Exceptional service in the national interest

SAND2016-6210PE

Sandia National Laboratories

Enhanced FCT Final Review

ASC L2 Milestone #5672

Provide Enhanced Feature Coverage for Assessment of Application Specific Model Credibility Along with Suitable Tool Verification and Testing

June 1, 2016

SAND2016-XXXXX

U.S. DEPARTMENT OF ENERGY

NNSA National Nuclear Security Administration

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000. SAND NO. 2011-XXXXP

Introductions

Committee:	Allen Robinson (1443) [chair]
	Tom Russo (1355)
	David Lo (1554)
	Dean Dobranich (1514)
	Jay Dike (8259)
Sierra:	Kevin Copps (FCT)
	Kendall Pierson, Jim Cox (SM)
	Paul Wolfenbarger (DevOps)
	Garth Reese (SD)
	Lisa Mondy (T/F)

Introduction

BACKGROUND OF THE FCT

What is the Feature Coverage Tool?

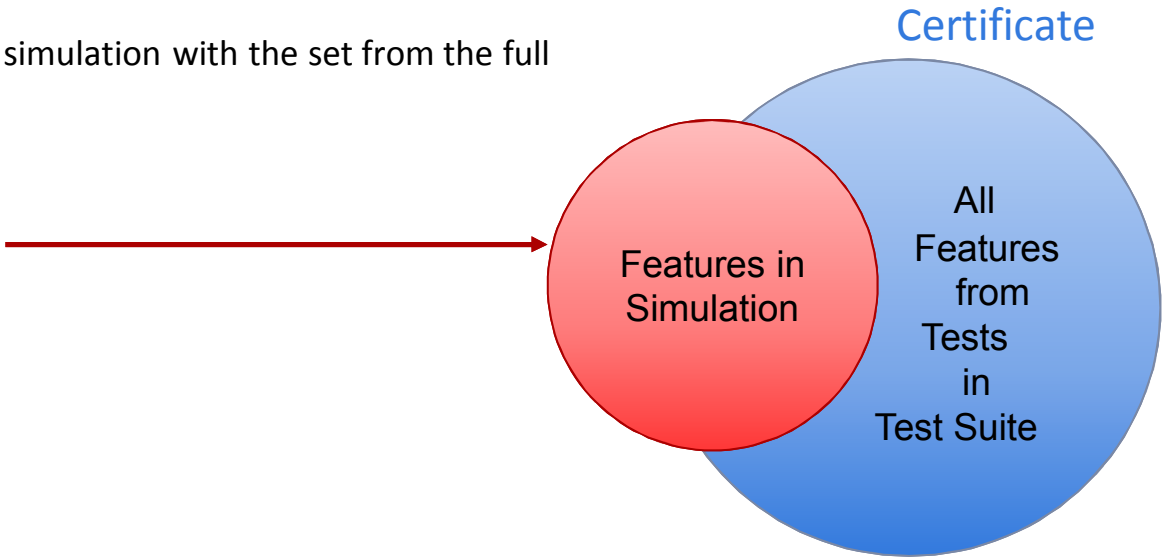
An auxiliary software tool to increase the quality of code simulations by making code testing **relevant** and **transparent.**

- Compiles a *coverage certificate*, containing a summary of all test information tied to a specific version of the code.
- Automatically intersects the features used in a simulation with the features in the test suite.

What is the Feature Coverage Tool?

Intersects the set of features used in a simulation with the set from the full suite of nightly and verification tests.

Analyst's simulation (input file)
annotated by the FCT

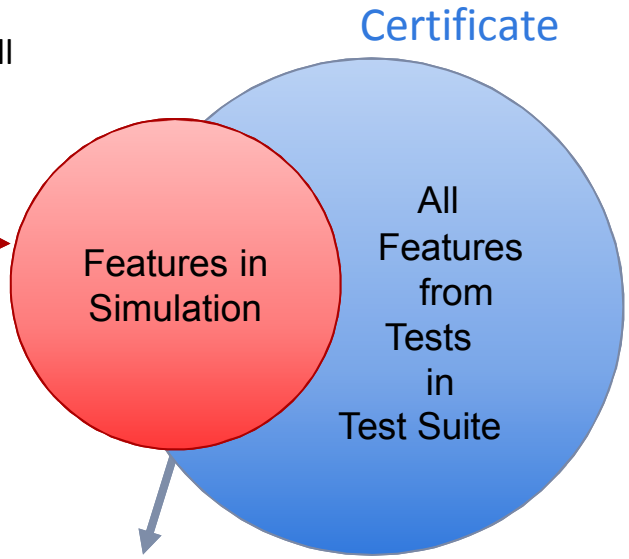


5

What is the Feature Coverage Tool?

Intersects the set of features used in a simulation with the set from the full suite of nightly and verification tests.

Analyst's simulation (input file)
annotated by the FCT



Report percentages and annotate the simulation input on intersection

```
verified

  * one-way: 58%
  * two-way: 27%

tested

  * one-way: 71%

untested

  * ignored
```

6

What is the Feature Coverage Tool?

Intersects the set of features used in a simulation with the set from the full suite of nightly and verification tests.

Analyst's simulation (input file)

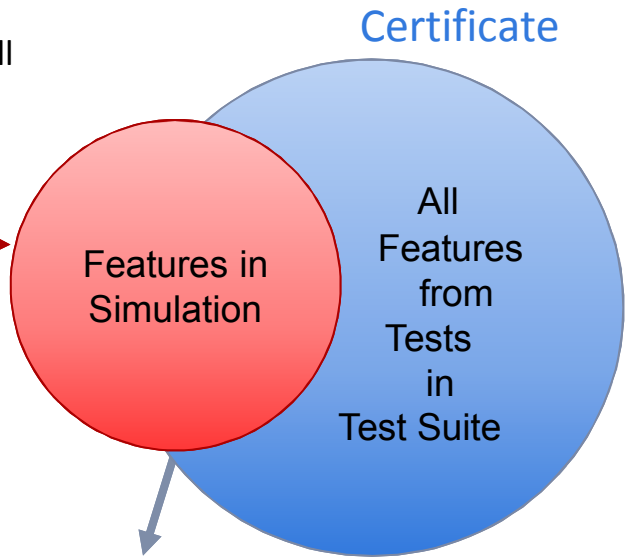
annotated by the FCT

```
begin sierra myJob 1214 +
  # Define gmres linear system solver, this should not need to be changed.
  begin trilinos equation solver gmres 113 +
    param-int AZ_kspace value 1000 4 +
    param-real AZ_ilut_fill value 2 55 + # changed from 5.0
    maximum iterations = 1000 55 +
    solution method = gmres 29 +
    preconditioning method = dd-ilut 7 +
    #matrix scaling = row-sum # added
    preconditioner subdomain overlap = 2 67 +
    residual norm tolerance = 1.0e-9 56 +

  end   begin trilinos equation solver direct_solver 113 +
              solution method = amesos-umfpack 10 +

        end  # Define the mesh and which mesh blocks correspond to which elements.
  Begin Finite Element Model rect 1214 +
    database Name = single_stack_3D.g 1213 +
    coordinate system is cartesian 79 +
    decomposition method = rcb 173 +
    omit block block_4 8 +
    use material cathode_collector for surface_2 33 +
    use material cathode   for block_3 33 +
    use material separator for block_2 33 +
    use material anode     for block_1 33 +

  End Finite Element Model rect  Begin Output Scheduler forResults 18 +
    At step 0, Increment = 1 16 +
```



Report percentages and annotate the simulation input on intersection

```
verified

  * one-way: 58%
  * two-way: 27%

tested

  * one-way: 71%

untested

  * ignored
```


7

What is the Feature Coverage Tool?

An auxiliary software tool to increase the quality of code simulations by making code testing **relevant** and **transparent.**

- Compiles a *coverage certificate*, containing a summary of all test information tied to a specific version of the code.
- Automatically intersects the features used in a simulation with the features in the test suite.
- Exposes the specific and relevant tests in the test suite *in the context of the simulation.*
- Distinguishes between *tested* and *verification* tested by percentages.
- Connect names of verification tests to *documentation*.

What is the Feature Coverage Tool?

An auxiliary software tool to increase the quality of code simulations by making code testing **relevant** and **transparent.**

- Compiles a *coverage certificate*, containing a summary of all test information tied to a specific version of the code.
- Automatically intersects the features used in a simulation with the features in the test suite.
- Exposes the specific and relevant tests in the test suite *in the context of the simulation.*
- Distinguishes between *tested* and *verification* tested.
- Connect names of verification tests to *documentation*.
- Help the analyst *deliver and package evidence* for the credibility of simulations on the area of code verification.
- Help developers see hot spots or *gaps* where testing could improve.

Are there downsides?

- What good is a tool if we don't use it? It better be painless.
- Metrics tend to be gamed, intentionally or unintentionally. Take metrics with a grain of salt and be aware of subtleties.
- Danger of focus on numbers of tests and percentages, not on quality or content, actual relevance, and deeper difficulty of writing good verification tests, and determining if a test is strong and relevant, which is expensive in human time.
- If testing gaps are found, what is the process for analysts to communicate this to developers? A test only shows lack of bugs and solver correctness in a tight closed context.

  "A wise software developer will look both ways when crossing even a one way street."
- If an analyst wants to see details of a test, how to get to it? Can they run the test themselves?

The goals of this milestone try to address some of these issues.

No matter how we good we make the FCT collect information and display it, we must still perform due diligence.

Project Strategy

GOALS OF FCT DEVELOPMENT

Review of Goals of the Milestone

To improve productionize the **SIERRA Feature Coverage Tool**:

- Increase robustness
- Improve usefulness of the results
- Make it easier and faster to use
- Expose categories of tests, distinguish different types and classes of tests—not only just **tested** or **verification tested**.
- Alternatively report coverage in terms of a PIRT (Phenomenon Identification and Ranking Table)
  - Tag and tie input features to physics and algorithm
  - Be able to create reports showing tests by physics and algorithm

Q1

Q2

Q3

Milestone Deliverables

Q1: **Productionize FCT:**

Implement nightly testing of FCT in Sierra. Produce verification tests that check FCT output for correctness.

Q2: **Sierra Test Credibility:**

Work with Sierra development teams to define quality scale and categories that apply to all regression tests. Modify FCT to format test coverage based on these scale/categories. Update verification test suite to tag tests with quality scale/categories.

Q3: **Integration with PIRT:**

Design and prototype an interface to specify features in a model associated with a PIRT. Generate FCT output specify to specified features for at least one Sierra application code.

Q4: **Workflow (Stretch Goal):**

Integrate FCT within SAW to enable analysts to launch basic FCT analyses.

Robustness Improvements

- Issue:
  - Running the FCT seemed to break periodically for various reasons and edge cases: Network issues, version mismatches, not displaying properly.
- Completed:
  - Separate report into two halves: *html* and *Javascript* data/functions.
  - Check static syntax of *html* reports so they pass standardized tests. Uses xmllint tool. Ensures cross browser compatibility.
  - Static syntax checking of *Javascript* data and functions.
  - FCT *Python* code should accord to *Python* PEP8 and *PyLint* standards checking with no warnings.
  - Warnings and error messages should report a best guess at how to resolve the issue.
  - The software tool is broken if the user ever sees a traceback.
  - Avoid pushing changes to released FCT tool until well tested.

Fix Annoyances

- Completed:
  - Gaps (commands that are not tested) can be summarized at top of report, with links to their location in the user input file. (option)
  - Do not touch analyst input files, or change creation date of input file.
  - Do not count repeated command results.
    Material model used 10 times now only counts once to percentage.
  - Fix aliases: Adagio and Presto Region are all Solid Mechanics Region. Without this coverage percentages are artificially low. Aliasing done on backend certificates and analyst inputs.
  - Fix parallel race conditions when creating certificates.
  - Improve error messages. For example,
    `Unable to determine Sierra version. Did you "module load sierra?"`
  - Simplify "`fct --help`". Separate out developer and esoteric options.
  - Question: do we really need to run FCT with every sierra command? For now, we keep it. (option)

More Improvements

- Completed:
  - Sped up FCT analysis execution time: **now half as long as before**.
    We are mostly limited because we call the Sierra code to parse for solid mechanics, aero, thermal cases (~5 seconds).
    Many efficiency improvements, including reading certificate in parallel while parsing the user input file.
  - Removed dependence on slow and cumbersome MS Excel Python.
    Produce smaller 2-way graph using html charting.
  - Reduced report file sizes,
    FCT report is now **10% of previous size on disk**.
    Use indexed lists to map between features and tests.
  - Transparency of test repository, so often requested by analysts:
    Links to test html directory of input files and meshes,
    Work with DevOps: now available on SRN soon on SCN.
  - Fix problems with some tests not included in certificate.
    Tests actually consisting of multiple subtests overwriting each other.
    Latest certificate only reveals 8 tests not included.

Current Status

PROVIDING TEST CREDIBILITY AND CATEGORIZATION

Analyst Interaction with Test Details

- Issue:
  - The FCT report does not indicate much about tests, only directory names and a link to a section in a large single PDF file.
    - Solid Mechanics: 15MB, Structural Dynamics: 18MB, Thermal: 3MB
  - We should provide an overall quick look at the classes of tests. We should deliver more details gradually as the analysts wants to dig deeper.
- Solution:
  - *Work with Sierra SM developers/analysts with a proposal we present now.*
  - FCT will read from test documentation (LaTeX) or from *.test description and create a table that open interactively in the FCT report for each feature.
    - Stored in certificate for each test.
    - Include verification category: a measure of test strictness.
    - Include verification quantities, primary or secondary phenomena.

Analysis Type	Implicit (Adagio) and Explicit (Presto) Dynamics
Element Types	Hex8
Element Formulations	Mean quadrature, Selective deviatoric, Q1P0, Fully-integrated
Strain Incrementations	Strongly Objective and Midpoint
Material Models	Hyperelastic, St. Venant-Kirchoff
Verification Category	Convergence
Verification Quantities	One-dimensional Displacement and Stress fields
Number of Tests	17
Keywords	dynamic, temporal integration, convergence, wave propagation

Test Credibility and Test Hierarchy

*A Sierra Solid Mechanics example of the usefulness of **primary phenomena** vs **type of test**.*

		Conservation	Symmetry	Sanity Check	Code-2-code	Discretization Error	Convergence	Error Quantification
	Totals	29	3	29	8	297	36	72
Contact	121	15	2	21	2	61	20	
Elements	314	9		6	5	206	16	72
Kinematic BCs	9		1	1		7		
Materials	14			1		13		
Loads	13	2			1	10		
MPCs	0							
Time Integrator	3	3						

- Tasks:
  - Define standards for keywords for tiers, categories, primary and secondary phenomena.
  - Work with code team representatives in SM, SD, and T/F to define standards and keywords for a hierarchy of code tests.
  - Add keywords to test definitions in the test suite.
  - Enable FCT reports to report the new type of all tests.

SM Team – Meeting Combined Results

- Preliminary Design Decisions:
  - test classification will be included in the .test file's description block or the LaTex documentation file
  - automated tier rating deemed undesirable
  - additional/alternative test categories should be considered
  - FCT should address two use cases:
    1) Google like search capability to, *e.g.*, find all contact convergence tests
    2) Coverage analysis by primary physics/capability with test classification data
- FCT Developers will:
  - extend data parsing to the .test description block and the LaTex documentation
  - develop capabilities to address both the above use cases
- SM Developers will:
  - include a test description in the .test file or more formal documentation in a LaTex file that provides the test classification
  - examine a process to make individual test documents for simpler linking to FCT

FCT Providing Categories and Test Details

- New Features:
  - Table of test data provided when analyst expands feature in report.
  - Table includes test type category and more.
  - Table sortable on any column.
  - New links to test directories.
  - DEMO table
- Still nice to have, possibly outside the scope of FCT
  - Searchable/sortable page on all tests, results table with same info/directories.
  - Faster loading individual test documents for verification tests.
  - Not all tests have documentation, provide an automatic landing page that is automatically created, showing input file and picture of mesh.
  - Demo test directory page

FCT Providing Categories and Test Details

Demo of the new test information summary in FCT reports.

Links go to the test documentation

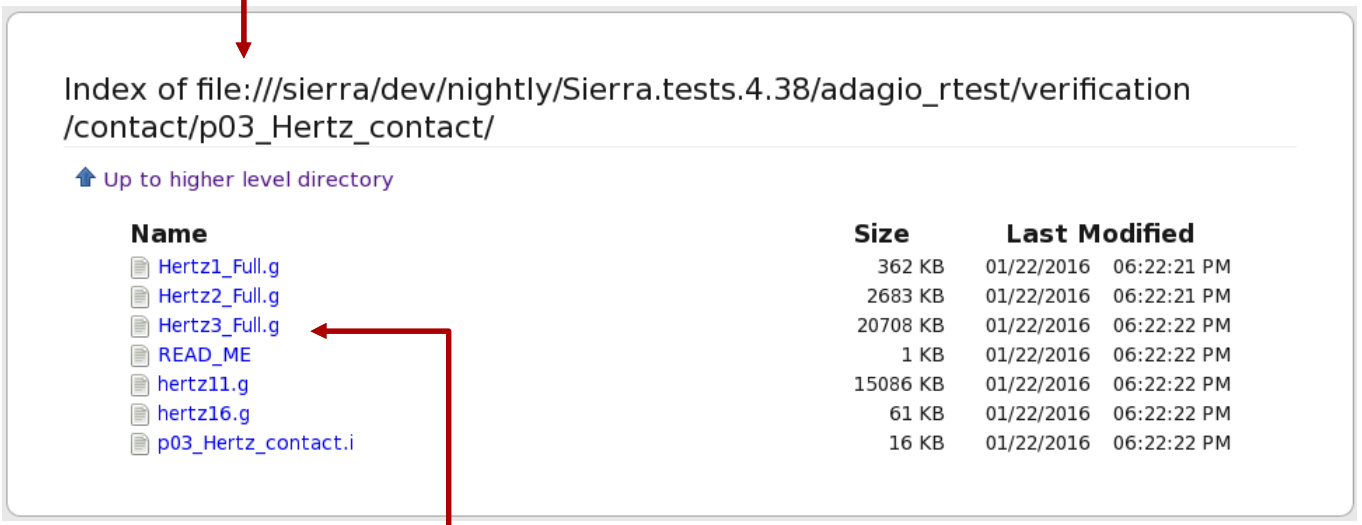
Name	Category	Analysis Type	Element Formulations	Element Type	Strain Incrementation	Material Model	Quantities	Directory
Contact Force Balance	Discretization Error	Quasi-statics		Hex8	Strongly Objective	Elastic	Contact Force	adagio_rtest/.../ContactForceBalance
Hertz Sphere-Sphere Contact		Quasi-statics		Hex8	Strongly Objective	Elastic		adagio_rtest/.../HertzSphereContact
Deriesiewicz Sphere-Sphere Contact	Discretization Error	Quasi-statics		Hex8		Elastic	Torque	adagio_rtest/.../ElasticSpheresInFrictionalTorsionalContact
Hertz Cylinder-Cylinder Contact	Convergence	Quasi-statics	Mean Quadrature, Fully Integrated	Hex8	Strongly Objective	Elastic	Boundary Displacement, Contact Force	adagio_rtest/.../HertzContactOf2Cylinders

Links go to the test webpages (see following slide)

FCT Providing Categories and Test Details

Demo of the web page for all tests in the repository.

Links to each test web page through the network file system work on all sierra platforms



Links to open or download all input files for the test

Physics and Algorithm Coverage

ORGANIZATION BY PHENOMENA (PIRT)

What is the PIRT?

- What is the PIRT?
  - Phenomena Identification and Ranking Table
  - Identifies gaps between technical requirements and models, code capabilities, and V&V activities
  - Focuses limited resources on prioritized activities that will most improve the predictive accuracy

Standard header of the PIRT

Phenomena	Importance	Adequacy			
		Math Model	Code	Validation	Material Data

What is the PIRT?

Example from fire and thermal simulation

PHENOMENA

CODE TEST CREDIBILITY

ID	*Phenomena*	*Imp*	**Model**	**Code**	**Val**	**Data**
A	**Radiative Heat Flux**	H				
A1	Gray emission/absorption, non-scattering	H	H	H	L	
A2	Non-gray, non-homogeneous radiative transport	M	L	L	L	
B	**Radiative Source Terms**	H				
B1	Emission: small scale, ~1cm, turbulent mixing (affects flame area)	H	M	L	L	
B2	Emissive flux: combustion chemistry (affects soot temperature)	H	M	L	L	
B3	Emissive flux: fine scale, ~1mm, turbulent strain (affects soot temperature)	U	L	L	L	
B4	Emissive flux: soot diffusive transport (affects soot mass fraction)	H	L	L	L	
B5	Emissive flux: soot formation chemistry (affects soot mass fraction)	H	M	L	L	
B6	Emissive flux: soot oxidation chemistry (affects soot mass fraction)	H	M	L	L	
B7	Emissive flux: gas emission	M	M	L	L	
B8	Absorption	M	H	L	L	
B9	Scattering	L	H	L	L	
C	**Convective Heat Flux**	H				
C1	Large-scale turbulent mixing affecting flame geometry and radiative view factors	H	M	M	L	
C2	Buoyant turbulence (affects combustion/radiation processes)	H	L	L	L	
C3	Convective heat transfer to the object	M	M	M	L	
C4	Fuel vaporization from surface of spilled fuel	M	M	L	L	

Standardizing Physics/Algorithms

MECHANICS	KEYWORD
Material models	mechanics.material.*
Element patch	mechanics.element.*
Element kinematics	mechanics.element.*
QS Limit for ITD+ETD	mechanics.?.*
Contact	mechanics.contact.*
Coulomb friction limit	mechanics.contact.friction.*
Dynamic Wave Propagation?	mechanics.dynamics.*
Boundary conditions	mechanics.bc.*
Global momentum	mechanics.momentum.*
Global energy	mechanics.energy.*
Rigid bodies	mechanics.contact.rigid-body.*
Hourglass forces	algorithms.material
Time integration	algorithms.time-integration.*
Element death	algorithms.element-death.*
Remeshing with NBTet	algorithms.remeshing.*
Contact?	algorithms.contact.*
Multipoint Constraints	algorithms.contact.mpc.*
Preload and Transfer	algorithms.preload-transfer.*
Solvers	algorithms.solver.*
Output and Postprocessing	algorithms.io.postprocess.*
Contact + elasticity	
NBTet remeshing + material model + …	
Fracture/failure modeling	
Hex, Tet Elements	
Shell Elements	

Input classifications for phenomena.

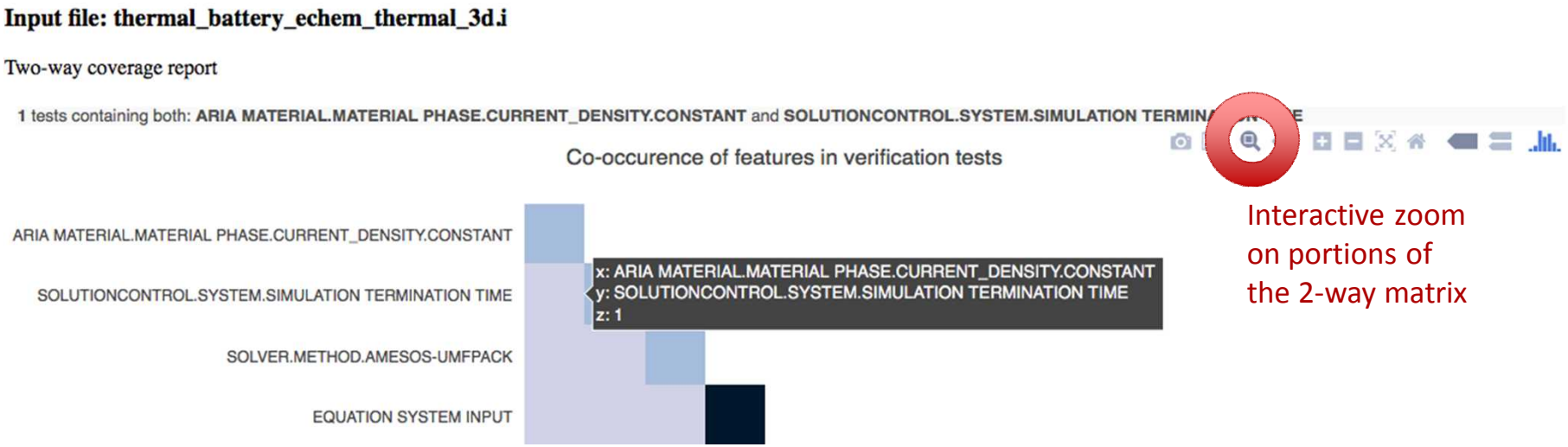
Some algorithms are important, but do not correspond with physical phenomena

Making FCT PIRT Relevant

- Issue:
  - The PIRT uses technical language specific to the physical processes being modeled. Why can't we presenting the FCT results in context with this common terminology. Not just as a user input file?
  - Solvers or other numeric algorithms, or I/O, are relevant but do not fit into PIRT.
  - This physics/algorithm taxonomy should be hierarchical.
- Completed:
  - Ensure a methodology in SIERRA apps to tie each feature to a physics/algorithm taxonomy
  - Process to also write taxonomy into Certificate
  - Offer second type of report organized by physics/algorithm
  - Show two-way coverage in terms of physics/algorithm
- TO DO:
  - Provide beta test of new FCT without breaking nightly certificates
  - Work with Salinas SD team to provide alternative without Sierra parser
- Auxiliary Tasks:
  - Cyclic improvement of standardizing and cataloging taxonomy (between SIERRA code teams, analyst representatives)
  - Maintenance of taxonomy in SIERRA code development teams

Improving two-way coverage

- Issue:
  - 2000 X 2000 sized matrix, which was too difficult to see and include in analysts' report.
  - Previous form in an Excel spreadsheet was cumbersome, and too large to tweak by hand into something suitable for written reports. Writing to Excel spreadsheet was very time consuming in our implementation in Python.
- Solution:
  - The new categorization in terms of physics/phenomena helps because it is a way to collapse features down into a manageable number of categories to be shown in the 2-way coverage
  - A new interactive heatmap matrix was developed that is shown in the FCT browser report



Interactive zoom on portions of the 2-way matrix

Summary

- Robustness, speed and user experience improvements.
  - Useful command line help, useful error messages
  - Runs faster, reports take up less disk space
  - Summary of gaps at top of report
  - Improved coverage accuracy from proper aliasing
- Transparency of test repository
  - Revealed types and primary purposes of tests in new interactive table
  - Instead of just seeing name of tests, analysts may dive on details of any test, including mesh, input
- Physics and Algorithms
  - Completed parsing of taxonomy
  - Producing preliminary 1-way and 2-way reports