Experiments with smart instruments at the Advanced Photon Source

Ian Foster

Argonne National Laboratory & The University of Chicago

Experiments with smart instruments at the Advanced Photon Source

Ian Foster

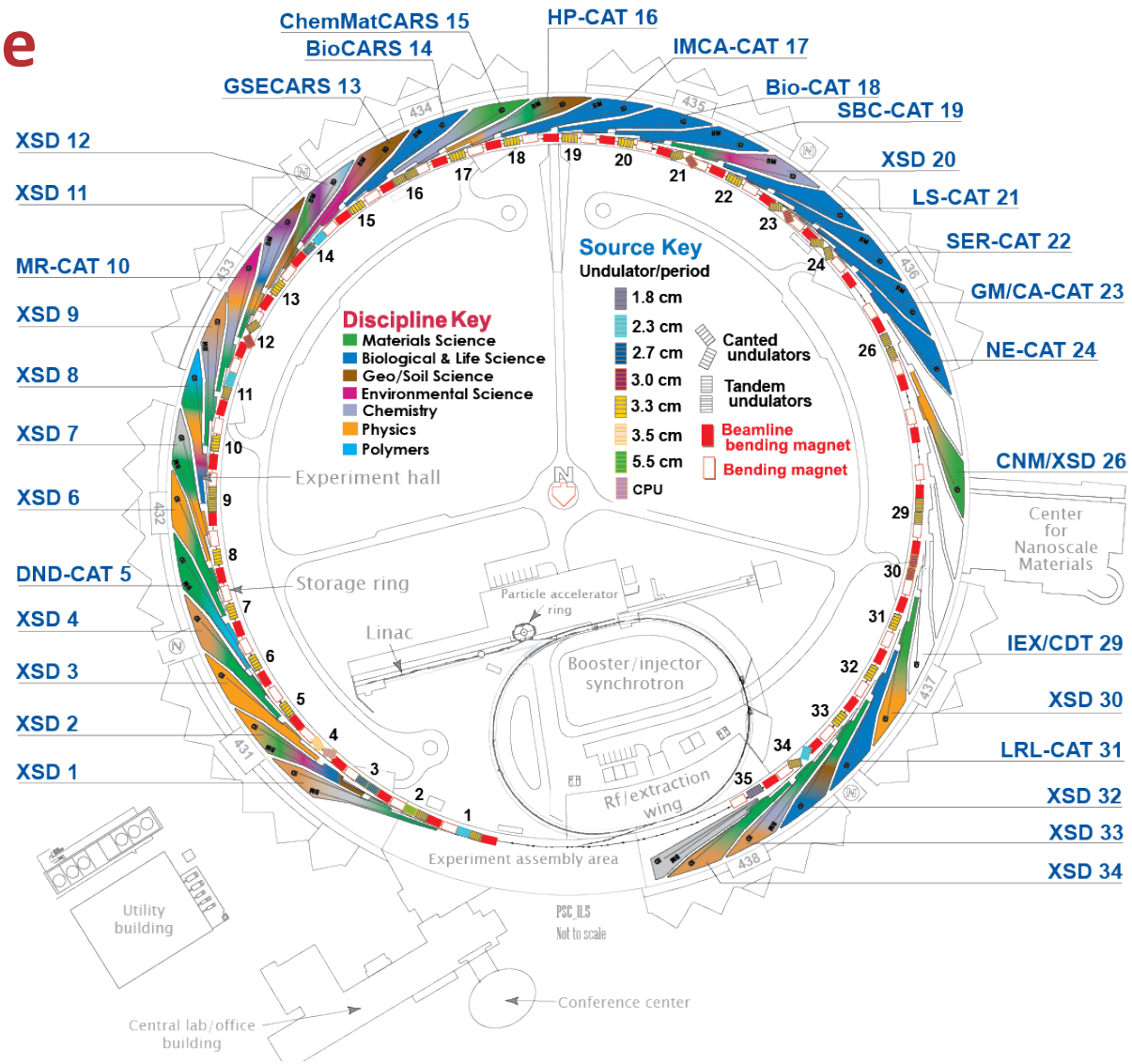
Argonne National Laboratory & The University of Chicago

Department of ENERGY NATIONAL LABORATORIES
symmetry
U.S. DEPARTMENT OF ENERGY
PNNL
NETL
INL
Sandia
Berkeley Lab
SLAC
Livermore
Sandia
Sandia
Los Alamos
Sandia
Sandia
Ames Lab
Argonne
Fermilab
Brookhaven Lab
NETL
PPPL
NETL
Jefferson Lab
ORNL
Savannah River
WASHINGTON
OREGON
CALIFORNIA
NEVADA
IDAHO
MONTANA
WYOMING
UTAH
ARIZONA
COLORADO
NEW MEXICO
NORTH DAKOTA
SOUTH DAKOTA
NEBRASKA
KANSAS
OKLAHOMA
TEXAS
MINNESOTA
IOWA
MISSOURI
ARKANSAS
LOUISIANA
WISCONSIN
ILLINOIS
MICHIGAN
INDIANA
OHIO
KENTUCKY
TENNESSEE
MISSISSIPPI
ALABAMA
GEORGIA
FLORIDA
SOUTH CAROLINA
NORTH CAROLINA
VIRGINIA
WEST VIRGINIA
PENNSYLVANIA
NEW YORK
NEW JERSEY
DELAWARE
MARYLAND
NEW HAMPSHIRE
VERMONT
MAINE
MASSACHUSETTS
RHODE ISLAND
CONNECTICUT
HAWAII
Pacific Ocean
Atlantic Ocean
Gulf of Mexico
Legend
DOE Laboratories
DOE Office of Science Laboratories

Advanced Photon Source

34 sectors
70 beamlines

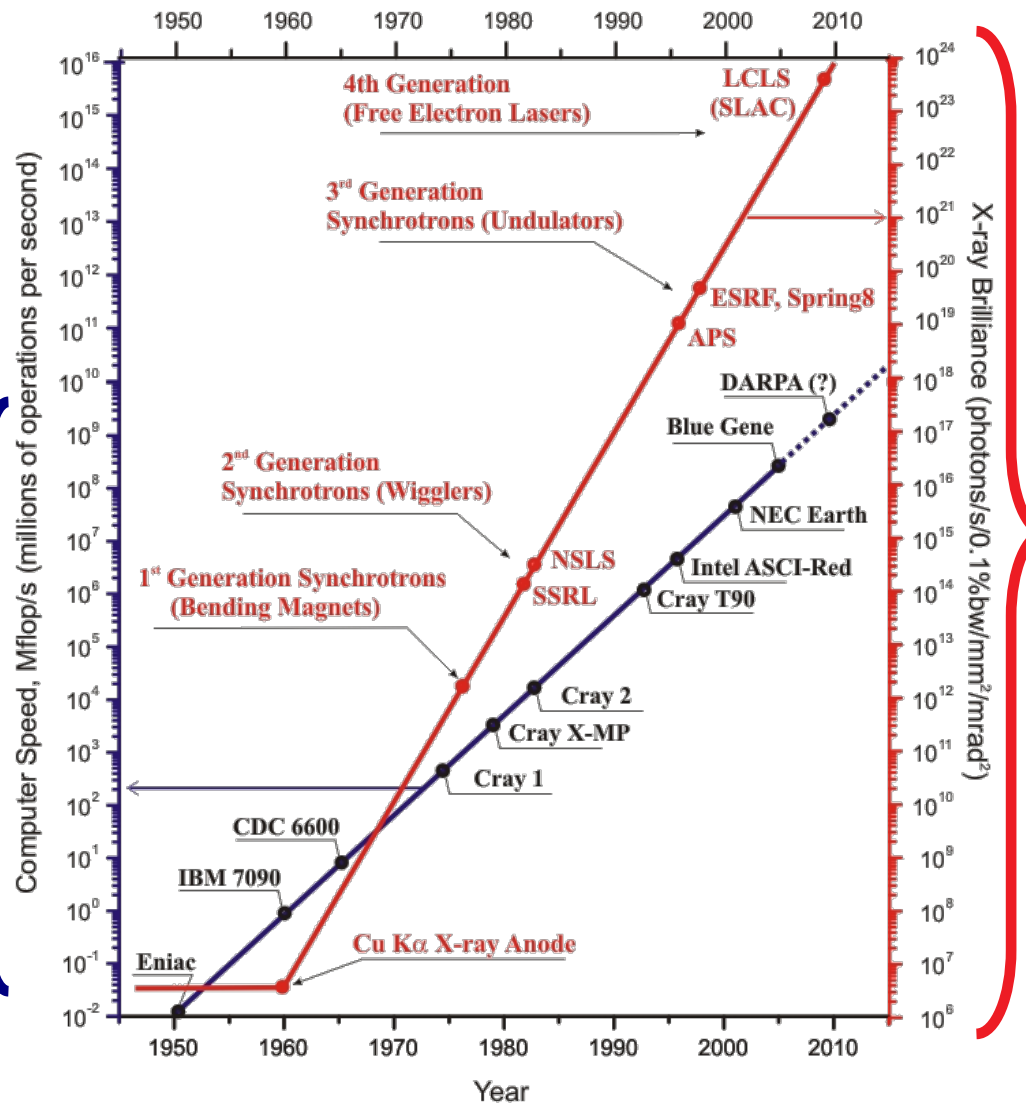
Imaging: 18 techniques
Spectroscopy: 16 techniques
Scattering: 49 techniques



Argonne NATIONAL LABORATORY

Light source brilliance is growing much faster than computer speed

12 orders of magnitude in 6 decades



18 orders of magnitude in 5 decades!

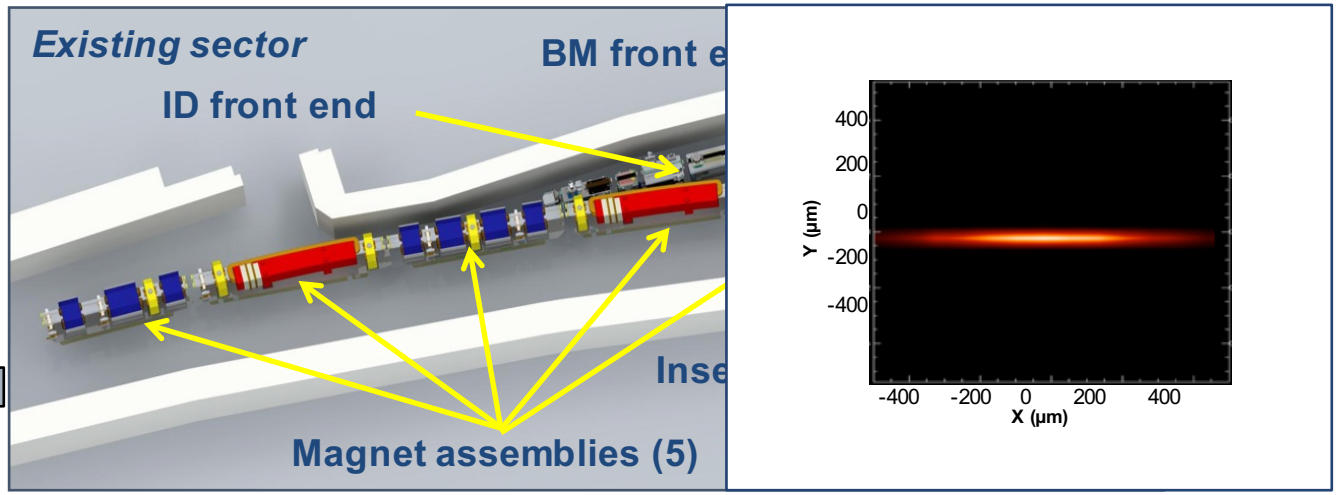
Argonne NATIONAL LABORATORY

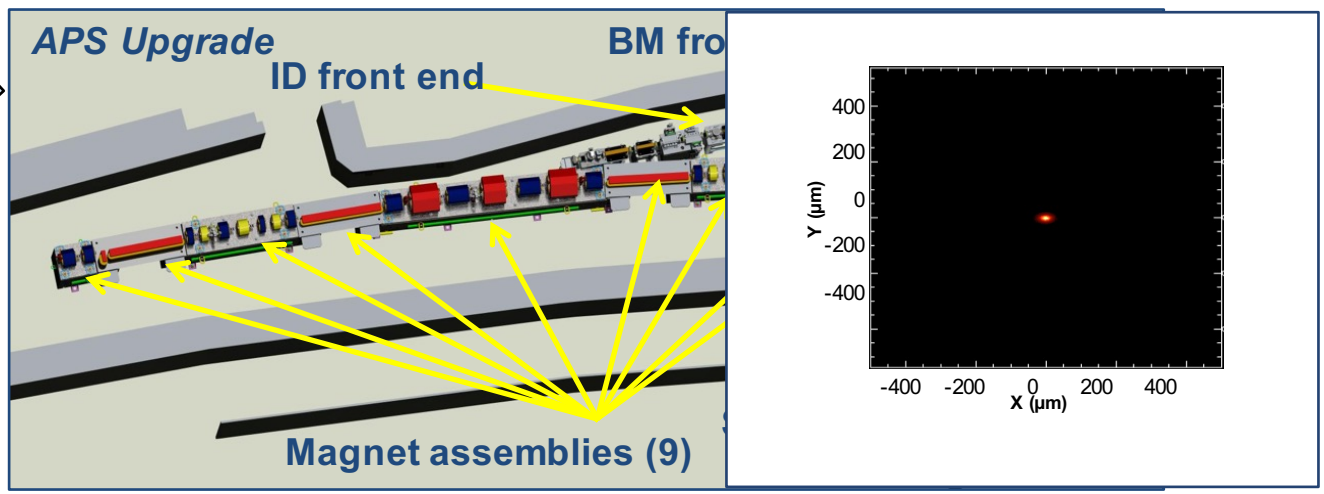
APS Upgrade: Existing vs. new storage ring

$$\varepsilon_x = C_L \frac{E^2}{N_D^3}$$

E = Beam energy
(E = 6 GeV for APS-U)

N_d = Dipoles per sector
(N_d = 7 for APS-U)





~50-fold reduction in horizontal emittance

Argonne NATIONAL LABORATORY

APS-U: The Ultimate 3D Microscope

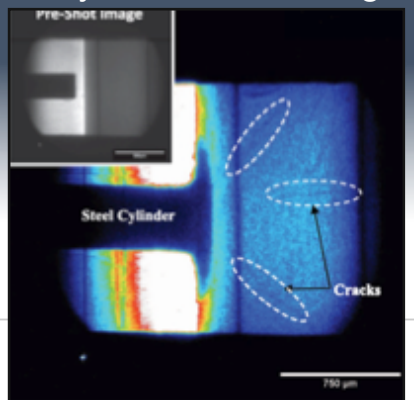
A new analytical tool to approach the supreme goal of measurement science:

To map any atom's position, identity and dynamics

High Energy

Penetrate bulk materials and operating systems

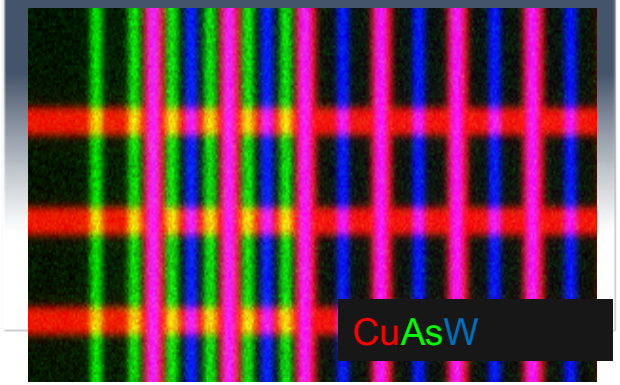
- World's brightest source of hard X-rays
- 3D mapping, deep inside samples
- X-ray cinematography in previously inaccessible regimes



Brightness

Provide macroscopic fields of view with nm-scale resolution

- Multi-scale imaging – connecting nanometer features across macroscopic dimensions
- Fast sampling with chemical, magnetic, electronic sensitivity

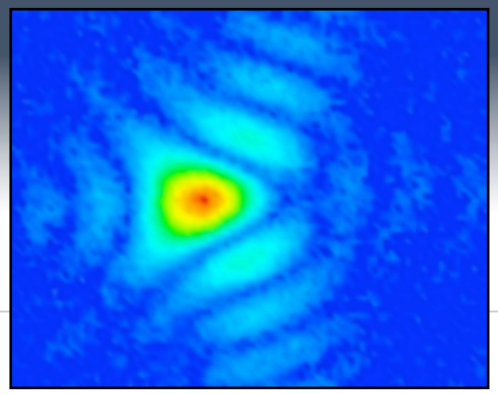


100 nm

Coherence

Enable highest spatial resolution even in non-periodic materials

- Extend lensless imaging to hard x-ray domain, resolution to <1 nm, localizing atoms
- Increase phase contrast for fast full field imaging
- Correlation methods improve by 10,000-1,000,000x



Major data and computation challenges arise across APS; Explode with APS-U

1) Huge data from new detectors, APS-U

- E.g., XPCS: Today: 2MB images @ 100 Hz; FY16: 1MB images @ 2000 Hz (x 10); Eiger: 2Mbyte @ 3000 Hz (x 3); APS-U +2-3 orders of magnitude

2) Complex, multi-modal data needs advanced computation for interpretation

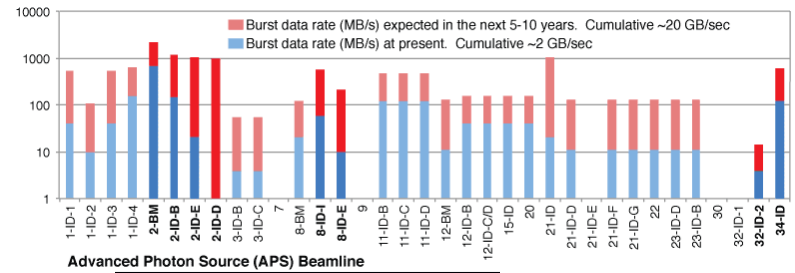
- E.g., ptychography+elemental mapping+visual images as a function of reaction conditions

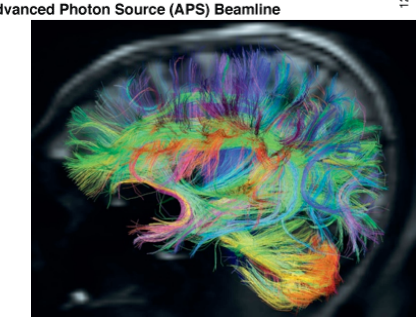
3) Advanced modeling and theory enable **fitting and co-optimization** of model and experiment

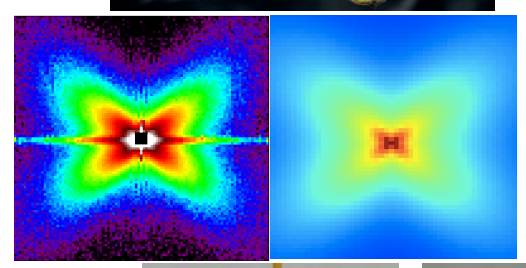
- Goal: Fit one model to all measurements

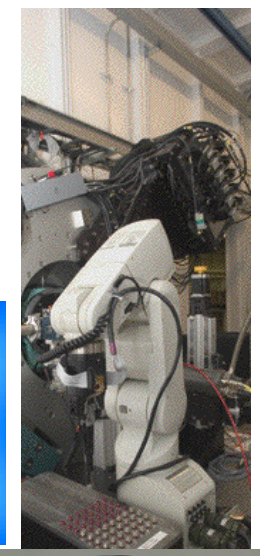
4) New user demographics → **automation**

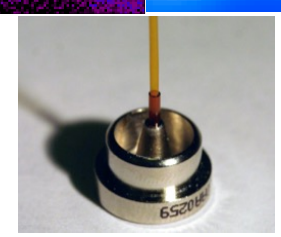
- Scale to more and different users, many with limited or no experience with light sources



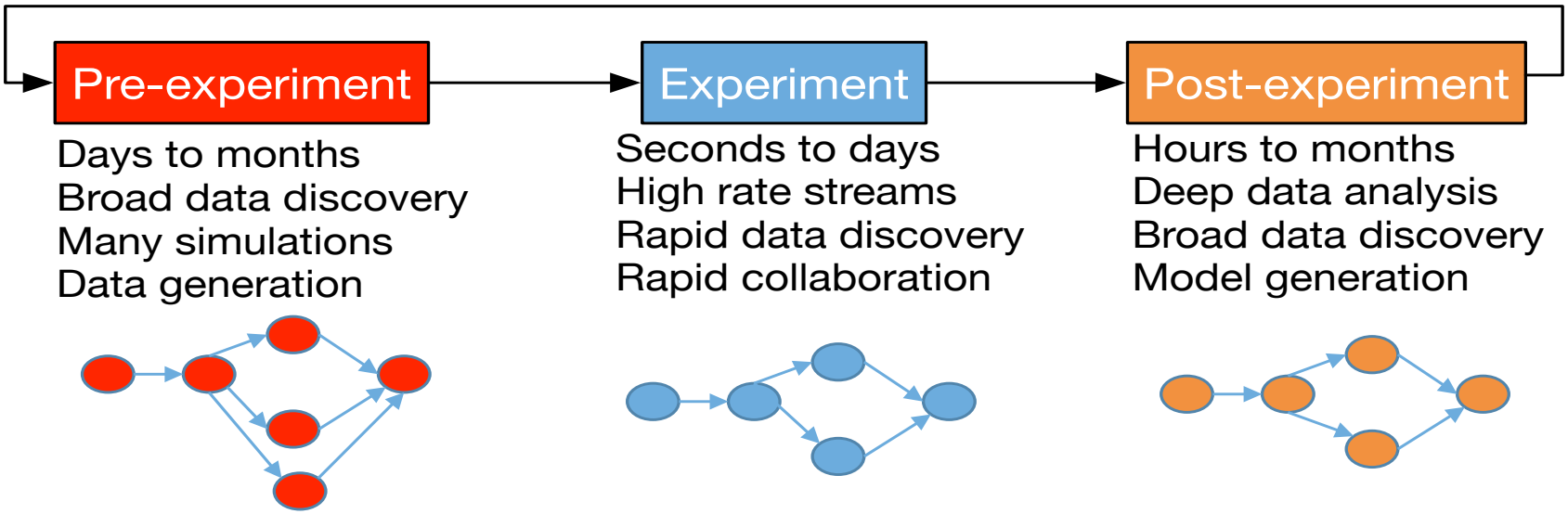








Light source computing: an end-to-end, multi-scale challenge



We seek **new experimental methodologies** that involve **rapid reconstruction and analysis** and **computer-in-the-loop control**

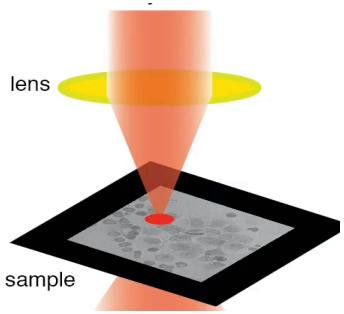
We also want to address other elements of the end-to-end pipeline: e.g., **experiment design** and **knowledge synthesis**

Experiments in advanced light source computing

(1) Parallel algorithms for rapid feedback

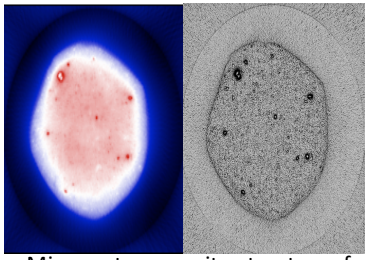
Ptychography: 360x speedup, from 7 hours to 72 s
[Deng, Vine, Chen, **Nashed**, Philips, Jin, **Peterka**, **Ross**, Jacobsen]
→Enable online analysis and use of fly scans

Coherence



Microtomography: Reduce reconstruction time from many days to 2 mins (32K cores)
[**Bicer**, **Foster**, Gursoy, **Kettimuthu**, De Carlo, Agrawal]
→Identify and correct experimental misconfiguration

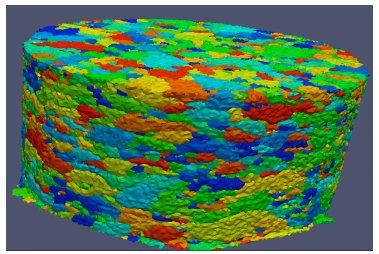
Brightness



Micrometer porosity structure of shale samples

High-energy diffraction microscopy: 5 hours on cluster reduced to 10 minutes (128K cores)
[Sharma, Almer, **Wozniak**, **Wilde**, **Foster**]
→ Zoom in on crack locations (switch far field → near field)

Energy

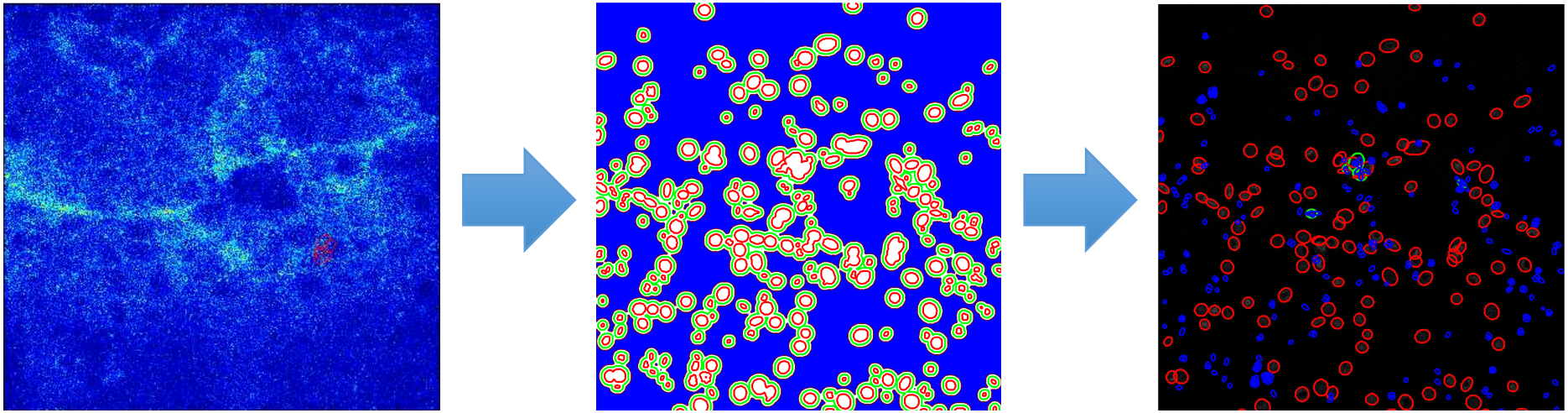


Microstructure of a copper wire, 0.2mm diameter

Experiments in advanced light source computing

(2) New algorithms, for example for feature detection

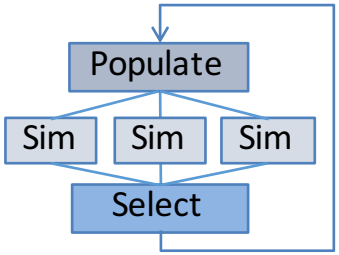
Fluorescence spectroscopy: identify cells and cell types
[Wang, Ward, **Leyffer**, **Wild**, Jacobsen, Vogt]



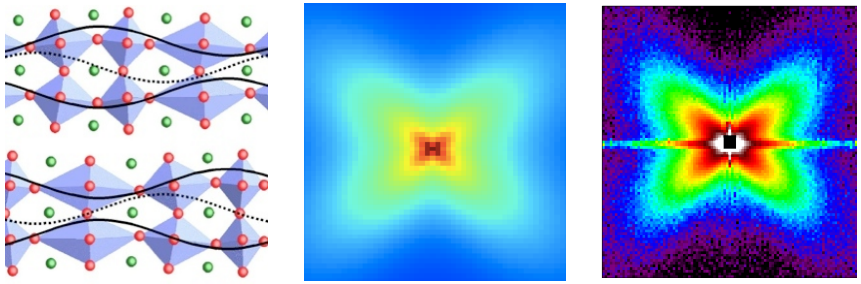
2000-channel fluoresence spectroscopy data.
Transformations from example element map, to cell delineation, to overlap detection and cell classification.
Green cells are algae, blue are yeast, and red are red blood cells.

Experiments in advanced light source computing

(3) Link experiment and simulation

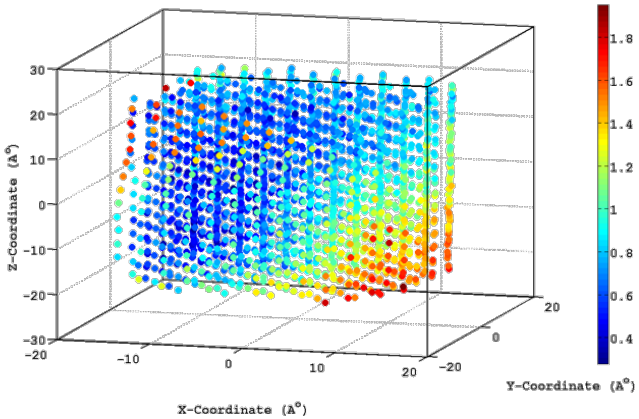


Diffuse scattering: molecular dynamics & evolutionary optimization to estimate disordered structure: 100K+ cores [Osborn, Jennings, Zapol, **Wozniak**, **Wilde**, **Foster**]



Candidate structure – simulated scattering – experimental scattering

Coherent diffraction imaging: molecular dynamics to interpret ultra-fast imaging [**Peterka**, **Ferrier**, **Leyffer**, **Munson**, **Sankaranarayanan**, **Wen**]



Relative displacement of gold atoms in ascorbic acid compared to initial lattice

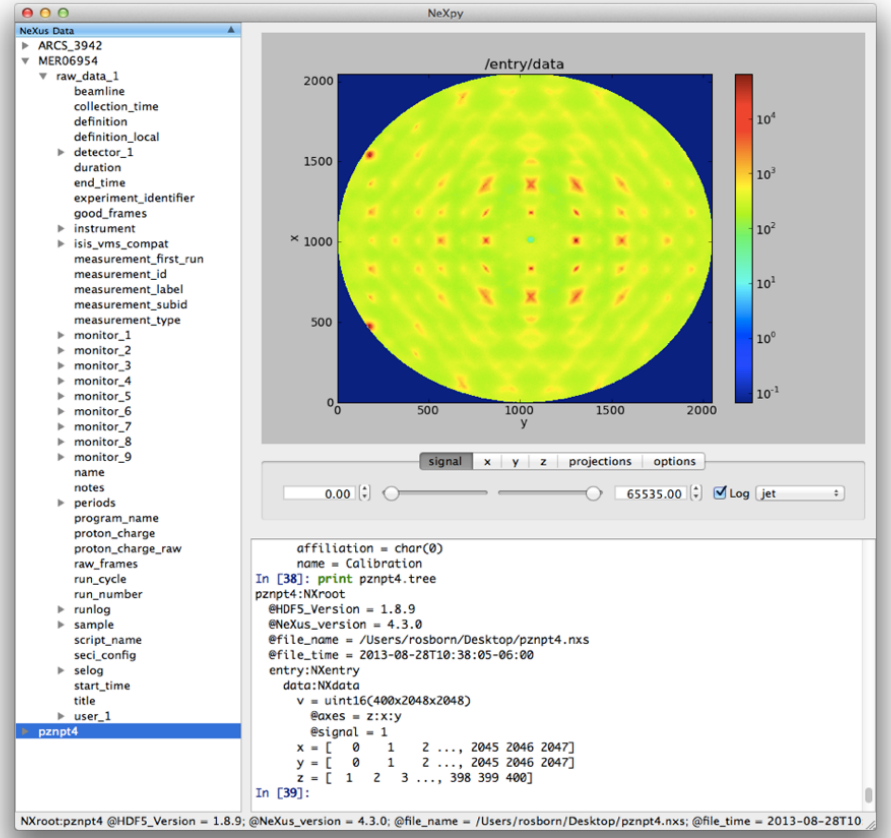
Experiments in advanced light source computing

(4) Data management architecture and methods

GB/s flows and PB data require new data architectures

We have demonstrated:

- GB/s streaming to data server
- Automated data reduction workflows
- Automated cataloging for discovery
- Interactive remote analysis
- Rapid data dissemination worldwide
- Automated publication

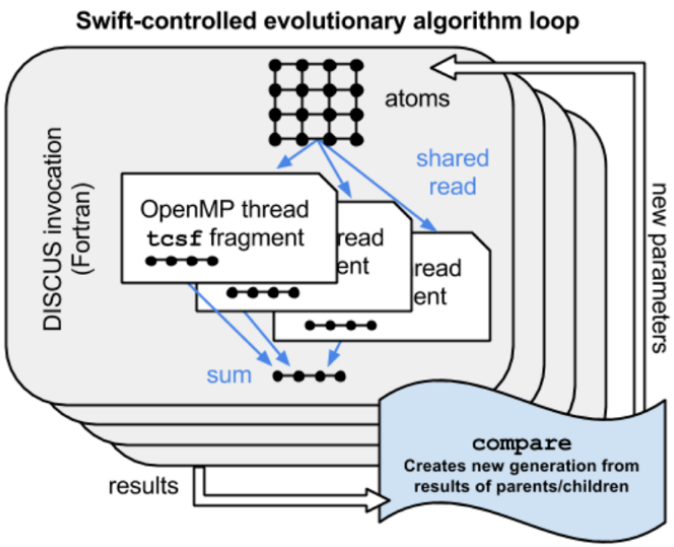


TomoPy tool with remote access capabilities for multi-TB datasets on Petrel (Osborn, Wozniak)

swift

for parallel programming

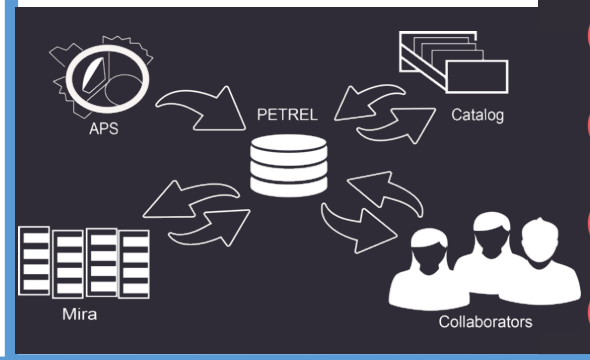
swift-lang.org



Petrel online store

petrel.alcf.anl.gov

94 Gbit/s Petrel—Blue Waters

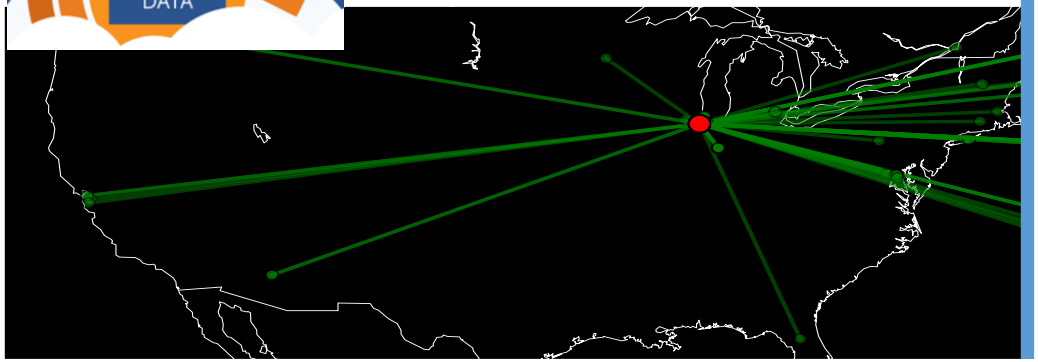


- 32 Nodes with 1.7 PB usable storage
- GPFS and Globus
- 100TB allocation per project
- Transfer and sharing data with collaborators
- Federated login
- Self-managed by PIs

globus.org

Globus service for data movement and sharing

globus.org



MATERIALS DATA FACILITY

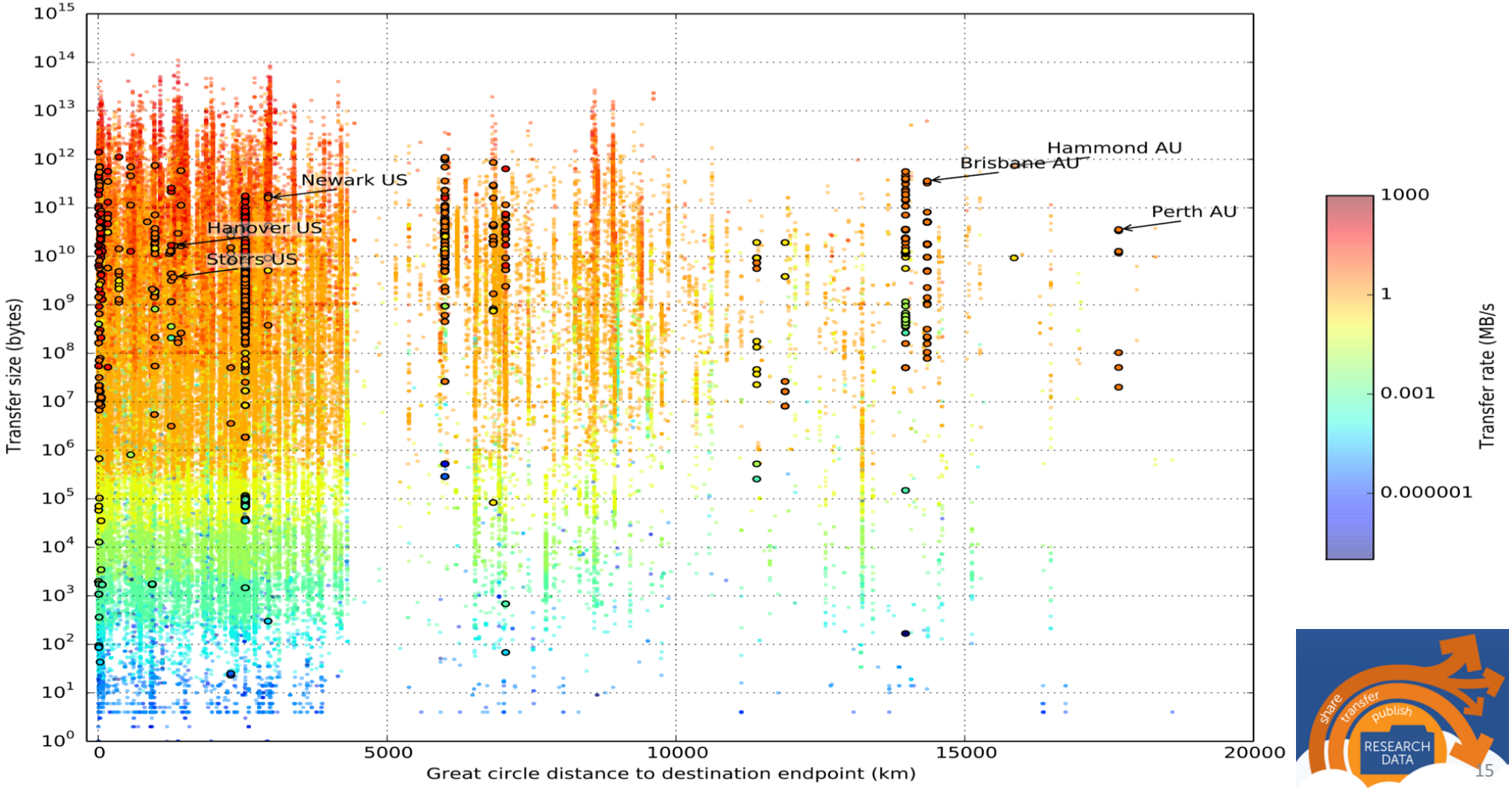
for data publication & discovery

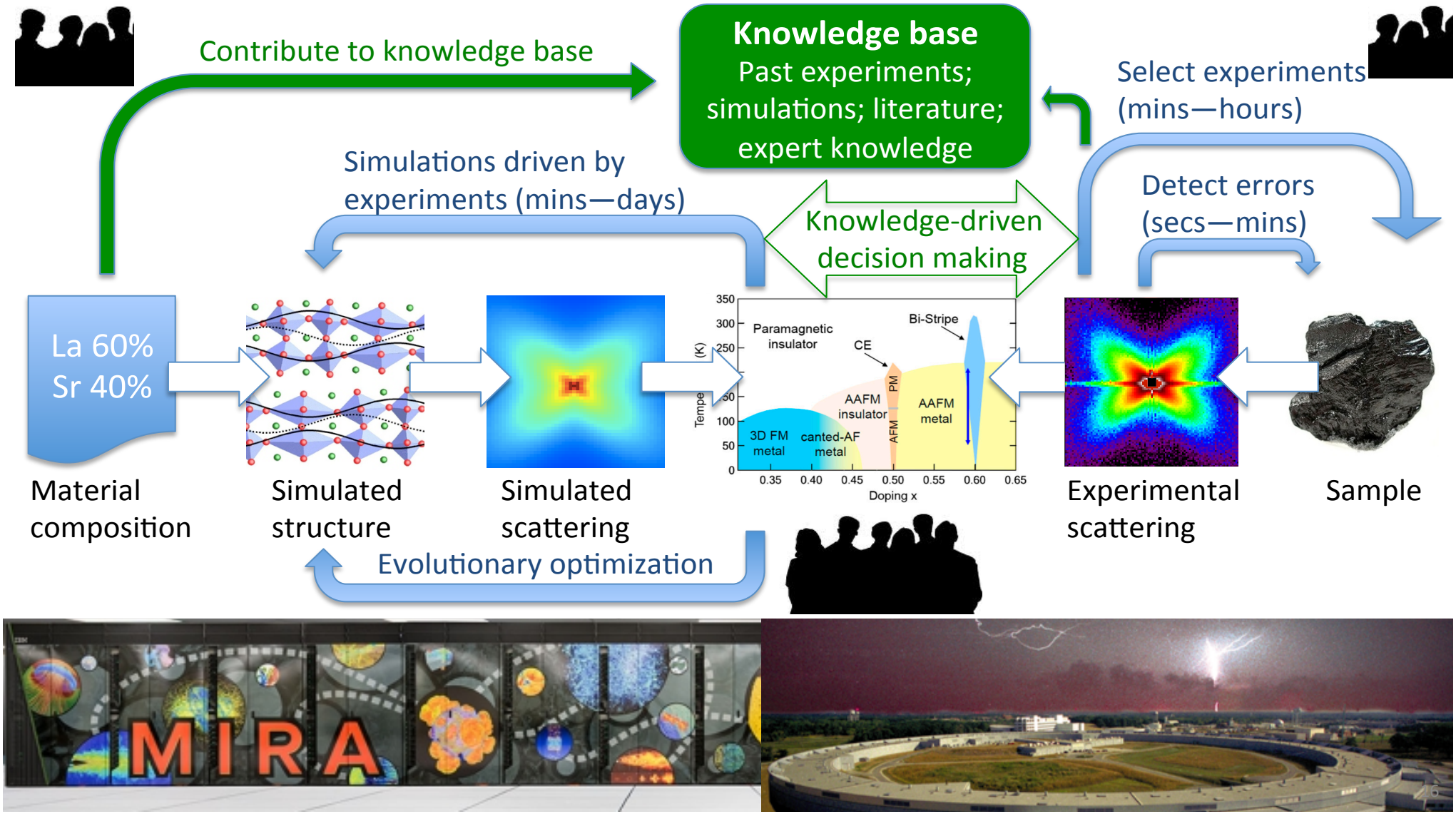
materialsdatafacility.org

Results 1-10 of 22

Issue Date	Title	Author(s)
15-Jun-2016	Twin-mediated Crystal Growth: an Enigma Resolved	*Shahani, Ashwin; Gulsoy, E. Begum; Poulsen, Stefan O.; Xiao, Xianghui; Voorhees, Peter W.*

Globus transfers, showing rate (via color) as a function of distance and volume.
The 1921 transfers from aps#clutch are highlighted.




Contribute to knowledge base
Knowledge base
Past experiments; simulations; literature; expert knowledge
Select experiments (mins—hours)
Simulations driven by experiments (mins—days)
Knowledge-driven decision making
Detect errors (secs—mins)
La 60%
Sr 40%
Paramagnetic insulator
CE
Bi-Stripe
AAFM insulator
PM
AFM
AAFM metal
3D FM metal
canted-AF metal
Doping x
Material composition
Simulated structure
Simulated scattering
Experimental scattering
Sample
Evolutionary optimization

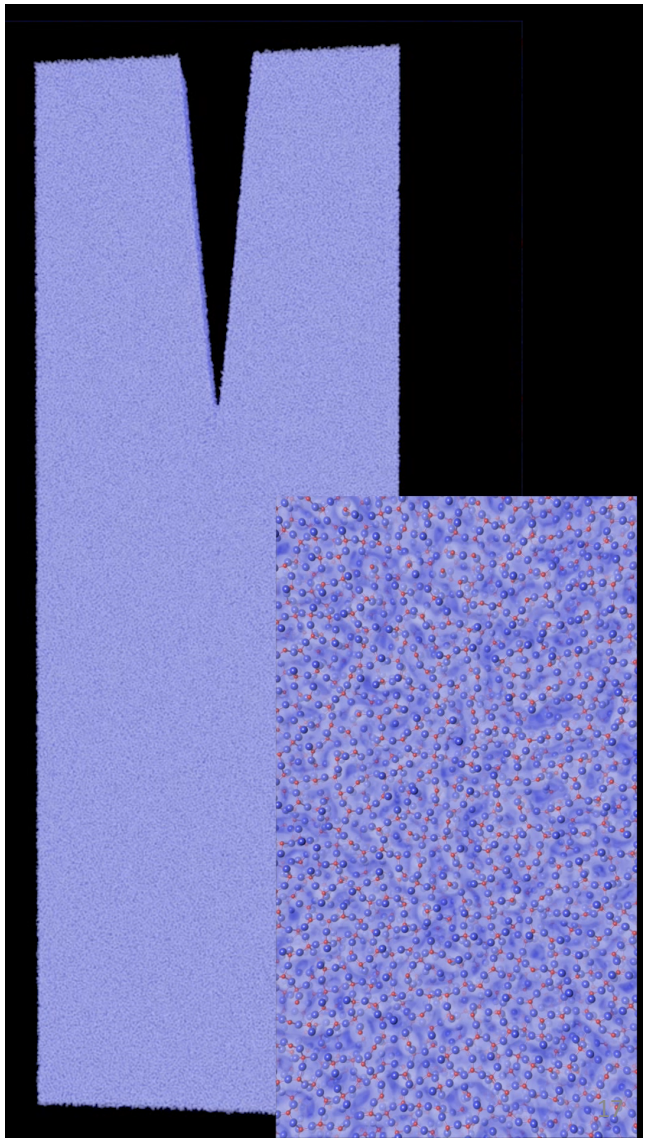
MIRA

An exascale challenge: Understanding materials failure

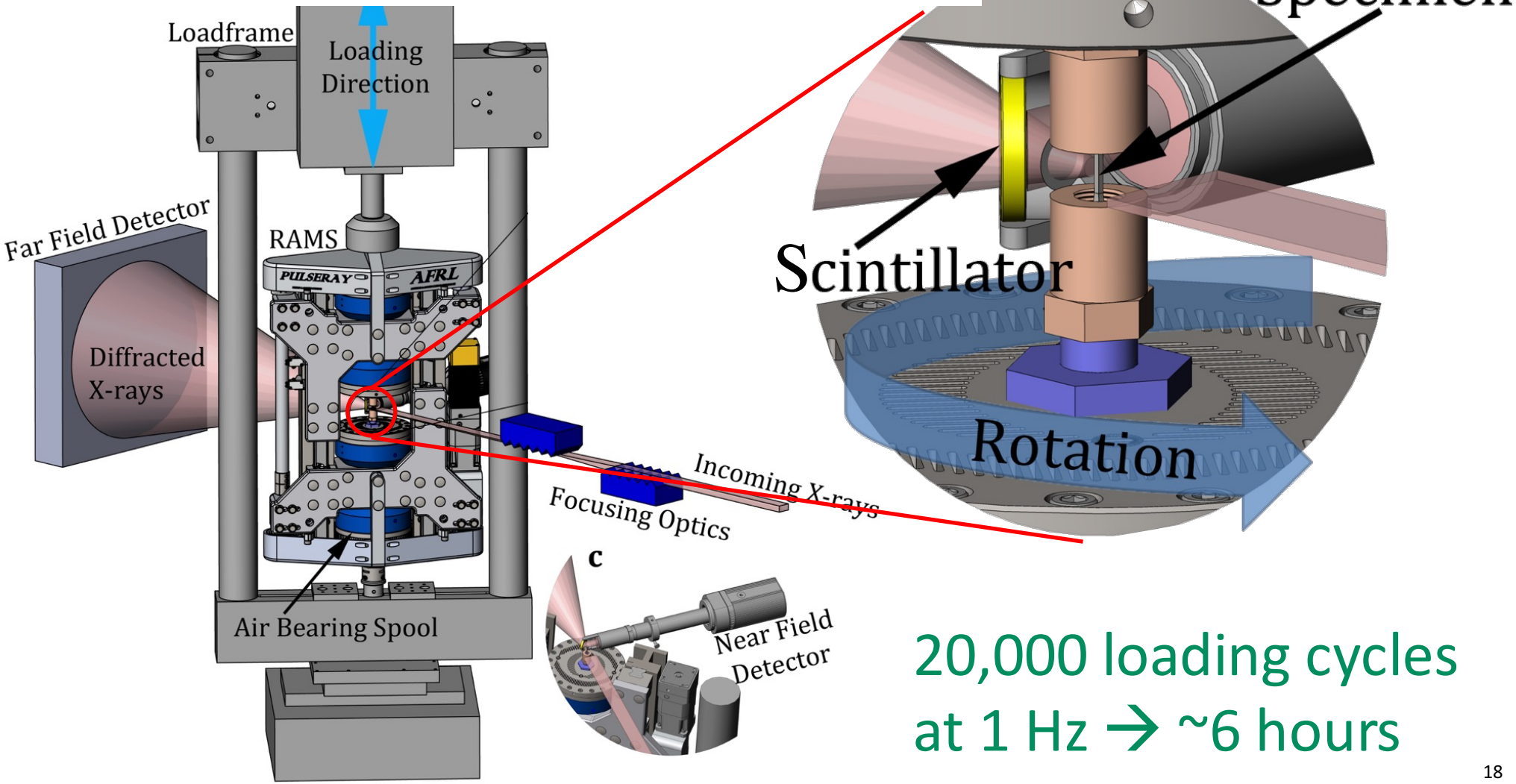
Multi-$B cost to US industry annually

Opportunity: Image failure as it occurs

- Use rapid imaging to track evolution of sample
- Use simulation to predict possible future paths
- Zoom in to likely fraction locations



AFRL Rotation and Axial Motion System device

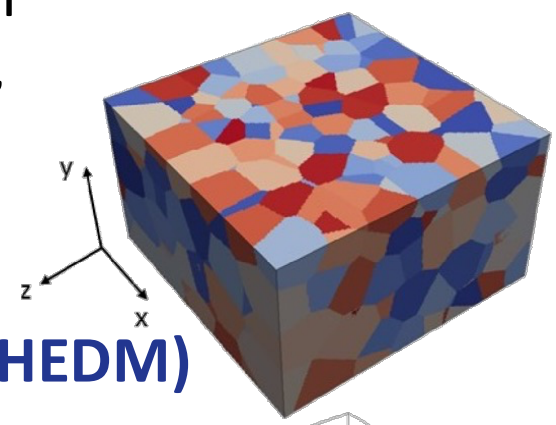


20,000 loading cycles at 1 Hz → ~6 hours

μ-CT, nf-HEDM, and ff-HEDM imaging

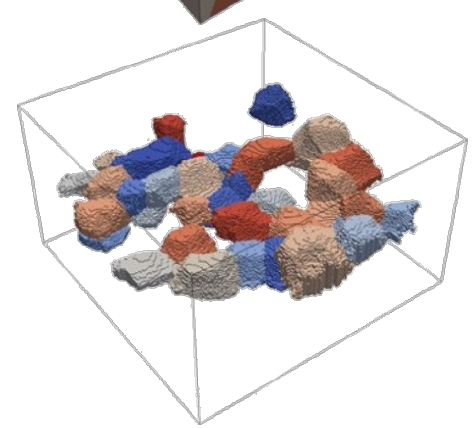
Microtomography (μ-CT) reveals inhomogeneity in electron density (e.g., presence of voids or cracks),

- Today: 4K×4K×1500×12 bits/s = 36 GB/s
- 2024: 20K×20K×8000×20 bits/s = 8 TB/s

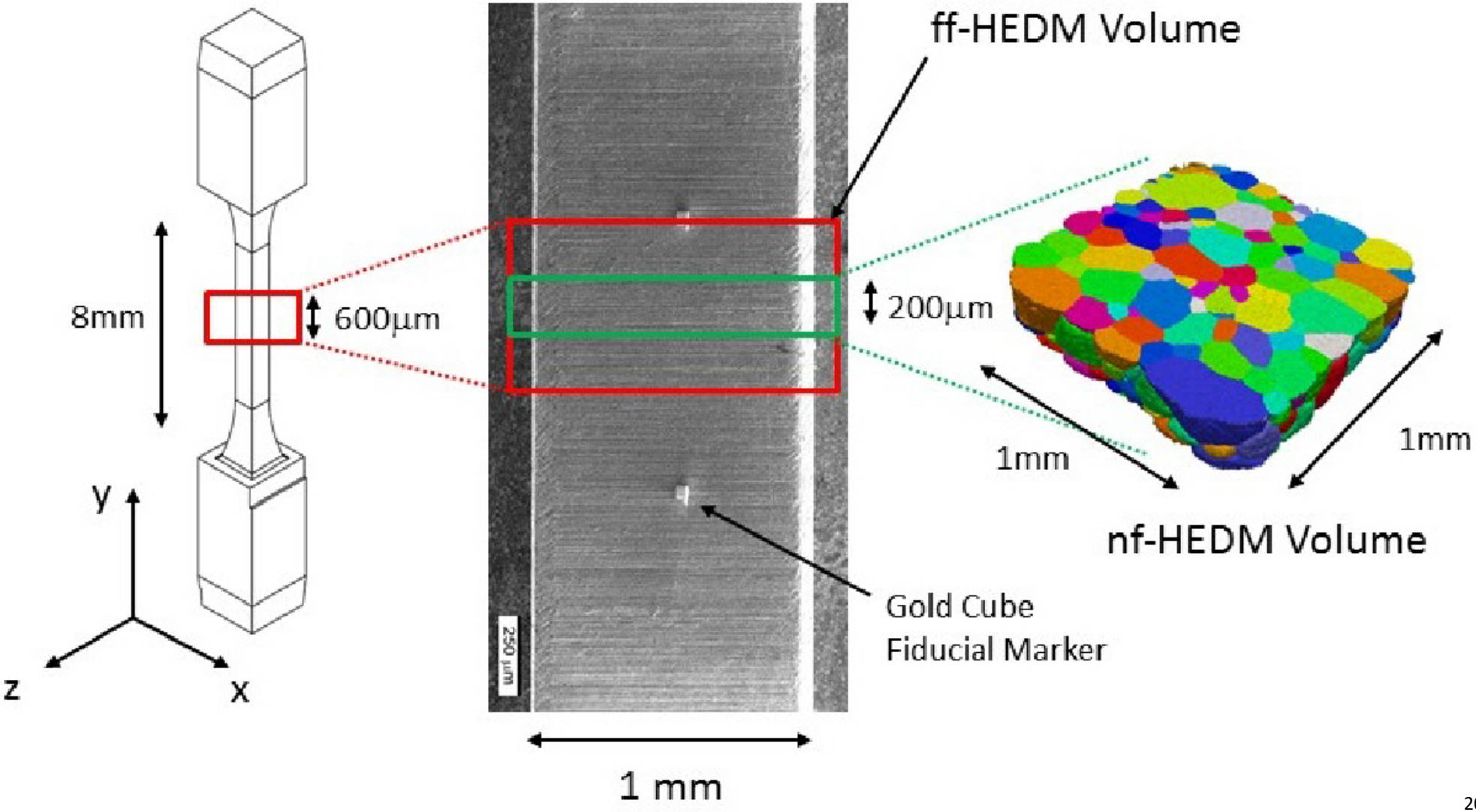


Near-field high energy diffraction microscopy (nf-HEDM) gives detailed shapes of each grain in a polycrystal

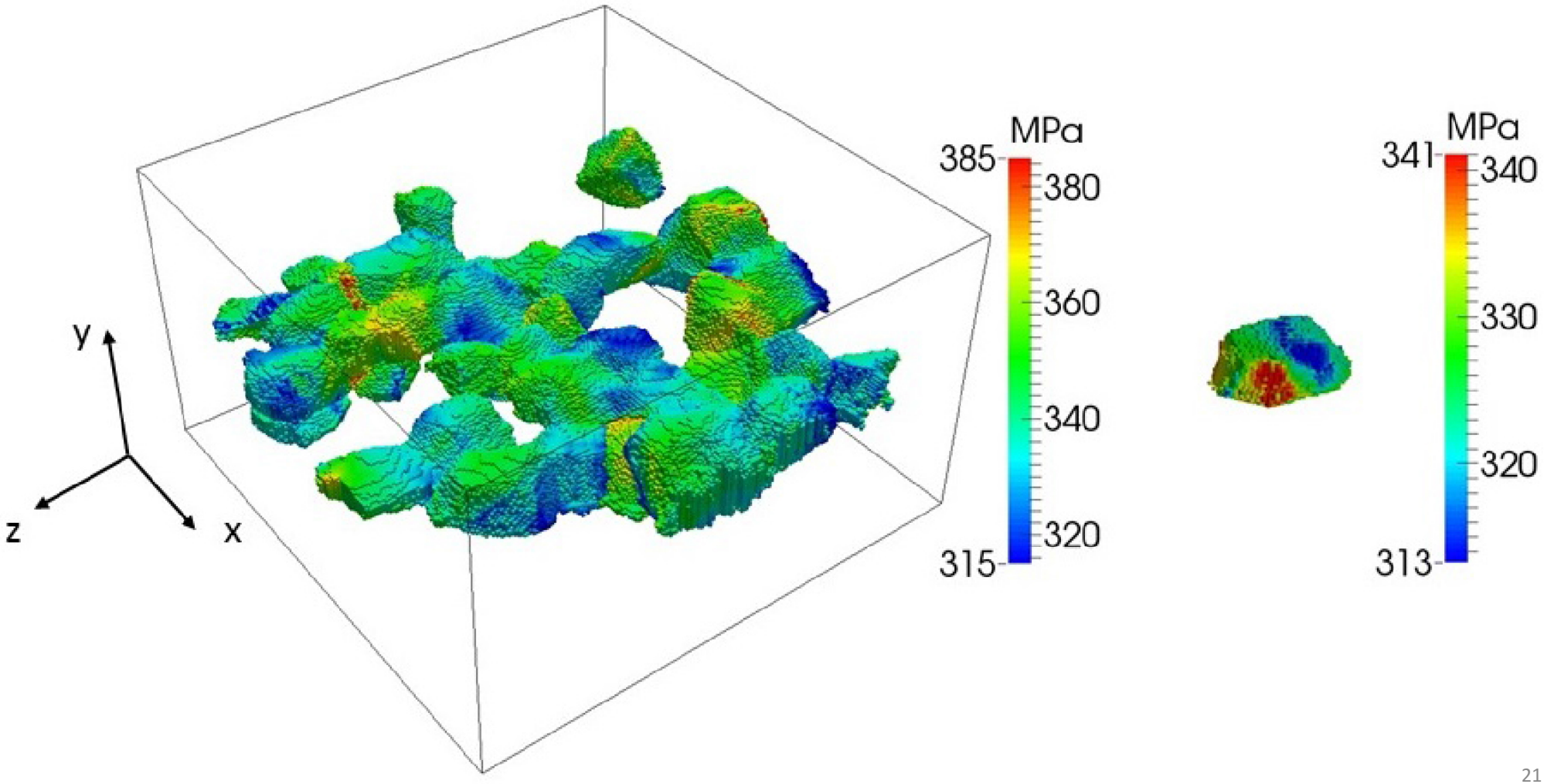
- Today: 10 μm resolution, 1000 grains → 10 GB dataset
- 2024: 0.1 μm resolution, 10,000 grains → 10^7 more data?



Far-field ff-HEDM (ff-HEDM) gives strains and orientations of each grain


8mm
600µm
y
z
x
ff-HEDM Volume
200µm
1mm
1mm
nf-HEDM Volume
Gold Cube
Fiducial Marker
250 µm
1 mm

Finite-element model of microstructure



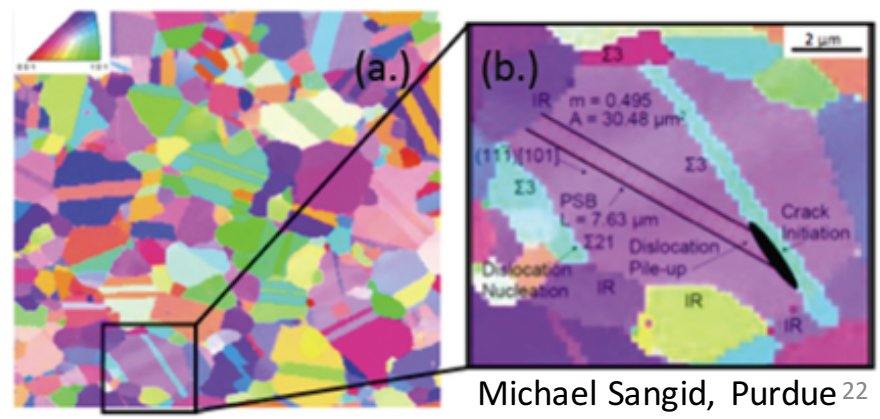
Construct, update, and run models

1) Construct an initial finite element (FE) mesh that fits the initial data
   - 100 to 10,000 grains, 1000 to 10,000 elements each → 10^5 to 10^8 elements

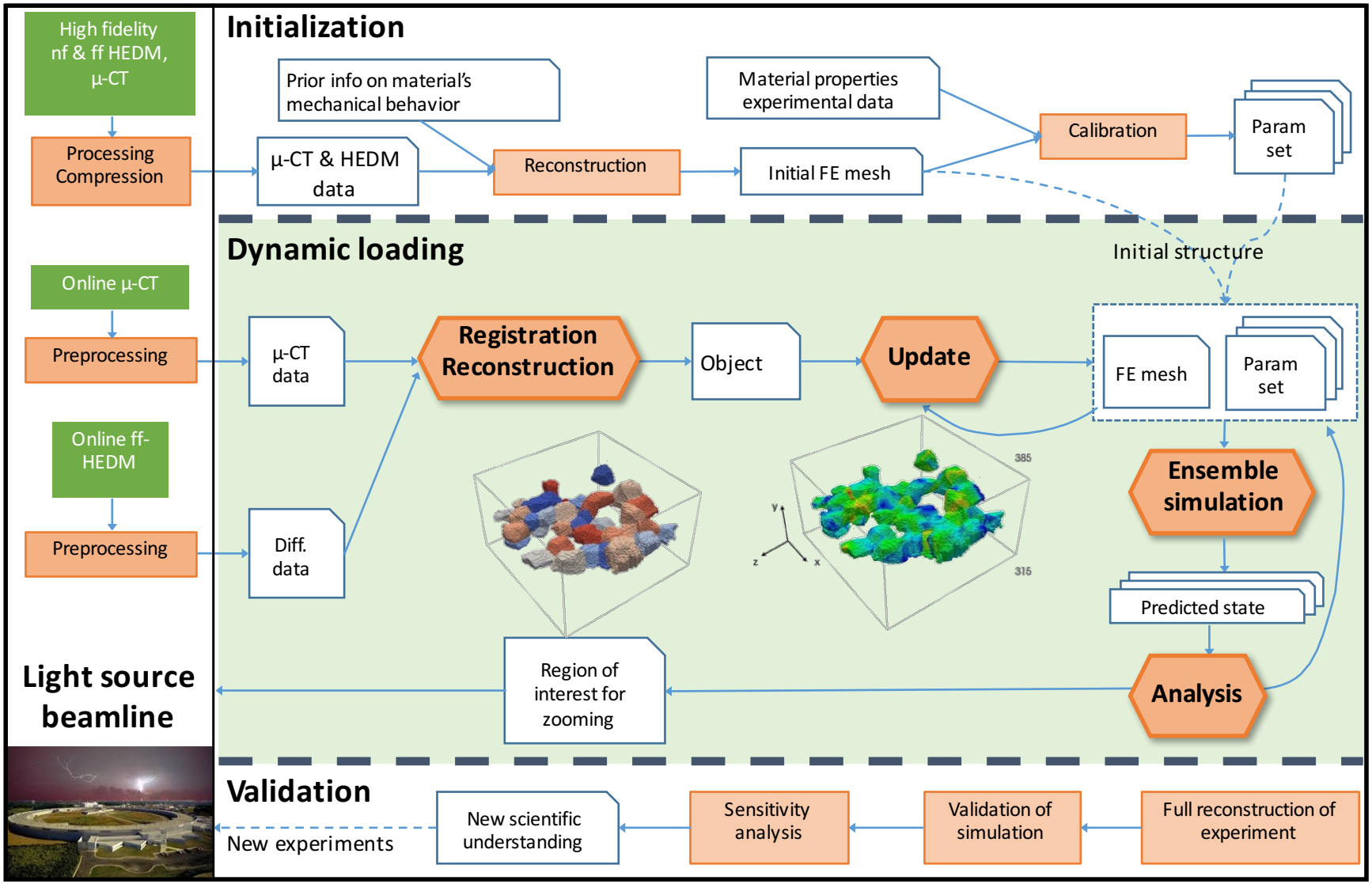
2) Estimate materials properties of grains
   - Initially, 5 or 6 parameters in all: later, could be per grain or even element

3) Update model based on μ-CT and HEDM reconstructions

4) Large ensemble FE simulations to explore possible trajectories
   - **Simulate 2+ loading cycles in < 1 second**



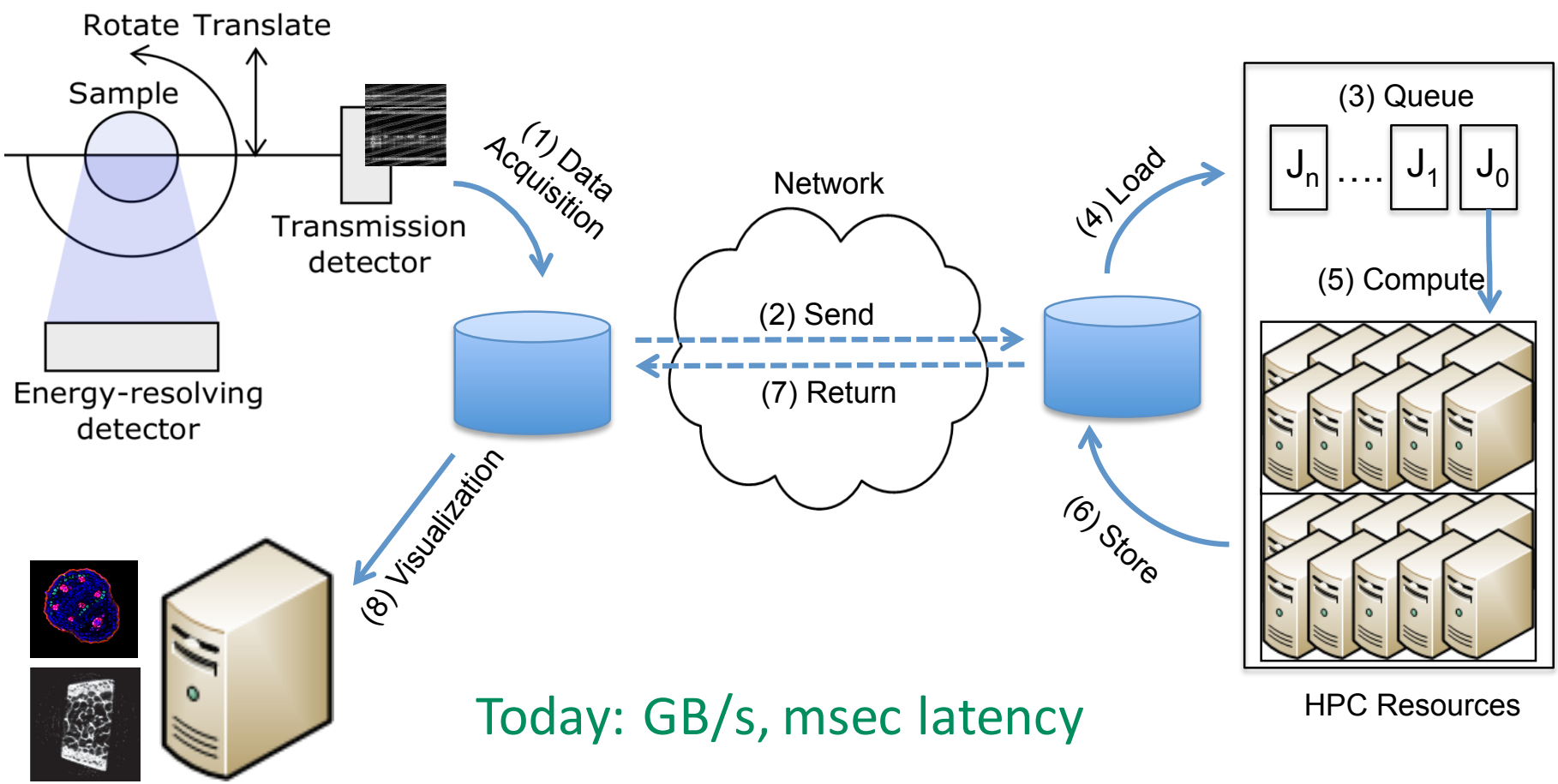
Michael Sangid, Purdue


High fidelity nf & ff HEDM, µ-CT
Processing Compression
Online µ-CT
Preprocessing
Online ff-HEDM
Preprocessing
Light source beamline
Initialization
Prior info on material's mechanical behavior
Material properties experimental data
µ-CT & HEDM data
Reconstruction
Initial FE mesh
Calibration
Param set
Dynamic loading
Initial structure
µ-CT data
Registration Reconstruction
Object
Update
FE mesh
Param set
Diff. data
Ensemble simulation
Predicted state
Region of interest for zooming
Analysis
Validation
New experiments
New scientific understanding
Sensitivity analysis
Validation of simulation
Full reconstruction of experiment

Computing requirements: Aiming for 1 Hz

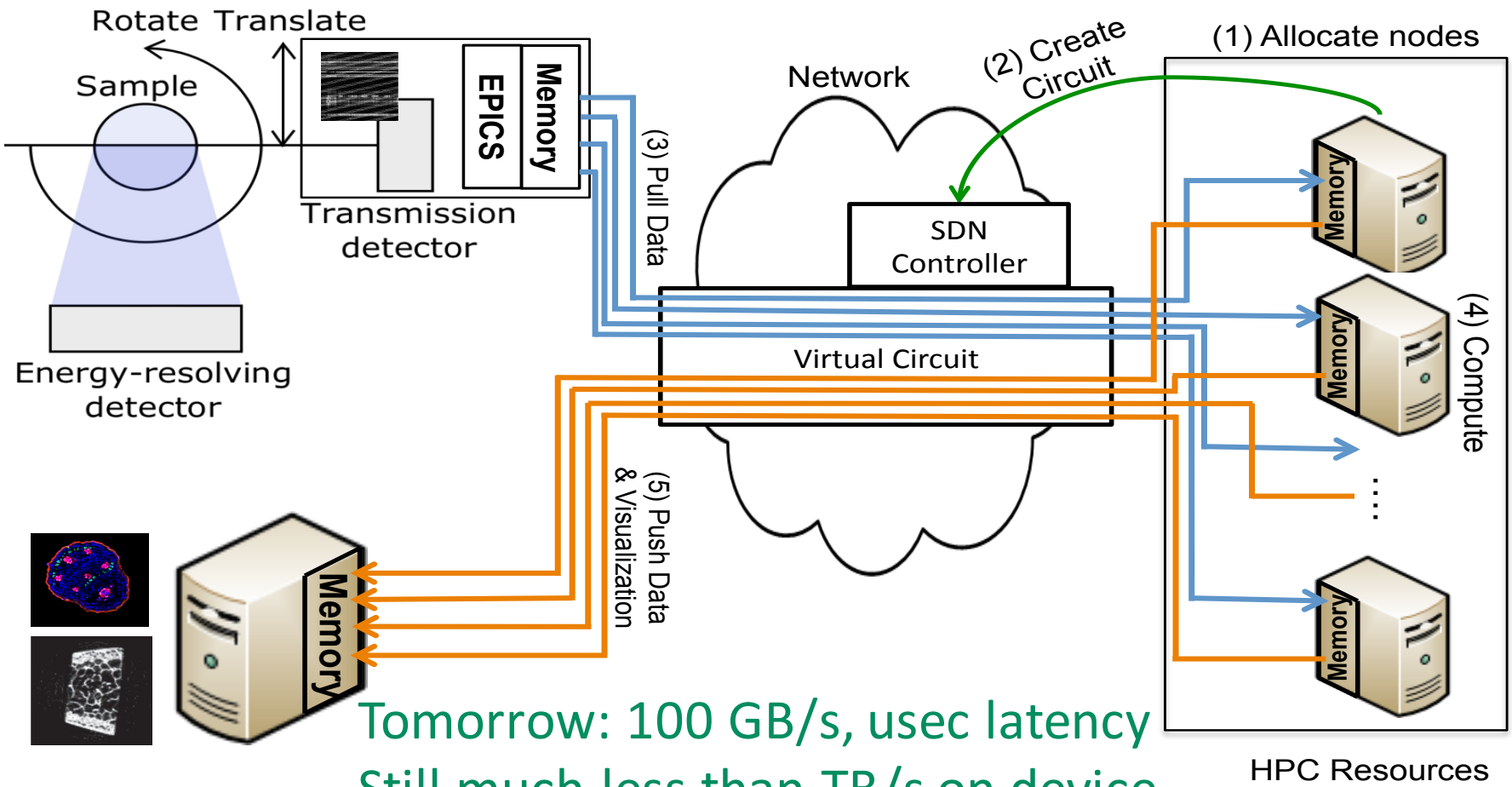
Reconstruction	**μ-CT**: $2K^3$ in 1 min on 32K cores using XXX. Need to further reduce reconstruction times and adapt methods to deal with partial data. **ff-HEDM**: 1000 grains in mins with 128K cores using MIDAS. Need to further reduce reconstruction times. **Registration**: Voronoi tesselation to estimate grain extents, determine strains, based on grain centroids and volumes from ff-HEDM
Update	AMR to reduce error in areas in simulation where large; data assimilation to update parameters defining simulation ensemble
Simulation	One loading cycle takes 150 hours on 30 cores. Want 2+ loading cycles in <1 sec → 10^6 speedup for one simulation instance.
Analysis	Extract features from ensemble of simulations, identify "regions of interest" for crack initiation, with error bars.

New architectures for high-speed data transfer



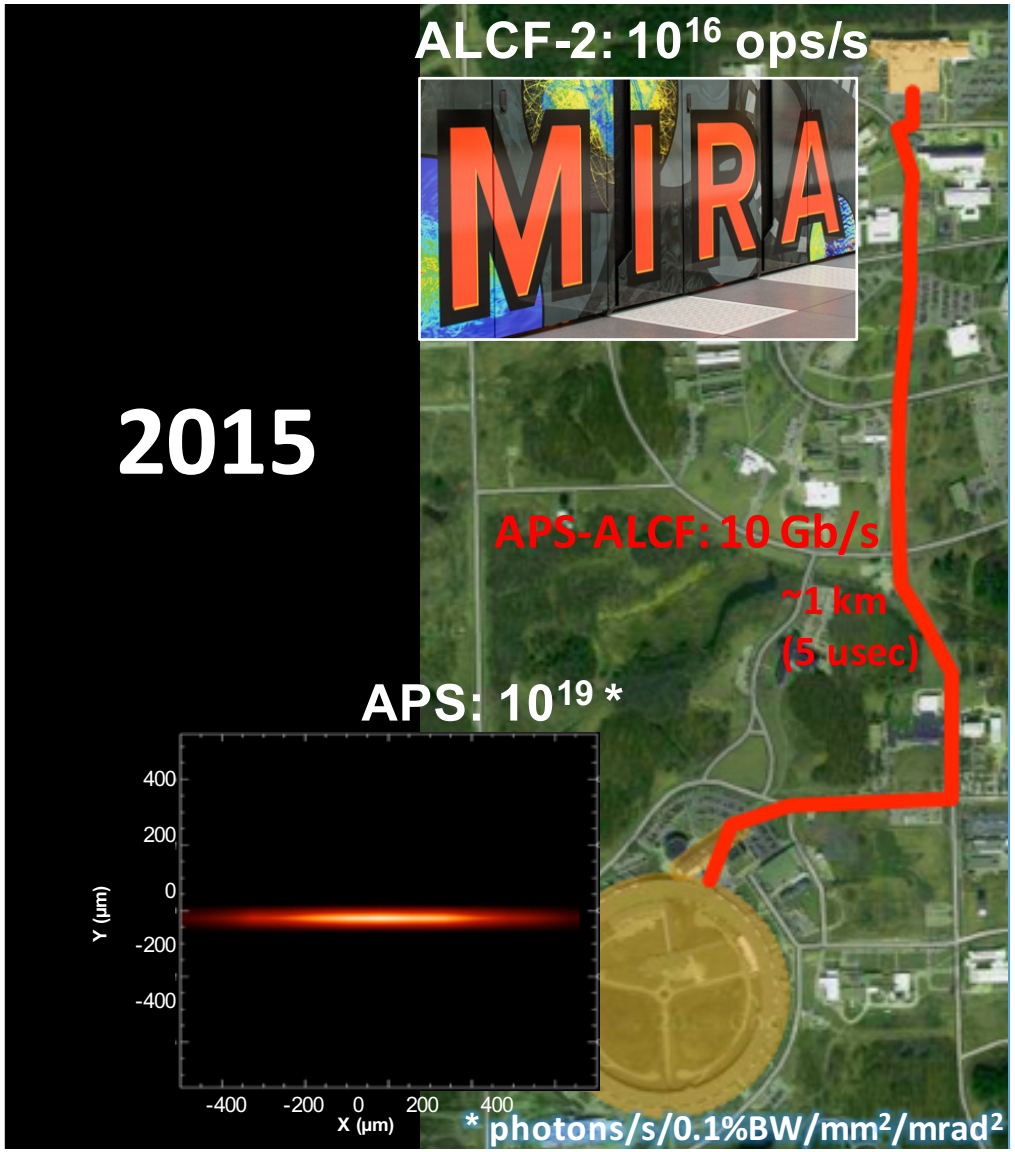
Today: GB/s, msec latency

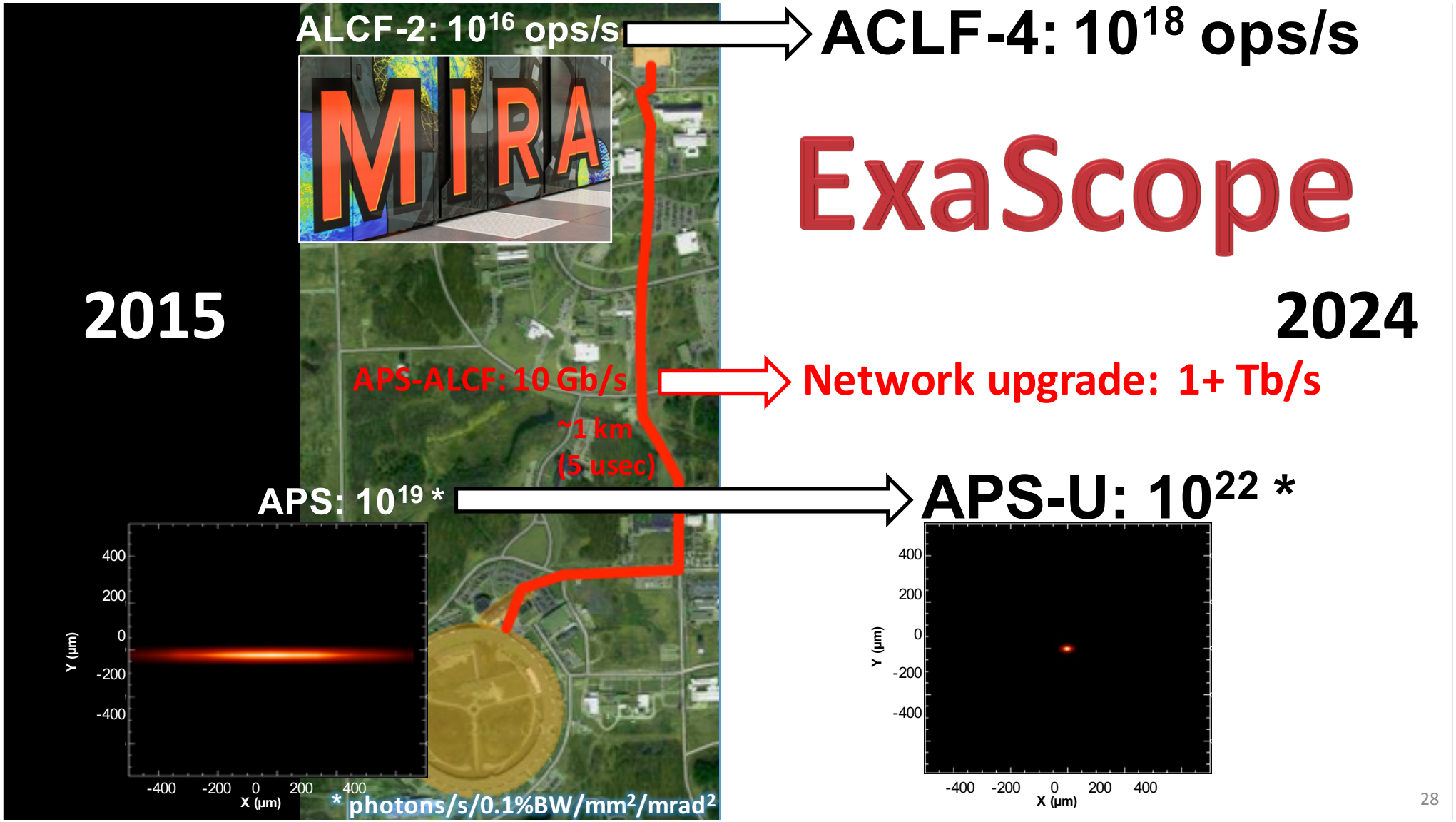
New architectures for high-speed data transfer



Tomorrow: 100 GB/s, usec latency

Still much less than TB/s on device


ALCF-2: 10^{16} ops/s
MIRA
2015
APS-ALCF: 10 Gb/s
~1 km
(5 usec)
APS: 10^{19} *
400
200
0
-200
-400
Y (μm)
-400
-200
0
200
400
X (μm)
* photons/s/0.1%BW/mm²/mrad²


ALCF-2: 10^{16} ops/s
ACLF-4: 10^{18} ops/s
MIRA
ExaScope
2015
2024
APS-ALCF: 10 Gb/s
Network upgrade: 1+ Tb/s
~1 km
(5 usec)
APS: 10^{19} *
APS-U: 10^{22} *
Y (µm)
X (µm)
* photons/s/0.1%BW/mm²/mrad²

Advances in instrumentation and experimental methodologies ➔ uniquely powerful DOE instruments can no longer function effectively without advanced computing

Integration of instruments and computers requires advances in numerical methods, algorithms, software, networks, and supercomputer h/w and s/w architecture

Thanks to: Jon Almer, Francesco de Carlo, Hemant Sharma, Brian Toby, Doga Gursoy, APS
Bill Allcock, Mike Papka, and others, ALCF
Rachana Ananthakrishnan, Tekin Bicer, Ben Blaiszik, Raj Kettimuthu, Todd Munson, Stefan Wild, Mike Wilde, Justin Wozniak, and others, MCS
Ray Osborn, MSD; Michael Sangid, Purdue; Paul Shade and Todd Turner, AFRL
Jamie Sethian, CAMERA
DOE ASCR, Argonne LDRD, NSF, NIST