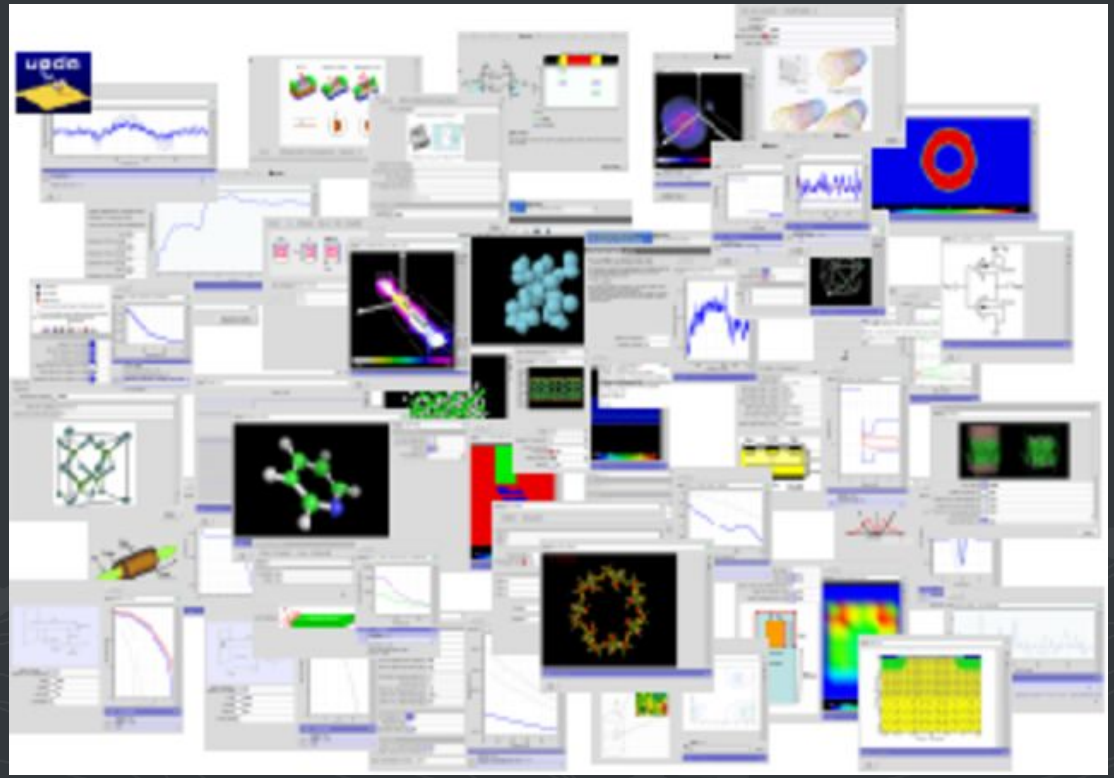
Evolution of the nanoHUB Tool

Steven Clark
nanoHUB.org
HUBzero.org
San Diego Supercomputing Center

Acknowledgement: This material is based upon work supported by the National Science Foundation under Grant 1664084.

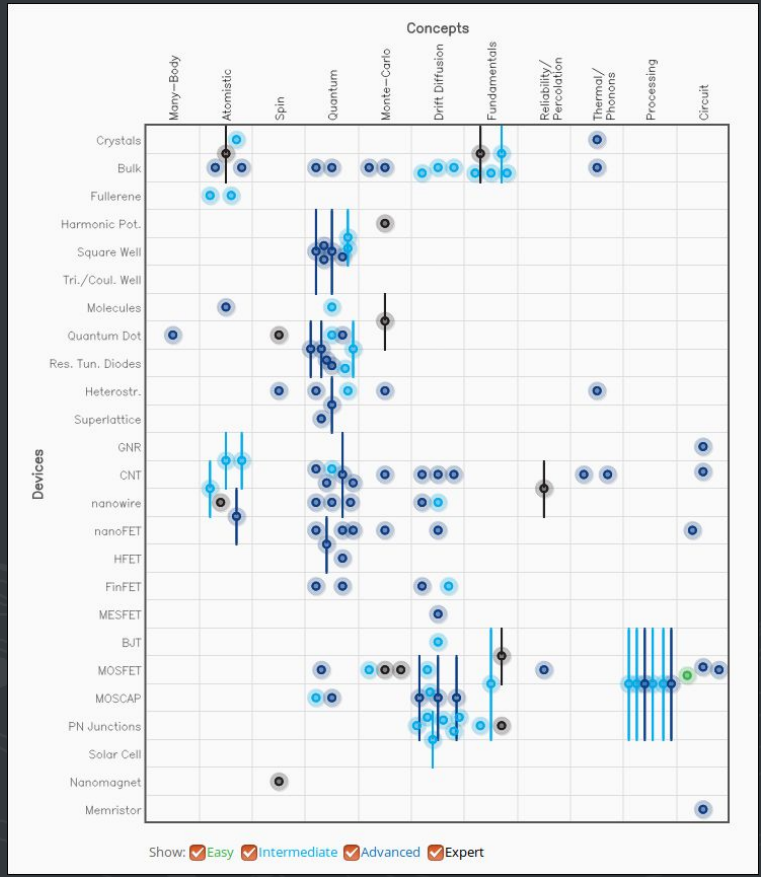
nanoHUB

- Collaborate
- Simulate
- Explore
- Learn



PLATFORM FOR COLLABORATIVE SCIENTIFIC COMPUTATION

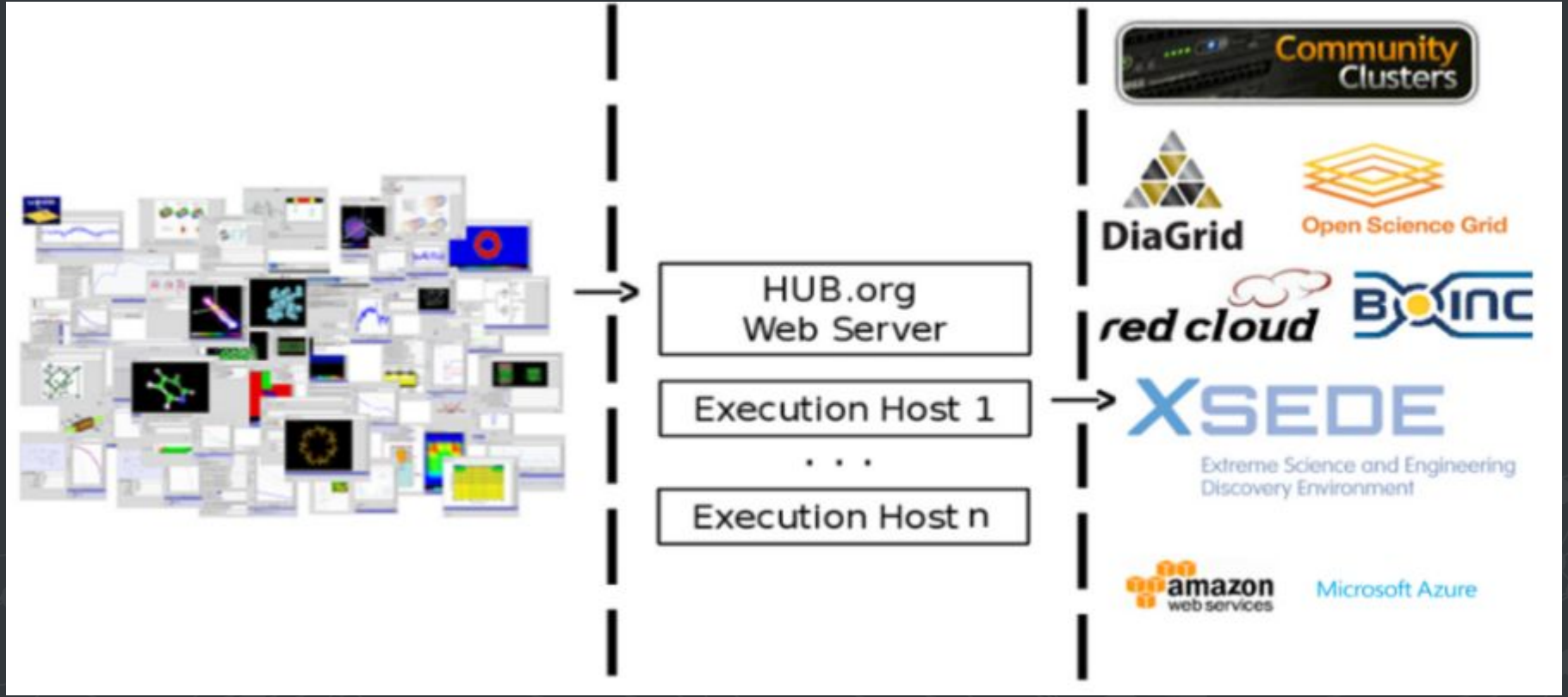
- 1.9 million users annually
- 220,000+ registered users
- 23,000+ simulation users
- 675+ published tools
- Varying degrees of skill required
- Wide range of topics covered
- Diagram highlights tools related to nanoelectronics, concepts vs devices



PLATFORM FOR COLLABORATIVE SCIENTIFIC COMPUTATION

- User perspective
  - Production level code
  - Powerful computing resources
  - No downloading, no compiling, ...
  - Automatically runs most updated version
  - Access regardless of location
- Developers perspective
  - GUI development environments
    - RAPPTURE
    - Jupyter Notebooks
  - Source code management - GIT/SVN/GitHUB
  - Rich development platform
  - Powerful computing resources

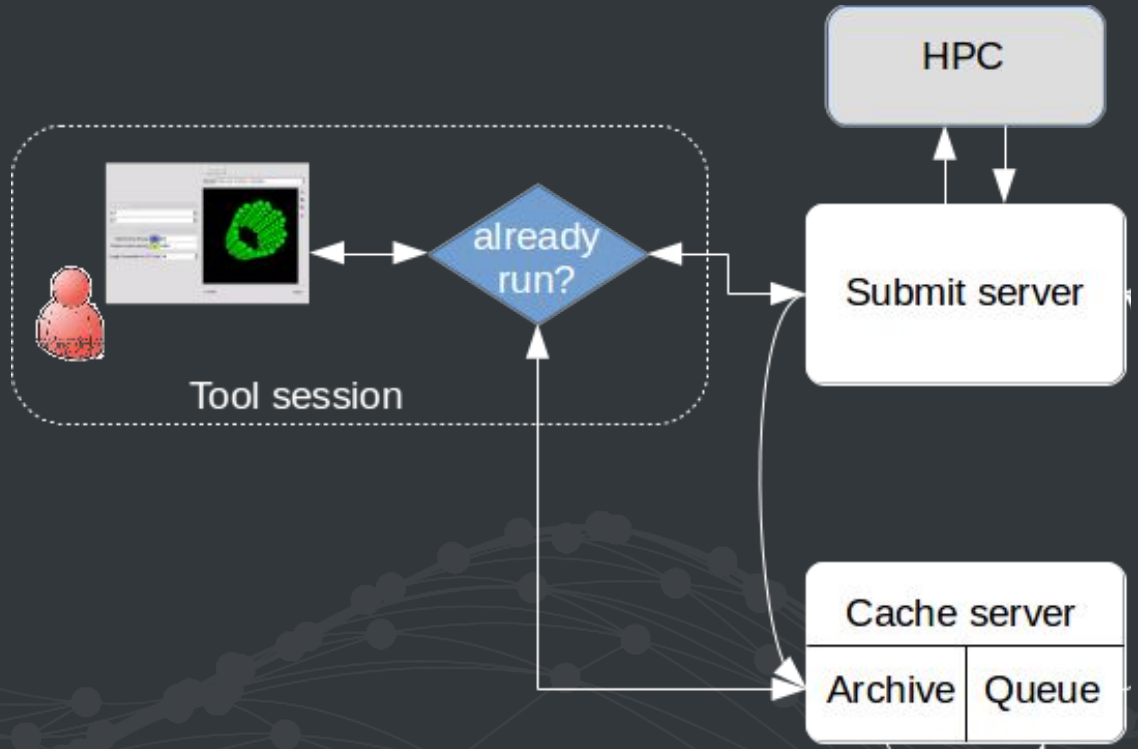
COMPUTATIONAL RESOURCES



SIMULATION USE CASES

- **On demand**
  - UI used to declare inputs for simulation
  - Command line
  - Single simulation or parametric sweep
  - If archive result exists no simulation is required simply pull the existing result
  - Faster response time provides better user experience
  - Generated results are immediately archived
    - Number of hits running slightly behind number of publications

Computation System - Architecture



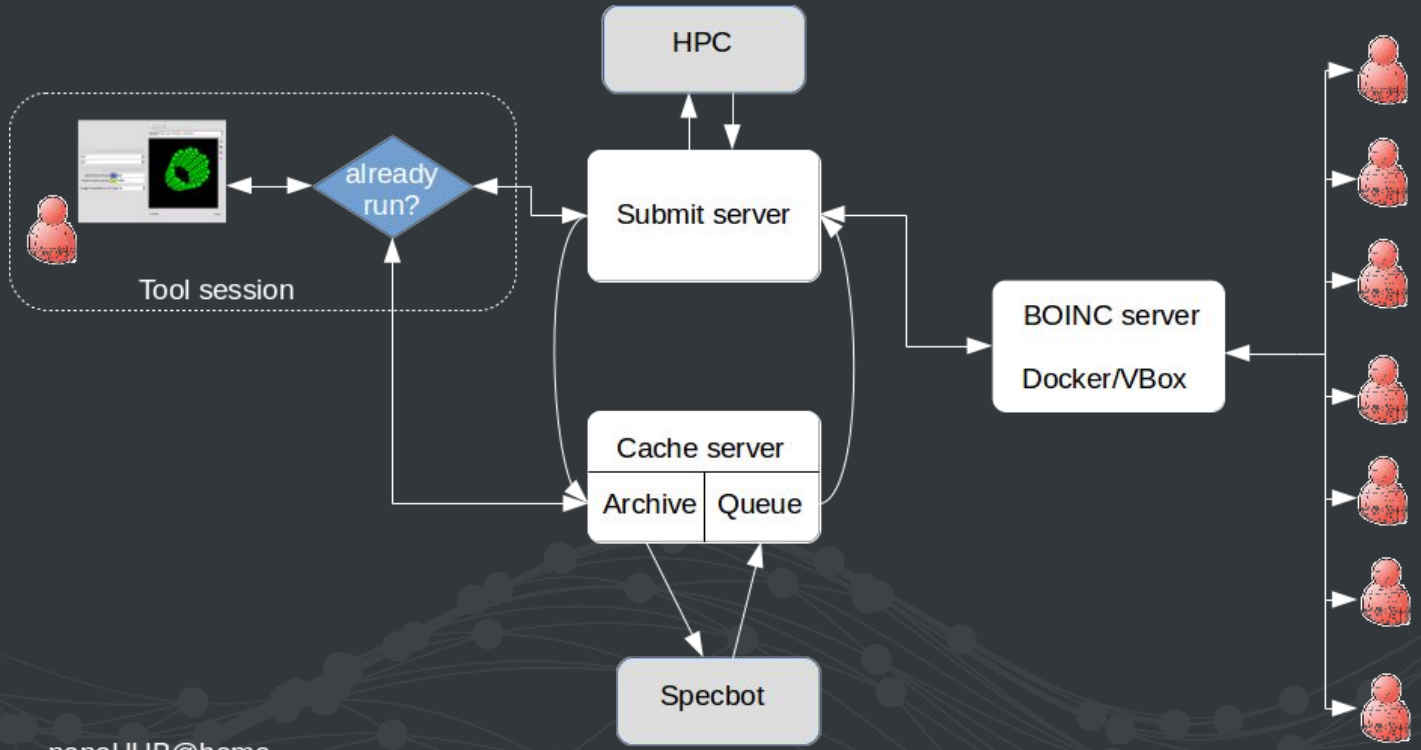
SIMULATION USE CASES

- **Uncertainty quantification**
  - Inputs declared as distributions
  - Statistical methods used to determine input samples
  - A simulation is run for each sample
  - Results include a response surface model which can be used to approximate simulation, sensitivity analysis, and probability distribution function (PDF) for outputs.
- **Exploratory simulation**
  - Automatically generate simulation input samples covering the space
  - Allow for interactive selection of multidimensional input space
  - Execute simulation for each sample

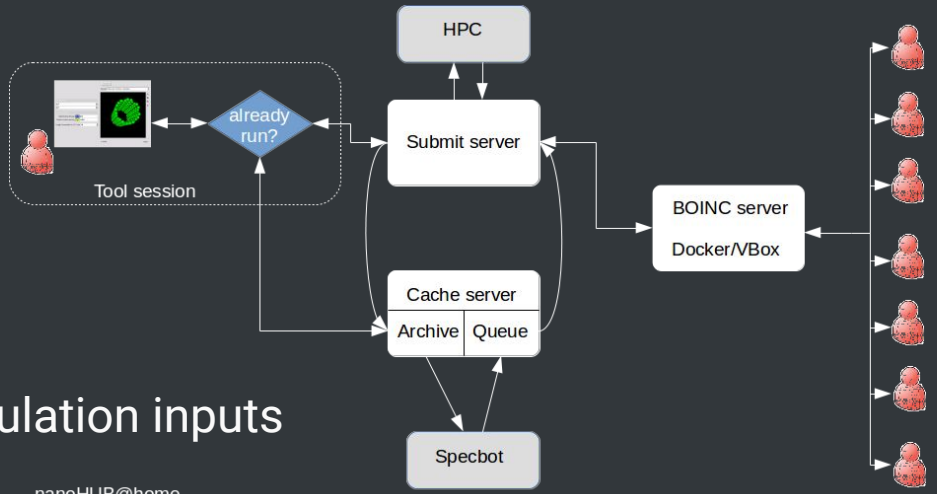
Exploratory Simulation - Motivation

- Probability of exact match with user input is small
- Enhanced visualization techniques
- Machine learning studies
- Surrogate model generation
  - Interpolate between archived results
  - Sensitivity analysis
  - Covariance analysis of results
- Online education services
- All of these require large numbers of simulations
  - The challenge is 10,000,000 simulations/yr

Computation System - Architecture

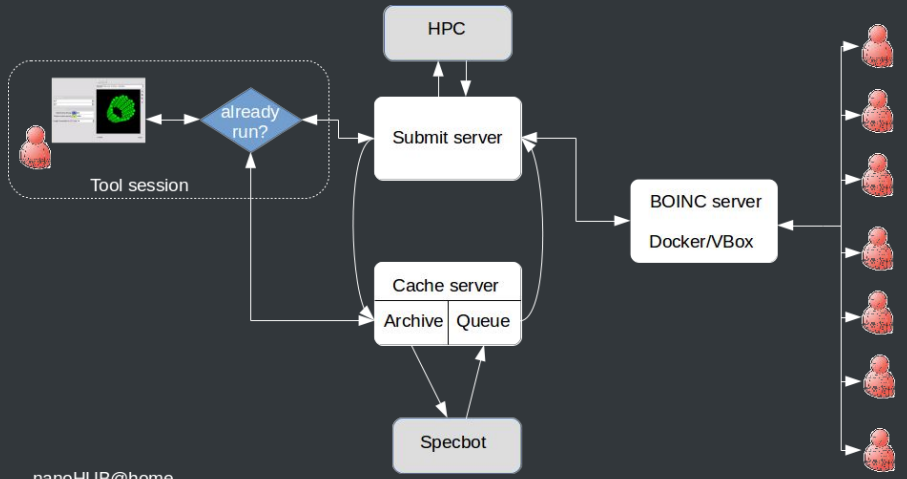


Exploratory Simulation - Generation



- Rappture XML schema defines simulation inputs
  - Scalar numeric inputs
  - Fixed lower and upper bounds
  - Automatically generate simulation input samples randomly distributed in the range
- Input samples are uploaded to cache queue

Exploratory Simulation - Execution

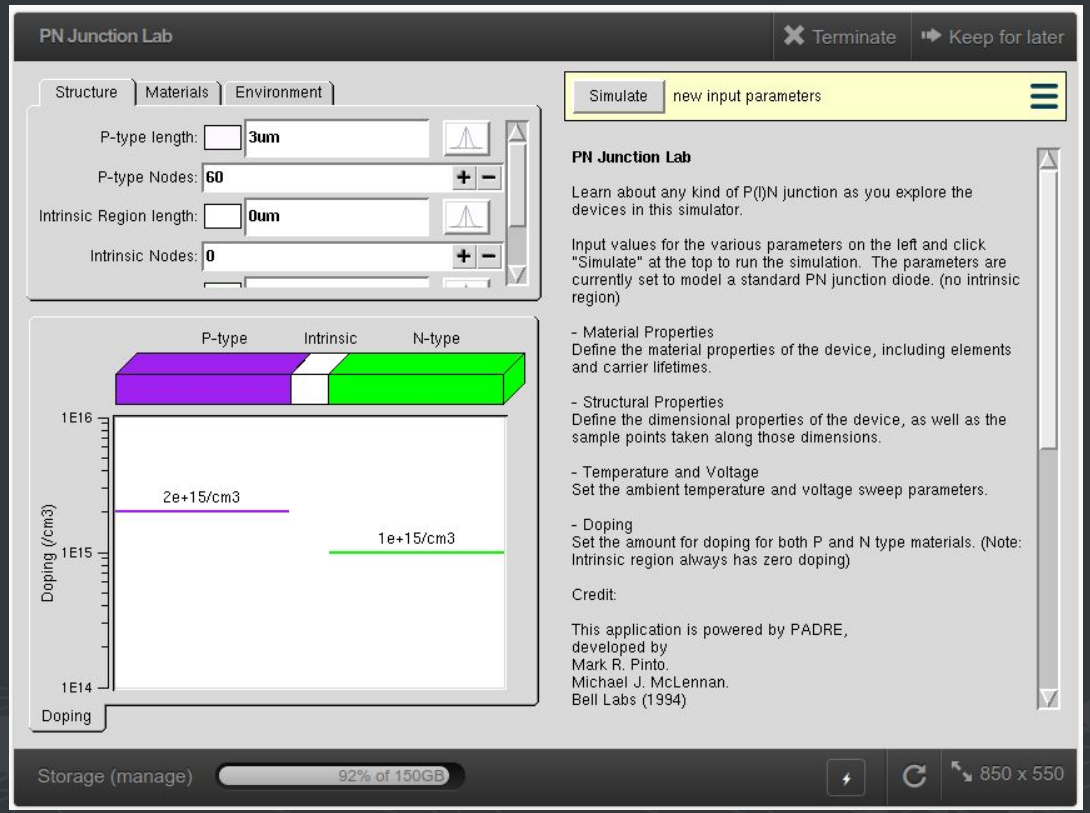


- Automated process pulls driver file inputs from cache queue
- Assemble driver input files into BOINC batches
- Employ submit to execute BOINC batches
- When BOINC batch is complete publish successful results to archive

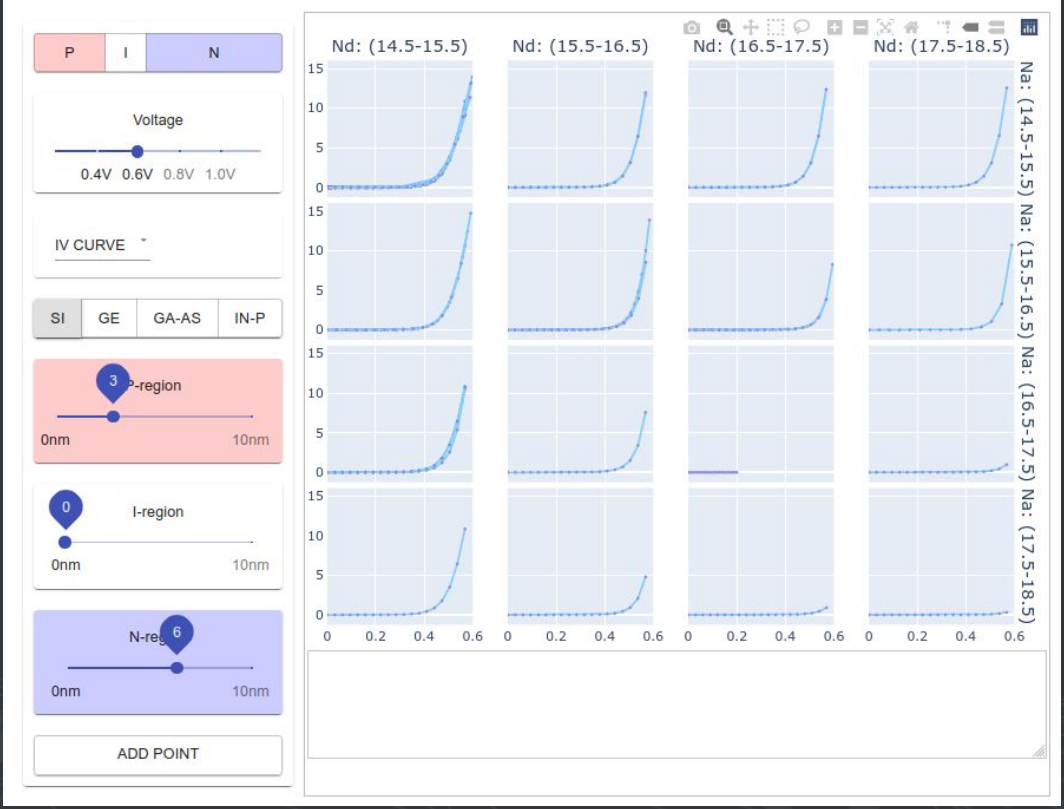
Exploratory Simulation - 10,000,000 jobs/yr ?

- Achieved late 2019
- Interrupted by HUBzero move to SDSC

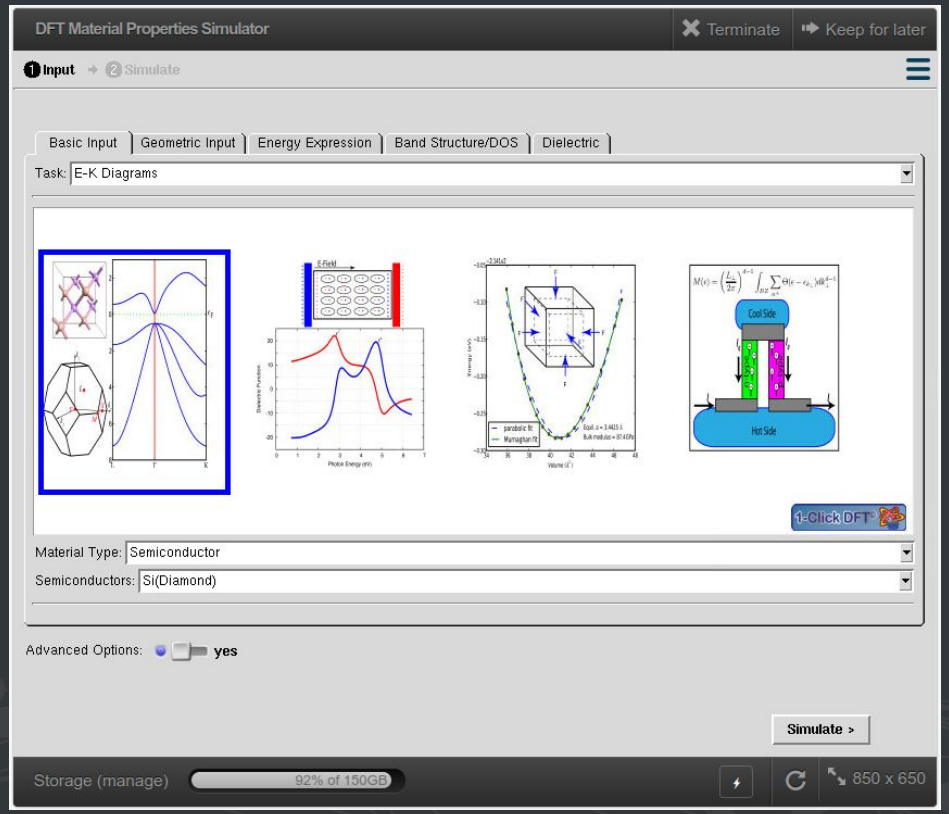
Traditional Simulation



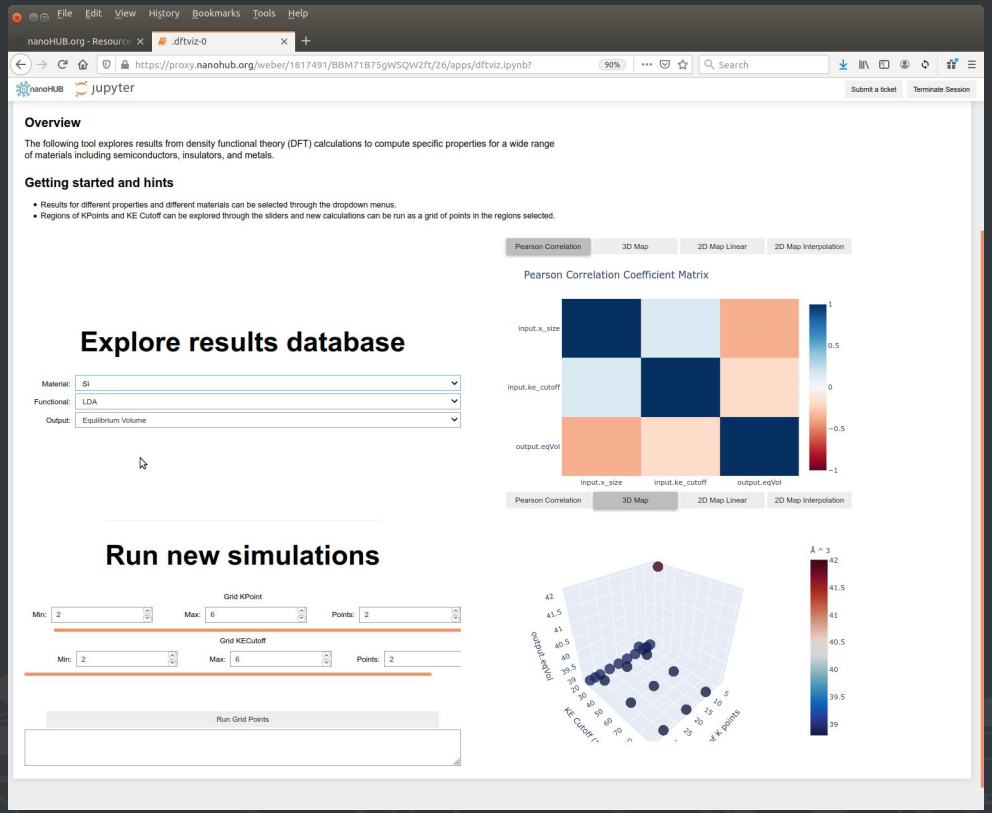
Exploratory Simulation



Exploratory Simulation



Exploratory Simulation



QUESTIONS & ANSWERS