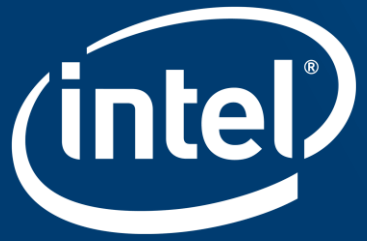


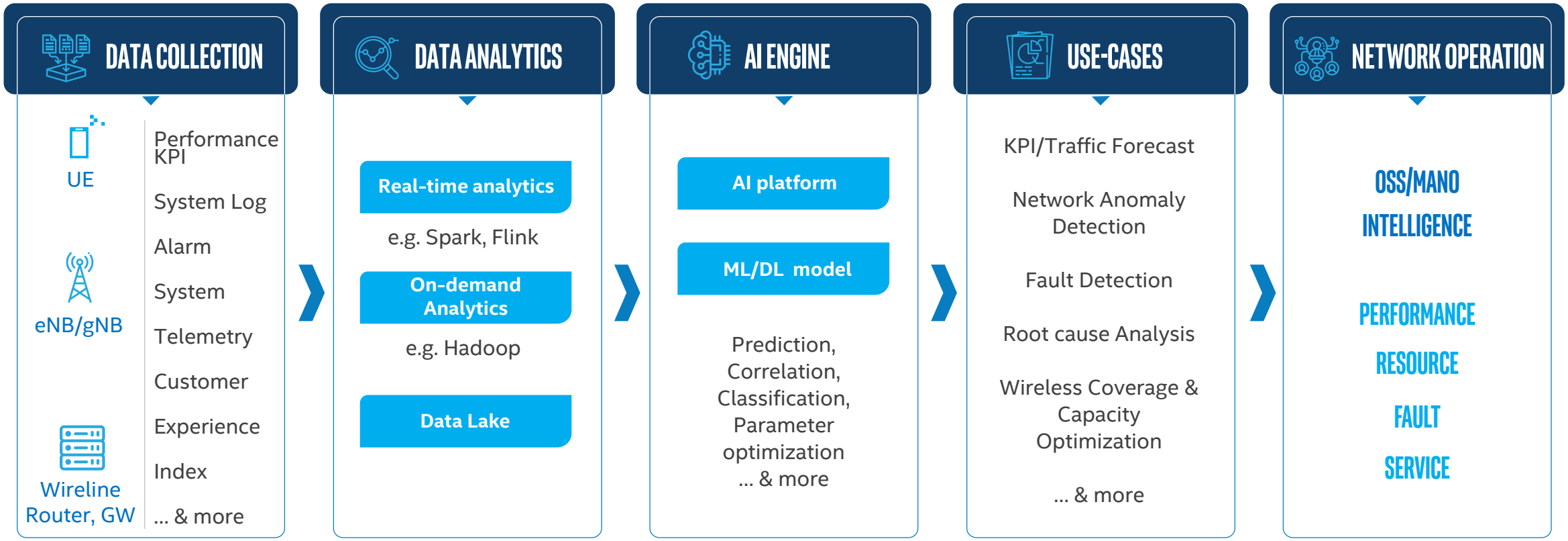
AI NETWORK ANALYTICS

Hongchan(Nate) Roh : Team leader/ Sr. Software Engineer, SK Telecom

Jason Dai : Sr. Principal Engineer, Intel

Huisuk(Kevin) Hong : Sr. Market Development Manager, Intel

AI NETWORK ANALYTICS & INTEL- CUSTOMER COLLABORATION

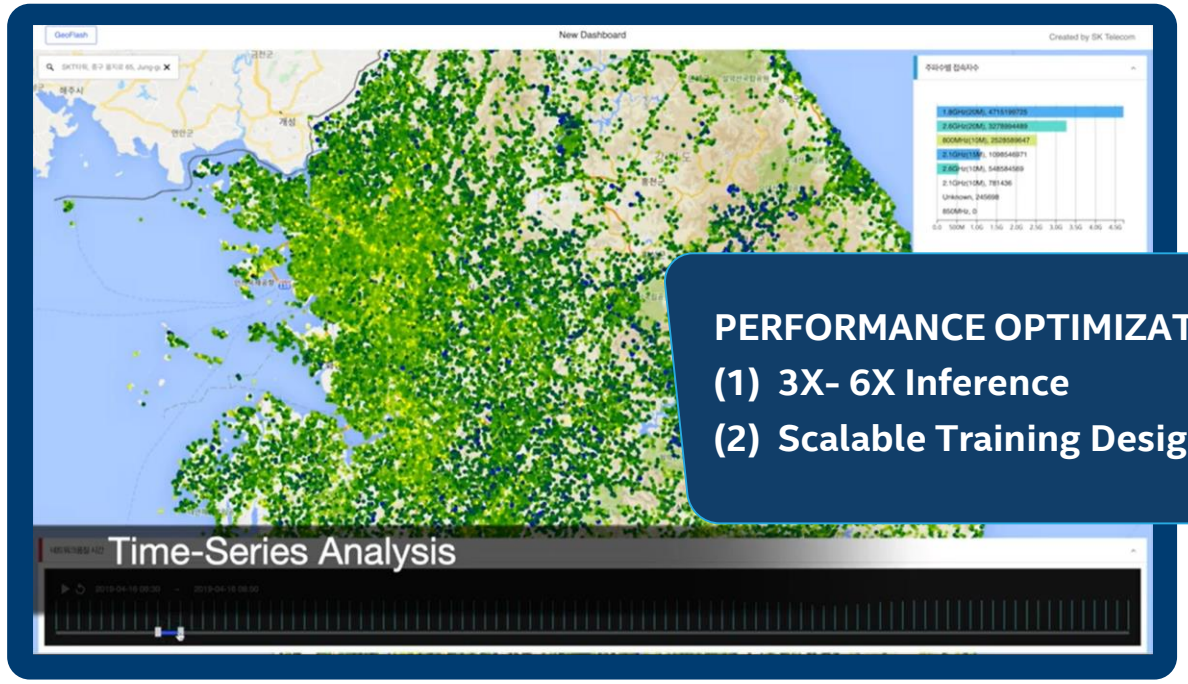


INTEL & CUSTOMER COLLABORATION

Build unified 'Analytics + AI' pipeline
AI Performance Optimization
Use-case Development with Reference AI Model
AI Hardware Platform Validation

SK TELECOM CASE STUDY - INTEL OPTIMIZATION TECHNOLOGIES

Analyze and Predict Wireless Network Quality Indicators
e.g. CQI, RSRP, RSRQ, SINR, and more



PERFORMANCE OPTIMIZATION
(1) 3X- 6X Inference
(2) Scalable Training Design

SK telecom | Lightning DB

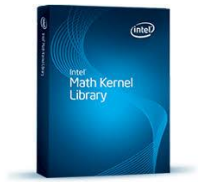
Unified Data Analytics + AI Pipeline

ANALYTICS ZOO

Intel® Optimization for Tensorflow

TensorFlow

Intel® Math Kernel Library for Statistics



2nd Generation Intel® Xeon® Scalable Processors

CQI : Channel Quality Indicator , RSRP : Reference Signal Received Power, RSRQ : Reference Signal Received Quality, SINR :Signal to Interference Noise Ratio

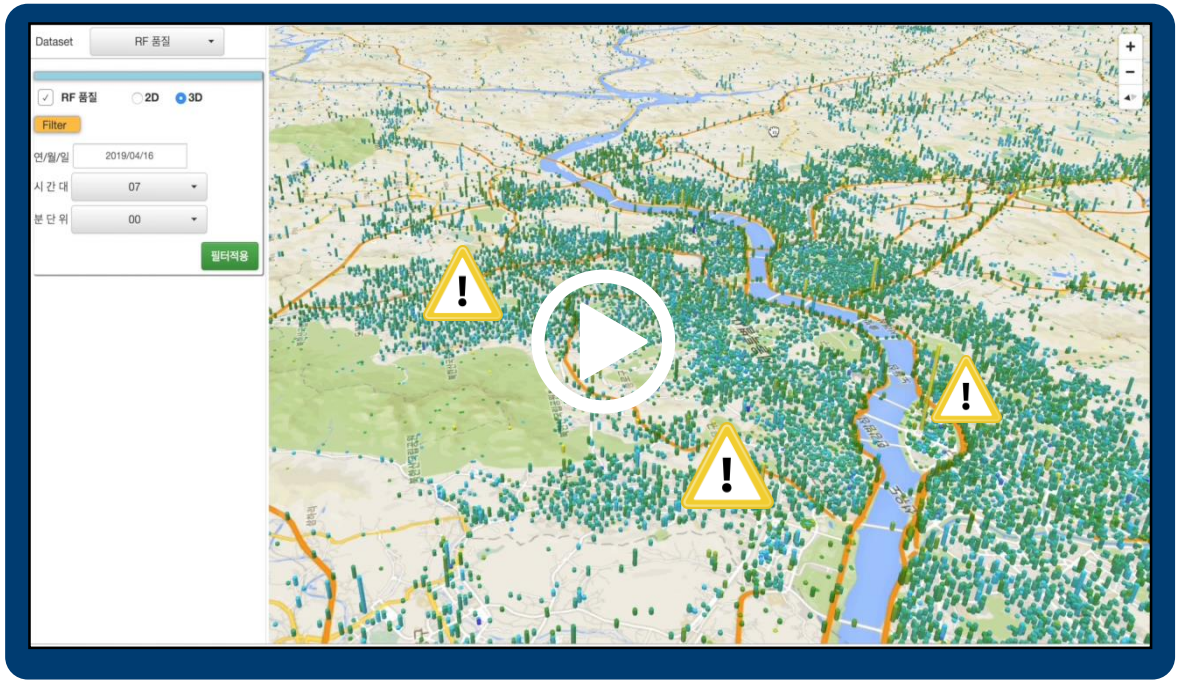
SK TELECOM USE-CASE

Hongchan (Nate) Roh: Team leader/ Sr. Software Engineer, SK Telecom

SK TELECOM 'LIGHTNING DB'

Network Quality Analytics, Visualization, and Predication

**Predict Network Quality Indicators (CQI, RSRP, RSRQ, SINR, ...)
for anomaly detection and real-time management**



- CQI : Channel Quality Indicator
- RSRP : Reference Signal Received Power
- RSRQ : Reference Signal Received Quality
  SINR : Signal to Interference Noise Ratio

SK telecom | Lightning DB

*Other names and brands may be claimed as the property of others.

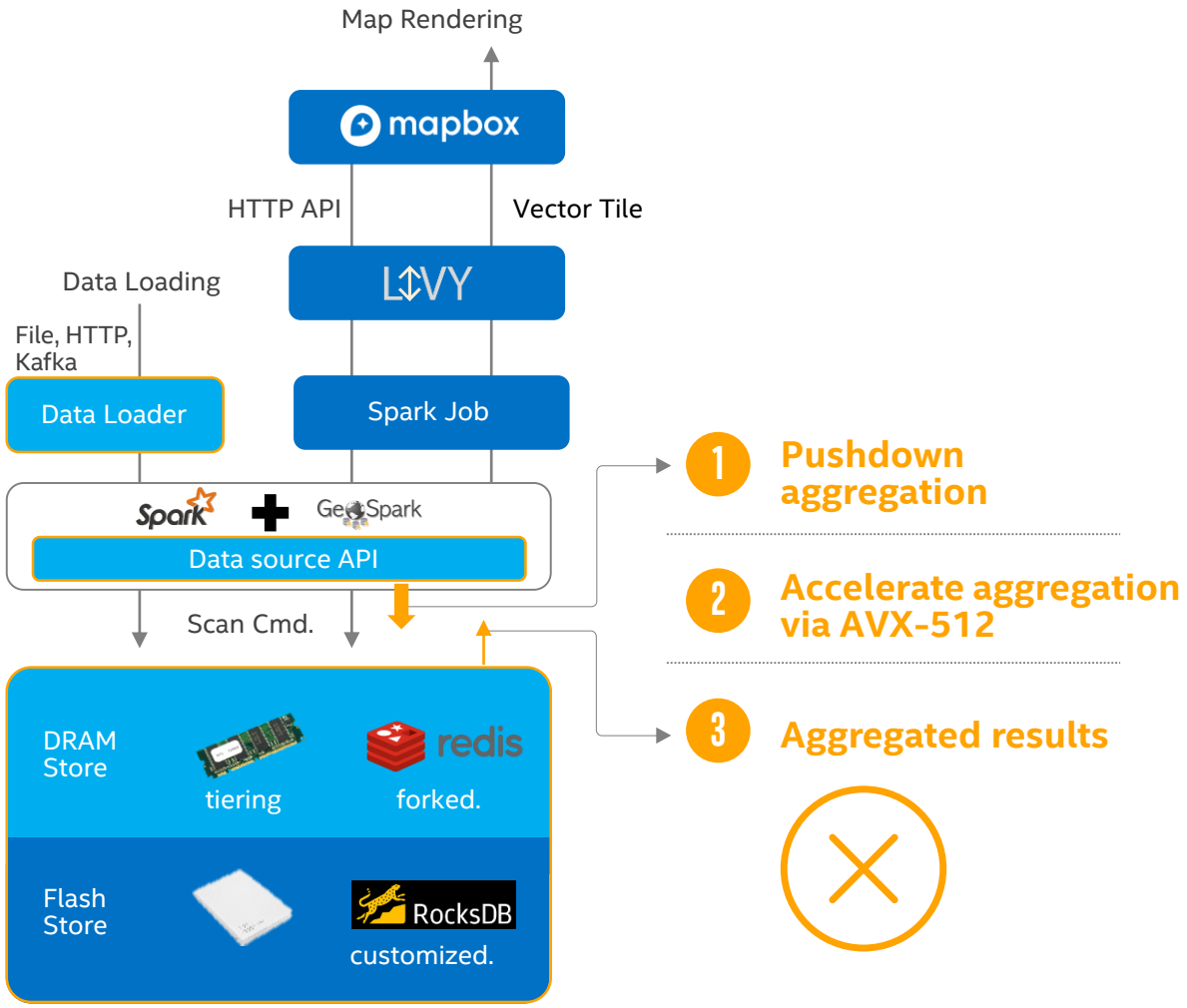
WHY GEOSPATIAL VISUALIZATION?

Geospatial Analysis

- Gathers, manipulates and displays geographic information system (GIS) data
- Requires heavy aggregate computations
  → **Good case to demonstrate real-time big data processing**
- Some companies demonstrated geospatial analysis to show advantages of GPU database over CPU database
  → **We could achieve the same level of performance with Spark & Lightning DB based on CPU**

*Other names and brands may be claimed as the property of others.

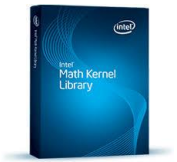
ARCHITECTURE OF GEOSPATIAL VISUALIZATION



OPTIMIZATION OF PERFORMANCE

1. Spark pushdowns aggregation to Lightning DB
   Lightning DB sends aggregated results to Spark
   → **Reduce Shuffle writing size and computation of Spark to 1/10.**

2. Lightning DB accelerates aggregation with vector-processing via Intel's AVX-512 (Intel Math Kernel Library)
   → **Upto 2x faster aggregation.**
   → **Upto 20 times faster than original GeoSpark**



Intel® Math Kernel Library

2nd Generation Intel® Xeon® Scalable Processors

*Other names and brands may be claimed as the property of others.

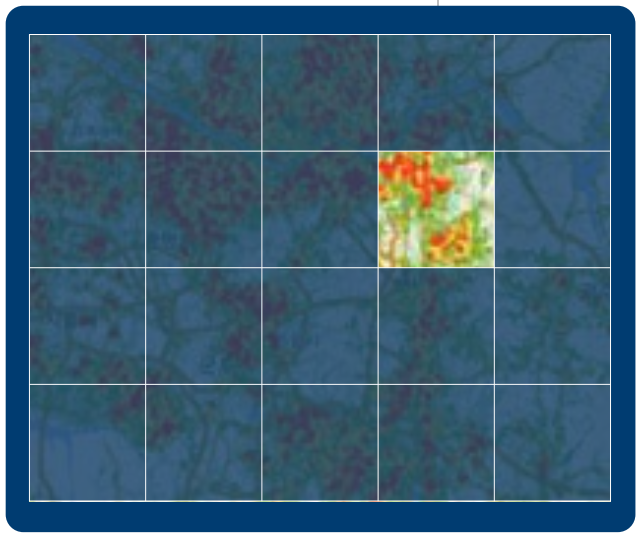
All performance testing and validation results were provided by SK Telecom Testbed. Intel does not control or audit third-party data.
You should consult other sources to evaluate accuracy.

OPTIMIZATION DETAIL

The Query building features of VectorTile

SELECT * FROM pcell WHERE **ST_VectorTileAggr('7,109,49', 'AVG')**

256
X
256 pix.



1. ST_VectorTileAggr(arg1, arg2)

- Custom predicate which contains aggregation information.
- arg1 : zoom level of map & tile pos (x, y) in Globe
- arg2 : aggregation type (SUM or AVG)

2. Define & Apply a custom optimization rule

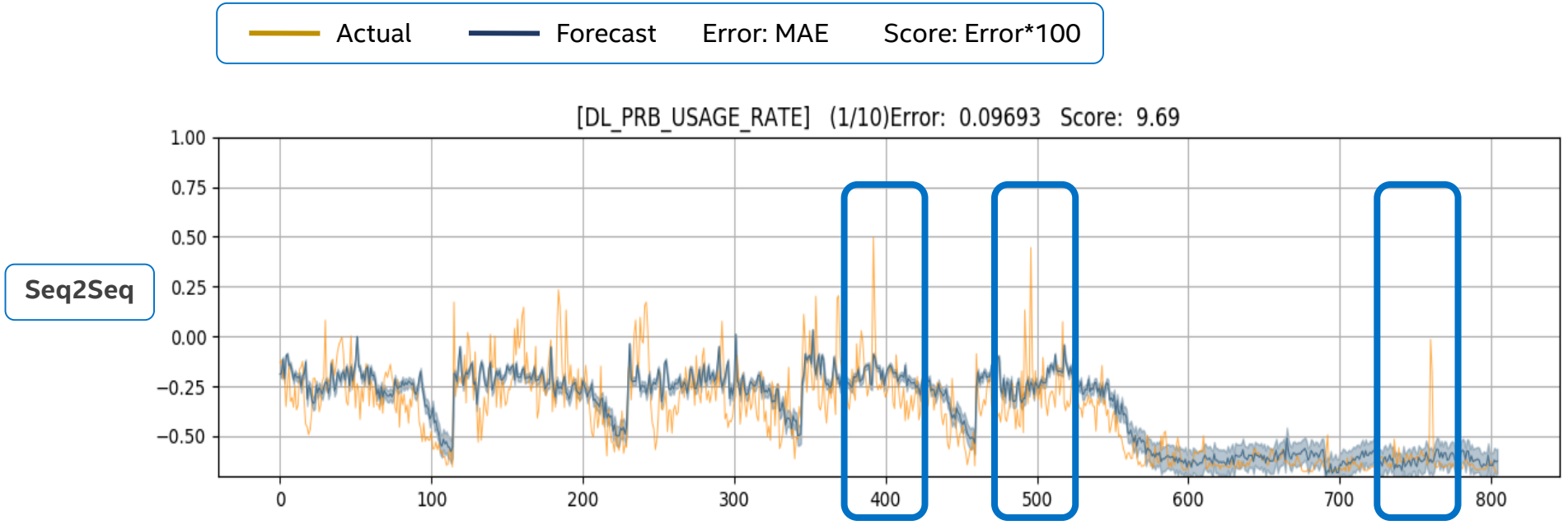
- Applied during optimization phase of query plan.
- Parse aggregation information from predicate and pushdown it to FlashBase

3. Aggregation in FlashBase

- Parallelized computation by FlashBase process count (Generally 100 ~ 200 process / node)
- Each process of FlashBase accelerates aggregation using Intel MKL.

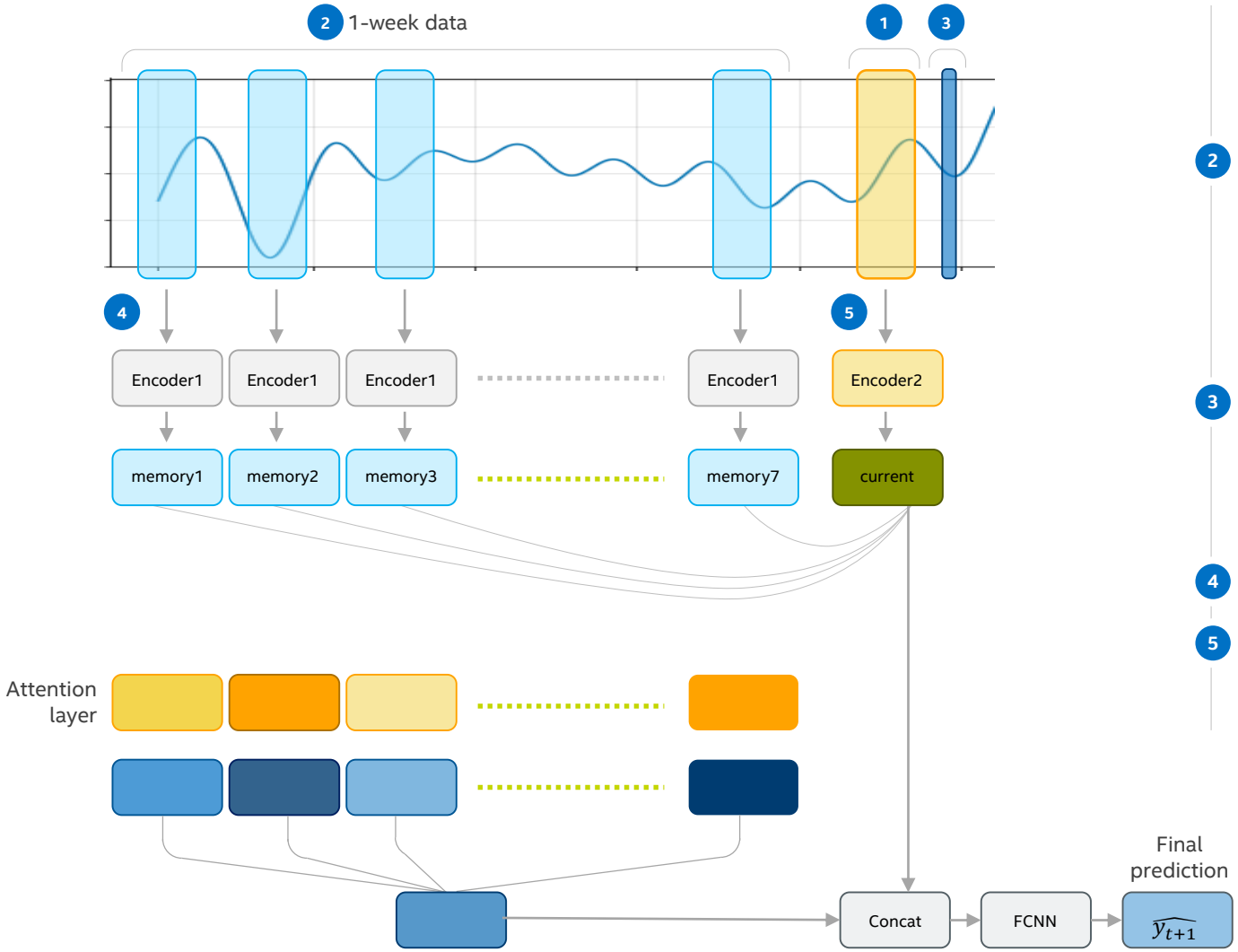
MODEL FOR NETWORK QUALITY PREDICTION - RNN+

RNN type model (Seq2Seq) is a common solution for time-series prediction.
But not suitable for our network quality prediction.



Cannot predict sudden change!

MEMORY AUGMENTED MODEL



1 **Current**

Recent 50 min data with 5-min period

2 **Memory**

Previous 7 day's historical data, each of which has same time band with current and target.

3 **Target**

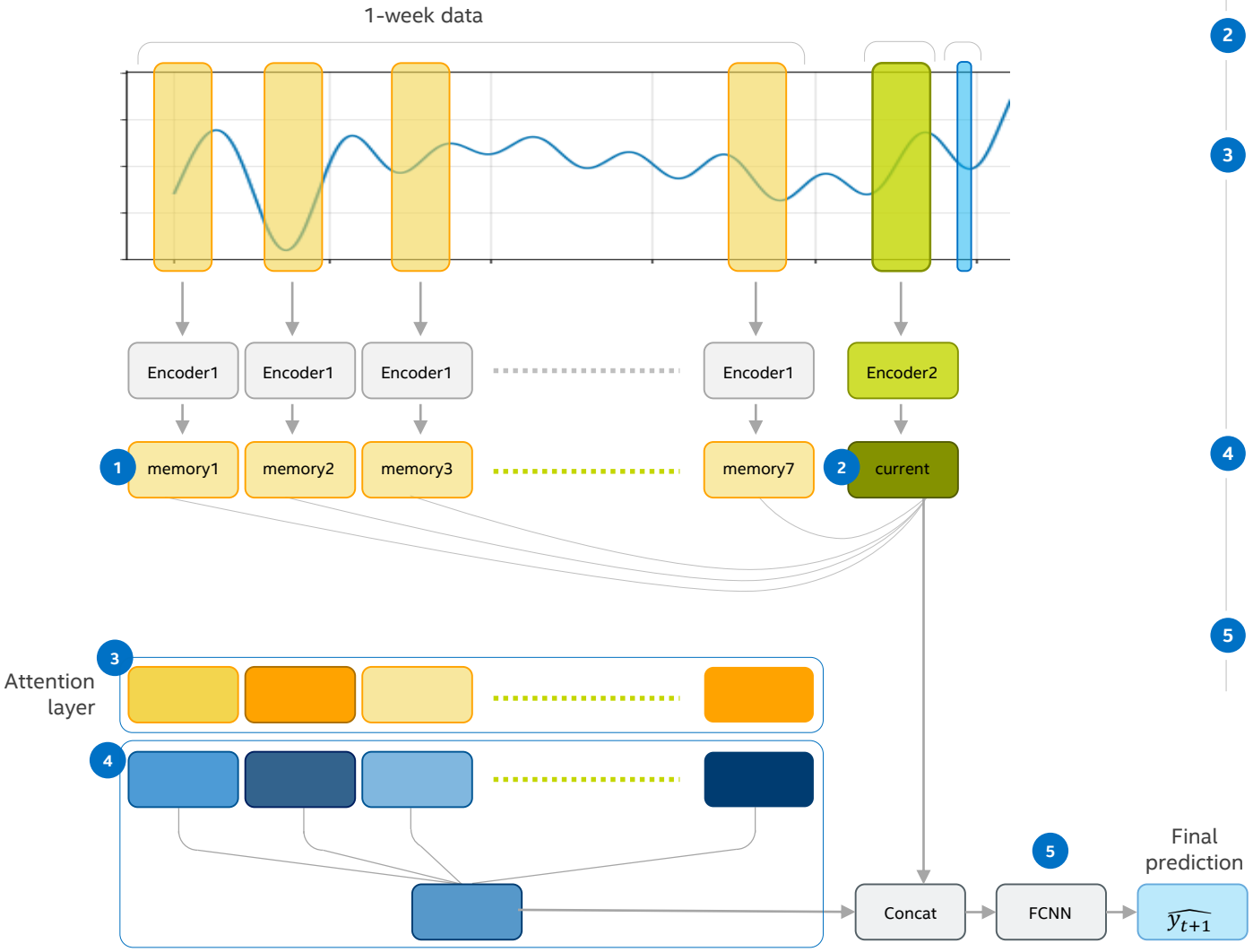
Network quality after 5 min

4 **Encoder** : 1-NN (Autoregressive term)

5 **Encoder1** : $h_t = c + w_1 y_{t\ -ndays-1} + \ldots + w_{11} y_{t-n\ days-11}$

Encoder2 : $h'_t = c' + w'_1 y_{t-1} + \ldots + w'_{10} y_{t-10}$

MEMORY AUGMENTED MODEL



1. $\mathbf{m_t}$ = memory for step t
2. c = current state
3. **Attention Layer** (1 layer neural-network)

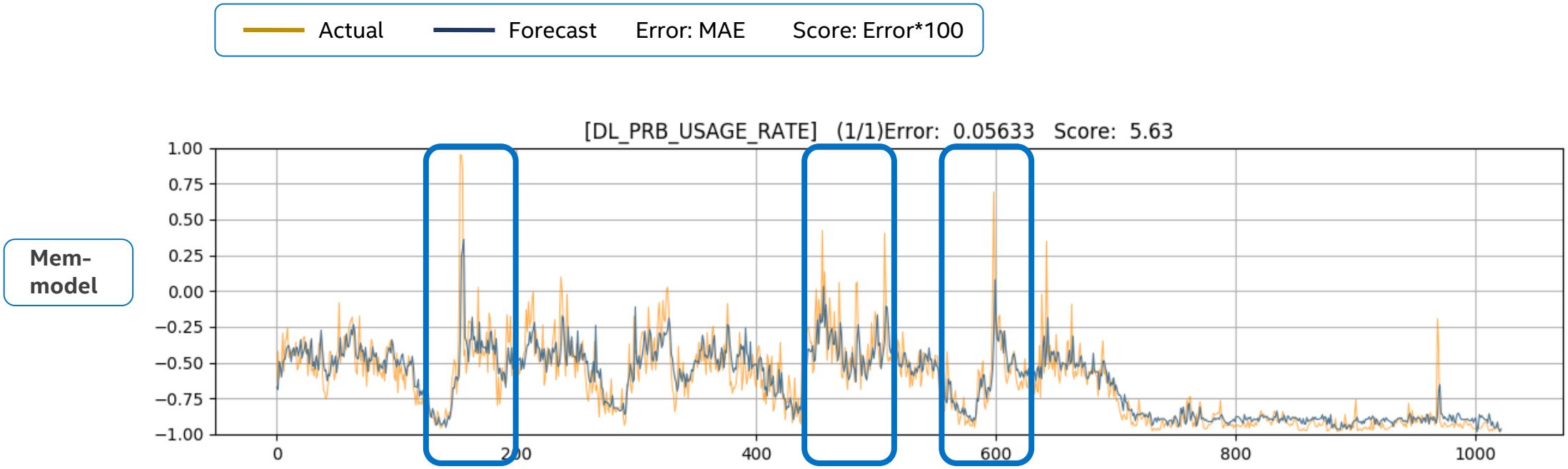
   $score_t = v^T \tanh(W_a[m_t; c])$

   (v, W_a: weight parameters)

   $\alpha_t = softmax(score_t)$
4. **Attention Vector**

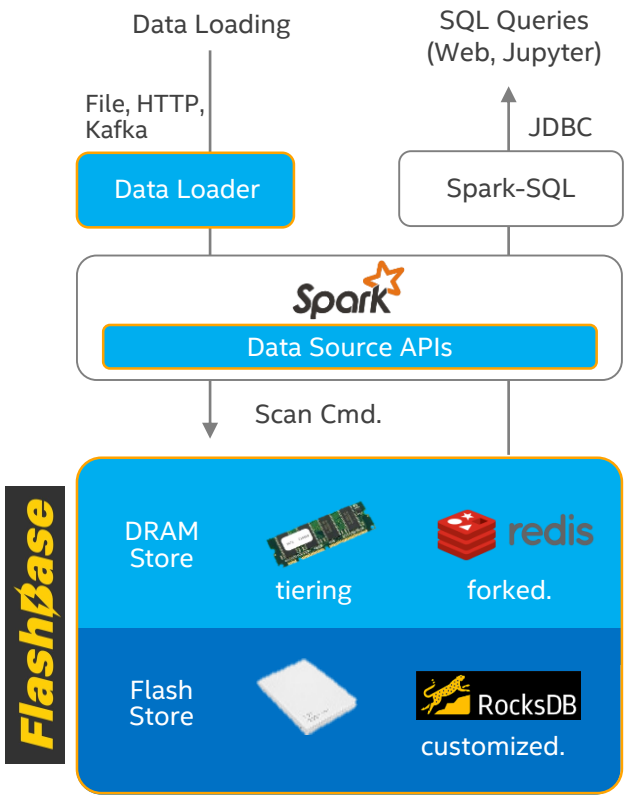
   Attention weighted summation of m_t
5. **Fully connected neural-network**

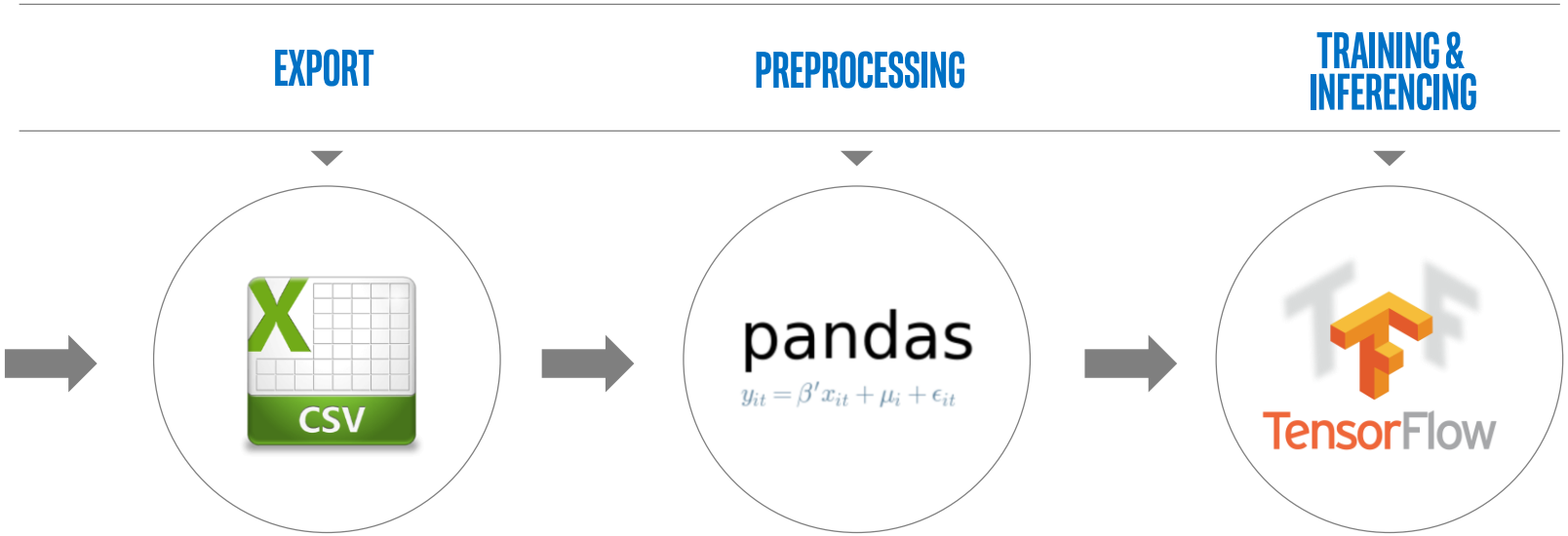
MEMORY AUGMENTED MODEL - TEST RESULT



Improved predictions for sudden change!

TRAINING & INFERENCE ARCHITECTURE - LEGACY



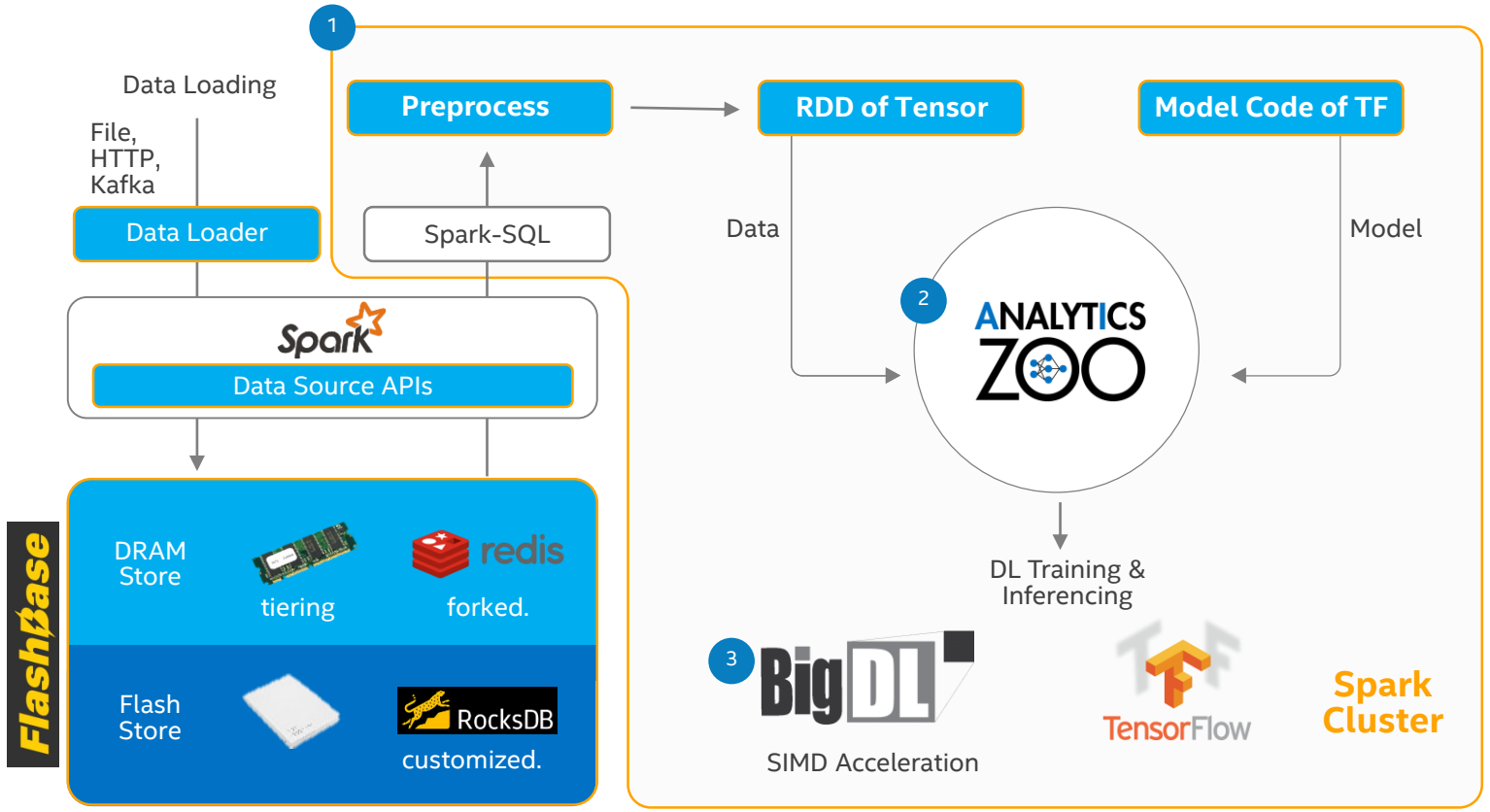


PROBLEM

1. **No in-memory Pipeline between data source and Deep-Learning layer.**
2. **Pre-processing & Inference & Training are performed in single server.**

TRAINING & INFERENCE ARCHITECTURE - NEW

1. Build in-memory Pipeline between FlashBase and Intel Analytics ZOO
   - Data Layer And Inferencing & Training Layer are integrated into the same Spark Cluster
   - Also share the same Spark session.
   - Source Code : https://github.com/mnms/ARMemNet-BigDL
2. Intel Analytics Zoo : Used to unify TF model into Spark Pipeline seamlessly.
3. Intel BigDL : inference & training engine
   - The processing of Inferencing & training can be distributed in Spark Cluster.



INTEL AI TECHNOLOGIES APPLIED INTO SK TELECOM CASE

Jason Dai: Sr. Principal Engineer, Intel

ONE INTEL ANALYTICS & AI PRODUCTS

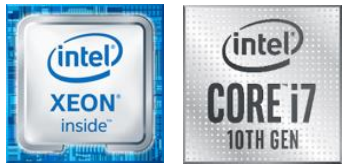
COMMUNITY

SOFTWARE

HARDWARE

WORKLOAD BREADTH — AI SPECIFIC

CPU | GPU | FPGA | CUSTOM



+

Intel AI Accelerator

Intel AI Accelerator

Multi-Purpose, Foundation for Analytics & AI	Data-Parallel AI, HPC, Media & Graphics	Real-Time & Multi-Function DL Inference	Edge DL Inference	Data Center DL Inference	Data Center DL Training

STORE

intel OPTANE MEMORY

intel OPTANE SSD

INTEL 3D NAND SSD

CONNECT

intel Ethernet

intel Silicon Photonics

BAREFOOT NETWORKS

All products, computer systems, dates, and figures are preliminary based on current expectations, and are subject to change without notice.

SPEED UP DEVELOPMENT

WITH OPEN AI SOFTWARE

TOOLKITS
App Developers

MACHINE LEARNING

ANALYTICS ZOO

DEEP LEARNING

ModelZoo

ANALYTICS ZOO

OpenVINO™

LIBRARIES
Data Scientists

Intel® Data Analytics Acceleration Library (DAAL)

Intel® Distribution for Python* (Sklearn*, Pandas*)

R (Cart, Random Forest, e1071)

Distributed (MLlib on Spark, Mahout)

Frameworks

TensorFlow, BigDL, Caffe, ONNX, mxnet, PyTorch

More framework optimizations in progress...

Intel Tools

NAUTA

RL Coach

NLP Architect

NN Distiller

KERNELS
Library Developers

Intel® Math Kernel Library (Intel® MKL)

Intel® Machine Learning Scaling Library (Intel® MLSL)

Intel® Math Kernel Library for Deep Neural Networks (Intel® MKL-DNN)

nGraph

CPU ▪ MEMORY

CPU ▪ GPU ▪ FPGA ▪ ACCELERATOR

Visit: *www.intel.ai/technology*

1 An open source version is available at: 01.org/openvinotoolkit

*Other names and brands may be claimed as the property of others.

Developer personas show above represent the primary user base for each row, but are not mutually-exclusive

All products, computer systems, dates, and figures are preliminary based on current expectations, and are subject to change without notice.

INTEL TECHNOLOGIES - ANALYTICS ZOO

Unified Data Analytics and AI Platform

MODELS & ALGORITHMS

- Recommendation
- Time Series
- Computer Vision
- NLP

ML WORKFLOW

- AutoML for Time Series
- Automatic Cluster Serving

INTEGRATED ANALYTICS & AI PIPELINES

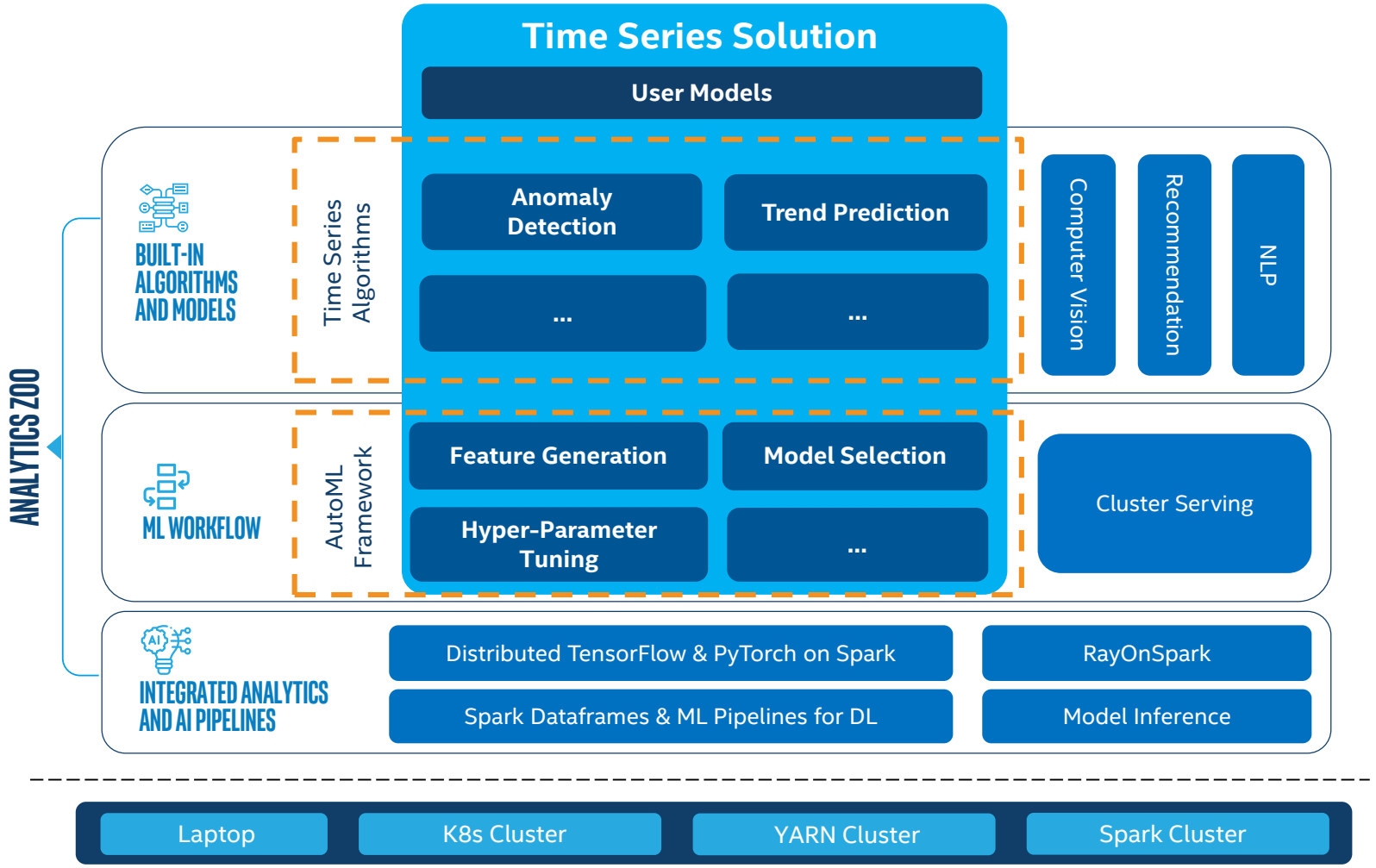
- Distributed TensorFlow & PyTorch on Spark
- RayOnSpark
- Spark Dataframes & ML Pipelines for DL
- Model Serving

LIBRARY & FRAMEWORK

- Distributions (Cloudera/Databricks/....)
- Distributed Analytics (Spark/Flink/Ray/...)
- DL Frameworks (TF/PyTorch/...)
- Python Libraries (Numpy/Pandas/...)

https://github.com/intel-analytics/analytics-zoo

INTEL TECHNOLOGIES - TIME-SERIES SOLUTION ON ANALYTICS ZOO



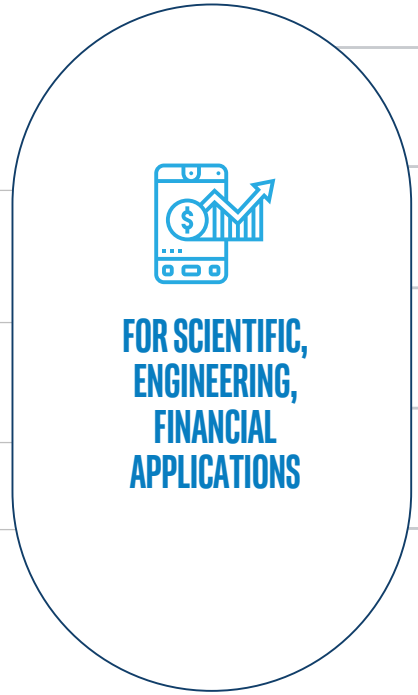
- **Rich models and algorithms** (statistical, neural-networks, hybrid state-of-art models, etc.)
- **AutoML** (for automatic feature generation, model selection, hyper-parameter tuning, etc.)
- **Seamless scaling** (with integrated analytics and AI pipelines)

FAST, SCALABLE CODE WITH INTEL® MATH KERNEL LIBRARY

The Fastest & Most-Used Math Library for Intel®-Based Systems[1]

Accelerate math processing routines, increase application performance & reduce development time

- Speeds computations by providing highly-optimized, threaded & vectorized math functions
- Provides key functionality for dense & sparse linear algebra, FFTs, vector math, summary statistics, splines & more
- Dispatches optimized code for each processor automatically without the need to branch code
- Optimized for single core vectorization & cache utilization
- Automatic parallelism for multi-core CPUs, scales from core to clusters



- DENSE LINEAR ALGEBRA
- SPARSE LINEAR ALGEBRA
- FAST FOURIER TRANSFORMS
- VECTOR MATH
- VECTOR RNGS

& MORE!

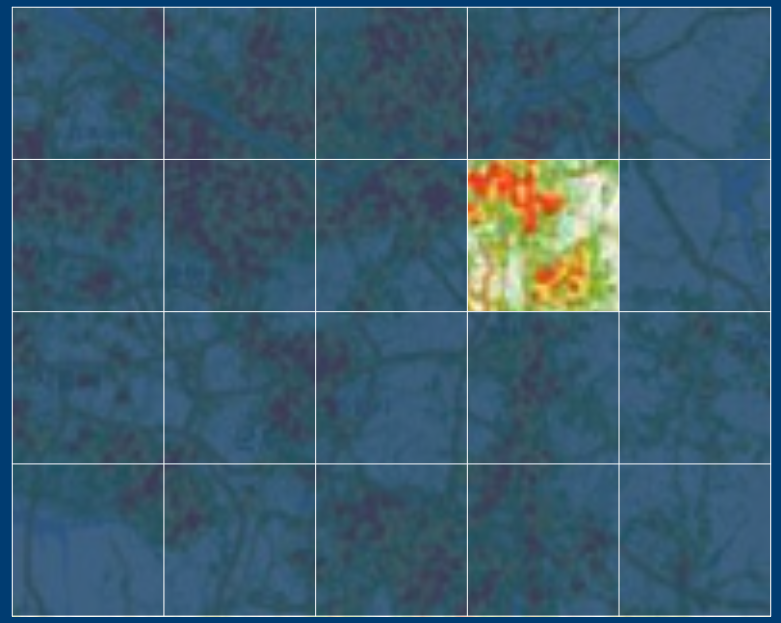
DOWNLOAD FREE INTEL® PERFORMANCE LIBRARIES: software.seek.intel.com/performance libraries

Also available in **Intel® Parallel Studio XE** & **Intel® System Studio**, or via YUM, APT-GET, Conda, Cloudera & NuGet

1. Data from Evans Data Software Developer surveys, 2011-2019

SKT PERFORMANCE IMPROVEMENT - (1) INTEL® MATH KERNEL LIBRARY

Building VectorTile to Visualize Network Quality Data

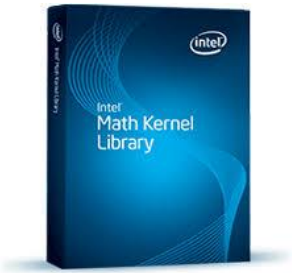


SK telecom | Lightning DB

Heavy Computation for Data Statistics Calculation and Aggregation

UP TO 2X FASTER AGGREGATION

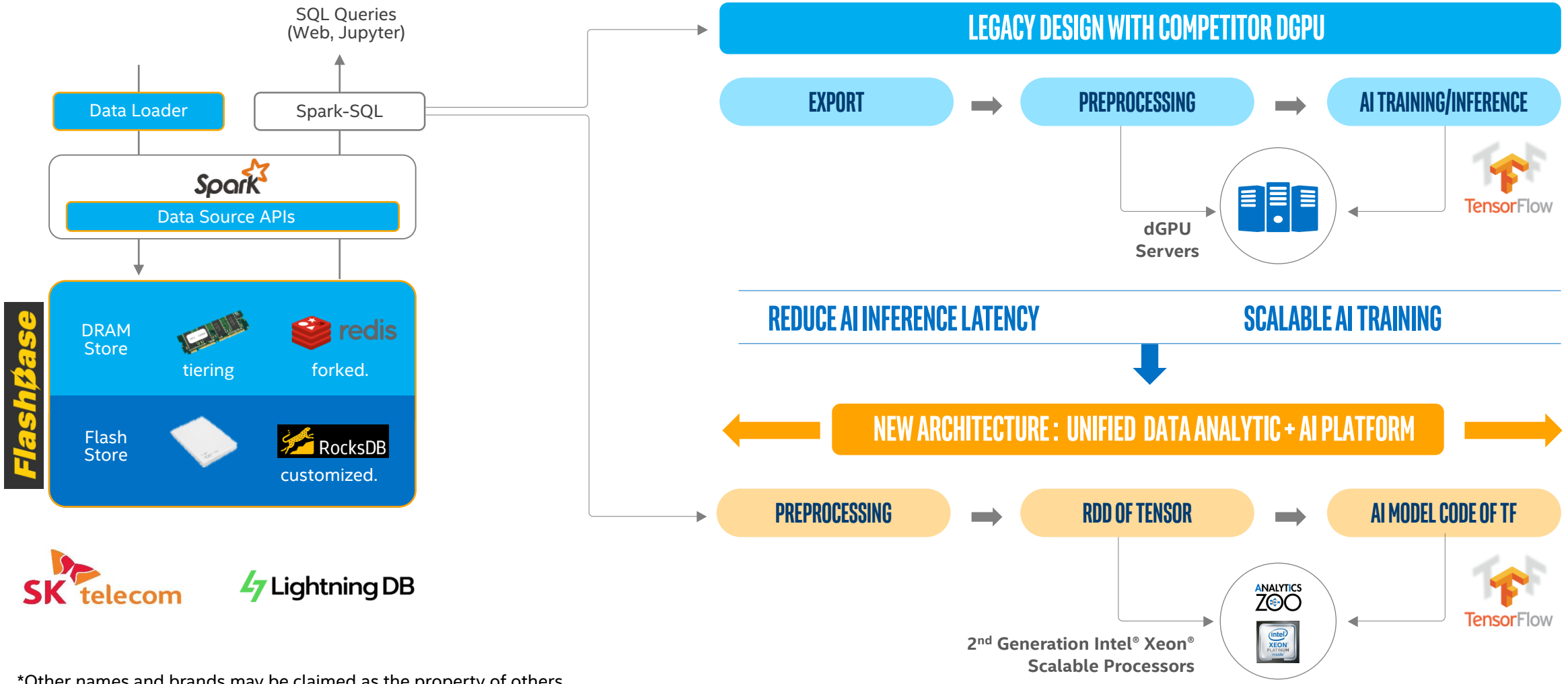
Vector Processing Optimization with AVX 512



Intel® Math Kernel Library

AVX-512
2nd Generation Intel® Xeon® Scalable Processors

SKT PERFORMANCE IMPROVEMENT - (2) ANALYTICS ZOO

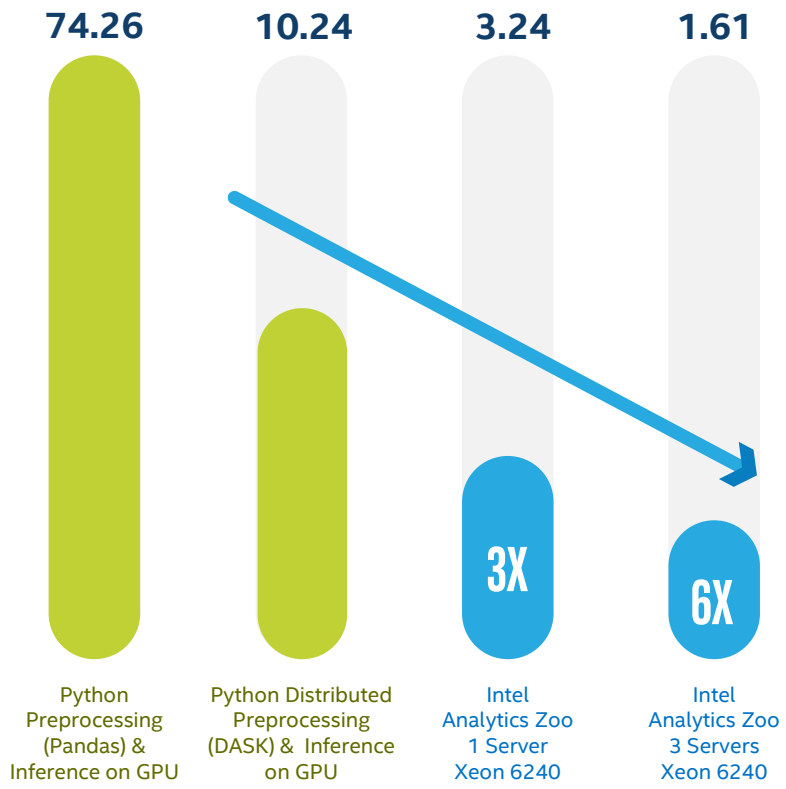


*Other names and brands may be claimed as the property of others.

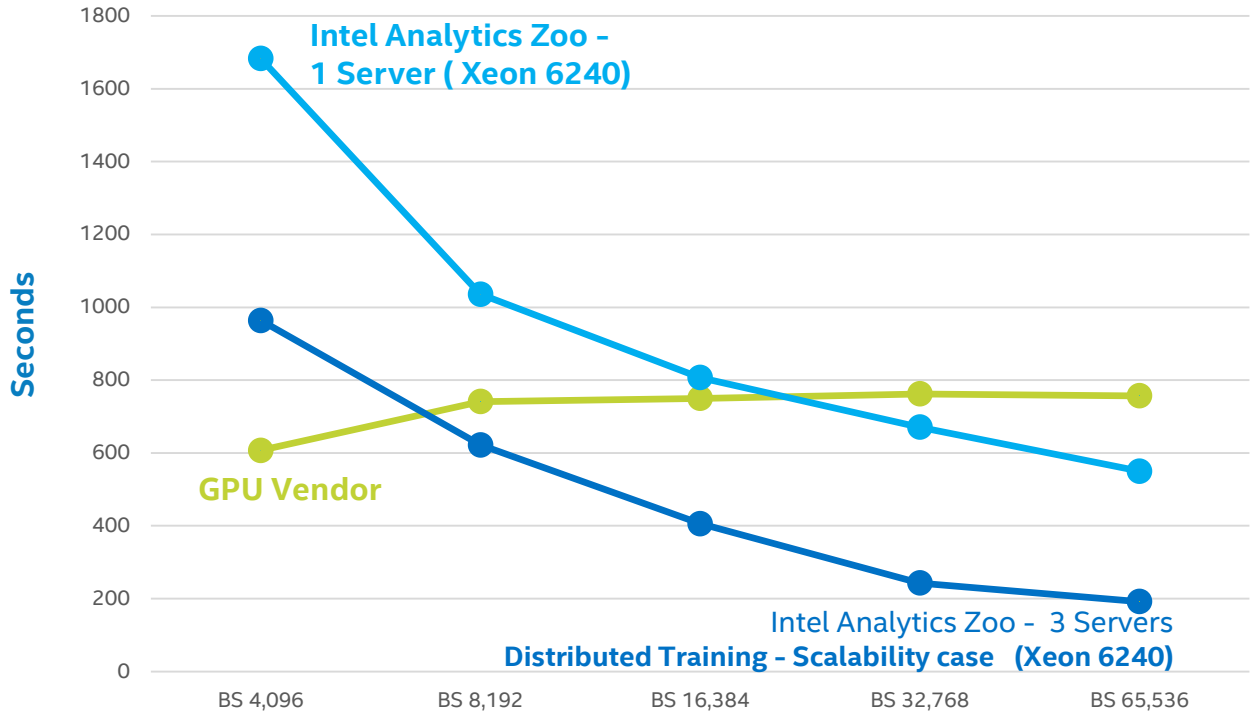
SKT PERFORMANCE IMPROVEMENT - (2) ANALYTICS ZOO

TCO OPTIMIZED AI PERFORMANCE WITH [1] ANALYTICS ZOO [2] INTEL OPTIMIZED TENSORFLOW [3] DISTRIBUTED AI PROCESSING

[1] PRE-PROCESSING & INFERENCE LATENCY



[2] TIME-TO-TRAINING PERFORMANCE



Test Data: **80K** Cell Tower, **8** days, **5**mins period, **8** Quality Indicator

Performance test validation @ SK Telecom Testbed

INTEL COLLABORATION

UNDERSTAND CUSTOMER PROJECTS

TARGET USE-CASE	KPI/Traffic Forecast, Anomaly Detection, Fault Detection, Root cause Analysis, more
AI MODEL, KPI, H/W PLATFORM	Model, performance KPI target (training/ inference), Target H/W
DATA ANALYTICS ARCHITRAVE	Analytics : Spark, Flink, Splunk, more Analytic + AI architecture
PROJECT KEY MILESTONES	PoC, Trial, Deployment milestone

IDENTIFY COLLABORATION AREA

- **Build unified 'Analytics + AI' pipeline**
- **AI performance optimization**
- **Use-case development w / AI model**
- **AI H/W platform validation**
- **Co-marketing**

POC /TRIAL / COMMERCIAL

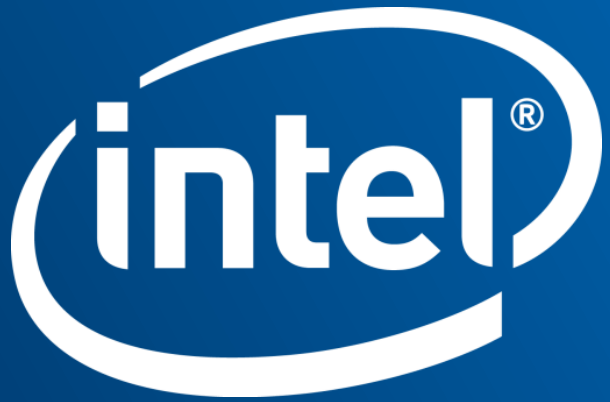
LEGAL NOTICES AND DISCLAIMERS

All performance testing and validation results were provided by SK Telecom Testbed. Intel does not control or audit third-party data.

No product or component can be absolutely secure. Intel technologies may require enabled hardware, software or service activation. Your costs and results may vary. Intel does not control or audit third-party data. You should consult other sources to evaluate accuracy.

Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors.

Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more complete information visit www.intel.com/benchmarks.


intel