

# Provenance-aware optimization of workload for distributed data production

*Tuesday 11 October 2016 15:00 (15 minutes)*

Distributed data processing in High Energy and Nuclear Physics (HENP) is a prominent example of big data analysis. Having petabytes of data being processed at tens of computational sites with thousands of CPUs, standard job scheduling approaches either do not address well the problem complexity or are dedicated to one specific aspect of the problem only (CPU, network or storage). As a result, the general orchestration of the system is left to the production managers and requires reconsideration each time new resources are added or withdrawn. In previous research we have developed a new job scheduling approach dedicated to distributed data production –an essential part of data processing in HENP (pre-processing in big data terminology). In our approach the load balancing across sites is provided by forwarding data in peer-to-peer manner, but guided by a centrally created (and periodically updated) plan, aiming to achieve global optimality. The planner considers network and CPU performance as well as available storage space at each site and plans data movements between them in order to maximize an overall processing throughput. In this work we extend our approach by distributed data production where multiple input data sources are initially available. Multi-source or provenance is common in user analysis scenario whereas the produced data may be immediately copied to several destinations. The initial input data set would hence be already partially replicated to multiple locations and the task of the scheduler is to maximize overall computational throughput considering possible data movements and CPU allocation. In particular, the planner should decide if it makes sense to transfer files to other sites or if they should be processed at the site where they are available. Reasoning about multiple data replicas allows to broaden the planner applicability beyond the scope of data production towards user analysis in HENP and other big data processing applications. In this contribution, we discuss the load balancing with multiple data sources, present recent improvements made to our planner and provide results of simulations which demonstrate the advantage against standard scheduling policies for the new use case. The studies have shown that our approach can provide a significant gain in overall computational performance in a wide scope of simulations considering realistic size of computational Grid, background network traffic and various input data distribution. The approach is scalable and adjusts itself to resource outage/addition/reconfiguration which becomes even more important with growing usage of cloud resources. The reasonable complexity of the underlying algorithm meets requirements for online planning of computational networks as large as one of the currently largest HENP experiments.

## Tertiary Keyword (Optional)

Reconstruction

## Secondary Keyword (Optional)

Distributed data handling

## Primary Keyword (Mandatory)

Distributed workload management

**Authors:** MAKATUN, Dzmity (Acad. of Sciences of the Czech Rep. (CZ)); Dr RUDOVÁ, Hana (Faculty of Informatics, Masaryk University, Brno, Czech Republic); Dr LAURET, Jerome (Brookhaven National Laboratory); Dr SUMBERA, Michal (Nuclear Physics Institute, Acad. of Sciences of the Czech Rep. (CZ))

**Presenter:** MAKATUN, Dzmity (Acad. of Sciences of the Czech Rep. (CZ))

**Session Classification:** Track 3: Distributed Computing

**Track Classification:** Track 3: Distributed Computing