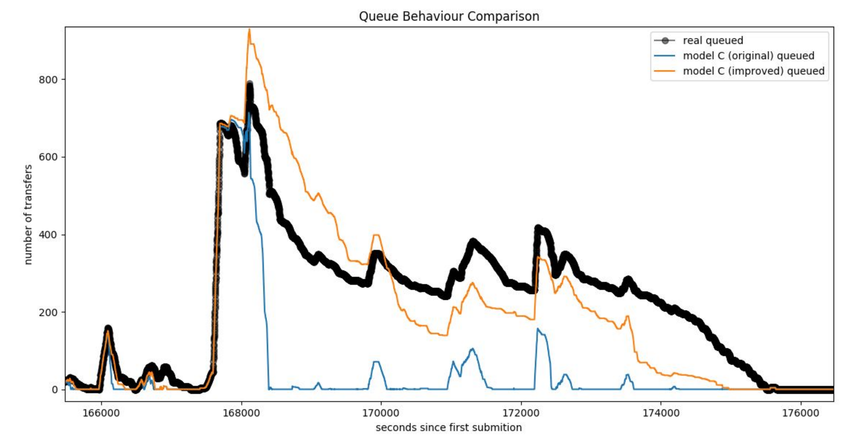
Estimating Time To Complete for ATLAS data transfers

Joaquin Bogado, Fernando Monticelli, Mario Lassnig, Cedric Serfon
on behalf of the ATLAS Collaboration
Contact: *rucio-dev@cern.ch*



About Rucio

- Rucio is the Distributed Data Management system in charge of managing all ATLAS data on the grid.
- The main purpose of the system is to help the Collaboration to store, manage and process LHC data in a heterogeneous distributed environment. Typical tasks are:
  - Transfer data between sites.
  - Delete data from sites.
  - Enforce the experiment computing model

The system manages **290 Petabytes** of physics data across more than **130 data centres** globally, with more than **830 million files**.

Rucio Extensions

- Rucio supports extensions since version 1.14.
- Through the use of extensions it is easy to add ancillary features to Rucio which are not core functionality.
- Transfers Time To Complete (T^3C) extension provides users and Rucio internal with transfer time estimation times at file, dataset, and replication rule level.
- This extension is modular and current algorithms can be replaced with more advanced or complex ones.

Transfers Time To Complete Extension

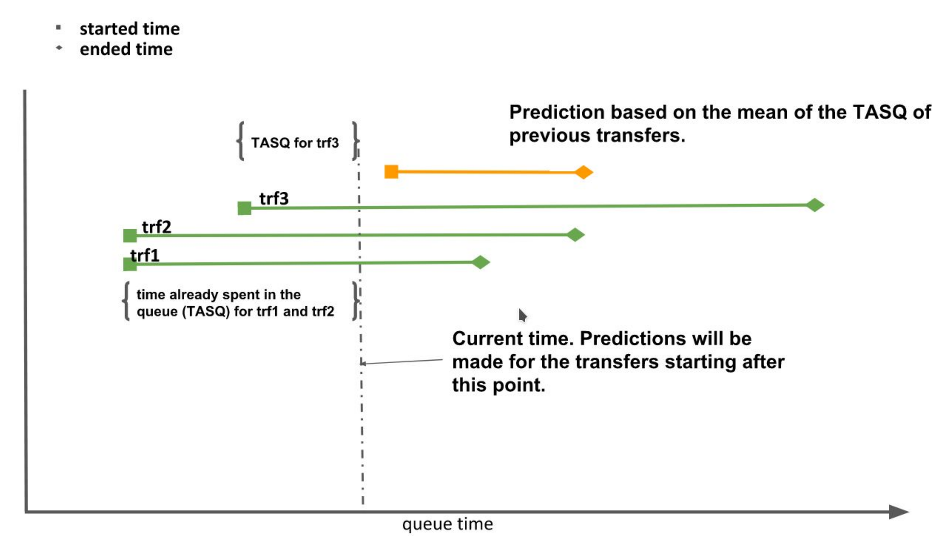
This is an extension aiming to predict the time a data transfer will take to complete. Transfer time is divided in two:

- Queue time: time the transfer waits it's turn in system queues before the actual transfer between sites start.
  - Not related with the size of the file.
  - Several activities have different priorities.
- Network time: time the file is in the network.
  - Can be approximated using the formula size/rate.
  - The average rate of a transfer is very difficult to estimate.

Addressing Queue Time

Queue Time predictions are made based on the Time Already Spent in the Queue (TASQ) metric.

TASQ Exponentially Weighted Moving Average of the active transfers in the link (source, destination pair) is calculated and used as a predictor for the next transfer queue time.



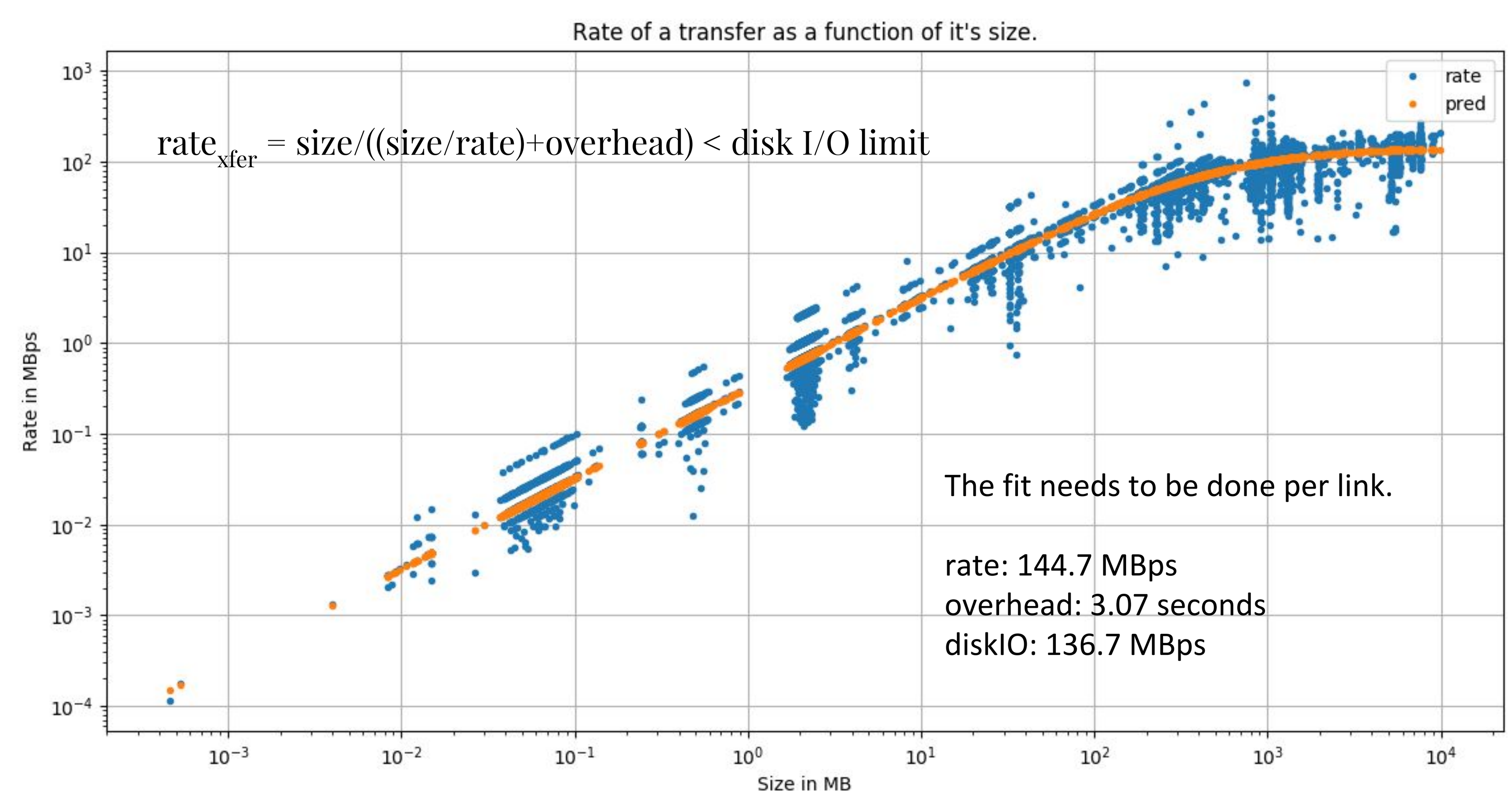
This is the best method found to predict the Queue times in the system, but still room to improve the accuracy of the model.

Addressing Network Time

Current model estimates the rate from the size of the transfer, fitting an ordinary least squares function of three free parameters (rate, overhead and disk read/write limit) over the distribution of rate / size of the transfers.

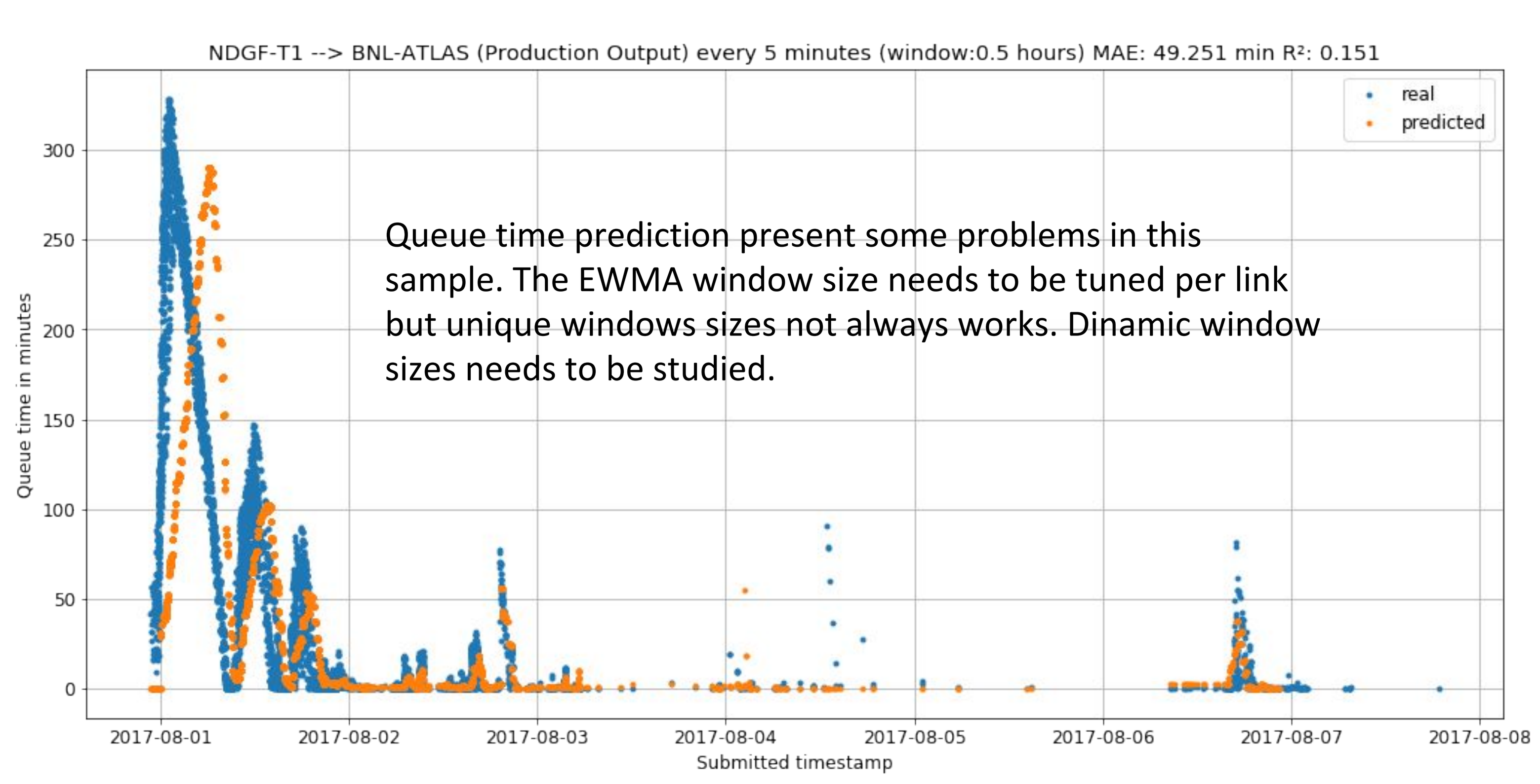
This fit needs to be done per link and being careful that it converge to reasonable values. A subsample of 500 data points usually works better than bigger samples.

This approach is used by other more sophisticated models still in develpment to improve the queue time predictions.



Results

- The extension implements the best model among a big set of models tested to predict the transfers time to complete.
- Network time predictions are in general more accurate than queue time predictions. The correlation of rate with size is important.
- Queue time proven to be difficult but efforts with the TASQ metric is a step in the right direction.



Next steps

- Fine tune the current model will allow to improve the accuracy of the predictions but is a difficult to automate task.
- More complex models for transfer time predictions are being studied.
- Queue time problem is being studied through the simulation of the number of queued transfers.