

Sommario del Workshop della HEP Software Foundation a Parigi

Dario Menasce (contributi da M. Jouvin e B. Hegner)
CCR, La Biodola, 20 Maggio, 2016



Meeting a LAL Orsay, Maggio 2-4, 2016

- 1st workshop at SLAC, January 2015

Agenda: <https://indico.cern.ch/event/496146/timetable/>

- Buon mix di discussioni generali con sessioni ad hoc
- Diverse sessioni hands-on
- Geant4 Technical Forum (evento co-locato, ospitato, al workshop)

Buona partecipazione: ~70 people

- Come a SLAC ma con persone in parte diverse
- Da comunità diversificate
- Buon bilanciamento tra Europa e Nord America (US)
- Rappresentanza da diversi esperimenti, in particolare Intensity Frontier e Belle II

Qualche mese fa è stato indetto un concorso per un Logo ufficiale

Abbiamo un vincitore!



(sottomesso da Joshka Lingemann; 17 voti in 37 risposte)

- Condividere expertise di calcolo scientifico ed infrastrutturale
- Aumentare la conoscenza su software e soluzioni già esistenti (forums)
- Catalizzare nuovi progetti (possibilmente in comune)
- Promuovere economia di scala e collaborazioni per massimizzare l'uso delle scarse risorse umane
- Aiutare sviluppatori ed utenti nello scoprire, creare, utilizzare e sostenere progetti di software di base comuni
- Sostenere lo sviluppo di carriera per specialisti di software e sviluppatori
- Creare un ecosistema a supporto degli sviluppatori di progetti innovativi
- Facilitare l'ampliamento delle connessioni tra sviluppatori: apertura a discipline scientifiche (contingue e non)

- 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

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

Un programma piuttosto denso di progetti

[DIANA-HEP \(USA\)](#)

Interessante sviluppo di un ecosistema di strumenti per l'analisi di grandissime moli di dati, centrato attorno a ROOT (GooFit, RooStat, uso di GPU e tecniche di vettorializzazione)

[AIDA2020 \(EU\)](#)

Iniziativa analoga agli scopi di HSF, centrata essenzialmente sulle problematiche dell'online, DAQ e software di supporto all'ecosistema dei rivelatori (non necessariamente quelli molto grandi, spazio per esp. tipo Gr. II e III)

[Future Conditions Database](#)

Progetto congiunto ATLAS/CMS per lo sviluppo di *condition-database* di prossima generazione
Belle-II ed LHCb lo stanno prendendo in considerazione; nessuna conclusione definitiva (per ora)

[HEP S&C Knowledge Base](#)

Collettore di informazioni (in lettura ma anche in scrittura) circa I progetti SW esistenti (**già in produzione ora**)

[WikiToLearn](#)

Piattaforma (sviluppata altrove) suggerita ad HSF per programmi di Training&Tutoring
Una comunità di circa 40 sviluppatori e 160 editors (fornitori di contenuti)

HSF è rivolto allo sviluppo di nuovi progetti SW e si avvale della collaborazione di quelli già avviati ed esistenti (es ROOT, Geant,...)

- Cosa significa essere un progetto supportato da HSF? Che benefici vengono al progetto stesso e ai suoi utenti? E agli sviluppatori? Concetto di “*umbrella organization*”, dibattito tuttora in corso...

Due progressi concreti nell'anno trascorso

- Documento sulle Software Best Practices ([draft](#))
- Project Starter Kit (`hsf_create_project.py` in <https://github.com/HEP-SF/tools>)

Entrambi i documenti e lo starter kit considerati favorevolmente come un buon approccio

Per il futuro

- Essere di aiuto per aumentare/migliorare la visibilità ed il riconoscimento dei progetti
- Lavorare sul lato interoperabilità dei progetti (laddove utile/necessario) ⇒ molto lavoro da fare
Esplorare i sistemi di automazione dei test di integrazione

Durante il 2015, il Packaging WG ha effettuato una review/survey del panorama degli strumenti di packaging dei SW in uso corrente

- Quali sono , chi li usa e come funzionano, dentro e fuori dalla nostra comunità?
- Quali sono i requirements dei diversi stake-holder? Esistono commonalities? Quali?
- C'è una prospettiva di ridurre lo Zoo di strumenti esistenti da mantenere?

Identificato un candidato molto promettente: **Spack**

- Sviluppato in ambito HPC: il requirement originale era la manutenzione di versioni multiple dello stesso package
- Vivace discussione (~30 persone) in una sessione dedicata, (Patrick Gartung, FNAL)
- Molto interesse: un buon strumento per un vasto range di casi, anche se non li copre tutti (difficile possa essere efficiente per grandi esperimenti con stack SW molto complesso, ma probabilmente più che adatto per comunità più piccole, meno complesse e meno demanding)

Futuro

- Continuare investigazione e sperimentazione in ambito HEP
- Manutene un repository “hep-spack” per SW specifico nella ns comunità
- Nessun investimento per ora per funzionalità specifiche

Bioconductor project (W. Huber; EMBL)

Progetto di portale Open source con intero ecosistema SW



- Supportato da un core team dedicato

Peer review del software considerato importante

- Sviluppo peer-reviewed, open, su github
- Per ogni articolo di ricerca, viene prodotto il corrispondente paper sul SW
- Timore iniziale degli sviluppatori che le loro idee/lavori venissero “trafugati”: timore superato, ora le loro carriere ne beneficiano (non è chiaro a me come)

Netherlands eScience Center - D. Remenska

netherlands eScience center

Un'organizzazione che avvicina la comunità scientifica a quelle delle computing infrastructures

- Core team di esperti in vari campi del computing
- Finanziato da NeSC: permette di avere expertise in casa e fare “technology transfer”

Depsy (J. Priem)



Piattaforma/infrastruttura per promuovere le carriere nello sviluppo del software scientifico

- [Depsy.org](https://depsy.org)
- Finanziato dalla NSF
- <http://www.nature.com/news/the-unsung-heroes-of-scientific-software-1.19100>

Analisi automatizzata di papers e software per uso/citazioni

- Text-mining nei papers
- Code-mining nei GitHub repositories
- Focalizzato su Python ed R
- Include crediti transitivi (quando il codice viene ri-usato da altre comunità)

Non ancora provato/applicato su codice in C++

Può essere complesso: non esiste un catalogo centralizzato di repository da monitorare

½ giornata: varietà di talk su Machine Learning con focus su SW e Tools

- Sommario su OpenLab Workshop (industria)
- TMVA updates
- Sfide per il Machine Learning (es. HiggsML)
- Deep Learning
- Common Tracking Software (ACTS)

RAMP (Rapid Analytics and Prototyping) Hands-On Session on Anomaly Detection (3h)

- Ricerca di anomalie (distorsioni) in un dataset opportunamente modificato (HiggsML)
- Inter-experimental Machine Learning Working Group (IML)
- WG inter-esperimenti dedicato alla condivisione di lavoro ed idee su ML in HEP
- Endorsed da tutti gli exp LHC, ma con partecipazione da DUNE, NOvA...
- Meeting mensili aperti a chiunque interessato
- HSF forum for Machine Learning

Organizzata da V. Innocente (CMS), P. Calafiura (ATLAS) e M. Schulz

Contributi da ALICE, ATLAS, CMS, GeantV, ROOT, Art/LArSoft, e la comunità di Astroparticle

Seguita da una sessione di discussione

- Su quale HW concentrarsi? Commodity vs. GPU vs. HPC
- Ancora inesistente una comprensione condivisa su cosa aspettarsi per il futuro e come affrontare il problema in modo organico
- Possibile disaccoppiare l'ottimizzazione low-level (esperti) da codice di alto livello (fisici) tramite opportune librerie? Come organizzarsi a tale scopo? Quale overseeing board?
 - ✓ Interessante input dagli astrofisici

Poste più questioni che risposto a domande (Dettagli nelle [meeting notes](#))

Sommario

Occorre organizzazione, visibilità e lavoro di coordinamento (HSF?)

- Concreto follow up nel SW Technology Forum (un tempo noto come Concurrency Forum)

Un importante input alla Community Whitepaper (dettagli dopo)

HSF ha sviluppato alcune iniziali attività collaborative, ma per affrontare le sfide poste di fronte alla comunità (e.g. HL-LHC) occorrono risorse dedicate

- Proposta per una Community roadmap che descriva queste sfide di HEP S&C e suggerisca soluzioni e metodi (in analogia al P5 per gli esperimenti HEP negli US)
 - Idea proposta/supportata da molti colleghi US e presentata da P. Elmer
- È stato suggerito di costituire una serie di workshop dedicati, sotto l'egida HSF, con una conclusione per la metà del 2017
- Occorre un organismo di coordinamento, un feed-back e un sostegno dalle funding agencies

Risultato della discussione

- LHCC ha chiesto un documento analogo: WLCG incaricato di preparare un TDR per HL-LHC
 - Scala temporale simile/compatibile con il CWP proposal
- **Generale consenso che si tratti di questione ad elevata priorità**
- Sollevata la possibilità di co-locare un kick-off workshop per il CWP al prossimo CHEP

- Una nuova proposta da colleghi tedeschi e da Springer
 - Presentata ad HEPiX a metà Aprile
- Una nuova rivista refereed, abstracted, indexed su tematiche specifiche di HEP computing
 - Che costituisca un archivio di riferimento centralizzato ed autorevole
 - Che sia di supporto al problema di riconoscimento delle carriere
 - Non ristretto alla sola HEP community: aperta alla Data-Intensive Physics.
 - Settori organizzati come collaborazioni centrate su esperimenti a grande scala
 - Lo scope copre tutti i campi del computing: da infrastrutturale a daq, a data analysis...
 - Pubblicazione continua (periodica): no paper/volume
- Diverse questioni aperte
 - Il focus e la dimensione più opportuni
 - Il business model: ibrido (OA access per articolo, su scelta/necessità dell'autore) vs. puro OA
 - Puro OA richiede sponsorship (supporto).
 - Tipologia: normali articoli, reviews, advanced tutorials, no letters e no proceedings
 - Struttura editoriale

HSF gode di buona salute

- Progressi significativi ed umentato interesse/motivazioni rispetto all'anno passato
- Community White Paper riconosciuto come ottimo incentivo a progredire verso una visione condivisa dei progetti e delle sfide future

HSF e la comunicazione: esplorare l'uso di StackExchange (nuovo branch)

- Open forum identificabile come riferimento standard per domande/discussioni sul HEP computing
- Un'alternativa al forum basato su mailing-lists?

Sviluppo/aumento del supporto dato da HSF ai progetti SW

Cercare un “*official blessing*” ad HSF da organismi quali ECFA/ICFA (et al)

Iniziata una discussione sulla possibilità di rendere HSF una entità legale

- Ancora lontani da consenso condiviso, ma concordi nell'esplorare almeno le possibilità e le implicazioni assieme alle Funding Agencies (e ai loro organismi legali)
- Scopo iniziale: configurare la IPR (proprietà intellettuale) di HSF su base analoga a quella della Apache SW Foundation (IPR transfer)