Report of the MINERvA Data Management Review Committee

Version 4, April 24, 2014
Mousumi Datta, Mike Diesburg, Stu Fuess, Mike Kirby, Adam Lyon, Emily Maher, Marc Mengel, Andrew Norman, Jaewon Park, Cheryl Patrick, Gabe Perdue, Phil Rodrigues, Heidi Schellman, Saba Sehrish, Erica Snider (chair)

Charge

The MINERvA spokespersons in collaboration with the Scientific Computing Division (SCD) senior management require an understanding of the resources needed to upgrade the MINERvA data management and workflow systems to effectively use the SCD supported services with the intent to provide more robust and reliable operations. The focus of this review is on official offline production activities and on user specific analysis jobs for both data and simulation. In particular this review has examined:

- The Minerva experiment's requirements for a data handling
- The current (as built and deployed) Minerva data management and workflow systems, including:
    - Input and output file characteristics
    - Account management
    - Services, tools and technologies currently deployed for Minerva
- The current operations model
    - Manpower availability from Minerva for these efforts
    - Manpower availability from SCD for these efforts

The review addresses the overall system needs and the dramatic improvements that must be implemented to meet the production processing that Minerva requires to meet experiment milestones. This report addresses the feasibility of a Minerva upgrade given the available resources from the experiment and from SCD and of incremental improvements where full upgrades are not feasible.

The review process

Members of the Minerva collaboration representing the Production Group, the Calibrations Group, the Physics Coordinators, and the Offline Coordinators presented the current details of Minerva's data management workflow.

The Introduction
http://minerva-docdb.fnal.gov/cgi-bin/ShowDocument?docid=9807

Physics Goals and Schedule
http://minerva-docdb.fnal.gov/cgi-bin/ShowDocument?docid=9806

Overview, Detailed Examples, and Review of Currently Used Products
http://minerva-docdb.fnal.gov/cgi-bin/ShowDocument?docid=9797

“Keep-up” Production (reformatting of the binary-raw data)
http://minerva-docdb.fnal.gov/cgi-bin/ShowDocument?docid=9793

Calibrations (deriving the constants)
http://minerva-docdb.fnal.gov/cgi-bin/ShowDocument?docid=9805

Physics Simulations
http://minerva-docdb.fnal.gov/cgi-bin/ShowDocument?docid=9809

MINOS Inputs to Production
http://minerva-docdb.fnal.gov/cgi-bin/ShowDocument?docid=9808

Calibration (applying constants) and Reconstruction Production
http://minerva-docdb.fnal.gov/cgi-bin/ShowDocument?docid=9804

User Analysis Ntuple Production
http://minerva-docdb.fnal.gov/cgi-bin/ShowDocument?docid=9795

End-user Analysis Production (Results Histograms Production)
http://minerva-docdb.fnal.gov/cgi-bin/ShowDocument?docid=9812

The data handling spreadsheet (also contains scripting requirements)
https://docs.google.com/spreadsheets/d/1VlvMfzXAs07b7P3vJo3pJm1onpGG2lDR09fnjhCg0ak/edit?pli=1#gid=0

Findings

(F1) The Minerva experiment has approximately 2.4 FTE of effort devoted to production activities. This effort is spread across 8 individuals. The effort comes from 2 university faculty members, 4 post-doctoral research associates and two graduate students. This effort is split between different tasks including Monte Carlo production, data processing, calibration and other steps in the Minerva analysis chain. Assignment of effort to tasks does not follow a formal plan developed by an offline coordinator or set of coordinating individuals. (R1)

(F2) The timescale for a full reprocessing of the Minerva low energy data is on the order of one month. There is an anticipated a factor of six increase in file size for the medium energy data due to increased cross section and intensity. Additionally, the medium energy run is planned to run longer than the lower energy run. The limiting factor in the speed with which the data can be processed is the I/O associated with retrieval job input and copy-out of job output. CPU is not a limiting factor in the Minerva reprocessing. (R2) (R4)

(F3) The Minerva experiment requires approximately 90-100 TB of storage space to house the output of a special Monte Carlo simulation production with a very short time horizon. This space must be accessible to stages of processing that run after the initial Monte Carlo generation. The Monte Carlo generation needs to be started no later than April 15, 2014 in order to meet the need of the experiment to present results at a Fermilab hosted "Wine & Cheese" seminar on May 5, 2014. (R3) (R4) (R5)

(F4) The Minerva experiment requires an additional 100 TB of storage to hold the analysis of their data for another high priority analysis. This space must be available in time to perform analysis that will be presented in the middle of May, 2014. (R3) (R5)

(F5) The Minerva experiment has a Monte Carlo/Analysis release titled "Titan" which is currently hosted on the Bluearc central disk systems. This release currently occupies 65 TB of storage. This release has also been written to archival storage and is available for retrieval via the dCache/Enstore restoration mechanisms. This release is not considered "essential" for the current physics analysis activities targeted for presentation in May 2014. (R3) (R5)

(F6) As of the time of this review, the Minerva experiment has approximately 32 TB of free space on their BlueArc disk. (R3) (R5)

(F7) The Minerva experiment uses a single instance of a File Transfer Service (FTS) for both their online data acquisition systems and for their offline production. The version of the File Transfer Service currently in use does not support optimizations that have been made to increase compatibility the dCache system. The Minerva experiment does not currently have personnel that are *trained* in the configuration, monitoring and maintenance of their FTS instances, although several members have limited experience writing configurations and with monitoring FTS operation. (R6)

(F8) The Minerva experiment uses at least six different scripts written in at least three different programming languages to control creation of their offline calibration constants. These scripts implement similar functionality which can be summarized as: (R7)

a. selection of specific files to process
b. the creation of an "options" file that controls the configuration of the job

c. a call to start the Gaudi job.

The scripts also manage the interface to job submission and to the SAM catalog. In addition to the calibration scripts, there is a large Python framework of more than one dozen scripts that manage full production. The number of people in the Minerva experiment who are familiar with the full chain of production scripts and workflow, or who are fluent in all the scripting languages used is extremely limited.

(F9) The Minerva experiment has a large amount of data that was taken before the institution of the current dCache shared read/write pool. This data was copied directly to the Enstore tape system for archival storage through the Small File Aggregation (SFA) system. These data need to be accessed from analysis jobs run on Fermigrid. To perform this analysis, the data will need to be restored from the tape archive and promoted directly into the dCache disk pools. The standard restore mechanism is slow due to limitations in the SFA system (R8)

(F10) The SFA restoration process currently in place for transfers between Enstore and dCache has been identified as having performance bottlenecks, which can result in extremely long latencies under conditions of high load or contention for disk/tape mover components. (R8) (R9)

(F11) The Minerva experiment has a race condition in the scripts that manage production file processing, which permits two grid jobs to concurrently process a single data/MC file when an adverse race outcome occurs. This race condition is due to the manner by which the bookkeeping for a production pass is run. Processing a single data/MC file twice results in a conflict during copyback and registration of the file to the SAM data catalog system. (R10)

(F12) The Minerva experiment is currently using the central BlueArc disk services as the source and destination for their production processing. The central disk service can provide a sustained, aggregate bandwidth of approximately 300 MB/s across all clients. In order to prevent large job clusters from overloading the available bandwidth, a maximum of five concurrent file transfers are permitted between the central disk and all Minerva analysis/production jobs. This has the effect of partially serializing large scale production at the data transfer stages. (R2) (R11)

(F13) Minerva has experienced difficulty with creating files with the proper ownership and permissions, and with some users storing very large data sets on the BlueArc disks. The collaboration would like to institute user quotas, but would like to keep the production-role accounts (e.g., minervapro, minervacal, minervadat, etc.) uncapped. (R19)

(F14) Minerva uses a file naming and cataloging scheme that is based on directory hierarchies and meta descriptions of file contents as part of the file name. This naming scheme does not guarantee filename uniqueness across the Minerva namespace. In particular, when end users run jobs to produce analysis ntuples (which are the input to the production of analysis histograms) the output file names are repeated. This collision of file names is incompatible with the SAM data catalog. (R12)

(F15) Minerva sets up `jobsub` in its framework setup script in exactly the same way for all users. The product setup portion of this is not versioned and is integrated identically across software releases and installations, making it difficult to change the script without affecting all users and sometimes leaving the collaboration in a position where the jobsub script can't be updated because of conflicts with versioned instances of the production Python framework.

(F16) Minerva has challenges using SAM for processing that requires or benefits from deterministic file delivery. Analysis job scripts exploit the current system that provides deterministic delivery to encode the input file names into the output file name. This is not practical if there are many, randomly chosen input files. MC jobs overlay multiple data files and MC data. Reproducibility therefore requires that the order of input overlay files be reproducible

(F17) The Minerva experiment may be using direct file I/O to open files hosted from the central BlueArc disk server. This can cause overload of the BlueArc system. (R2) (R13)

(F18) The Minerva experiment uses the older ".py" format descriptions for declaring metadata to the SAM data catalog for some processing stages, but also uses the ".json" format for declaration of newer data and MC. (R14) (R17)

(F19) The production workflow is complicated by several "by-hand" validation checks, all of which are compatible with batch execution on a Fermigrid worker node. Part of the validation being performed requires direct read IO to the files being produced. The validation checks are compatible with or duplicated by the standard operating functions of the File Transfer Service (FTS) or SAM. (R15)

   a. Minerva is performing validation and bookkeeping audits of their large scale production jobs "by-hand". The scripts being used to perform these validations traverse (walk) the directory structures on the storage system to account for missing/failed output files.

(F20) Minerva expressed reluctance to the idea of using SAM to catalog analysis ntuples for two reasons: (1) fear that ntuples produced by users often contain mistakes and are not "good enough" to be cataloged, and (2) because the analysis workflow for histogram production involves walking over the BlueArc area and building a text file list of the ntuples, then using that list to build a TChain for analysis. Users like having this list be stable, and when working in this way, need the files to be easy to find. (R16)

(F21) The Minerva experiment has at least six different group accounts under which data acquisition, software management and production data processing are accessed or run. In particular the following accounts have been identified: (R19)

a. **minervapro** is used for production
b. **minervaana** is used by analyzers (initially) to produce final analysis ntuples (analysis histograms derived from the user ntuples typically are created by the user in interactive ROOT sessions, but if they are run on the grid, they would be run under minervaana).
c. **minervasam**
d. **minervacal**
e. **minervadat**

(F22) The Minerva job that performs the conversion of the binary-raw data to the offline RawDigit format runs very slowly compared to what is expected given the active channel count and complexity of the Minerva hit data. This module has been identified as "DecodeRawEvent" in the offline framework. The module has not been profiled or analyzed to identify potential performance bottlenecks.

(F23) The Minerva experiment has difficulty performing periodic audits of its central BlueArc disk volumes. The difficulty has been identified as being due to the lack of tools with which to calculate the size of the Minerva disk volumes quickly. In particular Minerva has been using the Linux command "du" in a recursive manner to index their volumes. This process is reported to take days to complete.

(F24) Due to the size of overlay files that are copied to the local storage of worker nodes, Minerva may exceed the available space limitations of some worker nodes when copying multiple input files to a single node for processing.

(F25) Minerva needs to archive files of widely varying sizes, ranging from options files (needed to reproduce the job configurations) which are a few kilobytes in size to output data files in the POOL/ROOT format which are on the order of 500 GB in size. The current file families and associated small file aggregation policies have not been optimized to handle this diversity. (R9) (R11)

(F26) The Minerva experiment has difficulties identifying and recording the exact configurations and geometries with which its simulations are generated against. In particular the experiment has identified a situation in which the geometry for the ArgoNeuT detector had been included in their simulation but not recorded correctly. (R22)

Recommendations

(R1) The Minerva experiment should establish an official coordinator position to handle the organization of offline activities. (F1)

a. The offline coordinator should be familiar with the both the Monte Carlo production and data processing activities that the Minerva experiment is engaging in
b. The offline coordinator should understand the effort that is required for the production tasks
c. The offline coordinator should have capability and experience to develop a prioritized work plans for different portions of the offline activities
d. The offline coordinator should have the authority to redirect effort within the collaboration to match the prioritized needs of the offline activities
e. The offline coordinator should have the permissions needed to cancel any Minerva grid job so as to allow enforcement of priorities as needed.

(R2) To meet the long term data volume that the medium energy running will generate and to remove the current I/O bottlenecks associated with copy-in/out of data, the Minerva experiment should move all stages of their large scale production efforts to be compatible with a higher storage capacity system with higher aggregated I/O bandwidth than is currently provided by the central disk services. (F2) (F12) (F17)

a. In particular the experiment should modify, test and verify that all current stages of production and Monte Carlo generation can read, operate with and write files to the dCache systems.
b. Access to dCache should be made through the recommended and supported tools that the Scientific Computing Division has provided.
c. Recommended data paths and protocols should be used which minimize the use Bluearc disk system and the associated use of concurrency locks to access it.
d. The Minerva software and personnel should comply with the required and recommended access patterns to the storage system.

(R3) Combine the existing 32 TB of free space currently available with the approximately 65 TB that would be available after removal of the Titan replica to provide a 90-100 TB storage area on central disk that could be used for Monte Carlo generation during the period during which the Minerva code base is modified and tested for use with dCache. (F3) (F4) (F5) (F6)

(R4) Use approximately 100 TB of existing dCache volatile space (total size 266TB) to temporarily stage the output of the Monte Carlo Generation and perform any required validation of the data prior committing the output to tape-backed storage, (F2) (F3)

a. Write the output of the Monte Carlo generation stage after validation to non-volatile dCache

    b. Write a second copy of the output to the central disk to accommodate stages of production that are not yet compatible with dCache
    c. For stages that do not yet support reads from dCache, expunge the input of the previous stage from central BlueArc after a given production step is completed, so that there is at most one stage of input files on BlueArc disk at any given time, and rotate the files as required to accommodate each incompatible stage).

(R5) Expunge the BlueArc replica of the "Titan" release from the Bluearc system. (F3) (F4) (F5) (F6)

(R6) Modify the FTS configuration as follows: (F7)
    a. Update to Latest FTS to take advantage of dCache interfaces
    b. Institute a break between FTS handling raw data (DAQ) and offline systems
        i. One FTS for DAQ Files
        ii. One or more FTS for Production/Offline File handling
    c. Minerva should fully own their FTS instance for DAQ
    d. Minerva should own/have access to their FTS instances for offline

(R7) Modify the calibration scripts as follows: (F8)
    a. Unify job scripts for calibration
        i. Reduce number and language flavors of scripts
    b. Use common methodologies for editing/creating options files
    c. Factorize the job submission and production layers of scripting. The job submission pieces should then be "productized".

(R8) Perform a one time bulk restore of the low energy data set in a manner similar to what NOvA has done to bypass the standard SFA restore. (F9) (F10)

(R9) Adopt the following procedures to manage large datasets in order to ensure efficient resource utilization: (F10) (F25)
    a. Work directly with the data storage group to enumerate the types and sizes of files that the experiment will store with the Enstore system. Establish file families for classes of files with dramatically different sizes or restore requirements. Continue to adjust the SFA policies, based on new optimizations from the data storage group, for each file family based on these sizes and recommendations from the storage group. Bypass the SFA system for files above the new thresholds established by the storage group to avoid future performance problems with restores from tape
    b. Perform staging requests through the SAM data catalog system or through special arrangements with the data storage group to ensure optimized restore of files from tape and to minimize the number of mounts per physical tape
    c. Reduce contention for disk/tape mover resources through administrative scheduling of tasks. Notify and schedule with designated CS-liaisons and the data storage group any large scale restore or reprocessing efforts prior to their start.

(R10) Adopt the following procedures to manage production job bookkeeping: (F11)
    a. Minerva should do "missing/failed job" bookkeeping using a system that is not susceptible to the identified race condition in their existing

   production scripts. The SAM data handling system is immune to the condition. It should be used where possible.
   b. Remove the race condition that allows multiple jobs to process the same file and produce output files that are not identical. [possibly completed]
   c. Prevent overwrite of files that get multiply processed
      i. Rework the file locking system to use robust bookkeeping patterns

(R11) Let FTS handle file movement and replicas. FTS should also handle the file ownership issues when copying data to its final destination. (F12) (F25)

(R12) Establish a hard file naming convention that ensures filename uniqueness across the entire Minerva namespace. (F14)
   a. Enforce this convention across all production stages and personnel

(R13) Perform the following audits of file access methods: (F17)
   a. Minerva should conduct an audit of their jobs to detect and enumerate all open file descriptors. This can be done through the use of "lsof" or "strace" on a given worker node while it is running a Minerva job.
      i. Remove all instances of direct file access and replace with appropriate access patterns
   b. Minerva should conduct an audit of their jobs to detect and enumerate all direct calls of I/O locking to ensure they are not capable of causing stalls and deadlocks. In particular the experiment should purge all hard coded calls to the deprecated "CPN" utility. Proper I/O throttling and locking patterns should be followed and number of locks the jobs take should be consolidated to take advantage of bulk transfers.

(R14) The data handling group within SCD should provide the following capability: (F18)
   a. Provide an interface for SAMWeb to understand the older .py SAM files for backwards compatibility with existing Minerva production files.
   b. A conversion utility to convert metadata between .py and .json formats should be developed to allow for porting of older code.

(R15) "By-hand" validation should be replaced by checks made on the worker node at the end of a job and by rules for the FTS. (F19)

(R16) Build in the capability to declare analysis ntuples and analysis histograms to SAM with full parentage information, but do not (initially) require its use. Ideally develop scripting tools to declare existing nutple and histogram sets that need to be archived permanently (in support of a paper or important conference, for example) to SAM and copy them to the dCache read/write pools for eventual tape archival. (F20)

(R17) Minerva should ensure that all future stages of production use the .json format for metadata declaration. (F18)

(R18) Minerva should expand the team of experts able to understand and edit the production script infrastructure and reduce the number of scripting languages used. (F8)

(R19) Analysis and production jobs need to be able to obtain separate credentials for these roles that allow the transfer of data owned by these group accounts out of the dCache and central disk systems. (F13) (F17) (F21)
   a. Analysis and production jobs also need to be able to obtain credentials that allow for the writing of data to the dCache/Enstore system.

   b. Jobs run by individual (non-group account) users need to be able to obtain credentials which allow for the reading of files from the dCache/Enstore system and allow for writing data files to the central disk and volatile dCache systems in a way that ensures that production files cannot be overwritten.

(R20) Minerva should utilize the monitoring tools included with the FTS to monitor the migration of data to tape. (F7)

(R21) To improve the ability of Minerva to keep track of the configurations under which its files are processed, the experiment should:
   a. Insert version and build tags into the production executables
   b. Record these tags in the metadata declared to SAM for files processed by these executables.
   c. Record any other configuration specific information in the metadata for the files (i.e. a filed for some identifier for the config) (F26)

(R22) Ensure that the exact geometry with which files are generated is recorded in a manner that it can be used to select or filter data or simulation datasets used for analysis. An example of this procedure is to create a hash of the geometry files used in the simulation, then embedding that hash as a parameter in the metadata for the output files. (F26)

Actions

Corrective Actions

Large scale grid jobs which bypass bandwidth throttling/locking controls on the central disk services can adversely affect the performance of the storage system for all users at the Lab. The Minerva code base and production infrastructure must be modified to:

1. Remove all direct file read/write operations from/to the BlueArc systems. Replace all such I/O operations found with the prescribed pattern of safe file copies to the worker node's local storage, followed by an open of the local file, or a safe write of the file from local disk to central storage, as appropriate.
2. Remove all instances of the use of the deprecated "cpn" locking package and replace them with the equivalent calls to the supported "ifdh" patches.

The use of a Minerva personnel's personal credentials for the operation of *data handling services* should be curtailed. The Minerva experiment must:

1. Obtain proper special use principals for shared production accounts.
2. Generate appropriate service certificates and proxies for production and data transfer activities.
3. Configure the FTS system used by Minerva to use a proper service certificate.

Medium Term

Minerva has a need for a small, specialized set of grid slots for testing and debugging the job submission and production scripting framework. M. Kirby has submitted a ticket to create this special queue and is investigating the option to use nodes dedicated to testing where job duration was limited to a very short time..

The committee

MINERvA collaborators: Mousumi Datta, Emily Maher,, Jaewon Park, Cheryl Patrick, Gabe Perdue, Phil Rodrigues, Heidi Schellman, Erica Snider

SCD members: Mike Diesburg, Stu Fuess, Mike Kirby, Adam Lyon, Marc Mengel, Andrew Norman, Gabe Perdue, Saba Sehrish, Erica Snider