



# Accounting and monitoring tools enhancement for Run 3 in the ATLAS distributed computing

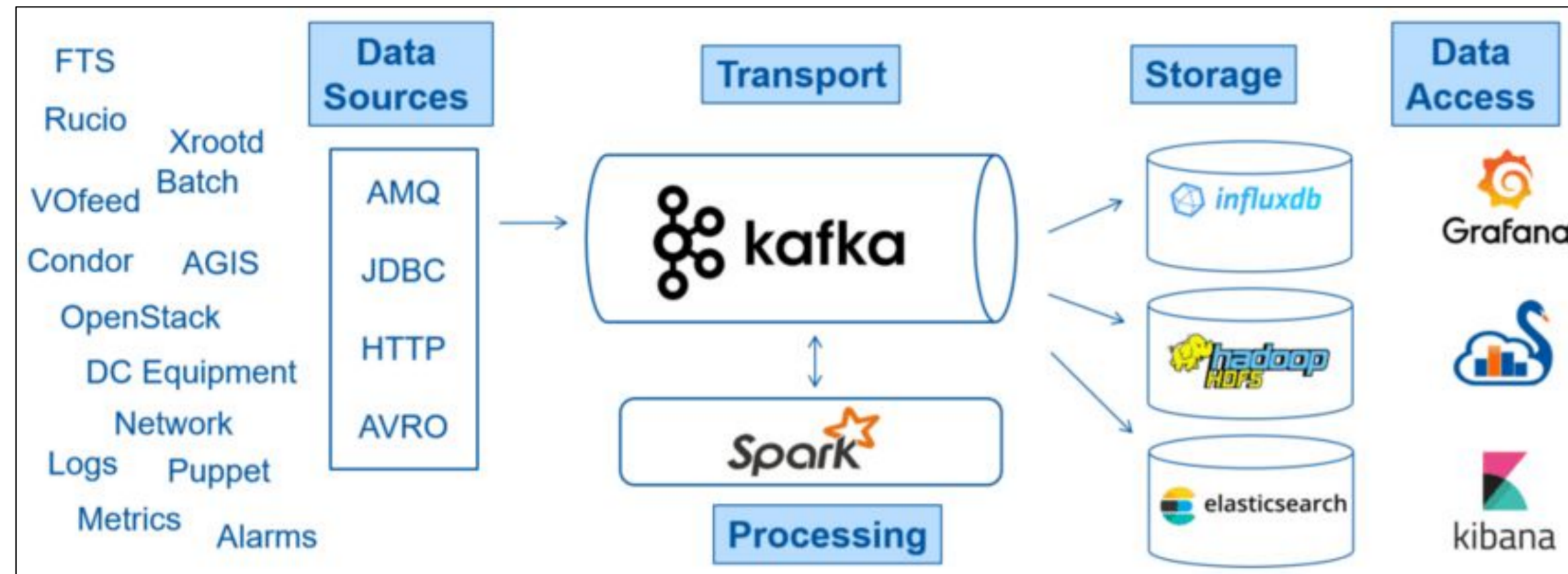


8-13 May 223  
Norfolk, USA

Aleksandr Alekseev, Dario Barberis, Michal Svatos  
on behalf of the ATLAS Computing Activity



## Unified monitoring infrastructure (UMA)



## Introduction

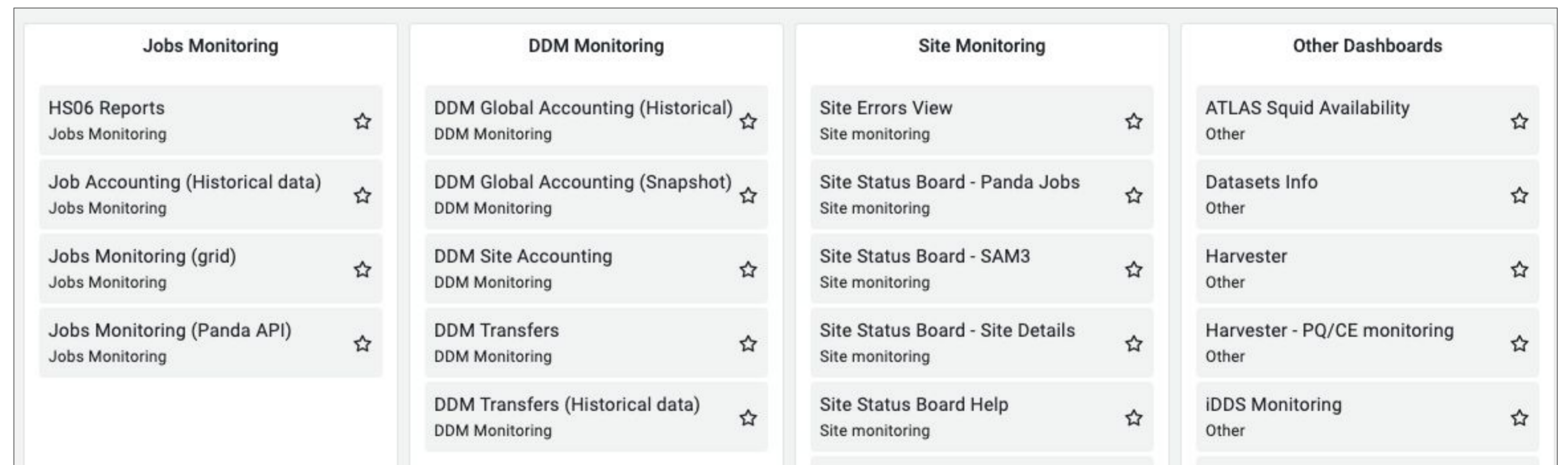
The ATLAS experiment at the LHC utilizes complex multicomponent distributed systems for processing (PanDA WMS) and managing (Rucio) of data. The complexity of the relationships between components, the amount of data being processed and the continuous development of new functionality of the critical systems are the main challenges to consider when creating monitoring and accounting tools able to adapt to the dynamic environment in a short time. To overcome these challenges, ATLAS uses the unified monitoring infrastructure (UMA)\* provided by CERN-IT since 2018, which accumulates information from distributed data sources and then makes it available for different ATLAS distributed computing user groups. The monitoring framework is based on modern open source solutions: Kafka, Spark, ElasticSearch, Kibana, Grafana.

## ATLAS Monitoring and Accounting

The information is displayed using MONIT Grafana. There are several main groups, including tools that are actively used by ATLAS users to identify and fix problems:

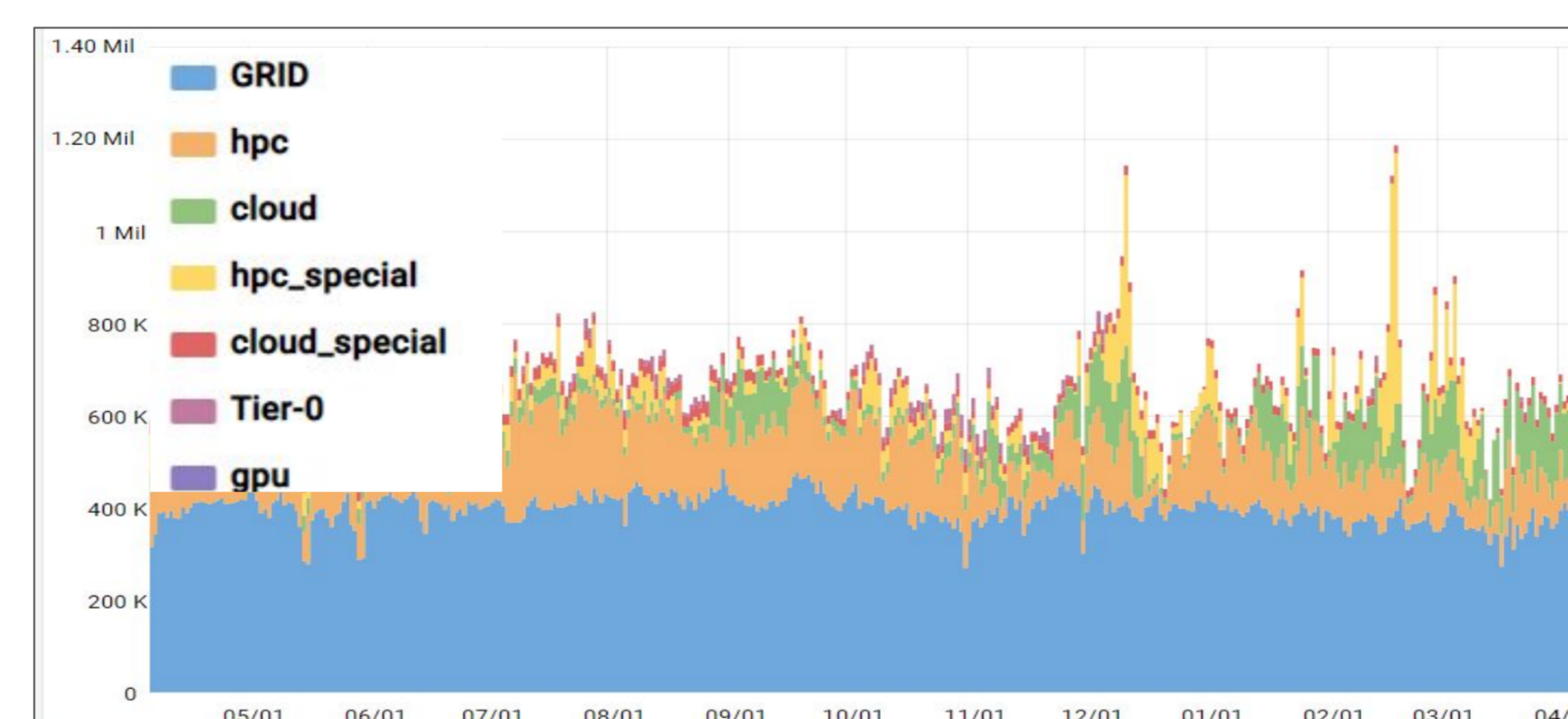
- **Jobs Monitoring and Accounting.** Contains monitoring and accounting tools which provide information about processing computing jobs. It uses snapshots of several tables in PanDA database.
- **DDM Monitoring.** Tools providing information about all operations with data or about data storages. It uses information from Rucio system.
- **Site Monitoring.** The tools from this group combine information from the PanDA and Rucio systems to enable monitoring of the data processing infrastructure.
- **Other.** This group includes other monitoring utilities that provide information about the health of the services in ATLAS infrastructure, or analytics (e.g. data popularity — a priority direction for the development of monitoring tools).

## Centralized space for ATLAS Monitoring and Accounting tools in MONIT Grafana



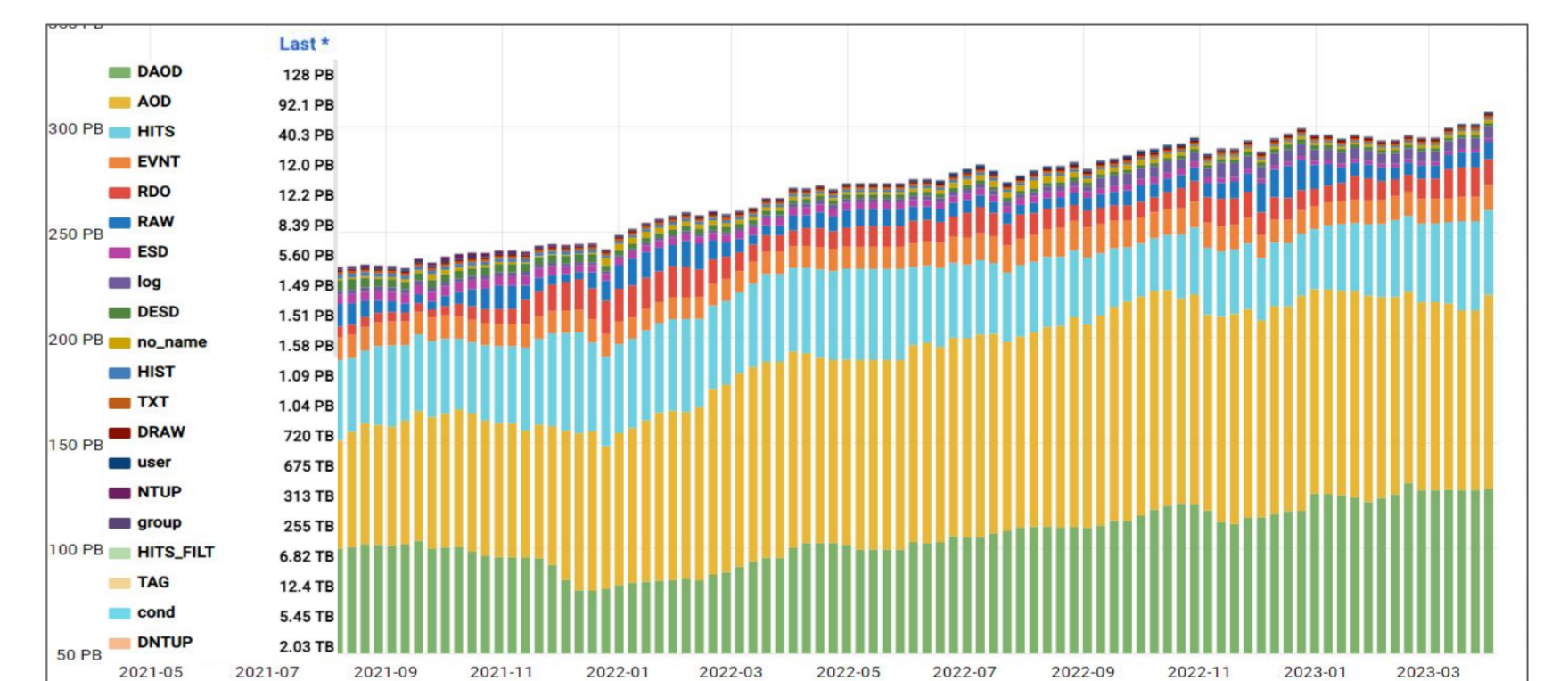
### Jobs Monitoring and Accounting

CPU usage in ATLAS by different payloads

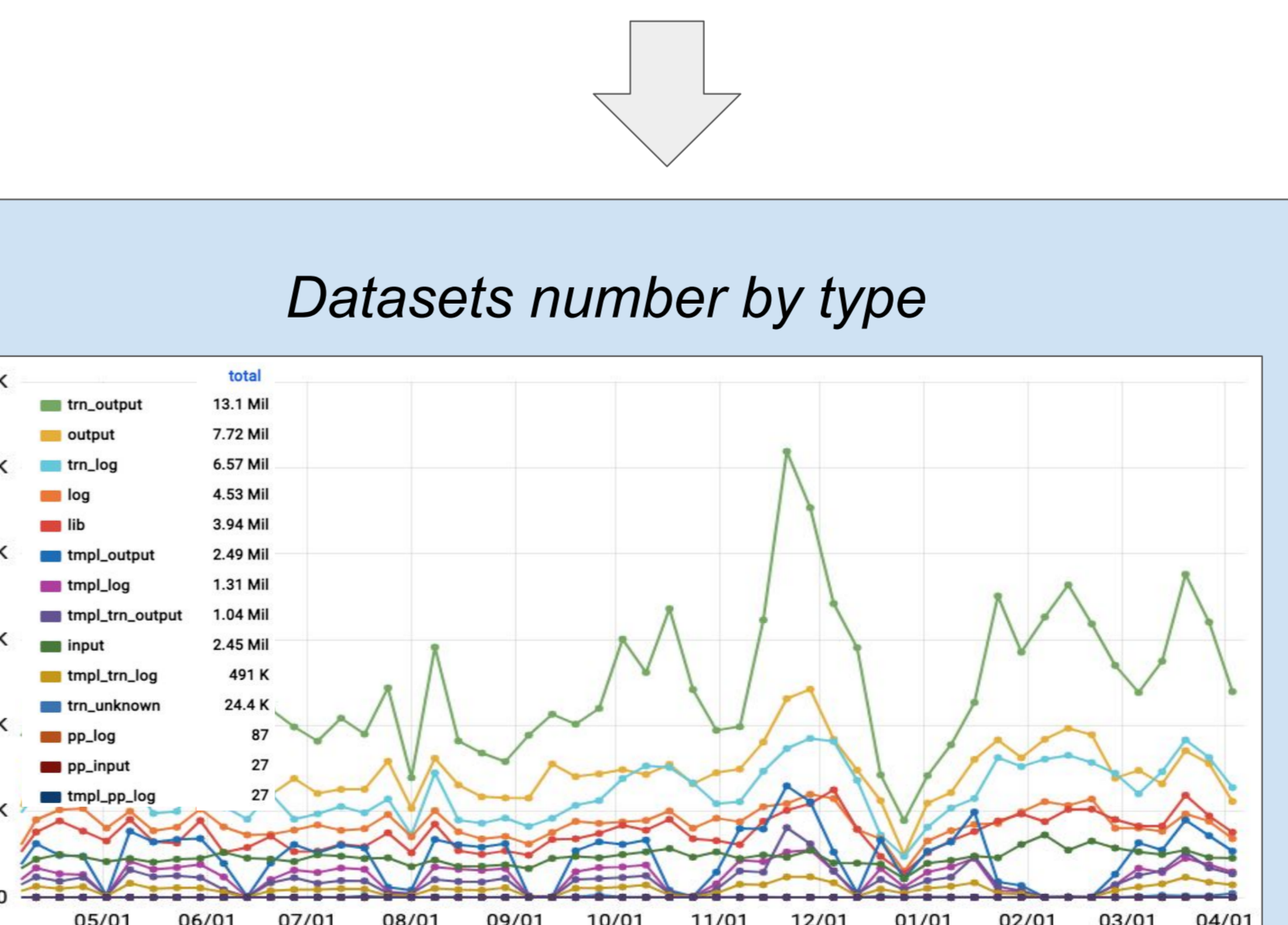


### DDM Monitoring

Data evolution in ATLAS by data format



### Other

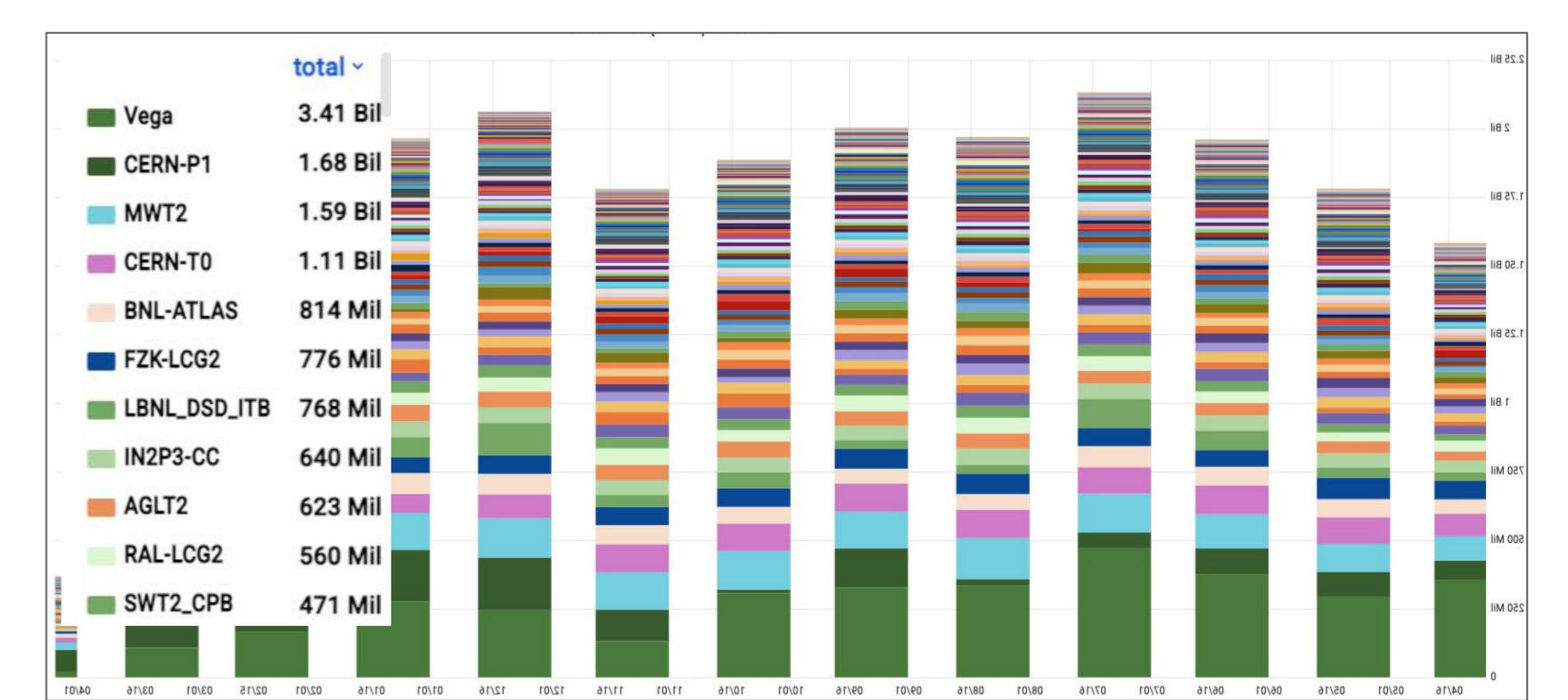


Dynamics of processing input files of the datasets



### Site Monitoring

CPU Consumption: All jobs in Seconds (from Jobs Accounting)



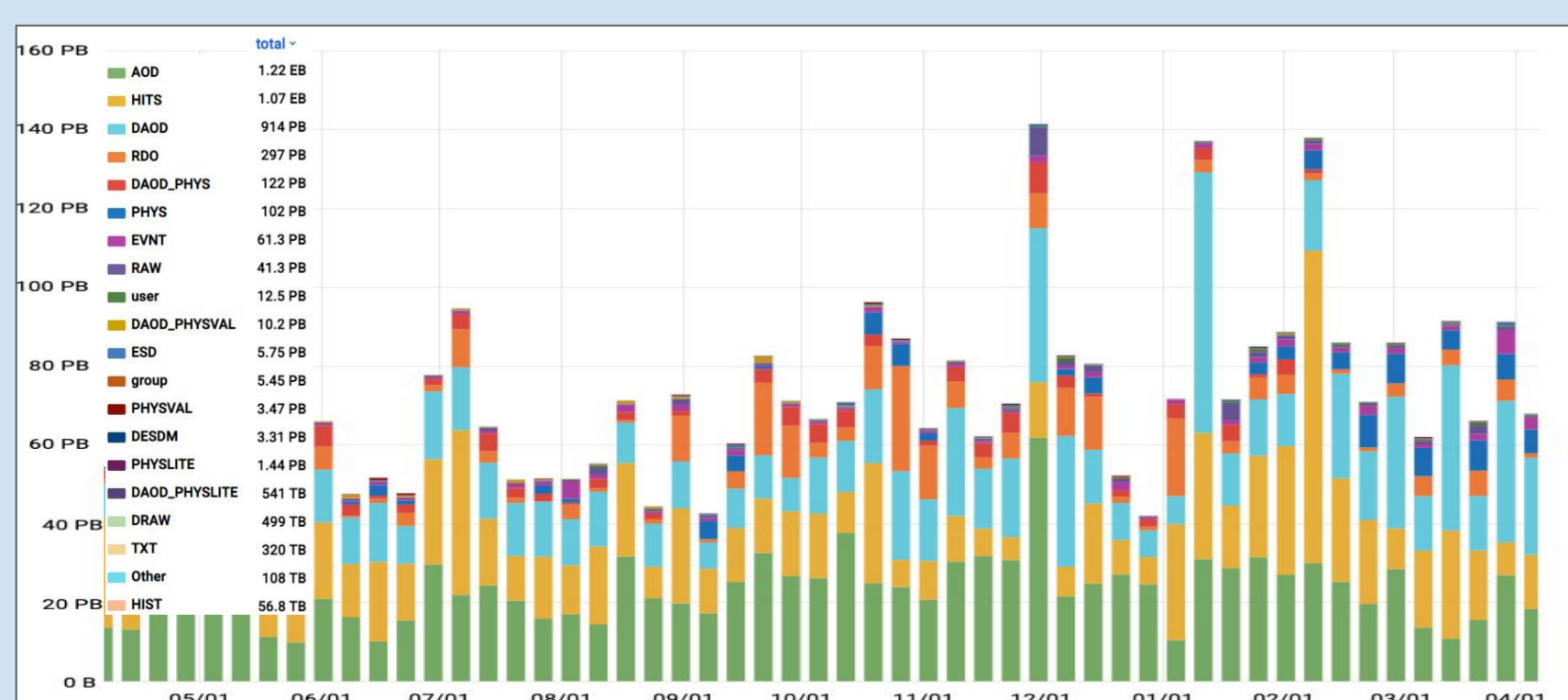
Transfer Efficiency by Site (from DDM Monitoring)



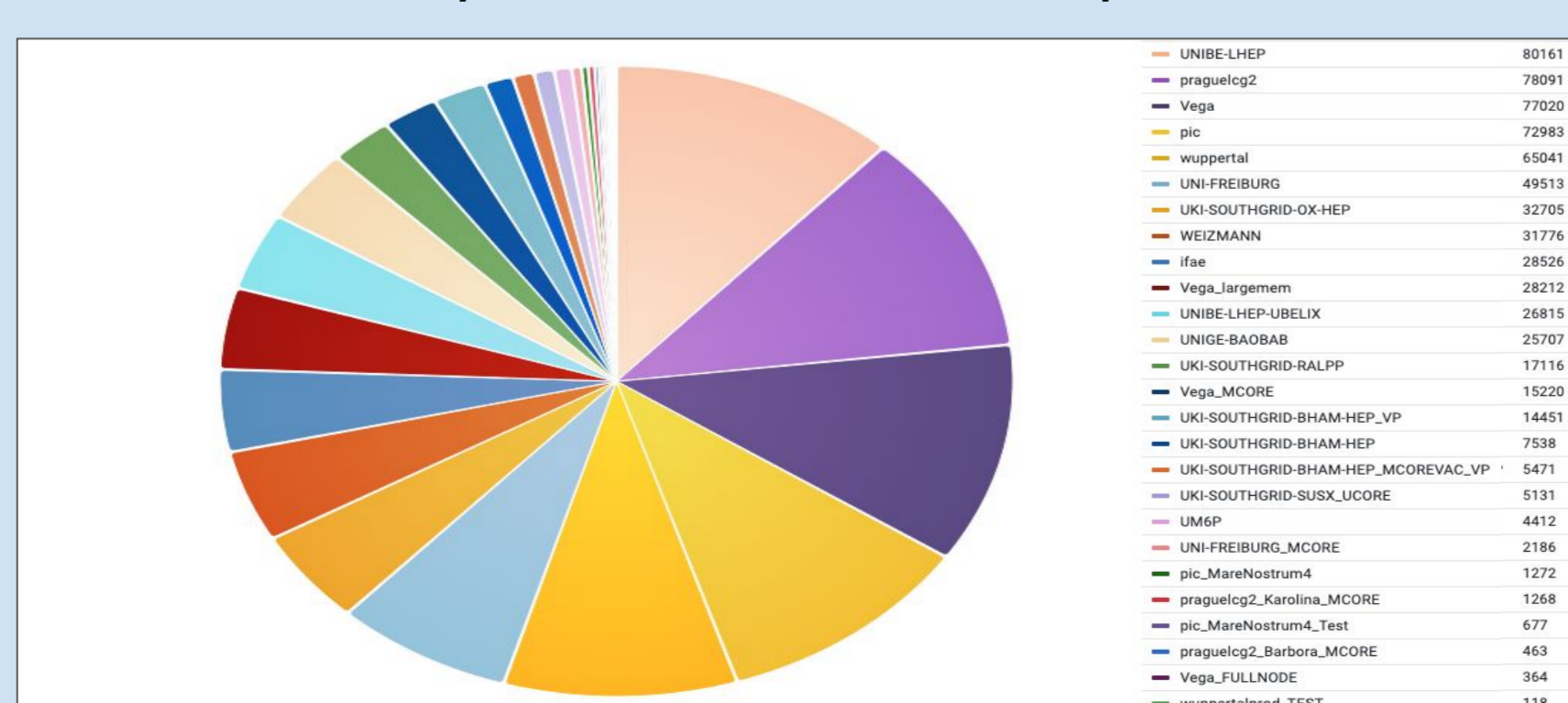
## Dataset Information Tool

This is a new monitoring tool based on metadata about datasets in use by the PanDA system. This makes complete statistics available for the different types of dataset that circulated in the system and the tasks that used or generated them. This provides combined information about tasks, datasets and dataset files information.

Input data consumption by PanDA system



Number of processed tasks on specific sites



\* MONIT: Monitoring the CERN Data Centres and the WLCG Infrastructure. EPJ Web of Conferences 214, 08031 (2019) CHEP 2018