Does Scale Really Matter?:

Ultra-Large-Scale Systems
Seven Years after the Study

Linda Northrop
Chief Scientist
Software Solutions Division
Software Engineering Institute
May 24, 2013
Software Engineering Institute (SEI)

- Department of Defense R&D Laboratory
- Created in 1984
- Part of Carnegie Mellon University
- Headquartered in Pittsburgh, Pennsylvania; Offices in Washington, Los Angeles, and Frankfurt
- **Mission:** To advance the technologies and practices needed to acquire, develop, operate, and sustain software systems that are innovative, affordable, trustworthy, and enduring.
My Comfort Zone
Ultra-Large-Scale (ULS) Systems
My Talk

ULS System Study Reprise

Current Climate

Experiences with Systems at Scale

ULS Systems-Related Research

Reflection

Interaction
Beginning of the

ULS System Journey
Seven Years Ago

- NASA’s Mars Reconnaissance Orbiter enters Mars orbit
- Windows Vista released
- Saddam Hussein sentenced to death and executed
- Beyoncé Knowles releases second consecutive No.1 album and fourth No.1 single in the US
- BlackBerry users numbered 4,900,000 in March, 2006
- Pittsburgh Steelers win Super Bowl XL
Societal Problems

Climate change and the environment
Powering our civilization
Disease, epidemics, and health care
Livable megacities
Safety and security
Transportation
Society’s Dependence on Software
Trend Toward Increasing Scale-1

- enormous web service and computing infrastructure
- supply chain systems
- software-based engineering systems
Trend Toward Increasing Scale - 2

Healthcare Infrastructure

Homeland Security

Military Systems

Networked Automobiles

Saving the Environment
Increasing Scale In Military Systems

Increasingly Complex Systems

- ultra-large, network-centric, real-time, cyber-physical-social systems
  - thousands of platforms, sensors, decision nodes, weapons, and warfighters
  - connected through heterogeneous wired and wireless networks

- Transient and enduring resource constraints and failures
- Continuous adaptation
- Sustainable - legally, technically, politically
“...How can future systems, which are likely to be a billion lines of code, be built reliably if we can’t even get today’s systems right?”

Intended outcomes:
- ULS System Research Agenda
- Program proposal
- Collaborative research network

About the Effort
- Funded by the Army (ASA ALT)
- Created and led by the SEI
- Staffing: 9 member SEI team
  - 13 member expert panel
- Duration: one year (04/05 -- 05/06)
The Journey
Inspiration: Open Source and Cooperative Communities
More Inspiration

Game Theory

Networks

Economics

Statistical Mechanics
ULS Systems Research Study Report

http://www.sei.cmu.edu/uls/
ULS Systems Research Agenda

Describes

- the characteristics of ULS systems
- the associated challenges
- promising research areas and topics

Is based on a new perspective needed to address the problems associated with ultra-large-scale systems.
What Is an Ultra-Large-Scale (ULS) System?

A ULS System has unprecedented scale in some of these dimensions:

- lines of code
- amount of data stored, accessed, manipulated, and refined
- number of connections and interdependencies
- number of hardware elements
- number of computational elements
- number of system purposes and user perception of these purposes
- number of routine processes, interactions, and “emergent behaviors”
- number of (overlapping) policy domains and enforceable mechanisms
- number of people involved in some way
  
  ULS systems are interdependent webs of software-reliant systems, people, policies, cultures, and economics.
Consequences of Scale

Characteristics of ULS systems arise because of their scale.

- Decentralization
- Inherently conflicting, unknowable, and diverse requirements
- Continuous evolution and deployment
- Heterogeneous, inconsistent, and changing elements
- Erosion of the people/system boundary
- Normal failures
- New paradigms for acquisition and policy
Approaches to Software Development

The Engineering Perspective

The Agile Perspective
A New Perspective is Required

“The older is not always a reliable model for the newer, the smaller for the larger, or the simpler for the more complex...Making something greater than any existing thing necessarily involves going beyond experience.”

Henry Petroski

*Pushing the Limits: New Adventures in Engineering*
Analogies are Useful
“Cities are places of massive information flows, networks, and conduits, and myriad transitory information exchanges.”
Howard Rheinegold: *Smart Mobs*
Think Ecosystem

Diverse users with complex networked dependencies and intrinsic adaptive behavior

Has:

- Robustness mechanisms: achieving stability in the presence of disruption
- Measures of health: diversity, population trends, other key indicators
Think Socio-Technical Ecosystems

Socio-technical ecosystems include people, organizations, and technologies at all levels with significant and often competing interdependencies.

- dynamic communities
- interaction between and among all entities – roles, responsibilities, and information flows
- competition for resources
- rules, incentives, and adaptation
Challenges

ULS systems will present challenges in three broad areas:

- Design and evolution
- Orchestration and control
- Monitoring and assessment

“There are challenges associated with ULS systems that today’s perspectives are very unlikely to be able to address.”
Research Portfolio

6.1 Human Interaction
6.2 Computational Emergence
6.3 Design
6.4 Computational Engineering
6.5 Adaptive System Infrastructure
6.6 Adaptable and Predictable System Quality
6.7 Policy, Acquisition, and Management
What We Learned

There is an unstoppable trend toward increasing scale in many systems important to our society.

Scale changes everything.

These changes undermine the assumptions we routinely make in traditional software engineering approaches.

Manifestations of scale and its attendant complexity arise in many disciplines, and can be understood as a phenomenon in its own right.

New, interdisciplinary perspective and new research in building ultra-large-scale systems is long overdue.
Our Assertion

“Fundamental gaps in our current understanding of software and its development at the scale of ULS systems present profound impediments to the achievement of mission objectives. These gaps are strategic, not tactical. They are unlikely to be addressed by incremental research in established categories. We require a broad new conception of both the nature of such systems and new ideas for how to develop them.”
Early Post-Study Observations

- We never suggested that all systems of the future will be ULS systems. Clearly, they won’t be.
- What you call it (system of systems, ULS system, complex net-centric system) is really unimportant.
- It is important that ULS system characteristics are recognized.
- Systems engineering does not have all the answers.
- Not having a research area on network security was a lightening rod.
- The research identified in the ULS system study has a positive impact on systems that are not ULS.
Seven Years Later
Since Then

Demonstrators in Cairo's Tahrir Square on February 8, 2011

Tsunami flooding on the Sendai Airport runway

Taylor Swift Dominates Billboard Music Awards With 8 Wins (May 2013)
Societal Problems

Climate change and the environment
Powering our civilization
Disease, epidemics, and health care
Livable megacities
Safety and security
Transportation
Software is Ubiquitous and Often Transparent
Software-Reliant Systems: What HAS Changed?

Increased connectivity

Challenges

- scale and complexity
- decentralization and distribution
- “big data”
- increased operational tempo
- mismatched ecosystem tempos
- vulnerability
- collective action
- disruptive and emerging technologies
More Fuel for Scale
The “Crowd”

**DARPA BAA 11-64: Social Media in Strategic Communication (SMISC)**

Research to investigate innovative approaches that enable revolutionary advances in science, devices, or systems for strategies to:

1. Detect, classify, measure and track the (a) formation, development and spread of ideas and concepts (memes), and (b) purposeful or deceptive messaging and misinformation.
2. Recognize persuasion campaign structures and influence operations across social media sites and communities.
3. Identify participants and intent, and measure effects of persuasion campaigns.
4. Counter messaging of detected adversary influence operations.
Some ULS Systems Buzz

~21,266 downloads and hardcopy
75+ citations in refereed publications
Presentations
Workshops
Blog, journal, and twitter references
Initiatives and degree programs

Health IT as an Ultra Large-Scale System
Dr. Doug Fridsma
Chief Science Office and Director
Office of Science and Technology
Office of National Coordinator
US Health and Human Services

Health IT Buzz
February 21, 2013
More Buzz

**Ultralarge Systems:** Redefining Software Engineering?
Greg Goth
March/April 2008 IEEE Software

although the ULSS report focused on challenges faced by the United States Department of Defense in engineering software intensive systems, “its description of how the fundamental principles of software design will change in a global economy … is finding wide appeal.”

**Managing Scale and Agility:** Transformational Architecture for the Smart Grid
Wayne Longcore

“We are creating the first true instantiation of a high-functioning Ultra-Large-Scale System—the Smart Grid.”

**Notes on ultra-large-scale systems**
http://blog.johnrooksby.org/post/132967033/notes-on-ultra-large-scale-systems
John Rooksby
University of Glasgow

“True to national stereotypes the Americans were asking how can we build the biggest systems in the world? The British were asking how can we stop screwing up when we try to build the biggest systems in the world?”

Software Engineering Institute | Carnegie Mellon

Does Scale Really Matter?: ULS Systems Seven Years Later
Linda Northrop: May 24, 2013
© 2013 Carnegie Mellon University
Upon Reflection

STE (Social-Technical Systems)

CPS (Cyber-Physical Systems)

ULS Systems
ULS System Perspective
Build Time

Run Time
Including Tactical Settings

ULS System

Time
Selected Experiences with Systems at Scale: Nibbling at the Edges
The eXtreme Science and Engineering Discovery Environment (XSEDE) enhances the productivity of scientists and engineers.

Department of Energy: Smart Grid

Diagrams courtesy of Wayne Longcore
Consumers Energy

Healthcare Analytics
Smart Grid – A ULS System

Diagrams courtesy of
Wayne Longcore
Consumers Energy
Specific Problem and Technical Approach

Problem
Create a capability to discover if an intruder is executing foreign code in the systems running US critical infrastructure (e.g., Stuxnet).

Approach
Exploit known performance characteristics of critical devices (timing profiles) and monitor run-time behavior for deviations.

Intelligent Electronic Devices (devices deployed to control field equipment) exhibit several desirable characteristics
• are real-time systems
• are deployed in known, stable configurations
• react to a reasonably small number of kinds of stimulus

A timing violation occurs when
• job execution is too short or too long
• job release period is too short or too long
Broader ULS System Impact

The expansion of communication among diverse devices being seen in the Smart Grid is also happening in other ULS systems and raises the same concerns for a capability to detect this class of intrusions.

Other real-time systems with knowable timing profiles where the technique could be used to enhance intrusion detection include

- sensors
- fire control systems
- vehicle and engine controllers
- avionics systems
- ..
The eXtreme Science and Engineering Discovery Environment (XSEDE) enhances the productivity of scientists and engineers.

XSEDE is the framework for a national cyber-infrastructure ecosystem, serving as a platform for multiscale cyber-infrastructure integration for scientific collaboration.
XSEDE’s innovative, open standards-based architecture facilitates an unparalleled level of integration.

Enabling this architecture are XSEDE’s professional systems engineering approach and technology insertion efforts, which ensure robustness and security while continuously incorporating new technologies.
Technical problem: XSEDE needs well-defined software development and software management practices across the XSEDE partner network before it can embrace practices more appropriate for a socio-technical ecosystem.

Approach:

• Identify and coach XSEDE community in adopting a variant of architecture-centric practices suitable for their collaborative ecosystem.

• Apply automated text and social network analyses to data gathered from XSEDE with the goal of providing automated infrastructure support for ensuring that the right people get the right information at the right time.

Impact: Enabling the evolution of the nation’s scientific computing grid via architecture-centered practices and serving as an exemplar for socio-technical ecosystems.
Healthcare Analytics

**Problem:** Define healthcare analytics from technical and organizational perspectives needed to achieve intelligent healthcare.

**Healthcare Infrastructure**

**Approach:** Define an analytics framework.
Organizational Dynamics

Context:
- Standards
- Market
- Vendors
- External Regulations
- Technology
- External Research

Ecosystem

Organization

Business and Mission Goals

Operating Organization

Organizational Outcomes

Direct

Feedback

Provides yardstick

Feedback

Feedback

Context:
- Culture
- Talent/People
Adding Analytics

- Standards
- Market Competitors
- Vendors
- External Regulations
- Technology
- External Research

Context:
- Culture
- Talent/People

Ecosystem
Organization

Business and Mission Goals
Operating Organization
Organizational Outcomes

- Measure Feedback
- Direct Predict
- Feedback Predict
- Provides yardstick
- Measure and support/guide improvements

Feedback Produces Measure and support/guide improvements

Predict Measure and support/guide improvements

Feedback Provides Measure Feedback

Does Scale Really Matter?: ULS Systems Seven Years Later
Linda Northrop: May 24, 2013
© 2013 Carnegie Mellon University
Best in Class Analytics Organizations: Learning Organizations

- Influence ecosystem
- Measure and shape
  - Training
  - Technology change management

- Context:
  - Standards
  - Market
  - Competitors
  - Vendors
  - External Regulations
  - Technology

- Research Ecosystem

- Organization
  - Culture
  - Talent/People

- Business and Mission Goals
  - Operating Organization
  - Organizational Outcomes

- Feedback
  - Direct
  - Predict

- Explicit analytics strategies

- Architecture
  - Organizational structure
  - Governance
  - Enterprise architecture
    - Business processes
    - Applications
    - Data
    - Technology infrastructure
    - Interoperability mechanisms

- Work process execution
  - Embedded data collection
  - Internal research
  - Domain-specific
  - Analytics

- Financial
- Operational
- Domain-specific

- Profit
- Customer satisfaction
- Market leader
- Innovations
Best in Class Healthcare Analytics Organizations: Learning Organizations

- Influence ecosystem
  - Patient safety and quality
  - Operational efficiency
  - Financially competitive
  - Regulatory compliant
  - Research leadership

- Measure and shape
  - Physician profiling
  - Training
  - Technology change management

- Explicit analytics strategies
  - Provides yardstick
  - Measure and support/guide improvements

- Work process execution
  - Clinical care
  - Administration
  - Embedded data collection

- Organization Outcomes
  - Internal research
    - Medical science
    - Clinical effectiveness
    - Analytics

- Operating Organization
  - Business and Mission Goals
  - Direct Feedback
    - Measure and Predict
  - Feedback
  - Produce

- Architecture
  - Organizational structure
  - Governance
  - Enterprise architecture
    - Business processes
    - Applications
    - Data
    - Technology infrastructure
    - Interoperability mechanisms

Context:
- Standards
- Market
- Competitors
- Vendors
- External Regulations
- Technology
- External Research
- Ecosystem
- Culture
- Talent/People
- Organization

Does Scale Really Matter?: ULS Systems Seven Years Later
Linda Northrop: May 24, 2013
© 2013 Carnegie Mellon University
Selected

ULS-Systems-Related Research
Sample Published Work

- **Contextual Design**
  - Collaboration and Coordination in Large-Scale Socio-technical Ecosystems
  - Social Network Analysis
  - James Herbsleb et al, Carnegie Mellon University

- **Ecosystem Modeling**
  - John McGregor et al, Clemson University

- **Machine Learning**
  - Socio Linguistics
  - Natural Language Processing

- **Crowdsourcing Requirements**
  - Stakeholder Analysis

- **Data Intelligence**
  - Data Privacy
  - Data Heterogeneity

- **Self* computing**
  - Self-Coordinating Systems
  - Self-Adaptive Systems
  - Dynamic Adaptive Systems

- **Architecting ULS Systems**
  - End-User Architecting
  - Middleware for ULS Systems
  - Domain-Specific Engineering
  - Model-based Approaches to ULS Systems
  - Multi-Product Lines
  - Multi-Sided Markets
  - Data-Intensive Large-Scale Systems
  - Cloud Computing in the ULS Space

**NOTE:** References at the end
Domain Specific Work

- **Climate Modeling**
  - NASA JPL

- **Climate Informatics**
  - Steve Easterbrook, University of Toronto

- **Disaster Management**
  - Martin Griss, Carnegie Mellon Silicon Valley

- **Intelligent Transportation**
  - Intel, University of Taiwan

- **Health Information Systems**
  - Kevin Sullivan, University of Virginia

- **Financial Markets**
  - Dave Cliff, University of Bristol

- **Intelligent River®**
  - Clemson University

- **Software Defect Analysis in Smart Grid Applications**
  - M. Ancaari, Norwegian University of Science and Technology
Selected SEI Research
Targeted at ULS Systems:
More Nibbling
Some SEI Research In ULS Systems

- Socio-Adaptive Systems Using Computational Mechanism Design and Adaptive QoS
- High-Confidence Cyber-Physical Systems
- Augmented and Virtual Actors for Threat Abatement Readiness
- Architecture in ULS Systems Context
- Concurrent Crowdsourcing of Requirements and Architectures for Socio-Technical Infrastructure Improvement

Edge-Enabled Tactical Systems
Architecture in ULS System Context

ULS system characteristics inspire key questions about systems at scale.

- What new quality attributes arise due to scale?
- What types of analyses are required to understand and design systems with these characteristics?
- What new architecture design principles needed?
  - E.g., synergy of concerns instead of separation of concerns?
- What are the associated architectural tactics, patterns, mechanisms?
- What types of analyses and design strategies are needed to design all levels of systems at scale?
  - E.g., population dynamics, connectedness/communication
- And what expertise is required for this design?
Edge-Enabled Tactical Systems (EETS)

Investigates architectures and technologies that adapt new generations of mobile devices and sensors to support humans operating in demanding edge environments.

Mobile technologies can enhance the manner in which people operate in tactical environments:

- Local data caching with reach back when available
- Cyber-foraging to enhance handheld and sensor device capabilities
- Flexible deployment and rapid adaptation for new missions
- Context-aware computing to reduce cognitive load and conserve resources
- Local, edge analytics to provide rapid data analysis
- Increased use of autonomy (drones, robots, sensors)

What architectures and technologies support soldiers and other edge users in customizing systems to unique needs, finding information that matters, and to continue processing in uncertain computing environments?
Reflection:

The Future is Here
So, Where Are We?

The report has been widely distributed via the web and hard copy. Relevant research is being conducted all over the world. The confluence of new technology is making ULS systems today.

- with a profound impact on the way society is structured and how society behaves
- substantial engineering challenges are becoming widely recognized if still poorly understood.
  - reliance on *autonomous* behavior
  - increased interaction and interdependence of socio-technical ecosystems
  - increased tempo of change across the spectrum of human behavior – driven by human demand

There is a wide range of technical and non-technical perspectives and approaches that can be brought to bear.
Climate Change (term used by David West at Code Freeze 2013)

Characteristics of ULS systems arise because of their scale.

• Decentralization
• Inherently conflicting, unknowable, and diverse requirements
• Continuous evolution and deployment
• Heterogeneous, inconsistent, and changing elements
• Erosion of the people/system boundary
• Normal failures
• New paradigms for acquisition and policy

These are real.
Opportunities and Threats
Implications for How We Do Our Work

Systems must be:

- Responsible
- Responsive
- Adaptive

- **Control**
- **Incentivize**

NORMATIVE | PARTICIPATORY

© 2013 Carnegie Mellon University

Does Scale Really Matter?: ULS Systems Seven Years Later
Linda Northrop: May 24, 2013
© 2013 Carnegie Mellon University
Putting Technology to Work: a Few Take Aways

Context is key
• Computation needs to be organic with human incentives and human workflow
• Technology standards do not ensure interoperability
• Multi-disciplinary approach is essential

Big data and machine learning don’t help without an analytics framework, feedback loops, and analytics-driven sensing
Humans, computational, and autonomous entities are peers
Research Progress – My Assessment

- 6.1 Human Interaction
- 6.2 Computational Emergence
- 6.3 Design
- 6.4 Computational Engineering
- 6.5 Adaptive System Infrastructure
- 6.6 Adaptable and Predictable System Quality
- 6.7 Policy, Acquisition, and Management

Progress has been made on all these fronts and others.
And yet...there is a fast growing gap between our research and reality.
Summing It Up

Scale Matters

• ULS systems are in our midst and the changes to our social fabric and institutions are significant.
• In hindsight, we were probably too conservative in our report.
• Recent technologies have exacerbated the pace of scale growth – allowing us to transcend time and space.
• There are great opportunities.
Food for Thought

- Is our research a match?
- Do we have the right incentives and mindset for the needed multi-disciplinary approach?
- Will we, the software engineering research community, make a difference?
Thanks To Many Who Made The Study Possible

Report Author Team:

Peter Feiler, Richard P. Gabriel, John Goodenough, Rick Linger, Tom Longstaff, Rick Kazman, Mark Klein, Douglas Schmidt, Kevin Sullivan, Kurt Wallnau, Bill Pollak (Chief Editor), Daniel Pipitone (Information Designer)

2006 Support System:

Hon Claude Bolton, Paul Nielsen (SEI CEO), Clyde Chittister (SEI COO), Hal Stevens (2006 SEI/Army Liaison), Jim Linnehan (2006 Army/SEI Liaison)
Thanks To The Entire ULS System Study Team
Thanks To My SEI and Campus Colleagues

In particular:

Felix Bachmann, Jeff Boleng, Gene Cahill, James Edmondson, Ian Gorton, Jim Herbsleb, Scott Hissam, Carolyn Kernan, John Klein, Mark Klein, Mike Konrad, Grace Lewis, John McGregor, Gabriel Moreno, Ed Morris, Daniel Pipitone, Bill Pollak, James Root, Mary Shaw, Soumya Simanta, Kurt Wallnau
A Special Thank You and Tribute
Contact Information

Linda Northrop
SEI Fellow
Chief Scientist
Software Solutions Division
Telephone: 412-268-7638
Email: lmn@sei.cmu.edu
Website: http://www.sei.cmu.edu/uls/

U.S. Mail:
Software Engineering Institute
Carnegie Mellon University
Pittsburgh, PA 15213-3890

SEI Fax: 412-268-5758
References for Slides 65-66
References - 1


References - 2


References - 3


References - 4


References - 5


References - 6


References - 7


References - 8


References - 9


References - 10


References - 11


References - 13


References - 14


References - 15


References - 16


This material is based upon work funded and supported by the Department of Defense under Contract No. FA8721-05-C-0003 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center.

Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the United States Department of Defense.

NO WARRANTY. THIS CARNEGIE MELLON UNIVERSITY AND SOFTWARE ENGINEERING INSTITUTE MATERIAL IS FURNISHED ON AN “AS-IS” BASIS. CARNEGIE MELLON UNIVERSITY MAKES NO WARRANTIES OF ANY KIND, EITHER EXPRESSED OR IMPLIED, AS TO ANY MATTER INCLUDING, BUT NOT LIMITED TO, WARRANTY OF FITNESS FOR PURPOSE OR MERCHANTABILITY, EXCLUSIVITY, OR RESULTS OBTAINED FROM USE OF THE MATERIAL. CARNEGIE MELLON UNIVERSITY DOES NOT MAKE ANY WARRANTY OF ANY KIND WITH RESPECT TO FREEDOM FROM PATENT, TRADEMARK, OR COPYRIGHT INFRINGEMENT.

This material has been approved for public release and unlimited distribution except as restricted below.

The Government of the United States has a royalty-free government-purpose license to use, duplicate, or disclose the work, in whole or in part and in any manner, and to have or permit others to do so, for government purposes pursuant to the copyright license under the clause at 252.227-7013 and 252.227-7013 Alternate I.

This material may be reproduced in its entirety, without modification, and freely distributed in written or electronic form without requesting formal permission. Permission is required for any other use. Requests for permission should be directed to the Software Engineering Institute at permission@sei.cmu.edu.

DM-0000065