Highlights of the BNL NPP Discussion on ePIC Collaboration Tools and Technologies

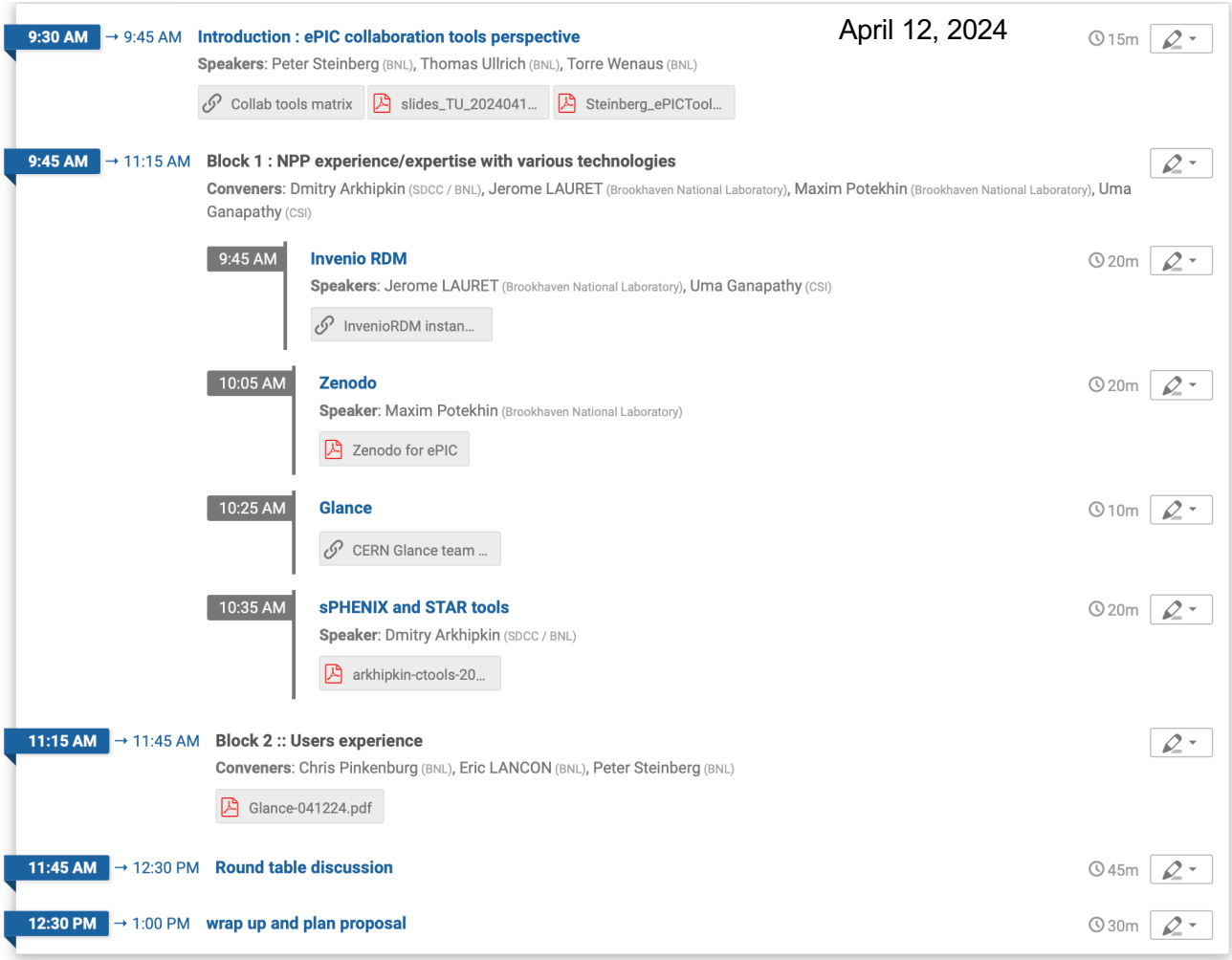
Alexei Klimentov

Brookhaven National Laboratory

April 18, 2024

Agenda

April 12, 2024



9:30 AM → 9:45 AM **Introduction : ePIC collaboration tools perspective** 15m
Speakers: Peter Steinberg (BNL), Thomas Ullrich (BNL), Torre Wenaus (BNL)
Collab tools matrix | slides_TU_2024041... | Steinberg_ePICTool...

9:45 AM → 11:15 AM **Block 1 : NPP experience/expertise with various technologies**
Conveners: Dmitry Arkhipkin (SDCC / BNL), Jerome LAURET (Brookhaven National Laboratory), Maxim Potekhin (Brookhaven National Laboratory), Uma Ganapathy (CSI)

- 9:45 AM **Invenio RDM** 20m
  Speakers: Jerome LAURET (Brookhaven National Laboratory), Uma Ganapathy (CSI)
  InvenioRDM instan...
- 10:05 AM **Zenodo** 20m
  Speaker: Maxim Potekhin (Brookhaven National Laboratory)
  Zenodo for ePIC
- 10:25 AM **Glance** 10m
  CERN Glance team ...
- 10:35 AM **sPHENIX and STAR tools** 20m
  Speaker: Dmitry Arkhipkin (SDCC / BNL)
  arkhipkin-ctools-20...

11:15 AM → 11:45 AM **Block 2 :: Users experience**
Conveners: Chris Pinkenburg (BNL), Eric LANCON (BNL), Peter Steinberg (BNL)
Glance-041224.pdf

11:45 AM → 12:30 PM **Round table discussion** 45m

12:30 PM → 1:00 PM **wrap up and plan proposal** 30m

25+ participants

Indico page and presentations
Live notes

CDS (CERN Document Server) is the CERN official repository for publications, articles, reports and multimedia content
The Glance system, a generic mechanism for accessing any database, acts as an intermediate layer isolating the user from the particularities of each database. It retrieves, inserts and updates the database independently of its technology and modeling.

Outline

- A set of requirements for collaborative tools have been discussed by the ePIC collaboration some as far ahead as 2020
  - The ePIC Ad-Hoc Collaborative Tools Committee will come up with an updated list of requirements and possibly a conclusion
- None of the available today solutions meet all the requirements
- ePIC needs these tools today, even if they provide very basic functionality
- All agreed that we need to be pragmatic (address ePIC TDR needs) and agreed with the proverb
  - *A bird in the hand is better than a pie in the sky*
- All agreed that *we can only make a proposal, it's up to ePIC to decide*
- Torre's table summarizes the current state of available tools
- The following tools were discussed in the context of available expertise, readiness and user experience

Membership database

- **"Dmitry's phonebook"**
- ***Glance***

Document database

- **Zenodo + overleaf + github**
- Invenio RDM
- ***Glance + CDS***

Document development

- Zenodo
- Invenio RDM
- ***Glance + CDS***

- *Zenodo and InvenioRDM quite similar in many ways, and related*
- *Glance not suitable for TDR timeline*
- In a development we learned of after the meeting, the EIC Project has decided that Overleaf+GitHub will be used for the TDR. Overleaf licenses have been purchased

1. Workforce consideration (that is a valid consideration that could lead to a pragmatic choice)
2. *How is authorisation done?*
3. *How do Zenodo and InvenioRDM compare to the requirements ?*
4. *How to migrate content from Zenodo to Invenio (and the other way round) ?*
5. *What are the external dependencies ?*

In italic – not suitable for TDR timeline

Brookhaven National Laboratory

The following slides for the ePIC Collaboration to consider

Proposal : Membership database

- Adopt the system developed and maintained by D. Arkhipkin as the system to use now.
  - *Dmitry commits to supporting the membership DB components of the system for ePIC. The code is public in github.*
    - *BNL will find a way to hand over some of his RHIC tasks to another person*
    - *ePIC/BNL will find someone to work part-time with Dmitry.*

- Begin researching a strategy for the longer term.
  - Glance is a possibility, very popular with its users, but it is a 20 year old system written in php that has never been used with a non-Oracle DB or with infrastructure outside CERN. Is it practical and sustainable to pursue BNL-resident non-Oracle Glance.
  - Is there any possibility of using Glance at CERN by ePIC? It would still leave a lot of work, it takes an entire team to tailor an instance for each experiment, but it would greatly reduce the effort and uncertainty.
    - *We will be asking the glance teams about their interest in evaluating non-Oracle backends.*
    - *We will be asking the glance team about their plans to have code modular*

Proposal : Document Database

- Adopt zenodo.org as a system meeting immediate needs and available today without further development, as the system to use through the TDR period.
- Content put in zenodo.org will be able to be migrated to other InvenioRDM based systems in the future, should ePIC decide to do so.
  - *Do we need to demonstrate/evaluate content migration today ?*
- Review after the TDR whether to continue with zenodo.org or move to another solution.

Proposal : Document development

- In a development we learned of after the meeting, the EIC Project has decided that Overleaf+GitHub will be used for the TDR. Overleaf licenses have been purchased. So **the document development toolset through the TDR period is settled.**

- InvenioRDM remains an option for the longer term that integrates document development and management. EIC should follow the development of InvenioRDM, evaluating it periodically and helping to guide the development.

- We learned in the meeting that sPHENIX InvenioRDM's search problem has been addressed (drafts are now indexed). In light of this, reevaluate the performance and usability of sPHENIX InvenioRDM.
  - Assemble a list of essential requirements that remain to be implemented to satisfy ePIC needs.
    - Start evaluation/development of an interface between InvenioRDM and membership DB
    - Establish how to obtain DOIs
  - Translate the list into an effort plan and timeline, explicitly indicating effort coming from BNL and from the core InvenioRDM team.
  - Assess the plan and decide whether to plan for InvenioRDM adoption post-TDR.

- There is a 'new CDS' emerging, implemented with InvenioRDM. If it presents the opportunity to use CDS in the longer term, that will be an interesting option to evaluate.

Summary Table

Tool	Technology	Today's Status	Plans
Membership database	Dmitry's phonebook	Functionality needed now is there	Long-time support and integration with Document Database
	Glance		Continue discussion with CERN
Document Database	Zenodo.org + Overleaf&Github	Feature set looks good now, holes have been filled, e.g. private documents, multiple managers, community (e.g. experiment) support, excellent search, citeable software (GitHub integration)	BNL zenodo instance Document identifiers to be added
	InvenioRDM	Working version for sPHENIX with not all functionalities requested by ePIC; ePIC prototype@BNL	Federated ID and accounts to access to be evaluated
	Glance + CDS		continue discussion with CERN
Document Development Database	Glance + CDS		CDS → InvenioRDM (by CERN)
	Zenodo		Required development and support to be assessed
	InvenioRDM		Required development and support to be assessed

Acknowledgements

Thanks to D.Arkhipkin, J.Lauret, P.Steinberg, T.Ullrich and T.Wenaus for slides and materials

Thanks to all BNL NPP people for very fruitful discussion and excellent presentations