OAK RIDGE National Laboratory

Reproducibility with IO-500… and challenges

George S. Markomanolis

HPC Engineer

Oak Ridge National Laboratory

ORNL is managed by UT-Battelle, LLC for the US Department of Energy

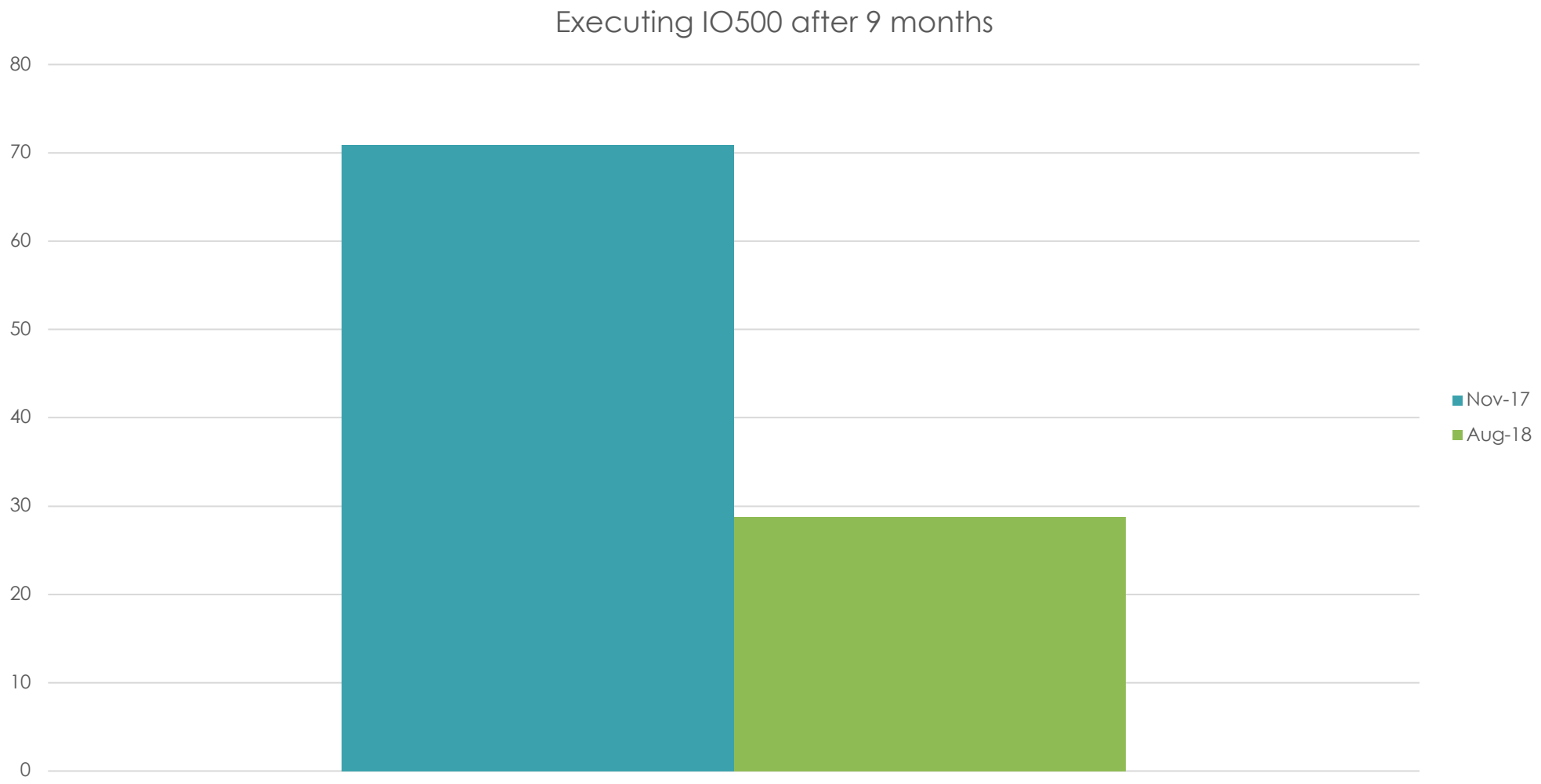
U.S. DEPARTMENT OF ENERGY

Resources

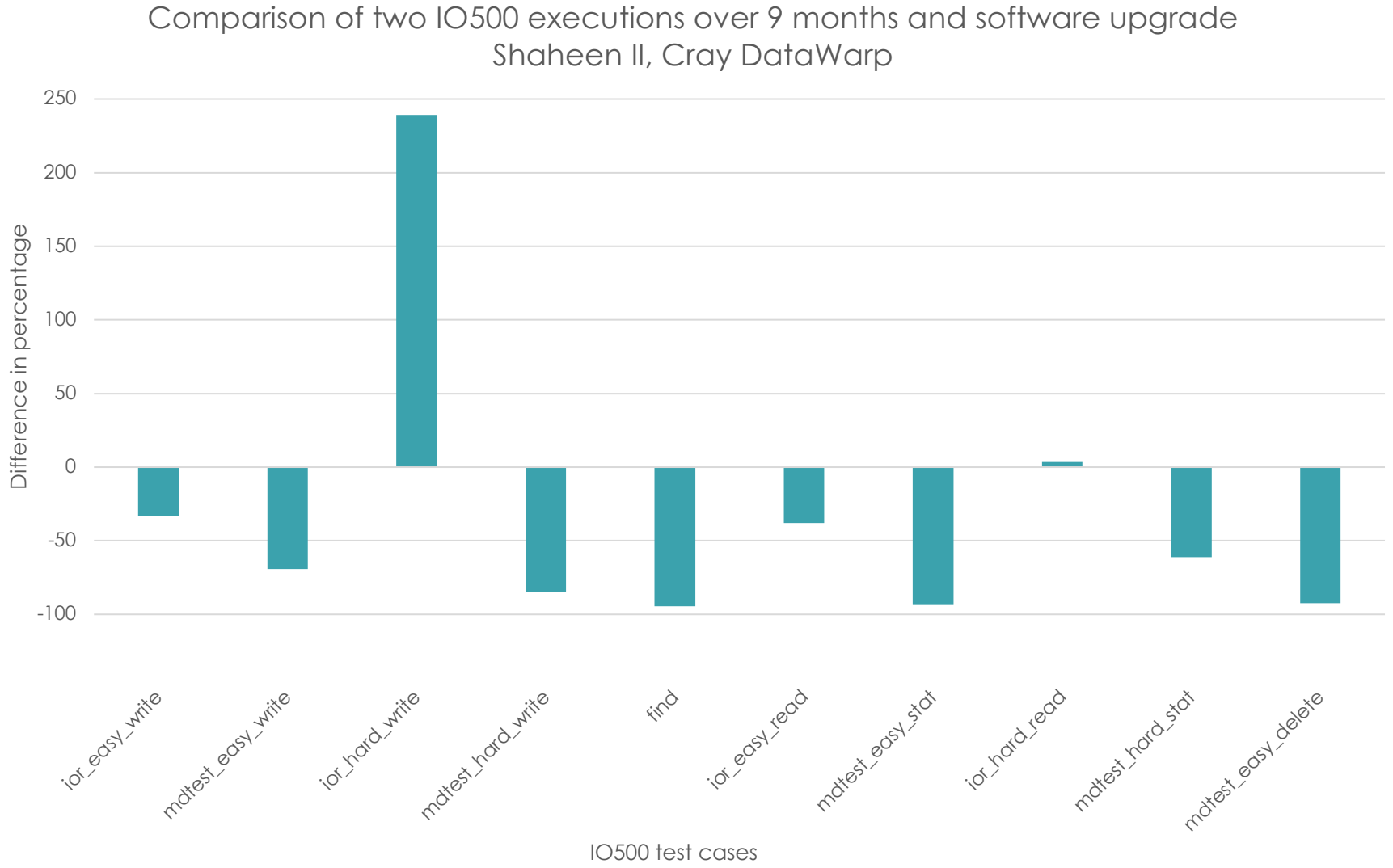
- Shaheen II, a CRAY XC-40 supercomputer
- Using 300 compute nodes
- 256 Cray DataWarp nodes with 2 SSD each one.

Executing IO-500 benchmark after long time

- Using the scripts from the previous submission and same resources
- Expecting valid results
- Hoping to have a submission ready with just one execution
- Job is starting…



The IO500 score was decreased by 60%!



Re-running IO-500 after 9 months, how difficult could it be?

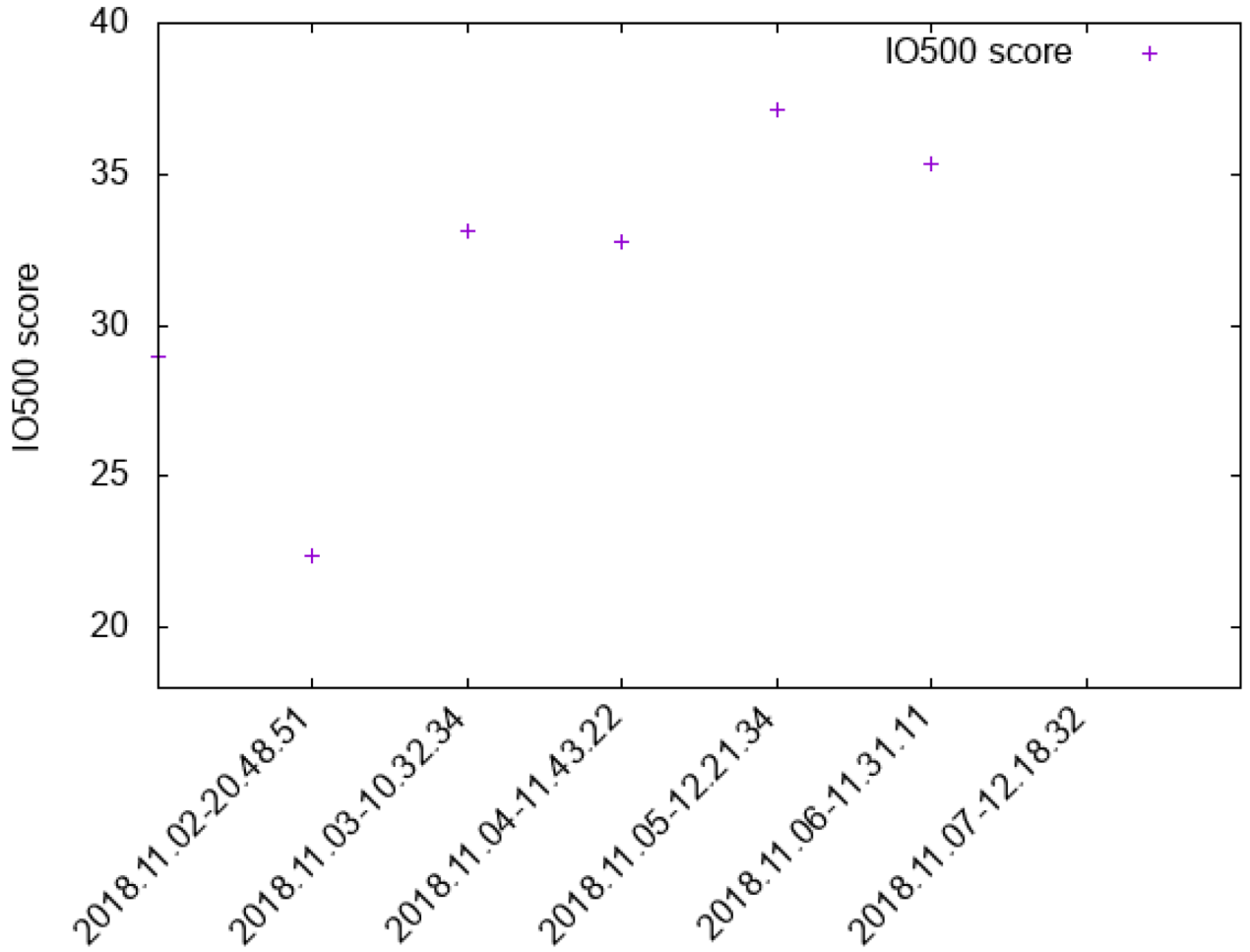
- What is changed on the system?
  - ShaheenII from CLE 5.02 to 6.05
  - New version of Suse Linux
  - New compilers
  - New MPI
  - DataWarp firmware

- Executing IO-500 on CLE 5.02 took 82 minutes
- On CLE 6.05 it didn't finish either in 500 minutes

Use IO-500 in your regression tests!

- Decide the size of the benchmark
- No need for 5 minutes limit for regression tests
- Be sure to use it before and after of any significant change on your system
- Track the results across the timeframe
- Identify which metric is causing problems.

Tracking performance of IO-500 benchmark over time



Acknowledgement

This research used resources of the Oak Ridge Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DE-AC05-00OR22725.

Thank you!

Questions?