Intelligent Cognitive Assistant for Emergency Landings using GPU Geometric Computing and Online Statistical Reasoning

W Randolph Franklin, ECSE Dept, Rensselaer Polytechnic Institute, mail@wrfranklin.org, https://wrf.ecse.rpi.edu/

Carlos Varela, CS Dept, Rensselaer Polytechnic Institute, cvarela@cs.rpi.edu, http://www.cs.rpi.edu/~cvarela/

This white paper will describe a new capability (powerful GPUs and massive memory), and then an important new use case (an intelligent cognitive assistant for emergency landings using GPU geometric computing and online statistical reasoning).

Capability: Powerful GPUs, multicore Intel CPUs, and massive local memory are new core technical capabilities that will permit massive geospatial software job to be run on inexpensive local machines and even good laptops. Tasks that had required supercomputers and cloud facilities can now, in many cases, be localized. Formerly infeasible tasks are now feasible.

The Nvidia Titan V Volta graphics card is a supercomputer on a card, with 12GB of memory running at a potential 110 teraflops. If the problem does not fit into the memory, the card can be attached to a server with 256GB of DRAM, for a total cost of well under $10,000. Smaller jobs can be run on a good laptop with 32GB of memory and a mobile GPU, costing perhaps $2000. While there are still jobs large enough to require distributed computing, message passing interfaces, and the cloud, they are fewer. This has several implications.

Jobs that were formerly batch-oriented, with turnaround times of hours, can now be processed in seconds, and in regions with poor internet connectivity. (You cannot do affordable cloud computing with a satellite phone modem.) That includes, by area, most of the world.

Since jobs run faster, the user can now model hypotheticals. E.g., if the task is viewshed determination and observer siting, then many possible scenarios can be compared. A mechanical parametric search might even be performed, to optimize over many choices. All this might be done on a laptop.

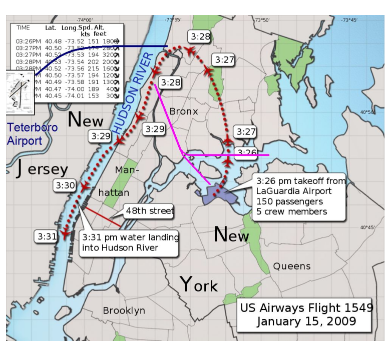
This capability will also allow community involvement in a more serious manner. Too many computer systems have a real but unstated function of persuading users that the preselected answer is correct. This might allow users to do serious modeling and choice comparison on their own.

However, utilizing this hardware imposes constraints on the SW designers. The above mentioned GPU has 5120 CUDA cores. However they are grouped by 32s into warps. All 32 cores in a warp execute the same instruction (or are idle). Ideally, the data for consecutive cores is stored consecutively in a global array. This has several implications for the SW designer.

Simple structures of arrays are much to be preferred. Pointers, trees, and recursion do not run well on this architecture. That obsoletes most of the favorite algorithms in the GIS community. Map-reduce and similar paradigms work well.

Use case: This project will develop new algorithms and software for a intelligent cognitive flight assistant to aid pilots to land their aircraft safely in emergency conditions, for example, after bird strikes over densely populated areas. In order to succeed in this endeavour, the project will investigate parallel computing techniques to preprocess terrain features and create initial trajectories from potential emergency locations to safe landing sites. We will also investigate online statistical methods to model uncertainties due to dynamic data, such as weather (wind, clouds, thunderstorms) and air traffic, and incrementally update pre-computed trajectories to account for emergency dynamics. For this project to succeed, offline computations must produce plans with efficient space footprints, and incremental plan update algorithms must be able to run in limited (aircraft-local) resources in real-time (within five seconds of emergencies) to be practical. This project will investigate the use of GPUs and multicore compute servers both for parallel offline computation and for on-board online real-time computation. Successful research results will turn into fewer lives lost and less property damage. Varela has been doing research on safer flight systems, in particular, on self-healing sensor data streams for more than five years. Varela is also an instrument-rated private pilot with complex and high-performance endorsements, and over 880 hours of flying experience. Franklin has extensive experience with parallel algorithms in geometry and with processing large terrain datasets.

This project will use actual data from the coterminous USA for terrain representation and classification, and from actual flight emergencies such as US Airways Flight 1549 (Hudson river ditching) to evaluate the new algorithms and software. The expected discoveries will reduce fatalities for both commercial aviation and general aviation, which suffered 435 fatalities in FY2014. Fast parallel algorithms to process terrain will also be beneficial to other branches of engineering. The software will be freely available for nonprofit research and education. Results will be incorporated into courses taught by the PIs on parallel computing, distributed computing, computer graphics, and introduction to engineering design.



US Airways 1549 after ditching in the Hudson river. The aircraft actual path is shown in red. Flying to Teterboro (blue) or back to La Guardia (purple) were alternative plans.