

Pure Storage Breakout session
11 09 2019 – 15.30

High Performance Analytics & AI Pipelines

Brian Carpenter



High Performance Analytics & AI Pipelines

- ▶ A production data pipeline is more than just training a deep learning model or transforming streaming datasets . It also includes 1) ingesting and running inference on new data, 2) transformation, processing, and cleaning new data to incorporate it into the training set, 3) continuously retraining to update and continue learning, and 4) experimental pipeline to test improvements outcomes. This presentation focuses on the importance of high-performance and highly-scalable storage that is needed to take advantage of ever-larger datasets in analytics and model training.
- ▶ We describe the common stages for input pipelines ranging from deep learning training to continuous improvement / continuous development and describe their resource requirements.

High Performance Analytics & AI Pipelines



Brian Carpenter

Sr. Dir. FlashBlade Technology Strategy

At Pure Brian enables customers' journeys to high performance outcomes with Artificial Intelligence, Advanced Analytics, and Platform as a Service. Brian's served in a variety of roles including heading up IT Departments, creating and fostering technology alliances and business development initiatives for the data storage industry. He's shared his passion for technology as a business enabler through mentoring programs for new hires, and advising start ups on IT strategy and value propositions to their organizations.