OPAALS

OPAALS PROJECT

Contract n° IST-034824

WP3: Autopoietic P2P Networks

Del 3.13 – Paper on evolutionary and interaction framework

Information Society Technologies

Project funded by the European Community under the "Information Society Technology" Programme

Contract Number: IST-034824

Project Acronym: OPAALS

Deliverable N°: D3.13

Due date: January 2010

Delivery Date: September 2010

Short Description:

This document introduces the evolutionary framework in a series of deliverables (D3.1, D3.2, D3.6, D3.10) that aim to describe the evolutionary frame of the infrastructural specifications in Digital Ecosystems research. As it has been shown in OPAALS the technological and social concerns in providing the necessary digital infrastructure are not treated as distinct but as part of the same continuum. This deliverable sets out the current ongoing relevant research on the evolutionary framework and offers a roadmap for future research on the area.

Author: Surrey, UNIVDUN

Partners contributed: Surrey, UNIVDUN

Made available to: All

Versioning		
Version	Date	Name, organization
0.1	**08/2010**	**A. Razavi (Surrey)**
0.5	**09/2010**	**A. Razavi (Surrey), A. Munro (UNIVDUN), P. Krause (Surrey)**
1.0	**23/09/10**	**P.Krause, N. Gilbert, R. Hoyle, J. McFadden, A. Razavi (Surrey)**
1.1	**26/09/10**	**P. Krause (Surrey)**
1.2	**02/10/10**	**P.Krause (Surrey)**

Quality check

Internal Reviewers: P. Dini (LSE); G. Briscoe (LSE)

Dependencies:

Achievements	Review of motivation behind species and cell-based approaches to the development of a theory of digital ecosystems. Successful funding bids for two UK EPSRC funded research programmes to take forward the study of complex adaptive systems, with particular reference to Industrial Ecosystems. Establishment of a joint research programme with a local data centre to build models of sustainable architectures for digital resource provision in industrial/knowledge economies.
Work Packages	WP12: Task 12.8 – OS principles of communication and collaboration, Task 12.9, 12.10 – Business models in DEs WP11: Task 11.1 – collaboration and innovation in the Knowledge Economy, Task 11.4 – social innovation networks WP10: Task 10.10 – visualisation of P2P infrastructure WP5: Task 5.6 – adding e-business support in current DBE, Task 5.7 – integration of P2P network services into the DE
Partners	TI, BCU, UniKassel, IPTI, NUIM, UL, IITK, TUT, UNIVDUN
Domains	Computer Science domain: P2P networks, interactions, long-running transactions, distributed systems and networks, redundancy, diversity, consistency, concurrency, design for failure, formal semantics, behaviour patterns, lock mechanisms, formal analysis, distributed identity, distributed trust. Social science domain: participation, power, control, technology infrastructure for competitive advantage, monopoly, lock-in, proprietary software / platforms, knowledge, openness, reciprocity, context-dependent trust, identity. Natural Science domain: simple reference to key analogies with ecosystems in nature and their key features found mostly in studies of biodiversity.
Targets	Computer, social and natural science researchers, SMEs, business analysts. Computer science communities: database, transactions, P2P networking and applications, formal methods. Social science: language, socio-economics, governance, power

	and control, identity and trust.
Publications*	Evolution and Resilience in Industrial Ecosystems, http://gow.epsrc.ac.uk/ViewGrant.aspx?GrantRef=EP/H021779/1 Models and Mathematics in Life and the Social Sciences, http://gow.epsrc.ac.uk/ViewGrant.aspx?GrantRef=EP/I000992/1 Krause PJ, Razavi AR, Moschoyiannis S, Marinos A, **Stability and Complexity in Digital Ecosystems**, 3rd IEEE International Conference on Digital Ecosystems and Technologies, Istanbul, TURKEY, 01 Jun 2009 - 03 Jun 2009. 2009 3RD IEEE INTERNATIONAL CONFERENCE ON DIGITAL ECOSYSTEMS AND TECHNOLOGIES. IEEE. 200-205. 2009 Kate Craig-Wood, Paul Krause, Nick Craig-Wood (2010) **Analysis of CPU & disk I/O utilisation in a real-world deployment of 950 virtual machines**. In *International Conference on Cloud Computing and Virtualization (CCV 2010)*. Kate Craig-Wood, Paul Krause, Aaron Mason (2010) **Green IT: Oxymoron or call to innovation**? In *International Conference on Green IT (Green IT 2010)*. Kate Craig-Wood, Paul Krause, Nick Craig-Wood (2010) **The case for medium-sized regional data centres,** GSTF International Journal on Computing, 1(1).
PhD Students*	Amir Razavi: Foundational work on adaptive DE architecture. Alexandros Marinos: Generative Business Information Systems as adaptive systems. Kate Craig-Wood: Data collection and analysis on performance and energy consumption of data centres. Aaron Mason: Analysis of energy consumption and CO2 emissions for data centres.
Outstanding features*	The setting up of the ERIE project as one of only three UK centres on *Complexity in the Real World*, is a significant outcome in terms of both sustainability and visibility of our work.
Disciplinary domains of	Prof. P. Krause (Computer Science) Prof. N. Gilbert (Sociology)

authors*	Prof. J. McFadden (Molecular Genetics) Prof. A. Munro (Molecular Oncology) Dr. R. Hoyle (Mathematics) Dr. A. Razavi (Computer Science) A. Mason (Computer Science) K. Craig-Wood (Computer Science) A. Marinos (Computer Science) N. Craig-Wood (Physics)

TABLE OF CONTENTS

EXECUTIVE SUMMARY

This report discusses the motivation behind continued use of biological systems at both the macro and micro level to inform the theoretical study of Digital Ecosystems as complex adaptive systems. It then presents a roadmap for continued research in this area over the next six years, with reference to two funded research programmes that are now active. We also present some results from a concrete intervention into the Digital Economy; the development of a medium scale regional data centre that has been set up as a model business that offers a more agile and less centralised approach to the supply of commodity computing.

AIMS & OBJECTIVES OF THE REPORT

The naming of the concept, *Digital Ecosystems*, and the use of the term *Autopoiesis* in its study both point to the use of biological systems at macro- and micro-scales as metaphors. This report provides a collection of papers that capture our current understanding of the usage of these metaphors in the study of the evolution of and interactions within digital ecosystems, and then outline our roadmap for the continued study of digital- and industrial-ecosystems over the next 6 years.

Chapter 1 of the report is the reproduction of a paper that was presented at the 2009 IEEE DEST conference. It focuses on resilience and adaptation in natural- and social-ecosystems. We explore the Holling Adaptive Cycle, which can be used to an extent to justify the need to maintain diversity in any complex adaptive system. The paper stops short of exploring some of the other issues that are important in the study of digital ecosystems. Can we, for example, obtain early warnings of impending transformations or regime shifts in digital ecosystems? And can we identify and put in place institutions and institutional change management processes that allow us to control the impact of, or perhaps even facilitate where the outcome may have a net benefit, regime shifts in digital ecosystems? Work by Brock and Carpenter (2009) and Ostrom (2005) are representative of approaches that may be relevant here.

Chapter 2 develops this work a little further, and then extends this to consider the lessons that may be learnt by extending the biological metaphor to the molecular level (the regulation of gene expression), and the cellular level. All this work is necessarily speculative at this stage, although theoretical work within the OPAALS community (Dini *et al*, 2010) and others (the late Robin Milner's work (2009) stands out in this regard) is beginning to demonstrate formal connections between interactive computer systems and biological systems theory. In order to strengthen these connections between modelling activities within the social and life sciences, the University of Surrey has secured funding for two major programmes of interdisciplinary research work. The plans for these are outlined in Chapters 3 and 4. ERIE, Chapter 3, brings together Mathematicians with Computer, Social, and Environmental Scientists in order to further the study of Industrial Ecosystems as complex adaptive systems (our use of the term Industrial Ecosystems being reflective of the extent to which the Digital Economy is now embedded in daily life). An extensive survey of foundational material on modelling socio-economic systems as Complex Adaptive Systems can be found in (Norberg and Cummings, 2010).

ERIE has some clearly defined objectives, although being a six-year programme, the work plan for the second half of the project is necessarily less precisely defined at this stage. MILES (Chapter 4) in contrast, is quite open with its objective of stimulating interactions across all the life and social sciences through intensely interdisciplinary studies of models and model construction processes. Although not specifically targeted at the Digital Economy, we expect to see many of its results informing the study of Digital/Industrial Ecosystems as complex adaptive systems.

Finally, we are beginning to engage with, and monitor some specific interventions that we hope will impact the continued evolution of the Digital Economy. We are working closely with a local data centre to study in depth the performance and cost effectiveness of medium scale regional data centres. Some

preliminary research findings are presented in Chapter 5 through the reproduction of two published papers. The first of these provides a detailed analysis of Memset's data centre architecture. The second is a more lightweight contribution to the literature on "Green IT". Although it does not go into a great deal of technical depth, it does capture some important issues in the current debate on "Green IT".

References:

W.A. Brock and S.R. Carpenter, "Interacting Regime Shifts in Ecosystems", *Ecological Monographs*, 2009.

P. Dini, Egri-Nagy, A, Nehaniv, C L, Schilstra, M J, Van Leeuwen, I, Munro, A J and Lain, S (2010). *Mathematical Models of Gene Expression Computing*, OPAALS Deliverable D1.4, European Commission, available at http://files.opaals.eu/OPAALS/Year_4_Deliverables/WP1.

R. Milner, *The Space and Motion of Communicating Agents*, CUP, 2009.

J. Norberg and G.S. Cummings, *Complexity Theory for a Sustainable Future*, Columbia University Press, 2008.

E. Ostrom, *Understanding Institutional Diversity*, Princeton University Press, 2005.

Support Publications

P.J. Krause, A. Razavi, S. Moschoyiannis, A. Marinos (2009) "Stability and Complexity in Digital Ecosystems". In *3rd IEEE Conference on Digital Ecosystems Technologies 2009* (DEST 2009).

K. Craig-Wood, P.J. Krause, N. Craig-Wood (2010) "The case for medium-sized regional data centres", *GSTF International Journal on Computing*, 1(1).

K. Craig-Wood, P.J. Krause, A. Mason (2010) "Green IT: Oxymoron or call to innovation?", *Proceedings of Green IT 2010*, Singapore.

1 STABILITY AND COMPLEXITY IN DIGITAL ECOSYSTEMS (SURREY)[1]

Although we have used a modification of the title of Lord May's landmark treatise in theoretical ecology [15], we do not intend any comparison to be made with regard to the potential impact of this paper. Rather, we use the title to highlight the importance of progressing from the use of the term "Ecosystems" as a metaphor in the study of Digital Ecosystems, to the development of a theoretical foundation of Digital Ecosystems. In this paper we will explore a number of theoretical and empirical advances in the study of natural- and social-ecosystems, and draw out a number of hypotheses that can be explored to assess their relevance to the study of digital ecosystems.

Boley and Chang [2] proposed a definition of a digital ecosystem, based on analogy with their own definition of ecosystem in the purely biological sense. In their view, a digital ecosystem was:

"an open, loosely coupled, domain clustered, demand-driven, self-organising agent environment, where each agent of each species is proactive and responsive regarding its own benefit/profit (…) but is also responsible to its system."

A key feature of this is that agents join freely and of their own volition. This is in contrast to a tightly coupled organisation in which the agents have pre-defined roles. The focus is very much on the autonomy of individual agents, and hence the global properties and institutions of such an ecosystem emerge (primarily) through self-organisation. This is a radical and potentially disruptive concept. However, we will argue later in this paper that current digital infrastructures constrain the emergence of properties and institutions, and instead impose such from outside the ecosystem.

To this end, we would like to work from a definition of an ecosystem that is at a higher level of abstraction. This is purely for the purposes of the discussion in this paper, and not intended as an alternative to the above definition, which we fully support.

We take as our starting point the Arthur Tansley's 1935 definition of an ecosystem as, [24]:

"*An interactive system established between living creatures and the environment in which they live*."

With regards to a digital ecosystem, we want to hold back any discussion around "living" for the moment, and so work with the following:

"*A digital ecosystem is an interactive system established between a set of active agents and an environment within which they engage in common activities.*"

Note that, consistent with this definition, it is possible for a given agent to interact in multiple digital ecosystems. "Agents" include (but might not be limited to) providers of software services, information

[1] This paper was presented at IEEE DEST 2009, and received a best paper award: P.J. Krause, A. Razavi, S. Moschoyiannis, A. Marinos (2009) Stability and Complexity in Digital Ecosystems. In *3rd IEEE Conference on Digital Ecosystems Technologies 2009* (DEST 2009).

sources, and human agents. The environment is a combination of a socio-economic context and a digital infrastructure. We will argue that the nature of the latter, the digital infrastructure, can impact (undesirably at present) on the properties that emerge in the ecosystem.

Our view is that the concept of *species*, even in natural ecosystems, has some difficulties. Nevertheless, it is difficult to engage in any dialogue on ecosystems without some reference to species, and it can be useful of course. Boley and Chang referred to species as “types of agents” [2]. We concur with that, but it still leaves open the question of how we define a taxonomy for digital ecosystems?

We don’t want to address that question in this paper. However, we would like to work with a fragment of a taxonomy for the purposes of discussion. In commerce, a business enterprise can be viewed as providing a set of services. In that such an enterprise provides a higher level of organisation to an individual service, we might consider (business) enterprises as species within an ecosystem. This equating of enterprise with species will be particularly useful when we come to discuss response diversity in digital ecosystems.

In the next section we illustrate the importance of innovation processes in sustaining ecosystems. After that we introduce some key concepts in ecological modelling, following [12]. We then outline some of the issues with current digital ecosystems before returning to the modelling of social-ecosystems. Finally, we conclude with some pointers to a research agenda for the science of digital ecosystems.

1.1 The K-T Extinction Event

The non-avian Dinosaurs and certain marine creatures were subject to a massive extinction event at the Cretaceous-Tertiary (K-T) boundary, about 65 Million Years ago. The current preferred theory seems to be that this was triggered by a massive meteorite impact at Chicxulub in what is now North-east Mexico. There is little doubt that this meteorite impact did cause widespread disruption to the global environment. However, it is less clear that this was the sole “impactor” on the inhabitants of the earth. Even more importantly, evidence is emerging that this did not even coincide with a period in which there was a net loss of biodiversity on the planet. Rather, the reductions in certain populations were being countered by ongoing speciation activities in other groups of species.

The first point to note is that the meteor impact was not the only major event that was effecting the physical environment of the time. In addition, the K-T boundary of 65 million years ago marked the final stages of the break up of the supercontinent Pangea into the broad continental structure that is now seen. This was a gradual process that started some 225 million years ago. The period from 135 million years ago (Jurassic Park time!) to 65 million years ago saw the separation of the plates that form North America and Eurasia, and the plates that form South America, Antarctica and Australasia. As well as this fragmentation of what were formerly massive continental areas, the continuing plate movements were beginning to form many of the topological features (mountains and high altitude plateaus) that are now homes to large pools of endemic species.

As an example, the huge sauropod footprints that were discovered in Croatia by a team led by

Michael Caldwell of the University of Alberta (but published by Mezga et al, *Cretac. Res.* 27, 735-742, 2006 – such is the cutthroat world of scientific research!) were laid down about 95 million years ago. As well as the important paleontological significance, these also represent the most recent signs of life on what is known as the Adriatic-Dinaric platform before is sunk below sea level 94 million years ago. This is clear evidence of niche removal for certain dinosaur species in one specific area.

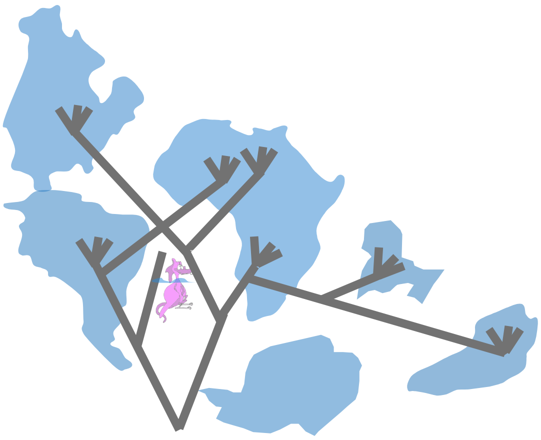

Fig.1 A major speciation event occurred prior to the dinosaur extinction, perhaps triggered by continuing changes in the Earth's geomorphology.

Simultaneous to this niche destruction process that appears to have provided an adverse selection pressure on non-avian dinosaurs, recent evidence indicates that there was on going speciation amongst both mammals [1] and the avian fauna [3]. Bininda-Emonds et al, [1] have constructed a near complete species phylogeny for the extant mammals. This includes estimates of divergence times using a combination of gene alignment techniques and fossil calibration points. A key result was their identification of two periods of diversification. The first was from 100-85 Million years ago. Again note that this coincide with the final stages of the break up of the supercontinental landmasses, although at the moment this is only a statement of coincidence and needs further theoretical investigation. The second appeared in the early Eocene, yet at approximately 50 millions years ago this was significantly after the K-T boundary. A key conclusion of [1] is that this "challenges the hypothesis that the end-Cretaceous mass extinction event had a major influence on the diversification of today's mammals".

Our proposal is that the K-T extinction event was a (final) perturbation in a long-term species innovation process that was driven by the continued break up and change to the topologies of the supercontinental landmasses. By the time of the final extinction of the dinosaurs, landmasses were significantly fragmented, separated by large bodies of water and undergoing significant topological changes due to tectonic activity. Overall, this led to acceleration in the Birth Death Innovation Model that underpins the evolution of biodiversity on this planet. Finally, and most importantly in terms of its relevance to our claims about the importance of support for SMEs in a DE, the long tail of rare species played a critical role as a pool of innovation during this adaptation of the natural ecosystem in response to this long-term period of niche creation and destruction.

1.2 Models of Social-Ecosystems

A primary goal behind our studies of ecological modelling for digital ecosystems is the avoidance of premature commitments to specific technological solutions or directions; the art of engaging with new communities is to leave your baggage and preconceptions behind. Instead, we will adhere to two guiding assumptions:

- Firstly, our core research should focus on generative and disruptive, rather than “sterile”, technologies;
- Secondly, minimizing the imposition of centralised control or organisation will enhance the ability of the Digital Economy to empower or transform a given socio-economic context.

The two assumptions are closely related, and could be summarised as a target of transforming the Digital Economy into a complex adaptive system, where institutions emerge and evolve through local interactions and memory. These are, of course, assumptions. Their validity must be established through both empirical and theoretical studies.

1.2.1 Adaptive Cycles, Resilience and Transformation

The concept of “ecosystem” is increasingly being used as a metaphor in business and systems thinking. In itself, this is a more inclusive view than the more traditional hierarchical, or value-chain models – it provides greater acknowledgement of the importance of a community as a whole in sustaining value creation. However, one can go much further in using the metaphor to drive the development of models that have real value in understanding and facilitating transformation in the Digital Economy.

The *adaptive cycle* concept emerged from studies of regional development and ecosystem management that were conducted in the 1980’s and 90’s [11]. The concept of a *climax community* is well understood in natural ecosystems. This corresponds to the *K* or *Conservation* phase of the adaptive cycle in Figure 1. But established business or stable socio-economic ecosystems also have an analogous K phase in which interventions and controls are typically chosen to maintain or streamline the delivery of the ecosystem’s outputs.

At some time, a change in either an internal or an external variable may trigger a transformation to the Ω or *release* phase (e.g. a forest fire, in a natural system, or the shift from a product oriented to a service oriented market place in the case of business ecosystems). The system may then respond with *reorganisation* before moving back through *exploitation* into a new (and possibly qualitatively different) conservation phase.

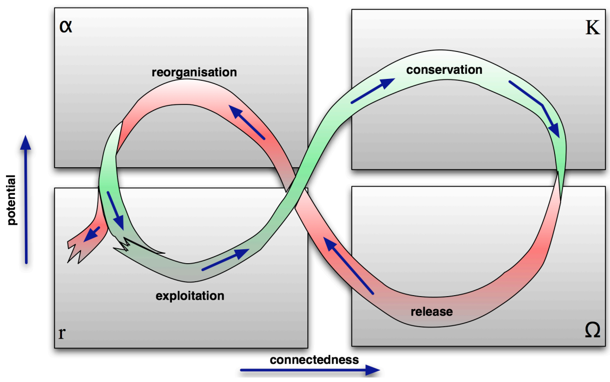

Fig.2 The adaptive cycle

The above adaptive cycle works well as a metaphor for many social-ecological systems. It also provides a useful framework for posing research questions and structuring simulations/experiments in Digital Ecosystems [12].

An important extension of this is that in most systems there is a hierarchy of such cycles (usually working on different timescales). A change in system variables in one level can trigger change in other levels. Typically a large business will hope to work in "K" sector. But a change in the higher-level business environment can trigger transition to Ω and then α. Note that if the response is not effective, the cycle can exit at "r".

Both natural and social/business ecosystems will demonstrate a degree of *resilience* to external change or shocks whilst in the *K* sector. In part a certain level of response diversity in the face of external change can provide this. Contrast the coral reefs of Great Barrier Reef (GBR) to those of the Caribbean. There is much higher species diversity in GBR. This in turn means that the various (trophic) functional groups in GBR have significantly enhanced response diversity compared with the coral reef ecosystem in the Caribbean. Broadly, this is what is currently providing the GBR with greater *resilience* than the coral reefs of the Caribbean (where there has been an 80% decline in hard corals).

This metaphor of *response diversity* in business ecosystems suggests that SMEs are a major factor in maintaining resilience of the ecosystems in the face of extreme events. This hypothesis is not established, but it does raise interesting research questions that can be addressed through simulations and comparative analyses of real-world case-studies. (It is appropriate that this paper was written in a time of financial uncertainty where resilience will be tested to the full.)

In general, resilience is a property that emerges from a range of features in any complex system. The mistake is often to use simple interventions to try to control a system in the face of some external pressure or shock, without considering, or even being aware of, all the variables that impact on the stability of the state of a complex ecosystem.

A number of case studies on the resilience of a range of social-ecosystems can be found in [12]. Some key lessons can be drawn from this:

Interventions are often planned without taking into account, or even being aware of, one of the key

variables of the ecosystem;

Interventions are planned on individual components of a complex system, without modelling how the system may adapt as a whole to these changes of individual components;

Ill-conceived interventions can impact on the *resilience* of an ecosystem to withstand extreme events;

Without carefully modelling the impact of an intervention, we may trigger an unrecoverable transition from one *stable state* to another (for example, the transformation of species rich sawgrass communities to single species stands of cattail, in the Florida Everglades).

In the case of business ecosystems, we believe SMEs play a critical role in maintaining the response diversity in the face of global change. However, as will be discussed shortly, SMEs are *not* fully engaged in the Digital Economy, and in many cases are seriously disenfranchised from it.

The importance of understanding the dynamics arising from the complexity of ecosystems and "panarchies" of ecosystems is not just in understanding how to maintain a valued ecosystem. It also helps us to understand the variables that impact on our ability to *transform* an ecosystem into a new, more desirable, state.

This has provided a short introduction to a number of key concepts. Before extending this to cover sustainability and the important enhancements of *memory* and *intention* that are needed for the study of social ecosystems, we will provide an overview of the current situation with digital ecosystems.

1.3 An Inclusive Digital Ecosystem for SMEs

In the EU-27, the percentage of enterprises' total turnover from e-commerce via the internet doubled between 2004 and 2007, passing from 2.1% (2004) to 4.2% (2007) of total turnover [8]. However, for B2B transactions among SMEs, on average only about 11% of SMEs use software solutions or internet-based services for e-procurement. Moreover, there is a massive gap between the percentage of SMEs placing at least some orders online (53% of total) and those that use special software for this (only 11% of total). SMEs without special software place orders mainly through websites or extranets of suppliers [9]. The result is a lack of digital back-office integration of procurement-related processes among European SMEs.

As discussed in [7] a major barrier to the expansion of use of B2B amongst SMEs is that the current standards for web service coordination require a central coordinator. Even where there is a natural hub for the business activities, this can provide a threat to the *local autonomy* of the participating SMEs – the web service coordinator will have access to business state, business logic and data about traffic that can help in enforcing a level of governance on the participating SMEs, and also inform acquisition policies of the coordinator. A distributed coordination model for long-running business transactions that aims to alleviate such concerns has been proposed in [17,20].

However, these are not the only barriers to adoption of digital innovation amongst SMEs. Our own studies in the Cambridge region indicate more fundamental issues need to be addressed, including:

Intense frustration at the lack of interoperability and loss of core business time; An already low ICT take up.

This situation is frustrating a potential expansion of the digital economy that would be of major significance. The focus on networking, collaboration and cooperation continues to develop in modern business processes [14], and is especially important amongst communities of SMEs. Greater inclusion into a technology enabled business landscape will enhance their capacity for achieving a form of “competitive co-evolution” [18]. We believe this model is especially relevant to the information-based industries that are becoming increasingly important in the redevelopment of the UK economy.

The tourism industry provides an example of the strong contrasts in inclusivity of the Digital Economy [4]. Increasingly, consumers rely on the Internet to search for information, and book their vacation. Yet, while many of the more satisfying (and socio-ecologically sensitive) vacation experiences can be found and booked through “micro-tourism” websites, these are easily eclipsed by the mega-tourism sites such as *Travelocity* and *Expedia*.

The barriers to adoption are not just technological. Evans *at al* [10] found that:

- Micro and small tourism enterprises don’t identify themselves with the mainstream tourism industry;
- The lifestyle choice of owner-operators often militates against their take up of ICT.

In addition, tourism SMEs tend to rely in intermediaries for their marketing and so have limited bargaining power in the distribution channel [6]. Our own work with the UK Technology Strategy Board and Regional Development Agencies indicates that these problems arise repeatedly across many SME sectors.

The key lesson we want to draw from the above is that the development of digital ecosystems needs to take into account the social perspectives of its members. To that effect we now return to ecological modelling and expand this to discuss inclusion of the socio-economic context in the models.

1.4 Response and Sustainability

So far, we have emphasised the importance of *response diversity* in providing a degree of resilience to shocks on an ecosystem. However, this may not be sufficient to guarantee recovery of an ecosystem. In addition, it is not only shocks that may trigger the transition of an ecosystem from the K (conservation) phase to the Ω (release) phase of the adaptive cycle, and its subsequent reorganisation into a qualitatively different regime. There is good empirical evidence from natural ecosystems that in certain circumstances they can respond catastrophically to a gradual increase in stress.

1.4.1 Catastrophic responses

We will summarise the discussion from Scheffer *et al.* in this sub-section [23]. For simplicity we will only consider one ecosystem state variable, and one impacting state variable. A common assumption in ecosystem management is that some increase in stress on that ecosystem will lead to a steady and gradual change in its state (figure 3, below).

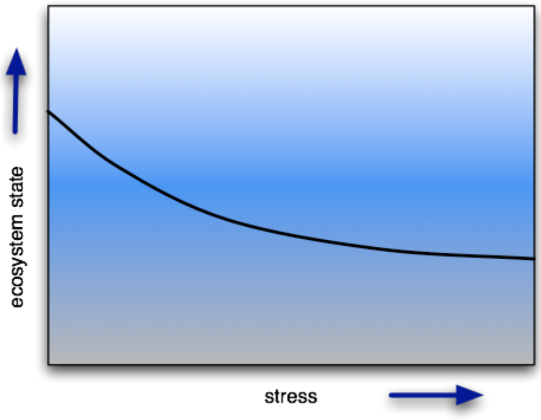

Fig.3 Naive model of ecosystem response to stress

This supports a comfortable feeling that if the ecosystem state deteriorates below an acceptable level for certain stakeholders, then it may be recovered simply by reducing the stress. That is, a K to Ω transition may be reversible.

Unfortunately, as we will describe in a concrete example shortly, a given ecosystem may support (at least) two stable equilibrium states over a range of environmental conditions. In such a scenario, the ecosystem may be relatively inert to change in stress over a certain range of conditions, but then respond with a catastrophic switch to a new stable state once the environmental stress exceeds a certain threshold (point F_2 in figure 4 above).

Two key issues arise. Firstly, given the relatively benign initial response of the ecosystem to the environmental stress, the catastrophic switch to a new stable state will typically be hard to predict and prepare for. Secondly, given the hysteresis effect of such a response curve, a very significant reduction in stress level will be needed before the ecosystem will transition back to the upper branch (point F_1 in figure 4). Indeed, it can happen that the point F_1 is to the left of the origin, and hence not reachable through a simple reduction in the stress factor that lead to the transition to the lower branch.

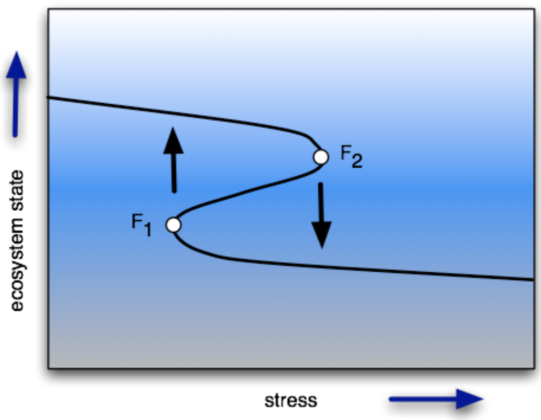

Fig.4 Stress may induce catastrophic change

It is worth including a concrete example to keep this paper self-contained. Both clear lakes, and arid ecosystems fit well to the above model, but given the greater experience of linking them into socio-economic systems, we will use the example of clear lakes following [5]. Fertiliser run-off from agricultural communities, and wastewater from other human activities can impact clear shallow lakes. The resulting increase in nutrients ("stress") stimulates the growth of phytoplankton. The resulting green *turbidity* reduces light intensity and hence impacts the characteristically lush plant growth on the bottom of the lake. This impacts on the small animals that feed amongst this vegetation. The fish species that feed on such animals, attach eggs to such plants or use submerged plants for shelter will then also decline. Finally, in response to the reduction in fish species, the number of birds visiting the lake will decline.

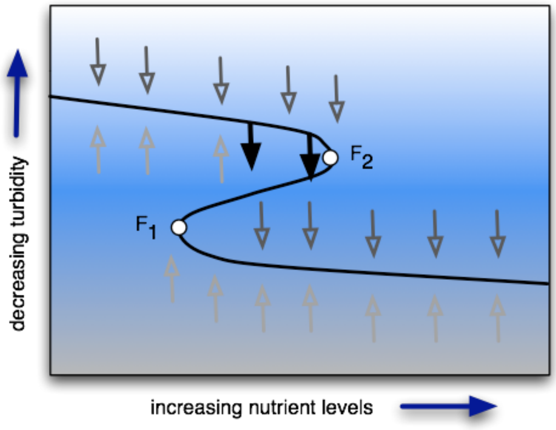

Fig.5 Resilience is minimal near the tipping points F_1 and F_2

A disturbance, or shock, to the ecosystem can be represented by a vertical displacement from the equilibrium line. Such a disturbance in this case might be a sudden increase in nutrient levels due to prolonged heavy rain, or a significant fish kill. Once the disturbance ceases, the system will return to equilibrium (the grey arrows). But note that if a large disturbance or shock occurs (represented by the heavy black arrows) near to the point F_2, it may tip the system into the second highly turbid stable state.

1.4.2 Resilience and Sustainability

It is still hard to gain a consensus on the semantics of sustainability as a concept. During the lead up to the Local Agenda 21 actions at the end of the last century, the Brundtland Commission defined sustainable development as: "development that meets the needs of present generations without compromising the ability of future generations to meet their needs" [26]. This is worthy, but a little hard to operationalise; does it mean that we should wind down oil consumption in case a future generation has a greater need for it, for example?

An alternative approach is to consider the sustainability of a social-ecosystem in terms of its capacity

to absorb shocks and chronic stress [13]. The concepts of *resilience* and the *adaptive cycle* then become important in guiding studies to aid the effective monitoring and management of complex (socio-economic, digital and natural) ecosystems. Provided we can identify all the key variables in an ecosystem, we may be able to combine theoretical and empirical analyses to develop a model of its response to change. As we have seen in the previous section, such models may need to be non-linear or even chaotic [5].

The adaptive cycle reminds us that *sustainability* should not be equated with *stasis*. A natural ecosystem may have a need to periodically cycle through instances of the adaptive cycle in order to avoid moving into a situation where a small disturbance could trigger a dramatic or catastrophic switch into an alternative stable state. A well-known example of this is the impact of fire control in temperate forests. Small scale, local, fires (Ω, release, phase) clear forest debris and lead to new growth, regeneration and gradual transformation back to a climax vegetation ($\alpha \rightarrow r \rightarrow K$). Attempts to control this natural cycle lead to a widespread and deep accumulation of forest debris. The potential for a catastrophic forest fire then builds up leading to at best a release phase that is hard, costly and extremely risky for humans to control, and at worst opens the risk of transformation into a new stable state (perhaps regeneration is dominated by a non-native water greedy tree species, that locks out recovery by the previous vegetation).

Things become more complex (and more relevant to digital ecosystems) when we include human social and economic interests into the ecosystem models. At this point we will quite often (perhaps usually) move away from closed cycles around the adaptive figure of eight. The release phase may be triggered by a factor internal (“revolution”) or external (“remember”) to the system. The ecosystem then needs (as humans are now involved, aspects of *intention* become relevant) to reorganise into a different state in order to succeed in progressing through the exploitation phase into a new (quasi-stable) conservation phase.

Gunderson and Holling [12] use the example of the Florida Everglades to demonstrate the institutional response to a series of ecological crises over the last century. Essentially the tensions there are between water management to support increasing economic development in a natural ecosystem where widespread flooding (release phase) was an important part of the latter’s adaptive cycle. To summarise briefly, we list here the institutions that were generated, together with the crisis that triggered the reconfiguration:

- 1903, Everglades Drainage District (Flood)
- 1947, Central and Southern Florida Flood Control District (Flood)
- 1971, South Florida Water Management District (Drought)
- 1983, Everglades Coalition (perception of switch to new ecological state and fear of pollution through high rainfall)
- 1989, Everglades National Park Protection and Expansion Act (Costly water quality lawsuit)

A full description of the evolution of the Everglades social-ecosystem can be found in [25]. The important point here is that the Florida Everglades social-ecosystem went through a sequence of “release” and reorganisation phases as memories of the consequences of external shocks triggered the

emergence of new institutions. These can be thought of as representative of new states in the history (ontogeny) of a complex adaptive socio-economic system. Collective memory in the system (hopefully) prevents reversion to an earlier, less resilient, state. Collective intention, in so far as it exists, helps the system to reorganise following release phases into states that are (hopefully) more resilient to external and internal tensions.

The development of well-informed models of social-economic systems is absolutely essential for continuing to guide the reorganisation of complex social-ecosystems following a release phase.

1.5 Lessons for Digital Ecosystems

The motivation behind the writing of this position paper was to raise awareness of an extensive corpus of empirical and theoretical studies of social-economic ecosystems that has shown real benefit in informing decisions in the management of such systems. Our thesis is that much of this work can form the foundation for similar studies of digital ecosystems.

We can already draw some provisional conclusions that are informing our own work and that of other partners in the OPAALS Network of Excellence [20,21]. We have seen that a key impactor on resilience of an ecosystem is the ability of the species that interact in that system to be able to maintain key functions as conditions vary [19]. The concern with current digital ecosystems is that they are often, perhaps always, focused around a single or a small number of centralised hubs. This provides a fundamental limitation on the *response diversity* of digital ecosystems. As with any natural ecosystem [13], we believe that the “functional diversity of species that support critical structuring processes” [5] is critical to the resilience of any digital ecosystem. Referring back to our view of enterprises as (one kind of) species within digital ecosystems, we believe that this requires us to develop digital infrastructures that respect and support the autonomy of small to medium sized enterprises (SMEs); the long tail of diversity in any resilient digital ecosystem [22].

Whilst some of the modelling from social-ecosystems may be directly relevant to digital ecosystems, there are some important distinctions too. For example, we mentioned that an agent might interact with more than one digital ecosystem. This is because in general, an agent may take multiple roles – a human agent, for example, may interact in a work related ecosystem, and an interest related ecosystem without cross-over between these roles. However, it is possible that one or more agents may identify a potential join between two previously separated ecosystems. As a concrete example, at the time of writing we are working to arrange funding to support a join between tourism and environmental/nature conservation digital ecosystems in India. This is an important form of innovation and can see the merging of two previously disjoint ecosystem ontogenies. We do not believe such merges have so far been studied.

There is a more private agenda too. We believe that effective use of digital ecosystems could lead to changes in lifestyles and working practices that could have significant environmental benefits (reduction in business travel, rural regeneration, for example). Our hope is that bringing an understanding of the dynamics of natural ecosystems into the science of digital ecosystems could help raise awareness of the importance of and lifestyle benefits to be gained by, a more holistic way of

living.

REFERENCES

[1] O.R.P. Bininda-Edmonds et al, "The delayed rise of present-day mammals," *Nature,* vol.446, 2007, 507-512.
[2] H. Boley and E. Chang, "Digital Ecosystems: Principles and Semantics", in Proceedings of the 2007 Inaugural IEEE Conference on Digital Ecosystems and Technologies, 2007, 1-6.
[3] J.W. Brown, R.B. Payne and D.P.Mindell, "Nuclear DNA does not reconcile 'rocks' and 'clocks' in Neoaves", *Biol. Lett.*, 2007.
[4] P. Braun, E-Commerce and Small Tourism Firms, In: Marshall S., Taylor, W. and Yu X. (eds), *Encyclopedia of Developing Regional Communities with Information and Communication Technology*, Idea Group, 2006.
[5] W.A. Brock, K. Mäler and C. Perrings, "Resilience and Sustainability: The Economic Analysis of Nonlinear Dynamic Systems", in: L.H.S. Gunderson and C.S. Holling, q.v., 2002, 261-289.
[6] D. Buhalis, "Information Technology for Small and Medium-Sized Tourims Enterprises: Adaptation and Benefits," *Information Technology and Tourism*, **2**, 79-95, 1999.
[7] P. Dini, G. Lombardo, R. Mansell, A. Razavi, S. Moschoyiannis, P. Krause, A. Nicolai and L. Leon, Beyond interoperability to digital ecosystems: regional innovation and socio-economic development led by SMEs, *International Journal of Technological Learning, Innovation and Development,* **1:3**, 410 – 426, 2008.
[8] Eurostat, Community Survey on ICT Usage and E-commerce in Enterprises, Brussels, 2007.
[9] European Commission, *The European e-Business Report*, e-Business Watch, Luxembourg, 2007.
[10] G. Evans, J. Bohrer and G. Richards, "Small is Beautiful? ICT and Tourism SMEs – A comparative European Study", *Information Technology and Tourism*, **3**, 139-153, 2001.
[11] L.H.S. Gunderson, C.S. Holling and S.S. Light (eds), *Barriers and Bridges in the Renewal of Ecosystems and Institutions*, Columbia University Press, New York, 1995.
[12] L.H.S. Gunderson and C.S. Holling, Panarchy: Understanding transformations in human and natural systems, Island Press, 2002.
[13] C.S. Holling, D.W. Schindler, B.W. Walker and J. Roughgarden, "Biodiversity in the functioning of ecosystems", in C.A. Perrings *et al.*, *Rights to Nature,* 1995, 44-83.
[14] J. Marceau and M. Dodgson, *Systems of Innovation*, Proc. National Innovation Summit, Melbourne, 2000.
[15] R.M. May, *Stability and Complexity in Model Ecosystems,* Princeton University Press, Princeton and Oxford: 2001.
[16] M. Mazzanti and R. Zoboli, "Economic instruments and induced innovation: The European policies on end-of-life vehicles", *Ecological Economics*, vol.58, 2006**,** 318-337.
[17] S. Moschoyiannis, A. Razavi and P. Krause, "Transaction Scripts: making implicit scenarios explicit," in *Proc. ETAPS 2008 – FESCA'08*, ENTCS, Elsevier, 2009. *To appear*
[18] A. Ordanini, "Infomediation and Competitive Advantage in B2B Digital Marketplace", *European Management Journal*, vol.19, 2001, 276-285.
[19] C. Perrings, "Ecological resilience in the sustainability of economic development," *Economie Appliqueé*, vol.48, 1995, 121-142.
[20] A. Razavi, S. Moschoyiannis and P. Krause, "A Coordination Model for Distributed Transactions in Digital Ecosystems," in *IEEE Digital Ecosystems and Technologies* (*IEEE-DEST'07*), 2007.
[21] A. Razavi, S. Moschoyiannis and P. Krause, "A Scale-free Business Network for Digital Ecosystems," in *IEEE Digital Ecosystems and Technologies* (*IEEE-DEST'08*), 2008.
[22] A. Razavi, S. Moschoyiannis and P. Krause, "A Self-Organising Environment for Evolving Business Activities," in *Computing in the Global Information Technology, ICCGI'08.* 2008, 277-283.
[23] M. Scheffer, F. Westley, W.A. Brock and M. Holmgren, "Dynamic Interaction of Societies and Ecosystems," in: L.H.S. Gunderson and C.S. Holling, q.v., 2002, 195-239.
[24] A. G. Tansley, "The use and abuse of vegetational concepts and terms," *Ecology,* vol 16, 1935, 284-307.
[25] B. Walker and D. Salt, *Resilience Thinking*, Island Press, 2006.
[26] World Commission on Environment and Development, *Our Common Future*, Cambridge University Press, UK, 1987.

2 BIOLOGICAL FRAMEWORK & DIGITAL ECOSYSTEM (UNIVDUN)

2.1 INTRODUCTION

This section is, of necessity, speculative. It is designed to give some indication, based on work carried out within the OPAALS project, on possibly fruitful directions for future research. The concept of the digital ecosystem has its origin in the natural sciences and it is only fitting that future directions of research into digital ecosystems should continue to be informed by developments in the natural sciences.

Biology operates, and can be interpreted, at a variety of levels. In ascending order of size, we move from molecules, to cells, to organisms, to populations, to species and, ultimately, to the biosphere as a whole. The previous sections have dealt with parallels between commerce and biology at the species level and several key themes have emerged: resilience and diversity, for example. In this section we will briefly expand on this approach before considering the problems and opportunities that arise when attempting to relate digital ecosystems to biology operating at somewhat different levels: the molecular level (the regulation of gene expression); and the cellular level.

2.2 COMPETITION AND COOPERATION

One key to the success of any digital ecosystem is the balance between competition and cooperation. Elements of competition are essential so that individual entities (agents, SME's) can evolve. Without competition there would be no need to hone skills, or consider alternative ways of operating; stasis would rule and no progress would be made. One problem is that, if all entities are engaged in remorseless struggle, this competition becomes an end unto itself and vast amounts of energy are expended to no great effect. If all SME's within a region were directly competing with one another it is difficult to imagine how that region might develop effectively and efficiently. This is where cooperation (mutually beneficial interaction) comes in. If there is an appropriate balance, between antagonism and mutual support, amongst agents in an ecosystem then that system is more likely to possess the key attributes - resilience and diversity - that will enable it to flourish.

Up until very recently antagonistic and mutually beneficial interactions in biological ecosystems have been studied separately. However a recent paper [1] has, perhaps for the very first time, studied the two aspects simultaneously [2]. The conclusions reached have important implications for the architecture of digital ecosystems. Put in simple terms, the architectures associated with cooperative systems (mutualistic interactions) are nested structures [3] whilst competitive systems are associated with a more compartmentalised design.

2.3 WHY CELLS MIGHT HELP…

The cell may be the biological structure that most usefully serves as a metaphor for an SME. The networks of interactions between cells are tractable in number, time and space. So, by taking the cell

as the unit of study, we may be able to avoid some of the problems associated with the use of species (as in classical ecosystems) or of molecules and gene expression (as in some of the work undertaken within WP1 of the OPAALS project). It does not really matter whether we consider the living world to have arisen as a result of chance, contingency, selection and a Blind Watchmaker or whether we see the Hand of God at work. What matters is the complexity of living systems and our imperfect ability to understand how they work now (rather than how this intricate web of interactions has arisen). What matters is that living systems represent optimal solutions to common problems.

One of the great virtues of cells is that they do a lot of complicated systems biology for us. The states of a cell reflect its complex internal metabolic activity, the overall results of the trafficking of molecules, signals and receptors and the balance of the expression of numerous genes. Computationally this extreme complexity is very difficult to model using an approach that moves up in scale from genes and molecules. Although Systems Biology can help resolve some of this complexity, there is, perhaps, a more straightforward way of viewing things: simply to accept that the behaviour or state of an individual cell reflects the sum total of all that is going on within it. The precise mechanisms and structures involved need not concern us. All we need to know in a cell-based model is the state of each given cell at a series of points in time. This knowledge is sufficient to help us towards an understanding of the architecture and dynamics of the tissue to which those cells contribute.

Cells form tissues and tissues have architectures that, through evolution, may represent optimal solutions to the problems of growth and form, function and structure. By analogy, SMEs are the structural components of the regional economy and perhaps some of the lessons learnt as we move from cells to tissues may apply as we move from individual SMEs to the architecture of sustainable regional development. The advantage of this approach is that we can pull in current research in several areas in the natural sciences. Two main strands of research are particularly relevant to this question: the human physiome project; and, the construction of artificial tissues.

2.4 The Human Physiome Project

The human physiome project is, in the words of its investigators http://www.physiome.org.nz:

> "… a worldwide public domain effort to provide a computational framework for understanding human and other eukaryotic physiology. It aims to develop integrative models at all levels of biological organisation, from genes to the whole organism via gene regulatory networks, protein pathways, integrative cell function, and tissue and whole organ structure/function relations. Current projects include the development of:
>
> - ontologies to organise biological knowledge and access to databases;
> - markup languages to encode models of biological structure and function in a standard format for sharing between different application programs and for re-use as components of more comprehensive models;
> - databases of structure at the cell, tissue and organ levels;

- software to render computational models of cell function such as ion channel electrophysiology, cell signalling and metabolic pathways, transport, motility, the cell cycle, etc. in 2 & 3D graphical form;
- software for displaying and interacting with the organ models which will allow the user to move across all spatial scales."

This places this project firmly at the interfaces between biology, mathematics and computer science and the relevance to the OPAALS project and its subsequent development is obvious. In simple terms, the mathematical and computational methods developed in the human physiome project, as it moves from cells to tissues and organs, will provide useful tools and analogies for use within and beyond the OPAALS project as it moves from individual SMEs, through a consideration of digital ecosystems and peer-to-peer architectures, and towards supporting sustainable regional development.

We already know a great deal about how cells behave in a complex tissue such as the human colon [4, 5], and how these behaviours will produce and sustain a particular architecture. We also have a keen sense of how this particular architecture is suited to function and how it permits dynamic adaptations within an ever-changing environment. We know the nuts and bolts: for example, how to model mathematically the gradients within the Wnt-catenin pathway that control the replacement of cells along the intestinal crypt and how to integrate this model into a more complete description of intestinal renewal [6].

2.5 Artificial Tissues and Organs

The other strand of biological research that may directly inform work carried out on beyond the OPAALS project concerns the cell-based construction of artificial organs and tissues [7]. This time we would be considering a real, rather than a virtual, environment as we move from individual cells to an organised, functional, tissue. The rules and constraints that govern the building of a viable tissue from a series of cellular components are likely to be very similar to those that would be usefully employed as we move from a collection of SMEs towards a sustainable regional economy. With the increasing availability of stem cells, and with the move to using artificially created tissues both for therapy and for research, the time is right to consider how we might use some of the principles involved in these systems for designing infrastructures to support SMEs and regional development.

2.6 Why cancer is a useful model

Cancer is characterised by disordered architecture and disordered function. Normal cellular behaviour in a complex organism was described over 150 years ago, by Rudolf Virchow, as a cell state in which every cell is a citizen. In this view cancer becomes an example of the breakdown of democracy and a descent into anarchy. We can then consider the question of reverse engineering. By considering the features that characterise the transformation from normal to malignant we may be able to define the properties of a cell that keep it in order and which sustain harmonious and productive relationships within a tissue -properties that we could then consider as worthy of consideration within any digital ecosystem.

The features of cellular behaviour that are associated with malignant transformation have been studied for over 70 years [8] and were most recently summarised in a classic paper by Hanahan and Weinberg published in Cell in 2000 [9]. The features that are associated with malignancy (together with their inversions - that is desirable properties for agents in a stable digital ecosystem) are summarised in the following table.

2.7 "Hallmark" of Cancer	2.8 Desirable property for agent
Invasion	Should respect local boundaries
Metastasis	Should not disseminate
Recruits its own blood supply	Should not be parasitic (diverting others' resources)
Does not require growth factors	Should only grow/expand when signals indicate that this is appropriate
Resists apoptosis	Ceases trading if internal/external conditions mean that it would be inappropriate to continue
Immortality	Has finite life-span

2.9 Conclusion

This paper and the previous one have argued, respectively, for the study of cell structures and resilience in social ecosystems as being important metaphors for the study of digital ecosystems. The challenge now is to collect together the interdisciplinary skills needed to progress from models and metaphors that can explain behaviours that we observe, to more formal models that can be used to predict outcomes of specific interventions, and be used to inform policy making in digital-, and more generally, industrial-ecosystems. The next two sections will detail the proposals that we have put together to move this agenda forward. The ERIE (Evolution and Resilience of Industrial Ecosystems) project is already running at the time of writing. This is a 6 year, £4M project involving the Departments of Computing, Mathematics, Sociology and the Centre for Environmental Strategy at the University of Surrey. To broaden the skills set, we also have set up MILES (Models and Mathematics in

Life and Social Sciences). This is a 3 year, £0.8M "Bridging the Gaps" project involving the Departments of Computing, Mathematics, Sociology and Microbial Sciences at the University of Surrey.

2.10 References

1. Thebault E, Fontaine C. Stability of Ecological Communities and the Architecture of Mutualistic and Trophic Networks. Science. 2010 Aug 13;329(5993):853-6.

2. Bascompte J. Ecology. Structure and dynamics of ecological networks. Science. 2010 Aug 13;329(5993):765-6.

3. Pascual-García A, Ferrera A, Luque B, Bascompte J. The architecture of mutualistic networks minimizes competition and increases biodiversity. Nature. 2009 Jan 1.

4. van Leeuwen IMM, Byrne HM, Jensen OE, King JR. Crypt dynamics and colorectal cancer: advances in mathematical modelling. Cell Prolif. 2006 Jun 1;39(3):157-81.

5. van Leeuwen IMM, Edwards CM, Ilyas M, Byrne HM. Towards a multiscale model of colorectal cancer. World J Gastroenterol. 2007 Mar 7;13(9):1399-407.

6. van Leeuwen IMM, Mirams GR, Walter A, Fletcher A, Murray P, Osborne J, et al. An integrative computational model for intestinal tissue renewal. Cell Prolif. 2009 Oct 1;42(5):617-36.

7. Placzek MR, Chung I-M, Macedo HM, Ismail S, Mortera Blanco T, Lim M, et al. Stem cell bioprocessing: fundamentals and principles. J R Soc Interface. 2009 Mar 6;6(32):209-32.

8. Foulds L. The experimental study of tumor progression: a review. Cancer Res. 1954 Jun 1;14(5):327-39.

9. Hanahan D, Weinberg RA. The hallmarks of cancer. Cell. 2000 Jan 7;100(1):57-70.

3 ERIE (Evolution and Resilience of Industrial Ecosystems)

3.1 Aim and Objectives

ERIE is one strand of our work in extending the study of digital- and industrial-ecosystems (with the pervasiveness of the Digital Economy, the two are becoming increasingly difficult to separate). It is a six-year programme, funded under the UK EPSRC's *Complexity in the Real World* theme. The material from this chapter is taken from the ERIE *Case for Support*.

Using a complexity science perspective, ERIE will formulate, develop and test tools and techniques for the analysis of industrial sectors, regarded as 'ecosystems', in order to facilitate the steering of such systems to achieve a variety of policy goals. The programme will:

- Develop a framework for generating and testing policy and governance strategies for complex industrial ecosystems.
- Develop complexity science models of such systems, and analytical and computational tools for investigating these models.
- Apply them to specific ecosystems focusing on real world policy problems.

3.2 Vision

ERIE will address a series of fundamental questions relating to the application of complexity science to social and economic systems. To a large extent the rapid growth of interest in complexity science has been driven by the success of the statistical physics theory of interacting particle systems. It has been tempting to make analogies between the networks of interactions in these models and those that occur in social systems. But social systems are not composed of 'simple' particles and cannot always be expected to behave in similar ways.

Biological ecosystems provide another class of models that can be regarded as lying between inanimate physical systems and those studied in the social and economic sciences. Here feedback and adaptation play a central role and the emphasis is on questions concerning sustainability, resilience, diversity etc. Analogies between these systems and social systems are again very tempting and have resulted in, for example, the development of 'Industrial Ecology' as a new paradigm for socio-economic systems. But again the extent of the appropriateness of the analogy between biological and social systems needs to be explored and tested. For example the role of 'reflexivity' in social systems – the extent to which the behaviour of social agents is modified by their perception of the 'emergent' properties of the system – needs to be much better understood.

The ERIE programme of research aims to embed cutting-edge complexity science methods and techniques within prototype computational tools that will provide policymakers with realistic and reliable platforms for strategy-testing in real-world socio-economic systems. We propose an ambitious, coherent programme encompassing the gathering of data from case studies, the development and application of appropriate theoretical and computational techniques, simulation using agent-based

models and the incorporation of all these elements into 'serious games' for use by policymakers. Unusually for a programme strongly focused on technical aspects of complexity science, we will put engagement with stakeholders right at the heart of our research: we will study the negotiation of policy goals and options, explore the role of models in policymaking and involve policymakers in the design and testing of our strategy tools.

The ERIE programme will focus on a crucial aspect of the UK economy: the way in which firms are interdependent on each other, with the interrelationships being multi-level and multi-valued. Within an industrial 'ecosystem', there are relationships of supply and demand; the transfer of knowledge; competition for labour; the transfer of materials down supply chains; negotiation over standards; collaboration in trade associations and unions; and innovation, product differentiation and branding. Thus a typical industrial ecosystem, such as those to be examined as case studies in the ERIE programme, will exhibit links at the levels of material flow, information, human labour, economics, and governance as well as being located within spatial, environmental, political, and cultural constraints.

Industrial and urban policy, as well as policies to reduce environmental impact, increase resilience and enhance wealth creation, have almost always treated the various material, environmental, social, political and economic 'levels' separately. The complexity science approach provides an exciting opportunity to do better.

An industrial ecosystem is an open system that interacts with its environment. It contains feedback loops. It has a history – two ecosystems in the same state, but with different histories (especially, perhaps, if cultural and social aspects are included) will evolve to different next states (that is, they will exhibit path-dependency). Industrial ecosystems are nested within higher-level structures/ecosystems, and may have other ecosystems nested within them. Transformations may be triggered by interactions with higher-level or lower-level ecosystems. Higher-level structures may emerge from the interaction between lower-level structures.

The ERIE programme will use mathematical and computational approaches to model these layered, nested, multiscale systems, where the links between actors are dynamic and the exchanges between them are unpredictable, fluctuating and perhaps sporadic. Within this context, ERIE will examine concepts and measures of resilience, emergence and immergence. This leads us to some of the most intriguing open questions of complexity science. What is emergence and how do we measure emergent properties? What are the characteristics of stochastic dynamics of layered networks? ERIE will seek answers inspired by the real-world network structure and dynamics of industrial ecosystems as captured in our case studies. Using the insights gained, ERIE will build realistic agent-based models to provide the engine of "serious games", bringing to policymakers the most recent results in complexity science within an intuitively comprehensible tool that they can use to explore alternative policy strategies.

The ERIE vision is to provide models of multi-level socio-economic systems that are useful for decision-makers aiming to 'steer' towards policy-relevant goals. We recognise that different stakeholders may have quite different goals, and/or place different degrees of importance on the same

goals. For example, there may be a different values placed on equity compared with sustainability (and these concepts may have contested definitions). Thus it is not our intention to provide 'the' policy solution to policy problems (specifically, it is not our intention just to show how particular ecosystems may be made more resilient or more sustainable), but rather to provide a suite of tools which will allow decision makers and their representatives to investigate alternative scenarios given a set of assumptions and initial conditions.

The ERIE programme will apply the methods of data assimilation, largely developed in the context of weather forecasting, to incorporate the inevitably patchy and incomplete data from case studies into agent-based models, on an ongoing basis, with the aim of providing 'predictive' tools that are continually updated with real-world data. By 'prediction' here we mean the identification of alternative scenarios along with estimates of the probability that each will be realised over given time frames, and estimates of the sensitivity of these to uncertainties in the data and underlying model. There is no implication of determinism. ERIE will develop a formal language to express the gradual elaboration of high-level strategic goals of industrial ecosystems in terms that can be checked against model outputs and in so doing, equip the models with a formal underpinning. In so doing, the models will be equipped with a formal underpinning. We will also use stochastic bifurcation analysis in order to map out where in the space of model parameters different scenarios are expected to be seen, and when and how they diverge from each other. This high-level analysis will be available to policymakers within the "serious game" tools.

In socio-economic systems, one cannot cleanly separate the observer from the system observed. The ERIE programme recognises that policymakers are themselves part of the system. It is therefore an integral part of our research to study – and involve – stakeholders. Specifically, one research stream in the programme is concerned with studying those with a stake in the system, as controllers, decision makers, customers, workers, etc., their goals, policy options and their links with the industrial ecosystems that they are interacting with.

The ERIE research programme is divided into four streams. Each stream will continue through the life of the programme, although with activity waxing and waning in relation to its role in the programme as a whole. In order to apply complexity science in a useful way to the real world, the development of a common language and shared perspectives between disciplines will be vital. Thus each stream consists of a number of cross-disciplinary projects, each of which will be supervised jointly between two of the four participating centres: the Departments of Mathematics, Computing and Sociology, and the Centre for Environmental Strategy. The four streams (with the approximate percentage of overall resources assigned to it in brackets) are:

1. Analytical and computational tools for complex socio-economic systems (35%)
2. Simulation and modelling of industrial and business ecosystems (30%)
3. Business and industrial ecosystem case studies (20%)
4. Stakeholder analysis and engagement (15%)

Each of these is described in more detail in the following pages.

3.3 Stream 1: Analytical and computational tools for complex socio-economic systems

The aim of this stream is to provide mathematical and computational tools for the analysis of industrial ecologies. The focus will be on four particular aspects:

Providing data assimilation tools for the interactive refinement of simulation scenarios using Case Study data and the quantification of uncertainty in model outputs

Developing model-checking techniques to validate and evaluate agent-based models

Developing theory and simulations of stochastic flows on networks with dynamic contact structure

Developing stochastic bifurcation techniques for the analysis of industrial ecosystems.

3.3.1 Data assimilation

Data from stream 3, the case studies of industrial ecosystems will be patchy and incomplete, as is usual with socio-economic data. There will also be uncertainties in the agent-based simulation models (ABMs). Existing data assimilation techniques have largely been developed in the context of weather forecasting, and involve representing the uncertainty in input data and model so that the likelihood of the various output simulation scenarios can be quantified (Williams et al 2005; Griffiths and Nichols 2001). They further aim to provide estimates of parameters and missing initial data required for the simulations and the uncertainties inherent in these estimates. This will be crucial in validating the models and evaluating the robustness of simulation outcomes (Judd and Smith 2004). In weather forecasting, probability distributions of atmospheric conditions are stepped forward in time using the forecast model. Similarly, ERIE will construct probability distributions of agent states, according to the uncertainties inherent in our Case Study observations and step these forward in time using ABMs. As ABMs are themselves stochastic in nature (meaning that they involve an element of randomness or unpredictability), we will average over an ensemble of realisations of the model in order to capture the evolution of the first few lower moments of the relevant probability distributions. Iterative updating of parameter values and missing initial data, using Bayesian (probabilistic) methods and the quantified mismatch between empirically measured and simulated values of observables at various times during the history of a simulation, will lead to more accurate predictions from the simulations as data and observations are incorporated and adjusted, in the same way that weather predictions become more reliable for a given date as that date approaches. In this context, 'prediction', means the identification of alternative scenarios, along with estimates of the probability (likelihood) that each scenario will be realised over a given time frame, and estimates of how sensitive these alternative scenarios and probabilities are to uncertainties in the data and model parameters.

3.3.2 Complex adaptive networks

We expect that the networks formed by contacts between entities in industrial ecologies will be highly dynamical: the contact structure will change over time with nodes joining and leaving, moving in space, links being formed, broken and reformed, changing connection strengths and vulnerabilities.

The networks will also be nested and layered and the flows on them will be stochastic. Questions of stability and resilience for such highly complex networks remain unresolved and are common to complex systems in many different contexts from communication networks to neuroscience and epidemiology (Lim et al 2007; Gross eta l 2006; Albert et al 2000; Callaway et al 2000). In particular, we will use insights gained from the Case Studies and agent-based models to construct various different measures of resilience for the dynamics on a network that might be relevant to real-world applications, such as, for example, the tendency of the dynamics to relax to a complex adaptive cycle and the continued involvement in the dynamics of a certain fraction of nodes or links in the network. We will investigate these issues both from a more formal mathematical perspective and also using numerical simulations in order to draw out such general properties as exist for such loosely structured networks. These results will have wide applicability beyond the specific context of industrial ecosystems. We will also focus in more detail on their relevance to the agent-based simulations of the Case Studies and in particular to project 2.2 on emergence and immergence.

3.3.3 Path-dependency analysis and compositional scenarios

This project will develop an automated tool for analysing the modality of the interdependencies between actors in satisfying a goal – a prescriptive statement of intent whose satisfaction requires the cooperation of agents. Such high-level requirements will be framed dynamically as aggregates of multi-valued state variables. The aim is to produce the set of possible transformations in terms of different courses of action available from a given configuration. Using the findings of the ABM simulations from the Case Studies, we will formalise the path dependencies between agents at multiple levels, taking into account the histories and sets of possible next states of the coupled ecosystems. We will develop a formal language to express high-level, strategic concerns (such as ‘improve resilience’ or ‘increase innovation capacity’) in industrial ecosystems in terms that can be operationalised into agent actions and then checked against the model outputs.

This work will draw on advanced model-checking techniques (Kwiatkowska, Norman, Parker, 2004) and goal-oriented methods (Letier, Kramer, Magee, Uchitel, 2008) that will be extended to account for properties whose definition is not pre-set but can take different forms in different contexts. An important aspect of this work will be to address concurrent and nondeterministic interactions that feature in the multi-level compositional scenarios that can satisfy goal refinement structures in industrial ecosystems.

Building on the representation of the connections between agents at several levels in the ABMs, and the associated quantitative data and probability distributions of agent states, the tool will produce all possible alternative scenarios that may arise due to (emergence and immergence in) the interplay between global and local interactions. It will frame the space of different courses of action available and the corresponding alternative scenarios, with levels of (un)certainty attached. Successful development of the technique will result in a powerful tool for exploring different courses of action in socio-economic systems and infrastructures comprising multiple actors that stand in layered, nested, dynamic interrelationships, and are influenced by the environment (constraints) in which they operate.

3.3.4 Stochastic bifurcation analysis for industrial ecosystems

ERIE will use an equation-free method to perform bifurcation analyses of agent-based models of industrial ecosystems, where ensembles of realisations of an ABM are used as a timestepper to find steady states and perform pseudoarclength continuation. This will allow us to analyse systematically how the structure of alternative outcome scenarios in the model depends on input parameters, thus mapping out where in parameter space different simulation scenarios are expected to be seen, and how and where they diverge from each other. For example, we might want to understand how various measures of resilience of an industrial ecosystem depend on such quantities as the average duration of a link between any two firms.

Equation-free methods have recently been applied to chemical kinetics (Makeev, Maroudas & Kevrekidis, 2002) and simple models of individual-based biological dispersal (Erban, Kevrekidis & Othmer, 2006). Their potential for systematic analysis of complex social systems is as yet unexplored and represents a considerable technical challenge. Successful development of such equation-free methods will provide a powerful tool with potentially general application to other complex real-world social and economic systems.

3.4 Stream 2: Simulation and modelling of industrial and business ecosystems

The aim of this stream is to create a series of generic models of industrial ecosystems, and to parameterise these to fit the Case Studies researched in stream 3. The models will allow the simulation of scenarios of interest to policy makers and will be suitable for analysis using the tools and techniques developed in stream 1. In carrying out this aim, the work will:

- Review, integrate and develop existing agent-based models of innovation networks, industrial clusters and ecosystems
- Create, validate and apply a series of agent-based models of industrial ecosystems of increasing sophistication
- Convert one or more of these models to 'serious games' and investigate their value in interacting with stakeholders

3.4.1 Multi-level models of industrial ecosystems

This project will develop one or more agent-based models of an industrial ecosystem, based on data from the South Humber case study (project 3.1). The simulation will then be applied to data from the Food Supply Chains study (project 3.2). The design of the simulation will draw on existing work on industrial dynamics, innovation networks, and industrial districts. An important aspect of the model will be its representation of the linkages between actors at several levels, including materials flows, energy, information, economic transactions, and collaboration/competition. These linkages may span different spatial and temporal scales. The research question to be answered is how such inter-level linkages should best be represented; the project will develop some general methods for doing so that are applicable not only to industrial ecosystems, but also more widely.

3.4.2 Emergence and immergence in industrial ecosystems

While the model developed in project 2.1 will focus on the global effects of local interactions

between actors, the prime focus in this project will be to understand better both the emergence of macro effects and the consequences of these on actors through 'downward causation' and immergence. This is a topic where the basic conceptual distinctions are not yet well understood, yet the domain of industrial ecosystems will provide a fertile ground for tracing through upward and downward influences and formalising the findings in simulation models. An example is the way that norms and regulations are developed within an industrial sector and then used to impose constraints on the actor's behaviour. The research questions for this project are how emergence and immergence may be modelled and how the modelling may be used to inform or critique the substantial philosophical literature on the subject.

3.4.3 Using serious games for industrial policy

As noted elsewhere in this proposal, if the ERIE research programme is to fulfil its aims, it needs to engage with policy-makers and those who make decisions within and about industrial ecosystems. One approach that is rapidly gaining ground for engaging with professionals is the development of 'serious games' (Bergeron, 2006; Kelly et al, 2007). The idea is to create agent-based models in which some of the agents are in fact people, who are thus immersed in the simulation. The effect is the analogue of a flight simulator, with the user able to take decisions and explore their consequences in simulation. While such "serious games" have been used effectively in domains such as health care, business operations, security, and social change, they have rarely if ever been applied to complex socio-economic systems such as being modelled in the ERIE programme. In order to develop such a serious game, the user/gamer must be provided with an interface that will make the game and his or her options for action intelligible. The prime research questions for this project are to develop principles for designing serious games for complex socio-economic systems, and to show by example that such games are effective for engaging with policy-makers.

3.5 Stream 3: Business and industrial ecosystem case studies

The aim of this research stream is to provide the data on complex societal systems that are to serve as the empirical foundation and test beds for the development of the ERIE programme's approach to modelling and analysing complex societal systems. It will:

- Identify the structure and dynamics of the systems including the identification of agents, their relationships and the interactions in and between various levels (e.g. individual, intra-organisational, inter-organisations, institutional/political)
- Collect appropriate information on the 'performance' of these systems in the dimensions of interest (e.g. environmental, economic, social effects) under different scenarios.

3.5.1 Case Study: Development of an Industrial Ecosystem on the South Humber Bank, North Lincolnshire

The South Humber Bank, situated on the Humber Estuary, is an active industrial area comprising a diverse set of industries ranging from fish and other food processing to oil refining and chemical and bio-chemical production facilities. The Port of Grimsby and Immingham is the UK's largest port by tonnage, the area providing infrastructure for 20% of the UK's gas landing, and export facilities for the

region. The Humber Estuary is also a UK Site of Special Scientific Interest (SSSI) and Special Area of Conservation, and the area has significant challenges in terms of flood risk management. Residential communities in the neighbouring area (including those in Immingham and Grimsby) face significant socio-economic challenges including unemployment and fuel poverty. Industrial Symbiosis (i.e. the sharing of materials, energy, assets, logistics and expertise between organisations) has been identified as a means to promote context-sensitive economic development, the efficient use of resources, the reduction in environmental impacts (including those associated with carbon emissions), job creation and the alleviation of fuel poverty through use of excess industrially generated steam for district heating. The eco-development of this area appears to be hampered by barriers due to national and regional government policies and associated legislation, corporate policies and structures, and local governance structures and cultures (including two local councils and one Regional Development Agency with potentially conflicting policies and objectives).

This case area is ideal for the exploration of linked biophysical, industrial, economic, social and governance systems, populated by a large number of diverse agents. To identify and characterise its structure and dynamics, a range of techniques will be used. These include historical analysis of the development of the area, mapping of current resources assets (biophysical, industrial, infrastructure) via GIS, stakeholder interaction through interviews, workshops to facilitate understanding and scenario development, and ultimately interactive model building and verification.

A foundation for this project exists though strong links to Link2Energy, an SME and the delivery partner for the UK National Industrial Symbiosis Programme (www.nisp.org.uk) in the Yorkshire and Humber Region, through Surrey Engineering Doctorate project (2008-2011). We shall use the advice, contact list and knowledge of Link2Energy under a small consultancy contract with one of their staff.

3.5.2 Case Study: Resilient Food Supply Chains to Support Food Security and Global Sustainable Development

Food security has become one of the main priorities of UK and international policy makers. Although the UK is 60% self-sufficient in all foods (Defra, 2008), it is a net importer of food and relies on global food supply chains. Maintaining supply chains that are resilient to short terms shocks and crises, as well as long-term structural change is a core part of food security (ibid). This case study will focus on developing tools that will enable policy makers to visualise national, regional and international food supply chains including the physical systems in place, their environmental effects and the (role of) stakeholders directly and indirectly involved in these supply chains. A key element of this work will be understanding the governance structures of these supply chains and how they affect resilience as well as the economic, environmental and social performance of the supply chains.

For example, in the UK, the majority of food supply to consumers is controlled by four retailers most of whom operate flexible and just-in-time supply chains. Although this may have financial benefits to consumers, reduce wastage and provide some resilience against short-term shocks, it can have significant knock-on effects in developing countries (leading for example to highly insecure markets for developing country farmers and associated employment and income insecurity and variability for farm workers). There is the increasing societal awareness of the potentially negative environmental and

social impacts of globalised food chains.

Key questions include: What are the structure and dynamics of local, regional and global supply chains? What are the key economic, environmental, social and geopolitical uncertainties facing these supply chains? What is the influence of governance structures and policies on these supply chains and their performance? What would be the effect of alternative governance structures and policies?

These questions will be addressed by focusing on a set of relevant foods sources locally, regionally and globally. A foundation for this work exists in a range of studies done by the Centre for Environmental Strategy and RESOLVE including life cycle assessment studies (to identify the environmental impacts of food supply chain from 'cradle-to-grave') and social value chain analysis studies (to identify the distribution of economic benefits and the influence of governance structures) for food and related products (primarily vegetables and energy crops). Furthermore, CES hosts the Food Climate Research Network, (www.fcrn.org.uk) which provides access to an extensive network of food related stakeholders and associated activities and resources.

3.6 Stream 4: Stakeholder analysis and engagement

Too often, major research resources have been invested in the development of models intended to inform policy-making that turn out to be addressing the wrong issues at the wrong time and in the wrong way. ERIE will avoid this, both by focussing on tools and techniques appropriate to the analysis of complex socio-economic systems, and by bringing engagement with stakeholders right into the heart of the ERIE programme. The latter is the aim of this research stream. The research has two, linked and overlapping projects. The first examines the social, political and institutional processes by which policy goals and options are formulated, negotiated, selected and implemented. The second project will develop a better understanding of the actual or potential role of models in policy-making processes. The research will use the case studies of stream 3 as starting points.

3.6.1 The negotiation of policy options and goals

This project will clarify the way in which policy is made in practice in contemporary UK local and regional government. It will study how issues are placed on the policy agenda, how stakeholders are identified, what kinds of lobbying and persuasion are conducted and by whom, the types of organizations that are involved, how decisions are made (or not made) and how they are implemented. The state of the art will be reviewed and then applied to the developing agenda in a case study site, to see the extent to which current knowledge 'fits' the case. This will be done by choosing one or two example issues and following these through, writing a history based on interviews with the major actors, documentary evidence and observation of public meetings. The aim will be to develop an understanding of actual policymaking processes so that it is possible to anticipate the types of policy options and scenarios that need to be modelled and analysed in streams 1 and 2.

3.6.2 The role of models in policy making

In contrast to the first project, which will mainly be backwards looking, that is, tracing the history of decisions already made or soon to be made, this project will be forward looking. It will hypothesise how the tools, techniques and simulations developed in the ERIE programme could be integrated into

policy making processes, and test this using a combination of trials (e.g. small scale experiments with prototype ‘serious games’), discussions with the people involved in the policy process concerned with one selected issue about how they see models fitting into their work, and workshops with stakeholders about likely policy goals and options that they would want to explore. The results of this project will be fed back to streams 1 and 2. An example of the type of problem that this project will address is the ‘Catch-22’ problem of using models: if a simulation suggests outcomes that are already expected, policymakers question the value of using a simulation; if it suggests unexpected outcomes, policymakers doubt the validity of the simulation.

3.7 Integration and cross-linkages

The ERIE programme has been planned to have strong links between the streams and with disciplinary inputs spread across the streams. In summary, the case study projects in stream 3 will provide data as a basis for the creation of agent-based models in stream 2. These models will be validated and their properties analysed using the tools and techniques developed in stream 1. Requirements about the range of policy options to be simulated using the agent-based models, and about the way in which the models could be used in the policy-making process, will be developed in stream 4 and will inform the work of streams 1 and 2. Stream 4 will use data about the case studies collected in stream 3 as examples of policy-making.

Streams 1 and 2 will primarily draw on the mathematical and computer sciences, stream 3 on engineering, ecology and social sciences, and stream 4 on sociology, regional geography and economics, but while these are the primary disciplines involved, there will be extensive cross-disciplinary links, reinforced by the involvement of at least two Investigators, from different disciplinary backgrounds, in the work of each stream.

3.8 Brief Biographies of the Named Investigators

3.8.1 Professor Nigel Gilbert (Principal Investigator)

Nigel Gilbert is Director of the Centre for Research in Social Simulation (CRESS) at the University of Surrey. He is experienced in managing large collaborative projects, having been co-ordinator of a very successful €5 million European Commission FP5 project, FIRMA (Fresh water Integrated Resource Management with Agents) and is currently the co-ordinator of the FP7 QLectives Integrated Project (€7 million, with 8 partners). He has been the Principal Investigator of over 30 projects, funded by the Research Councils, industry and the European Commission.

Nigel was one of the first to realise the potential of agent-based models (ABM) for the social sciences and has written two textbooks on social simulation as well as many papers. His main theoretical contribution has been on ‘emergence’ in social phenomena; his main methodological focus has been on appropriate levels of abstraction for policy-relevant modelling; and his main substantive interests have been on modelling innovation. He is the founding editor of the Journal of Artificial Societies and Social Simulation and past-president of the European Social Simulation Association.

Nigel is a Fellow of the Royal Academy of Engineering and an Academician of the Academy of Social

Sciences. From 1998 to 2005, he was a Pro Vice-Chancellor at the University of Surrey.

3.8.2 Dr Lauren Basson (Co-Investigator)

Lauren Basson is a member of the Centre for Environmental Strategy (CES) which has a well-established international reputation for working in multidisciplinary teams allowing collaboration campus-wide, nationally and internationally on strategic and policy orientated research projects. Lauren holds BSc, MSc and PhD degrees in Chemical Engineering which provide a strong foundation in industrial and environmental systems analysis. She is an active member of the International Society for Industrial Ecology and was co-organiser of the Complexity Stream of the Conference of the International Society for Industrial Ecology (June 2007, Toronto, Canada) and associated Complexity Workshop. Lauren was also co-editor of the Special Issue on "Industrial Ecology and Complexity" of the Journal of Industrial Ecology, 13(2) (April 2009). Lauren is supervising two EngD students with the UK National Industrial Symbiosis Programme (NISP). Her close relationship with NISP (and particularly Link2Energy which is the NISP delivery partner working in the South Humber Bank) and access to information on industrial and regional resource flows and governance structures will be drawn on for the Case Studies.

3.8.3 Dr Angela Druckman (Co-Investigator)

Angela Druckman is a Research Fellow in the Centre for Environmental Strategy at the University of Surrey and a member of the ESRC Research Group on Lifestyles, Values and Environment (RESOLVE). Angela read Engineering at Cambridge University and has a PhD from University of Surrey (Environmental Strategy). She is a Chartered Engineer and has been awarded two patents for her work in electronics. She has had a wide variety of experience, including lecturing in electronics, commercial sector management for a thriving IT company, and non-governmental advocate on climate change (Friends of the Earth).

Angela's research interest is in developing an evidence base and tools for devising policies for sustainable consumption and production. She is currently involved in developing an energy-economic accounting framework capable of mapping both the socio-economic and energy/carbon dimensions of UK lifestyles.

3.8.4 Dr Rebecca Hoyle (Co-Investigator)

Rebecca Hoyle is a Reader in the Department of Mathematics at the University of Surrey. Her research focuses on modelling and nonlinear systems approaches to problems in the natural and social sciences. Her publications include a graduate textbook on pattern formation in nonlinear systems and numerous research papers. She currently holds a Leverhulme Trust Research Fellowship to work with the Systems Biology group in the Faculty of Health and Medical Sciences on 'Dynamic modelling of the switch to dormancy in Mycobacterium tuberculosis', which involves the equation-free bifurcation analysis of stochastic flows on a complex network of biochemical reactions. Her other research interests include dynamics in coastal morphology, molecular motors and surface chemistry and modelling the impact of school league tables. She also has experience in commercial consultancy of diverse modelling problems from the analysis of golf swings to the design of an eye surgery machine. She has been an investigator on six grants from the Research Councils and charities (five as Principal

Investigator).

3.8.5 Professor Paul Krause (Co-Investigator)

Paul Krause is Professor of Software Engineering at University of Surrey. He has over twenty years' research experience in software engineering, in both industrial and academic research laboratories. Prior to moving into software engineering, he worked as a research physicist for 10 years. Currently his research work focuses on distributed systems for the Digital Ecosystem and Future Internet domains. This is currently supported by the FP6 funded OPAALS project, and was supported by the FP6 Digital Business Ecosystems project before that. His current work on response and stability of digital ecosystems draws on his mathematics, physics and computer science background, as well as his life-long volunteer work as a natural historian and ecologist.

3.8.6 Dr David Lloyd (Co-Investigator)

David Lloyd is a Lecturer at the University of Surrey in the Department of Mathematics. His research interests are in the development of theoretical and numerical methods for the understanding of dynamics in large-scale systems. His most significant contribution has been in the development of a bifurcation theory for multi-dimensional localized patterns in general reaction-diffusion systems with immediate application to physical experiments.

3.8.7 Dr Sotiris Moschoyiannis (Co-Investigator)

Sotiris Moschoyiannis is a Research Fellow (starting as a Lecturer on 1 September 2009) in the Department of Computing and a member of the Digital Ecosystems research group. His research focuses on modelling the behaviour of communicating systems and his contributions range from foundational work on computing as something that happens by and through interaction to providing formal underpinnings for more pragmatic approaches to software design.

Sotiris has been actively involved in the Digital Ecosystems research initiative through the EU FP6 Digital Business Ecosystem (DBE) project and the multi-disciplinary OPAALS Network of Excellence, which brings together computer, social and natural scientists interested in both the development of a technological infrastructure and the socio-economic context in which it operates. In OPAALS he is the chair of the Integration Coordination Team which aims to identify synergies between disciplines and oversee their implementation.

3.8.8 Dr Anne Skeldon (Co-Investigator)

Anne Skeldon is a Lecturer at the University of Surrey. She has extensive experience of research on dynamical systems from fundamental work on the bifurcations of patterns with symmetry to relating theoretical and computational work to real experiments. Notable publications have been on patterns with planar Euclidean symmetry and on mechanisms for pattern selection in parametrically excited systems. Her other research interests include modelling of oxygenation of cancer tumours. From September she will be supervising a PhD student funded by NERC through the data assimilation theme of the National Centre for Earth Observation led by Professor Ian Roulstone at Surrey studying models and data assimilation in the carbon cycle.

3.9 REFERENCES

[27] R. Albert, J. Jeory and A Barabasi (2000) *Nature* 406 378-382.

[28] B. Bergeron, (2006). *Developing Serious Games*, Charles River Media.

[29] D.S. Callaway, M.E.J. Newman, S.H. Strogatz, D.J. Watts (2000) *Phys Rev Lett* 85 5468-5471.

[30] Defra, Ensuring the UK's Food Security in a Changing World, HMSO, 2008

[31] R. Erban, I. Kevrekidis and H. Othmer, *Physica D*, 215,1-24 (2006).

[32] A.K. Griffith and N.K. Nichols (2001) *J. Flow, Turbulence and Combustion*, 65, 469-488

[33] T. Gross, C Dommar D'Lima, B. Blasius (2006) *Phys Rev Lett* 96, 208701-4.

[34] K. Judd and L.A. Smith (2004) *Physica* D, 196 224-242.

[35] H. Kelly, K. Howell, E. Glinert, L. Holding, C. Swain, A. Burrowbridge, M. Roper (2007) *Comm. ACM* 50(7) 45-49.

[36] M.Z.Kwiatkowska, G. Norman and D. Parker (2004) *Softw. Tools for Technology Transfer* 6(2):128-142, Springer.

[37] E. Letier, J. Kramer, J. Magee, S. Uchitel (2008) *Autonomous Software Engineering* 15:175-206, Springer.

[38] M. Lim, D. Braha, S. Wijesinghe, S. Tucker, Y. Bar-Yam (2007) *Euro Phys Lett*, 70, 58005-6.

[39] A. Makeev, D. Maroudas and I.G. Kevrekidis, *J. Chem. Phys*., 116, 10083-10091 (2002)

[40] M. Williams, P.A. Schwartz, B.E. Law, J. Irvine and M.R. Kurpius (2005), *Global Change Biology* 11, 89-105.

4 MILES: MODELS AND MATHEMATICS IN LIFE AND SOCIAL SCIENCES

4.1 OVERVIEW

The ERIE project has a specific focus on industrial ecosystems and food security. However, as reflected in the first two chapters of this report, the study of complex adaptive systems is heavily influenced by the use of analogies and metaphors in the hope that these will lead to the development of novel and insightful new models. In this chapter we describe a parallel project that has a specific focus of stimulating creative interactions across a wide range of disciplines.

Mathematical contributors to the social and life sciences typically aim to provide insight by building a model and solving it using ingenious techniques and computer programmes whose details can be inaccessible to researchers in the host discipline. Mathematicians equally can be surprised to discover that these disciplines themselves use quite different kinds of models, in unfamiliar ways to seemingly perplexing ends. Revolutionary progress comes when researchers from all the relevant disciplines create new kinds of models together, owned and exploited by them all. To this end we have obtained three years of funding for a programme of networking, idea generation and collaboration activities focused on modelling approaches to the life and social sciences, their synergies and dissonances. This EPSRC grant provides funding over three years for: pump-priming research fund prize pools; the running of a monthly *Café Scientifique*; Networking Events; Multidisciplinary workshops; and a Visiting Scholar programme.

We shall explore the different types of model used in different disciplines, the extent to which the models themselves, modelling methods (ways in which models are created and validated), methodologies (the philosophy of science behind modelling) and ways of using models (e.g. to promote understanding or for prediction) can be transferred between disciplines, and when and how to create entirely new modelling frameworks.

Analogies between biological and social systems are very tempting and have resulted in, for example, the development of “industrial ecology" as a new paradigm for socio-economic systems. Mathematicians see network models as widely applicable in both contexts, from metabolic networks to ecosystems to maps of interactions between agents. However, the extent of the appropriateness of these analogies and models remains to be fully explored. We need to appreciate better which aspects of a particular system are of special interest or value to the researchers who study it, and why. For example the role of “reflexivity" in social systems – the extent to which the behaviour of social agents is modified by their perception of the “emergent” properties of the system – is of great interest to social scientists. We need to understand the extent to which this differs from global feedback in physical or biological systems, and why, or indeed whether, that matters.

Within our broad framework, we shall emphasise three interrelated themes of particular interest to research groups in Surrey University, where different disciplines have distinct perspectives. These are outlined now.

4.2 SUSTAINABILITY

Societal and biological pressures meet head on in the quest for sustainable living. This is a broad area ripe for modelling. Three examples involving current research at Surrey illustrate some of the questions we might address:

4.2.1 THE CARBON ECONOMY

Carbon governance and footprinting, and the psychology of energy behaviours are studied by the RESOLVE project at Surrey. Meanwhile, Dr Anne Skeldon (Mathematics) is working on data assimilation for carbon cycle models that are controlled by biological and socioeconomic factors, Drs Alfred Thumser (Biochemistry) and Claudio Avignone-Rossa (Microbial Sciences) are working on microbial fuel cells in order to use waste to produce electricity efficiently and Dr Lauren Basson (Centre for Environmental Strategy) studies industrial symbiosis: the sharing of materials, energy, assets, logistics and expertise between organizations in order to reduce environmental impact. MILES would bring together these and other interested researchers to share perspectives and create new modelling approaches.

4.2.2 WATER:

All four faculties at Surrey are involved in water-related research, including fluid dynamics and meteorology in Mathematics, flooding and hydrology in Engineering, water supply safety (actual and perceived) in Psychology, Economics, Engineering, CES, Chemistry and the Postgraduate Medical School, and water policy, regulation and management in Law, Sociology and CES. Modelling is already central to some of these research activities, while in others there is scope for its greater involvement. Models in one area may be highly relevant to those in another, e.g. as boundary conditions. However there is, as yet, little interaction between the models, or modellers, in the different areas. Surrey is currently building water-related activity through initiatives such as a *Water Research Network*, focused on the provision of “safe” water, in particular for disaster relief, the appointment of a new chair in Water (Engineering) and an advertised chair in Environmental Flow Modelling (Engineering/Mathematics). MILES is especially timely in bringing together existing water modellers and water researchers in need of models from across this research theme.

4.2.3 INDUSTRIAL ECOLOGY

This multidisciplinary field draws an analogy between biological ecosystems and the web of relationships between firms, industries, people, environment and governance structures in the economy. The ERIE project, including Dr Lauren Basson (CES), Dr Hoyle and Profs. Gilbert and Krause, focuses on agent-based and network models of industrial ecosystems and global food supply webs. MILES will extend the existing activity in this area, by examining the relationship of this social science approach to the kinds of models that are typically applied to true ecology, for example, deterministic and stochastic differential equations and evolutionary game theory, an area where Dr Hoyle is working with biologists. From another point of view entirely, researchers in Microbial Sciences might see food supply and security a high priority of UK and international policy makers – as primarily involving the control of livestock and crop pathogens. It will be interesting to investigate the interplay between these two outlooks.

4.3 THE IN SILICO CELL

The phrase, "the *in siloco* cell" encapsulates the idea of modelling in a computer programme all the biochemical reactions and processes that occur in a living cell. This approach is useful because *in silico* experiments are faster and more flexible in design than those that take place in real cells and in the case of pathogens, are also much safer.

Prof. McFadden and his group have constructed the first genome scale computational model of the metabolism of Mycobacterium tuberculosis (MTB), the bacillus that causes tuberculosis in humans, and it is currently being used to identify potential novel drug targets. Recently, Dr Hoyle has joined this project contributing her expertise to the analysis of dynamic models of the gene regulatory networks in MTB and other microganisms. Professor McFadden has also, with Prof. Ravi Silva (Advanced Technology Institute), recently founded a bionanotechnology laboratory within the University, an interdisciplinary venture between engineering and biosciences currently focused on development of novel chemotherapeutic agents for cancer, that would benefit from greater mathematical and modelling expertise. Also in Microbial Sciences, Profs. Michael Bushell and Colin Smith and Dr Avignone-Rossa have made pragmatic use of genome scale metabolic network models for bioprocess design of antibiotics manufacture. As the genome sequence of each species is published, the theoretical metabolic repertoire of that species becomes available. BBSRC funded studies have led to a new industrial grant (from Eli Lilly) to scale the technology to process level and to recent EPSRC/BBSRC funding for the development of a new generation of biofuel systems. However, the mathematical/computational bases of these models are relatively simple and we have reached the limits of their functionality. What is now needed is detailed cross-faculty engagement to develop the next generation of in-silico cells. Other areas for possible cross-disciplinary collaboration are stochastic models of molecular interaction network dynamics and statistical analysis of the complex datasets resulting from high throughput functional genomics studies.

Networks, stochasticity and data analysis are core to all these projects and also to the analysis of complex social systems, such as those that are being studied in the ERIE project. Thus it would be natural to consider how stochastic biochemical networks compare to those found in the social sciences and investigate whether exchange of modelling technologies might be fruitful. This is an example of a synergy between systems biology and complexity science, both of which are currently growing at a phenomenal rate, both internationally and at Surrey, and where there is great potential for cross-fertilisation.

Modelling of whole tissues can also aid understanding of important clinical problems. Prof. Chris Fry (Postgraduate Medical School) studies the abnormal spread of electrical signals in cardiac and smooth muscle, which can lead to dysfunction such as cardiac arrhythmias and the overactive bladder. Characterisation of the biophysical properties and mapping the flow of electrical current in these complex tissues, using experimental measurements and mathematical modelling, is key to the analytical approach and may benefit from the engagement of modelers with experience in complex systems.

4.4 Mathematical and computational techniques

In addition to focusing on application areas our programme will also turn the spotlight back onto mathematics and computer science *per se* to discover what can be learnt about core concepts, such as those below, from the many different models in biology and social science that incorporate them in varied manner.

Data. In systems biology, there is so much data that it is hard to know how to use it well. In social science, data might take a form – e.g. a narrative – that makes it hard to incorporate into a model. In both cases, it may be incomplete, sparse in places. MILES will explore the emerging solutions to these problems in relation to mathematical, statistical and computational approaches to data assimilation and analysis.

Stochasticity, discreteness, spatial variation and rare events. Traditional ordinary differential equation models of biological and social systems are being replaced by a suite of models that capture a whole host of awkward (and often interrelated) real-world effects, such as randomness (stochasticity) in the timing of events and the choice of interactions (e.g. in biochemical reactions or social exchanges), interacting entities that may be present in small numbers (discreteness) and/or be unevenly distributed in space (e.g protein and mRNA molecules in the cell), and rare events with important consequences (e.g. stock market crashes, DNA mutations). We aim to encourage this stimulus to new theoretical development in our programme by pairing theorists with researchers who have technically challenging real-world problems to address.

Network dynamics. The interplay between the dynamics on a network and the network structure is highly sensitive to the details of both. Thus one can expect network models arising in specific real world contexts to generate new insights into complex network dynamics that we would like to capture. Within this context we hope to gain some insight as to how emergence and reflexivity in social complex systems relate to global phenomena and feedback in the network dynamics of physical and biological systems, for example in neuroscience and telecoms, and hence why (or whether) the former are inherently social phenomena.

Biomimetics and sociomimetics. Evolutionary pressures drive towards increased efficiency in biological organisms. The idea of borrowing Mother Nature's designs is thus appealing. In computing, this approach has led to artificial neural networks and neurons and swarm intelligence. In space science it is used to develop low power and energy mechanisms such as a drill based on wasps and locusts (Dr Y. Gao, Surrey Space Centre). Similarly "sociomimetics" was coined to describe the design of electronic information systems following the patterns of communication in natural societies. A similar principle can be used, for example, to design footpath layout in communal spaces following the pattern made by many pairs of feet in natural usage. MILES will act as a conduit for the influence of biological and social design on mathematics, computing and engineering.

In all three cases, the MILES programme aims to stimulate ideas that exploit the synergies identified between modelling approaches in the biological and social sciences, and also clarify the essential differences that must be respected. The themes are deliberately broad, and we have described a great

variety of potential avenues of investigation to illustrate the wide range of interesting research questions in this area. Of course no-one can predict where the next brilliant idea will come from, and the more transformative it is, the more surprising it is also likely to be. The purpose of MILES is to generate projects and collaborations that we cannot foresee, so the themes are intended as talking points to intrigue and attract clever, creative people, bringing them together in new ways and combinations, rather than as descriptions of research projects that we expect to emerge.

5 AGILE DATA CENTRES FOR THE DIGITAL ECONOMY

5.1 INTRODUCTION

This section reproduces two recent papers that have come out of the on-going collaboration between the Department of Computing at the University of Surrey, and Memset Ltd. Memset is the UK's first carbon neutral data centre, and is also distinguished in having developed an architecture that enables them to compete favourably in terms of cost and performance with large-scale data centres. We believe that stimulating development of regional data centres is an important activity in support of the evolution of the digital economy, in order to retain and open and fair marketplace for * as a Service.

The first of the two papers introduces Memset in more detail, and provides an extensive evaluation of their architecture. The second paper addresses some of the environmental aspects of IT and datacentres.

5.2 THE CASE FOR MEDIUM-SIZED REGIONAL DATA CENTRES

Kate Craig-Wood
Memset Ltd, Surrey University
25 Fredrick Sanger Rd.
Guildford, UK
+44 1483 608012
kate@memset.com

Prof. Paul Krause
Dept. of Computing
Surrey University
Guildford, UK
+44 1483 689861
p.krause@surrey.ac.uk

Nick Craig-Wood
Memset Ltd.
25 Fredrick Sanger Rd.
Guildford, UK
+44 1483 608013
nick@memset.com

ABSTRACT

Cloud computing is widely associated with major capital investment in mega data centres, housing expensive blade servers and storage area networks. In this paper we argue that a modular approach to building local or regional data centres using commodity hardware and open source hardware can produce a cost effective solution that better addresses the goals of cloud computing, and provides a scalable architecture that meets the service requirements of a high quality data centre.

In support of this goal, we provide data that supports three research hypotheses:

1. **that central processor unit (CPU) resources are not normally limiting;**
2. **that disk I/O transactions (TPS) are more often limiting, but this can be mitigated by maximizing the TPS-CPU ratio;**
3. **that customer CPU loads are generally static and small.**

Our results indicate that the modular, commodity hardware based architecture is near optimal. This is a very significant result, as it opens the door to alternative business models for the provision of data centres that significantly reduce the need for major up-front capital investment.

Keywords

Cloud computing; Virtualisation; Data centres; Utility computing; Infrastructure as a Service.

5.2.1 BACKGROUND AND RATIONALE

A key factor in the dramatic increase in interest in Cloud Computing has been the offering of Cloud services by major keystones in the digital economy: Amazon, Google, Microsoft, Sun. Other less publicly known providers, such as Rackspace, have joined this space. However, there is a widespread perception that "Cloud" means a flexible rental scheme for compute resource from a single mega-data centre. This view is well articulated in the Berkeley take on Cloud Computing where it is argued that:

> "the advantages of the economy of scale and statistical multiplexing may ultimately

lead to a handful of Cloud Computing providers who can amortize the cost of their large datacenters over the products of many "datacenter-less" companies." (Armbrust et al, 2009)

One of the arguments for this is by analogy with the semiconductor industry's move to a small number of high cost fabrication lines supporting a large number of "fab-less" semiconductor innovators. However, there is a key distinction in the form of business relationship in the two cases. In the semiconductor industry, the relationship between requestor (fab-less semiconductor innovator) and supplier (mega fabrication line) is asynchronous. The customer can, within limits, absorb an unexpected delay on the supplier side. In addition, the customer could mitigate this risk by contracting more than one supplier: there is a standard "interface" between designer and manufacturer. This provides a critically important duality. The supplier can time-multiplex the use of expensive fab lines in order to secure return on investment. In return, the consumers can choose and transfer their relationships between a choice of fab plants in event of loss, or reduction in quality, of service.

In contrast, the relationship between a "datacentre-less" service innovator and a cloud provider is synchronous. As witnessed by the 22-hour downtime of Microsoft Azure on 13/14 March 2009, an interruption or loss of quality on the supply side will immediately and directly impact on the revenue generated by the customer's service. In addition, the interface between supplier and consumer is typically non-standard. We are not really seeing the resilience of a true cloud. Instead, the consumer sees some important benefits of flexibility and scalability, but has a very tight dependency on their chosen cloud provider – the relationship is highly asymmetric in the degree of empowerment to mitigate risk.

In fact, it turns out that the arguments for economy of scale are largely fallacious. We challenge the vendor-led view that expensive blade servers and/or storage area networks are needed for a virtualised platform. The use of commodity hardware, combined with the use of open source software for the platform can deliver cost effective utility computing.

Armbrust et al (2009) provide the following cost comparison (quoting (Hamilton, 2009)) between a medium-sized data centre and a mega data centre:

Technology	Cost in Medium-sized DC	Cost in Very large DC	Ratio
Network	$95 per Mbit/sec/month	$13 per Mbit/sec/month	7.1
Storage	$2.20 per GByte / month	$0.40 per GByte / month	5.7
Administration	≈140 Servers / Administrator	>1000 Servers / Administrator	7.1

Our figures from Memset, a successful example of a small-sized data centre, show that the economies of scale can be realised through the use of commodity hardware rather than powerful blade servers.

Technology	Memset (small DC)
Network	£14 ($21) per Mbit/sec/month
Storage	£8.83 ($13.30) per **TByte** / month
Administration	≈385 Servers / Administrator

The only area where there is a significant economy of scale (only a factor of 2.6) for a mega data centre is in the number of servers that can be maintained by a single administrator.

Furthermore, a significant body of data supports the effectiveness of a data centre architecture based on commodity hardware. Hourly samples of CPU utilization, disk read/write data rate and disk transactions per second (TPS) were measured for 950 individual virtual machines (VMs) on 45 virtual machine hosts (VMHs) in Memset's (see below) real-world deployment. The users and applications of the VMs ranged across all sizes and types of business. The VM platform was managed using open source Xen 3.2, and the VMHs were low-cost single-CPU commodity server hardware.

Analysis of the data shows that CPU availability is very rarely a limiting factor; directly challenging the widely-held view that CPU resources are the main limiter in VM deployments. The paper also shows that TPS, while rarely limiting, are more frequently limiting in the architecture used. This supports the hypothesis that one should aim to keep the TPS-to-CPU ratio high, supporting the practice of using a large number of commodity servers with local disks.

Overall, this paper argues that the move to utility computing is better served by a move to more "agile" local or regional data centres, built in a modular fashion. The data provided in the paper provides strong supporting evidence for this claim.

5.2.2 MEMSET

Memset, a managed hosting company, started renting virtual servers (Miniserver VM®) to consumers and business in late 2002, originally using User Mode Linux, but in 2005 moved to the open source Xen hypervisor (now using version 3.2). This practice has been popularised by a growing acceptance of virtualisation as "industrial strength", and is now more commonly referred to as Cloud Computing or Infrastructure as a Service (IaaS).

Over the last 7.5 years Memset has accumulated a set of best practices for optimal environments for hosting large numbers of virtual machines. However, many of those practices go against widely held beliefs, and this research is intended to analytically test those views.

Despite its leadership in the field of virtual servers, Memset is a relatively small company. They are turning over £2m, profitable, debt-free and growing at roughly 30% per annum. Their share of the approximately £1bn UK hosting market is therefore 0.2%. However, despite their size they are able to offer the highest quality services; voted UK's Best Web host 4 years running by PCPro magazine, and consistently in the top 10 global rankings of hosting company reliability according to the Netcraft survey. They also currently boast the lowest prices for virtual server rental in the UK, compared against the open information on other providers' Web sites. This makes them an ideal test case against the growing view that IaaS can only be delivered at commodity prices from mega-scale data centres. Memset's servers are housed in a 10,000 square foot 2MWatt data centre facility.

5.2.3 COMMODITY HARDWARE AND OPEN SOURCE SOFTWARE PLATFORM

Accepted wisdom of the Memset systems administrators is that central processor unit (CPU) resources are not normally limiting (research hypothesis #1), but rather disk I/O transactions per second (TPS) are more often limiting (hypothesis #2). RAM is a statically provisioned resource in the Memset architecture, and is fast-becoming the unit of comparison between VM providers.

Their rationale is that the limiter is normally the speed at which the disk head can physically seek[2], measured in TPS, rather than CPU which appears to be the general belief, based on conferences and debates attended by the authors.

As a consequence of the beliefs of their systems administrators, the Memset virtual machine (VM) hosting architecture does not follow vendor-led wisdom of expensive blade arrays for huge CPU capacity, with storage on even-more expensive disk arrays accessed over fibre channel. Instead, Memset uses commodity 1U Dell PowerEdge servers, with two disks in a RAID(1) configuration.

The latest VMH specification is a R300, with single quad-core 2.5GHz Intel Xeon processor (minimum spec), 2 x 2,000GB 7,200 RPM 3.5" SATA disks software RAID(1) and 6 x 4GB DIMMs (total value approx. £1,200). Faster (eg. 15k RPM 2.5") disks are not employed because they are prohibitively expensive , are only an option in larger, multi-disk chassis, and overall give a much less cost-effective TPS-to-RAM ratio.[3]

Each VMH runs the Xen hypervisor. Its domains (VMs) run their disk off a local disk array using the Logical Volume Manager on a RAID1 array. Their CPU scheduling parameters are set and their RAM allocations are fixed to make them as independent as possible. CPU is fairly shared among the VMs in proportion to the amount of RAM allocated to them. Disk TPS is shared fairly, but is not shared in proportion to RAM, using the Linux kernel's CFQ ("completely fair scheduler") I/O scheduler.

One oft-cited reason for using a shared storage array for the VM disk images is to facilitate real-time migration of VMs between front end machines. Memset have not found there to be a requirement for this among their customers, however, since those that need resilience simply rent two VMs on different VHMs with fail-over between them. Additionally, the CPU load profiles for most customers appear to be quite static and small (research hypothesis #3), which further reduces the need for live VM migration.

To summarise, Memset use commodity hardware and local VM images for five reasons:

1. It is cheaper.
2. Costs scale in smaller, more manageable increments.
3. Disk TPS are believed to be limiting, thus local RAID(1) should be better.

[2] A 7,200 RPM disk will rotate 120 times per second. On average, a seek will take half a rotation. Therefore, when using RAID1 which effectively doubles the transaction capacity, the peak TPS is expected to be a little above 480 due to intelligent caching and request sequencing.

[3] Memset has been experimenting with Solid state (flash) drives (SSDs) as an alternative to rotary disk drives since they would appear to be the answer to TPS-bound applications. In their tests, however, they found SSDs performed no better than high-end rotary disk drives in seek-intensive operations. The issue appears to be the large, fixed 128kByte read/write block size of SSDs, meaning that even for a tiny read operation an entire 128kB block has to be read, and that additional bandwidth clogs up the serial interface. Memset anticipates that the issue will be resolved in time, perhaps with placement of SSD on the PCI bus, with better caching or with a reduction in the SSD read/write block size itself.

4. 2x 7.2k disks chassis (vs. 4/6/8 15k disks) gives the most cost-effective TPS-to-RAM ratio.
5. There is no demand for live/hot VM migration.

5.2.4 DATA INCLUDED IN STUDY

VM host (VMH) servers are used exclusively for either Linux or Windows VMs, there being 36 and 17 respectively of each type in the study. Due to unreliability with the open source paravirtualised Windows drivers (found during Memset's testing), only the Linux VMH use paravirtualisation. There is, therefore, expected to be an additional performance overhead associated with the virtualisation layer on the Windows hosts compared with the Linux hosts.

Until April 2009 the VMs were provisioned with swap partitions enabled. This greatly exacerbated the disk I/O issues, and it was frequently (and painfully) obvious that the host servers were I/O bound with disks "thrashing". Since then, with the rapid reduction in RAM prices making it more feasible, swap has been disabled on the Linux VM range (not on Windows VMs and VMHs)

This study only looks at the newer VMH deployed since April 2009. To compensate for the lack of swap, and in anticipation of disk I/O being less loaded, Memset increased the RAM in the host servers to 16GBytes. This also increased the number of VMs deployable on each host server.

During the course of this study it was noted that even with the increased number of VMs per host the servers' resources were being significantly under utilized. Therefore Memset increased the RAM in new VMH to 24GByte during the study (mid-January 2010), allowing 50% more VMs to be hosted on each VHM. This study examines the 21 large (24GB) VMHs and 36 small (16GB) VMHs independently where suitable since it is anticipated that they have different usage patterns.

Generally, Memset's VM customers are motivated more by cost savings than over-specification of the server, otherwise they would rent a physically dedicated server from the company. Therefore, this user group should represent a higher workload for a given amount of computer resource than, for example, a virtual machine deployment servicing an internal corporate demand.

5.2.5 METHODS

5.2.5.1 POWER AND BANDWIDTH DATA COLLECTION

From bench testing of server equipment, Memset has also shown that Dell PowerEdge servers power draw only changes appreciably from the idle state with increased CPU utilization, and not greatly with increased disk activity. A Dell R300 like those used as VMHs (but with less RAM in this

case) draw 87 Watts when idle, 131 Watts when at maximum CPU, and 137 Watts when both CPU and disks are at maximum. Those results suggest a 50.6% power draw increase possible due to increased CPU activity and a possible 6.9% increase from increased disk activity.

This knowledge, combined with the observation by Memset systems administrators that there is little diurnal variation in their power requirements, suggests a crude test of CPU utilization. Therefore, the hourly power draw across all servers was recorded, and was compared against the total bandwidth use across all servers as a very crude measure of activity.

Most of Memset's customers are UK-based,

and their customers' users are also mostly UK-based. The majority of the applications hosted by Memset are Web sites or Web applications, therefore bandwidth should be a reasonable approximation of average load across the servers. This data includes fully dedicated servers as well as VHMs, and also includes older VHM machines.

Both sets of data were already being collected via the Memset infrastructure and stored in their central database, and was extracted with simple SQL queries. If CPU utilization is generally low and not limiting little-to-no correlation between bandwidth and power draw is expected (hypothesis 3).

5.2.5.2 VM RESOURCE UTILIZATION DATA COLLECTION

For the main research project (testing hypothesesies 1 and 2), a set of scripts were written, using Python and some Linux shell, to query the VHMs and log resource utilization data hourly.

The Xen hypervisor provides accurate accounting data for CPU usage since it is involved in scheduling the Virtual CPUs. Its user space tools from the master domain were used to collect CPU usage figures (both spot and cumulative).

The Linux kernel on the master domain accounts for the disk traffic to and from the LVM disk array. Which domain is attached to which logical disk and then use standard Linux userspace tools to account for the disk traffic. This includes total disk reads and writes, spot disk I/O and spot transactions per second (TPS). Table 1 shows the data fields collected.

Field	*Description*	*Source*
created	The date and time the info was collected.	
host	The VHM that this info was collected from.	
name	The Miniserver VM name for which this data was collected (hidden in results).	
id	The xen domain id - a number just used for identification purposes	xm list
mem	MBytes of RAM for the VM. This is a fixed allocation without ballooning.	xm list
vcpus	Number of Virtual CPUs assigned to this VM.	xm list
stat	Spot status of the VM - whether it is running or not.	xm list
cpu_time	Total amount of CPU time the VM has used (since last reset)	
cpu_pc	Instantaneous percentage of CPU the VM is using.	xm top
tps	Average disk transactions per section.	iostat
read_kbps	Instantaneous kBytes/s read from the disk.	iostat
write_kbps	Instantaneous kBytes/s read from the disk.	iostat
kb_read	Total kBytes read from the disk.	iostat
kb_written	Total kBytes written to the disk.	iostat

TABLE 1: DATA FIELDS COLLECTED AND THEIR SOURCES.

5.2.5.3 DATA ANALYSIS

The VHM were grouped into four classes (Linux 16GB, Linux 24GB, Windows 16GB and Windows 24GB) so that comparisons were with-like. These groups are referred to as the VMH-groups.

Of the variables available, only CPU percentage, TPS, and spot read KBps (R-KB/s) were examined in detail (the core metrics). Memory utilization profile and VM count was also noted for some data sets.

In total, a little under one gigabyte of data was collected over 8 months. 5.1 million individual VM-hour samples from 1,416 different VMs across 53 VMHs. This volume of data necessitated a programmatic approach to the analysis. Therefore, the data was organized in a SQlite database, and a set of Python programs were designed and written which could break down the data sets and produce the following:

- **Averages**: Simple average of resource utilization for each core metric, within each VMH group.
- **Load bands distribution**: Average time spent at 5 percentile ranges for an individual core metric averaged across all VMH within each group which were at least 75% full (by RAM).
- **Find peaks**: Peak search for each core metric among each VMH-group. Where VMHs share the highest peak, finds one with most counts of peak.
- **Single VMH detailed**: Detailed data for each core metric on each VM for a specified VHM around a specified date-time (usually a peak). One chart per core metric, each showing that metric over time for each VM individually.

Names of individual VMs were removed for security and customer privacy since host names are discernable from the Memset VM naming scheme.

5.2.6 RESULTS AND ANALYSIS

5.2.6.1 Power and bandwidth

A plot of power draw (instantaneous) and bandwidth (averaged over the previous hour) against time over a period of five consecutive days (Friday to Tuesday) is shown in figure 1.

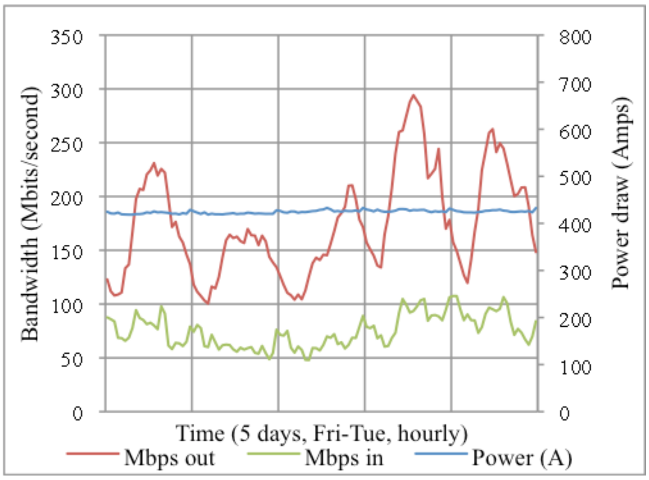

FIGURE 1: BANDWIDTH AND POWER OVER TIME

The diurnal bandwidth variance for the measured set of servers was significant, with the outbound data rate dropping below 75 Mbits/second (Mbps) at night and peaking at over 200Mbps during the day.

From figure 1 it is not possible to discern any relationship between activity (bandwidth) and power draw, therefore a second scatter plot of power draw at varying bandwidth levels was rendered (figure 2). The trend line on figure 2 (dotted, red) show the expected upward trend in power consumption with bandwidth, however it is a very small effect, only visible thanks to the large data set, which includes a total of 12 days' readings. Even taking the most extreme samples, the power variance with bandwidth is only 4.1%, suggesting that the CPUs are very lightly loaded and not appreciably moving out of an idle state even with the manifold increase in bandwidth activity.

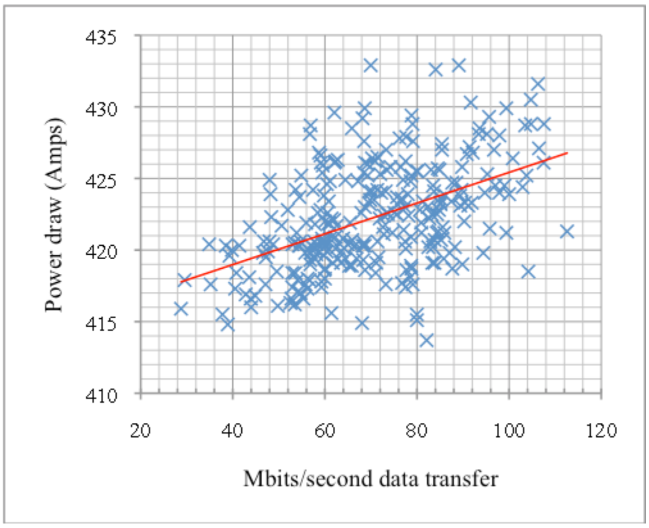

FIGURE 2: POWER AGAINST BANDWIDTH

5.2.6.2 Averages

Table 2 shows the average utilization of each core metric across the entire VMH set.

TABLE 2: AVERAGES

	Lin	**16GB**	*Lin*	**24GB**
	Raw	*%*	*Raw*	*%*
Host #	26		14	

CPU	46.3	11.6%	96.1	24.0%
TPS	46.2	9.2%	63.3	12.7%
R-KB/s	196	0.5%	362	0.9%
Mem MB	14,969	93.6%	22,262	92.8%
VM #	18.9		21.1	

	Win	**16GB**	***Win***	**24GB**
	Raw	*%*	*Raw*	*%*
Host #	10		7	
CPU	58.4	14.6%	81.8	20.4%
TPS	-	-	-	-
R-KB/s	315	0.8%	210	0.5%
Mem MB	14,264	89.1%	21,483	89.5%
VM #	14.8		16.8	

For calculation of CPU percentage utilization, 100% was taken to be was taken to be 400. For TPS 100% was taken to be 500 (the theoretical maximum) and for read KBytes/sec the observed maximum of 40,000 was used as the 100% figure.

As expected, the average resource utilization is overall higher on 24GB VMH compared with 16GB VMH, which can be attributed to a larger number of VMs on the 24GB hosts.

5.2.6.3 Peaks

Part of the analysis was to find peak states in the data. As well as locating useful time-slices on peaking VMH for further examination (see 3.6) this also yielded some differences between the Linux and Windows hosts.

5.2.6.3.1 Disk I/O statistics under Windows and Linux

There is a clear and large difference in the peak TPS and R-KB/s figures between Windows and Linux servers. This is because the Linux VMH have a paravirtualised driver, but Memset does not use one under Windows. Therefore when streaming from the disk each 512 byte block appears as one transaction, limiting the usefulness of TPS statistics on Windows VMHs. Under Linux, the kernel is more efficient, grouping adjacent blocks together to reduce the number of transactions when streaming data. That efficiency is clearly working through the Xen layer too. However, that does still suggest an unforeseen weakness in our measurement system. Even under Linux the TPS increased with data read rate, and streaming data results in high TPS readings, making the TPS readings less useful in supporting hypothesis 2.

5.2.6.3.2 Effects of paravirtualisation

There is also an apparent difference between the CPU peaks under Linux and Windows, with Windows VMH peaking at only 330.7%, well under the 400% theoretical maximum. Some of the difference can be seen in dom0's CPU utilization. The Linux hosts CPU utilization peaked at 399.7%.

Under Windows the device model (disk controller, ethernet driver emulation etc) requires some dom0 CPU, hence the 26.4% and 16.4% CPU utilization (6.6% and 4.1% of total) by dom0 in those instances. Because the Linux VMs are paravirtualised and there is no device model almost 100% of CPU is available to the VMs, with dom0 only using 0.1% of total CPU capacity. The ~50% CPU unaccounted for under Windows is believed to be the Xen Hypervisor performing the interprocessor communications between device model and VM kernels. This would not show in dom0's CPU utilization, since it is just another domain as far as Xen is concerned. The unaccounted for CPU under Windows is being directly used in the hypervisor.

5.2.6.4 Load band distribution

Figures 3 shows the average time spent at each 5-percentile CPU load bands for each VMH group.

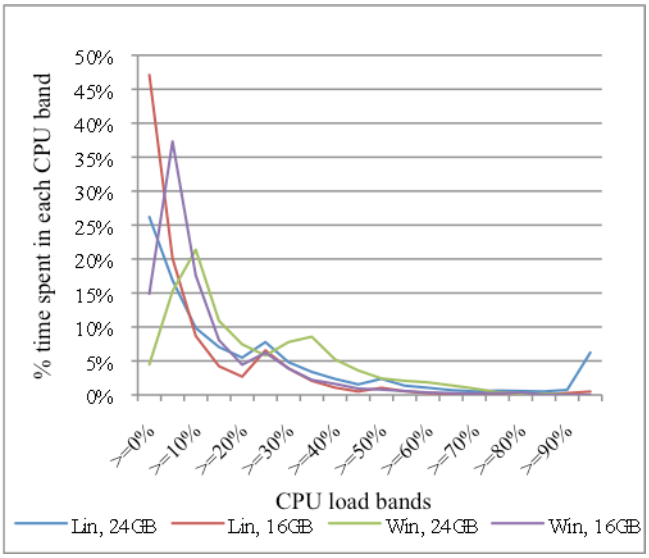

FIGURE 3: % OF TIME SPENT AT EACH 5% LOAD BAND.

The CPU load curve can be considered accurate, since the peak is definitive (400%), but the effects of paravirtualisation (see 3.3.2) should be kept in mind since this in-effect reduces the maximum CPU utilization under Windows. The sub-peak at 25% is symptomatic of VMs or processes maxing out individual cores on the host machines. This effect was marginal, however on the CPU-peak 24GB Linux VMH (figure 4) a small increase in time spent at higher CPU loads. In total, the 24GB Linux VMHs spent 10.1% of their time above 90% CPU load, whereas the 16GB Linux VMH spent only 1.3% of the their time above 90%.

The CPU load profile of hosts running Windows VMs shows them to be generally a little more loaded, which would fit with the lack of paravirtualisation, The sub-peak at low CPU utilization for Windows VMH could also be indicate a generally higher background CPU utilization of Windows operating systems. A sub-peak at >=95% utilization for 24GB Linux VMH suggests that with the increased RAM CPU is occasionally now becoming a limiter.

5.2.7 Detailed examination of peak states

Figures 4 and 5 show a selection of peak CPU states in detail, with each line representing an individual VM. In each case the peak being observed occurs at the time mid point, with the time range spanning 4.5 days before and after the peak moment. The different vertical scales on Linux (400%, figure 4) and Windows (100%, figure 5) should be noted. Both charts show 24GB VMH peaks, but the ones for the 16GB hosts were similar.

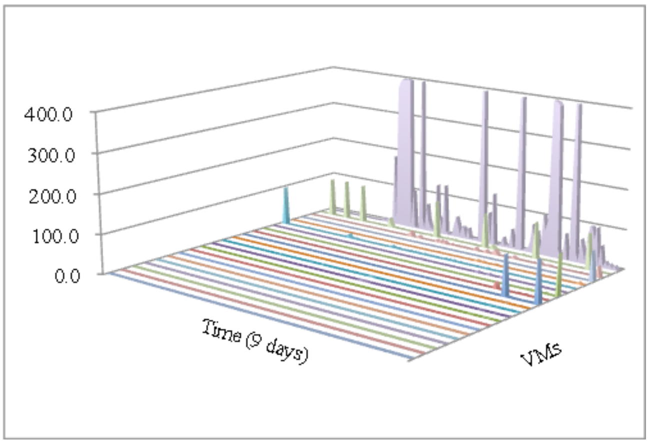

FIGURE 4: CPU PEAK, LINUX 24GB

For Linux VMH (figure 4 was typical of others examined) the CPU peak can be attributed to the activities of just one VM. The peak-finding algorithm was designed to find the most intensive (longest lasting) peaks even where many VMHs reached the same level, suggesting that (for the Linux deployment) coincidental VM CPU peaks are vanishingly rare. Those charts also demonstrate the general low CPU activity levels among the majority of VMs.

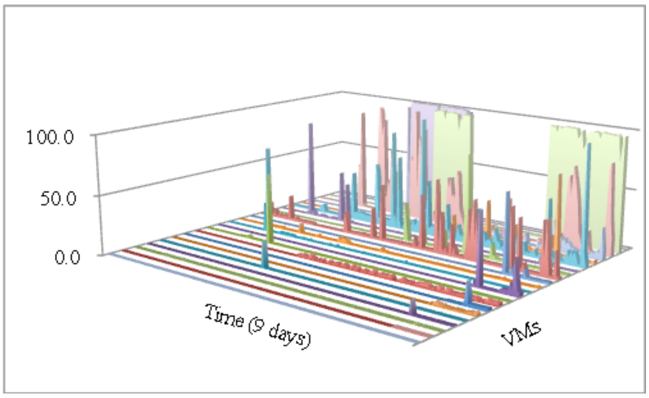

FIGURE 5: CPU PEAK, WIN 24GB

The Windows VMH CPU peak shown in figure 5 was also typical of other Windows host peaks, and is the result of a coincidence of individual VMs maxing out a single CPU-core (of which there are four available). This is partially an artifact of the Memset architecture since it was only towards the very end of the study that they augmented their systems to allow Windows VMs to have more than one virtual CPU core.

5.2.8 CONCLUSIONS

5.2.8.1 Hypothesis #1: CPU is rarely limiting

According to the averages and load band distribution charts, CPU is indeed rarely limiting, confirming Hypothesis #1. Further, the observed CPU peaks were due to only one VM in the case of Linux VM and due to only a few coinciding VM CPU peaks in the case of Windows VMs. The peaks were sufficiently rare that even when in a limited state, only the few VMs bursting were affected; the vast majority of VMs used inconsequential amounts of CPU resource.

The small increase in time at high CPU loads for the 24GB Linux VMH compared with the 16GB Linux VMH suggests that the ratio of 24GB RAM to one quad-core Intel Xeon processor is near optimal. At no time were VMs deprived of provisioned resources (the peaks were as a result of VM CPU bursting beyond guaranteed limits), but a higher ratio could result in a restriction on the helpful ability to burst CPU.

5.2.8.2 Hypothesis #2: Disk TPS is more limiting

The data collected was an insufficiently accurate representation of what is believed to be the limiting factor (disk head seeks) due to data streaming causing increased apparent transactions and masking seek-bound states. This research has, however, enabled enhancements to the measurement systems to be designed, which should enable assessment of the limiting effects of disk TPS in future work. However, the very low average TPS does suggest that, as with CPU, TPS is at most rarely limiting for Memset's architecture. This also suggests that Memset's approach of maximizing TPS-to-CPU ratio is valid since they have largely avoided fully loading the VMHs, even when provisioned with maximum RAM.

With an architecture that was based on storage area networks (SAN) rather than local disks one would expect TPS to be more often limiting, necessitating a disproportionate investment in very high-performance SAN equipment, further supporting the case for a commodity platform such as Memset's.

5.2.8.3 Hypothesis #3: Most customer CPU loads are fairly static and small

The very small size of changes in power draw with activity levels (as measured by bandwidth) shown in figure 2 strongly supports the research hypothesis that most of the CPU loads among Memset's user base are fairly static and small. Additionally:

- We see low average CPU utilizations, and load distributions towards the low end of the spectrum;

- The observed CPU load peaks are largely attributable to only a small number of busy VMs.

This supports hypothesis #1 and the view that there is little need for the bulk of VM users to have the ability to migrate to other VMHs in real time in order to balance CPU load across the VMH grid.

5.2.8.4 RAM

While this paper predominantly examines CPU and TPS as limiters in virtual machine deployments it is worth remembering that the key resource considered when VMs are provisioned is normally RAM. RAM is fast becoming the unit of comparison of VMs from different providers, and also now accounts for at least half of the cost of the VMH in Memset's architecture, and therefore presumably some others.

RAM is also an additional power drain, and since it is constantly being refreshed regardless of utilization level it is expected that the power drain changes little in relation to utilization (unlike CPU). Therefore it should be considered a resource to use efficiently, perhaps even more so than CPU which does at least partially auto-regulate its power consumption with activity levels. At present varying RAM availability in response to demand is difficult since operating systems (especially Linux) tend to efficiently make use of any spare RAM, giving misleading results for RAM utilization.

However, hot-(un)plugging VM RAM is possible under Xen, and following this research Memset intends to explore the practicalities of such an approach in order to maximize efficiency. It is expected that, at present, any liberated resources could only be employed in disk-I/O-light activities (e.g. a compute grid) since TPS would otherwise still become a limiter.

5.2.9 Smaller, more modular data centres, using commodity hardware

Overall, the results clearly demonstrate the validity of the Memset VMH architecture. Following the increase to 24GB, which was as a direct result of this study, one can also argue that their approach is now near-optimal, as demonstrated by the occasional maxing out of resources; if limits were never reached that could be considered wasteful, but to very occasionally “bump” them suggests a good balance.

Further, the facts that Memset, a relatively small but also very financially sound company, has managed to achieve such industry-leading quality of service and value for money with this approach strongly supports the assertion that using commodity hardware for utility / cloud computing platforms is sensible. When combined with other benefits of this approach, in particular the reduced capital expenditure and modular cost scaling, the validity of the widely accepted wisdom that VM platforms should consist of blade arrays and SANs is seriously questioned.

The findings also seriously question the widely held belief that commoditised cloud computing services can only be delivered from mega-scale data centres by huge corporations able to make investments in the hundreds of millions of dollars range. This is important news for smaller, regional data centres since it demonstrates that they can compete with the international computing utilities (such as Amazon EC2) on price. Indeed, localised data centres can provide better quality of service to customers in the region thanks to cultural similarities and reduced connectivity latency.

5.2.10 References

Armbrust M., Fox A., Griffith R., Joseph A.D., Katz R.H., Konwinski A., Lee G., Patterson D.A., Rabkin A., Stoica I. and Zaharia M. 2009. *Above the Clouds: A Berkeley View of Cloud Computing*, Technical Report No. UCB/EECS-2009-28, University of California at Berkeley.

Hamilton, J. Cooperative Expendable Micro-Slice Servers (CEMS). In: Conference on Innovative Data Systems Research (CIDR ’09) (January 2009)

5.3 GREEN ICT: OXYMORON OR CALL TO INNOVATION?

Kate Craig-Wood‡, Paul Krause and Aaron D. Mason

Department of Computing
University of Surrey University
Guildford GU2 7XH, UK
(‡and Memset Ltd.)
kate@memset.com; {p.krause, a.d.mason}@surrey.ac.uk

Abstract—**"Green ICT" has often been referred to as an oxymoron. In contrast, we argue here that technological developments have enabled data centres to make significant reductions in their energy usage over the past 3-4 years. In addition, strategic use of ICT can facilitate energy reduction through behavioural change such as travel replacement. The use of ICT to support dematerialisation is more equivocal, however.**

Keywords-datacentres; virtualisation; travel replacement; dematerialisation; energy use; data transfer

5.3.1 INTRODUCTION

Taken in isolation, it is very hard to argue for there to be anything environmentally friendly about ICT provision. Apple's iPhone and iPad sales continue to push the boundaries for purchase rates of consumer goods, disposal of waste materials from short-lived pc's and mobile phones continue to hit the headlines and the carbon footprint of the much loved data centres of Amazon, Google and their fellow keystones of the Digital Economy will shortly exceed that of the airline industry.

Action can be, and is being, taken to mitigate the environmental impact of ICT provision. Careful architecting and the use of carbon offsetting have allowed two of this paper's authors to establish the first carbon neutral data centre, for example. In this paper, we will show that the use of local data centres can lead to environmental benefits through the efficiencies of virtualisation.

ICT can also lead to potential reductions in carbon footprint through its support for behavioural changes. The use of video conferencing to reduce the need for business travel is a well-known example. Our work with a local Small to Medium sized Enterprise, the Surrey Wildlife Trust, is another example of where the introduction of some quite straightforward technology is leading to very significant reduction in work-related travel around the county. Indeed, the World-Wide Fund for Nature see ICT as having very significant potential to reduce carbon footprint in a number of ways :

- Travel replacement
- Sustainable consumption: de-materialisation
- Sustainable community/City planning

However, there are still open questions here. The "dematerialisation of society", for example, has had relatively little impact. The move from gadgets to remote services is not taking place. Certainly, usage of remote services has expanded dramatically. However, this is currently being used to leverage a massive increase in sales of smart phones and tablet like devices.

In this paper, we will first present some concrete figures for the impact of the use of ICT for travel reduction. We will then move on to analyzing the energy reductions that have been made possible in the ICT industry itself. Following that, we will return to the issue of dematerialisation, and then conclude by arguing that increasing the utilization of the compute resource in gadgets may be a more effective mode of energy reduction than trying to wean people away from them.

5.3.2 ICT – PART OF THE SOLUTION, NOT PART OF THE PROBLEM

The ICT sector is regularly harangued about the "2%" figure – the amount of global carbon emissions attributable to ICT according to a Gartner report [5]. That figure is often quoted alongside real dirty polluters such as the airline industry (who pump CO_2 straight into the upper-atmosphere, bypassing many of the natural ground-level sequestration mechanisms), but what is often forgotten is that in exchange for our emissions (2-3% of total in the UK) we are contributing roughly 10% of UK GDP and 15% of national trade.

Further, the ICT sector has its own house well in order and have committed to reducing its own emissions as we will describe shortly. However, of much greater important is what the intelligent use of ICT can do to reduce emissions in other sectors, as highlighted by several groups including GeSI:

> "ICT can reduce annual global emissions by 15 per cent by 2020 and deliver energy efficiency savings to global businesses of over EUR 500 billion" [6]

Even the dedicated conservation organization, the *World Wide Fund for Nature*, highlights in [7] (and other reports) the capability of intelligent application of ICT as being key to reducing our collective carbon emissions:

> "'Green IT' is an oxymoron, until you consider use of IT to 'green' business and society." [8]

5.3.2.1 Example: Transport Avoidance

Perhaps the most obvious way that ICT can help is in transport avoidance. As David MacKay illustrates [9], personal transport in the form of driving cars and flying in jet aeroplanes are two of the worst activities in the developed world for energy consumption and carbon emissions. Together they contribute to over 40% of the UK's total energy consumption, for example.

Cars are the worst offender, consuming 40 kilo Watt-hours (kWh) per day per person (to put that in perspective, we use about 4 kWh/d each on lighting). Even with electric cars we still have to get the energy to run them from somewhere, and there are simply not going to be enough renewables to go around at current usage levels. The only way to significantly reduce the energy consumption attributable to cars & planes is to use them less, and that is where ICT comes in; for example by enabling home working (tele-working), even if just one day a week, and reducing travel to meetings with telepresence technologies.

For the "average" person in the developed world, air travel is only just behind car use in energy consumption. Based on one intercontinental flight per year being "average", air travel consumes about 30 kWh/d of energy per person [9]. Each additional trip adds a further 30 kWh/d to the successful business person's energy budget. Counter to that, each use of teleconferencing to avoid a business trip reduces the energy budget of that successful business person by a very substantial 30 kWh/d – the energy consumed by the teleconferencing facility is second order to this – and saves that person's business a very significant amount of money.

We will come back to this issue of "greenness" making sound business sense in the next section when we look at ICT's "own house" before reverting to looking at ICT as a facilitator of energy reduction.

5.3.3 Keeping our own house in order

Although we can help reduce carbon emissions elsewhere, we absolutely must do so in a sustainable manner, which is why we in the ICT industry are putting lots of effort into keeping our own house in order. Last year, Intellect UK (Britain's high-technology trade association) released their *High-Tech: Low-Carbon* report, which articulates an action plan on how the UK technology sector is going to reduce its emissions.

Further, Digital Europe (formerly EICTA) has committed to reduce Europe's ICT-related carbon emissions by 20% by 2020. Many think that target is achievable by 2015, but how can we be so sure of dramatic carbon savings when our collective use of ICT is increasing constantly?

A lot of the existing inefficiencies of the sector lie in the data centre, and that is also where we expect to see the largest efficiency gains. The UK, in particular the *BCS Data Centre Specialist Group*, has taken a global lead in advancing the field of energy efficiency within the data centre, and was instrumental in developing the European Union's Code of Conduct for data centres, which stipulates a range of best practices for every layer of the IT service delivery stack (from mechanical & electrical to software selection).

Memset (a UK based data centre of which one of the authors – Craig-Wood – is Managing Director) recently became the first British Web hosting provider to become a participant to the Code of Conduct, and we encourage others to follow suit (which many already are). The Code is free, is not hard to do (it took a day to implement the Code at Memset) and the best practices contained in it are designed to improve efficiency which means saving money, so it is also good business sense.

5.3.3.1 CO_2 SAVINGS THROUGH VIRTUALISATION

However, there is a much bigger effect that incremental improvements to data centre design, and that is the combination of Moore's Law with virtualisation technology. The work done per Watt by servers has been increasing roughly in line with Moore's Law, i.e. doubling every 18 months, and is expected to continue to do so. Now that virtualisation has reached the main stream it is being deployed en-masse, allowing legacy servers to be shut down and replaced with vastly more efficient virtual systems, usually consolidating physical machines by a factor of more than 10 to 1.

Take us as an example; in 2009 Memset deployed roughly 1,000 virtual servers. Each virtual machine (VM) would otherwise have been a physical server (or in many cases used to be before it was migrated to Memset), and in fact many people are still using cheap old tower PCs for cheap hosting, although that practice is dying out. A normal server or PC uses around 90-120 Watts continuously, whereas one of Memset's Xen-based Miniserver VMs uses 5-10Watts, but does the same work. Taking into account cooling and other data centre inefficiencies let us estimate this as 100Watts saving in round numbers:

1,000 VMs x 100 Watts = 100,000 Watts

x 30.4 days x 24 hours = 73,000 kWh / month

x 430g / kWh = 31,400 kg CO2 / month

So, this one small-scale data centre in just one small location in the world has helped avoid over 30 tonnes per month, or 360 tonnes per year, of carbon dioxide emissions. To put that in context, each British citizen is responsible for about 9 tonnes of CO_2 emissions per year.

5.3.3.2 BEING GREEN IS JUST GOOD BUSINESS SENSE

When it comes to ICT services, especially in the data centre, the two things that cost you the most money also cause the most carbon emissions; manufacturing the hardware (the servers / computers) and electricity to run them. In short:

Green = Efficient = Lower costs

There really is no excuse for us as an industry not to improve our energy-and-carbon efficiency, and companies that don't will end up with higher cost bases and ultimately will be driven out of business by their more efficient competitors. But let us now look into the energy budgets of Servers and PCs in a little more detail.

5.3.4 EMBEDDED ENERGY OF SERVERS AND PCS

Over the last two years there has been a lot of debate about what the embedded energy of a PC or server is compared with how much power it

uses. We believe that the figure for a server is about 1,000,000 Watt-hours (1,000 kWh or 1MWh). We will outline the calculation of this figure now, and explain why we conclude that the energy efficient strategy is frequent replacement of servers, but to prolong the life-times for PCs.

Some figures from a study by Williams [1] are relevant to this. However, the paper includes CRT (old-style monitor) production in with the figures and this is difficult to factor out. However, the paper does provide a table listing the electricity, fossil, and total energy use in computer production. A quick bit of analysis: The total estimated cost of production is 6,400MJ, and if we remove the CRT-specific bits, we take off:

CRT manufacture/assembly: 255MJ

bulk materials – CRT 800MJ

printed circuit boards: 20MJ (est)

electronic chemicals: 200MJ (est)

other processes: 400MJ (est)

Total: 1,675MJ

So, from the paper a PC's production is about 4,700MJ, which is 1,300kWh. Fujitsu (pers. comm.) asserted in 2008 that their range of green PCs took 730kWh to make (materials, production & distrubution) using their latest fabrication plant. If the numbers in [1] are correct, that is an impressive improvement in 4 years, but Fujitsu have been working hard in the area.

As an aside, this is very interesting from a recycling point of view. Most PC manufacturers, be it Fujitsu, Dell or IBM will proudly telling us about less than 2% goes to landfill, but surely the only energy that can be "reclaimed" from manufacture would be the bulk materials; all the energy of making chips, assembly, PCBs, transport etc is entirely lost. Therefore, in reality one could at most hope to recover perhaps 800-1,000MJ of the original energy-cost (i.e. about 20%).

A server is just a PC with a slightly different set of components (an extra disk & more RAM, but less additional cards like graphics & audio), so it is reasonable to assume the energy costs are similar. Therefore, we pick a figure half way between what we have deduced from the paper (1,300 kWh) and the only figure we have been able to obtain from a vendor (730 kWh) and have gone for 1,000 kWh in our estimations.

5.3.4.1 "Sweat" the desktops

So what about the fabrication energy vs. utilisation? The 81% fab, 19% use lifetime cost that is estimated in [1] is probably no longer accurate. First, [1] assumes 3 hours usage per day on average, which is far too low now given the current volume of office PCs and the often intensive use of family PCs. Second, a 3 year lifetime is too low – the indications are that most people use their PCs for much longer (or they get passed down / re-used rather than thrown away) – the Fujitsu figure of 6.6 years for home users seems much more realistic.

However, the figure of 128W for PC+screen in [1] are most likely still valid – the gains we have made in LCD screen efficiency have been outweighed by power-hungry CPU-intensive machines in recent years, although that trend is reversing. Fujitsu's figure is 80W for their "green" PC in full power mode, and an average LCD screen uses about 20W (about half a similar CRT).

So, an updated estimate (based on an average of PC & home use) is:

120W * 5 hours/day * 365 * 5 years ≈ 1,100 kWh

If we assume LCD screens are as energy intensive as CRTs and go with the figure of 1,700 kWh for production from [1] then the ratio is 61% fab : 39% use.

Using Fujitsu's figures, we have 730kWh in fabrication, plus ~300kWh for a screen (a guesstimate – it is about 465 kWh for a CRT), giving about 1,000kWh fabrication then the embedded

vs. use energies are almost equal.

If one then does the calculation based on an office PC usage pattern and a 6.6 year lifetime, then even with more energy efficient PCs the ratio is more like 35% fab : 65% use.

Therefore, we can conclude that the ratio of production energy to usage energy for a PC (with or without screen – the proportions seem about the same) range widely from something like (35% fab : 65% use) to (70% fab : 30% use), and that the main determining factor is the usage pattern of the PC, which is also the one bit of data that we probably have the worst grasp on. Either way, though, less energy will be used overall if the life of a desktop PC is stretched as far as possible (i.e. “sweat the desktops”).

5.3.4.2 Replace the servers

The situation is very different for a server, however. A typical modern 1U pizza-box server will use 80W when idle and 140W when working hard. Most of the time they are not straining, so call it 100W:

100W * 24 hours/day * 365 * 1.25 PUE ≈ 1,100 kWh per year

In other words, a server uses about the same amount of energy as was required to create it every single year, and the same amount that a PC with a fairly average usage pattern uses in 5 years.

Because of this it is worthwhile to replace servers with more efficient models on a fairly regular basis. Moore’s Law (that transistor density doubles every 18 months) means that server work capacity per Watt is increasing by a factor of 4 every 3 years. This means that provided you are using the servers properly (virtualisation etc) and consolidating onto a smaller number of newer machines, if you replace a 3 year old server its 1,000 kWh embedded energy cost will be saved by the 3 you are turning off (4:1 consolidation) in only 4 months.

5.3.5 Energy Consumed per MByte of Data Transfer

Let us now return to the use of ICT to facilitate energy reduction. We have briefly looked at the use of ICT for travel reduction. A second potential area for energy reduction is through “dematerialisation”, or the substution of services for material goods.

Before we look at a specific instance of dematerialisation, we will need an estimate of the energy consumed in the transfer of data from a remote data centre to an end user. A number of estimates are extant (e.g. [3]), but here we base our calculations on actual power consumption figures from the data centre Memset Ltd.

Sampling power consumption (servers and in-datacentre network infrastructure) and bandwidth output over a two-week period at Memset yielded a figure of 2.58 Watt-hours (Wh) per Megabyte (MB). Adding 50% for cooling and infrastructure losses (this latter is an upper bound estimate) and we have 3.87 Wh/MB.

We next need to take into account the embedded energy of the servers. We will be performing a more detailed analysis of this later in the year, but currently our upper bound estimate is 1Wh/MB.

After a number of discussions with ISP colleagues, it appears that the only significant part of the network delivery energy will be the DSLAMs (a ***D****igital* ***S****ubscriber* ***L****ine* ***A****ccess* ***M****ultiplexer* links many DSL connections to a single ATM line) and power for the home phone line. We estimate that the DSLAMs need about 1 Watt per user (again, a figure that we are in the process of refining) and the phone lines probably need about 2 Watts. Counting the phone line, and even the DSLAM, could be argued as unnecessary since they are likely to be there anyway.

If counting them, then we have, say, 3 Watts continuous (2,190 Wh) which gets used for about

2GBytes per month (a low average usage estimate), which gives another 1.1 Wh/MB.

In summary, we have the following estimate for the energy consumed per MByte of data transferred from a regional data centre to an end user.

TABLE I. SUMMARY OF ENERGY CONSUMED IN DATA TRANSFER

Servers and data centre networking	2.6 Wh/MB
Data centre cooling and losses	1.3 Wh/MB
Energy embedded in data centre hardware	1 Wh/MB
DSLAMs and phone line	1.1 Wh/MB
Total	**6 Wh/MB**

This figure is consistent with the estimate of 9-16 Wh/MB for the energy consumed in transferring data in the USA in 2006 which Taylor and Koomey published in [3]. It is possible that our lower figure is due to efficiency gains over the last four years, but more detailed analysis is needed in order to validate this.

5.3.6 DEMATERIALISATION THROUGH ICT?

Using a figure very similar to ours, Weber, Koomey and Matthews [4] provided some estimates to compare the energy required to deliver music to end users in a range of scenarios. The most energy intensive scenario was the method of selling music on CDs at retail outlets. They estimated this as costing approximately 15kWh (kilowatt-hours) per album. This contrasted with the most energy efficient scenario of straight download of the music file with no burning of the music to CD by the end user. The energy cost of this was estimated at approximately 2kWh per album.

This 7-fold reduction in energy usage looks promising at first sight. However, there is a difficulty emerging. The ease and lower cost of music downloads is working towards an increase in consumption. This is an example of *Jevons Paradox*, which was first put forward in 1865 [10]. That is, increases in the efficiency of usage of a resource tend to increase rather than decrease usage of that resource – through an increase in demand that outweighs the efficiency gains.

More recently, although still 40 years ago, this paradox was formalized through a simple mathematical identity first proposed by Paul Ehrlich and John Holdren [12]. The Ehrlich equation simply recognizes that the impact (I) of a human activity is the product of three factors:

Population size (P);

(Relative) Level of affluence (measured as income per capita) of the target society (A);

A technology factor that measures impact per unit of spend (T), or *technological intensity*.

We then have:

I = P x A x T

This formula can be used to demonstrate that in a growing economy, efficiency gains have to be significant enough to outweigh growth in population size and “material well-being” (expressed as income), in order for their to be a net reduction in environmental impact. This formula is used in [13] to illustrate the very stringent targets for reduction in carbon intensity of output per $ spend of GDP if we are to meet global impact targets for 2050, whilst still accepting a fair and just increase in GDP for the developing nations.

We can provide a very simple reinterpretation of this formula for specific consumer products or services (we will use the generic term *resource*) if we take

Customer base – or population of users (P);

Average spend per customer on a specific resource (A);

Technological intensity of impact of that resource (T).

We believe the Ehrlich formula is still valid in this reinterpretation, although it is harder to model the consumer dynamics as the technological intensity is reduced. However, qualitatively this usually entails a cost reduction, and hence increase in resource usage per consumer, and an increase in the customer base as the cost brings usage of the resource within reach of more consumers. The consequence is, and this is a fundamental requirement for a growth-based capitalist society [13], is that technological optimization and innovation will always tend to increase the absolute impact of any consumer resource. We need to model the situation in more detail with music downloads, but the market trends appear at the moment to be confirming this.

At the time of writing we remain a little skeptical of the potential for significant sustained overall energy reduction through dematerialisation, unless some mechanism was put in place to limit demand (and enforcing limits on demand for music, for example, seems somehow counter to a basic human right). The potential for use of ICT for energy reductions through behavioural change, as exemplified in the figures above for travel reduction, is far more significant.

5.3.7 Green ICT – A Call for Innovation

Let us for the moment withhold judgement on the potential for significant energy or carbon reductions through dematerialisation. Could we instead exploit the proliferation of gadgets by extending the use of virtualisation to make optimal use of this pervasive compute resource?

One approach that we are currently exploring is *Community Cloud Computing* (C3) [11]. It is the goal of C3 to exploit the immense power of combined compute resource that is rapidly evolving around us. It is a generalisation of the current NIST definition of Community Cloud to become a Cloud infrastructure that utilises spare capacity from (ideally) all the compute resources within a community that has shared concerns.

The key feature of the C3 concept is that it puts the ownership of both content and infrastructure resources in the hands of the community. This means that if supported by C3, the hardware resources available to support an application such as Wikipedia would scale as its community of users and contributors scaled. In addition, a Facebook-like social networking tool would not have to be monetized through the commercialisation of the personal data it captures. Instead, the entire application stack, including compute resource, would be provided by and owned by the community it supports. Hence, that community retains full ownership of its collective data.

C3 is designed to stimulate open innovation in applications that focus on social and community aspects; applications that are difficult to provide with current centralised clouds while respecting the privacy of user data. We are currently trying to gain funding for three living laboratories that will provide concrete examples of such applications, and will both be important stimulations of innovation in the concept, and tests of the validity of C3's claimed potential.

So, how could the C3 infrastructure be realized? At a conservative estimate, by the middle of this new decade we should see on a global basis:

2 billion PCs;

7 billion smartphones;

A smart grid for managing power that could be 1000 times larger than the Internet;

100 million servers from data centres around the world;

The beginnings of a growth in the use of Body Area Networks that could also match the volume of smartphones.

If we can learn how to exploit the collective

capacity of this Community Cloud as an application run time, then we will have a compute resource that scales with population and the level of activity of that population. C3 offers an architecture that aims to address this vision, created by combing the Cloud with principles from Digital Ecosystems, sustainability from Green Computing, and paradigms from Grid Computing, while remaining true to the original vision of the Internet. It will facilitate a new wave of innovation, with applications that mash up resources on a global scale and evolve to the needs of global communities.

The word ‘ambient’ describes our concept of a massively distributed collection of compute resources collected from a globally connected pool of machines with diverse capabilities for computation, storage and networking. The exponential rate of growth of these resources is what opens the potential for processing, networking and storage power that will support globally scalable applications. However, the second key feature of this vision is that this Cloud is not a product of a single vendor, committee or standards body. Rather, it is held together by informal, tacit and negotiated relationships. Hence our use of the name 'Community Cloud' to emphasise the spirit of collaboration that is required to realise the vision.

In order for the continuously changing pool of infrastructure resources to be made available in real-time, they will need to be traded in ways similar to the automated trading in financial markets. A key distinction, however, is that the trading will be supported by a ‘Community Currency’ that will enable peers to (automatically) access resources in the cloud in exchange for fairly providing (spare) capacity into the cloud.

Interestingly, the Community Cloud has potential to provide greater resilience and privacy than conventional data centres, but with lower storage demands. Sufficient layers of indirection can be put in place so that it is virtually impossible for anyone other than the original owner (or user authorised by the owner) to access data that is stored in C3. C3 will put the user back in full control of their data, and in so far as they choose to share it with a community, that community has full control over its collective data.

The hypothesis is that the community cloud will have a significantly smaller carbon footprint that conventional clouds – through the extended use of virtualisation that is already facilitating significant energy reductions as we have seen. But realizing this vision is a serious challenge. The concept itself is somewhat disruptive to conventional business models in the consumer electronics industry – it advocates increasing the efficiency of usage of “boxes”, rather than the selling of more boxes. It remains to be seen as to whether a community led pull could sustain this alternative model.

5.3.8 Conclusion

There is still a relative sparsity of concrete data on the impact of technological solutions to green IT. However, we have shown in this paper that there are some areas (virtualisation and travel reduction) where significant energy savings can be made. There are other areas, dematerialisation, where the benefits are more equivocal. However, we do hold open the possibility that other more innovative solutions may continue very positive moves towards the use of ICT as a solution, and not as part of the problem.

5.3.9 References

[1] E. Williams, “Energy Intensity of Computer Manufacturing”, Environ. Sci. Technol., 38, pp. 6166-6174, 2004.

[2] N. Hogg, and T. Jackson, “Digital media and dematerialisation”, Journal of Industrial Ecology, 13(1), pp. 127-146, 2008.

[3] C. Taylor and J. Koomey, “Estimating energy use and greenhouse gas emissions of internet advertising.” Working paper prepared for IMC2, 2008.

[4] C. Weber, J. Koomey, and H. Scott Matthews, “The energy and climate impacts of different music delivery methods”, Journal of Industrial Ecology, in press.

[5] Gartner, “Gartner Estimates ICT Industry Accounts for 2 Percent of Global CO2 Emissions”, Gartner Newsroom, http://www.gartner.com/it/page.jsp?id=503867, 2007.

[6] Global e‐Sustainability Initiative (GeSI), SMART 2020: Enabling the Low Carbon Economy in the Information Age, June 2008.

[7] D. Pamlin, and S. Pahlman, "Outline for the First Global IT Strategy for CO2 Reductions", WWF, 2009.

[8] S. Mingay, Gartner Expo, Las Vegas, 2008.

[9] D. MacKay, "Sustainable Energy - Without Hot Air", UIT, 2008.

[10] B. Alcott, "Jevons' Paradox", Ecological Economics, 54(1), pp. 9-21, 2005.

[11] A. Marinos, and G. Briscoe, "Community Cloud Computing", CloudCom 2009, LNCS 5931, pp. 472-484, 2009.

[12] P. Ehrlich, "The Population Bomb", New York: Buccaneer Books, 1968 (reprinted 1995).

[13] T. Jackson, "Prosperity without Growth: Economics for a Finite Planet", earthscan, 2009.

6 CONCLUSIONS

We have looked at both species and cell modelling approaches to building a foundation of digital ecosystems in OPAALS. These are indeed complementary, with both strands informing the development of general theories of self-organisation, resilience, adaptation and transformation in complex adaptive systems. Whilst other deliverables have been primarily focused on technical results from OPAALS, this one has taken a broader view of placing our work into the context of the science of social-ecosystems, and showing how we propose to further develop our research work over the next six years.

An important aspect of this plan is the repositioning of our work in the emerging science of Industrial Ecosystems. This is an important change, and reflects the CAS "system of systems" perspective, where we accept and embrace the fact that digital-ecosystems are embedded within and influenced by a higher-level socio-economic context, *and that these coupled systems must be constrained within our global ecological and resource boundaries*.

It would be a tough call to expect a global transformation to result from one project. However, we believe that OPAALS has made a significant and sustainable contribution to a constituency of research that is now close to being able to define environmentally and socially sustainable computing.