2013 R&D 100 Award Winner
ADIOS named one of top 100 tech products by R&D Magazine.

ADIOS easy-to-use, fast, scalable, and portable I/O

ADIOS: Creating a sustainable I/O framework

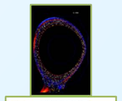
Scott Klasky, Qing Liu, Norbert Podhorszki: Oak Ridge National Laboratory, Manish Parashar, Rutgers
Greg Eisenhauer, Karsten Schwan, Matthew Wolf: Georgia Tech, Nagiza Samatova: ORNL/NCSU
Tahsin Kurc, Joel Saltz: ORNL/Stony Brook

Create a collaborative framework for scientist around the world to contribute and work with Data Intensive Science

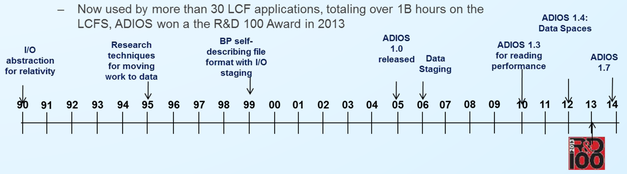
ADIOS Timeline

How did we come about this approach?

- **Problem**
  - Before ADIOS, application writers had trouble achieving high-performance I/O for self-describing data
- **Solution**
  - Working with many leading DOE applications we developed a new framework for I/O with an API to abstract the implementation from the API
  - Burst Buffers, a simplified version of I/O-staging, becoming the de-facto standard for exascale I/O was created as part of the ADIOS framework
  - The first fully developed DOE I/O framework developed for sustainable I/O on LCFs
- **Impact**
  - Applications using ADIOS demonstrated input/output results more than 10 X faster than previous implementations
  - Now used by more than 30 LCF applications, totaling over 1B hours on the LCFS, ADIOS won a the R&D 100 Award in 2013



Using ADIOS, I/O for the XGC code went from 4,000 secs/hour using HDF5 to 252 secs/hour on Titan on 24K nodes



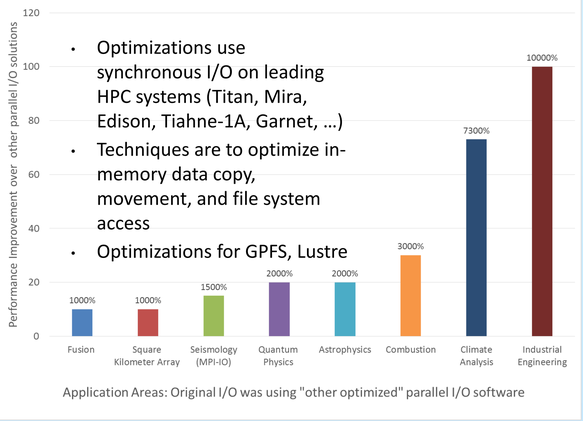
I/O MiniApps

- ADIOS output files can use "skel" family to **extract metadata from simulations to re-create output**
  - Allows researchers to understand and replay performance bottlenecks from any run
  - Information is stored in ADIOS metadata and can be extracted and moved to allow "I/O experts" to re-create the I/O"
- Metadata from codes allow users to re-generate I/O kernels
  - SKEL address the issue of code changes leading to I/O kernels being data, with Mini-I/O Applications (MiniApps)
  - MiniApps offer all of the benefits of I/O kernels including allowing I/O optimizations to focus on useful I/O patterns

ADIOS Write/Read optimizations improve I/O

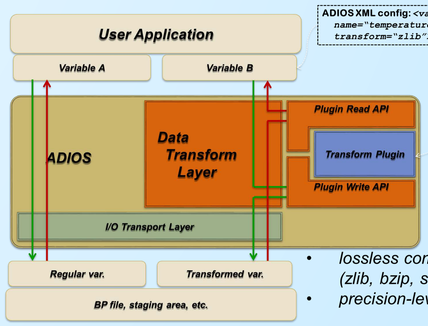
ADIOS contributions in bridging the data gap for high fidelity science

- Optimizations use synchronous I/O on leading HPC systems (Titan, Mira, Edison, Tiahne-1A, Garnet, …)
- Techniques are to optimize in-memory data copy, movement, and file system access
- Optimizations for GPFS, Lustre



ADIOS 1.6-1.7 New Features *new*

ADIOS Transforms Framework



- *lossless compression (zlib, bzip, szip, ISOBAR)*
- *precision-level-of-detail encoding (APLOD)*

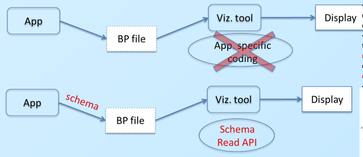
- Topology aware writing on Titan
- Dataspaces on Blue Gene-Q
- Dataspaces running as a service
- Improved Useability (cmake, etc.)
- skeldump, skel replay

Coming soon

- Indexing and Queries
- WAN Staging
- Improve ease-of-use
- ADIOS "best" method
- GRIB2 file format

Visualization Schema

- Why do we need a schema
  - Bridge the gap between application and viz.
  - Keep it consistent
- Current status: application/data format specific reader in visualization tools



Schema Example

Uniform Mesh

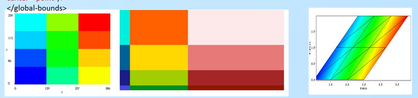
```
<mesh name="uniformmesh"
type="uniform">
    <nspace value="2" />
    <dimensions value="
nx_g,ny_g" />
    <origin value="O1,O2" />
    <spacing value="S1,S2" />
</mesh>
<global-bounds
dimensions="nxg,nyg"
offsets="off_x,off_y">
<var name="data"
type="double"
mesh="uniformmesh"
dimensions="nx_l,ny_l"
center="point"/>
</global-bounds>
```

Rectilinear Mesh

```
<mesh name="rectilinearmesh"
type="rectilinear"
    <nspace value="2" />
    <dimensions value="nx_g,ny_g" />
    <coordinates-multi-var value="X,Y" />
</mesh>
<global-bounds
    dimensions="nx_global,ny_global"
    offsets="off_x,off_y">
    <var name="data" type="double"
     dimensions="nx_l,ny_l"
     mesh="rectilinearmesh"
     center="point"/>
</global-bounds>
```

Unstructured Mesh

```
<mesh name="trimesh"
type="unstructured"
    <nspace value="2" />
    <points-single-var value="points" />
    <uniform-cells count="num_cells"
     data="cells" type="triangle" />
</mesh>
<global-bounds dimensions="npoints"
offsets="op">
    <var name="N" type="double"
     dimensions="lp"
     mesh="trimesh" center="point"/>
</global-bounds>
```



Topology-Aware Methods

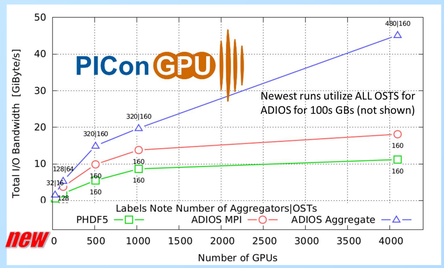
Addresses I/O challenges on LCF systems

- High communication cost
- Complex routing on Titan
- Small data size per core

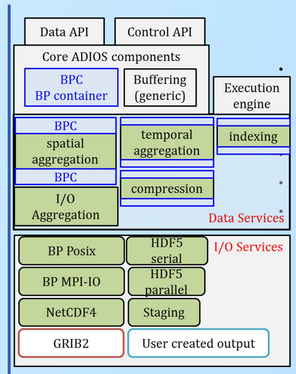
Techniques

- Topology-aware data movement that takes advantage of BGQ/Cray topology
- Minimize data movement
- Properly align data when being written to disk
- Allows I/O to reach 100s GB/s on Mira, Titan
- Currently being optimized for Titan, released for BGQ systems

New application performance



ADIOS framework

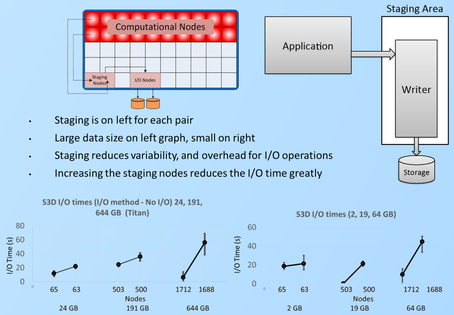


- An I/O abstraction framework : API is abstracted away from the method
- I/O componentization framework for Data-at-Rest and Data-in-Motion
- Provides portable, fast, scalable, easy-to-use, metadata rich output
- Change I/O method on-the-fly
- http://www.nccs.gov/user-support/center-projects/adios/
- Need to provide solutions for "90% of the applications"

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

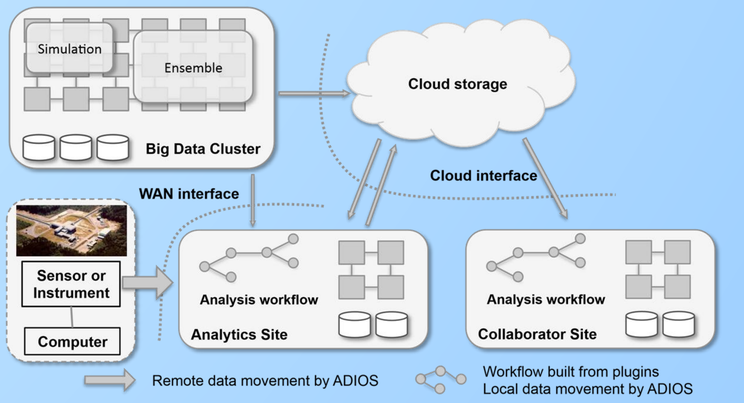
I/O Staging with ADIOS 1.7

- Dataspaces and FlexPath released in ADIOS
- Simple staging allows codes to utilize ADIOS to use APIs to work with data-in-memory
- Current work allows staging on the (same node, different nodes, different machines)
- Graphs shown below illustrate staging benefits for S3D (full code) 6 runs/data point



- Staging is on left for each pair
- Large data size on left graph, small on right
- Staging reduces variability, and overhead for I/O operations
- Increasing the staging nodes reduces the I/O time greatly

ADIOS framework extensions



SDAV Scalable Data Management, Analysis, and Visualization

U.S. DEPARTMENT OF ENERGY | Office of Science