



LSST simulator and data

Challenges for LSST type data sizes

Andy: Simulator to prepare us for the data statistically.

What about computationally?





LSST simulator and data

- Working with LSST will data be different?
 - Can we still have casjobs (web+sql)?
 - Downloading data chunks to your PC?
 - 30 year old f77 code still doable, data size matters
 - 32bit versus 64bit
 - Will data locality be required?





LSST data bottlenecks

dB sizes (alone) will be as large as 10PB?
60PB of images after 10 years.

- Several geographically separate data centers
- Casjobs type system is possible?
 - Query scales almost linear with size
 - Derive own catalogs and joins (as today)
 - Variable objects?
 - Already have experience with Stripe 82 type data





LSST data bottlenecks

- Data Mining/Machine Learning are N^2 or worse
 - Classification/clustering, density estimation, etc
 - Inverting large non-sparse matrices...
- Do we need data locality?
 - SciDB will have R interface for “expert users”
 - Lets aim for python (numpy, scipy, Rpy, mlabwrap)

OR

- If I’ve (somehow) downloaded 500TB of data from casjobs can I work with it in the year 2020?





LSST data bottlenecks

- When does one need Data Locality?
 - Only worth moving if need 100,000 CPU cycles per byte of data (Bell, Gray, Szalay 2006)
 - e.g. Seti@HOME, cryptography, etc
 - Most science apps are more in-line with Amdahl's ROT and are more info intensive (CPU:IO \ll 10,000:1)
 - For PB sized datasets we must co-locate computer power with dB
 - But the naïve person says →



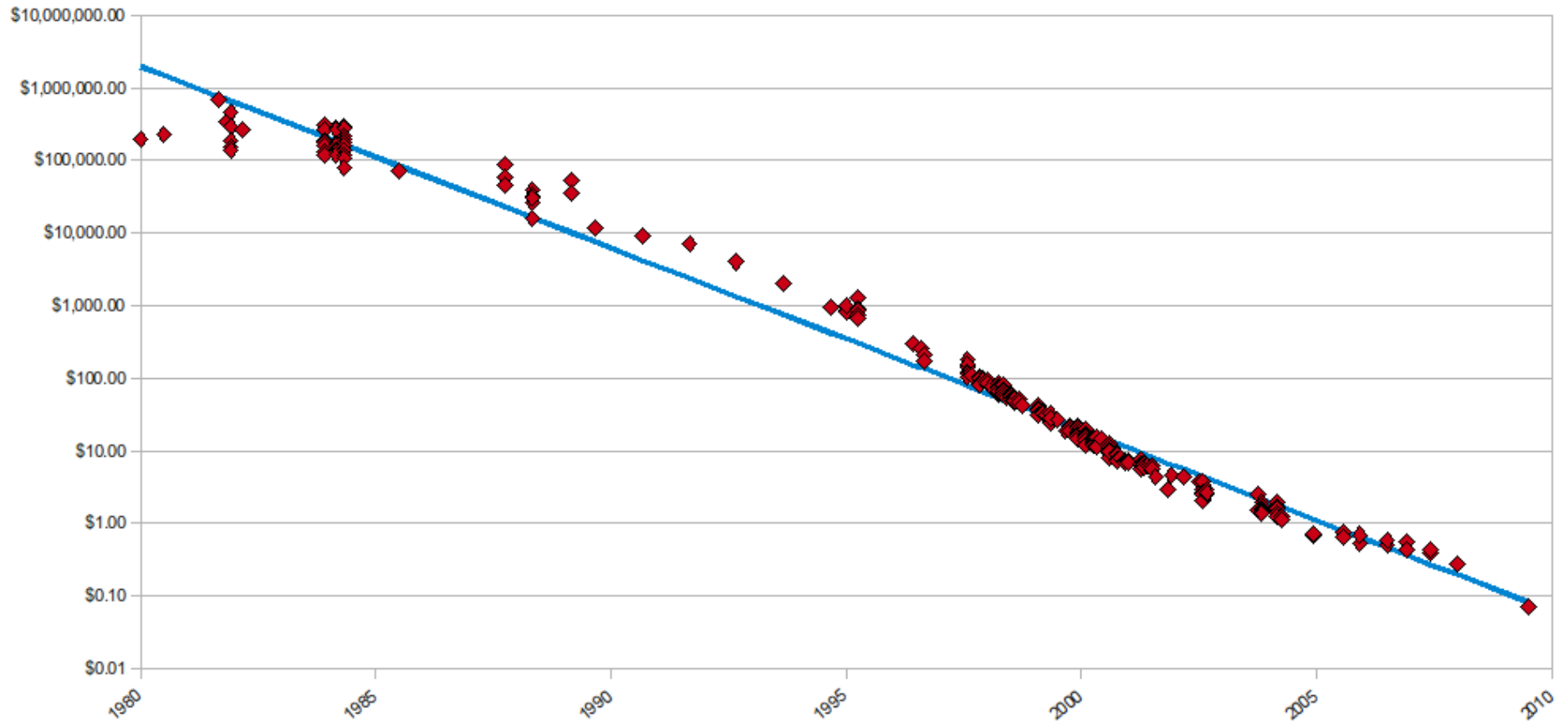


LSST data bottlenecks

- What if I really want 1PB on my desktop?
 - Today: 1PB at 1Gbs \Rightarrow 100 days
 - 1PB Cost today:
 - Raw drive: $\$80/2\text{TB} \times 500 \text{ drives} = \$40,000/\text{PB}$
 - By 2020: 1PB at 10Gbs \Rightarrow 10 days (not bad!)
 - 1PB cost by 2020? If historical trends continue (e.g. <http://www.mkomo.com/cost-per-gigabyte>)
 - 2000: $\$19.00000/\text{GB} * 1000 = \$19,000/\text{TB} = \$19 \text{ million}/\text{PB}$
 - 2010: $\$00.06000/\text{GB} * 1000 = \$62/\text{TB} = \$62,000/\text{PB}$
 - 2020: $\$00.00002/\text{GB} * 1000 = \$0.2/\text{TB} = \$200/\text{PB}$
 - 1PB on my desktop by 2020 is possible?



Hard Drive Cost per Gigabyte 1980 - 2009



- <http://www.mkomo.com/cost-per-gigabyte>
- $Cost = 10^{-0.2502(year-1980)+6.304}$





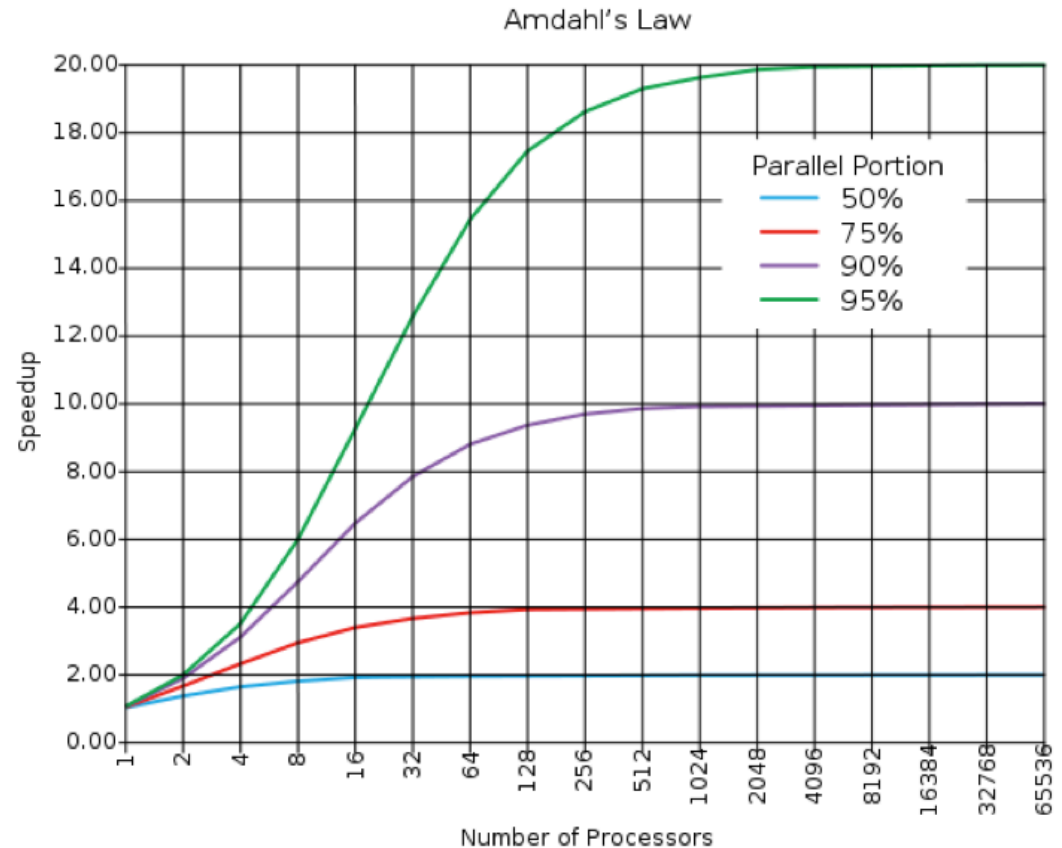
LSST data bottlenecks?

- Time machine back to 1998
 - SDSS data was also considered impossibly large
 - Yes, you can get a local copy of the entire SDSS dB today, but is it worth the trouble?
 - Maybe, but how many people have one?
- Will we experience the same thing with LSST?
 - I/O, I/O, I/O is the issue today, not back in 1998...



Lets remember some of Amdahl rules of thumb for a balanced system

1.) Amdahl's "Law": Max speed-up for a computation with Sequential (S) and Parallel (P) parts is: $S/(S+P)$



2.) Bandwidth (BW):

one bit of IO/second per instruction/second

3.) Memory: $\alpha=1=MB/MIPS$

One byte of memory per one instruction/sec

Problems today...

- CPU Heavy, but I/O poor
 - High Performance Comp have I/O & storage issues

Using Amdahl's laws for a PB scale balanced system:

OPS	RAM	Disk I/O Bytes/s	# Disks for that BW at 100MB/s/disk	Disk Byte Capacity	# disks for that capacity at 1TB/disk
Giga/10 ⁹	GB	10 ⁸	1	10 ¹¹	1
Tera/10 ¹²	TB	10 ¹¹	1000	10 ¹⁴	100
Peta/10 ¹⁵	PB	10 ¹⁴ (100TB/s)	1,000,000	10 ¹⁷	100,000
			500,000 (200MB/s)		

100TB/s of I/O bandwidth required

Takes 1,000,000 disks to deliver (at 100MB/s/disk)

100,000 disks to store 100PB produced/used by machine



LSST data bottlenecks?

- Even if the cost comes down how can I pay for the electricity and space?!
 - We're not Google – and even they have problems with cpus waiting for I/O
- Now you do believe in data locality?





LSST data bottlenecks?

We will deal with the data

- We will probably have to learn new ways to handle computation remotely on the databases
 - SciDB+R and hopefully SciDB+python et al.
- Tremendous opportunities with LSST simulator
 - Imagine how much **better** the early SDSS papers would have been with something similar

