



VIRGO computing

approaches to GW computing models

general GW computing



- 3 detectors
- Online analysis happens on the measurement site
- Data is stored on site in a circular buffer (~6 month)
- All data is transferred/stored to CCIN2P3, CNAF, Ligo sites
- Offline analysis requires constraints for each workflow
 - Virgo uses its two main CCs
 - LIGO uses dedicated Condor clusters and XSEDE supercomputer
- No unified job scheduler is used over the collaboration
- Analysis code - once reviewed - is frozen

Expect Data Acquisition Rate (DAR) ~ 27 MB/s

Data storage ~ 592 TB

Pledge cpus ~ 25000 HS06

10 HSE06 = 1 core

Virgo Data facts

3^o most demanding experiment @ CNAF

Low num of workflows + cpus limited tasks

Throughput to our CCs

CNAF ~65 MB/s

Lyon ~50 MB/s

Ligo model

- Shared fs per farm
- same scheduler type
- Different farms
- Data replication
 - $t = \text{Bandwidth} * \text{szData} \ll \text{JobTime}$
- Accounting/Priority per user
- Testing other models: pilots...

processorType	n.Cpu
Opteron 275	662
Opteron 2376	328
Xeon E5-2670	96
Xeon E3-1241v3	200
Tesla K10	64
Opteron 2376	504
Xeon E3-1241v3	100
Opteron 2376	504
Xeon E3-1241v3	100
Xeon X3220	1,176
Xeon E3-1220v3	1,820
...	...
Total	8747



Data Transfer

- Online data
 - Only inside EGO
 - Two fluxes, only one used
- RDS
 - ~ 1/10 of Bulk Data
 - Sent to CNAF, IN2P3, Ligo & received from Ligo
- Bulk data
 - Sent to CNAF, CCIN2P3
- Throughput problem (Equilibrium between recovery and transmission)
 - $Bw \geq 2 \cdot DAR$
 - CCIN2P3



Analysis

- Online
 - On site
 - Low-latency (few seconds)
 - Used to alert astrophysical partners
 - 13 Xeon E5-2650
- InTime
 - Few hours latency
 - On CCs
- Offline
 - During run/End of run
 - On CCs (mainly CNAF)
 - Burst & Continuous Waves

CNAF model

- Grid like environment
- Used by Virgo workflows
- Porting workflow not easy
 - on-going activity
- Different solutions are possible
 - Emulating Ligo Farm (ahistorical)
 - Writing new tools (Better solution)
- Cost
 - CNAF pledge
 - CCIN2P3 pay for real use

Hostname	HSE06	Core	Slot
wn-200-01	172	16	16
wn-200-03	172	16	16
wn-200-04	172	16	16
wn-200-08	172	16	16
wn-201-01	172	16	16
wn-201-03	172	16	16
wn-201-04	172	16	16
wn-201-07	172	16	16
wn-201-08	172	16	16
wn-205-01	182	12	16
wn-205-03	182	12	16
wn-206-01	182	12	16
wn-206-03	206	24	24
wn-206-04	206	24	24
wn-206-07	206	24	24
wn-206-08	206	24	24
wn-200-11	354	16	24
wn-200-10	354	16	32
wn-200-13	354	16	32
N Host 20	3980	324	376



Pilots

- Farm → Grid
- Well known objects
- Already tested at CNAF
- Basic job retrieve real task from a condor pool
- Executables are exported via cvmfs
- DataFile downloaded everytime
- If jobTime >> retrievalTime
- Required centralized resources & data storage
- Partial workflow adapting



Future (and GPUs)

- Philosophically is a distributed system
 - Cpu, gpu and dsp
- Cpus need!
- Exploring gpus (CCIN2P3)

	2015	2016	2017	2018
Disk(TB)	445	592	656	720
CPU(HS06)	10000	25000	36000	78000