

# sPHENIX/NPPS planning

[sPHENIX MDC2 planning meeting googledoc](#)

## sPHENIX MDC2 planning meeting Jan 13 2021

- Conditions DB
  - Client new release v0.6.0 <https://github.com/BNLNPPS/xpload/releases>
    - A test program that tries to simulate random calls to the CDB and a condor template to submit multiple jobs
    - An addition of a parameter to the API fetch() function to constrain the selection payloads by payload type. <https://github.com/BNLNPPS/xpload/blob/main/include/xpload/fetch.h#L17>
  - On the server side, attention is shifting to deployment and performance, it would be good to understand the timeline of the production hardware for the persistent database (currently NFS). Deployment will move to an apache front end to the Django application.
- Distributed computing
  - PanDA is there to be used. Ball is in sPHENIX court.

## sPHENIX MDC2 planning meeting Dec 16 2021

- Conditions DB:
  - Server side updates, main read interface remains the same
    - Added interfaces to extract information for PayloadList and PayloadIOV objects by Id
    - Added interface to set the GT's status (locked/unlocked).
  - Scalability tests continue. Apache Benchmark was installed (by John DeStefano) at spool0001/spool0002 machines to test parallel requests.
  - Tested xpload client read functions against OKD Server instance with permanent storage
  - Performance tests with multiple jobs to determine the need for caching on the client side
  - The client interface to resemble a version controlled additions into DB (as discussed in [sPHENIX meeting](#))
- Distributed Computing
  - Jobs run now under sphnxpro - the sPHENIX production account, verified the file ownership is correct
  - Queue was set up with 10GB max memory, 48h cpu limit, 1 core which is sufficient for now

## sPHENIX MDC2 planning meeting Dec 2 2021

- Conditions DB:
  - New version of the client: <https://github.com/BNLNPPS/xpload/releases/tag/v0.5.0>
    - [Installed and tested for read functionality](#)
  - Server side also updated, read interface remains the same
    - Deployed in OKD, 1 pod for the Django application, 1 pod for the DB (Postgres with 10GB of permanent storage)
  - Scalability tests starting, firstly against a trivial DB content to study service scalability, and then add realism (increase the amount of information)
- Distributed software
  - CERN-based PanDA:
    - in use by sPHENIX, to be used in MDC2 from the beginning (yes?), no issues NPPS knows about
    - We of course have more to do, if/when sPHENIX wants to use other sites like NERSC, use pchain for complex workflows, add sPHENIX-specific monitoring, etc

- The NPPS group meeting tomorrow will focus on sPHENIX MDC2 deliverables and also the longer view, in particular data carousel application to sPHENIX
- BNL-based PanDA:
  - Doug said a couple of weeks ago that he was going to set up Postgres on OKD for PanDA (and other purposes). Tadashi hasn't heard further.

## sPHENIX MDC2 planning meeting Nov 18 2021

- NPPS update
  - Conditions DB
    - *OKD service & ephemeral database storage → issues/ticket?*
      - was solved quickly and permanent database storage now works
      - Work done by Tejas (SDCC): "I have created few Persistent volumes that can be used for this purpose. The default storage class is `standard` and all the newly created Persistent volumes are using the `standard` storage class. Since this is a test OKD cluster, the Persistent volumes are NFS based on a shared storage. You should get very good performance from the shared storage in general. The production OKD cluster will be using a dynamic storage provisioner (Trident) and will have a dedicated storage appliance."
    - Bugfixes and updates to the CDB service side APIs and schema, making payload lists first class objects for easier conditions management.
    - Documentation, testing/hardening the OKD based service, scaling tests
      - Working on automating the deployment of the DB service in OKD in order to provide robust and predictable development and test environments. We are also working on automating the tests to make sure any API changes are rolled out in sync with the corresponding changes on the client side.
  - Distributed software
    - *Scale of available resources for sphenix on the google cloud?*
    - *Progress on identifying a postdoc to help?*
    - Sergey and Xin found out the reason why the pilot running on the shared condor pool was not able to talk to the PanDA server. Jobs have been showing up successfully in PanDA after a configuration parameter was set in the pool to copy job's X509 credentials.
    - Helm charts of the PanDA server and JEDI are ready to try at BNL OKD.
  - Looking ahead...
    - if sPHENIX plans to use NERSC, work needs to begin soon, to establish an allocation and to set up there
    - and if sPHENIX plans to use NERSC (and any other remote site), Rucio prep needs to move ahead
      - What is the Rucio timeline? This is mainly SDCC, NPPS has expertise that can help, but SDCC has taken on setting up sPHENIX Rucio, and we will need to work closely with the effort, integrating PanDA-Rucio

## sPHENIX MDC2 planning meeting Nov 4 2021

- a. NPPS updates
  - i. Conditions DB
    1. On the [client side](#) we have a prototype (c++) providing basic functionality for fetching payloads for a given global tag and timestamp, using the CDB service in OKD.
    2. The current OKD service uses ephemeral database storage, this needs to be changed to permanent storage but there were problems with using that initially.
      - A ticket has been opened [HERE](#) (may not be visible for non-SDCC, I can't see it)
    3. Next focus on tuning and debugging the client/service interface.
    4. Stress testing using e.g. apache bench or Locust - <https://locust.io/> (an open source load testing tool) will wait until permanent database storage is used.
  - ii. Distributed software

1. Chris has PanDA submission going, jobs are succeeding but success is not fed back to PanDA yet, diagnosing the reason
    - iteration time can be long, we really really need a postdoc with the time to be banging on the system and tightening up the bugfix iteration time
  2. Doug helping with PanDA@OKD, consulting with Tadashi et al
  3. Good progress on containerizing PanDA, JEDI central services
  4. PanDA monitor is working with postgres, meaning all the pieces are working with postgres now (to first order)
  5. Configuring BNL PanDA queue for sPHENIX cluster, pointing to the right gatekeeper etc, needs to go to shared pool under sPHENIX quota
  6. PanDA team prepared to set up queue for sPHENIX on the Google cloud (we have resources there through an ATLAS/Google collaboration), in case extra resources will be needed for MDC2
    - Scale of available resources?
  7. Amol's production monitoring task is agreed and [described](#) so he can get started
- b. PanDa
- i. Run actual sPHENIX tasks via iDDS (sPHENIX contact available?)
    1. cf above, debugging getting submissions directed at proper sPHENIX resources. Needs a tight iteration loop with an sPHENIX person submitting jobs, need a postdoc to offload Chris.
- c. Conditions DB
- i. Implement the sPHENIX-side CDB client service in their framework (sphenix contact available?)
    1. Client service alpha version ready to be integrated and tried in the framework
  - ii. Documentation, testing/hardening the OKD based service, scaling tests
    1. Underway

## NPPS meeting focused on sPHENIX developments Oct 29 2021

- [Meeting notes](#)

## sPHENIX MDC2 planning meeting Oct 21 2021

- NPPS update
  - Conditions DB
    - Service is up on OKD at SDCC
    - Populating it with test conditions data - Done October 21st
    - Dmitri Smirnov has familiarized himself with the system (brought up a test service instance on his laptop), with next steps
      - test the API (via curl) to the SDCC instance
      - develop the C++ API and library interfacing the fun4all based framework to the conditions DB
    - As of Oct 21 morning, Dmitri has a first version
      - <https://github.com/plexoos/xpload/releases/tag/v0.1.0>
    - Conditions DB channel in mattermost will be the primary rapid-iteration communication channel as development proceeds
      - <https://chat.sdcc.bnl.gov/npps/channels/conditions-db>
  - PanDA
    - CERN DOMA instance ready when Chris is ready :-)
    - Chris has asked sPHENIX for a postdoc to eventually take over, good! Surely this is high enough priority to make sure one is found?
    - Extremely slow progress towards establishing BNL instance, still don't have VMs

- New distributed software channel in mattermost will be the primary rapid-iteration communication channel as development proceeds
  - <https://chat.sdcc.bnl.gov/npps/channels/distributed-software>

# sPHENIX MDC2 planning meeting Oct 7 2021

- [Meeting notes](#)
- NPPS update
  - What is the timeline for implementing/testing PanDa and condition-DB?

## ■ PanDA

- The plan and timeline as established at the [Sep 17 NPPS meeting](#):
- Proceed towards MDC2 making use of the available, proven, scalable PanDA DOMA services
  - The DOMA instance is of sufficient scale and capability to support this, scalability shown by Rubin usage (~250k concurrent jobs)
  - DOMA instance is usable now by sPHENIX; sPHENIX authentication has been supported on it for ~1yr
- In parallel establish BNL PanDA services backed by postgres so this development can proceed concurrently, while off the critical path
  - Postgres port is well advanced, with the PanDA core working, and other services including IDDS and harvester working already with postgres. Remaining work to do is in completing the porting of monitoring, and testing scaling.
  - PanDA team specified the resources needed as follows, and sPHENIX requested them of SDCC
    - JEDI and the panda server: 2 VMs with 8 cores + 16 GB RAM + 200 GB disk on each
    - Postgres: 1 VM with 8 cores + 16 GB RAM + 1 TB disk
    - Harvester+condor: 1 VM with 4 cores + 8 GB RAM + 200 GB disk
    - Pandamon: 1 VM with 8 cores + 16 GB RAM + 200 GB disk
- When BNL PanDA services are complete and validated, proceed with a smooth switch-over from DOMA to BNL PanDA.
  - If postgres completion and validation proceeds quickly, this switch-over could happen before MDC2 (an ideal scenario, but possibly not a realistic one).
- Timeline
  - ~2 weeks [ie now]: hello world and real payloads with DOMA
    - run hello-world using prun on DOMA PanDA BNL-OSG queue
    - run real payload on BNL-OSG through direct task submission to PanDA/JEDI

- [PanDA is ready for this, [sPHENIX engagement is needed to do it](#)]
  - 1 month: workflow management for real sPHENIX tasks
    - run actual sPHENIX tasks via iDDS
    - [\[we can try this as soon as we have sPHENIX engagement to implement sPHENIX tasks\]](#)
  - 2 months: production tasks at BNL and remote
    - sPHENIX using prun/pchain for script-based production submission
      - [pchain (which uses iDDS) is ready, and proceeding directly to this step in lieu of (ii) above may be the most efficient course]
    - run actual sPHENIX tasks at NERSC
      - using an sPHENIX harvester instance there cloned from the ATLAS one, probably Globus Online for data transport with BNL, and preferably using Rucio
      - [access to and services at NERSC have to be arranged]
- **Conditions DB**
  - Demonstrator prototype of the conditions DB has been implemented, as a generalization and technology update of the Belle II conditions DB, reported in [Sep 3 NPPS meeting](#), with next step to establish the service at SDCC
  - OKD server made available at SDCC to host the CDB, **and working** on Oct 5 the CDB developer Ruslan Mashinistov got the CDB running on the OKD server
  - Developer for the sPHENIX CDB, Dmitri Smirnov, brought into the project this week as the project is now ready for him (APIs defined, SDCC based service available)
  - In next weeks Dmitri will implement the sPHENIX-side CDB client service in their framework
  - Priorities for the CDB team are documentation, testing/hardening the OKD based service, and scaling tests
  - [sPHENIX engagement is needed to guide the client service design and framework integration, and to populate the CDB with data](#)
- What are the hardware needs for large-scale testing, and on what timescale?
  - As discussed above. With DOMA PanDA available, BNL PanDA resources requested (see below regarding issue with how they've been provided), and BNL CDB resources requested and delivered, the hardware is there for MDC2 scale operation.
  - BUT see comments below on OKD based resources.
- **SUMMARY of the status of NPPS services:** [they are at the stage of requiring direct and sustained sPHENIX engagement to advance them from 'available services' to 'MDC2-ready sPHENIX services'.](#)
  - **Need a postdoc to act as liaison/develop and working with npps**

- **IMPORTANT late addition:**

- Comment from Tadashi, PanDA project lead and lead developer, this morning with regard to getting going on the SDCC PanDA instance, for which OKD test cluster resources have been provided:
  - “Container images of panda components are not very matured except for harvester.
    - Right now, there is no manpower in the panda team to work on the container stuff as we don't put a higher priority, so I don't expect that it will be ready in a couple of months.
    - It would be much easier if there are VMs up and running, and panda components are installed directly there using sudo. (which is done w/ atlas at cern and in other instances)
- I (Torre) strongly second his comment (not having appreciated how different OKD based resources are). We are operating with *no* supported sPHENIX specific manpower. Expect to add months if we cannot implement sPHENIX on a smooth familiar path.
- Possible to set OKD same as for atlas?
  - Eric: First time hear that OKD is not the right solution... **Need some time to propose a solution**
  - Torre: we asked specifically for VM (cose was used in the past , and what's we were familiar with) → not exposed today
  - Doug: at purchasing review, because of cyber security policy at BNL (which is different than CERN)...
  - Eric: also didn't realize that sudo is OK → **have to define which sudo**
    - **TW - Tadashi on Sep 9 when he was asked in a thread with Chris: “ATLAS uses sudo.”**

## New developments for sPHENIX - Sep 17 2021 NPPS meeting

- <https://indico.bnl.gov/event/12550/>
- [Meeting notes](#)
  - Developments and needed actions since the Sep 3 meeting
  - CDB - Ruslan spent the week working on deployment on openstack at CERN (until OKD becomes available) and instructions. Paul L has been testing the instructions and giving feedback, got as far as getting the service to start in apache (django service, with postgres as DB). DB population scripts and service testing scripts now also available for testing. Ruslan also looking at cleaning up the code repository to move into NPPS github.
  - Workflow/workload management. The draft plan towards MDC2:
    - Proceed towards MDC2 making use of the available, proven, scalable PanDA DOMA services
      - DOMA in shape to support this, scalability shown by Rubin

- In parallel establish BNL PanDA services backed by postgres so this development can proceed concurrently, while off the critical path
  - Postgres port is well advanced, with the PanDA core working, and other services including iDDS and harvester working already with postgres. Remaining work to do is in completing the porting of monitoring, and testing scaling.
- When BNL PanDA services are complete and validated, proceed with a smooth switch-over from DOMA to BNL PanDA.
  - If postgres completion and validation proceeds quickly, this switch-over could happen before MDC2 (an ideal scenario, but possibly not a realistic one).
- Preliminary timeline
  - ~2 weeks: hello world and real payloads with DOMA
    - run hello-world using prun on DOMA PanDA BNL-OSG queue
    - run real payload on BNL-OSG through direct task submission to PanDA/JEDI
  - 1 month: workflow management for real sPHENIX tasks
    - run actual sPHENIX tasks via iDDS
  - 2 months: production tasks at BNL and remote
    - sPHENIX using prun/pchain for script-based production submission
    - run actual sPHENIX tasks at NERSC
      - using an sPHENIX harvester instance there cloned from the ATLAS one, probably Globus Online for data transport with BNL, and preferably using Rucio

## Distributed computing plans Sep 2021

- See Sep 17 NPPS meeting notes for the distilled version
- Workplan on workflow and workload management
  - PanDA for sPHENIX operating initially on the DOMA instance
    - Work start: Sep 1 2021
    - Initial workflow objective: simulation
    - Ultimate objective: all workflows from raw data processing to data carousel based analysis
  - Transition to BNL/Postgres instance preferably before MDC2 (if doable without compromising MDC2 readiness)
    - This work will involve Misha + core PanDA team
    - Contingent upon SDCC providing resources in response to an sPHENIX request
  - Basis for workflow management: adaptation of existing components/capabilities. Components which may be drawn on:
    - Light version of ATLAS production system
    - iDDS/DAG based workflow support
    - pchain (Tadashi will discuss this on Friday)
  - Initial project duration: 3 months. Deliverables:
    - sPHENIX Production System version 1 according to sPHENIX specs. The system will be ready for MC production in December, and it will be ready for Data Challenge tests in Jan 2022
    - Documentation and code repository
    - Tutorial for sPHENIX (if needed)
  - Physics is aware that we may ask for an extension. The need for an extension will be evaluated in November and possibly again during/following MDC2. Extension deliverables may include

- Working through any delays/extensions of the MDC2 timeline and/or production system implementation
- Implementation of new features based on experience with V1
- Tuning and optimization based on MDC2 performance
- sPHENIX request to SDCC for resources sufficient to migrate from CERN/Oracle DOMA instance to BNL/Postgres instance
  - PanDA/JEDI, iDDS, Harvester, PanDA monitor, CRIC and associated Postgres databases
  - NPPS will handle setting up the services (including Harvester?), we need SDCC to provide the VMs ASAP
- Define sPHENIX specific monitoring needs beyond what PanDA monitor provides
  - Establish workplan and effort - Amol, Tatiana?
  - on Friday Tatiana will describe experience with/customizations for COMPASS
- Identify opportunities for Amol to contribute,
- Rucio role, objectives, requirements for sPHENIX
  - Includes management of distortion correction files; manage them same as data files
  - Rucio timeline: for MDC2, or later?
  - Identify effort to work on it. Covered by SDCC personnel, working closely with NPPS? Direct NPPS involvement?
- sPHENIX request to SDCC for Rucio service support

## SDCC PanDA service VMs request Sep 2021

----- Forwarded message -----

From: **Tadashi Maeno** <tmaeno@cern.ch>

Date: Wed, Sep 8, 2021 at 9:45 AM

Subject: Re: Server request to SDCC for sPHENIX PanDA etc

To: Torre Wenaus <wenaus@gmail.com>

Cc: Chris Pinkenburg <pinkenburg@bnl.gov>, Alexei Klimentov <aak@bnl.gov>

Hi all,

JEDI and the panda server: 2 VMs with 8 cores + 16 GB RAM + 200 GB disk on each

Postgres: 1 VM with 8 cores + 16 GB RAM + 1 TB disk

Harvester+condor: 1 VM with 4 cores + 8 GB RAM + 200 GB disk

Pandamon: 1 VM with 8 cores + 16 GB RAM + 200 GB disk

Tadashi

## New developments for sPHENIX - Sep 3 2021 NPPS meeting

- <https://indico.bnl.gov/event/12550/>
- [Meeting notes](#)

# NPPS tech meeting on conditions DB - Mar 12 2021

- <https://indico.bnl.gov/event/10999/>
- [Meeting notes](#)