



mongoDB

MongoDB at CERN CMS

paul@10gen.com

XLDB 2012 Lighting Talk
September 12, 2012

Acknowledgement

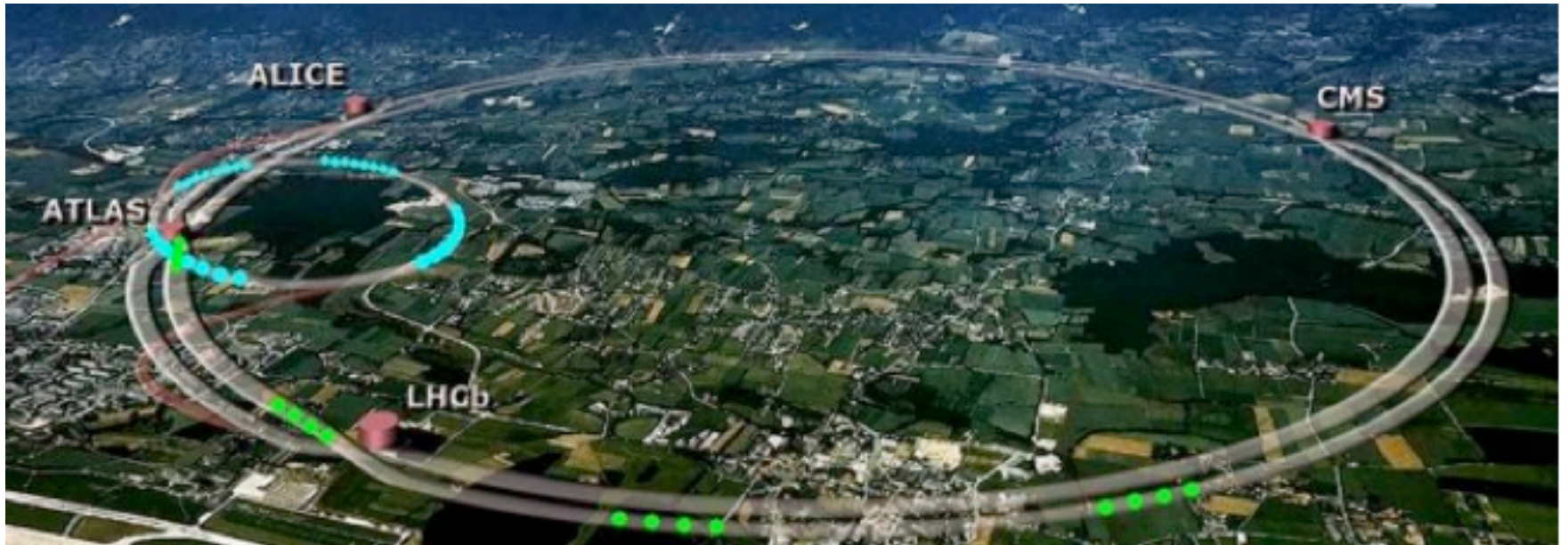
“The CMS Data Aggregation System”, ICCS 2010

Valentin Kuznetsov (Cornell University)

Dave Evans (Fermilab)

Simon Metson (Bristol University)

LHC

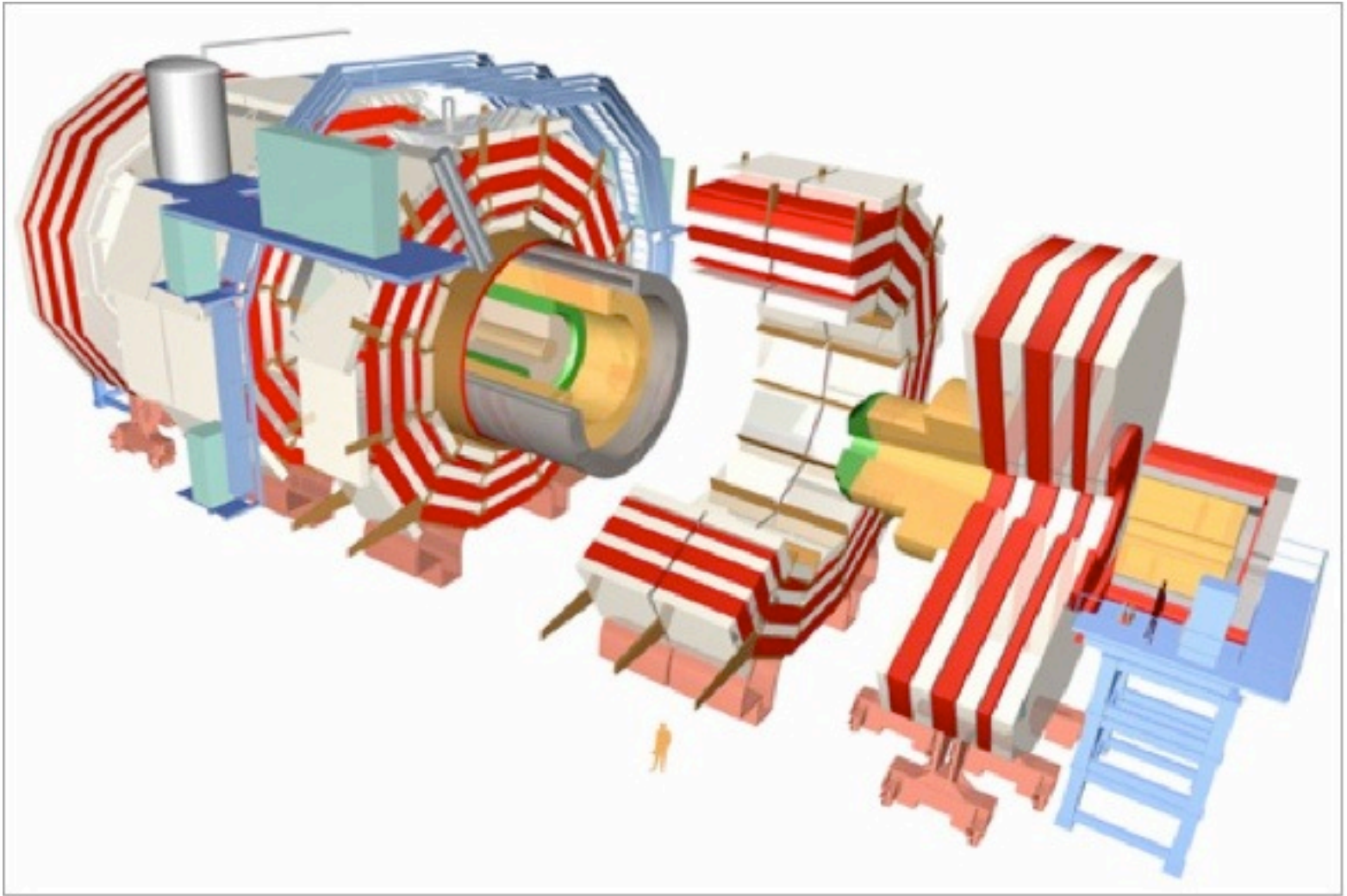


LHC

- Inside the accelerator, two beams of particles travel at close to the speed of light with very high energies before colliding with one another.
- The beams travel in opposite directions in separate beam pipes – two tubes kept at ultrahigh vacuum.
- They are guided around the accelerator ring by a strong magnetic field, achieved using superconducting electromagnets.

CMS

(Compact Muon Solenoid)



CMS

Compact Muon Solenoid

More than 3000 scientists collaborate in CMS, coming from 183 institutes in 38 countries.

The CMS experiment uses a general-purpose detector to investigate a wide range of physics, including

- search for the Higgs Boson (announced 7/4!)
- extra dimensions
- supersymmetry (SUSY)
- particles that could make up dark matter

CMS Data Scale

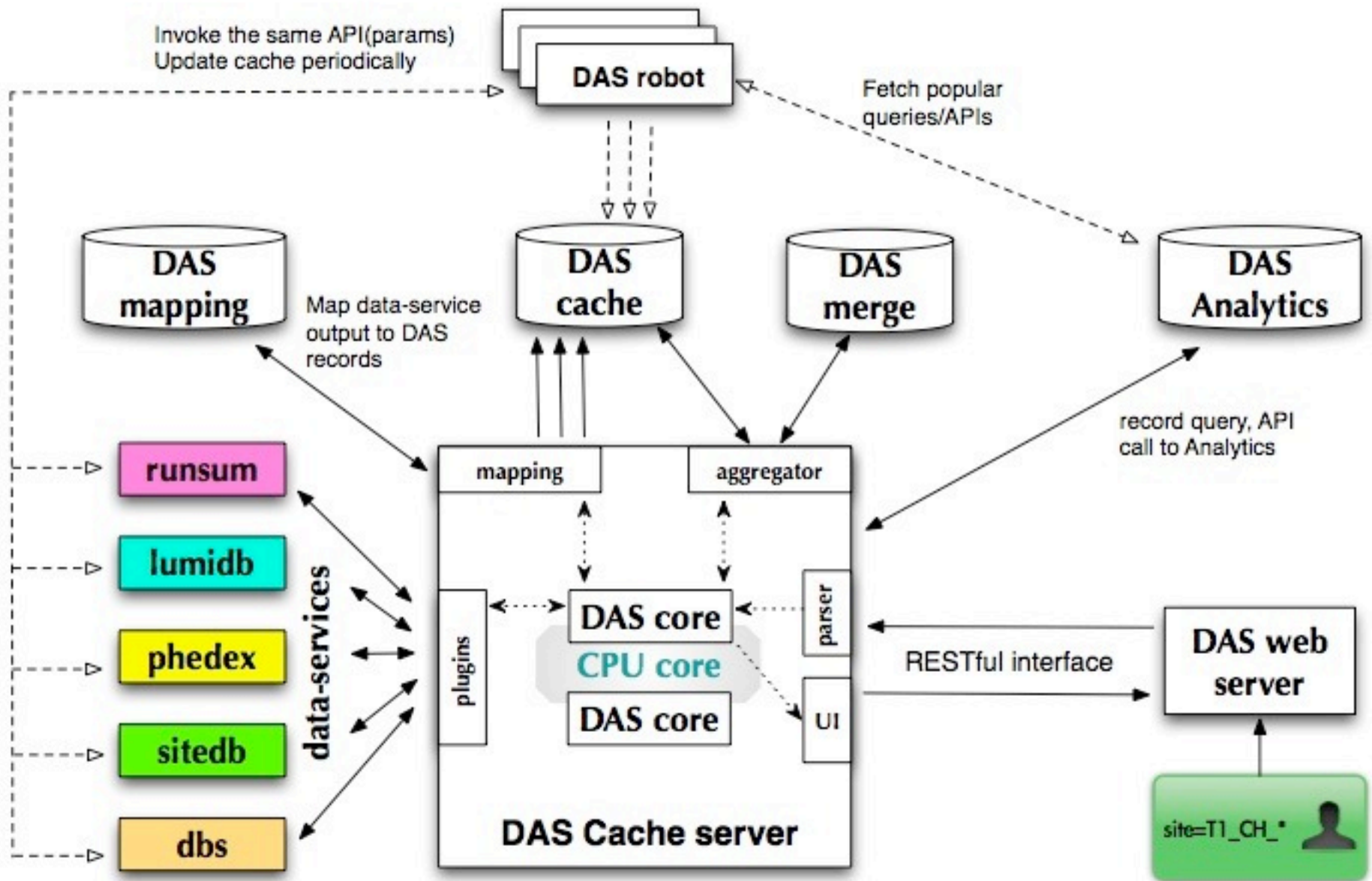
Approximately:

- 2-6 PB of raw data / year
- 500 GB metadata / year

mongoDB stress testing shows:

- 20K docs / sec insert raw
- 10K random reads / sec

CMS Data Aggregation



8

runsum = run summary db

lumidb = luminosity db

phedex = physics experiment data export

sitedb = db of sites

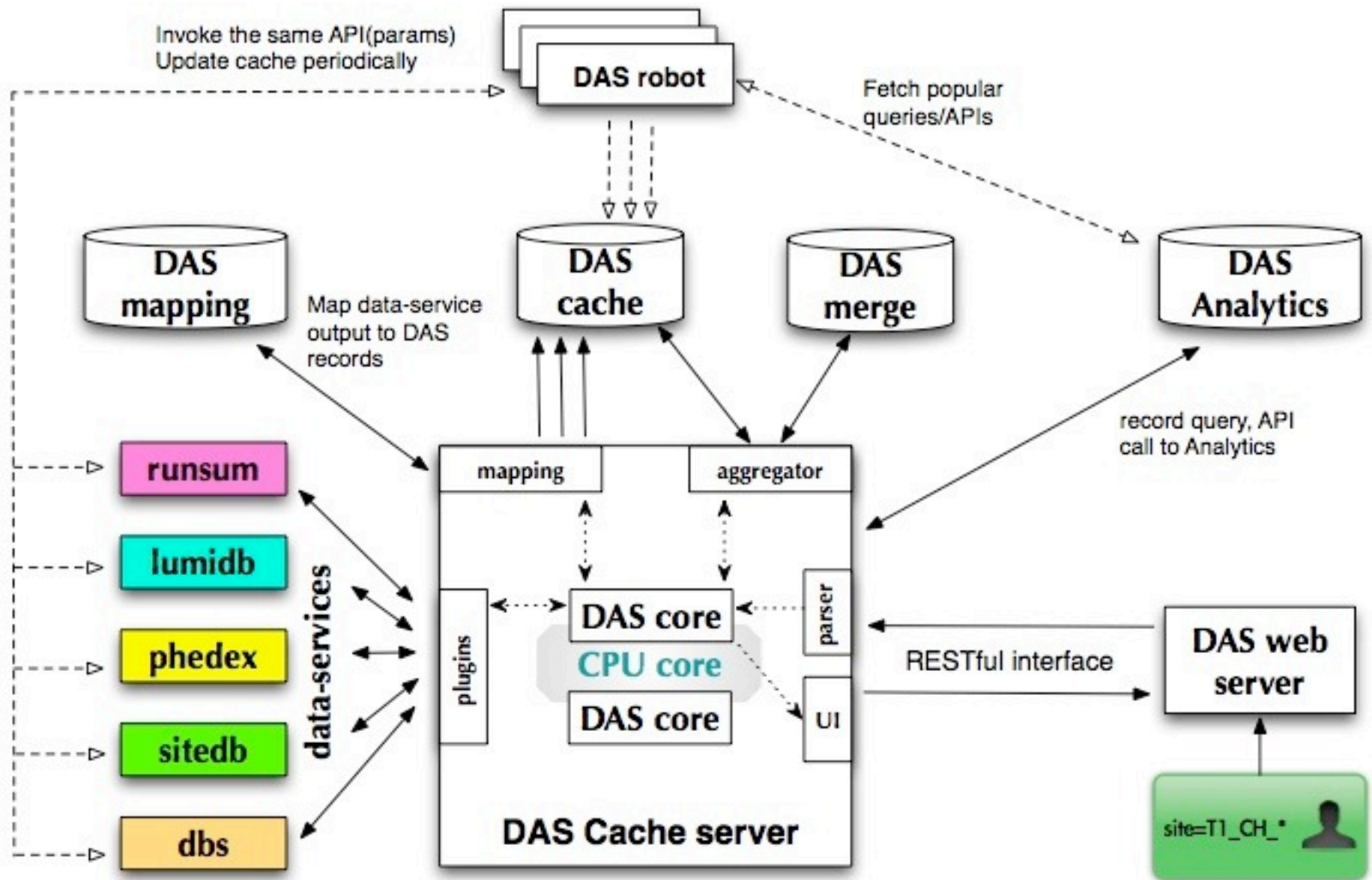
dbs = dataset bookkeeping system - indexes event-data data for the CMS Collaboration. The primary functionality is to provide cataloging by production and analysis operations and allow for data discovery by CMS physicists.

DAS: Why MongoDB?

- Fast database with plenty of native language drivers
- Storage of JSON documents via BSON (binary JSON)
- Simple querying via flexible Query Language (on par w/ SQL)
- Support multiple indexing
- Data organized in collections (a la databases)
- Replication and sharding (key-range partitioning)
- Open source, commercial support via <http://www.10gen.com/support>

Time check / Thank you

CMS Data Aggregation



11

runsum = run summary db

lumidb = luminosity db

phedex = physics experiment data export

sitedb = db of sites

dbs = dataset bookkeeping system - indexes event-data data for the CMS Collaboration. The primary functionality is to provide cataloging by production and analysis operations and allow for data discovery by CMS physicists.

DAS Core Library

Responsible for data retrieval on demand.

It has the following components:

- input query validator and parser
- query analyzer
- workflow builder
- execution engine
- data-service plug-ins

DAS Caching System

The DAS caching system is used to dynamically fetch and aggregate data upon user requests:

- DAS cache
 - store the raw results coming out from data-services
 - expiration timestamps
- DAS merge
 - contains aggregated records

DAS Mapping DB

The Mapping DB is used to track services known to DAS and how DAS should interact with them.

- collect information about data-service APIs
- translate API input/output parameters into DAS keys
- map DAS records to their UI representation

DAS Analytics DB

The Analytics DB collects information on user requests:

- users request
- mapping of requests onto data service API's
- time taken by remote data services to process API calls

Objectives:

- plan pre-fetch strategies for common queries
- identify issues in remote data services
- cross check that DAS resolves requests deterministically

Thank you