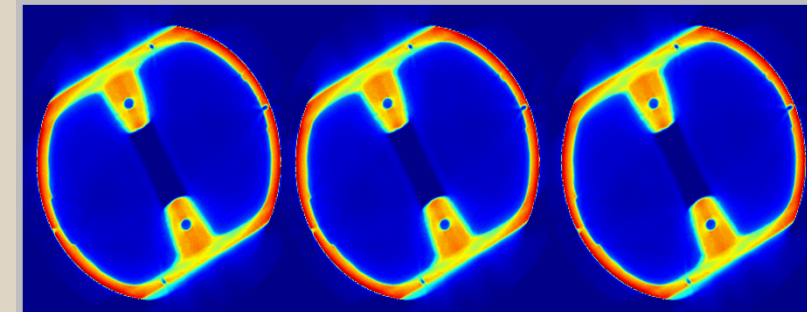
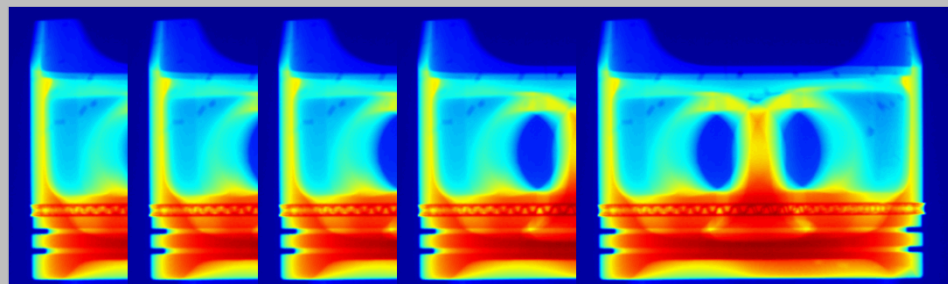


Exceptional service in the national interest



Computed Tomography: New Technology Applied to New Problems

Dr. Edward S. Jimenez
Sandia National Laboratories
Software Systems R&D

Before We Begin...

- I'd like to thank my team:
 - Kyle R. Thompson
 - Elizabeth Lopez
 - John Korbin
 - Stephen Attaway
 - Noelle Collins
 - Madison Devonshire
 - Erica Holswade

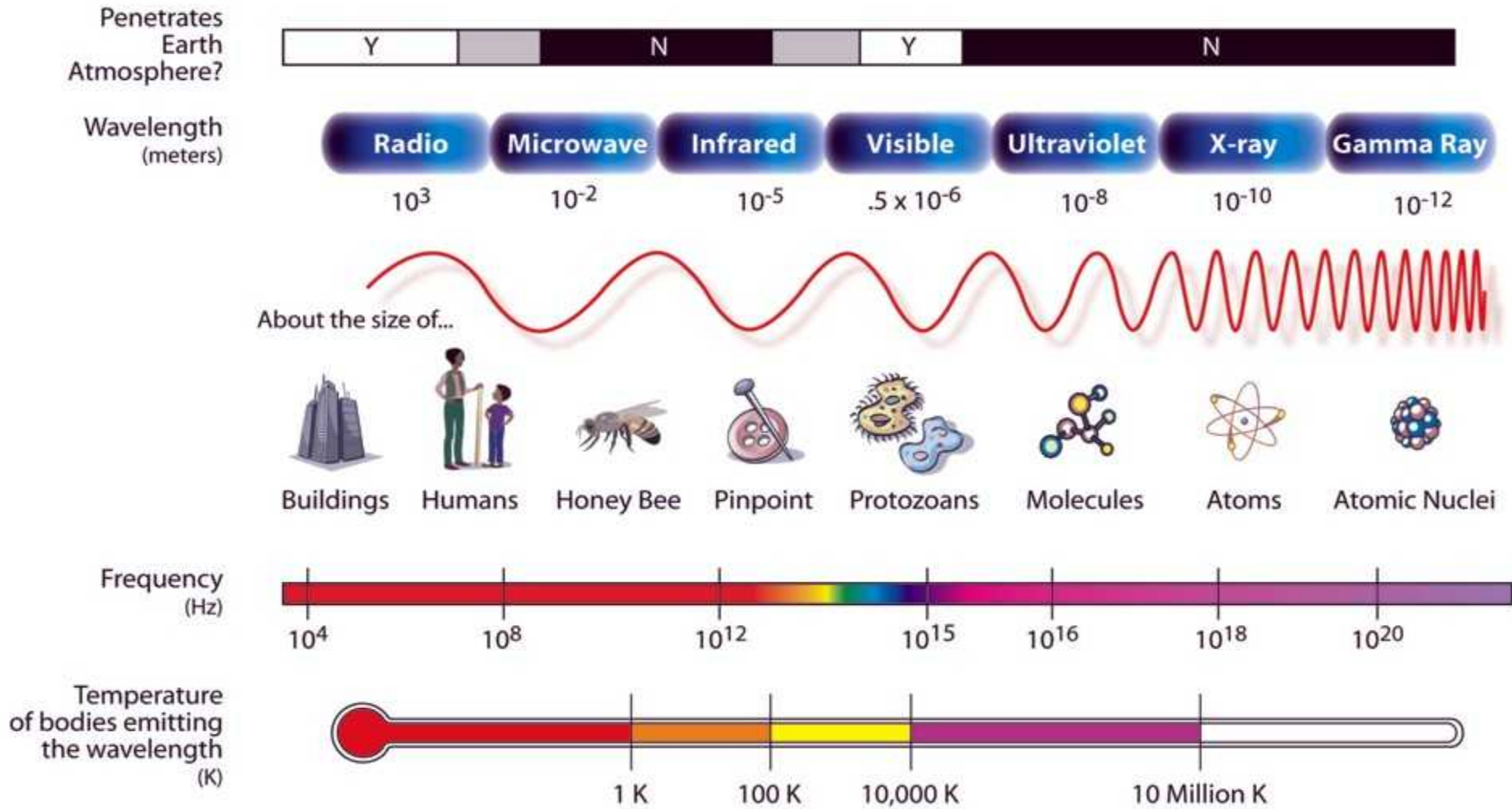
What are X-rays?

- 1895: Wilhelm Conrad Röntgen discovers X-rays by accident.
- He noticed a cathode-ray tube in his lab would excite phosphorescent materials
- He also noticed these mysterious rays would pass through soft materials but not other materials such as bone and metal.
- One of his first experiments involved his wife Bertha as a test subject.
- The first application of Radiography was not medical, but actually industrial.

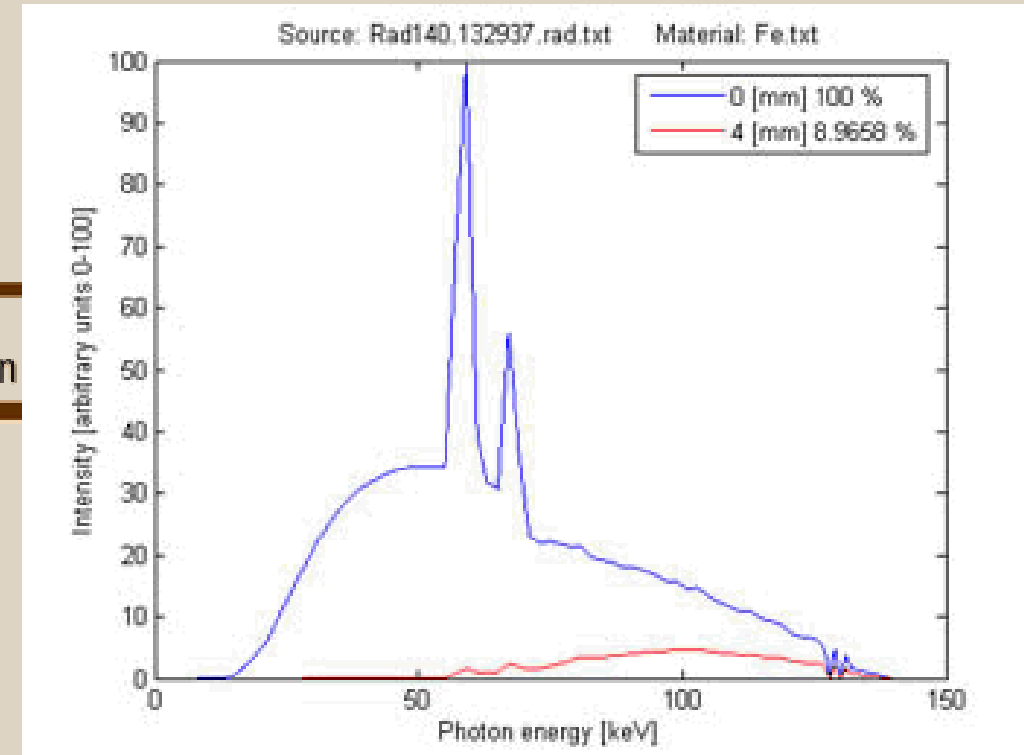
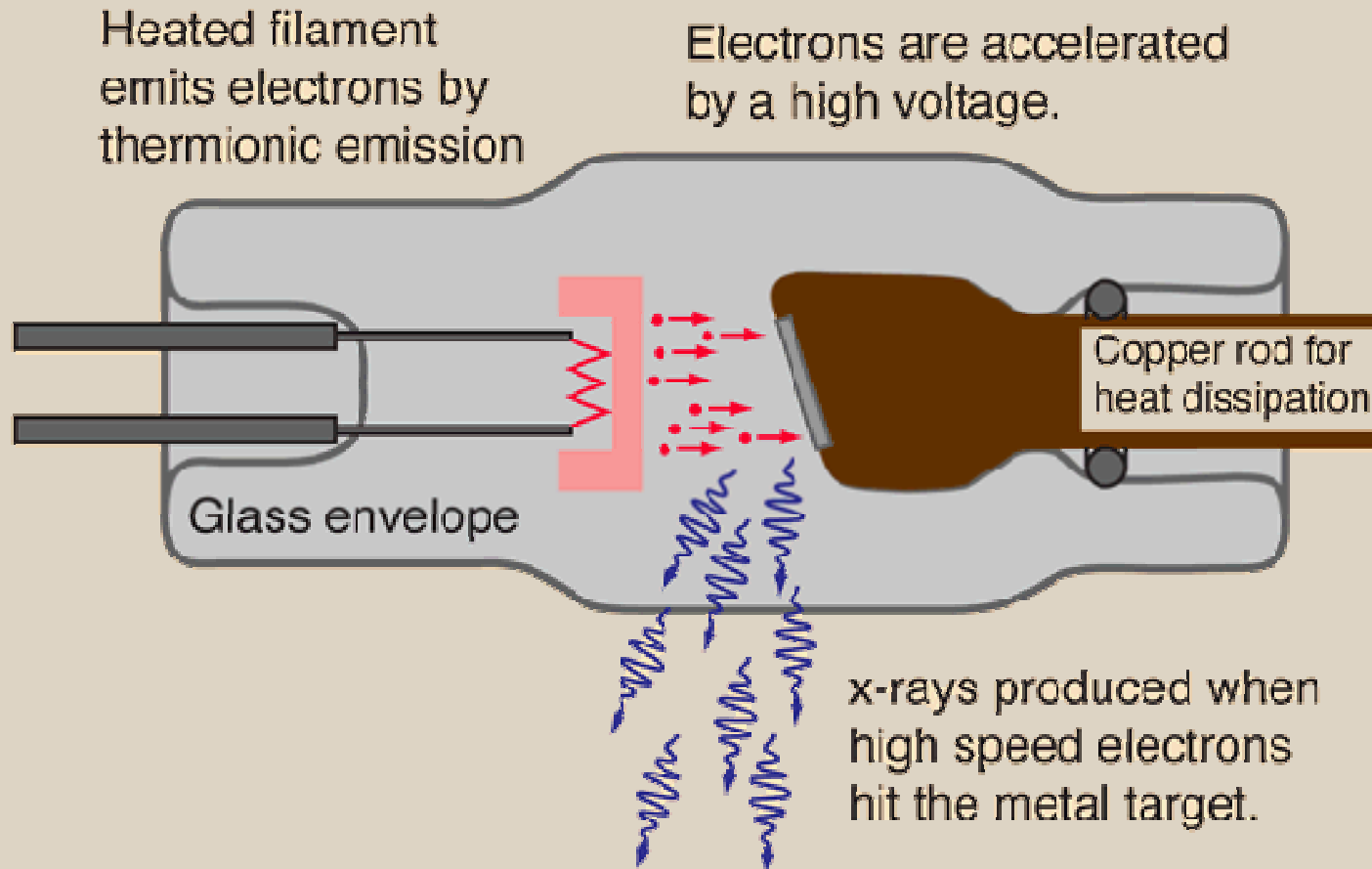
Wilhelm Conrad Röntgen



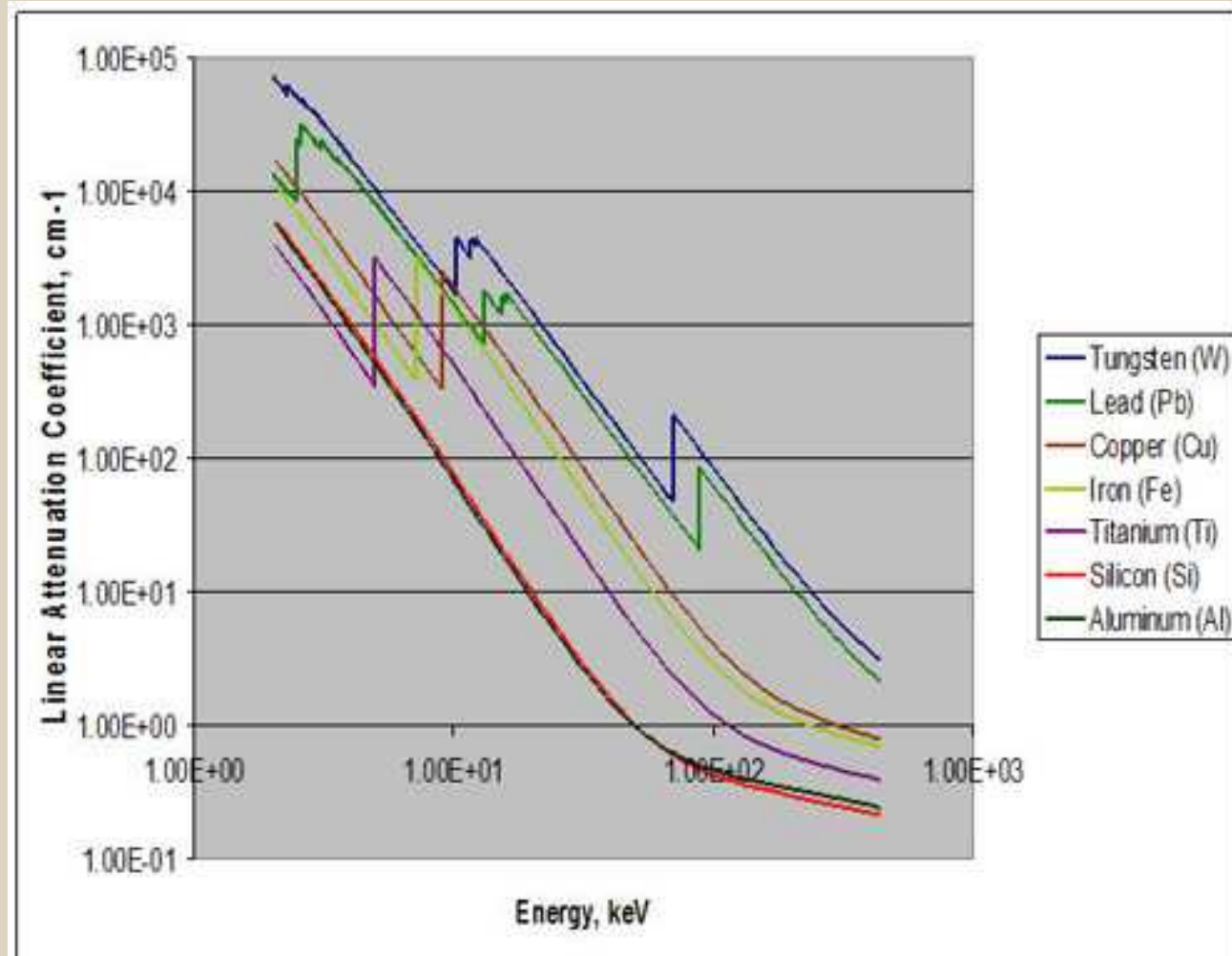
THE ELECTROMAGNETIC SPECTRUM



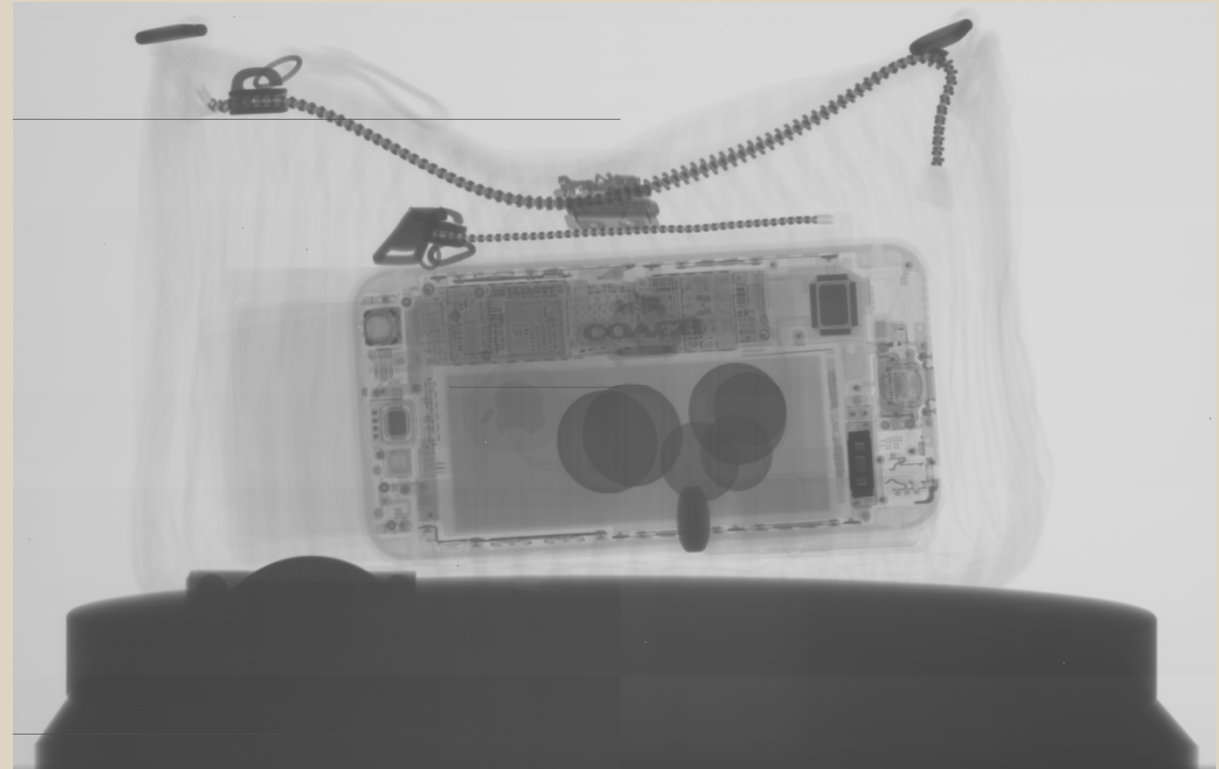
How are X-Rays Created?



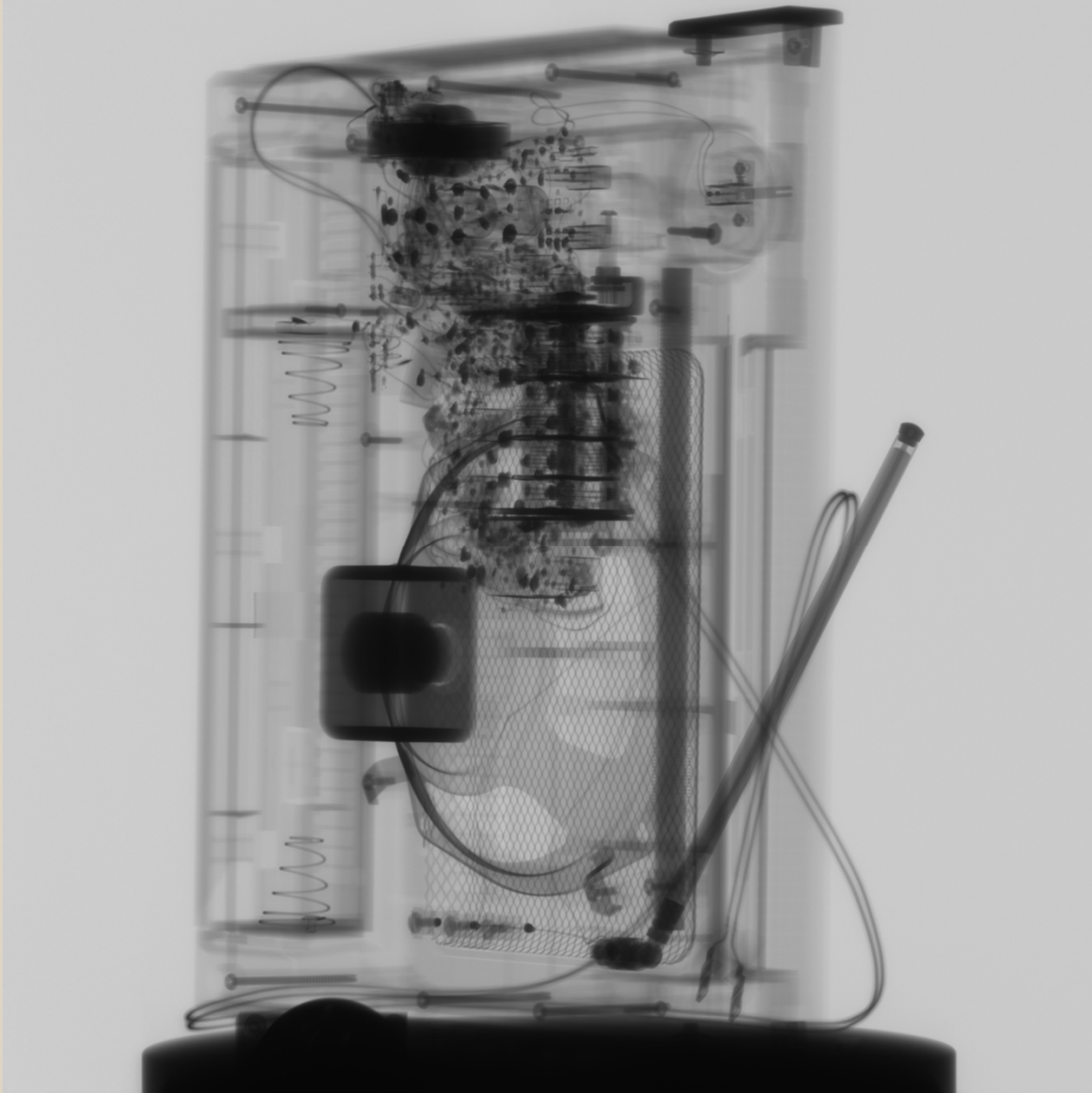
How they interact with materials?



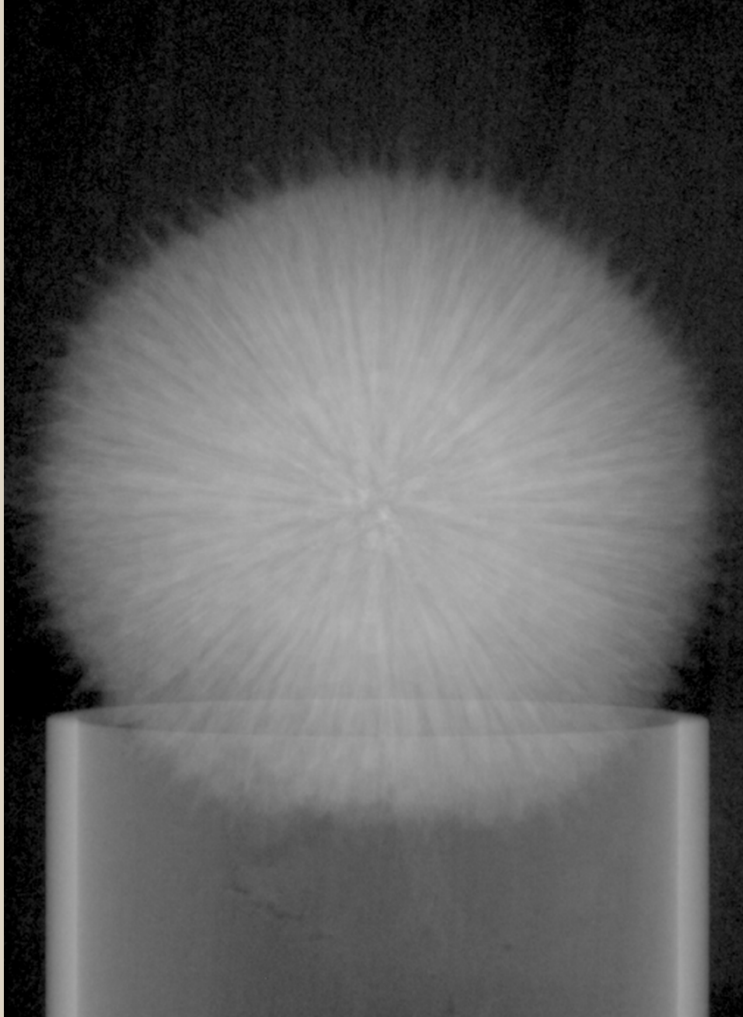
Some Examples



More Examples



One More...



SNL Applications

- Current Radiography technology leveraged at SNL
 - 3D Computed Tomography
 - Digital Radiography
 - Computed Radiography
 - Flash Radiography

- Applications
 - Defect detection
 - Anomaly detection
 - Materials characterizations
 - Feature extraction

Industrial Computed Tomography

- Computed Tomography (CT) is an indirect 3D imaging technique.
- Input: Set of X-ray images acquired about a center of rotation.
- Output: Three-dimensional approximation of internal and external structure
- Reconstruction: Convolution- Backprojection Algorithm (Feldkamp-Davis-Kress)
- Geometry and Configuration of CT System determines magnification
- Reconstruction algorithm is $O(n^4)$

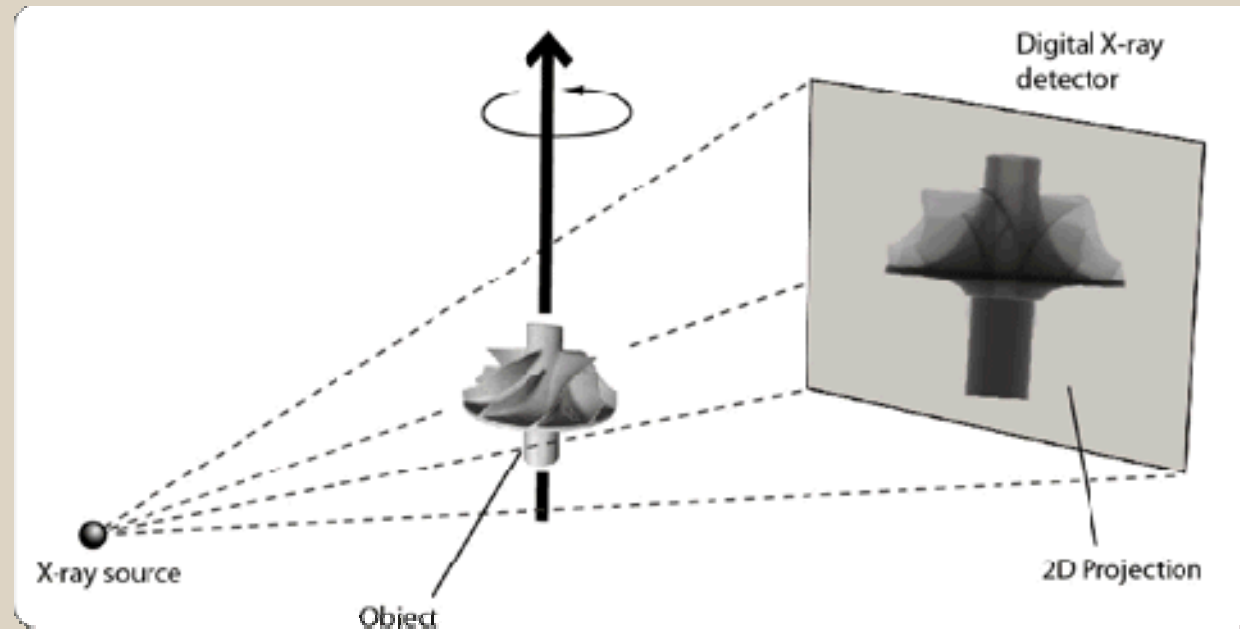
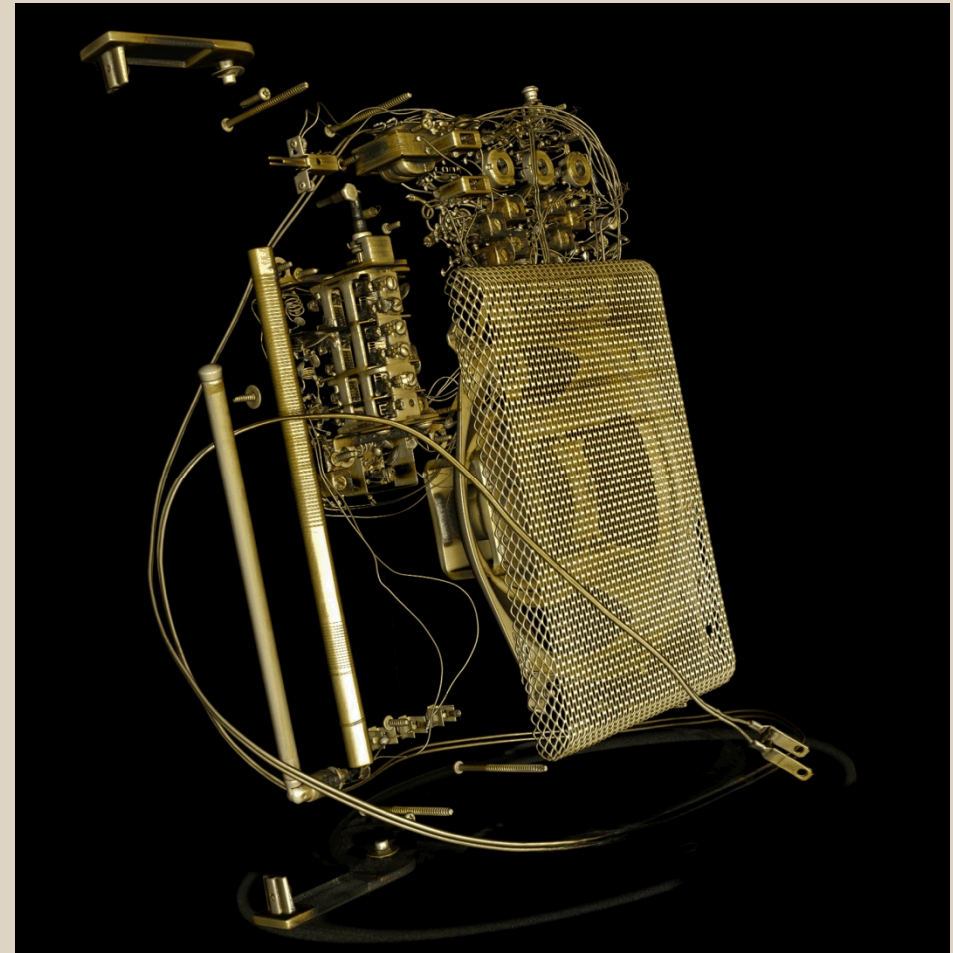
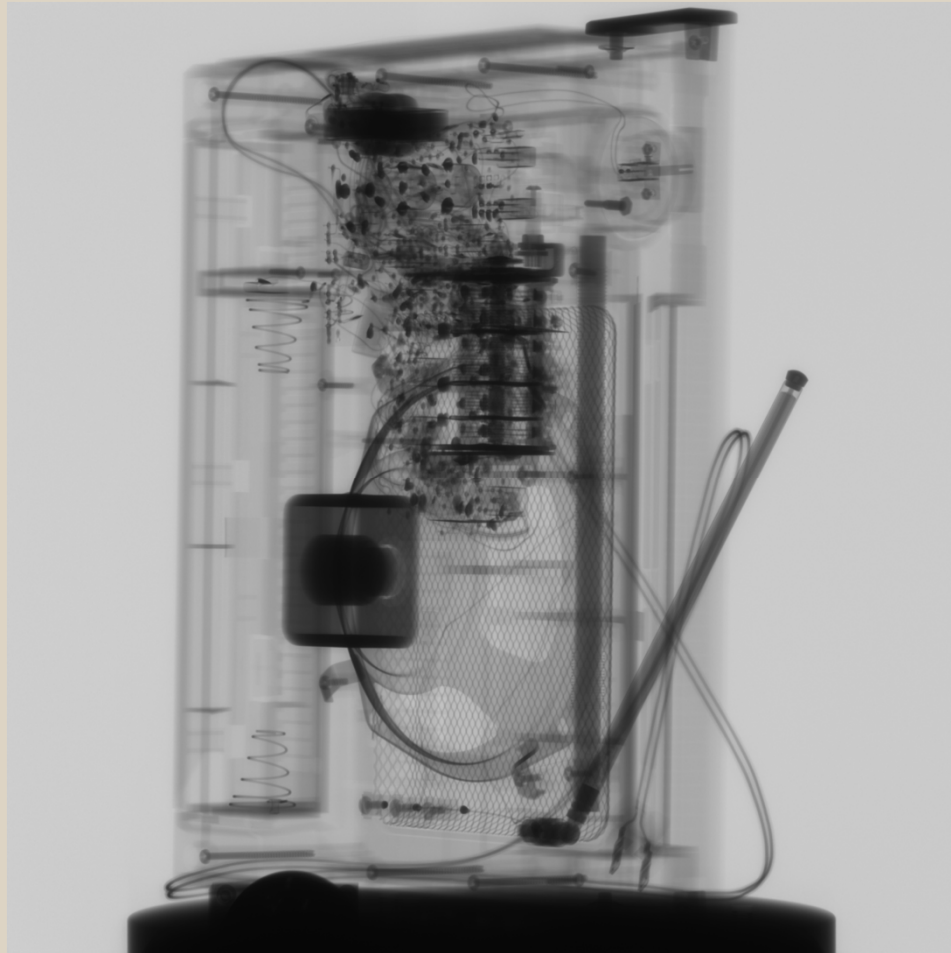


Image Source: <http://www.xviewct.com/assets/images/how-ct-works.gif>

Why Do We Care About CT?



Industrial Computed Tomography

- Why is this hard?
 - Non-Medical Datasets are orders of magnitude larger
 - Usually between 6GB to 4TB!
 - Hours to Years to Process!
 - Disk and Memory I/O
 - Complex Algorithms
 - Medical GPU-based methods take days to reconstruct large volumes
 - GPUs have a very different architecture from traditional CPUs.
- My First Research Project at Sandia
 - Create a HPC Multi GPU-based approach to reconstruction
 - Data movement will be optimized via Multi-Resolution Parallelization
 - Overlap Data movement and computation
 - Will be 20 times faster than traditional CPU approaches

Challenges

1. Code must be flexible and efficiently handle varying problem sizes for various GPU models.
2. Must be able to utilize multiple GPUs
 1. Must ensure GPUs are not starved for work.
 2. Must ensure GPU resources are maximized.
3. File I/O is now a major bottleneck in performance
4. A Supercomputer Algorithm that functions on desktops, workstations, and clusters.
5. Must handle future sized datasets.

Discoveries

- CT Reconstruction becomes an Irregular Problem
 - Discovery of Irregular nature led us to new approaches.
 - Developed Approach to Maximize Cache Utilization.
- Kernel Optimization
 - Scientific Papers only utilized texture memory and threads.
 - Discovered implementation that improves kernel performance.
 - 10x improvement of cache hit-rates.
- Utilization of MIMD/Multi-Resolution Implementation

Evaluation

- 1 Teravoxel (1 Trillion voxels) Reconstruction
 - 10K x 10K x 10K voxel volume
 - 10,000 Projections
 - 4TB input and 4TB output
 - “Future-proof” dataset.

Milestones

- Original Proposal:
 - Concurrent Data movement and computation
 - Data Queuing (Concurrent Data/Computations)
 - Dynamic Task Allocation (Maximize GPU Resources)
 - Pinned Memory (Improve Memory Transfers to GPU)
 - Will be 20 times faster than traditional CPU approaches
 - Single Thread: 260-3313 times faster
 - Multi-Thread: 33-328 times faster
 - GPU-Port: ~10-40 times faster
 - Impact: Minutes-to-seconds/Hours-to-minutes/Years-to-hours
 - Overall Impact
 - Utilized not just at SNL, but other labs as well for various National Security Problems

Results: Performance

| Algorithm | Time(hours) | Speedup | Energy(kWh) |
|---------------------|-------------|---------|-------------|
| CPU (Single-Thread) | 45581 | N/A | * |
| CPU (Multi-Thread) | 2576 | 17.7x | 1362.43 |
| Medical 8 GPU | 114.4 | 398.4x | 164.34 |
| SNL-Irregular 8 GPU | 23.2 | 1964.7x | 46.96 |
| SNL-Modular 8 GPU | 20 | 2779.1x | 38.72 |

Results: Scalability

| GPUs | SNL-Irregular | SNL-Modular | Medical |
|------|---------------|-------------|---------|
| 1 | 1x | 1x | 1x |
| 2 | 1.87x | 1.86x | * |
| 4 | 3.29x | 3.25x | * |
| 6 | 4.38x | 4.32x | 1.72x |
| 8 | 5.28x | 4.96x | 1.58x |

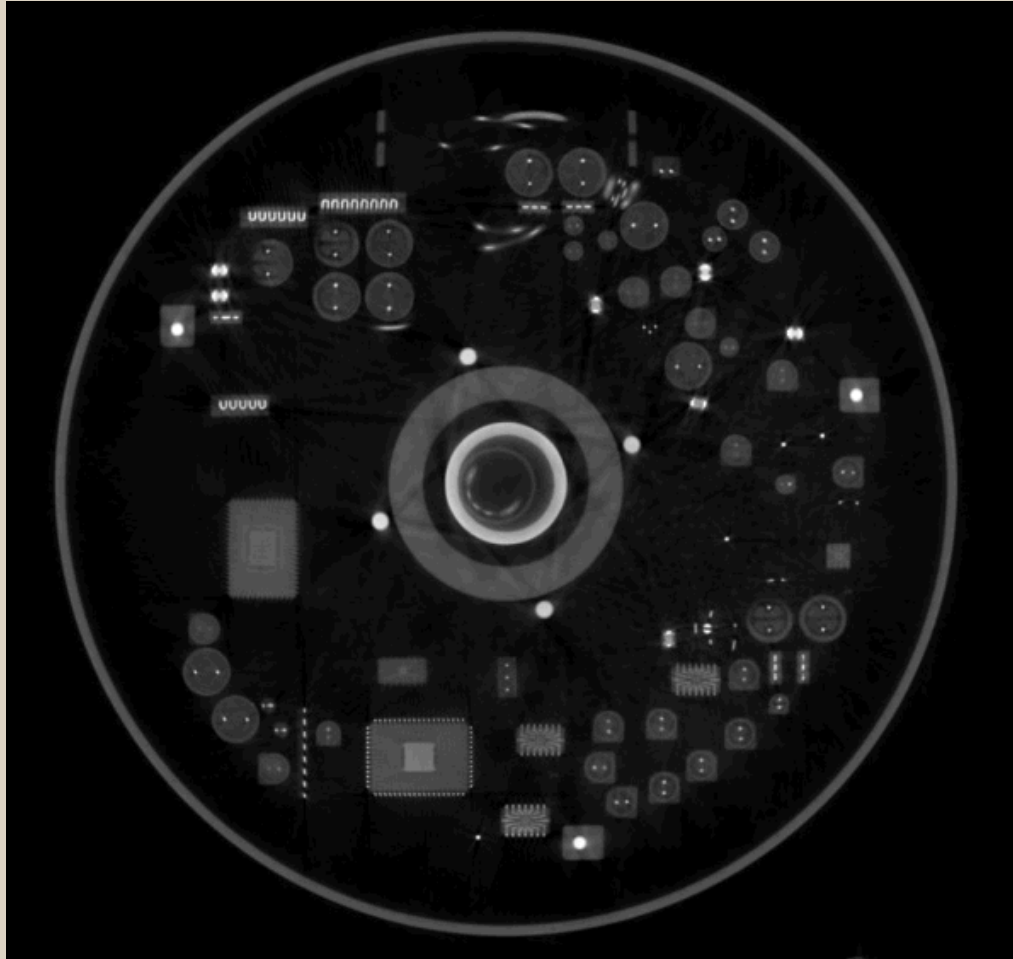
Conclusions for this First Project

- A five-year computation down to less than a day
- Can run on a laptop to a scientific cluster.
- Potentially reconstruct unthinkably high-resolution CT volumes.
- Tremendous insight into materials and fluids characterization.
- CT could be integrated as a regular quality control tool.
- This work could be applied to Green Computing and Exascale efforts.
- What other algorithms could benefit from this work?

What's Next?

- As great as CT is, it still is not perfect...
- CT Reconstruction makes some assumptions that can create errors.
 - Assumption: Bremsstrahlung Radiation is ignored or near monolithic.
 - Constant or Fixed Attenuation
- Bad x-ray images will lead to a bad reconstructions
 - X-ray Images have noise due to x-ray photon scatter
 - Not having enough X-ray images for reconstruction will degrade volumetric quality

CT Artifacts



Looking to the Future...

- Create an advanced reconstruction tool to identify slight nuances of as-built parts.
 - Small differences can lead to large difference in outcomes for simulated chaotic systems.
 - Better understanding in how well a part is made or how accurate a simulation can be achieved.
- CT for Non-Destructive Evaluation
 - Not ideal, various types of artifacts
 - Artifacts located exactly at regions of interest (interfaces, material boundaries, etc.)
 - CT datasets are hand segmented and imported into simulation tools.
 - Big Data for input and output (Typically ~12GB-4TB input and output)
- Solution: Multi-Energy Iterative Reconstruction (MEIR)
 - Energy resolved iterative reconstruction for CT
 - Greatly reduce artifacts and noise in volumetric reconstruction.
 - Investigate techniques that combine innovative detector technology with intelligent data science approaches.
 - Very Big-Data (~25GB-100TB input and output)

What are Energy Resolved X-Ray Images?



Why now? Why this Project?

- CT limitations have been reached.
 - SNL has a growing and urgent need beyond those limitations.
 - Component inspection, materials testing, advanced components testing, etc.
 - Potential components could include: Neutron Generators, Detonators, Thermal Batteries, etc.
 - Limitations due to penetrating power, resolution, magnification, etc.

- Iterative Reconstruction
 - MEIR has never been attempted at non-medical scale.
 - Medical Scale: 1 – 6 Gigabytes
 - SNL Scale: 12 Gigabytes to 4 Terabytes (each input AND output)
 - Avoided due to immense acquisition and post-processing.
 - How do we handle such large amount of data?!

- Proposed experimental acquisition
 - Leverage nascent technology and scanning techniques.
 - Very high risk, has never been attempted for this big of a problem.

- If successful, then SNL can transform many areas.
 - NDT, big-data sciences, advanced diagnostics, information science, and SNL's world class radiography capabilities.

Consequences of using bad data



Goals and Success

- Deliver generalizable Scientific Big-Data Approaches
 - Intelligently and efficiently manage 100s of terabytes of data for numeric processing.
 - Optimize information flow through scalable data reuse and transfer minimization

- Deliver optimized advanced reconstruction tool
 - Capable of reconstructing SNL relevant engineering big data.

- Deliver a big data optimized semi-autonomous volumetric segmentation algorithm with suitable importation to SIERRA/CTH.
 - Fairly well studied for CT reconstructions (low risk).

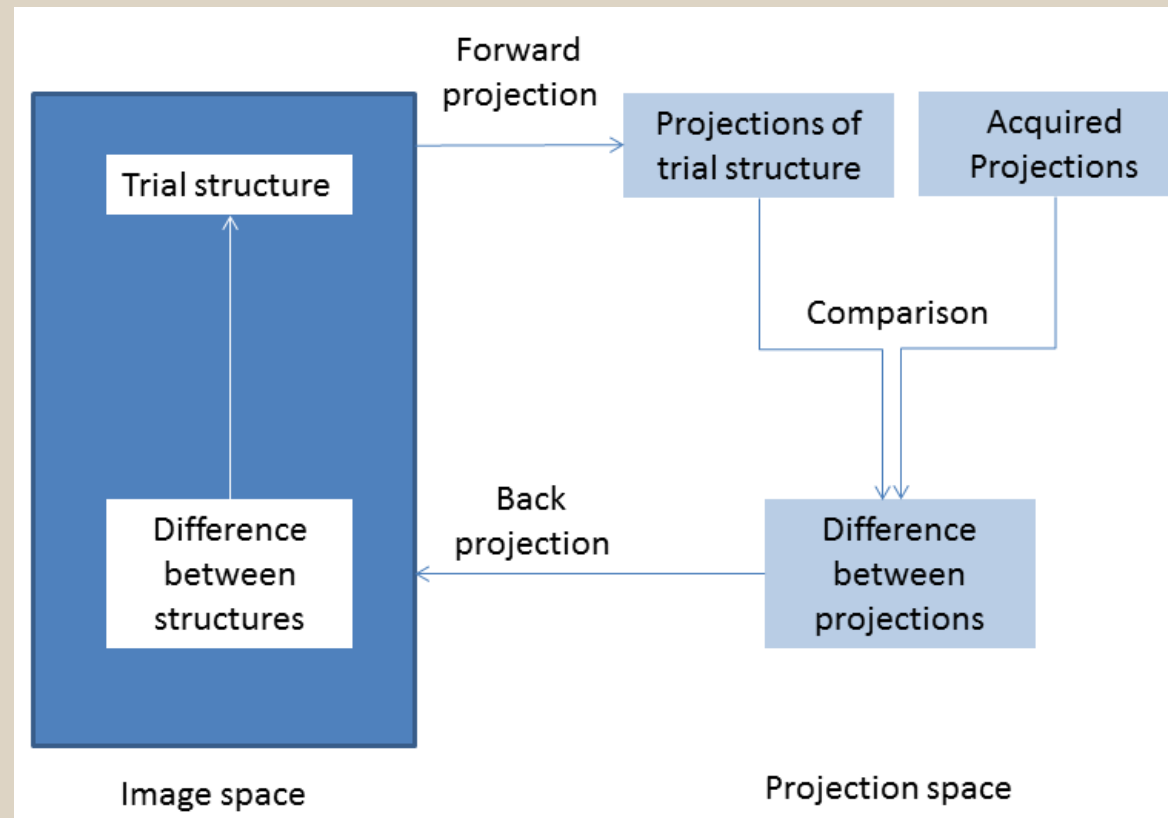
- Develop novel X-ray sampling techniques to supplement MEIR.

Parallel Efforts

- This work contains two main efforts
 - Algorithm R&D (Big-Data processing)
 - Experimental acquisition techniques (Big-Data acquisition)
- Neither effort is completely dependent on the other; however much synergy exists.
 - If Algorithms effort failed: Multi-Energy acquired datasets can still leverage the traditional multi-energy direct reconstruction.
 - If Experimental effort fails: The algorithms can still leverage traditionally acquired dual and multi-energy datasets.
- Algorithmic efforts will start with traditional dual-energy datasets which can be easily acquired with existing SNL hardware.
 - Mitigates risk
 - Algorithmic work won't be delayed by slips from the experimental effort.

MEIR (Iterative Reconstruction)

- Iterative algorithms:
 - Iteratively changing a trial structure until its projections are consistent with the original projections of the unknown structure.
 - Typically an iterative sequence of forward and back projection with constraints.
 - **We have decided to implement the Landweber Algorithm as it will give us the best flexibility with respect to information movement approaches.**



MEIR (Iterative Reconstruction)

- Many candidate algorithms for MEIR exist
 - Our first investigation revolves around Landweber-type algorithms

$$\hat{\mathbf{f}}^{k+1} = \hat{\mathbf{f}}^k + \mathbf{H}^T \left[\mathbf{g} - \mathbf{H}\hat{\mathbf{f}}^k \right]$$

- If successful, this algorithm will reduce artifacts due to
 - Beam hardening around edges and interfaces
 - Sparse sampling/Incomplete Data
- Landweber can be used on traditional datasets or multi-energy datasets.
 - **This was an originally proposed risk mitigation.**
 - **Best approach to create modular multi-use algorithms.**
- Risk: Very computationally intense and data intense problem.
 - Traditional single energy Reconstruction on Big Data: 1 hour to 3 days (using multi-GPUs).
 - Iterative energy resolved big data (theoretical): 10 hours to 300 days!

Iterative Reconstruction

iterative step = 1



Experimental Work Status

- Detector hardware was DIY
 - SNL-Designed configuration (TA coming soon).
 - Blueprint inconsistencies from vendor created challenges.
 - First of its kind for industrial R&D applications.
- The detector is now operation and data acquisition is ramping up.
 - Stages for CT hardware set up is still being executed.
- Have identified classified and unclassified components to test.
 - Ongoing team effort to identify other components to image.
 - Dependent on how system is characterized.
- Fortuitous Risk: Detector hardware has been calibrated for 300keV, able to penetrate several inches of steel efficiently.
 - Could no longer use originally planned lead cabinet.
- Currently underway to perform material identification.

The Power of Energy Resolved Imaging

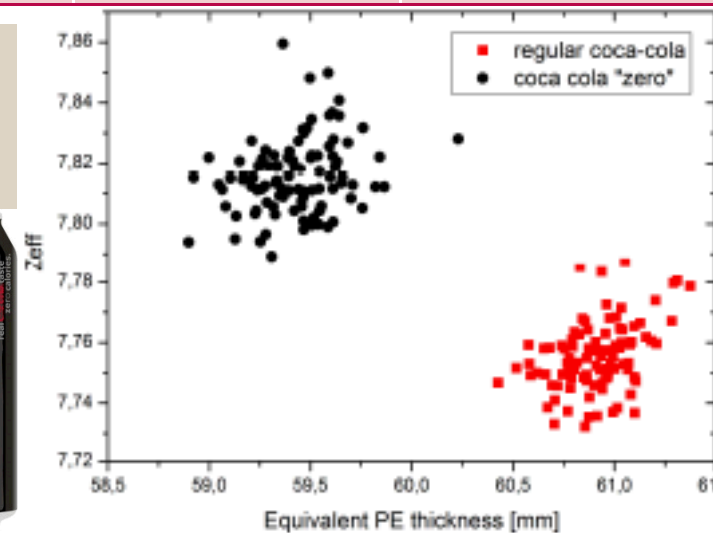
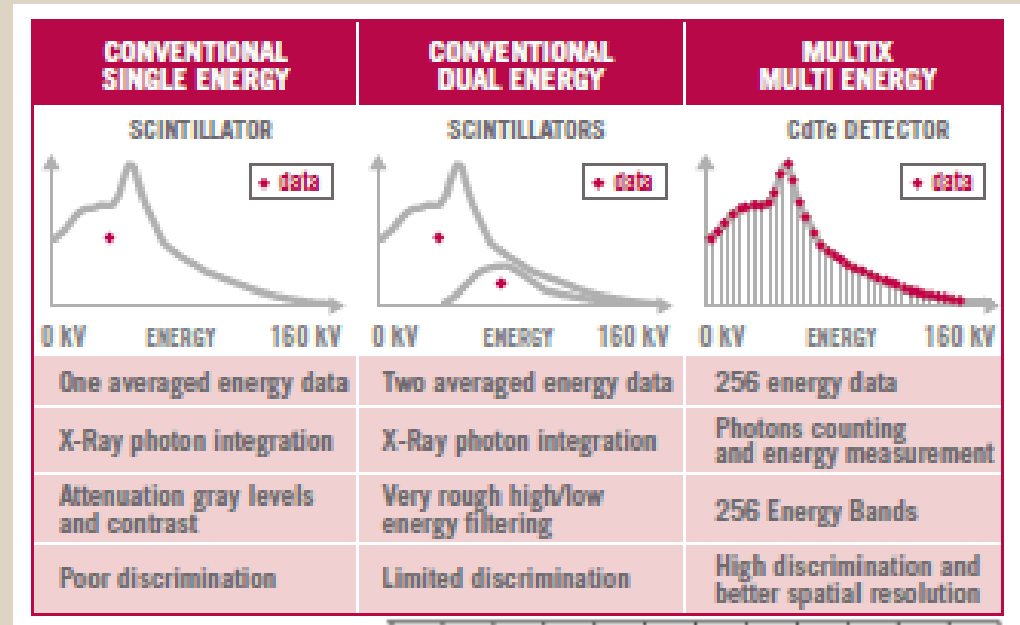
- We will use the MULTIX detector.
 - Best fit for industrial applications.

- Virtually mono-energetic energies.
 - Approximately 3.5keV bin width, up to 300 keV maximum energy.
 - Results in cleaner data for analysis and segmentation.

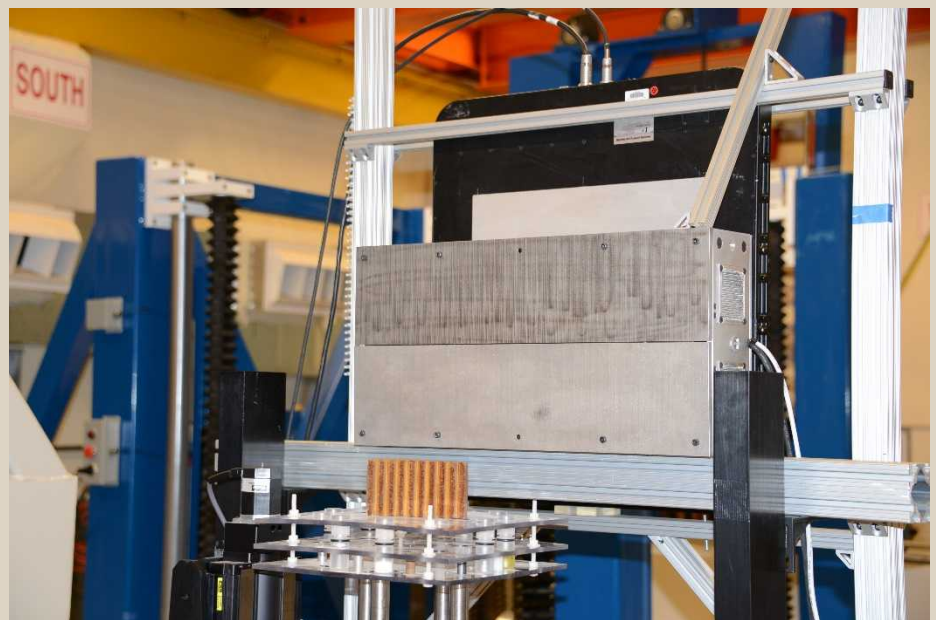
- Investigate material identification approaches.

- Risk: Detector is DIY
 - Detector assembled
 - GUI in development

- High Risk: Very new technology;
 - Large payoff for many engineering, science, and security applications that could significantly evolve each area.



System Configuration



Thank you

- Our team appreciates the opportunity to present our work.
- Questions?