

# An Exploration of SciDB in the Context of Emerging Technologies for Data Stores in Particle Physics and Cosmology

*Thursday, 8 September 2011 14:25 (25 minutes)*

Traditional relational databases have not always been well matched to the needs of data-intensive sciences, but efforts are underway within the database community to attempt to address many of the requirements of large-scale scientific data management. One such effort is the open-source project SciDB. Since its earliest incarnations, SciDB has been designed for scalability in parallel and distributed environments, with a particular emphasis upon native support for array constructs and operations. Such scalability is of course a requirement of any strategy for large-scale scientific data handling, and array constructs are certainly useful in many contexts, but these features alone do not suffice to qualify a database product as an appropriate technology for hosting particle physics or cosmology data. In what constitutes its 1.0 release in June 2011, SciDB has extended its feature set to address additional requirements of scientific data, with support for user-defined types and functions, for data versioning, and more.

This paper describes an evaluation of the capabilities of SciDB for two very different kinds of physics data: event-level metadata records from proton collisions at the Large Hadron Collider, and the output of cosmological simulations run on very-large-scale supercomputers. This evaluation exercises the spectrum of SciDB capabilities in a suite of tests that aim to be representative and realistic, including, for example, definition of four-vector data types and natural operations thereon, and computational queries that match the natural use cases for these data.

**Primary authors:** Dr MALON, David (High Energy Physics Division-Argonne National Laboratory (ANL)); Mr WEINSTEIN, Jack (Argonne National Laboratory); Dr VAN GEMMEREN, Peter (Argonne National Laboratory)

**Presenter:** Dr MALON, David (High Energy Physics Division-Argonne National Laboratory (ANL))

**Session Classification:** Thursday 08th - Computing Technology for Physics Research

**Track Classification:** Track 1: Computing Technology for Physics Research