Data analysis platform for stream and batch data processing on hybrid computing resources*

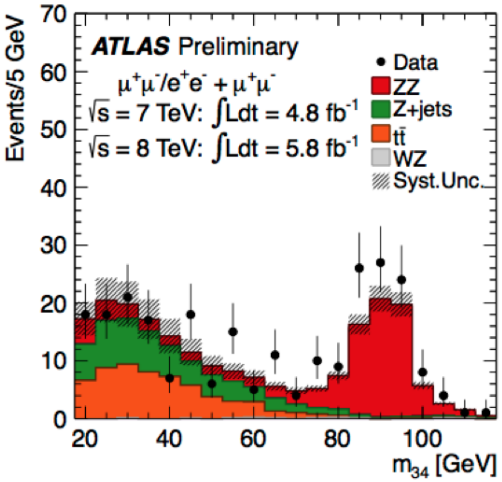
I. Kadochnikov[1,2], S. Belov[1,2], V. Korenkov[1,2], R. Semenov[1,2], P. Zrelov[1,2]

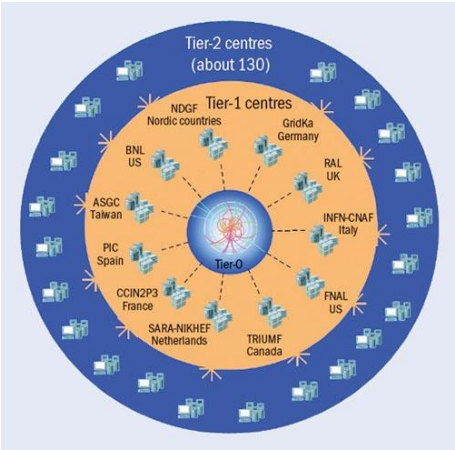
1 JINR Meshcheryakov Laboratory of Information Technologies
2 Plekhanov Russian University of Economics

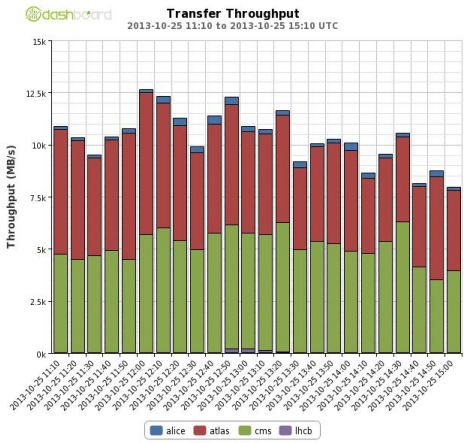
* The study was carried out at the expense of the Russian Science Foundation grant (project No. 19-71-30008)

Physics data and incidental data

- Incidental can still be critical:
  - Monitoring experiment infrastructure
  - Monitoring distributed processing
  - Monitoring computing infrastructure
  - Overview of scientific publications
  - Actual scientific analysis?
- Batch processing for deep analysis and machine learning
- Stream processing for using the new knowledge on-line

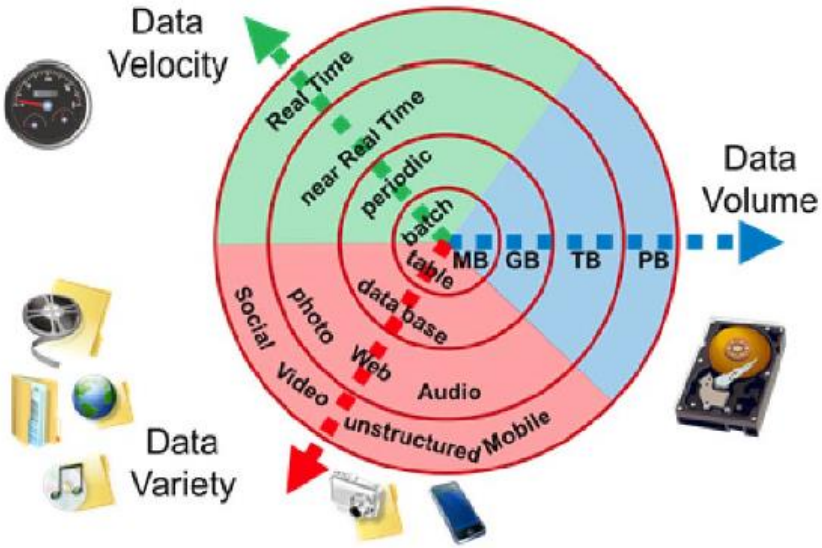


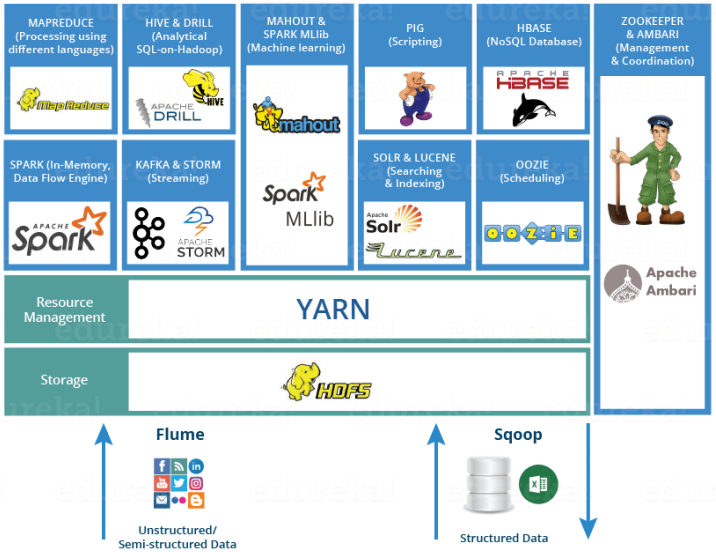




Big data platforms

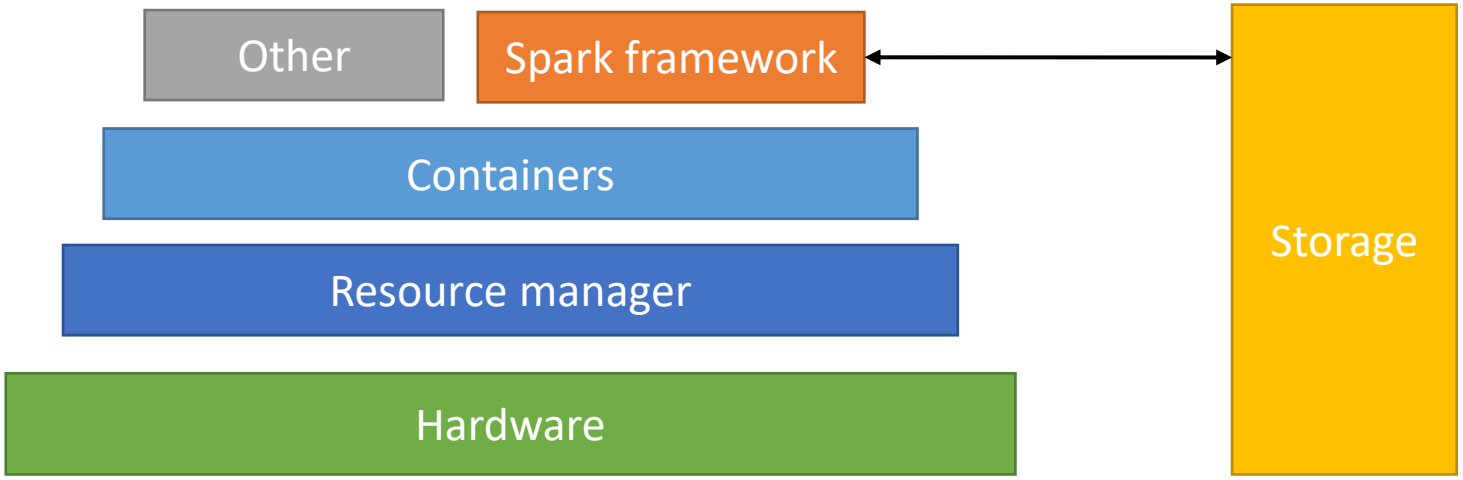
- Robust, general-purpose tools
- Scalable in practice
- Convenient for common data formats and analysis paradigms





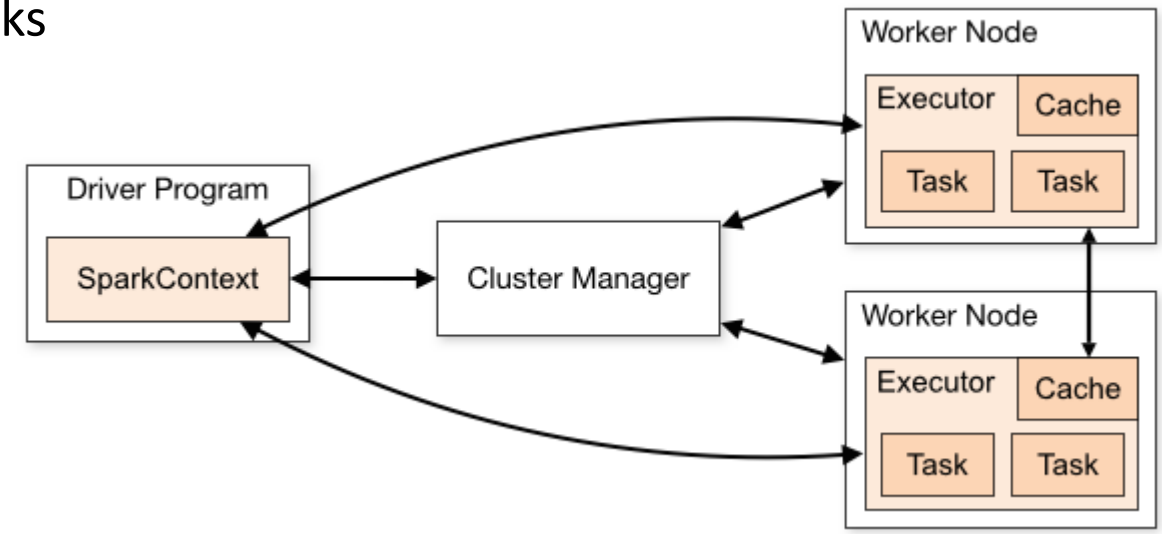
Platform architecture

- Core: Apache Spark
- Basis: resource manager
  - Mesos -> Kubernetes



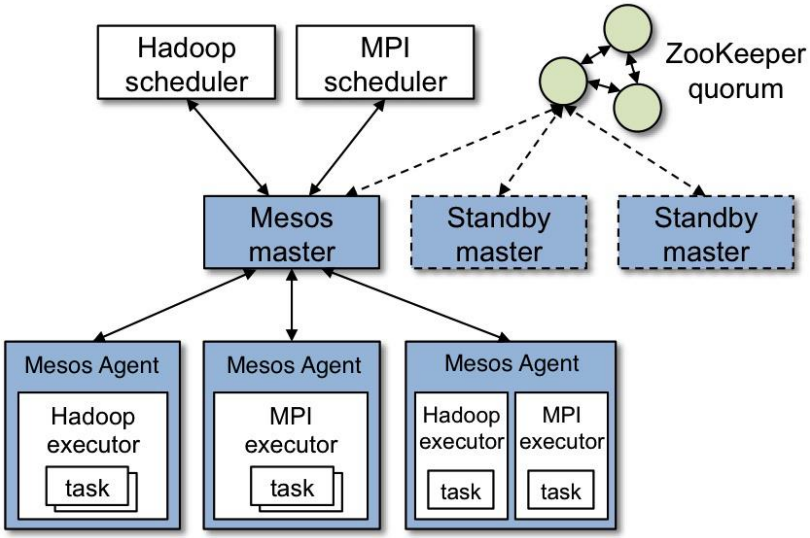
Apache spark

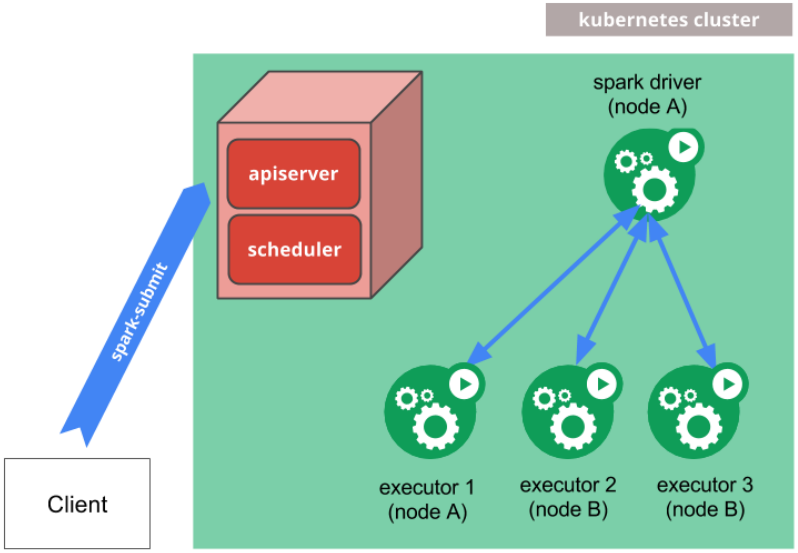
- Distributed Big Data analysis
  - Batch AND stream mode
  - Machine learning libraries
- Manage data shuffle between nodes:
  - Less detailed and frequent than MPI on HPC
  - More than independent GRID jobs
- Manage jobs and tasks
- Manage data in/out
- Libraries
- Multi-language



Role of resource manager in a Spark cluster

- Dedicated coordination service
- Share resources between instances
- Different versions
- Non-spark jobs
- Service containers





GPU support in Spark

- Standalone cluster
  - simple
  - no benefits of resource management
- Mesos
  - No GPU resource support yet
- YARN
  - Hadoop ecosystem
- Kubernetes
  - Container management for other platform services
  - Need to install NVIDIA components to add GPU support to Docker

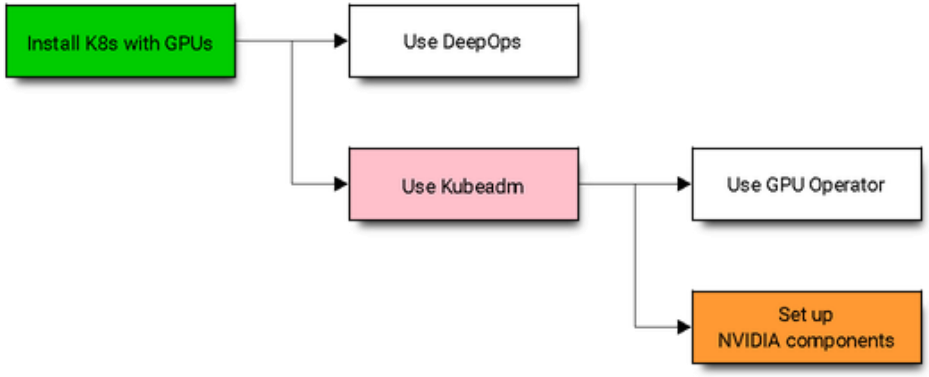
Nvidia container components

- Nvidia-docker
  - libnvidia-container – the main functionality of GPU support in container
  - nvidia-container-toolkit – script using libnvidia-container cli to inject GPUs into the container pre-start based on the container's config.json
    - --gpus option in docker 19.03+ uses toolkit directly
  - nvidia-container-runtime – thin wrapper around runC that adds nvidia-container-toolkit as a prestart hook
  - nvidia-docker2 – installs nvidia-container-runtime into docker's daemon.json
    - Adds option --runtime=nvidia to docker
    - This is the one needed for Kubernetes GPU support

Deepops repository

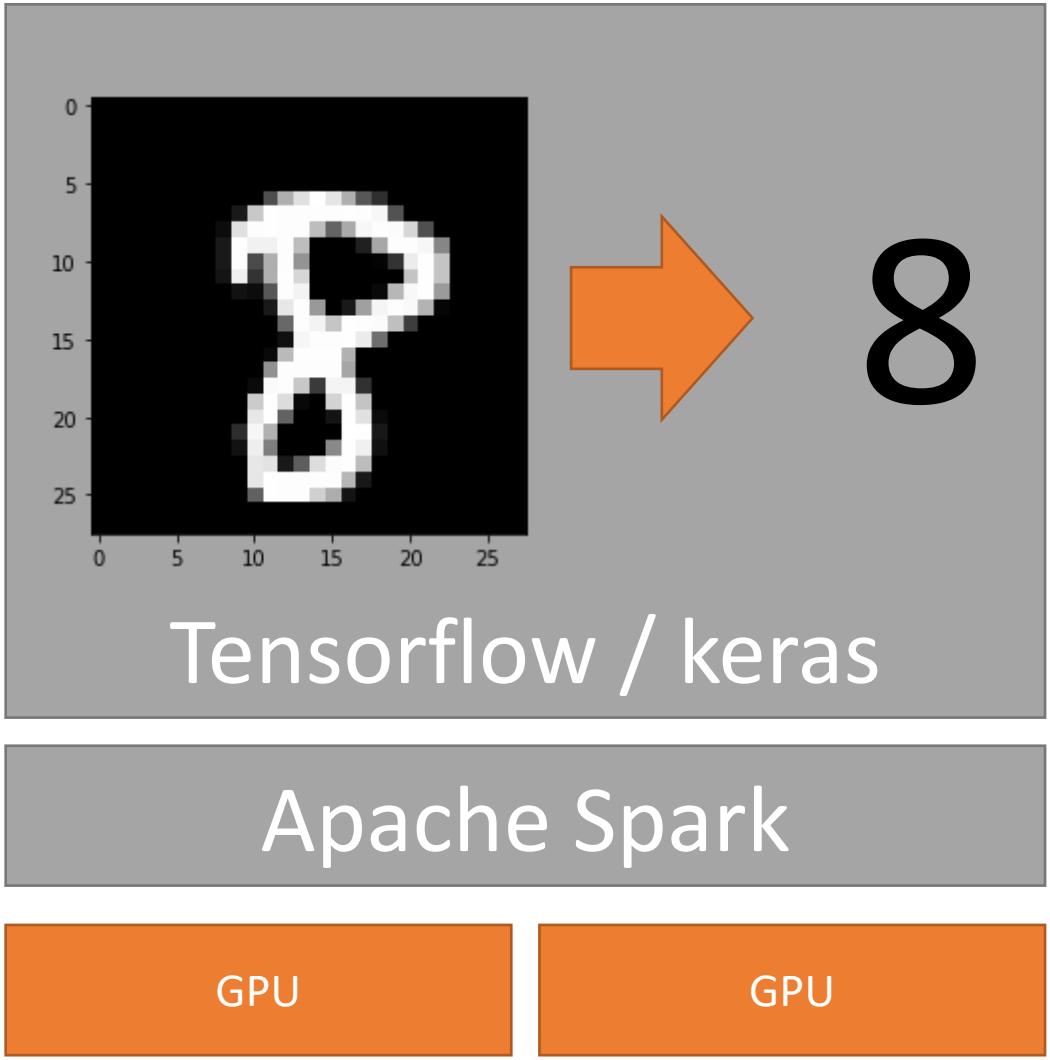
- Nvidia-authored set of Ansible playbooks to deploy clusters of GPU servers
- Kubernetes cluster is an option
- Works well out of the box
- Ongoing cluster support not easy
- Changes not easy

- Python for Ansible
- Hostnames, ssh
- Drivers, k8s plugin
- Kubernetes



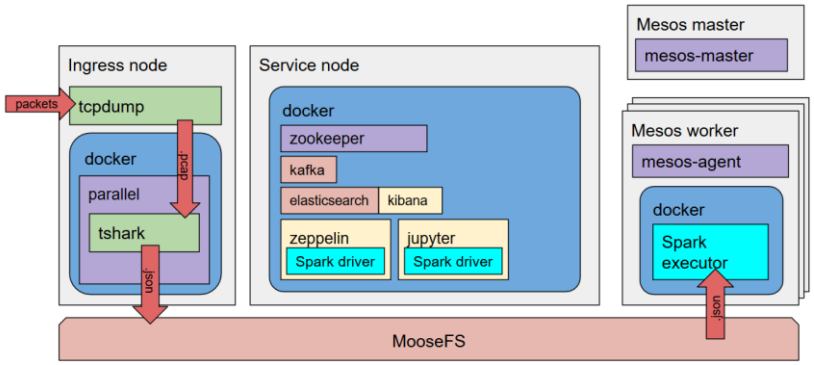
Tensorflow over spark test

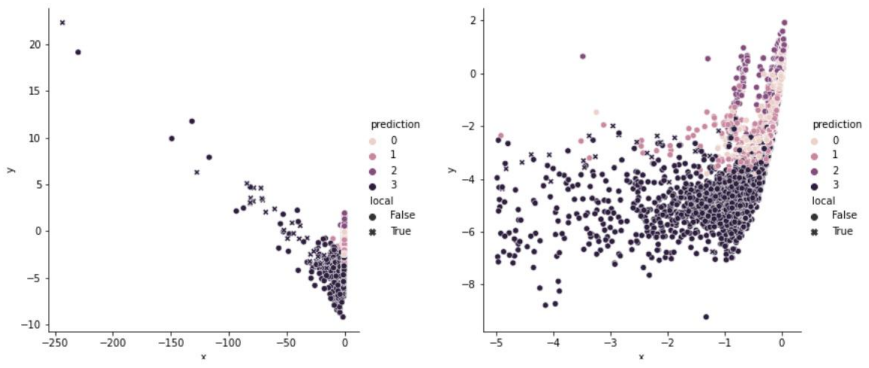
- Distributed learning parallelized on the Tensorflow level
- Spark is just a “pilot” / job manager
- Nvidia Tesla T4



https://github.com/tensorflow/ecosystem/tree/master/spark/spark-tensorflow-distributor

Network packet analysis





- One day of packets from one network subnet
  - 383.3 Gb of packet headers in pcap format
  - 2.2 Tb of json records compressed to 213 Gb by gzip after tcp and ip metadata extraction
    - 8.3 hours by 7 parallel streams of tshark | gzip
- 835 Mb of data in columnar parquet format after selecting only relevant fields
  - 3.5 hours on 3 Spark workers
- TCP/IP feature extraction into AGgregate and Mean (AGM) form, feature encoding into NAGM and final clusterization
  - tens of minutes on 3 Spark workers

Summary

Resources	CPU nodes	CPU nodes + Nvidia T4
Analysis core	Spark 2.4	Spark 3.1
Container repository	Docker Hub	Gitlab.com
Resource management	Mesos	Kubernetes
Configuration management	None with plans for Puppet	Ansible with plans for Puppet
Supporting services	Separate Docker swarm	The same Kubernetes cluster
Coordination	Zookeeper	etcd
Authentication	None	FreeIPA in progress
Notebook service	Zeppelin and Jupyter	Zeppelin and Jupyter

Thank you