Nuclear physics: The STAR experiment has continued the use of data movement capabilities between its established Tier-1 and Tier-2 centers and namely, between BNL and LBNL (Tier-1), Wayne State University and NPI/ASCR in Prague (two fully functional Tier-2 centers). A new center, the *Korea Institute of Science and Technology* Information (KISTI) have joined the STAR collaboration as a full partnering facility and resource provider in 2008 and activities surrounding the exploitation of this potential have taken a large part of STAR's activity in the 2008/2009 period.

The RHIC run 2009 had been projected to bring to STAR a fully integrated new data acquisition system with data throughput capabilities going from 100 MB/sec reached in 2004 to a 1000 MB/sec. This is the second time in the experiment's lifetime STAR computing has to cope with an order of magnitude growth in data rates. Hence, a threshold in STAR's Physics program was reached where leveraging all resources across available sites is essential to success. Since the resources at KISTI have the potential to absorb up to 20% of the needed cycles for one pass data production in the early 2009, efforts were hence focused to bring the data transfer from BNL to KISTI to a constant 1 Gb/sec data transfer. It was projected (section 3.2 of the STAR computing resource planning[1]) that such a rate would sustain the need up to 2010 after which a maximum of 1.5 Gb/sec would cover for the currently projected Physics program up to 2015. Thanks to the help from ESNet, Kreonet and collaborators at both end institutions this performance was reached (see [2] and [3]). At this point in time, baseline Grid tools are used and the OSG software stack has not been deployed just yet. Later plan will include a fully automated job processing and return of the results using the STAR data transfer tool of choice since 2001, BestMan/SRM (the Berkeley's implementation of SRM server).

Encouraged by the progress on the network tuning for the BNL/KISTI path and driven by the data flood from Run-9 brings and need to consolidate all of our site interconnect bandwidth, the computing team is in the process of re-addressing all of its network data transfer connectivity: especially, the transfer between BNL and NERSC as well as the transfer from BNL to MIT is being actively addressed. MIT has been a silent Tier-2, a site providing resources for local scientist's research and R&D work but not providing resources to the collaboration as a whole. MIT has been active since the work made on Mac/X-Grid reported in 2006, a well spent effort which has evolved in leveraging more standards Linux based resources as well. Data samples are routinely transferred between BNL and MIT. It is noteworthy to mention that the BNL/STAR gatekeepers have all been upgraded and all data transfer services are being re-tuned based on the new topology. Initially planned for the end of 2008, the strengthening of the transfers to/from well established sites was a delayed milestone (6 months) to the benefit of the BNL/KISTI data transfer.

---

[1] The STAR Computing resource plann, STAR Notes CSN0474, http://drupal.star.bnl.gov/STAR/starnotes/public/csn0474

[2] http://www.bnl.gov/rhic/news/011309/story2.asp , From BNL to KISTI: Establishing High Performance Data Transfer From the US to Asia

[3] http://www.lbl.gov/cs/Archive/news042409c.html, ESnet Connects STAR to Asian Collaborators

At Prague / Bulovka, data transfers are also handled using a BeStMan SRM client but in interoperability mode with a *Disk Pool Manager* (DPM) SRM door. Xrootd remains the low-human cost middleware of choice for STAR and its Tier-2 center storage aggregation strategy but sites such as Prague typically rest on components such as DPM, already deployed within the context of other grid projects. Data rates between BNL and Prague, reaching 300 Mb/sec at the moment, are sufficient to sustain the local needs. Local data access in Prague rests on the use of the *STAR Unified Meta-Scheduler* (SUMS) offering users a common interface for job submission. STAR's approach provides a transparent submission interface to both Grid and non-Grid resources and SUMS remains at the heart of STAR's strategy to migrate an entire class of jobs to Grid resources. We would like to mention that analysis of data sets now entirely relies on access to Scalla/Xrootd data aggregation at BNL (since 2006) and DPM/rfio access at Prague (2007/2008). Users make extensive use of SUMS abstraction to seamlessly launch jobs on the respective farms; the same job description works on both farms. STAR has plans to utilize the Prague resources for opportunistic Monte-Carlo event processing by mid to end of 2009.

One activity which will carry both the data management program and the network tuning to fruition is a research activity carried between STAR and the computer science department at Prague. We intended to study and test a multi-site data transfer paradigm, coordinating movement of datasets to and from multiple locations (sources) in an optimal manner that is, using a planner taking into account network and site performance. This project relies on the knowledge of file locations at each sites and a known network data transfer speed as initial parameters (as data is moved, speed can be re-assessed so the system is a self-learning component). The project is beyond initial simulation at this stage and results are more than encouraging and show outstanding gain in the makespan of a plan over a standard peer-to-peer approach for data transfer. Although this activity is remotely related to OSG, we would be using the OSG infrastructure to test our implementation and prototyping which will be under-taken during the end of summer 2009. To this aim, we paid close attention to protocols and concepts used in Fast Data Transfer (FDT) as its streaming approach has non trivial consequence and impact on TCP protocol short-comings.

STAR has continued its use of and helps consolidate the BeStMan/SRM implementation and have engaged in active discussions, steering and integration of the messaging format from the *Center for Enabling Distributed Petascale Science* (CEDPS) Troubleshooting team, in particular targeting the use for BeStMan client/server troubleshooting for faster error and performance anomaly detection and recovery. At the time of this report, tests and a base implementation is underway to pass BestMan based messages using syslog-ng; several problems were already found and bootstrap leading to better and more robust implementations. We believe we would have a case study within months and able to determine if this course of action represents a path forward to distributed message passing. STAR has finished developing its own job tracking and accounting system, a simple approach based on adding tags at each stage of the workflow and collecting the information via recorded database entries and log parsing. The work was presented at the

CHEP 1009 conference[4]. STAR SBIR Tech-X/UCM project, aimed to provide a fully integrated *User Centric Monitoring* toolkit has reached its end-of-funding cycle. The project is being absorbed by STAR personnel at this stage with an aim to deliver a workable monitoring scheme at application level. The library has been used in nightly and regression testing to help further development (mainly scalability, security and integration into Grid context). The knowledge and a working infrastructure based on syslog-ng may very well provide a simple mechanism for merging UCM with CEDPS vision.

STAR grid data processing and job handling operations have continued their progression toward a full Grid-based operation relying on the OSG software stack and the OSG Operation Center issue tracker. The STAR operation and support team has been efficiently addressing issues and stability. Overall the grid infrastructure stability seems to have increased. To date, STAR has however mainly achieved simulated data production on Grid resources. Since reaching a milestone in 2007 and one hand, it has become routine to utilize non-STAR dedicated resources from the OSG for the Monte-Carlo event generation pass and to run the full response simulator chain (requiring the whole STAR framework installed) on STAR's dedicated resources. On the other hand, the relative proportion of processing contributions using non-STAR dedicated resources has been marginal (and mainly only the Fermi-Grid resources reported as used in 2007). This disparity is explainable by the fact that the complete STAR software stack and environment, which is difficult to impossible to recreate on arbitrary grid resources, is necessary for full event reconstruction processing and hence, access to generic and opportunistic resources are simply impractical and not matching the realities and needs of running experiments in Physics production mode. In addition, STAR's science simply cannot suffer the risk of heterogeneous or non-reproducible results due to subtle library or operating system dependencies and the overall workforce involved to ensure seamless results on all platforms exceeds our operational funding profile. Hence, STAR has been a strong advocate for moving toward a model relying on the use of Virtual Machine (see contribution at the OSG booth @ CHEP 2007) and have since closely work, to the extent possible, with the CEDPS Virtualization activity, seeking the benefits of truly opportunistic use of resources by creating a complete pre-packaged environment (with a validated software stack) in which jobs will run. Such approach would allow STAR to run any one of its job workflow (event generation, simulated data reconstruction, embedding, real event reconstruction and even user analysis) while respecting STAR's policies of reproducibility implemented as complete software stack validation. The technology has huge potential in allowing (beyond a mean to reach non-dedicated sites) software provisioning of Tier-2 centers with least workforce to maintain the software stack hence, maximizing the return to investment of Grid technologies. The multitude of combinations and the fast dynamic of changes (OS upgrade and patches) make the reach of the diverse resources available on the OSG, workforce constraining and economically un-viable. This activity reached a world-premiere milestone when STAR made used of the Amazon/EC2 resources, using Nimbus Worskspace service to carry part of its

[4] *Workflow generator and tracking at the rescue of distributed processing. Automating the handling of STAR's Grid production*, Contribution ID 475, CHEP 2009, http://indico.cern.ch/contributionDisplay.py?contribId=475&confId=35523

simulation production and handle a late request (see articles from iSWGTW [5], Newsweek[6], SearchCloudComputing[7] and HPCWire[8]). This was the very first time cloud computing had been used in the HENP field for scientific production work with full confidence in the results. The results were presented during a plenary talk at CHEP 2009 conference where others presented "tests" rather than actual use (Belle Monte-Carlo testing was most interesting as well). We believe this represents a breakthrough and have since, actively engaged in discussions with the OSG management for the inclusion of such technology into the program of work (present or future) of the OpenScience-Grid project.

We would like to remind that since all of STAR's physics results are built on a foundation of comparison and verifications via simulated data, all STAR's physics publications acknowledge the resources provided by the OSG along those provided by other institutions.

[5] Feature - *Clouds make way for STAR to shine*, http://www.isgtw.org/?pid=1001735
[6] *Number Crunching Made Easy* **-** *Cloud computing is making high-end computing readily available to researchers in rich and poor nations alike* http://www.newsweek.com/id/195734
[7] *Nimbus cloud project saves brainiacs' bacon* http://searchcloudcomputing.techtarget.com/news/article/0,289142,sid201_gci1357548,00.html
[8] *Nimbus and Cloud Computing Meet STAR Production Demands* http://www.hpcwire.com/offthewire/Nimbus-and-Cloud-Computing-Meet-STAR-Production-Demands-42354742.html?page=1