

Cambridge Feb 2016

Eramus+

Informació del document			
Nom de l'arxiu	Cambridge Feb 2016		
Versió	2		
Elaborat per	Carles Perarnau	Data d'elaboració	22-02-2016
Aprovat per		Data d'aprovació	
Confidencialitat			

Control de versions			
Versió	Descripció dels canvis	Elaborat per	Data d'elaboració
2	Introducció de continguts	Carles Perarnau	27-03-2016

Responsable del document: Carles Perarnau

Índex

[Situació](#)

[Objectiu](#)

[Preparació](#)

[Visita HPC Service](#)

[Equipaments:](#)

[Darwin:](#)

[Wilkes:](#)

[Emmagatzemament:](#)

[Altres](#)

[UIS](#)

[Organització UIS:](#)

[Descripció del l'organigrama](#)

[Reunió Agustin Fernandez](#)

[Visita EBI](#)

Situació

L'any 2015, conjuntament amb el Servei d'Informàtica, el DCEXS crea el Servei Científico-Tècnic d'Informàtica Científica (Scientific IT, SIT) per donar resposta a les necessitats informàtiques del departament en l'administració de l'equipament informàtic d'altres prestacions que és utilitzat per a la recerca. El SIT s'integra també en el conjunt de serveis científico-tècnics que ja disposa el departament (Genòmica, Proteòmica, Confocal,...) i pren la seva filosofia de servei. Aquest es basa en que una part del servei es pagada està subvencionada i l'altra se'n fa càrrec usuari del servei.

La Direcció del Departament i del servei estableixen unes tarifes que tenen en compte els costos derivats de cada un dels serveis que s'ofereixen. També s'estableix un reglament d'ús, etc...

Objectiu

Donada la innovació en la gestió dels recursos informàtics destinats a la recerca i aprofitant la convocatòria dels ajuts Erasmus+, em proposo conèixer altres llocs on, amb problemàtiques semblants, que pugui servir per fer una comparació de models.

Per dimensions de les universitats, àmbits temàtics tecnològics i possibilitats d'intercanvi, faig la sol·licitud de beca per visitar el HPC Service (High Performance Computing Service) de la Universitat de Cambridge a través del Cap de Departament. És dona el cas que dins d'aquest servei hi ha un grup de recerca, el Biological Computing Lab, el cap del qual és Ignacio Medina, col·laborador d'alguns dels recercadors del Departament.

Preparació

La preparació ha estat un mal de cap. Per una banda, el temps donat per poder fer la sol·licitud, entendre les demandes i poder-les satisfer és molt curt. Per fer la sol·licitud, cal aportar unes dates concretes de visita, les quals es fa difícil d'establir si la sol·licitud no ha estat acceptada. El retard de la informació per tal de tirar endavant la preparació i la proximitat de les dates d'estada amb resolució de la sol·licitud ha fet que si s'ha tirat endavant ha sigut una qüestió de sort.

Tenint clars els objectius de la visita, vaig voler establir un calendari d'activitats a fer durant l'estada. Això ha estat impossible per diferents problemes.

També s'ha volgut aprofitar l'estada per visitar altres centres del nostre àmbit. A Cambridge hi ha:

- The European Bioinformatic Institute (EBI, que és part de l'European Molecular Biology Laboratory)
- Sanger Institute <http://www.sanger.ac.uk/>
<http://www.sanger.ac.uk/science/groups/information-communications-technology>

S'ha pogut fer la visita a EBI, però no ha estat possible el Sanger.

Visita HPC Service

El High Performance Computing Service és un Servei del UIS. És tractat com a tal i per tant té independència amb la resta d'organització. (Veure més endavant)

La persona de contacte amb el servei ha sigut: Ignacio Medina¹: Head of Computational Biology Lab, High Performance Computing Service



External Scientific Collaborator, EBI Variation, EMBL-EBI

El primer que em va sobtar d'aquest servei és que tenen un grup de recerca dins. I sí. És així. (model BSC)
El servei d'HPC, com a tal, està obligat a publicar un parell de papers a l'any. Per allò de mantenir la dinàmica de recerca dins dels mateixos serveis de la universitat.

El Lab de l'Ignacio està dins d'un projecte d'investigació que ha d'analitzar 100000 genomes i la seva recerca és per a la millora i modelació d'algorismes i programaris per a l'anàlisi d'aquestes quantitats de genomes.

A part del Nacho, el HPC Service, està compost per:



Paul Calleja²: Responsable del Servei d'Alta Computació de Cambridge. És l'artífegs de l'existència del HPC Service. Tenim funcions equivalent, però ell treballa pels tots els departaments que necessiten HPC. Té un conveni amb DELL que aconsegueix uns preus molt competitius. Aquest fet ha sigut clau per aconseguir fer veure als investigadors el benefici d'unificar o centralitzar el servei.

Stuart Rankin: Sysadmin dels clusters per a la recerca de la Universitat de Cambridge. Gestiona dos clusters, Darwin i Wilkes (veure especificacions més endavant).



¹ <http://www.bigdata.cam.ac.uk/directory/ignacio-medina>

² <http://www.bigdata.cam.ac.uk/directory/paul-calleja>



Wojciech Turek: Administrador de Lustre. Porta 5 instal·lacions de lustre (cada instal·lació amb un mínim de 6 servidors i un munt de cabines)

La direcció de la Universitat a declarat aquest servei com estratègic per a la Universitat -només amb equipament de primera és podrà fer recerca de primera- i ha apostat pel servei. Prova d'això és l'expansió que tindrà el servei en 4 mesos. Per una banda s'espera la contractació de 4 persones per ajudar en les tasques d'administració de còmput i emmagatzemament.

Per una altra, es vol crear una infraestructura amb openstack per tal de proveir sistemes tipus cloud per a les necessitats dels lab i oferir-los com un servei més. Per fer-se càrrec d'aquesta infraestructura també tenen previst contractar entre 2 i 4 persones més.

També s'estan fent proves amb hadoop per tal de posar en marxa un cluster i poder resoldre certes demandes, les quals són més eficient amb aquest tipus de cluster. I finalment està treballant en la renovació del cluster Darwin, ja que actualment té 4 anys.

Equipaments³:



Darwin:

Cluster de computació del servei. El primer Darwin va ser posar en producció al febrer del 2007 i des de llavors ha tingut dues actualitzacions (octubre del 2010 i Juny 2012) Al seu dia va ser el 20è en capacitat de càlcul. S'espera fer una actualització aquest any. Hores d'ara està compost per 600 nodes amb doble Intel Sandy Bridge, 64GB of RAM

per node i Infiniband FDR. (una arquitectura molt semblant a l'actual cluster del SIT. El SIT només té 30 nodes). Aquest equipament dona servei de còmput internament a 5 departaments de la Universitat i també forma part de STFC Dirac facility⁴.

³ <http://www.hpc.cam.ac.uk/services>

⁴ <http://dirac.ac.uk/> Agrupació de recursos de supercomputació de diferents institucions.



Wilkes:

Cluster de computació amb GPUs. Posat en marxa del novembre de 2013, està compost per 128 nodes and 2 GPU per node i connexió infiniband. Al seu dia va ser el 166 del Top500. També va ser el núm 2 de la llista Green500 a l'estar refrigerat per aigua. Va ser sponsoritzat per Rolls Royce i Mitsubishi.

Emmagatzemament:

El servei administra 5 instal·lacions de Lustre, cada una amb un mínim de 6 servidors i un total de 4,5PB. Aquests sistemes d'emmagatzemament, hores d'ara, no s'ofereixen com a servei, sinó que són dels diferents departaments. Són administrats pel personal del HPC Service. Estan directament connectats amb els clusters del servei.

Altres

El servei utilitza la gestió d'usuaris del UIS per donar accés als seus serveis. Utilitzen un sistema anomenat Raven per l'autenticació (no em va quedar clar si raven és el sistema d'edició de web o també el sistema d'autenticació).

Horas d'ara no estan aplicant cap ISO, tot i que s'han compromès a aplicar-ne una per la gestió d'accés a dades confidencials.

El servei HPC utilitza un sistema de tickering diferent al del servei. Utilitzen JIRA, amb el Service Desk i Confluence. Això els permet veure en un sol board, les tasques a fer que provenen dels tickets i les que provenen dels projectes. També els permet gestionar les compres, i les tasques fetes amb checklist. (eina de pagament que prové d'un sistema de gestió de projectes de software que s'ha anat ampliant amb diferents mòduls)

El servei de còmput es paga per endavant després d'acordar les necessitats de cada grup.

Per descriure el servei tenen els conceptes de Service Level (nivell de servei) i Quality of Service (Qualitat de servei)

Han començat a utilitzar Ansible per a la configuració de nodes del cluster
(Ansible és una eina semblant al Puppet o al Salt

Està previst posar en marxa en un breu temps un cluster d'OpenStack per tal
d'oferir "lloguer de hardware" i virtualitzar el servidors que tenen.

UIS

University Information Services és el servei d'informàtica de la Universitat. És un servei centralitzat, molt semblant al nostre model.



Disposa d'un edifici exclusiu pel servei on hi ha la gran majoria del personal, on es fan cursos de formació,.... Aquest servei està ubicat al Roger Needham Building, West Cambridge Site⁵. (campus situat a 2km del centre on hi ha diferents edificis amb serveis, departaments, ... de la Universitat)

⁵ <http://www.westcambridge.co.uk/>

Organització UIS:

En la meua visita em vaig limitar a fer una foto de l'organigrama del servei i entrevistar-me amb un dels treballadors del servei, a fi de tenir una mínima visió del seu



funcionament.

Descripció del l'organigrama:

1. Director UIS

a. Build & Development

- i.* Administration & Project Support
- ii.* Project and Programme Office
 - 1. Programme/Project
 - 2. Business Analysis
 - 3. Delivery Office
- iii.* System Assurance
 - 1. QA and Testing Services → Agustin Fernandez
 - 2. User Centric Design
- iv.* Enterprise & Mid Range System
 - 1. Finance System
 - 2. Information System
 - 3. HR System & Pensions
 - 4. Teaching & Learning System
 - 5. Functional Analysis
 - a.* Student Systems
 - b.* Research Systems
- v.* Student System Technical
- vi.* Development Services
 - 1. Project Development
 - 2. Development Services
 - 3. Web Development Services

b. Digital Transformation

c. Information Management

d. Service Operations

- i.* Networks
 - 1. Network Installation
 - 2. Network Systems
 - 3. Network Services
 - 4. Telecons
 - 5. Granta Backbone Network
- ii.* Infrastructure
 - 1. Infrastructure System Manage
 - 2. Enterprise and Platforms Service Manager

- a. Platforms
 - b. Enterprise Apps
 - 3. System Integration Manager
 - iii. Frontline Services
 - 1. First line Services Manager
 - 2. Service Desk
 - a. *Bussines Systems*
 - b. *Reception*
 - c. *Technical*
 - d. *Card Services*
 - 3. End User Computing
 - a. *Manager*
 - b. *Development & Integration*
 - c. *Client & Applications*
 - d. *Administrator*
 - 4. Training, Software Distribution, User Admin
 - a. *Training*
 - i. *Services*
 - ii. *Delivery*
 - iii. *Admin & Operational Support*
 - b. *Software Distribution*
 - c. *User Administrations*
 - d. *CERT*
 - 5. Operational Officer and Divisional Administrator
 - a. *Operational Officer*
 - b. *Divisional Administrator*
 - c. *Assestive Technology*
 - d. *Videoconfierence Service*
 - e. *Print Room*
 - f. *PANDIS*
 - 6. Data Centers
 - a. *Critical Services*
 - b. *Operations*
 - iv. Architecture
 - 1. Architects
 - v. Research Institutional Services
 - 1. High Performance Computing Services**
 - 2. Intitutional Strategy
 - vi. Departamental Operations
 - 1. Communications
 - 2. Finance
 - 3. Information Provisson
 - 4. Administration
 - 5. Building and Maintenance

- 6. Planning and Performance
- vii. Education, Administration, & Student Services

Reunió Agustin Fernandez

L'Agustin és un espanyol que està treballant al servei de qualitat i testing del System Assurance del Build & Development. Em va estar explicant en anglès la seva feina, les eines que utilitzaven per testejar i fer proves de càrregues de les aplicacions.

Resum de les eines que utilitza:

- Functional Testing
 - SeleniumHQ.org
 - Cucumber.io
 - Veure videos
 - https://www.google.co.uk/search?q=simon+stuart&ie=utf-8&oe=utf-8&qws_rd=cr&ei=a-7NVt5ExYZo7qKIoAM#q=simon+stewart+webdriver&tbm=vid
 - https://www.google.co.uk/search?q=simon+stuart&ie=utf-8&oe=utf-8&qws_rd=cr&ei=a-7NVt5ExYZo7qKIoAM#q=selenium+conference&tbm=vid&tbs=qdr:y
 - <https://www.youtube.com/watch?v=p6cqqXrBvBU>
- Performance Testing
 - https://aws.amazon.com/free/?sc_channel=PS&sc_campaign=acquisition_UK&sc_publisher=google&sc_medium=cloud_computing_b&sc_content=aws_core_e_control&sc_detail=aws%20web%20services&sc_category=cloud_computing&sc_segment=73311487129&sc_matchtype=e&sc_country=UK&s_kwcid=AL!4422!3!73311487129!e!!g!!aws%20web%20services&ef_id=VstL3gAAACvZoU7U:20160224180956:s
 - https://aws.amazon.com/free/?sc_channel=PS&sc_campaign=acquisition_UK&sc_publisher=google&sc_medium=cloud_computing_b&sc_content=aws_core_e_control&sc_detail=aws%20web%20services&sc_category=cloud_computing&sc_segment=73311487129&sc_matchtype=e&sc_country=UK&s_kwcid=AL!4422!3!73311487129!e!!g!!aws%20web%20services&ef_id=VstL3gAAACvZoU7U:20160224180956:s
 - <https://reinvent.awsevents.com/>
- Grafana
- Collaboration
 - Redmine
 - Atlassina Confluence

Va ser una reunió interessant, ja que no conec una funció igual al nostre servei, però se m'escaven molts conceptes.

Visita EBI:

Dijous, havent dinat, visito a en Pete Jokinen <pete@ebi.ac.uk>, responsable de la infraestructura de EBI (European BioInformatics Institute)



l'institut està a uns 10km de Cambridge al Genome Campus. En aquest lloc també hi ha el Sanger Institute.

Aquesta és una altra lliga i el sistema de funcionament també. Bàsicament, l'equip del Pete, donen la infraestructura necessària als investigadors, un cop sabudes les seves necessitats i ells se'n fan càrrec.

Això és així, gràcies a la manca de recursos econòmics. Certs investigadors no els agrada aquest model per què no poden intervenir i participar a la definició i administració de la infraestructura.

També em vaig poder entrevistar amb en Jordi Valls <jvalls@ebi.ac.uk>, administrador de sistemes que està treballant a l'EBI.

