// Learning GRID

A newsletter form the Kaleidoscope Learning GRID SIG

Chief Editor: Nicola Capuano

Learning GRID

A newsletter form the Kaleidoscope Learning GRID SIG **Chief Editor: Nicola Capuano**

Newsletter #8: May 2006

Deliverable D 12.2.10

Prepared for the European Commission
under Contract No. NoE IST-507838 as a deliverable from

Task: 12.2: Networking and Dissemination
Date of issue: 31/05/2006
Version: 1.0
Distribution: Public

In this Issue:

Editorial

Welcome to the eight issue of the Kaleidoscope Learning GRID SIG newsletter and the second issue of 2006.

The struggle to understand how Grid systems can be utilised within learning frameworks presents a real and continuing challenge to everyone who wishes to see the Grid technology deployed fully and effectively.

In this issue of the newsletter we present overviews of two active projects aimed at identifying information on how knowledge can be gathered, processed and consequently distributed and utilised within Grid environments.

By gaining an understanding of these learning processes the resultant knowledge can be distributed within the Grid community thereby enhancing the knowledge pool and providing excellent opportunities for informed learning processes.

In short, if we can understand the nature of the knowledge that is required to use a Grid effectively, then this information can be incorporated within learning programs thereby producing informed and knowledgeable Grid users.

The first article focuses on the work of Steven Johnstone, who presents a self-organising peer-to-peer Grid middleware network model in support of intelligent knowledge pooling for collaborative learning environments.

The second article reports on the work of Jawad Yaqub. In this article the problems associated with the development of knowledge models for Grid systems is discussed. Bascially, if we do not have a firm grasp on what knowledge (information) is required to develop applications for a Grid system then (clearly) poor quality applications will be produced resulting in a loss of confidence in the Grid in general.

Both of these articles argue a case for possessing a greater understanding of how Grid knowledge can be gathered and utilised. An ultimate goal would be the development of self aware Grid systems that could extract knowledge from their environments (in other words learn about their environments) and as a consequence process users requirements with greater efficiency and effectiveness.

In the *Technology Watch* section Blanca Jordan presents an overview of the Sakai project. This project is concerned with the development of a new Collaborative and Learning Environment. You can read more about this on pages 15 and 16.

The News section contains many interesting features including, e.g., information on a special issue of the International Journal on Applied Artificial Intelligence dealing with learning Grid Services, the Development of Next generation Web Service Specifications, and the release of Gridscape II.

Finally, the *When-What-Where* section provides an overview of forthcoming events (Calls for Papers, Conferences, Workshops, etc.) that should be of interest to our Grid Learning community.

We hope that you will enjoy this issue of the newsletter.

Paul Sage & Peter Milligan

Learning GRID SIG Members

Self-organising grid topologies for collaborative knowledge distribution

This article presents a self-organising, peer-to-peer, grid middleware network model in support of intelligent knowledge pooling for collaborative learning environments. Peers share information, such as usage history and centres of interest, to self organise into semantically clustered groups and intelligently elect super peers using a neural network. Super peers can efficiently route queries based on a 'small world model' and demonstrate a strong self organisational capability using ontology based interest representations. Preliminary results show that the neural network super peer election architecture can be trained in an acceptably low amount of time resulting in successful autonomous election of efficient and reliable super peers.

1. Introduction

Peer-to-peer (P2P) systems have recently attracted considerable attention in the research community due to the success of file sharing systems such as KaZaA [1]. P2P systems offer very few or no guarantees about the lifetime of the nodes which are transient by nature [2]. To ensure a high quality of service the P2P network must be tuned towards efficient message routing with tolerance of change in, or failure of, participating nodes.

A class of P2P systems known broadly as distributed hash tables (DHTs) offer a hashing functionality where content is assigned an identifier which is hashed with the result being a network storage location. In the case of Chord [3], objects can be located in O(logn) hops. However, current DHTs support search by object identifier only, limiting the ability to perform complex searches. Further, they do not consider the heterogeneity of peers inherent in P2P networks.

Peers vary greatly in their processing capability, connection times, available bandwidth and shared resource types. The super peer topology exploits this heterogeneity through assignment of greater responsibility to a subset of high quality nodes known as super peers.

As super peers handle all query traffic, their failure can leave the peers for which they handled queries effectively isolated, necessitating efficient message routing strategies. Given the main use of current super peer systems is for the application of file sharing, any such routing strategy should allow the user to form expressive queries. This paper considers an approach that allows both of these major design challenges to be met. The iXChange model utilises the information available from participating peers to: cluster based on shared interest, intelligently elect high quality super peers, recover from super peer failure and efficiently route messages.

Peers cluster, based on their shared interest, resulting in the formation of peer groups. Using exchanged performance data the peer group can intelligently elect a super peer to manage the peer group based upon superior performance characteristics. The routing strategy is a greedy one where super peers attempt to route queries as closely as possible to a location where they may be answered by exploiting the small world characteristics that occur in peer-to-peer file sharing communities [4].

2. Related Work

General design issues for super peer networks are discussed in [5]. [6] and [7] propose more adaptive self-organising approaches for P2P architectures. [6] makes topology decisions based on history. Peers maintain connections with nodes from which they have received successful responses in the past. [7] makes topology decisions using a gossip-based protocol and introduces the requirement for the selection of super peers based on capacity. Both present metrics by which this capacity can be measured although neither discuss an intelligent super peer selection mechanism as described here.

[8] proposes to segment a P2P network based on the topics the peers in the system are interested in and queries can be efficiently routed using a global routing scheme based on the work described in [9]. However, [8] relies on global knowledge of the topics stored within the network and they suggest a global routing technique based on Symphony [9]. The approach taken in [10] describes super peer networks using schema based content descriptions to facilitate efficient query routing, through the use of meta data indexes, stored at super peers that keep a record of metadata descriptions of content stored at other sites. The schema approach is beneficial as it allows a high-grained description of peers' interests, which should allow the creation of clusters that are highly representative of peers' interests. A recently proposed architecture, ICN [11], clusters peers using a more limited level of meta data than the schema approach. ICN forms clusters based on single sub-categories of large content categories.

3. Super Peer Networks for Large Scale Information Retrieval

A study of the Gnutella and Napster [12] systems highlights that the user population of both systems included significant heterogeneity in the available bandwidth of participants and concluded that this heterogeneity should be exploited by assigning peers with greater capabilities an increased level of responsibility. In the case of the totally unstructured systems, the potential of the overloading of low capability peers is removed by not requiring these peers to handle query traffic. For hybrid systems there is no longer a single point of failure due to the removal of the requirement of a single object location server.

3.1. Point of Failure

Super peer failure means peer group isolation and since peers are transient by nature, connection is of an indeterminate duration. Current recovery techniques are not adequate. For example, in JXTA, a node left unconnected to a super peer for a set length of time, automatically becomes a super peer. This leads to the creation of unsuitable super peers, introducing performance bottlenecks due to limited capabilities or poor connection behaviour. As the global nature of a large scale, P2P, interest clustered network cannot not be accurately specified at design time, super peers cannot be predetermined. Therefore, the network requires the ability to identify and subsequently elect or reject super peers. This should be guided by issues relating to the physical capability of peers to take advantage of node heterogeneity. Additionally, the behaviour of a peer, while it is connected to the network, should be a factor when selecting super peers.

3.2. Super Peer Message Routing

A peer issues a query to its super peer. If the super peer can not handle the query then the query is propagated to other interconnected super peers. Queries are assigned a time to live (TTL) value. If no super peer can handle the query, it does not continue to propagate around the network indefinitely but will be terminated.

3.2.1. Semantic Organisation.

To facilitate efficient object location, and as a means by which peers may dynamically form and join peer groups, the P2P overlay is organised according to the interests of the participating peers, ensuring close proximity of peers sharing similar interests. If a peer frequently requests content of a particular type, high quality results should be returned quickly if the object can be located within the peer's group.

3.2.2. Small World Characteristics of File Sharing Communities

Through a study of Gnutella, KaZaA and a scientific collaboration network [4] demonstrates that file sharing communities exhibit small world characteristics: firstly a small average path length, where path is defined as the shortest node to node path and secondly, a high clustering coefficient, independent of network size, a measure of how many of a node's neighbours are connected together [13].

IXChange limits this small world to a set of super peers. The intuition is that six degrees of separation may be found by utilising the notion of shortcuts or long distance links. By making use of the information in the system IXChange creates links between super peers based upon content similarity within their respective managed groups.

4. Ontology based Peer Characterisation

Peers are organised semantically to create peer groups of similar content facilitating routing of queries to sets of peers most likely to respond. In iXChange peers' interests are represented using a standard approach to ensure that the structure of peer characterisation data is consistent, meaning that the semantics of the data can be extracted easily.

4.1. Concept Trees for Interest Representation

Central to the creation of interest centred peer groups is the use of semantic web standards to cater for differing approaches to peer interest classification. The iXChange model requires that classification is carried out using ontologies written using the OWL [14].

Taking into account the fact that classification techniques can produce results of varying quality, iXChange only takes account of the concepts or ontology classes representing a peer's interests.

4.2. Concept-Based Group Creation

The use of concept hierarchies facilitates the creation of peer groups and helps reduce the search space by allowing the creation of sub groups. The process of group creation and group membership in iXChange is now described:

1. Peers generate an XML peer advertisement including the concept hierarchy tree
2. An attempt is made to discover peers matching the knowledge domain at the base of the tree
3. If not successful a new domain group is created

4. If domain search is successful compare concepts on this peer with discovered peer.
5. Traverse the class hierarchy to the most specific sub class that still yields matching peers
6. If no peers share specific concept traverse up hierarchy one level for a more general concept and attempt to find matching peers. Repeat until peers are found or the base of the hierarchy has been reached
7. Join or create a group with the discovered peers sharing common concepts

The above grouping process yields some useful properties. Firstly, the process does not restrict peers to a single ontology for a given domain. For example, if the same concepts were structured by different ontologies in different levels of hierarchy, the groupings will still be achieved. The concept hierarchy is stored within peers in the form of a tree, traversed using a breadth first traversal algorithm. If at any stage during the traversal there are no matching subclasses, the traversal does not stop.

It cannot be assumed that the trees on different peers share the same hierarchy structure. The whole tree is traversed when comparing peer interests until a match has been found or the top of the tree has been reached. Due to the top down nature of the traversal, it is possible that a peer will find matches with other peers at more general levels of its hierarchy that do not share the more specific concepts. These matches facilitate the notion of group/sub-group network segmentation. The result of this is a network segmentation structure consisting of groups and sub-groups, resembling the concept tree structures of the member peers.

5. Intelligent Super Peer Selection

The decision of whether or not a peer is a super peer candidate is left to the user. If the group requires a super peer election to take place, peers can query for candidates according to the following scenarios:

1. If a super peer is still present in the group, peers can query this super peer for a list of other peers that are possible election candidates
2. If no super peer is present for the group the peers will check the local advertisement cache for candidates. (peers will periodically try to discover nodes in the group to maintain a view of the group)
3. If the local cache does not contain candidates, peers will send discovery requests for super peer candidates.
4. In the case where there are no super peer candidates, peers can connect to one of a set of reliable public super peers until suitable candidates become available

On gaining a connection to a super peer a node will periodically test for its presence. If the super peer fails to respond in a suitable amount of time the node will send an 'Election Request' message to the rest of the group.

5.1. iXChange Super Peer Selection

The super peer selection approach aims to reduce the impact of the point of failure problem by reducing the probability of failure through the use of only high quality super peers. For election purposes, iXChange peers share information in a number of categories: Available Bandwidth, System Memory, Processor speed, Shared Storage Available and Average Uptime. Using these criteria, a combination of physical capability and past behaviour are used to make a decision for super peer selection.

Although work such as [5] offer some rules of thumb for the design of super peer networks there has been no real study of what features are ideal in a super peer. [2] shows that there is a trend that most peers in the file sharing systems studied do not have long lifetimes but that around 10% stay in the system for long periods of time. Such peers would be ideal super peer candidates. However, there is no evidence to suggest how important physical capabilities of these systems are to super peer performance. iXChange uses a back propagation neural network model for best candidate selection. The output from the back propagation network can be used to compare peers' suitability for election to super peer status. The aim of the network is to predict if, based on the input data, a peer would make a good super peer. The output of the network represents a quality measure of the super peer's likely performance.

5.1.1. Training the Neural Network

When a node first starts as a peer in the system it will have to train its neural network to learn the properties of a good super peer. It will discover super peers and gather the criteria data from their peer advertisements. To calculate the 'quality' the round trip time (rtt) for the query, the bandwidth utilisation (u), maximum available bandwidth (B) and the average uptime (t) for the super peer are determined. An output value for the super peer is determined using equation 1, which ensures that super peers with high line utilisation, low RTT times and high average uptimes will achieve the highest output values.

$$O = (U(B) + t)/ rtt \qquad \text{equation 1}$$

It is the task of the neural network to determine a mapping between the input criteria and the output value calculated from actual super peers in the system.

5.1.2. Utilising the Neural Network

When the neural network has been successfully trained, a peer can take part in super peer elections within peer groups. During the election a peer will receive a list of candidates for promotion to super peer. The input criteria can be gathered from the candidate peer's advertisement. The output represents the likely performance of the peer if elected to be a super peer. The candidate with the highest output value will be the selected candidate for election.

5.2. Voting Process

When a peer has selected a candidate a single 'Vote' message is sent to it. As with any election process there is a set time for voting to take place. After this time has elapsed candidates will not process any further vote messages.

The voting process concludes with the vote counting scheme. Prior to the start of the voting process the peers agree upon a group ID. This group ID will be used to temporarily segment the candidates that received votes from the rest of the peers in the group. When the vote counting stage has been reached, candidates holding votes join this temporary group. Each member of the temporary group will have the totals of the other members. When the candidate with the most votes has been identified it will start itself as a super peer for the group.

6. Super Peer Interaction

With peer groups created according to context information in the form of interest centres, the super peers can utilise this information to store a representation based on a combination of the knowledge domain and key concepts represented within the group. Queries are propagated via super peers to super peers most likely to be able to respond. This model forms the basis of a routing mechanism for efficient query propagation through the super peer network.

If a peer queries its super peer for content similar to the group's, should the super peer be unable to give a response, it will forward the query to its neighbouring super peers sharing similar content; hence making use of its short distance links. Alternatively, if a peer issues a query for content outside of the scope of the interest centre of the group, the super peer should make use of long distance links to forward the query to a super peer most likely to be able to respond. To maintain a list of relevant links super peers will periodically discover other super peers in the system and obtain their interest centres and using this both long distance and short distance links can be obtained.

6.1. Groups and Query Matching

Group representation and query matching is based around the vector space model [15]. Super peers in the iXChange model store a representation of the concept hierarchies represented by the group they are managing. These hierarchies are converted into a term vector and stored locally at the super peer. As the concepts in the peer group change, the term vector can be updated accordingly. When a user issues a query it is stemmed and stop words removed. The resultant query string is then represented as a term vector, which is routed to the peer's super peer. When the query arrives at a super peer it examines each of the terms in the vector to see if they match any of the concepts in the peer group. If they do the query will be passed on to the relevant peers. If not, the super will attempt to send the query first to super peers connected via short distance links and then those connected via long distance connections.

7. Experimentation and Results

The first stage of implementation is a simulation to verify the performance of clustering and super peer selection. Presented in this section are simulation results demonstrating the quality of peer clustering or group formation and the training of the super peer selection neural network based on peer information exchange.

7.1. Peer Groups

Table 1 shows the concept hierarchies used to carry out simulation tests to verify the peer grouping mechanisms. All of the concepts are selected from the music domain. Concept hierarchy E represents a peer that has been classified using two different paths through the concept tree. A peer with hierarchy E is interested in, or has content relating to, both fast music and heavy metal music. It may just be that the peer is interested in heavy metal that has also been classified as fast music but it is desirable to separate these so that the peer may be identified as one which may respond to queries both about specifically heavy metal music or fast music. Hierarchy G represents a peer that can only be classified as having an interest in music.

Table 2 shows the results from taking a set of 30 simulated peers and allowing them to self organise into groups. This process was repeated with different sets of concept hierarchies represented according to the list in the leftmost column of Table 2. The subscript number shows the number of peers representing each hierarchy. When peers are represented by a single hierarchy such as 'A', the peer group formed is based around the heavy metal concept, as all the peers in the set share this as their deepest point in the concept hierarchy.

Concept Hierarchy	*Depth of Hierarchy (highest -> lowest)*			
A	Heavy Metal	Rock Pop	Genre	Music
B	Rock Pop	Genre	Music	
C	NewOrleans	Jazz	Genre	Music
D	Fast	Tempo	Music	
E	Fast HeavyMetal	Tempo Rock Pop	Music Genre	 Music
F	AFlatMajor	A	Tone	Music
G	Music			

Table 1. Concept Hierarchies

Concept Hierarchies	No. of Created Groups	Group Name (s) – deepest common concept
A_{30}	1	HeavyMetal
A_{15},B_{15}	2	HeavyMetal, Rock Pop
A_{1},C_{29}	2	NewOrleans, Genre
D_{29},E_{1}	1	Fast
A_{14},D_{1},E_{15}	2	Heavy Metal, Fast
B_{15},D_{14},E_{1}	2	Rock Pop, Fast
B_{14},C_{1},D_{14},E_{1}	3	Rock Pop, Fast, Genre
A_{15},F_{15}	2	Heavy Metal, AflatMajor
A_{1},F_{1},G_{28}	1	Music

Table 2. Generated Peer Groups

Multiple peers of different hierarchies result in the generation of multiple groups as shown with the case of the 'A, B' set, where two distinct groups are generated according to the deepest tree level shared by peers. Interesting cases are those where there is only a single instance of a certain hierarchy such as the 'A, C' combination. Here the single 'A' peer cannot find any matching peers to its most specific concept. Due to the multiple 'C' instances a 'NewOrleans' group has been created but, as the 'A' peer does not have in interest in the specific genre of 'New Orleans jazz', it is forced to request the generation of a more general group, representing 'genres' as this is the highest point in common it shares with others.

This result shows the success of the creation of groups and sub groups. Essentially the 'NewOrleans' group is a sub group of the more general 'Genre' group. Those 'C' peers will automatically become members of the 'parent' group. Such multiple group membership is also achieved in the 'B, D, E' case from Table 2. Here, due to the multiple interest classifications on the 'E' peer, it can join two groups each representing its interests. The 'A, F, G' case shows that even with the usage of complex ontologies, to define a domain if a peer cannot be classified accurately or it cannot find precisely similar peers, all that is required to join or create a peer group is the general knowledge domain that represents the peer's interest. This is shown in the case mentioned by the creation of the most general 'music' group.

7.2. Neural Network Selection of Super Peers

Implementation of the neural network was achieved using the Java based neural network framework Joone [16]. The resultant networks can be easily incorporated into applications. As the approach described in Section 5.1.1 describes the need to compare the desired output value with the output produced by the neural network in order to train it, it was trained using a back propagation network. The Joone framework utilises the RPROP or Resilient Propagation training technique for back propagation networks, which improves training times. The back propagation network was designed with three layers.

The first layer in the network is the input layer with a linear transfer function to the hidden second layer and consists of 5 neurons. This layer uses a sigmoid activation function. Differing numbers of neurons in this second layer were tested in order to achieve fast convergence. The output layer consists of a single neuron and also uses a sigmoid activation function. Table 3 shows the results obtained through training the neural network on three different systems.

Since P2P networks contain significant heterogeneity, tests were performed on 3 different PCs with CPU speeds of 350 MHz, 600 MHz and 2.4 GHz. The size of the training data set was varied using a small data set of 25 and a larger set of 200 to compare overall error and training time.

These training results are promising in that a low error of 0.04 was achieved in a good time using the larger data set. For accuracy of the resulting neural network it would be preferable to use the larger data set and, as can be seen from Table 3, even on a CPU speed of 350 MHz the neural network can be trained in just under 14 minutes. When using the larger of the two data sets, increasing the number of neurons in the hidden layer appears to have no effect on the overall error but it takes a little longer to train whereas there is a noticeable difference when using a smaller data set.

CPU (MHz)	*Hidden Layer neurons*	*Training Set Size*	*RMSE (10,000 epochs)*	*Training Time (s)*
350	3	25	0.01	124
	3	200	0.04	810
	10	25	0.003	129
	10	200	0.04	904
800	3	25	0.01	90
	3	200	0.04	710
	10	25	0.003	100
	10	200	0.04	736
2400	3	25	0.01	14
	3	200	0.04	122
	10	25	0.003	15
	10	200	0.04	150

Table 3. Backpropagation Training

8. Conclusions and Future Work

Peer groups can adaptively elect super peers from candidates within the group intelligently based upon a back propagation neural network approach where peers gather their own training data from existing super peers in the system. Simulation results have shown that peers can successfully form peer groups, based around their interest centres, generated through the use of standard ontology class/concept hierarchies and results show that the proposed neural network super peer selection approach can be trained successfully with an acceptably low error and in an acceptable amount of time for a range of system capabilities.

The iXChange model makes use of information exchange regarding peers' interest centres to achieve efficient query propagation between super peers in the system, based on the small world characteristics that have been observed in knowledge sharing communities. Implementation of the routing mechanism is currently underway with simulation work to be carried out prior to a full implementation on a P2P platform for testing.

References

[1] Leibowitz, N., Ripeanu, M., Wierzbicki, A., "Deconstructing the Kazaa Network". *3rd IEEE Workshop on Internet Applications (WIAPP'03)*, San Jose, CA, June, 2003

[2] Sen, S., Wang, J., *"Analyzing* Peer-to-Peer Traffic Across Large Networks", *IEEE/ACM Transactions on Networking*, Vol. 12, No. 2, April 2004

[3] Stoica et al, "Chord: A Scalable Peer-to-Peer Lookup Protocol for Internet Applications". *IEEE/ACM Transactions On Networking*, Vol. 11, No. 1, Feb. 2003

[4] Iamnitchi et al, "Small-World File-Sharing Communities". *InfoCom 2004*, Hong Kong, March 2004.

[5] Yang & Garcia-Molina, "Designing a Super-Peer Network". *International Conference on Data Engineering*. Bangladore, India, Mar. 2003.

[6] Condie T., Kamvar S., Garcia-Molina H., "Adaptive Peer-to-Peer Topologies", *4th IEEE International Conference on Peer-to-Peer Computing*, Zurich, Switzerland, Aug. 2004

[7] Montresor A., "A Robust Protocol for Building Super Peer Overlay Topologies", *4th IEEE International Conference on Peer-to-Peer Computing*, Zurich, Switzerland, Aug. 2004

[8] Bawa et al, "SETS: Search Enhanced by Topic Segmentation". *26th Annual International ACM SIGIR Conference*, Toronto, Canada, July 2003

[9] Manku G., Bawa M., Raghava P., "Symphony: Distributed Hashing in a Small World", *4th USENIX Symposium on Internet Technologies and Systems (USITS'03)*, Seatle, WA, USA, March 2003

[10] Nejdl et al, "Super-peer-based routing strategies for RDF-based peer-to-peer networks", Journal *of Web Semantics: Science, Services and Agents on the World Wide Web 1*, p177-186, 2004

[11] Bai X., Liu S., Zhang P., Kantola R., "ICN: Interest-based Clustering Network", *4th IEEE International Conference on Peer-to-Peer Computing*, Zurich, Switzerland. Aug. 2004

[12] Saroiu, S., Krishna, P. & Gribble, S.D., "A Measurement Study of Peer-to-Peer File Sharing Systems". *SPIE Multimedia Computing and Networking (MMCN'2002)*. San Jose, California, USA, Jan 18, 2002.

[13] Watts, "Small Worlds: The Dynamics of Networks between Order and Randomness", *Princeton University Press*, 1999.

[14] McGuiness D., Van Harmelen F., Editiors, "OWL Web Ontology Language Overview", *WC3 Recommendation*, Feb 2004, http://www.w3.org/TR/owl-features/

[15] Berry M., Drmac Z., Jessup E., Matrices, "Vector Spaces and Information Retrieval", *SIAM Review*, 1999, vol. 41(2), pp 335-362

[16] Joone, http://www.jooneworld.com

S. Johnstone

Ph.D. Student at School of Electronics, Electrical Engineering & Computer Science, Queen's University, Belfast

Research Project Focus: Grid knowledge modelling

Traditionally, users of parallel systems have fallen into two groups: those who wish to have a detailed knowledge of parallel systems (perhaps for the purpose of developing systems software such as compilers) and those who wish to use the machines but do not want to have to bother with detailed knowledge of architectures, parallelisation techniques, etc. (in other words, typical computer users who want the benefits of parallelism but do not want to have to deal with technicalities).

While these views have migrated to the newer world of Grid and Peer-to-Peer systems there remains a need to have a clear understanding of how Grid systems function and how they can be used effectively.

In other words, the nature of the knowledge required to use a Grid environment effectively must be determined and publicised widely.

This project started with an examination of existing knowledge models (such as the KATT system proposed by McCollum et al) demonstrating that the model proposed in KATT was defective, and proposes a new, novel model for describing the nature of the knowledge required to utilise Grids effectively.

The goal is the provision of a knowledge modelling system that can be used within a computational Grid environment. If successful, the model can be distributed to a wider base of Grid users and expanded to deal with different Grid applications. The knowledge model will provide a useful learning paradigm for those working within Grid environments.

1. Introduction

The purpose of this work is to create an automated knowledge modelling system. No such system has been demonstrated to date, and yet would have widespread application in many areas of computer science, most especially Artificial Intelligence where knowledge management is often a key component in intelligent, adaptable behaviour.

The knowledge modelling system proposed is based on the theory that the universe of discourse maybe divided into two parts, which are given the labels of Pattern and Awareness. It is impossible to model awareness/consciousness which is the base context of all patterns, and so a knowledge modelling system will focus patterns within the universe of discourse, using them as material for knowledge models. These patterns are extracted through a set of pattern acquisition algorithms which query the domain of operation and provide patterns which are then indexed.

It is normal to think of knowledge as static forms of information which are available for manipulation however a key principle of this work is that *knowledge is behaviour*, or a process. The points of reference which we normally call knowledge or information exist at stable points of interaction between the knowledge processes. For example, a symbolic alphabet (such as A,B,C, or 1,2,3, 0,1) consists of symbols with no intrinsic meaning. The meaning arises from the behaviour engaged with these symbols.

The symbol '1' means single unit only when there is a mathematical process which is queries it in such a way, i.e. it was set down as a 1 because it would be reliably interpreted that way. There is direct behaviour-to-behaviour mapping, with the symbol as a nexus. Thus when generating knowledge models the major criterion for their success is not the form of representation, but rather their development of utilitarian behavioural responses within the context of operation. This *process feature* frees a software system from many anthropic constraints one would normally associate with a modelling task, and is the central platform for the effectiveness of the modelling system under construction.

This new modelling system is embedded within an intelligent agent working on a collaborative learning Grid environment.

2. Related Work

The Semantic Web, [1] specifies mainstream grid architecture for use by e-Science and forms the template for the computational grid architecture employed here where task integrity is given equal importance to information retrieval. In [1] there is a definite focus on the role of knowledge in various aspects of the grid, the role of ontologies and their justification in the context of inter-agent operation. Important issues facing current grid developers are specified, and knowledge modelling is highlighted as a major one.

Shakshuki et al [5] present a suitable agent-based architecture for co-operative information systems. The approach highlights the reduction of complexity for the user and does so through user controlled agents which in turn work through brokerage agents to bridge to resources. Agent architecture is then presented in which knowledge is key.

The advantages of the approach are manifold: it is streamlined, elegant, and meets the needs of those users who are interested in the information exchange. However it does not meet the requirements for a task-to-resource system as the agents do not provide the capability to maintain task integrity, nor do they have a knowledge modelling system with sufficient degrees of freedom to deal with the complex ontological manipulation necessary for the desired level of autonomy which would guarantee the complexity reduction the authors seek.

The major influence in terms of artificial intelligence has been the recent paradigm shift away from symbolic AI towards what is increasingly being referred to as Embodied Cognition [8].

Andersen provides an excellent overview of the new field and it's emphasis on 'situated awareness' wherein the AI's appropriate response is due to a consideration of the context in which it is working rather than through an internal symbolic engine. Knowledge modelling and representation play a key role in this, and this has obvious implications for intelligent agents and their behaviour within grid environments.

Chrisley took issue with some aspects of Andersen's review [9] raising issues of representational necessity, to which Andersen [10] responded with an acknowledgement and clarification that while representation is still necessary it is of a distributed nature throughout the domain of operation rather than simply being centralised within the AI.

Non-Symbolic Knowledge Modelling Agent: Miao et al [6] in contrast to [5] are concerned with the specifics of agent behaviour and those capabilities of reasoning, modelling, and decision making. This they do through a presentation of an agent which performs each of these tasks numerically. This again circumvents the ontological requirement, as well as the traditional symbolic approaches and has the advantage of being very fast, amenability to mathematical modelling, and being much smaller.

However while the agent is able to make knowledge objects and specify causal relations between them, as well as map inputs values to object values, it does so in a rigid fashion with a fixed set of rules to guide it's behaviour, and thus it's autonomy is limited, and it's flexibility lacking for the much more dynamic grid environment sought here.

Nevertheless, the proof of concept of a non-symbolic approach to knowledge modelling in which causal knowledge relations between objects are established without recourse to semantically accessible representation shows that such an approach is possible in the modelling system under development here.

The necessary heterogeneity of the Grid implies that new types of agents will appear within a reasonable timeframe and some will almost definitely have updated or even entirely different communication languages than the ones the rest of the system possess. As communication is a vital part of the architecture, the effect of the multiplicity of languages is studied in Kamarova et al [4] and it is discovered that far from settling on the most widely used of the languages, the given set converge to a better communication state with a blend of the available languages. This raises issues of expressibility, ambiguity and representation and communication of knowledge effectively.

The emphasis in Azoulay-Schwartz et al [3] is on achieving a stable form of information interchange in a neutral and competitive agent environment. It shows that with a relatively limited set of rules it is possible to achieve a stable communication network even if agents choose to refuse requests on the basis of past performance of requesting agents. It is important then, that any developed agent should either converge upon these stable strategies of communication or discover new ones through simulation and experimentation, and thus some sort of heuristic search or knowledge modelling capability is implied.

2.1. Goals and Negotiation

Each agent will have goals in relation to the tasks it must perform, and each agent will have need of services in order to meet those goals. As agents may bid for services, broker services, or establish communication links, negotiation is all but unavoidable. Thus Shaheen et al [11] propose an agenda based framework for pursuing single and multi-issue negotiations, and Freund [12] proposes a system to revise preferences to improve rational communication, both of these combine with Stewart's [13] analysis of the complexity of task maintenance and achievement to give an overview of the terrain an agent must traverse in order to achieve a reasonable number of goals within a realistic timeframe. All of these systems require the accumulation and application of knowledge models to pursue their ends.

2.2. Ontologies

Ontologies are important templates for the exchange and manipulation of complex forms of knowledge between systems. Ontologies are usually very difficult to alter and translate effectively and thus Gruber's [18] to specifying portable ontologies provides for more flexible way to exchange knowledge between the varieties of agents in the grid. While standardisation would be preferable it worth noting that even experts who create viable solutions do so with ontologies that do not match which is shown clearly in the paper by Hameed et al [2] where they cite several forms of ontology mismatch amongst ontologies clearly in use and effective in real world environments.

The implication is that ontologies constructed by users and supplied to agents are highly unlikely to match exactly with each other and thus some form of translation software will be required for knowledge maintenance. This situation is worsened by the findings in Silva et al [7] which indicate that ontologies can have severe limitations, some of which this knowledge modelling system will seek to address.

3. Background and Theoretical Framework

A recent shift away from traditional purpose built parallel machines towards a distributed Grid/Peer-to-Peer model has provided a new approach to parallelisation, and presents new challenges for those wishing to deploy parallel applications in a heterogeneous and dynamically changing platform. To meet these challenges an agent based service architecture has been proposed [1] for resource management and user interfacing called the **Semantic Grid**. Within this architecture the user is shielded from the heterogeneity and dynamically shifting resource base of the Grid environment by software agents with varying degrees of autonomy. Some software agents will display no autonomy as they may simply provide listings of appropriate services or run direct queries and refer decisions to their owners. Other agents enjoy limited autonomy for negotiation protocols when bidding for services. Ideally as agent should undertake service acquisition within user specified parameters and the user be left with control over significant decisions pertaining to a potentially costly servicing. Fully autonomous agents have specific roles i.e. managing task integrity or re-routing in the event of a resource failure. While the user influences the operational domain of the agent, the agent is most effective when autonomous due to the real-time nature of decision making. All such agents obtain and manage knowledge their environment, each other, and user requirements (in the form of requests for information or jobs to be processed).

This entails development of knowledge modelling and reasoning systems within agents. A typical grid is a group of distributed computing resources connected for running applications which would otherwise be impractical or even impossible to run on a single isolated machine [1]. This work is aimed at the *intersection* between grid technology, multi-agent systems, knowledge modelling and also, implicitly, artificial intelligence.

The project is centrally concerned with the development of the knowledge modelling system for task-to-resource management which will be prototyped on a coarse-grained computational grid (as opposed to an informational grid) whose primary purpose is the execution of linear algebra applications (e.g. LAPACK). A typical task list for an agent will be a specially formatted FORTRAN file with clearly marked task sections which the agent extracts into a series of *tasks* to be run through the grid platform most likely consisting of large scale matrix multiplication and/or matrix addition. Each task is composed of *Processes*. Processes are matched to *Resources* which are either addressed directly through local machine resources or through agents which *broker* services from external sources such as information retrieval, calculation, storage, or any other currently available on the platform for which local needs are not suitable.

The list of services an agent offers may update with time, as can the cost of using a service. The reliability of the service, and the quality and speed of processing and communication of the results are also issues as the agent is responsible for maintaining at least a reasonable quality-of-service for the jobs it accepts [1]. All of these roles require the acquisition and management of knowledge pertaining to the resources, other agents, the tasks, processes, inter-process communication, negotiation for services [5], and scheduling of resource allocation to complete the job. The situation is further complicated by the fact that Agents communicate with each other for information about services offered or known about, in passing results, in co-operation and co-ordination of tasks, as well as in competition with each other while bidding for services.

Thus one can divide the roles of the agent into 3 categories:

1. User Management

2. Inter-Agent Communication

3. Task Management

Each has its own particular knowledge needs as shown below:

3.1. User Management

This role involves knowledge of the user, user preferences, the types of tasks the user has run in the past, an accumulated profile of the types of services most likely to match those tasks and present results in a satisfactory time [3]. The user's membership in groups, their authentication, which services they are able to access or designate available, which they are qualified to use also fall into this category, as do security concerns.

3.2. Inter-Agent Communication

As most resources on the Grid are offered as services through agents inter-agent communication is unavoidable [1, 5, and 13]. The agent must match the processes which need to be run with the services available, and if it is unable to meet the requirements of the task, it must request information on further services of which its peer agents are aware, and may even go to broker agents for this purpose.

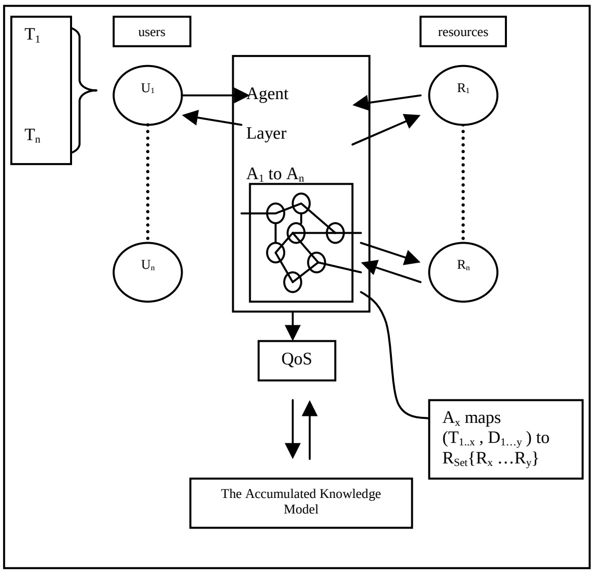
Agents may be uncooperative of the user and the institution, and they will usually compete for resources in limited time frames, so knowledge of agent's capabilities, their dispositions, their associated institutions, the likely services they request, the services they offer, the costs of those services, the quality of their brokerage all fall into this category.

3.3. Task and Resource Management

This category holds knowledge needed to maintain task integrity and complete the jobs assigned to the agent the time allotted. Tasks are composed of processes, they have owners, they have applications, they have a timeframe in which to run and maximum cost for negotiation for services, they are prioritised amongst the services available and affordable to the agent and thus must be scheduled. Each of these represent an area of knowledge crucial to task and management. Tasks present results which must be kept for the user, and they also communicate with each other, and this knowledge too must be collected. The agent must maintain a reasonable running speed for the task through utilization of the knowledge base in decision making.

Users (U) either individually or as part of a group or institution submit jobs consisting of Tasks (T) to Agents (A). These in turn communicate with each other through the agent layer to determine which resources (R) are available, and which match most closely the requirements of the tasks. Then the user's agent will request costs for the services from a list of providers, and select the most appropriate balance between price and speed and quality of service. Agents are not necessarily co-operative[13], and may not always provide answers even if available as their purpose is to maximise utility for their users or groups and thus they must compete for resources and the agent will there for have to converge upon stable strategies for information exchange [3].

Parallel application tasks intercommunicate, requiring intermediate storage for results thus communication (of very large amounts of processed data) will be a service to which a cost may be attached, becoming a factor for economical execution. Intermediate storage almost certainly outlives at least one of the processes which make use of it, and resources upon which these processes are running are able to change thus the user agent becomes the natural choice for preserving and passing results between processes.



4. The Knowledge Modelling System

Developing suitable informational and behavioural models is essential to the agent's role. As a grid is constantly in a state of flux no single knowledge model can accommodate the multiplicity of resources, ever evolving agent types, newly introduced services, or behaviours. Thus a knowledge modelling *system* suited to information rich environments and with a generic platform for creating knowledge models is required.

There currently exists no such system and so the development of such a system will be novel and is a major feature of the work and also a major contribution to the research field.

4.1. Patterns of Knowledge

Work done in recent years has lead to the development of several models for the representation of parallel environments [16]; there was a shift away from a specific model to a generic modelling system for the parallelization process [17] and was thus developed to overcome these issues within information rich environments. This system (shown below) consists of a *universe of discourse* (defined as the environment from which information and knowledge may be drawn) which provides *patterns of knowledge* (information and knowledge brought together from the universe of discourse for and applied for a purpose) through normal *pattern acquisition operations* (such as queries to databases, statistics on agent usage, and information requests to other agents) in the first layer. Pattern Acquisition Operations simply collect and index all the available unique patterns in the universe of discourse from which a model may be constructed. A pattern is any data-structure an agent may use within its domain such as a *list of other agents*, *services*, *storage directories*, *processor specifications*, *a query*, *an organisation name*, *costs for various types of actions*, *a task*, *a process*, *or a job* which are provided through direct acquisition via the agent's knowledge elicitation algorithms, deductive inference from previously elicited/acquired knowledge, or knowledge given through communication with other agents or users. These all make up the complete *world* in which the agent lives, and thus any model the agent generates must at base be rooted in these patterns.

4.2. Mental Models

After patterns of knowledge about the system are acquired these are grouped together in terms of informational content, relevance to each other or to goals and intentions of the user in structures known as *mental models.*

An example of a mental model is a collection of users, agents, services (with associated costs and reliability factors), and previously run tasks that together form the conceptual structure of an *institution.* This conceptualization into one behavioural unit based on relevance of information gained through indexing of typical grid information through the *ownership* pattern allows for better decision making when trying to allocate a task to services run by various institutions - some of which may be associated with the task managing agent's owning institution and thus offer preferential service rates, or may be hostile or indifferent to the agent's agenda and thus may be a costly and wasteful expenditure of effort as a service negotiation candidate.

The sum of all models in play is the ***working knowledge model.*** Previous models which become necessary in the future remain dormant and are available for incorporation, query, and development for agent use. They are the accumulated knowledge of the agent throughout its working history.

4.3. Ontologies

An ontology is an *explicit specification of a conceptualization* and is described as a set of content-specific knowledge representation primitives: classes, relations, functions and object constants [18]. A collection of these ontologies when richly instantiated with domain related content enable interoperability between systems, facilitate communication between agents and organisations, and make existing knowledge shareable and re-useable [2]. For these reasons ontologies are vital to grid computing.

They blueprint mental models. However, as Hameed [2] points out, often there are mismatches between ontologies (even among those created by human experts) which include: conceptualisation mismatches (including class and relation mismatches), and explicit mismatches (concept, term, definitions and their combinations). Not due to individual ontologies being *incorrect*, as one may assume, rather this is due to multiple conceptualisation and classification approaches employed by experts working any given field. Ergo, *ontologies are constructed on the basis of behavioural priorities* which are reflected in their final structure. While an ontology serves the behaviour for which it was instantiated its structure remains intact; however if this behaviour alters, or the priorities which specified the structure alters (as may occur in the grid environment) then restructuring is required.

Several important implications arise for the knowledge modelling system. Firstly it may be possible to employ a ***variety of ontologies for a particular purpose***. Secondly the mismatches between various ontologies may require the generation of ***translation software***, or even ***intermediate ontologies*** to allow suitable communication. Thirdly the ***automatic generation of new ontologies*** and subsequent modification becomes as issue.

Ontologies forming templates for static mental models must be dynamic to ensure continued relevance for the agent's knowledge and actions. This requires a completely new approach to knowledge modelling system architecture, as may be seen in the diagram below:

This architecture has several functions aside from simple pattern acquisition and modelling. Most vital of which are create and modify ontologies through a variety of restructuring algorithms residing in the *Dynamic Ontology* layer. These algorithms evolve to determine stable collections of patterns which form models to serve the agent role. The initial patterns collected are restructured to form mental models which are deployed as representations of knowledge or as operational procedures for the agent.

The evolution of the ontology restructuring algorithms is controlled through the Quality of Service Matrix which evaluates the effect of the ***working knowledge model***. This evaluation provides feedback for the next generation of algorithms, with the most valuable and stable features of each algorithm surviving, with eventual convergence on an optimal (or at least reasonable) modelling repertoire. Agent goals reside in the Purpose Profile which defines the role of the agent. This final, topmost layer is the driving engine behind the whole agent, and can function so because it represents intelligent behaviour as a reaction to entropic forces which prey upon the agent.

The enabling of architecture had led to progress in Knowledge Theory, The Knowledge Modelling System and has stimulated work on publications related to the work.

4.4. Knowledge Theory

The relationships between knowledge and formal systems, representation and consequently patterns have been further researched over the last year. Knowledge is a process and consequently a touchstone for the transfer of behaviour between related higher order processes with certain key characteristics. The knowledge itself is generally independent of the representation system; however there are remarkable consistencies in the symbolisation processes across domains where knowledge is generated. There are also symbol derivative behaviours which are icon specific and consequentially architecturally native. An example of this can be seen in the relative expressiveness of various formal systems, where certain concepts are difficult if not impossible to translate across easily, i.e. ð in binary. The automatic generation of scaleable and practical knowledge models is directly constrained by reliance on human interpretation of symbolic level interactions, and consequently a shift towards the evaluating the effective meaning of symbolic interactions must be programmed into modelling systems to prevent their obsoletion.

4.5. Knowledge Modelling System

The architecture of the knowledge modelling system was designed to address the needs raised by the knowledge theory above. Each of the five components of the architecture use novel techniques to create a knowledge model which is maintained with minimal user interaction. This allowed the specification of a prototype system throughout the five components of the system. Once the specification was completed the Pattern Acquisition level and the Mental Model level were designed, and implementation began. An agent environment was chosen and the agent layer will be programmed in Java as many development tools already exist and so work can focus on the implementation. The final implementation should be completed within the coming year. Tests for the various capabilities of the system have been isolated and test harnesses will be applied to each level of the system before doing a system wide test to establish overall performance.

References

[1] The Semantic Grid, University of Southampton.

[2] Detecting Ontological Mismatches among experts' ontologies acquired through knowledge elicitation – Hameed, Sleeman, Preece, Knowledge-Based Systems 1 (2002) 265-273.

[3] Stable repeated strategies for information exchange between two autonomous agents – Azoulay-Schwarz, Kraus, Artificial Intelligence, 154 (2004) 43-93.

[4] "Convergence of Agent Languages" – Kamarova, Niyogi, Artificial Intelligence 154 (2004) 1-42.

[5] An architecture for co-operative information systems –Shakshuki, Ghenniwa, Kamel, Knowledge-Based Systems 16 (2003) 17-27.

[6] Agent that models, reasons and makes decisions – Miao et al, Knowledge-Based Systems 15 (2002) 203-211.

[7] On the insufficiency of ontologies: problems in knowledge sharing and alternative solutions – Silva et al, Knowledge-Based Systems 15 (2002) 147-167.

[8] Embodied Cognition: A field guide, Michael L. Anderson, Artificial Intelligence, Volume 149, Issue 1, September 2003, Pages 91-130.

[9] Embodied artificial intelligence, Ron Chrisley, Artificial Intelligence, Volume 149, Issue 1, September 2003, Pages 131-150.

[10] Representations, symbols, and embodiment, Michael L. Anderson, Artificial Intelligence, Volume 149, Issue 1, September 2003, Pages 151-156.

[11] An agenda-based framework for multi-issue negotiation, Shaheen S. Fatima, Michael Wooldridge and Nicholas R. Jennings, Artificial Intelligence, Volume 152, Issue 1, January 2004, Pages 1-45.

[12] Michael Freund, On the revision of preferences and rational inference processes, Artificial Intelligence, Volume 152, Issue 1, January 2004, Pages 105-137.

[13] The complexity of achievement and maintenance problems in agent-based systems, Iain A. Stewart, Artificial Intelligence, Volume 146, Issue 2, June 2003, Pages 175-191.

[14] Controlling the learning process of real-time heuristic search, Masashi Shimbo and Toru Ishida, Artificial Intelligence, Volume 146, Issue 1, May 2003, Pages 1-41.

[15] Multiobjective heuristic state-space planning, Ioannis Refanidis and Ioannis Vlahavas, Artificial Intelligence, Volume 145, Issues 1-2, April 2003, Pages 1-32.

[16] A Meta-Heuristic approach to the auto parallelisation of legacy codes – PhD Thesis, Queens University of Belfast, B. McCollum 2000

[17] An Advanced Meta-Heuristic Approach to the Auto Parallelisation of Legacy Codes – MPhil Thesis, Queens University of Belfast, J. Yaqub 2003

[18] A translator approach to portable ontology specifications, T.R Gruber, Knowledge Acquisition 5 (1993) 199-220

Jawad Yaqub,

Ph.D. Student at School of Electronics, Electrical Engineering & Computer Science, Queen's University, Belfast

Technology Watch

Sakai: Collaboration and Learning Environment for Education. Free to Use, Free to Develop, Freedom for Education.

1. About Sakai™ Project

"The Sakai Project is a community source software development effort to design, build and deploy a new Collaboration and Learning Environment (CLE) for higher education"[1]

The Sakai Project is mainly composed by USA partners (high education institutions mainly). It was born at the universities of Michigan and Indiana with the development of open source code to replicate and enhance the functionality of their existing content management systems.

Next institutions joining the initiative were the MIT, Standford and the Open Knowledge Initiative (OKI) and the uPortal consortium in 2003. This initiative has a main goal to deliver an application framework and tools that can be used as content management system.

The Sakai Project's primary goal is to deliver the Sakai application framework and associated CMS tools and components that are designed to work together. These components are for course management, and, as an augmentation of the original CMS model, they also support research collaboration. The software is being designed to be competitive with the best CMSs available.

2. Sakai Product and Technologies

The Sakai product tries to satisfy the actual goal of any CMS: easy to use, scalable, flexible, portable and reliable.

It has been conceived to work either in academic environment or in enterprises environments. The offered solution is suitable to cover the learning/teaching process and also to facilitate the research collaboration. The Sakai product offers ad hoc communication tools.

The Sakai solution is based in SOA (Service - oriented Architecture) and it is structured in a two-layered approach. Within this architecture there are defined some services' interfaces that allow the communication and interaction of the different components.

The technology framework used by Sakai is the Java enterprise environment. And the goals that have been defined to achieve the requirements expressed before are among others:

– Component –based expandability

- Integrated application
- Presentation based on portals and HTML.
- Support for web services

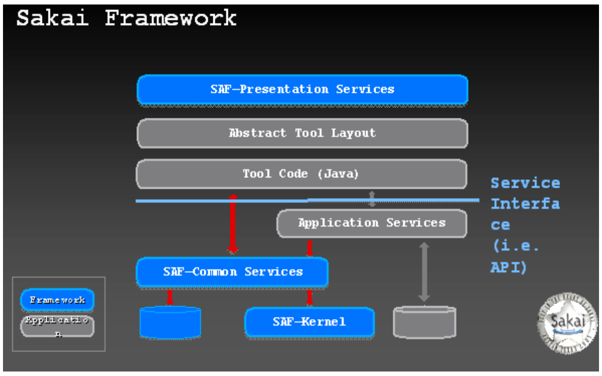


Figure 1. Sakai Framework

The complete Sakai's framework and applications are represented in figure 1.

As it was said previously, Sakai is oriented to offer web services support and web applications. The way how it has been implemented follows the structure that is represented in the following figure:

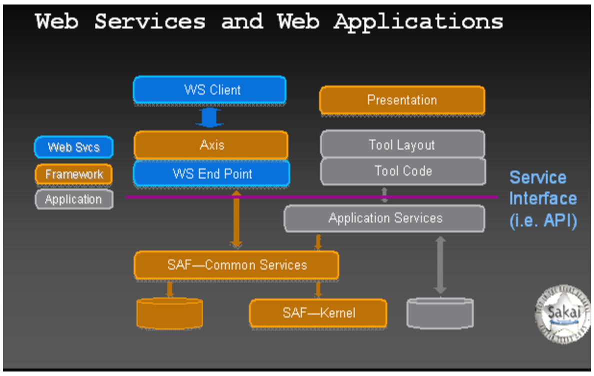


Figure 2. Web Services and Applications

The kernel component is a crucial part for the system mainly because it contains, among others, the functionality of synchronising the web services and the web applications. This kernel was totally rewritten for Sakai's 2.0 version. Currently, the last release available is 2.1.

3. Sakai partnership and future

Sakai initiative is currently composed by 85 institutions all over the world and is in permanent growing. These institutions are public and private and from academia or industry. Some of the most important partners related worldwide to e-Learning initiatives and also to Sakai are IMS Consortium, OKI Initiative, SUN, UNISYS, IBM, etc. The adoption of the solution by the different institutions is planned for the coming years in a clear implementation schedule.

The Sakai initiative seems to have a good projection and an increasing position in the market. The expected position is represented in figure 3.

There is also foreseen some structural changes for the Sakai Community, from what it is now – a group of interested institutions working towards the same direction – to an integrated solution that may cover the market niche of a complementary solution to existing CMSs (not a competitor).

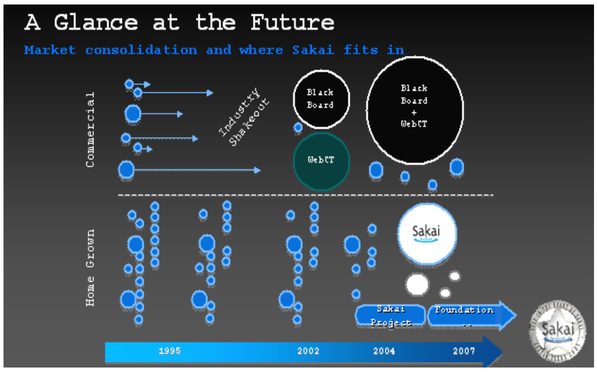


Figure 3. Sakai's Future

4. Learn more about Sakai

The description offered in this article is just a part of what Sakai is and means. Sakai initiative has huge potential and also application possibilities that are not already covered.

The Sakai Community also offers two interactive groups: The Sakai Working Groups and the Sakai Discussion Groups. These two components have the final aim of fostering the discussion among all the researchers and work towards the same direction. They contribute to the generation and discussion of new ideas and paradigms for the Sakai Community.

Finally the Sakai Project has launched the Sakai Foundation in October 2005. This is a non-profit corporation to support, sustain and promote the project and the research community. The governing bodies of this foundation have planned the future not only on the technical aspects but also in the economical ones ensuring the continuity of the results. Sakai may be a self-sustained initiative that will also find its place in the market.

In order to know the Sakai project in deep it is necessary to visit it web site and also the list of references that appear in the References section (These links are recommended by the Sakai group to know about themselves, their products and other related tools in deep).

References

[1] http://www.sakaiproject.org

[2] http://bugs.sakaiproject.org/confluence/display/winterconf2005/home

[3] http://www.sakaiproject.org/release

[4] http://collab.sakaiproject.org

[5] http://source.sakaiproject.org/svn

Blanca Jordan
ATOS Origin

News

On-line a Special Issue on Learning Grid

February 2006

The International Journal on Applied Artificial Intelligence published a Special Issue on Learning Grid Services whose Guest Editors are Stefano A. Cerri, Guy Gouarde' res and Roger Nkambou. The Journal addresses concerns in applied research and applications of artificial intelligence and also acts as a medium for exchanging ideas and thoughts about impacts of AI research.

The special issue was published in November and December 2005 and now it is completely reachable on line. The table of content of the special issue on Learning Grid is reported below.

- *Toward Learning Grid Infrastructures: An Overview of Research on Grid Learning Services*. Roger Nkambou, Guy Gouardères, Beverly P. Woolf.
- *Customizing the Instructional Grid*. Beverly Woolf and Chris Eliot.
- *Competence Ontology for Domain Knowledge Dissemination and Retrieval*. Bernard Lefebvre, Gilles Gauthier, Serge Tadié, Tran Huu Duc, Hicham Achaba.
- *Services, Semantics, and Standards: Elements of a Learning Grid Infrastructure*. Colin Allison, Stefano A. Cerri, Pierluigi Ritrovato, Angelo Gaeta, Matteo Gaeta.
- *Collaboration in the Semantic Grid: A Basis for E-Learning*. Kevin R. Page, Danius T. Michaelides, Simon J. Buckingham Shum, Yun-Heh Chen-Burger, Jeff Dalton, David C. De Roure, Marc Eisenstadt, Stephen Potter, Nigel R. Shadbolt, Austin Tate, Michelle Bachler, Jiri Komzak.
- *Ontology-Based User Modelling and E-Portfolio Grid Learning Services*. Liana Razmerita, Guy Gouardères, Emilie Conté.
- *Interactive Knowledge Construction in the Collaborative Building of an Encyclopaedia*. Philippe Lemoisson And Stefano A. Cerri.
- *The Strobe Model: Dynamic Service Generation On The Grid*. Clement Jonquet And Stefano A. Cerri.
- *Analysis And Synthesis Of Learning Agent's Communicative Behaviour*. Nik Nailah Binti Abdullah And Stefano A. Cerri.

- *The Grid-E-Card: Architecture To Share Collective Intelligence On The Grid*. Guy Gouardères, Mansour Saber, Roger Nkambou, Roland Yatchou

More info at: http://www.tandf.co.uk/journals/titles/08839514.asp

Source: The ELeGI Newsletter.

A Joint Effort to Develop Next Generation Web Services Specifications

March 2006

Four of the biggest names in technology have banded together to advance Web services standards, a move that could benefit grid computing. HP, IBM, Intel and Microsoft announced on March 15 that they plan to work together to develop a common set of Web Services specifications for resources, events, and management that can be broadly supported across multiple platforms.

They proposed reconciling two competing approaches: the Web Services Distributed Management (WSDM) family of specifications, including Web Services Resource Framework (WSRF) and WS-Notification (WS-N), supported by IBM, HP and others; and the WS-Management family of specifications, including WS-Transfer, WS-Eventing, and WS- Enumeration, supported by Microsoft, Intel and others.

The published roadmap suggests that the new specifications that are to be developed will include essentially all of the core concepts introduced back in 2001 in the Open Grid Services Infrastructure (OGSI) and subsequently incorporated in WSRF/WS-N.

The new announcement "promises to overcome this final hurdle to industry-wide standards for Web services-based systems management," Foster wrote. "According to the roadmap, essentially all core concepts in the first OGSI draft of 2001 (and subsequently in WSRF/WS-N) are to be included in the reconciled specifications. We have WS-Addressing Endpoint References (EPRs) as names for state elements, document-based inspection of state, lifetime management of state elements, and notification of changes to state. We also have a commitment to building higher-level system management specifications on this foundation."

Source: GRID Computing Planet.

Sun Grid Compute Utility: How to Buy Computing Power Over the Web

March 2006

Sun Microsystems is making its Sun Grid Compute Utility available over the Web in the U.S., charging customers $1 per hour per CPU through billing fulfilment provider PayPal.

Developers and researchers in several fields will be able to access the utility with a click-through license through a simple portal. Customers, who must be using AMD Opteron or Sun UltraSparc machines to procure the processing power, will be billed through PayPal. The computer power will be provisioned within 24 hours of an order.

The move is the next step from the company's commercial plan, in which Sun secured specific contracts with organizations such as Applied BioSystems, CDO2 and Princeton.

These contracts allowed companies in financial services, oil and gas and biological sciences to let their customers dial up or dial down the computing power they required as a utility, according to their company needs. Customers were not billed if they were not using the computing power.

The new offering is essentially the same system already proposed by SAN, but accessible through a portal. U.S. customers can procure the computing power through PayPal here. Sun will announce international availability later this year.

Sun, along with rivals IBM and HP, has been a leading proponent of grid computing. But to this point, Sun has been the only vendor to offer the technology as a utility that customers can control through their laptops at the touch of a few keystrokes.

Analysts tend to agree the technology has a great deal of value in the flexibility it affords customers, but not all agree on whether or not the adoption required to keep the CPUs flowing freely will take root.

More info at:
http://www.sun.com/service/sungrid.

Source: GRID Computing Planet.

Europe and China Researcher and Students Connected Through ORIENT

April 2006

Communication and collaboration among 45 million researchers and students across Europe and China will be greatly facilitated by a new Sino-European high-speed network connection. Co-funded by the European Union, China and European National Research and Education Networks, the €4.15 million ($5.67 million)-ORIENT (Oriental Research Infrastructure to European NeTworks) project will benefit all Sino-European research, including radio astronomy, sustainable development, meteorology and Grid computing, by helping to step up the flow of information between Europe and China. The ORIENT project is supported by the EU's 6th Research Framework Programme.

“ORIENT enables truly international research co-operation, by making geographic location almost irrelevant,” commented Information Society media commissioner Viviane Reding. “Access to applications such as telemedicine, digital libraries and e-learning will help the general public, as well as the research community, to build academic and cultural links between Europe and China and an open exchange of opinions and expertise between Chinese and European researchers. And at the same time, ORIENT will bring together the world's best minds to tackle global challenges such as climate change and sustainable development.”

ORIENT will connect Europe's GÉANT2, the world's most advanced international research and education network and the Chinese research networks CERNET and CSTNET on an overland route via Siberia. Scheduled to go live later in 2006 it will link over 200 Chinese universities and research institutions, at speeds of up to 2.5 Gbps. It will be coordinated by research networking organisation DANTE in Europe and the CERNET network in China.

The ORIENT launch follows Commissioner Reding's meeting with Xu Guanhua, Minister of Science and Technology of the People's Republic of China, in Beijing in January. At the meeting, both sides agreed that e-infrastructure (advanced networking and Grid technologies) and its applications would play a key role in economic and social development and acknowledged the high importance of bilateral cooperation.

Source: GRID Today

The Learning GRID SIG Will Organise a Workshop in Barcelona Inside e-2006

April 2006

The goals of e-2006 (eChallenge conference) are to stimulate take-up of applied ICT research results by industry – in particular SMEs – and the European public sector, to promote knowledge sharing between industry, government and the research community, to exchange experiences about the current state of eAdoption at a sectoral, national or regional level, support International Cooperation and open up the European Research Area (ERA) to the rest of the world.

The conference will take place in Barcelona, Spain from the 25 to the 27 October 2006. The second day will see the presence of our SIG organising a workshop on “Next Generation in Technology Enhanced Learning” together with the e-LeGI project under the umbrella of the Academy Industrial Digital Alliance (AIDA) of Kaleidoscope. The ELeGI-Kaleidoscope Joint workshop aims to bring together researchers, academics, professors, educational scientists and technologists in all areas of ICT and e-learning who are interested in exploring methodologies and systems for an effective learning based on the new technologies.

The workshop will primarily focus on the following aspects: experience in use of existing grid applications within an educational context; instantiation of new pedagogical approaches in Grid-based solutions for learning; advanced service-oriented Grid based software architecture; Virtual Laboratories and Virtual Scientific Experiments for learning; learner models and knowledge representation; enhanced presence, collaboration and conversational processes in e-learning; semantic and knowledge based systems for learning; Virtual Learning Organisations and Communities.

The main objective of the workshop is on exploring both theoretical and practical aspects of TEL in order to valorise those applications that can actualize learner centred approach. The contributors will present recent advances in TEL research, both from pedagogical and technological points of view, and their innovative application in GRID-Aware e-learning environments. The proposed outcomes are to show the most innovative research results, applications and the best practices, to summarize the current status of the relevant work and to identify trends for future researches in the field of the methodologies and systems for an effective learning and on the TEL.

More info at http://www.echallenges.org/e2006.

GRIDS Lab Releases Gridscape II

May 2006

The GRIDS Lab and the Gridbus Project at the University of Melbourne released the Gridscape II software, which manages the gathering of information from arbitrary, heterogeneous and distributed sources and presents them together seamlessly within a single interface. It also leverages the Google Maps API in order to provide a highly interactive user interface. Gridscape II is simple and easy to use, providing a solution to those users who don't wish to invest heavily in developing their own monitoring portal from scratch, and also for those users who want something that is easy to customize and extend for their specific needs.

Gridscape aims at providing a high-level, user-friendly and highly customizable portal interface in order to present the status of Grid resources. It interacts with existing technology so that no additional installation or configuration of Grid resources is required. Major improvements over the previous implementation of Gridscape are that it supports the integration of multiple arbitrary information sources through an extensible design; it provides a simple customization mechanism to allow it to be enhanced to meet the specific needs of each individual Grid portal. Other improvements are integration with Google Maps, simplified portal administration and the use of portlet-based Web components which means it can be plugged into other Grid portals to complement them.

The key features of Gridscape II are the following: it manages diverse forms of resource information from various types of information sources; it provides a simple framework for introducing new information service types; it provides simple portal management and administration; it provides a clear and intuitive presentation of resource information in an interactive and dynamic portal via Google Maps.

It has a flexible design and implementation such that core components can be reused in building new components, presentation of information can be easily changed and a high level of portability and accessibility (from the Web browser perspective) can be provided. Gridscape II currently comes with implementations for MDS2 (Globus 2) and MDS4 (Globus 4) information service types.

Source: GRID Today

When	What	Where
19-23 June 2006	**15th IEEE International Symposium on High Performance Distributed Computing** The Fifteenth IEEE International Symposium on High-Performance Distributed Computing (HPDC) will be a forum for presenting the latest research findings on the design and use of parallel and distributed systems for high end computing, collaboration, data analysis, and other innovative applications More information at: http://hpdc.lri.fr.	**Paris, France**
26-30 June 2006	**8th International Conference on Intelligent Tutoring Systems** The conference provides a leading international forum for the dissemination of original results in the design, implementation, and evaluation of intelligent tutoring systems and related areas. The conference draws researchers from a broad spectrum of disciplines ranging from artificial intelligence and cognitive science to pedagogy and educational psychology. More information at: http://its2006.org.	**Jhongli, Taiwan**
5 July 2006	**6th IEEE International Conference on Advanced Learning Technologies** Life-long learning becomes much more achievable as versatile infrastructures for mobile communication and portable Web access emerge. The conference will bring together people who are working on the design, development, use and evaluation of technologies that will be the foundation of the advanced technologies for life-long Learning. More information at: http://www.ask.iti.gr/icalt/2006.	**Kerkrade, The Netherlands**
3-6 July 2006	**International Advanced Research Workshop On High Performance Computing and Grids (HPC 2006).** The aim of the Workshop is to discuss the future developments in the HPC technologies, and to contribute to assess the main aspects of Grids, with special emphasis on solutions to grid computing deployment. The HPC Advanced Workshops in Cetraro have been well established and two of them (1992 and 1996) were sponsored by NATO. More information at: www.hpcc.unical.it/hpc2006.	**Cetraro, Italy**
4-7 July 2006	**Workshop on P2P Data and Knowledge Sharing (P2P/DAKS)** The aim of this half-day workshop, within ICDCS 2006, is to explore the P2P area of data and knowledge sharing [P2P/DAKS] for information intensive applications. The objective is to blend concepts and ideas from the database and AI domains in order to produce a sound basis for an effective P2P/DAKS facility. More information at: http://www.takilab.k.dendai.ac.jp/conf/DAKS/2006.	**Lisboa, Portugal**
23-28 July 2006	**International Conference on Engineering Education** Organized as an information dissemination and exchange forum, the International Conference on Engineering Education emphasizes the dissemination of information on state-of-the-art advances in education and research, especially innovative approaches that link the two activities in addressing the education goals of new engineers of the 21st century. ICEE programming also emphasizes the importance of trans-national cooperation and cross-cultural networking to advance the effectiveness of engineering education. It is a conference series that highlights the worldwide progress and experiences in engineering education; it is also a platform for forging mutually beneficial collaborative efforts. More information at: http://icee2006.uprm.edu.	**San Juan, Puerto Rico, USA**

When	What	Where
27-29 September 2006	**Interactive Computer Aided Learning 2006** - This interdisciplinary conference aims to focus on the exchange of relevant trends and research results as well as the presentation of practical experiences gained while developing and testing elements of interactive computer aided learning. Therefore pilot projects, applications and products will also be welcome. The conference will be organized by the Carinthia Tech Institute in cooperation with the European Distance and E-learning Network (EDEN), the Federal Ministry for Education, Science and Culture of Austria, the IEEE Education Society, the Austrian Computer Society (OCG), the University of Klagenfurt and the FernUniversität Hagen. More information at: http://www.icl-conference.org.	**Villach, Austria**
13-17 October 2006	**E-Learn 2006** The World Conference on E-Learning in Corporate, Government, Healthcare, & Higher Education is an international conference organized by the Association for the Advancement of Computing in Education (AACE) and co-sponsored by the International Journal on E-Learning. This annual conference serves as a multi-disciplinary forum for the exchange of information on research, development, and applications of all topics related to e-Learning in the Corporate, Government, Healthcare, and Higher Education sectors. More information at: http://www.aace.org/conf/Cities/Honolulu.	**Waikiki Beach, Honolulu, Hawaii**
25-27 October 2006	**e-2006** (eChallenge conference) The goals of e-2006 are to stimulate take-up of applied ICT research results by industry and the European public sector, to promote knowledge sharing between industry, government and the research community, to exchange experiences about the current state of eAdoption at a sectoral, national or regional level, support International Cooperation and open up the European Research Area (ERA) to the rest of the world. The second day will see the presence of The Learning GRID SIG of Kaleidoscope organising a workshop on “Next Generation in Technology Enhanced Learning” together with the e-LeGI project under the umbrella of the Academy Industrial Digital Alliance (AIDA) of Kaleidoscope. More information at: http://www.echallenges.org/e2006.	**Barcelona, Spain**
28-31 October 2006	**Frontiers in Education Conference Borders: International, Social and Cultural** The 2006 Frontiers in Education Conference (FIE 2006) continues a long tradition of promoting the widespread dissemination of innovations that improve computer science, engineering, and technology (CSET) education. FIE is a major annual international conference devoted to improvements in CSET education. It is an ideal forum for sharing your ideas, learning about new developments in CSET education, and interacting with your colleagues. More information at: http://www.fie-conference.org/fie06.	**San Diego, California**