Cloud computing for AGATA data analysis

Damian Ralet*

Centre de Sciences Nucléaires et de Sciences de la Matière

*Supported by the P2IO excellence center

For a GSI experiment with around 20 crystals: in average,
10 To of data, including calibrations

For a GSI experiment with around 20 crystals: in average, 10 To of data, including calibrations

1 To of useful data: 12 h for a PSA replay with a 8 core computer (multi-threading)

For a GSI experiment with around 20 crystals: in average, 10 To of data, including calibrations

1 To of useful data: 12 h for a PSA replay with a 8 core computer (multi-threading)

GANIL experiment with about 24 crystals: 28 To of data for one experiment in July 2015

For a GSI experiment with around 20 crystals: in average, 10 To of data, including calibrations

1 To of useful data: 12 h for a PSA replay with a 8 core computer (multi-threading)

GANIL experiment with about 24 crystals: 28 To of data for one experiment in July 2015

20 To of useful data $\rightarrow$ 10 days on an 8 core computer

Storage of the 28 To of data in an institute

Storage of the 28 To of data in an institute

Data replay in an acceptable time

Storage of the 28 To of data in an institute

Data replay in an acceptable time

Solution: increasing the number of core

Utilisation of the grid:

Utilisation of the grid:

Advantage: data are already stored on the grid

Utilisation of the grid:

Advantage: data are already stored on the grid

But not an easy solution:

- grid access point
- modification of the replay code (femul) to launch a process per crystal
- femul and NARVAL are not ready for this
- verifications: need to transfer the result on a local computer

Utilisation of "Cloud" computing: project started by some laboratories on the Orsay campus (CSNSM, LAL, ...)

Utilisation of "Cloud" computing: project started by some laboratories on the Orsay campus (CSNSM, LAL, ...)

An easy solution:

- Can be used like a standard computer
- Possibilities to have big computers (40 cores, 90 Go of RAM)
- Utilisation of femul with multi-threading: easy
- Disk space available (?)

- Replay example

Developed for data analysis of AGATA@GANIL data
(and GSI):

Creation of a virtual machine

Developed for data analysis of AGATA@GANIL data
(and GSI):

- Ubuntu LTS 14 operating system
- AgataSofware libraries
- GammaWare
- VAMOS libraries for the experiment e672
- root 5 and 6, compiled with python option
- Geant4 (version?)
- Some monitoring tools

Developed for data analysis of AGATA@GANIL data (and GSI):

- Ubuntu LTS 14 operating system
- AgataSofware libraries
- GammaWare
- VAMOS libraries for the experiment e672
- root 5 and 6, compiled with python option
- Geant4 (version?)
- Some monitoring tools

An computer science student generalised the virtual machine

Cluster of virtual computers

Cluster of 4 virtual machines
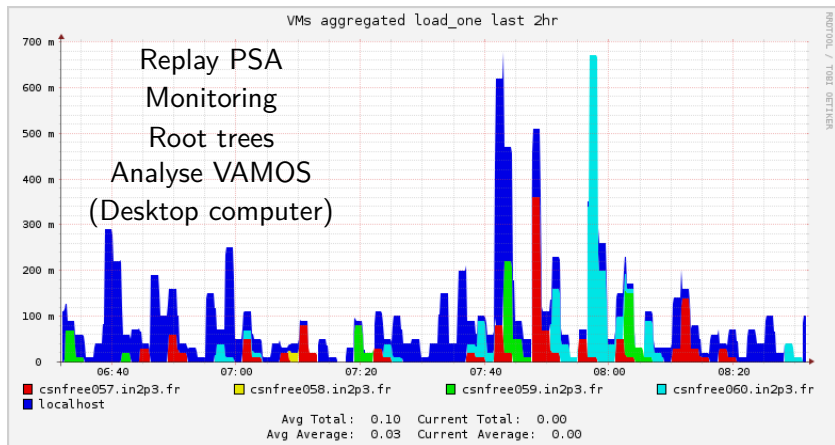
One computer for PSA replay

PID	USER	PRI	NI	VIRT	RES	SHR	S	CPU%	MEM%	TIME+	Command
14154	ralet	20	0	14.4G	12.5G	16632	S	3925	14.1	1h04:31	femul Topology.conf
14187	ralet	20	0	14.4G	12.5G	16632	R	92.2	14.1	1:05.09	femul Topology.conf
14175	ralet	20	0	14.4G	12.5G	16632	R	90.3	14.1	1:04.46	femul Topology.conf
14204	ralet	20	0	14.4G	12.5G	16632	R	84.4	14.1	1:05.84	femul Topology.conf
14176	ralet	20	0	14.4G	12.5G	16632	S	82.5	14.1	1:05.12	femul Topology.conf
14205	ralet	20	0	14.4G	12.5G	16632	R	81.8	14.1	1:06.03	femul Topology.conf
14201	ralet	20	0	14.4G	12.5G	16632	R	76.0	14.1	0:56.35	femul Topology.conf
14158	ralet	20	0	14.4G	12.5G	16632	R	75.3	14.1	1:02.73	femul Topology.conf
14286	ralet	20	0	14.4G	12.5G	16632	R	75.3	14.1	0:54.50	femul Topology.conf
14174	ralet	20	0	14.4G	12.5G	16632	R	70.8	14.1	1:06.82	femul Topology.conf
14228	ralet	20	0	14.4G	12.5G	16632	R	70.1	14.1	0:59.17	femul Topology.conf
14200	ralet	20	0	14.4G	12.5G	16632	R	69.5	14.1	1:04.98	femul Topology.conf
14313	ralet	20	0	14.4G	12.5G	16632	R	69.5	14.1	0:52.27	femul Topology.conf
14177	ralet	20	0	14.4G	12.5G	16632	R	68.8	14.1	0:58.39	femul Topology.conf
14188	ralet	20	0	14.4G	12.5G	16632	S	68.2	14.1	1:04.90	femul Topology.conf

One computer for PSA replay

Computer load

Computer load
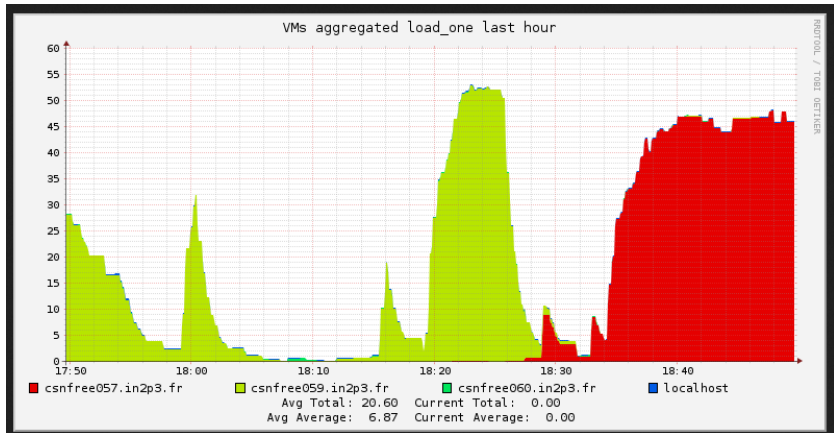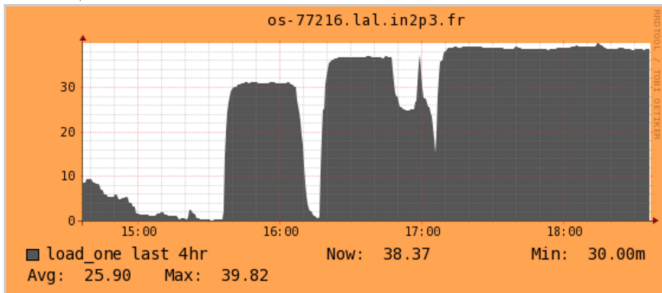
Cloud computing is handy to use

Cloud computing is handy to use

Hard part (as an user): install analysis software on a
virtual machine

Cloud computing is handy to use

Hard part (as an user): install analysis software on a virtual machine

With a proper infrastructure: real plus compared to grid computing