

# Journal on Software and Computing for Big Science

Michel Jouvin

CNRS/IN2P3/LAL

[jouvin@lal.in2p3.fr](mailto:jouvin@lal.in2p3.fr)

HEPiX, Budapest, 24/4/2017



# Prehistory

- Idea presented by Y. Kemp (DESY) during HEPiX in Berlin (spring 2016)
  - Originally proposed by Günter Quast (CMS, KIT, Karlsruhe) in 2013 during a HEP conference in Mainz (Germany)
  - Took 2 years for maturing the idea and come up with the proposal presented by Yves
- Spring 2016: idea presented in different forums like HSF workshop in Orsay, GDB...
  - For example by Markus Elsing (CERN):  
<https://indico.cern.ch/event/394783/contributions/2188799/attachments/1287140/1915139/GDBTalk.pdf>
- General feedback always positive: clearly addressing something that is missing
  - Main controversial topic was the perimeter as discussed at HEPiX: physics versus data-intensive science
  - At the end, Big Science (see Wikipedia): science based on big experiments requiring a long planning, collaborations, a lot of money... and a lot of computing!
  - Good mobilization in France around the project

# Motivation

- Several scientific communities produce and analyze large/huge amount of data ( $O(\text{PB})$ ) in experiments organized as large (international) collaborations
  - HEP with LHC experiments is currently the archetype but coming astrophysics and cosmology experiments (CTA, SKA, LSST), NP experiments (FAIR) on the same line
  - Similar trends in photon science, biology...
- Sharing many computing technologies and facing common computing challenges
  - We know this at HEPiX! Regular presentations by other communities... like Alf in Berlin
  - Not only on the computing resource side but also on the software application side: in particular the challenge of parallelization and efficient use of new processor/machine architectures
  - No community can address these alone: collaboration is a requirement, publication a prerequisite
- In addition, the challenge of managing/distributing large amount of data: several frameworks/infrastructures already developed
  - But often not really known outside of the experiments that develop them
  - Sharing the information about it would avoid to reinvent the wheel for these complex systems

# Publication Challenge

- (Almost) no place to publish about our R&D and solutions
  - Too computing-related to be published where our scientific results are published
  - Not enough research (too much engineering) to be published in computing science reviews
  - CHEP proceedings is the main publication but not frequent enough (18 months) and difficult to share with other communities
  - What is published is too dispersed: difficult to identify and thus failing to contribute to knowledge sharing
- Importance of a validated, peer-reviewed journal
  - Goal: be a reference archive for knowledge sharing in data-intensive sciences/big sciences about software and computing challenges and solutions
  - Means that contents must be validated and its value recognized: impact factor, even if low
  - Indexing required for impact factor : 25-30 articles/year during 2 years before it is considered
  - Help with career recognition of « Research SW Engineers », in particular the young ones

# Journal SW&C for Big Science



- Online journal, continuous publication, in partnership with Springer
  - Christian Caron (Executive Publishing Editor @Springer) associated with the project since its early stage (present at HEPiX in Berlin) and very motivated
  - Springer provides all the manuscript handling system to help with an efficient peer-review
  - Hybrid business model rather than pure Open Access until potential inclusion into SCOAP<sup>3</sup>
    - Too late for including in phase 2 (2017-2019) but positive preliminary contacts for next phase, even if not strictly HEP-focused
    - In fact, Springer decided to allow free access during the 2 first years... so in fact Open Access for now
- Types of articles: topical articles, advanced tutorials, project/technology reviews
  - No proceedings: may cause the journal to be excluded from indexing, other places like CHEP
  - Do not duplicate what is published somewhere else: would be damage the journal image

# Topics: Everything related to SW&Computing

- Infrastructures for large-scale, high-throughput computing
- Middleware development
- Data processing, hosting and sharing
- Distributed data analysis
- Software development Infrastructures
- Software benchmarking and Performance Assessment
- Frameworks and software integration
- Novel algorithms for efficient data reconstruction and filtering
- Deep learning algorithms
- Event and object classification
- Online/Offline data quality monitoring
- Data visualization



# Journal Status



- **Journal has been launched this winter:** <http://www.springer.com/journal/41781>
- Scientific Advisory Board and Editorial Board created
  - <http://www.springer.com/physics/particle+and+nuclear+physics/journal/41781?detailsPage=editorialBoard>
  - S.A.B.: recognized people like E. Elsen (research director, CERN), John Harvey (CERN/SFT)...
  - First Editorial Board meeting at CERN mid-March 2017
- Article submission is open: see the web site
  - First article received one week ago, currently being reviewed
  - Think about what you could publish! A lot of things presented at HEPiX would be worth an article
  - Flexible format/length for articles: from a few pages to hundreds! As the journal is online, no real constraint
- Waiting for a few articles (~5) representative of the journal goals for a more active advertisement
- Feel free to inform your colleagues and all the people that could contribute!

# This is Your Journal



- Editorial Board: several HEPiX members
  - 3 editors-in-chief: Günter Quast (Germany), Markus Elsing (CERN), Volker Beckmann (France)
  - Associate editors: ~15 from all continents/horizons including G. Chen, Yves Kemp, H. Meinhard, M. Jouvin
- No journal without readers... but also without authors!
  - You are among the potential authors
  - Contact us (at E.B.) to discuss your ideas... or submit them directly!
- Journal: an opportunity to increase collaboration with other communities
  - HEPiX has a role to play as this is something we have done for a long time