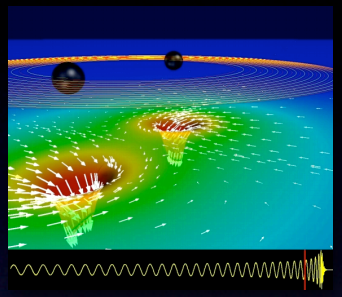
Engineering the Data Processing Pipeline

Mark Stalzer
Center for Advanced Computing Research
California Institute of Technology
stalzer@caltech.edu

October 29, 2009

A systems engineering view of computational science & engineering and how this might apply to the data processing pipeline...

CACR Mission & Partners



PES
Center for Plasma Edge Simulation

bnmc

cell center

LIGO

Computational Infrastructure for Geodynamics

US CMS

UltraLight

JPL
Jet Propulsion Laboratory
California Institute of Technology

Seismological Laboratory

DANSE

GALCIT
MECHANICS: FLUIDS, SOLIDS, BIOSYSTEMS, AND SPACE

NVO
NATIONAL VIRTUAL OBSERVATORY

ASC™

CALTECH
PSAAP

Accelerating scientific discovery & engineering through advanced computation, collaboration and research

CACR

The (Tsunami) Wave of the Future

- Now: data streams of ~ **0.1 TB / night, ~ 10 - 10^2 transients / night** (SDSS, PQ, various SN surveys, asteroid surveys)
- Forthcoming on a time scale ~ 1 - 5 years: ~ **1 TB / night, ~10^4 transients / night** (PanSTARRS, Skymapper, VISTA, VST…)
- Forthcoming in ~ 5 - 10 years: LSST, ~ **20 TB / night, ~ 10^5 - 10^6 transients / night**

A major, qualitative change!

- Observational follow-up needs:
  - Rapid photometric/positional monitoring
  - Rapid spectroscopy
  - Information/computation infrastructure

Transient classification technologies are essential

Source: George Djorgovski Caltech

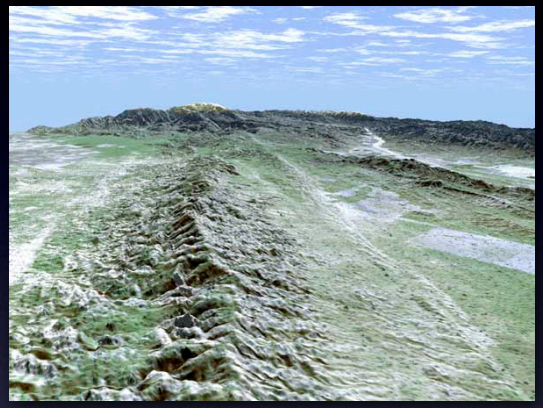
Sample CSE Systems Design Questions

In a systems view, the questions affecting computer resources are different -

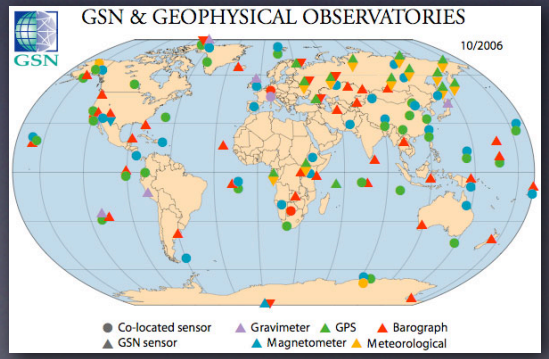
- How best to deploy resources (experiments, model development, computation) to certify a given design to 99.99%?
- What's the next best experiment?
- *Which earthquakes best resolve Earth's structure?*
- *How does experiment sensitivity change with computer resources?*

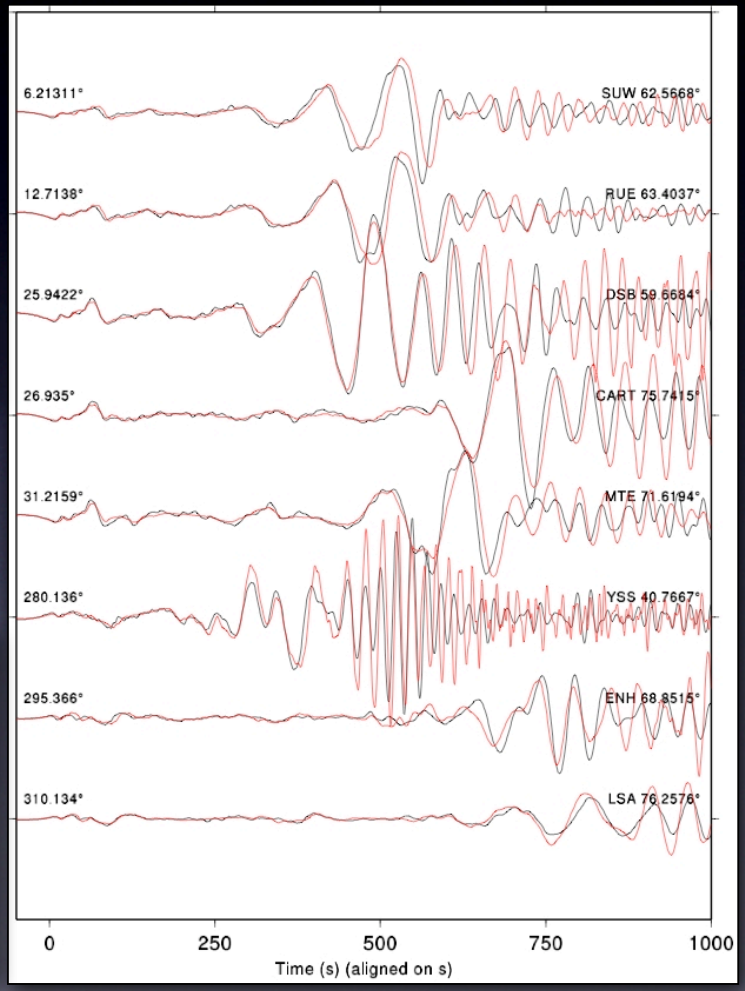
CACR

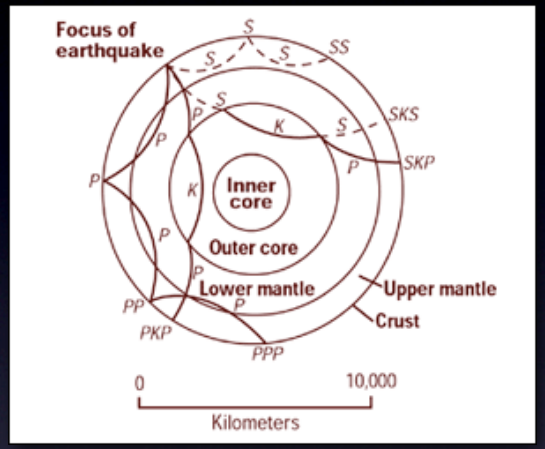
Earth Tomography



Earthquake recorded globally and stored.



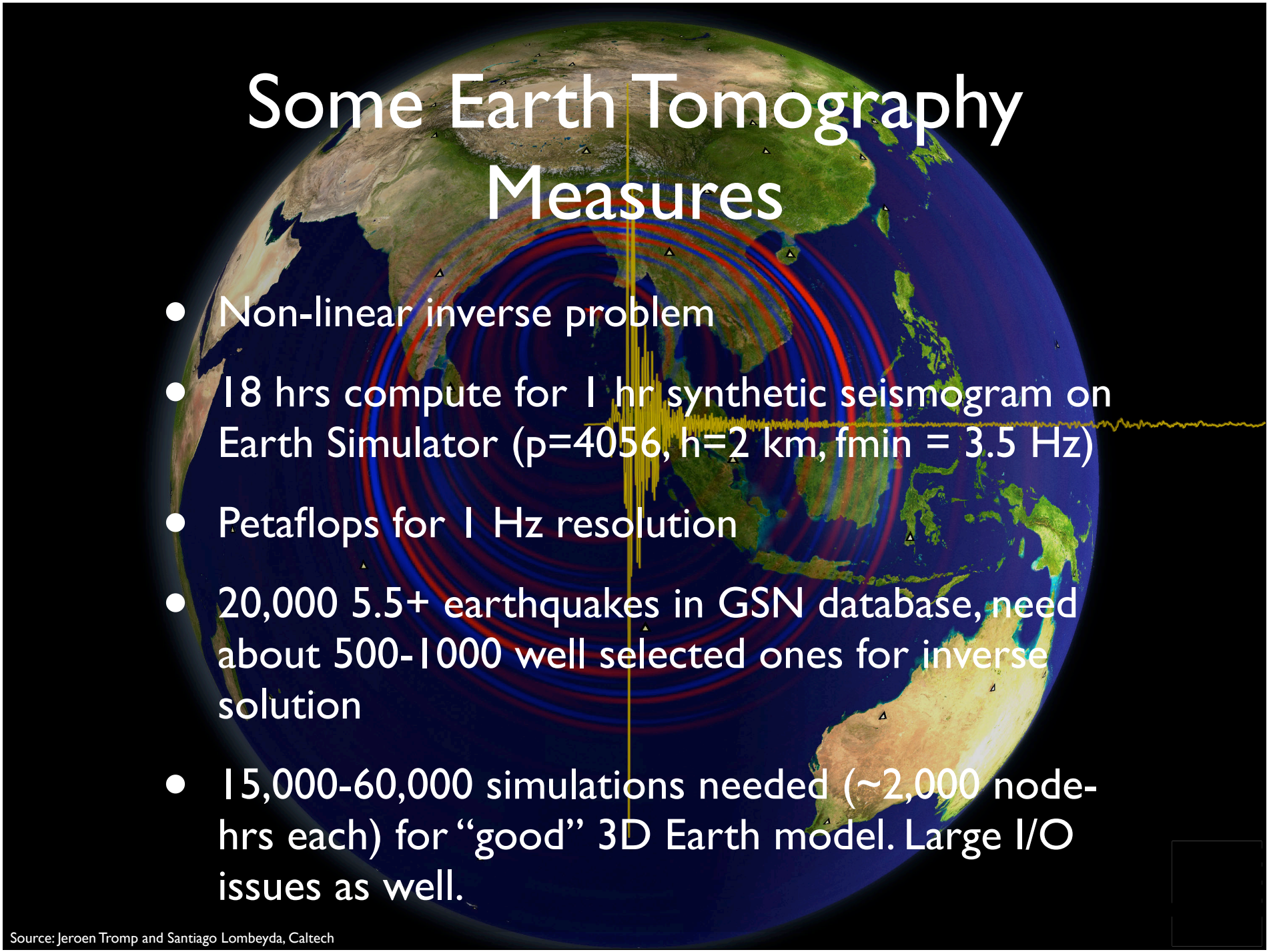




3D Earth model and sources used to compute synthetic seismograms. These are compared with recorded data and used to refine the Earth model.

CACR

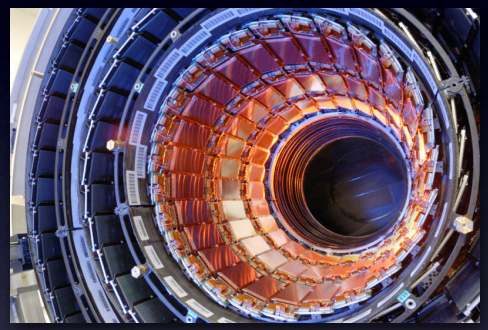
Some Earth Tomography Measures



- Non-linear inverse problem
- 18 hrs compute for 1 hr synthetic seismogram on Earth Simulator (p=4056, h=2 km, fmin = 3.5 Hz)
- Petaflops for 1 Hz resolution
- 20,000 5.5+ earthquakes in GSN database, need about 500-1000 well selected ones for inverse solution
- 15,000-60,000 simulations needed (~2,000 node-hrs each) for "good" 3D Earth model. Large I/O issues as well.

Source: Jeroen Tromp and Santiago Lombeyda, Caltech

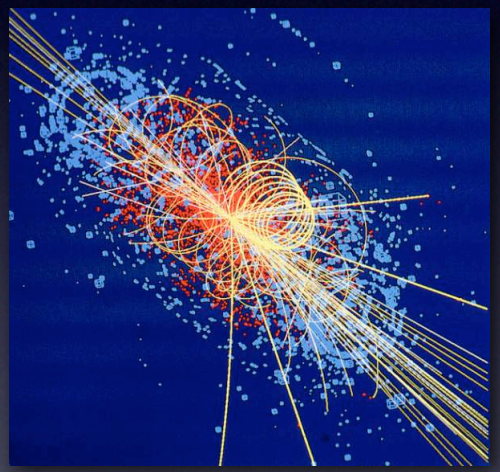
LHC/CMS Data Analysis

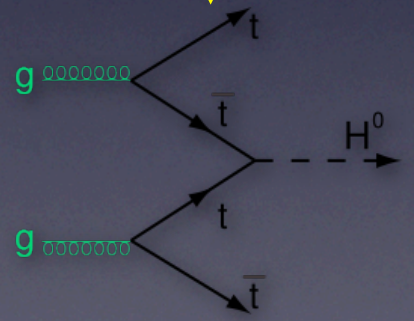


Tier0

Tier1_1
Tier1_2
.
.
.

Tier2_1
Tier2_2
Tier2_3
Tier2_4
.
.
.

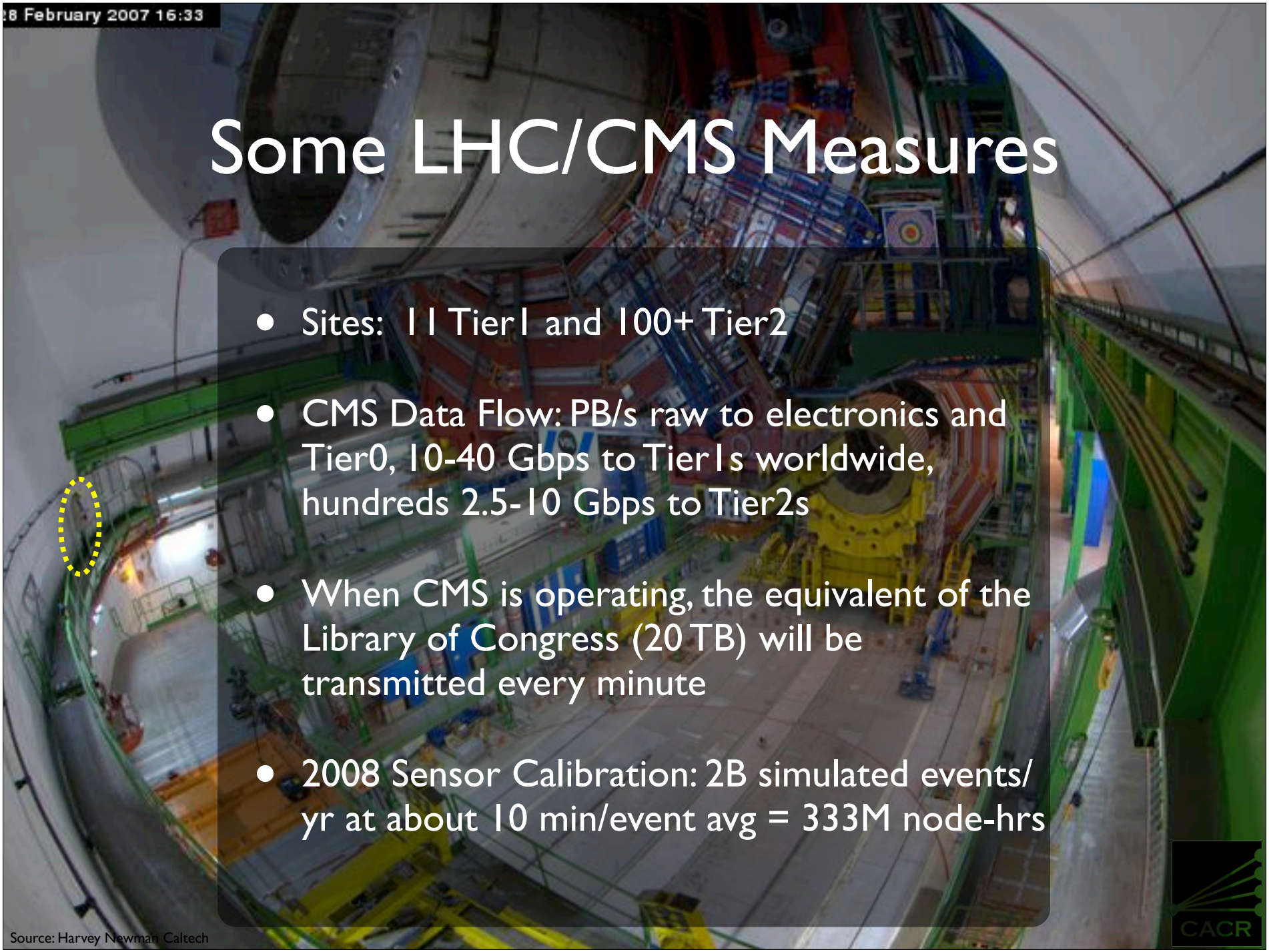




Virtual Detector
-Design
-Calibration
-Transport
-Classification

New physics?

CACR



Some LHC/CMS Measures

- Sites: 11 Tier1 and 100+ Tier2
- CMS Data Flow: PB/s raw to electronics and Tier0, 10-40 Gbps to Tier1s worldwide, hundreds 2.5-10 Gbps to Tier2s
- When CMS is operating, the equivalent of the Library of Congress (20 TB) will be transmitted every minute
- 2008 Sensor Calibration: 2B simulated events/yr at about 10 min/event avg = 333M node-hrs

Source: Harvey Newman Caltech

CACR

Commonalities

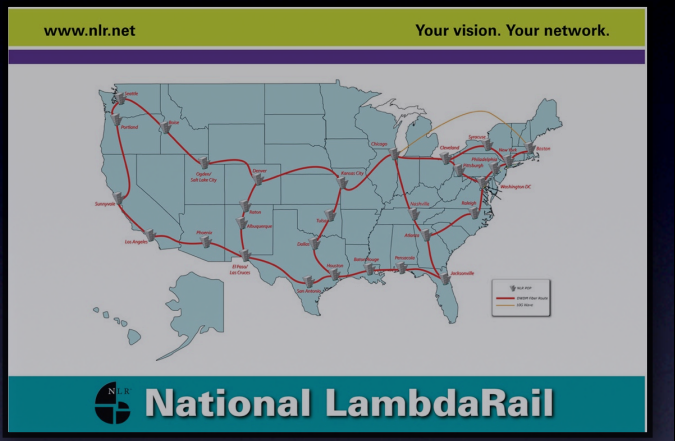
- Science driven; engineering enabled
- Complex social (virtual) organizations
- ***Integrated*** experiments & computation - near real-time
- Large simulations as "inner loops" & sensor calibrators
  - Importance of algorithms...
- Massive data sets due to sensors that are large and complex, or cheap and numerous, that require transportation and storage

It used to be that people built the best instruments and then figured out the results; now people build the best sensor architectures and then hope that computers figure it out.

CACR

Making Costs Explicit (Amazon EC2)

+



Instances

United States Europe

Standard Instances	**Linux/UNIX**	**Windows**
Small (Default)	$0.10 per hour	$0.125 per hour
Large	$0.40 per hour	$0.50 per hour
Extra Large	$0.80 per hour	$1.00 per hour
High CPU Instances	**Linux/UNIX**	**Windows**
Medium	$0.20 per hour	$0.30 per hour
Extra Large	$0.80 per hour	$1.20 per hour

Sources: New York Times & National LambdaRail & Amazon

CACR

Applications Costs for EC2

Certification: $5.4B (3 months at 100 Pflops; weapons labs ~$6B/yr)

Single reaction: $2.7x10^{-11} (1 ms/r; how much to simulate a human?)

New Earth model: $7.5M

LHC/CMS calibration for 2008: $33M (LHC & detectors ~$5B)

Not having to run a supercomputing center: *priceless*

Explicit costs allow design trades

(See Walker, The real cost of a CPU hour, *Computer,* 35-41, April 2009. Also, Strand, KEYWORD: EVIL: Google's addiction to cheap electricity, *Harpers Magazine,* March 2008)

CACR

Pipeline Data Stream Measures

Look at optical as typical case

Earth land surface is 150,000,000 sq km

Satellite images are 60 sq km & 600 MB in 2.5m band

Fly-over time is about 2-3 days

Implies *10 TB/day* data stream ((150,000,000/60^2)*600/2.5)

1 TB disks are ~$150 (also have SAR & LIDAR & other optical bands)

Will need OC192 (10 GbE) networks to stream

CACR

Pipeline Processing Measures

Baseline: Image registration takes 2 hours per image on a core

High resolution optical could have 20,000 images/days

Many other processing loads:

- Other sensors/bands
- Real-time (cross look on close orbits)
- Sensor fusion?

At least 10,000 full cores, would be *~50 on TOP500 List*

CACR

What are Some Systems Engineering Qs for the DPP?

- How much data is needed for change detection & what is the processing load for detection & fusion?
  - Driven by science missions
  - Can benchmark LSST data storage & processing
  - *Little computers: How can we improve the best known algorithms?*
- Distributed (e.g. CMS) or central architecture?
  - Scalable architecture for new detectors (10 years?)
- What does the virtual organization look like?
- Costs?

Workshop is timely, just inside boundary of deployable resources

CACR