NSF Sustainability Initiative and Big Data

Dr. Krishna Kant
Directorate of Computer and Information Sciences
National Science Foundation

Sustainability Challenge

- Increasing natural & anthropogenic pressures
  - Accelerating climate change
  - Unsustainable population concentrations
  - Diminishing resources (fresh-water, forest cover, …)
  - Increasing frequency of extreme events

Science of Sustainability

Sustainability:

The interactions between natural and social systems, and how those interactions affect the challenge of… meeting the needs of the present and future generations while substantially reducing poverty and conserving the planet's life support systems

Scientific investigation of the nexus of societal needs and behavior, environmental impact, and economic demands

Numerous Areas of Challenges

Renewable Energy Systems

Science, Engineering, and Education for Sustainability (SEES)

Vision: To advance science, engineering, and education to inform the societal actions needed for environmental and economic sustainability and sustainable human well-being

Support interdisciplinary research and education to facilitate move towards global sustainability

Build linkages among existing projects and partners and add new participants in the sustainability research enterprise

Develop a workforce trained in the interdisciplinary scholarship needed to understand and address the complex issues of sustainability

GOALS

Prior NSF SEES Activities

NSF sponsors 5 climate focused solicitations in FY10

- Earth Systems Modeling (EaSM)
  - $50-$35M total awards including DoE & USDA
- Ocean Acidification (OA)
- Dimensions of Biodiversity (DB)
- Water, Sustainability, and Climate (WSC)
- Climate Change Education (CCEP)

NSF issues Dear Colleague Letter

- Issued on January 2011 (NSF 11-022)
- Informs community on the planned future of SEES activities

NSF continues to support SEES efforts through FY11

- Research coordination networks (RCN)
- Couple natural human systems (CNH) solicitation
- Planning workshops, fostering international collaboration, and more

SEES Master page: www.nsf.gov/SEES

Current SEES Solicitations

CISE Sponsored

ArcSEES
Arctic Sustainability

SEP
Sustainable Energy Pathways

SRN SEES
Sustainability Research Networks

RCN SEES
Research Coordination Networks

PIRE SEES
Partnerships for Int. Rsrch & Edu

SEES Fellows
Preparing New Researchers

Others – Current & past

WSC
Water Sustainability and Climate

EASM
Modeling Earth Systems

CNH
Coupled Natural Human Systems

DBD
Dimensions of Biodiversity

OA
Ocean Acidification

CCE
Climate Change Education

New

New

New

Role of Data in Sustainability

- Understanding built & natural systems
- Forecasting
  - To identify challenges and opportunities for change
- Mitigation & Adaptation
  - Reduce environmental impact
  - Adapt to variable or diminishing resources
  - Reduce Human impact (Behavior change)
- Ensure Quality & Robustness
  - Availability, reliability, intrusion, degradation,

Data Challenges

- Monitoring & data collection
- Data Quality
- Configuration Management
- Human Aspects
- Others
  - Coupling between infrastructures
  - Analytics & Forecasting
  - Security & Privacy

Data Collection & Sustainability

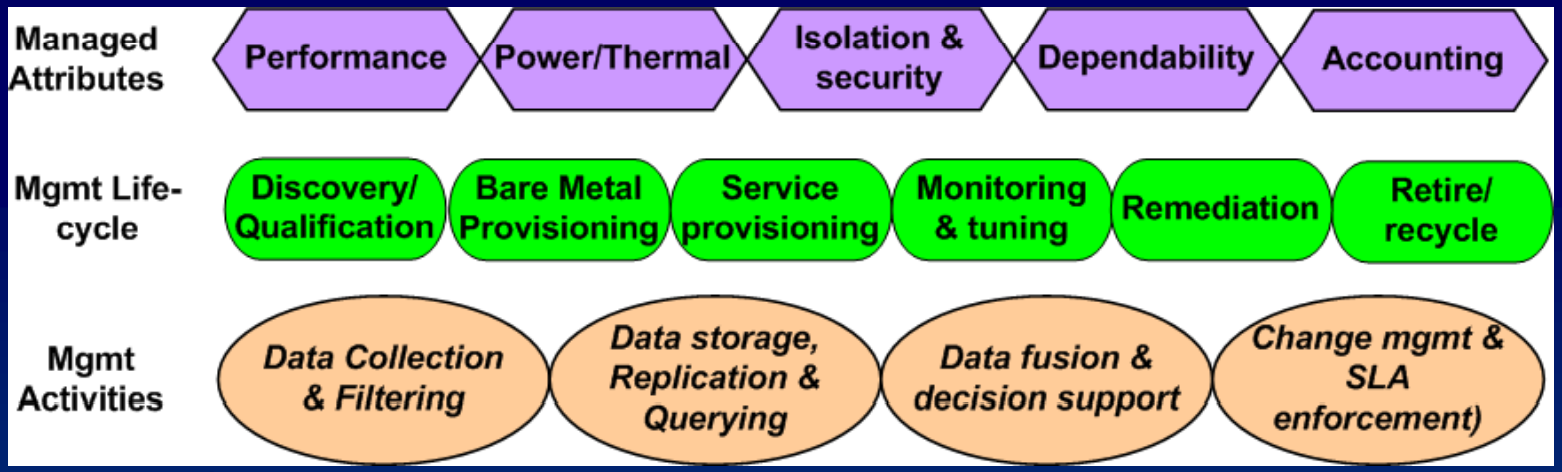
- A wide range of environments to monitor
  - Harsh environments
    - Arctic, wild-fire, underwater, disaster scenarios
    - Impact of extreme events & infrastructure coupling
  - Privacy, intrusiveness, legal limitations
- Challenges
  - Sensor design and deployment
  - Data evacuation challenges
    - Storage/BW limitations, Intermittent or periodic link
  - Poor data quality (missing or low precision, …)

Data Quality

- Data with anomalies
  - Inaccurate measurements (permanent), Sensor malfunction
- Partial data
  - Missing data points (dead or nonexistent sensors)
  - Unknown metadata (location or time of data collection, id of the entity that the data belongs)
- Uneven granularity
  - Spatial (e.g., some data at finer granularity than others)
  - Temporal: uneven frequency, or data from multiple sensors at different times

Configuration Management

- Traditional systems mgmt
  - Fault, Configuration, Availability, Performance, & Security (FCAPS model)
  - Need to add energy management (FCAPSE)
  - Cradle to grave automated mgmt
- Applies equally well to smart infrastructures



Why Config Mgmt is Hard?

- Limited sensing resources
  - Limited battery, network BW, & communication range
- Difficult maintenance
  - Inaccurate or missing data due to failures, degradation, mobility.
  - Easily captured, but need lightweight security
- Emerging Issues
  - Multiple sensor types deployed at large scale
  - Virtualization turns hard assets into “soft” assets

Human Aspects

- Privacy & intrusiveness
  - Limit on granularity & type of monitored data
  - Limited points of collection
- Trustworthiness
  - Trust related to use of supplied data
  - Trustworthiness of humans as data suppliers
- How to effect behavior change?
  - Pricing: well studied, but mostly human as rational actor
  - Feedback: How, when, what are crucial
  - Long term sustainability: Need to remember Jevon's paradox

Example: Arctic Sustainability

- Extremely harsh environment
  - High stress, inaccessible, failures due to cold, wind, etc., burst data transmission, …
  - Slumping permafrost, storm avoidance, constant relocation, …
- What does it take to deploy smart infra. there?
  - Need to deal with damage/relocation to sensors & actual infrastructure.
  - Sharing & adaptation of energy

Concluding Remarks

- Numerous data challenges in sustainability
- Mitigation & Adaptation
  - Key aspects of sustainable systems
  - Adaptation in data center context being studied
  - But we really need them for smart infrastructures in general
  - Coupling between infrastructures & impact of extreme events need to be accounted
- NSF's actions on big data: stay tuned

Thank you!