GPU-Powered WRF in the Cloud for Research and Operational Applications

John Manobianco, Chief Scientist
Don Berchoff, Chief Technical Officer
john@tempoquest.com, don@tempoquest.com

2017 Modeling Research in the Cloud Workshop
Boulder, Colorado
31 May 2017

Introduction
Approach
GPUs, GPU-Powered WRF
Benchmarks, Challenges
Summary, Q&A

TempoQuest

Who is TempoQuest, Inc. (TQI)?

- Scientific and Visualization Software Company
  - Port applications from CPU hardware to NVIDIA Graphics Processing Units (GPUs)
  - Accelerate processing, ultimately enhancing precision and accuracy
    - First Product (AceCAST): Weather Research and Forecasting (WRF) Model
    - Target: 3x – 10x acceleration, enable finer grid spacing, more ensembles, better physics, etc.
    - Parallel effort on accelerated visualization software (WSV3 Professional)
    - Future plans include other models, data assimilation, visual analytics, machine learning
- Platforms
  - Client data center and/or commercial cloud
  - Software As a Service (SaaS) in the commercial cloud

Partners

- Equity Investors (founders, venture capitalists)
- Space Sciences and Engineering Center (SSEC; University of Wisconsin, Madison)
  - Software development / testing
- NVIDIA
  - Investor
  - Access to development hardware
  - Subject matter expertise
  - Software testing, integration, benchmark
  - Marketing, sales, and distribution assistance for GPU forecast products

TQI Approach

User 1 → Internal HPC
User 2 → Shared HPC
User 3 → Internal HPC
• • •
User N → Internal HPC

MIGRATE USER EXPERIENCE

Research & operational applications

Client 1
Client 2
Client 3
• • •
Client N

AceCAST Support S/W Apps and User Interface
(engineering tools to kick off model cycle)

GPU WRF
GPU / CPU Hybrid
CPU WRF

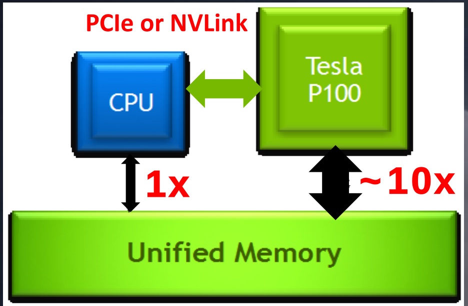
Commercial cloud
Custom cloud
Private cloud

Provide secure, scalable, reliable, intuitive, and flexible execution

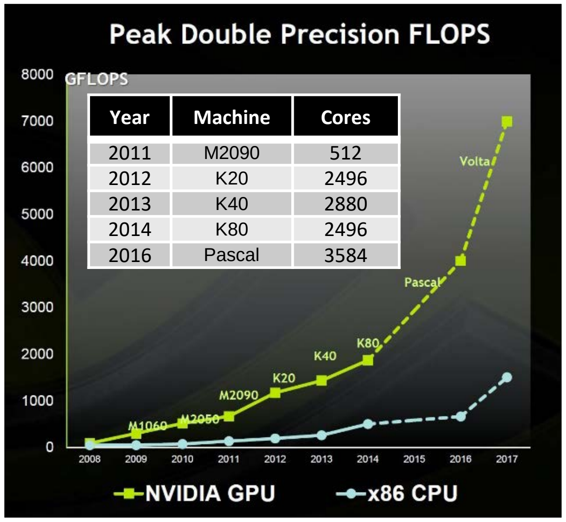
TempoQuest

NVIDIA Graphics Processing Unit (GPU)*

GPU Introduction



- Co-processor to the CPU
- Threaded Parallel (SIMT)
- CPUs: x86 | Power | ARM
- HPC Motivation:
  - Performance
  - Efficiency
  - Cost Savings



Year	Machine	Cores
2011	M2090	512
2012	K20	2496
2013	K40	2880
2014	K80	2496
2016	Pascal	3584

IMAGE: Facebook's new Big Sur GPU server
http://venturebeat.com/2016/08/29/facebook-gives-away-22-more-gpu-servers-for-a-i-research

*Original slide contents from Stan Posey, HPC Program Manager, NVIDIA

Programming Strategies for GPU Acceleration*

Applications

GPU Libraries	OpenACC Directives	Programming in CUDA (API)
Provides Fast "Drop-In" Acceleration	GPU-acceleration in Standard Language (Fortran, C, C++)	Maximum Flexibility with GPU Architecture (Code rewrite / restructure)

Increasing Development Effort
Increasing Acceleration

NOTE: Many applications include combination of these strategies

*Original slide contents from Stan Posey, HPC Program Manager, NVIDIA

Benchmarks with WRF CUDA Modules

- Module timing
  - Write driver routine for each module
  - Compare timing on CPU versus GPU
  - Examine accuracy of results
- Integrate one or more modules in full WRF model
  - Run realistic test cases on representative HPC hardware
  - Compare timing for CPU versus GPU+CPU versions
  - Compare output for scientific validation
    - Graphic fields
    - Dependent and derived variables averaged in space, time
    - Integrated quantities

Individual WRF Module Testing

Module	CUDA V3.6.1	GPU Speed Up (w/wo I/O)	CUDA V3.8
Kessler MP	√	70x / 816x	
Purdue-Lin MP	√	156x / 692x	
WSM -3-class MP	√	150x / 331x	
WSM 5-class MP*	√	202x / 350x	
Eta MP	√	37x / 272x	
WSM 6-class MP*	√	165x / 216x	√
Goddard GCE MP	√	348x / 361x	
Thompson MP*	√	76x / 153x	√
SBU 5-class MP	√	213x / 896x	
WDM 5-class MP	√	147x / 206x	
WDM 6-class MP	√	150x / 206x	
RRTMG LW*	√	123x / 127x	√
RRTMG SW*	√	202x / 207x	√
Goddard SW	√	92x / 134x	
Dudia SW*	√	19x / 409x	√
MYNN SL	√	6x / 113x	
TEMF SL	√	5x / 214x	
Thermal diffusion LS	√	10x / 311x	√
YSU PBL*	√	34x / 193x	√
Betts Miller Janjic CP		———	√

*Previous NVIDIA-SSEC project

Integrated WRF Module Testing

- 4 CUDA module integration w/ WRF V3.8
  - WSM6, Thompson MP, RRTMG SW, RRTMG LW
- NVIDIA PSG cluster (12 nodes), high speed InfiniBand
  - 384 CPU cores (Haswell chips, 16-cores per chip, 2 chips per node)
  - 48 GPUs (P100s, 4 per node)
- Tornado case (upper mid west and Ohio valley)
  - 2100 UTC 12 Jun 2013
  - 3-hour benchmark
  - 625 x 625 x 50 levels, 3-km grid
- Hybrid WRF (select physics modules on GPU, all other code on CPU)
  - 35% speed up WSM6, RRTMG SW, RRTMG LW
  - 38% speed up Thompson, RRTMG SW, RRTMG LW
- Use wrfdiff to compare statistics on U, V, W, P, etc. between CPU and GPU runs

TempoQuest

Challenges Moving Forward

- Software
  - Porting all code to GPUs
    - Avoid communication penalties: CPU ⇔ GPU
    - Not all modules benefit from conversion to CUDA (use OpenACC to get resident on GPU)
  - Sustainment (latest versus historical versions of WRF; why important?)
  - Portability and customization for specific generation of hardware
- Cloud
  - Various offerings, pricing, etc. from different vendors
  - Elastic computing (resources that are not permanent, but scalable)
  - Data and software management including storage
  - Reduce data transfer (bring applications and software to cloud) – what about proprietary algorithms, sensitive, or classified data?
  - Lack of high speed interconnection (Infiniband)
  - Viable business model (ROI analysis, product pricing, profitability)

Schedule and Milestones

2016 | 2017 | 2018 | 2019

Hybrid WRF
- 8 CUDA physics modules
- Target 25% acceleration

AceCast v1
- 8 CUDA physics modules
- Dynamics (scalar advection in CUDA)
- Beta client testing

AceCast v2
- 20 CUDA physics modules
- 100% dynamics
- 70%-80% of client namelist needs
  - Agriculture
  - Renewable energy
  - Fire weather
  - Transportation
  - Surface weather (weather companies)

AceCAST Support S/W Apps
- Initial SaaS offering Mar 2019
- Supporting tools and infrastructure

Summary

- Introduction to TempoQuest (TQI)
- TQI plans and progress to enable GPU-powered WRF in the cloud
  - Target 3x – 10x acceleration
  - Broad user base with initial focus on 5 verticals (research and operations)
  - Provide SaaS with interface and application layers
  - Achieve 70%-80% of client namelist requirements by March 2019
- Software and Cloud Challenges
- Comments and Questions?