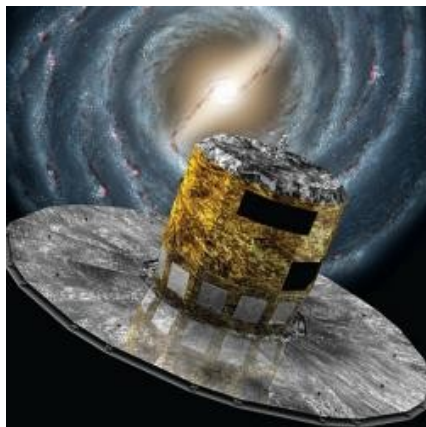


Report on ESA Gaia activities & VLDB 2010

Pilar de Teodoro
Gaia Science Operations database administrator
European Space Astronomy Centre (ESAC)
Madrid
Spain



The Milky Way

'Our Sun'



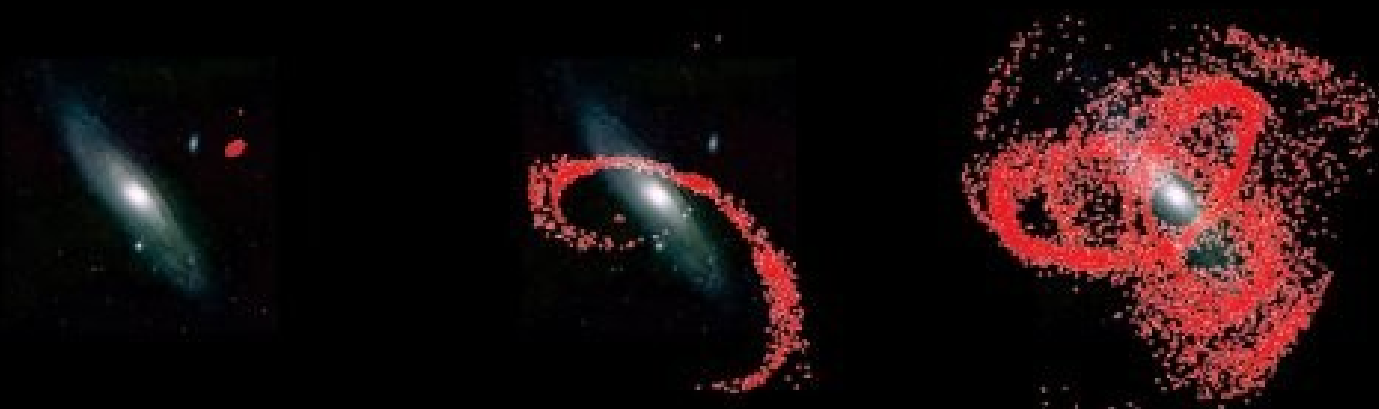
Gaia's main aim: unravel the formation, composition, and evolution of the Galaxy

Key: stars, through their motions, contain a fossil record of the Galaxy's past evolution

Gaia's key science driver

Stellar motions can be predicted into the future and calculated for times in the past when all 6 phase-space coordinates (3 positions, 3 velocities) are known for each star

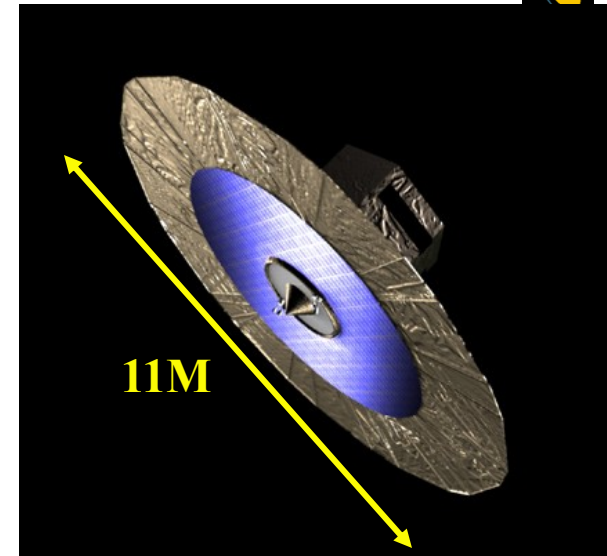
The evolutionary history of the Galaxy is recorded mainly in the halo, where incoming galaxies got stripped by our Galaxy and incorporated



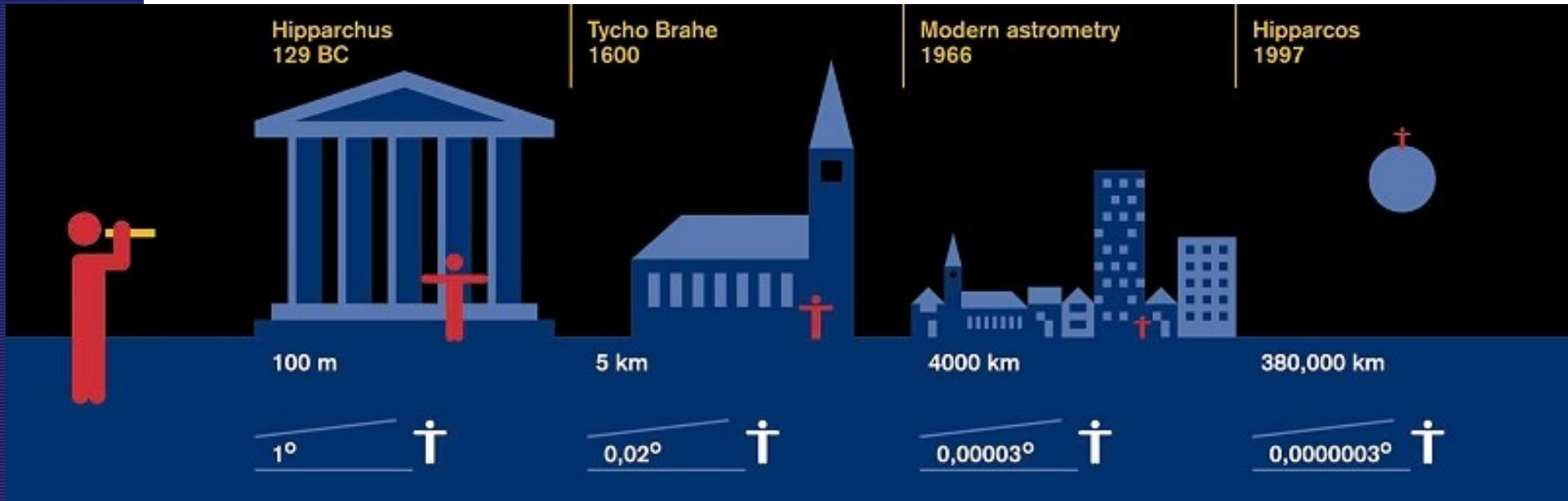
In such processes, stars got spread over the whole sky but their energy and (angular) momenta were conserved. Thus, it is possible to work out, even now, which stars belong to which merger and to reconstruct the accretion history of the halo

Satellite and Mission

- Mission:
 - Stereoscopic Census of Galaxy
 - μ arcsec Astrometry $G < 20$ (10^{10} sources)
 - Radial Velocities $G < 16$
 - Photometry millimag $G < 20$
- Catalogue due 2020
- Status:
 - ESA Corner Stone 6
 - ESA provide the hardware and launch
 - Mass: 2120 kg (payload 743 kg)
 - Power: 1631 W (payload 815 W)
 - Launch: 2012 to L2 (1.5 Million Kms)
 - Satellite In development /test (EADS/Astrium)
 - Toros complete
 - Many mirrors, most CCDs



Hipparchus to ESA's Hipparcos



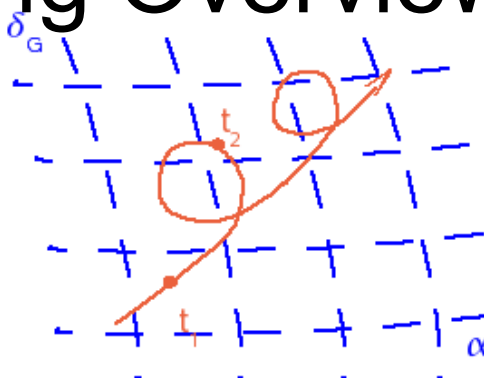
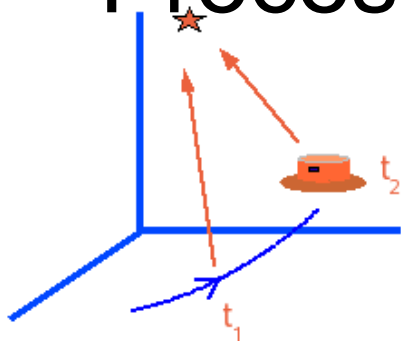
Gaia will take us to the next order of magnitude the microarcsecond.

e.g. A euro coin on the moon viewed from earth

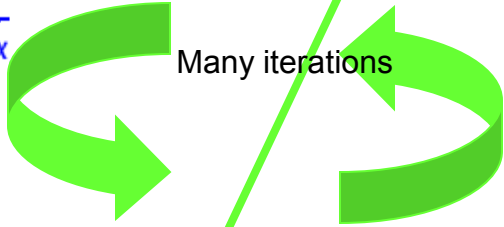
Data Flow challenges

- Daily data flow not a problem
 - ~50GB per day over standard internet
- Using Aspera/ FASP for now
- The MainDB updates may be a problem
 - 100 Mega bit line => 1 TB in one day
 - Have done this from Marenstrum for simulation data
 - 1 Giga bit line => 10TB in one day
 - ESAC now has gigabit – have not saturated it yet
 - Ok initially but 100TB means 10 days
 - Cost effectiveness of faster lines ?
 - Should we ship ?
 - Should we put it all in the cloud ?
 - Will decide later ..

Processing Overview (simplified)



Catalogue



+Many wrinkles !

Initial Data Treatment
 Turn CCD transits into source observations on sky
 Should be linear transform **CU3**

Astrometric Treatment
 Fix geometrical calibration
 Adjust Attitude
 Fix source positions **CU3/SOC**

Photometry Treatment
 Calibrate flux scale
 give magnitudes **CU5**

Spectral Treatment
 Calibrate and disentangle
 provide s spectra **CU6**

Solar System **CU4**

Variability **CU7**

Astrophysical Parameters **CU8**

Non Single Systems **CU4**



Databases and Security

- CUs and DPC relatively independent. Can see
 - Oracle, HBASE, MySql, Postgress, Intersystems Cache
- At ESAC Oracle since 2005
 - Not impressed with support
 - Moving some parts to Intersystems Cache
 - Excellent support and faster write times
- There are no “users” in this system.
- JDBC provides sufficient network access to data
 - thin Data Access Layer (DAL) on top
- So no need for
 - complex security schemes, Access control lists
 - Encryption etc ...

Some numbers

- We'll have ~1000,000,000 (10^9) Sources
- 100,000,000,000 (10^{11}) Observations
- 1000,000,000,000 (10^{12}) CCD transits
- For every Source we have to determine a number of parameters:
 - Position, velocity
 - Color, brightness
 - Type, age
- Final MDB (2017) may reach 1PByte

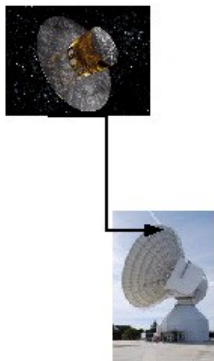
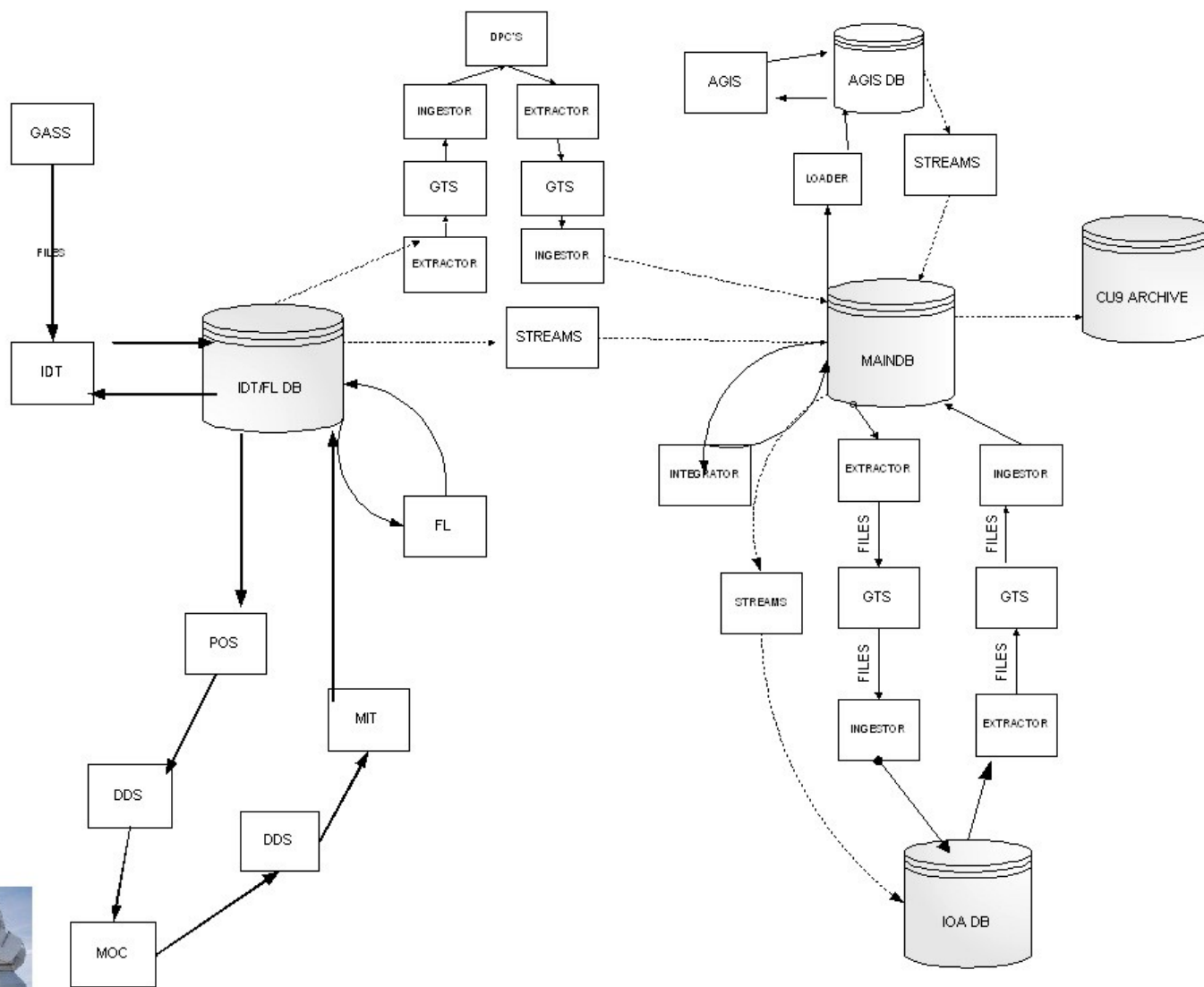
What we need

- Need to read and write large amounts of data fast enough
- Want to saturate our hardware (disks, CPUs, NW).
- We want to ingest or extract all the MDB data in one or two weeks max.
- We have different degrees of reliability:
 - IDT Dbs will have to be highly reliable
 - Others (MDB, AGIS) not so much if we can regenerate them quick enough
- store the data compressed?

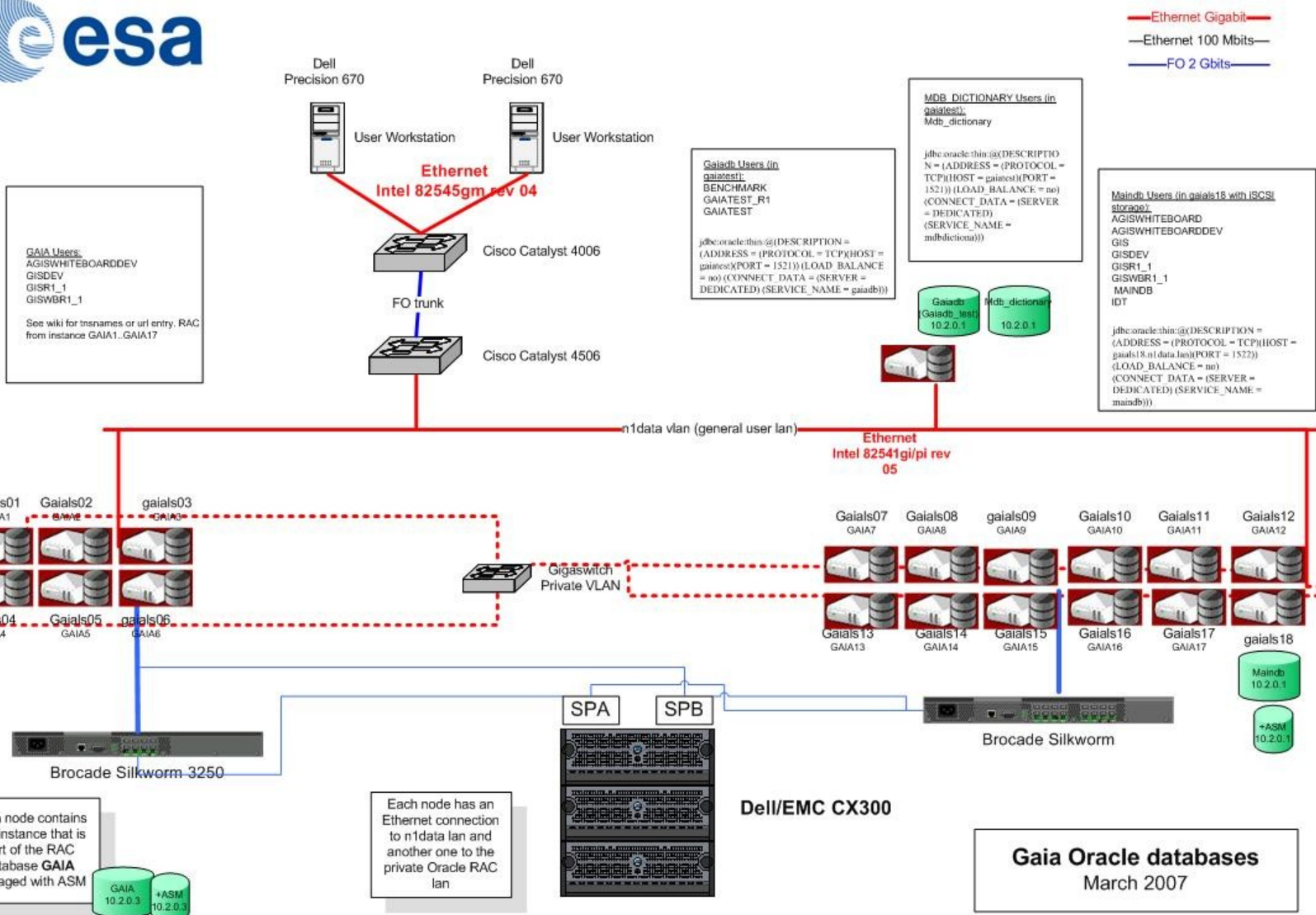
AGIS on the cloud

- Took ~20 days to get running (Parsons,Olias).
 - Used 64Bit EC2 images Large, Extra Large and High CPU Large
 - Main problem DB config (But oracle image available)
 - Oracle ASM Image based on Oracle Database 11g Release 1 Enterprise Edition - 64 Bit (Large instance) -ami-7ecb2f17
 - Also found scalability problem in our code (never had one hundred nodes before)
 - only 4 lines of code to change
- It ran at similar performance to our in house cheap cluster.
 - E2C indeed is no super computer
- AGIS image was straightforward to construct but was time consuming – better get it correct !
 - Self configuring Image based on Ubuntu 8.04 LTS Hardy Server 64-Bit (Large, Extra Large and High CPU Large Instances) - ami-e257b08b
- Availability of large number of nodes very interesting
 - not affordable in house today.
- **With 1000 nodes however we have new problems.**

DPC Dataflow



Gaia Oracle Databases 2007

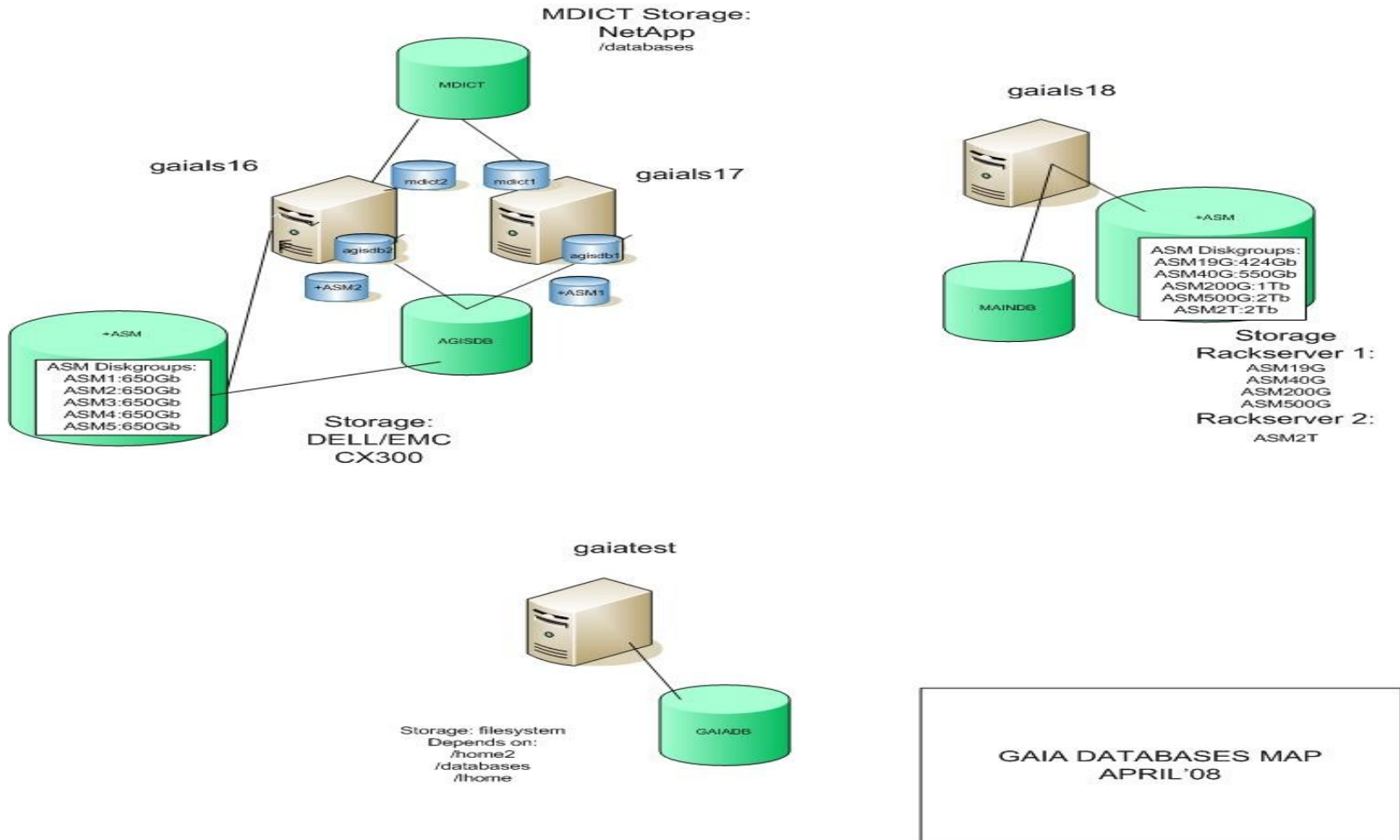


Each node contains one instance that is part of the RAC database GAIA managed with ASM

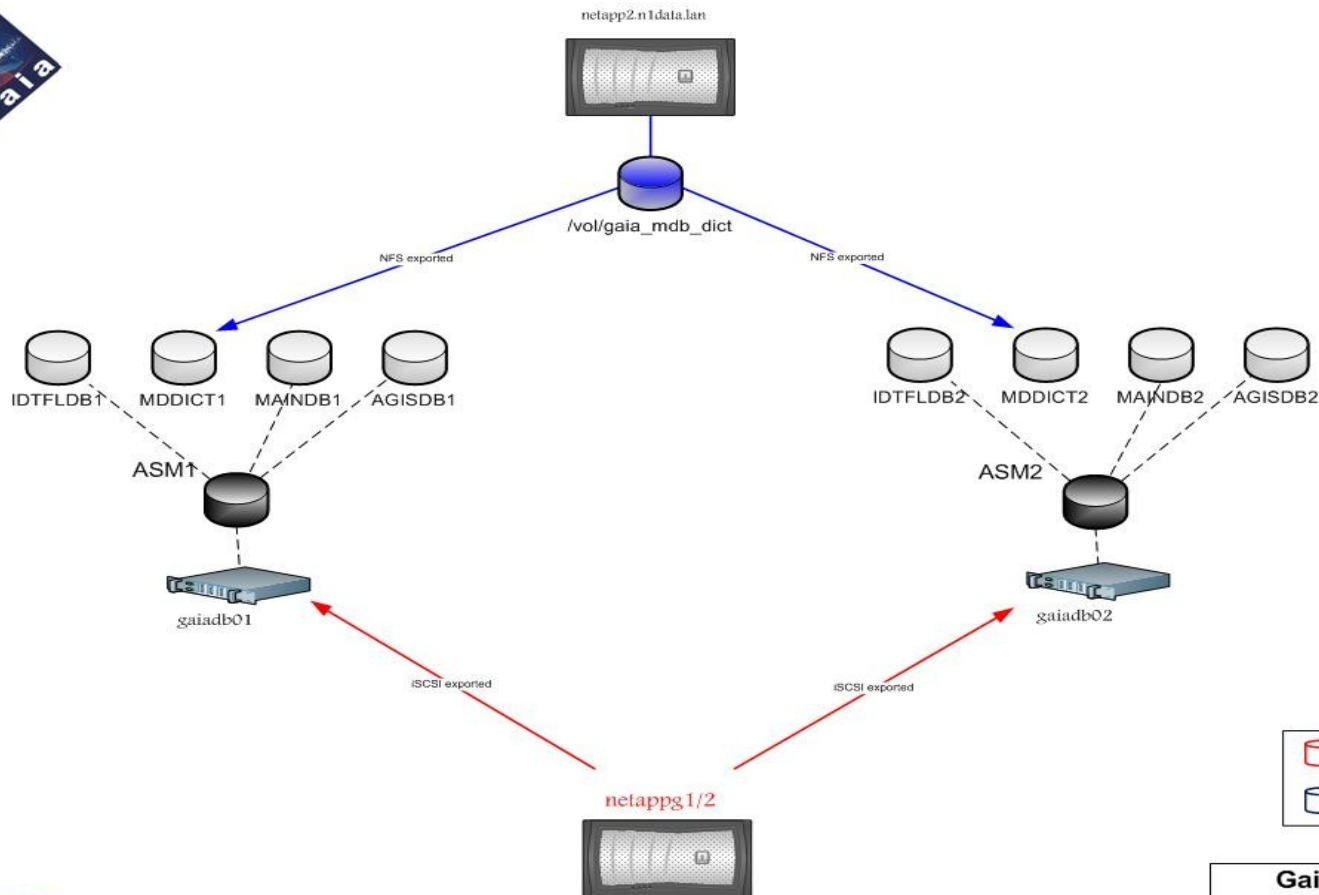
Each node has an Ethernet connection to n1data lan and another one to the private Oracle RAC lan

Gaia Oracle databases
March 2007

Gaia Oracle Databases 2008

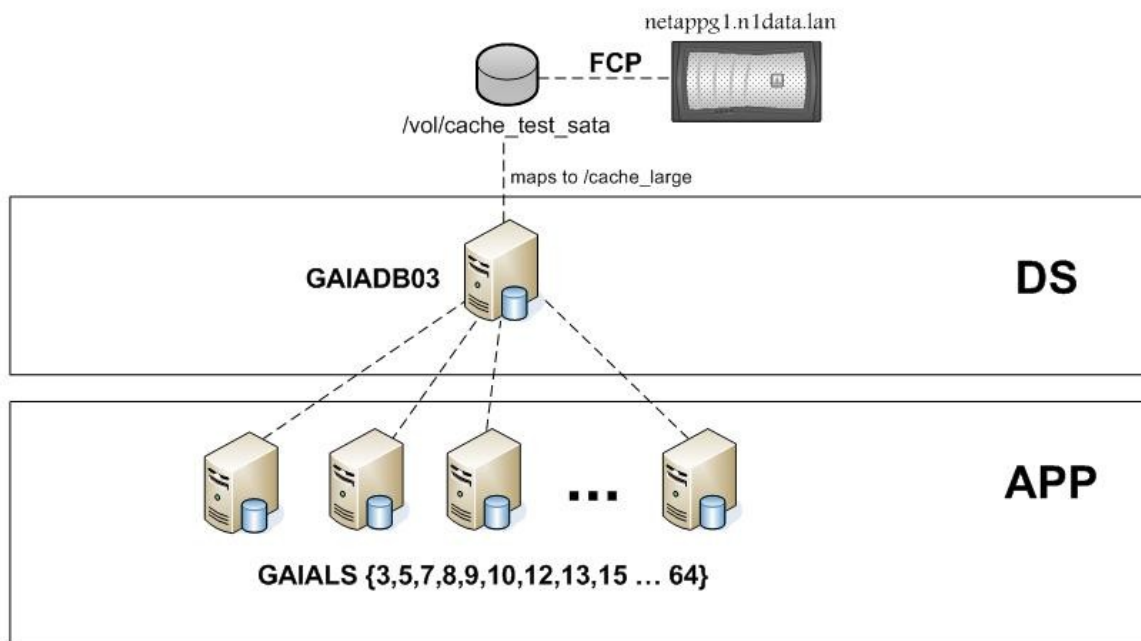


Gaia Oracle Databases 2010



Gaia Oracle environment
September 2010

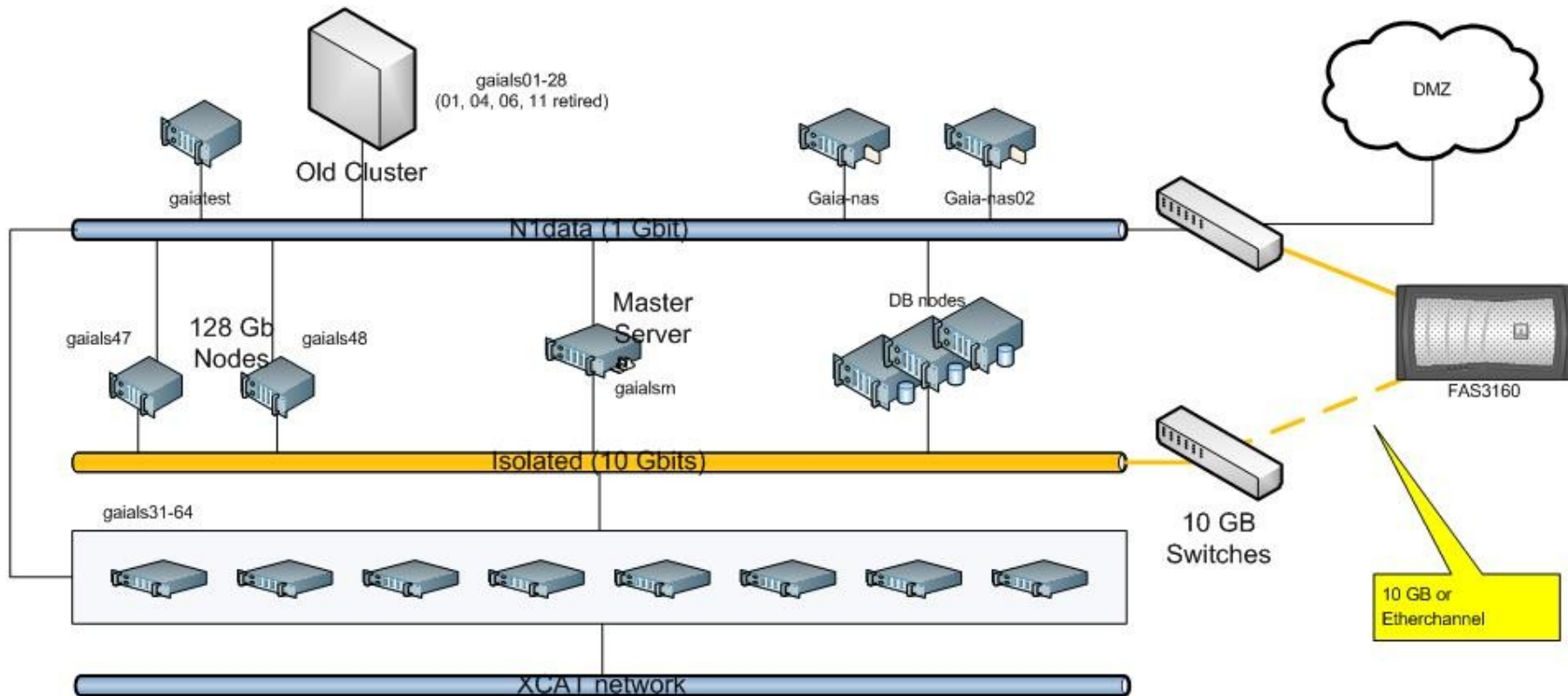
Gaia Cache Database 2010

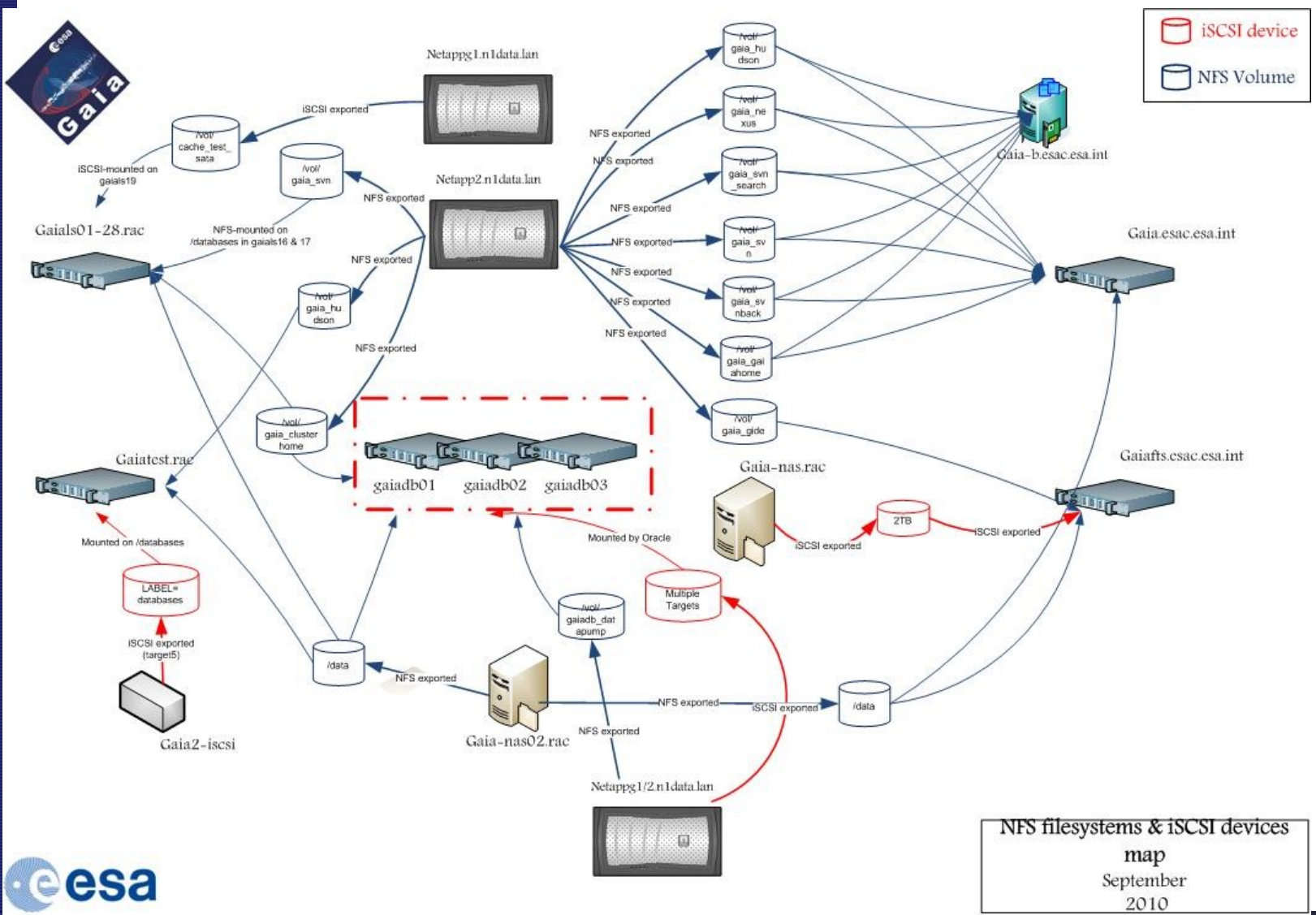


Gaia cluster architecture

Gaia new cluster network layout

September 2010





More db's

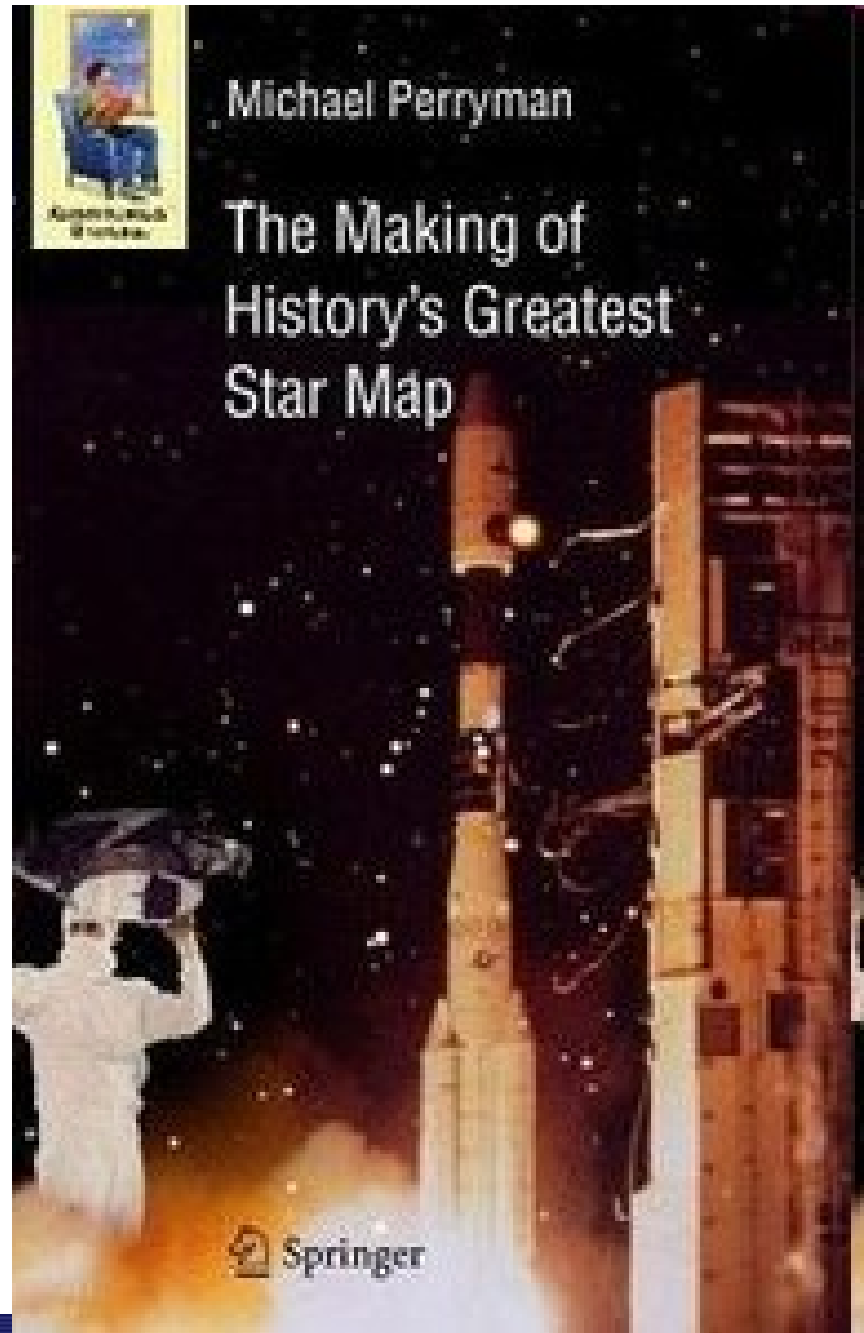
- Validation db in RAC for configuration tests:

Flashback recovery, ASM tests, Streams, Goldengate...

- MITDB for MIT
- SDB for spacecraft database

In a nutshell

- Build a complex piece of hardware
- Build LOTS of software to process the Data
- Launch hardware into space on a rocket
- Carefully get it to L2 - 1.5 Million KM
- Start Processing data
- Keep Gaia in a 300,000 Km orbit
- Keep processing the data
- Finally make that phase space map
 - That will be about **2020** !



Pilar de Teodoro



VLDB 2010

Singapore 13-17 September 2010



VLDB Conference

Premier annual international forum for data management and database researchers, vendors, practitioners, application developers, and users.

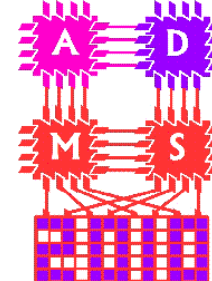
The conference features research talks, tutorials, demonstrations, and workshops.

Covers current issues in :

- **data management,**
- **database**
- **information systems research.**

Data management and databases remain among the main technological cornerstones of emerging applications of the twenty-first century.

ADMS



First International Workshop on Accelerating Data Management Systems Using Modern Processor and Storage Architectures (ADMS*10)

"NAND Flash as a New Memory Tier" by Neil Carson, CTO of Fusion-io:

Implement a cut-through design that eliminates the need for legacy storage architecture and protocols.

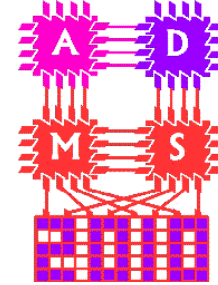
They connect the flash directly to the system bus, similar to RAM.

Partners: hp, IBM, Dell.

40x faster in BI systems.

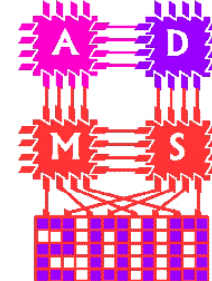


ADMS



- **Building Large Storage Based On Flash Disks.** (Technische Universität Darmstadt):
 - SSD and RAID configurations: strongly affected by current RAID controllers that are not designed for the characteristics of SSDs and rapidly become performance and scalability bottlenecks.
- **Buffered Bloom Filters on Solid State Storage.** (IBM)
 - Probabilistic data structure, efficient queries, small memory area, lots of false positives.
- **Towards SSD-ready Enterprise Platforms.** (Intel)
 - the majority of platform I/O latency still lies in the SSD and not in system software.
 - file systems may gain from organizing meta-data in a way that take into account the special characteristics of SSDs.
- **Panel Session on "The Impact of Modern Hardware on DB Design and Implementation"** moderated by Dr. C. Mohan, IBM Fellow.

ADMS



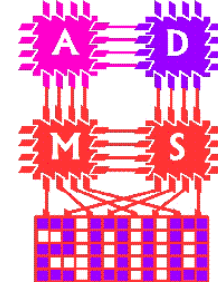
Panel Session on "The Impact of Modern Hardware on DB Design and Implementation" moderated by Dr. C. Mohan, IBM Fellow.

Panelists include

- Professor Anastasia Ailamaki (EPFL/CMU),
- Professor Gustavo Alonso (ETH Zürich),
- Sumanta Chatterjee (Oracle Corporation),
- Pilar de Teodoro (European Space Agency)



ADMS



1. Are future databases going to be mostly in-memory? If so, what are the DB design and implementation implications?

Locality optimizations, Persistency/Recovery issues, Availability of PCM-based TB main memory systems

2. Is there an ideal programming methodology for the current and future DB processing?

Map-Reduce, OpenCL based Hybrid processing, Streaming, SQL Extensions..

3. Is there an ideal processor architecture for database workloads? (Or what would you like in a DB processor)?

More cache, more cores, wider SIMD engines, specialized hardware cores?

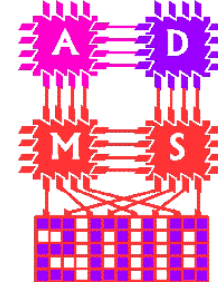
4. What are implications of the new (FLASH SSD) and upcoming (e.g., PCM) memory solutions on the DB implementations?

Buffer pool management, Cost models, Data layout, Query operators?

5. What would a future DB hardware system look like?

Hybrid accelerator/appliance based on GPU or FPGAs? Cluster of small cheap processors (e.g., ARM)? Interactions with SSD, PCM storage subsystems?

ADMS



1. Are future databases going to be mostly in-memory? If so, what are the DB design and implementation implications?

Depending on the money

2. Is there an ideal programming methodology for the current and future DB processing?

NO

3. Is there an ideal processor architecture for database workloads? (Or what would you like in a DB processor)?

Nothing ideal

4. What are implications of the new (FLASH SSD) and upcoming (e.g., PCM) memory solutions on the DB implementations?

They will be used when fully understood.

5. What would a future DB hardware system look like?

cheap



INTERESTING PAPERS AND TALKS

- HADOOP++: Making a yellow elephant run like a cheetah (Without It Even Noticing)
 - Injecting the technology at the right places through UDFs only and affecting Hadoop from inside
- HALOOP: Effective Iterative Data Processing in Large Clusters:
 - dramatically improves their efficiency by making the task scheduler loop-aware and by adding various caching mechanisms
- High-End Biological Imaging Generates Very Large 3D+ and Dynamic Datasets
- Big Data and Cloud Computing: New Wine or just New Bottles?
- Panel Session: Cloud Databases: What's New?
 - examined whether we need to develop new technologies to address demonstrably new challenges, or whether we can largely re-position existing systems and approaches.
 - covered data analysis in the cloud using Map-Reduce based systems such as Hadoop, and "No SQL" systems: hadoop. Cassandra. hvpertable,

The logo for NoSQL, featuring the letters 'N', 'S', 'Q', and 'L' in a bold, black, sans-serif font. A red five-pointed star is positioned between the 'N' and the 'S'.

Questions?

