Parallel Map Projection of Vector-based Big Spatial Data

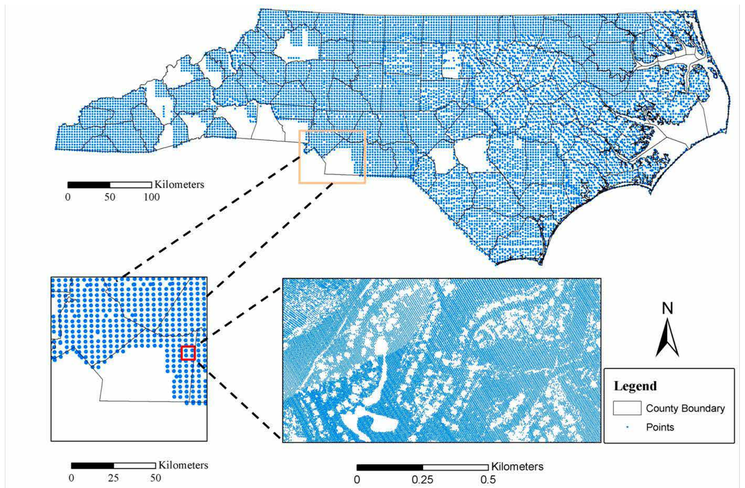
Coupling Cloud Computing with Graphics Processing Units

Wenwu Tang, Wenpeng Feng

Center for Applied GIScience and Department of Geography and Earth Sciences
University of North Carolina at Charlotte, Charlotte NC, 28223 Email: WenwuTang; wfeng5@uncc.edu

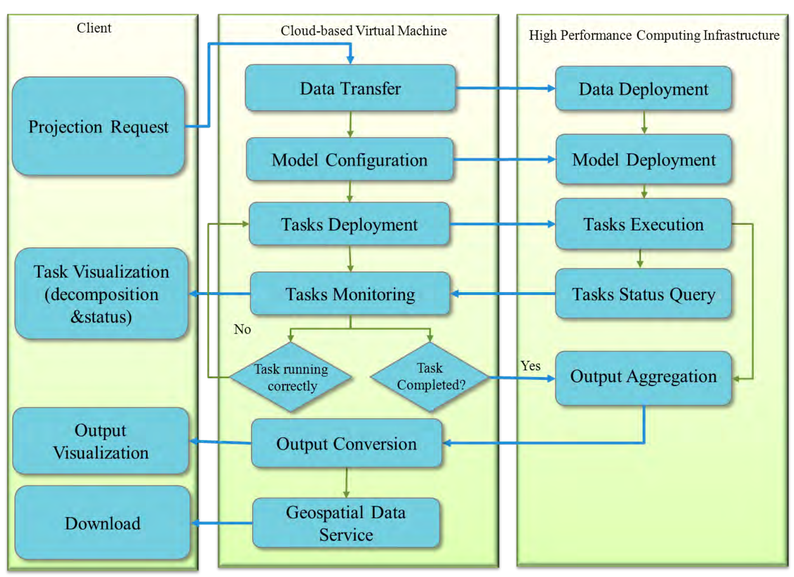
Objective & Study Area

- Objective: Develop a parallel spatial computing framework for the map projection of **vector-based big spatial data**
- Study area: North Carolina, USA
- Data:
  - Bare earth LiDAR (Light Detection And Ranging)
  - Total 230 **Gigabytes**
  - 13,596 sub-datasets (in file format)



Workflow

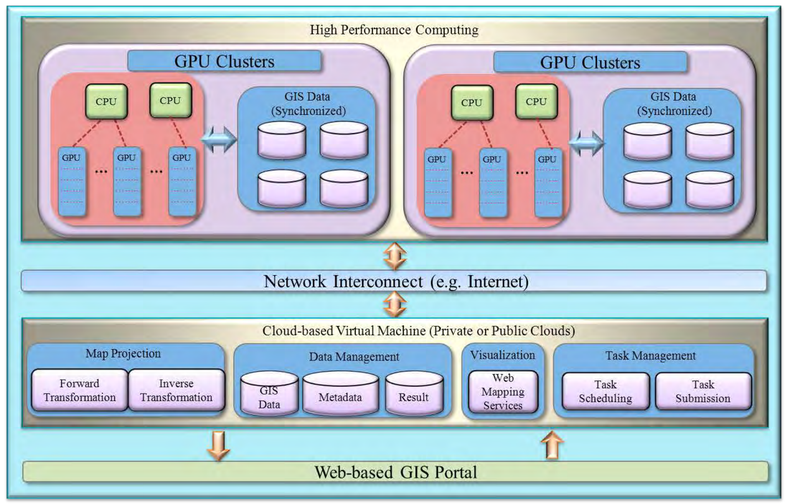
The workflow with respect to interactions between layers and the specific functionality of each layer



Methodology

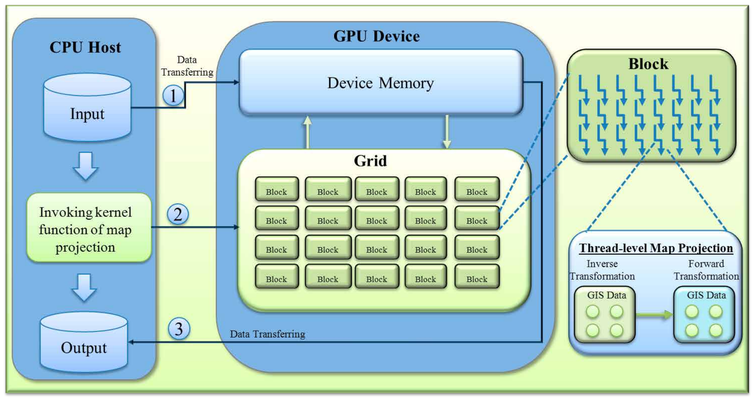
Framework

- Integration of three functionality layers: **High-performance computing**, **cloud-based virtual machine**, **and web-based GIS portal**
- The advantage of the layered architecture: Blending capabilities of **GPU-enabled high-performance computing**, **cloud computing**, and **Internet GIS**

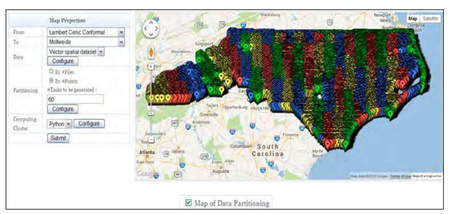


GPU-accelerated high-performance computing

- GPU-enabled parallel **map projection algorithms** were developed based on CUDA platform.
- Heterogeneous parallel computing requires cooperation between CPUs and GPUs.

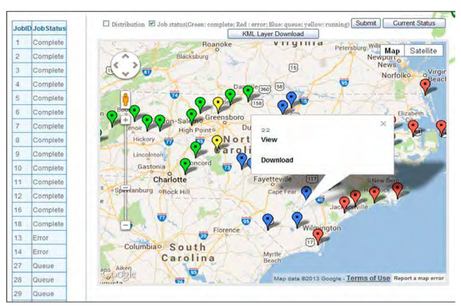


Web GIS portal



Task configuration

- Specify map projection
- Configure partitioning strategies
- Determine number of tasks



Job monitoring

- Submit jobs
- Query status of jobs
- Download results

Results

#GPUs	Computing Time (with transfer time)			Computing Time (without transfer time)		
	Load Balancing		Time difference	Load Balancing		Time difference
	Before	After		Before	After	
20	30.72	27.11	3.61	4.56	3.78	0.79
30	19.64	19.03	0.61	2.97	2.58	0.39
40	15.32	15.28	0.03	2.27	2.02	0.25
50	12.51	12.26	0.25	1.96	1.64	0.33
60	10.91	10.78	0.13	1.70	1.42	0.29

Results of computing time for parallel map projection using multiple GPUs (time unit: seconds; the entire North Carolina LiDAR dataset was used).

Conclusions

- The transforming time of entire LiDAR dataset is decreased from 3 hours to 11 seconds.
- The larger the size of input data (i.e., problem size), the higher speed up that we gain in terms of pure computing time for map projection
- The use of load balancing leads to increase in acceleration factors.
- This integrative framework provides substantial support for best leveraging each component that complements to each other with respect to the map projection of big spatial data.

Acknowledgement

- CAGIS (Center for Applied Geographic Information Science)
- UNCC, URC (University of North Carolina, Charlotte, University Research Computing)