Wrangler:

A Transformational Data Intensive Resource for the Open Science Community

Dan Stanzione, Director and PI
dan@tacc.utexas.edu
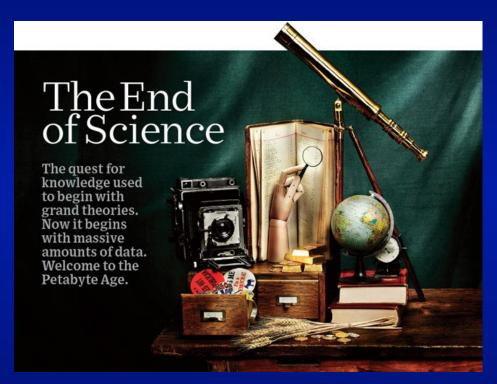Niall Gaffney
Director for Data Intensive Computing
ngaffney@tacc.utexas.edu

THE UNIVERSITY OF TEXAS AT AUSTIN
TEXAS ADVANCED COMPUTING CENTER

Not Just Simulation Any More…

- Remarkable success with large scale computational simulation has led to an explosion of new data.

- Modern science and engineering therefore is about **managing** and **analyzing** this data as well

- So, TACC's strategy is about computation, data, and data analysis.

The End of Science

The quest for knowledge used to begin with grand theories. Now it begins with massive amounts of data. Welcome to the Petabyte Age.

Goals of the Wrangler Project

- Respond directly to recent NSF and PCAST reports, which recommend that:

  - *High end ... data ... and sensor-based systems and the associated user support are needed for transformative science; ... networking, interoperable data systems and mining, including a focus on sustainability and extensibility.*

  - *[The NSF should] Serve scientific communities' data service requirements ... across a range of data types ... Such a service should NOT exclusively focus on large-scale or what could be referred to as "petabyte data" but rather include mid/small-sized research [and] working with the research community to positively and actively promote open access ...*

  - *The NSF should fund national facilities for at least short-term storage and management of data to support collaboration, scientific workflows, and remote visualization*

Goals of the Wrangler Project

- Our analysis of community needs indicated we needed:
  - To address the data problem in multiple dimensions
    - Big (and small), reliable, secure
    - Lots of data types: Structured and unstructured
    - Fast, but not just for large files and sequential access. Need high transaction rates and random access too.
  - To support a wide range of applications and interfaces
    - Hadoop, but not *just* Hadoop.
    - Traditional languages, but also R, GIS, DB, and other, perhaps less scalable things.
  - To support the full data lifecycle
    - More than scratch
    - Metadata and collection management support
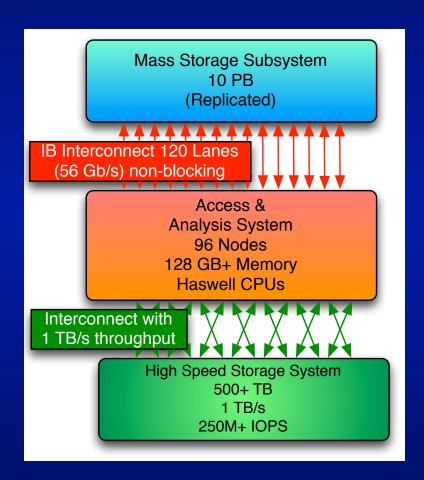- Wrangler is designed with these goals in mind.

Wrangler Team

- Funded by a $6M grant from the National Science Foundation

- Academic partners:
  - TACC
  - Indiana U. ; hosting and end-to-end network tuning.
  - U. of Chicago: Globus Online integration, high speed data transfer from user and XSEDE sites.

Wrangler Hardware @TACC

Mass Storage Subsystem
10 PB
(Replicated)

IB Interconnect 120 Lanes
(56 Gb/s) non-blocking

Access &
Analysis System
96 Nodes
128 GB+ Memory
Haswell CPUs

Interconnect with
1 TB/s throughput

High Speed Storage System
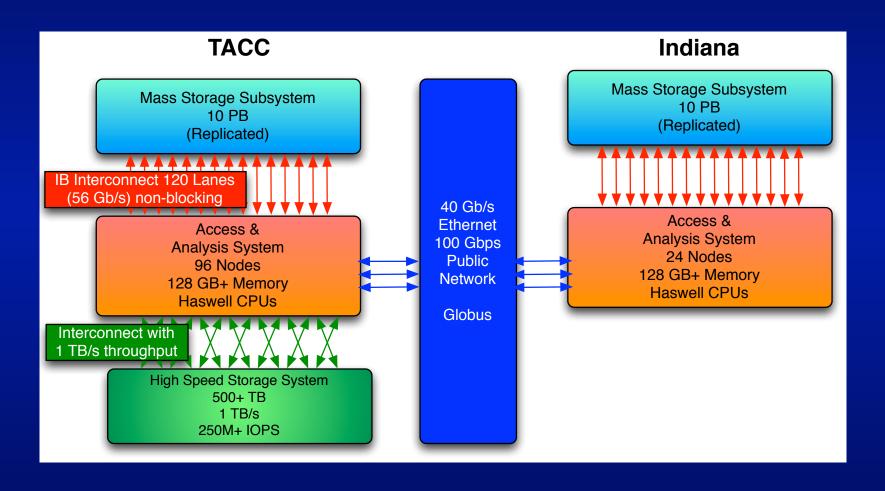500+ TB
1 TB/s
250M+ IOPS

Three primary subsystems:
- A 10PB, replicated disk storage system.
- An embedded analytics capability of several thousand cores.
- A high speed global object store
  - 1TB/s
  - 250M+ IOPS

Wrangler – The Whole Thing

Storage

- The disk storage system will consist of more than 20PB of raw disk for "project-term" storage.
  - Geographically replicated between TACC and Indiana (more reliable than traditional scratch).
  - Ingest at either site.
  - Exposed as Globus endpoint
  - Exposed to users on the system as a traditional filesystem

Analysis Hardware

- The high speed storage will be directly connected to 96 nodes for embedded processing.

  – Each node will have 24+ Intel Haswell cores, and at least 128GB of RAM.

  – An additional 24 nodes will be at the replica site for data ingest, mirroring, processing on the bulk storage.

Need for Globus

- Complex Storage system needs Globus
  – Need to manage 10 PB of user storage
    • Users managing terabytes of data
    • Collaboration, sharing, and data discovery
  – Two replicated 10 PB file systems accessible at both sites
  – Scheduled High Performance Storage system with storage needs tied to job schedules

External Connectivity

- Wrangler will connect externally through both the existing public networks connections and the 100Gbps connections at both TACC and Indiana

- Fast network paths will be available to Stampede and other TACC systems for migration of large datasets

- Globus Online will be configured on Wrangler on day 1.

Wrangler Services

- Support for complex projects via XSEDE Extended Collaborative Data Support Services
  - Working directly with XSEDE staff
  - Focus on specific data management challenges
  - Work with teams to improve data workflows
  - Work to integrate tools into existing infrastructure
  - Transition of a projects processing to Wrangler

Wrangler Services (2)

- Curation & Autocuration Service addressing
  - Leveraging Globus Online Dataset Services
  - Data Organization tying research project to facilitate tracking of a projects data assets
  - Data Provenance leveraging emerging standards to ensure data accessibility and reusability
  - Data Fixity by automating extraction of technical metadata for files in the system
  - Data Usage by tracking overall size of a projects data and (internal and external) utilization information
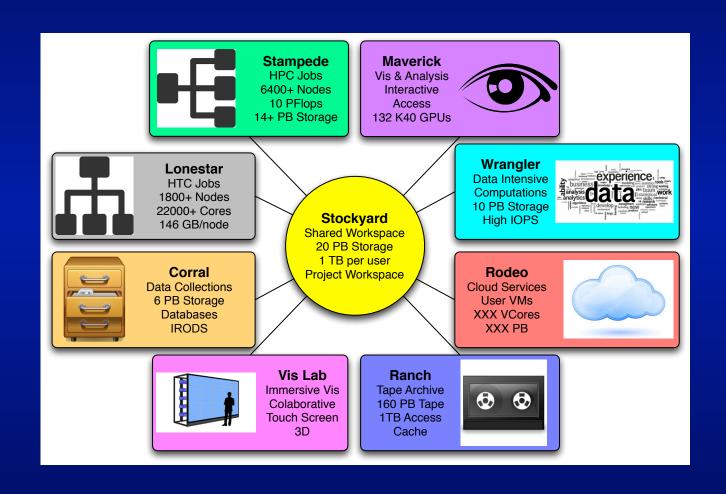
Wrangler Services (3)

- **End to End Network Performance Tuning**
  - Leveraging I^2 and Globus based Perfsonar
    - XSP support to allocate bandwidth on demand
    - Software Defined Networking support via I2's AL2S
    - Network Performance Tools to monitor for bottlenecks and react to congestion

- **Data Dock capability for the "last mile problem" where limited by network capabilities outside of I^2.**

TACC Ecosystem

Stampede
HPC Jobs
6400+ Nodes
10 PFlops
14+ PB Storage

Maverick
Vis & Analysis
Interactive
Access
132 K40 GPUs

Lonestar
HTC Jobs
1800+ Nodes
22000+ Cores
146 GB/node

Wrangler
Data Intensive
Computations
10 PB Storage
High IOPS

Stockyard
Shared Workspace
20 PB Storage
1 TB per user
Project Workspace

Corral
Data Collections
6 PB Storage
Databases
IRODS

Rodeo
Cloud Services
User VMs
XXX VCores
XXX PB

Vis Lab
Immersive Vis
Colaborative
Touch Screen
3D

Ranch
Tape Archive
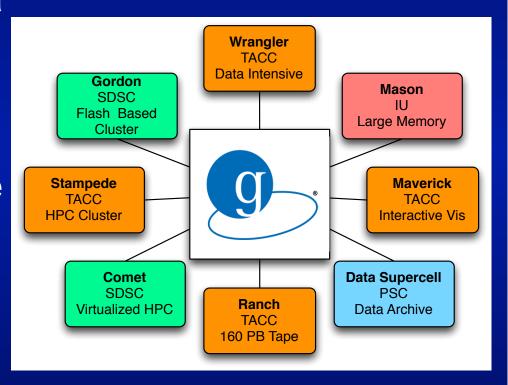160 PB Tape
1TB Access
Cache

Wrangler in the TACC Ecosystem

- Wrangler will be closely connected (in a network, software, location and policy sense)

- Access the TACC cloud system to deploy VMs to host gateways or other permanent sites for direct access to Wrangler.

- The TACC global filesystem will be visible on Wrangler and other TACC resources.

- Optimized network connections to move data from Stampede, Maverick, or other TACC systems.

Wrangler in the XSEDE Ecosystem

- First system for data needs, not just storage, not just compute.

- Allocation of storage beyond tape

- Integrated using Globus to work with all XSEDE systems

Wrangler Software and Use Cases

- The high speed storage will be visible in several ways:
  - As a traditional filesystem, for traditional applications
  - As an HDFS filesystem, for Hadoop and other Map Reduce applications.
  - As a SQL database
  - As an object store with a native API, for novel data applications
- In addition to our "traditional" HPC stack, we will support R, databases, NoSQL databases, and the full Hadoop "suite".

Schedule

- Full production is scheduled for January, 2015.

- An "early user" period will begin a couple of months before.

- We expect to invite some select "very early users" to access a prototype system late next spring to begin porting of key applications
  - Contact us if you would like to be in one of these groups, we are recruiting early users now!

Thanks!

Questions?

ngaffney@tacc.utexas.edu