



# HOW TO BUILD A DATA HUB

*FOR THE UNKNOWN IN MODERN DATA*

# LEGAL DISCLAIMER

© 2017 Pure Storage, Inc. All Rights Reserved. Pure Storage and the Pure Storage Logo are trademarks or registered trademarks of Pure Storage, Inc. or its affiliates in the U.S. and other countries. Other names may be trademarks of their respective owners.

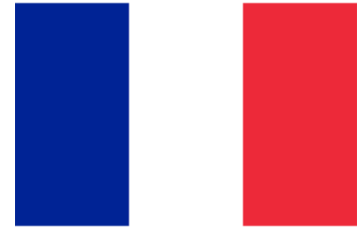
This presentation contains forward-looking statements regarding Pure Storage's business roadmap. These forward-looking statements are subject to known and unknown risks, uncertainties and other factors that may cause our actual results, levels of activity, performance or achievements to differ materially from results expressed or implied in this presentation. Such risk factors include those related to: general economic conditions; maintaining customer and partner relationships; the anticipated growth of certain market segments, particularly with regard to storage; the competitive environment in the software and hardware industry; changes to operating systems and product strategy by vendors of operating systems; fluctuations in currency exchange rates; the timing and market acceptance of new product releases and upgrades; the successful development of new products and integration of acquired businesses, and the degree to which these products and businesses gain market acceptance. Actual results may differ materially from those contained in this presentation. We assume no obligation, and do not intend, to update these forward-looking statements as a result of future events or developments.







# Official Languages Act (1969/1988)



**Translation  
Bureau**



## 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, 2016 5:14 PM ET | Last Updated: Feb 04, 2016 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 compare 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 Neale 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
- Postcards
- Twitter
- Alerts
- Newsletter

THE LAND REVER RANGES FROM \$60'S \$27 MONTHLY WITH \$-VR SERVICING\* 3.9% FIN. RATE ABOVE & BEYOND

### Weather

Ottawa (Kananaskis - Ottawa)	Gatineau	Pelissier	Corwall	Kingston
20°C	20°C	18°C	n/a	18°C

More Weather |

### Latest Ottawa News Headlines

- Ottawa police plans to add hires 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-

Industry News

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

by *Marion Marking* on October 28, 2016

- Facebook
- Twitter
- LinkedIn



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.

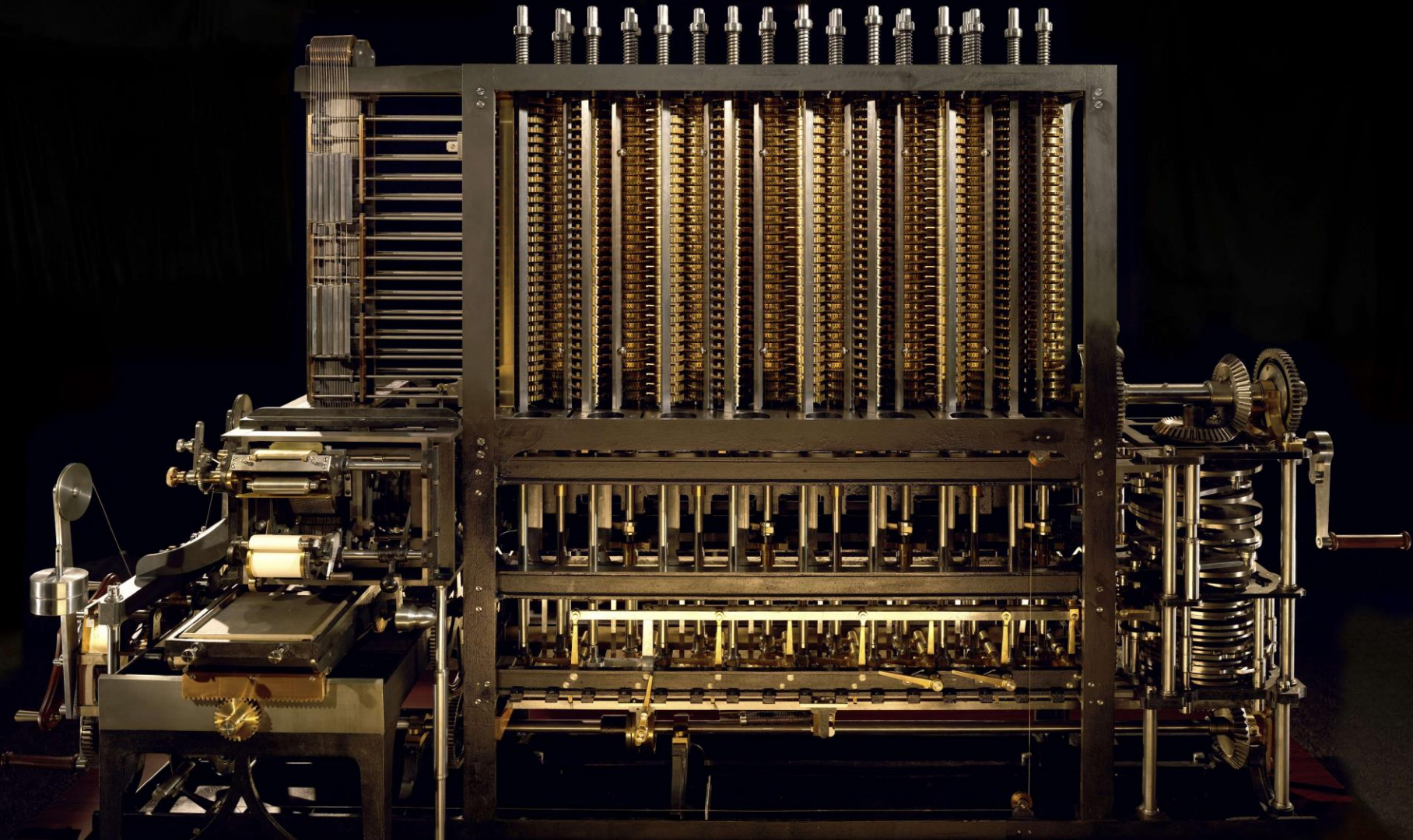
ADVERTISEMENT

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.





# Analytical Computing


$$1 + 1 = 2$$



# Analytical Computing

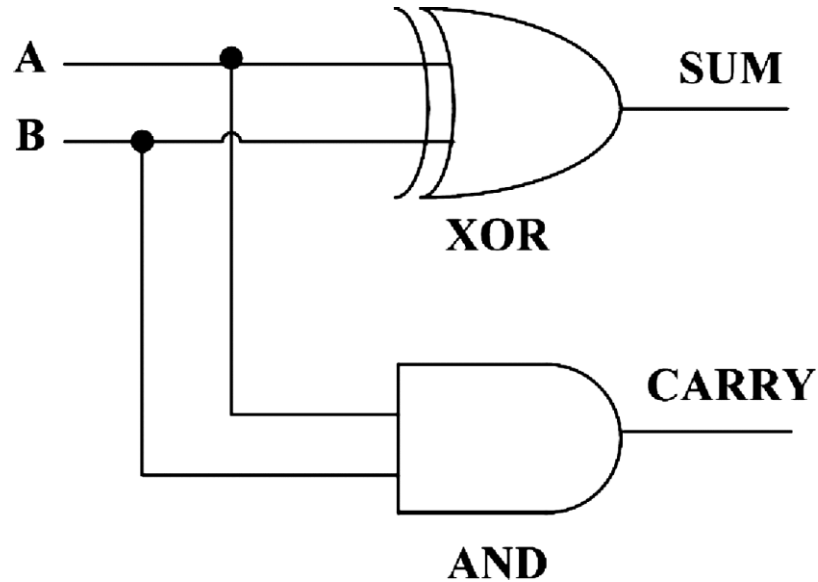
$$1 + 1 = 2$$

Algorithm



# Analytical Computing

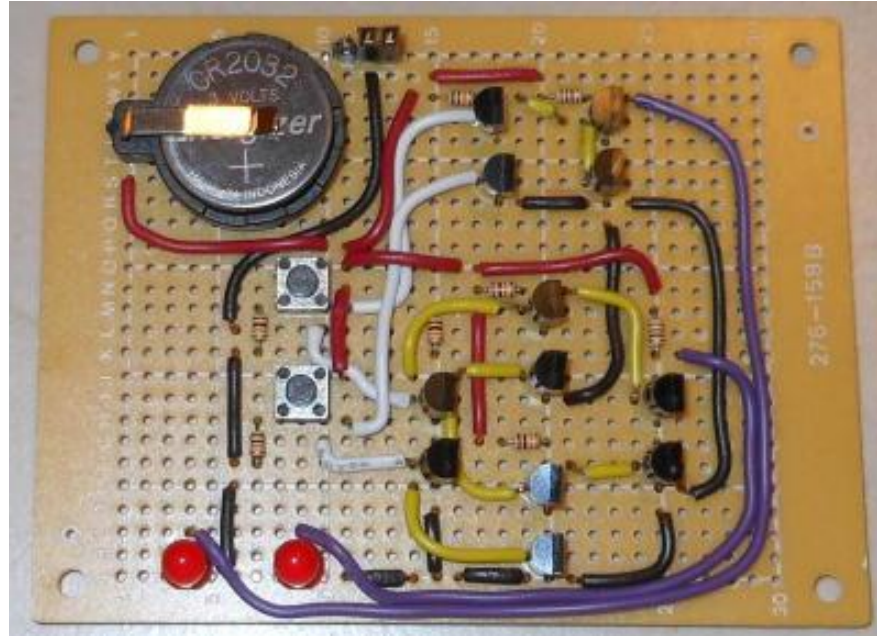
*Logically*



Source: Drexel University

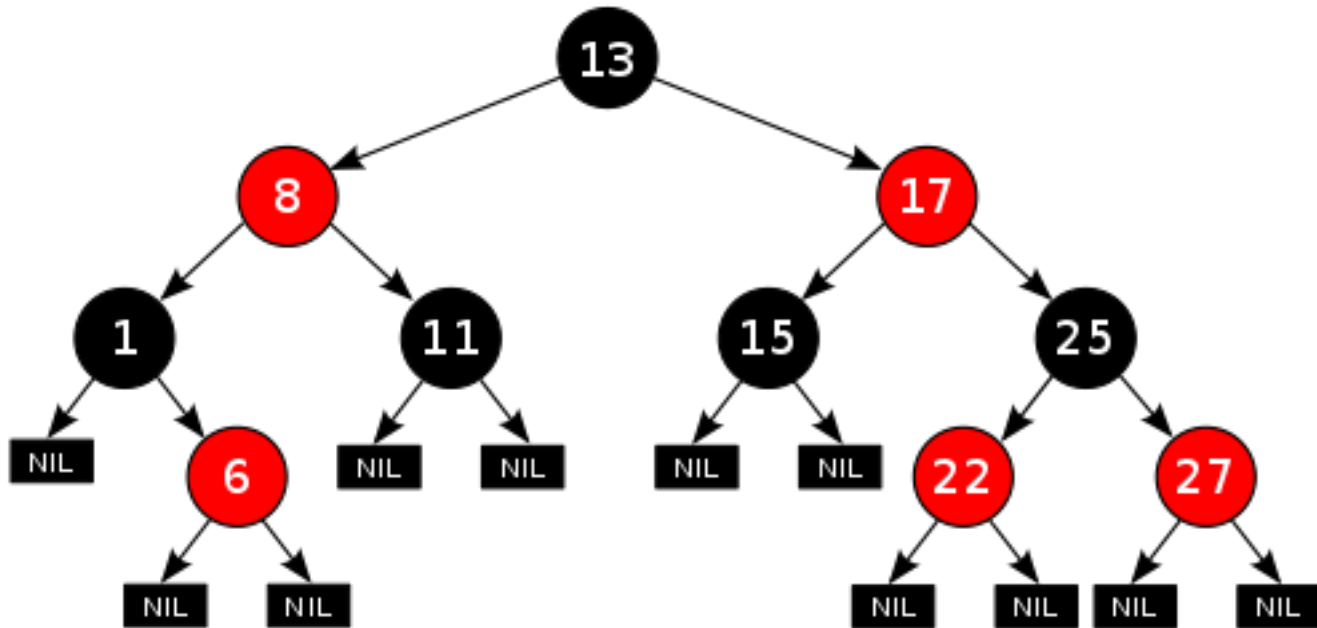
# Analytical Computing

## *Transistors*

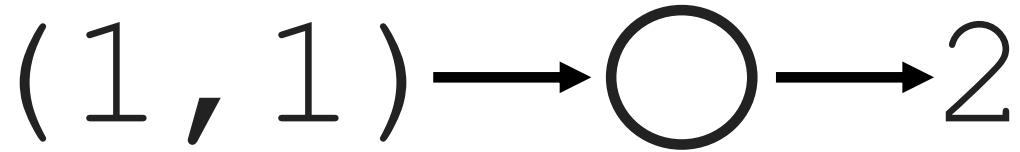


Source: Michael Kohn

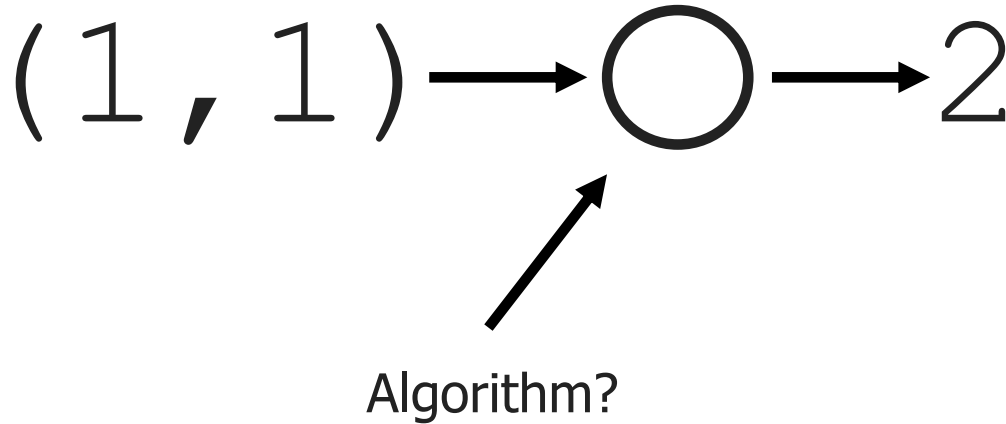




# Empirical Computing

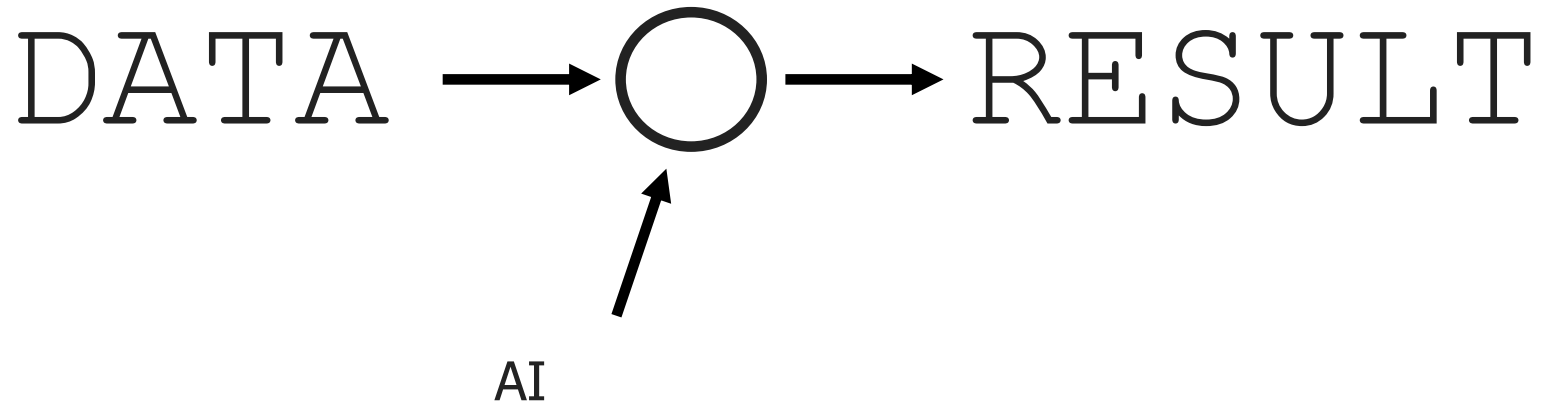


# Empirical Computing





# Empirical Computing



DATA

# DAWN OF 4<sup>TH</sup> INDUSTRIAL REVOLUTION

## BIG DATA, AI DRIVING CHANGE IN EVERY INDUSTRY



### 1<sup>st</sup> Revolution

1760-1820's  
Steam Power  
Rural to Industrial



### 2<sup>nd</sup> Revolution

1870-1914  
Electricity  
Industrial to Mass  
Production



### 3<sup>rd</sup> Revolution

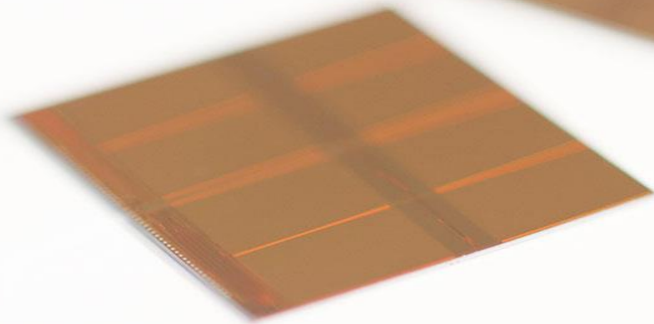
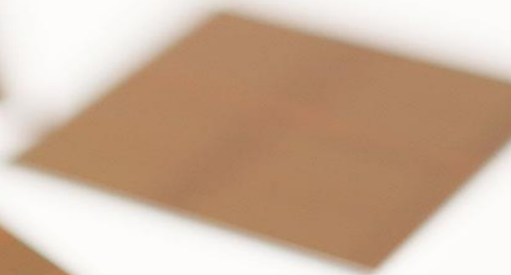
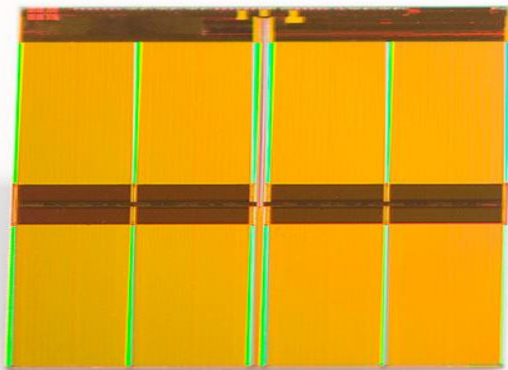
1980-2010  
PC  
Mass production to Digital



### 4<sup>th</sup> Revolution

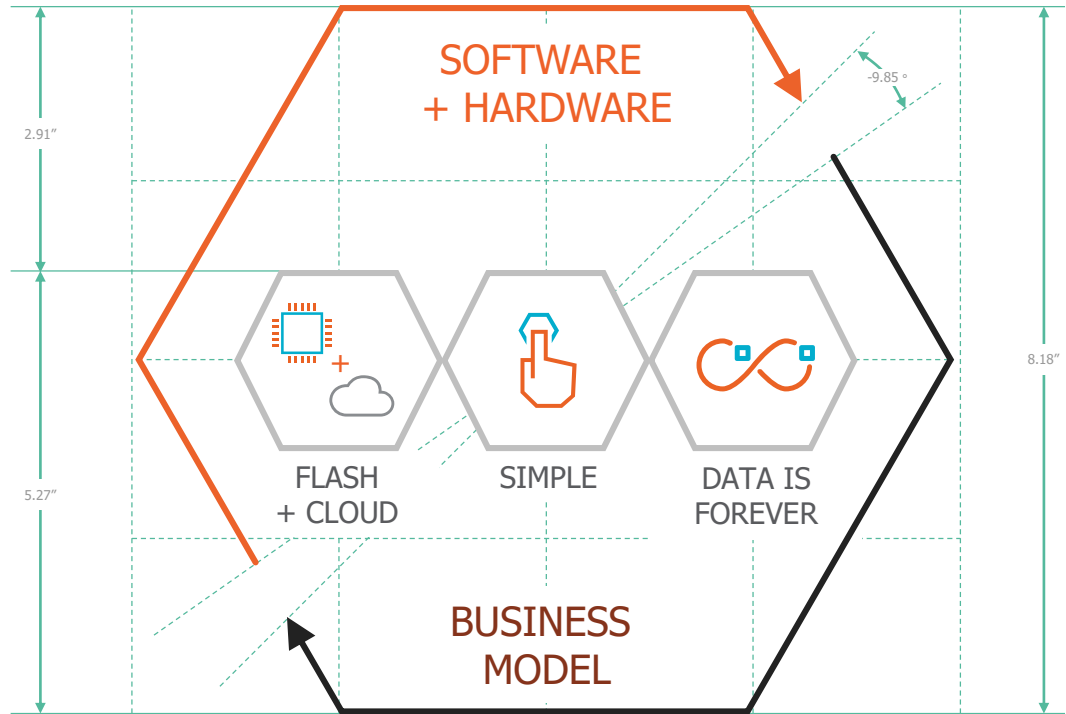
2010-now  
AI, Big Data & IoT  
Digital to Intelligence



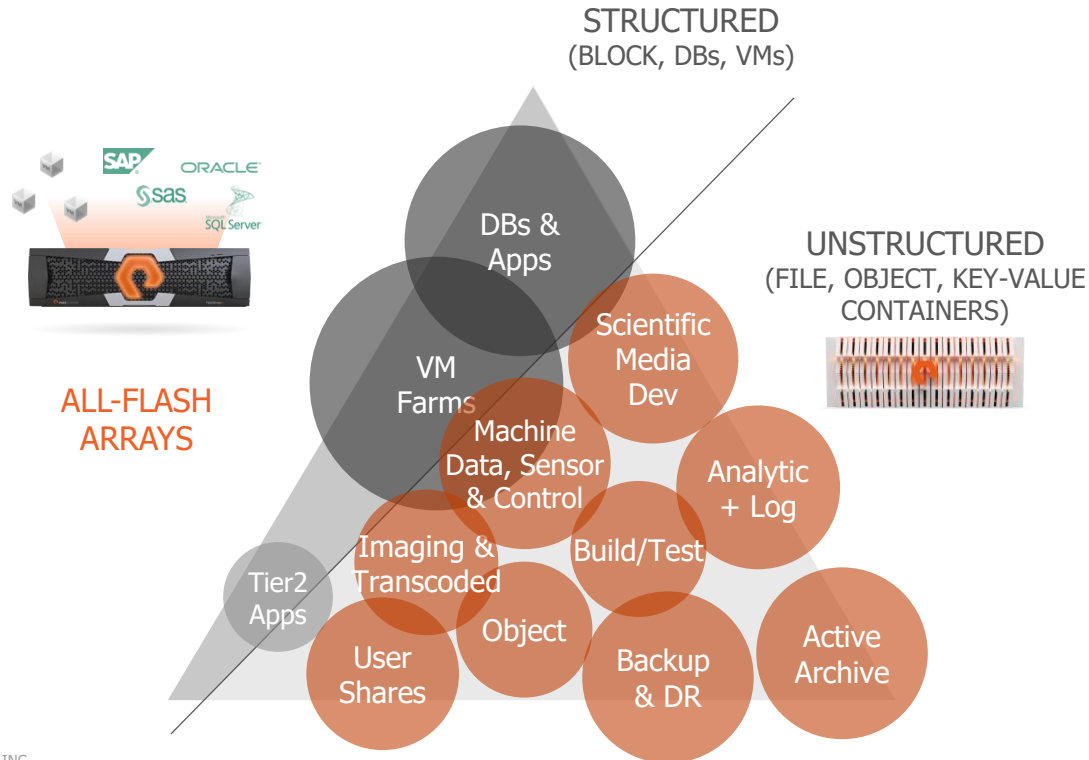


# The Pure Storage Vision

PLATFORM FOR  
**STORAGE  
INNOVATION**

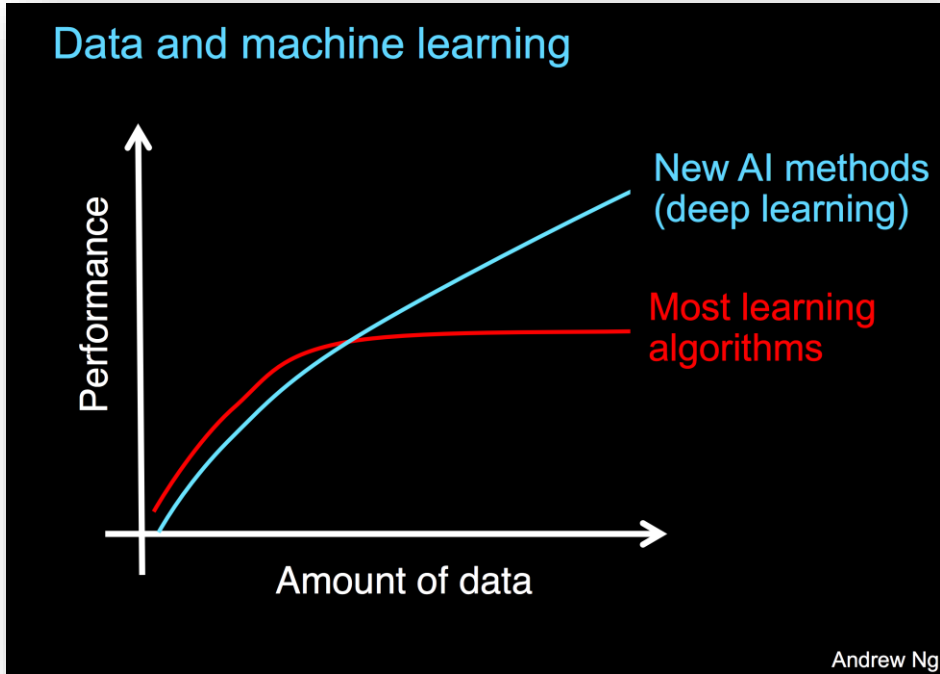


# OUR VIEW OF THE WORLD 4 YEARS AGO

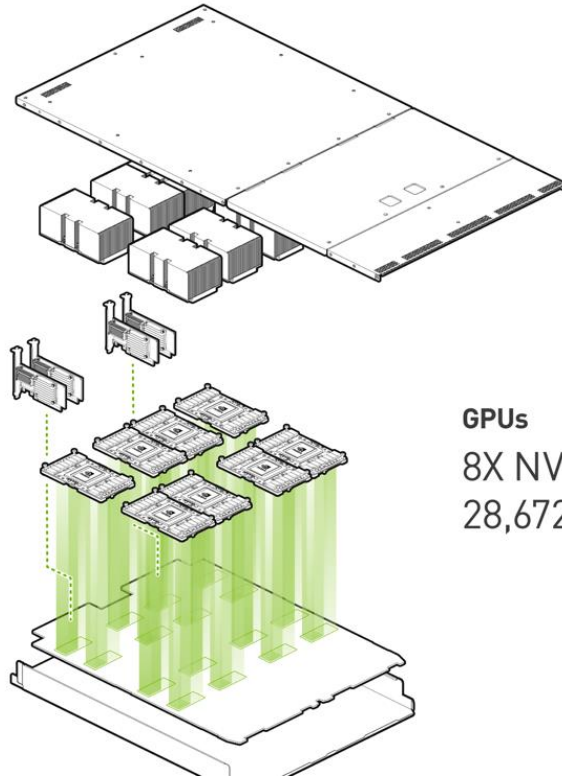


# DATA IS VITAL TO MACHINE LEARNING

OBSERVATION BY PROF. ANDREW NG, AI LUMINARY



# SOUL OF DGX-1 IS PARALLEL



## NETWORK INTERCONNECT

4X InfiniBand™ 100 Gbps EDR  
2X 10GbE

## GPUs

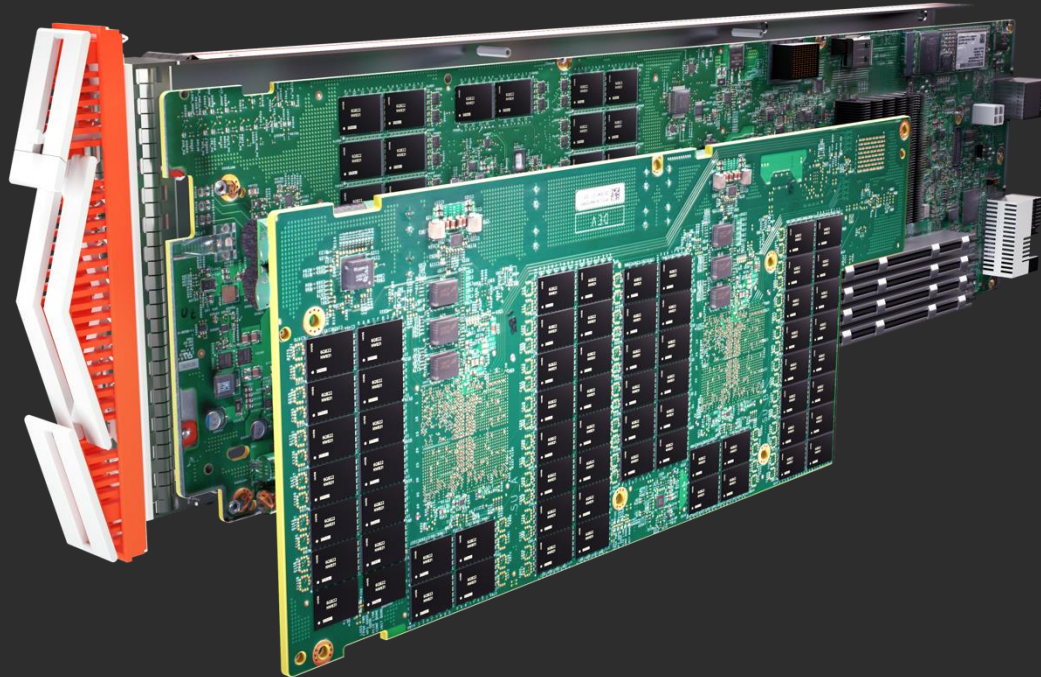
8X NVIDIA Tesla® P100 16GB/GPU  
28,672 Total NVIDIA CUDA® Cores

## GPU INTERCONNECT

NVIDIA NVLink™  
Hybrid Cube Mesh

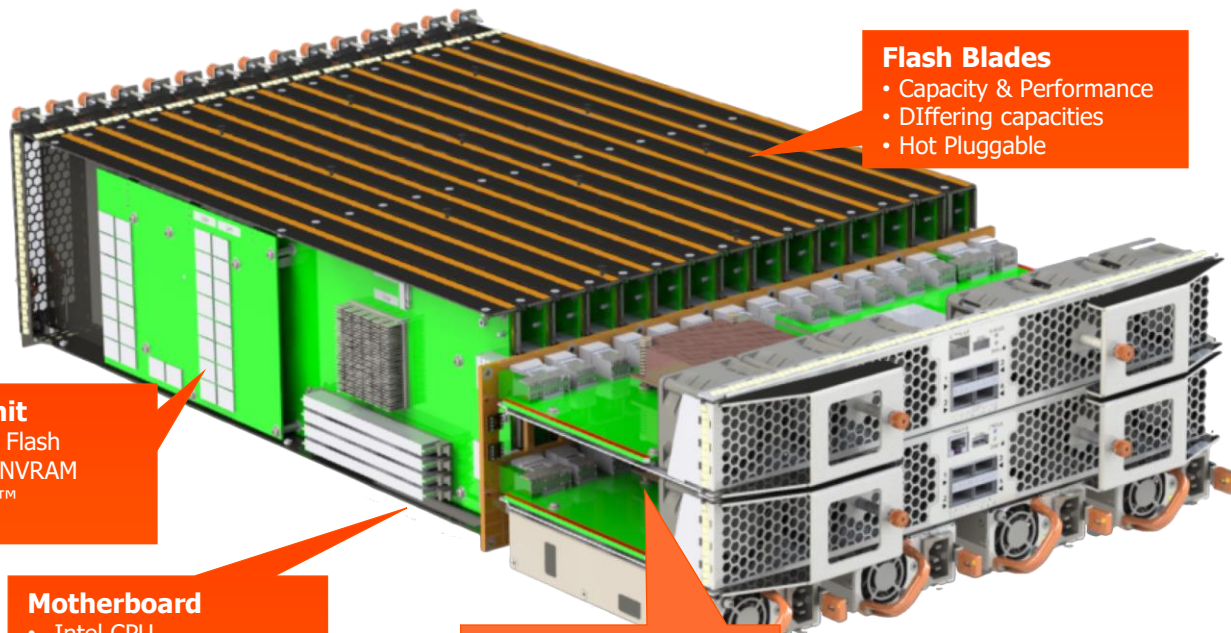
IN 2016 WE PIONEERED A NEW APPROACH WITH FLASHBLADE

DirectFlash





# MASSIVELY PARALLELIZED STORAGE



**Flash Blades**

- Capacity & Performance
- Differing capacities
- Hot Pluggable

**Storage Unit**

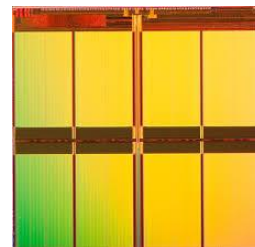
- Up 52TB of Flash
- Embedded NVRAM
- DirectFlash™

**Motherboard**

- Intel CPU
- Dual 10Gb ports (SoC)
- Scale-out Object Store

**Fabric**

- Dual redundant switches
- Cluster communications
- External connectivity

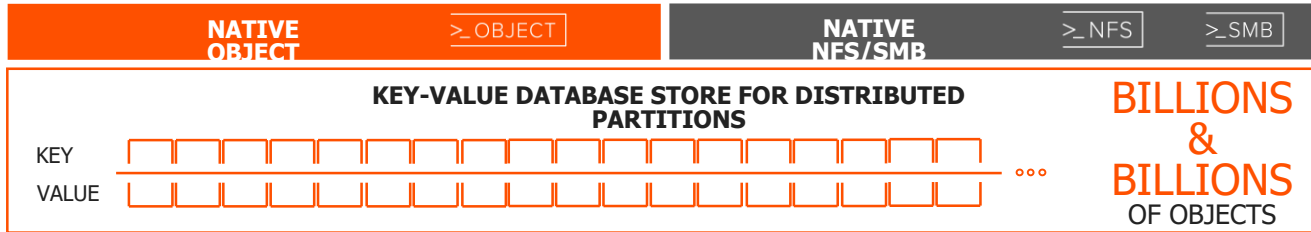


**12,480 NAND Die**

# SOUL OF FLASHBLADE IS PARALLEL

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

Purity **FB** 2



# THREE ESSENTIAL THINGS FOR AI

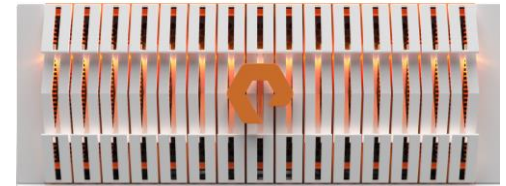
## FRAMEWORKS & APPLICATIONS



## COMPUTE FROM CPU TO GPU SERVERS

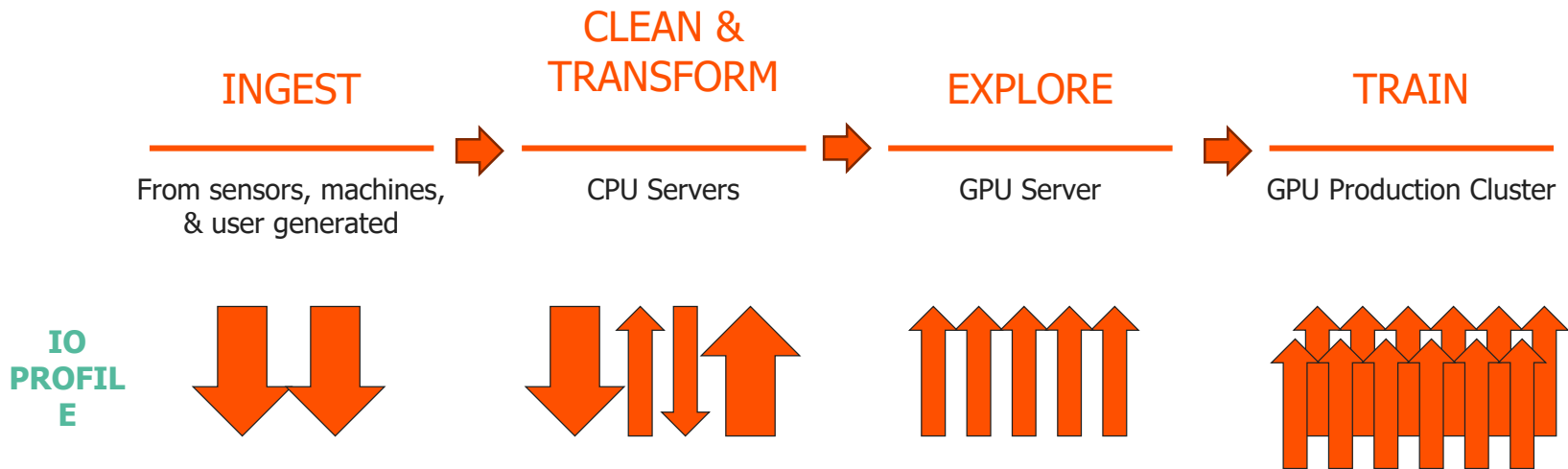


## STORAGE POWER ENTIRE AI PIPELINE

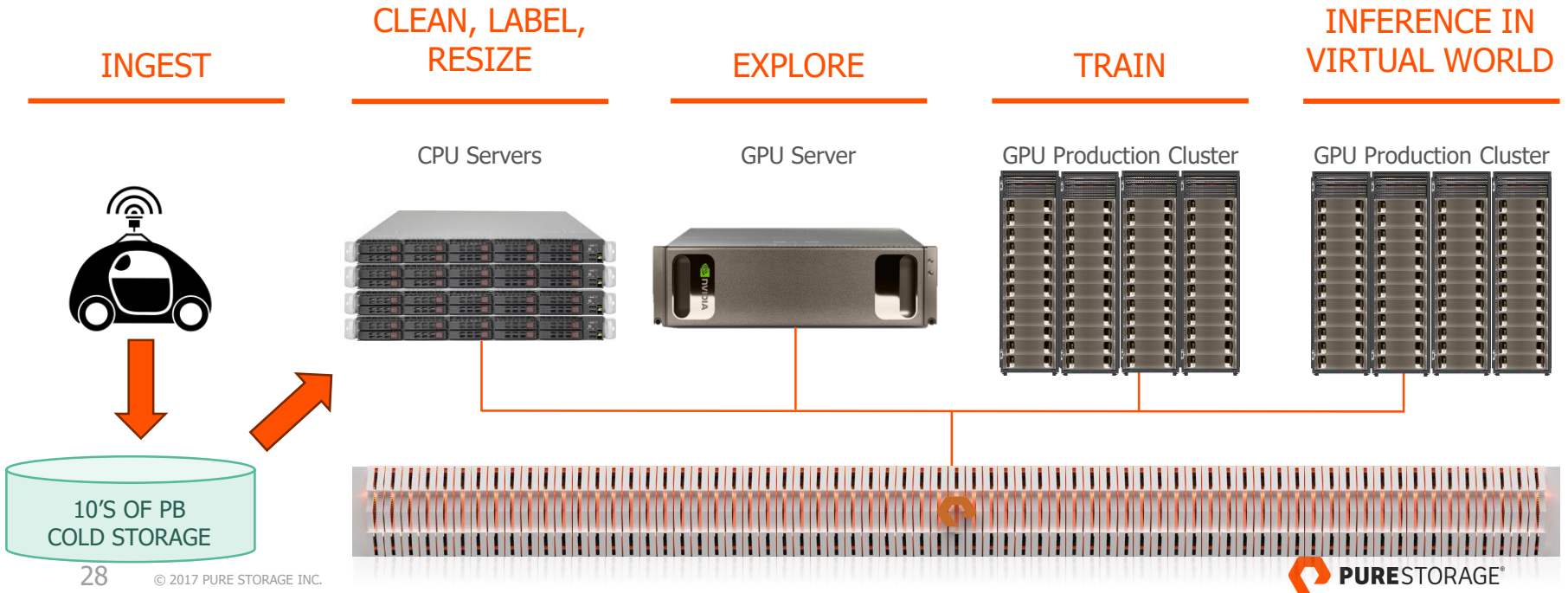


# WIDE RANGE OF NEEDS IN THE PIPELINE

SIGNIFICANT CHALLENGE TO LEGACY STORAGE



# REAL WORLD PIPELINE IN AN AUTONOMOUS CAR COMPANY



# AI TRAINING SYSTEM

GOAL IS TO KEEP THE GPU<sub>s</sub> 100% BUSY

## FULL TRAINING WORKFLOW



## BENCHMARK SETUP

### GPU ONLY



Setup #1: Synthetic Data from System RAM into GPU<sub>s</sub>

### I/O + CPU + GPU



Setup #2: Real Image Data from FlashBlade into DGX-1



# AI SYSTEMS DESIGN PATTERNS

HOW MUCH PERFORMANCE PENALTY DUE TO SHARED STORAGE?

FULL TRAINING  
WORKFLOW



BENCHMARK  
SETUP



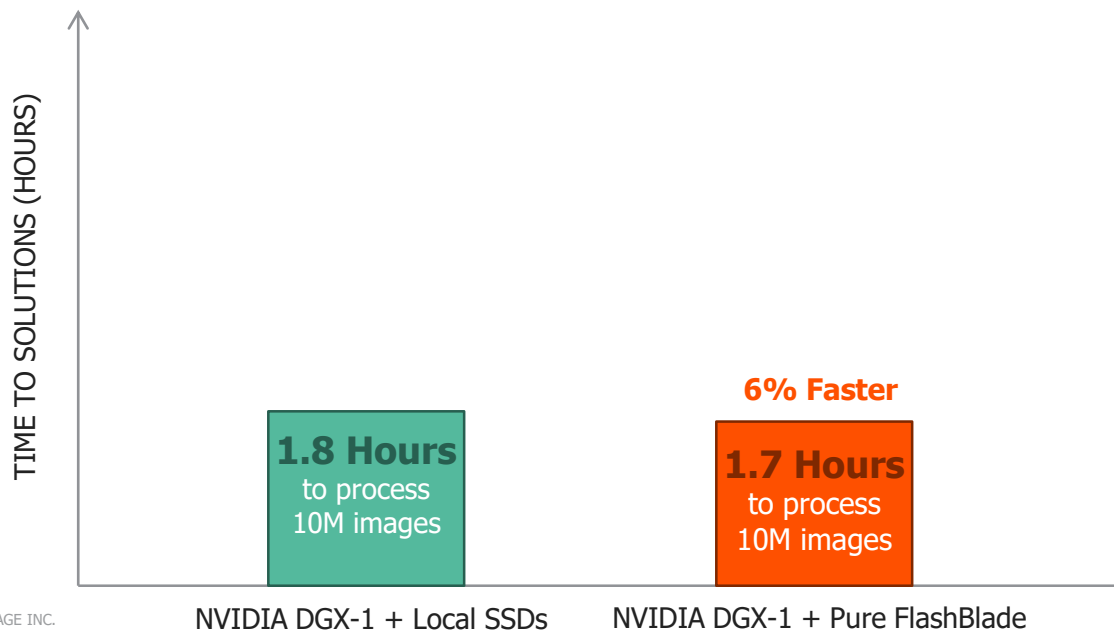
Setup #1: DGX-1 with 4x Local SSDs



Setup #2: DGX-1 with 1x FlashBlade

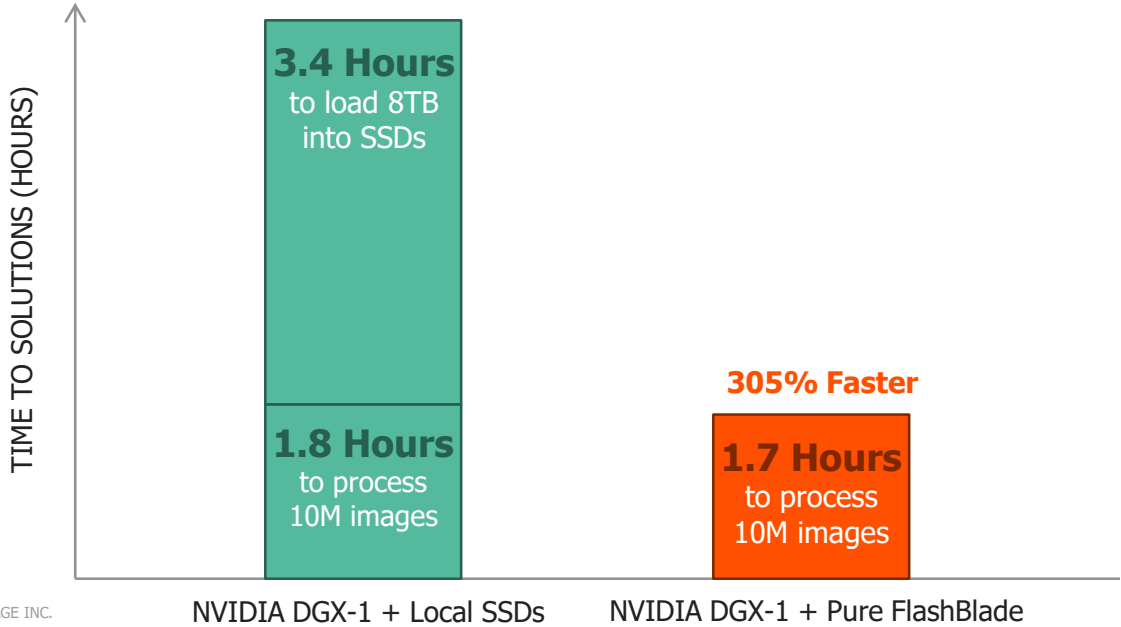
# RESULT: FLASHBLADE vs LOCAL SSDs

## TENSORFLOW TRAINING BENCHMARK WITH RESNET-50



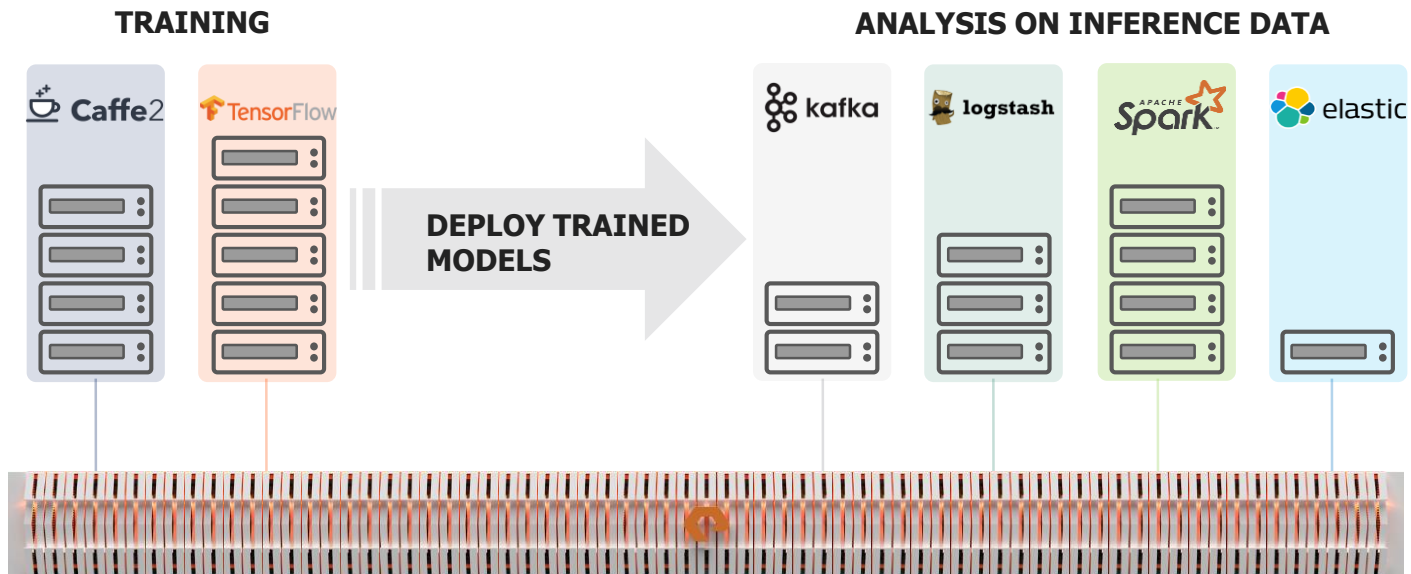
# RESULT: 3X FASTER END-TO-END

## TENSORFLOW TRAINING BENCHMARK WITH RESNET-50

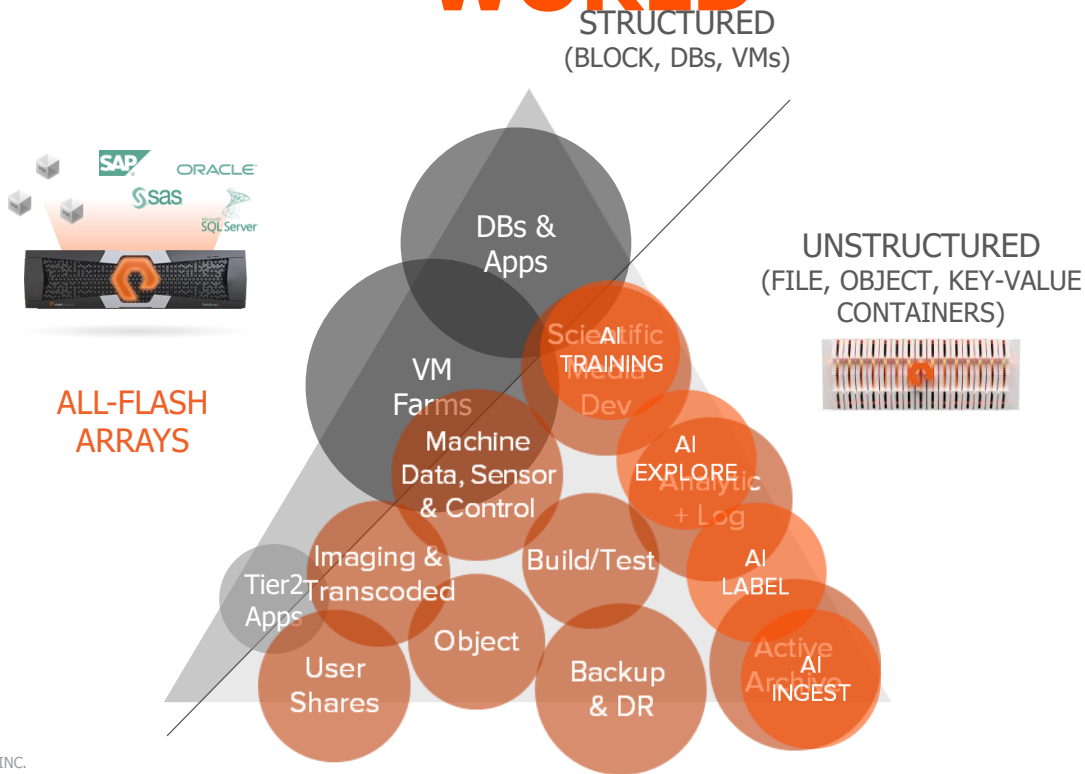


# ANALYTICS FOR PRODUCTION DATA

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



# AI EXPANDED OUR VIEW OF THE WORLD



**The  
Economist**

MAY 6TH-12TH 2017

Theresa May v Brussels

Ten years on: banking after the crisis

South Korea's unfinished revolution

Biology, but without the cells

# The world's most valuable resource



**Data and the new rules  
of competition**



