EIC_NET & CNAF

June 10

Informazioni

- E’ stato pubblicato lo Yellow Report dell’EIC:

  Science Requirements and Detector Concepts for the Electron-Ion Collider: EIC Yellow Report
- E’ stato pubblicato il CDR dell’EIC:

  Conceptual Design Report (CDR) 2021
- In marzo è stata pubblicata la call for detector proposals con scadenza al primo dicembre 2021

  Call for Collaboration Proposals
- Si sono formate tre proto collaborazioni (INFN è in ATHENA):
  - ECCE (EIC COLLIDER EXPERIMENT)
  - ATHENA (A Totally Hermetic Electron-Nucleus Apparatus)
  - CORE (The COmpact detectoR for the Eic)

Informazioni 2

- E’ stato molto recentemente formato il Computing Working Group dell’EIC, con il task iniziale di controllo/gestione delle risorse messe a disposizione per le simulazioni delle protocollaborazione
- Composizione: Jerome Lauret, Graham Heyes and me
  - + representatives from the proto-collaborations (up to two)
    - For Athena is Wouter
- First action meeting on Monday:
  - Regular contacts with ECCE already started
  - For Athena will start soon
  - For CORE no news
- First estimates from collaborations received. Will need extra resources from the labs

Resource request from ECCE and ATHENA

CPU and storage estimates

Background justification

Short time to produce the samples necessary for the detector proposals. Some of the workflows envisioned are conservative (do with what we have as tools, framework, … KISS). Estimates should also fold the reality of a changing model (not optimized in many ways)

.ATHENA (IP6)

- CPU: 10 M CPU hours estimated, done by September ⇔ 3,500 cores at BNL, 5k-7k cores total
- Storage: 2 PBytes needed
- Will fit in 2 GB memory / job

CORE

- So far, no feedback provided

ECCE

- CPU: 2k-3.5k cores estimates, 4 months usage at 75% capacity (3 months + 30% contingency)
- Storage: 2 PBytes of simulation output, 100 TB of analysis - estimated based on 1 Billion events (sPHENIX simulations @30 TB per 100 Million events)
- Will fit in < 2 GB memory / job

Overall needs

- CORE unknow - assuming ½ CPU and 30% less storage (revised storage estimates on the next slide include 25% contingency)
- **CPU is a non-issue**
  - 2x2k CPUs (BNL and JLab) - **4k available**
  - Assuming CORE @ 1,000 CPUs, **8k core possible** modulo ***recovering 4k slots from the OSG***
  - Memory is standard - wide available of off-the-shelf HTC nodes on the Grid
- Storage - Revised ECCE estimates [3PE approach] indicates a 1.5 Pbytes / collaboration need
  - **Assuming 4 PBytes needed total** (1.5x2 + 1 CORE)
  - 1 PBytes available at BNL, 1 PBytes being investigated at JLab, **we will be short by 2 PBytes**
  - Possible squeeze: some sample may be stored on tape (tape will need to be purchased) **MINIMUM additional not accounted for +1 PBytes**

Estimates verifications by ECCE (storage and memory)

- Topics
  - SIDIS acceptance and unfolding matrix, full Geant 4 simulations, 70 x-Q^2 bins [YR Fig 7.50], 16 Phi bins, 10k events per bins => ~ 100 M events (*low end*)
  - HF physics and background signal - 100 M events, G4 simulation to qualify fast background simulation
  - Exclusive process - 1 Milion events
  - Overall event counts = 200 M events
- Cardinality: x3 simulation campaigns x 1-3 configuration options => assuming 1 Billion events
  Detailed plan
  - First iteration/May : 1 detector option - in progress
  - Second iteration/June-July : 3 detector options
  - Third iteration/July-Sept: 2 detector options
- Verifications
  - The first event sample produced (1 M events) show a size of 300 GB - simplest configuration possible, 100 M == 30 TB
  - 100 million events of p+p events in sPHENIX requires 130 TB of storage (pp multiplicity higher than ep but no forward detectors …)
  - Mean memory was measured to be 1.GB with an stdev of 200 MB.- a standard 1.5 GB of memory will be good
- Assumption: 1 Billion events required
  - **2 Pbytes is remarkably consistent between ECCE and ATHENA initial evaluation** (assuming most likely value)
  - **ECCE provides two more points - an optimistic lower bound and a pessimistic value**
  - **Assuming a triangular distribution and still a 25% contingency, 1.5 Pbytes seems needed per proto-collaboration**
  - Note: perhaps 30% could be archived on tape (this will be considered in our minimal estimates for a purchase but tape will then need o be purchased)

Stato dell'OSG entry point al CNAF (ultimo mail)

From Jerome LAURET <jlauret@bnl.gov>
Subject **Re: EIC configuration**
To Federico Versari <federico.versari@cnaf.infn.it>, emfajardohernandez@physics.ucsd.edu, andrea.bressan@ts.infn.it, Domenico Elia <Domenico.Elia@cern.ch>
Cc Cnaf Support <user-support@lists.cnaf.infn.it>

Kind greetings Federico,

The way the OSG would support this is through a data Federation (Xrootd / XCache based ...) so any instance can be used seamlessly. We had slides from Edgar a while back (will find them back). We are also trying to understand if we can leverage "S3 like" storage (we have storage at BNL with a S3 interface in front). The possible advantage here is a "thin" client (no need for special infrastructure component) but we are not all sure yet this is the proper path forard - we agreed with Graham to finalize our thoughts by next week. The submit host is much more straight forward ...

Hope this answers your question (will send you more info soon) and that my Email will find you all well.

On 2021-02-04 09:58, Federico Versari wrote:

> In particular, for what concerns the file transfer, how do you expose EIC data at BNL? So that we can implement the best way to transfer data between our sites.

--

```
      ,,,,,
     ( o o )
--m---U---m--
    Jerome
 <he-him-his>
```