Informatics Computing Plan 2019 (draft)

School of Informatics

March 25, 2019

Contents

1. Long-term vision and strategic objectives

Informatics Computing serves around 350 staff (250 teaching and research), 440 research students, 715 taught postgraduate students, 1380 undergraduates, and 290 visitors and associates.

The aim of the Informatics Computing staff is to ensure that members of the School of Informatics (staff, students and visitors) receive computing services necessary for their research, teaching and knowledge transfer activities. These services should be efficient, fit to users' requirements, good value for money and use open standards. Appendix A outlines the evaluation processes that we have established to ensure that we are fulfilling this aim.

Strategic objectives

We have five principal aims underpinning the Informatics Computing Strategy:

S1 Maintenance, review and update of a computing environment fit for the purposes of all members of the School.

S2 Providing added value over services offered by College and IS.

S3 Maintaining an optimum level of interoperability of Informatics Computing with College and IS services.

S4 Provision of expertise to support the teaching and research activities of the School.

S5 Engagement with international best practice.

We have specific objectives relating to the computing infrastructure and to the activities of the School: research, teaching and knowledge transfer :-

Infrastructure   We are committed to providing an infrastructure that ensures that members of the School get those services that they need. These services may be provided by the School, by IS or by external organisations.

I1 Review and evaluate computing infrastructure change taking account of changing user needs and general computing trends.

I2 Development of new services.

I3 Provision of Informatics know-how and technologies to college and university level, and beyond.

Research   In addition to providing a flexible, responsive environment for research in the School, we must meet the specific research requirements across our research institutes, and structure research computing support to be well-matched to the ways researchers propose and carry out research projects.

R1 Continued development of lightweight, responsive support for research computing that is fully compatible with full economic costing of research

R2 Ensuring that Informatics users get efficient, responsive access to high performance research computing and storage facilities

R3 Provision of support for interdisciplinary and collaborative research projects

R4 Development of prototype services from R&D projects

Teaching In addition to providing a stable environment for the School's teaching activities, we shall

T1 Support research-led teaching by providing support for the transfer of research tools to our standard teaching platform.

T2 Support appropriate assessment of students (eg online examinations).

T3 Provision of expertise to support teaching activities

Management Information We shall support the ISS business processes. We also aim to support planning and decision making through the timely and effective maintenance and provision of Management Information.

Interaction with IS

We shall focus :-

- on being early adopters of services that may or may not become commodity
- on developing new services that are specific to, or inspired by, our environment

We shall use IS services wherever possible, unless there are sound academic or technical reasons for not doing so. However, we shall take a careful approach when considering migration from a School service to the equivalent IS service.

2. Pressing issues for 2019

We have identified a number of strategic strands for consideration in 2019. We shall :-

- produce a strategy for improving the efficiency of teaching admin support. This will cover the effective use of central systems, identifying which of our remaining manual processes can be automated and how we might support the admin processes for online teaching.

- produce a strategy for promoting the use of the VLE (and other related tools) and to support the transition of courses to the VLE. This will, at least, define the scope of our Learn pilot, identify potential tools (other than Learn) and consider whether we have good reason not to use Learn as a "home" for teaching related systems (as done in other Schools).

- produce a strategy for resourcing the compute and data intensive needs of the School. The demand for compute resources from compute and data-intensive courses continues to grow, particularly with increasing interest in machine learning. There is also increasing demand for compute resources from the research postgraduate community.

- consider the School's network in the light of the University Network Strategic Review

3. Report on 2018

Mandatory Goals

1. **Goal** Migrate the Robotics teaching and research and Commercialisation to the Bayes building

   Progress Completed.

2. **Goal** Produce a register of medium and high risk data and a mechanism for users to self populate the register

   Progress We have decided, in the light of little apparent progress on the College register, to produce our own simplified system : leveraging existing Theon database technology. The expectation is that data entered into this system can be migrated to any future College or University register.

3. **Goal** Take remaining steps to implement College security action plan

   Progress The one remaining step is to introduce encryption for DICE desktops. We plan to develop and deploy this as part of the next major platform upgrade. This is expected to be in 2020.

4. **Goal** Continue to engage with University Network Strategic Review

   Progress Two of our staff are involved. Alastair Scobie represents the College on the project board and George Ross has been heavily involved in the tender preparation and evaluation.

5. **Goal** Prepare for GDPR

   Progress We have identified where we are processing data and have produced DPIAs and LIAs where appropriate. Still have some ongoing work to improve compliance. We believe that more could have been done by the University (eg. producing University wide LIAs and central location for DPIAs).

6. **Goal** Upgrade administrative staff desktops to Windows 10

   Progress We have started this process. Since mid 2018, all new desktops have been installed with Windows 10.

7. **Goal** Migrate off Staffmail

   Progress A pilot migration of a few senior academics was unsuccessful. This has led to the School's management becoming nervous. Subsequent migrations of admin, computing and technical staff were performed without significant problems. Academic staff will be migrated after Semester 2.

8. **Goal** Implement separate backup streams for MHR and non-MHR data to meet differing retention policies

   Progress This project was delayed by a requirement to upgrade the underlying backup technology. This has now been done and the project should be complete by mid 2019.

High priority goals

1. **Goal** Consideration of how best to make use of the new central RDM services

   Progress No progress.

2. **Goal** Continued consideration of appropriate use of central data storage facilities

   Progress The migration of all administrative data from AFS to DataStore is close to completion. One problem that we have identified is that if a researcher has donated, say 200GB, of their allocation to their research group, that group loses that space if the researcher leaves the University.

3. **Goal** Engage in requirements capture for and design of proposed central archiving service

   Progress We never received an offer to engage in this.

4. **Goal** Produce a strategy for improving the efficiency of teaching admin support

   Progress A working group was formed and the first meeting identified a list of quick wins which will be implemented.

5. **Goal** Produce a strategy for supporting the transition to use of VLE (and other tools)

   Progress A working group has been formed and so far has met once.

6. **Goal** Produce a strategy for resourcing the School's compute and data intensive requirements

   Progress A working group has been formed and the first meeting identified existing issues and various information required to produce a strategy. This information is now being captured.

7. **Goal** Migrate existing institute web sites off Plone CMS service to School Drupal service (based on IS Drupal distribution)

   Progress Four institute sites have been migrated to Drupal, leaving two to do.

8. **Goal** Investigate whether the existing network file system is still appropriate for the School's requirements and identify possible alternatives

   Progress No progress.

9. **Goal** Perform a review of the future of the DICE desktop platform

   Progress Little progress. We are changing the focus of this review to address how best to deliver the School's computing environment for taught students to both on-site and distance students.

10. **Goal** Implement upgrade of AV facilities in Forum G.07/G.07A, in collaboration with IS LTS

    Progress The upgrade of AV in G.07 has been completed.

11. **Goal** Introduce Docker under DICE

    Progress No progress.

12. **Goal** Produce a production Slurm service

    Progress No progress.

13. **Goal** Produce a production Gluster (or equivalent) service

Progress No progress.

14. **Goal** Establish a target of maintaining energy consumption (by computing equipment) at 2016 level, or lower. Possible approaches include increased use of cloud services for computation work

    Progress This has not been achieved largely because of the substantial growth in the number of GPU based servers. Given the growing demand for GPU compute, it is difficult to see how this goal can be met in the near to mid future.

15. **Goal** Review policies with respect to self-managed servers (due to energy and security concerns)

    Progress No progress.

16. **Goal** Improved management of Commercial tenants, with respect to network provision and user support

    Progress A network and support agreement for commercial tenants of Bayes has been developed. We intend to extend this agreement to our commercial tenants in Appleton Tower.

17. **Goal** Continue to improve end-of-life account management.

    Progress We have enabled account deletion which has resulted in a considerable number of old accounts being deleted. We now automatically identify and disable inactive accounts.

18. **Goal** Consider how to deal with growing server estate, given limited scope for increasing server room space

    Progress We have introduced space management for the server rooms - including booking space before procurement, periodically reviewing tenancy and eviction.

19. **Goal** Investigate and deploy technology to improve certificate management

    Progress We have identified LetsEncrypt as a suitable technology and deployment is in beta test.

20. **Goal** Complete the audit of all research data within the School

    Progress A snapshot has been completed. This data will be entered into the School data register once that is complete. The School needs to develop a mechanism to keep this data up-to-date.

21. **Goal** Implement agreed steps to improve security of system administrator accounts

    Progress There has been little progress other than some prototyping.

22. **Goal** Produce user security training materials (supplementing IS material)

    Progress A project has started but not yet made significant progress.

23. **Goal** Migrate to https

    Progress No progress.

Discretionary goals

Only those goals which have had any substantial progress are listed below.

1. **Goal** Produce a production Hadoop service

   Progress Done.

2. **Goal** Improve security of LCFG profile access

   Progress The necessary code and configuration to achieve this has been completed and will be rolled out in summer 2019.

3. **Goal** Produce a Limesurvey service

   Progress We have subscribed to the Qualtrics Research Platform, as used in CAHSS.

Recurring goals

1. **Goal** Aim for a minimum of 20% of development time to be dedicated to user submitted projects

   Progress Not achieved.

2. **Goal** Further promote School developed solutions to the rest of the University and beyond

   Progress No progress this year.

3. **Goal** Ring-fencing 5% of individual computing staff's time for staff development, including user support staff.

   Progress There has been patchy progress.

4. **Goal** Consideration of ways to minimise our energy footprint, eg identifying under-used research servers

   Progress No progress.

5. **Goal** Assess system security and identify potential improvements

   Progress No progress this year.

6. **Goal** Further consideration of migration to central services (big ticket items only)

   Progress None.

7. **Goal** Review impact of University activities wrt. teaching - VLEs, Distance Learning (including MOOCS) , EUCLID developments (as assessment).

   Progress Two Computing Strategy Group working groups are discussing VLEs and teaching admin (including with respect to EUCLID)

Collaboration with others

We would like to register our continued appreciation of the assistance of Angus Rae and Victoria Dishon, who have acted as very responsive and effective interfaces with IS.

We are very keen to collaborate with other Schools and IS on development and even service delivery.

1. We continue to provide the base LCFG Linux platform to other schools (via IS) and host the monthly LCFG deployers meeting.
2. Network Strategy and Procurement projects

3. We worked with other Schools and IS to deliver the Bayes building

4. Revised plan for 2019

Each project has a cost effort estimate, where small is 1 to 3 FTE weeks, medium is 4 to 7 weeks and large is 8+ weeks.

Information on computing projects is available at http://computing.projects.inf.ed.ac.uk

Mandatory goals

1. Produce a Data Register and a mechanism for users to self populate that Register
   Who: School, Cost: small, Project(s): 307

2. Populate the Data Register with details of medium and high risk data
   Who: School, Cost: medium, Project(s): 307

3. Continue to engage with University Network Strategic Review
   Who: School, Cost: small, Project(s): 404

4. Consider the School's network in the light of the University Network Strategic Review
   Who: School, Cost: small, Project(s): 478

5. Complete the upgrade of administrative staff desktops to Windows 10
   Who: Admin, Cost: medium, Project(s): 477

6. Migrate off Staffmail
   Who: School, Cost: small, Project(s):

7. Complete the implementation of separate backup streams for MHR and non-MHR data to meet differing retention policies
   Who: School, Cost: medium, Project(s): 346

High priority goals

1. Improve processes for research data management - including production of Data Management Plans and recording, in the Data Register, details of all research data.
   Who: Research, Cost: small, Project(s):

2. Consideration of how best to make use of Information Services' RDM offerings, producing use-case guidance to researchers, in partnership with IS Research Services
   Who: Research, Cost: small, Project(s): 386

3. Produce a strategy for improving the efficiency of teaching admin support
   Who: Teaching, Cost: small, Project(s):

4. Produce a strategy for supporting the transition to use of VLE (and other tools)
   Who: Teaching, Cost: small, Project(s):

5. Consider how to replace the School's ageing course work submission system
   Who: Teaching, Cost: small, Project(s):

6. Consider how to facilitate the use of Piazza (or similar system)
   Who: Teaching, Cost: small, Project(s):

7. Scope requirements for managing School teaching content, including audit
   Who: Teaching, Cost: small, Project(s):

8. Consider how best to deliver the School's computing environment for taught students, considering the needs of both local and distance learning students
   Who: School, Cost: small, Project(s): 379

9. Produce a strategy for resourcing the School's compute and data intensive requirements
   Who: Teaching/Research, Cost: small, Project(s):

10. Procure and commission a GPU cluster for research postgraduate use
    Who: Research, Cost: medium, Project(s): 472

11. Migrate existing Institute web sites off Plone CMS service to School Drupal service (based on IS Drupal distribution)
    Who: School, Cost: medium, Project(s): 388

12. Investigate whether the existing network file system (AFS) is still appropriate for the School's requirements and identify possible alternatives
    Who: School, Cost: medium, Project(s): 445

13. Research the possibility of re-basing the DICE desktop platform on an more frequently updated Linux distribution than Redhat
    Who: School, Cost: medium, Project(s): 474

14. Improve the working knowledge of Office365 tools amongst particularly the administrative and front line computing support staff
    Who: School, Cost: small, Project(s):

15. Produce more sustainable and performant compute/GPU clusters - focusing on job management (slurm) and filesystems for 2019
    Who: Research and Teaching, Cost: medium, Project(s): 463

16. Investigate the feasibility of supporting Docker under DICE for end user use
    Who: Research and Teaching, Cost: medium, Project(s):

17. Upgrade the AV facilities in the Forum (other than G.07/G.07A), in collaboration with IS LTS
    Who: University, Cost: small, Project(s):

18. Improve procurement and purchasing processes
    Who: School, Cost: small, Project(s):

19. Procure and commission replacement UPS system for Forum server rooms
    Who: School, Cost: small, Project(s): 473

20. Review policies with respect to self-managed servers (due to space, security and energy concerns)
    Who: Research, Cost: small, Project(s):

21. Produce user security training materials for users of self-managed machines (supplementing IS material)
    Who: School, Cost: small, Project(s): 403

22. Consider how to deal with growing server estate, given limited scope for increasing server room space
    Who: School, Cost: small, Project(s):

23. Implement improved management of Commercial tenants, with respect to network provision and user support using Bayes model
    Who: Research, Cost: small, Project(s):

24. Continue to improve end-of-life account management. Investigate feasibility of automatically removing non home directory file space associated with expired user accounts.
    Who: Infrastructure, Cost: medium, Project(s): 478

25. Rollout use of LetsEncrypt certificate technology across all services
    Who: School, Cost: small, Project(s):

26. Implement agreed steps to improve security of system administrator accounts
    Who: School, Cost: medium, Project(s): 399

27. Investigate implications of, and produce a plan for, a wholescale move to HTTPS
    Who: School, Cost: small, Project(s): 454

28. Develop and document policies and procedures for physical security
    Who: School, Cost: small, Project(s): 394

29. Identify a more sustainable procurement path for GPU servers
    Who: Research and Teaching, Cost: small, Project(s):

Recurring goals

1. Aim for a minimum of 20% of development time to be dedicated to user submitted projects
2. Further promote School developed solutions to the rest of the University and beyond
3. Ring-fencing 5% of individual computing staff's time for staff development, including user support staff.
4. Consideration of ways to minimise our energy footprint, eg identifying under-used research servers
5. Assess system security and identify potential improvements
6. Further consideration of migration to central services (big ticket items only)
7. Review impact of University activities wrt. teaching - eg. VLEs, Distance Learning (including MOOCS)

Activities to be considered for de-commissioning

- Legacy web sites
- Twiki
- Wordpress (blog.inf)
- coltex

IS services critical to Informatics

As far as we are aware, these are the IS services that are critical to the School. It is possible that there are other IS services that are widely used by our users; presumably IS maintains usage statistics that could be used to identify these.

1. EdLAN / eduroam - delegated, fine-grained control would make IS management of the School's internal network more attractive
2. Office365 mail (IMAP compliant)
3. VLE (Learn)

4. Lecture recording service
5. Managed Windows Desktop
6. DataStore
7. Phones and Access Control
8. Central authentication and directory services
9. Central administrative services (and feeds from) such as EUCLID/HR/BIS/PURE/EUGEX etc.
10. ECDF subversion service
11. Public PC labs - we would like to investigate more effective use of these labs for our 1st and 2nd year students.
12. ECDF GPU cluster (based on expected use)
13. ECDF Eddie cluster
14. EdWeb distribution
15. EdWeb hosted service
16. WIKI
17. blog.ed
18. IDM (Identity management system)
19. IS printer service (including pcounter and cloud)
20. SSL certificate signing service
21. MOOC
22. Software purchasing
23. Visitor Registration Service

Additional services we would like

1. Improved computational facilities for taught students
2. Provision for data archiving and, perhaps, curation of non research data.
3. We are interested in the proposed ECDF cloud virtualisation service, particularly for taught students.
4. Additional programmatic interfaces to central administrative systems, as documented in Colin Higg's note on "Arguing for Authorised APIs". eg SOAP to EUCLID, SAT and VRS.
5. An API to "upload" assessment data to central systems instead of manual copy/paste processes.
6. The ability to feed into Grouper, from our School Database, would reduce the barrier to the School making more use of central IS services
7. The ability to make more use of centrally provided group data but this remains dependent on the quality and accuracy of the data and suitable APIs
8. Replacement for ESISS scanning service
9. A more pervasive attitude, particularly with respect to increasing provision and access, to the use of video conferencing across the University
10. For ECHO360 to be available in all teaching and seminar spaces, in particular Forum G.07 and AT 6.06
11. We would like for it to be easier for people to adopt our research software whilst maintaining IPR.
12. An API to the VLE (Learn) - for example, to upload and maintain content and to extract tutorial groups (and membership of).

13. User authenticated, POSIX compliant, remote filesystem access to DataStore from multi-user Linux machines

In order to achieve the aforementioned improvements and additional services, we are very keen to collaborate with both IS and other CSE schools on development and even service delivery.

5. Plan for 2020

Goals

1. Any required actions resulting from network file system review
2. Any required actions resulting from IS authentication and authorisation review
3. Any required actions resulting from replacement of EdLAN
4. Any required actions resulting from strategy to improve teaching admin support
5. Any required actions resulting from strategy to transition to VLE
6. Any required actions resulting from strategy to resource School's compute and data intensive requirements
7. Any required actions resulting from consideration on how to best deliver computing environment for taught students
8. Any required actions resulting from consideration of DICE platform rebase
9. Deliver a replacement for the School's ageing course work submission system
10. Full disk encryption on DICE (remaining step for College Security policy)
11. Any required actions resulting from consideration of services to be decommissioned

De-prioritised areas

- To be identified

A. Evaluation

We have established a number of evaluation processes, to ensure that we are delivering a service in line with our strategic objectives.

- **Fit to requirements** User requirements are captured using various mechanisms. Any member of the School can submit a project proposal via a web form. Each project is categorised into one of three prioritised categories - Mandatory, Strategic (meets one of the goals in Section 2 of this document) or Objective (meets one of the Strategic Objectives in Section 1). Projects are resourced in priority order when effort becomes available. There is a target of 20% of development time to be dedicated to user submitted projects. Teaching software requirements are met through a stable and well established system. Specific requirements are also captured in depth via focused innovation meetings, which all members of the school may attend.
- **Value for money** This is a criterion for the annual review document, and is related to transparent support for research computing, centralised procurement that remains close to academic needs, and official audits of various research project expenditure.
- **Objective evaluation** Each computing team provides a triannual report, which reports on activities in the past four months and future plans for work in the forthcoming four months. It includes figures on effort spent on development activities.

The School's Computing Strategy Group (Head of School, Director of Research, Director of Teaching, Head of Computing, Deputy Head of Computing, Director of Professional Services) provides oversight of the above processes.

B. Staffing and Resources

The school employs 23 computing staff (22.8 FTE). This figure includes 3 vacant, but approved, posts. This includes an additional post for 2019 which is targetted at improving support for our growing GPU cluster estate.

There are 643 managed DICE (Linux) desktops; 259 personal machines for staff and research students, and 384 in 10 undergraduate teaching labs and 12 tutorial rooms. There are a further 75 managed Windows desktops for administrative staff.

In addition there are several hundred self-managed Linux, Mac OS and Windows desktops and laptops.

There are 406 managed DICE (Linux) servers (215 physical, 191 virtual) and a further 88 cluster nodes. There are an additional 110 physical self-managed servers. Many servers are used to host GPUs - we now have over 500 GPUs with over 1,200,000 cores.

Trends :-

- The number of DICE desktops used by staff and research students has continued to decline. This is probably largely due to research students increasingly self-managing their desktop as the Redhat platform, on which the DICE desktop is based, is becoming out-of-date, and also due to to academic staff, who largely use Apple Macs for commodity computing, now using the remote DICE desktop service for developing teaching materials.

Our servers are housed in 5 air-conditioned machine rooms, with a total area of around 170 m^2. (Figures as of 15/02/2018).

C. Categories and activities

Support

| Procurement | Online Exams | General Web | Documentation | Front Line |
| | Exam prep | Inventory | AV and VC | User accounts |

Compute

| DICE compute |
| Grant servers |
| Hadoop |
| Gridengine |
| CDT clusters |
| MSc cluster |
| GPU clusters |

Desktop

| DICE desktop |
| Teaching s/w |
| Research s/w |
| MDP |
| DICE virtual desktop (XRDP) |
| Virtual DICE image |

Misc

| ssh servers |
| Printing |
| Mailing lists |
| Licence servers |
| Postgresql |
| Coltex |
| Forums |

Admin

| Theon |
| ISS desktop |
| IGS desktop |
| HR desktop |
| ISS and KM RT |
| Room Booking |

Web

| Web (CVS) |
| EdWeb |
| Wiki |
| Group web |
| Homepages |
| Plone |
| Legacy |
| Wordpress |

File

| AFS |
| Samba (admin) |
| Backups (user) |
| iFile |
| SVN |
| GIT |
| Gluster |

Infrastructure

Network	Directory	Kerberos	LCFG	Linux platform	Server rooms	Virtualisation
Network services	Authorisation	Cosign + Certs	pkgforge	S/W repository	Consoles	
System backups	Prometheus	Monitoring				