ORNL visit

Wednesday, January 25, 2017   4:54 PM

Prep: add synchronous SOS aggregation call to TAU.

Process:
- configure / build SOS
- configure / build TAU (configured with SOS)
- configure / build ADIOS (all deps)
- configure / build ADIOS TAU wrapper
- configure / build SOS examples
- configure / build skel examples

Meeting

Jeremy overview of Skel
Jeremy overview of Diesel
- Model -> DSL -> outputs
- Templates -> DSL -> outputs
- Model is defined in YAML
    o model.dsl
    o define all subcommands (like skel)
        ▪ source
        ▪ makefile
        ▪ xml
        ▪ etc.
- Templates written with variable subsitition

Post-lunch

Paper plan
3 past paper examples for the need to avoid introducing network jitter.
- FlexIO - Fang Zheng
- DataTap - Hasan Abbasi
- Manish Parashar (something related to jitter)
Target:
        o Clouds and Distributed computing
        o Data analytics, vis and storage

- Manish Parashar (something related to jitter)

Target:
- SC? - not an I/O track. March 20
  - Clouds and Distributed computing
  - Data analytics, vis and storage
- Cluster? - April 24 / May 1
- SC Workshops?
  - ISAV
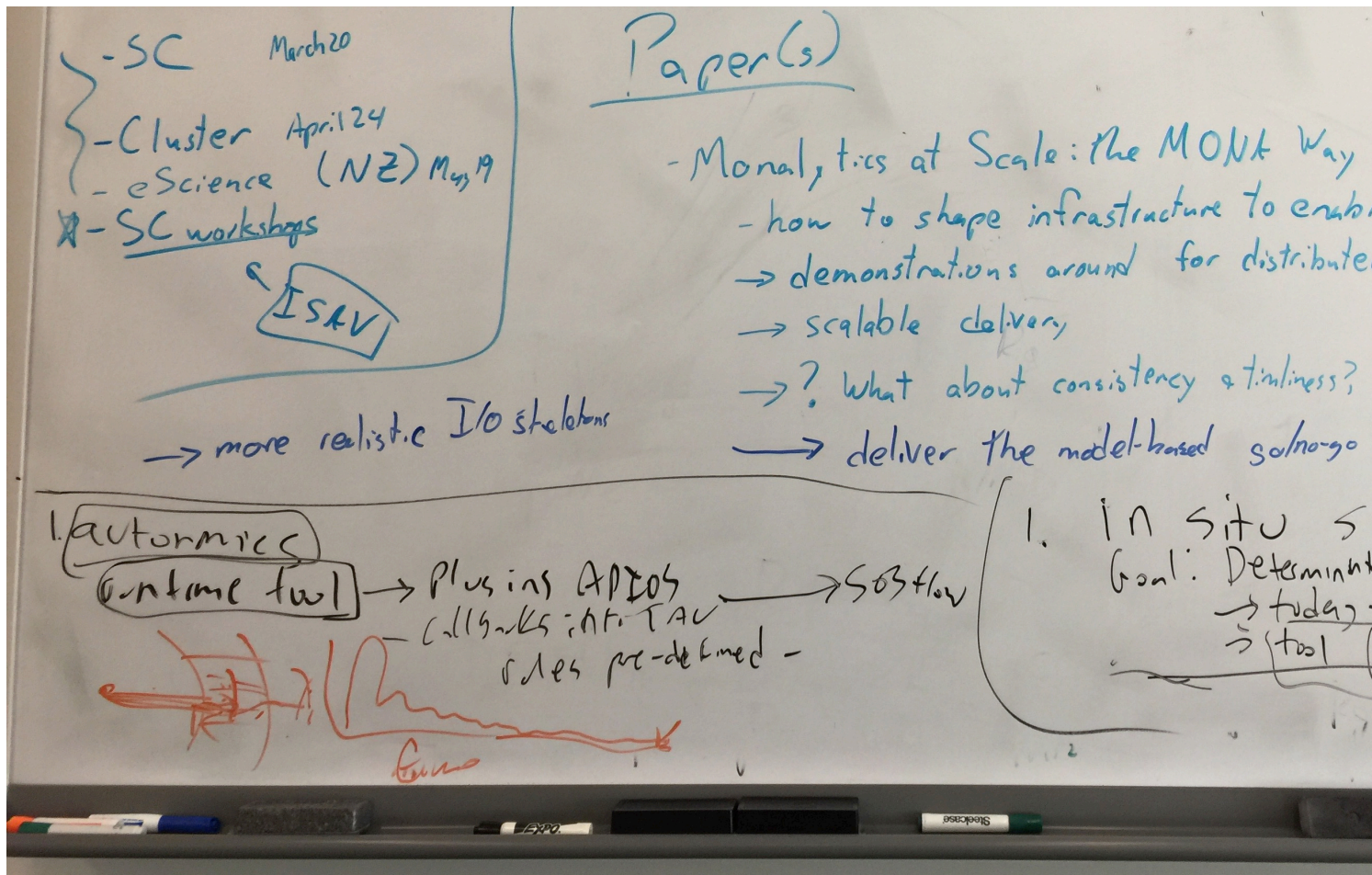- Middleware - could be a hard sell.
- EScience - 19 May

Story:
- Monalytics at scale: The MONA way
  - Infrastructure focus
  - how to shape infrastructure to enable requirements
  - NOT a framework paper
  - demonstrations around the different resource choices for distributed store of performance data
  - scalable delivery
    - What about consistency?
    - Timeliness
  - Actually delivering the model-based go / no-go analysis
    - preventing jitter
    - clock synchronization in the database
    - need to implement the scheduling pull
  - *More realiistic I/O skeletons*

Discussion with Scott Klasky:
- What are the ***metrics*** of success for the MONA project
 - Demonstrable effect on application performance

Captureboard

-SC   March 20

-Cluster April 24
- eScience (NZ) May 19
-SC workshops
    ISAV

→ more realistic I/O skeletons

1. autonomics
    runtime tool → Plugins ADIOS → SOSflow
        - callbacks into TAU
        rules pre-defined -

Paper(s)

- Monalytics at Scale: the MONA Way
- how to shape infrastructure to enabl...
→ demonstrations around for distribute...
→ scalable delivery
→? What about consistency & timliness?
→ deliver the model-based solno-go

1. in situ s...
    Goal: Determinat...
        → today...
        → tool

Generation of quadchart

DAY 2

Discussion of
synthetic workflow integration (see photo)

wrap () £
    TAU Start
    Discover
    adios
    Enddef   APP   TAU wrap   TAU
    Tau Stop
        ADIOS   SOS
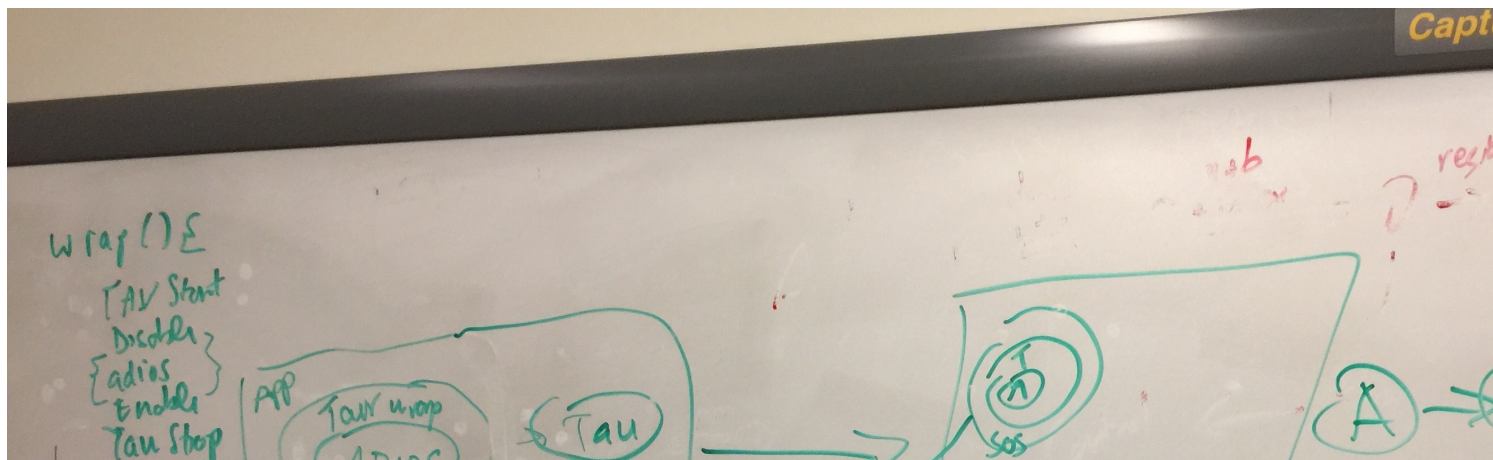
le requirements
d store of perf. data

analysis WDM

er vice appro.
un of performance
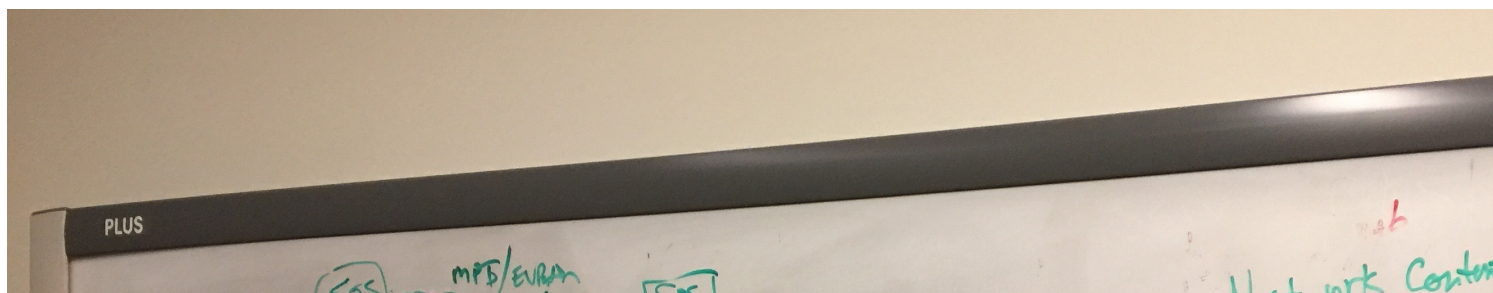- currently difficult
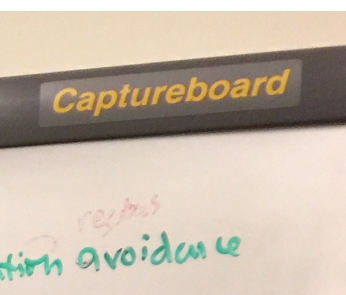S OS flow → web par
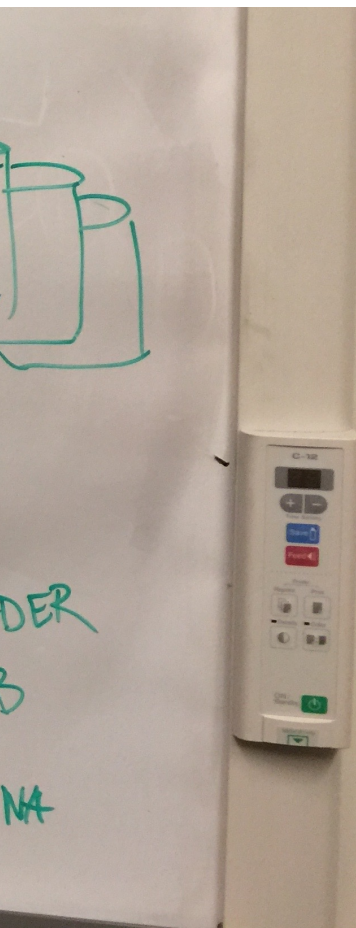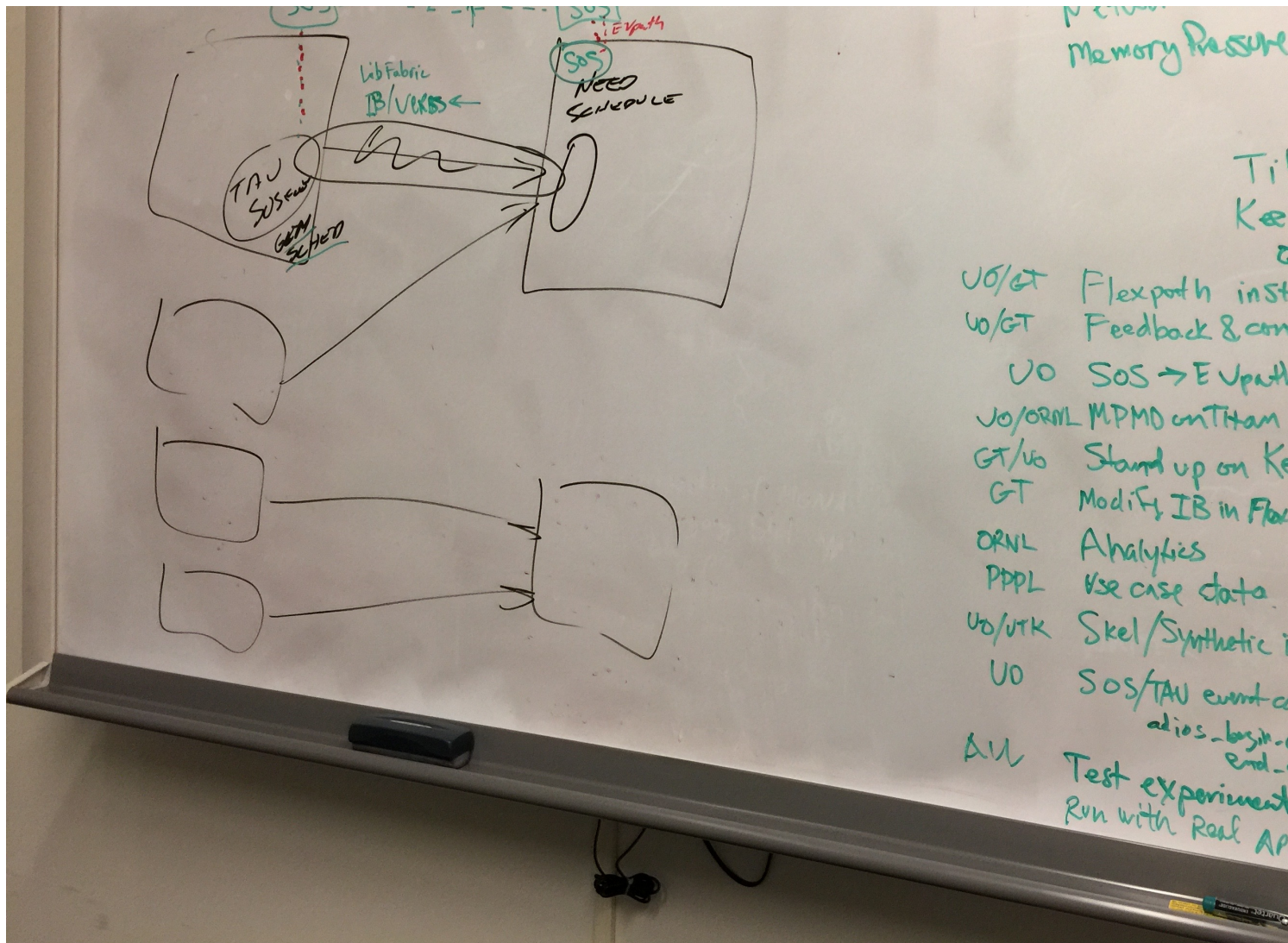. Use care

ureboard

ms

r

B
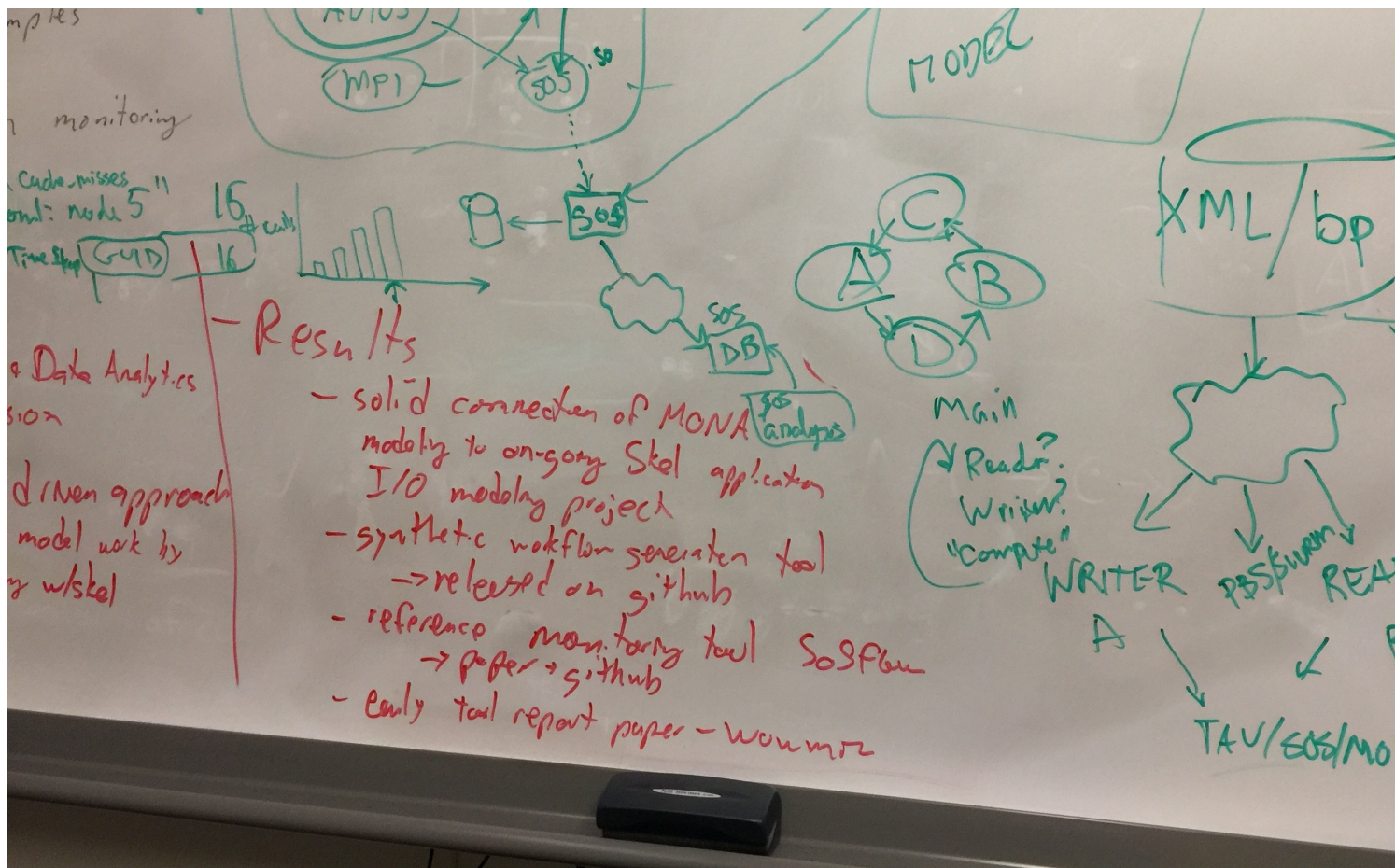
Integration plan
- reduced jitter / timed pulls / a la datatap
- "generic" test harness for doing analysis development and testing
- datatap code is in ADIOS
- AMR, load balancing codes are tricky, because of irregular behavior
- Gyrokinetic codes are tricky because of multiple, interacting periodicity.
- I/O time is not taken into account - need to consider multiple periodicities (writing vectors, writing particles)
- Parameters: loop length, offset within the loop, length of good window, realtime offset, I/O frequency
- GTS, LAMMPS examples

DER

B

NA

C-12

Captureboard

result
tion avoidance

Top whiteboard

Lib Fabric
IB/VERBS ←

SOS
Evpath

SOS
NEED
SCHEDULE

TAU
SOS...

GEN
SCHED

Memory Pressure

Ti...
Ke...

UO/GT	Flexpath inst...
UO/GT	Feedback & com...
UO	SOS → Evpath
UO/ORNL	MPMD on Titan
GT/UO	Stand up on Ke...
GT	Modify IB in Flo...
ORNL	Analytics
PPPL	Use case data
UO/UTK	Skel/Synthetic ...
UO	SOS/TAU event c...
	adios_begin...
	end...
All	Test experiment...
	Run with Real AP...

Bottom whiteboard (Captureboard)

net
regions

Network Contention avoidance
Memory Pressure

Falcon

Falcon

tan

↓

nland/Kids 80
         +40

DDR IB

-rumentation

trol

n

eonland

path for Scheduling

integration

apture of MPI
          Collectives
-rmputation
-rmputation

S

P WORKFLOW

Titan ↓
Keenland/Kids 80
+40
QDR IB

UO/GT	Flexpath instrumentation
UO/GT	Feedback & control
UO	SOS → Evpath
UO/ORNL	MPMD on Titan
GT/UO	Stand up on Keenland
GT	Modify IB in Flexpath for Scheduling
ORNL	Analytics
PPPL	Use case data.
UO/UTK	Skel/Synthetic integration
UO	SOS/TAU event capture of MPI Collectives
	adios_begin_computation
	end_computation
All	Test experiments
	Run with Real APP WORKFLOW

Paper plan
- planning postponed until phone calls

Call with Stephane
- Current holdup - waiting on jobs on Edison to validate ADIOS integration in GTS update.
- Matt suggesting Kids/Keenland/Falcon as backup plan?
- GTS is running, needs PetSc

Call with Stephane

~~Current holdup - waiting on jobs on Edison to validate ADIOS integration in GTS~~
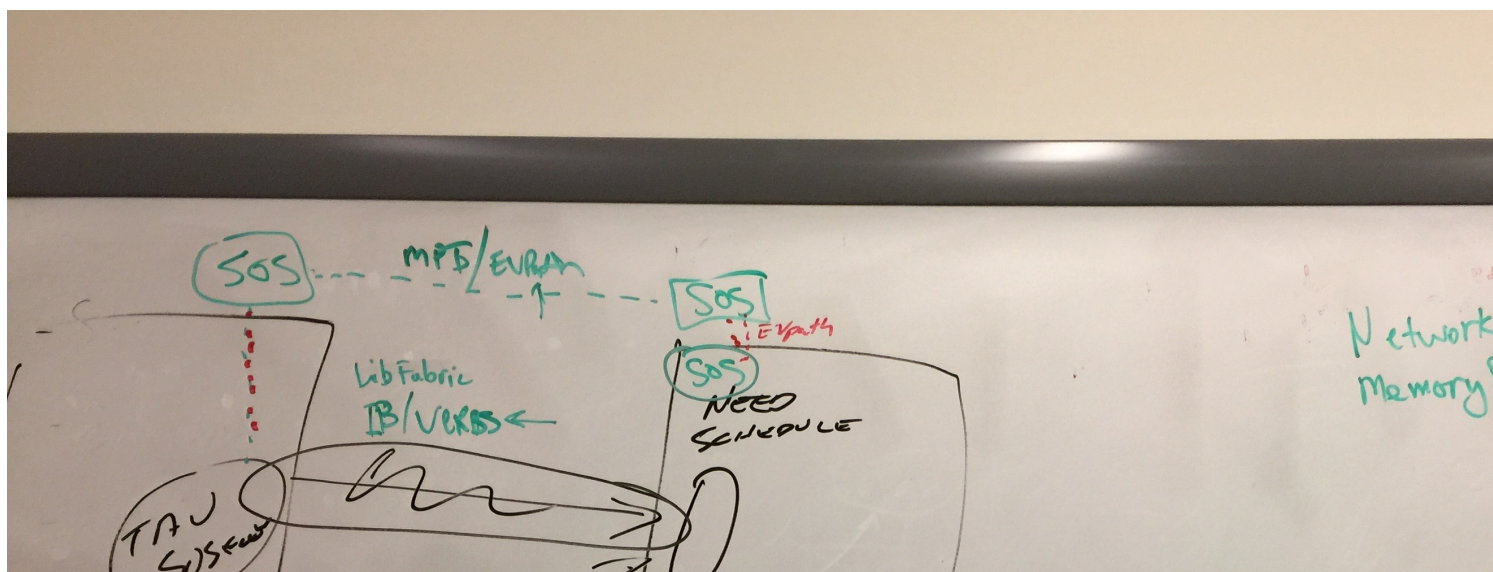update.

- Matt suggesting Kids/Keenland/Falcon as backup plan?
- GTS is running, needs PetSc
- After GTS is running, need analysis component
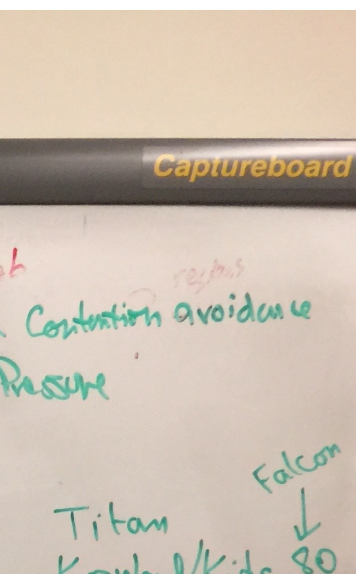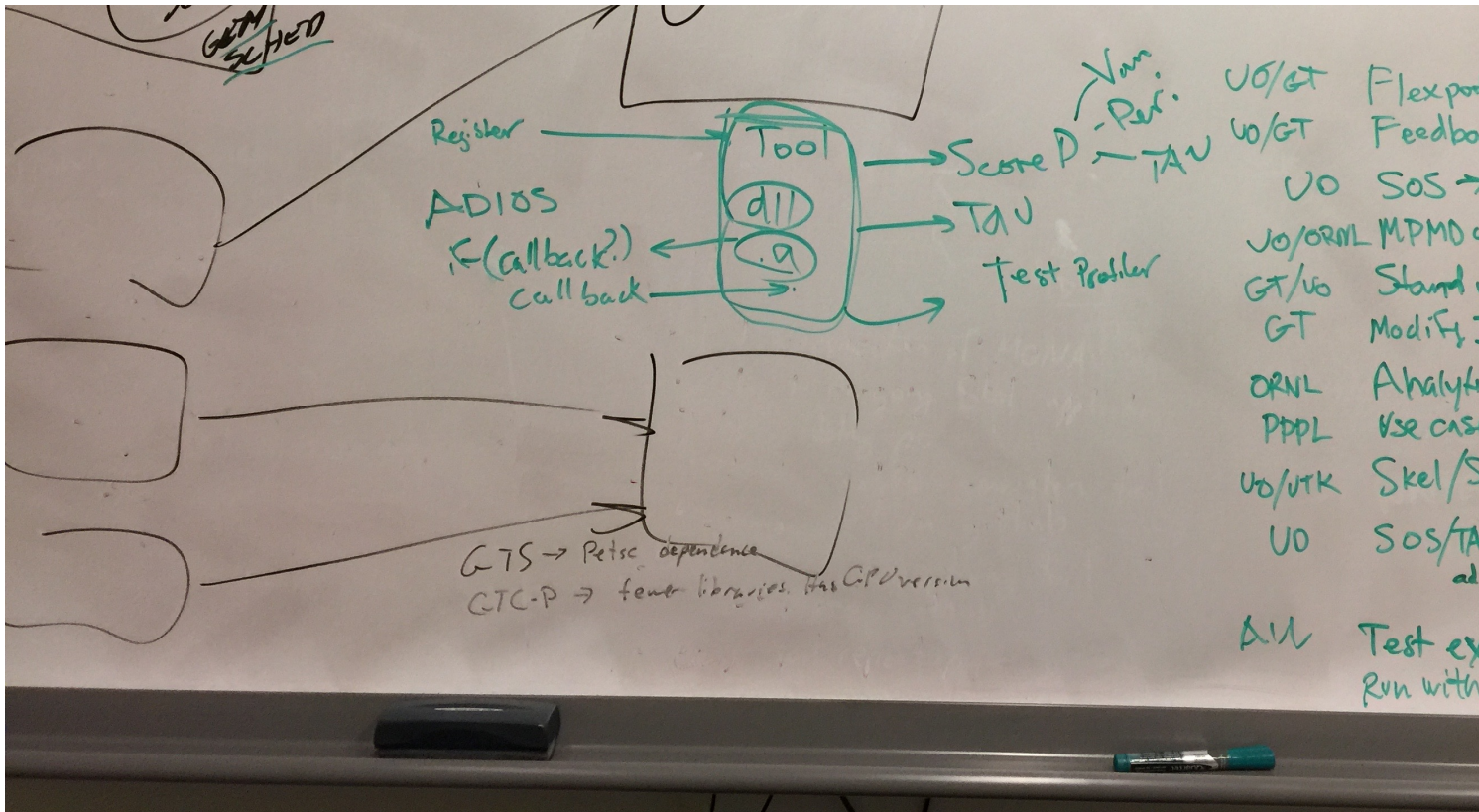- Then...could add Visit as a third component

Lunch
Data science group meting

Meeting with Norbert

- What in ADIOS should be measured by TAU - if we were going to have a more "permanent" instrumentation
- Persistent capability to observe ADIOS behavior (like Vampir tracing)
- is External API sufficient?
- What about internal transport information?
- Bytes written, metadata size, bytes in memory.
- callback interface - register a tool, the tool does the timing/measurement
- We care about:
    o interval timers
    o memory footprint (especially max size)
    o bytes transferred
    o synchronization delays
- Work with Norbert on creating callback API in ADIOS
- Would be useful for CODAR "always on" data collection, Interaction with Score-P / Vampir

Captureboard

...b

...rs *restus*

... Contention avoidance

...Pressure

Falcon
↓

Titan

Kawla.../Kids 80

CODAR overview

- Center for Online Data Analytics and Reduction
- one of 4 codesign centers for ECP - all others are math/methods
- EVERYONE needs to deal with the I/O bottleneck on exascale machines
- Divided into 3 teams
  - Analysis
    - Broader scope than reduction
    - Not graph/topological analytics
    - Focused on another type of reduction (ML, statistical inference, summarization)
  - Reduction (compression, some lossless, mostly lossy)
    - Guidance on which compression methods to use for which kinds of data
  - Infrastructure
    - How do we deliver this?
    - If things are good, what is the platform for delivery?
    - How do users evaluate the solution for themselves?
    - Testing harness that drives initial tables and evaluation processes of the raw data and compression methods and algorithms.
      - How to extract performance, throughput, memory footprint, etc
    - Complications - temporal compression, buffering techniques

Keenland/KIDS +40
QDR IB
th instrumentation
ck & control
E Vpath
nTitan
up on Keenland
IB in Flexpath for scheduling
ies
e data
Synthetic integration
W event capture of MPI
Collectives
ios_begin-computation
end_computation
xperiments
Real APP WORKFLOW

- - - - How do users evaluate the solution for themselves?
  - Testing harness that drives initial tables and evaluation processes of the raw data and compression methods and algorithms.
    - Something that looks a lot like Skel, ADIOS running on SwiftT
    - How to extract performance, throughput, memory footprint, etc
  - Complications - temporal compression, buffering techniques
  - Not just single timestep, single processor. The complex examples are the hard things.
  - ADIOS people, Justin Wosniak (SwiftT / ANL), Kerstin's group (provenance, Brookhaven NL)
- Expected to have crossover with other application, software project groups