Collaborative e-Science Architecture for Reaction Kinetics Research Community

Tran Vu Pham[1], Dr. Lydia MS Lau[1], Prof. Peter M Dew[2] & Prof. Michael J Pilling[3]

[1] School of Computing, University of Leeds, Leeds, UK

[2] Informatics Institute, School of Computing, University of Leeds, Leeds UK

[3] School of Chemistry, University of Leeds, Leeds, UK

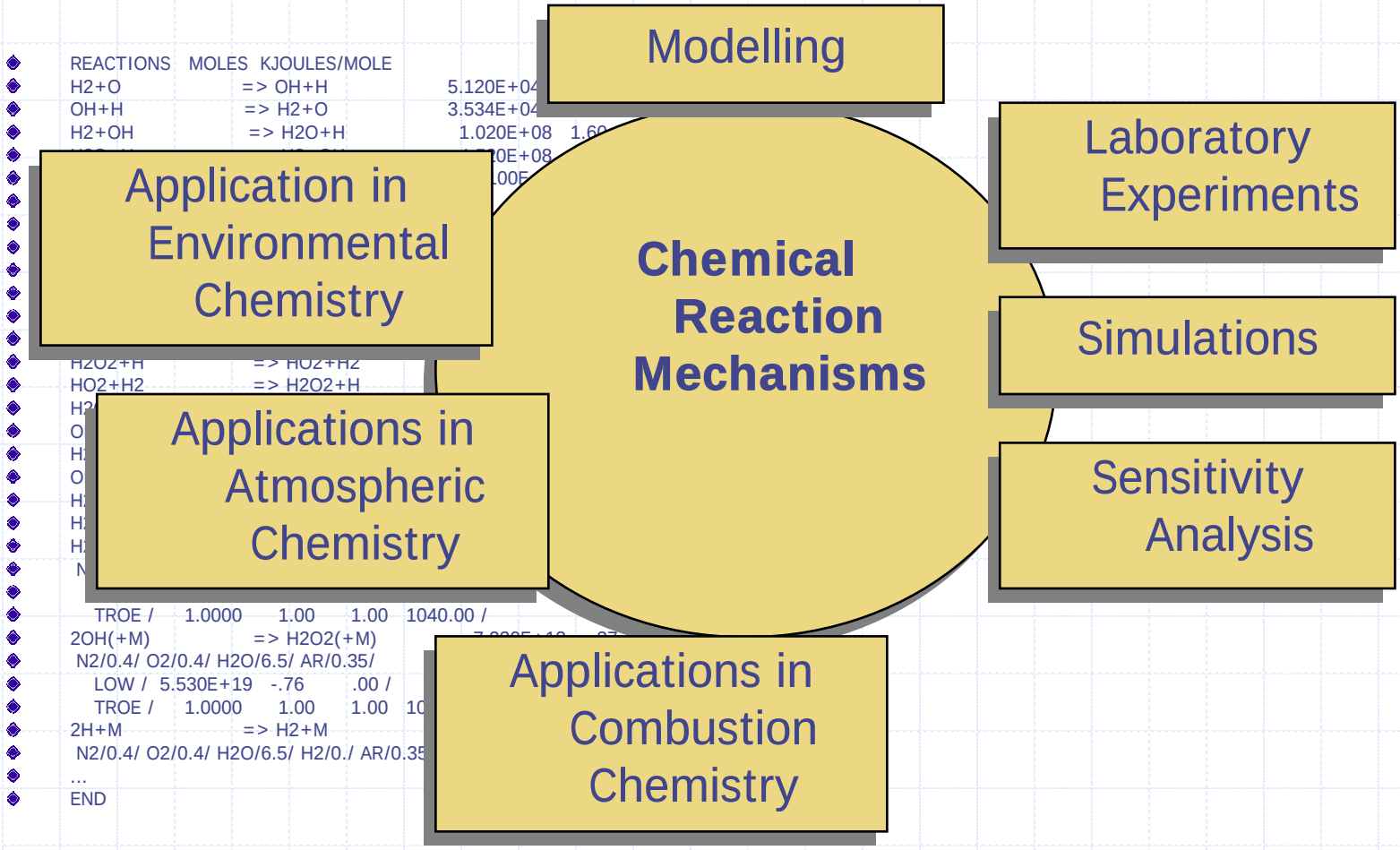
Content

- What do the scientists do?
  - A Case study: The Reaction Kinetics research community
- Current popular architectures
  - Web-based, Grid-based
- The Collaborative e-Science Architecture
- Prototype
- Experiment & evaluation
- Conclusion and future work

Case Study: The Reaction Kinetics Research Community

- Reaction Kinetics is a multidisciplinary research subject
- The research community spans across Combustion Chemistry, Atmospheric and Environmental studies
- The research community is highly distributed, consisting of members from different part of the world

Case Study: Main Research Activities



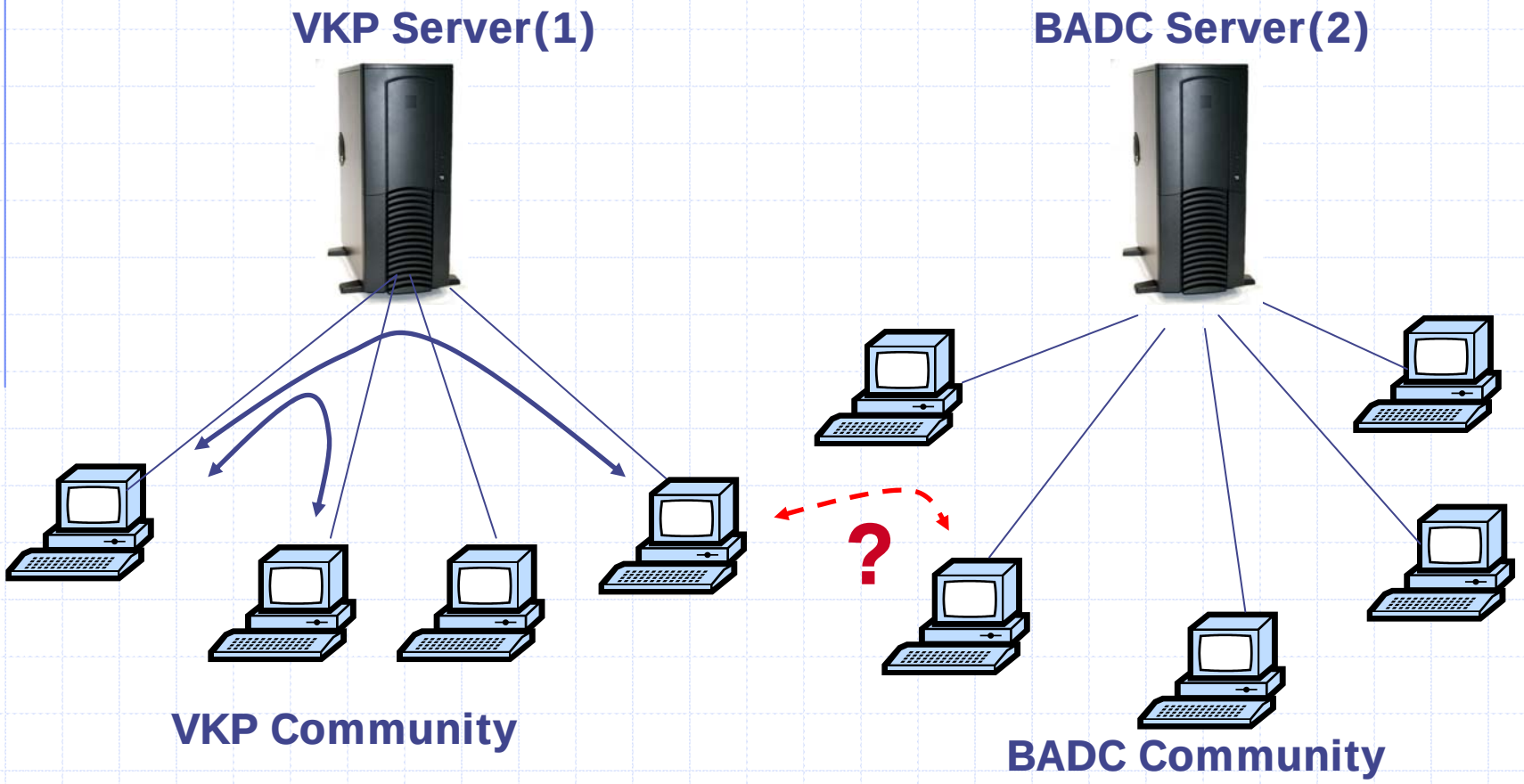
Case Study: User Requirements

- Facilities to allow scientists who are working on the same or similar research activities to dynamically form ad hoc working groups
- Efficient support for timely collaborations within and across working groups in the community for sharing expert knowledge, day-to-day working data, such as experimental data, chemical reaction mechanisms and related input data for reaction modelling.
- Easy access to computational intensive resources for time and resource consuming simulations and analyses and for storage of large amount of experimental data.

Case Study: Challenges

- These requirements are related to two challenging issues in e-Science:
  - How to provide the scientists with an integrated collaborative environment
  - How to provide the scientists easy access to computationally intensive resources from a desktop computer

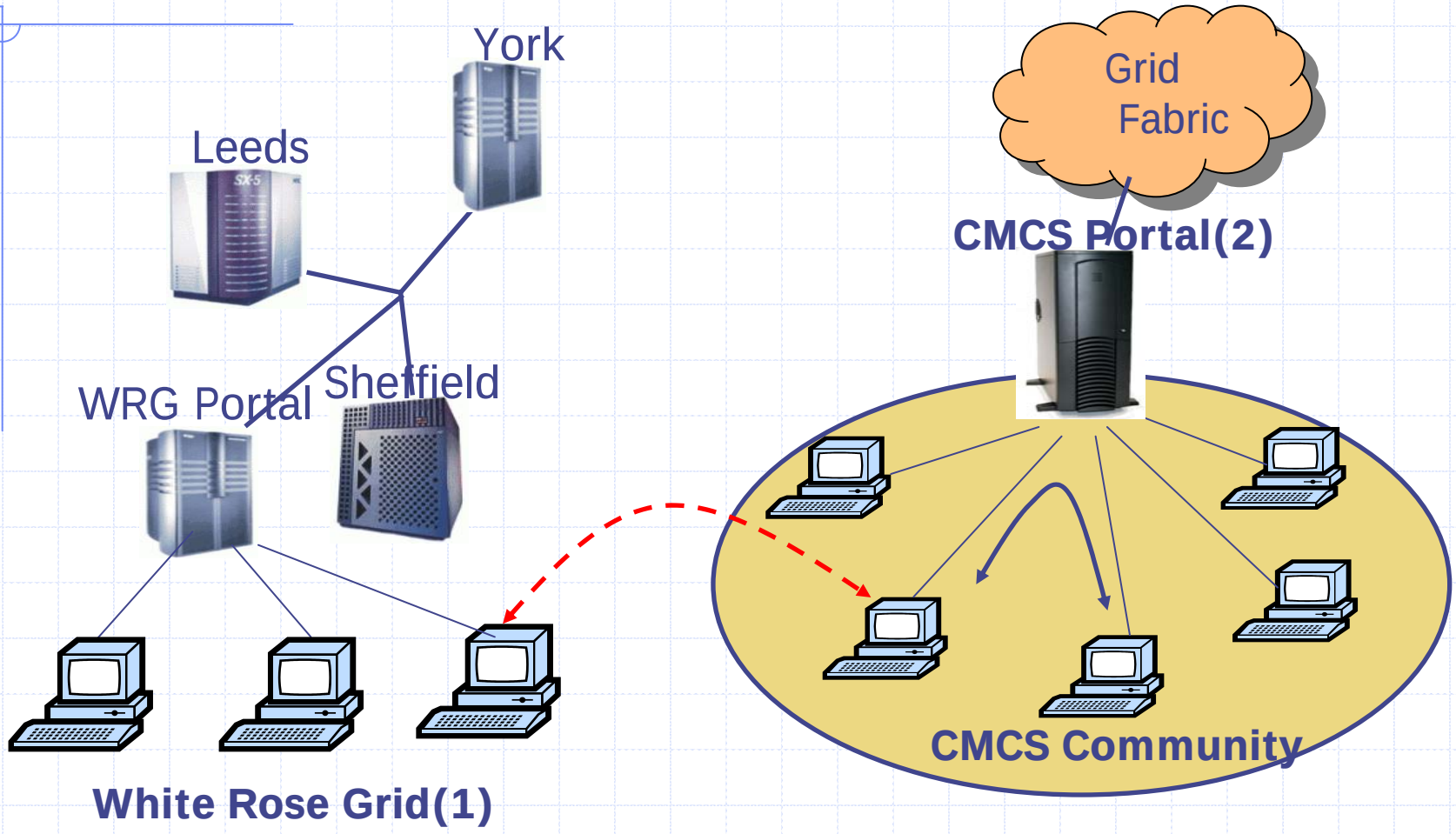
Current Popular Architecture: Web-Based



(1) The Virtual Knowledge Park, http://vkp.leeds.ac.uk

(2) British Atmospheric Data Centre, http://badc.nerc.ac.uk

Current Popular Architecture: Grid-based



(1) The White Rose Grid, http://www.wrgrid.org.uk

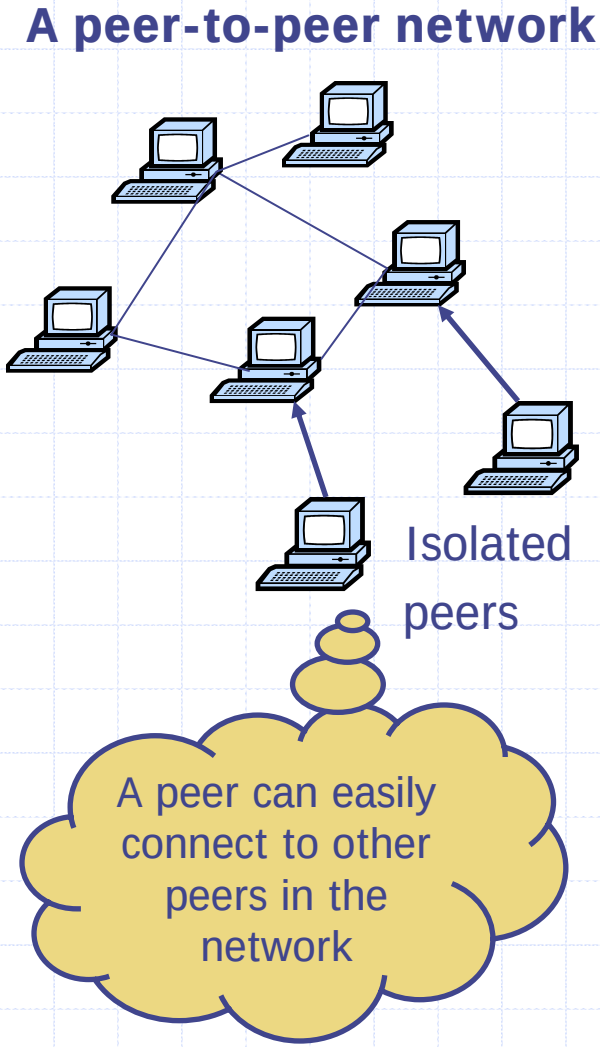
(2) Collaboratory for Multi-Scale Chemical Science, http://cmcs.org

Current Popular Architecture: Summary

- Grid architecture is good for dealing with the need for computational resources and data storage
- Collaborations amongst users are based on centralised web-based architecture. This approach has a few limitations:
  - Direct collaborations within such a community are limited (it is possible to use email, but this method is not suitable for sharing large data files)
  - Across community collaborations are limited
  - It is hard to form ad hoc working groups, which consist of members from different communities

Potential of Peer-to-Peer Computing

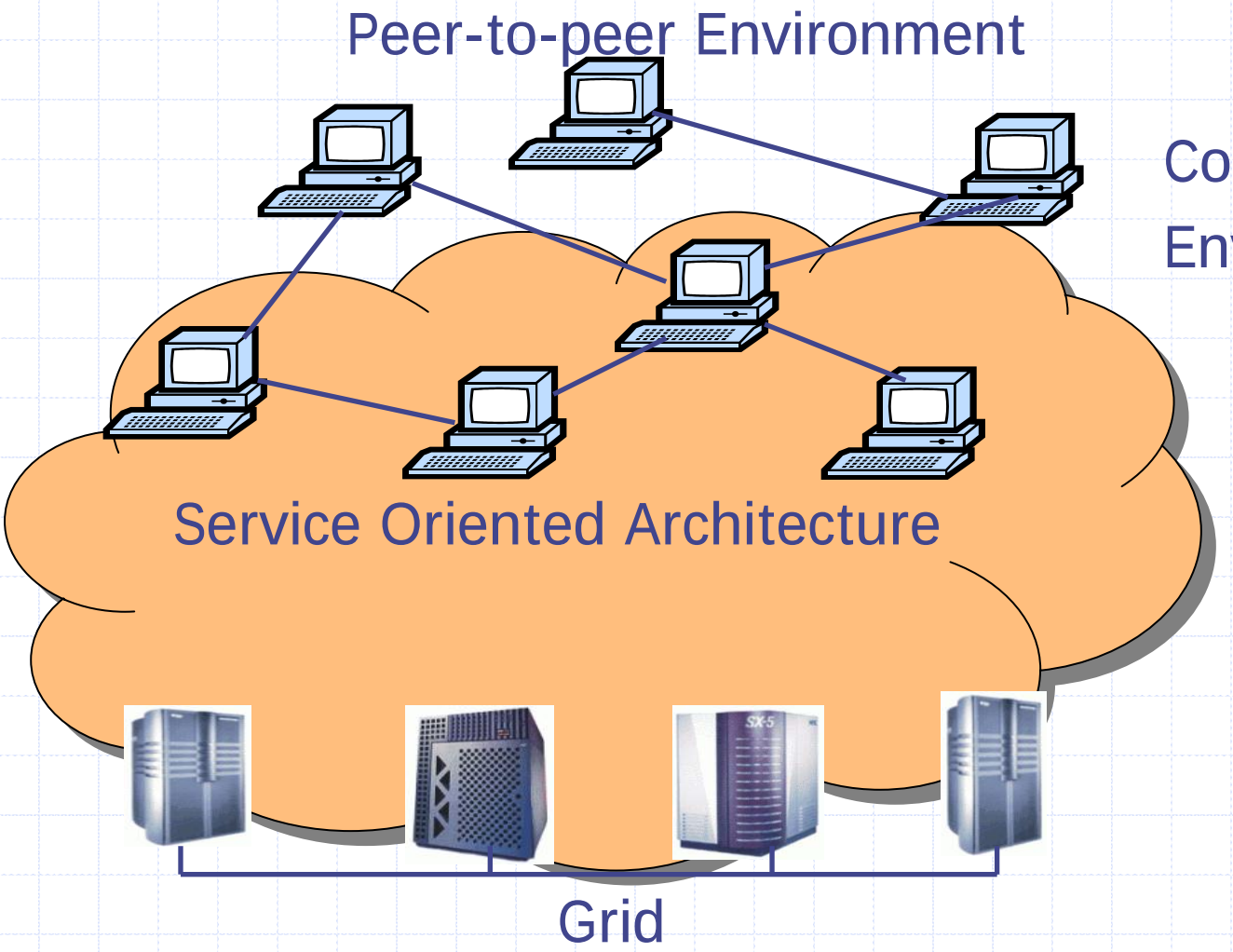
- Peer-to-peer Computing
  - Direct communication of peer users
  - Bring end users closer to their communities and shared resources
  - Sense of privacy and ownership over shared resources
  - Ad hoc group can be formed easily to support collaborative work



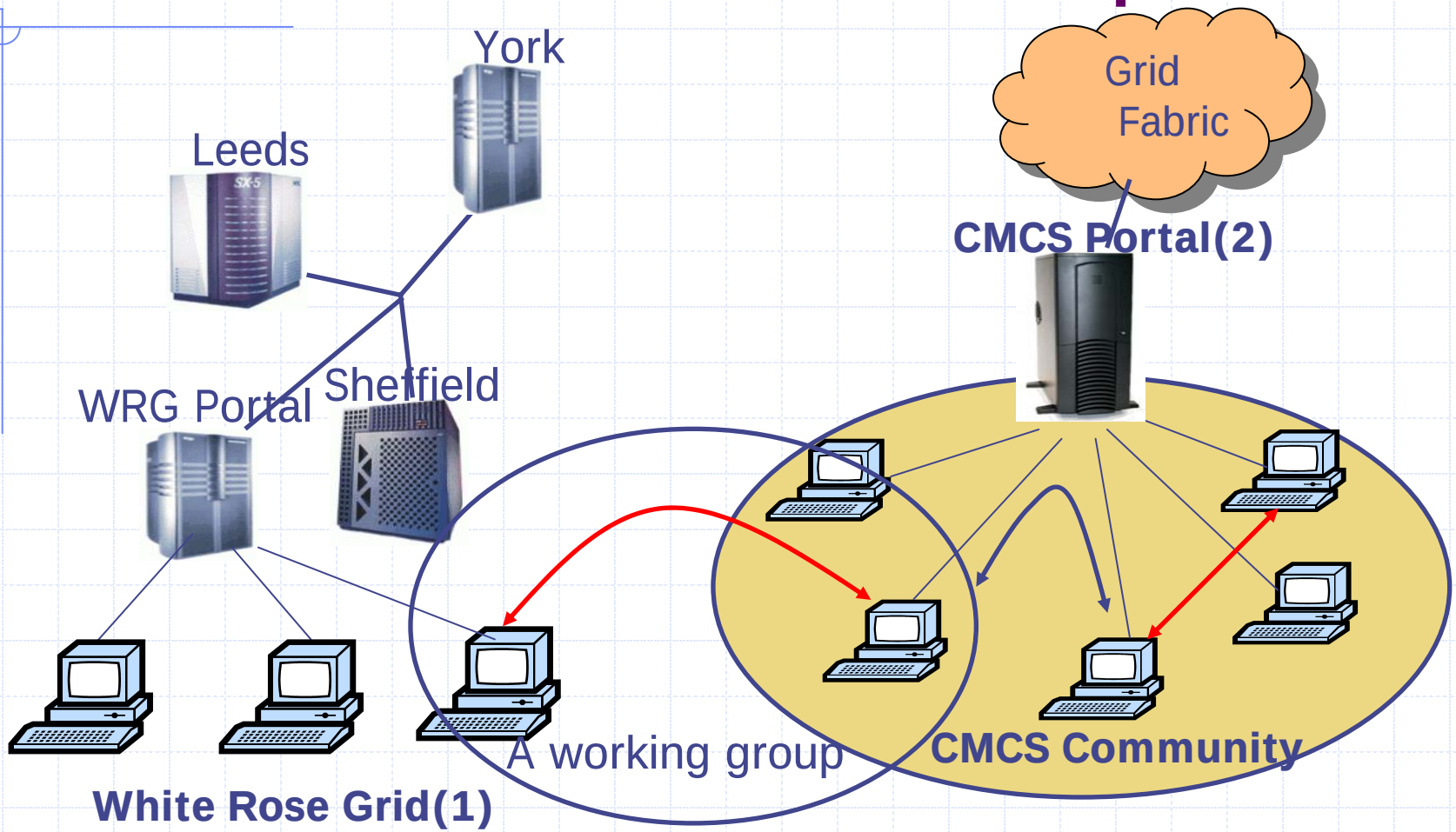
The Collaborative e-Science Architecture (CeSA)

- Goals:
  - To provide an integrated collaborative environment to better support users in distributed communities
  - To provide scientists with easy access to computation intensive resources large-scale storage

The Collaborative e-Science Architecture (CeSA)



The Collaborative e-Science Architecture: An Example



(1) The White Rose Grid, http://www.wrgrid.org.uk

(2) Collaboratory for Multi-Scale Chemical Science, http://cmcs.org

The Collaborative e-Science Architecture (CeSA)

- Functionalities of peer-to-peer environment:
  - An integrated environment for user lightweight collaborations: chatting, file sharing, discovery of shared resources, etc
  - Tools for users to form virtual working groups
  - User interfaces for executing services from grids
  - Publication and discovery of services from the grids
- Grids are providers of computation and data intensive resources (e.g. large datasets)
- Management of user community in peer-to-peer environment is separated from the management within grids

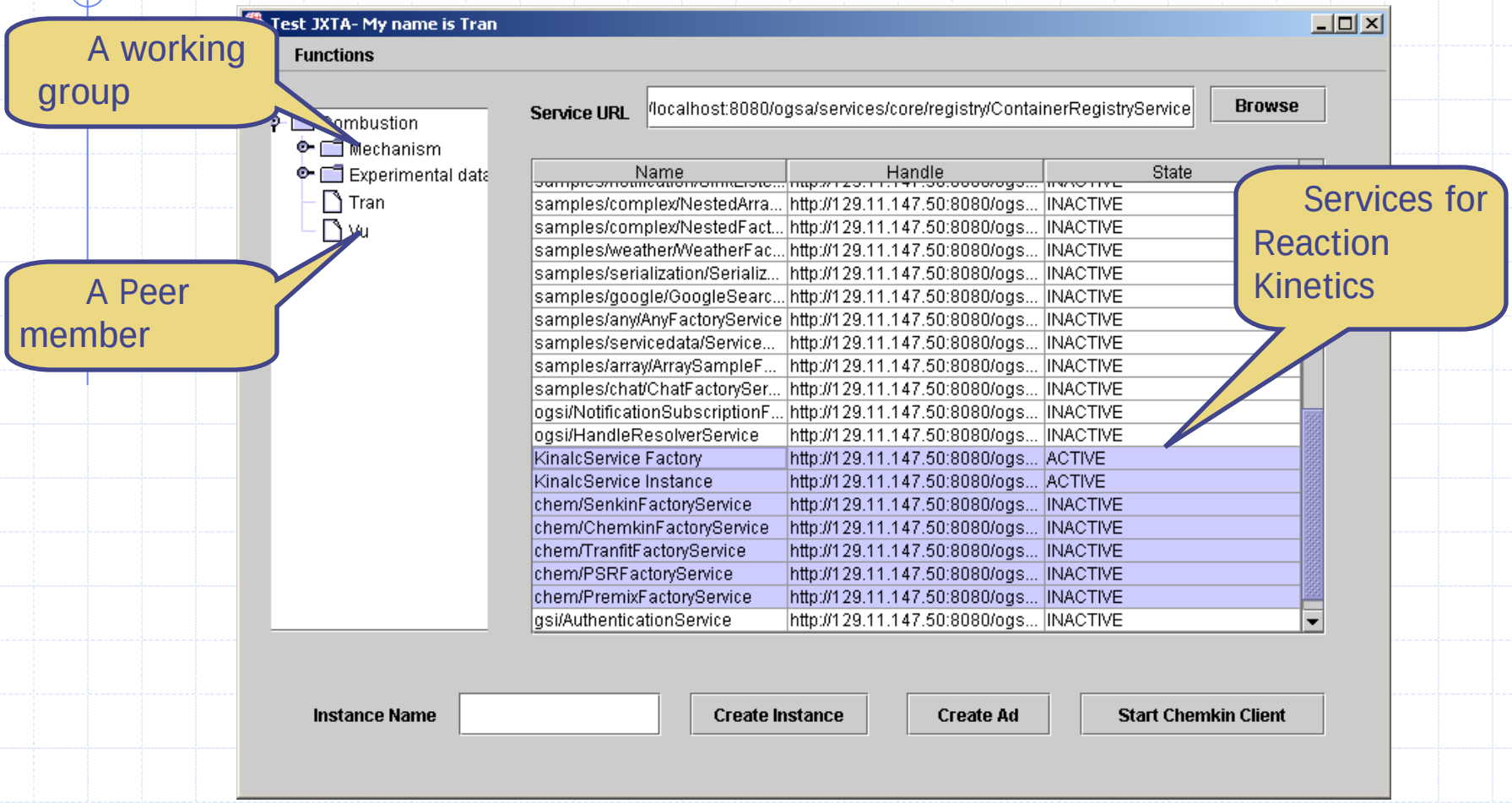
Prototype: An Instance of the CeSA

- To have an insight into the technical challenges as well as to test the applicability of the architecture for the user community
- Using requirements from the Case Study of the reaction kinetics research community
- Technologies involved:
  - JXTA was used to build the peer-to-peer application
  - Globus Toolkit version 3.0.2 (GT3) was used to develop services
  - Java as the programming language

Prototype: Steps

- Wrapping a few command-line programs for simulations and analyses of reaction mechanisms into Grid Services, then, deploy them in GT3 service container
- Developing the peer-to-peer application using JXTA, which also included a Grid Service client to execute Grid Services published from grids

Prototype: An Example Screen Shot



Browsing service registry from a peer-to-peer application

Experiment and Evaluation: Objectives

- To evaluate the effect of using peer-to-peer environment provided by the CeSA in a realistic user environment
- To assess how users can benefits from the access to remote simulations and analyses provided by grids via Grid Services
- To capture user general attitudes to the new collaborative infrastructure

Experiment and Evaluation: Method & Process

- Data collection method:
  - Questionnaire
  - Participants were provided with mixture of closed and open questions
  - A collaborative scenario was also provided
- The experiment process
  - Three scientists involved in the experiment at the same time
  - They used the prototype as guided by the scenario to collaborate with each other
  - Their feedback was recorded in the questionnaire

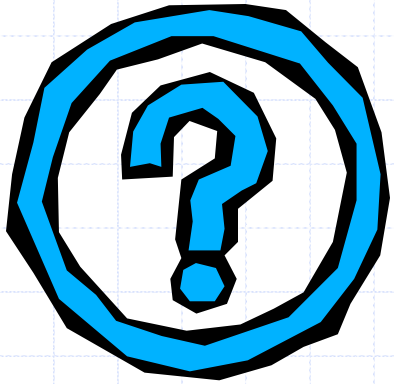
Experiment and Evaluation: The Results

- Generally positive
- Participants expressed their interests on using the prototype system, here is some feedback:
  - "A fully working system would benefit the atmospheric chemistry group provided it was widely accepted by the whole community"
  - "I think that our group would certainly use such a system if it proved to be the way forward in e-Science (which I feel it is) and the community embraced the use of such a system"
- However, there were also some worries about security

Conclusion & Future Work

- The positive result has shown the potential of the CeSA for collaborations in a scientific community
- The CeSA is being specified in more detail, especially the unified interface for a wider range of services to be used will the peer-to-peer application
- Experiment with a wider user community is also needed

Questions?



Email: tranp@comp.leeds.ac.uk