

MAGIC Meeting: Storage Management Discussion

Dries Kimpe

dkimpe@mcs.anl.gov

5 March 2013

Context

(About me & what I'll be talking about)

- Assistant Computer Scientist at the **Mathematics and Computer Sciences** division of **Argonne National Laboratory**
- Current storage-related projects:
 - **Triton**: Next-Generation Storage System
 - **NoLoSS**: Integrating In-System Storage into the Storage Stack



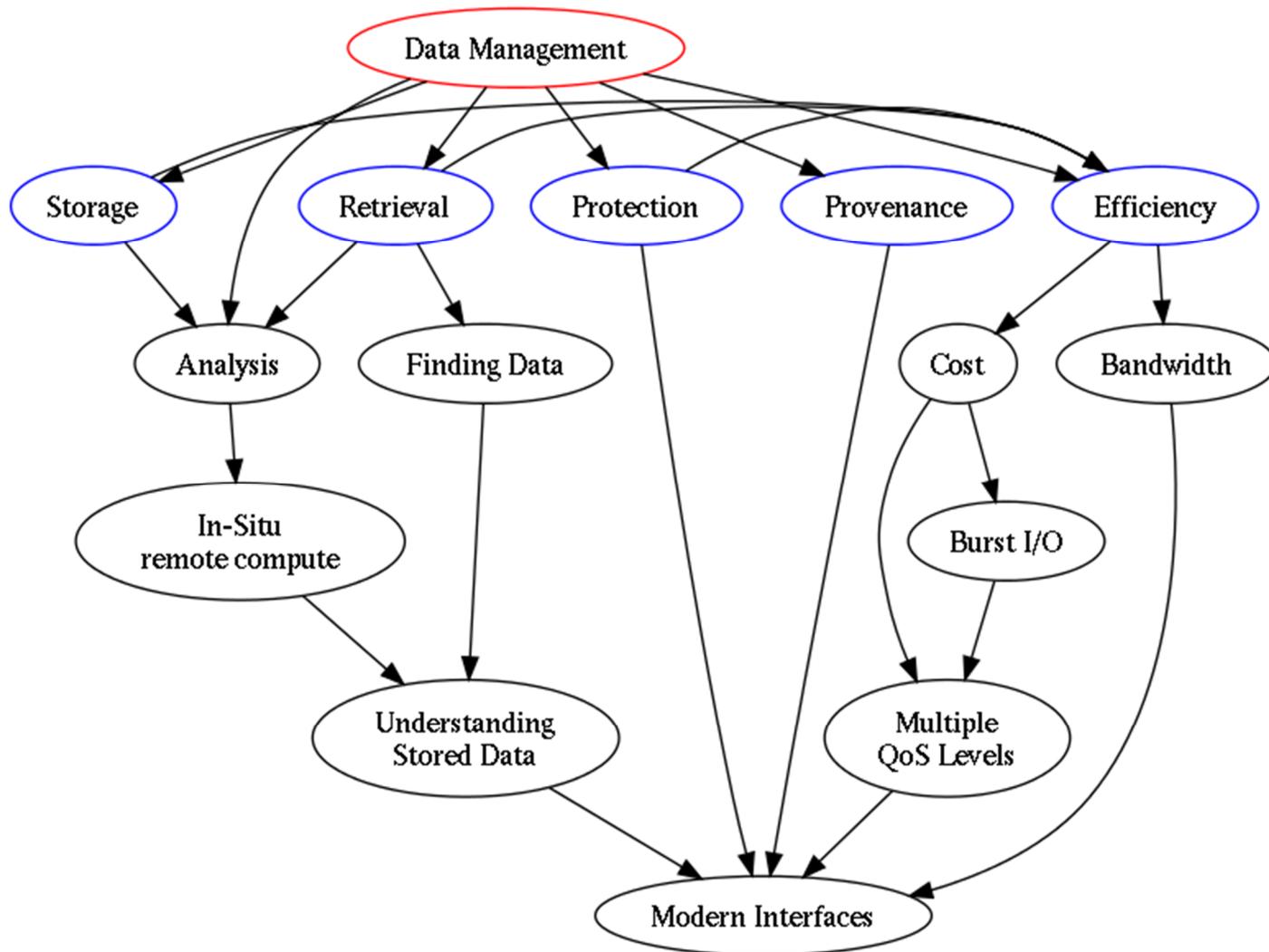
Always seeking collaboration, feedback, advice or other input!

What follows (next 5-10 minutes):

- Briefly listing a collection of data management related topics in no particular order
- A few will be discussed further
- Based on personal opinion and experience
- Topics seen from a storage researcher's point of view
- Very high level overview



Data Management?



Efficiency

Faster *and* Cheaper!



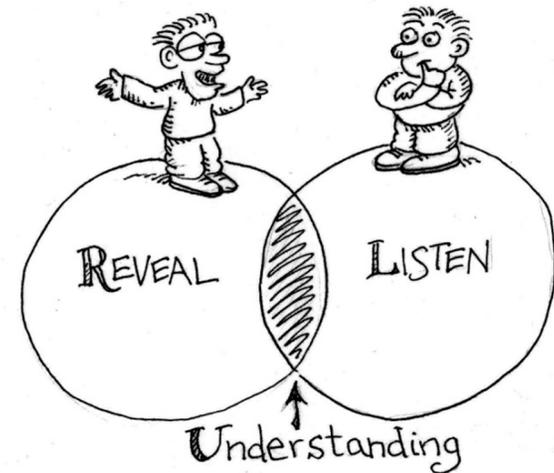
- Increasing data output, but especially **increased demand for bandwidth** have made storage a major component of the overall data management cost.
 - In **HPC** environments, increasing gap between compute and I/O have made **capacity practically free**, while **paying for peak bandwidth**.
 - Bandwidth limited areas (cloud storage), **encoding efficiency** has a considerable influence on profit margins.
- This sparked the creation of **multiple QoS levels**, sometimes exposed to the end user.
 - **Microsoft Azure**: classify data as hot or cold and tune accordingly
 - **Amazon Glacier**: Makes choice explicit
 - **Burst Buffers** in HPC Systems: Provision for **average** bandwidth
- **Caution**: we need more efficiency, where efficiency refers to **the overall goal**, not some small component
 - This can mean **disruptive change** instead of small incremental gains!



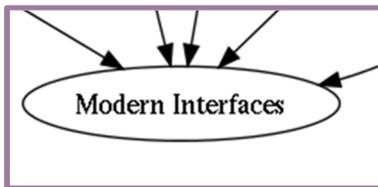
Understanding Data

The Need for Modern Interfaces

- The storage system needs to **understand** what it is storing for reasons of
 - **Performance**: understand how it will be accessed
 - **Provenance**: understand where it came from
 - **Indexing**: understand how to identify
 - **Archiving**: understand dependin
- This requires a **modernization** of our **I/O APIs**
 - Have been unchanged for many **decades** (>25yr!)
 - **Self describing formats** exist, but only **above** storage system (e.g. HDF5, NetCDF, ...)



- **Problem**: it is hard to get people to **accept changes**
 - A lot of **legacy**, both in terms of **code** and in people's **heads**!
 - See what happens if good old '**cd**' goes away...
 - Research exists but never makes it into production



We have the data. What now?

Finding Data

- Traditional directory organization **does not scale** (low efficiency, computer **and** human time)
 - Some recent papers about how to **find** a file name on a large HPC storage system
 - Namespace work (e.g. *SANE*)
- Databases a solution, but does not suffice for everybody
 - **Index partitioning**: search email, databases, directories, ...
- The storage system needs to be more involved in indexing and search

Analyzing Data

- Changed cost equation/system architectures force reconsideration of data handling
 - Energy to **move** one bit (as opposed to **compute**)
 - New (and old) class of algorithms becomes viable
 - **Recompute** is cheaper than fetch in some cases
 - Data-flow languages
- This creates new scientific workflows
 - In-Situ analysis
 - Active-Storage and other remote code approaches



Questions?



Questions? (10 minutes)

Follow-up questions/discussion:
dkimpe@mcs.anl.gov

Acknowledgements

- **Collaborators** at Argonne
 - R. Ross, K. Harms, P. Carns, K. Iskra, R. Latham, D. Goodell, J. Wozniak and others
- ... and other people and institutions in the storage community
- **DoE, DoD & NSF** for making all this possible.
- **MAGIC & NITRD**: Thank you for the invitation and opportunity!

PS. We're always looking for bright people to join our team!
(postdoctoral researchers, software developers) Please contact us if interested!

