Enabling Access to High-Resolution LiDAR Topography through Cyberinfrastructure-Based Data Distribution and Processing

C. J. Crosby[1], J R. Arrowsmith[1], E. Jaeger-Frank[2], V. Nandigam[2], H. S. Kim[3], J. Conner[1], A. Memon[2], N. Alex[1], C. Baru[2]

[1]School of Earth and Space Exploration, Arizona State University
[2]San Diego Super Computer Center, University of California, San Diego
[3]Department of Computer Science and Engineering,, University of California, San Diego

Digital data acquisition technologies such as LiDAR (Light Distance And Ranging) topography have resulted in an increase in the volume and complexity of scientific data that must be efficiently managed, distributed and processed in order for it to be of use to the scientific community. Capable of generating digital elevation models (DEMs) more than an order of magnitude more accurate than those currently available, LiDAR data offers the opportunity to study earth surface processes at resolutions not previously possible yet essential for their appropriate representation.

Unfortunately, access to these datasets for the average user is difficult because of the massive volumes of data generated by LiDAR. The distribution and processing of large LiDAR datasets, which frequently exceed billions of data-points, challenge internet-based data distribution systems and readily available desktop software. Figure 1 shows the conceptual workflow required to produce results for scientific analysis using LiDAR.

Our approach to the distribution and processing of LiDAR data capitalizes on cyberinfrastructure developed by the GEON project (http://www.geongrid.org) to harness distributed computing resources (Figures 2 and 3). We utilize a workflow-based solution, the GEON LiDAR Workflow (GLW), which begins with user-defined selection of a subset of point data and ends with download (including dynamically generated metadata) and visualization of DEMs and derived products. Users perform point cloud data selection, interactive DEM generation and analysis, and visualization all from an internet-based portal. Users may experiment with DEM resolution and DEM generation algorithms so as to optimize terrain models for their application. By using cyberinfrastructure resources, this approach allows users to carry out computationally intensive LiDAR data processing without having appropriate resources locally.

We are currently in the process of migrating the system from its current proof of concept implementation to a fully robust, production level, community data portal. In addition to selection and processing, the GLW (through the GEON portal) now includes job management tools for users to view and modify previously submitted jobs, and to monitor the status of existing jobs. As of April 30, 2007, the GLW contained 4 datasets comprising about 10 billion individual spatially indexed points. A total 1184 jobs had been submitted for processing of 12.3 billion points by 90 active users.

References:

Crosby, C.J., Arrowsmith, J R., Frank, E., Nandigam, V., Kim, H.S., Conner, J., Memon, A., Baru, C., Enhanced Access to High-Resolution LiDAR Topography through Cyberinfrastructure-Based Data Distribution and Processing, *Eos Trans. AGU, 87*(52), Fall Meet. Suppl., Abstract IN41C-04, 2006.

Crosby, C.J., Arrowsmith, J R., C.J., Frank, E., Conner, J., Memon, A., Nandigam, V., Kim, H.S., Alex, N., Wurman, G., Baru, C., A Geoinformatics Approach to LiDAR Data Distribution and Processing: *in preparation*

Efrat Jaeger-Frank, Christopher J. Crosby, Ashraf Memon, Viswanath Nandigam, J. Ramon Arrowsmith, Jeffrey Conner, Ilkay Altintas, Chaitan Baru, A Three Tier Architecture for LiDAR Interpolation and Analysis, *Lecture Notes in Computer Science*, Volume 3993, Apr 2006, Pages 920-927, DOI: 10.1007/11758532_123.

Thesis

Christopher Crosby, M. S., A geoinformatics approach to LiDAR data distribution and processing with applications to geomorphology, Arizona State University, August, 2006.

Links:

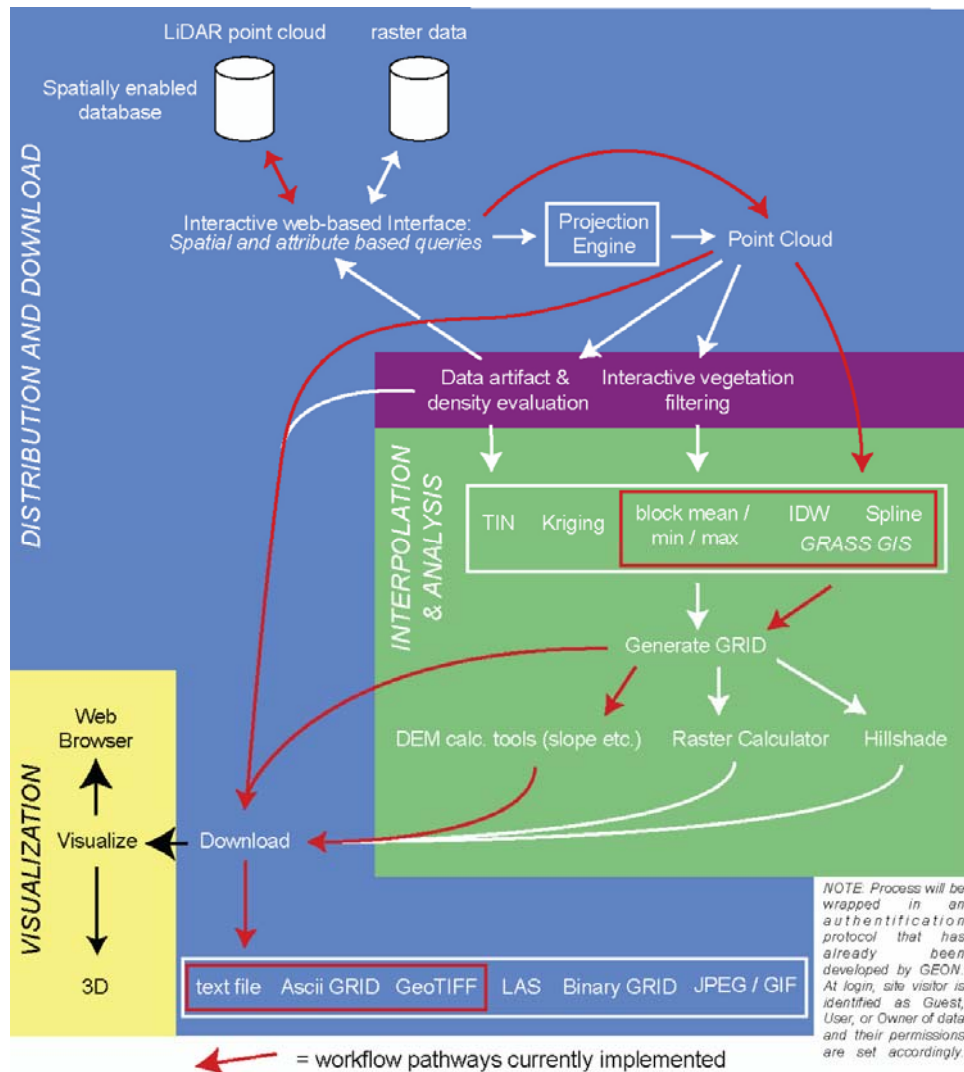http://www.geongrid.org/science/lidar.html; http://lidar.asu.edu/ (ASU LiDAR research).

Figure 1. Conceptual GEON LiDAR workflow. This workflow capitalizes on cyberinfrastructure available via the GEON project to offer an end-to-end LiDAR pont cloud distribution and processing toolset. The workflow begins with point cloud data (and raster imagery) selection and includes DEM generation and analysis as well as product download functionality. Pathways shown in red have been implemented in our proof of concept. Pathways in white or black represent future development directions.
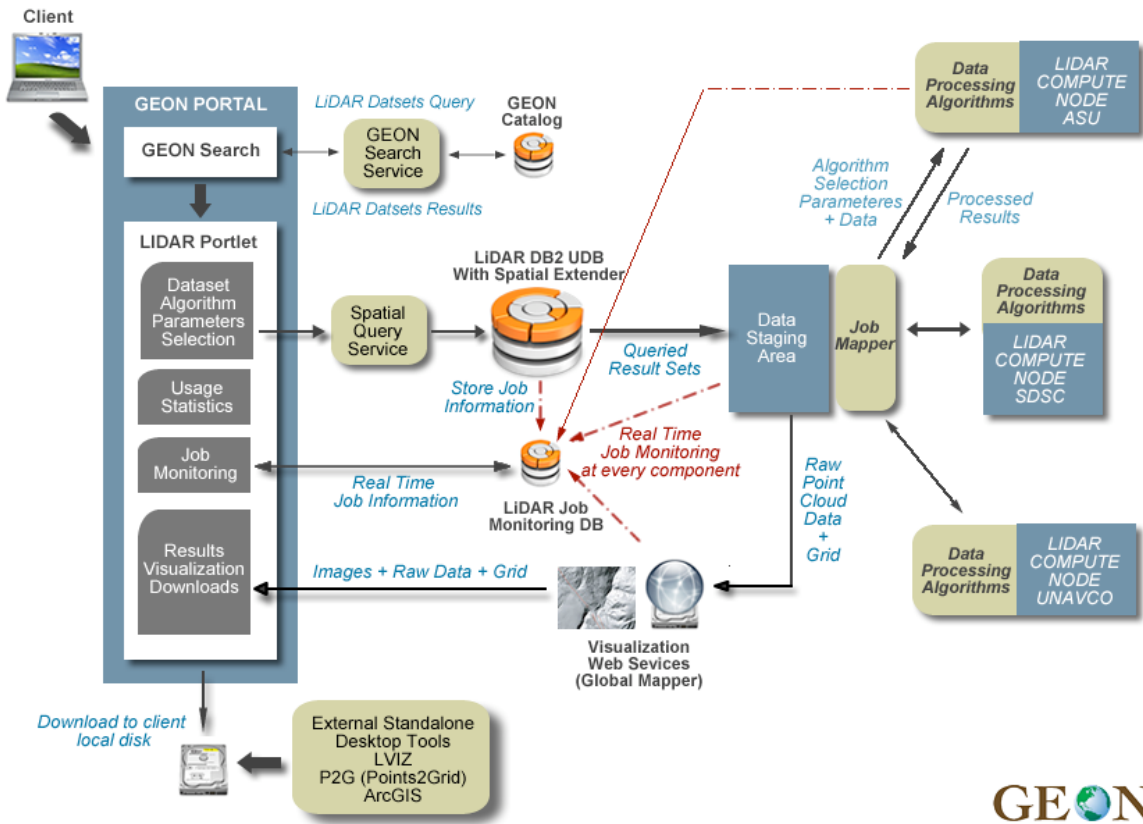
Figure 2. GEON LiDAR Workflow architecture illustrating the distributed cyberinfrastructure utilized to perform data distribution and processing tasks. It also enables real time job monitoring for feedback to users. Note the modular, web service-based architecture which allows the GLW to be scaled to include other database, tools, or computing resources.
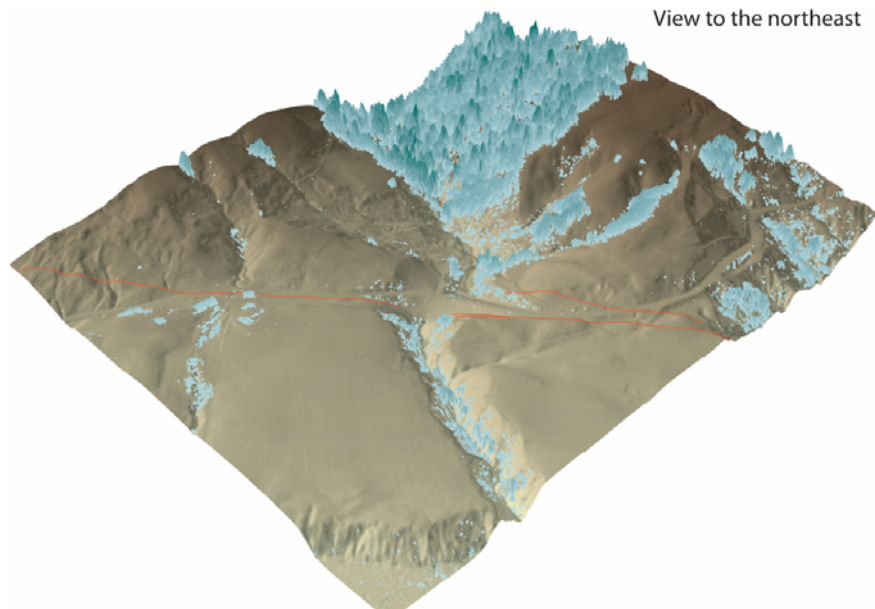

Figure 2. Example view of gridded LiDAR data selected and processed in the GLW using the NASA/USGS dataset for the Northern San Andreas Fault. In this ~2 m digital elevation model of Mill Gulch, the green areas show the vegetative canopy taller than about 50 cm overlain on the topography. Canopy height is colored in shades of green with darker colors indicating taller trees. The red lines show the discontinuous traces of the San Andreas Fault which broke in the 1906 earthquake.