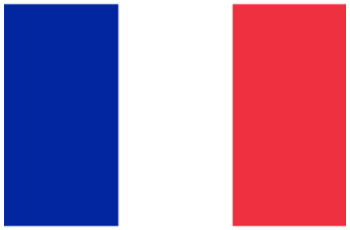
HOW TO BUILD A MODERN AI
FOR THE UNKNOWN IN MODERN DATA
PURESTORAGE®

Official Languages Act (1969/1988)



Translation Bureau

© 2017 PURESTORAGE INC.

CBC news | Ottawa

Home | Opinion | World | Canada | Politics | Business | Health | Entertainment | Technology & Science | Video

Canada | Ottawa

AUDIO | Online translator helps federal workers 'do their job,' say defenders

French-English translator developed by National Research Council over 10 years

CBC News Posted: Feb 04, 2018 5:14 PM ET | Last Updated: Feb 04, 2018 5:14 PM ET

Donna Achimov, the Translation Bureau's chief executive officer, says the tool is meant to help workers wade through technical documents and compose minor communications between staff, such as emails and texts, which aren't normally sent to the bureau for translation. (CBC News)

75 shares

Facebook | Twitter | Reddit | Google | Share | Email

Related Stories

- MP wants to reconsider launch of French-English online translator for federal workers

The federal departments behind a translation tool that has been criticized as "clumsy" are defending the software as a way for employees to gain confidence working in both official languages.

- **Online French-English translator for federal workers erodes bilingualism, MP says**

The online tool was developed by the National Research Council and will be operated by the government's Translation Bureau when it is launched for 350,000 workers across the country on April 1.

It was tested in a pilot project with 300 public servants last summer.

The tool comes as the bureau's workforce has shrunk from 2,000 to a little more than 1,300 employees over the course of five years.

Translation Bureau CEO Donna Achimov said the tool is not meant to replace human translators but to help government workers build confidence crafting short emails in both official languages.

"Over a million uses a week of Google Translate happen in government without government terminology," she told CBC Ottawa's Alan Neal on *All In A Day*.

"We're sitting on hundreds and thousands and millions of words and phrases that we've already translated. And so what we wanted to do is to provide a tool for comprehension."

The program, called Portage Statistical Machine Translation, was developed over the course of 10 years, said Joel Martin, director of research and development at the NRC's information and communications technologies division.

Stay Connected with CBC News

Mobile | Facebook | Podcasts | Twitter | Alerts | Newsletter

Weather

Ottawa (Kanata - Orléans)	Gatineau	Petawawa	Cornwall	Kingston
20°C	20°C	18°C	n/a	19°C

More Weather |

Latest Ottawa News Headlines

- Ottawa police plans to add hijabs to inclusion policy
- Homicide victim's son tells court he watched man chase his mother with a gun
- PSAC calling on government to expand coverage of contraceptives
- College strike halts courses for students without high school diplomas
- Accused triple murderer assaulted victim year before her killing, trial hears

Top News Headlines

- 'It's not a coercive law,' Quebec says amid criticism over face-

SLATOR
Language Industry Intelligence

Visibility, collaborat – all in **real-time!**

NEWS ~ | PRESS RELEASES | SPONSORED CONTENT | EVENTS | JOBS | DIRECTORY

Industry News

Woes of Canada's Translation Bureau Compounded By Minister's Response

by Marion Marking on October 28, 2016

Reactions ranged from "disappointing" and "lacking" to "insulting," "shocking" and "jaw-dropping." The Canadian government, already under fire from several quarters due to, among other things, the downsizing of the Translation Bureau's workforce, the rollout of its homegrown machine translation tool Portage, and the planned use of a vetting system that favors the lowest-bidding freelancers, is embattled once again.

The heat is specifically directed at the Department of Public Services and Procurement, which oversees the Translation Bureau. The reason, Minister Judy Foote's response to the recommendations of the House of Commons Standing Committee on Official Languages regarding the Translation Bureau, which falls under Foote's purview.

ADVERTISEMENT

Your tool for translation...

Foote, the most popular MP candidate in last year's federal elections, topping it with an 82% landslide, heads the Department of Public Services and Procurement. A Liberal, Foote was sworn into the Cabinet along with Prime Minister Justin Trudeau in November 2015. Supported by, among others, Canada's labor unions, the Liberal Party won last year's elections for the first time in nine years by a huge margin.

© 2017 PURESTORAGE INC.
PURESTORAGE®

© 2017 PURESTORAGE INC.

DAWN OF 4TH INDUSTRIAL REVOLUTION

BIG DATA, AI DRIVING CHANGE IN EVERY INDUSTRY

1st Revolution
1760-1820's
Steam Power
Rural to Industrial

2nd Revolution
1870-1914
Electricity
Industrial to Mass Production

3rd Revolution
1980-2010
PC
Mass production to Digital

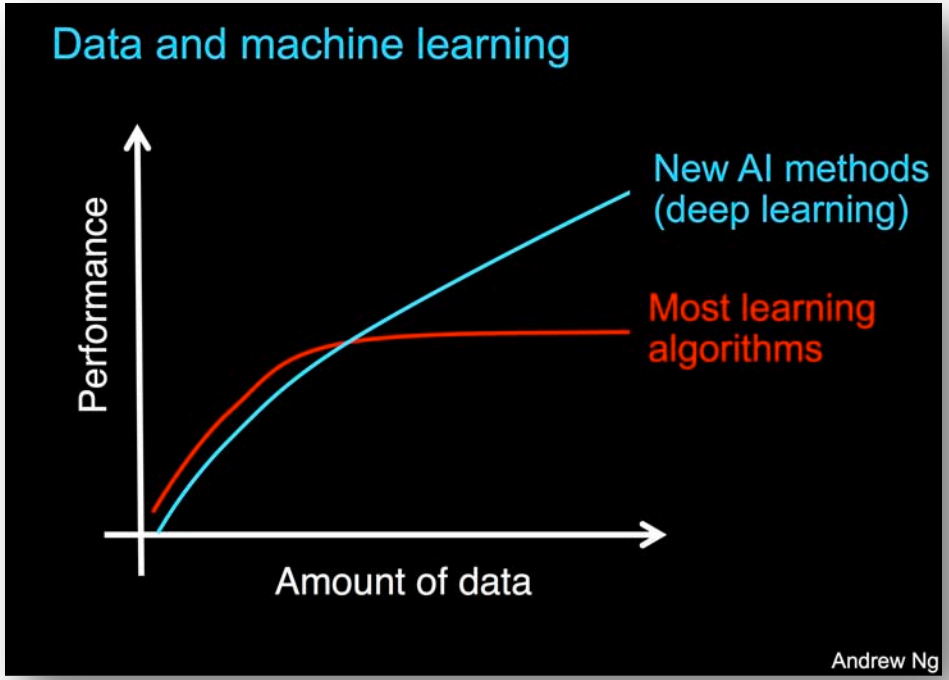
4th Revolution
2010-now
AI, Big Data & IoT
Digital to Intelligence

© 2017 PURESTORAGE INC.

PURESTORAGE®

DATA IS VITAL TO MACHINE LEARNING

OBSERVATION BY PROF. ANDREW NG, AI LUMINARY



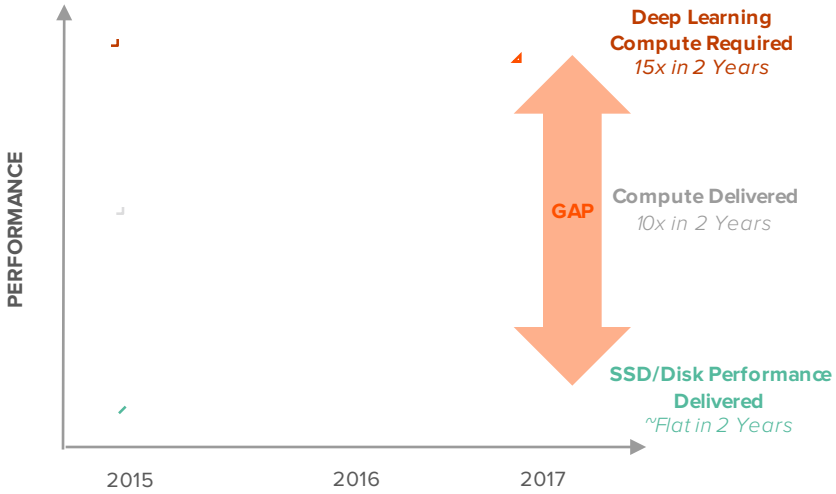
© 2017 PURESTORAGE INC.

VALUABLE DATA STUCK IN NEUTRAL

LEGACY, RETROFIT STORAGE BUILT ON SERIAL TECHNOLOGIES, PERFORMANCE GAP GROWING

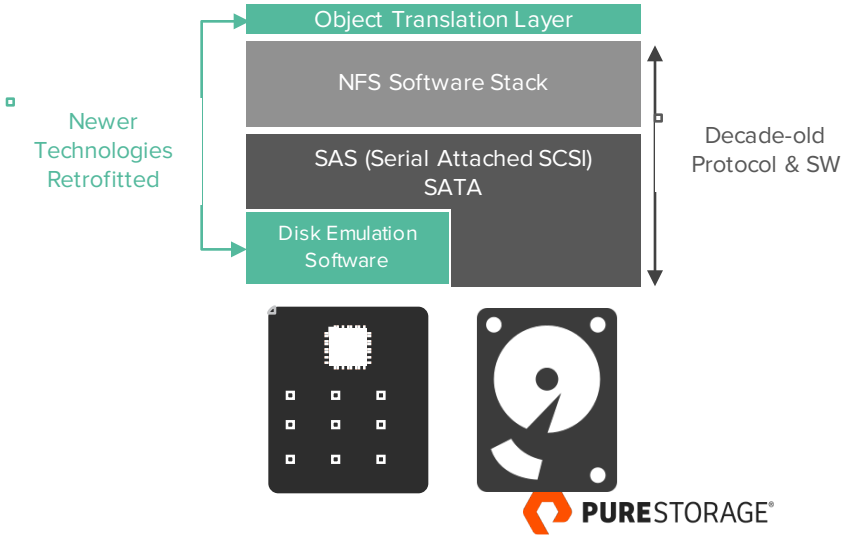
STORAGE TECHNOLOGY NOT KEEPING UP

Gap Will Only Grow Worse



LEGACY & RETROFIT STORAGE

Built on Decade-Old Serial Technology



© 2017 PURESTORAGE INC.

PURESTORAGE®

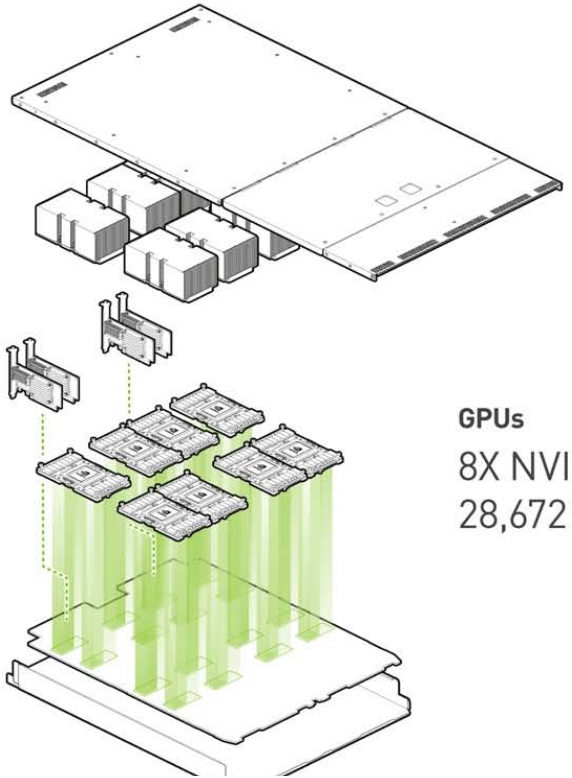
STORAGE FOR AI:
BOTH A TRUCK AND A RACE CAR

CAPACITY
LARGE FILES
THROUGHPUT
SEQUENTIAL ACCESS

CONCURRENCY
SMALL FILES
LATENCY
RANDOM ACCESS

© 2017 PURESTORAGE INC.

SOUL OF DGX-1 IS PARALLEL



NETWORK INTERCONNECT
4X InfiniBand™ 100 Gbps EDR
2X 10GbE

GPUs
8X NVIDIA Tesla® P100 16GB/GPU
28,672 Total NVIDIA CUDA® Cores

GPU INTERCONNECT
NVIDIA NVLink™
Hybrid Cube Mesh

© 2017 PURESTORAGE INC.

SOUL OF FLASHBLADE IS PARALLEL

POWERING 75 BLADE-SCALE IN SINGLE IP WITH PURITY FOR FLASHBLADE

Purity FB 2

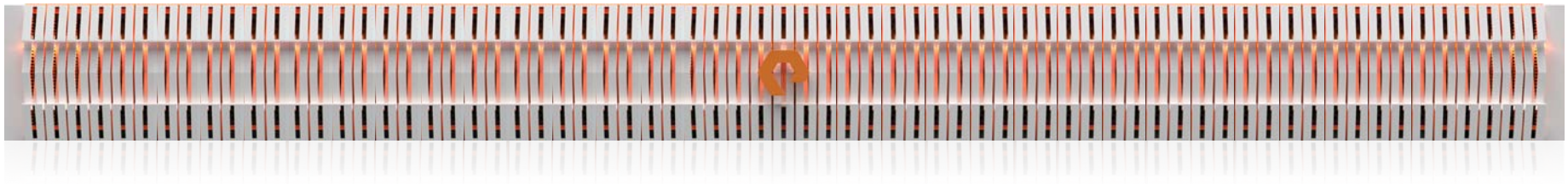
NATIVE OBJECT >_OBJECT

NATIVE NFS/SMB >_NFS >_SMB

KEY-VALUE DATABASE STORE FOR DISTRIBUTED PARTITIONS

KEY
VALUE

BILLIONS & BILLIONS OF OBJECTS



© 2017 PURESTORAGE INC.

PURESTORAGE®

75 blade feature is subject to GA release

THREE ESSENTIAL THINGS FOR AI

FRAMEWORKS & APPLICATIONS

Caffe2

mxnet

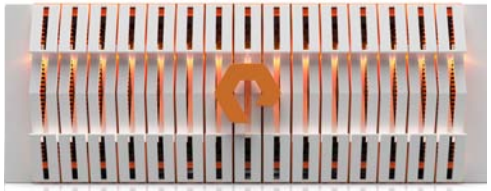
PYTORCH

TensorFlow

COMPUTE
FROM CPU TO GPU SERVERS

NVIDIA

STORAGE
POWER ENTIRE AI PIPELINE



PURESTORAGE

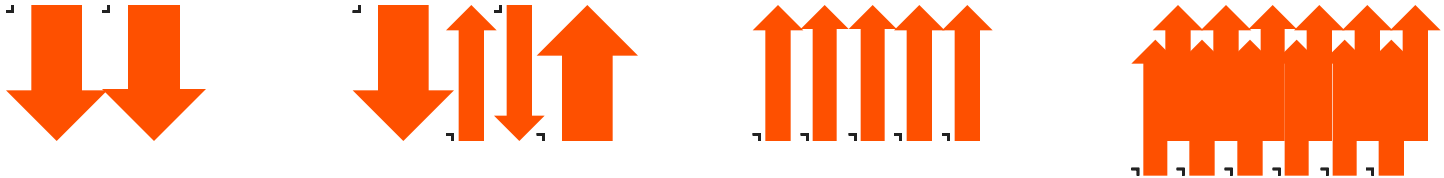
© 2017 PURESTORAGE INC.

WIDE RANGE OF NEEDS IN THE PIPELINE

SIGNIFICANT CHALLENGE TO LEGACY STORAGE

INGEST	CLEAN & TRANSFORM	EXPLORE	TRAIN
From sensors, machines, & user generated	CPU Servers	GPU Server	GPU Production Cluster

IO PROFILE



© 2017 PURESTORAGE INC.

PURESTORAGE®

REAL WORLD PIPELINE IN AN AUTONOMOUS CAR COMPANY

INGEST

CLEAN, LABEL, RESIZE

EXPLORE

TRAIN

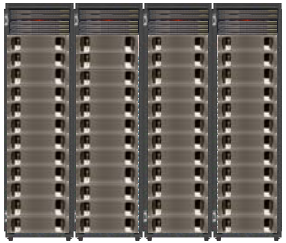
INFERENCE IN VIRTUAL WORLD

CPU Servers

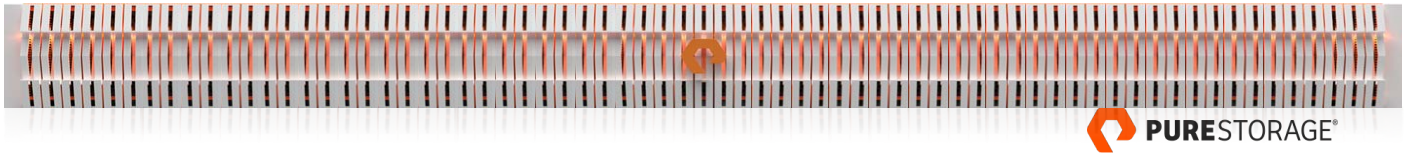
GPU Server

GPU Production Cluster

GPU Production Cluster



10'S OF PB
COLD STORAGE



© 2017 PURESTORAGE INC.

PURESTORAGE®

AI SYSTEMS DESIGN PATTERNS

GOAL IS TO KEEP THE GPUs 100% BUSY

FULL TRAINING WORKFLOW

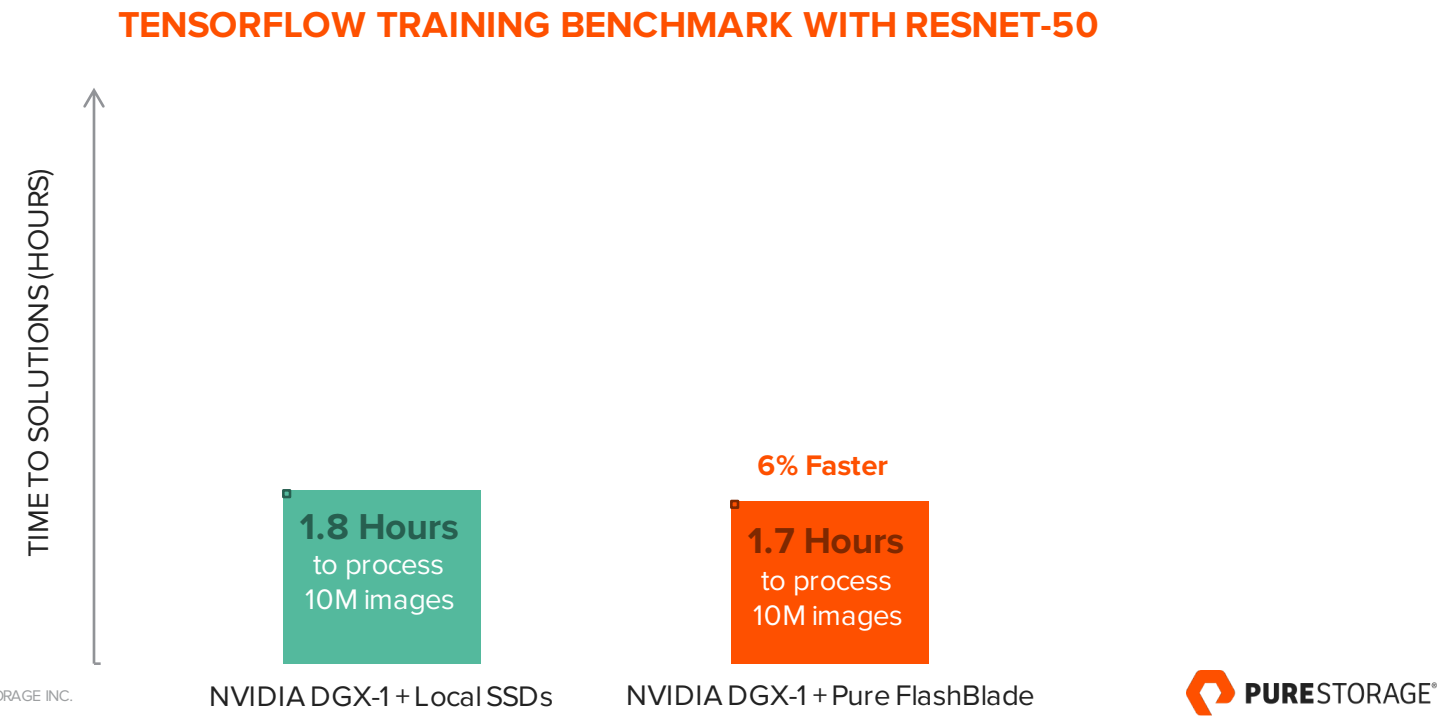
BENCHMARK SETUP

Setup #1: DGX-1 with 4x Local SSDs

Setup #2: DGX-1 with 1x FlashBlade

© 2017 PURESTORAGE INC.

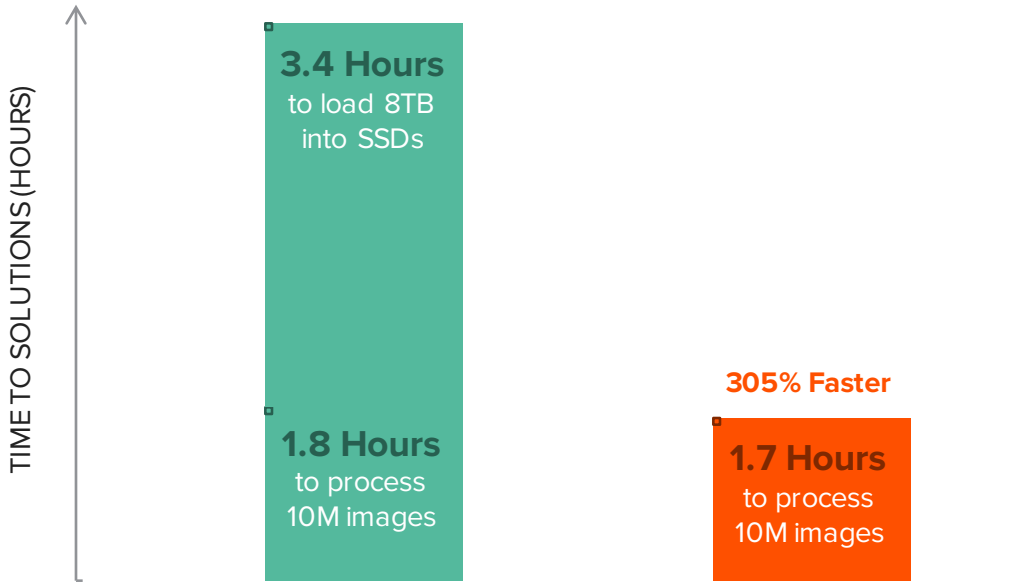
RESULT: FLASHBLADE vs LOCAL SSDs



© 2017 PURESTORAGE INC.

RESULT: 3X FASTER END-TO-END

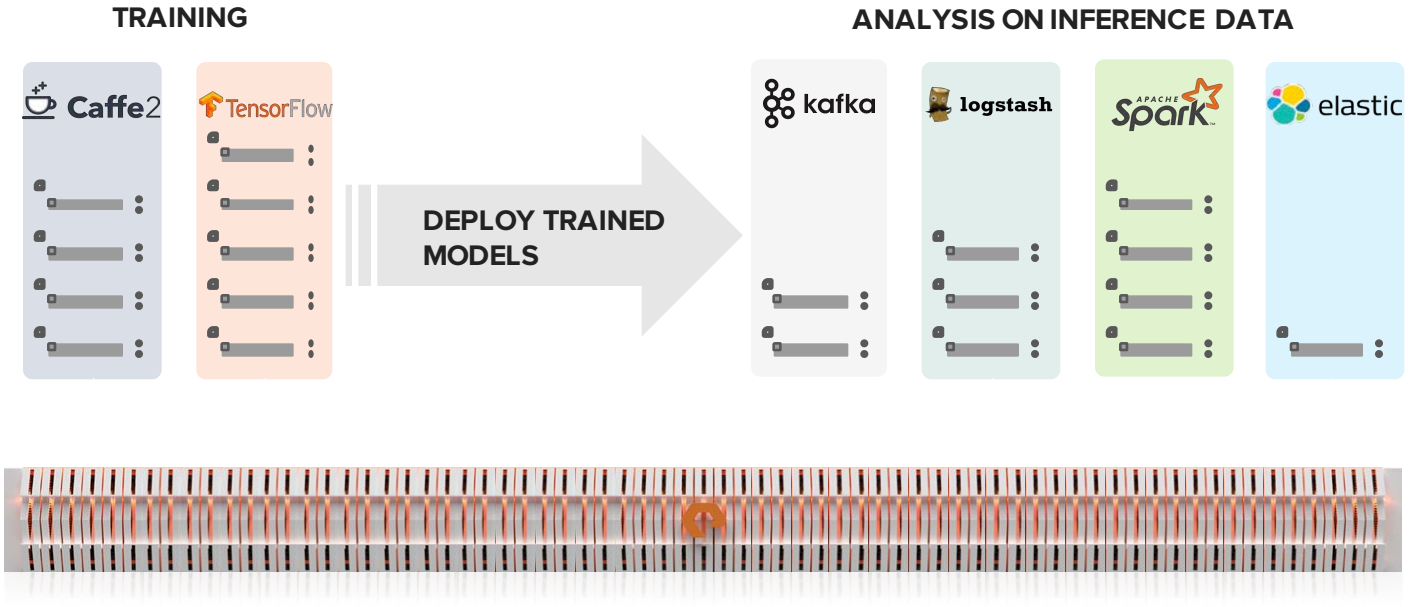
TENSORFLOW TRAINING BENCHMARK WITH RESNET-50



© 2017 PURESTORAGE INC.

ANALYTICS FOR PRODUCTION DATA

"TUNED FOR EVERYTHING" DATA PLATFORM FOR BOTH TRAINING AND INFERENCING WORKLOADS



© 2017 PURESTORAGE INC.

PURESTORAGE®

ONE OF THE
WORLD'S LEADING AI
SUPERCOMPUTERS

WORLD'S FASTEST
RACE CARS

WORLD'S FASTEST
WEBSCALE DATABASE
RESTORES

WORLD'S MOST
ADVANCED PLANES,
TRAINS, AUTOMOBILES,
AND ROCKETS

Make it real.

MAKING AUTONOMOUS CARS POSSIBLE BY 2021

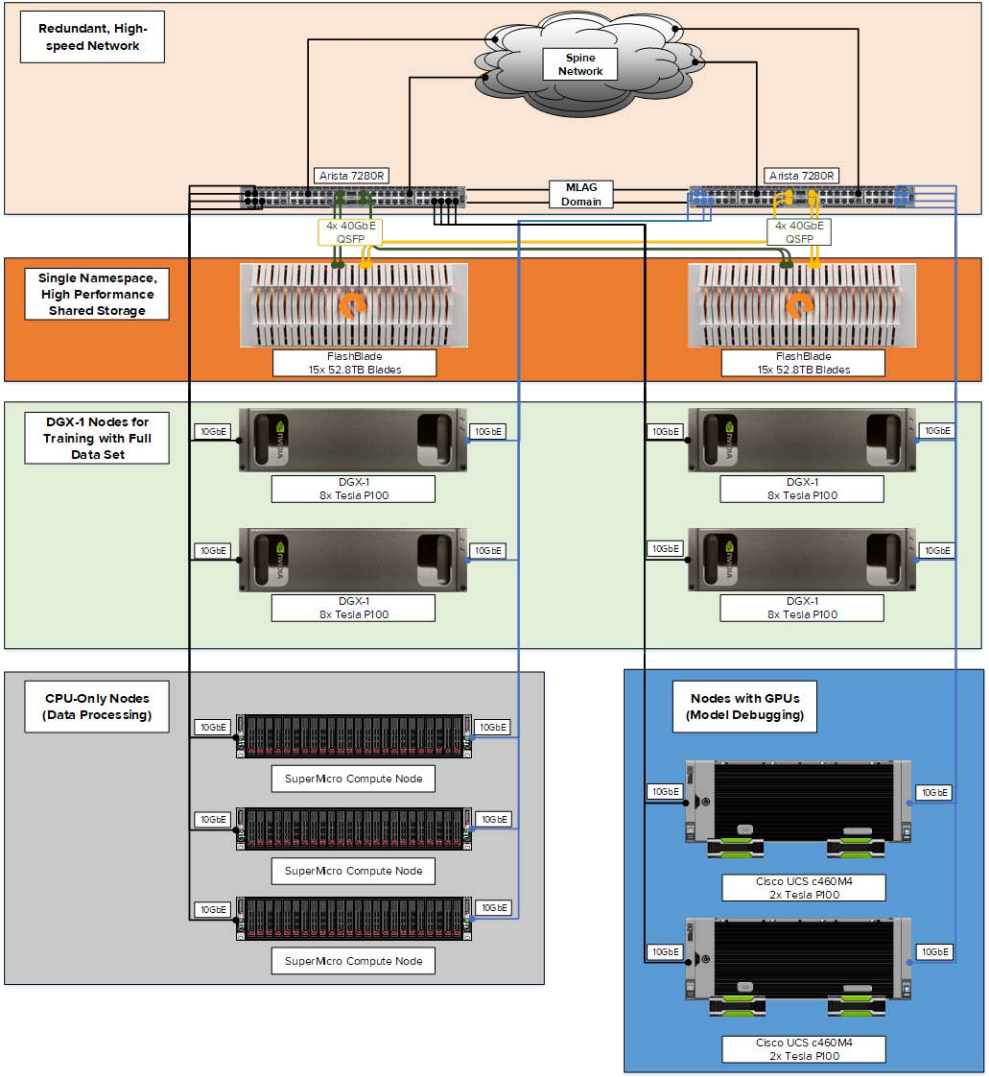
Zenuity, a joint venture of Volvo and Autoliv, aims to build autonomous driving software for production vehicles by 2021. They chose to build their deep learning infrastructure with NVIDIA DGX-1 servers and Pure FlashBlade systems to accelerate their AI initiative.

AUTONOMOUS VEHICLE SOFTWARE COMPANY

4 NVIDIA DGX-1
DL Training Cluster

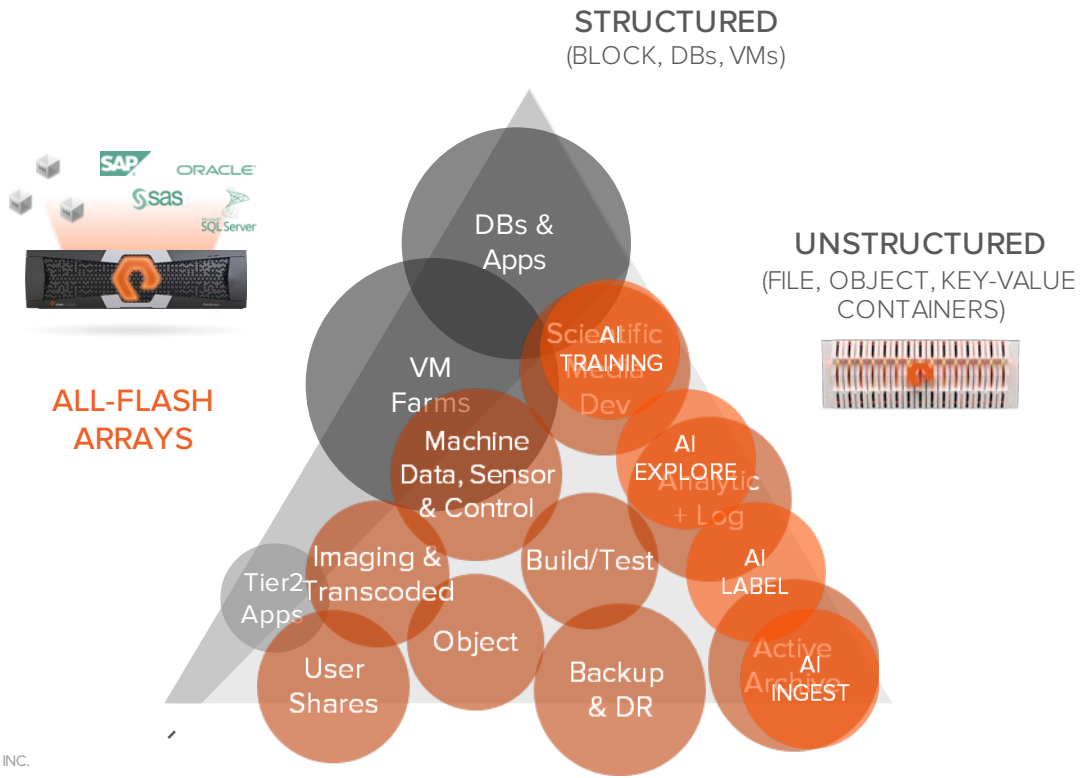
2 PURE FLASHBLADE
Over 1 PB of training data w/ performance headroom

ENTIRE AI PIPELINE ON A SINGLE HUB
Preprocessing, Exploring, and Training on FlashBlade

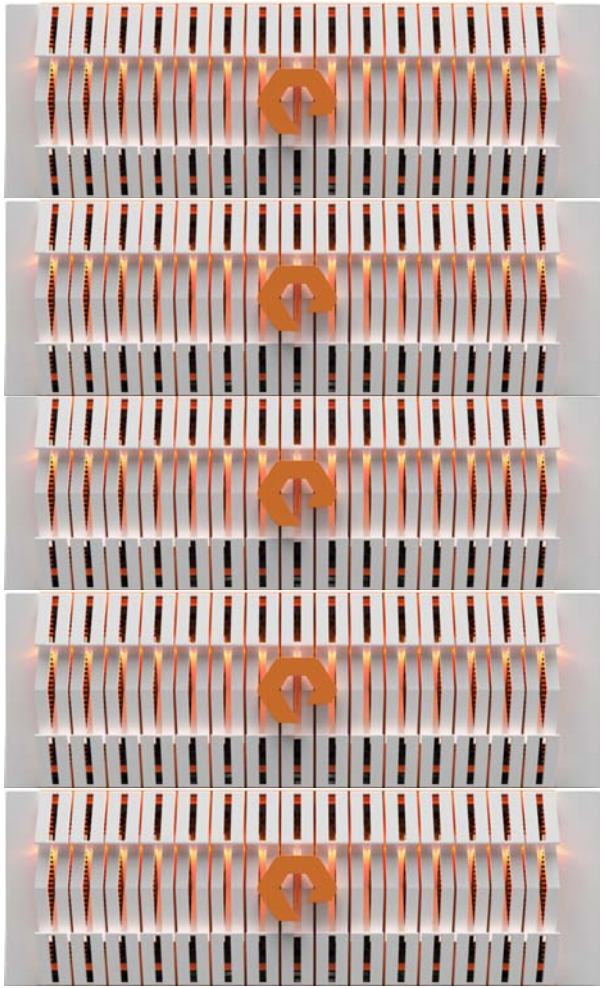


© 2017 PURESTORAGE INC.

AI EXPANDED OUR VIEW OF THE WORLD



© 2017 PURESTORAGE INC.
PURESTORAGE®



FLASHBLADE

INDUSTRY'S FIRST DATA HUB PURPOSE-BUILT FOR AI & DEEP LEARNING

BLADE

Powerful, Elastic Data Processing & Storage Unit

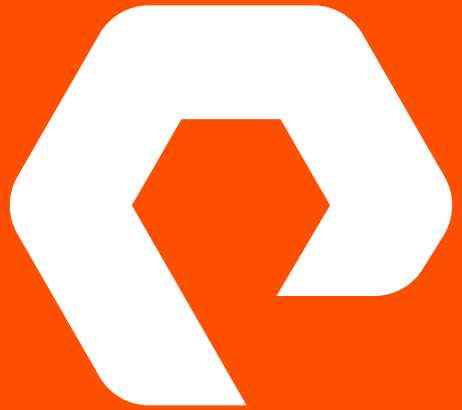
FB 2

PURITY

Massively Distributed Software for Limitless Scale

SCALE-OUT FABRIC

Software-defined fabric that scales linearly with more data & clients

PURESTORAGE®

75 blade feature is subject to GA release