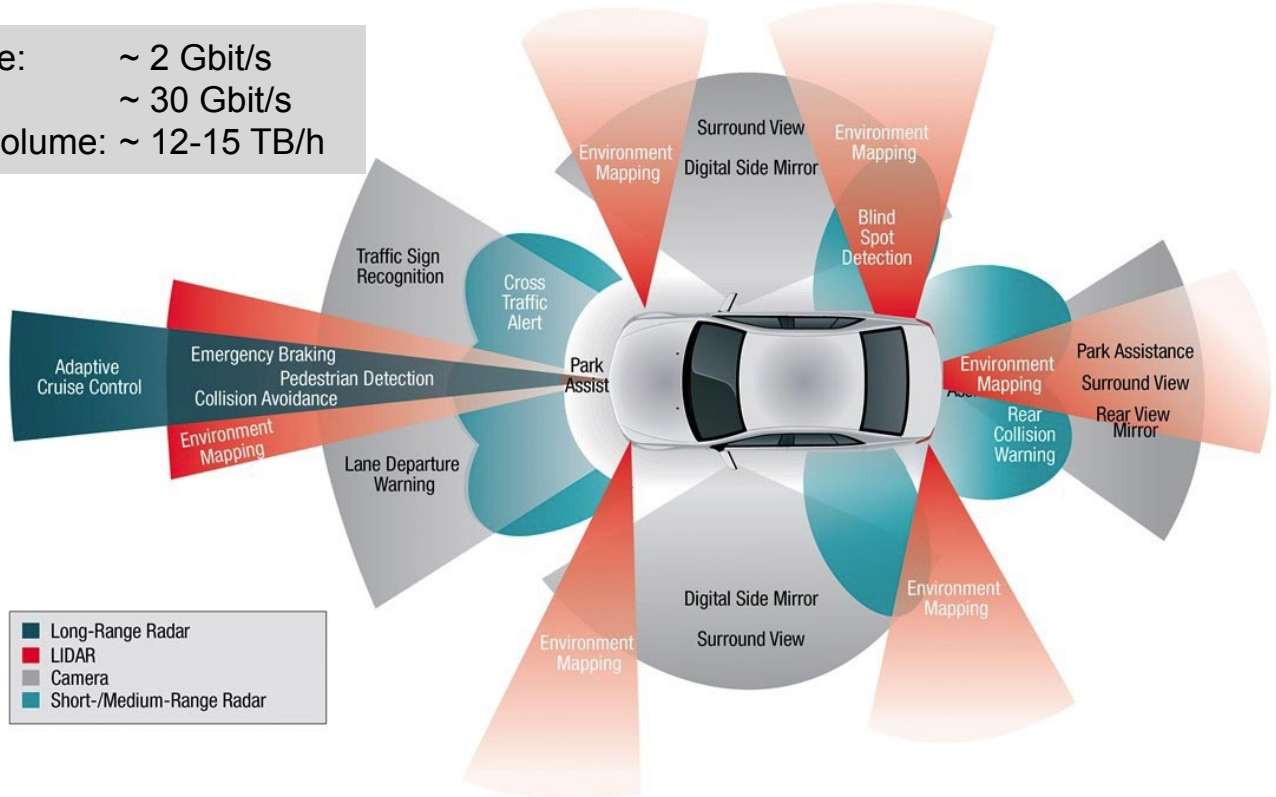
Autonomous Transportation – Phase 2

Clodoaldo Barrera

IBM Storage Systems

Automotive Sensor Setup for AD

Each data source: ~ 2 Gbit/s
Sensors sets: ~ 30 Gbit/s
Data collection volume: ~ 12-15 TB/h



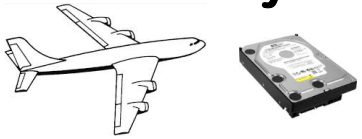
http://currencyobserver.com/2017/12/global-automotive-sensors-market-2017-2022/

Data Management for AV Projects is a challenge

Test Drives

- Europe
- USA
- China
- Japan
- Asia
- Africa

50TB / day / car

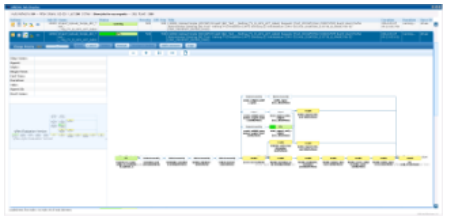


R&D Labs: tagging

R&D Labs: developing & testing & (re-)simulation & AI training

> 5PB / car model (project)

Major IT Challenges for ADAS



1. How to implement & operate an efficient storage, workflow and management system?
2. How to distribute data globally within an enterprise and partners?

3. How to preserve digital data for decades with optimized costs?

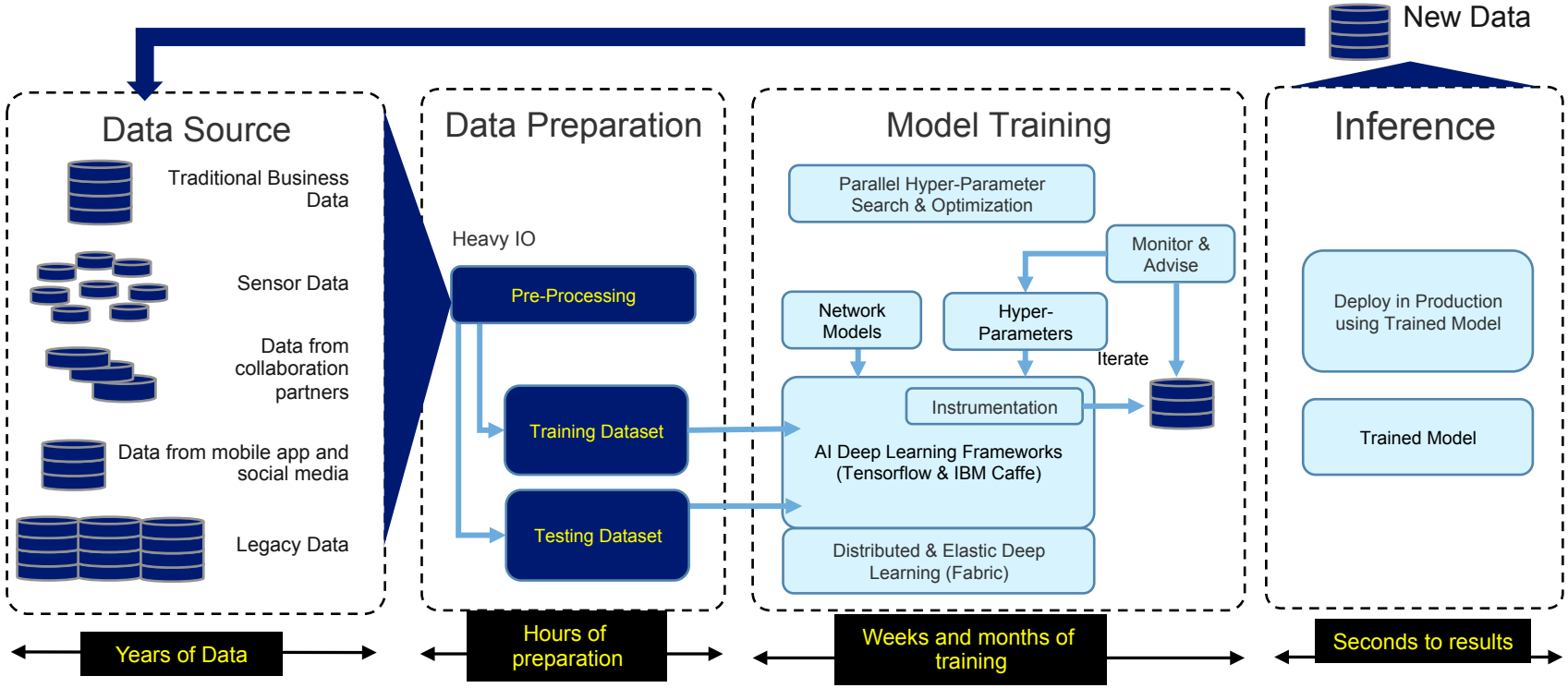
4. How to analyze sensor and video data with fast analytics and modern BigData tools?
5. How to run Machine Learning (ML) and AI training with Nvidia GPU technology at scale?

6. How to do efficient IT workload and resource scheduling?

7. How to embed analytics/data management into R&D Environment?
8. How to run massive workloads on large topology Clusters with data centric workloads?

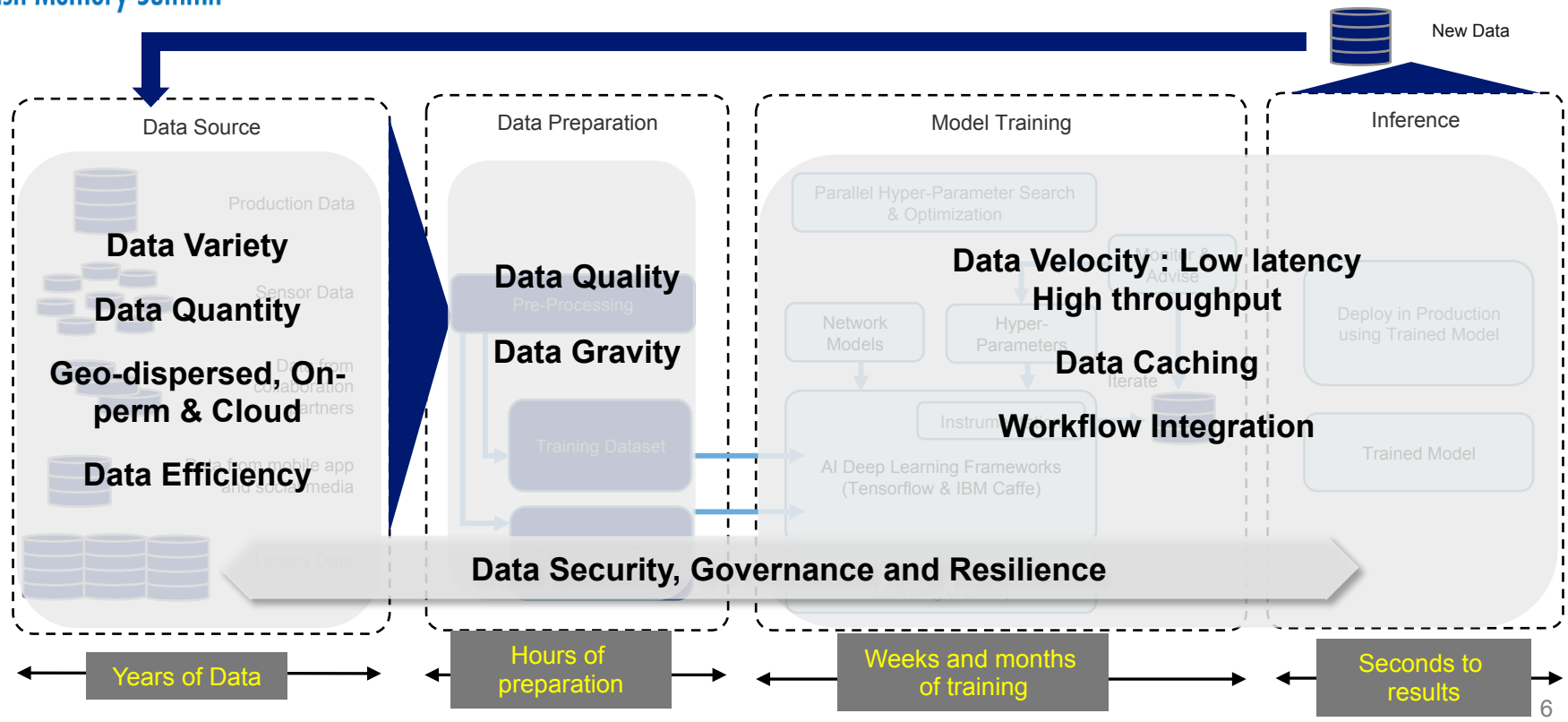
Workload and data flow for AI flow is complex

Machine Learning - Workload flow and data flow



5

Data requirements vary significantly



Data architecture for AI workloads

Capacity Tier

Scale-out
To scale capacity and performance linearly

Flexible deployment
To deploy as an appliance, as software only on commodity hardware or as cloud services

Multi-protocol
To house various data types, structured, unstructured, semi-structured

Geo-dispersed
To facilitate collaboration across long distances

Efficient
To store data across various media including Flash, Disk and Tape; to apply data reduction when possible; to store efficiently using Erasure Coding

Data Preparation

High throughput

To curate massive amount of data including read and write operations

To feed multiple GPU's

example, each GPU could require 1-2 GB/sec; assuming 4 GPU's per server, each server could require 4-8 GB/sec

Performance Tier

Low Latency for Inference

To quickly read random and sometimes small data

Distributed accelerated cache for Training

To read the same data over and over again 'reference data'

Secure:
To protect data in-flight and data at risk

Governance:
Managing and curating data through different stages of processing

Resilient:
To ensure data's availability against failures

The Storage Hierarchy

Scale out File on Flash (HOT)

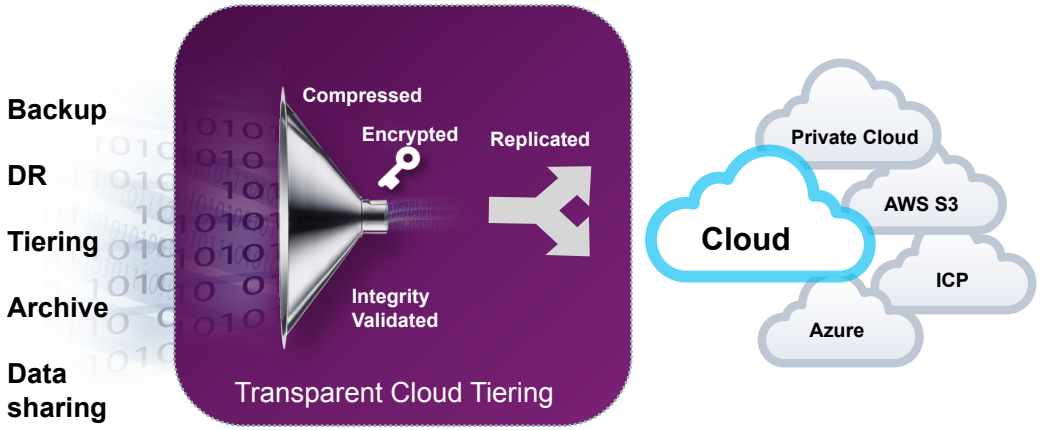
- File based storage with Object & HDFS support
- High End I/O performance
- Information Lifecycle Management (ILM)
- Sub Micro-seconds access time

Cloud Object Storage (S3) (WARM)

- Site Fault Tolerant
- Geo Dispersed and WW scale
- Easy to Deploy
- Milli-seconds access time

Archive & Tape (COLD)

- Lowest TCO
- Tape ILM target – especially frozen archive
- Long term retention and Minutes access time
- Access as files via LTFS
- Reduced floor space requirements and energy consumption
- Up to 260PB native capacity in a single Tape Library



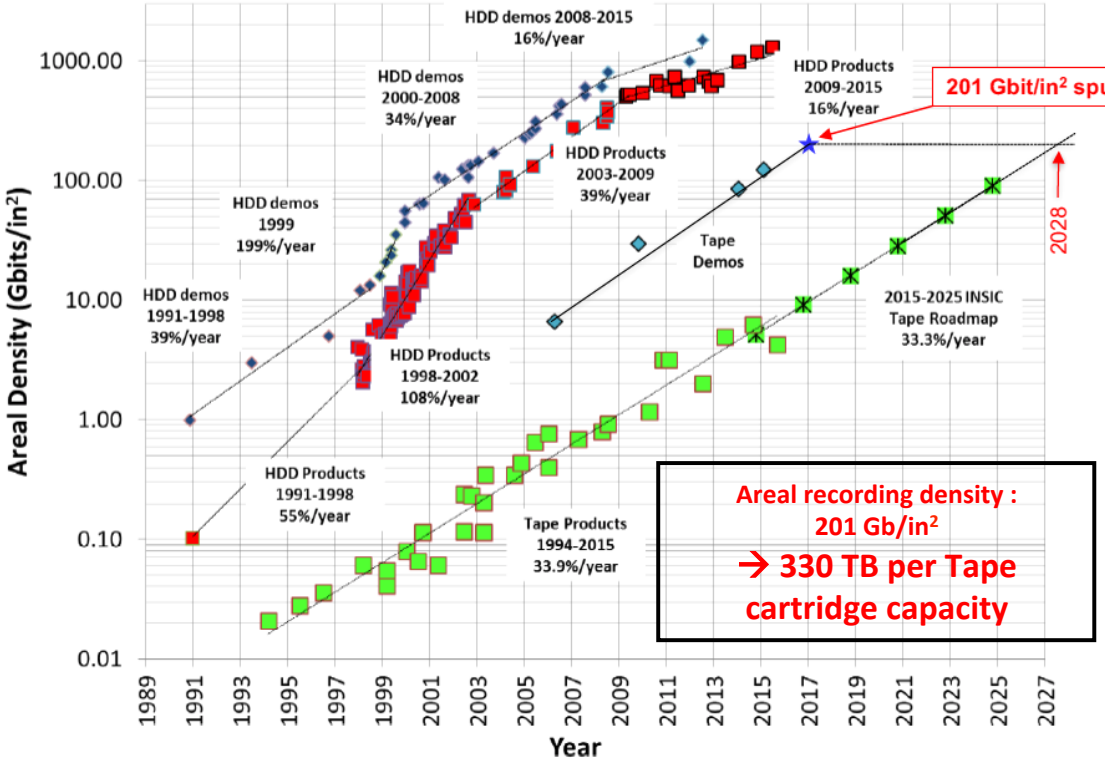
- Tiering from flash, to disk, to tape, to cloud.
- Cloud appears as external storage pool.
- Auto Tiering & migration.
- High performance Read/Write operations.
- Public cloud-ready.
- Support of multi cloud environments.

Trends for Magnetic Recording Areal Density

https://phys.org/news/2017-08-team-magnetic-tape-storagemakes-competitive.html

2015: IBM-FujiFilm demonstration of 123 Gb/in^2 on BaFe tape

2017: IBM-Sony demonstration of 201 Gb/in^2 on Sputtered Tape



It appears that tape will maintain its **5x** to **10x** cost advantage over other storage technologies for at least another decade.

IBM Zürich Research Lab