EVOLVE

Applications

Vassilis Spitadakis v.spitadakis@neurocom.lu

EBDVF21,   29 November 2021

Seven EVOLVE pilots

Automotive Industry

Space Hellas
- **Maritime Surveillance using** observation data, historic metadata and classification models

Thales Alenia
- **Change detection on satellite images**

Cybeletech
- **Optimising agro production yield** using numerical models and massive historic data

Automotive Industry

AVL
- **Predictive Maintenance** Application

Koola
- **Data driven vehicle engineering** processes support

Transport

Tiemme/MEMEX
- **Public Transport** planning and operations

### BMW
- **Demand Mobility Services**

Applications Stack Profiles

| Pilot Application	Technology Stack
Maritime Surveillance	Spark, Tensorflow, GPUs, external visualisation and Weblyzard
Change detection on satellite imaging	Kafka, Tensorflow, GPUs, DASK
Optimising agri-production yield	Tensorflow, GPUs, MPI, Memoscale
Improvement of bus transportation service	Kafka, Spark, external visualisation and Weblyzard
Advanced mobility services optimization	Kafka, Tensorflow, GPUs
Predicitive Vehicle Maintenance	Tensorflow, GPUs, R
Data-assisted automotive service development	Spark, R, plotpy and Weblyzard

Goals

- **Performance**

- **Flexibility to adapt**

Change Detection
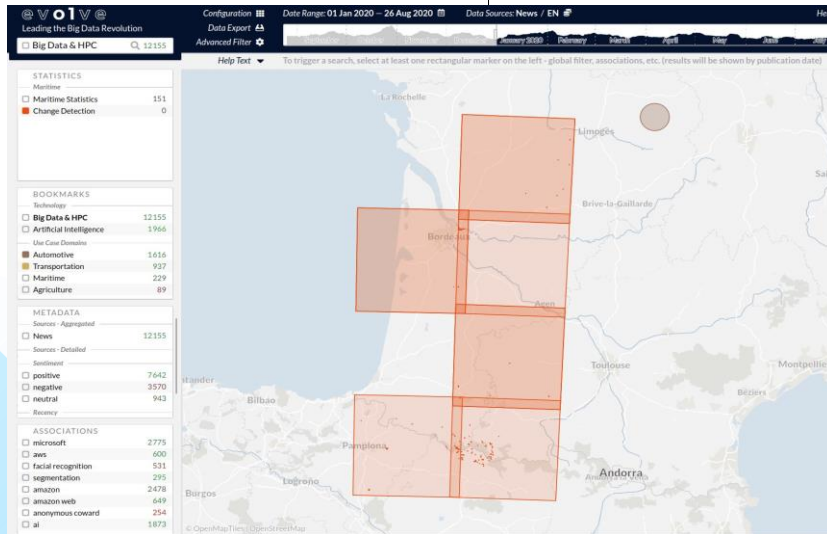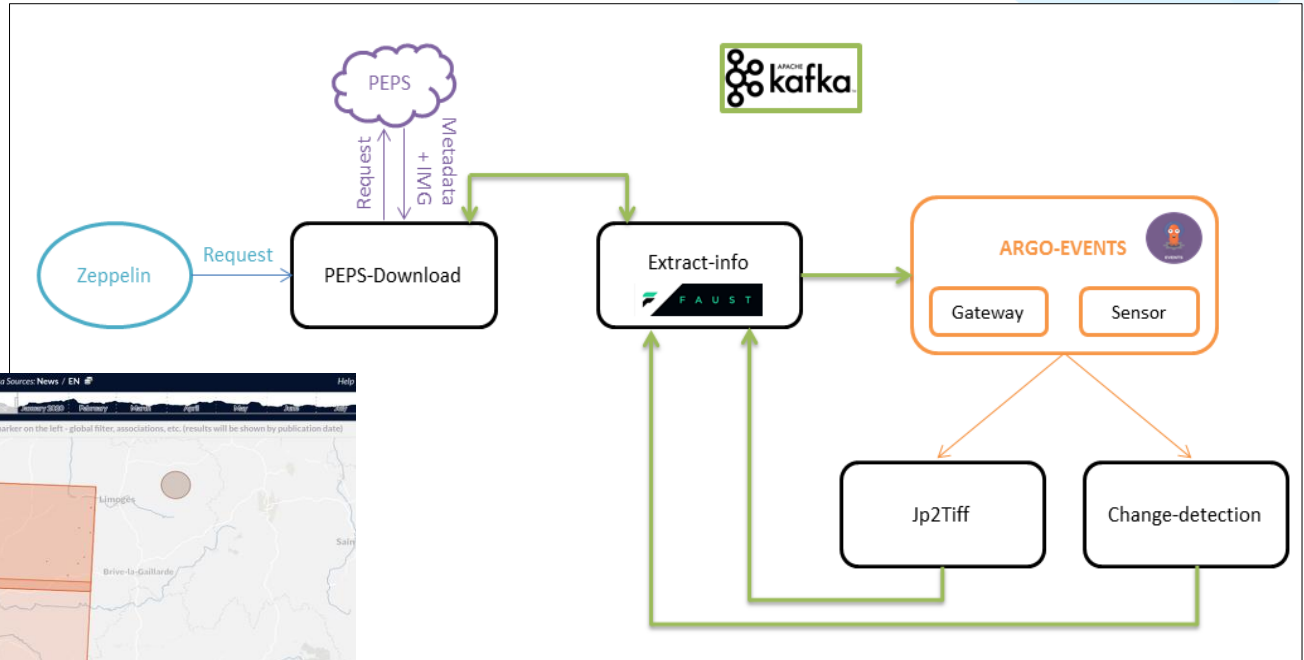
A data pipeline exploiting microservices, GPU and EVOLVE visualization technologies.
**> 6x and wider geo scale**

# Change Detection
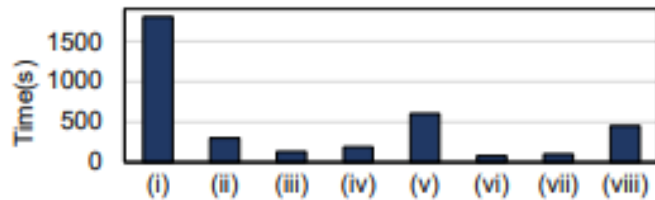
8 different versions of this application:

i)   CPU-only,
ii)   iCPU-only using TensorFlow(TF),
iii)   iGPU and TF,
iv)   DASK,
v)   DASK and GPU,
vi)   KAFKA and DASK in a small dataset,
vii)   KAFKA and GPU in small dataset,
viii)   KAFKA and DASK in a big dataset.

KPI Description	EVOLVE Impact (speedup)
Average time to generate one change detection map from a couple of Sentinel-2 images with the tile generation step running on CPU and DASK is used for parallel computing	x10 faster
Average time to generate one change detection map from a couple Sentinel-2 images with tile generation running on GPU	x15 faster
Average time to generate one change detection map from **200** Sentinel-2 images	x120 faster for the CPU+Dask version

When ?	Processing time	Number of results
Before Evolve	**> 20min**	**1** CD map
Dask	**3min15s**	**1** CD map
Dask + Kafka	**6min30s**	**5** CD maps

Maritime Surveillance
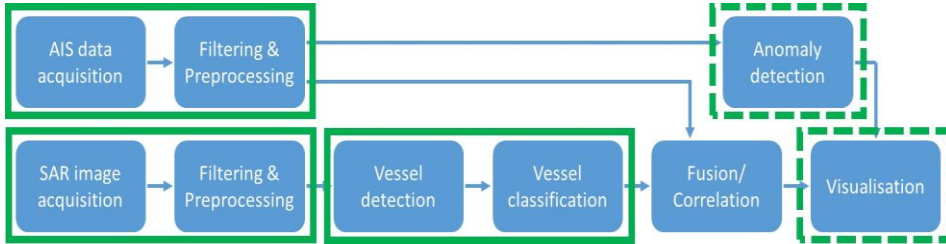
IME, microservices and stream processing, GPU, Zeppelin visualization brought substantial performance benefits
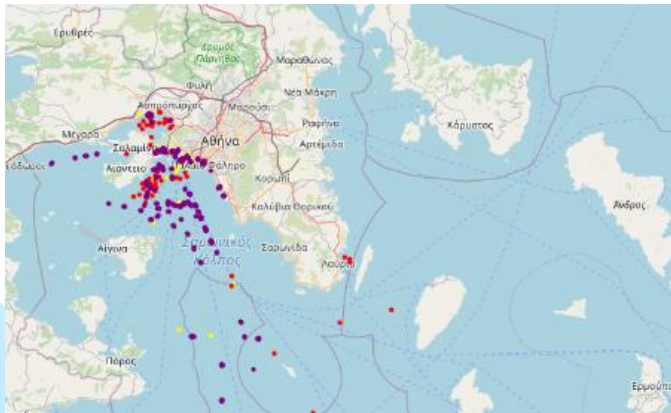

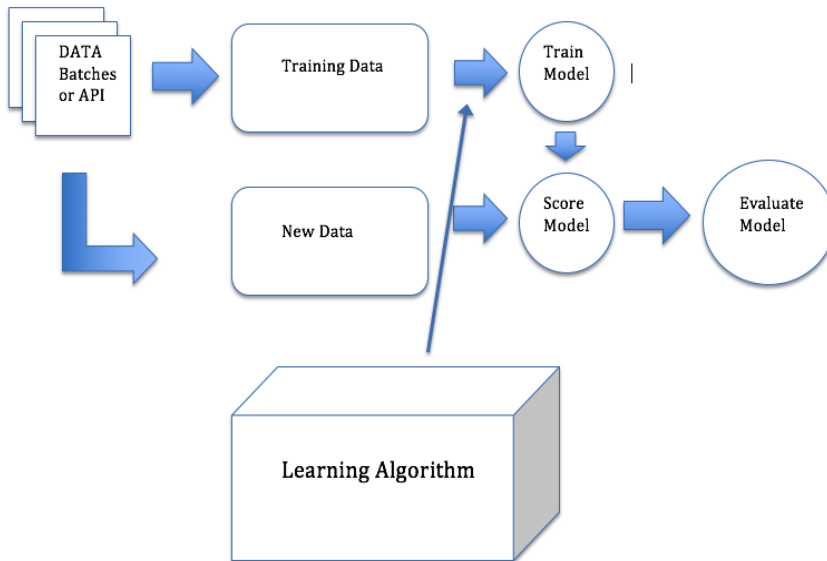| Data Volume	
• Data processed per batch	
Vessel Detection	1 SAR Image ~ 1Gb
Anomaly Detection	8 records for testing, 4096 for training
• Time for batch processing	
Vessel Detection	~60 secs/image
Anomaly Detection	~0.015 secs/batch with GPU
• Batches Required	
Vessel Detection	depends on user criteria
Anomaly Detection	895 batches/epoch, 500 epochs
Comparison with previous situations	
Vessel Detection	~210 secs (previous) VS ~60 secs (now)
Anomaly Detection	>8 hours (previous) VS <2 hours (now) – training phase

Maritime Surveillance
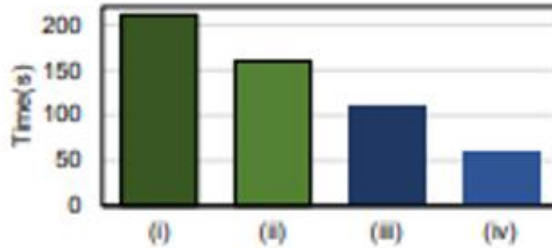
SAR scene processing
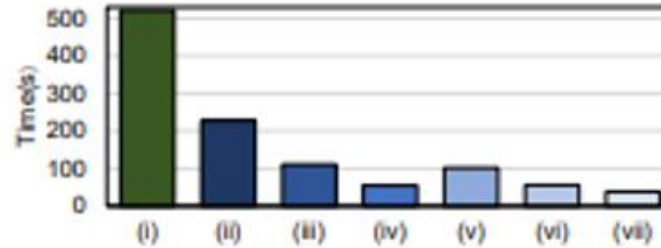
Anomaly Detection based on AIS data

KPI Description	EVOLVE impact
Vessel detection time	71% reduction
AIS anomaly detection training time	89% reduction
Detection Accuracy	5% increase
Operational spatio-temporal window increase	10x

Predictive Maintenance

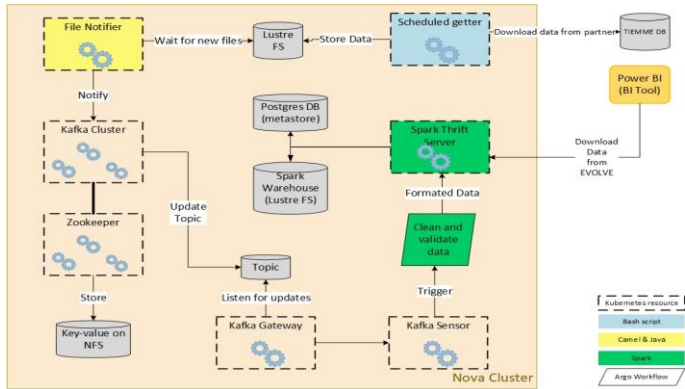| KPI Description	EVOLVE impact
Preparation, Training and Serving Data models time	2x – 3x speedup (without GPUs) More expected via GPUs (in progress M36)

Different set of Machine Learning libraries are assessed with different and extensive datasets. Models deployed on EVOLVE are showing significant speedups (from 2x to 3x)
in four different algorithms (logistic regression, linear discriminant analysis, classification and regressions trees, support vector machines)

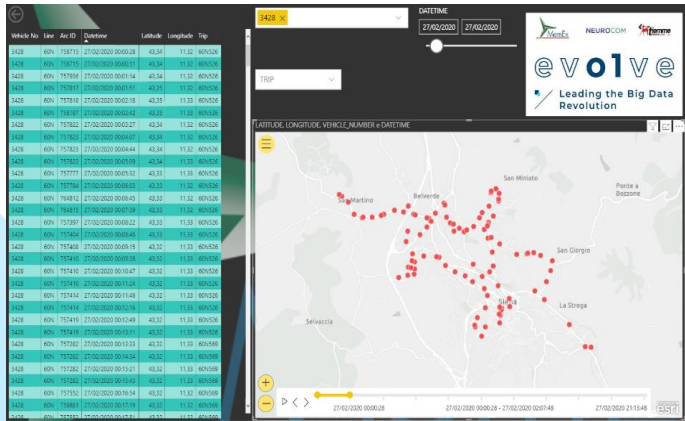More to expect with tensorflow and GPU

Public Transport

Two data workflows (**batch** and **real-time)**, are exploiting IME and Spark plus Spark streaming engine over EVOLVE to improve decision support in public transport operations.

Finally exporting data to external Business Intelligence tools and WebLyzard visualisation

Past analysis made on 1-month data ➔ with EVOLVE can be made on yearly data

Past analysis only on historical data ➔ with EVOLVE real-time ability was easily deployed

Public Transport

| KPI Description	Before Evolve	After Evolve
Time period covered by the analysis (for the offline workflow context), which can be considered by the management authority as possible to analyse	1 month data involving around 220 bus shifts	1 year and related all the 4 areas served by Tiemme fleet
Listing of trips operated in a period of interest along with the extracted relevant bus events (for the offline context) Indicative KPI : The time required to extract 86370 records (corresponds to 2 days service) from SQL	20 min and 35 seconds	>20x speedup Less than one minute
Daily query: summary of transits time analysis carried out within a period of interest: 30-60-90 days of service (for the offline context)	For 30 days data: 2 min 15 seconds For 90 days data: 29 minutes and 49 seconds	For 30 days data 11,21 seconds For 60 days data 13 seconds For 90 days data 16 secods From 10x to 120x speedup
Necessary time to visualize the real time status of 2 or 3 selected lines/routes (for "Real Time" context) and related information, on map	Not supported	Under 6 seconds to analyse 2 or 3 lines.
Arcs number which can be processed and visualized simultaneously with related updated information (for "Real Time" context) on the map	Possibility to analyze only a few arcs in an acceptable interval	Possibility to process 21K arcs in under 20 seconds

Agri production optimisation
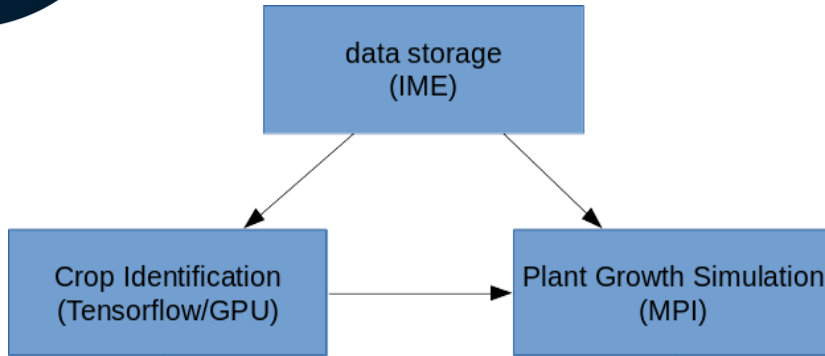
Integration of MPI within EVOLVE

Use of IME, GPUs and stream analytics engine by tensorflow on 20 Gb datasets. Memoscale compression stage at data transfer phase.

KPI Description	EVOLVE impact
Crop Detection Rate - The success rate of crop detection is improved with larger learning datasets	90% → 97%
Learning dataset size	x4 Geographcal area coverage enlarged
Plant Growth Simulation – Computation time reduction	:4
Number of simulations	x2
Overall process execution time	:5, More robust, stable and easy to deploy

Demand Mobility Services

Exploiting EVOLVE HPC aims to accelerate three different on-demand mobility services (allocating cars to requests)
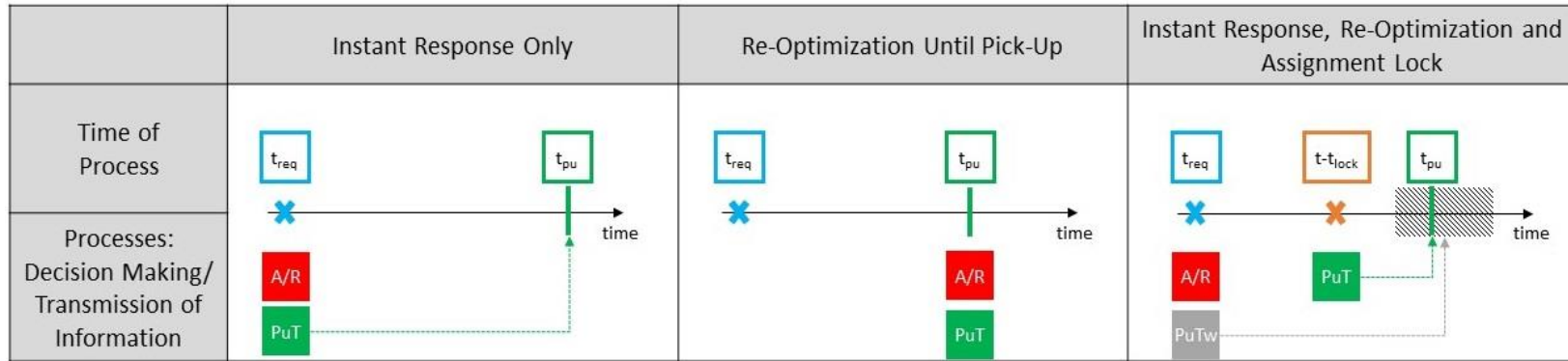
Evident speedups based on New York demand datasets.
Three simulations are end-to-end possible on EVOLVE, assessed over different distributed parallel schemes of Branch-and-Bound algorithm using MPI.

Conclusions

► **Remarkable Performance Improvement:** The use cases are using technology components, such as parallelization frameworks, stream analytic engines, fast storage, and hardware accelerators that automate, optimize, and boost the performance of applications.

► **Business-level achievements**
- Increase of quality and quantity business indicators
- Development efficiency ➔ lower efforts and costs, quicker time-to-market:

The Consortium

DDN STORAGE
www.ddn.com

BULL
www.atos.net

IBM
www.ibm.com

FORTH
www.ics.forth.gr

OnApp
ww.onapp.com

Institute of communications and computer systems
www.microlab.ntua.gr

MemoScale
www.memoscale.com

webLyzard technology
www.weblyzard.com

LOBA
www.loba.pt

Thales Alenia Space
www.thalesgroup.com

Space Hellas
www.space.gr

CybeleTech
www.cybeletech.com

Neurocom Luxembourg
www.neurocom.eu

MemEX
www.memexitaly.it

Tiemme SPA
www.tiemmespa.it

Virtual Vehicle
www.v2c2.at

AVL List GmBH
www.avl.com

BMW AG
www.bmw.com

KOOLA
www.koola.io

www.evolve-h2020.eu
info@evolve-h2020.eu

The Use Case Providers

DDN STORAGE
www.ddn.com

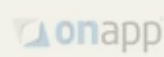**BULL**
www.atos.net

IBM
www.ibm.com

FORTH
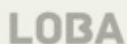www.ics.forth.gr

OnApp
ww.onapp.com

Institute of communications and computer systems
www.microlab.ntua.gr

MemoScale
www.memoscale.com

webLyzard technology
www.weblyzard.com

LOBA
www.loba.pt

Thales Alenia Space
www.thalesgroup.com

Space Hellas
www.space.gr

CybeleTech
www.cybeletech.com

Neurocom Luxembourg
www.neurocom.eu

MemEX
www.memexitaly.it

Tiemme SPA
www.tiemmespa.it

Virtual Vehicle
www.v2c2.at

AVL List GmBH
www.avl.com

BMW AG
www.bmw.com

KOOLA
www.koola.io

www.evolve-h2020.eu
info@evolve-h2020.eu

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 825061