

From MATS to MATS-NIH: 4X speedup with 4 lines of code

Multivariate Analysis of Transcript Splicing (MATS) is an open tool for transcript slicing that is commonly used by NIH Biowulf users. The MATS package accepts a number of pairs samples of RNA-Seq data to detect differential alternative splicing events. For four samples pairs, the package took around 5 days to generate its results. An NCI investigator wanted to use the package to analyze multiple groups of 32 sample pairs which might take over a month to complete. The HPC/DM group was approached to enable such analysis.

In this blog, we will show how we were able to get an average of 4x speedup in the total runtime of the MATS package which we pushed on Biowulf under the name MATS-NIH.

[CLICK HERE](#) for the full report.