

Institution: University of Derby
Unit of Assessment: 11
Title of case study: Knowledge Transfer of Innovative Cloud Computing Technologies
<p>1. Summary of the impact</p> <p>This case study reports our work on the development, application and dissemination of innovative cloud-based technologies to industrial problem domains. First, decentralised scheduling is implemented within federated Clouds, to facilitate the new drug discovery process for a global pharmaceutical company. Second, multi-objective approaches to the management and optimisation of video processing and analysis workflows in distributed environments is described in the context of an SME organisation that is developing new products, services and markets. Both of these examples have attracted, and continue to attract, commercial funding, and demonstrate the efficacy of knowledge transfer into industry from University of Derby (UoD) research.</p>
<p>2. Underpinning research</p> <p>The work commenced in 2003, in the context of the Large Hadron Collider Grid (LCG) based Data Analysis project as a collaboration with scientists at CERN, Geneva. Large-scale scientific experiments produce massive volumes of data that require resources for storage, analysis and transformation. The scientific community responded by developing Grid Computing architectures to harness the collective computational power and storage of many, discrete, economical, computational resources.</p> <p>Anjum's (who joined the UoD in 2011) PhD thesis in 2003, proposed a Data Intensive and Network Aware (DIANA) meta-scheduling model, to facilitate analysis in data intensive applications [3.1][3.2]. DIANA makes intelligent scheduling decisions about the analysis of the data, when the data is distributed in nature, and when the performance of local computational resources is limited. DIANA takes into account a cost based mechanism to map jobs, algorithms and queries against resources, when making scheduling decisions.</p> <p>In 2011, Anjum extended the core scheduling approach in DIANA to utilise pilot jobs to implement decentralised scheduling [3.4]. This not only reduces the queuing times that can be quite high in local schedulers, but it also reduces the instances of failed jobs. Moreover, this approach makes the decision-making process distributed, cooperative and fault tolerant, as scheduling decisions are optimised to make efficient use of resources. This process therefore significantly reduces the overall execution time of a workflow by reducing the scheduling and data access latencies in a decentralised way.</p> <p>Roche, a global pharmaceutical company, through an industrial partnership (2012-2015) with the UoD, has been exploiting this approach for the development of Cloud-based services for their clinical trial management system. Such services can be dynamically orchestrated and optimised as they are executed, enabling the clinical trials management lead times to be reduced and resources to be better utilised for their drug discovery process.</p> <p>Further development was achieved during 2013 [3.3], as DIANA was augmented to use machine learning algorithms to optimally plan workflows prior to scheduling. The potential for applications to efficiently manage and utilise resources was illustrated when [3.3] demonstrated a multi-objective approach to the management and optimisation of scientific workflows in distributed environments.</p>

The approach is particularly relevant to real world applications, specifically those where the evaluation of objectives may be computationally expensive and an extensive evolutionary search may not be feasible [3.5][3.6]. Such an example is the exploitation of this research by XAD Communications as 'Stream Cloud', through a Technology Strategy Board (TSB) funded Knowledge Transfer Partnership (KTP) with UoD in 2012.

The aim of this project is to capture and automatically process video streams for security and surveillance domains, by combining consumer-grade video cameras with real-time video stream analysis, using DIANA. Users who need to analyse video remotely have three options: (1) video files moved to the user's location; (2) analysis algorithms moved to the location of the video, or (3) both video and algorithms are moved to a location that has sufficient network bandwidth and computational resources. The DIANA model enables each option to be evaluated, and then the optimum schedule is constructed to minimise turnaround time.

During 2013, UoD was awarded the IT as a Utility (ItaaU), EPSRC funded project 'Video Analytics as a Service' that will develop the Stream Cloud software and applications into a Cloud Service for its use in communities beyond the security and surveillance domains. This project will make use of the augmented DIANA approach to find events of interests in video streams, by optimally processing and mining the data.

In addition to this, the work has been the subject of two PhD theses [5.3][5.4] and a number of high quality journal (6) and peer-reviewed conference publications (8).

3. References to the research

- 3.1. Anjum, A. et al. "Bulk Scheduling with DIANA Scheduler", IEEE Transactions on Nuclear Science, 53(6):3818-3829, 2006. DOI: 10.1109/TNS.2006.886047
- 3.2. McClatchley, R., et al, "Data Intensive and Network Aware (DIANA) Grid Scheduling", Journal of Grid Computing, Springer Verlag, 5(1):43-64, 2007. DOI: 10.1007/s10723-006-9059-z
- 3.3. Habib, I. et al, "Adapting Scientific Workflow Structures Using Multi-Objective Optimisation Strategies", ACM Transactions on Autonomous and Adaptive Systems (TAAS), ISSN: 1556-4665, 8(1), 2013. DOI: 10.1145/2451248.2451252
- 3.4. Hasham, K. et al, "CMS Workflow Execution using Intelligent Job Scheduling and Data Access Strategies", IEEE Transactions on Nuclear Science (IEEE TNS), ISSN: 0018-9499, 2011. DOI: 10.1109/TNS.2011.2146276
- 3.5. McClatchley, R. et al, "Intelligent Grid Enabled Services for Neuroimaging Analysis", Elsevier NeuroComputing, ISSN: 0925-231, 122(25):88-89, 2013. DOI: 10.1016/j.neucom.2013.01.042
- 3.6. Anjum, A. et al. "Glueing Grids and Clouds together: A Service-oriented Approach", International Journal of Web and Grid Services (IJWGS), 8(3):248-265, 2012. DOI: 10.1504/IJWGS.2012.049169

4. Details of the impact

The principal strategy of the School is to directly apply the outputs of research to industrial contexts. Using a combination of the School's expertise in Cloud Computing, together with specific research into the applicability of Grid scheduling techniques - such as DIANA - to Cloud Computing, the School has developed innovative architectures, algorithms and applications that

demonstrate the wide scope of this work.

With this in mind, two specific examples illustrate the breadth of impact that this work illustrates. Firstly, the extension of the DIANA system has been transferred to Roche, a global pharmaceutical company, in order to streamline the drug discovery process. This is achieved by using the decentralised scheduling extension to DIANA to enhance the existing Enterprise Knowledge Exchange workflow systems at Roche.

Evidence of Impact

- Knowledge transfer into Roche pertaining to the use of workflow scheduling to orchestrate cloud-based services [5.1].
- £55K Research and development funding contract with Roche.
- Employment of 1 full-time employee (FTE), with another 1 FTE to start in January 2014 [5.1].

The potential impact of this upon a global organisation such as Roche, which dedicates considerable resources to the management of intellectual capital and Enterprise Knowledge Management, is significant.

The second example utilises further development of the algorithms to pre-schedule video processing and mining in a distributed federation of compute Clouds. XAD Communications Ltd., is an SME that has partnered with the School of Computing and Mathematics to apply our knowledge and develop new products and services. The company is developing video surveillance products that make use of low-cost, consumer grade video cameras, as well as cloud-based video processing and streaming services for customers who already have a physical video capture infrastructure.

Evidence of Impact

- Knowledge transfer into XAD Communications Ltd. [5.5][5.9].
- £129K funding from the Technology Strategy Board.
- Employment of 1 full-time employee (KTP Associate), plus buy-out of time for UoD Academic (Anjum)[5.8]. Additional 1 FTE to be recruited by XAD Communications Ltd [5.2].

With regard to profit from existing customers, the margins are likely to increase. Currently the margins are quite narrow and they mostly come in the shape of add-ons that the company has developed, which are around £500 per server node in a video data centre. An important source of the increase in margins is less reliance on third party tools. The algorithms and cloud library is likely to enable the company to use their own data store and analysis platform instead of buying this from third parties.

The application areas that have already been identified have the potential for massive impact upon society. Traffic surveillance for pro-active management, and pre-emptive policing, all become more feasible as a direct result of being able to process large volumes of video content in real-time[5.5][5.6]. Similarly, application in the security domain can utilise the proliferation of consumer video cameras, which can be deployed much cheaply and widely.

The detection and correct attribution of evidence of crime will also be significantly enhanced as the technologies developed in Stream Cloud significantly reduce the manual labour required to analyse video footage [5.7]. Additionally, the development of associated Cloud-based services has the potential to facilitate the creation of new business models that utilise inexpensive video capture.

Additionally, subsequent follow-on investment from the IT as a Utility (ITaaU) network (EPSRC funding) of excellence, recognises the potential impact of this work by awarding a 0.5FTE academic secondment to UoD, commencing September 2013.

As a result of his successful engagement with these projects and the Grid Computing research community, Anjum was awarded the role of 'European Grid Initiative (EGI) Champion'. This role recognises expertise in the use of Grid Computing infrastructure, and in particular is a mechanism by which this knowledge is conveyed as widely as possible. To date, Anjum has engaged with the following dissemination activities:

- Delivered a talk to the biomedical and security community in the EGI community forum, Manchester, April 8-12 2013
- Reviewed the EC deliverable D6.9 (April 2013) that provides report on the Heavy User Communities Tools and Services.
- Co-Chaired (with Antonopoulos, Gillam) ITAAC 2011 and 2012 workshops, held in association with Utility and Cloud Computing (UCC) Conferences.
- Edited two journal special issues, (with Antonopoulos, Gillam), Journal of Cloud Computing: Advances, Systems and Applications (JoCCASA) in 2012 and 2013.
- Invited talk at ICOSST 2012, Lahore, The Role of Cloud Computing in Video Streaming Applications.

The following activities were agreed and arranged prior to 31st July 2013, and indicate further planned dissemination activity:

- Provided expert opinion to the EGI Technical Forum (September 16-20, 2013), Madrid, participants about the latest Cloud Computing trends in Biomedical and Security domain.
- Member of Expert Panel - Trust Management models and approaches (October 2013).
- Delivered a keynote talk, Kings College London (October 2013) to European Commission, Future of Clouds in the healthcare domain.
- Co-Chair (with Antonopoulos, Gillam) workshop, in association with Utility and Cloud Computing (UCC) Conferences (December 9-12, 2013).

5. Sources to corroborate the impact

- 5.1. Letter of support from Roche Global R&D, Basel, Switzerland.
- 5.2. Letter of support from XAD Communications Bristol, UK.
- 5.3. PhD Thesis, 2007, Data Intensive and Network Aware Grid Scheduling
- 5.4. PhD Thesis, 2011, Multi-objective Optimisation of Compute and Data Intensive e-Science Workflows.
- 5.5. Security System Project Is One Worth Watching, <http://www.derby.ac.uk/news/security-system-project-is-one-worth-watching#>
- 5.6. Experts say high-tech surveillance system is one to look out for, <http://www.thisisderbyshire.co.uk/Experts-say-high-tech-surveillance-look/story-16826678-detail/story.html#axzz2MitAyBjF>
- 5.7. Security system is in the cloud, <http://www.vision-systems.com/articles/2012/09/security-system-is-in-the-cloud.html>
- 5.8. KTP Proposal for Stream Cloud Project.
- 5.9. Letter of support from Dr Neil Grice, KTP Advisor, UK.