Synapse:
Facilitating large-scale data management in research contexts

Gerrick Teague (feebleone@gmail.com) & Daniel Andresen (dan@ksu.edu)

Kansas State University

Daniel Andresen, Ph.D.

Daniel Andresen, Ph.D. is a professor of Computing & Information Sciences at Kansas State University, Michelle Munson-Serban Simu Keystone Research Scholar, and Director of the Institute for Computational Research. His research includes embedded and distributed computing, biomedical systems, and high performance scientific computing. Dr. Andresen coordinates the activities of the K-State research computing cluster, Beocat, and advises the local chapter of the Association for Computing Machinery (ACM). He is a National Science Foundation CAREER award winner, and has been granted research funding from the NSF, the Defense Advanced Research Projects Agency (DARPA), and industry. He is a member of the Association for Computing Machinery, the IEEE Computer Society, the Electronic Frontier Foundation, the American Society for Engineering Education, and has been an XSEDE Campus Champion since 2011.

Cognitive and Neurobiological Approaches to Plasticity (CNAP) Center Overview

Kimberly Kirkpatrick (kirkpatr@ksu.edu)
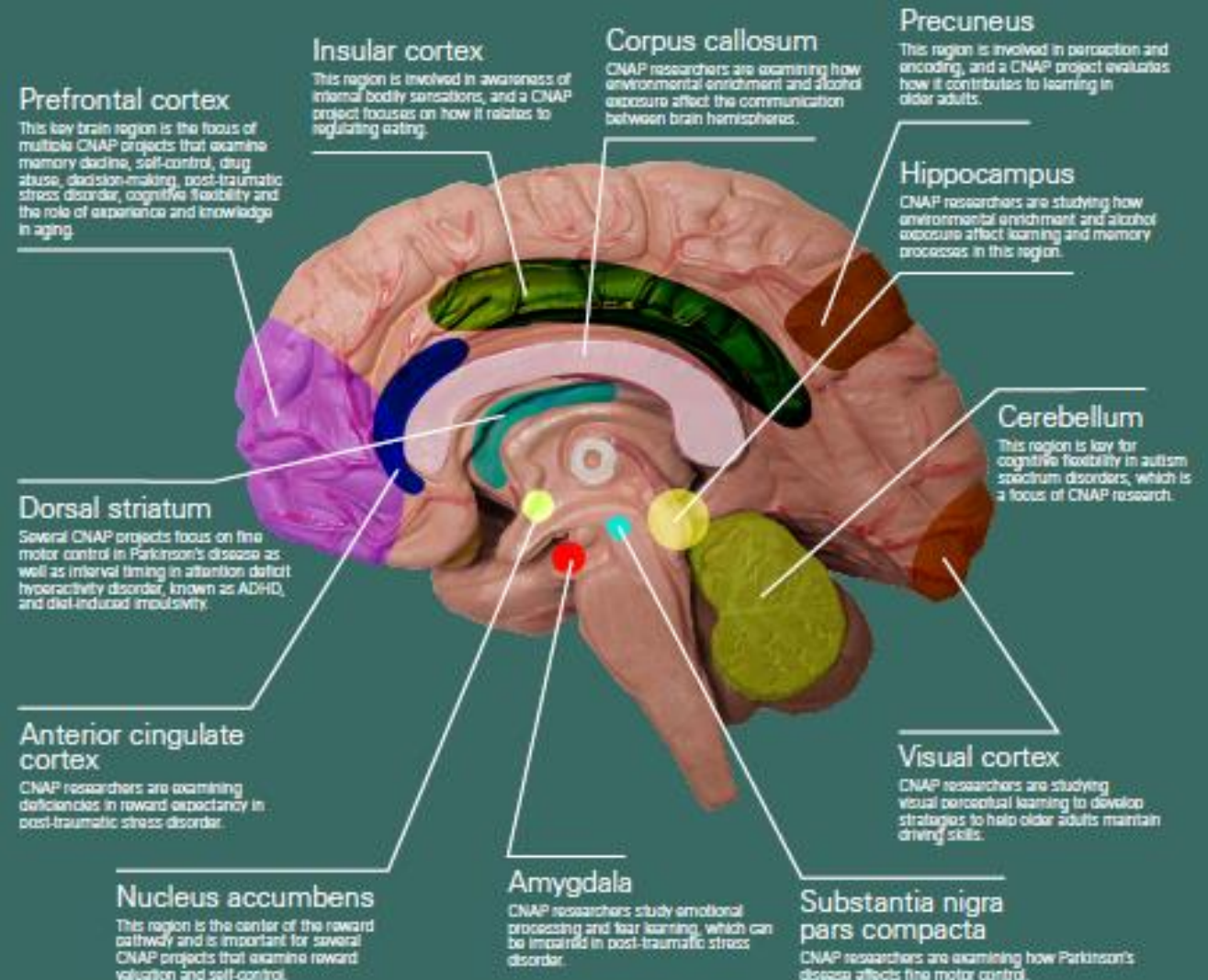
Kansas State University

What is CNAP?

- COBRE = CENTERS OF BIOMEDICAL RESEARCH EXCELLENCE
- CNAP is funded by a $10.6M, five-year award from the IDeA program and administered by NIGMS
- CNAP is early in Year 5 of Phase 1
  - Renewal is available for two further 5-year phases
- COBRE Program Goals
  - Faculty development
  - Infrastructure investments
  - Growth of thematic research

CNAP Research

- Alcohol and substance abuse
- Obesity
- Autism spectrum disorders
- Parkinson's disease
- Hearing disorders
- Aging
- Post-traumatic stress disorder
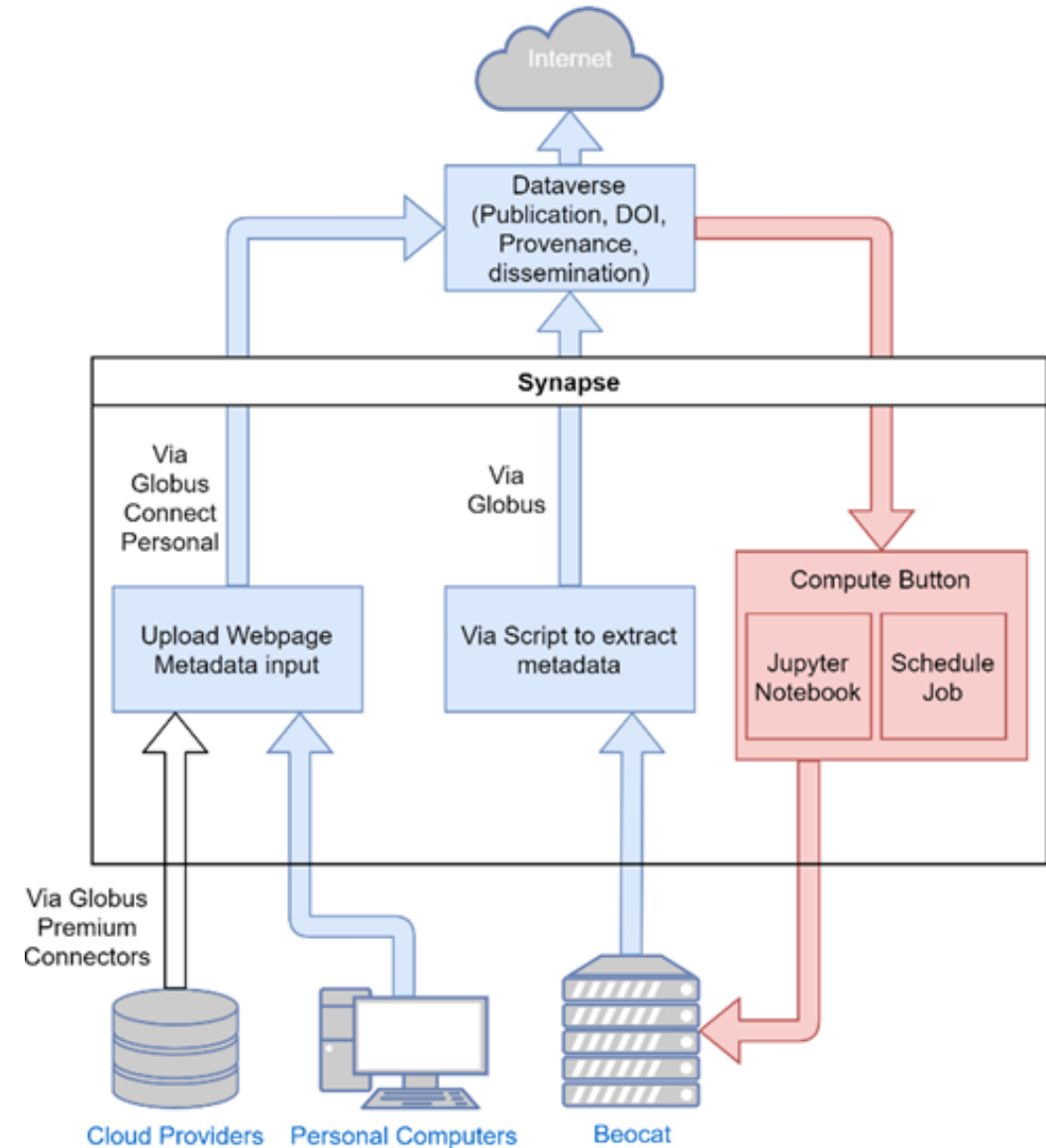- Attention deficit hyperactivity disorder

The brain

Research at the Kansas State University Cognitive and Neurobiological Approaches to Plasticity Center, or CNAP, addresses many regions of the brain. CNAP involves more than 63 collaborators working on a variety of projects in these brain regions.

Prefrontal cortex
This key brain region is the focus of multiple CNAP projects that examine memory decline, self-control, drug abuse, decision-making, post-traumatic stress disorder, cognitive flexibility and the role of experience and knowledge in aging.

Insular cortex
This region is involved in awareness of internal bodily sensations, and a CNAP project focuses on how it relates to regulating eating.

Corpus callosum
CNAP researchers are examining how environmental enrichment and alcohol exposure affect the communication between brain hemispheres.

Precuneus
This region is involved in perception and encoding, and a CNAP project evaluates how it contributes to learning in older adults.

Hippocampus
CNAP researchers are studying how environmental enrichment and alcohol exposure affect learning and memory processes in this region.

Cerebellum
This region is key for cognitive flexibility in autism spectrum disorders, which is a focus of CNAP research.

Dorsal striatum
Several CNAP projects focus on fine motor control in Parkinson's disease as well as interval timing in attention deficit hyperactivity disorder, known as ADHD, and diet-induced impulsivity.

Anterior cingulate cortex
CNAP researchers are examining deficiencies in reward expectancy in post-traumatic stress disorder.

Nucleus accumbens
This region is the center of the reward pathway and is important for several CNAP projects that examine reward valuation and self-control.

Amygdala
CNAP researchers study emotional processing and fear learning, which can be impaired in post-traumatic stress disorder.

Substantia nigra pars compacta
CNAP researchers are examining how Parkinson's disease affects fine motor control.

Visual cortex
CNAP researchers are studying visual perceptual learning to develop strategies to help older adults maintain driving skills.
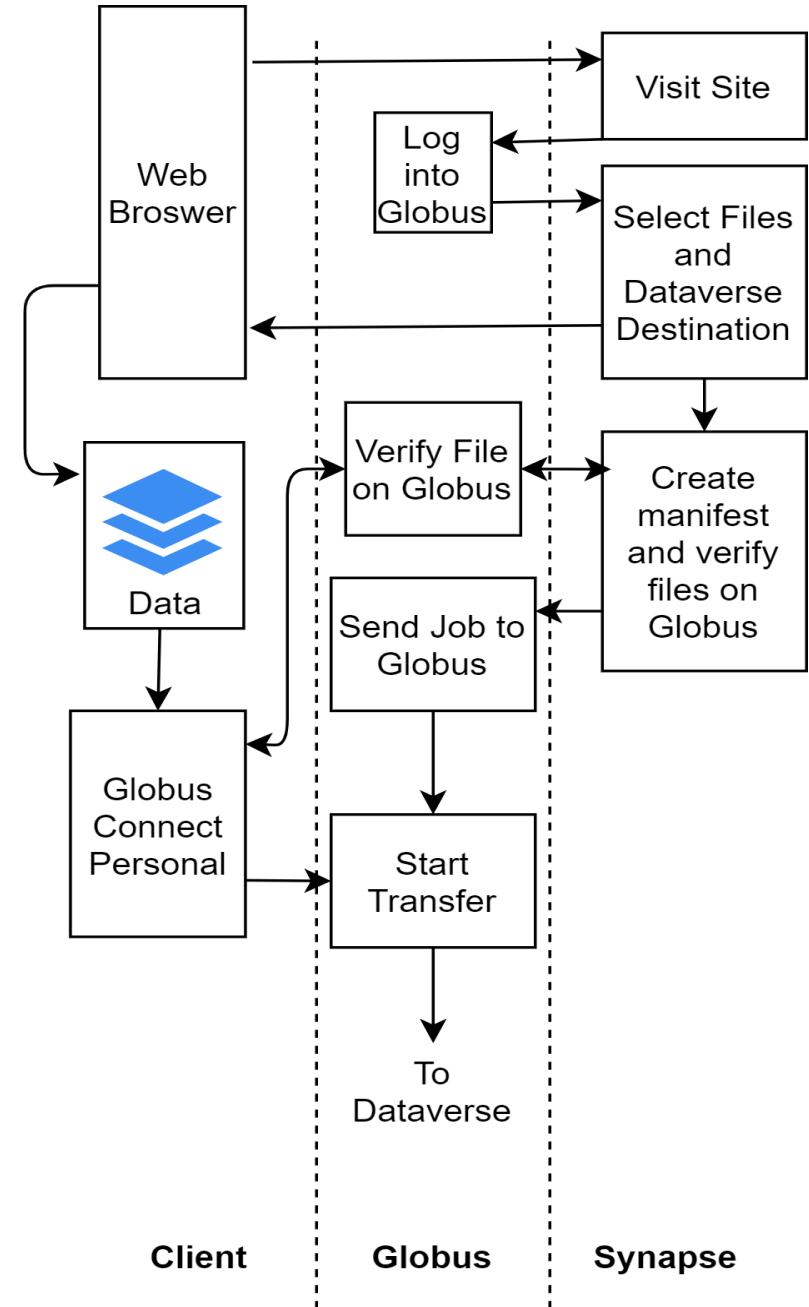
Synapse architecture

- Automated data collection/publication

- High-performance
  - using Globus on backend

- Easy-to-use GUI

- Built on industry standard components

Synapse workflow

- User-controlled data upload/download
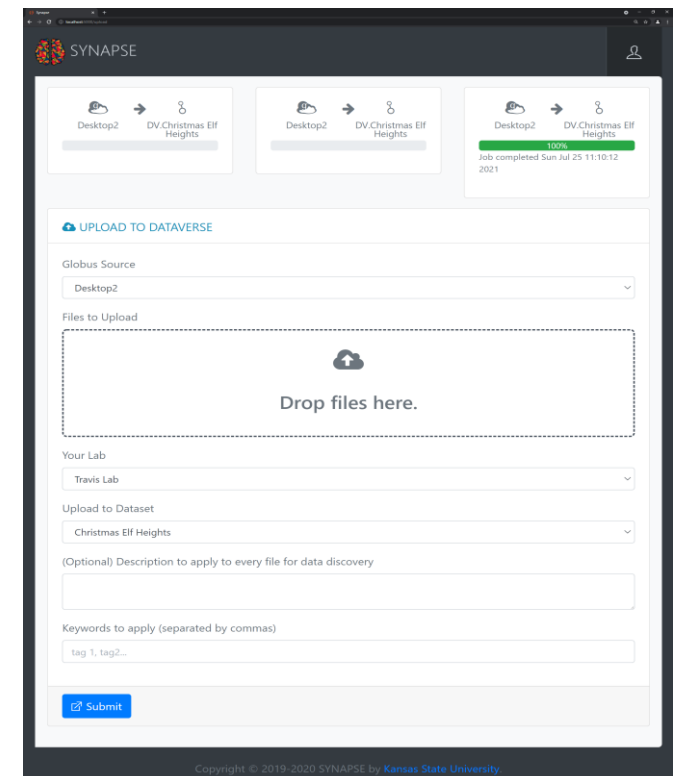- Separation of metadata and data transfer
- OS-independent backend

Synapse UI

- Automatic metadata extraction
  - Pluggable modules to customize automatic metadata on a per-lab basis
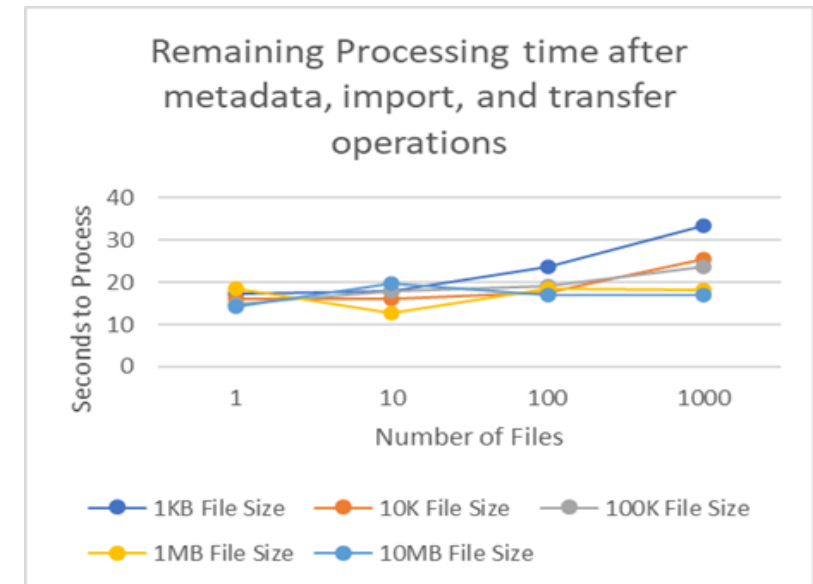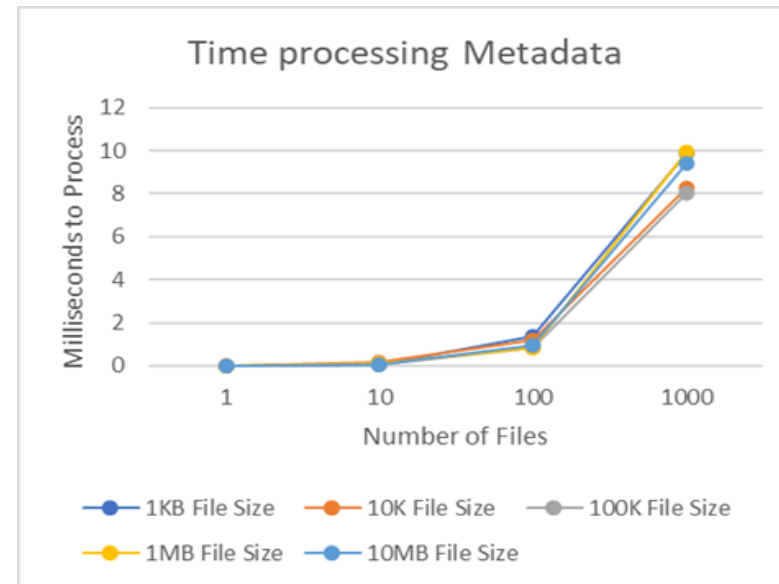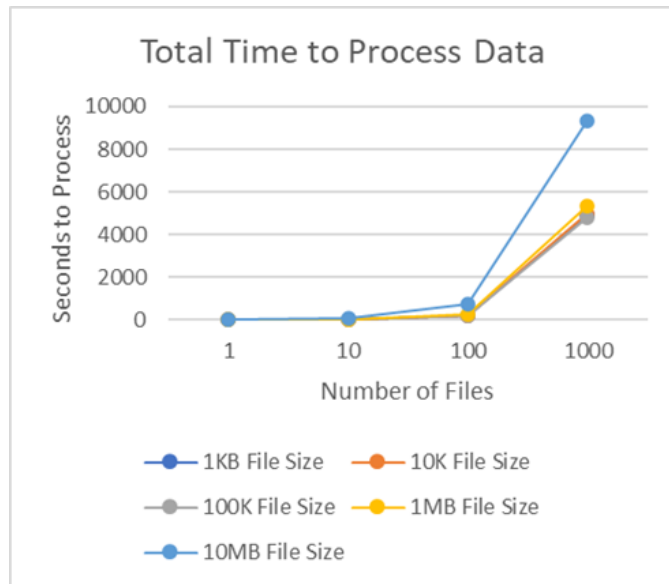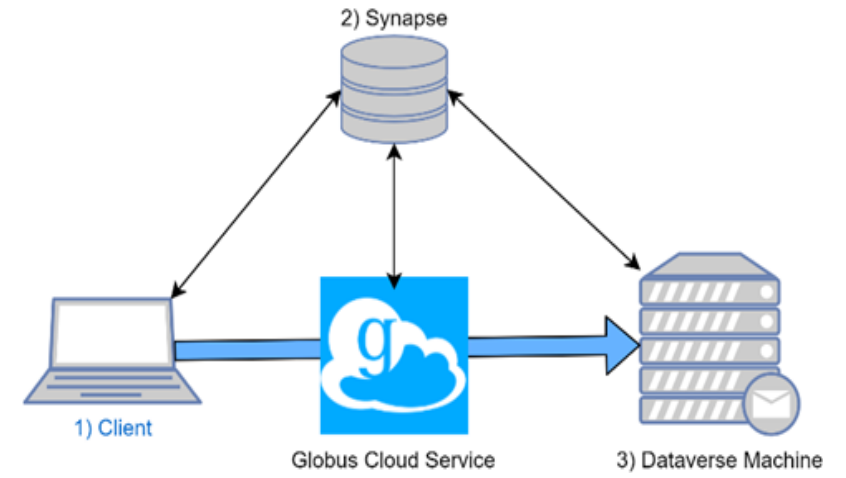- Familiar GUI – drag & drop
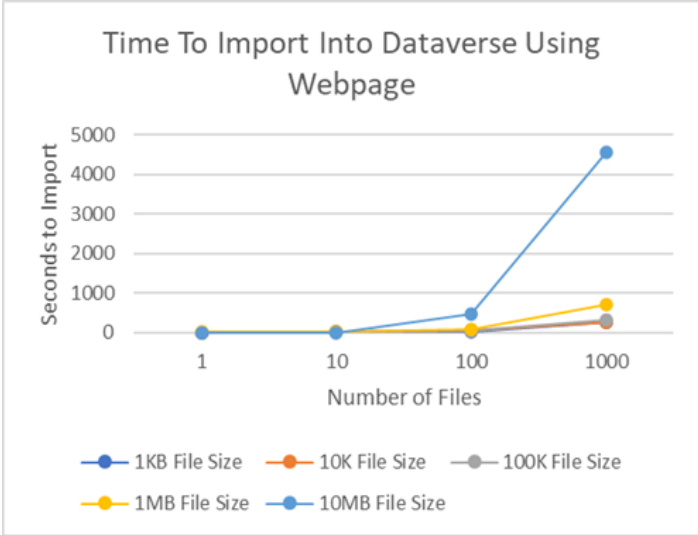

Dataverse data with and without metadata.

Evaluation – Synapse

- Performant for expected size and quantity of files
- Low overhead for metadata/transfer operations
- Globus usage provides efficient point-to-point transfers

Total Time to Process Data

Time processing Metadata

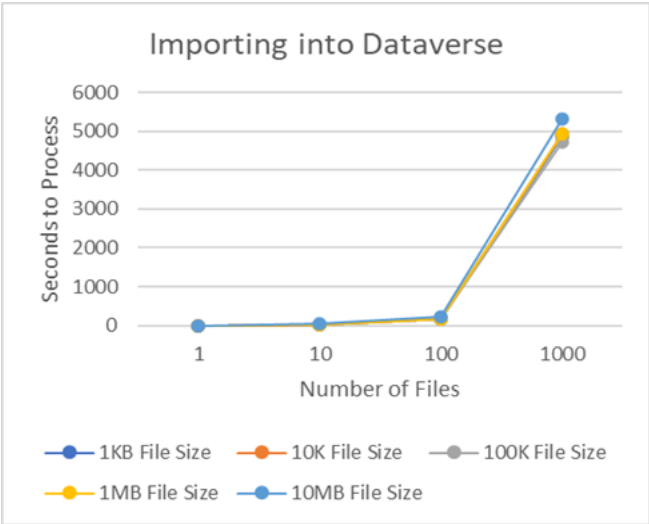Remaining Processing time after metadata, import, and transfer operations
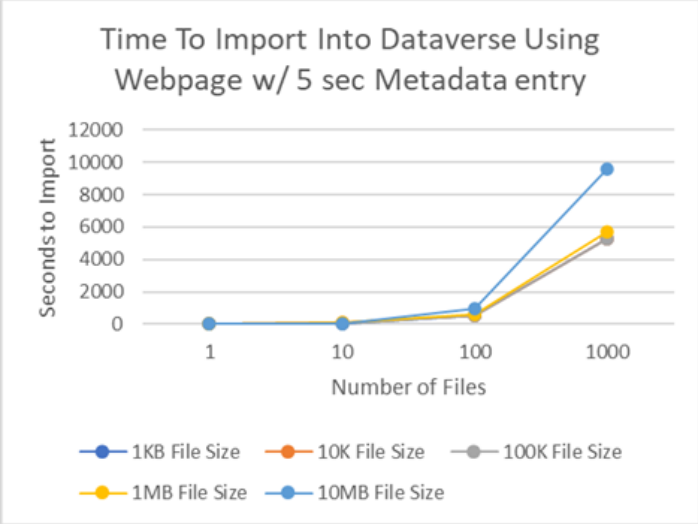
Evaluation – Dataverse Web UI
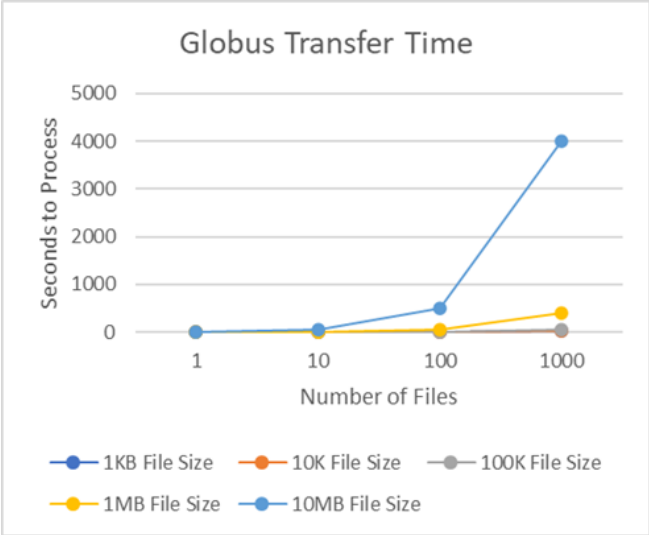
Globus transfer time.

Time to import data using Dataverse website.

Seconds to import into Dataverse after data is transferred.

Dataverse website entry with metadata time allowance.

Conclusions

- Synapse – web-based tool to store, manage, process and publish data using familiar technologies

- Automated metadata extraction

- Resilient to network faults

- Competitive performance to native Dataverse app in our setting

- Code available at:

  - https://github.com/cnap-cobre/synapse-globus