Scalable Data Middleware for Smart Grids
Jian Yin, Ian Gorton

Motivation
• The future power grids will become highly dynamic.
• Large amount of data from a large number of sensors
• Scalable near real time data middleware is needed
  • High throughput data ingestion
  • Near real time data access
  • Predictable performance

Observations
• Power grid data is structured
• High temporal locality
• Many insertion and deletion operations, less or no modification operations

Our Approach
• Built on block device interfaces
  • Eliminate indirection
  • Efficient
  • Predictable performance
• Compact format for power grid data
  • Application specific compression
• Log structure data storage
• Specialized index to speed up data retrieval
• Aggregate main memory and SSD for index
• Reduce data retrieval to one disk access

Preliminary results
• Ingestion is significantly speeded up because data are generated in temporal order
• Retrieval can be significantly speeded up for common access patterns of power grid applications
• Predictable performance is achieved

SOSP 2011 Poster