GPU performance comparison for accelerated radar data processing

Chris Fallen[1], Beau Bellamy[1], Greg Newby[1], and Brent Watkins[2]

[1]Arctic Region Supercomputing Center

[2]Geophysical Institute

University of Alaska Fairbanks

2011 Symposium on Application Accelerators in High Performance Computing (SAAHPC'11)

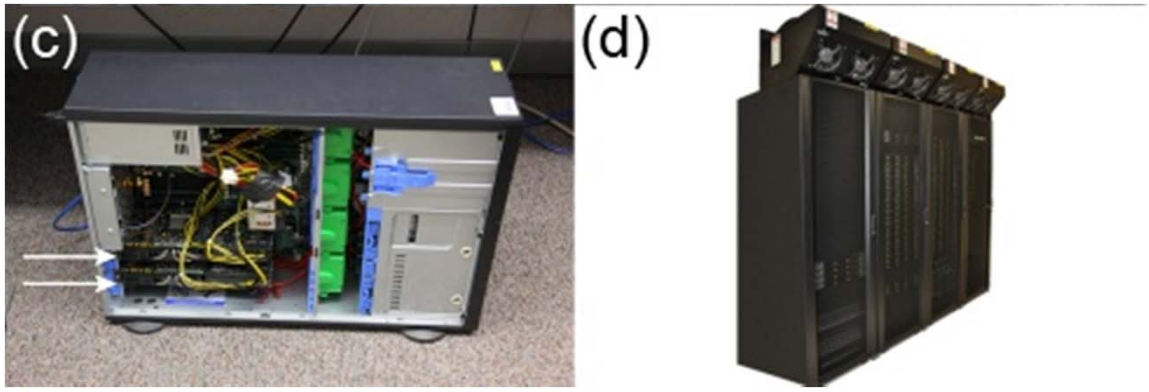
19-20 July 2011, Knoxville, Tennessee

Outline

- Overview
- Motivation
- Experiment description
- Results
  - Preview of preliminary findings

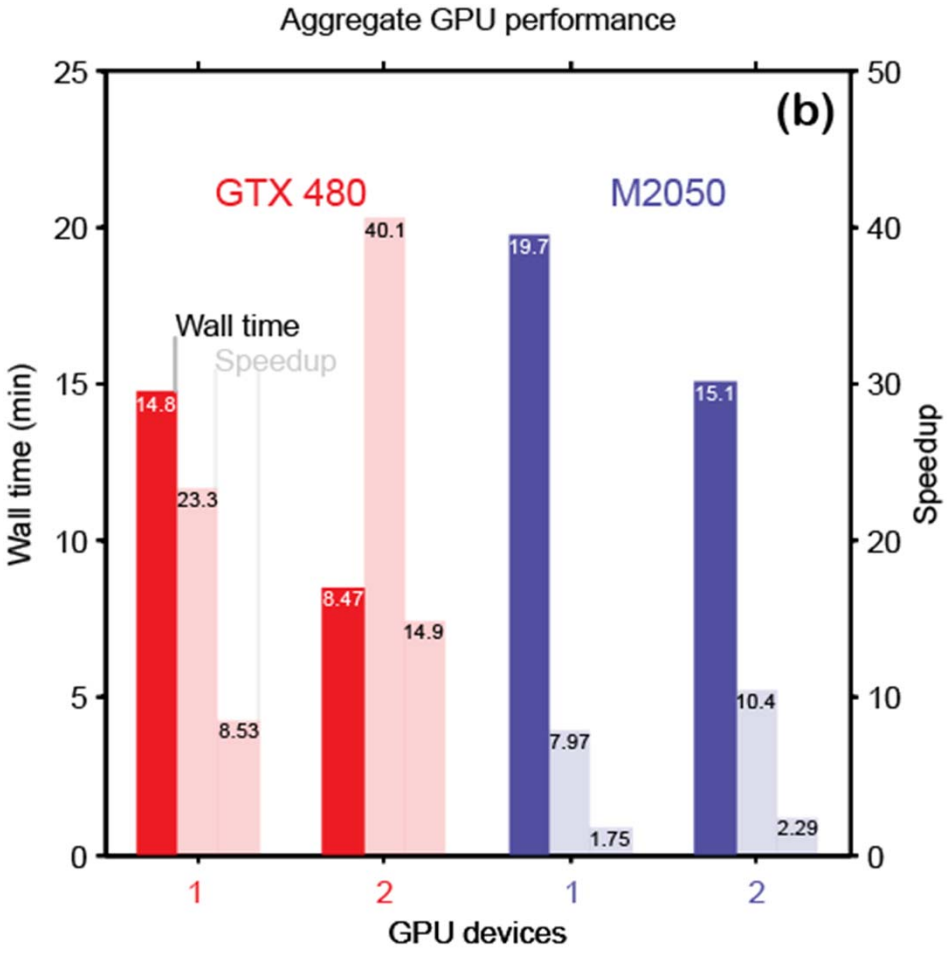
Overview

- Compared performance of two systems equipped with either dual NVIDIA GTX 480 or M2050 GPU accelerators for a single-precision radar data processing application



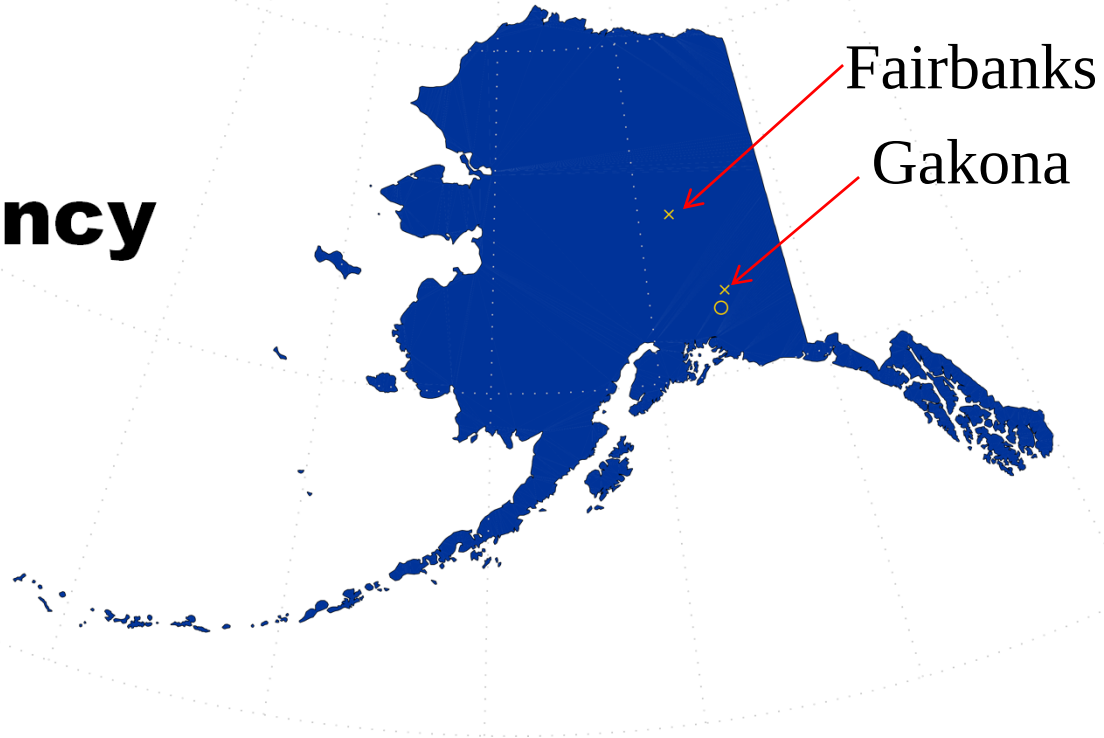
Overview

- Main result:
  Both the GeForce GTX 480 and Tesla M2050 systems exceed performance objectives, but the GTX 480 is a better choice given single-precision input data



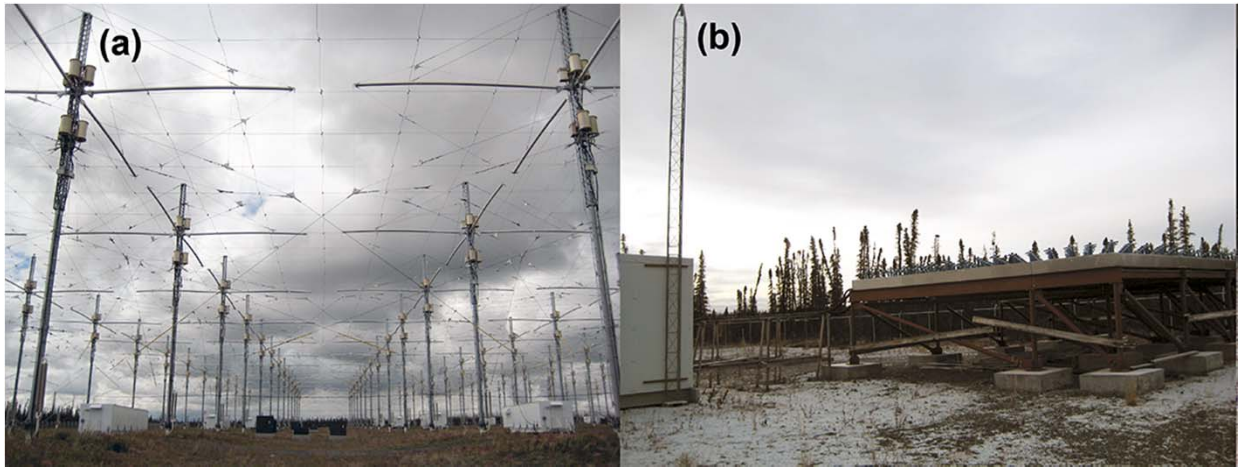
Motivation

- HAARP
  High-frequency
  Active
  Auroral
  Research
  Program



  - DoD observatory in Gakona for pure and applied ionosphere research

“HF Pump” **“Diagnostic probe”**



3.6 MW high-gain HF transmitter heats and modifies the ionosphere

Up to ~10 MW net power supplied by on-site diesel generators

Diagnostic instruments suite includes **MUIR**

Modular **U**HF **I**onosphere **R**adar detects strong HAARP-generated plasma waves in the F-region ionosphere


UP CONVERTER

Poker Flat Incoherent Scatter Radar (PF-ISR)

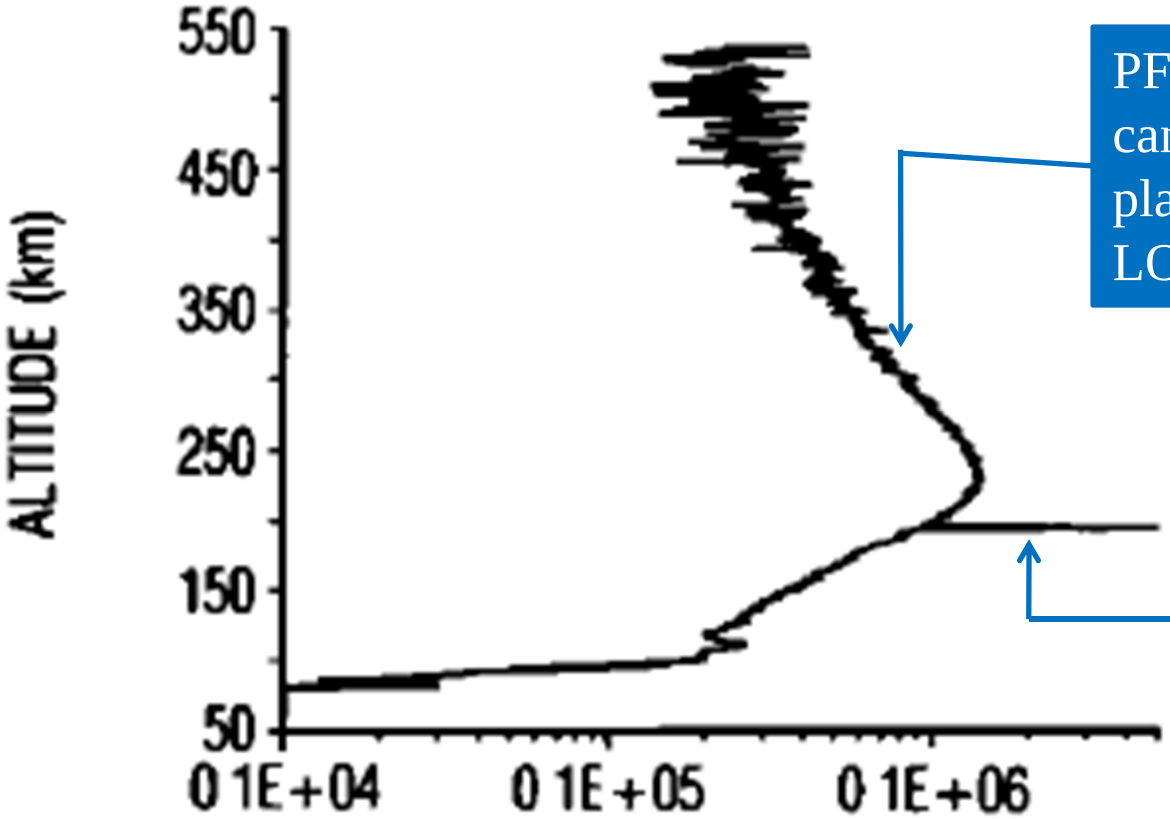
MUIR: 16 panels

PFISR is an Advanced Modular Incoherent Scatter Radar (AMISR)
128 panels,
32 antenna element units (AEUs) each.


AEU

PI: Craig Heinselman
SRI International

ISR vs. diagnostic radar



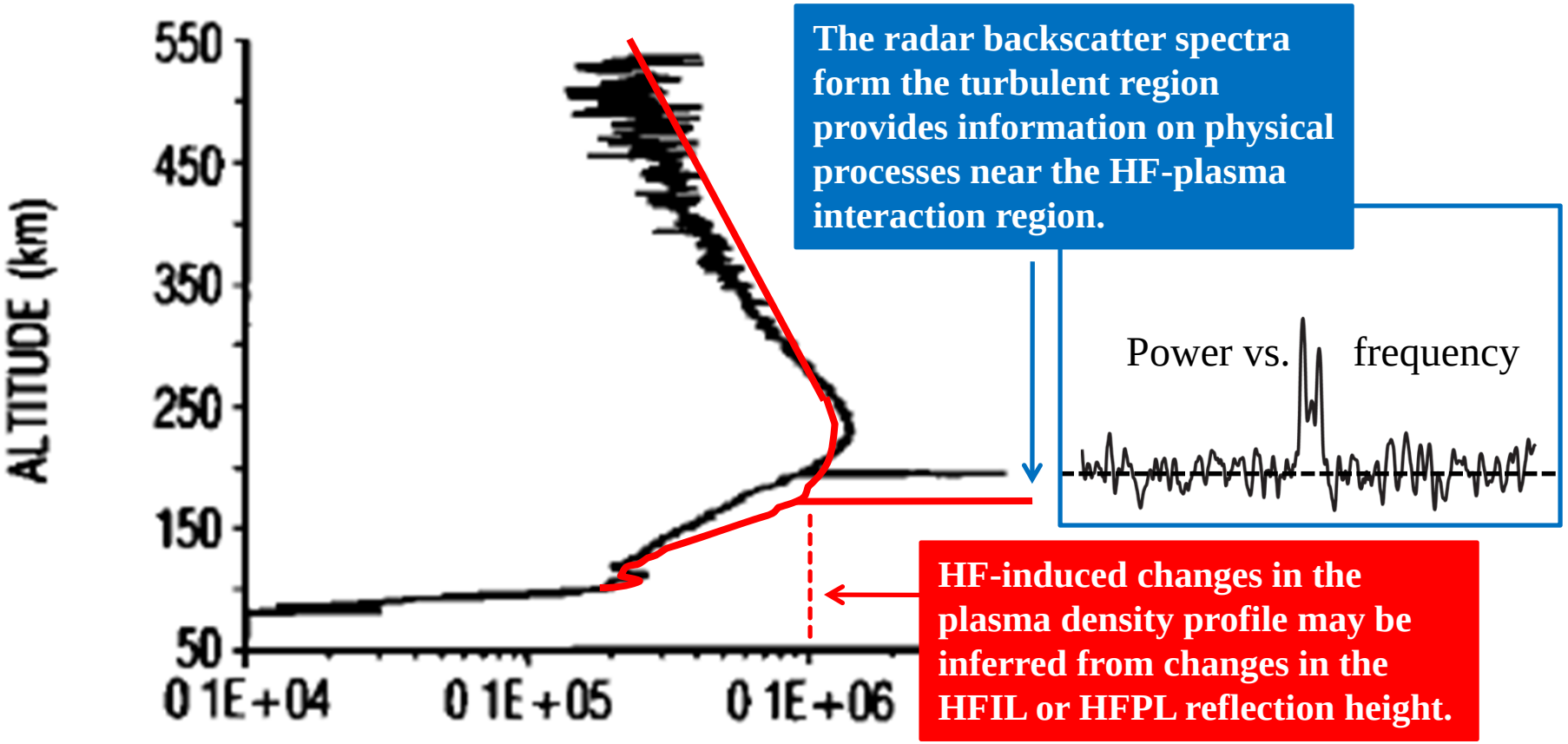
PF-ISR (700 m^2 antenna aperture) can measure altitude profiles of plasma density, temperature, and LOS velocity

MUIR (100 m^2 antenna aperture) at HAARP can measure the position and spectra of signals scattered from HF-enhanced ion-line or plasma-line turbulence

Range-corrected backscatter UHF radar power measured at Arecibo

[Djuth, F. T., et al. GRL 1987]

Experiments with MUIR diagnostic radar at HAARP

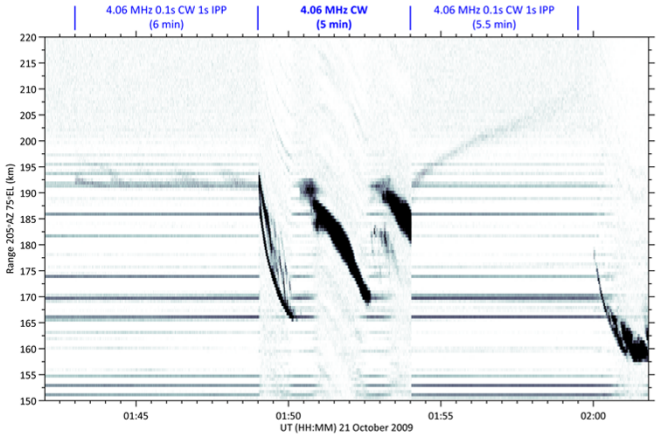


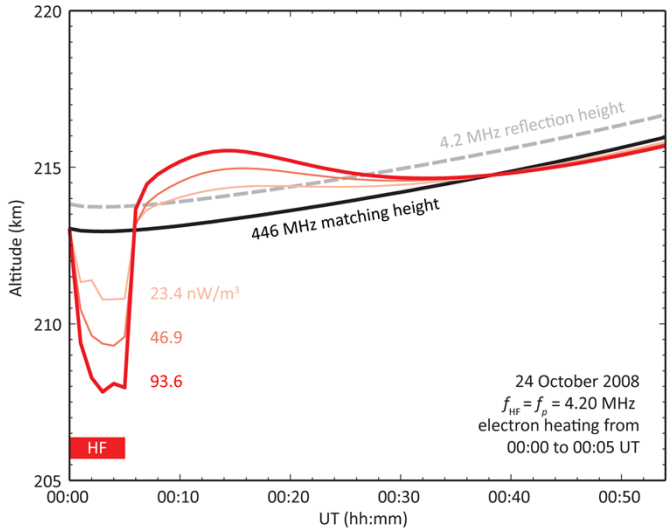
Range-corrected backscatter UHF radar power measured at Arecibo

[Djuth, F. T., et al. GRL 1987]

Movie

MUIR applications



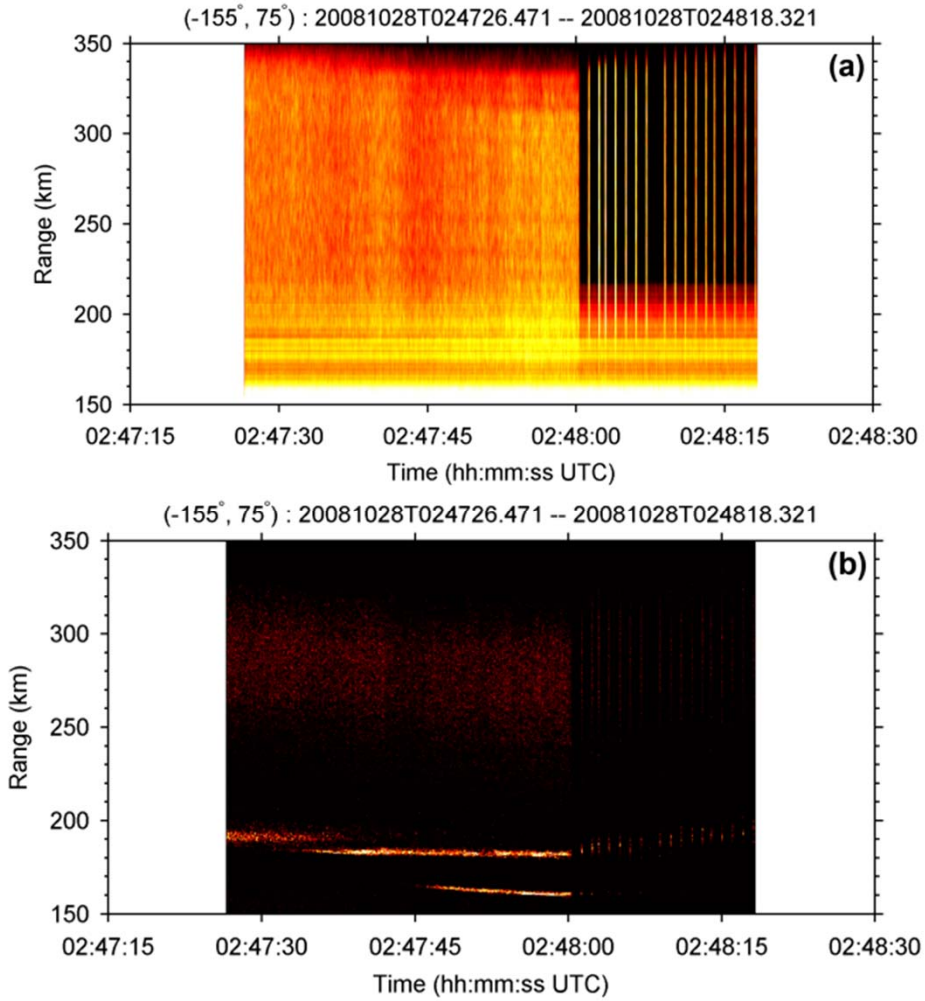


- Ionosphere model validation
- Measurement of HF reflection heights with high time and spatial resolution
- Proxy measurements of artificial airglow emission altitudes

Coded long pulse processing

- Target radar cross section is small so a long pulse (1 ms) is Tx to improve SNR at the expense of range resolution (150 km)

- Phase-coding techniques can improve radar range resolution to (0.6 km) *with significant post-processing*.



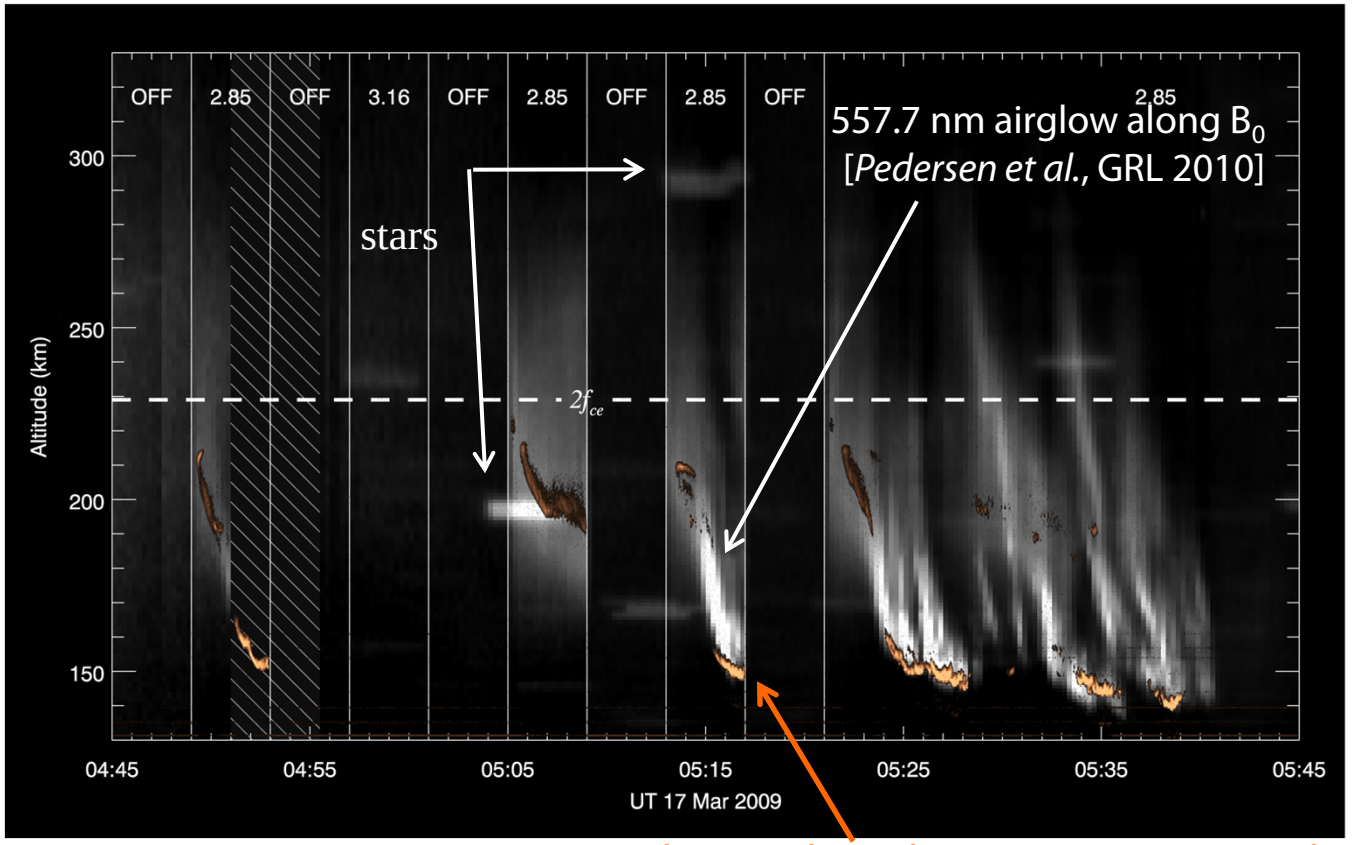
GTX 480 System

- TAU
  - Linux workstation from Penguin Computing
    - 64-bit 2.4 GHz AMD Opteron quad-core CPU
    - 4 GB RAM
    - Red Hat Enterprise Linux 5.6
  - **GeForce GTX 480 accelerator (x2)**
    - **480 CUDA cores**
    - **1.5 GB dedicated RAM**
    - **177 GB/sec bandwidth**
    - **1350 gigaflop peak single-precision performance**

M2050 System

- PACMAN
  - Linux cluster from Penguin Computing with two GPU nodes
    - Two 64-bit 2.4 GHz Intel Xeon quad-core CPUs
    - 64 GB RAM
    - Red Hat Enterprise Linux 5.5
  - **Tesla M2050 accelerator (x2)**
    - **448 CUDA cores**
    - **3.0 GB dedicated RAM**
    - **149 GB/sec bandwidth**
    - **1030 gigaflop peak single-precision performance**

Test collection: 71-files (one hour) of MUIR data



HFIL intensity along B_0 measured with MUIR [Fallen, PhD thesis 2010]

Radar data test collection

- One hour of MUIR coded long pulse data from a typical HAARP experiment
  - 71 files of ~44 MB each
  - 360,000 frames (includes Tx pulse and Rx)
    - 5000 frames per file
    - Each frame contains 1100 complex-valued single-precision samples
    - ~390,000,000 total range-time bins to process
  - CPU processing time
    - ~32 cores needed for “real time” processing
    - ~1000 sec single-core processing per 50 sec of radar data
- CPU processing is not sufficient to use high-resolution MUIR data on-site during experiment campaigns (when the data is most valuable)

Implementation

CPU

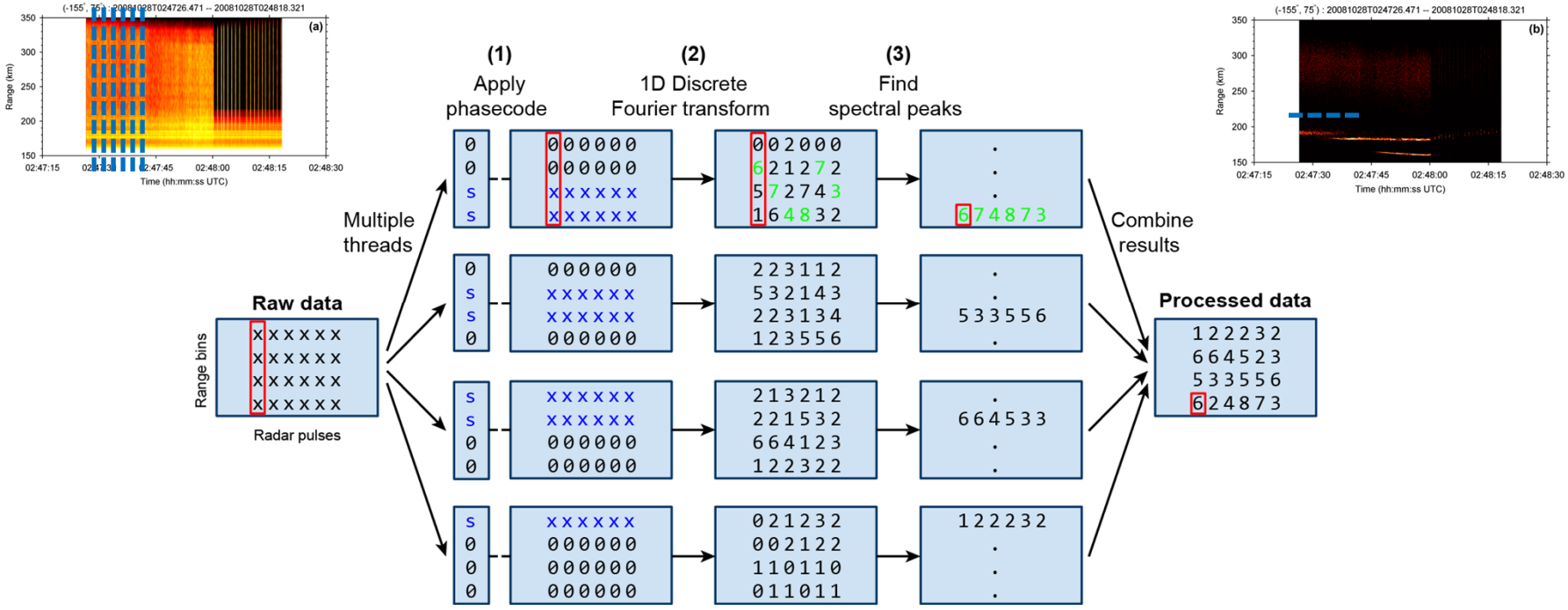
- Multithreaded computation with OpenMP
  - Each thread calculates one row of the final image
  - Discrete Fourier transform calculated with FFTW
  - Ad hoc methods for “phase code” and “peak find” steps.

GPU

- One thread per GPU device accessed via OpenCL
  - Each thread operates on a separate data file
  - DFT calculated with Apple OpenCL_FFT
  - Ad hoc kernels for “phase code” and “peak find” steps.

Algorithm

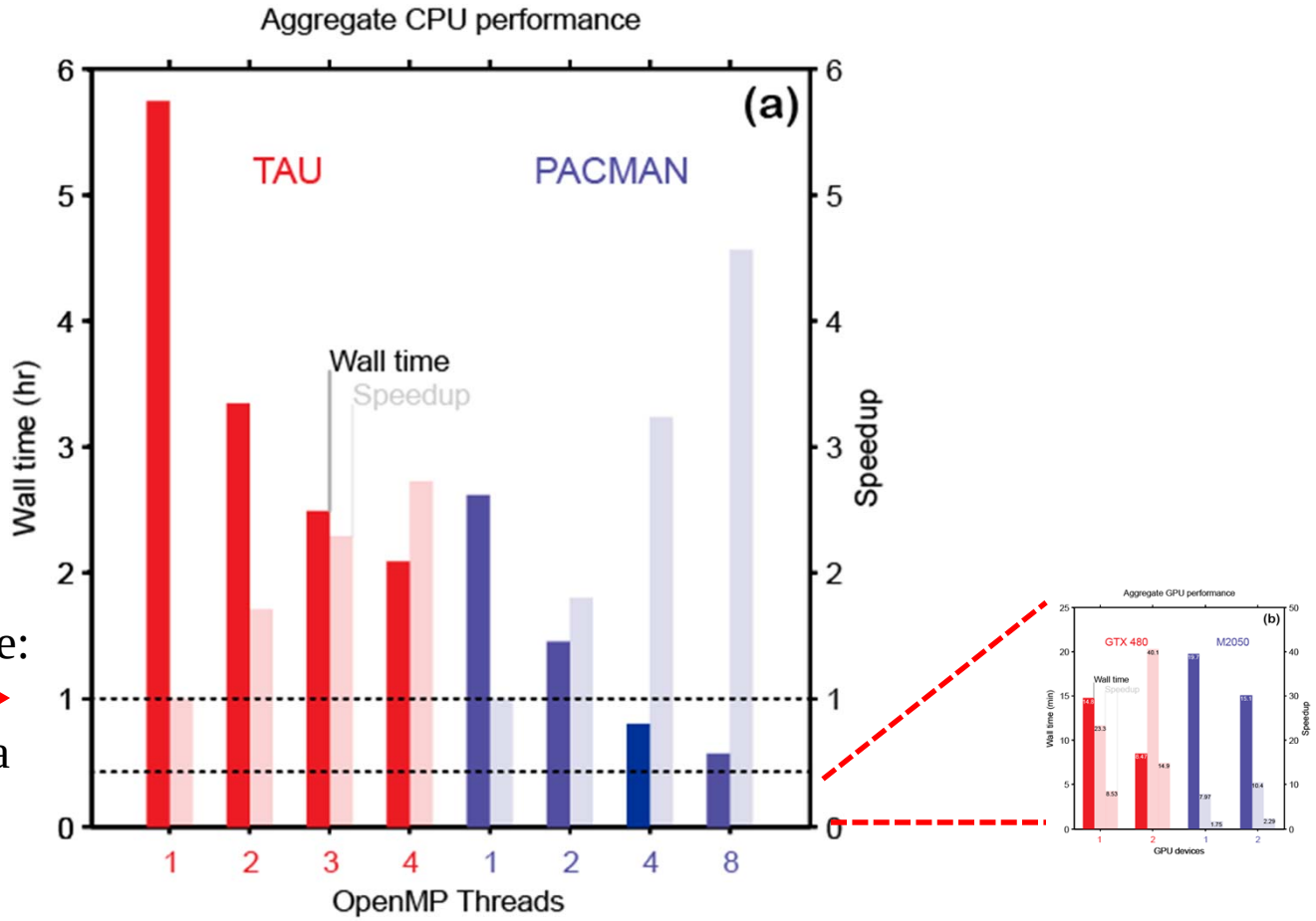
1. “Phase code” (Array multiply)
2. Calculate power spectra (FFT)
3. Identify spectral peak (array power, max-find)



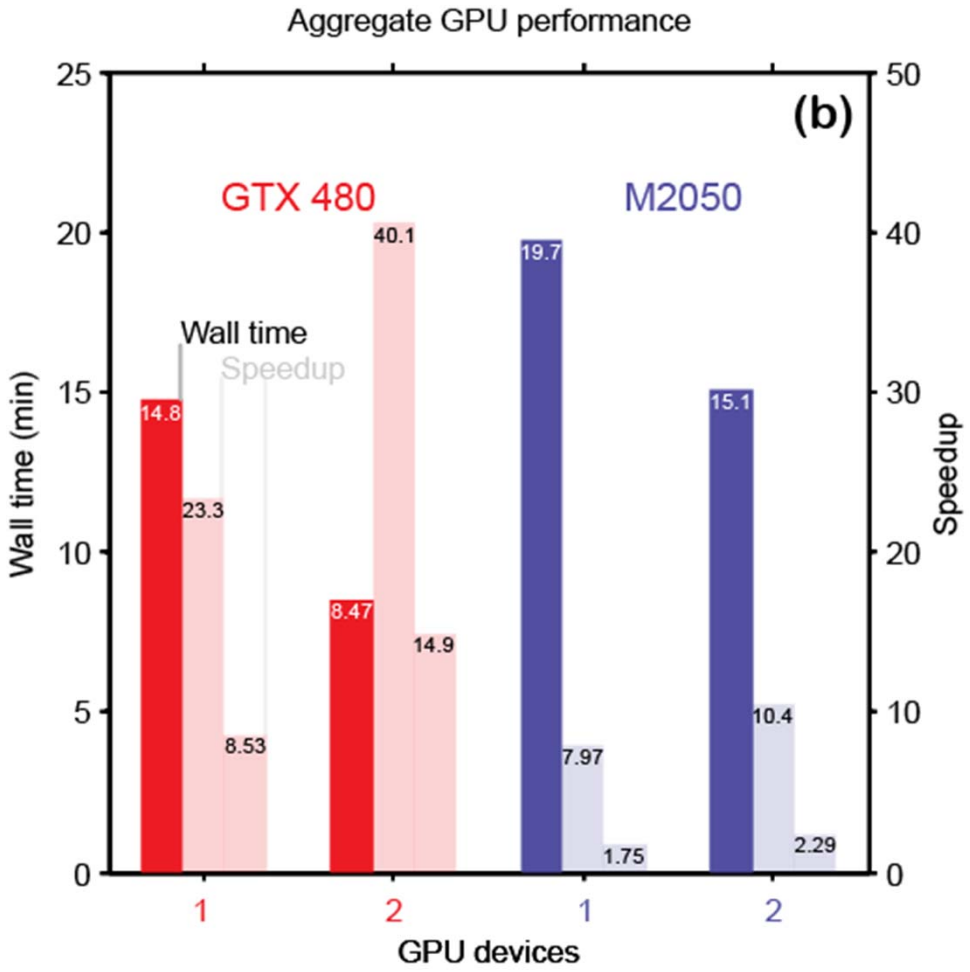
Results: CPU processing time

Performance objective:

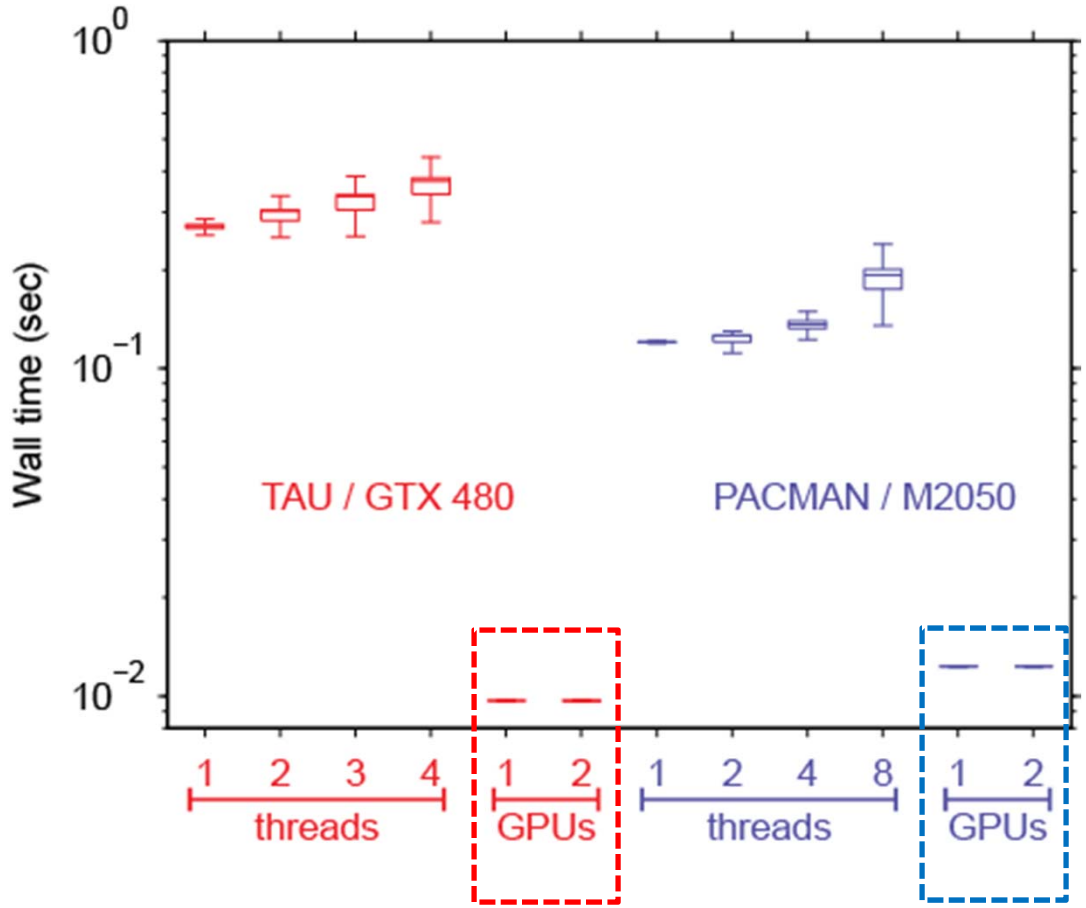
Process 1 hour of data in 1 hour or less



GPU processing time



Performance variability



multithreaded
~×2 penalty
implies potential
for CPU
performance
improvement

Task-level speedup

Time (per thread) to complete each step on one 5000-element row

Machine	Experiment	Setup (0)		Phasecode (1)		FFT (2)		Power (3)		Peakfind (4)		Row total	
		Time (ms)	Speedup	Time (ms)	Speedup	Time (ms)	Speedup	Time (ms)	Speedup	Time (ms)	Speedup	Time (ms)	Speedup
TAU	1 thread	53.76	1.00	129.51	1.00	45.30	1.00	29.61	1.00	0.01	1.00	258.19	1.00
	2 thread	63.68	1.69	140.05	1.85	53.66	1.69	29.67	2.00	0.01	1.88	287.07	1.80
	3 thread	72.98	2.21	153.81	2.53	57.75	2.35	30.22	2.94	0.02	1.63	314.79	2.46
	4 thread	85.18	2.52	168.65	3.07	63.98	2.83	31.31	3.78	0.02	2.48	349.13	2.96
	1 GTX 480	0.00	N/A	0.17	783.11	0.86	52.66	2.71	10.94	5.94	-617.56	9.67	26.70
	2 GTX 480	0.00	N/A	0.17	1567.90	0.85	106.10	2.71	21.88	5.95	-309.28	9.67	53.38
PACMAN	1 thread	31.85	1.00	33.78	1.00	26.95	1.00	24.36	1.00	0.01	1.00	116.94	1.00
	2 thread	36.24	1.76	33.88	1.99	27.02	1.99	24.40	2.00	0.01	1.52	121.55	1.92
	4 thread	42.60	2.99	37.31	3.62	27.25	3.96	24.49	3.98	0.02	0.90	131.67	3.55
	8 thread	67.79	3.76	57.45	4.70	33.52	6.43	24.56	7.94	0.09	0.49	183.40	5.10
	1 M2050	0.00	N/A	0.22	153.47	1.25	21.47	3.55	6.87	7.28	-1374.01	12.31	9.50
	2 M2050	0.00	N/A	0.22	307.08	1.26	42.93	3.55	13.73	7.28	-686.55	12.30	19.01

Speedup is calculated relative to single-thread CPU performance

Summary

- GPU acceleration for ionosphere radar has promising potential ...
  - But post-spectra processing tasks may erase performance gains
- Single-precision performance is still important for some applications

Next steps: Heterogeneous processing?