



AT A GLANCE

Accelerate AI Data Pipeline – FlashStack™ for AI



THE CHALLENGE: SCALABLE INFRASTRUCTURE FOR THE AI PIPELINE

Artificial Intelligence workloads process massive amounts of data from structured and from unstructured sources. High performance, architecturally optimized solutions that can harmoniously exist within existing data centers is just one of the infrastructure factors IT teams need to easily address without increasing complexity.

No data pipeline is alike and data scientists developing artificial intelligence and machine learning based applications depend on dedicated platforms to manage their complex data pipeline. They need an infrastructure that can manage every element of the pipeline from data ingestion and cleansing to modeled results. They require high performance engines dedicated for workload compute processing, with a storage platform that easily scales and protects the raw and subsequent processed data.

INTRODUCING: FLASHSTACK FOR AI

Announcing **FlashStack for AI**, a holistic solution designed to deliver an end-to-end AI pipeline solution. Jointly developed by Cisco and Pure Storage®, this flexible scale-out solution is powered by Cisco® UCST™ server and Pure Storage FlashBlade™.

FlashStack for AI combines the compute scale of the Cisco UCS C480 ML in particular, and Cisco UCS server portfolio in general, with Pure's Data Hub. This **fully converged infrastructure for both ML and DL applications** offers:

BENEFITS OF FLASHSTACK FOR AI

- Powers the data pipeline at every stage with flexible compute and storage
 - Easily scales at each ingestion point when more performance is needed
 - Reduces deployment risks with a FlashStack proven and validated architecture
 - Easily adopts AI/ML without re-architecting the data center
- **Cisco UCS C480 ML system:**
A purpose-built rack server for deep learning and machine learning that includes eight NVIDIA Tesla V100 Tensor Core GPUs with NVLink Interconnect for faster model training.
 - **Pure Storage FlashBlade:**
An elastic scale-out storage platform that delivers performance that scales linearly as data grows. FlashBlade, designed for next-gen analytics and AI workloads, enables enterprises to seamlessly share data across the modern AI pipeline.
 - **Cisco Nexus switches:**
Provide the massive network bandwidth required by AI and ML applications.

FlashStack for AI provides flexibility in terms of CPUs, memory, storage and network. Customers have a broad portfolio available for AI/ML applications with C480 MLs and C220/C240s with GPUs. Depending on the stage of their journey and the task being handled, customers can choose a server with a single GPU or a set of C480 MLs with 10s/100s of GPUs.





THE FLASHSTACK FOR AI ADVANTAGE

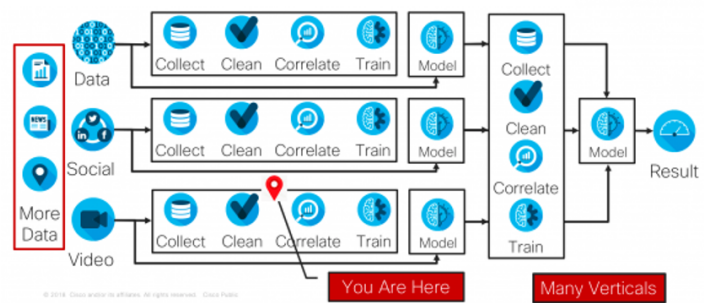
FlashStack customers can easily adopt AI/ML for their business without re-architecting their data center. Here are just a few use cases for FlashStack for AI converged infrastructure.

- A 24/7 light out warehouse that uses automated, intelligent robotics ready to meet high demand customer shipment requirements
- A virtual concierge that responds to customer online queries and suggest purchases based on customer choices
- A more secure banking transaction program that uses intelligent algorithms to detect fraud patterns.

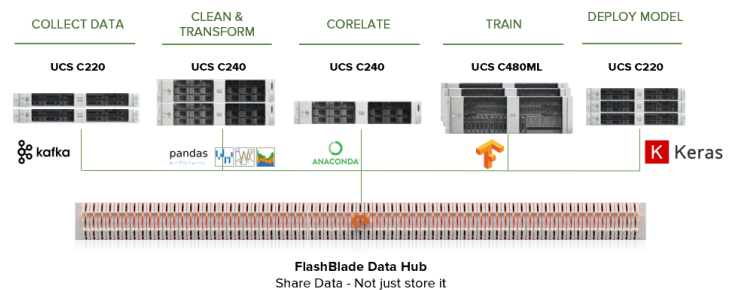
A CLOSER LOOK AT FLASHSTACK FOR AI FOR THE DATA PIPELINE

As artificial data pipelines grow, the data sources become more sophisticated. Data scientists often work with multiple data sources, each with their own pipeline for collection, cleansing, and correlation. The inferred data from each of these data sources is then fed to a second stage of learning gathering the information from multiple data sources. This type of data pipeline with multiple data sources can be found in many verticals, from security to targeted marketing campaigns. Data scientists wonder whether additional data sources, such as structured data, live news feed, geolocation data, and other sources, should be added to the mix to gain deeper insights into the business problem at hand. The focus on the data and its operations prevents the data scientists from focusing on the infrastructure that is needed to run the data pipelines.

DATA PIPELINE FOR MULTIPLE DATA SOURCES



SIMPLIFY YOUR DATA PIPELINE - POWERED BY FLASHSTACK FOR AI



With **FlashStack for AI workloads**, IT organizations can deploy a validated architecture for AI workloads that reduces the design risks in building a data, compute, and storage infrastructure for the AI data pipeline so data scientists stay focused on delivering AI solutions for better business outcomes. As the amount of data that needs to be processed for AI is massive, high performance, architecturally optimized solutions that can harmoniously exist within your existing environment is critical. IT teams need to easily address these new workloads without increasing complexity. FlashStack for AI workloads provides a proven, validated design for these highly data intensive workloads.

PURE STORAGE, INC. 650 CASTRO STREET, MOUNTAIN VIEW, CA 94041
www.FlashStack.com



© 2019 Pure Storage, Inc. Pure Storage, the "P" Logo, and FlashStack are trademarks or registered trademarks of Pure Storage, Inc. in the U.S. and other countries. Cisco and the Cisco logo are trademarks or registered trademarks of Cisco and/or its affiliates in the U.S. and other countries. Intel, the Intel logo, Xeon, and Xeon Inside are trademarks or registered trademarks of Intel Corporation in the U.S. and/or other countries. All other trademarks are the property of their respective owners. PS-FS-AAG-IntroducingFSforAI-0219-v1