Hardware and networks for Gaia data processing

William O'Mullane

Gaia Science Operations Development Manager

European Space Astronomy Centre (ESAC)

Madrid

Spain

http://www.rssd.esa.int/Gaia

Collaboration

- DPAC is a big collaboration !
- I speak here about activities in DPCs, the DPC managers are :
  - Mathias Beck
  - Francesca De Angeli
  - John Hoar
  - Michele Martino
  - Xavier Passot
  - Jordi Portell
- They have all contributed to this talk.

Hardware

- There is lots of hardware you all know about ..
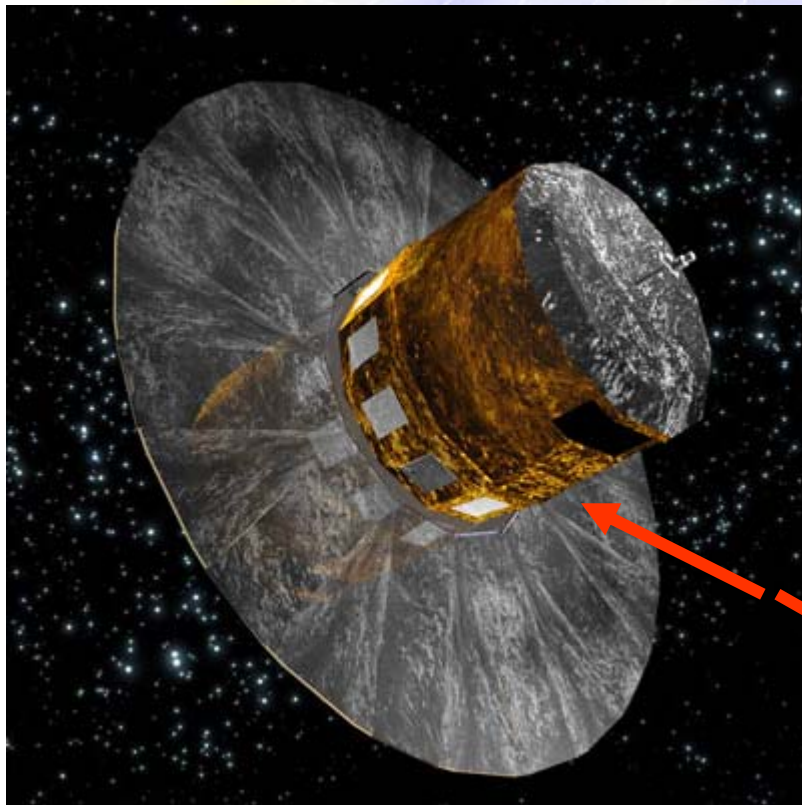- You hear plenty about the satellite….

Image Astrium

And there is probably some software you rarely think of ..

Of course we will use some dishes

You Know these ones…

Image WOM

Image ESA

Cebreros
35 Meter Diameter
40 Meters High, 650 tons
Ka-Band 31.8 – 32.3 GHz
X-Band 7.1-8.5 GHz

New Norcia
35 Meter Diameter

But did you know ….

During Launch and Early Orbit we will use 15M antennae:

Kourou '93   Perth   Maspalomas

Images ESA

The other hardware

- Of course you really wanted to know about computers…

- We have six data processing centres

- Each with its own plan for hardware purchasing

- They are all connected to some backbone ..

- Next we will speak of these…

DPAC : DPCs

DPCs underpin and support CUs

– Software support and production

– Operation of processing system(s)

- DPCB (CU2,3) Barcelona (Portell)

- DPCC (CU4,6,8) CNES Toulouse (Passot)

- DPCE (CU1,3) ESAC, Madrid (Hoar)

- DPCG (CU7) ISDC Geneva (Beck)

- DPCI (CU5) IoA Cambridge (De Angeli)

- DPCT(CU3) Torino (Martino)

Architecture

- **Highly distributed**
  - Multiple independent DPCs and CUs

- **Want/need decouple**
  - Reduce
    - dependencies
    - risk

- **Hub and spokes**
  - Max flexibility for CUs and DPCs
  - Minimum ICDs
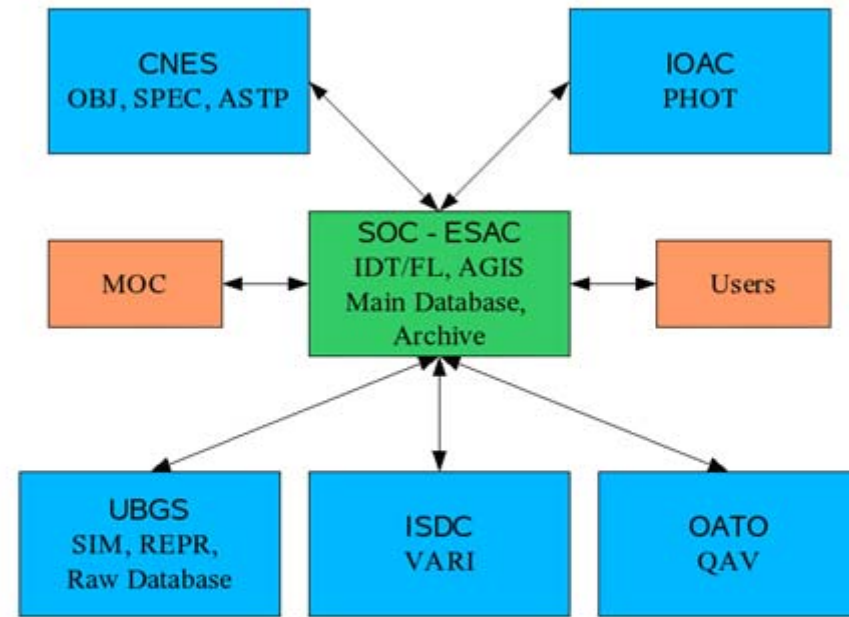
- MDB Version each Processing cycle
  - Estimate ~1000TB in final
- GTS push data to other DPCs
  - Governed by single ICD

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 Marenostrum 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 ..

Current state of interconnect

- Barcelona   1 Gbps to Anella Científica
- Cambridge   1 Gbps to academic network
- CNES   10 Gbps to ISP
- ESAC   1 Gbps to REDIRIS
- Geneva   1 Gbps
- Torino   100Kb to be Upgraded to 1Gb

- Currently testing these links with Aspera

Aspera Console - Dashboard - Mozilla Firefox

File   Edit   View   History   Bookmarks   Tools   Help

oogle Docs... | SummerPart... | Sonatype N... | dvd | GaiaAdmin: ... | Preliminary ... | new norcia ... | ESA - ESA S... | Asper...

aspera console

Logged in as admin. Logout   search

Dashboard   Activity   Nodes   Users   Groups   Transfer   Reports   Preferences   Config

Active Transfers

Name	Contact	ETA	Status
DPCI_DPCE	admin (console)	3:09pm	2%
DPCE_DPCI	admin (console)	2:00pm	90%

Scheduled Transfers

Recent Transfers

Name	Contact	Ended	Transferred
DPCE_DPCI	admin (console)	1:21pm	7.2 GB

Problem Transfers

Map

DPAC Gaia

Bandwidth

INBOUND
320.0 Mb/s
280.0 Mb/s
240.0 Mb/s
200.0 Mb/s
160.0 Mb/s
120.0 Mb/s
80.0 Mb/s
40.0 Mb/s

OUTBOUND
320.0 Mb/s
280.0 Mb/s
240.0 Mb/s
200.0 Mb/s
160.0 Mb/s
120.0 Mb/s
80.0 Mb/s
40.0 Mb/s
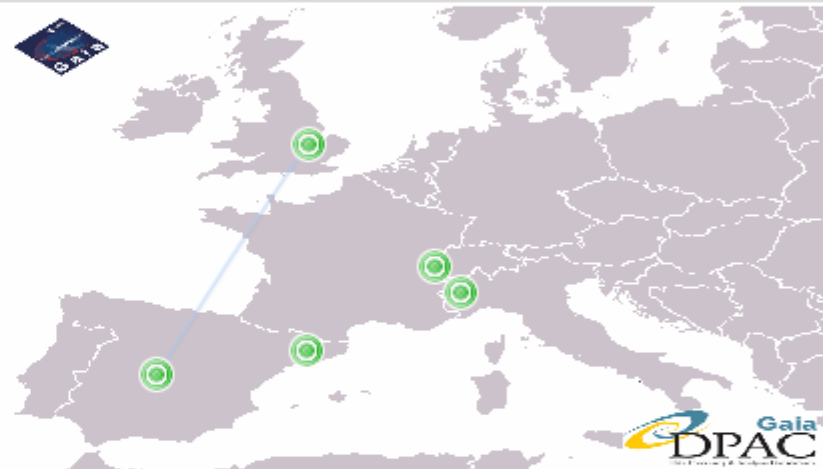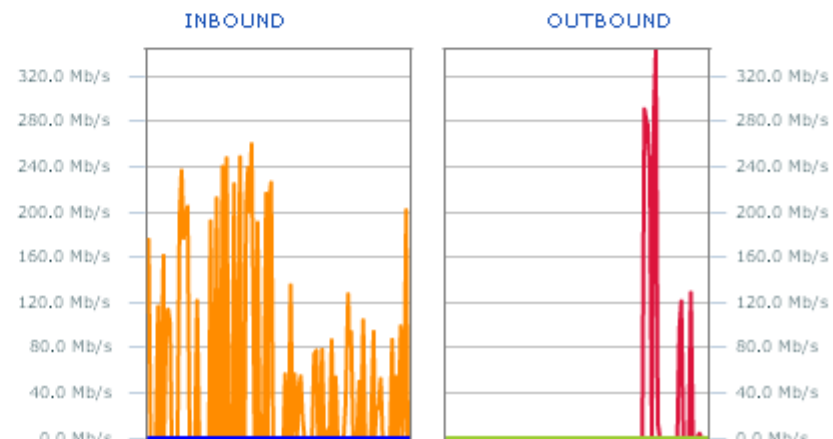
Machines

- Generally going for distributed architecture on cheap machines
  - Not super computers (one exception)
- Almost all have Intel Xeon clusters
- All are aiming to buy as late as possible
  - Some purchases will be mid mission or later
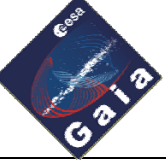
DPCB - Barcelona

- **CESCA**
  - 2 clusters: 4x33 Opteron (Cadí), 8x28 Xeon (Prades)
  - Hi-perf NFS definitely OK, old EVA disks retired
  - Expecting upgrade of Xeon cluster
- **BSC: MareNostrum**
  - Waiting for Interface Server (Aspera + NFS to central disk)
- Disks:
  - 3TB at CESCA
  - ~10TB at BSC (or more, if needed)

Most beautiful Supercomputer

(lots of) IBM blades and SUSE Linux …

Copyright 2005. Barcelona Supercomputing Center - BSC

DPCI Cambridge

- DPCI Hadoop cluster has been purchased and installed. The cluster is composed of 10 data nodes:
  - Intel Xeon, 16 cores total
  - 12 Gb memory
  - 12 disk bays currently filled with 10 1TB SATA disks 8 of which used for the Hadoop DFS.
  - All packaged in 2U and requiring ~0.5 kW of cooling
  - 1Gb network connection between the nodes

DPCC CNES

- A scalable cluster, tuned to the Gaia progressive needs
- Includes Operational and Validation configurations

CNES Cluster for Gaia	2010	2012	2014	2016
Computing nodes (minimum allocated for Gaia)	10	20	~80	~150
Number of cores (nb_nodes x proc/node x core/proc)	80	160	640	1200
Memory in Gbytes (minimum) (nb_nodes x memory/node)	10 x 24	20 x 32	~80 x 32	~150 x 48
DB servers Quadcores 48GB	2	4	4	4
DBMS system	PostgreSQL	Oracle or Caché or ?	Oracle or Caché or ?	Oracle or Caché or ?
Mass storage disks (TBytes)	6	60	~200	~300
Fast disks (Tbytes) **(300 to 500 MB/s)**	2	4	4	4
SAGA supervision servers	2	2	4	4
GIDE GTS telecom servers	2	2	2	2

DPCE ESAC

- Built up since 2005 (all red hat linux)
  - ~0.4 TFLOP (Intel Xeon, Gbit Ethernet), ~60Tb storage (NAS-type systems generally, 1 SAN for AGIS)
- Now till Launch +X (3 years < X > 6 months)
  - 100 Tb Netapp storage
  - 32-node cluster based on latest generation of Intel Nehalem EP Xeons, with two large-memory nodes dedicated to First Look
  - 10 Gbit Ethernet between nodes
  - May add more storage

What do you get if …

- Ever wonder what other projects are doing.
- Get your Google out …put NetApp Aspera
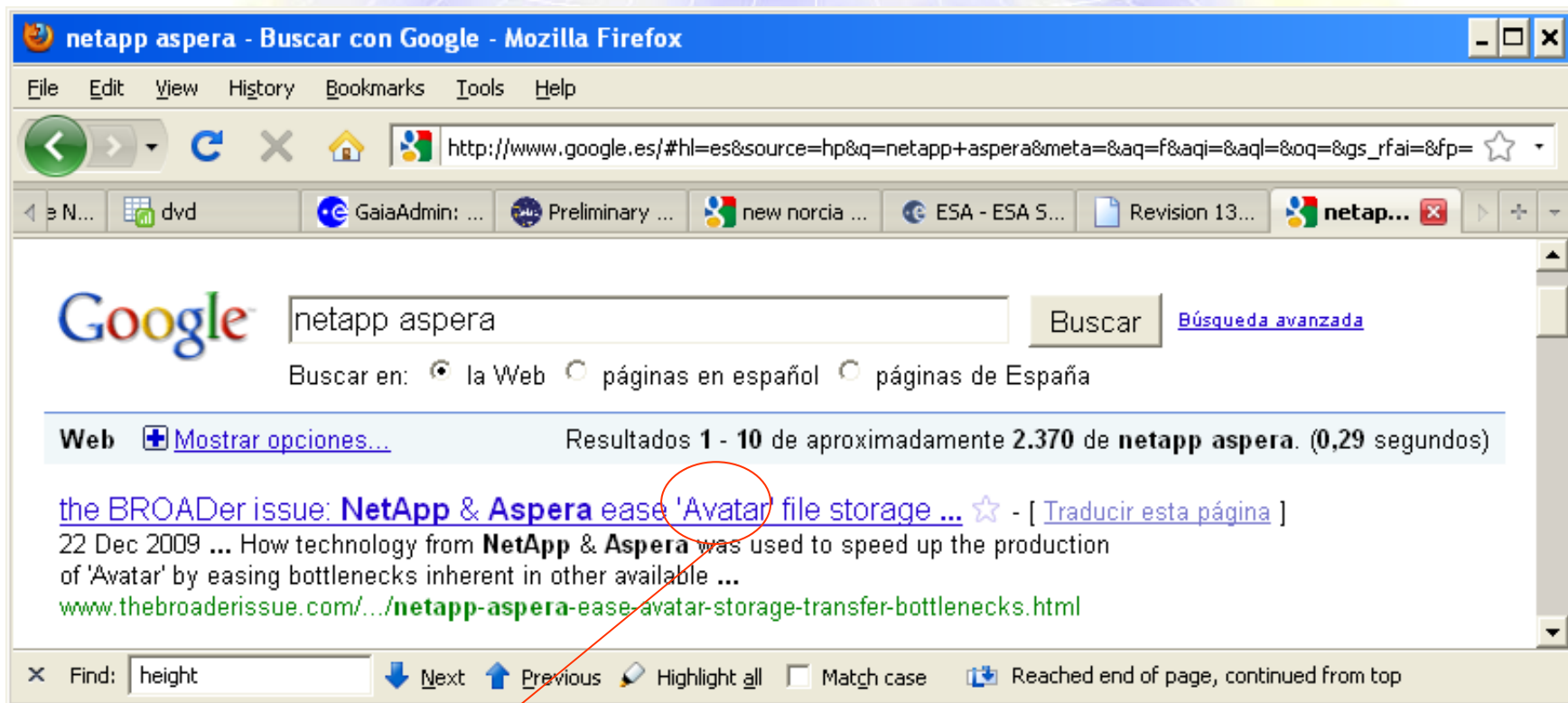
- 13 x 4 core Xeons @ 2.5 GHz,

- 8 GB RAM, 1.8 TB HDD

- 1 head node, 1 FTP node, 2 DB nodes, 8 compute nodes (flexible)

- Looking to add one more DB node
  - Possibly with Solid State Disk

DPCT - Torino

- Nodes
  - 4 HP DL580 servers with Intel Xeon Quadcore processor, 16 GB RAM, 2x72GB HDU, 2 HBA, 3 Gbit Ethernet network cards, Redundant power supply and fans, Red Hat Linux
  - 1 SGI ALTIX 450, 192 GB RAM, double processor, redundant power supply, Suse Linux

- Storage:
  - HP EVA 4400 (fully redundant power supply, controller and embedded SAN switches with 10x8Gbps FC interfaces)
  - Enclosure with 12 disks – 450 GB 15K rpm for a total of 5.4 TB raw

Databases …take your pick

- Oracle
  - ESAC, CNES (perhaps) and Cambridge
  - Neither ESAC nor Cambridge very happy ☹
- HBASE
  - ESAC some time back, Cambridge recently
  - Not mature enough
- PostgreSQL
  - Geneva and CNES
  - Some performance problems
- Intersystems Cache
  - ESAC, still testing – looks good
- MySql and Derby
  - Used mainly for testing

Key Challenge remains I/O

- disk -> memory -> core
- We have a lot of Data
- We need to look at most of it - Frequently !
- Hence constant attempts to
  - isolate algorithms from data access
  - Go for lots of distributed processing ..

- And we always have time to process vs cost trade off

New technology ..

- Both Geneva and ESAC are looking at Solid State Disks
  - These are now affordable and fast.
- HADOOP is being employed by Cambridge and investigated by Geneva.
  - Geneva add adaptation to GaiaTools Framework was relatively easy
- Amazon E2C is in use by ESAC ..
  - Others also looking at this

A final cloudy thought

- When Gaia started (2005)
  - Looked at leasing (expensive)
  - Looked at GRID (complex)
  - Decided to buy hardware cluster/cheap
- Now Amazon and others allow per hour leasing at much more favourable rates.
  - Seriously considering this for AGIS runs
    - Do it faster with more nodes
  - CU5 also considering it – they are out of space!
  - It takes 6-8 months to get a physical machine
  - I can have one on Amazon in <5 minutes..

Conclusion

- We are gearing up for Astronomy's biggest computational challenge.

- We are starting small and building up during the mission (take advantage of Moore's law)

- There are many many choices to be made

- For sure there are problems ahead

- We will solve them one way or another

Questions ??

Ariane V188 carrying Herschel and Planck (May 14 2009)