Pacific Northwest
NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

Top Challenges of Extreme-Scale Visual Analytics

PAK CHUNG WONG

Pacific Northwest National Laboratory

NIST Data Science Symposium 2014, Gaithersburg, MD 20899

Extreme-Scale Visual Analytics

- Extreme-Scale – Petabyte (10^{15}) to Exabyte (10^{18})
- Visual Analytics
  - "Science of analytical reasoning facilitated by interactive visual interfaces"
  - http://en.wikipedia.org/wiki/Visual_analytics
- Extreme-Scale Visual Analytics
  - "Applying visual analytics to extreme-scale data"
  - Three critical elements:
    1. Size
    2. Inclusion of visual and analytical
    3. Active involvement of a human

Top Ten Challenges of Extreme Scale Visual Analytics

1. In situ Analysis
2. Interaction and User Interface
3. Large Data Visualization
4. Database and Storage
5. Algorithm
6. Data Movement, Data Transport, and Network Infrastructure
7. Uncertainty Quantification
8. Parallelism
9. Domain and Development Libraries, Frameworks, and Tools
10. Social, Community, and Government Engagements

Pak Chung Wong, Han-Wei Shen, Christopher R. Johnson, Chaomei Chen, and Robert B. Ross, "Top Ten Challenges in Extreme-Scale Visual Analytics," *IEEE Computer Graphics and Applications*, 32(4):63-67, July 2012.

Top Ten *Interaction* Challenges of Extreme Scale Visual Analytics

1. In situ interactive analysis
2. User-Driven Data Reduction
3. Scalability and Multi-Level Hierarchy
4. Representation of Evidence and Uncertainty
5. Heterogeneous Data Fusion
6. Data Summarization and Triage for Interactive Query
7. Analytics of Temporally Evolving Features
8. The Human Bottleneck
9. Design and Engineering Development
10. The Renaissance of Conventional Wisdom

Pak Chung Wong, Han-Wei Shen, and Chaomei Chen, "Top Ten **Interaction** Challenges in Extreme-Scale Visual Analytics," *Expanding the Frontiers of Visual Analytics and Visualization*, pages 197-207, Springer, UK, 2012.

Emerging Trend and Opportunity

- Usability and User Interface Design
- Scientific Data Management, Analysis, and Visualization at Extreme Scale 2, DOE FOA 14-1043, 2014
  1. "…truly usable requires deep understanding of the culture, expectations, and processes of science in the fields to be supported."
  2. "True usability is not tacked on at the end by providing a web interface to an application."
  3. "…usability must be designed and built in from the beginning, helping to drive functional choices at every level."

Other Hot Topics and Resources

- Different versions of the V-challenges
  - Volume, velocity, and variety
  - Volatility, validity, and viscosity
  - Volume, velocity, variety, and veracity
  - Volume, velocity, variety, variability, and value
- IEEE Symposium on Large Data Analysis and Visualization (LDAV) 2014, Paris, France, Nov 2014, http://www.ldav.org/
- DOE Office of Science Advanced Scientific Computing Research (ASCR), http://science.energy.gov/ascr/

Acknowledgment

- This work has been supported by the U.S. Department of Energy Office of Science Advanced Scientific Computing Research under award number 59172, program manager Lucy Nowell. The Pacific Northwest National Laboratory is managed for the U.S. Department of Energy by Battelle under Contract DE-AC05-76RL01830.

- Contact information:
  - pak.wong@pnnl.gov
  - http://www.linkedin/in/pakchungwong