Enabling the use of e-Infrastructures with Microsoft HPC and the Matlab Distributed compute server

Shah Zeeshan Ali, Erwin Laure. PDC – HPC

Mikael Sterner. Microsystem Technology Lab KTH

www.eu-egee.org

e-infrastructure

Agenda

- **Agenda**
  - Overview
  - Solution
  - Feedback from Users
  - Future improvements
  - Conclusion
  - Q&A

Overview

- **Scientists with Small scale problems**
  - Could benefit with e-infrastructures but have little experiences with shell based Linux environment, Instead
  - They often use windows platform and higher level packages like MATLAB, COMSOL, etc..
- **Microsoft and MathWorks**
  - Integration of MSHPC and MATLAB distributed server is a good use-case to enable e-infrastructures for those scientists.

Overview contd.

- **MathWorks**
  - MATLAB Desktop with Parallel Toolbox.
    - Parallel Computing Toolbox™ lets you solve computationally and data-intensive problems using MATLAB® and Simulink® on multi core and multiprocessor computers. *
  - MATLAB Distributed Computing Server (MDCS).
    - You can use the toolbox to execute applications on a single multi core or multiprocessor desktop. Without changing the code, you can run the same application on a computer cluster or a grid or cloud computing service with MDCS. *
    - It includes a basic scheduler and directly supports Microsoft Windows HPC Server 2008 and other schedulers. **

* *http://www.mathworks.com/products/parallel-computing/*

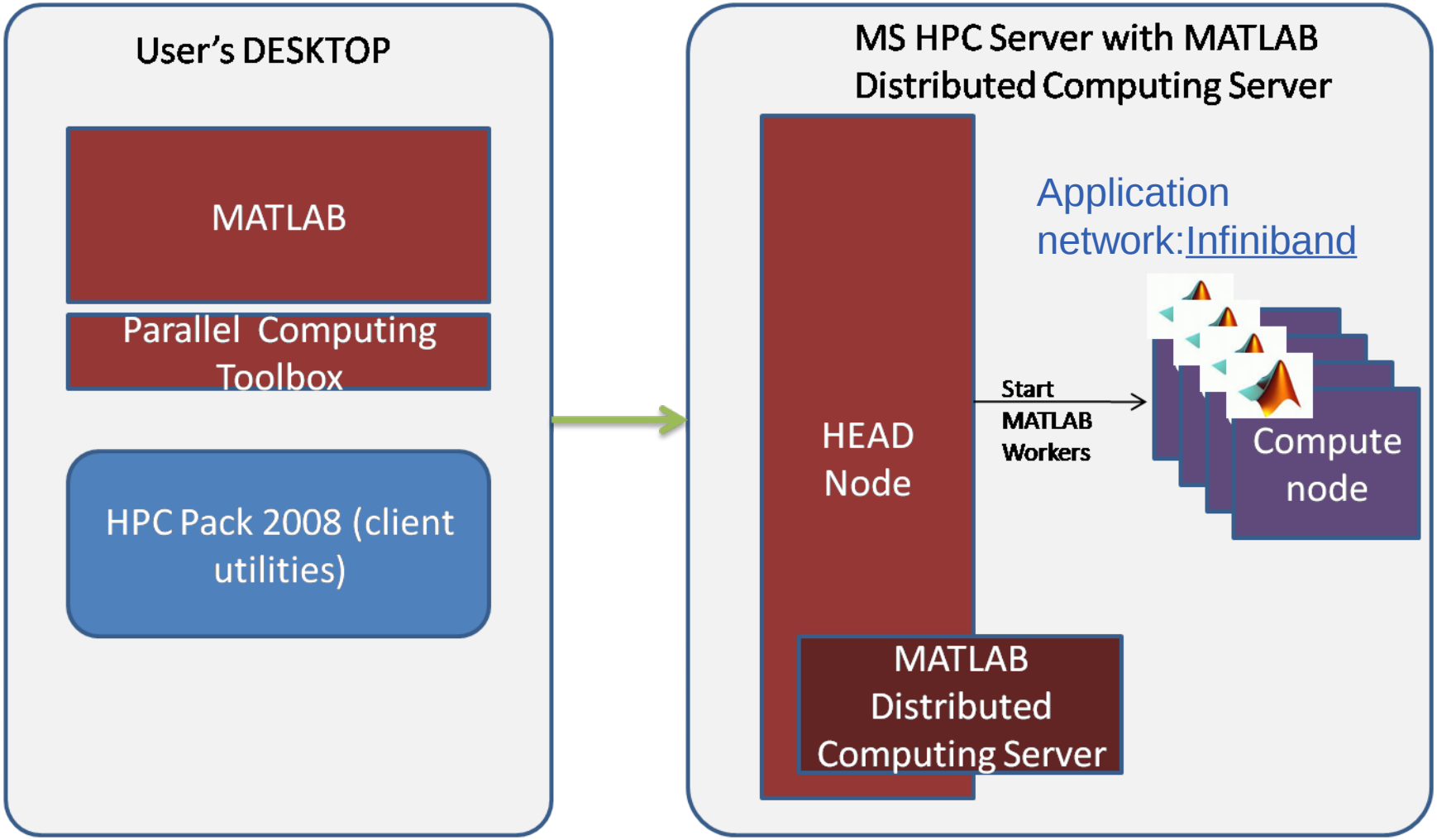
** http://www.mathworks.com/products/distriben/

Overview contd.

- **Microsoft ***
  - HPC Server 2008
    - Windows 64-bit Server platform with out-of-the-box functionality to improve the productivity, and reduce the complexity, of your HPC environment.
    - A cluster consists of :
      - *Head Node*
        - The single point of management and job scheduling for the cluster.
      - *Compute Node*
        - Carries out the computational tasks assigned to it by the job scheduler.
      - *Job Scheduler*
        - Queues jobs and their associated tasks. It allocates resources to these jobs, initiates the tasks on the compute nodes; and monitors the status of jobs, tasks, and compute nodes.

* *http://www.microsoft.com/hpc/en/us/default.aspx*

Solution contd.



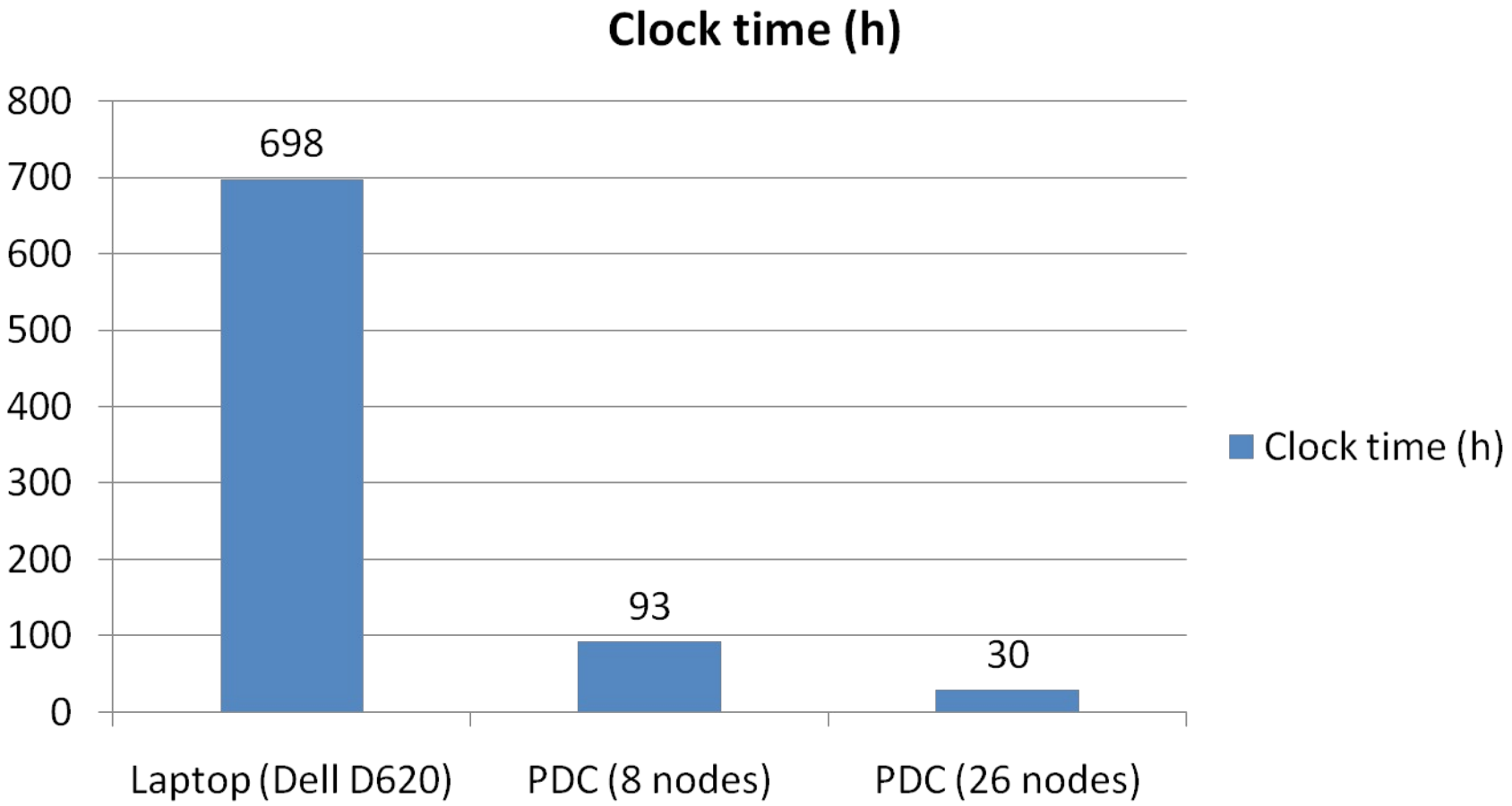
Feedback from User-1

- **Application**
  - Multiphysics simulation of Microsystems.
- **Related Areas**
  - Optics, solid mechanics, Electromagnetics, Chemistry, Fluid Dynamics, Heat transfer.
- **Pre-HPC way**
  - Simulations on laptops or on self-managed standalone multi-processor machines.
- **Getting started with HPC**
  - Very easy to get started in MATLAB, once all necessary software updates were in place. Only trouble was with caching of credentials.

Feedback from User-2

- **Client Installation**
  - Update to XP SP3, install MS HPC Pack (client utilities), MATLAB Parallel toolbox, credentials for HPC client.
- **Gain**
  - Not needing to purchase and manage the hardware and HPC server ourselves.
  - Extensible computing power.
- **Desired Extensions**
  - More pre-installed simulation software.
  - COMSOL Multiphysics is the most important one.

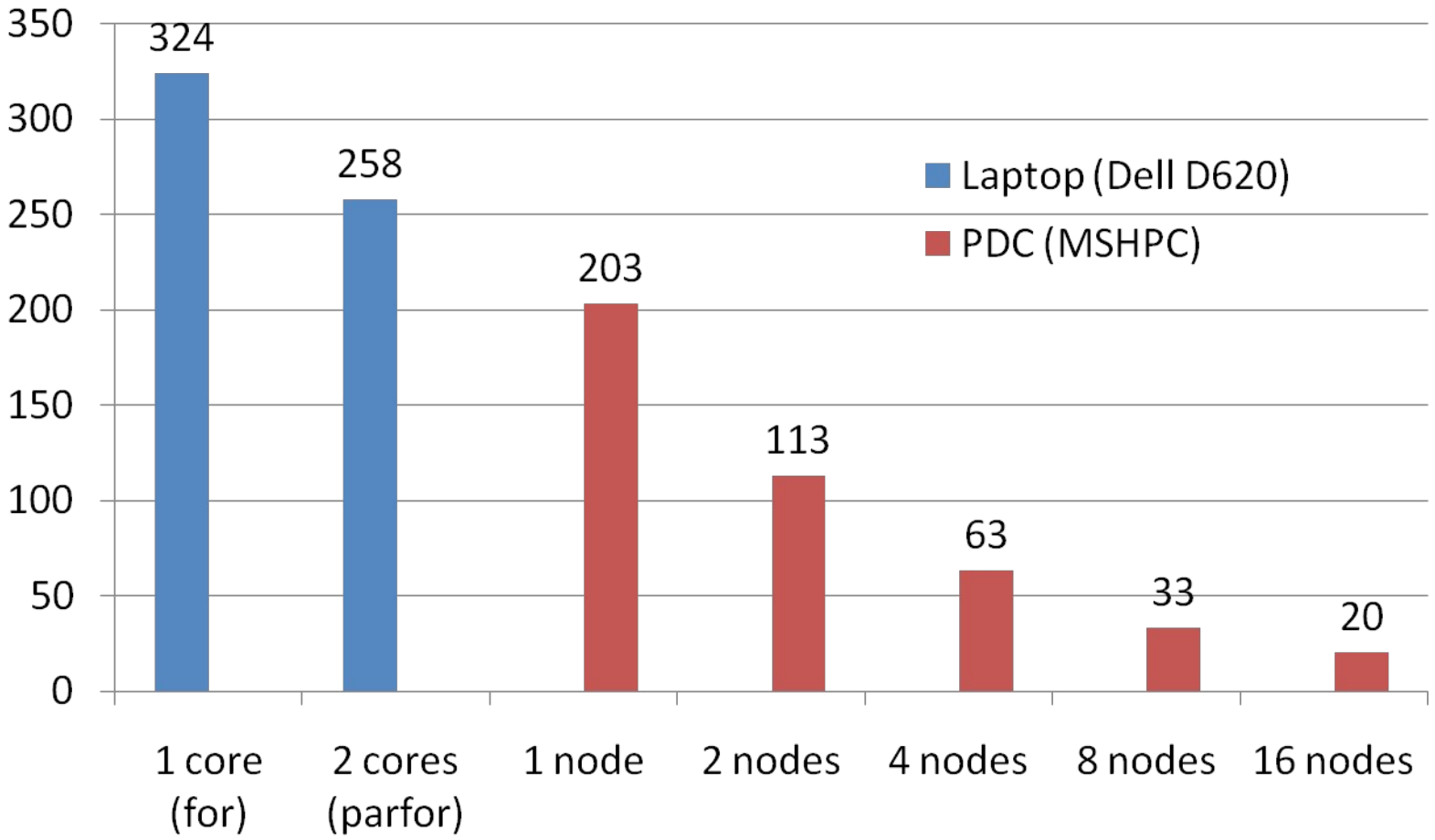
CAMFR performance



Parameter sweep of CAMFR simulation on optical structure

PDC-HPC : Zeeshan Ali Shah zashah@pdc.kth.se

MATLAB performance



Runtime in seconds of simple MATLAB benchmark

PDC-HPC : Zeeshan Ali Shah zashah@pdc.kth.se

Future development

- Expand the prototype locally by increasing number of resources
- Integrating with EGEE infrastructure
- More applications to support like Comsol , etc..

Conclusion

- **Good opportunities for small scale applications to benefit from e-science resources.**
- **Integration of MATLAB and MSHPC enabled researchers to easily use the Parallel resources from there Desktop**
- **Porting of new applications on HPC cluster opened new ways to improve research**
- ***In brief we should lower the technology barrier so that Users can focus on their research problems directly.***

Thanks

Questions / Comments ?