(개념검증) 단계에서 AI를 실험하지만, 상당수가 실제 운영까지 이어지지 못하고 있습니다. 이는 데이터 품질, 운영 프로세스, 인프라 성숙도 등 기반 역량의 부족이 주요 원인입니다.

이러한 현실은 명확한 시사점을 줍니다. AI 도입의 성공 여부를 결정하는 요소에는 전략과 기술력도 있지만, 무엇보다 이를 뒷받침할 인프라의 전략적 준비가 중요합니다. 단순히 예산을 늘리는 것만으로는 충분하지 않으며, AI의 가치를 실질적인 성과로 연결하기 위해서는 인프라부터 정교하게 설계되어야 합니다. 실제로 선도 기업들은 인프라를 더 이상 비용 절감을 위한 수단이 아니라, AI 기반 비즈니스 혁신의 핵심 자산으로 재정의하고 있습니다. 과거 비용 중심의 IT 예산이 이제는 AI, 클라우드, 보안 등 핵심 인프라 기술에 대한 선제적 투자로 전환되고 있으며, 이는 디지털 경쟁에서 우위를 확보하기 위한 전략적 선택으로 자리잡고 있습니다.

이 글에서는, AI 확산 시대에 왜 인프라가 기업 경쟁력의 본질로 부상하는지 그 배경을 짚고, 실제 데이터와 사례를 통해 인프라 투자 전략의 중요성을 분석합니다. 이어서 성공적인 AI 도입을 위한 인프라 요건과 이를 고려한 전략적 투자 방향을 제시함으로써, 인프라를 비용이 아닌 성장의 원천으로 전환하려는 기업에게 인사이트를 제공하고자 합니다.

인프라가 AI 워크로드에 최적화되기 위해 갖춰야 할 조건

AI 워크로드에 최적화된 인프라는 단순히 장비 성능이 뛰어난 수준을 넘어, 데이터 흐름, 운영 민첩성, 보안 통제까지 포함한 통합 역량을 요구합니다. 특히, 기업이 AI를 단기간의 프로젝트가 아닌 지속 가능한 혁신 자산으로 운영하기 위해서는 아래의 네 가지 조건을 반드시 갖춰야 합니다.

강력한 연산 능력: AI에 맞는 하드웨어 기반 확보

AI 모델 학습과 추론은 전통적인 시스템보다 수십 배 이상의 연산 성능을 요구합니다. 최근 전체 AI 서버 투자 중 70% 이상이 GPU 기반 가속 서버에 집중되고 있으며, 이 비중은 앞으로 더 확대될 것으로 보입니다. 복잡한 모델을 빠르게 학습하고, 실시간 응답까지 수행하려면 전용 연산 자원 확보는 반드시 전제되어야 합니다.

막힘없는 데이터 흐름: 고속 스토리지·네트워크 최적화

GPU 성능만으로는 AI 시스템이 완성되지 않습니다. AI가 고속으로 학습하고 추론하기 위해서는, 데이터를 얼마나 빠르게 공급할 수 있는지가 핵심입니다. 이를 위해 고성능 스토리지, 네트워크, 데이터 전송 경로까지 전체 인프라가 유기적으로 최적화되어야 합니다. 실제로 최신 인프라를 적용한 기업들은 데이터 처리 속도가 약 2배 향상되었으며, 이는 AI 모델 학습 속도와 서비스 민첩성 향상으로 직결되고 있습니다.

유연한 확장성: 클라우드와 온프레미스의 전략적 조화

기존 인프라는 예측 가능한 수요에 맞춰 고정된 자원을 구축하는 방식이었지만, AI 프로젝트는 데이터와 연산 요구가 급격히 증가하는 특성상 이러한 접근만으로는 대응이 어렵습니다. 하이브리드 클라우드는 필요한 순간에 외부 클라우드 자원을 빠르게 확장하고, 민감한 데이터는 온프레미스에서 안전하게 처리할 수 있어, 민첩성과 보안을 동시에 확보할 수 있는 대안이 됩니다. 실제로 AI를 도입한 기업 3곳 중 2곳은 하이브리드 클라우드를 채택하고 있으며, 이는 AI 운영의 확장성과 유연성을 뒷받침하는 핵심 전략으로 자리잡고 있습니다.

신뢰 기반 운영체계: 보안·규제·거버넌스를 아우르는 설계

AI 인프라는 고성능 못지않게 신뢰성과 통제력이 중요합니다. 특히 규제가 엄격한 산업일수록 데이터 보호, 감사 추적, 모델 설명 가능성이 핵심 요건으로 작동합니다. 최근에는 AI 시스템의 투명성과 책임성을 보장하기 위한 거버넌스와 보안 체계 도입이 AI 인프라 설계의 기본값이 되고 있습니다.

고성능·고확장 인프라 도입이 기업에 주는 효과

AI 워크로드에 최적화된 인프라에 대한 투자는 단순한 기술 업그레이드를 넘어, 기업이 생산성 향상, 비용 절감, 민첩한 확장, 그리고 혁신적 비즈니스 가치를 실현하는 기반이 됩니다.

Dell Technologies와 Nutanix 등의 사례를 중심으로, 고성능·고확장 인프라가 기업에 가져다 주는 효과를 살펴보겠습니다.

AI 처리 성능 향상과 시장 대응 속도 가속화

최신 세대의 Dell PowerScale은 이전 세대 대비 스트리밍 쓰기 속도가 220% 향상되었고, 읽기 속도는 99% 빨라져 AI 개발 일정을 단축하고 있습니다. 이처럼 고성능 인프라는 AI 개발 주기의 단축, 서비스 출시 가속화로 이어져 AI 모델 학습 및 추론 속도를 획기적으로 개선하고, 시장에서 빠르게 경쟁 우위를 확보할 수 있게 합니다.

IT 생산성 제고와 운영 효율 개선

최신 인프라는 시스템의 가용성과 운영 안정성을 높여, IT 부서의 일상 업무 효율을 크게 향상시킵니다. Dell의 인프라 구독 모델을 통해 서버와 스토리지를 주기적으로 최신화한 기업은 예기치 않은 다운타임을 최소화하고, 시스템 운영의 연속성을 확보할 수 있었습니다. 또한 자동화된 운영관리 서비스는 IT 인력의 반복적인 유지보수 업무를 줄이고, 기술팀이 더 전략적인 프로젝트에 집중할 수 있는 환경을 만들어 줍니다. 이러한 변화는 조직 전체의 IT 대응 속도와 유연성을 향상시키는 기반이 됩니다.

DELL PowerFlex 도입 기업이 얻을 수 있는 총 경제적 효과 (연간)

*IDC / DELL, 2023

724만 (달러)

비즈니스 생산성 향상: 328만 달러

IT인력 생산성 향상: 262만 달러

인프라 비용 절감: 92만 달러

리스크 완화 및 사용자 생산성 개선: 42만 달러

TCO 절감과 ROI 향상

운영 효율성과 안정성을 바탕으로 한 인프라 혁신은, 결과적으로 재무적 가치 창출로 이어집니다.

Dell PowerFlex를 도입한 기업은 3년간 평균 276%의 ROI, 40%의 TCO 절감, 88%의 다운타임 감소를 달성했습니다. Nutanix 역시 하이퍼컨버지드 인프라(HCI)를 통해 별도 하드웨어 구성 없이 고성능 환경을 구현함으로써 IT 인프라의 총비용을 62% 줄이고, 477%의 ROI와 61%의 생산성 향상을 기록한 바 있습니다. 이처럼 고도화된 인프라는 단순한 기술 업그레이드를 넘어, 비용 구조 개선과 투자 효율 극대화라는 실질적인 재무 성과를 만들어냅니다.

유연한 확장성과 민첩한 대응력 확보

고확장 인프라는 예측 불가능한 수요 변화나 AI 트래픽 증가에 빠르게 대응할 수 있는 기반입니다. Dell은 최대 96개 GPU 탑재가 가능한 PowerEdge 서버, 스토리지·컴퓨팅 자원을 독립적으로 확장할 수 있는 PowerFlex, 그리고 GPU 자원을 탄력적으로 할당할 수 있는 Run:AI 플랫폼 등을 통해 리소스 확장과 AI 오케스트레이션의 민첩성을 제공합니다. Nutanix는 온프레미스, 퍼블릭 클라우드, 엣지를 아우르는 통합 플랫폼을 통해 AI 워크로드의 위치에 상관없이 일관된 성능과 보안을 보장하며, 유연한 하이브리드 확장을 지원합니다.

혁신 가속과 ESG 가치 실현

AI 인프라는 기술 혁신의 속도를 높이는 동시에, ESG 목표 달성에도 기여합니다. Dell은 생성형 AI, 디지털 트윈 등 다양한 활용 사례를 빠르게 구현할 수 있도록 지원하고 있습니다. Dell의 고성능 인프라를 도입한 한 글로벌 에너지·제조 전문기업은 동일한 성능을 유지하면서도 에너지 소비를 70% 절감하며, 기술 혁신과 ESG 목표를 동시에 실현한 바 있습니다. Nutanix를 도입한 기업들 또한 평균적으로 물리적 공간 70%, 에너지 소비 50%를 절감하며, 지속 가능한 인프라 전환을 가속화하고 있습니다.

AI 인프라 전략 결정 시 고려해야 할 4가지

AI 도입이 전 산업으로 확산되며, 이에 최적화된 고성능·고확장 인프라에 대한 기업들의 투자도 본격화되고 있습니다. 하지만 단순한 기술적 비교나 비용 검토만으로는 인프라 전략을 결정할 수 없습니다. 특히 기업의 의사결정자는 단기 효율성과 장기 전략, 기술 리스크와 규제 리스크 간 균형을 고려해야 하며, 이를 위해 아래 4가지 관점에서 종합적인 검토가 필요합니다.

총소유비용(TCO)과 투자수익률(ROI)의 균형

단기적인 장비 구매비용뿐 아니라, 향후 3~5년간의 인프라 운영비용, 증설 비용, 에너지 소비, 인력 운용비 등을 종합적으로 고려해야 합니다. 초기 자본적 지출(CAPEX)이 부담될 경우 Dell APEX와 같은 운영비용(OPEX) 전환 모델을 검토해 현금 흐름을 최적화할 수 있으며, 인프라 도입에 따른 비용 절감(다운타임, 유지비 절감 등)과 가치 창출(생산성 향상, 신사업 기회 등)을 계량화하여 ROI 관점의 설득력 있는 비즈니스 케이스를 마련해야 합니다.

확장성과 유연성 확보

AI 워크로드의 급격히 증가하거나 방향이 급변할 수 있으므로, 모듈형 설계 및 스케일 아웃 구조 기반의 인프라가 필요합니다. 퍼블릭 클라우드와의 연계 고려한 하이브리드 전략, 멀티클라우드 간 이동성을 보장하는 포터블 아키텍처, 쿠버네티스 기반 컨테이너화 등의 기술도 검토 대상입니다. 장기적으로는 클라우드 네이티브를 기반으로 온프레미스와 클라우드를 유기적으로 통합할 수 있는 구조가 필요합니다.

운영역량과 기술인력 확보

아무리 우수한 인프라를 도입하더라도 이를 운영할 조직 체계와 인력이 뒷받침되지 않으면 효과를 내기 어렵습니다. 현 조직의 IT 운영 역량, 데이터 사이언스와 MLOps 수준을 진단하고, 부족할 경우 외부 전문가 투입, Dell 등의 기술 파트너를 통한 매니지드 서비스 활용, AI CoE 조직 신설 등도 검토해야 합니다. 인프라는 기술만이 아니라 '사람과 프로세스'의 조화 속에서 가치를 만들어냅니다.

보안·규제 준수와 리스크 관리

AI 인프라는 민감한 데이터와 고도화된 모델을 동시에 다루기 때문에 보안 위협과 컴플라이언스 리스크에 노출될 수 있습니다. AI 모델 보안, 접근 제어, 데이터 주권, 유럽연합 개인정보보호법(GDPR) 등 글로벌 규제, AI Act 및 국내 윤리 기준 등을 모두 검토하여 인프라 설계에 반영해야 하며, 재해복구와 백업 체계 또한 필수입니다. 특히 금융·의료 등 고규제 산업은 온프레미스·퍼블릭 클라우드 선택 자체도 규제 기준에 따라 신중히 판단해야 합니다.

인프라를 '비용'이 아닌 '경쟁력의 기반'으로

이제 인프라는 비용 절감의 대상이 아니라, 디지털 전환과 기술 경쟁력의 핵심 기반입니다. AI 선도 기업들은 고성능·고확장 인프라를 통해 단순 효율을 넘어 변혁적 성과를 창출하고 있으며, 이는 시장 내 뚜렷한 수익과 성장 격차로 이어지고 있습니다.

특히 민첩한 인프라(Agile Infrastructure)는 급변하는 시장에 빠르게 대응하며, 새로운 AI 기술을 유연하게 적용할 수 있는 역량을 제공합니다. 반면 인프라가 뒤쳐진 기업은 변화에 발맞추지 못하고 기회를 잃을 위험이 큼니다.

결국, 지금의 인프라 결정은 향후 수년간 우리 기업의 혁신 역량과 생존 가능성을 좌우합니다. 단기 비용보다는 장기 경쟁력에 기반한 인프라 전략이 중요해지고 있습니다.

과감한 인프라 혁신으로,
AI 시대의 기회를 비즈니스 성과로 연결할 수 있습니다.

References

OpenAI. AI and Compute, 2018.

IDC. Worldwide Semiannual Artificial Intelligence Infrastructure Tracker®, 2025.

Red Hat. Two out of Three Enterprises Use Hybrid Cloud to Power AI Initiatives, 2023.

IDC (sponsored by Dell Technologies and Intel). The Business Value of Dell PowerFlex, 2023.

Dell. Transforming AI Data Pipelines with Dell and NVIDIA, 2025.

IDC. Organizations Leverage Nutanix Enterprise Cloud as Scalable, High-Performing, and Cost-Effective Infrastructure Foundation, 2020.

Contributors

Ko, Jung-woo | Lee, Ji-hye | Joo, Hye-rim

About RockPLACE

Rockplace is a leading IT solutions provider driving digital transformation through cloud, open-source technologies, data analytics, and AI-powered automation. We deliver end-to-end services—from infra and data optimization to tailored AI solutions—to maximize business efficiency. By leveraging open-source technologies, we offer flexible and cost-effective solutions, while automation ensures seamless operations. As a trusted partner, we support our clients’ sustainable growth and enhance their competitiveness in the global market.

About Metanet

Metanet is Korea’s largest full-stack IT service partner, leading digital innovation through AI and cloud-based solutions. We drive sustainable growth through integrated processes from digital strategy to IT implementation and operations. Our industry-leading expert groups collaborate closely on a unified platform to deliver innovative solutions and optimal support, empowering businesses to grow and strengthen their competitiveness.

