An Active, Ontology-driven Network Service for Internet Collaboration

David Lewis[1], Kevin Feeney[1], Thanassis Tiropanis[2], Simon Courtenage[3]

[1]Trinity College Dublin, [2]Athens Information Technology Centre, [3]University of Westminster

Dave.Lewis@cs.tcd.ie, Kevin.Feeney@cs.tcd.ie, ttir@ait.edu.gr, courtes@westminster.ac.uk,

Abstract

Web portals have emerged as an important means of collaboration on the WWW, and the integration of ontologies promises to make them more accurate in how they serve users' collaboration and information location requirements. However, web portals are essentially a centralised architecture resulting in difficulties supporting seamless roaming between portals and collaboration between groups supported on different portals. This paper proposes an alternative approach to collaboration over the web using ontologies that is de-centralised and exploits content-based networking. We argue that this approach promises a user-centric, timely, secure and location-independent mechanism, which is potentially more scaleable and universal than existing centralised portals.

1. Introduction

In the short time that it has been in existence, the web has revolutionised the way people find and use information. However, from the perspective of allowing users to gather information that is relevant to their work and way of working, the web presents problems in finding and keeping track of useful and timely information relevant to the user interests and user location. Bookmarks go stale, are often stored on one desktop machine and must be managed separately to the user's own information output, which may be spread over many physical locations.

Collaborative web portals provide enhanced support for users within the resticted domain of a collaborating community or an organisation. However, the growth of web-based communities outside of formal organisations (i.e. communities with no legal or financial basis), the emergence of virtual organisations and the growth of dynamic team-based working within organisations means that users increasingly have to manage their involvement in a number of separate web portals as well as with the publicly available Web.

Many of these problems can be traced back to the underlying architecture of the web, in particular, the client-server model of communication between web browser and web server, and to the presentational emphasis of document structure. The highly competitive web portal market also contributes, in that it encourages proliferation and differentiation of features within web portal platforms, making cross portal activities more complex and impractical for the user.

Problems of accurately in matching user information search requirements, i.e. queries, are being addressed by the Semantic Web initiative by the W3C (the main standards body for the Web). The aim of this initiative is to develop standards for the representation of the semantics of information such that they can be automatically processed. This offers immediate advantages in collaborative web portals, since potentially shared ontological models can be quickly agreed (or imposed) within the fixed communities they support and can then readily support information acquisition and maintenance [lara]. The semantic web initiative at the W3C has defined a standard description-logic based ontology language for the Semantic Web [owl]. However, the Semantic Web will remain essentially a large distributed data repository with little direct support for individual or collaborating knowledge workers in maintaining and growing the information relevant to them in a manner that is as flexible and adaptable as they themselves wish to be.

In this paper, we describe a flexible and scaleable architecture that allows communication and collaboration between users operating in fluid organisational environments. The aim is to make information on the web available to the user as a personalised, automatically maintained *information space*, which can be used as the basis for collaboration with others. This is in contrast to the centralised approach of existing collaborative portals, which hinders seamless involvement in multiple communities. It also contrasts with the simply structured data-space that WWW users perceive through their bookmarks, which requires constant updating, is not dynamically sharable and which is typically left bound to specific user desktop. However, this architecture does not aim to replace existing web sites and collaborative portals, but to complement them. It does this by providing users with a unified means for acquiring and maintaining knowledge on the web and for communicating knowledge within communities, but in a timely, accurate, secure and location independent manner. A holistic, Internet-scale approach is envisaged that makes efficient use of network resources.

The communications mechanism underlying the architecture is essentially a hybrid push/pull mechanism based on knowledge-driven content-based publish/subscribe. This brings together the accuracy, flexibility and extensibility of queries that is possible through the use of ontologies and the timely, efficient and adaptive delivery of information provided by Content-Based Networks (CBN). We discuss initial results in implementing this architecture, in particular in designing mechanisms for aligning open domain queries with CBN routing and in controlling access to information spaces in dynamic organisational environments.

2. Background

This section establishes the requirement for our architecture and described the relevant technologies upon which it builds.

2.1 User Requirements

Knowledge workers spend a significant proportion of their working hours acquiring and managing information that they assemble into their own skill and knowledge sets. In most professions, the development and progression of knowledge published on the web is increasing rapidly, so acquiring and managing knowledge efficiently is increasingly important. However, users are increasingly producers of information as well as consumers. Users who wish to provide information for the world at large are already well served by the existing web infrastructure, e.g. homepages and web logs. Of more relevance to most people, however, is exchanging knowledge within professional or social networks. These range from the networks established as part of working life, e.g. engineering teams, professional bodies, ongoing customer relationships, to ones that we establish in other areas of life, e.g. soccer team supporters clubs, support groups for health conditions, parent-teacher associations etc. Currently these groups communicate through an increasing variety of electronic means, e.g. collaborative web portals, chat rooms, bulletin boards, mail lists. To reduce the overhead of involvement in several such on-line communities, we aim to provide a common mechanism for receiving targeted notifications about current postings or conversations. User mobility requires directing adapted notifications and information delivery to the most convenient terminal device, e.g. an Internet kiosk in a public library, the voice mail box of a blind user or a teenager's SMS inbox. However, the user must remain in ultimate control of the information delivered to him or her through the knowledge-based specification of what information is relevant (and what is not) in different contexts and within different communities.

The integration of powerful handheld devices and ubiquitous wireless access at a variety of quality and cost levels liberates the now untethered user in deciding when and where to conduct their working, citizen, family and social roles. However, such freedom may also mean that the user is overwhelmed with demanding sources of attention, which must be navigated via terminals that have varying display, processing and communication resources. The proposed architecture therefore crucially complements the freedom brought by mobile wireless devices, allowing users to manage information spaces in the network so that relevant information is not just available anyplace-anytime, which can disruptive to the user, but delivered at the right place, at the right time and in the right presentation format.

The success of communications within any community relies on the trust that participants place in that community and therefore in the information they receive from it and in the use others make of information they provide. This trust can be built most directly by allowing the community itself to transparently define the rules under which information is posted and access to information is allowed, rather than this being the remit of separate security professionals. The proposed platform includes such a community-based access control mechanism that allows access control rules to be set in a number of ways, e.g. by mandating a trusted subgroup to do so, or by common consent.

In the longer term, the collaborative information spaces and tailored asynchronous information delivery supported by this architecture will present a key baseline technology in the roll-out of ubiquitous computing environments. The knowledge

acquisition techniques we propose have the potential to support the massive explosion of data sources, as these move from web sites presenting human generated knowledge to the output of sensors, actuators and processors embedded in the physical environment [crowcroft]. A semantic information space infrastructure will help filter and control the reception of the potentially vast amounts of information available from the ubiquitous computing environment to both mobile users and the agents who serve them. The collaborative aspect of the proposed architecture must therefore support the dynamic grouping of ubiquitous computing devices, agents and their users presents in any given ubiquitous computing environment at one time.

2.2 Technical Challenges

The challenge presented is therefore to make the content of the Web available to the user as a personalised, automatically maintained information space, which can be used as the basis for collaboration between knowledge workers so that many different patterns of work and collaboration can be accommodated. The information space is assembled and maintained for the user based on high-level, knowledge-based queries. This is done without the user having to be aware of the numerous different servers and portals on which is stored the information they need for a particular task.

The publish/subscribe networking paradigm allows users to make subscriptions to the network about the kind or type of information they wish to receive, while publishers describe to the network the type of information they want to make available or publish. Publish/subscribe is not a new paradigm in itself (for example, Tibco/Rendezvous implements a hierarchical channel-based publish/subscribe, in which messages are published to channels and users subscribe to receive what is published on a particular channel), but combined with content-based routing, it provides a powerful and expressive way for users to select the information they want to receive based on the properties of the information itself. A network that routes information according to its content makes routing decisions based on the (structured) content of messages, matching message content against user subscriptions in order to send messages to interested users. A typical application of this kind of networking is an event notification system, such as Siena [carzaniga], Elvin [segall] or IBM's Gryphon [strom].

By decoupling publishers and users, a content-based publish/subscribe approach to networking allows users to build information spaces independent of individual publishers, so that the information received by a user in response to a subscription represents a collation of information from a number of sources. It also allows virtual routing paths to be set up between publishers and the information received by a user, so that as new information is made available by a publisher, it can be routed immediately to interested subscribers. In effect, this allows a hybrid push/pull mechanism to implemented, in which the first time a user makes a subscription acts as the trigger to the pull phase by relevant publishers. Subsequently, new relevant information is pushed to the interested user.

Content based networking provides a mechanism for matching subscriptions to advertisements of notification types provided by an information source. Recent work

on the SIENA project [carzaniga] and the Cambridge Event Architecture [pietzuch] have demonstrated increasingly scaleable architectures for CBN in terms of the numbers of publishers and subscribers and their join/leave turnover. The specific applicability of CBN to mobile application has also been demonstrated. Our proposed architecture builds on this work by integrating it with the capabilities of the Semantic Web.

The vision of the Semantic Web entails providing rich meta-data for information present on the web, to it more amenable to machine processing and making the mean of its content more precise. This uses ontologies to specify meta-data types, instances of which represent specific elements of information. The Semantic Web envisages ontologies being developed and published at various level of abstraction by user communities and domain experts. The W3C has already begun standardising languages for the expression of ontologies, building on the Resource Description Framework (RDF) in developing the description logic-based Web Ontology Language (OWL) [owl]. We propose the explicit support of OWL for specifying mappings between related concepts in different ontologies to extend the currently limited expressive capabilities of CBN. Semantic mark-up of web content is already visible in the agreement and use of the RDF Site Summary (RSS) [rss] specification, that allows RDF statements of new web site or portal information to be transmitted on channels. These are collected by aggregator functions with appropriate channel subscriptions and filters which then provide users with dynamically updated information feeds. Though these mechanisms do not use CBN for efficient delivery, they demonstrate the effectiveness of Semantic Web meta-data to asynchronous information applications.

Access control of shared information on the web has traditionally been performed by access control lists. However these have proven inefficient in complex situations with large numbers of users, all of whom need individual management of their access control rights. Role-based access control introduces the grouping abstraction of a role against which rules or policies determining access rights are granted, e.g. nurse or doctor roles in a hospital information system [sandhu]. Though this makes managing access control for large numbers of users more efficient, it requires a static analysis of the roles present in an organisation, requiring skilled changes to this analysis to reflect any organisational changes. In the collaborating, internet-based communities we aim to support, groups will be created and dispersed rapidly to meet the needs of specific projects, groups may be autonomously self structuring and decision making will be more collective than in traditional hierarchical organisations. This is often driven by the lack of a fixed, high-cost infrastructure, which are the traditional drivers of centralisation and hierarchical patterns of responsibility. Our architecture must therefore support a mechanism of access control that is flexible enough to respond to dynamically changing organisation structures, while at the same time allowing integration with the information space infrastructure.

This makes the management of that information space more closely related to the knowledge worker's core activities. Instead of being just a list of URLs, the information space is an evolving map of ontological terms that express the user's knowledge acquisition requirements, both on an ongoing and a task by task basis. The interface also provides a single point for collaborating with multiple communities and

thus for managing links between related interest in those communities. The information space is therefore less vulnerable to the movement or removal of specific pieces of information and is automatically updated when new relevant information is made available. The approach also makes the delivery of information more robust, since information is delivered to the user at their convenience, in a manner appropriate to the device via which they currently have access.

The user's working preferences (e.g. presentational and navigational), task context and location context in the notification and delivery of information is used to address the need for automatic context-sensitivity, user profiling and personalisation. We therefore put the emphasis on the user, as both information sink and information source, being the centre of such adaptive behaviour.

The distinction with current collaborative information tools is that this system is built as a managed, application-layer overlay across existing information sources, so that information providers, at most, have to define access control rules for their published information, but are freed from uploading information to a specific server or subscribing to a specific collaborative portal for it to be part of a communal information space.

3. Architecture

In this section we outline an architecture that aims to:

- Securely and flexibly support the acquisition and dissemination of information between members of collaborative groups working across the Internet.
- To use ontology-based queries for defining the information being sought and shared, so that supported application domains reflects the growth of ontologies defined for the Semantic Web.
- To use content-based networking techniques to efficiently deliver information meta-data as well as to support user mobility and the maintenance of the architecture's semantic matching capability.

3.1 Overview

Broadly, the architecture consists of a content-based network connecting a number of different types of servers. Each server supports a different set of resource types, where resources are anything with an URI, the meta-data of which is described using RDF. Servers provide standard HTTP pull access to resources, and support push capabilities for resource meta-data. The pushing of resource meta-data is mediated by the content based network using the publish-subscribe paradigm. Permissions to advertise and subscribe to meta-data are addressed by the architecture. The access control to the actual information on the servers must be consistent with publish-subscribe

permissions, but the mechanisms for maintaining this consistency are beyond the scope of this paper.

The content-based network service we propose exhibits two novel features. Firstly subscriptions are dealt with as queries, so any resource meta-data matching a query-based subscription will be immediately pushed to the subscriber, while subsequent matches are pushed as they become available. Secondly, queries and the information meta-data against which they are matched are expressed using OWL. By using general purpose OWL reasoners to perform matching the CBN is not limited to any fixed domain of concepts and can encompass any domain for which OWL ontologies are available. Thus the CBN can be dynamically applied to any domain simply by making suitable ontologies available and publishing corresponding meta-data advertisements for domain information. For these reasons we refer to the CBN in this architecture as a Semantic Query Based Network (SQBN).

3.2 The User's Perspective

From a usage perspective, users of the service are free to organise themselves into communities into which they may bring resources. Information resources belonging to community members and access control of meta-data notification streams related to these resources are managed by the community (with the relevant members responsible for corresponding pull-access to information resources). Information resources may also be open, in that read access to the meta-data notification and the information is public. The primary resource managed by a community is an Information Collaboration Space (ICS). This is a set of semantic queries that defines a structured set of concepts, about which the community has an interest in acquiring and sharing information. These queries that are submitted to the SQBN on behalf of community members so that they receive the meta-data for the relevant information. This meta-data, called an information profile, includes the URI of the information resource.

A user community defines access control rules determining which users, within and external to the community, may publish and subscribe to their ICSs. A community is free to define sub-communities, specifying to which ICSs the members of that sub-community can publish and subscribe. Alternatively, the parent community may specify that the sub-community has authority to define the publish and subscribe permissions for specific ICSs, potentially related to communities the sub-community may itself later spawn.

3.3 SQBN Clients and Operation

The different types of server in the architecture are defined by the types of information they serve. Each server is a client of the Semantic Query Based Network either as a publisher, a subscriber or both as follows:

- Content Servers: These hold information resources and only publish information profiles, i.e. information resource meta-data, to the SQBN.

- ICS Servers: These hold ICS Profiles and publish changes to these. They may also optionally subscribe themselves to ICS profiles so that may log all matching information profiles. This supports users who may be restricted to pull access to the ICS server.
- User and Community Servers (UCS): These hold Community profiles, detailing resource access control rules and sub-community mandate profiles for each community. They also hold user profiles which track users' preference for interacting with the resources, their community membership and their current location. UCSs publish changes to community and user profiles.
- SQBN Edge Servers (SBQNES): These are essentially edge servers of the SQBN and mediate all interactions between users and the SQBN. They are responsible for verifying user identity and publishing user mobility events (connection to and disconnection from the SQBN). They subscribe to the user's User Profile, relevant Community Profiles and ICS Profiles. The user informs the SBQNES of the ICSs in which he or she is currently interested. The server subscribed to these ICSs on the user's behalf and is able to adaptively present these to the user, using its knowledge of of access link and terminal capabilities, combined with presentational preferences from the user profile.

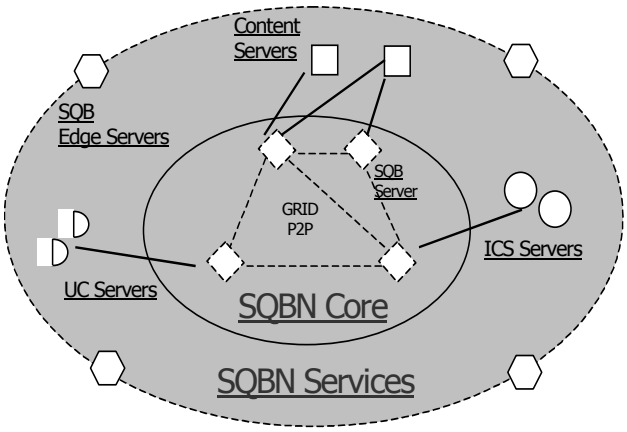

Figure 1: Network Layer Architecture

A SQBN is made up of Semantic Query Based Routers (SQBR) which form the SQBN Core to which any of the server types may connect. The other server types connected to the SQBN Core (including the SQBN Edge Servers) belong to the SQBN Service. The elements of SQBN are illustrated in Figure 1. We assume shared trust between all SQBRs and other server types. Therefore, given SBQNES verify user identity and community membership, publish and subscribe access control is implemented by including community membership credentials in relevant advertisements and subscription queries that the SBQNES injects on the user's behalf.

Community membership is resolved against community-related permissions as part of the process of matching subscription queries to advertisements.

The SQBN decomposes subscription queries to common, sub-expression based queries. These are disseminated across the network and matched against existing sub-queries to avoid duplication of routes. The integration of queries and the matching of queries to meta-data advertisements in general will require access by a SQBR to semantic matching function, typically co-located with the SQBR for efficiency. This will exploit the language features of OWL that identify class and property subsumption and equivalent and disjoint classes, properties and instances. Other, more sophisticated matching algorithms, based on ontology structure, may be employed as efficient implementations become available. The matching function will rely heavily on human authored mapping between ontologies. However, by allowing the semantic matching functions to act as SQBN clients themselves, they may subscribe to the publication of ontologies and ontology mappings relevant to the concepts that are currently required to map. In this way the range of ontological information the semantic matching function must store is tailored just to that it currently needs. In addition, the publication of new ontological information is automatically delivered to all the semantic matching functions that require it.

3.4 Operational Scenarios

This section aims to illuminate the operation of the architecture through examining scenarios where an ICS is established, where a document is published, where a user subscribes to an ICS and where a user moves around the network.

3.4.1 Setting up an ICS Profile

A community dedicated to celebrity gossip has an account on an ICS server with which it creates an ICS Profile, which it names ‘goss-ics’. This profile contains the terms in which the community has an interest as well as restricting the subscribers to group members only and publishers to another know group of reliable sources. E.g.

```
{?X rdf:about celeb:footballers.
AND ?X rdf:about celeb:singers.
AND ?X relations:marriedTo ?Y.
AND ics:publisher ics:memberOf ‘goss-group’.
AND ics:subscriber ics:memberOf ‘reliableSources’.}
```

Where ‘celeb’ is an ontology defining concepts representing current celebrities and ‘ics’ is ontology providing semantics of query predicates specific to our scheme. An ICS Profile can contain multiple queries and can be subscribed to as a whole. An ICS profile can specify community-based access control constraints that apply to all its constituent queries (i.e. will be ANDed to each query). An open ICS access policy would therefore give default values of:

```
{ics:subscriber ics:memberOf ics:world} and
```

```
{ics:publisher ics:memberOf ics:world}.
```

An ICS can also contain logging rules. If logging is activated then the ICS server itself submits a query corresponding to the ICS Profile. Logs are themselves resources, so changes to log can be subscribed to (subject to ICS access rules), or the ICS server could provide pull access for users unable to connected to the SQRN.

3.4.2 Publishing a document

A user puts a document on a content server and then publishes the corresponding meta-data as an information profile to the nearest SQBNES. The information profile must include the publisher's group membership credentials, which are used to match any ICS profile restrictions within the SQBN. The SQBNES must validate any credentials before injecting the profile advertisement into the SBQN. It is up to the information publisher to ensure that the pull access restriction for the actual information resource on the content server is not more restrictive than that defined in any matching ICS profile.

3.4.3 Subscribing to an ICS Profile

A SQBNES will obtain, for a suitably authorised user, their relevant ICS profile and allow them to select which they want to subscribe to at any one time. SQBNES get selected ICS profile and subscribed the user to them. Again, the SQBNES has the role of verifying user credentials against submitted ICS Profile subscriptions. The SQBN then routes matching published ICS profiles to the user via the SQBNES. The SQBNES may perform an aggregation function presenting the information to the user in form dictated by user preference as well as link and terminal capabilities which the SQBNES is able to monitor. For instance the SQBNES may be able to pass user notification messages to an application on the user terminal, or it may just generate web pages organised to the users preferences, which the user simply accesses via a browser on his/her terminal. This is analogous to how some ISP or search engine service providers allow management of personal home pages using RSS channels.

Note that when a user connects to a SQBNES as a publishers and/or subscriber, they may know the URI of ISC Servers and User and Community Servers where they have accounts or profiles. However, in more advanced roll-out scenarios where ISC, User and Group profile may migrate between servers to support reliability and efficiency, the SQBNES can locate this information with a suitable subscription for profile information related to this user. This therefore allows the SQBNES to quickly react to community membership changes, granting and revocation of permissions and ICS profile changes.

3.4.4 User Mobility

Wireless network technology allows use of multiple access networks, potentially from different network providers, each hosting their own SQBNESs which are part of the global SQBN. When a user roams from one access network to another, they may therefore connect to a new SQBNES. This subscribes to the user's User Profile which

gives group memberships and ICSs of interest, profiles of which the SQBNES also subscribes to. The SQBNES may also subscribe to any SQBNES that are serving that user currently thus proving the potential for co-ordinating access link load balancing or other adaptive sharing of ICS profile notifications. Suitable levels of trust between such SQBNES means that user identify verification can be shared between them, allowing the user to avoid multiple authentications.

Users can disconnect from a SQBNES explicitly or be timed out according to rules in the user profile, and/or SQBNES operator policies.

4. Business Models and Roll-out Scenarios

The rollout of an ICS services requires the identification of business roles and the specification of the business relationships (obligations and rights) among them. It is also essential to clearly identify the formation of value chains in ICS service provision.

Three business roles can be initially outlined:

- The SQBN Provider; which is the entity responsible for the operation of one or more SQBRs that form the core of the dissemination network. SQBN Providers establish agreements with each other and they have to ensure efficient and secure interaction among their SQBRs. Since SQBRs are infrastructural elements of the proposed architecture, ideal stakeholders to implement this role could be Internet Service Providers (ISP) or various access network providers, building on existing federation and settlement arrangements.
- The Information Providers; which is the entity that will provide and maintain information content servers to be made available to the SQBN. There is a large number of such providers already in operation. They will have to be connected to a SQBR and to establish business relationships with a SQBN provider for that purpose.
- The Information & Collaboration Space (ICS) Provider; identifies the business role responsible for the efficient management of ICS to ensure timely delivery of relevant meta-data (integrated from information sources) to the users respecting their user profile preferences. ICS providers maintain business relationships with SQBN Providers for access to the dissemination network.

The maintenance of user and community profiles could be a responsibility of an ICS provider or a specific User Profile (UP) Provider. Alternatively the maintenance of user profiles could be a sole responsibility of the users.

Business relationships need also to be established between wireless/fixed access network providers and SBQN Providers to provide users with access to their information and collaboration spaces anywhere anytime.

We envisage a number of evolutionary scenarios where an information space service may be commercially deployed, including the following:

- Intra-domain ICS service: Large knowledge-based organisations, such as multi-national technology firms and universities, could deploy an ICS service to support the exchange of information and collaborative working within the organisation as an alternative to a centralised collaborative web portal. Open interfaces will allow this service to interwork with those later deployed at collaborating organisations or made available as commercial services.
- Collaborative ICS services: Commercially available services managing information spaces specifically for collaborating individuals and organisations.
- Public ICS services: A range of competing ICS services are operated by different providers, who use open interfaces to pass semantic routing information between users, communities and information sources, with associated financial settlements for roaming users. Global mobile access to information spaces is thus supported.
- Ubiquitous computing infrastructure: The ICS service becomes universally supported by Internet service providers and private network operators and becomes an assumed and invisible part of the intelligent, ubiquitous Internet.

We envisage open GRID standards are being a promising means for deploying initial implementations the various servers defined in the architecture, subject to a suitable profile being established for asynchronous web service interactions.

5. Initial Results

Initial results in support of the proposed architecture include work toward a suitable subscription query language and an initial roll-out of the community-based access control scheme.

The query language at the heart of the SQBN in is based on FEL (Functional Event Language) [courtenage]. FEL is a language for content-based publish/subscribe networks and is used for building subscriptions that collate information into a composite structure from a variety of sources using a set of simple operators. The structure of expressions in FEL relates directly to the structure of the composite information space being built, which allows routing paths across the content-based publish/subscribe network to be derived automatically. The structure of the original information is therefore preserved within the composite information structure. Currently, FEL is being used to build a pure event notification system, and as such, the operators used to build composite information spaces are temporal operators, which relate to when one event occurs in relation to another. However, we are developing FEL further to include other operators relating to adaptive data collation and selection.

The limitations of role-based access control, restricts its use in fluid community structures so we have endeavoured to establish a more flexible abstraction for policy-based access control. This has resulted in a model for community-based policies, where communities, rather than roles, become the central abstraction use to specify access control. The power of this approach is that a community may mandate authority to both access resources and to author new access control policies for certain resources to a sub community. In this way the authoring of policies is distributed through a organisation to the group considered best qualified to make those policies. In this way an organisation can organically grow and change, perhaps starting from a single group with all authority, and decomposing into subgroups with mandated authority as the need arises naturally in the lifecycle of the organisation. Conflicts between policies authored by separate groups are automatically reported to the nearest mutual parent of the two groups. This group, having mandated policy authoring authority, is the best placed to decide how to resolve any resulting conflict.

The semantics of the community-based policy abstraction and of how authority is mandated to sub-communities has been established [feeney]. A community policy management system has been implemented using Ponder, an existing policy based framework [sloman], and found to operate satisfactorily. This implementation has been applied to the access control of a CVS code repository being used by an internet community and usage tests with that community are about to commence. This will provide us with initial user acceptance results with respect to the use of community-based access control

6. Related Work

In comparing this architecture with existing semantic web portal approaches it is useful to use the evaluation scheme described in [lara]. The architecture thus is similar to existing web portal in that information access is via HTTP, however, the user interface requires sufficient intelligence on the user terminal to handle asynchronous notification and interaction with user profile and adaptive presentation of information via aggregator functions. Related to the information processing element of the evaluation model, our architecture essentially allows creation, publication maintenance of information can be kept local to servers more convenient to the information publishers, while the organisation and access of information become a de-centralised, user-centric feature of the ICS service. The information and collaboration spaces are the core mean of collaboration through information exchange. However the service can also be used to notify users of relevant audio/video conference, chat session, bullet-ins board posting or other prompts for collaborative communication, some of which may be best conducted outside the SBQN, e.g. a/v streaming, HTTP pulls. In terms of grounding technologies our architecture is similar in its use of HTTP and ontologies, but essentially add the use of content-based networks.

Our architecture differs from existing CBN approaches, in that it addresses the scaleability of the application domain for content-based networking. By basing the

matching of advertisements and subscriptions on ontology-based semantics the architecture is open to any application domain with published semantics without the need for modification. New published ontologies will be automatically disseminated. The proposed approach also differs in attempting to combine query decomposition with CBN subscriptions in building up routing paths through the network. The use of ontology-based semantics makes the querying aspect potentially very powerful, as it is not restricted to meta-data that has been specifically modelled for a CBN. This approach therefore provides a comprehensive information space maintenance service through combination of query-based pull and change driven push mechanisms.

There has been little work in providing access control to CBNs with the notable exception of [belokosztolszki]. However this approach, together with most approaches to web pull access control adopt a role based approach. However, roles are statically defined artefacts requiring skilled analysis in their identification and modification. Our novel approach to this problem is to use the community as the base abstraction for access control. This allows us to better reflect the increasingly dynamic, flattened, team-oriented structure of organisations, a trend particularly visible in Internet communities. By aligning the authority to define access control rules and spawn subgroups with the natural group structure of organisation, the need for role design is avoided, and the community model used tracks the natural evolution of organisation structure.

7. Conclusions and Further Work

In this paper we have proposed an architecture that addresses the needs for finding and maintaining information for mobile users in collaborative settings. Our approach is a novel combination of content-based networks, ontology-based queries and community-based access control. In concert with broadening CBN applicability, this architecture provides a route to a secure, seamlessly managed collaborative communication infrastructure for mobile users and the agents that serve them.

The proposed architecture raises several issues that require further investigation in order to assess usability and scalability of this architecture for deployment on the Internet.

We must perform a more detailed assessment of the performance possible with existing ontology-based matching algorithms, though in the long term we expect this to benefit from the kind of optimised software and hardware support for OWL that has already emerged for XML processing. One possible optimisation that will reduce the reasoning load on SQBR will be to decompose ICS queries based on known routes prior to submitting them as subscription queries to the SQBN.

We plan to extend the community based management scheme to express resources and rules in terms of ontologies following emerging semantic web rule language proposals. This will allow use to the same semantic specification of information resources as the rest of the architecture, which should in turn improve the detection of access control policy conflicts.We also plan to examine the use of semantic web service definitions [mcilraith] as part of meta-data available to information spaces.

This would allow users to be notified of new or changing services, as well as of instances of a service, e.g. a voice over IP call from a particular colleague.

8. References

[bazinette] Bazinette, V et al "An Intelligent Notification System", IBM Research Report, RC 22089, June 2001

[belokosztolszki] Belokosztolszki, A., Eyers, D.M., Pietzuch, P.R., Bacon, J., Moody, K,. "Role-based access control for publish/subscribe middleware architectures". In International Workshop on Distributed Event-Based Systems (DEBS03), ACM SIGMOD, San Diego, CA, USA, 2003. ACM.

[carzaniga] Carzaniga et al, "Design and Evaluation of a Wide-Area Event Notification Service", In ACM Transactions on Comuter Systems, Vol. 19(3), 2001, pp. 332-383.

[courtenage] Courtenage, S. "Specifying and Detecting Composite Events", In 1st International Workshop on Discrete Event-Based Systems, Vienna, 2002

[crowcroft] Crowcroft, J., Jean Bacon, J., Pietzuch, P., Coulouris, G., Naguib. H., "Channel Islands in a Reflective Ocean: Large-Scale Event Distribution in Heterogeneous Networks", IEEE Communications Magazine, 40(9):112-115, September 2002.

[feeney] Feeney, K., Lewis, D., Wade, V., "Policy based Management for Internet Communities", Proceeding of the 5th IEEE International Workshop on Policies for Distributed Systems and Networks, 7-9 June 2004, York Town Hieght, NY, USA, to appear

[lara] Lara, R., Han, S-K., Lausen, H., Stollberg, M., Ding, Y., Fensel, D., "An Evaluation of Semantic Web Portals", IADIS International Conference on Applied Computing 2004, Lisbon, Portugal, March 2004.

[mcilraith] McIlraith, S.A., Son, T.C., Honglei Zeng, H., "Semantic Web Services", IEEE Intelligent Systems, 16(2), March/April 2001

[owl] OWL Guide, W3C, http://www.w3.org/TR/owl-guide/

[pietzuch] Peter R. Pietzuch and Jean Bacon. Peer-to-Peer Overlay Broker Networks in an Event-Based Middleware. In H. Arno Jacobsen, editor, Proceedings of the 2nd International Workshop on Distributed Event-Based Systems (DEBS'03), ACM SIGMOD, San Diego, USA, June 2003. ACM.

[rss] "RDF Site Summary Tutorial" http://www.mnot.net/rss/tutorial

[segall] Segall, B. et al, "Content-Based Routing in Elvin4", In Proceedings AUUG2K, Canberra 2000.

[strom] Strom et al., "Gryphon: An Information Flow Based Approach to Message Brokering", In International Symposium on Software Reliability Engineering 1998

[sandhu] Sandhu, R.S. et al., "Role Based Access Control Models", EIII Computer, vol. 29, no.2, 1996, p 38-47.

[sloman] Sloman, M., Lupu, E., "Security and Management Policy Specification", IEEE Network, vol.16 No. 2, March/April 2002. p 10