Throughput Computing

Miron Livny Vials Research Professor John P. Morgridge Professor of Computer Science Director UW Center for High Throughput Computing Technical Director of the OSG







Welcome to HTC23

Throughput Computing 2023

OSG All-Hands Meeting S CHC

HTCondor Week







- **1986** First deployment of (HT)Condor
- 1996 High Throughput Computing (HTC) formulated
- 2005 OSG Consortium established
- 2006 Center for High Throughput Computing (CHTC) established
- 2020 Partnership for Advanced Throughput Computing (PATh) funded by NSF
- 2023 First Throughput Computing event







"The Partnership to Advance Throughput Computing (PATh) project will expand Distributed High Throughput Computing (dHTC) technologies and methodologies through innovation, translational effort, and large-scale adoption to advance the Science & Engineering goals of the broader community."











What is a Feedback Loop?

Feedback Loop means using customer feedback to improve the service and products of a company. The process implies that any feedback received by a company can be worked on to drive improvements to the internal processes, products, or services offered to the customers.

www.zonkafeedback.com/blog/positive-feedback-loop









Throughput Commuting is all about scaling out and therefore about the rate of job handling

During the month of June 2023, the OSPool handled close to 18,000,000 jobs

- **600,000** jobs per day
- **25,000** jobs per hour
- **410** jobs per minute
 - 7 jobs per second



PATTE services and technologies enable federation of computing capacity

 More than 50 institutions contribute to the OSPool capacity provided by more than 70 sites

PATE services and technologies enable effective access to remote datasets

• More than **160** datasets federated by the Open Science Data Federation (OSDF)







Scaling out with federated capacity and datasets means a lot of file transfers

- In 2022, the OSDF executed more than 1B file transfers (32 transfers per second)
- In June 2023, Jobs executed by the OSPool required 320M file transfers (120 transfers per second)









"Broader Impact – We firmly believe in dHTC as an accessible computing paradigm which supports the **democratization** of research computing to include researchers and organizations otherwise underrepresented in the national CI ecosystem. Our work is founded on universal principles like sharing, autonomy, unity of purpose, and mutual trust."











Computing and data infrastructure—part of what the NSF terms *cyberinfrastructure*—is relevant to virtually all fields and disciplines and is essential for 21st-century science and research. Thus, progress (or lack of progress) in this domain is integral, impactful, and informative to any efforts to address the missing millions across the science, technology, engineering, and mathematics (STEM) workforce.







In June 2023 **236** researchers from 90 projects placed OSPool Jobs. What can we do more/different to scale out adoption of **Throughput Computing?**





