

Limits of the HTCondor Transfer System

Edgar Fajardo¹, Frank Wuerthwein¹, Richard Jones², Sandy Philpott³ and Kurt Strosahl³

¹University of California San Diego, ²University of Connecticut,

³Jefferson Laboratory



Motivation

GLUEX experiment wanted to know if they could use the HTCondor Transfer File mechanism to bring the output of their jobs back.

Expected Parameters

Parameter	Value
# Parallel running jobs	20000
Output sandbox	10-100 MB
Input Sandbox	1-10 MB
Job length	8h - 9h

$$O(n, l, s) = \frac{nJobs * size}{length} = \frac{20000 * 90}{9 * 3600} \approx 55.5 \frac{MB}{sec}$$

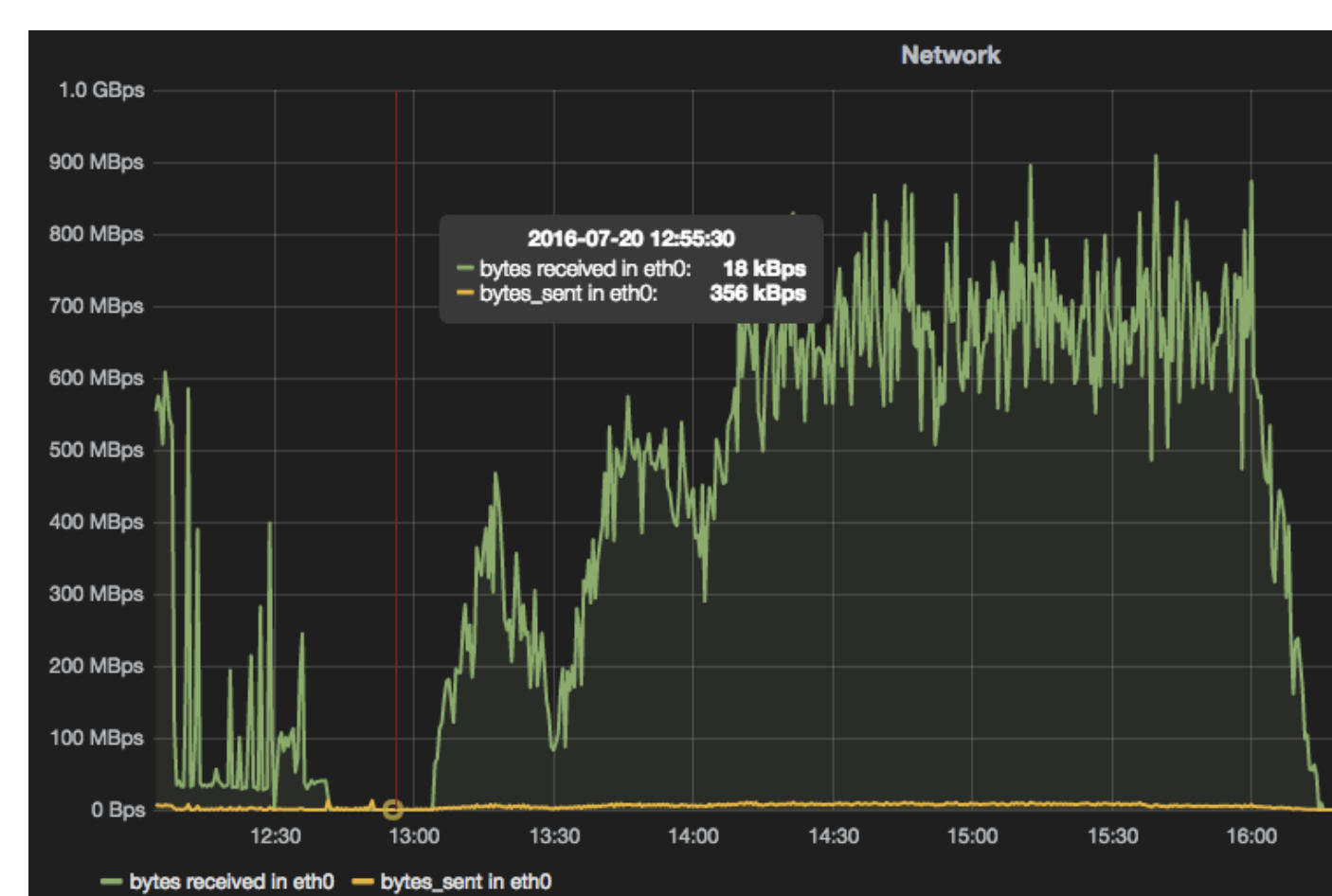
Predicted Output rate



Test SubmitPoint Specs

Hardware	Spec
Memory	128 GB
Network Card	10 Gbit Full duplex
Core count	40
Disk Setup	ssd + raid6

HTCondor can saturate the network



Parameter	Value
# jobs	4000
out Size	250 MB
job length	30 min
Expected throughput	~555.5 MB/s
Achieved throughput	~800 MB/s

The discrepancy is due to retries. HTCondor does not do partial file transfers.

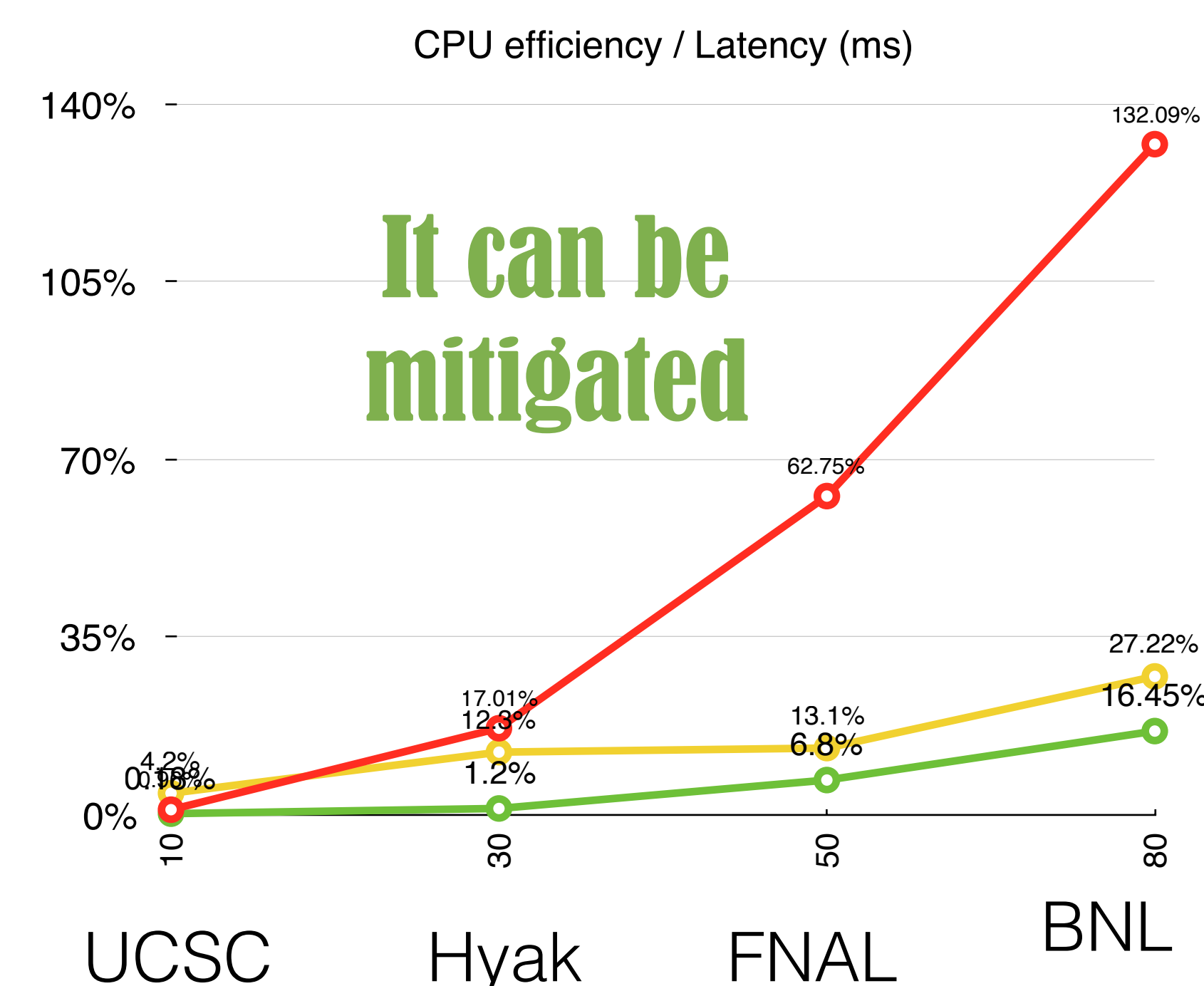
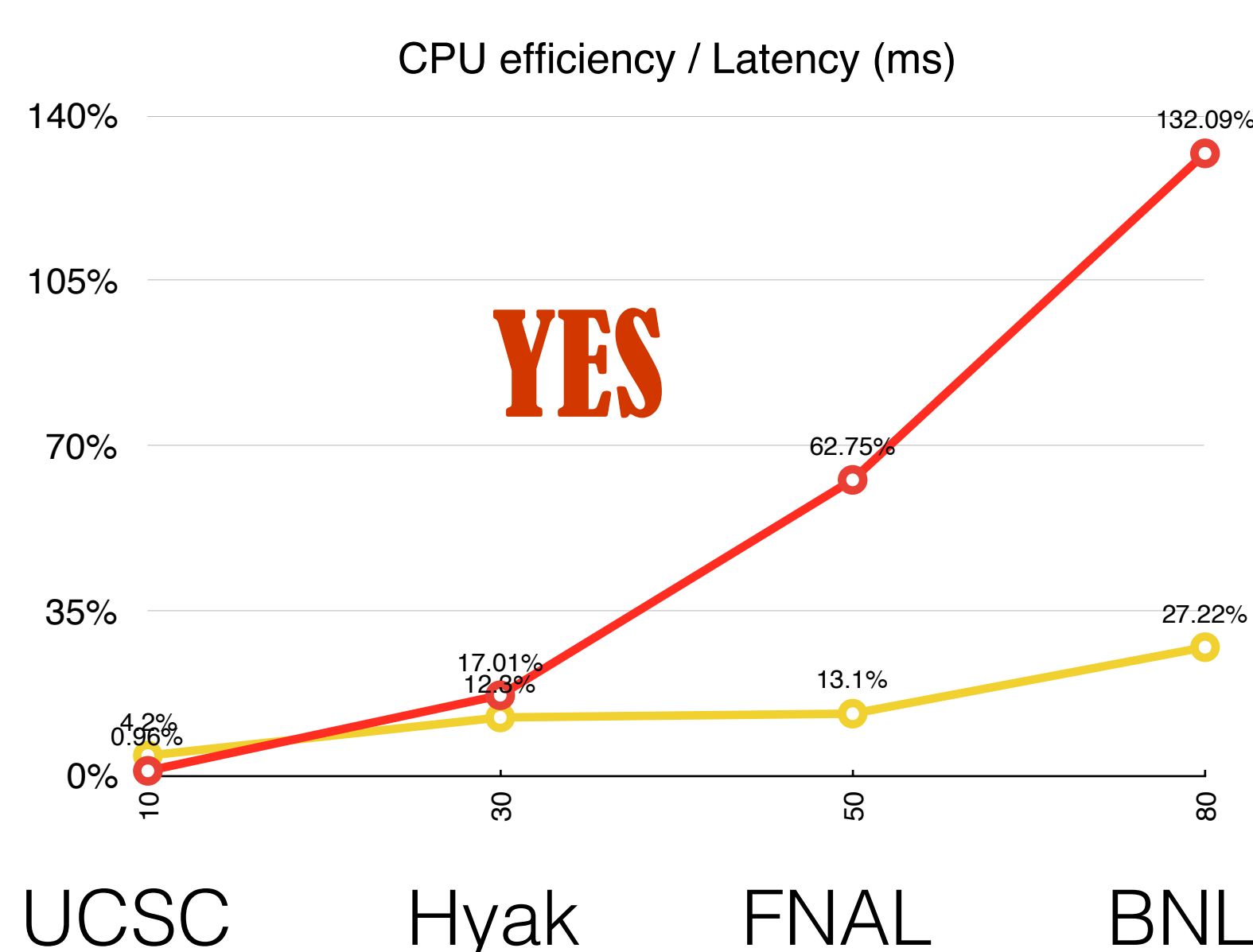
Conclusions

- At the proposed rates GLUEX will efficiently use its resources bringing their output through the HTCondor transfer mechanism
- It also holds true if GLUEX double the rates (either on shorter jobs or twice as much output size).
- Latency greatly influences the efficiency of the HTCondor transfer mechanism but it can be mitigated by tuning the submit host.

Disclaimer

- All tests were done with HTCondor 8.4. New stable release 8.6 is available but the output file transfer mechanism has not been greatly modified.

Does latency affect transfer and CPU efficiency?



- The red line is at two times the expected rate.
- The yellow line is at the expected rate.

The green line is after kernel changes