GPU developments for the WRF and MPAS Models

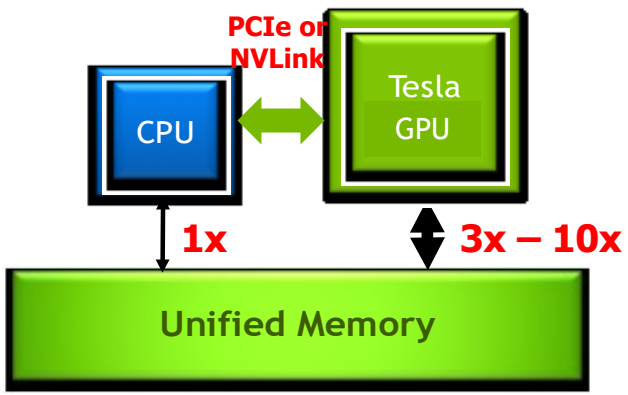
Jeff Adie, Carl Ponder, Alexey Romanenko, Stan Posey, NVIDIA, Santa Clara, CA, USA

TOPICS OF DISCUSSION

- GPU INTRODUCTION
- WRF DEVELOPMENTS
- MPAS DEVELOPMENTS

NVIDIA GPU: Introduction and Hardware Features

GPU Introduction



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

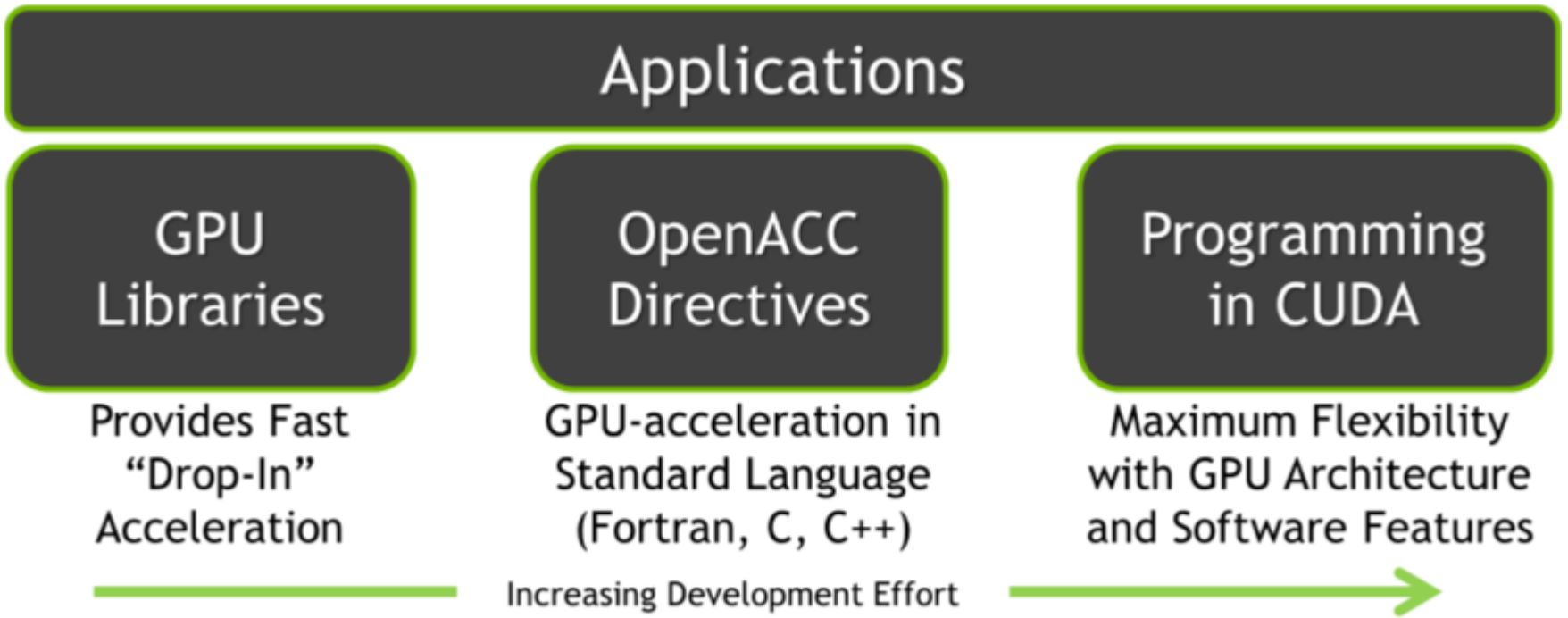
ORNL Summit: 4,600 nodes; 27,600 x V100 GPUs

NVIDIA GPU: Introduction and Hardware Features

	V100 (2017)		P100 (2016)		K40 (2014)
Double Precision TFlop/s	**7.5**	1.4x - 5.4x	5.3	3.8x	1.4
Single Precision TFlop/s	**15.0**	1.4x - 3.5x	10.6	2.5x	4.3
Half Precision TFlop/s	**120 (DL)**	~6x	21.2		n/a
Memory Bandwidth (GB/s)	**990**	1.3x - 3.4x	720	2.5x	288
Memory Size	**16 or 32GB**	2.00x	16GB	1.33x	12GB
Interconnect	NVLink: Up to 300 GB/s PCIe: 32 GB/s		NVLink: 160 GB/s PCIe: 32 GB/s		PCIe: 16 GB/s
Power	**300W**	1.00x	300W		235W

V100 Availability	DGX-1: Q3 2017; OEM : Q4 2017

Programming Strategies for GPU Acceleration



NOTE: Many application developments include a combination of these strategies

Examples

- IFS: FFT, DGEMM
- COSMO: Tridiag Solve

- FV3
- MPAS
- IFS
- ICON
- COSMO (Physics)
- WRFg

- ACME
- CAM-SE
- NICAM
- ICON
- NEMO
- UM/Gungho

- COSMO (Dycore)
- NUMA
- ICON
- NICAM (Dycore)

TOPICS OF DISCUSSION

- GPU INTRODUCTION
- **WRF DEVELOPMENTS**
- MPAS DEVELOPMENTS

WRF Current GPU Status - June 2018

- **NVIDIA**
  - Based on release 3.7.1 – COMMUNITY version; also some 3.6.1 code
    (Further details later in this presentation)

- **TQI**
  - Based on WRF release 3.8.1 – PROPRIETARY (see TQI for licensing)
  - NVIDIA alliance with TQI; NVIDIA engineering support provided
  - TQI early round investors include NVIDIA

- **NCAR**
  - Release 4.0 during Q2 2018
  - Development relationship with both NVIDIA and TQI
  - Have expressed interest and value on commercialization of WRF

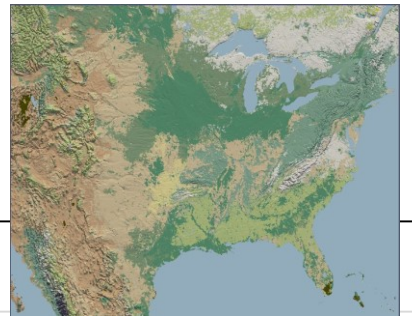
WRF Current GPU Status - WRFg - June 2018

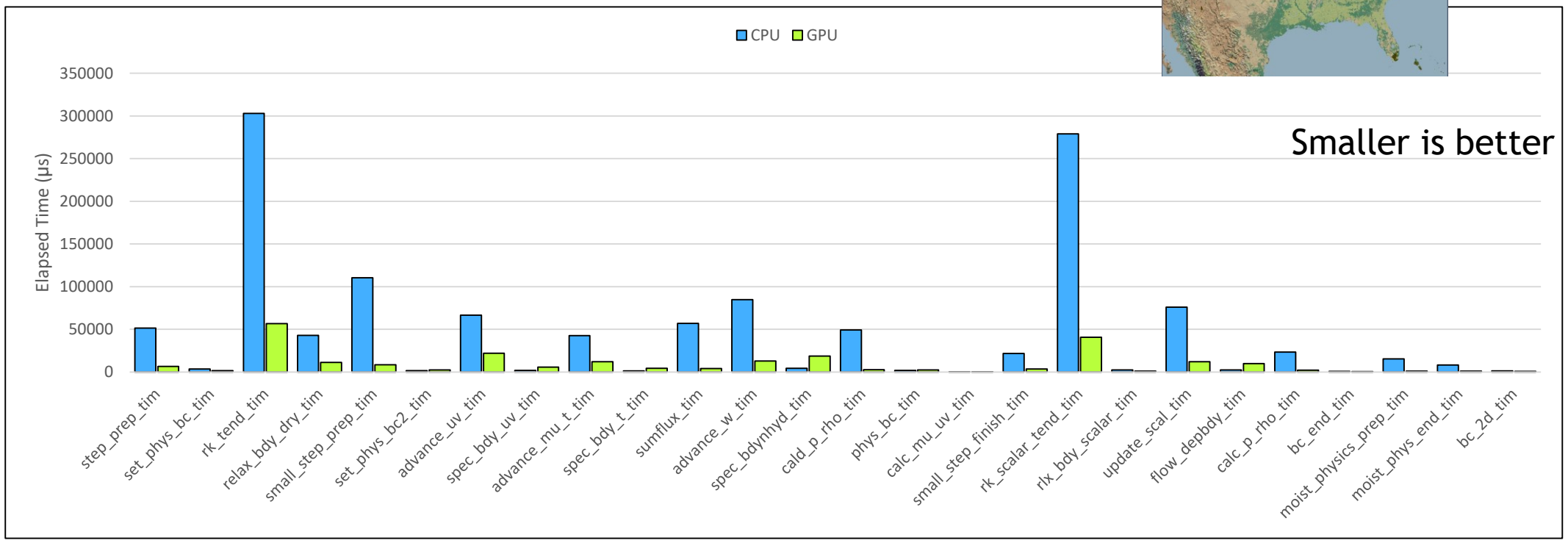
- **WRFg**
  - **Based on ARW release 3.7.1**
  - **Limited physics options:**
    - Enough for full model on GPU
    - Further development ongoing
  - **Community release:**
    - Freely available objects/exe
    - Restricted source availability
  - **Availability during Q4 2018:**
    - Benchmark tests – NOW
  - **Support – TBD; Roadmap – TBD**
  - **Performance demonstrations:**
    - ~4x speedup full model (CPU:GPU)

- **WRFg Physics Options**

Microphysics	Option
WSM6	6
Thompson	8
WSM5	4
Radiation	
RRTM(lw)	1
Dudhia (sw)	1
RRTMG_fast(lw,sw)	24
Planetary boundary layer	
YSU	1
GWDO	
Surface layer:	
Revised MM5	1
Land surface:	
5-layer TDS	1

WRFg Dycore Performance



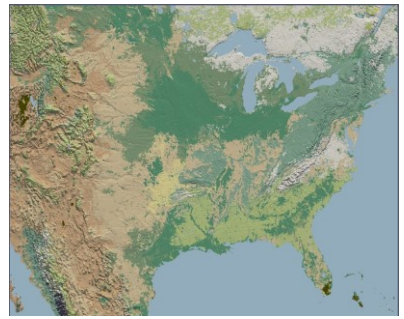


Average speedup = 5.5x, Best speedup for calc_p_rho = 18x

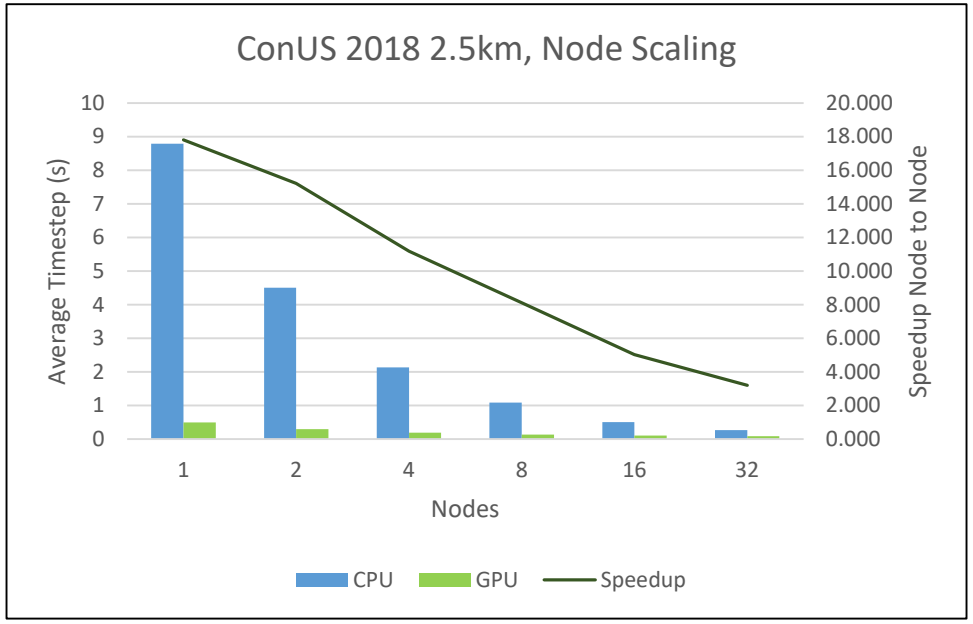
ConUS 2.5km 2018* Model, 2xE5-2697V3 CPU, 4xPascal P100 GPU

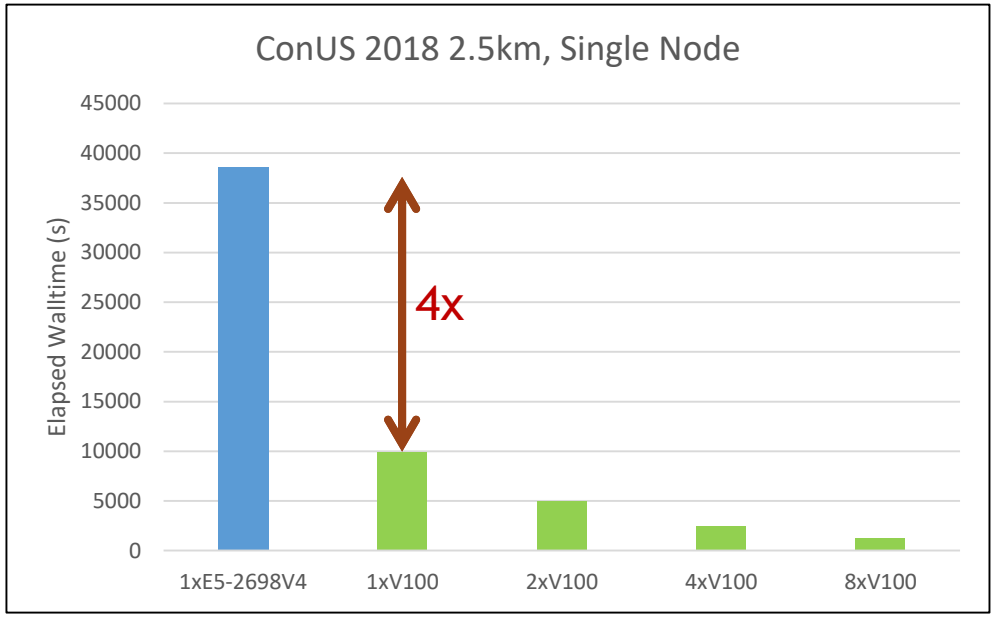
*Modifications to standard ConUS 2.5: WSM6, RRTMG, radt=3, 5-layer TDS

WRFg Full model Performance



Smaller is better





1x CPU: 10hr 42m, 1x GPU: 2hr 45m, 4x GPU: 41m, 8x GPU: < 21 m

ConUS 2.5km 2018* Model, 2xE5-2698V4 CPU, 8xVolta V100 GPU

*Modifications to standard ConUS 2.5: WSM6, RRTMG, radt=3, 5-layer TDS

Proposed: NVIDIA–TQI Development Agreement

NVIDIA has proposed an exclusive GPU WRF development agreement with TempoQuest

- **NVIDIA recognizes TempoQuest as the world leader in GPU WRF**
  - Proven GPU expertise at TQI for WRF development (see talk 5.5, Abdi)
- **TQI to lead validation of the community release "WRFg" model**
  - During Q3, NVIDIA will provide WRFg development code to TQI for testing
  - During Q4, WRF users will have a choice of solutions for GPU WRF from TQI:
    - WRFg – community available (free) software – moderate performance
    - AceCAST – commercial licensed software – high performance option
- **NCAR has endorsed an NVIDIA – TQI development alliance for WRFg**
  - TQI and NVIDIA synergies will strengthen overall GPU WRF deployments
- **Details of the final agreement will be announced during Q3 2018**

TOPICS OF DISCUSSION

- GPU INTRODUCTION
- WRF DEVELOPMENTS
- **MPAS DEVELOPMENTS**

GPU Developments for the MPAS Model

Weather and Climate Acceleration (WACA) Collaboration

Motivation

- Project Led by NCAR to prepare MPAS for accelerator evaluation
- Collaboration team at NCAR:
  - Rich Loft, CISL Director, WACA PI
  - Raghu Raj Kumar, CISL Scientist, GPU lead
  - Michael Duda, MMM Scientist, MPAS HPC lead
- Vendor technical teams:
  - PGI/NVIDIA: Brent Leback, Carl Ponder
  - IBM/TWC: Constantinos Evangelinos, Todd Hutchinson

Project Plan

- NCAR to focus on dycore initially, later physics for full model approach
- KISTI (KR) focus on critical physics, in collaboration with NCAR
- NCAR to complete re-integration of GPU-based dycore and physics for fully accelerated MPAS model
- IBM to test on P9 + V100 and NVLink

GPU Developments for the MPAS Model

Results Provided From an Invited Talk at NVIDIA GTC, March 2018:

MPAS on GPUs Using OpenACC: Portability, Scalability & Performance

Dr. Raghu Raj Kumar
Project Scientist I & Group Head
Special Technical Projects (STP) Group
National Center for Atmospheric Research

http://on-demand.gputechconf.com/gtc/2018/presentation/s8812-an-approach-to-developing-mpas-on-gpus.pdf

NCAR Talk at GTC 2018 (R. Kumar)

Single Node Performance: MPAS Dry Dycore

Dataset		Broadwell (Fully Subscribed, OpenMP Enabled, Intel compiled, Base code)	P100 with Haswell(1 GPU, PGI compiled, OpenACC code)	V100 with Haswell (1 GPU, PGI compiled, OpenACC code)	P100 with Power8(1 GPU, PGI compiled, OpenACC code)	Speedup Broadwell vs P100	Speedup Broadwell vs V100
120 Km (40K)	SP	0.40	0.28	0.19	0.26	1.54	2.16
	DP	0.88	0.40	0.29	0.35	2.51	2.99
60 Km (163K)	SP	1.90	1.02	0.69	1.01	1.88	2.74
	DP	3.80	1.54	1.12	1.41	2.70	3.40

Test Case: Baroclinic Instability Test

- Dry dynamics test-case produces baroclinic storms from analytic initial conditions
- Split Dynamics: 2 sub-steps, 3 split steps
- Current work: 60km resolution (163k grid points, dt=300s), 120km (40k grid points, dt=600s)
- Number of levels = 56
- **Double precision**
- Execution time for simulating 1 day

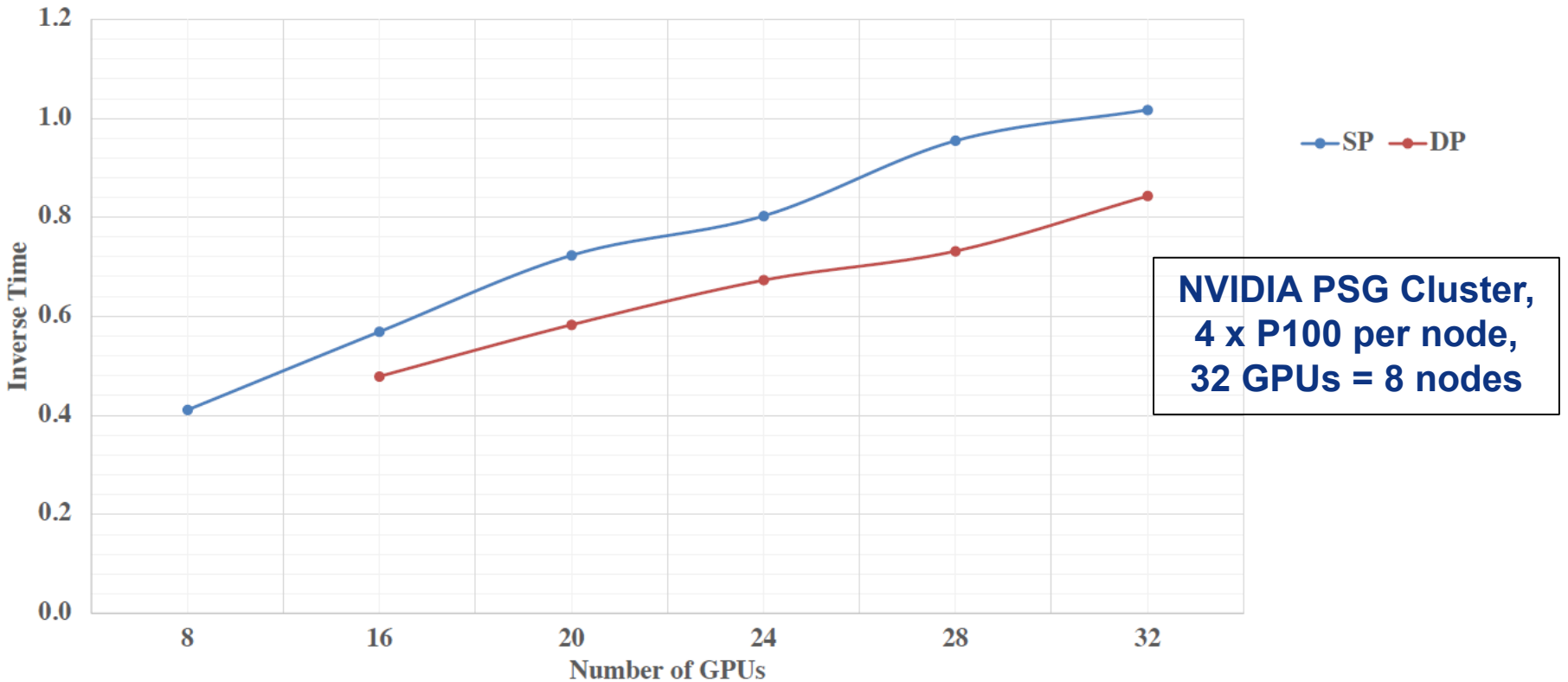
  144 timesteps (600 sec) for 120 km; 288 timesteps (300 sec) for 60 km

Source: MPAS on GPUs Using OpenACC: Portability, Scalability & Performance
http://on-demand.gputechconf.com/gtc/2018/presentation/s8812-an-approach-to-developing-mpas-on-gpus.pdf

NCAR Talk at GTC 2018 (R. Kumar)

MPAS Model for Prediction Across Scales — NCAR

Strong Scaling for MPAS Dry Dycore (SP & DP) for 15 Km (2.6M) on P100 GPU



Time per timestep, 4 GPUs per node, 1 MPI rank per GPU, Max of 4 MPI ranks per node, Intranode Affinity for MPI ranks, Uses OpenMPI, PCIe no NVLink, PGI 17.10.

Source: MPAS on GPUs Using OpenACC: Portability, Scalability & Performance
http://on-demand.gputechconf.com/gtc/2018/presentation/s8812-an-approach-to-developing-mpas-on-gpus.pdf

Summary: GPU developments for the WRF and MPAS Models

- **Good progress with GPU acceleration of the WRF and MPAS models**
  - NVIDIA and TQI to provide community WRFg – benchmarks available TODAY
  - TQI to provide proprietary AceCAST – contact TQI: www.tempoquest.com
  - NCAR and WACA alliance to provide MPAS on GPUs – contact PI, Dr. Rich Loft
- **Future work will develop more physics options based on user demand**
  - NVIDIA—TQI proposal for additional physics in WRFg based on user priorities
  - Outcome of NVIDIA—TQI development agreement announced during Q3 2018
- **Contact NVIDIA for details on GPU benefits to NWP and climate models**
  - HPC progress on several numerical models and AI models using deep learning
  - Contacts: Technical – jadie@nvidia.com; General – Stan Posey, sposey@nvidia.com

Questions?

Jeff Adie jadie@nvidia.com