UNIVERSITY OF LONDON
IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE

EXAMINATIONS 2004

MEng Honours Degree in Information Systems Engineering Part IV
MSci Honours Degree in Mathematics and Computer Science Part IV
MEng Honours Degrees in Computing Part IV
MSc in Advanced Computing
for Internal Students of the Imperial College of Science, Technology and Medicine

This paper is also taken for the relevant examinations for the Associateship of the City and Guilds of London Institute
This paper is also taken for the relevant examinations for the Associateship of the Royal College of Science

PAPER C428=I4.52

GRID COMPUTING

Thursday 6 May 2004, 10:00
Duration: 120 minutes

Answer THREE questions

Paper contains 4 questions
Calculators not required

***Section A** (Use a separate answer book for this Section)*

1a In the context of e-Science applications built using a Service Oriented Architecture explain what is meant by:

i) Workflow

ii) Ontology

In what ways can workflows and ontologies be used to support collaborative and communal activities in an e-Science community?

b A software vendor supplies two components: a modeling package, MP, and a visualisation package, VP. There are two customers: Customer A and Customer B. The perceived values of these packages to the two customers are given in the following table.

Value to User	MP	VP
Customer A	6	5
Customer B	8	2

i) What is the maximum price the vendor wishing to maximize revenue could set for each package in an open market if they are both sold separately?

ii) Explain why it is sometimes in the interest of vendors to *bundle* goods, i.e. sell them as a package. Given the perceived values given above could the software vendor gain greater revenue from his two customers by bundling the software packages? If so what is the maximum price he could set for the bundle?

Both parts carry equal marks.

2a What features of a component-based software technology make it particularly suitable for constructing Grid applications?

b List *four* factors that can influence the performance of an e-Science application executing in a Grid environment. For each factor show how a Grid application construction and execution infrastructure can be used to optimise the execution time of the application while remaining transparent to the user.

c Two parallel machines, Machine A and Machine B, are available to execute a fixed-size modeling application. The execution times, t, of the application on the two machines are given by the following formulae, where p is the number of processors used.

Machine A: $t = 200 - 0.4p$ *for* $0 <= p <= 499$

Machine B: $t = 400 - 1.0p$ *for* $0 <= p <= 399$

i) For what range of processor numbers, p, is it more efficient to use Machine A on this problem than Machine B?

ii) If it was decided to charge for the use of these machines at a rate relative to the number of processors used, explain how similar equations could be developed to guide the choice of which machine and how many processors to use.

All parts carry equal marks.

Section B *(Use a separate answer book for this Section)*

3a The Condor system represents an example of a working Grid infrastructure.

i) Describe the *architecture* of the Condor system describing the difference between submit, execute and central manager nodes. Different execution environments are supported in Condor through a 'universe'. Describe the differences between **two** of the Vanilla, Standard and Globus universes.

ii) Condor uses an information system based around ClassAds. Describe the structure of the ClassAd, their contents and explain how the information is matched and ranked.

b The Grid is built from resources owned by different organizations exposed through middleware. To ensure an 'end-to-end' quality of service it is necessary to negotiate 'service level agreements' (SLA) with each resource.

i) WS-Agreement (WSA) is one proposed mechanism to negotiating access to resources and defining an SLA. Describe the architecture and key principles behind WSA. Detailed description of the XML is **not** required.

ii) The exploitation of resources within a Grid is usually co-ordinated by a super-scheduler or resource broker. Describe the activities of the super-scheduler with respect to its stakeholders - the user, the resource provider and the manager of the virtual organisation. Explain how WSA would be used to reserve resources within the virtual organisation.

Both parts carry equal marks.

4 The Grid community has converged on the use of Service Oriented Architectures (SOA) to build Grid infrastructures.

a There is considerable discussion as to how state should be represented in the Grid. Describe how state is represented within **two** of the following architectures: Web Services, the Open Grid Services Infrastructure (OGSI), Web Services Resource Framework (WS-RF) and the Web Services Grid Application Framework (WS-GAF).

b The Web Services infrastructure is a SOA built upon three key specifications to discover, describe and communicate within services. Describe these specifications. Detailed description of any XML syntax is **not** required.

c The collection of services necessary to support a 'Grid' is defined within the Open Grid Services Architecture (OGSA). Describe **five** services that you might expect within the OGSA and their expected functionality.

Parts a, b and c carry 25%, 25% and 50% of the marks respectively.