FY 2016 Milestone: Identify materials data management and processing approaches and UQ methodology for the 9Cr steel family of alloys.

PI: Slava Romanov
romanov@netl.doe.gov

Summary:

- NETL data science initiative is a focused effort to support data-intensive research projects that will adapt and develop software tools so that it is possible to curate, archive and analyze large volumes of raw data, with the initial focus on fuel cell materials and advanced alloys, while they are used under extreme environments and power plant cycling conditions. Predictive materials degradation models will be validated against experimental data. Certain analytical models that could be used in simulation and visualization will be designed to account for the features and relationships uncovered using data analytics. The experimental data are gathered as part of the ongoing domain science research at NETL as well as through literature search and collaboration with external organizations. The developed analytical tools and user interfaces will allow the FE experts and technology developers to train the system and to extract hypotheses and promising technological approaches out of simulated and experimental data, in response to their questions and requests, through the integration of guided experimental research with computational sciences and engineering across time and length scales. It will provide the foundations of fundamental scientific understanding for advancing broad areas of science dealing with properties and behaviors of advanced materials and power plant components. Preliminary results of the alloys data analytics pilot [test temperature affecting tensile test outcomes; applied creep stress causing early rupture] are in line with domain knowledge expectations and [skewness of composition vs tensile test data] suggest that cluster analysis may be useful to organize the observed data into meaningful structures.

The project team develops and utilizes unique capabilities in parallel and distributed computing, not only SQL databases, data ingestion, integration and preprocessing (cleaning, parsing, organizing according to the tidy/structured data principles, managing, maintaining, and validating), machine learning, artificial intelligence, fast and scalable exploratory data analysis, statistical inference, feature discovery, mathematical and semi-supervised generalized higher-order system equation modeling to identify relationships and latent variables, advanced expert systems and knowledge-based technologies.

This represents NETL's emerging expertise in data-intensive, domain-guided statistical design for optimal manufacturing, computational process and materials engineering, with uncertainty quantification to support decision-making, and additional scientific insight into complex, noisy, high-dimensional, and high-volume data sets from experiments and simulations.
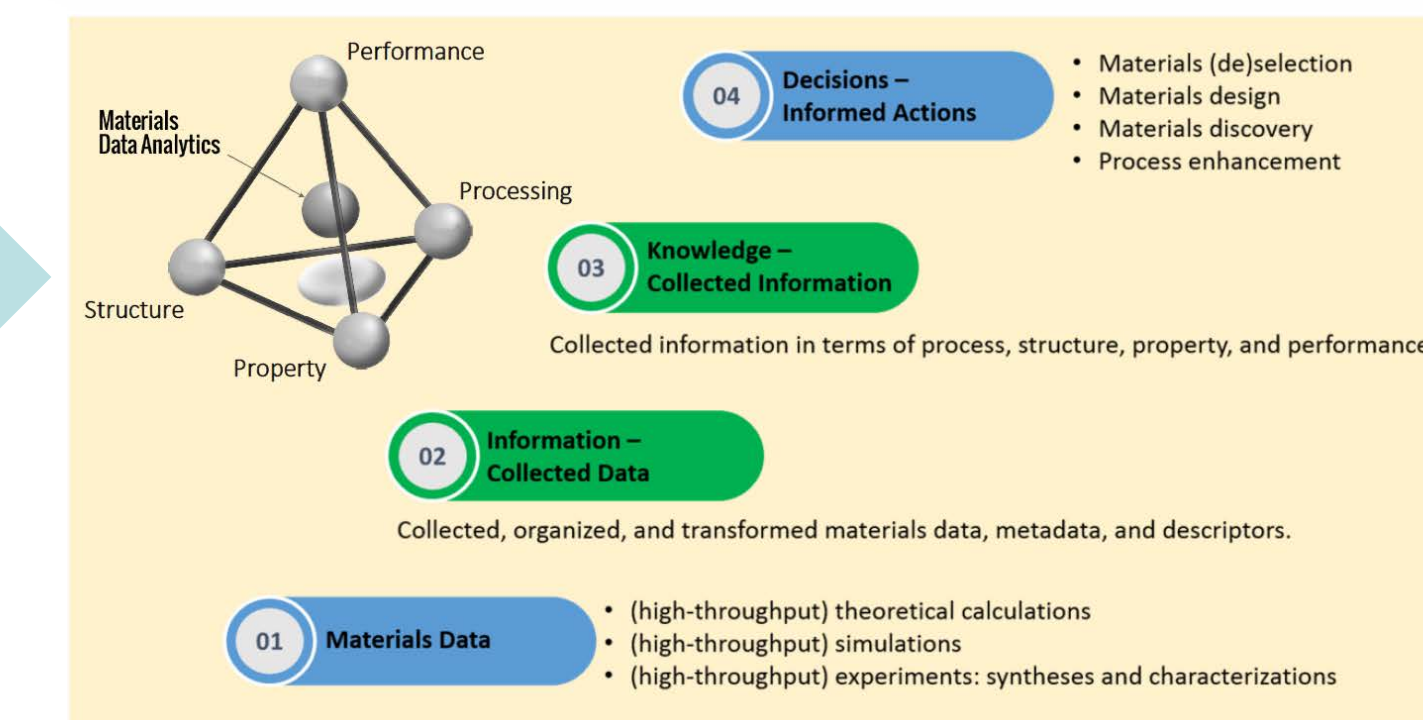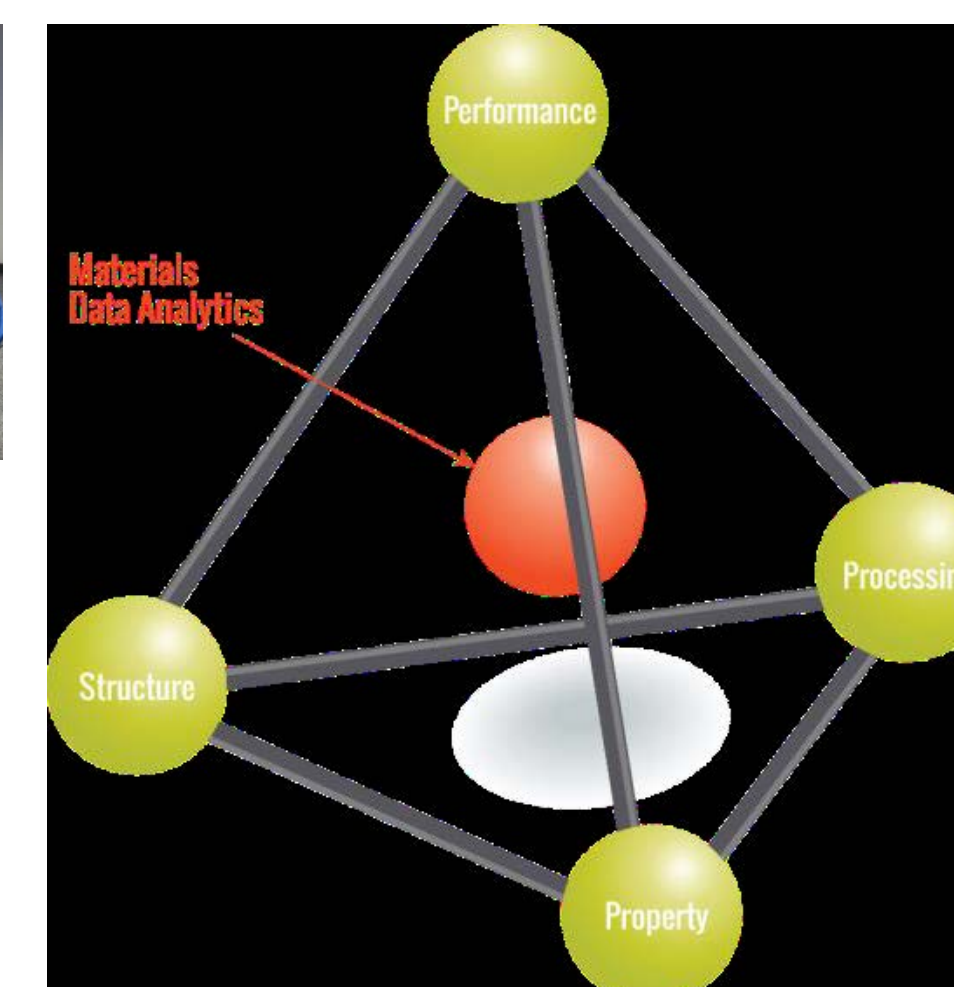

JOULE **Supercomputer (NETL)**

- 24,152 2.6GHz Intel Sandy Bridge cores
- 72,576 GBs of Total RAM
- Full bisection bandwidth QDR Interconnect
- 1 PB of primary storage
- 3 PB of local scratch storage
- Rmax: 503.2 TFLOPs
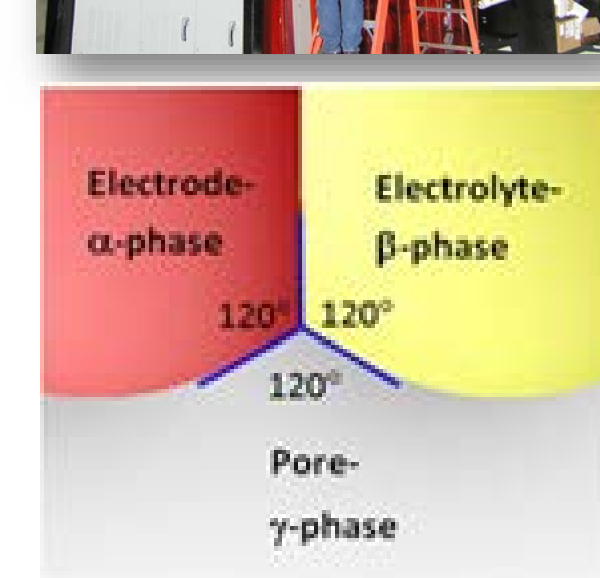- Rpeak: 413.5 TFLOPs
- Refreshment is planned

PILOT PROJECT FUEL CELLS
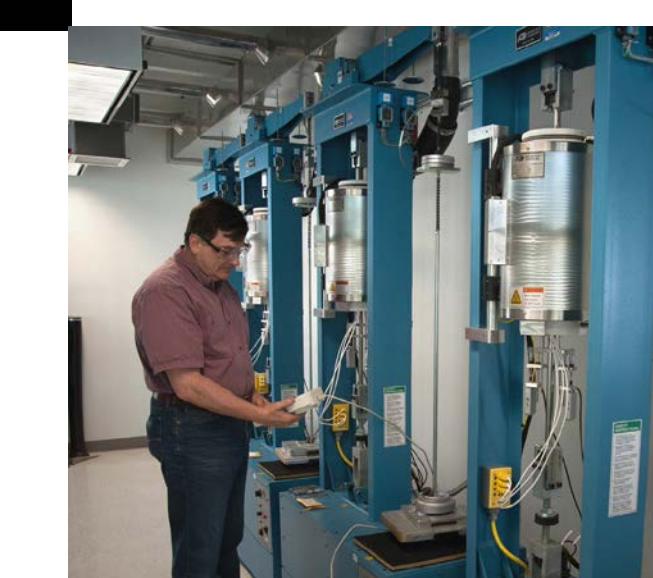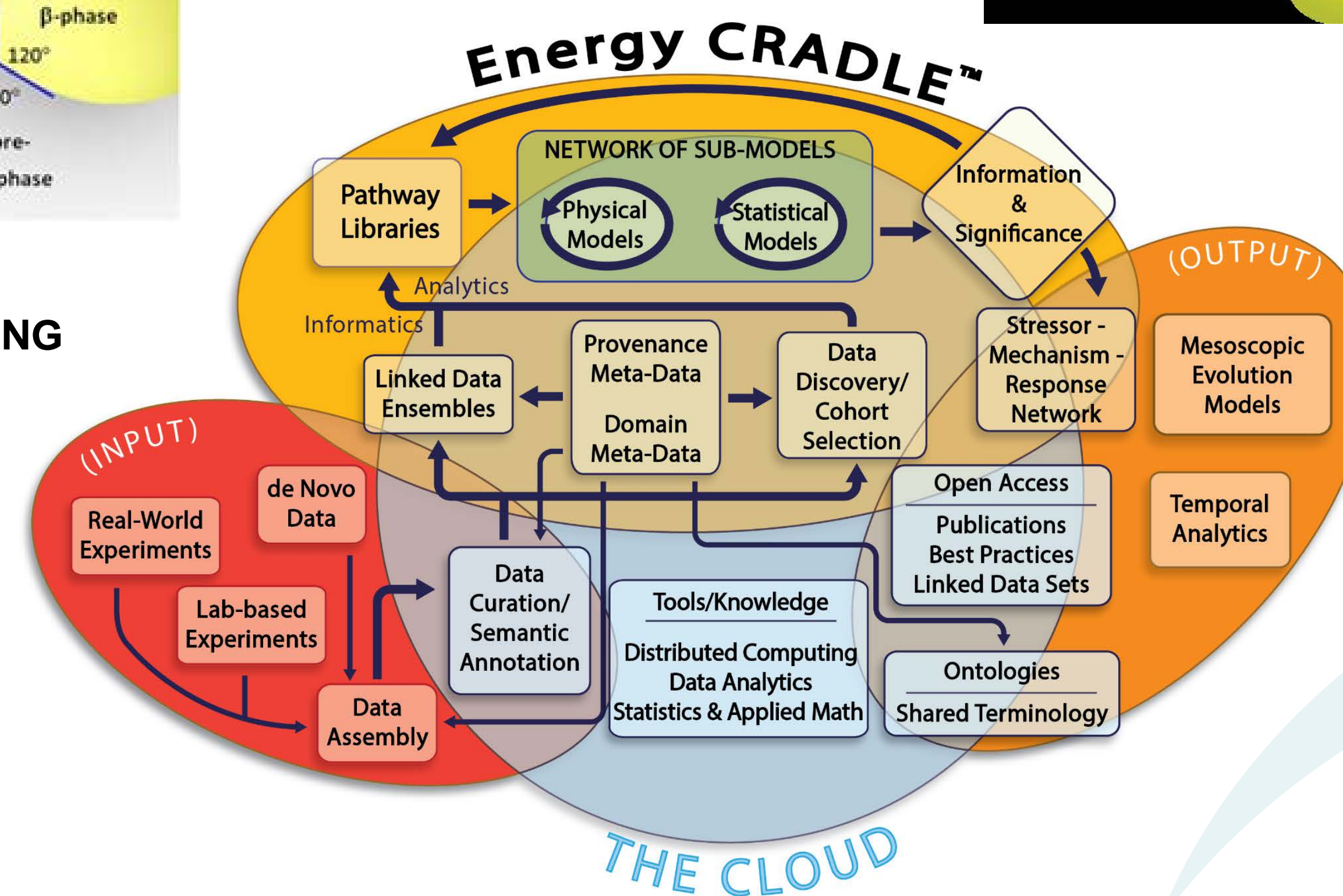
DATA ANALYSIS VISUALIZATION

MODEL VALIDATION

SCALE BRIDGING

Engagement in **MATERIALS GENOME INITIATIVE**

Energy-CRADLE v2.1 Hadoop VM Architecture

PILOT PROJECT RAPID QUALIFICATION of ALLOYS
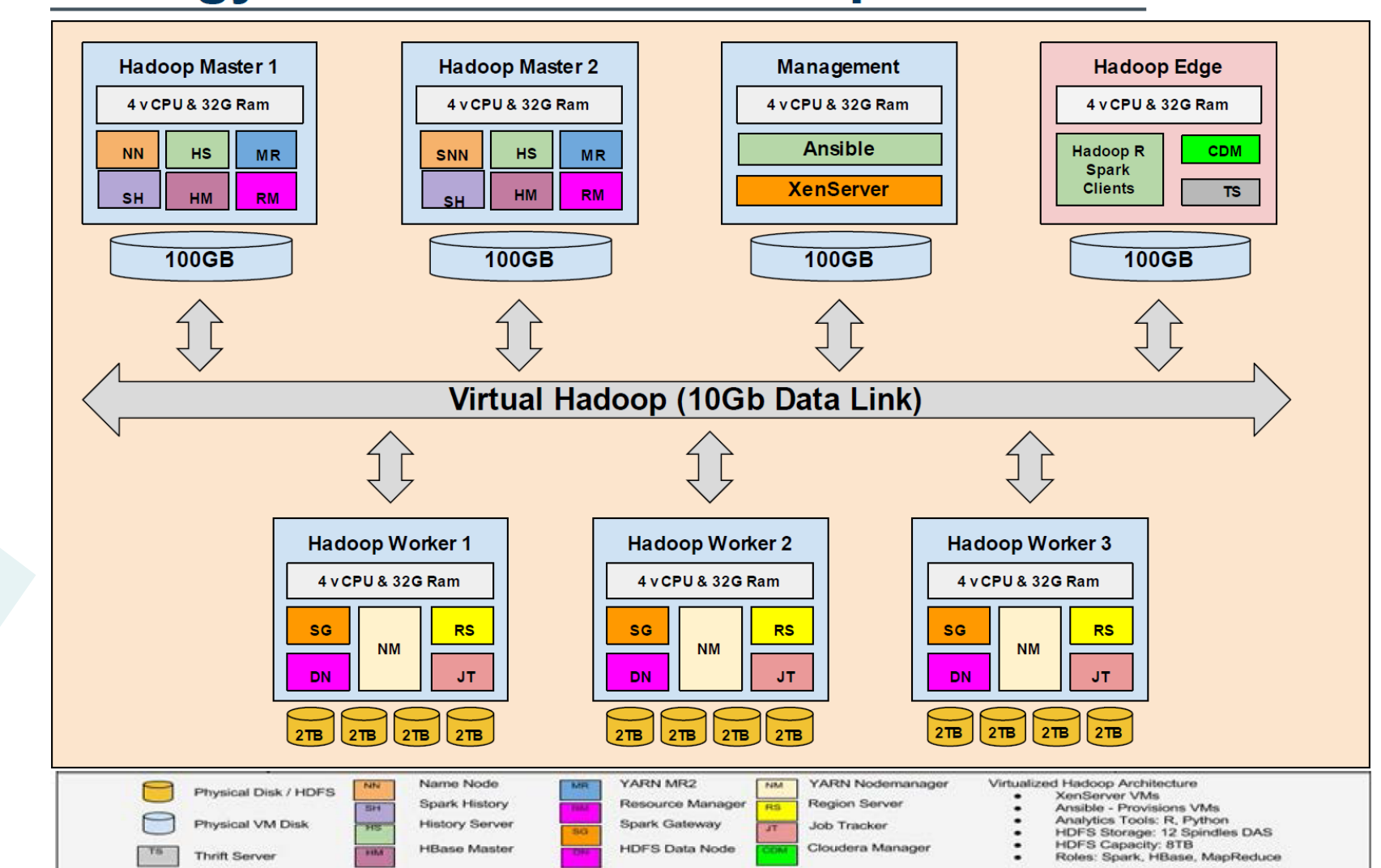
Collaboration with (subcontract award) CASE WESTERN RESERVE UNIVERSITY

CWRU / SDLE Center, VUV-Lab, Materials Science & Engineering Department, Roger H. French © 2012

U.S. DEPARTMENT OF ENERGY