



Examples

CGNS



CFD Standard

- CGNS = **C**omputational Fluid Dynamics (CFD) **G**eneral **N**otation **S**ystem
- An effort to standardize CFD input and output data, including:
 - Grid (both structured and unstructured), flow solution
 - Connectivity, boundary conditions, auxiliary information.
- Two parts:
 - A standard format for recording the data
 - Software that reads, writes and modifies data in that format.
- An American Institute of Aeronautics and Astronautics Recommended Practice





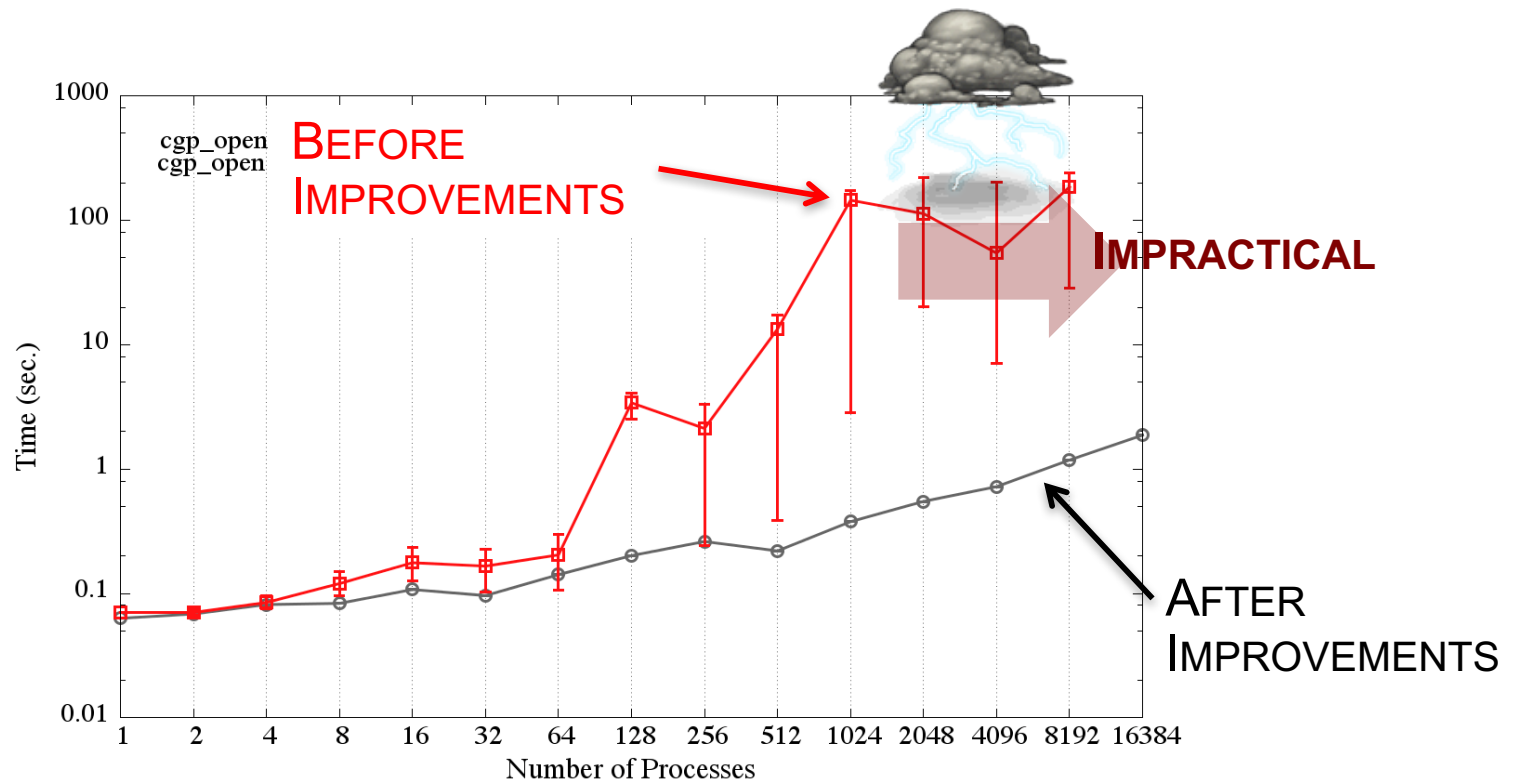
CGNS Storage Evolution

- CGNS data was originally stored in ADF ('Advanced Data Format')
 - ADF lacks parallel I/O or data compression capabilities
 - Doesn't have HDF5's support base and tools
- HDF5 superseded ADF as the official storage mechanism for CGNS
- CGNS introduced parallel I/O APIs w/ parallel HDF5 in 2013
 - The HDF Group has been involved with the CGNS library since 2014.



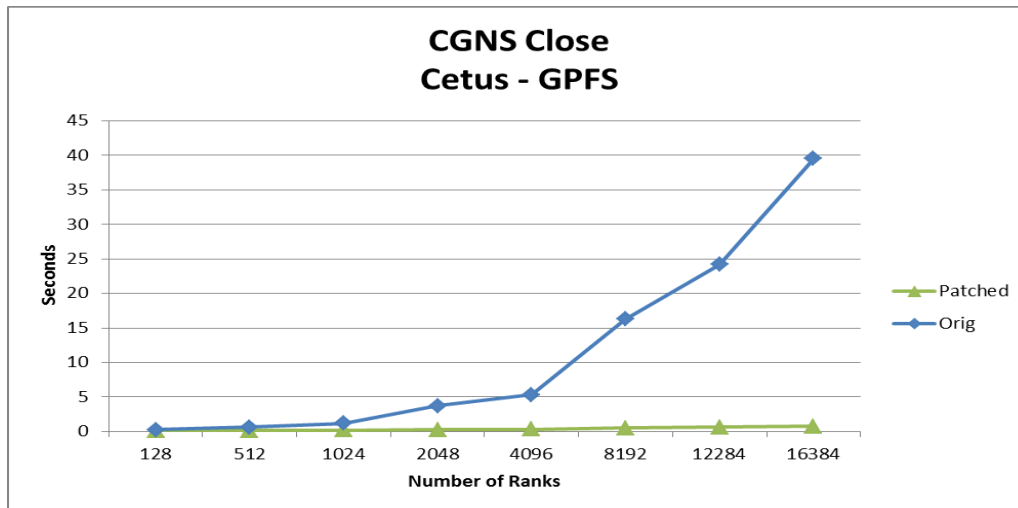
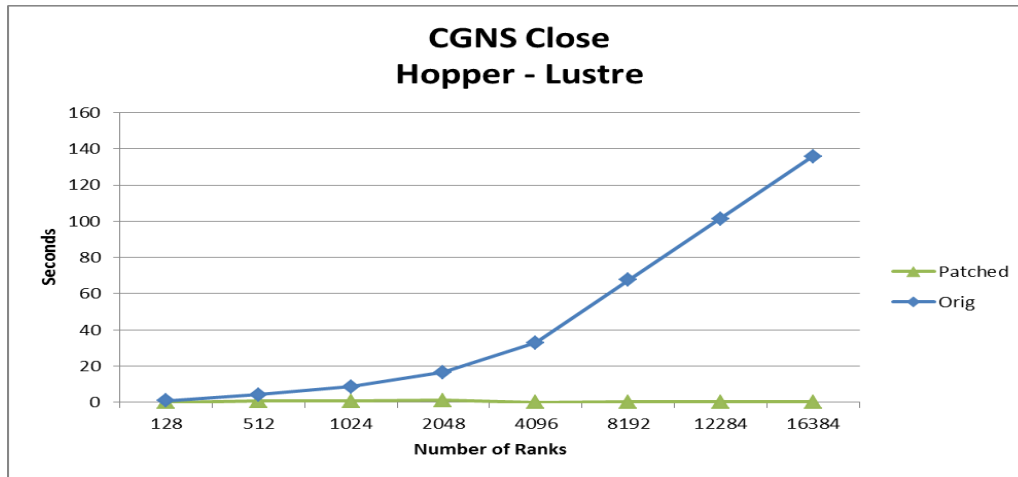
CGNS Performance Problems

- **Symptoms:** Opening an existing file was slow.
- **Diagnostics:** Loading a lot of (HDF5) metadata, where reads occur independently on ALL ranks competing for the same metadata.
- **Solution:** Use collective metadata HDF5 APIs.





Write Metadata Collectively!



- **Symptoms:** Many users reported that `H5Fclose()` is very slow and doesn't scale well on parallel file systems.
- **Diagnosis:** HDF5 metadata cache issues very small accesses (one write per entry). We know that parallel file systems don't do well with small I/O accesses.
- **Solution:** Gather up all the entries of an epoch, create an MPI derived datatype, and issue a single collective MPI write.



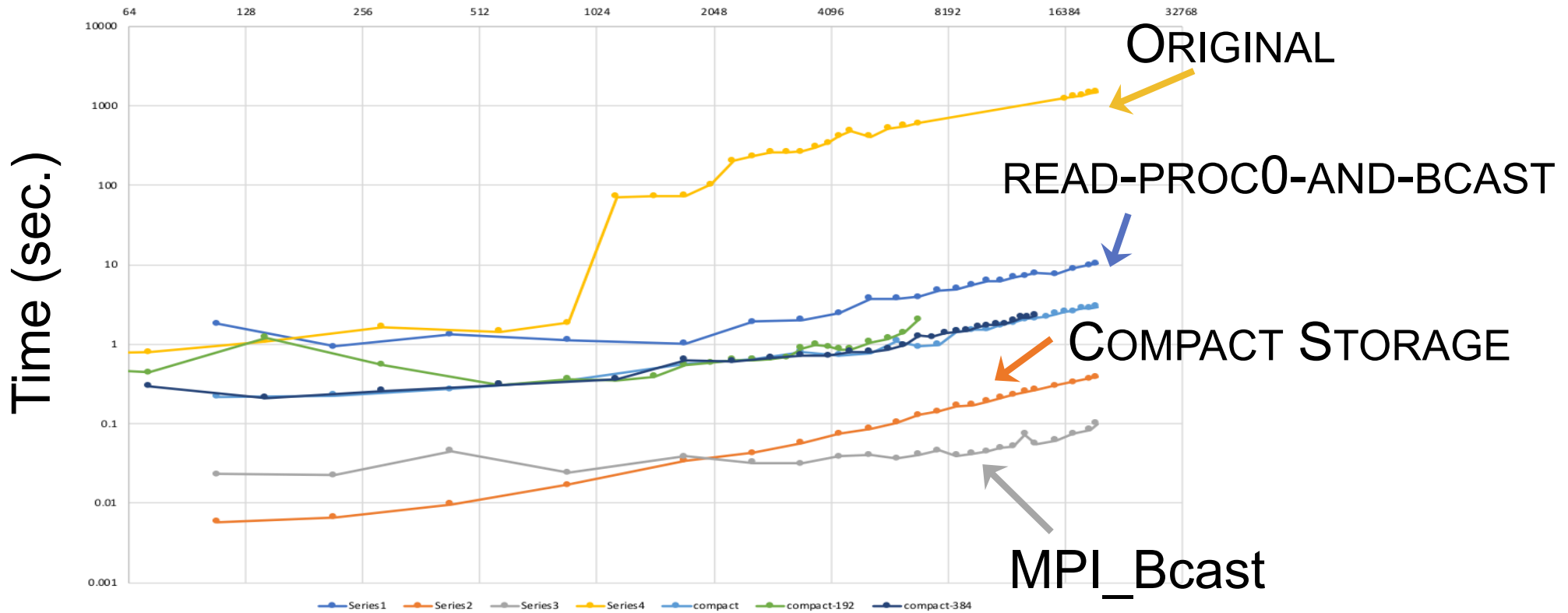
BROADCAST For READ ALL, By ALL

- **Symptoms:** CGNS reads the same CGNS format-specific metadata on all the processes.
- **Diagnosis:** We know that parallel file systems don't do well with small I/O accesses.
- **Solution:** Added to HDF5 detection for reading all the same data, triggering a single rank read and broadcasts of the read data.



BROADCAST For READ ALL, By ALL

Number of Processes



Greg Sjaardema, Sandia National Labs