ICEE Wide-area In Transit Data Demo

-- Real-Time Analysis of large experiments

SC14 Demos

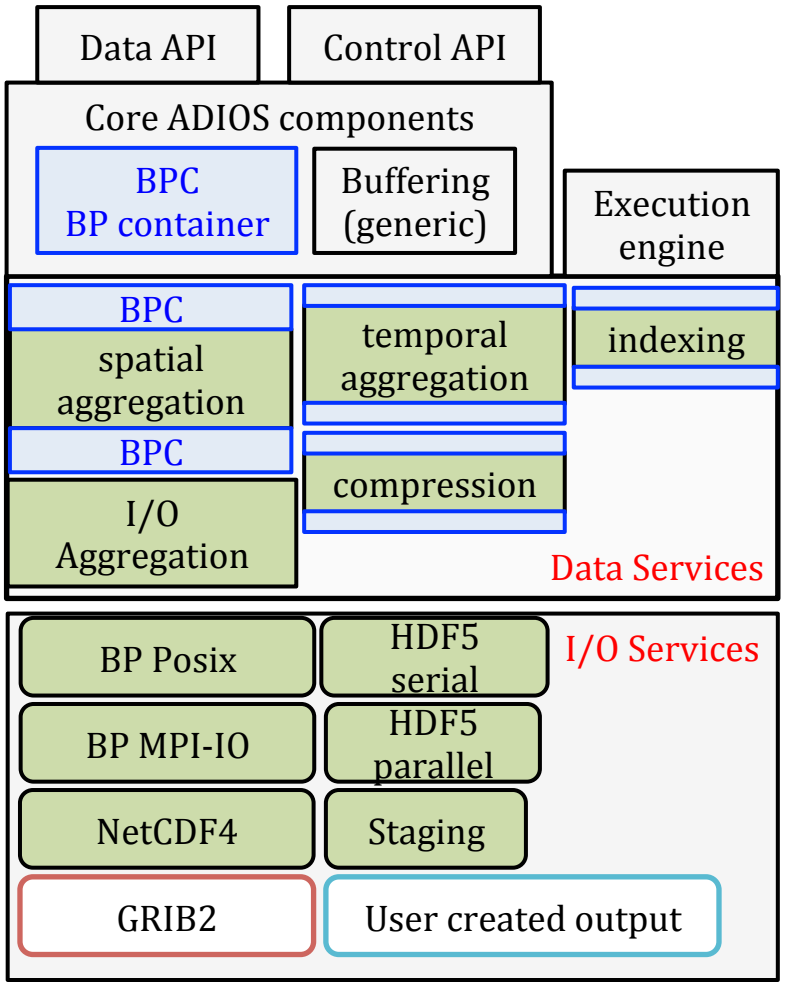
POC: John Wu <kwu@lbl.gov>

Scott Klasky <klasky@ornl.gov>

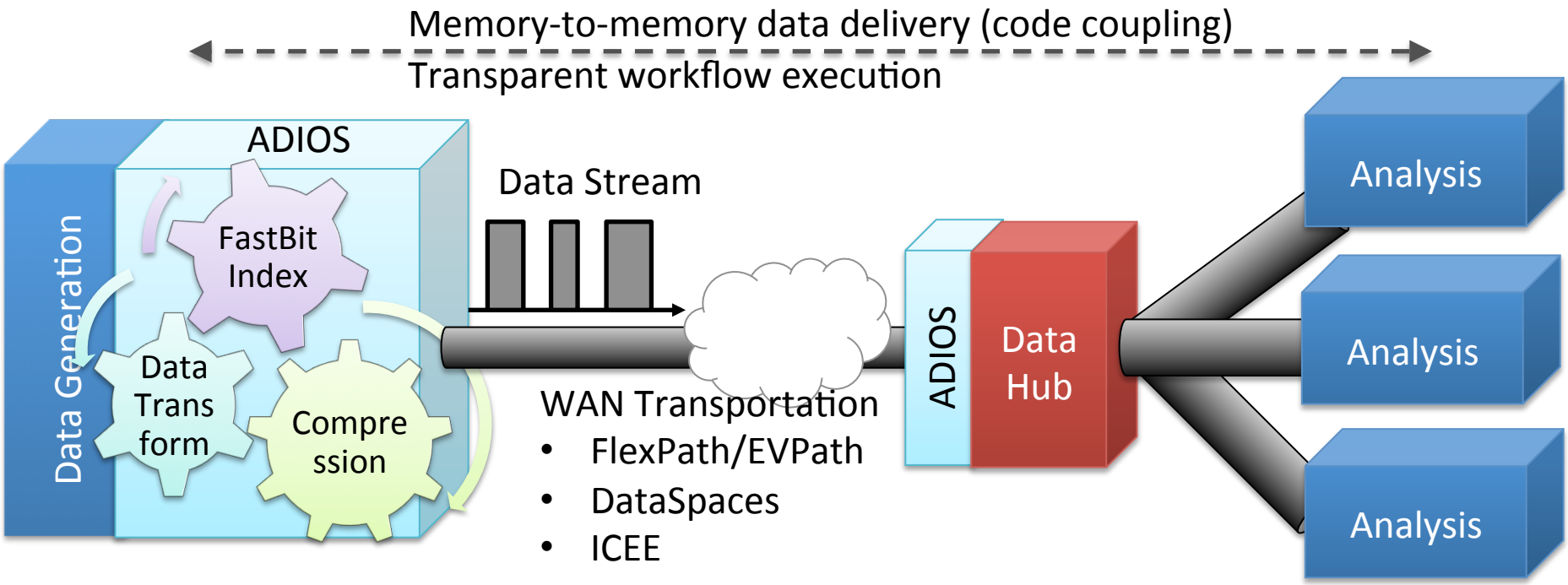
ADIOS Abstraction Unifies Local And Remote I/O

- I/O Componentization for Data-at-Rest and Data-in-Motion
- Service Oriented Architecture for Extreme scaling computing
- Self Describing data movement/ storage
- Main paper to cite

Q. Liu, J. Logan, Y. Tian, H. Abbasi, N. Podhorszki, J. Choi, S. Klasky, R. Tchoua, J. Lofstead, R. Oldfield, M. Parashar, N. Samatova, K. Schwan, A. Shoshani, M. Wolf, K. Wu, W. Yu, "Hello ADIOS: the challenges and lessons of developing leadership class I/O frameworks", Concurrency and Computation: Practice and Experience, 2013



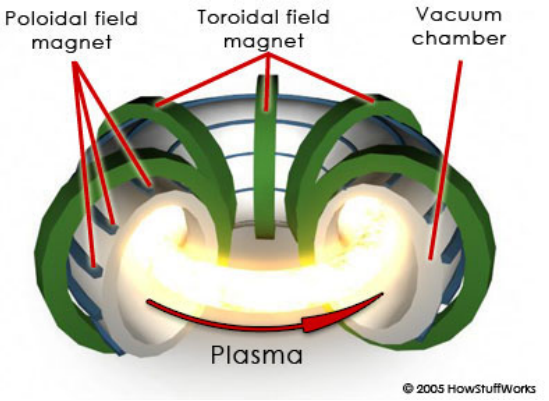
ICEE Method Enables WAN Staging

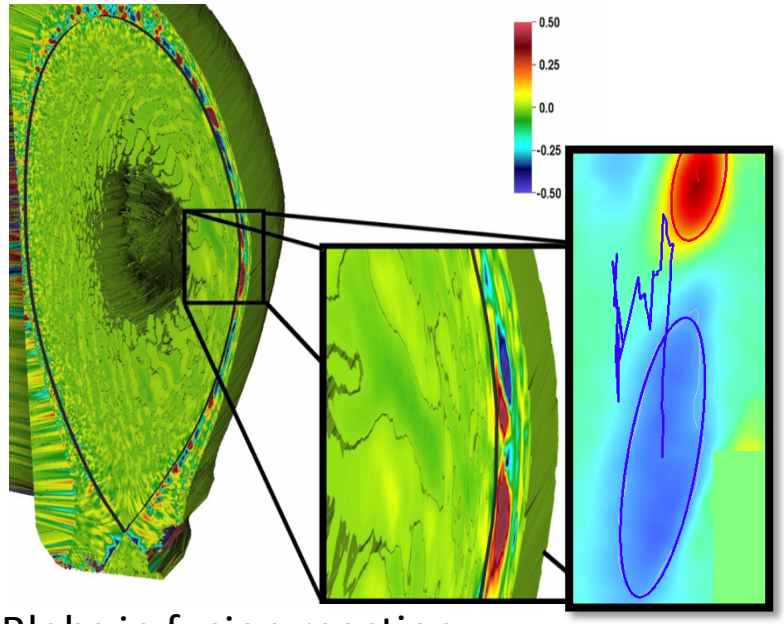


- Research on stream-based WAN data process
  - In-transit processing (supporting data-in-memory)
  - Data indexing & query to reduce network payload
  - WAN transportation: FlexPath (GATech), DataSpaces (Rutgers), ICEE (ORNL/LBNL)

SC14 Demo 1: Near Real Time Detection of Fusion Blobs

- Plasma blobs
  - Lead to the loss of stability and/or confinement of tokamak plasmas
  - Cause fast thermal and/or current quench
  - Could damage multi-billion tokamak
- The experimental facility may not have enough computing power for the necessary data processing
- Distributed in transient processing
  - Make more processing power available
  - Allow more scientists to participate in the data analysis operations and monitor the experiment remotely
  - Enable scientists to share knowledge and processes
- Lingfei Wu, Alex Sim, Jong Choi, M. Churchill, K Wu, S Klasky, CS Chang



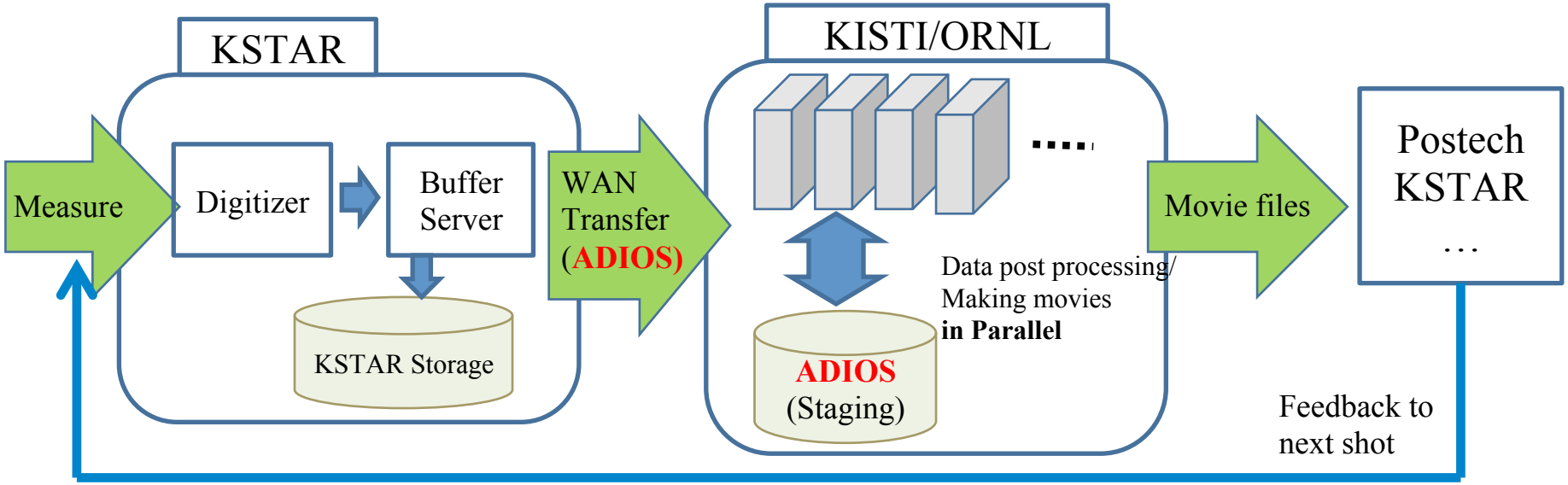


Blobs in fusion reaction (Source: EPSI project)

Blob trajectory

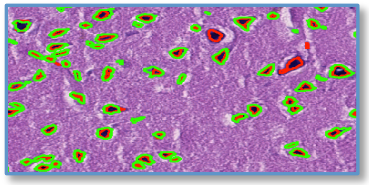
SC14 Demo 2: KSTAR ECEI Sample Workflow

- **Objective**: Making ECE-Image movies of tokamak plasma voltage
- **Input**: 3 cameras x (8x24)channels x 500KHz x 2Byte Voltage data x shot-duration + Metadata(about experimental setting)
- **Output**: Tokamak movie files, spectrogram of each channel
- **Requirement**: Finish making movies **within 15min** (inter-shot time)
- **Implementation**: distributed data processing with ADIOS ICEE method
- Eunkyu Byun, Soonwook Hwang, KISTI, South Korea

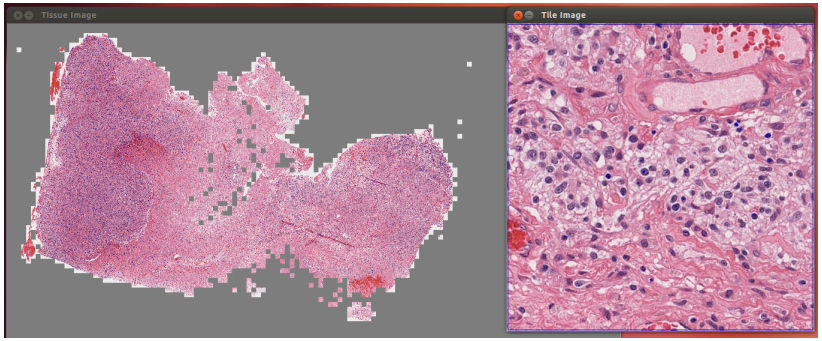


SC14 Demo 3: Microscopy Image Analysis

(Stony Brook University (SBU) & Oak Ridge National Laboratory (ORNL))



- **Significance:** Understanding of disease morphology at micro - anatomic level has tremendous potential for better diagnosis and better understanding of disease mechanisms.
- **Challenge:** Significant processing power to rapidly analyze tissue slide images (40Kx40K to 120Kx120K pixels in resolution) in order to quantitatively assess disease condition. Machines near slide repositories may not have sufficient power.
- **Approach:** Split processing workflow. (1) Partition slide image into tiles; (2) Process each tile at low-resolution to see if it contains information; (3) Send selected full-resolution tiles to a powerful cluster over wide-area network; (4) Perform analysis (nuclear segmentation and features) on full-resolution tiles on powerful cluster.
- **Technologies:** (1) SBU ADIOS for wide-area, efficient transfers; (2) Longbow for very fast, low-latency connection; (3) SBU RT for high-throughput, pipelined processing on cluster.
- **Demo:** Tissue slides on machine in Singapore. Analysis done on cluster at Georgia Tech. Segmentation results displayed on client machine.



Snapshot of adaptive processing of a remote slide (53Kx36K pixel resolution).

Work supported in part by NCI and NLM: 1U24CA180924, R01LM011119-01, R01LM009239 grants, and ORNL SDAV and JFA.

SC14 Demo 4: Molecular Dynamics – Design of Materials

- This demonstration is based off a scenario from materials scientists interested in understanding fracture in nano-structured materials
- It uses LAMMPS to simulate the block of nano-structured metal while under stress.
- Simulation proceeds until the first plastic deformation (start of fracture) is detected.
- At that first fracture, the system is fully characterized to understand where and, hopefully, why things broke.