The Convergence of HPC and Big Data

Pete Beckman

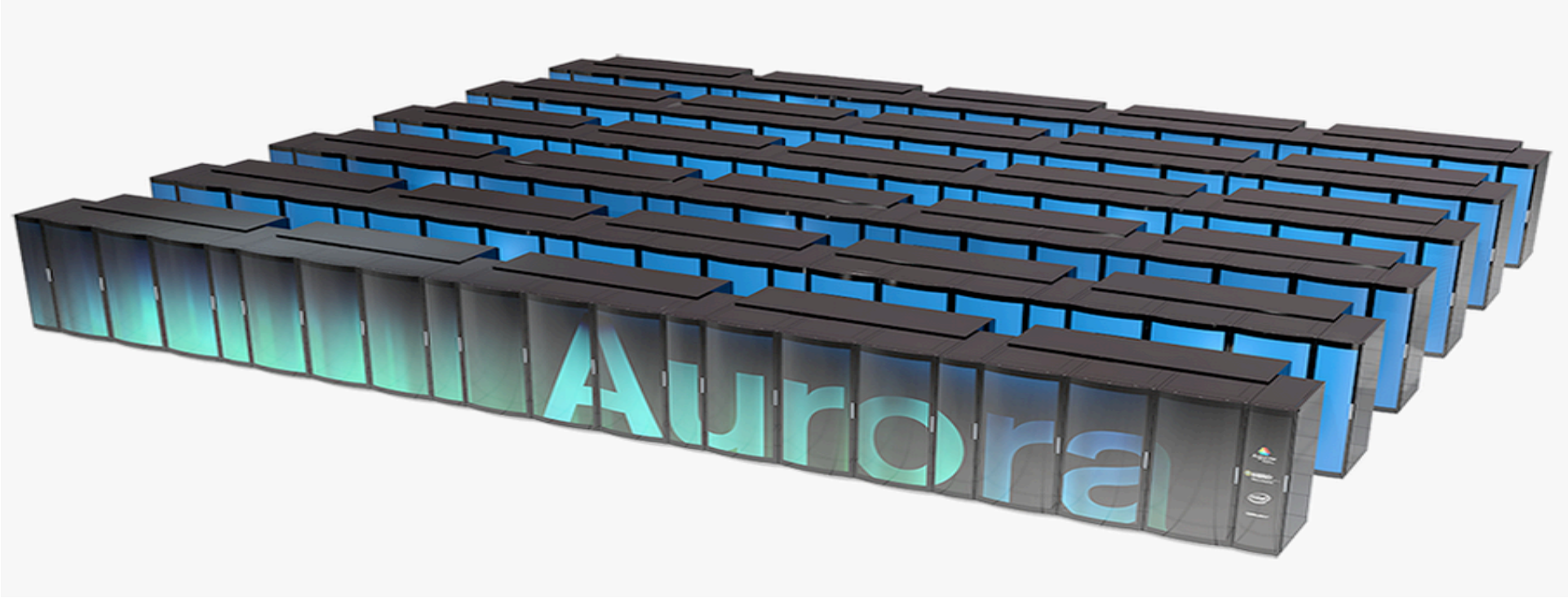
Senior Scientist, Argonne National Laboratory

Co-Director, Northwestern / Argonne Institute for Science and Engineering (NAISE)

Senior Fellow, University of Chicago Computation Institute


Argonne National Laboratory
$675M /yr budget
3,200 employees
1,450 scientists/eng
750 Ph.D.s

Argonne's Next Big Supercomputer: Aurora



Europe-USA-Asia

International series of Workshops on Extreme Scale Scientific Computing

Following the International Exascale Software Initiative
(IESP 2008-2012 ➔**Big Data and Extreme Computing** workshops (BDEC)
http://www.exascale.org/bdec/

Overarching goal:

1. Create an international collaborative process focused on the co-design of software infrastructure for extreme scale science, addressing the challenges of both extreme scale computing and big data, and supporting a broad spectrum of major research domains,
2. Describe funding structures and strategies of public bodies with Exascale R&D goals worldwide
3. Establishing and maintaining a global network of expertise and funding bodies in the area of Exascale computing

1 – BDEC Workshop, Charleston, SC, USA, April 29-May1, 2013
2 – BDEC Workshop, Fukuoka, Japan, February 26-28, 2014
3 – BDEC Workshop, Barcelona, Spain, January 28-30, 2015
4 – BDEC Workshop, Frankfurt, Germany, June 15-17, 2016

BIG DATA AND EXTREME-SCALE COMPUTING

U.S. DEPARTMENT OF ENERGY | Office of Science

ASCR Research Challenges for Convergence of Data Intensive and High-performance Computing

Summary Findings

DOE ASCR Data+Compute Convergence Summit

March 31, 2016

Rockville, MD

New discoveries in Cosmology require automation of data analysis combined with large scale simulation

Understanding climate change will require increased integration of data with simulations

Vignette: Developing better models of clouds and precipitation

- **How would you develop a better cloud model for climate simulations?**
  - We understand basic moisture physics at small scales but we can't model at that scale
  - How can we combine data from observations and small scale simulations to make a better model at climate simulation scales
- **What are some of the mathematical tools are needed here**
  - Efficient cloud-resolving atmospheric flow models
  - Bayesian parameter estimation methodology to calibrate parameters in those models
    - Reduced order modeling
    - Improved sampling methodologies
  - Multiscale techniques to derive model at larger scale from data and simulation at smaller scales
    - Models are likely to be stochastic, not deterministic
  - Numerical methods for stochastic differential equations

Panomics: addressing challenges in molecular biology

- **Panomics Challenges:**
  - Sequencing technology advancements have led to vast numbers of genomes
    - Yet the value of these data is undermined by exponential growth in discovery of protein family but linear increases in protein function validation
  - Genomics advances can now assign functional inferences to all genes in a single experiment through global analyses
  - Panomic data suffers from severe underutilization, undermining investments, and impeding DOE Missions
    - Most work done manually
  - Present databases/resources are generally either technique or organism specific
- **Computing Needs:**
  - Integrated data storage technologies for organization of disparate types of petabyte-scale and larger data
  - Computational panomics at a massive scale for deeper understanding of the impact of gene function and interactions
  - Advanced data analytics for trans-omic data, systematic knowledge discovery, and smart information synthesis
  - Computational platform for tight integration of heterogeneous data analytics and data-driven modeling & simulation

Modern sensor deployment plus HPC will lead to increased stability and predictability of the power grid

Planning and operation with more diversity in generation source and size, including increased penetration of renewables

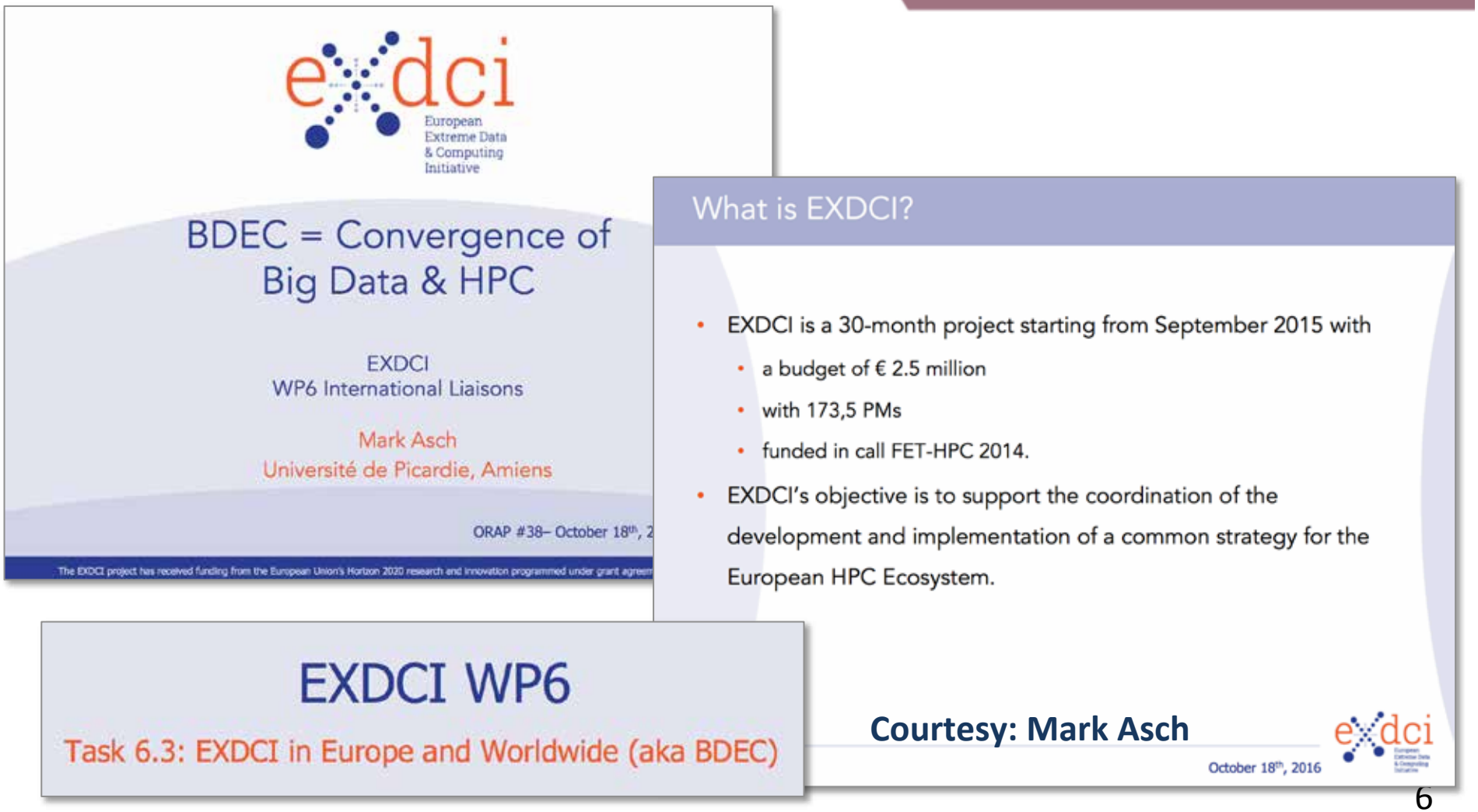
Prediction and mitigation of blackouts

Leverage increasing sensor deployment (PMUs, AMI,…) in grid management

Use of more advanced analysis, e.g. stochastic optimization for planning and operation, advanced data analytics for pattern and anomaly detection

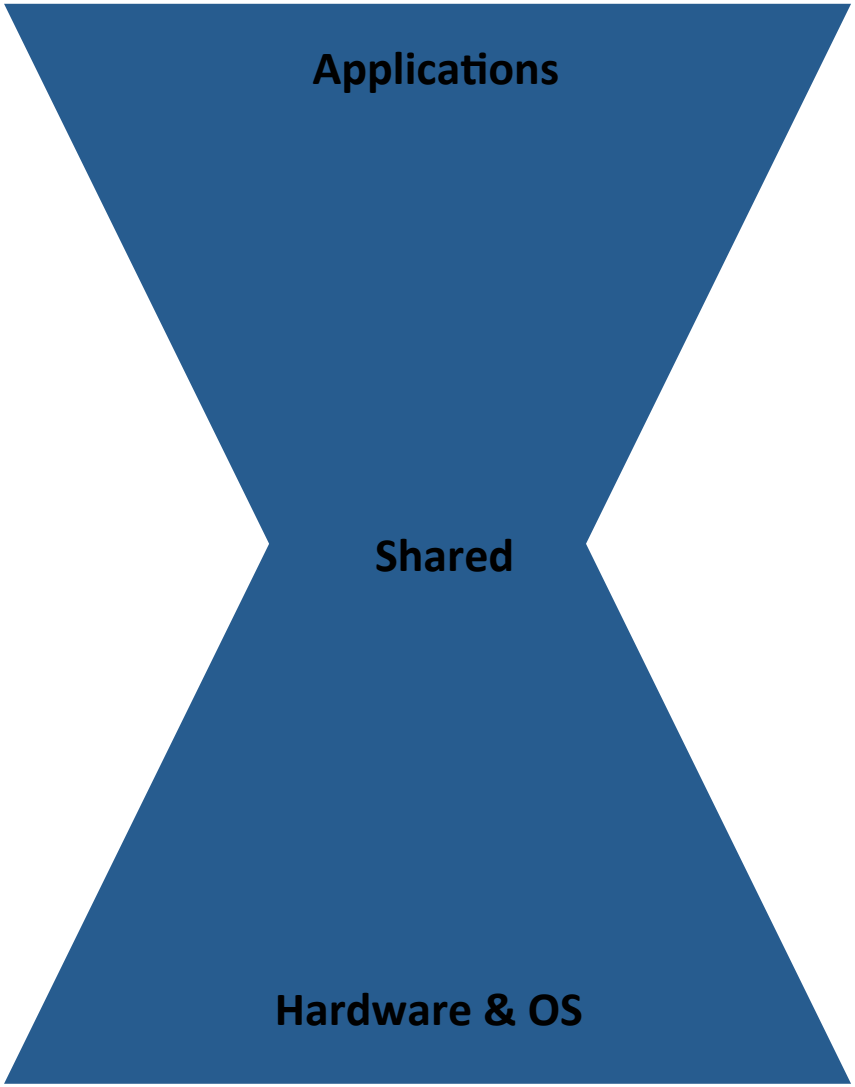
Dramatic increases in data collection rates from observation, experiments and simulation challenge math and computing

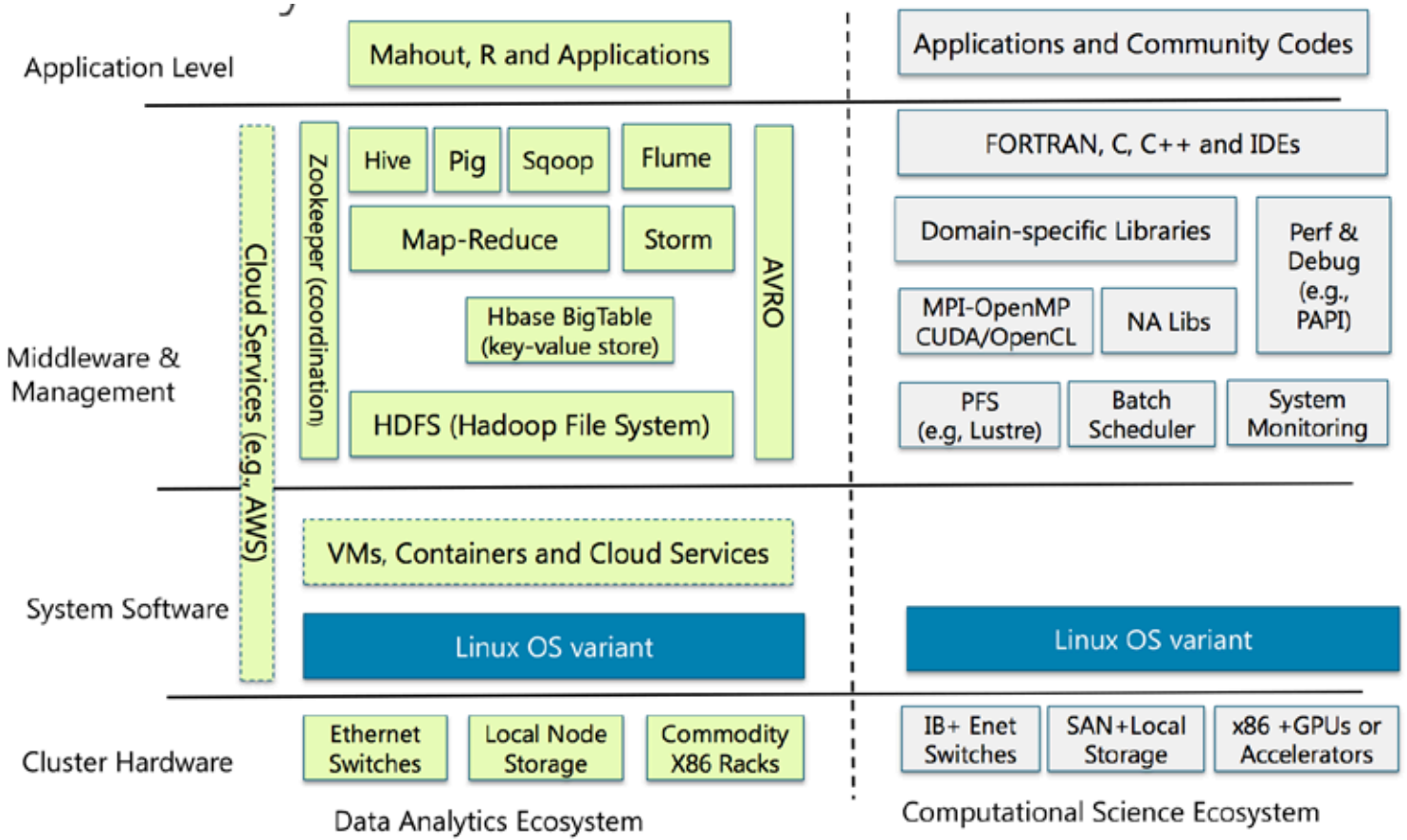
- **Address needs of new experimental techniques where the relationship between the signal the original measurements is complex**
  - Grazing-incidence small-angle x-ray scattering
  - Ptychography
  - Xray nanocrystallography
  - Fluorescence measurements in flames
- **Address increasingly complex systems where "theory" is expressed in terms of a model that can only be evaluated with HPC**
  - Relationship between cosmological parameters and structure of the universe
  - How chemical rate expressions affect the predicted emissions of a flame
  - How a cloud parameterization impacts predictions of tropical rainfall
- **Implement a modernized Data infrastructure for experimental and observational data at DOE**
  - Data-friendly computational facilities
  - Computing support for complex scientific


exdci
European
Extreme Data
& Computing
Initiative
BDEC = Convergence of
Big Data & HPC
EXDCI
WP6 International Liaisons
Mark Asch
Université de Picardie, Amiens
ORAP #38– October 18th, 2
The EXDCI project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreem
What is EXDCI?
EXDCI is a 30-month project starting from September 2015 with
a budget of € 2.5 million
with 173,5 PMs
funded in call FET-HPC 2014.
EXDCI's objective is to support the coordination of the development and implementation of a common strategy for the European HPC Ecosystem.
EXDCI WP6
Task 6.3: EXDCI in Europe and Worldwide (aka BDEC)
Courtesy: Mark Asch
October 18th, 2016
exdci

BDEC @ Europe

Courtesy: Mark Asch

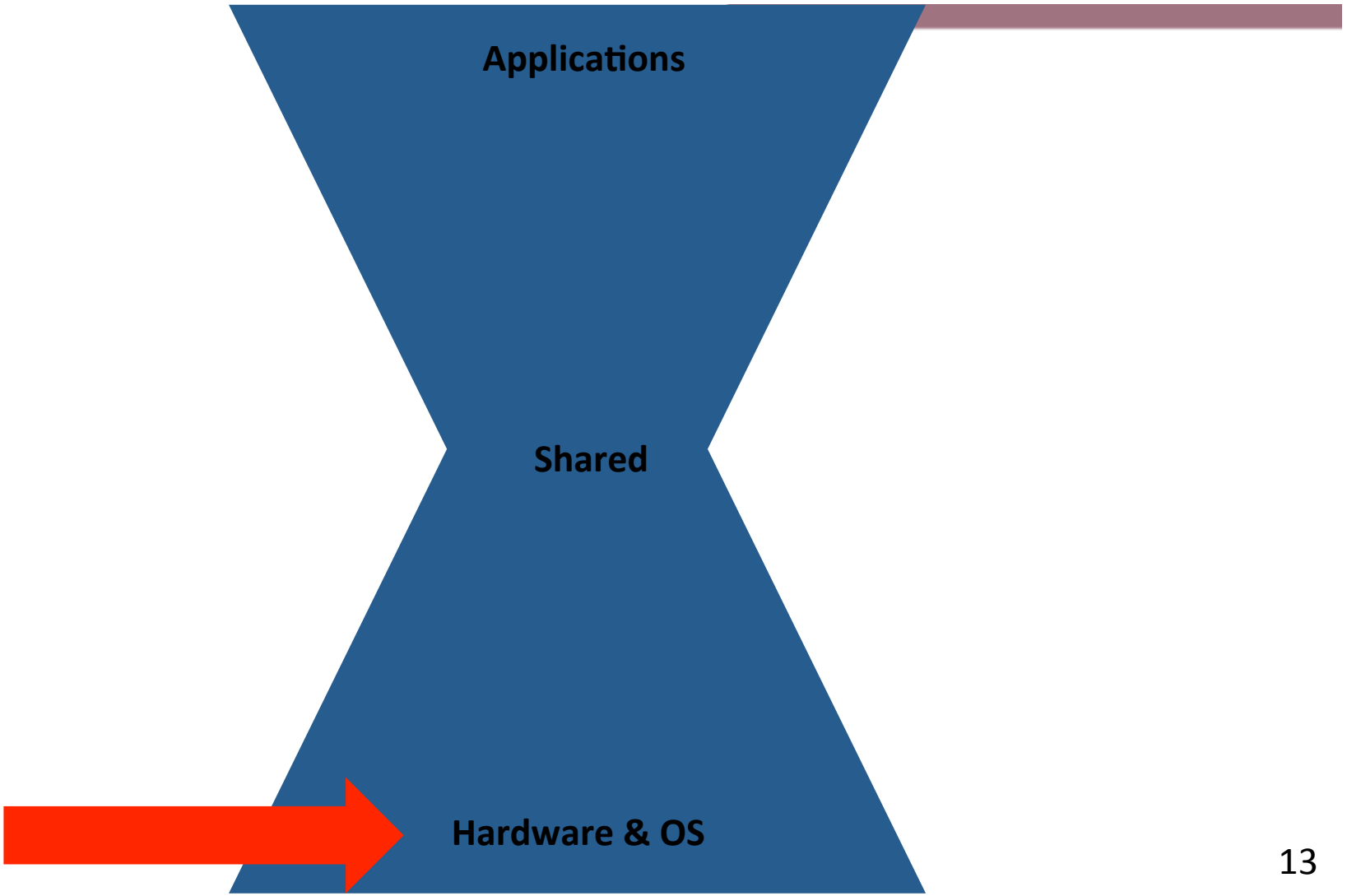

Applications
Shared
Hardware & OS


Application Level
Mahout, R and Applications
Applications and Community Codes
Middleware & Management
Cloud Services (e.g., AWS)
Zookeeper (coordination)
Hive
Pig
Sqoop
Flume
Map-Reduce
Storm
Hbase BigTable (key-value store)
HDFS (Hadoop File System)
AVRO
FORTRAN, C, C++ and IDEs
Domain-specific Libraries
Perf & Debug (e.g., PAPI)
MPI-OpenMP CUDA/OpenCL
NA Libs
PFS (e.g, Lustre)
Batch Scheduler
System Monitoring
System Software
VMs, Containers and Cloud Services
Linux OS variant
Linux OS variant
Cluster Hardware
Ethernet Switches
Local Node Storage
Commodity X86 Racks
IB+ Enet Switches
SAN+Local Storage
x86 +GPUs or Accelerators
Data Analytics Ecosystem
Computational Science Ecosystem

Why Converge?

Independent paths: More Cost, Less Science,

- $ multiple hardware software infrastructures
- $ developing software for two communities
- $ learning two computing models
- $ smaller discovery community, fewer ideas
- Less science


Applications
Shared
Hardware & OS

An Exascale Operating System and Runtime Software Research & Development Project

Developing vendor neutral, open-source OS/R software

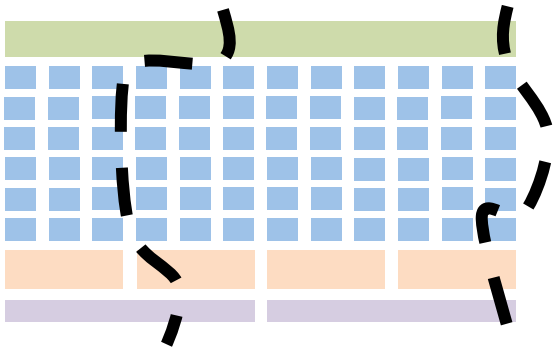
ANL: **Pete Beckman** (*PI*), Marc Snir (*Chief Scientist*), Pavan Balaji, Rinku Gupta, Kamil Iskra, Franck Cappello, Rajeev Thakur, Kazutomo Yoshii
LLNL: Maya Gokhale, Edgar Leon, Barry Rountree, Martin Schulz, Brian Van Essen
PNNL: Sriram Krishnamoorthy, Roberto Gioiosa
UC: Henry Hoffmann
UIUC: Laxmikant Kale, Eric Bohm, Ramprasad Venkataraman
UO: Allen Malony, Sameer Shende, Kevin Huck
UTK: Jack Dongarra, George Bosilca, Thomas Herault

See http://www.argo-osr.org/ for more information

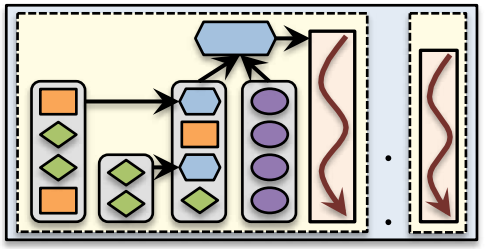
What OS/R Gaps ***Must*** We Address?

- **Extreme in-node parallelism**
  - Poor mechanisms for precise resource management (cores, power, memory, network)
  - Legacy threads/tasks implementations perform poorly at scale
- **Dynamic variability of platform; Power is constrained**
  - Poor runtime mechanisms for managing dynamic overclocking, provisioning power, adjusting workloads
  - No mechanisms for managing power dynamically, globally, and in cooperation with user-level runtime layers
- **Hierarchical memory**
  - Poor interfaces / strategies for managing deepening memory
- **New modes for HPC**
  - No portable interfaces for easily building workflows, in-situ analysis, coupled physics, advanced I/O, application resilience

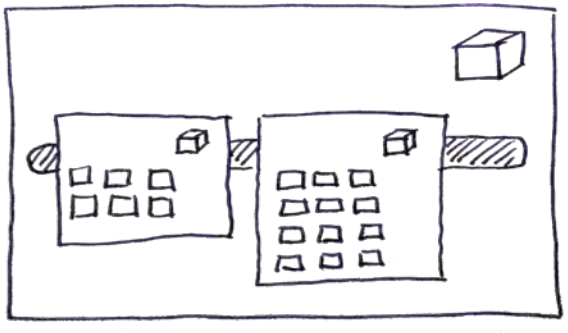
Argo Explorations to Address Exascale Gaps



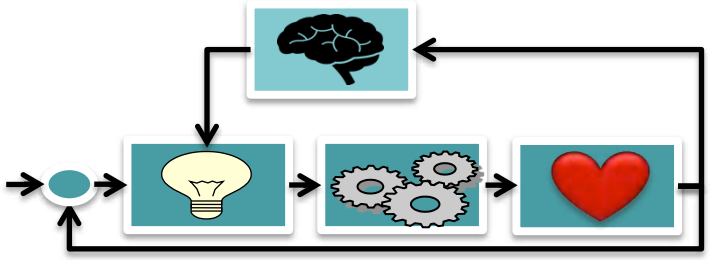
Elastic intranode containers with resource knobs



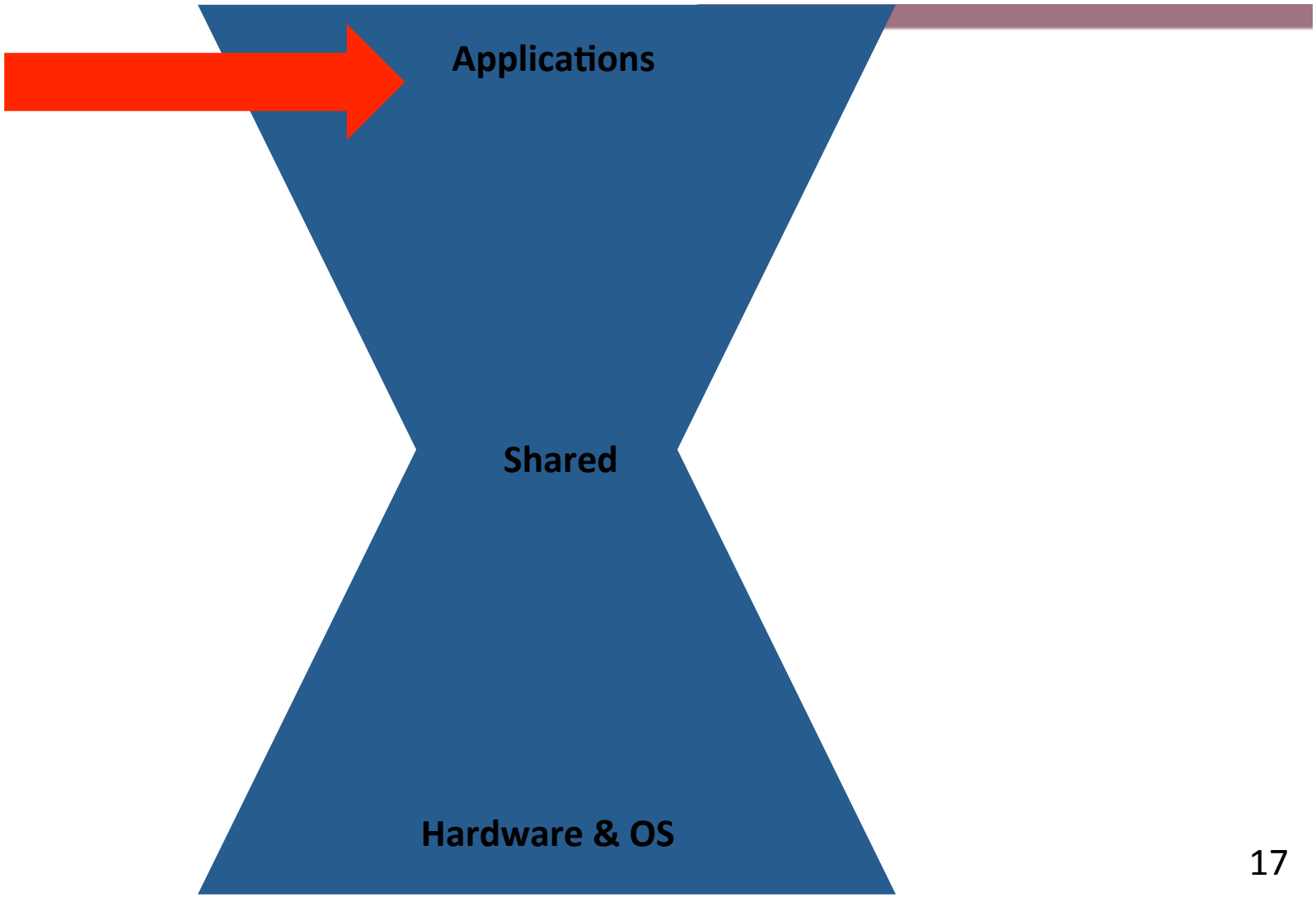
Lightweight thread/tasks designed for containers, messaging, and memory hierarchy



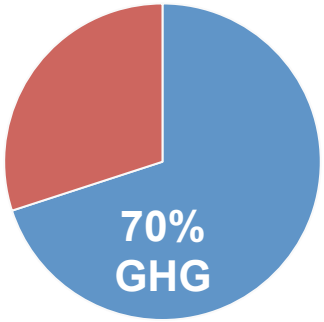
Hierarchy of *Enclaves* connected via a *Backplane*

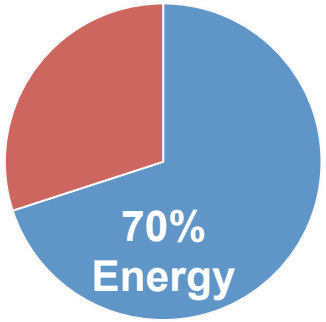


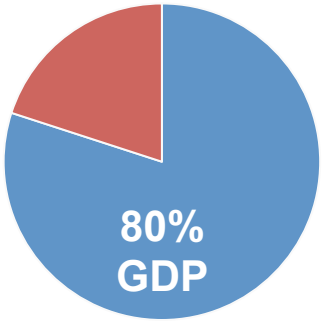
Adaptive, learning, integrated control system

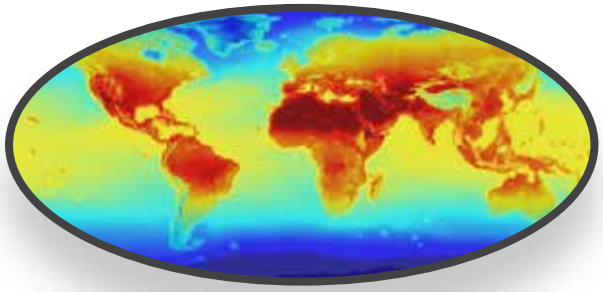

Applications
Shared
Hardware & OS

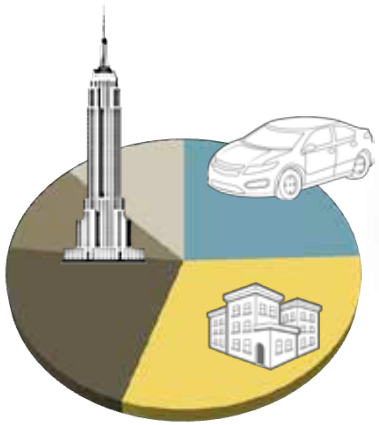
Understanding Cities



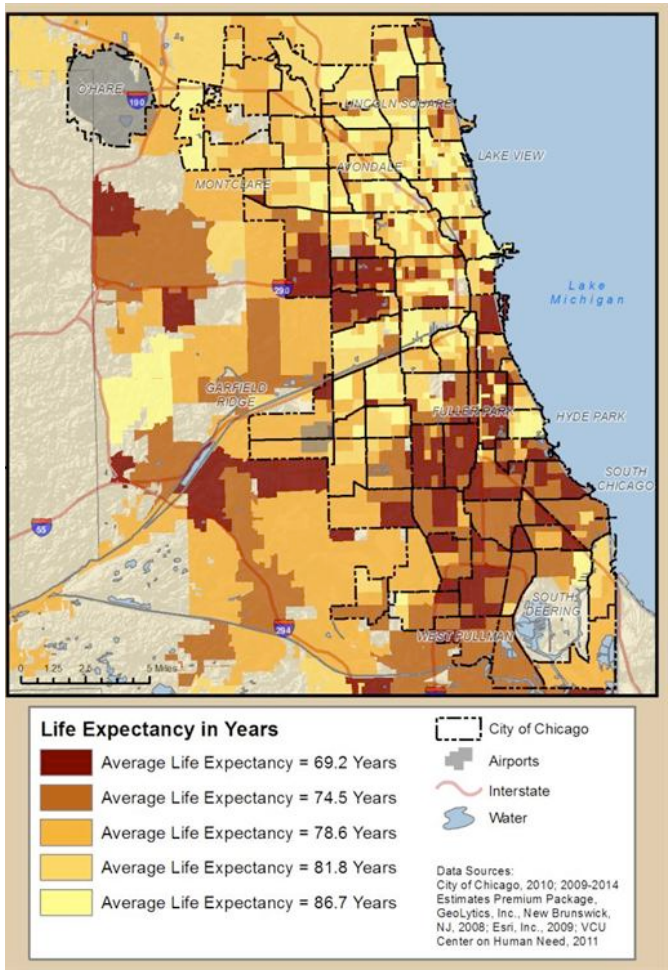




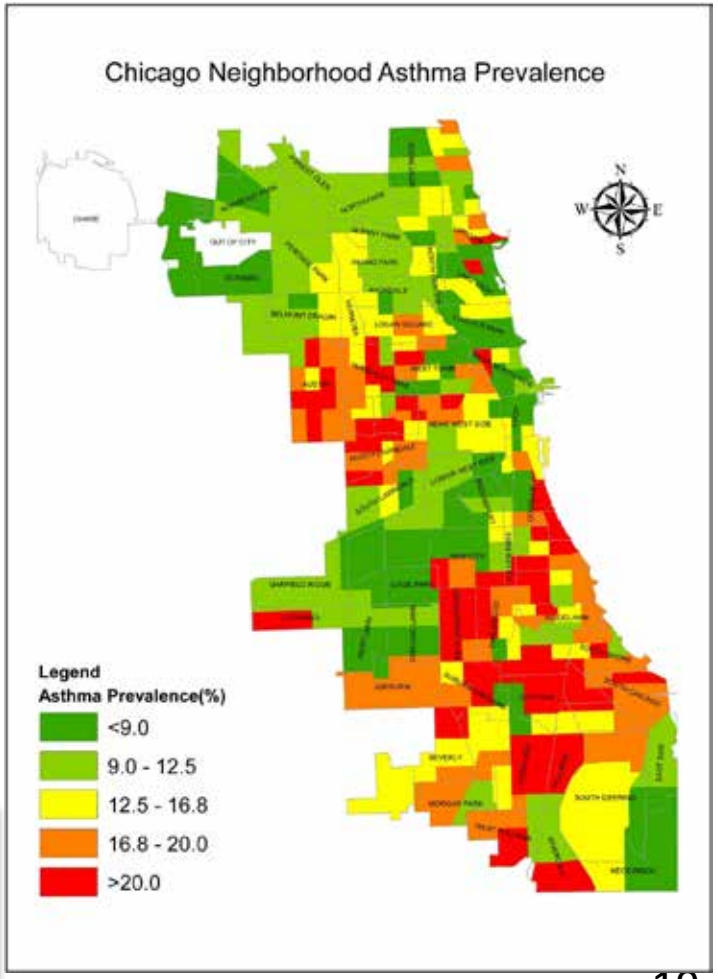


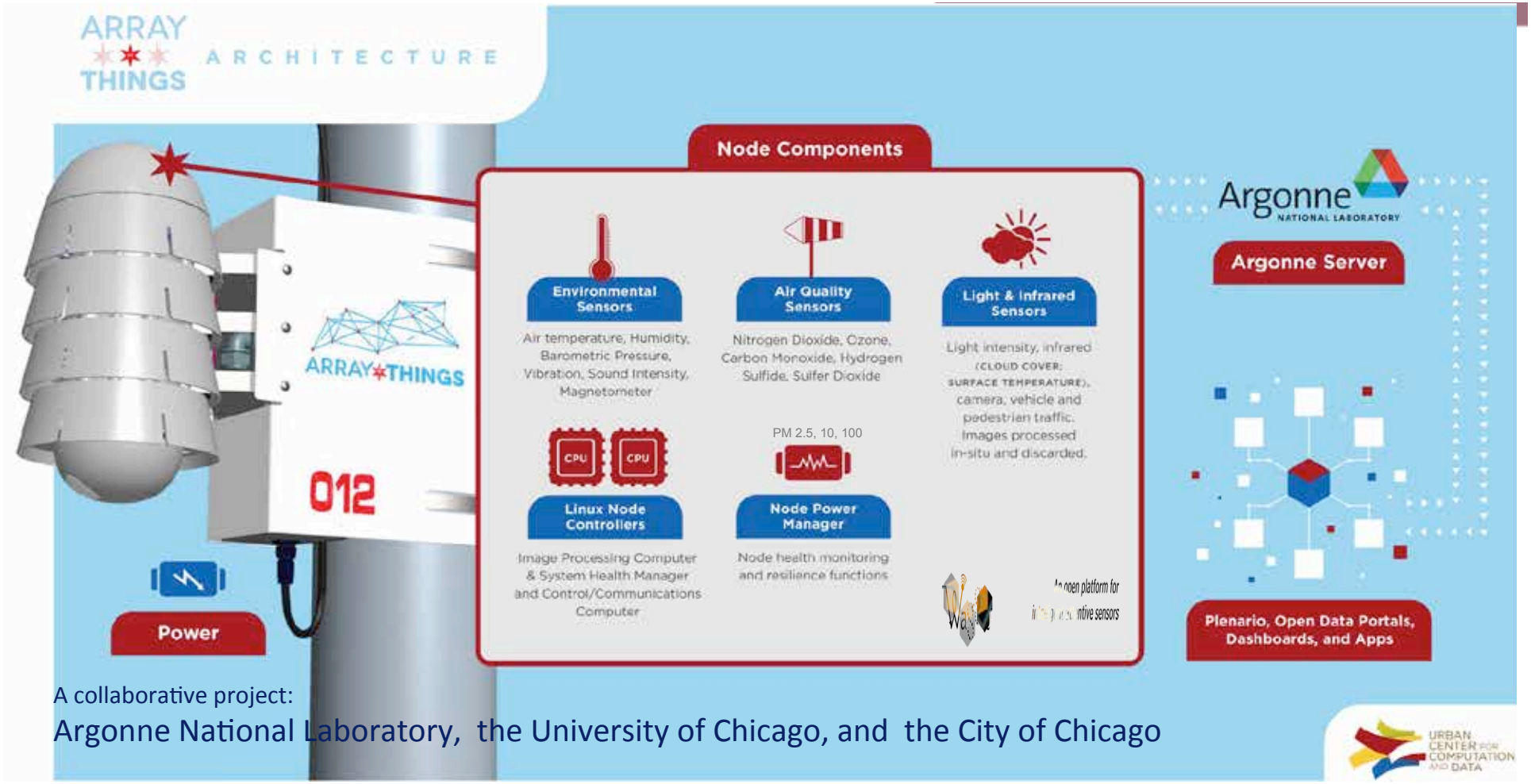


Why measure cities?



VISION ZERO
FOCUS CITIES
Austin | Boston | Chicago | DC | Fort Lauderdale | LA | NYC | Portland | San Francisco | Seattle





A collaborative project:

Argonne National Laboratory, the University of Chicago, and the City of Chicago

Supported by collaborating institutions and the U.S. National Science Foundation.
Industry In-Kind partners: AT&T, Cisco, Intel, Microsoft, Motorola Solutions, Schneider Electric, Zebra


THE UNIVERSITY OF
CHICAGO

Waggle: An Open Platform for ***Intelligent*** Sensors

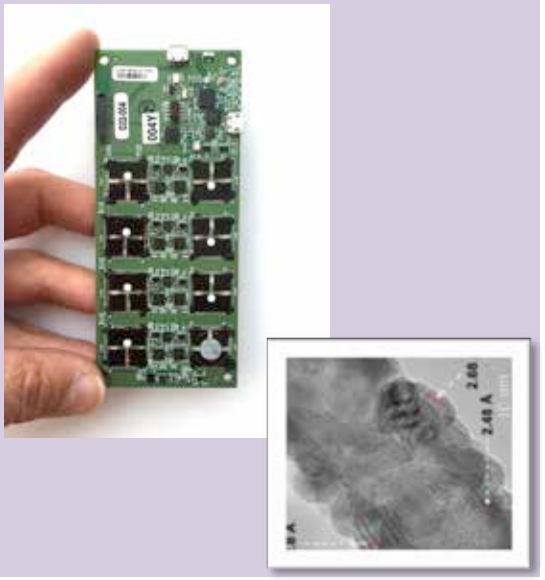
Exploiting Disruptive Technology, *Edge Computing*, Resilient Design

Machine Learning

Computer Vision

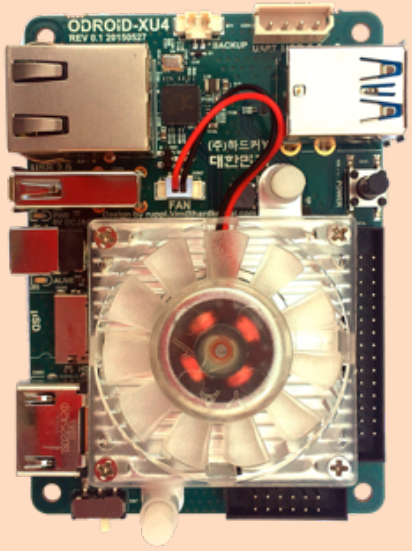
Novel Sensors

Nano / MEMS



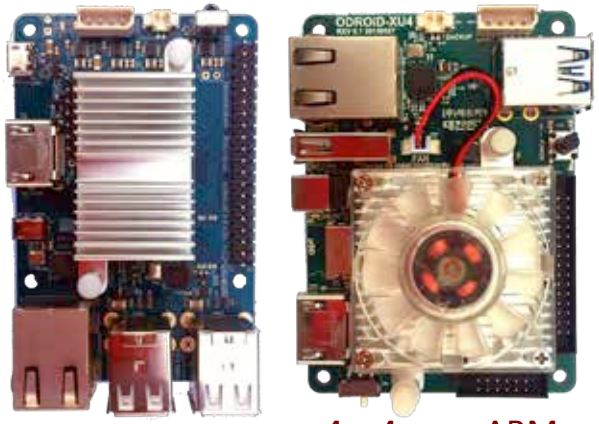
Low Power CPUs

GPU / Smartphones



Powerful, Resilient & Hackable

Multiple boot media (µSD / eMMC)



4-core ARM

4 + 4-core ARM
8-core GPU

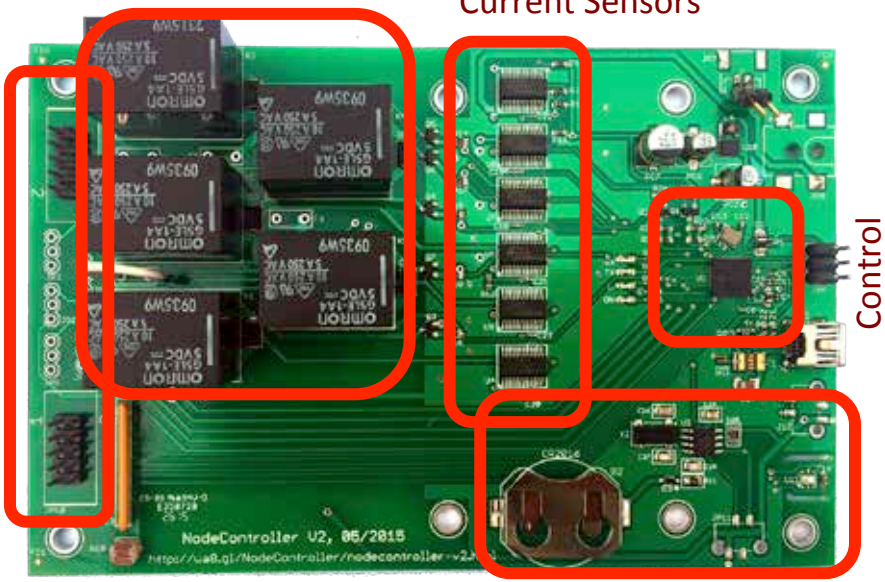
Node Control &
Communications

In-Situ / Edge
Processing

Linux Development Environment

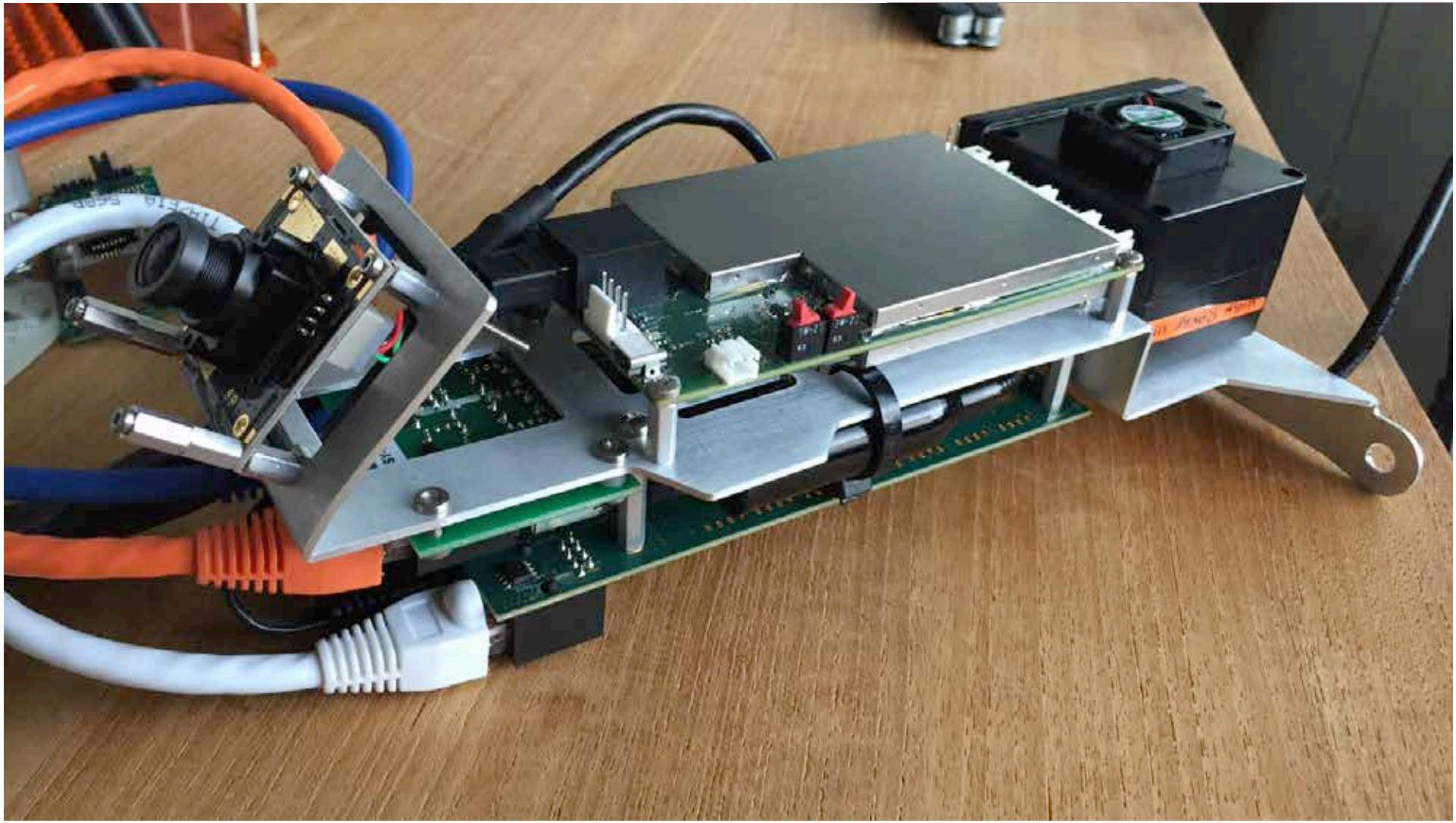
Relays

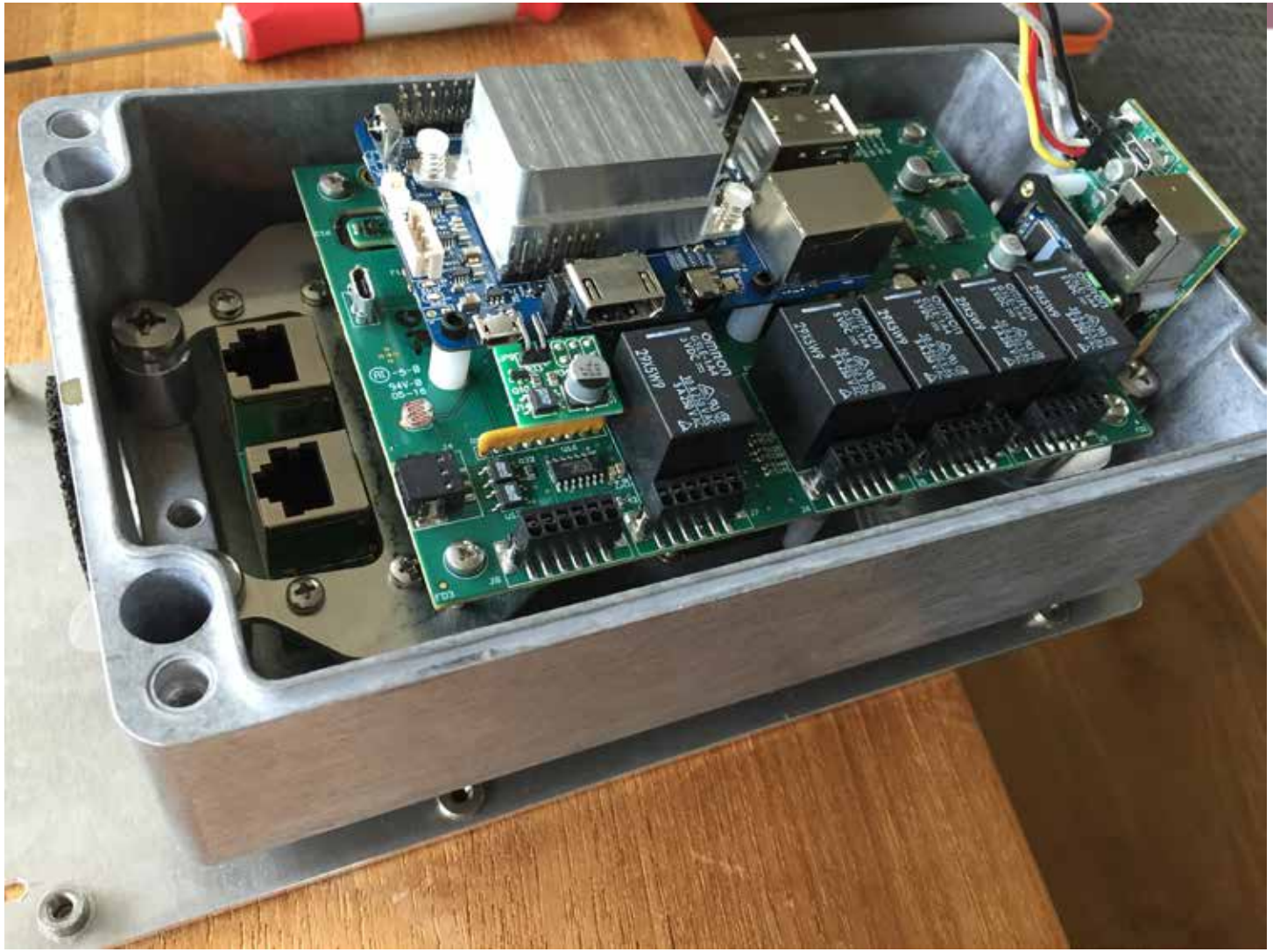
Current Sensors

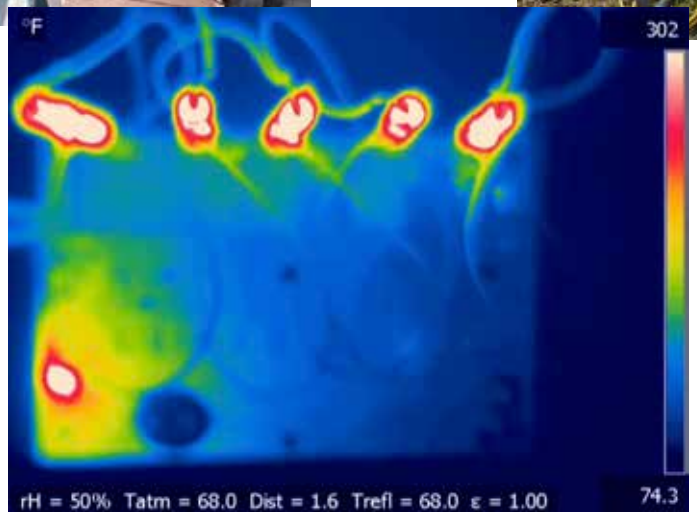


Real time clock & Internal sensors

"Deep Space Probe" Design







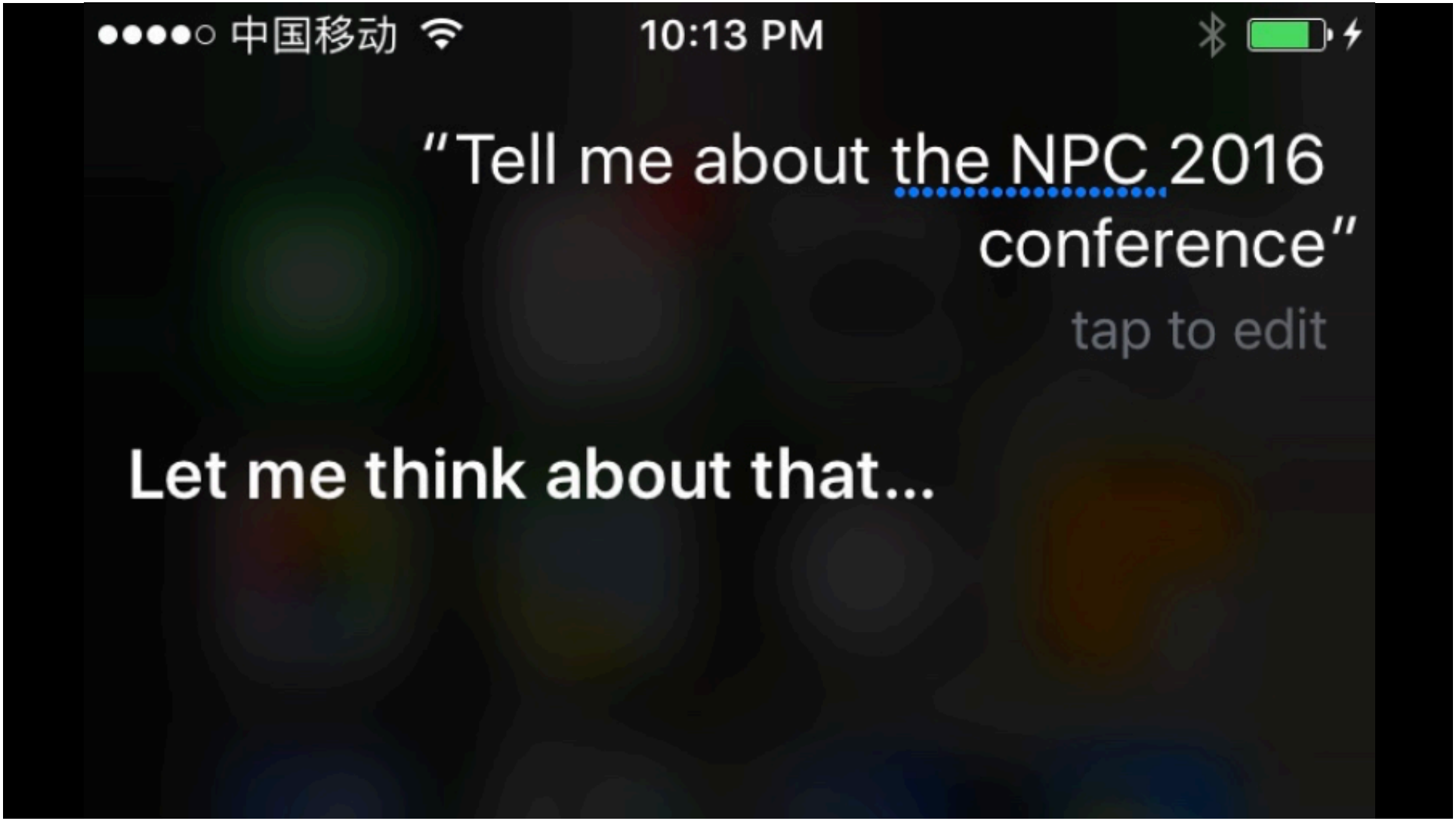
Waggle / AoT
Robust Testing

Edge Computing: Analysis and Feature Recognition

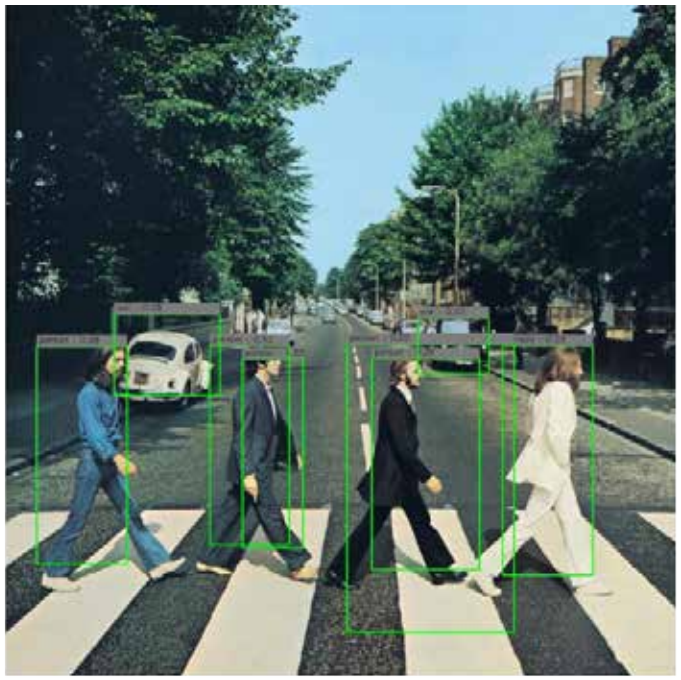
Preserving Privacy…

- Parallel Computing
- Open Platform
- Deep Learning


HDR
PANO
SQUARE
PHOTO
VIDEO
SLO-MO

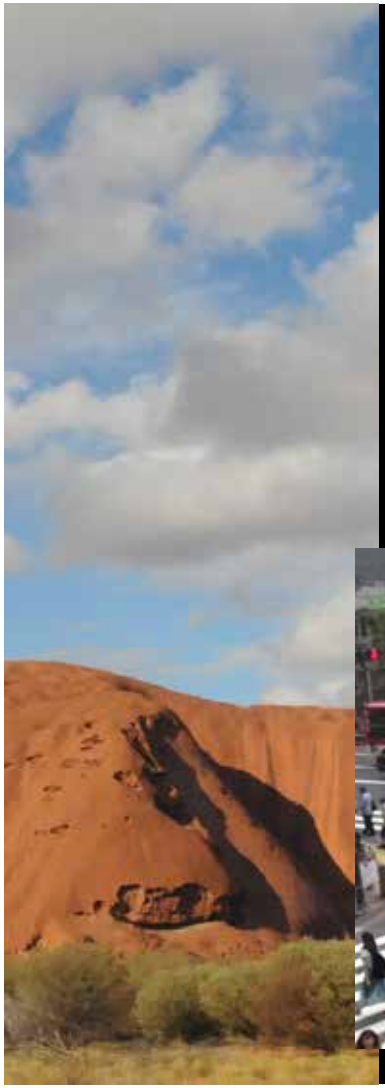

●●●●○ 中国移动
10:13 PM
“Tell me about the NPC 2016 conference”
tap to edit
Let me think about that...

Waggle Machine Learning & Edge Computing



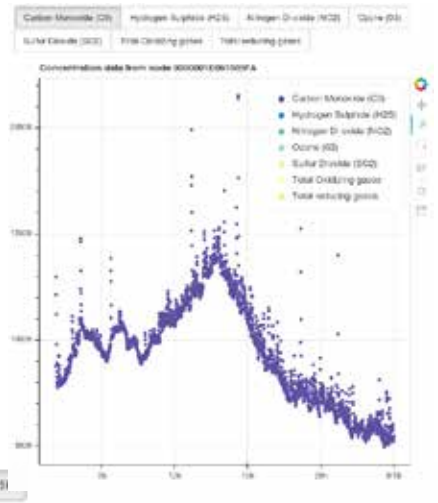
https://waggle-sensor.slack.com/files/noaholsman/F243LQL66/output.jpg

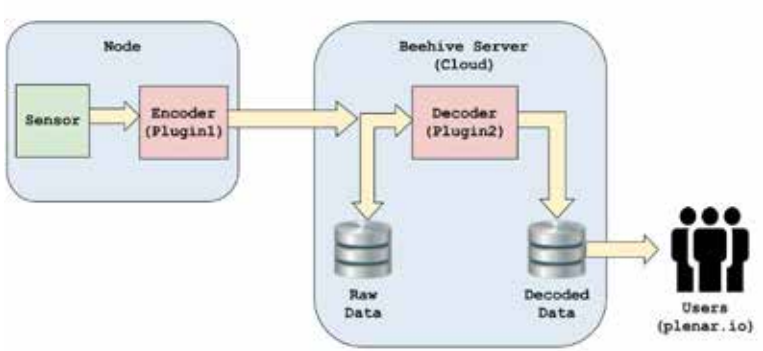
- We are exploring Caffe & OpenCV
  - Convolutional Neural Networks
- Training will be done on systems at Argonne
- Classification on Waggle

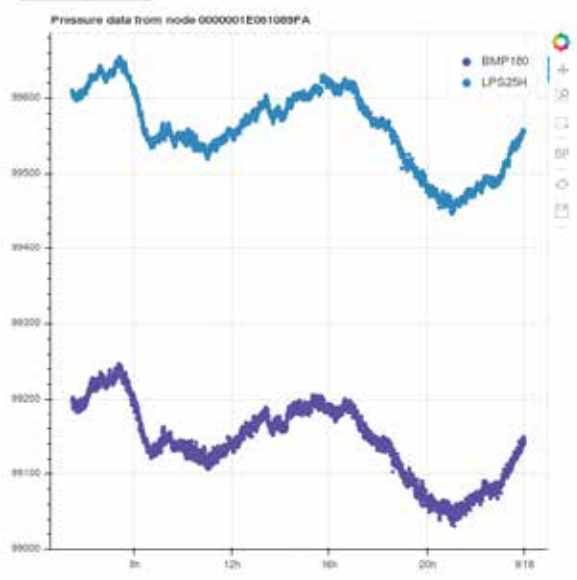


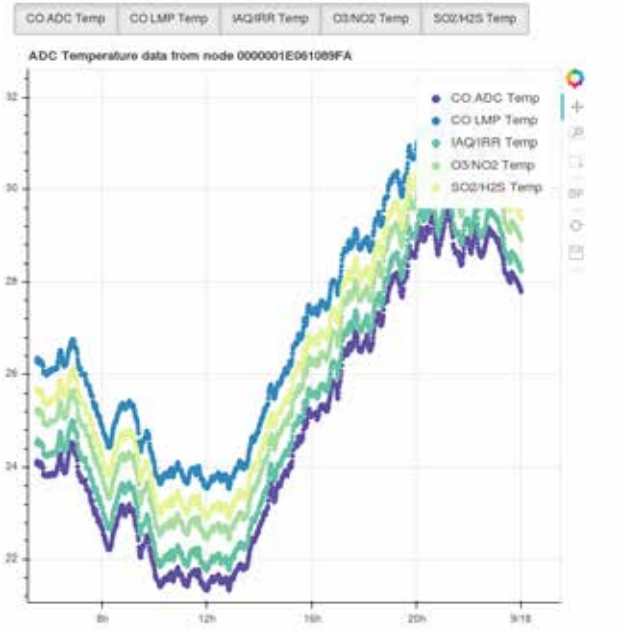

9 FT 6 IN

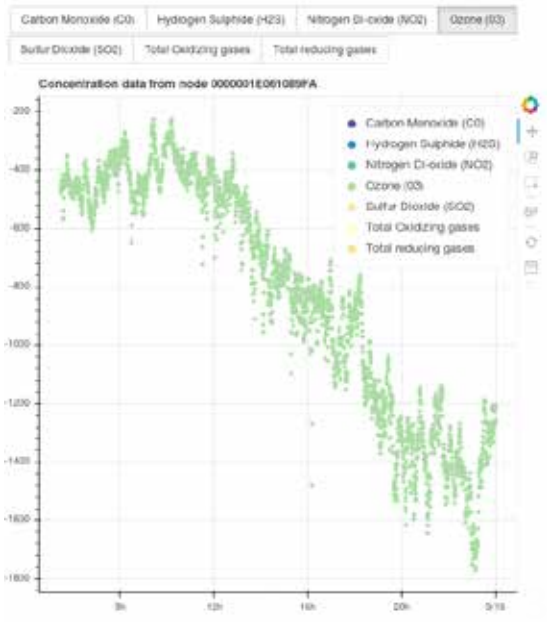
The Data



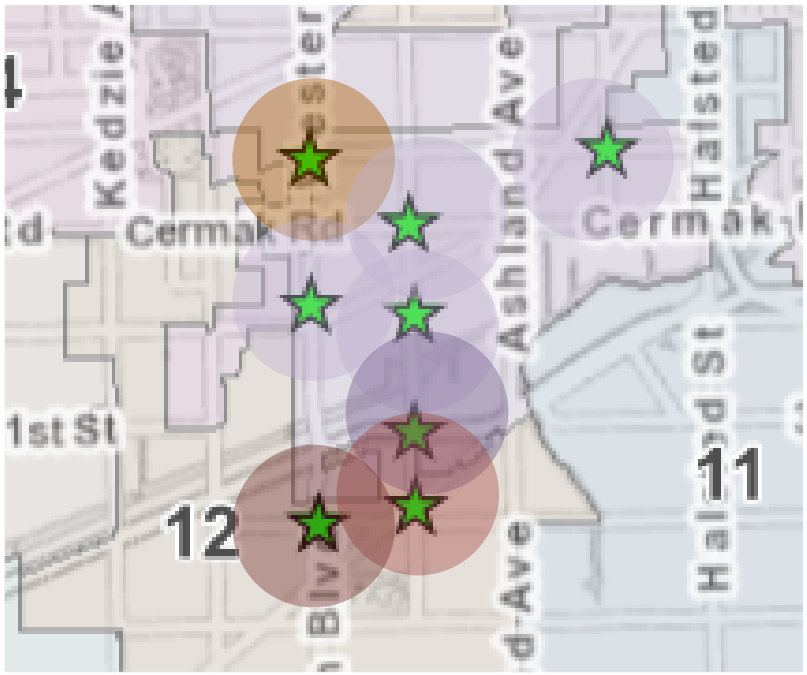

Node
Sensor
Encoder
(Plugin1)
Beehive Server
(Cloud)
Decoder
(Plugin2)
Raw
Data
Decoded
Data
Users
(plenar.io)


BMP180
LPS25H
Pressure data from node 0000001E061089FA


CO ADC Temp
CO LMP Temp
IAQ/IRR Temp
O3/NO2 Temp
SO2/H2S Temp
ADC Temperature data from node 0000001E061089FA

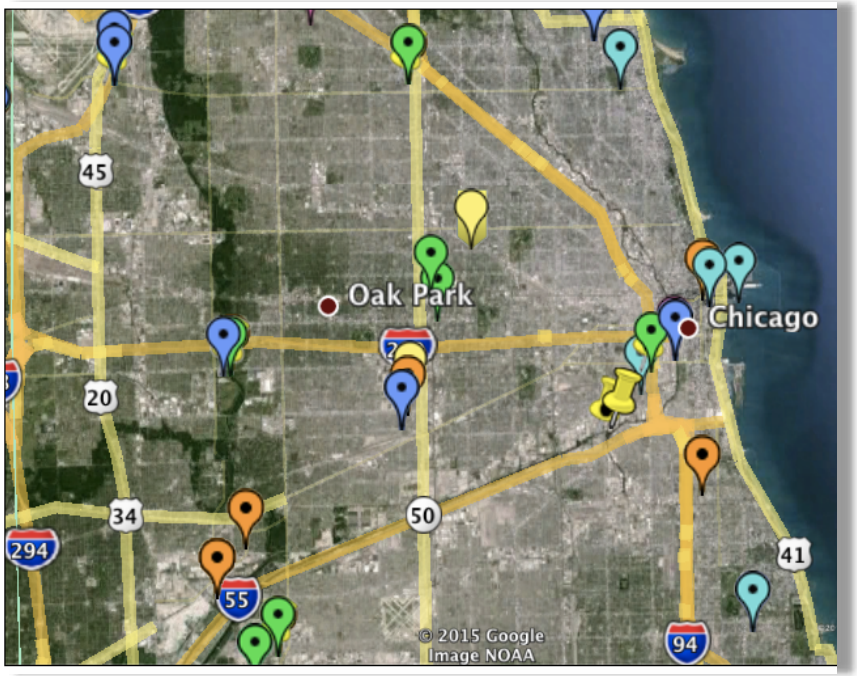

Carbon Monoxide (CO)
Hydrogen Sulphide (H2S)
Nitrogen Di-oxide (NO2)
Ozone (O3)
Sulfur Dioxide (SO2)
Total Oxidizing gases
Total reducing gases
Concentration data from node 0000001E061089FA

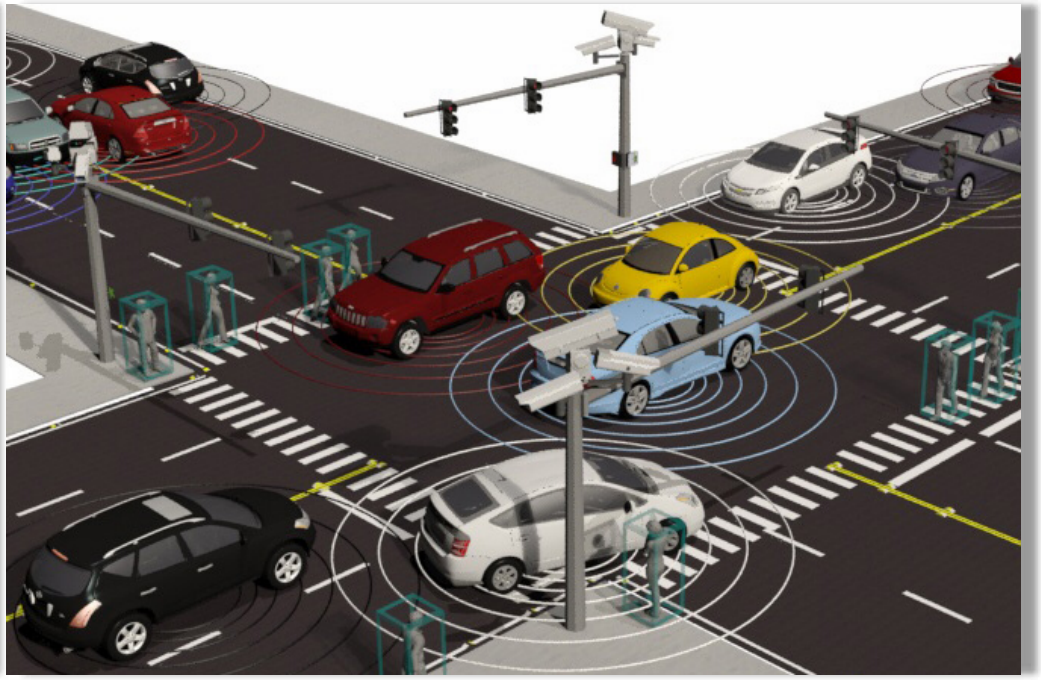

Damen & Ashland
PM 2.5 Alert
Power Outage


Cermak Rd
Cermak
Ashland Ave
Halsted St
1st St
12
11


CDOT


N State St


45
Oak Park
Chicago
20
34
50
41
294
55
94
© 2015 Google
Image NOAA



$2.00

THE NATION'S NEWS

MONDAY

USA TODAY™

08.29.16

JAIR CHAVARRIAGA

FALL MOVIE PREVIEW

'Birth of a Nation' leads the pack of **10 must-see films IN LIFE**

EXCLUSIVE

Chicago puts up sensors to track city's vitals

Fitbit-like tech to provide block-by-block environmental data and more

Aamer Madhani
@AamerISmad
USA TODAY

CHICAGO The Windy City has begun installing what sounds a whole lot like a Fitbit that can measure the vitals of a bustling metropolis.

Chicago, which partnered on gonne National Laboratory and several corporations, last week installed the first two of 500 modular sensor boxes. The devices will eventually allow the city and public to instantly get block-by-block data on air quality, noise levels, as well as vehicular and pedestrian traffic.

The project — dubbed the Array of Things and described by of-its-kind effort in the nation. Plans are in the works to replicate the project in the coming years in more than a dozen other cities, including Atlanta, Chattanooga, and Seattle. The Chicago project was funded with the help of a $3.1 million National Science Foundation grant.

"Five years out, if we're successful, this data and the applica-

Nick Stodony installs a modular sensor box that is part of Chicago's Array of Things project.

Waggle: A Platform for Research

- **Open Source / Open Platform**
  - Reusable, extensible software communities
- **Machine Learning: Computer Vision**
  - Data must be reduced in-situ
- **Novel Sensors**: Nano / MEMS / µfluidics
  - Explosion of nano/MEMS & imaging tech
- **Low-Power CPUs**: GPU / Smartphones
  - Powerful, low-power, smartphone CPUs

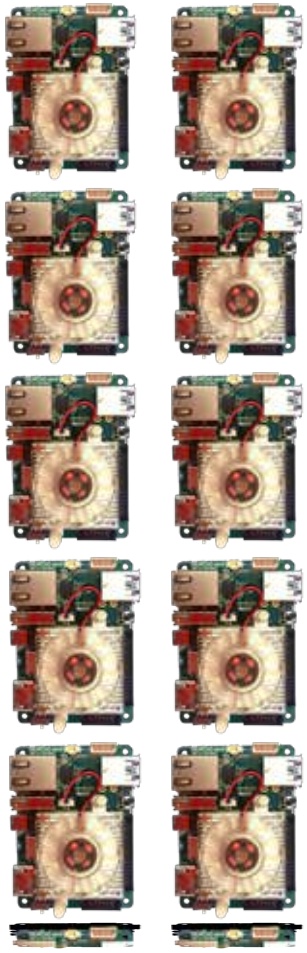
Opportunity: Big Data + Predictive Models

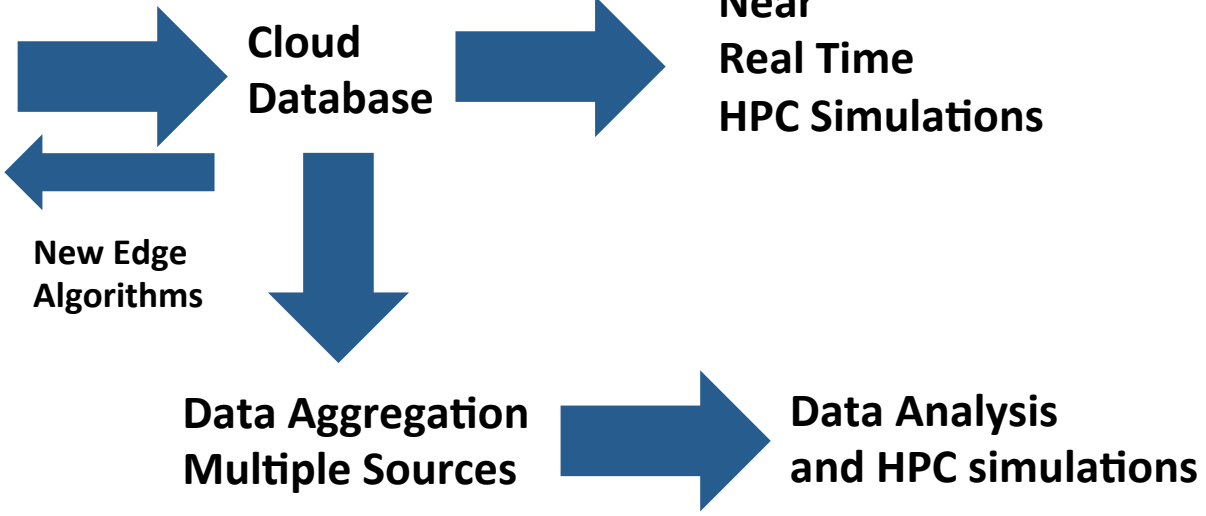
Smart Sensors + Supercomputers/Cloud Computing = predictions and analysis

Why HPC Geeks Should Care

- New sensors are ***programmable* parallel computers**
  - Multicore + GPUs & OpenCL or OpenMP
  - New algorithms for in-situ data analysis, feature detection, compression, deep learning
  - Need new progmod for "stackable" in-situ analysis (for sensors and HPC)
  - Need advanced OS/R resilience, cybersecurity, networking, over-the-air programming
- 1000s of nodes make a **distributed computing "instrument"**
  - New streaming programming model needed
  - New techniques for machine learning for scientific data required
    - Both for within a "node" and collectively across time series
- How will **HPC streaming analytics and simulation** be connected to live data?
  - Can we trigger HPC simulations after first approximations? (weather, energy, transportation)
  - Unstructured database with provenance and metadata for QA/collaboration
- Use novel HPC hardware to solve power issue?
  - Can we use neuromorphic or FPGAs to reduce power for in-situ analysis & compression?
- We are trading precision & cost for greater spatial resolution: What is possible?

Parallel Computation
at the Edge





RISC version of Convergence Story:

Start by enabling… remove roadblocks
(then everyone's wish list follows)

- Applications (Science Drivers): Software needs & workflow patterns
- **Operations**
  - Support real-time and streaming from fast networks
  - Support node sharing, long-lived services, storage requests for years…
- **Architecture**
  - Mothball current parallel file systems, replace with persistent storage services (databases, KV, etc.)
  - Accelerate move of storage into compute infrastructure
- **Software**
  - Linux software development environment.
  - Native support for low-level infrastructure: Docker, VMs, Mesos, etc.
  - New focus on QoS; Software Defined Storage, on-demand services, etc.

Pathways to Convergence: outline of the report (1)

- Introduction
- A Common Context for Planning Cyberinfrastructure Convergence
  - A More Analytical Model of Scientific Inquiry
  - Confronting the General Problem of Data Logistics
    - Alternative 1: Streaming
    - Alternative 2: Content Delivery Networks (Processing In-transit)
    - Alternative 3: Computing at the Edge (at the Source)
    - Other material relevant to this section:
  - Software Infrastructure
    - Stakeholders in Sof
  - What is to be gained f

Pathways to Convergence: outline of the report (2)

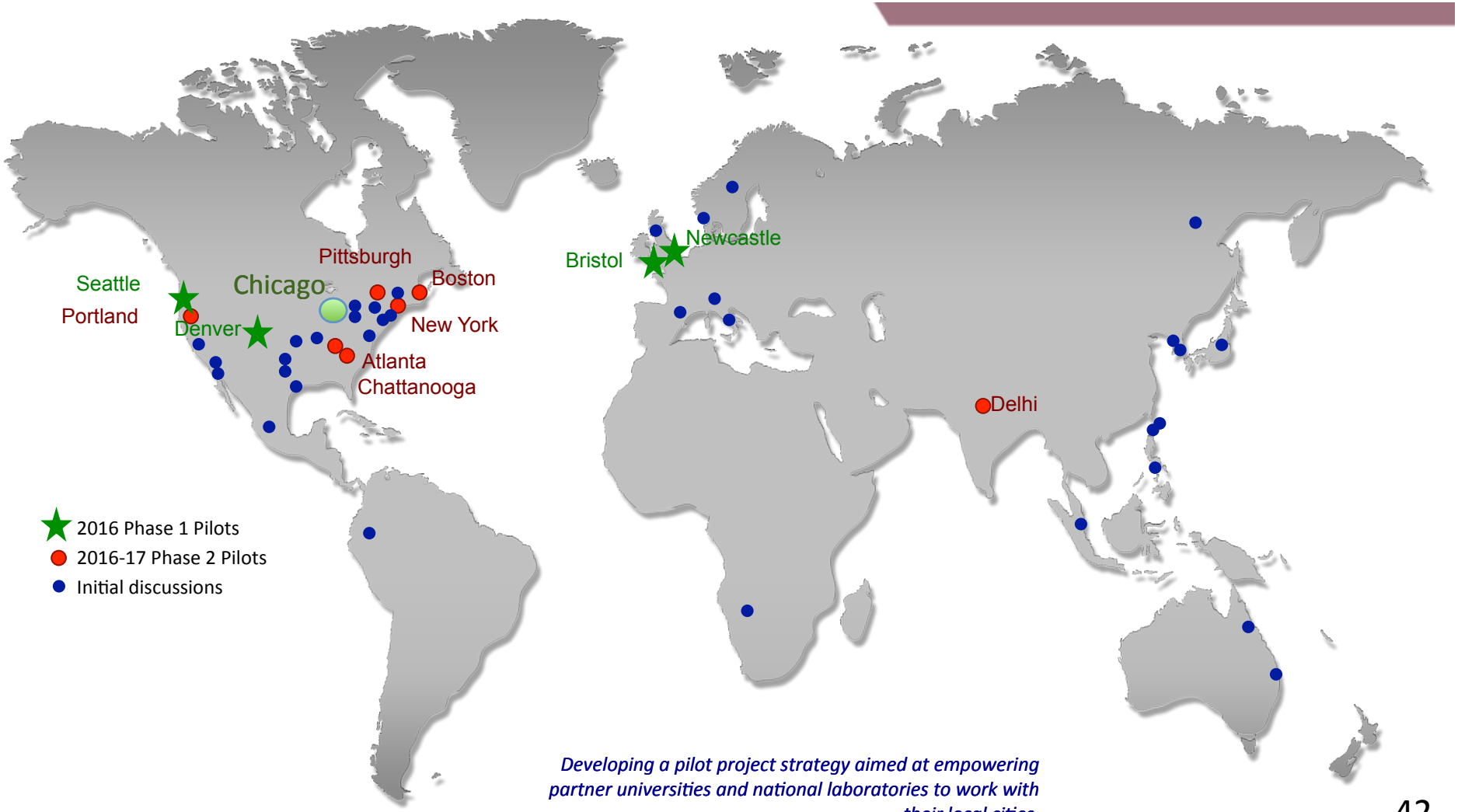
- Pathways to Convergence: What are the Alternatives?
  - The classical paradigm of a common infrastructure: The hourglass architecture.
  - BDEC target goal onto the BD stack; use HPC technology underneath
  - Continue in the Bifurcated world, but do containerization to achieve some commonality
- An Architecture for the New Stack
  - cture Section
  - ing of convergence: Building a new
  - e new "narrow waist"?
  - equate HPC and security capabilities

Pathways to Convergence: outline of the report (3)

- Trends in the Hardware Platform Substrate
- Applications in the Scientific Research Process ("Applications")
  - Taxonomy of Application/workflow patterns and templates
  - Science at the boundary of observation and simulation
  - Numerical laboratories
  - Multimodal sensing of the same reference object
- Layers and Components in the Scientific Software Stack
  - Math Libraries, Software Ecosystems for app development
- Operations and Systems Management
  - Radically improved resource management
  - QoS and Performance tools
- Conclusions and Recommendations

BOF SC2016

Courtesy: Mark Asch



Developing a pilot project strategy aimed at empowering partner universities and national laboratories to work with their local cities.

Questions?