Data Transfer Use Cases:

peer-peer (or many-many) or point-point. On demand versus scheduled

Richard P. Mount

June 16, 2010

SLAC NATIONAL ACCELERATOR LABORATORY

PPA Particle Physics & Astrophysics

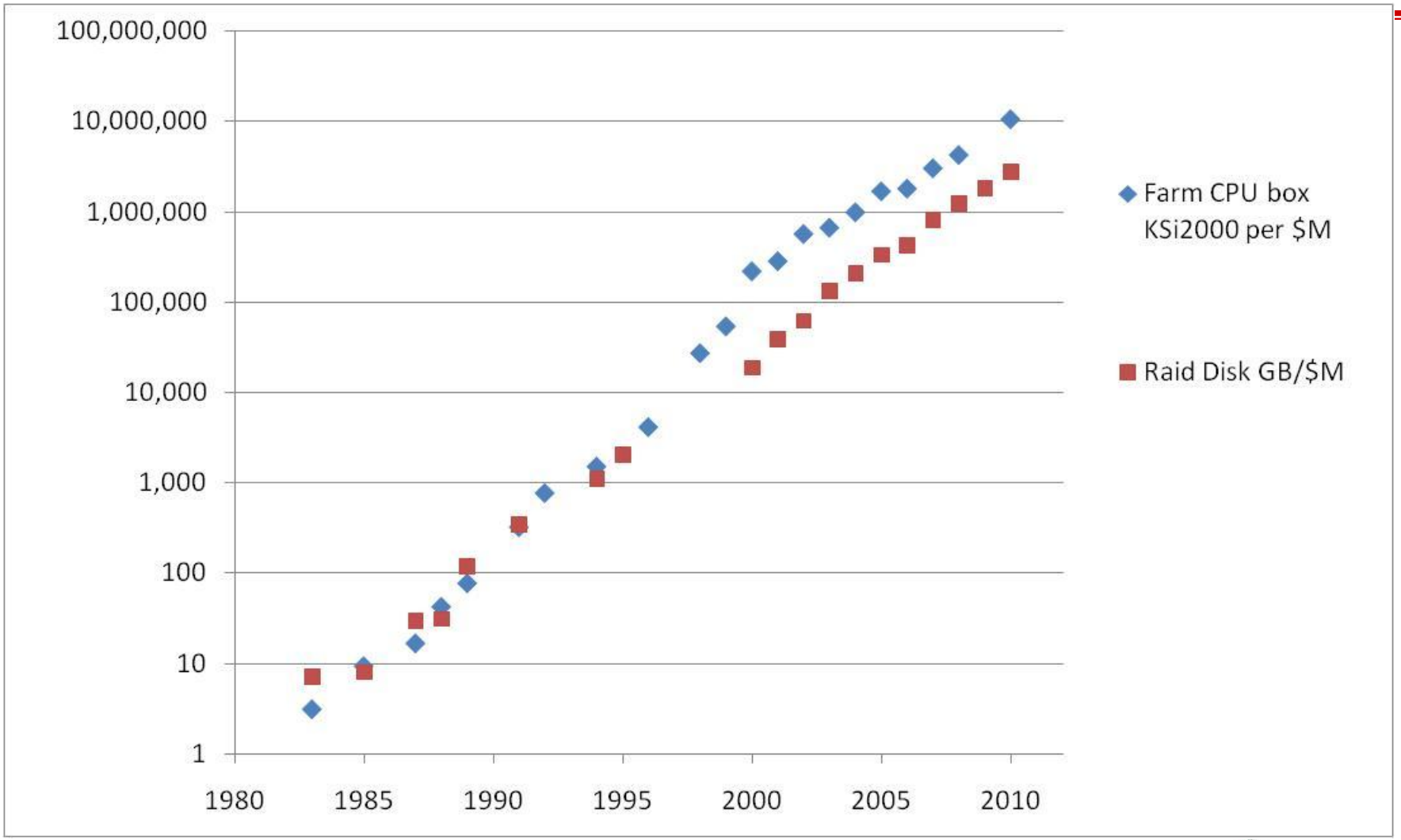
My Starting Point – Necessary Input

1. What Networking, storage and computing are technically and financially feasible?

2. How would we like to analyze our data in a perfect world?

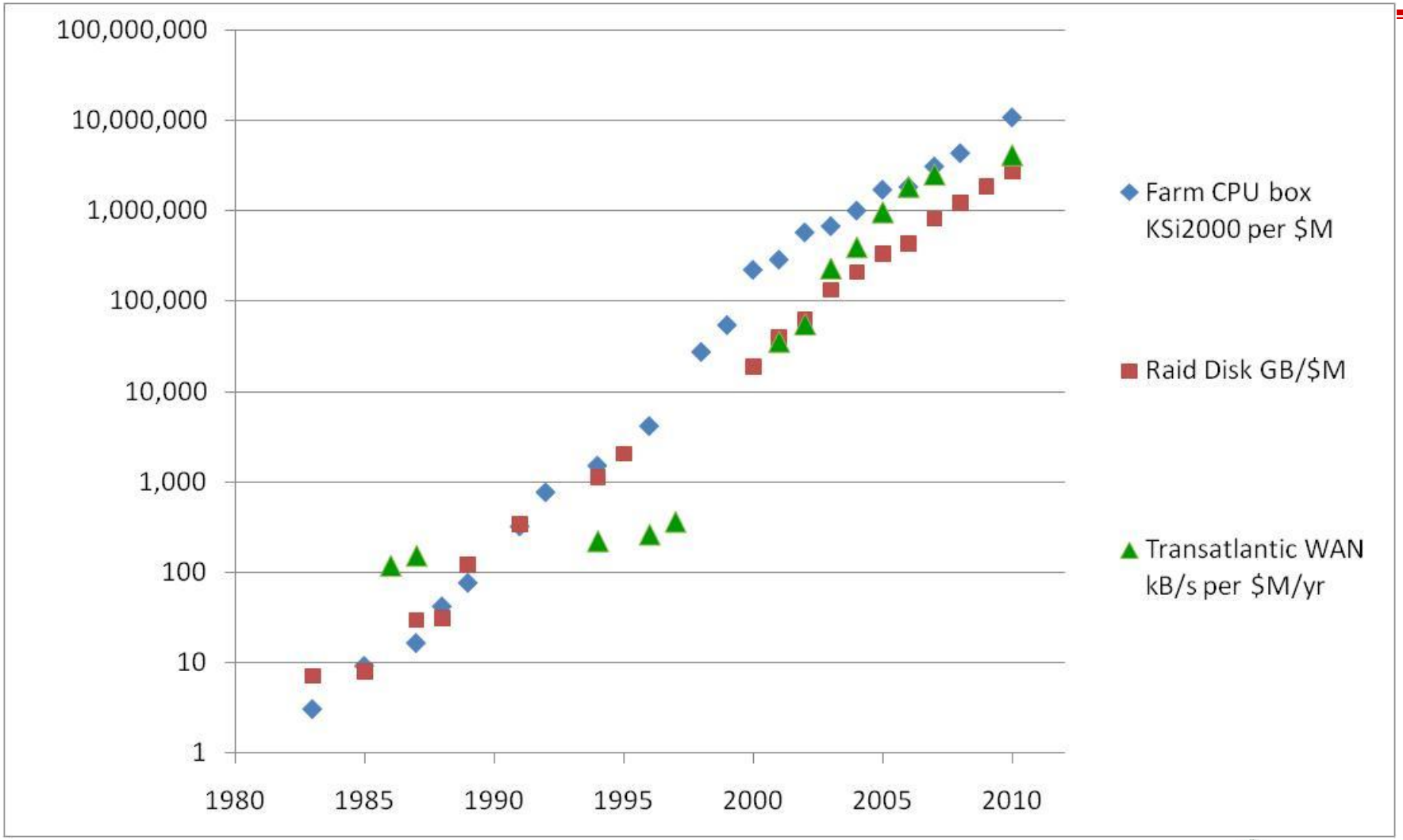
Technical and cost evolution changes physics use cases!

Some Technical and Cost Evolution
(aka stuff I bought)



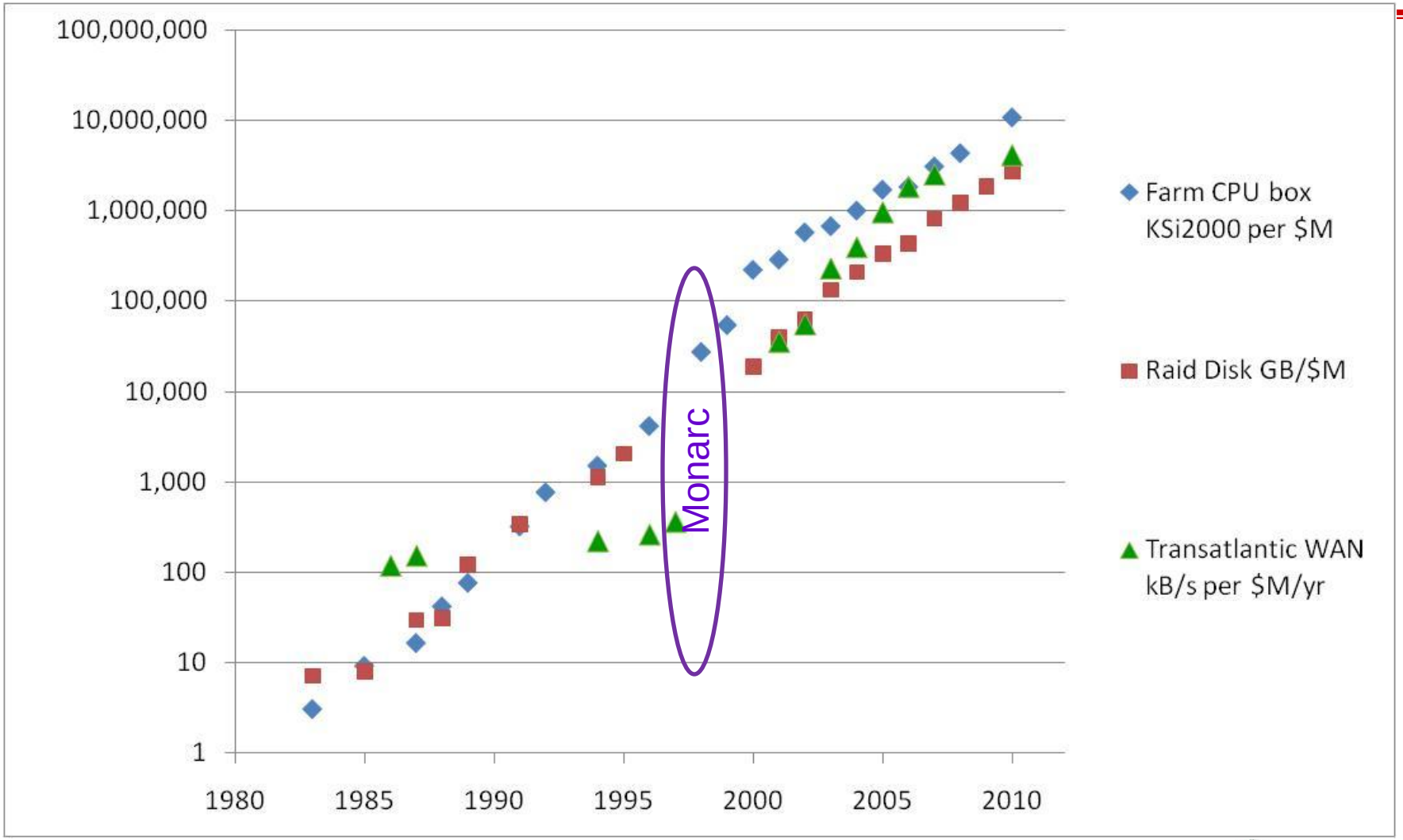
Some Technical and Cost Evolution

(aka stuff Harvey and I bought)



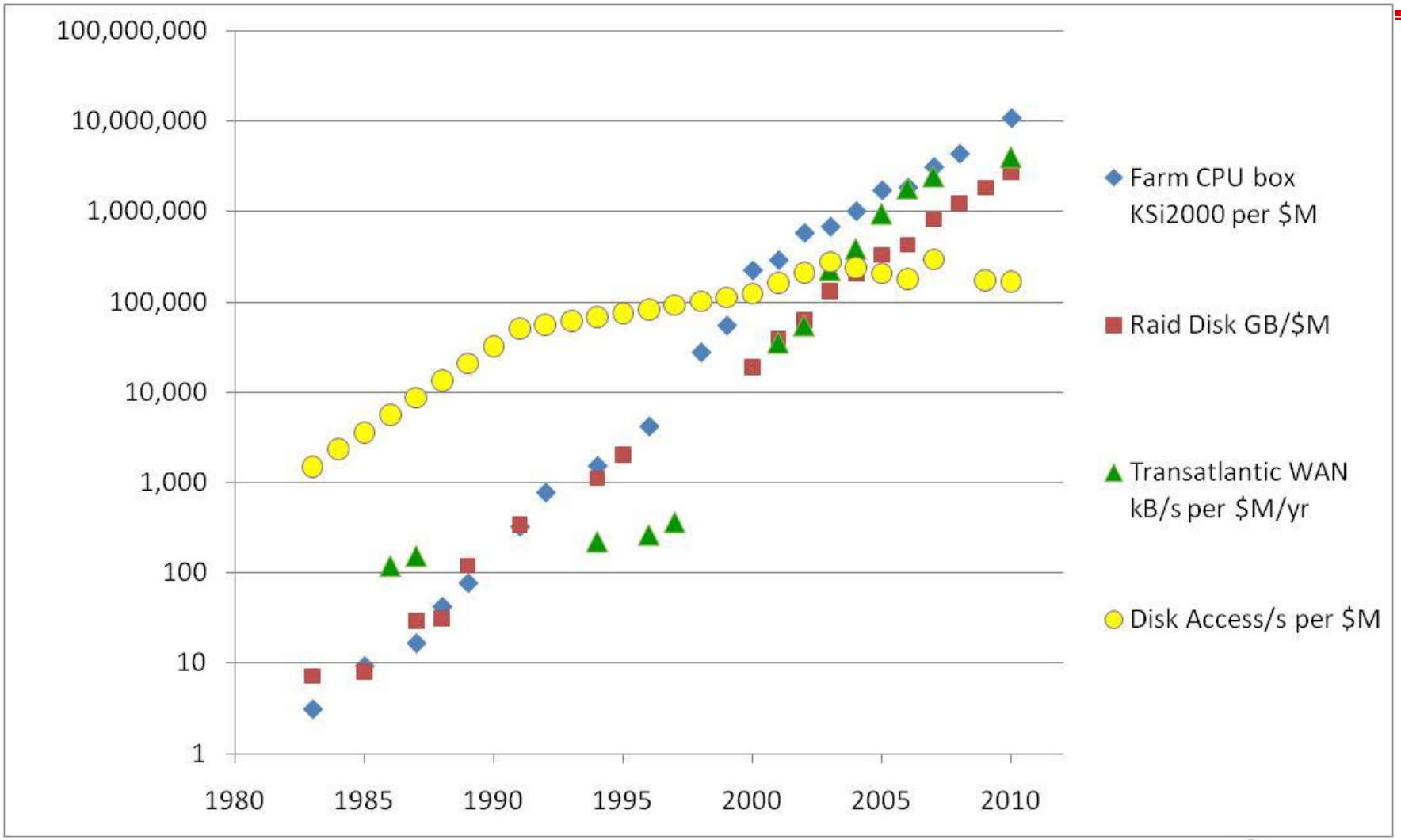
Some Technical and Cost Evolution

(aka stuff Harvey and I bought)



Some Technical and Cost Evolution

(aka stuff Harvey and I bought)



Sparse Data Access

- Disks offer ~100 accesses/s per device
- Disk streaming speed is ~100 MB/s
- A disk delivers sparse 1MB objects at 50% of streaming rate
- (Let's not think about 100 byte objects)

Ideal (=Unattainable) Analysis Environment

- Goal: facilitate detailed examination of anything and everything that helps understand backgrounds and systematics
  (Discovery = understanding the background!)

- [I believe that much of this requires]
  Doubly sparse access to data
  - Sparse selection of events
  - Sparse selection of objects within the events

- And that payload bits are delivered as fast as they can be consumed

Use Cases

- Current (ATLAS) use case:
  - Many/most physicists are doing what they are not supposed to do
  - Accessing the "wrong" datasets
    - ESD instead of AOD
    - Version *n* – *m* (*m*>2), where *n* = current version
  - I believe this provides a good future use case
- Disruptive use cases:
  - Specialized reconstruction study
    - Retrieve raw objects for one or two detector systems for few % of events for ~1 year of data
    - Few hundred core-weeks
  - Calibration
  - Understand background

Types of Event Data

- Production Data
  - Rigidly documented/provenanced
  - May be used by any collaborator to derive publishable results
- Analysis Group Products
  - Well documented/provenanced
  - May be used to derive publishable results
- Individual (or small group of) Physicist's Data
  - May be documented/provenanced
  - Very few people could use it to derive publishable results

Types of Event Data Access

- Access to Production Data
  - By Production Tasks
  - By Analysis Group Production Tasks
  - By Individual (or small groups of) Physicists
- Access to Analysis Group Products
  - By Analysis Group Production Tasks
  - By Individual (or small groups of) Physicists
- Access to Individual (or small groups of) Physicist's Data
  - By (the) Individual (or small group of) Physicists

Types of Event Data Access

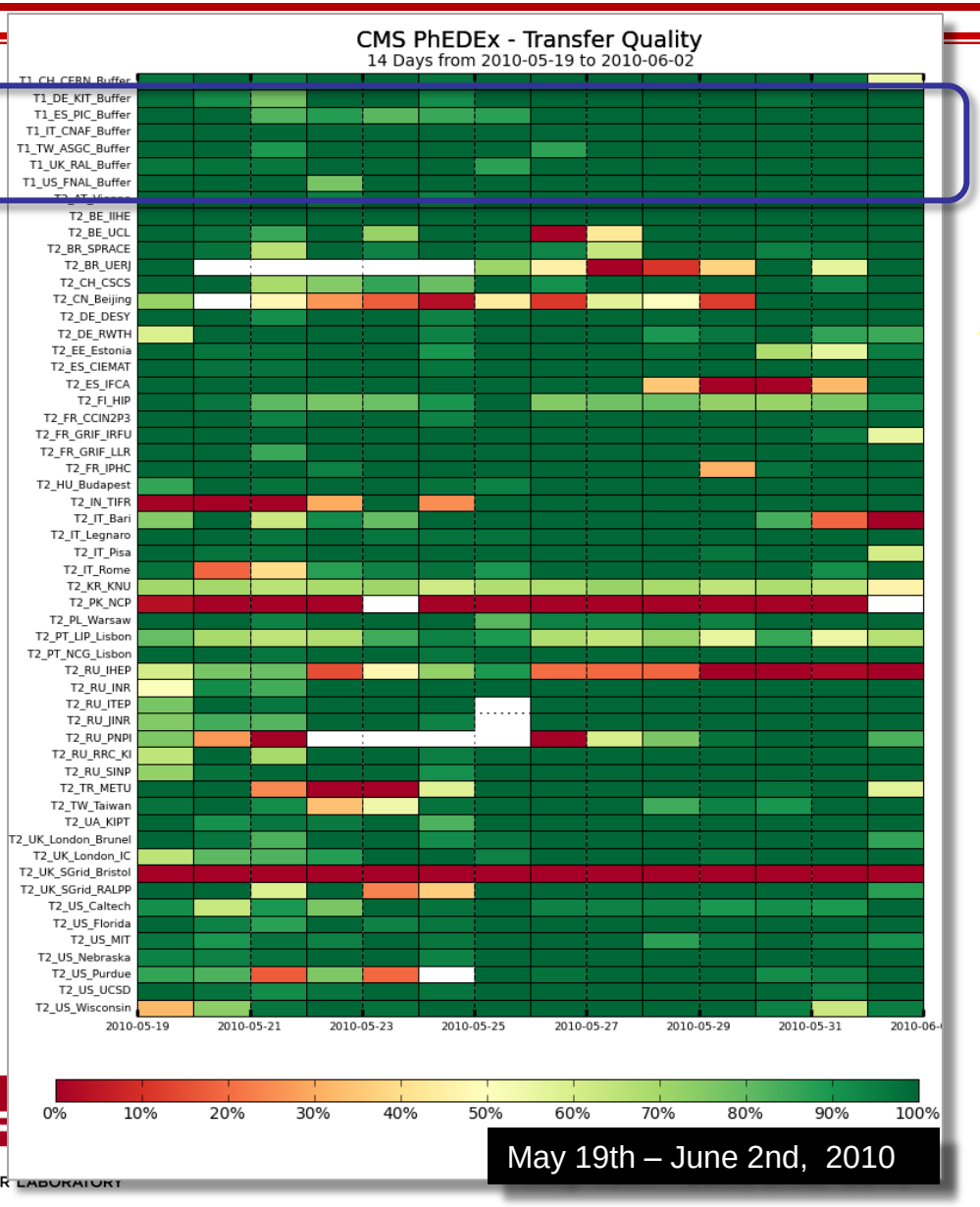
- Access to Production Data
  - By Production Tasks — Very High Volume
  - By Analysis Group Production Tasks — Very High Volume
  - By Individual (or small groups of) Physicists — Very High Volume
- Access to Analysis Group Products
  - By Analysis Group Production Tasks — High Volume
  - By Individual (or small groups of) Physicists — High Volume
- Access to Individual (or small groups of) Physicist's Data
  - By (the) Individual (or small group of) Physicists — Moderate Volume

- Dominant Need: Access to "official datasets"

Access to (Official) Datasets - Issues

- Network Cost, Bandwidth and Latency
  - Are we making enough use of the network?
- What is the optimum (minimum) number of copies of data?
  - keep more versions of data and more simulation
  - Place more demands on network transfers
- Can we hide the network latency?
- Managing data access and transfer:
  - Can we prioritize access when needed?
  - Can we avoid meltdown triggered by unexpected access patterns?
- Can we offer robust, low manpower cost, access in spite of flakey sites?

Slide stolen from Fabio Hernandez – CCIN2P3



Transfer quality for CMS data export from CCIN2P3 to other sites, as measured by the experiment.

CCIN2P3 exchange data with **50+ sites** all over the world.

The quality of every single channel is routinely monitored and human interventions by site experts are triggered when needed.

Source: CMS PhEDEx
http://cmsweb.cern.ch/phedex

SLAC NATIONAL ACCELERATOR LABORATORY

PPA Particle Physics & Astrophysics

On Demand versus Scheduled

- Spectrum includes
  a. Object accessed/transferred on demand
  b. Event on demand
  c. File on demand
  d. Dataset on demand
  e. Dataset scheduled transfer based on measured demand
  f. Dataset scheduled transfer based on imagined demand

- ATLAS has used “f” up to now
- “e” probably makes more sense
- But nothing makes sense without attention to cache management (aka deletion)

On Demand versus Scheduled

- Spectrum includes
  a. Object accessed/transferred on demand — killed by latency
  b. Event on demand — killed by latency?
  c. File on demand
  d. Dataset on demand
  e. Dataset scheduled transfer based on measured demand
  f. Dataset scheduled transfer based on imagined demand

- ATLAS has used “f” up to now
- “e” probably makes more sense
- But nothing makes sense without attention to cache management (aka deletion)