The Impacts of AI on the Data Center and IT

Scott Tease Executive Director, HPC and AI | October, 2017

Why the Time is NOW for Artificial Intelligence

More Data.

90%

of Fortune 500
companies have
Big Data initiatives

Source: University of Maryland, US

Greater Compute.

1,000,000x

faster; today's mobile vs.
all of NASA computers
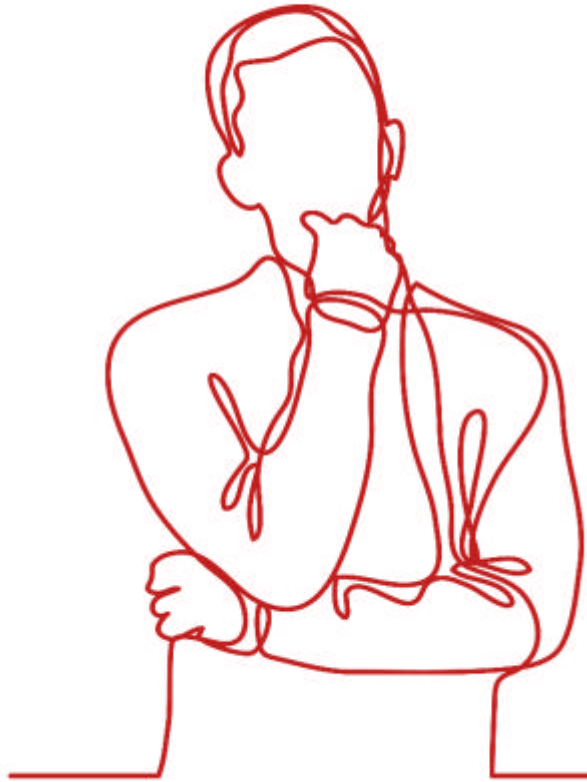from 1969 combined

Source: ZME Science Journal

Enterprise Interest.

$1.2 Trillion

Businesses adopting AI
will capture from less
informed peers by 2020

Source: Forrester

Typical Questions we hear on AI

1. How do I cut thru all the AI noise and understand my options?

2. What AI technology should I bet on and who is doing AI for my segment?

3. Everyone else seems to be deploying AI. Am I getting left behind?

A Better Way to Think of AI

1. AI is an enabler, not a destination

2. AI will be a journey - AI will deliver better efficiency and predictive capability over time

3. AI will bring a massive amount of change in technology in the coming years – openness is critical

4. AI unfolding different than previous evolutions

LENOVO AI VISION

GOOD AT
COMMON SENSE
INTUITION | CREATIVITY
EMPATHY | VERSATILITY

GOOD AT
LARGE DATA SETS
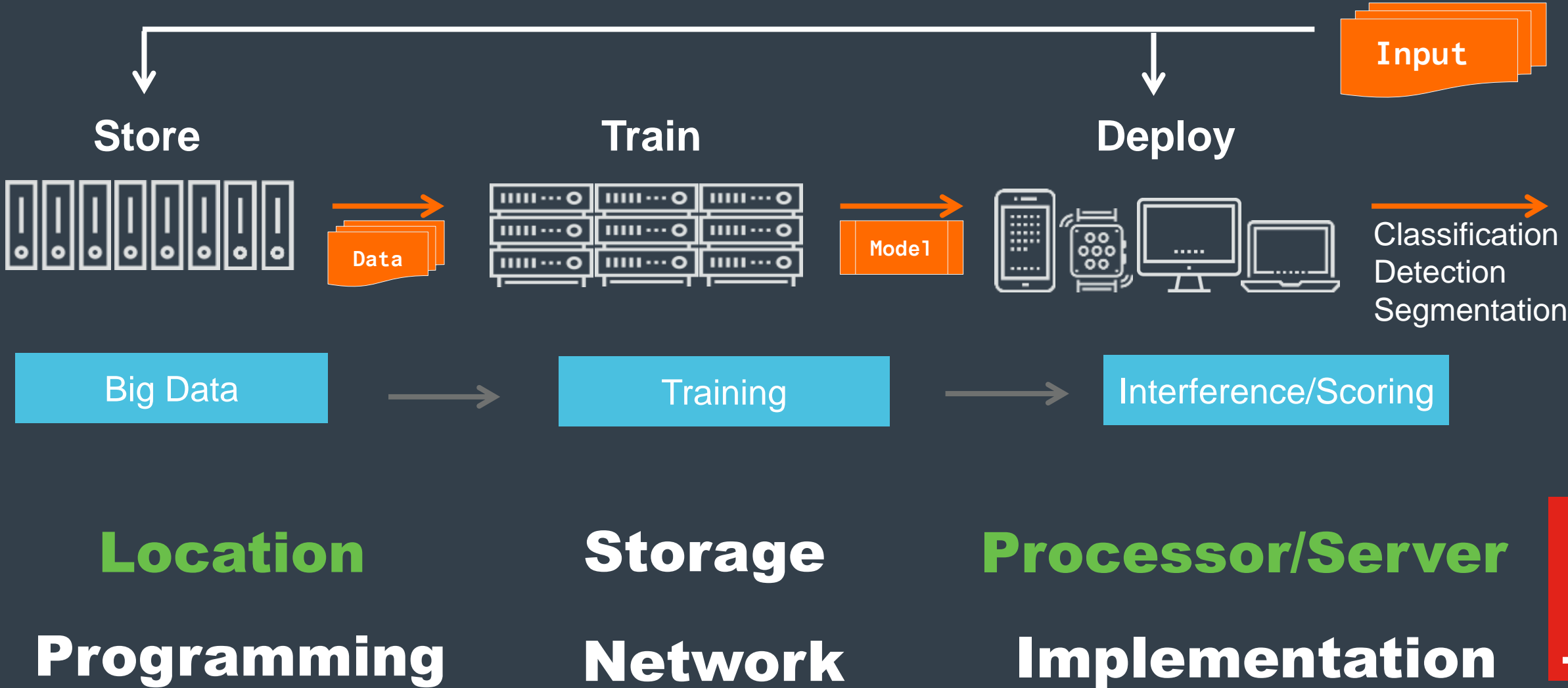COMPLEX CALCULATIONS
CONSTANT LEARNING

LENOVO AI VISION

+

SUPER EXPERTS – AUGMENTED INTELLIGENCE TO SOLVE HUMANITY's MOST COMPLEX CHALLENGES

AI Will Change Our Approach to IT

Store   **Train**   **Deploy**

Input

Data   Model

Classification
Detection
Segmentation

Big Data → Training → Interference/Scoring

Location **Storage** **Processor/Server**

Programming **Network** **Implementation**

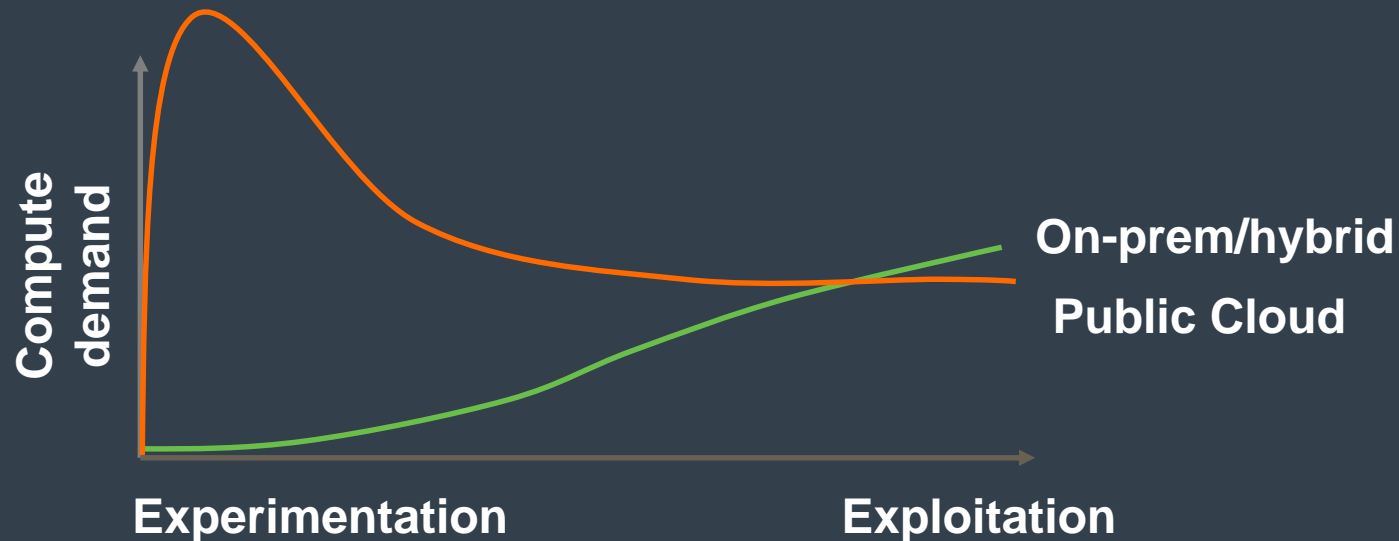# AI and Data Location will Drive Locality of Compute

Data Location
Moving large amounts of data is neither practical nor economical

IT Processes
AI will be integrated into existing business processes

Data Privacy
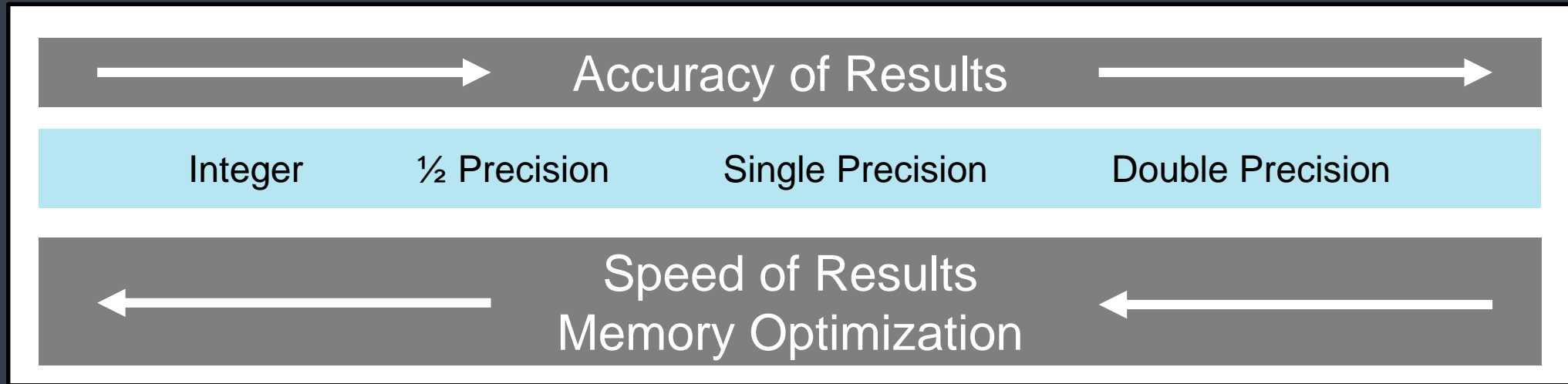Regulations/competitive reasons preclude some movement of data

Compute demand

On-prem/hybrid

Public Cloud

Experimentation

Exploitation

Finance

Healthcare

Manufacturing

Academia

Precision – is Being Close Good Enough?

Accuracy of Results →		→
Integer	½ Precision	Single Precision	Double Precision

Speed of Results
Memory Optimization
← ←

Deep Learning	**HPC**	**Enteprise IT**
Variable precision **As close as possible** **1+1 is about 2**	**Single precision** **Should be accurate** **1+1 should be 2**	**Double precision** **Ensures Accuracy** **1+1=2**
GPU, Nervana, Phi, FPGA	CPU, GPU, Phi	CPU

Will Lower Precision Technologies Win?

64-bit Memory Register

Definitely a Moose

Double

64-bit
Image 1

Single

32-bit **32-bit**
Image 1 Image 2

Accuracy

Half

16-bit **16-bit** **16-bit** **16-bit**
Image 1 Image 1 Image 1 Image 1

Most Likely a Moose

INT8

8-bit
Img 1 Img 2 Img 3 Img 4 Img 5 Img 6 Img 7 Img 8

Lowering the precision = more differentiated data in memory
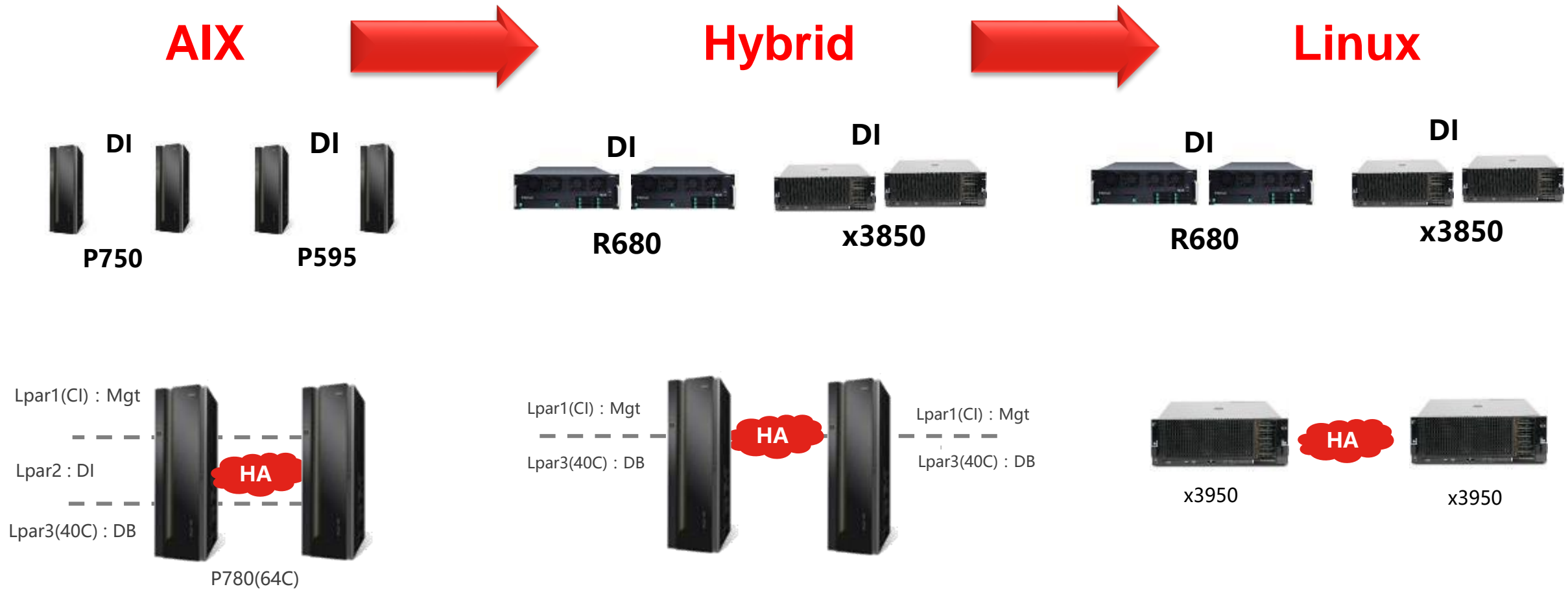
Reduce amount of compute required per task

AI In Use at Lenovo

"The video channel is definitely difficult [to analyze] and images are as well – we wanted to look at YouTube and Instagram so if someone posts a picture of a Lenovo desktop or a laptop that is broken or doing something, we want to see that."

Lenovo SAP ERP Migration Journey

HANA@ Lenovo

2011	**BW on HANA**	237X Performance Improvement
2012	**CRM on HANA**	10X Perf trade promotion/interactive report
2013	**CEO Dashboard**	Mobile, real time analysis, quick decision
2014	**Hybris on HANA**	Optimized B2C platform with holiday support for sales season
2015	**SCM on HANA**	x86 product roadmap planning HANA as platform, enabling self-development
Today…	**HANA platform consolidation**	

First in AP

86% Hardware Saving

50X Improvement of Data Loading

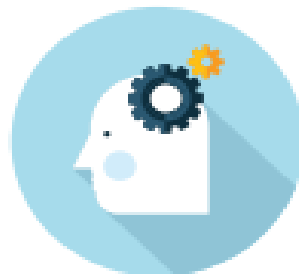# Whats Next For Lenovo with AI and SAP

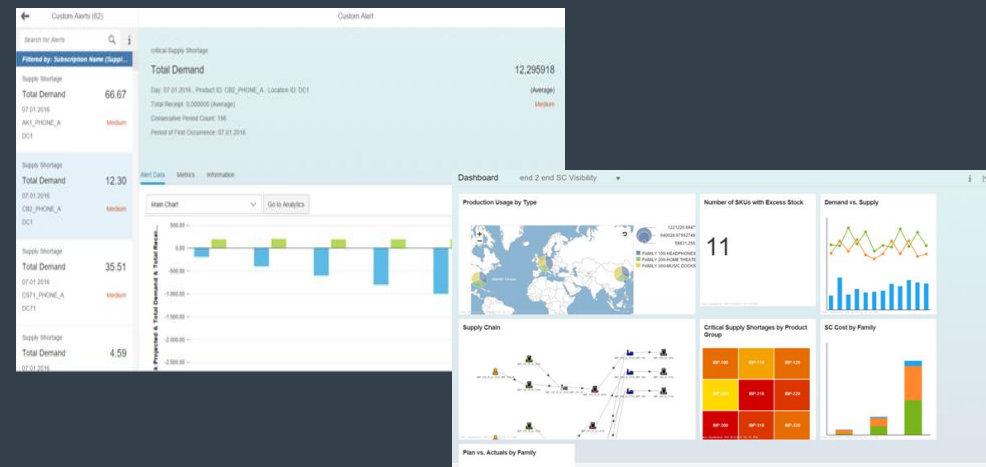**E2E Visibility**   **Analytics**   **Actionable Intelligence**

- Track and trace across silos

- Tell what happened

- What's happening now?
- What is at risk?
- Why is this happening?

- Fix situation prior to impacting business and customers
- Reduce decision making latency
- Continues improvement

Real-time insight:

o **Business KPI**

o **Pre-Alert**

o **Close loop mgmt**

o **View from**

anywhere

BARRIERS TO AI ADOPTION

✓ **DATA (STRUCTURED / UNSTRUCTURED)**

✓ **AI EXPERTS AND CONSULTANCY**

✓ **AI HW / SW RESOURCES**

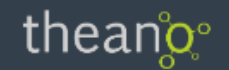# ANNOUNCING: LENOVO AI INNOVATION CENTERS

INFRASTRUCTURE

ThinkSystem

PARTNERS

RESEARCH INSTITUTIONS

TECHNOLOGY PARTNERS

ECOSYSTEM PARTNERS

FRAMEWORKS

Caffe2   torch   theano   mxnet

MORRISVILLE   STUTTGART   BEIJING

Lenovo