Review Report: The QoSDREAM Project

EPSRC Grants GR/M25872/01 and GR/M76195/01

Author: George Coulouris

Contents

Note to readers

This is a slightly modified version of a report that was submitted to meet the requirements of the projects sponsors. It is released for public use because it provides a reasonably concise summary of the project's current status. The somewhat idiosyncratic section headings reflect the sponsors' requirements.

Summary

This is an integrated report on two UK Engineering and Physical Science Research Council (EPSRC) grants bearing the short name QoSDREAM. The project was originally proposed in two parts to enable work to be carried out at separate institutions (Cambridge and Queen Mary, London), It was subsequently managed and operated in an integrated fashion at Cambridge following several staffing changes and the transfer of Prof. George Coulouris from Queen Mary to Cambridge.

The key achievements of the project are:

1. The development of an architecture for the construction and reconfiguration of distributed multimedia applications which enables consistency and quality of service constraints to be modelled and managed dynamically. This achievement relates to objective 1 of the original proposal.
2. The development and demonstration of a technology-independent approach to the handling of location information derived from a variety of location-sensing systems and its presentation to applications in a simple, unified form, with reference to a simple world model. This achievement relates to objectives 2 (as modified) and 3 of the original proposal.
3. The construction and internet distribution of a software framework to assist in the development of location-aware and multimedia applications. The framework incorporates the key achievements above and several other significant advances. This achievement relates to all of the objectives.

1. Chronological Overview

Staffing changes and their consequences

The project was originally conceived by Tim Kindberg and George Coulouris at Queen Mary London and John Bates in the Laboratory for Communications Engineering (LCE) at Cambridge. Their motivations were:

- To address a perceived need for software to support the development and management of applications that change in response to external events (sometimes called reactive applications).
- To integrate expertise in location-aware and sentient computing systems available in the Cambridge lab with the experience in the construction of object-oriented distributed systems and the preliminary work done on distributed multimedia at Queen Mary.

Following the award of the grants, Tim Kindberg left Queen Mary for an industrial lab in the US and George Coulouris retired from his full-time position there. The Distributed Systems group at Queen Mary was reduced to a size that was judged too small to provide a sound basis for the project and Prof. Coulouris was offered a part-time position in Cambridge in order to lead the project in partnership with John Bates. With the agreement of Queen Mary and the EPSRC, the Queen Mary grant was transferred to Cambridge and the project commenced in May 1999 with the appointment of two research associates (Spiteri and Mitchell) in the LCE.

The intended industrial partner, Acorn, had ceased trading early in 1999. This left the project with a need for a real-time operating system suitable for the implementation of distributed multimedia applications with QoS guarantees. The team undertook a substantial survey of the available commercial offerings [9] before concluding that none of them addressed those requirements adequately. However, SRT Linux, a recently-completed research development of Linux at Cambridge, was identified as meeting most of the requirements.

During the project's first year, John Bates became a founding partner in Apama, a start-up company whose business is the production and marketing of real-time event-processing software. This led to his departure from the LCE at the end of the project's first year and the subsequent departure of both of the RAs originally appointed.

Fortunately, the project was able to recruit another suitably-experienced RA, Hani Naguib, who joined the project in September 2000 and remained with it until its termination in November 2001. His appointment was judged sufficient, with an six-month extension to the project's duration, to complete the project, with the modification to objective 2 already mentioned.

It may be considered noteworthy that although all of the RAs had completed three or more years of PhD studies at the times of their appointments, none had completed their doctoral dissertations. Spiteri and Mitchell completed them during their appointments and were successfully examined. Naguib is completing his with support from the LCE's industrial funds following the project's termination.

We were fortunate in obtaining contributions to the project from a substantial number of students. This has included design studies and software components from four PhD students (Wharton, Katsiri, Rehman, Samugalingham, several of whom are EPSRC funded) and five summer interns (funded from industrial contributions to the LCE).

Work programme

Although scheduled as a two-year project, it was extended by six months to accommodate staffing changes. We report here on the progress of the work in two fifteen-month periods.

Substantial changes to the work programme set out in the proposal occurred due to staffing changes and insights gained during the project.

The first fifteen months

Much of the first year and a quarter was occupied with a hospital requirements study and with several preliminary investigations and prototypes on techniques for achieving the project's objectives.

An aim of the original proposal was to develop a hospital application as a testbed for the technologies that were to be investigated and developed. To this end, we carried out a detailed study of the work activities in an Accident and Emergency section of the Royal London Hospital (RLH) [1]. Information from this study did inform the design for the software framework developed by the project and a mock-up for a hospital-based follow-me video application was developed and demonstrated at a press open day, resulting in several reports in the media. However the plan to develop a hospital-oriented application was dropped for two reasons:

1 The requirements identified in our hospital study were quite basic – mainly for simple applications of location-sensing. There was no clear indication on the requirements for improved communication (audio and video) and we perceived a high risk that any attempt to introduce new technology could be counter-productive (indicated by the failed deployment of several hospital-oriented systems).

2. Although the clinical staff at the RLH cooperated well with our study, it was not clear how the cooperation of the administrators of the hospital could be obtained for a more intensive collaboration or a deployment. Nor were the project's resources judged to be sufficient for such a deployment.

Instead, a testbed application was developed for use within the LCE and this proved sufficient to enable an evaluation of the QoSDREAM framework to be completed. Subsequently, discussions were commenced with another potential end-user company and with several industrial research groups about the deployment of the QoSDREAM technology. These discussions are continuing and some potential partners are listed in Section 3 below. We have analysed the requirements in detail for an application scenario based on a broadcast TV studio [2] and more recently we commenced a study with British Airways of the requirements for location-aware multimedia applications for the Heathrow Terminal 5 project. Both of the application domains appear promising as targets for the application of QoSDREAM.

Another development in the first year and a quarter was the design and construction of a preliminary version of the proposed programming framework for multimedia applications. The framework adopted the two-level approach (composed of model and active components) first proposed in the precursor Djinn project at Queen Mary [7]. The new framework included the transactional support referred to in our objectives. A simple video conference application was constructed and used as a testbed. The purpose of the transactional support is to enable complex configuration changes to be constructed in the model layer and checked for consistency (using rules embedded in the model) before it is implemented in the active layer. If any of the checks fail, the model is rolled back to its previous state. This framework was then extended to include the automatic scheduling of 'smooth' reconfiguration updates to the active layer in order to minimize 'glitches'. The algorithm for this forms a substantial part of Mitchell's doctoral thesis [2].

Towards the end of this period, the design of a framework for handling location events (derived from location systems such as the Active Badge system deployed in the LCE) was studied in detail. An architecture was proposed that could accommodate a wide variety of location inputs (coming, for example from the active badge system, a vision-based location system developed in the LCE or from sources such

as infra-red beacons or data derived from knowledge of current logins). A preliminary implementation was constructed and demonstrated in a simple location-tracking application.

The final fifteen months

Despite an almost total turnover of staff towards the end of the first half, the second half of the project was quite fruitful. At around the halfway point, we decided to give priority to objective 3 – to complete the application framework and evaluate it in as far as could be achieved in the time available.

This was achieved in the form of a total re-design and re-implementation of the two preliminary frameworks as a single integrated framework together with a simple location-aware application for use within the LCE. The latter was instrumented to give numeric data on the reduction in event rates that could be achieved by the location information processing algorithms included in the framework. The key features of the resulting framework and the evaluation results are described in a series of three papers [4, 5, 6] published in international conferences in the second half of 2001.

Fortunately, the architectures for the multimedia and location management frameworks that were designed during the first half of the project proved basically sound, modulo some significant oversights and over-simplifications that had to be addressed in the re-implementation. Nevertheless, the redesign and re-implementation of the framework and the associated testbed application that was carried out by Hani Naguib with some student assistance, can be seen as a major achievement which enabled the project to be brought to a fruitful conclusion with quantified results.

The opportunity to re-engineer the software was taken and used to produce a framework with a highly modular structure, employing standard APIs wherever possible and this is seen as an important step towards its further development and possible exploitation.

2. Key Advances

Space does not allow us to report the technical results of the project in detail here. Most of the advances are described in the three recent papers already cited. Further publications are planned. Here we list some of the main results.

Two-level architecture for multimedia application frameworks

The basic idea is outlined in Section 2 above. A design and implementation for this were completed and validated in some simple applications developed during the project. This advance enables:

1. the ability to construct and evaluate changes to an application without disrupting its operation, employing the transactional approach mentioned earlier.
2. the ability to employ a reconfiguration-scheduling algorithm to minimize glitches on reconfiguration.
3. the integration of QoS management algorithms in the context of the application modelling layer.

Analytic solutions can in principle deliver near-optimal quality of service, but this has been achieved only in small-scale ‘proof-of-concept’ experiments and there is reason to believe that its achievement in larger-scale applications is likely to prove unrealistically complex. Despite this, the model-based approach to the construction of distributed multimedia applications developed in the project is considered a success because of the simple, high-level and intuitive model that it offers to application builders and its support for more limited consistency checks.

Region-based location modelling and its effectiveness in integrating diverse location sensing technologies

The key idea here is to adopt a simple world model that is adequate for the represention of the information deriving from a wide range of location sensing technologies and for presenting the information to applications in a concise and relevant form. The model contains just three abstractions:

Locators; which represent objects whose location can be determined by a given location technology. Active badges and Bats are examples of Locators.

Locatables; these represent objects whose locations need to be tracked. Examples include people and equipment. Locatables must have at least one associated locator, so that their location can be inferred from its locator.

Regions; Which include both static (geographic) regions and regions that define various location-specific characteristics of Locatables. For instance in the case of a person, they may include in addition to the person's location, their visual and audible fields.

The model leads to a hierarchy of location event types and algorithms to convert low-level events to higher-level ones. The key algorithm is a geometric analyser that combines the raw location events derived from sensors to generate 'region overlap events'. These are then filtered, adapted to the requirements of specific applications and distributed to applications according to their requirements. In some measurements on the testbed application, a ratio of 50:1 was observed between the arrival rate for raw sensor events and the rate at which relevant overlap events were generated and passed to the application.

This simple model seems to be adequate for handling a variety of location sensing technologies (we have carried out preliminary testing with five sources of location data and considered the requirements of several others) and can produce high-level representations of location data suitable for a wide range of location-aware applications.

3. Research Impact and Benefits to Society

Substantial interest in the QoSDREAM software emerged at three international conferences in late 2001. Intel Research, the Xerox Research Labs in Cambridge and Grenoble and British Airways are amongst the organisations that have expressed firm interest in deploying the software.

The AT&T Cambridge Research Lab. has taken an active interest in the research, including several interactions with their sentient computing research team. Further development of QoSDREAM will be funded for several months from their general grant to the LCE.

Preliminary discussions have taken place with Apama Limited concerning the joint development of a software product based on QoSDREAM.

Seminars:

During 2000 George Coulouris presented the project at two of the DTI/EPSRC Software Technology Outreach Programme workshops. He concluded that few of the attendees at these workshops were looking for software tools of the type being developed in the project.

Seminar presented by Hani Naguib for Opera Group, Computer Laboratory, Cambridge University, November 2001. Seminar presented by George Coulouris at Microsoft Research Lab, Beijing, December 2001.

4. Further Research and Dissemination

The project's web site (http://www-lce.eng.cam.ac.uk/qosdream/) offers the software for download under an open-source licence, with supporting documentation.

The preparation of further publications is in hand. A contribution has been submitted for a Special Issue of the ACM Multimedia Systems Journal. We hope to refine our evaluation and to publish a further journal paper based on the results.

We are actively working on a joint proposal with Dr. Jean Bacon and Prof. J. Crowcroft to explore protocols for large-scale wide-area event distribution. The experience and the software from this project will be fed into this new project.

We intend to explore the potential for developing the QoSDREAM framework into a research platform (in the sense defined in the recent International Review of UK Research in Computer Science [8]). The software embodies many of the results of about five years' work by a significant number of researchers, and the response to its presentation at international conferences suggest that there would be substantial interest in its continued development and support.

References

All except item [8] can be found at:

http://www-lce.eng.cam.ac.uk/qosdream/publications/publications.htm

1. QoSDREAM Report, Hospital User Study, *web publication only.*
2. Robert Scott Mitchell, *Dynamic Configuration of Distributed Multimedia Components*, Ph.D. Thesis, University of London, August 2000.
3. Scott Mitchell, Mark D. Spiteri, John Bates and George Coulouris. Context-Aware Multimedia Computing in the Intelligent Hospital, In *Proc. SIGOPS EW2000, the Ninth ACM SIGOPS European Workshop*, Kolding, Denmark, September 2000.
4. George Coulouris, Hani Naguib and Scott Mitchell, Middleware support for context-aware multimedia applications, In *Proc. DAIS'2001, The Third IFIP WG 6.1 International Working Conference on Distributed Applications and Interoperable Systems*, Krakov, Poland, September 2001.
5. Hani Naguib, George Coulouris.Location Information Management, In *Proc. UBICOMP 2001*, Atlanta, Ga. USA, October 2001.
6. Hani Naguib, George Coulouris, Towards automatically configurable multimedia applications, In *Proc. Multimedia Middleware workshop at ACM MM 2001*, Ottawa, Canada, October 2000.
7. Scott Mitchell, Hani Naguib, George Coulouris and Tim Kindberg, Dynamically Reconfiguring Multimedia Components: A Model-based Approach. *8th ACM SIGOPS European Workshop*, Lisbon, 7-10 September 1998.
8. Fred Schneider and Mike Rodd (Eds), International Review of UK Research in Computer Science, 2000. http://www.iee.org.uk/Policy/CSreport/cs_report.pdf
9. QoSDREAM Report, Realtime Operating System Review. *web publication only.*

18/2/02