Implementing the WRF Model on the German Grid

Jan Ploski, OFFIS

GEFÖRDERT VOM

Bundesministerium für Bildung und Forschung

WISENT: 8 Man-Years, 10/2005 - 09/2008

Technological features of the Grid infrastructure
OFFIS

Utilization of energy data
meteocontrol

Remote sensing data
DLR-DFD

Technical thermodynamics
DLR-PA

Environmental physics
DLR-TT

Research in Energy Meteorology
University of Oldenburg

OFFIS
Institute for Information Technology
Business Information Management Department

meteo control
meteocontrol GmbH
Energy & Weather Service

DLR
vIEM
German Aerospace Center e.V.
DLR-DFD, DLR-PA, DLR-TT

CARL VON OSSIETZKY universität OLDENBURG
University of Oldenburg
Institute of Physics, Energy and Semiconductor Research Dept. (EHF)
ForWind – Center for Wind Energy Research

Mesoscale modeling in WISENT

- **Task: near-real-time prediction of generated wind energy**
  - Includes research to improve prediction quality
- **Current means:**
  - WRFv2 model from NCAR $\rightarrow$ successor of MM5
  - High-performance cluster at OFFIS
- **Our (technical) goals:**
  - Expand from 1 cluster to distributed computing within the German Grid
  - Create an improved working environment for NWP researchers

Grid Computing and e-Science

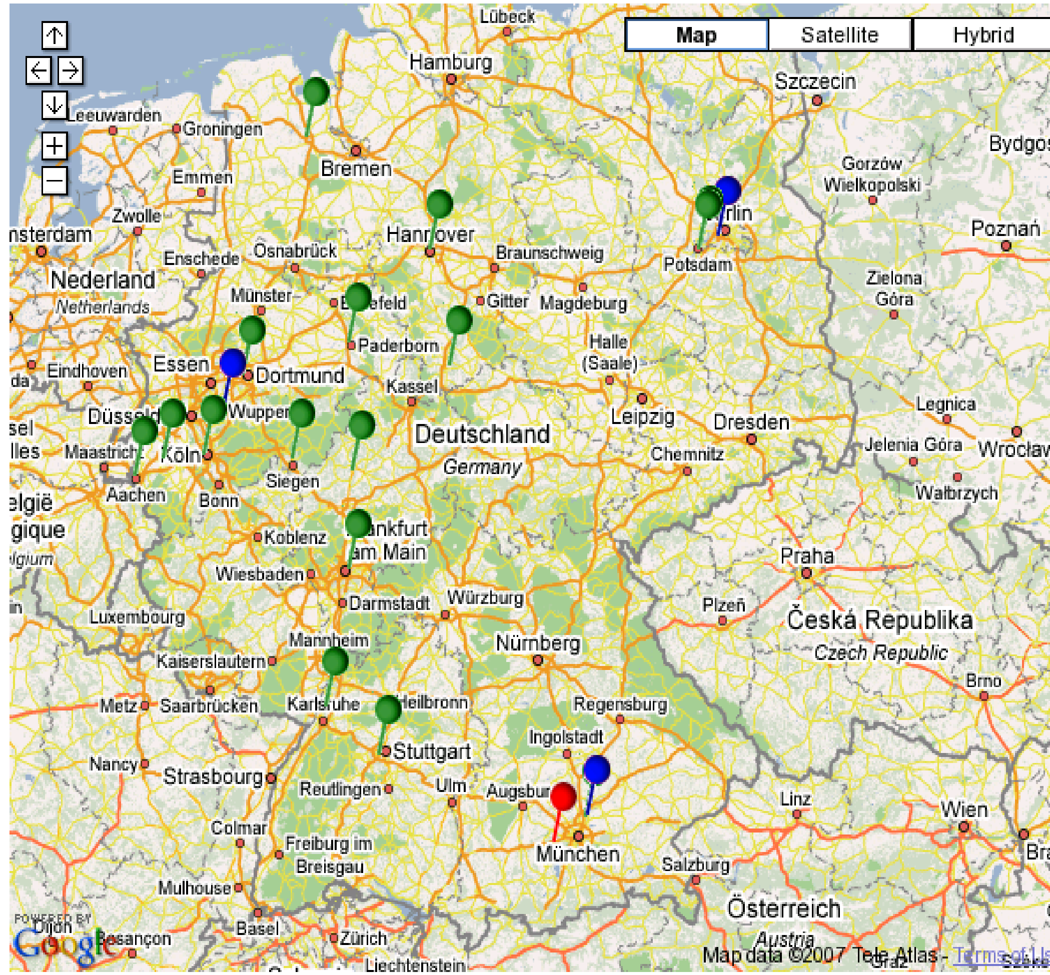
***Coordinated resource sharing and problem solving in dynamic, multi-institutional virtual organizations.* (Foster & Kesselman 2001)**

- **Resources:** **CPUs, disk space, memory, network bandwidth**
- **Sharing:** **not dedicated to a single user (group)**
- **Coordinated:** **with fair access, accountability, SLAs**
- **Problem solving:** **wind energy output prediction, research**
- **Multi-institutional VOs:** **collaboration between Uni OL, OFFIS**

- **The general idea of the Grid...**
  - connect multiple computing clusters
    - to solve bigger problems
    - to reach more users
    - to work more conveniently

The German Grid (D-Grid)

- **Multiple clusters, multiple providers**
- **Connected by DFN (transfer rates: 10-20 MB/s)**
- **4000+ CPU cores**
- **~2 PB total storage space (~40 TB for WISENT)**



- **Enough capacity to run hundreds of models in parallel.**

Benefits of Grid Computing for NWP research

- **Quality of service (compare platforms)**
- **Correctness (cross-check results)**
- **Community building**
- **Scalability/new business models**
- **Improved data storage and access**
- **Improved visibility for meteorological applications**

Benchmarking WRF performance

- **Little public data exists on performance tuning**
- **Benchmark results come from very large sites, tuned by experts**
- **“Small“ users must wonder...**
  - which compiler should I buy/use?
  - which compiler options?
  - which namelist parameters may affect performance?
  - OpenMP or MPI parallelization? or hybrid?
  - how many processors?

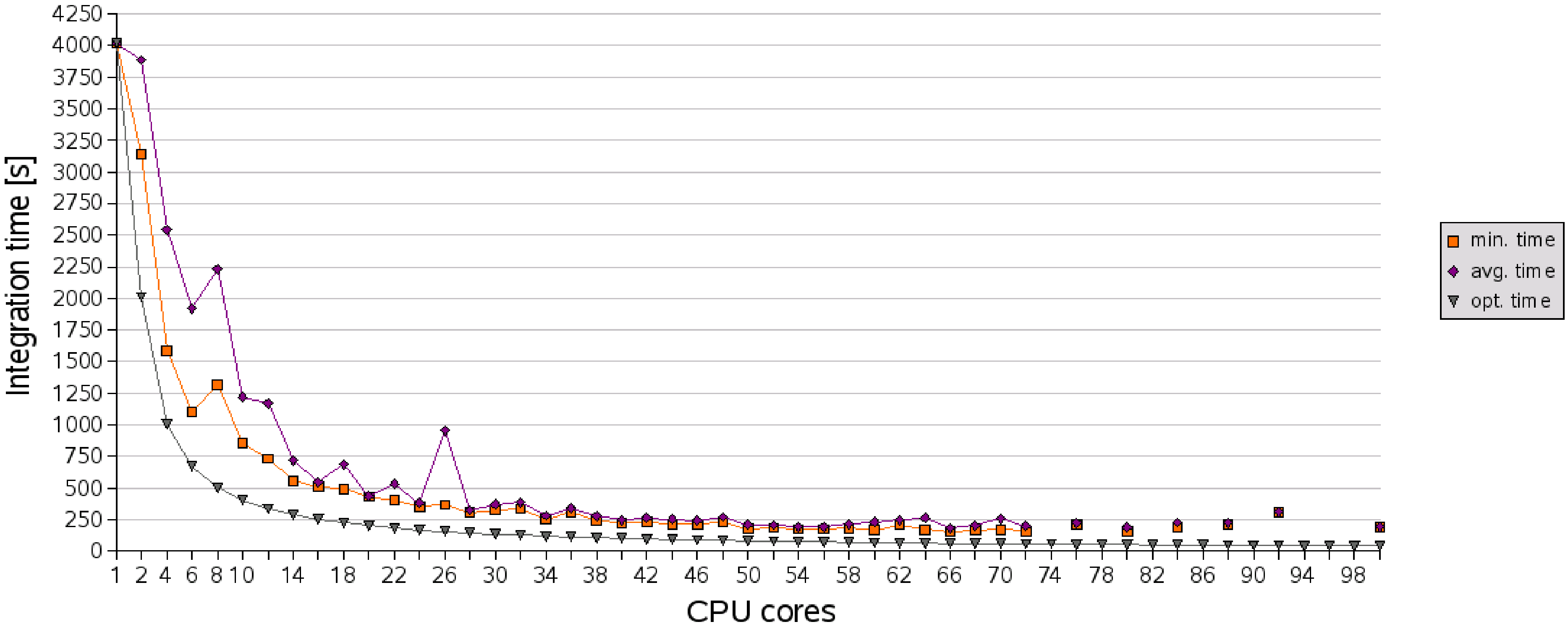
Benchmark cases

- **Case 1, ConUS2001_12km (from NCAR):**
  - **Single domain, 48 hour, 12km res., 425x300x35 grid, time step 72s**
- **Case 2, Katrina hurricane (from NCAR):**
  - **Single domain, 24 hours, 30 km res., 75x70x28 grid, time step 180s**
- **Case 3, Kyrill storm (our own):**
  - **Domain 1: 25 km res., 100x100x28 grid, time step 150s**
  - **Domain 2: 5 km res., 261x281x28 grid, time step 30s**

Detailed case descriptions, configurations and benchmark results soon at

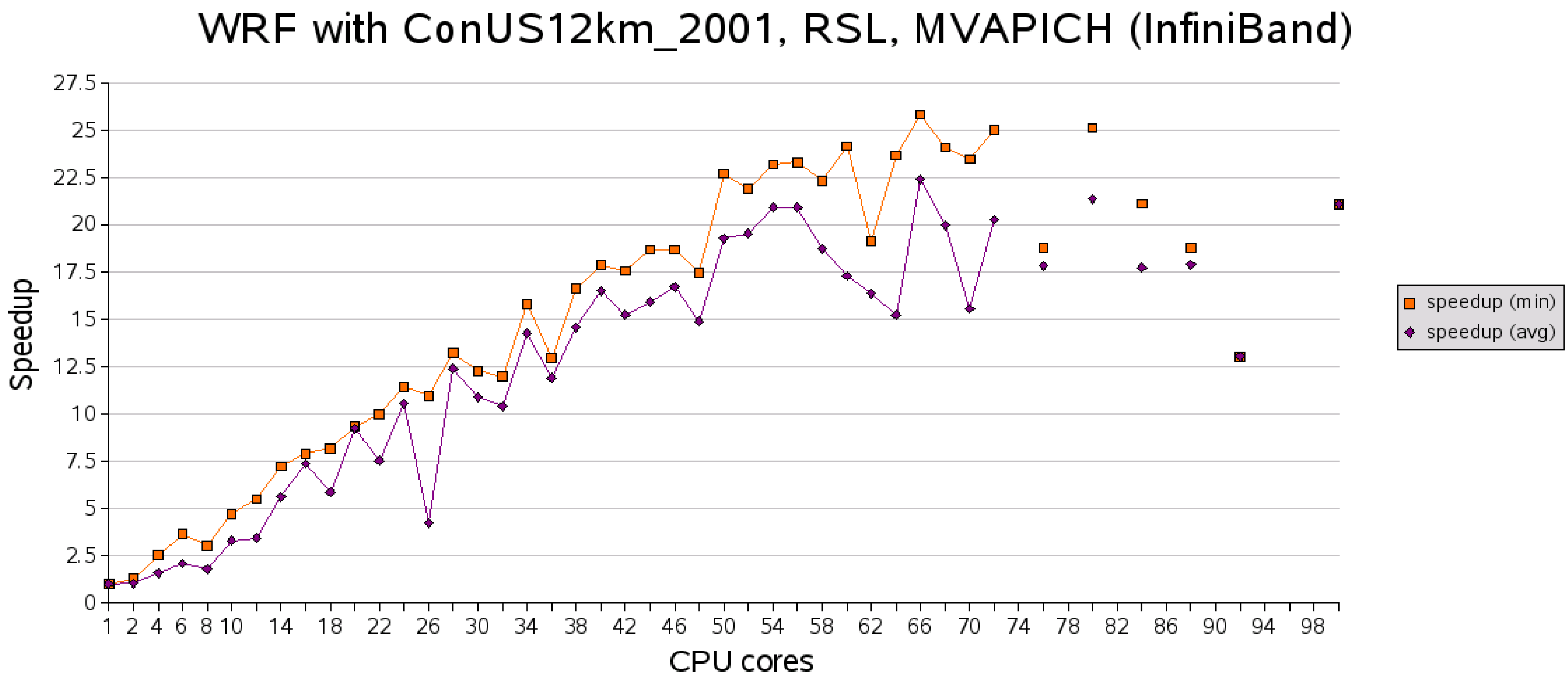
`http://wisent.d-grid.de`

WRF with ConUS12km_2001, RSL, MVAPICH (InfiniBand)



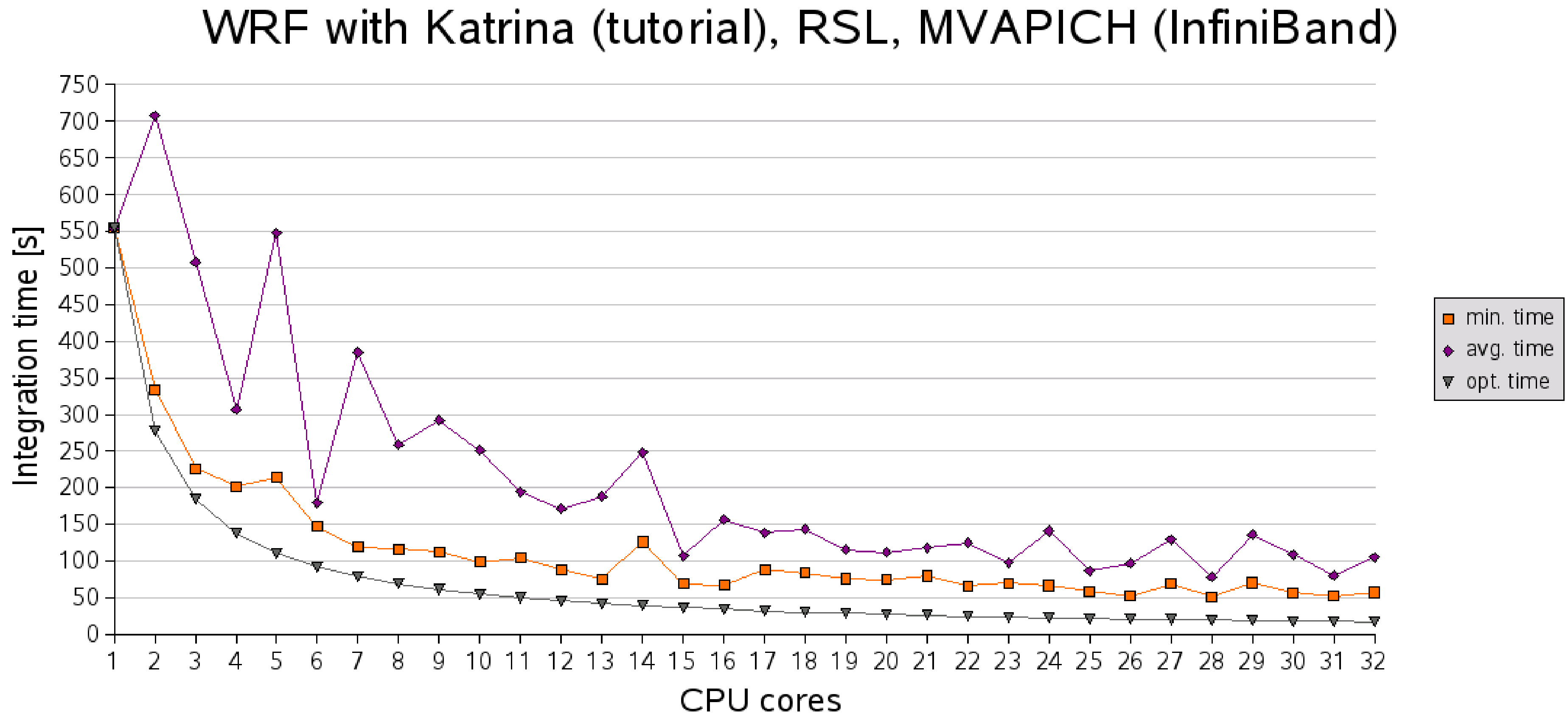
- **Standard benchmark case from NCAR**
- **Single domain, 48 hour, 12km res., 425x300x35 grid, time step 72s**
- **Benchmark period: 3 hours starting from end of hour 24 (restart)**

Hardware configuration: https://bi.offis.de/wisent/tiki-index.php?page_ref_id=91



- **Standard benchmark case from NCAR**
- **Single domain, 48 hour, 12km res., 425x300x35 grid, time step 72s**
- **Benchmark period: 3 hours starting from end of hour 24 (restart)**

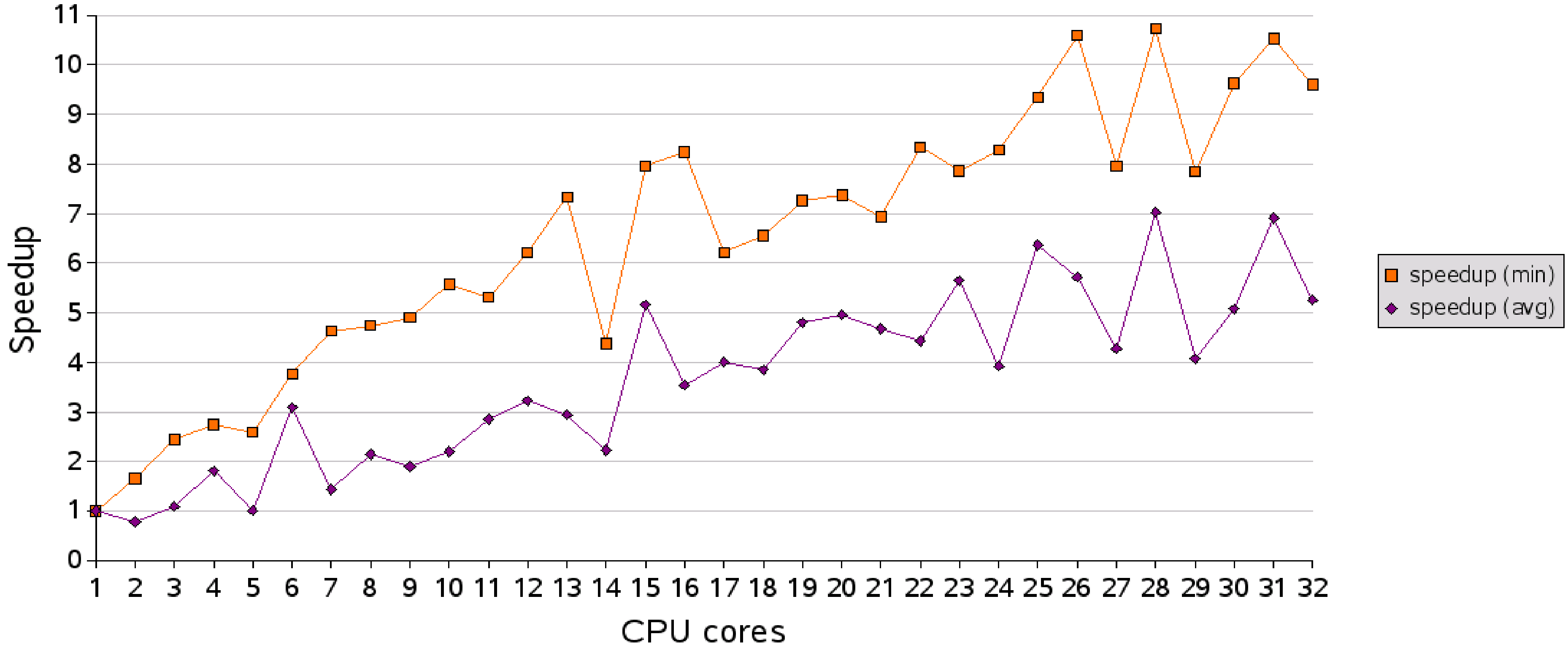
Hardware configuration: https://bi.offis.de/wisent/tiki-index.php?page_ref_id=91



- **Tutorial example from NCAR, hurricane Katrina simulation**
- **Single domain, 24 hours, 30 km res., 75x70x28 grid, time step 180s**

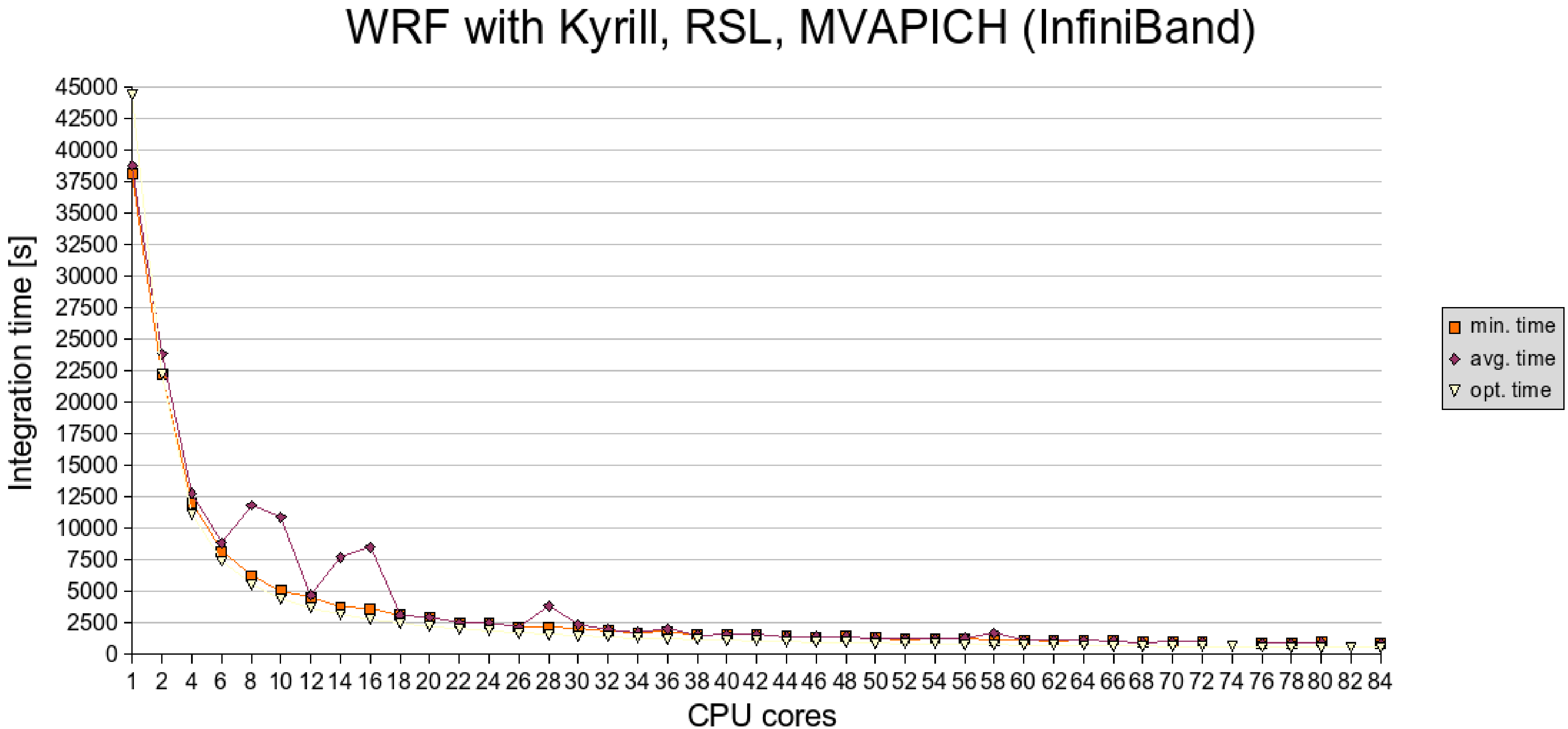
Hardware configuration: https://bi.offis.de/wisent/tiki-index.php?page_ref_id=91

WRF with Katrina (tutorial), RSL, MVAPICH (InfiniBand)



- **Tutorial example from NCAR, hurricane Katrina simulation**
- **Single domain, 24 hours, 30 km res., 75x70x28 grid, time step 180s**

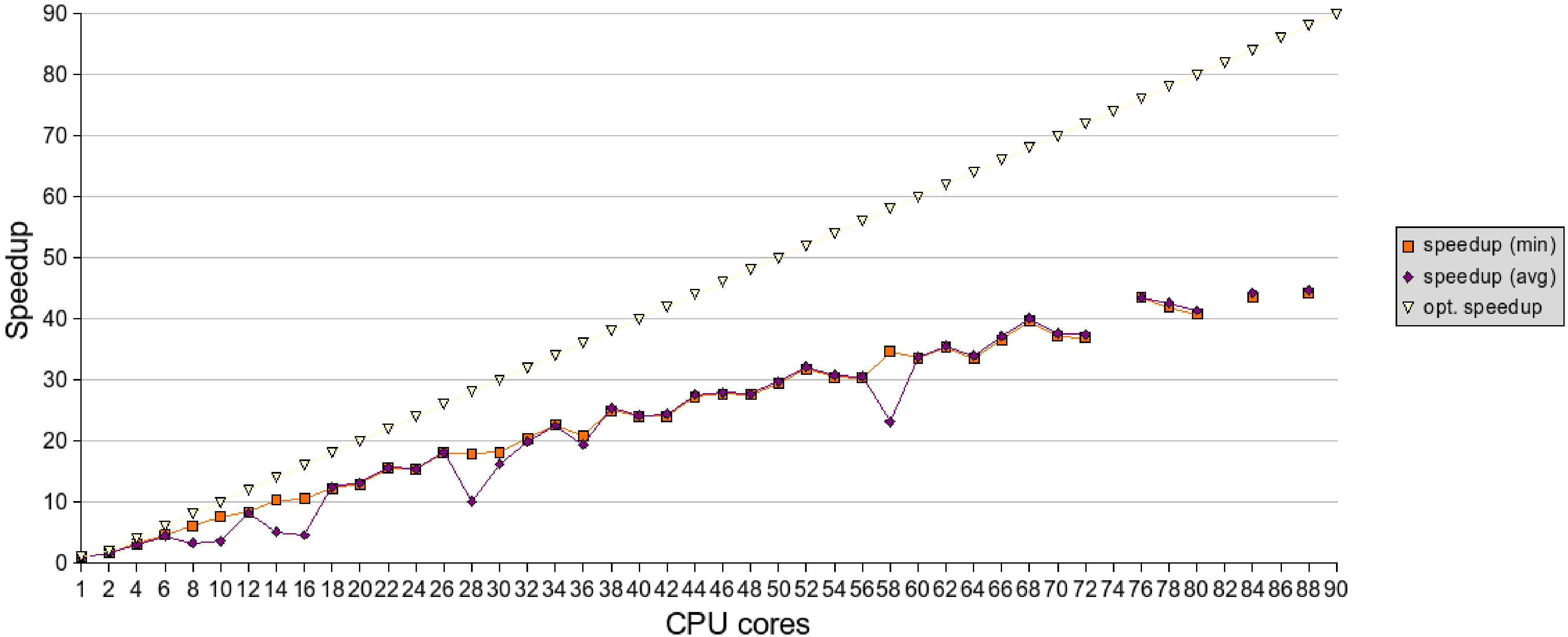
Hardware configuration: https://bi.offis.de/wisent/tiki-index.php?page_ref_id=91



- **Our own benchmark case, Kyrill storm, 24 hours**
- **Domain 1: 25 km res., 100x100x28 grid, time step 150s**
- **Domain 2: 5 km res., 261x281x28 grid, time step 30s**

Hardware configuration: https://bi.offis.de/wisent/tiki-index.php?page_ref_id=91

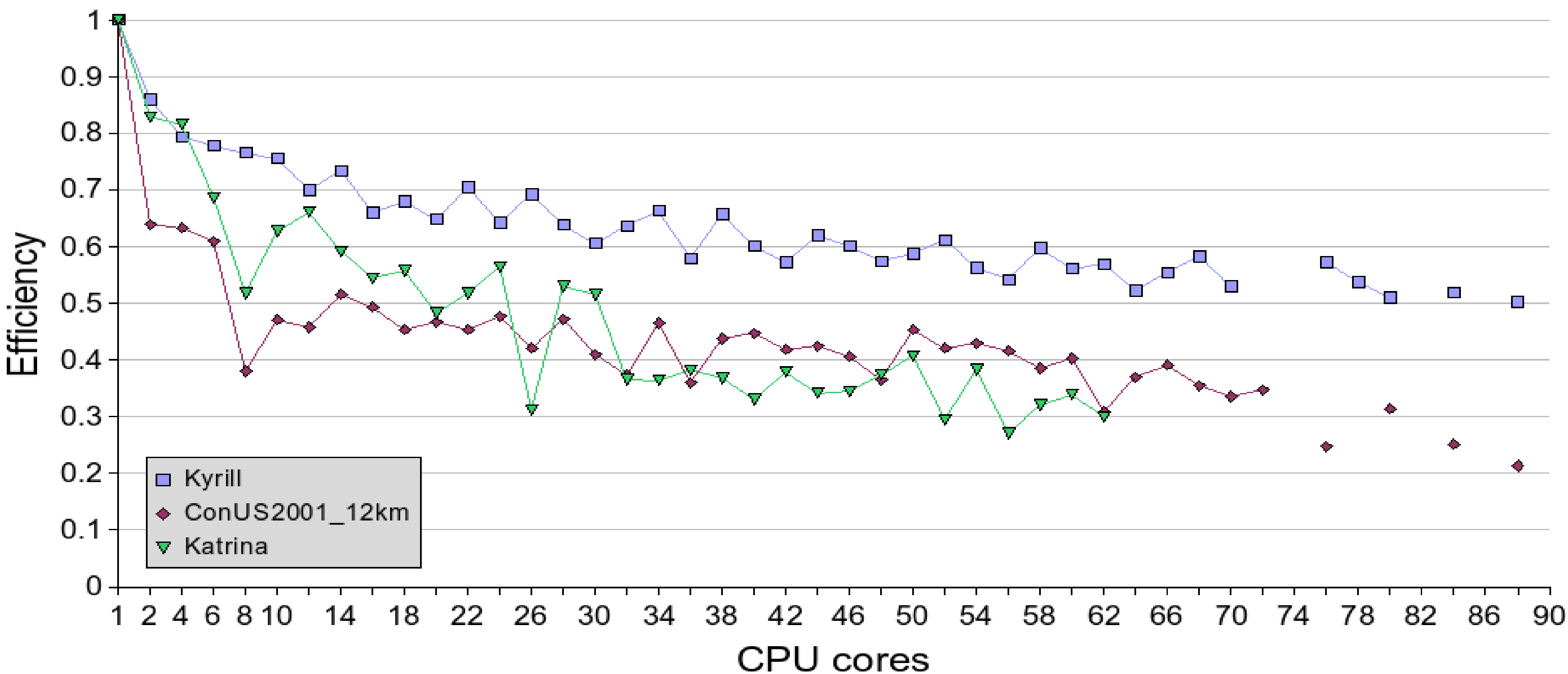
WRF with Kyrill, RSL, MVAPICH (InfiniBand)



- **Our own benchmark case, Kyrill storm, 24 hours**
- **Domain 1: 25 km res., 100x100x28 grid, time step 150s**
- **Domain 2: 5 km res., 261x281x28 grid, time step 30s**

Hardware configuration: https://bi.offis.de/wisent/tiki-index.php?page_ref_id=91

WRF Efficiency, RSL, MVAPICH (InfiniBand)



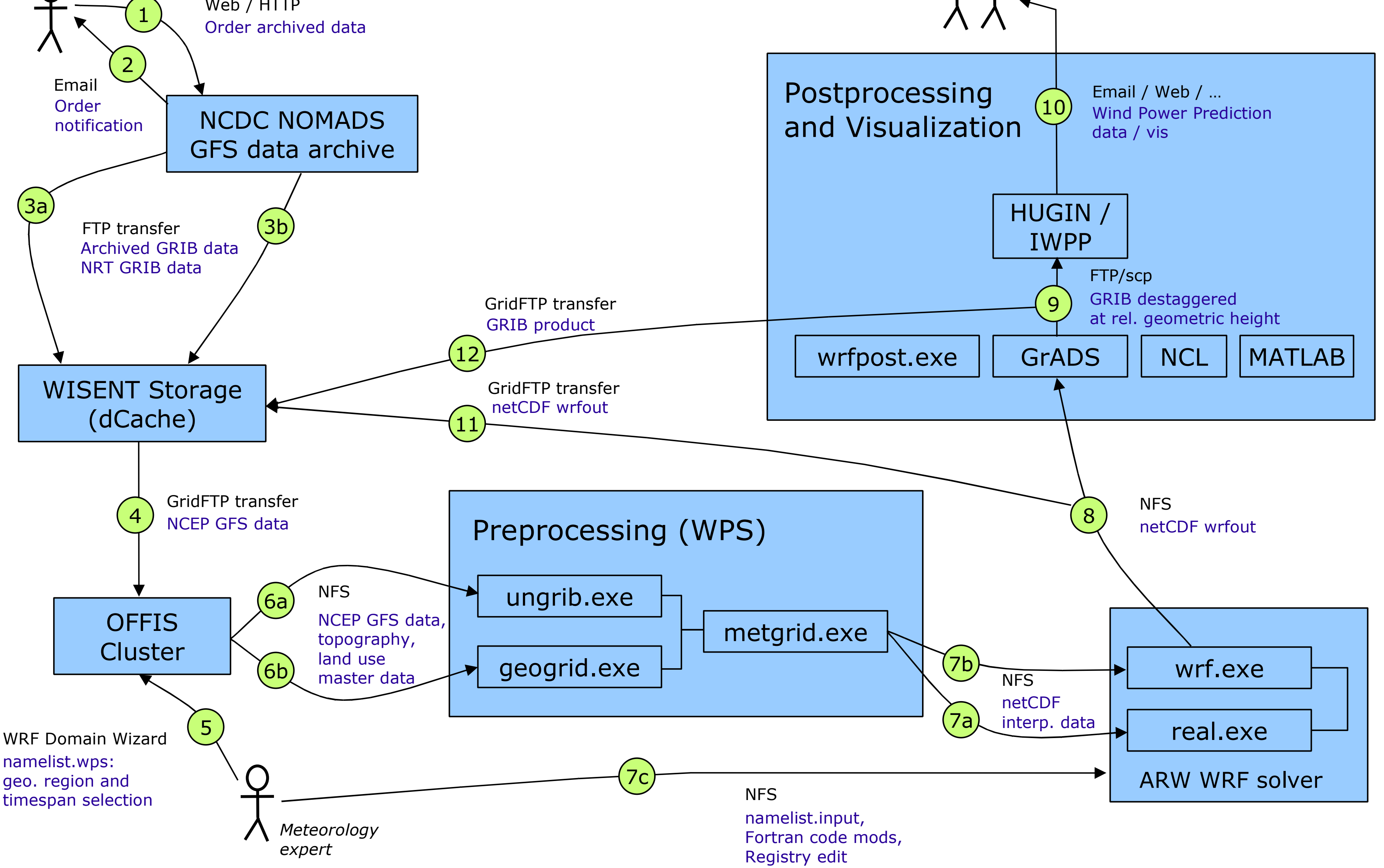
- **Kyrill:** **from 11 hours to 15 minutes using 90 CPU cores**
- **ConUS2001_12km:** **from 66 minutes to 2 minutes using 100 CPU cores**
- **Katrina:** **from 9 minutes to 1 minute using 32 CPU cores**

Hardware configuration: https://bi.offis.de/wisent/tiki-index.php?page_ref_id=91

Challenges for Grid computing

- **Usability**
- **Multi-user operation and access**
- **Software configuration managament**
- **Data management**

WRF Workflow



Usability

- **Users today struggle with software**
  - command line + Unix tools + visualization packages
  - multiple programs, from different vendors, not integrated well
  - the main interface is data files
  - configuration files, different syntaxes, symlinks, conventions...
  - several scripting languages
  - technical jargon that really *should not matter*
- **A uniform working environment is needed**
  - to hide away the unnecessary details
  - to lower entry barriers for new users
- **Pre-arranged workflows good for operational forecasting...**
  ...what about explorative research?

Multi-user operation and access

- **Everyone sets up their own model in `$HOME/somewhere`.**
- **Users don't share executables (~300 MB), data (~9.6 GB), custom tools.**

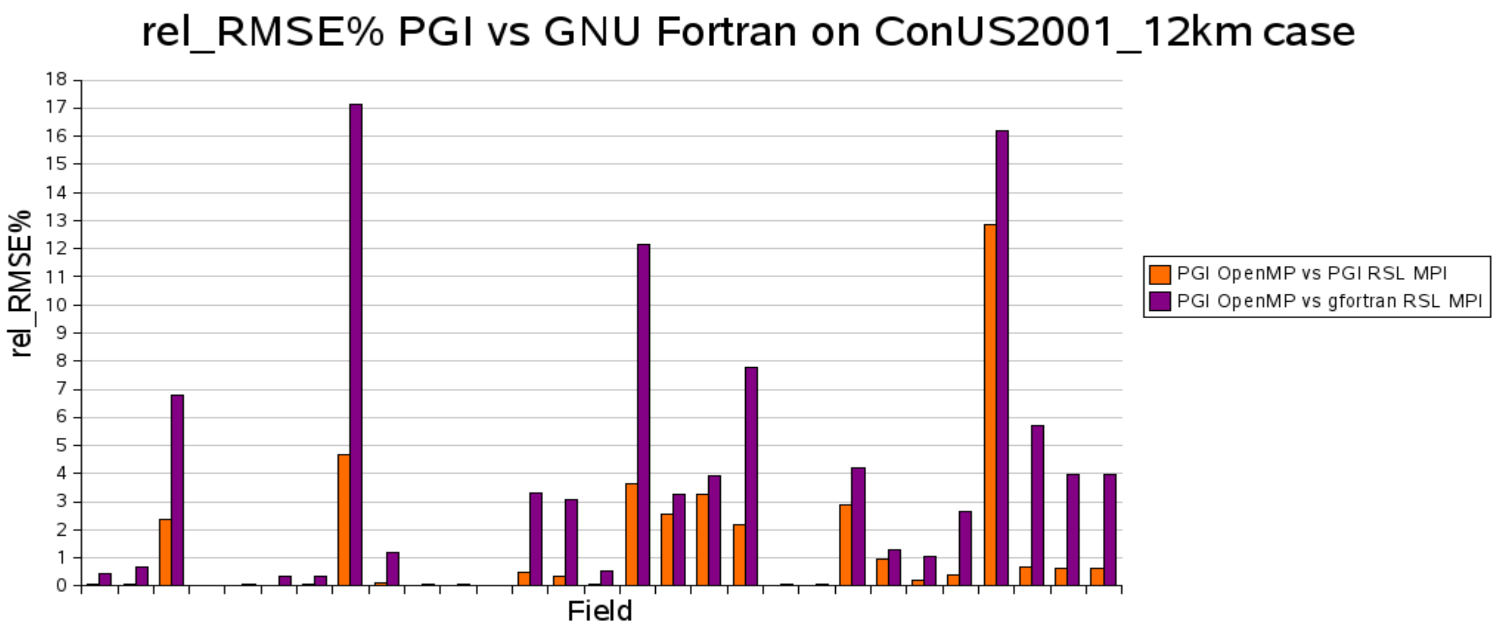
> ***Coordinated resource sharing and problem solving in dynamic, multi-institutional virtual organizations.***

- **There is very little coordination in the traditional approach.**
- **It does not scale up to multiple users and multiple Grid sites.**

Software configuration management

- **A full WRF "working environment" requires multiple software components.**
- **Deploying them manually once is a hassle...**
- **Doing it 10+ times (once per site) is really expensive.**
- **Sites differ in software configuration and availability.**

- **Sites synchronization and QA are needed afterwards...**



Data management

- **Input data: GRIB up to 3.2 GB/day.**
- **Need for reliable download, storage.**

- **A case study over 3 months → 600 GB of input+output data**
- **Afterwards, (some of) it has to be archived...**
  - where sufficient storage space is available
  - not necessarily where the model executes

- **Data management should be effortless...**
  - high-level interaction
  - fault tolerance
  - limited potential for mistakes

Summary

- **Grid Computing is a novel way to utilize Internet for the benefit of meteorological applications such as NWP models:**
  - Scalability and availability
  - Performance and result comparisons

- **WRFv2 – in the MPI over InfiniBand variant – scales well even for small cases. More results with different configs will follow.**

- **Grid Computing must address issues beyond computational power:**
  - Usability
  - Multi-user operation and access
  - Software configuration management
  - Data management