Case Study:
Clinical Research Imaging Centre,
University of Edinburgh and NHS Lothian Joint Venture

1. Background

The Clinical Research Imaging Centre (CRIC) is based within the University of Edinburgh's Queens Medical Research Institute (QMRI), based at Little France, in Edinburgh.

The Little France campus consists of the new Edinburgh Royal Infirmary, the University's Medical School, the QMRI and work will start shortly on the new Sick Children's Hospital and Department of Clinical Neuroscience.

The vision of a dedicated imaging centre for clinical research was instigated during the building of the QMRI and it was always hoped to house such a facility there, especially with the close location of the ERI

Initial groundwork was carried out by Elizabeth McDowell, IT Manager for Wellcome Trust Clinical Research Facility (WTCRF) and Trevor Carpenter Brain Research Imaging Centre (BRIC), formerly SBIRC.

They came up with the initial workflow and network concepts, ***(Diagrams 1 and 2, Appendix 1)***, created with much input from users and Clinicians.

Julian Sparrow, CRIC IT Support Manager, came on board in May 2008, to facilitate the implementation of the IT infrastructure for CRIC.

The role involved:

- reviewing all previous material
- creating a more granular workflow ***(Available on request)***
- creating a Memorandum of Understanding (MoU) ***(Available on request),*** which underpinned the liaison between the University of Edinburgh and NHS Lothian
- getting both University and NHS Lothian IT specialists together to discuss ways to create and implement the network infrastructure, to allow communication to both University and NHS Lothian networks ***(Diagram 3, Appendix 1)***
- liaising with vendors to facilitate the speedy installation of kit
- providing IT support and solutions to vendors and users in all areas of CRIC, during construction and installation, particularly in the Radiochemistry suite
- specify, in liaison with the Manager of the Image Analysis core, procure, install, set up and support all Image Analysis equipment
- specify, procure, install and support all other CRIC IT needs

2. Baselines

There were a few basic baselines to work from:

- CRIC was to be classed as a “no-mans land”; neither in nor out of either the University or the NHS
- Patient Identifiable data would, in all probability, be required to be accessed anywhere within CRIC, so no portion of CRIC could be deemed specifically University or specifically NHS, with the possible exception of the Radiochemistry suite.
- All data classified as research must be anonymised before it can leave the CRIC, and Anonymisation must adhere to the guidelines provided by both the NHS and the University with regards to the Safe Storage and Transfer of Research Data
- NHS to manage Firewalls
- Scanner usage (% Research/% Clinical): MRI 70/30
  CT 50/50
  PET 30/70

3. Practical Implementation

3.1 Network Infrastructure

CRIC would run on a University supplied subnet, primarily because 90% of resources would come from the University; email, applications, storage etc., but would reside behind NHS managed firewalls, so that the NHS Security protocols would be in effect on data passing to/from CRIC

It was decided that all network points within CRIC would feed back to one comms room and one set of patch panels. This meant that any port within CRIC can be placed onto any network, within reason. If you want an NHS port in a given location, that port would be patched directly to an NHS switch, with no requirement to pass through the firewall. A University port would be patched to a University switch, which then passes through the firewall and back to the University infrastructure.

Internally, you can create as many standalone networks as are desired, as long as each is contained by its own set of switches. This is how the Radiochemistry suite network is set up, requiring to be standalone for GMP reasons, but with the ability to extend that network anywhere in CRIC, should there be a requirement, e.g. accessing data from the Radiochemistry server on the standalone Radiochemistry network, from within the CRIC seminar room.

Wireless Access Points are not possible because this was felt to be a potential bridge between the University network and the NHS network and because the traffic cannot be monitored as it passes through the firewall.

3.2 Firewall

Because of the network requirements of both University and NHS Lothian, the set up of both networks is, naturally, different. The IP ranges used are different and to avoid confusion or conflict, it was deemed necessary to apply a double NAT configuration through the firewall, for connectivity of devices in CRIC to devices within the NHS, e.g. Reporting Workstations, PACS, HIS/RIS etc. This provides the ability to send data from a CRIC device to an NHS device, going from say a University IP address to a NHS IP address, which requires two additional sets of dedicated addresses, one for each side of the firewall. ***(Example shown in Appendix 2)***

The example shows that the correct set up of the firewalls is crucial. The simplest way would be to decide what should be connected to where, set the firewall rules, and then set up the kit at either side. This requires a close liaison between both support teams. At CRIC, with Julian onsite, this has been easier because it is one point of contact for all.

3.3 Data

3.3.1 In General

All data is pushed from the modality, there is nothing automated. This is due primarily to the requirement in some areas to send multiple copies of the data to various locations. Originally, the requirement was for the reporting clinician to pass relevant datasets on to PACS at RIE and subsequently National PACS (Note: All data received by RIE PACS is automatically forwarded onto National PACS), however, given the type of modalities CRIC has (3T MRI, 320-slice CT, 128-slice PET/CT) the datasets that are generated are larger than average, especially when carrying out research scans. It was known very early on that the size of datasets that reconstructed CT would generate, were in the range of 4-10GB for a full body scan. PACS capacity at RIE would not be able to handle such large datasets.

3.3.2 Clinical

For clinical work it was decided that all clinical data be sent to a local CRIC NHS Archive solution, in the form of a Carestream PACS system, set as a sub-node of RIE PACS with a one-way dataflow. Data can be queried and retrieved from CRIC NHS Archive to RIE PACS but no data can go from RIE PACS to CRIC NHS Archive.
Clinical MRI data is also passed directly to RIE PACS, as its size will have little impact on RIE PACS.
Clinical CT is straight to Archive, and reporting Workstation in CRIC, secondary reporting is done by Clinicians in RIE pulling the data from CRIC NHS Archive via web browser application.
(Addendum: Currently in the process of allowing data on CRIC NHS Archive to be queried directly by workstations in RIE and possibly sending directly

from scanner to workstations in RIE in time). The Reporting Clinician is responsible for sending relevant datasets to PACS/National PACS.
Clinical PET data is sent to NHS RIE Archive, Reporting workstations at RIE and Western General Hospital (WGH), also in Edinburgh, for secondary reporting and a predetermined dataset is sent directly to PACS/National PACS.

3.3.3 Research

All research data should pass through the CRIC Anonymisation process and be held offsite on NAS storage hosted by ECDF, the Edinburgh Compute and Data Facilities. From here it is made available for Image Analysis. Data is stored for the duration of a study and a copy is then passed to the PI. Unless stated otherwise, retention times are set at 15 years for research data. At the end of a study, it is up to the PI to decide if they want CRIC to archive and retain the data; however the University reserves the right to retain research data indefinitely, for future research purposes.
The ideal scenario is that at the end of the study, all study data is archived in one file and stored on the NAS or on an archive solution provided by ECDF, fit for purpose.

As all research subjects are passed through TRAK, NHS Lothian's HIS/RIS, thus all research scans must be reported on. To this end, depending on the reporting clinician involved in the study will depend on which reporting workstations the data is sent to, either in CRIC or RIE. Again, it is the responsibility of the reporting Clinician to send a relevant dataset to PACS/National PACS, to accompany the report.

3.4 Anonymisation

As previously stated, all research data must be held offsite, anonymously. Originally it was only stipulated that DICOM data was to be stored. In reality, this is not the only data generated. MR produces many different types of data and research studies can use many different protocols, producing such data types as fMRI, Spectroscopy, DTI and DICOM. Both CT and PET produce raw data, from which the reconstructed DICOM data is generated. Most facilities do not hold onto raw data, because it's the reconstructions that are used for analysis/reporting, however, it has become apparent in CRIC that raw data storage may become an issue, especially on research studies, because of the reconstructed data is wrong, having the raw data means you can apply different reconstruction techniques to yield better images for better analysis. At this time, only research CT raw data is saved, currently onto blu-ray disks, and held in a fire safe within CRIC, probably until the study is complete, then it will be passed to the study PI.

PET raw data, or List Mode data, has been requested to be kept for a couple of studies. In itself, this is anonymous, as it requires the PET acquisition console to be able to do anything with it, but some internal storage at CRIC has been given over to List Mode data. These files will be given pseudo

anonymous names by the PI, to be able to store them offsite at ECDF, as the internal CRIC storage is not currently backed up.

For DICOM data, from all three modalities, an anonymiser solution *DICOMConfidential* has been provided by David Rodriguez, through SINAPSE. *DICOMConfidential* comes in two flavours, the Receiver anonymiser and the Folder Anonymiser. Currently BRIC are using the Folder anonymiser effectively, Psychiatry is also using this version, which generates a pseudo ID to replace CHI number. However CRIC's requirements are a little different. CRIC uses, in conjunction with TRAK, the NHS Lothian HIS/RIS, a Patient Scheduling software package, developed by the WTCRF IT developer, Dickon Jackson, under Elizabeth McDowell. This is tied into the CRF Manager system for study management. This allows CRIC to interrogate the Patient Scheduler for a unique identifier for the actual scan date/time, to make sure the patient data being anonymised is the correct patient, further to this, based on patient data, CRIC can retrieve a unique identifier to replace the CHI number, so the Receiver anonymiser developed for CRIC runs as an automated product that on receipt of data, sorts the files into a suitable folder structure then does the anonymisation to CRIC specifications, before moving the data offsite to ECDF NAS storage.

Work is ongoing on this solution as *DICOMConfidential* is not a catchall for all three modalities' DICOM data at this point. For example, it does not deal well with fMRI data, possibly due to the sheer large volume of small data files. Spectroscopy data will require an alternative solution, as will DTI data.

Implementation is on three separate virtual servers, one per modality, to avoid conflicts if two modalities are sending data at the same time. Currently the data is sent straight to ECDF from the anonymiser, but it may be that CRIC decide to stage the data, temporarily in CRIC and run an overnight job to transfer it offsite. This is in case the links to ECDF fail, CRIC would still be in a position to supply anonymised data to PIs directly, until the link is re-established.

Alternative automated anonymisation methods are being investigated.

4. The Future

It seems likely that at this juncture, the DICOMConfidential anonymiser will be the main automated method for data anonymisation, whether it can be improved to deal with data other than DICOM cannot be answered at this time.

It is possible that further requirements are made to store raw data, and so anonymisation solutions would need to be sought to achieve this, or funding for managing and maintaining increased storage within CRIC may need to be sought. There is also the potential requirement to store any analysis software used in a study, along with the resultant output, which in some cases can be many times greater than the original data, something people tend to forget when creating grant proposals.

5. Contact Details

Should you require any further information, or to discuss individual elements in more detail, Julian will be pleased to assist in any enquiries.

Julian Sparrow, CRIC IT Support Manager
Clinical Research Imaging Centre, University of Edinburgh
Queens Medical Research Institute, 47 Little France Crescent
Edinburgh, EH16 4TJ
T: 0131 242 7766 F: 0131 242 7773 E: julian.sparrow@ed.ac.uk

Appendix 1

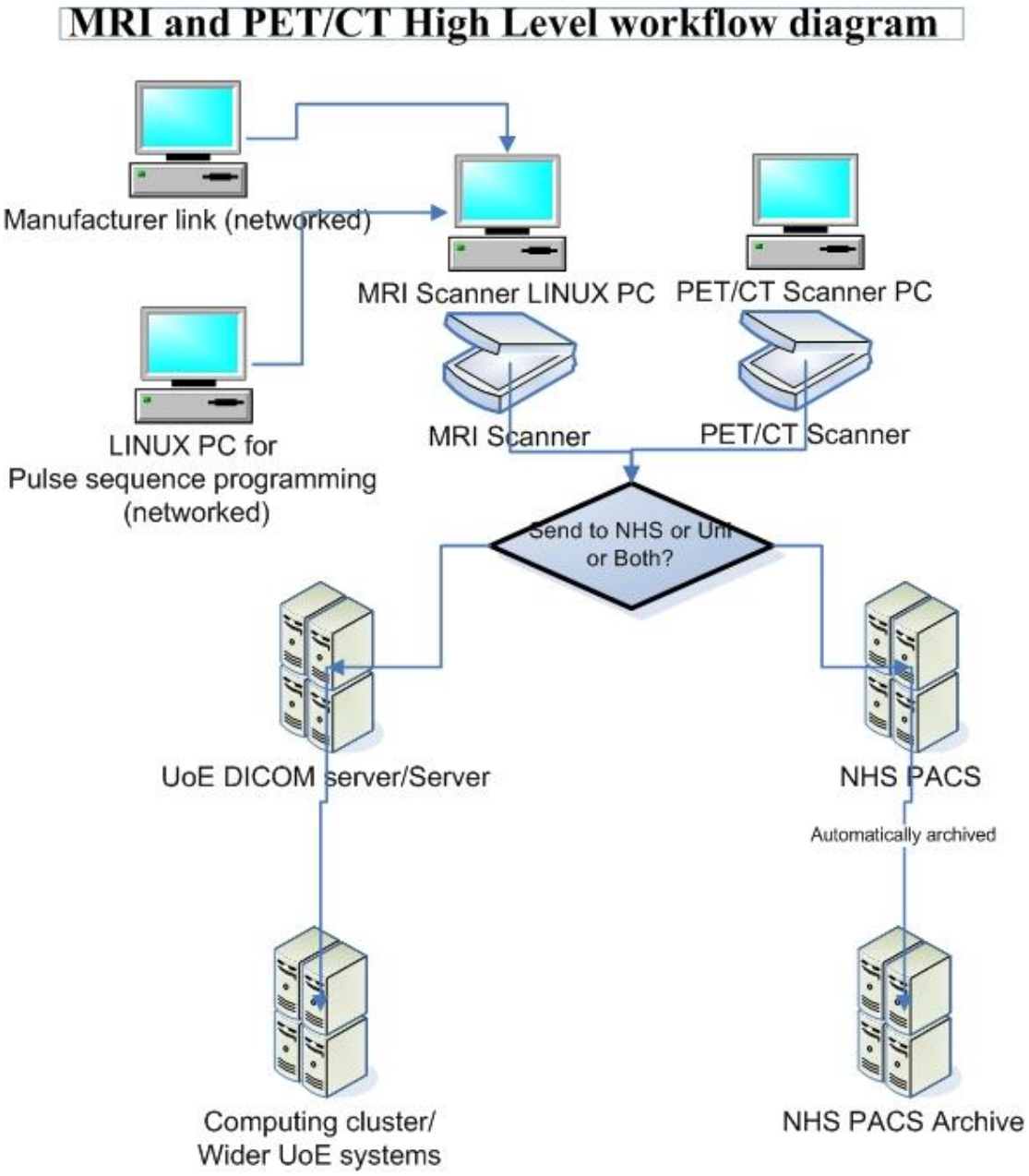

Diagram 1: Proposed High Level Workflow

Appendix 1

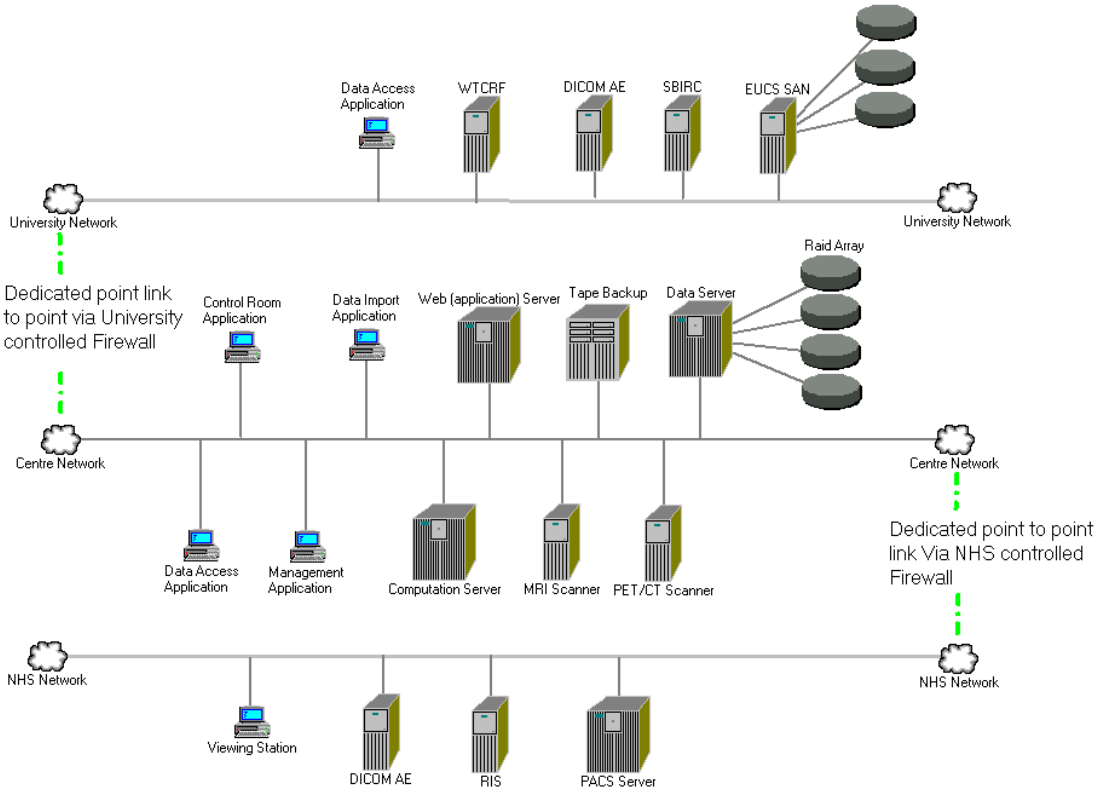

Diagram 2: Proposed Network Layout

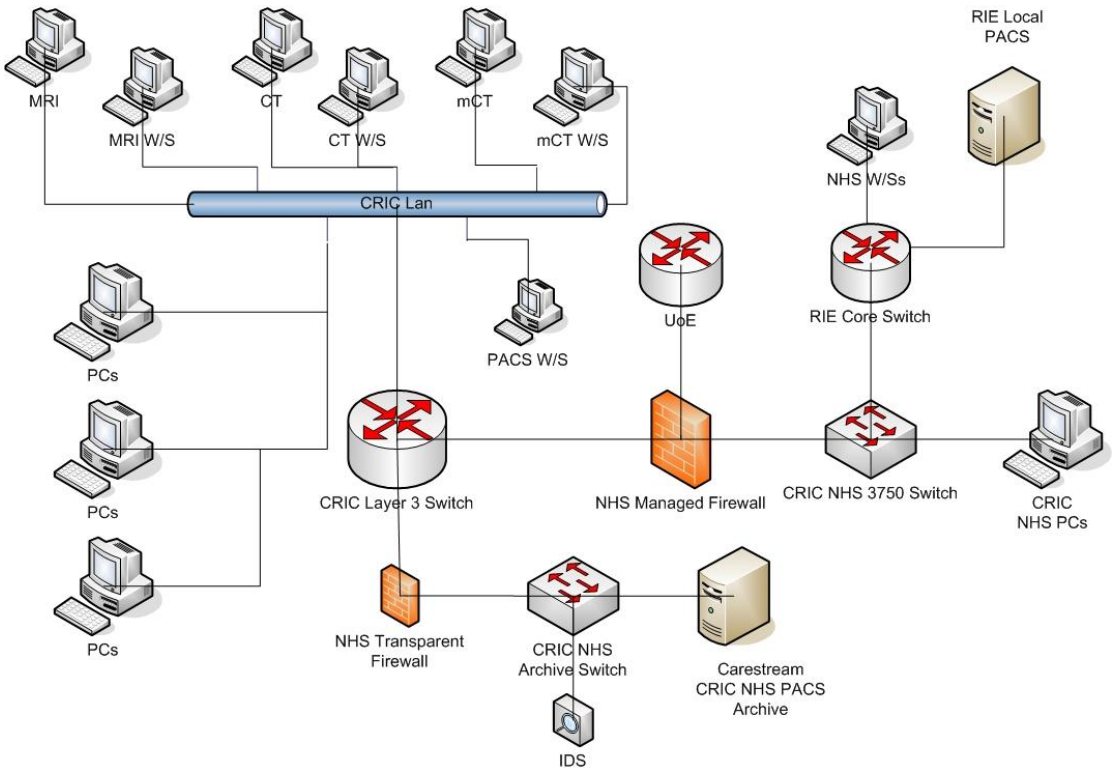

Diagram 3: Simple Finalised CRIC Network Layout

Appendix 2

Example of Joining University and NHS Networks via Firewall and NAT

To send data from CRIC MRI Scanner to NHS Lothian ERI PACS, we need to communicate from University Scanner Console IP **129.215.1.1** to NHS Lothian PACS IP **10.48.1.1** through the firewall, but we cannot just simple route this directly as there may also be a **10.48.1.1** on the University side, so using Network Address Translation at the firewall, we specify a University IP for the NHS PACS, pretending the PACS is within the University. So we would set the scanner console up to send to NHS PACS, with an IP of **129.215.100.1**, but the firewall would translate this to **10.48.1.1** on the NHS network and the data would go to PACS.
Visa versa, we cannot guarantee that NHS Lothian will not use a **129.215.1.1** address, so for PACS to communicate back to the scanner console e.g. requesting data from PACS, we need to pretend the scanner console is on the NHS network and give it an NHS address of **10.48.100.1**, which the firewall translates to **129.215.1.1**
Connection and send handshaking would be something like this:
Request to send data from scanner to PACS
Scanner knows PACS as 129.215.100.1
Scanner sends connection request from 129.215.1.1 to 129.215.100.1
Firewall receives request to connect to 129.215.100.1
Firewall translates 129.215.100.1 to 10.48.1.1 and says the request is from 10.48.100.1
PACS receives request from 10.48.100.1
PACS says 'Ok' to receive data from 10.48.100.1
Firewall receives 'OK' for 10.48.100.1 from 10.48.1.1
Firewall translates 10.48.100.1 to 129.215.1.1 ands says 'Ok' is from 129.215.100.1
Scanner receives 'Ok' from 129.215.100.1
Scanner sends data to 129.215.100.1
Etc.

Note: All IP addresses are for example only