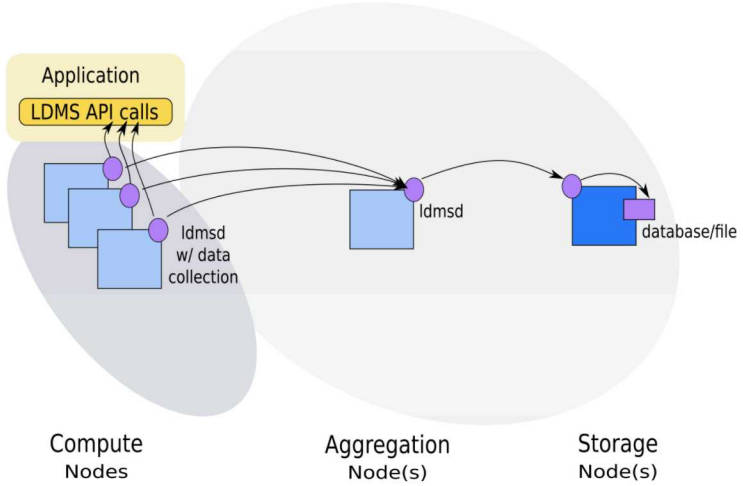
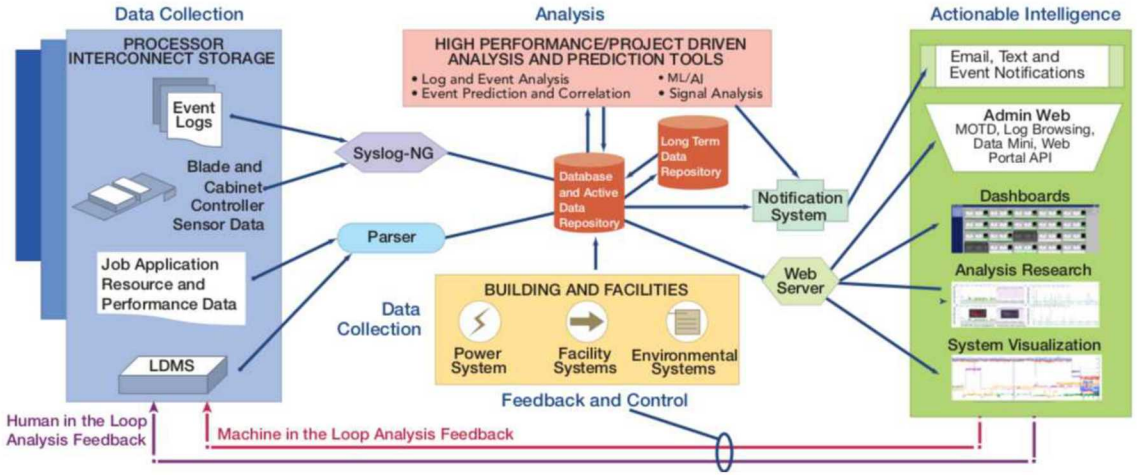
SAND2020-1592C



LDMS



HPC Operational Analytics: Monitoring, Analysis, and Feedback

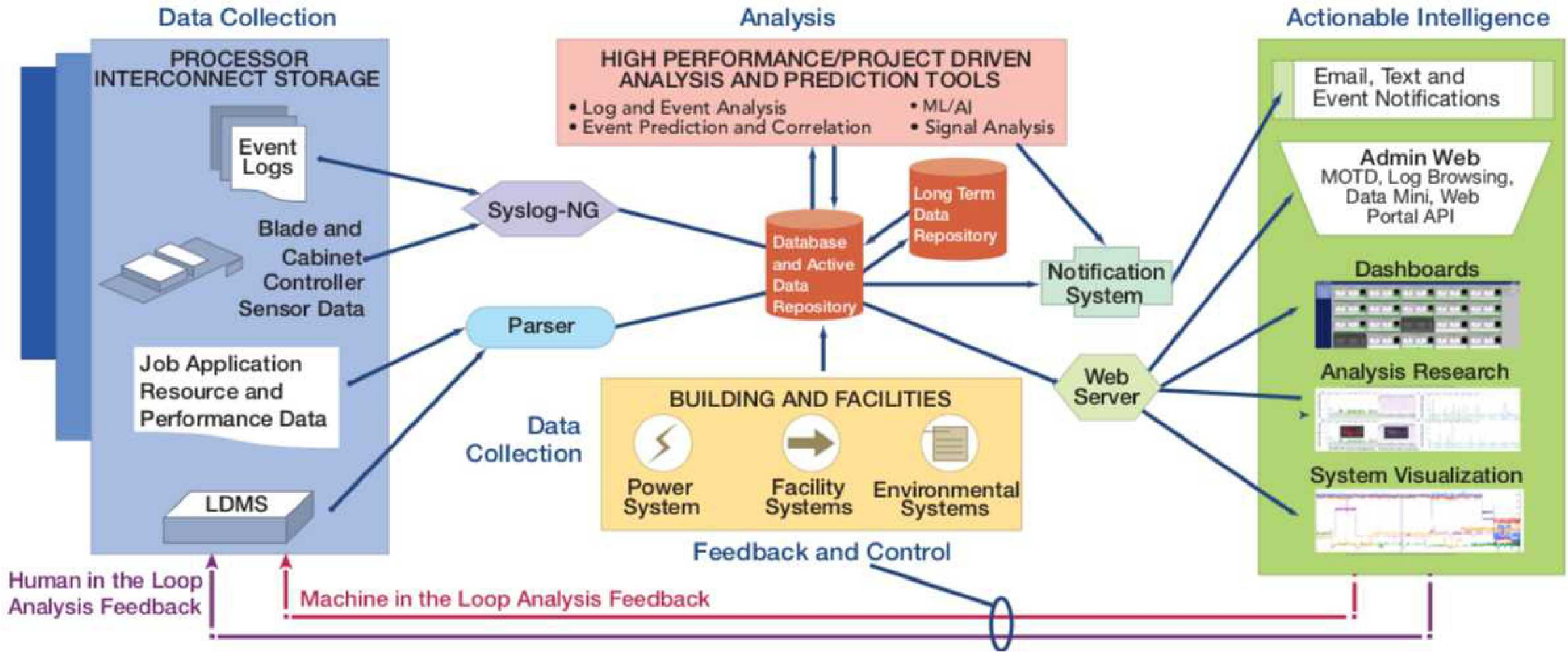
HPC Monitoring & Analysis at Sandia National Laboratories

Jim Brandt

U.S. DEPARTMENT OF ENERGY NNSA

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

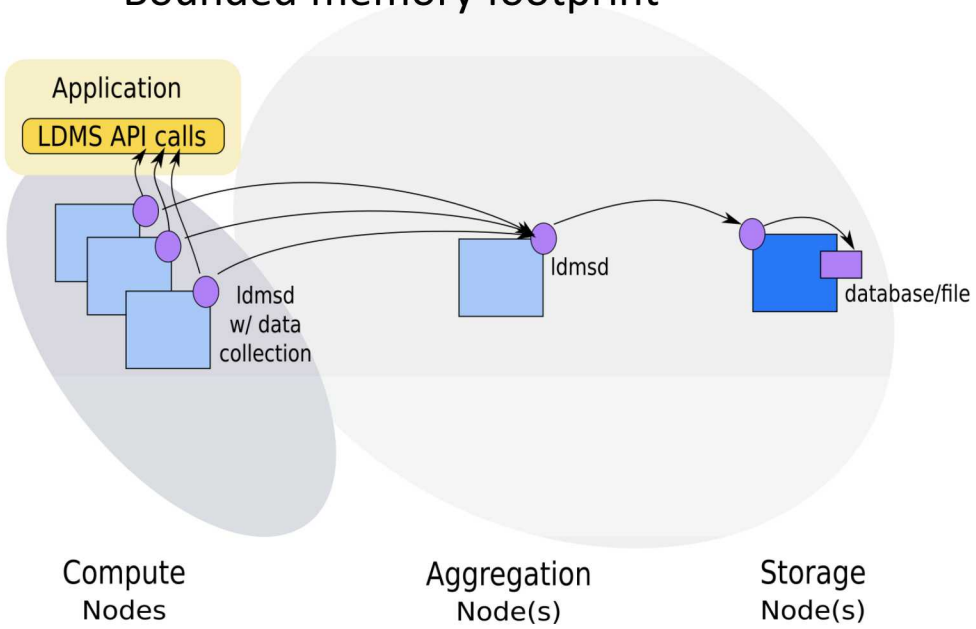
Holistic HPC Monitoring, Analysis, Feedback, and Response:
System Architectural Overview



From: Holistic Measurement Driven System Assessment – *ECP Annual Meeting 2019*

Lightweight Distributed Metric Service (LDMS)

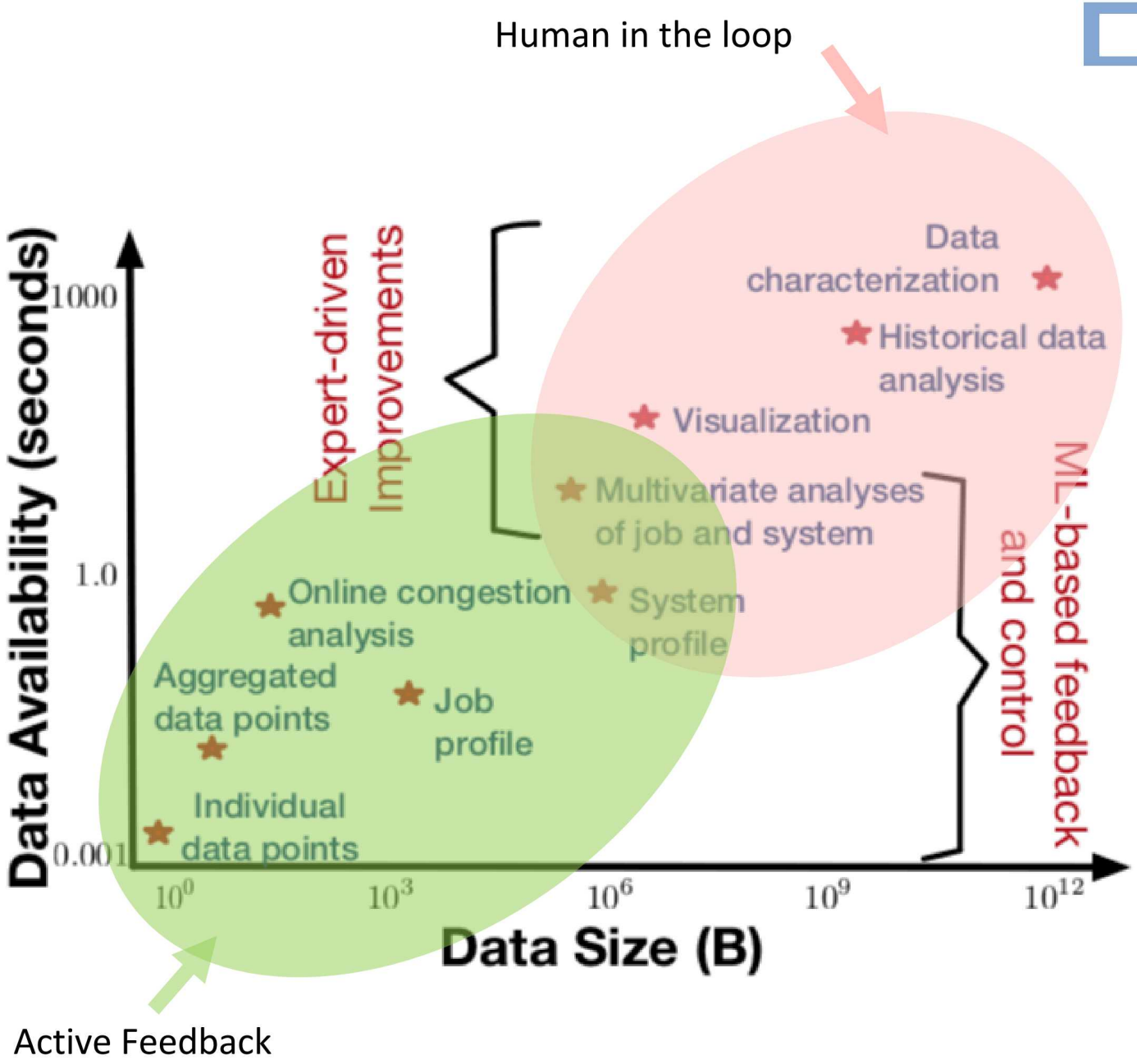
- LDMS – vendor independent data collection, transport, and interface for storage
  - System data collection (network, filesystem, OS counters etc)
  - Application data collection (e.g., HW performance counters)
  - Application-driven data collection (AppInfo, Kokkos profiling data, Application progress FoMs via LDMS Streams)
- Collect data at fidelities necessary for resolving features of interest:
  - Low overhead (~0.1% of a core)
  - High fidelity (down to < 10ms)
  - High fan-in (1000s:1)
  - Bounded memory footprint



Feature	Benefit
Optimized Data Structures	Low Overhead (cpu, network) and Bounded Footprint
Synchronized Data Sampling	System State “Snapshot” for Coherent Analysis
Resilience	Automated Failover & Recovery of Aggregators
Large Array of Transports	Run on IB, Ethernet, OmniPath, Aries, Gemini
Plugin Based	Easy to contribute to: samplers, stores, authentication
RDMA based PUB/SUB	Low Overhead Compared With Kafka, Rabbit
On-the-Fly Sampling Interval Change	High Fidelity On Demand
Sampling and Transport Rate Independent	Low Overhead and Network Resiliency
RDMA & Socket Communication	Lightweight Over Any RDMA Enabled Transport
Job & Application Info. With Data	Simplifies Job Based Analysis
Platform Independent	Run On X86, ARM, Power
Large Fan-in	Minimal Additional Infrastructure
Security	Range of authentication plugins (e.g., munge)

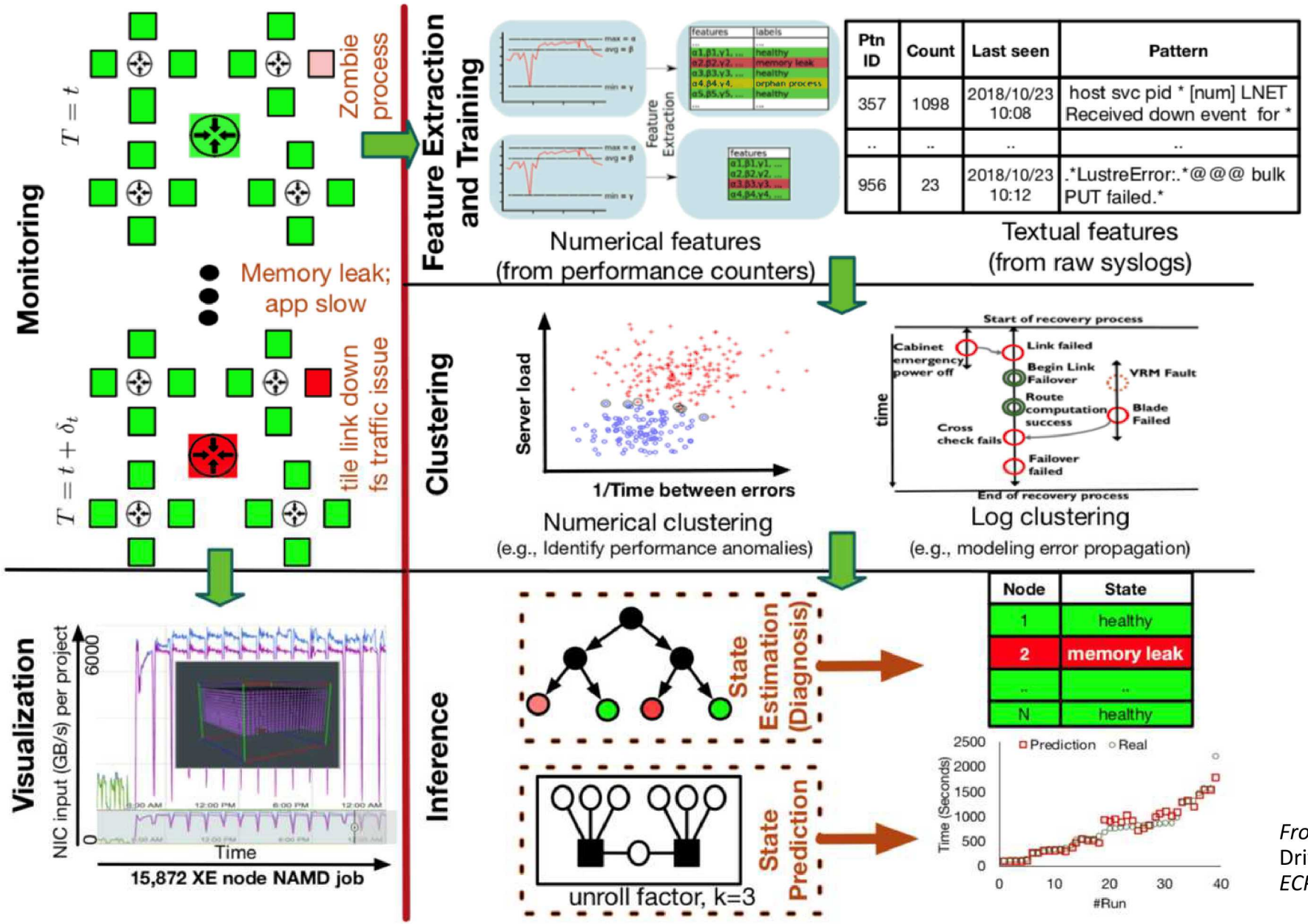
Actionable Data Analytics

- Data features:
  - High volume (~10s of TB/day)
  - High dimension (100s to 1000s of discreet variables) data
  - **Getting** data is *not* a challenge!
- Machine Learning (ML) areas:
  - Validated and explainable ML
  - Clustering
  - Feature extraction
  - Inference



From: Holistic Measurement Driven System Assessment – *ECP Annual Meeting 2019. Analysis Results from NCSA Blue Waters*

Analytics and Feedback (High Level)



From: Holistic Measurement Driven System Assessment – *ECP Annual Meeting 2019*

Activity Highlights

- Development and Deployment of LDMS-based Performance Analysis and Diagnostics:
  - Collaboration with LANL on Trinity to support full scale run-time diagnostics
  - Collaboration with Kokkos, Trilinos, and SNL application code teams to develop run-time capabilities for performance variation attribution (e.g., code changes vs. system conditions)
- Collaboration with HPE/Cray: developing analytics for efficient management and operations
- Open community development of LDMS on Github: https://github.com/ovis-hpc/ovis
  - Bi-weekly user group meetings
  - Annual users group meeting (LDMSCON20XX) with presentations and tutorials