

CPPM

Marlon Barbero, Arnaud Duperrin

ATLAS France [CAF users meeting](#)
December 9th 2021

Team

- **Composition of the team**
 - **14 physicists** / 10.4 FTE → 9.4 CNRS + 1 University
 - **27 ITAs** / 18.5 FTE → Soft/Computing: 1.2 (E. Knoop/D. Laugier/M. Biaut) – Calo 7.3 – Pixel 7.3 – FC 2.7
 - **8 PhDs** / 2 University + 2 cotutelles + 2 CNRS/region + 1 CERN + 1 AMU Amidex) [3 defense and 3 new PhD in 2021]
 - **2 PostDocs** (Tarna/Sur)
- **Total OTP under class 3 = 8 FTE** (from glance x2)
 - **Computing/Software** = 1.6 (Djama, Duperrin, Rozanov, Strebler, J. Toth)
 - **Analysis support** = 0.44 (Duperrin, Strebler)
 - **Data preparation** = 0.44 (Djama, Toth)
 - **Detector operation** = 5 (Aad, Bernon, Fortin, Hallewell, Knoop, Laugier, Monnier, Rozanov, Strebler, Tisserant, Vallier)
 - **Trigger** = 0.54 (Feligioni, Nagy)
- **Computing**
 - **Local T2 manager:** 1 ITA (0.5 FTE - Edith) + 1 PH (0.05 FTE)
 - + ITA from computing service involved (Carlos Carranza + Thierry Mouthuy)
 - + collaboration with LHCb (Andrei Tsaregorodtsev)

Involvements

- OTP from “software+analysis support” categories for class 1-2-3-4:

- **Evolution:** 2021 ([2.68 FTE](#)), 2020 ([3.86 FTE](#)), 2019 ([2.96 FTE](#)), 2018 ([3.93 FTE](#)), 2017 ([4.56 FTE](#)), 2016 ([5.71 FTE](#))

- Involvements

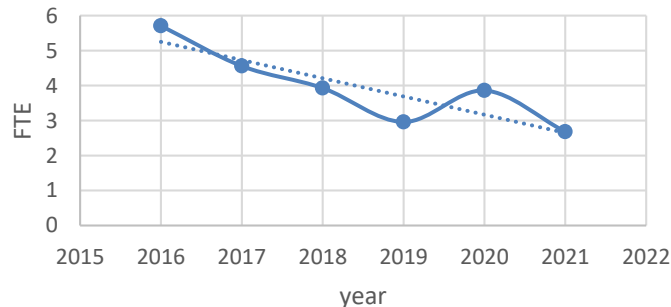
- **Analyses:**

- ttH multi-leptons (ttW), ttHbb l+jets
- coupled with BSM searches in similar final states (ttWML, $H^{++} \rightarrow W^+W^+$, EWK SUSY, RPV SUSY),
- $HH \rightarrow bb\gamma\gamma$ on Run 2 with BSM $S(bb)H(\gamma\gamma)$ searches + preparation for $Z(\rightarrow bb)H(\rightarrow \gamma\gamma)$ for Run 3
- Upgrade Physics
 - Convener of the Upgrade Physics group (Thomas Strebler)

- **Performance:** trigger+ tracking + electron-ID + b-tagging (Run 3 + Run 4)
 - Convener of FTAG reconstruction/software/validation subgroup (A.D.)

- **Detector and R&D:** ITk + LAr (Phase 1 + Phase 2)

S&C FTE evolution



CPPM Software involvement

LAr:

First Name	Last Name	Activity	System	Task	Alloc
Georges	Aad	Detector Operation	LAr	LAr online software	0.04
Florent	Bernon	Detector Operation	LAr	LAr online software	0.24
Mathieu	Biaut	Detector Operation	LAr	Phase-II LAr	0.50
Thomas	Calvet	Detector Operation	LAr	Phase-II LAr	0.70
Etienne	Fortin	Detector Operation	LAr	LAr online software	0.60
Etienne	Fortin	Detector Operation	LAr	Phase-II LAr	0.20
Edith	Knoops	Detector Operation	LAr	LAr online software	0.10
Edith	Knoops	Detector Operation	LAr	LCS (Linux & Configuration Support) experts	0.10
Daniele	Laugier	Detector Operation	LAr	LAr online software	0.30
Daniele	Laugier	Detector Operation	LAr	LCS experts	0.30
Daniele	Laugier	Detector Operation	LAr	Phase-II LAr	0.10
Jozsef	Toth	Data Preparation	LAr	Reconstruc,EDM,Validation	0.14

Pixel:

First Name	Last Name	Activity	System	Task	Alloc
Fares	Djama	Computing/Software	PIXEL	Software Development	0.10
Alexandre	Rozanov	Detector Operation	PIXEL	Software Development	0.10

CPPM Software involvement

Trigger:

First Name	Last Name	Activity	System	Task	Alloc
Lorenzo	Feligioni	Trigger	General Tasks	Bjet Software and Perf	0.24
Elemer	Nagy	Trigger	General Tasks	Bjet Software and Perf	0.30

Tracking

First Name	Last Name	Activity	System	Task	Alloc
Thomas	Strebler	Computing/Software	ID gen	Common Tracking Software Support	0.60
Thomas	Strebler	Computing/Software	Upgrade	ITk - ITk Offline Software	0.10
Thomas	Strebler	Computing/Software	Upgrade	ITK - Performance Studies	0.24

b-tagging:

First Name	Last Name	Activity	System	Task	Alloc
Arnaud	Duperrin	Analysis Support	General Tasks	Internal Software	0.22
Arnaud	Duperrin	Computing/Software	General Tasks	Reconstruction	0.24
Arnaud	Duperrin	Analysis Support	General Tasks	Run 4	0.10
Thomas	Strebler	Analysis Support	General Tasks	Run 4	0.10

Other tasks:

First Name	Last Name	Activity	System	Task	Alloc
Arnaud	Duperrin	Computing/Software	General Tasks	Validation of software release	0.08
Thomas	Strebler	Detector Operation	Upgrade	Convener Upgrade Physics Group	0.06
Thomas	Strebler	Detector Operation	General Tasks	Phase-II HGTD	0.24

CPPM Computing involvement

Class 3

CAF

First Name	Last Name	Activity	System	Task	Alloc
Arnaud	Duperrin	Computing/Software	General Tasks	Cloud Operation & Management	0.05

Class 4

Computing: T2 ATLAS

First Name	Last Name	Activity	System	Task	Alloc
Arnaud	Duperrin	Computing/Software	General Tasks	FR CPPM, Marseille	0.05
Edith	Knoops	Computing/Software	General Tasks	FR CPPM, Marseille	0.4
Institut	CPPM	Computing/Software	General Tasks	FR CPPM, Marseille	0.2

Computing resources in 2021-2022

- **Grid resources** (Tier 2-ATLAS, pledged in 2021)
 - **Storage** = 2 200 TB (2200 TB pledged in 2021 → +0% for 2022)
 - **Computing** = 24 000 HS06 (24k HS06 pledged in 2021 → +0% for 2022)
 - CPER/FederM3AMU University ~ 900 k€ (M3AMU) ended last year.
 - Due to lack of funding visibility, T2 pledged resources were not increased for 2022.
- Other “grid” resources (corresponding to this year non pledged resources)
 - **Storage** = 245 TB in 2021 on **LOCALGROUPDISK** (+0%)
 - **Computing** = 6k HS06 in 2021 (+0%)
- **Other local** (lab) resources (i.e. whatever is non grid)
 - 4 local-lab servers for ATLAS-CPPM activities (no new servers bought)
 - Storage: 192 TB
 - Computing: 170 cores @ 2.2-2.7 GHz
 - CPPM has GPU for developments/tests (not prod) shared among groups (used by Atlas for NN LAr studies)
 - Cloud computing model at CPPM under development to mutualize computing resources among groups

Analysis and needs

(see backup for details)

- Detector studies (local/CERN)
 - **Pixels & LAr studies**: resource level and usage similar to last year
- Performance studies (local/GRID)
 - **trigger+ tracking + electron-ID + b-tagging**: resource level and usage similar to last year
- Analyses studies (local/CERN/cc-in2p3/GRID)
 - **ttH+SUSY+HH**: resource level and usage similar to last year (mostly local/GRID)

Analysis and needs

(see backup for details)

- Machine learning specific needs
 - **At analysis level**, no specific needs (i.e. can be done on standard servers)
 - Ex: ttHbb (Neelam Kumari): investigate Deep Sets models, aiming at combining Higgs kinematic reconstruction in multi-jet events and STXS signal+background multi-classification in a single tool.
 - **Detector and Performance studies** (Run 4)
 - **LAr**: see for instance NN pour FPGAs [paper](#)
 - Firmware takes lot of memory (30 GB) and time (3h) to compile, need dedicated machines with large memory and fast CPUs (rather than many CPUs, multi threading not very adapted for this)
 - NN training is resource consuming (1-3 days to train). Using local GPUs and also GPU cluster at CERN (condor). More powerful GPUs here would help.
 - **b-tagging** (Maxime Fernoux):
 - development of the Flavour Tagging tagger using Deep Sets (DIPS) models and study its performance in the context of the ITk upgrade simulation for the HL-LHC.
 - set up and test the training procedures + optimization of the sets of inputs to DIPS

Near future

→ please address foreseen evolutions in near future (1-2 years or LS) which could affect software and computing (mostly if different from what was shown last year)

• Recent evolution of the team and activities

• team:

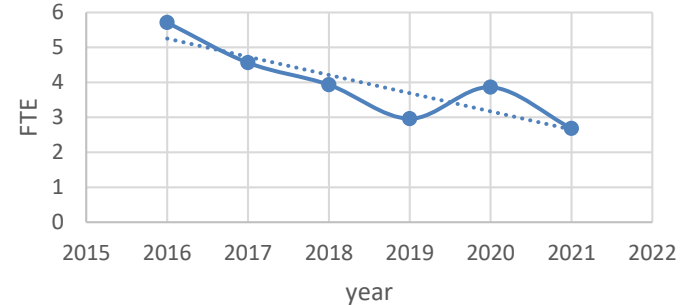
- -3 physicists / + 7 ITA (versus 2019 numbers)
- Smaller number of **PhD** students (-6 vs 2019)

• **software** involvement steadily **decreasing**

• Analyses/Performance

- Topics: **susy** + run 2 legacy paper for **ttH** (next year is last year involvement) + **HH** analysis for Run 3&4 (increasing interest in the group, but small group at the moment)
- All based on Run 2 dataset and release 21 frameworks
- Run 3 involvement (release 22) is currently limited to **performance work** (b-tagging)
- Run 4 involvement in performances will continue: tracking+b-tagging & Calo/Pixel
- Involved in **FR-ALPAMED storage federation** for DOMA/ESCAPE R&D (not covered in this presentation)

S&C FTE evolution



Near future

→ please address foreseen evolutions in near future (1-2 years or LS) which could affect software and computing (mostly if different from what was shown last year)

- Computing resources and needs

- **T2 ATLAS CPPM:**

- **Increase by ~20% every year** our CPU and disk, +LOCALGROUPDISK **is not possible** a.t.m
 - **Lab direction supports the development** of grid infrastructures for the coming new protocol
 - CPER is over, watching closely possibility of new funding in close collaboration with **University mésocentre** infrastructure

- Evolution of **local resources**:

- Moving to cloud model at CPPM
 - Problematic external low network connection (limited to 10 Gb/s... often saturating, a real limitation)
 - Several retirements from computing services are coming soon

BACKUP

CPPM Analyses: ttH

- Analysis: ttHbb leptons+jets (→ also involved in run 2 legacy paper)
 - **S&C contributions:** L1 common ntuple production, analysis
 - **Inputs:** DAOD TOPQ1 full Run 2
 - **Size:** prod ~3.2 TB by CPPM for a total shared of 25 TB for L1, 14 TB for L2
 - **Time:** ~1 week for bulk of samples for L1, few days for L2
 - **Where:** L1 on the grid, L2 on local servers
 - **Comments:** L1 prod very automated
- Analysis: ttHML+ttW (+derived H^{++} for Run 2) (→ involved in run 2 legacy paper)
 - **S&C contributions:** GFW1, GFW2 production, analysis
 - **Model:** DAOD HIGG8 full Run 2
 - **Size:** for full Run2 ~8 TB per round of analysis for L1, x TB for L2,
 - **Time to process:** 3 days for reduction to GFW1, 1 for GFW2
 - **Where:** reduction on grid, EOS+local storage, analysis: lxplus, local, (CC)
 - **Comments:** use of 1 TB at CC

CPPM Analyses: RPV SUSY and HH

- Analysis: SUSY RPV

- **S&C contributions:** production and analysis
- **Inputs:** DAOD TOPQ4 full Run 2
- **Size:** 13 TB for L1, 20 TB for L2
- **Time:** 1 week for L1, 2 weeks for L2
- **Where:** on the grid

- Analysis: HH \rightarrow $bb\gamma\gamma$

- **S&C contributions:** analysis
- **Model:** DAOD HIGG1D1 + mini-xAOD produced by the $H\rightarrow\gamma\gamma$ group
- **Size:** for full Run2 \sim 5 TB per round of analysis
- **Time to process:** 1 month to produced DAOD+mini-xAOD, hours for analysis
- **Where:** mini-xAOD stored on eos, some on local + CC

S&C Performance studies

- Tracking for Run 4

- **S&C contributions:** development of Athena Reco for various ITk layout
- **Model:** starting from EVNT, ESD, AOD, ...
- **Size:** ~2T
- **Time to process:** various (few seconds to 24h)
- **Where:** grid + local
- **Comments:** grid submission limited by memory limits for non-production jobs

S&C Performance studies

- **b-tagging**
 - **Code development and maintenance:** Athena (Reconstruction+Derivation) for Run 2 reprocessing and Run 3 (release 22)
- **b-tagging ITk**
 - **S&C contributions:** maintain+develop b-tagging soft infrastructure for phase 2 + production of ntuple for CP performance + b-tagging algo optimization (ex: track categories for IP3D)
 - **Model:** AOD → ntuple, then histos + from EVNT w/ low stat
 - **Size:** 9 TB for step 3.1 ttbar → 600 GB local cppm ntuple, 5 GB for histos
 - **Time to process:** 2 days for the complete chain
 - **Where:** on the grid, storage on T2 local group disk, analysis on local
 - **Comments:** all process has to be repeated several times per year
- **b-jet trigger calibration (→ not be continued at the moment)**
 - **S&C contributions:** calibration based on dilepton tt analysis
 - **Model:** FTAG2-DAOD → ntuple
 - **Size:** 26 TB per round of analysis
 - **Time to process:**
 - **Where:** reduction on the grid, storage at Desy, analysis at Desy+local

S&C Performance studies

- Electron-ID Run 2 SF

- **S&C contributions:** selection, analysis
- **Model:** DAOD→ntuple
- **Size:** 1 TB per round of analysis
- **Time to process:** 3 days per round
- **Where:** grid prod, storage+analysis on local

- Phase 2 DAQ Lar calorimeter

- **S&C contributions:** developing the main backend board with artificial intelligence algorithms (NN) on FPGAs
- **Model:** Firmware takes lot of memory (30 GB) and time (3h) to compile, need fast CPU rather many CPU
- **Where:** local dedicated machine at CPPM + will use local/CERN GPU platform
- **Comments:** project supported by University funds (A*MIDEX+ANR)

- Pixel

- Software Development/Maintenance and Physics Performance, including CMOS simu.
- On lxplus with eos storage.