The HEP Software Foundation

Benedikt Hegner (CERN)

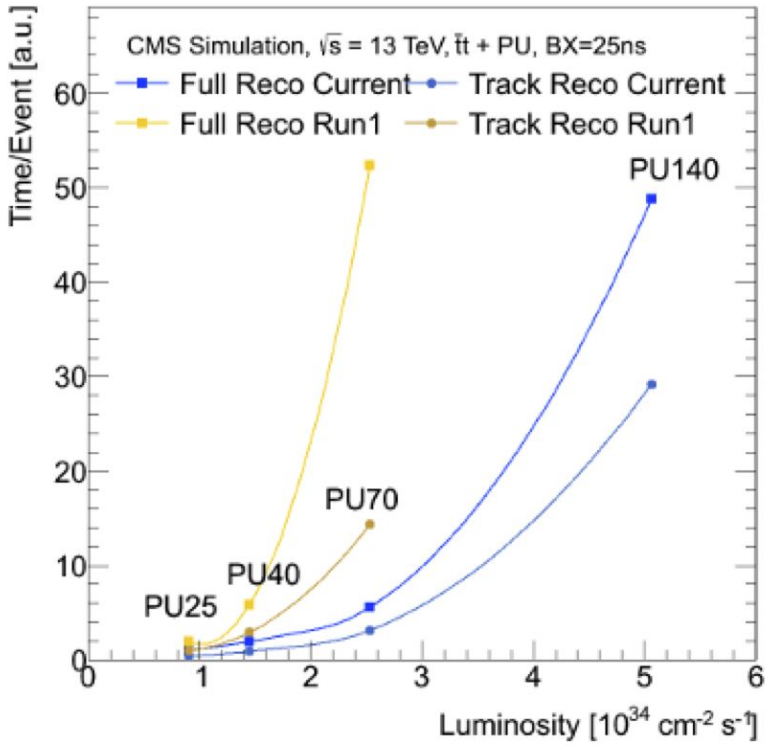
GridKA School 2016, 31.8.2016

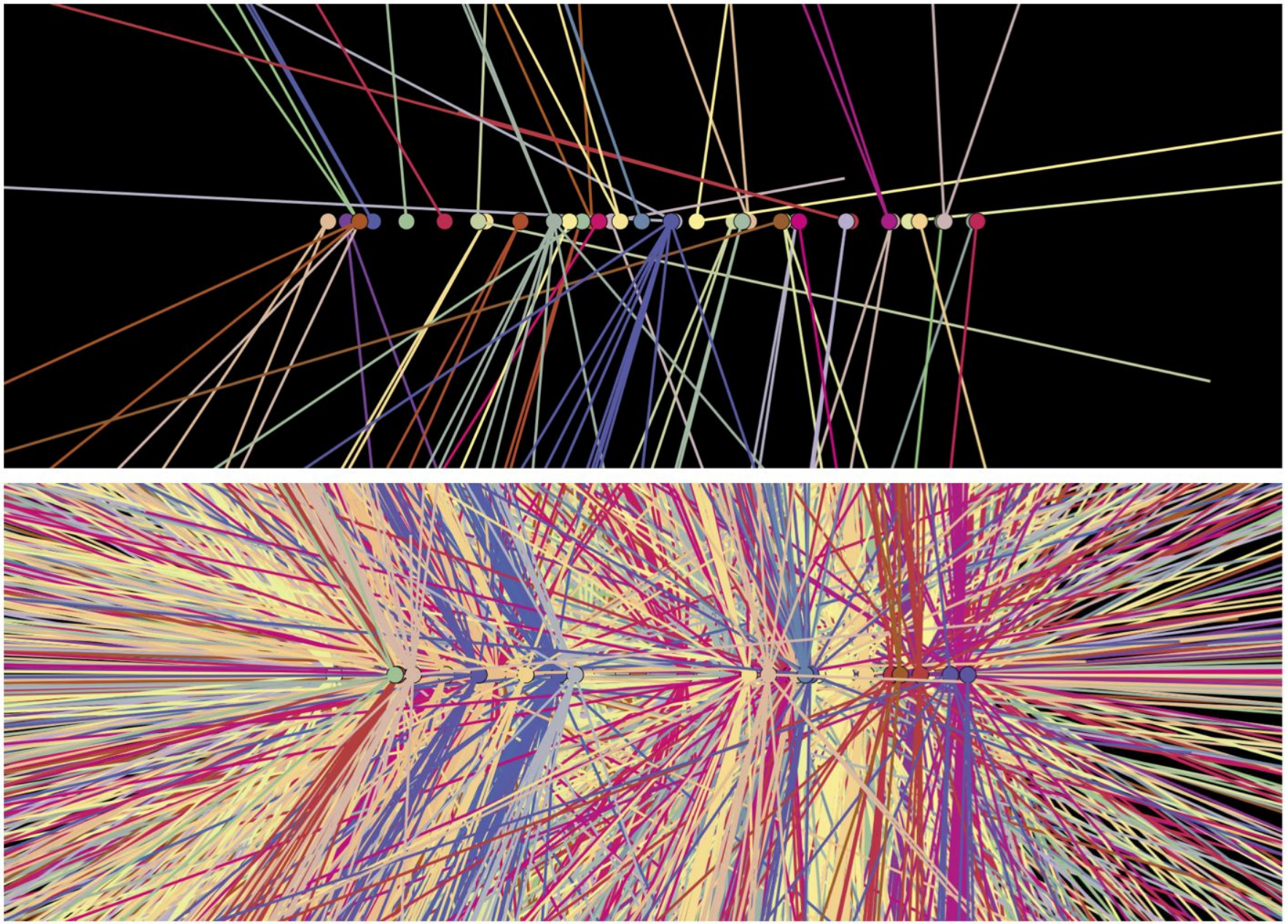
hepsoftwarefoundation.org

- Challenges of HEP Software
- Vision of the HSF
- Some HSF activities
- Hopes for the Future
- Discussion

- **Increasing complexity and multiplicity of physics events** a challenge for simulation and reconstruction performance

- The **free lunch is over** (no news!) and we have to actively address the problems of SW performance





- Community has to **re-learn many things**
  (evolution of CPUs, new architectures, HPC)
- **HEP community too small** to have every experiment
  evaluate solutions by themselves
- Make sure **experts have a career perspective**
- Have to seek **commonality and collaboration** in new developments

⇒ Improve both in SW efficiency and usage of people's time!

- This is **not the first time** HEP is trying better cooperation on software
  - An entire sub-project of the WLCG was dedicated to common software (LCG Applications Area)
    - almost all projects died in the meantime!
  - Some major successes from that time survived
    - Common simulation engine (Geant4)

- What changed since then:
  - **We know** that we cannot survive if we re-do everything from scratch
  - Around LCG start C++ expertise was rare
    → technology evaluation, rather than project development
  - New developers are used to open-source and its useful conventions
  - **Other scientific domains now heavily depend on SW workflows as well**
    - **Can learn from them!**

Bioconductor project

- Open source project portal for an entire SW ecosystem
  - Supported by a core team of people
- Peer review of software important
  - Development on github, open-peer review
  - Research papers get corresponding paper about the SW
  - Initially people feared their ideas/work are taken away, now their careers benefit from it
- Trade-off between standard settings / non-duplication and being inclusive/open

Depsy

- Platform / infrastructure to **promote credit for software** in science
  - Depsy.org
  - Funded by NSF
  - http://www.nature.com/news/the-unsung-heroes-of-scientific-software-1.19100
- Automatic analysis of papers and software for citations/usage
  - Text-mining in papers
  - Code-mining in GitHub repositories
  - Focussing on Python and R
  - Including transitive credits / indirect contribution
  - Relies on project catalogues (like PyPI)
- Wasn't applied yet/tried for C++ code bases
  - May be harder: no well-established central repository/catalogue for packages

- **HEP Software Foundation (HSF) as the umbrella for addressing the SW challenges together!**

- **Sharing expertise**
  - Schools, trainings and courses (not always easy to find), wikiToLearn
    - Build upon existing solutions like Software Carpentry
  - HEP S&C Knowledge Base
  - HSF Technical Notes
  - Topical fora and working groups in HSF

- **New hardware architectures and technologies**
  - Concurrency forum
  - What technologies are worth betting on?
  - Usage of resources provided by e.g. CERN's TechLab / Openlab
  - Porting efforts within the LHC experiments

- **Software performance and evolution**
  - Simulation: parallelisation of Geant4; GeantV
  - Reconstruction: HSF common tracking SW forum + Machine Learning Forum
  - I/O: parallel ROOT I/O, key-value-store evaluations
  - Mathematics: MetaLibm
  - Ad-hoc improvements and parallelization in various SW projects
  - Performance tools (e.g. igprof, FOM tools)

- **Efficient development**
  - Support in creating/discovering/using/maintaining software
    - *Catalyze common projects*
  - Advertising tools making developers' life easier
  - Build up more commonality (less SW plumbing)
  - Showing best-practices in action

Working Group	Objectives	Forum - Mailing list
Communication and information exchange	Address communication issues Build the SW&C knowledge base Technical notes	hep-sf-tech-forum
Training	Organization of training and education, learning from similar initiatives	hep-sf-training-wg
Software Packaging	Package building and deployment, runtime and virtual environments	hep-sf-packaging-wg
Software Licensing	Recommendation for HSF licence(s)	hep-sf-tech-forum
Software Projects	Define incubator and other project membership or association levels. Developing templates	hep-sf-tech-forum
Development tools and services	Access to build, test, integration services and development tools	hep-sf-tech-forum

- **HSF Forum**
  - http://groups.google.com/d/forum/hep-sf-forum
- **HEP S&C community list**
  - http://groups.google.com/d/forum/hep-sw-comp
  - General mailing list everybody in our field should subscribe to
- **Other specialized lists**
  - General HSF technical discussion forum - hep-sf-tech-forum
  - Lists of the various working groups

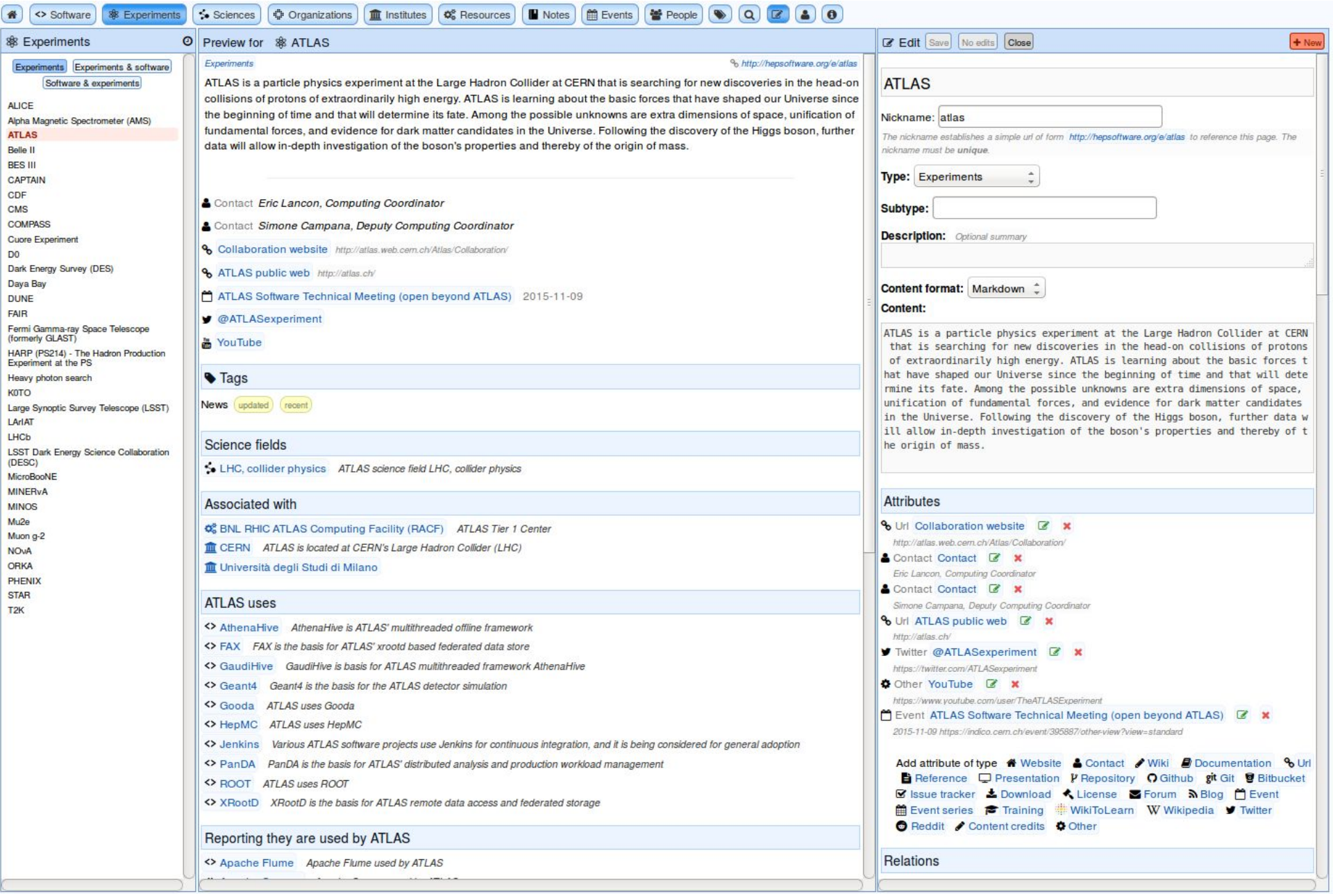
- See the 'Get involved' page on the website for details

- **Our community suffers from code duplication**
  - In some cases a matter of taste, not of technology ('software diva fallacy' or 'not invented here syndrome')
  - In most cases people learn too late about existing solutions

- **The HEP knowledge base is a catalog of software, science fields, communities, and events**
  - implementation is a browser-based app (javascript client, node.js server, json in between, MySQL)
  - authentication is via github, google, facebook etc.
  - **emphasizes easy adding/editing of content, and extensibility.**

    ⇒ Adding content should be fun.

- Available at http://hepsoftware.org
  - Comments/feedback are welcome!
  - If you find something missing → add it!


Software
Experiments
Sciences
Organizations
Institutes
Resources
Notes
Events
People
Experiments
Experiments
Experiments & software
Software & experiments
ALICE
Alpha Magnetic Spectrometer (AMS)
ATLAS
Belle II
BES III
CAPTAIN
CDF
CMS
COMPASS
Cuore Experiment
D0
Dark Energy Survey (DES)
Daya Bay
DUNE
FAIR
Fermi Gamma-ray Space Telescope (formerly GLAST)
HARP (PS214) - The Hadron Production Experiment at the PS
Heavy photon search
K0TO
Large Synoptic Survey Telescope (LSST)
LArIAT
LHCb
LSST Dark Energy Science Collaboration (DESC)
MicroBooNE
MINERvA
MINOS
Mu2e
Muon g-2
NOvA
ORKA
PHENIX
STAR
T2K
Preview for ATLAS
Experiments
http://hepsoftware.org/e/atlas
ATLAS is a particle physics experiment at the Large Hadron Collider at CERN that is searching for new discoveries in the head-on collisions of protons of extraordinarily high energy. ATLAS is learning about the basic forces that have shaped our Universe since the beginning of time and that will determine its fate. Among the possible unknowns are extra dimensions of space, unification of fundamental forces, and evidence for dark matter candidates in the Universe. Following the discovery of the Higgs boson, further data will allow in-depth investigation of the boson's properties and thereby of the origin of mass.
Contact Eric Lancon, Computing Coordinator
Contact Simone Campana, Deputy Computing Coordinator
Collaboration website http://atlas.web.cern.ch/Atlas/Collaboration/
ATLAS public web http://atlas.ch/
ATLAS Software Technical Meeting (open beyond ATLAS) 2015-11-09
@ATLASexperiment
YouTube
Tags
News updated recent
Science fields
LHC, collider physics ATLAS science field LHC, collider physics
Associated with
BNL RHIC ATLAS Computing Facility (RACF) ATLAS Tier 1 Center
CERN ATLAS is located at CERN's Large Hadron Collider (LHC)
Università degli Studi di Milano
ATLAS uses
AthenaHive AthenaHive is ATLAS' multithreaded offline framework
FAX FAX is the basis for ATLAS' xrootd based federated data store
GaudiHive GaudiHive is basis for ATLAS multithreaded framework AthenaHive
Geant4 Geant4 is the basis for the ATLAS detector simulation
Gooda ATLAS uses Gooda
HepMC ATLAS uses HepMC
Jenkins Various ATLAS software projects use Jenkins for continuous integration, and it is being considered for general adoption
PanDA PanDA is the basis for ATLAS' distributed analysis and production workload management
ROOT ATLAS uses ROOT
XRootD XRootD is the basis for ATLAS remote data access and federated storage
Reporting they are used by ATLAS
Apache Flume Apache Flume used by ATLAS
Edit
Save
No edits
Close
New
ATLAS
Nickname: atlas
The nickname establishes a simple url of form http://hepsoftware.org/e/atlas to reference this page. The nickname must be unique.
Type: Experiments
Subtype:
Description: Optional summary
Content format: Markdown
Content:
ATLAS is a particle physics experiment at the Large Hadron Collider at CERN that is searching for new discoveries in the head-on collisions of protons of extraordinarily high energy. ATLAS is learning about the basic forces that have shaped our Universe since the beginning of time and that will determine its fate. Among the possible unknowns are extra dimensions of space, unification of fundamental forces, and evidence for dark matter candidates in the Universe. Following the discovery of the Higgs boson, further data will allow in-depth investigation of the boson's properties and thereby of the origin of mass.
Attributes
Url Collaboration website
http://atlas.web.cern.ch/Atlas/Collaboration/
Contact Contact
Eric Lancon, Computing Coordinator
Contact Contact
Simone Campana, Deputy Computing Coordinator
Url ATLAS public web
http://atlas.ch/
Twitter @ATLASexperiment
https://twitter.com/ATLASexperiment
Other YouTube
https://www.youtube.com/user/TheATLASExperiment
Event ATLAS Software Technical Meeting (open beyond ATLAS)
2015-11-09 https://indico.cern.ch/event/395887/other-view?view=standard
Add attribute of type Website Contact Wiki Documentation Url Reference Presentation Repository Github Git Bitbucket Issue tracker Download License Forum Blog Event Event series Training WikiToLearn Wikipedia Twitter Reddit Content credits Other
Relations

- Technical Notes can be proposals, ideas, whatever people want to add
  - First TN was the one about the TN guidelines
- Repository and version control in GitHub

TN Number	Title	Authors	Download
HSF-TN-2015-01	HSF Technical Notes policy	A. McNab	PDF GitHub

Drafts in the acceptance process:

Draft TN Reference	Title	Authors	Download
HSF-TN-2015-LIC	(Draft) Software Licence Agreements HSF Policy Guidelines	J. Harvey et al.	GitHub
HSF-TN-2015-NAM	(Draft) HSF Platform Naming Conventions - A Proposal	B. Hegner	GitHub
HSF-TN-2015-PKG	(Draft) HSF Packaging Working Group Report	B. Hegner, L. Sexton-Kennedy	GitHub

- **WikiToLearn** is a wiki-based platform tailored at training and teaching
- Initiated in the context of italian universities
  - Basic idea was that students can improve and extend the material of their professors, while still being quality-controlled
- Now investing in providing interactive tutorials
  - think of the combination of jupyter style notebooks and a privately owned sandbox - start tutorial now, resume later
- This is only the shell, content has to come by the community (i.e. you!)
  - **for the next C++11 / 14 / .. tutorial of yours consider exploiting WikiToLearn**

- **Only one of many solutions we are currently exploring!**

HSF is built for and rely upon SW projects

- What does it mean being an HSF project? What's the benefit to projects and users?
- *HSF does not enforce any particular software process, project management or methodology, however packages should conform to some standards to facilitate integration*
  - *Apache Foundation seems a good example!*

Two concrete advancements in the last year

- Software Best Practices Document (draft)
- Project Starter Kit (*hsf_create_project.py* in https://github.com/HEP-SF/tools)
- Both the document and the starter kit considered good approach

Future

- Help with visibility of projects
- Project peer-review: will start with GeantV
- Interoperability of projects ⇒ should actively be worked on
  - Explore the possibility of automatic integration tests
- Get a better understanding within the field what SW licenses to use

- **There are quite a few (more or less) new cross-experiment collaborations, with involvement or moderation of the HSF**

- **GaudiHive**
  - Parallelization of the Gaudi framework
  - Developed by ATLAS, LHCb, FCC

- **Common Conditions Data Project**
  - Discussed between ATLAS, Belle II, CMS and LHCb
  - Repository

- **Projects in the context of the European AIDA2020 project**
  - DD4hep for detector description
  - PODIO for data model definitions

- **Cooperation on Reconstruction Software**
    - "Connecting the Dots" workshop extended by session about common tracking implementations

- **Common Software Build and Packaging Tool efforts**
  - Working group of HSF comparing HEP and non-HEP solutions
  - Starting point was LCG's Librarians and Integrators Meeting

During 2015 the HSF made a review of the packaging tool landscape

- Another field where we have the frequent re-invention of the square wheel
- What tools are around - inside and outside the community?
  - Looked at homebrew, Nix, conda, worch, cmsBuild, aliBuild, LCGCMake, SciSoft, contractor, ...
- What are the requirements of the various stakeholders?
- Is there a chance to reduce the zoo in HEP SW?

One very promising candidate identified: Spack

- From HPC world: one requirement was to support multiple version of the same package
- Great interest: a very good tool for a broad range of use cases… but probably not all

Current activities

- Continue work for adopting spack in HEP
  - Maintain add-on repo “hep-spack” for HEP specific SW
- Fruitful cooperation with spack maintainers
  - All open on GitHub

join the hep-sf-packaging-wg

The HSF has not only a technical part but a political one as well

To address the challenges ahead (e.g. HL-LHC) we need more and dedicated resources

- But where do we want to go?
- Preparing a community roadmap describing HEP S&C challenges and directions (similar to P5 for HEP experiments in the US)
- Similar input to computing TDR for HL-LHC needed
- Will have a set of HSF-branded workshops with a conclusion mid-2017
  - Getting a common understanding within the community
- Participation more than welcome!

There are ideas to make HSF a legal body to hold the intellectual property of software in high-energy physics

License situation in HEP is rather messy and non-understood

- Jan 2015: SLAC HSF workshop established concrete activities and next steps
- Apr 2015: HSF meeting at CHEP 2015 on progress, opportunities, next steps
- Mid 2015: Survey and discussions of tools & approaches in Packaging WG
- Sep 2015: Technical Notes policies published and TN series begun
- Sep 2015: HSF on WikiToLearn
- Oct 2015: Evaluation of new Knowledge Base finished, hepsoftware.org deployed
- Nov 2015: Second generation HSF website deployed
- Early 2016: Startup team meetings replaced with weekly open HSF meeting
- Early 2016: Intensive packaging activity around Spack
- April 2016: Asked to organize a review of GeantV
- April 2016: Refurbishing of the HSF website
- May 2016: HSF Workshop at LAL Orsay

- The HSF is a rather new effort in the HEP SW landscape
- You are the next generation of computing and software experts
  - Make sure your voice is being heard!
- Try common tools
  - If you don't like them, don't walk away silently
  - Give feedback and trust in cooperation
  - If you know a better one - help us making it the common one!
- Contribute with your ideas and software
  - Means much more investment than "here's the code" though
- If you think, things are utterly wrong...
  - … they well may be!
- Let's prepare for the future together