Fermi National Accelerator Laboratory
LDRD Project Data Sheet - FY16

Project ID: FNAL-LDRD-2016-032
Project title: Implement open source HEP NoSQL database
Principal investigator: Oliver Gutsche

Project description: (short description and explanation of cutting edge, high-risk, high-potential science or engineering)

As an alternative to current file-based data storage now employed by most High Energy Physics (HEP) experiments, NoSQL databases (such as used by Google, Facebook, and others) offer an alternative that has the potential to speed the overall computational time required to handle large data sets.

Tie to Mission: (explain the project's relevance or anticipated benefits to Fermilab's and DOE's missions)

In order to advance HEP experimental work, new capabilities in handling large data sets are required. Thus, this LDRD project has the potential to open up new capabilities in how experiments are designed and optimized to produce new science results. There is also a DOE Office of Science emphasis on developing advanced computing including handling of big data. If successful, this project will help establish Fermilab as a contributing leader in this expanding area.

Previous year's accomplishments: (as applicable) User-community prototypes were developed for CMS, DES, and MicroBooNE. CMS and DES implemented their use cases. The development of the framework was finalized with some changes to match the community use cases rather than a batch ingestion tool.

Work proposed for current fiscal year and anticipated / desired results:

The project is in the finishing phase with some additional work to support meaningful physics analysis based, in part, upon user's feedback. The framework will be made easier to deploy and use and documentation will be enhanced.

Project funding profile: (costs, budgets, projected budgets, and total)

Prior year(s) costs	FY16	FY17	FY18	FY19 (3 mos)	Total
N/A	46,206	130,272	128,364	46,636	351,478

Project Start Data: 1/01/2016   Total Approved Project funds: $ 395,000