A Brief History of the Data Network in the Department of Computer Science and the School of Informatics at the University of Edinburgh

George D M Ross

gdmr@inf.ed.ac.uk, gdmr@dcs.ed.ac.uk

January 26, 2016

This document relates the evolution of the data networks in the Department of Computer Science and the School of Informatics at the University of Edinburgh. It attempts to say why things ended up the way they did, as well as looking at what emerged. Note that diagrams may be simplfied slightly for clarity, though significant details are included. Related documents can be found on the Edinburgh Computer History site: <http://history.dcs.ed.ac.uk/>.

The Author was a member of the Department/School from 1974–1976 and from 1979–present.

Departmental Links

A significant step towards the development of a high-speed data network in the Department of Computer Science was the design in the early 1970s of the "Link" board. This provided a pair of independent serial connections, often but not always used as a half-duplex pair. The Links ran at up to 1Mbps over standard 75Ω coax cable. Each connection was 8-bit clean and flow-controlled, with a positive ACK signalled to the sender for each octet transmitted. Implementations existed for the Interdata backplane, the DEC unibus (PDP-11 and VAX-11) and the DEC PDP-9 and -15 machines.

By 1975 the Department had two Interdata clusters interconnected by Links, as well as a PDP-9 and a PDP-15/20. Machines were gradually added, so that by about 1980 the following were interconnected by Links (see figure 1):

- the "Filestore," an Interdata processor with two 70Mb disc drives and a number of Link boards, the purpose of which was to act as a central file storage repository for all the Department's other small machines

- the Interdata clusters

- the PDP-9 and PDP-15/20, now using the Filestore as virtual DECtape drives, as this turned out to be faster than using real physical hardware

- a PDP-15/40

- an ICL 7502 discless editing station, which ran VECCE and very little more

- the "LEGOS" systems: standalone Interdata machines with rather more memory than the small cluster machines

- the Edinburgh Remote Terminal Emulator ("ERTE"), a research project into performance testing
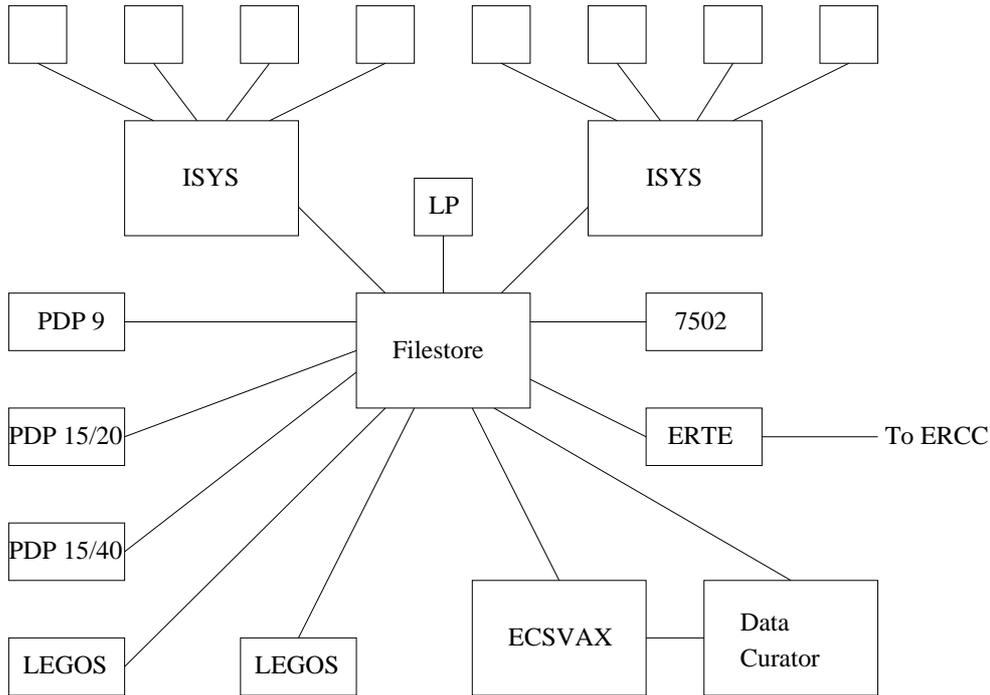
Figure 1: Network diagram from around 1980

- a DEC VAX-11/780 "`ECSVAX`" running VMS, which was the Department's main multi-user machine at the time

- the "Data Curator": a Perkin-Elmer 3220 used as part of another research project

- a lineprinter.

The Filestore was the hub of the network, and almost all traffic over the Links was filestore protocol (described elsewhere[1]), either directly manipulating the files in the case of discless machines, or to transfer them back and forth in the case of `ECSVAX` and the Data Curator. There were only two real exceptions to this.

One of these was the lineprinter. This was provided with a Link receiver (no need for a transmitter) and connected to its host by this means. Initially this was to one of the Interdata clusters, and latterly to the Filestore. The Link's inherent flow control meant that the despooler could simply throw characters down the line as space in the printer's buffers permitted.

The other exception was that there was a direct connection between the Data Curator machine and `ECSVAX`. This was originally intended to allow direct data transfer between the two machines. In fact, it was not used for this purpose, with file transfer being staged via the Filestore instead, but it did run a home-brew bidirectional remote login protocol for a while.

2Mbps Departmental "Ethernet"

There were a couple of drivers for the development around 1980 of the Department's 2Mbps ethernet-like network. The Metcalfe & Boggs paper[2] had come out, showing what could be done; and the sheer number of connections back to the Filestore were becoming unmanageable — indeed, latterly there was a multiplexor machine attached to the Filestore, partly to provide additional connectivity.

---

[1] For example, `http://history.dcs.ed.ac.uk/archive/os/APM-filestore/FS.1976/FSDOC.LAY.html`

[2] Linked from `https://en.wikipedia.org/wiki/Ethernet`

The Departmental ether was different in a few ways from standard ethernet: it ran at 2Mbps rather than 10Mbps; it had a slightly different addressing structure; and by default it presented as a *reliable* datagram service between configurable end-station ports, in contrast to standard ethernet's unreliable datagrams between stations, to allow filestore protocol to be layered directly on top. Ether stations were connected to their hosts (Filestore, LEGOS, ECSVAX) over serial Departmental Links, now 9-bit-clean and running at 2Mbps, and were designed to offload the ether protocol processing from the hosts. Like standard ethernet, the Departmental ether ran over 50Ω yellow "thick ether" cable, and used the same "vampire" clamps, though interfaced through purpose-built transceivers.

The arrival of 32-bit microprocessors, in particular the Motorola 68000 series, prompted several members of the Department to construct an experimental testbed, known as the "Fred Machine." This was a modular system, with processor cards, memory cards and graphics cards all plugging into a common backplane. In particular, there was an ether card, with a revised version of the Departmental ether station. These machines proved so successful that they were relabelled the "Advanced Personal Machine" ("APM") and gradually rolled out as the Department's main workhorse for staff and students (cite Gordon).

Unfortunately, the original version of the ether station as used by the Filestore turned out to be under-buffered, and unable to cope with the increasing load. Packets would be dropped quite frequently, resulting in broken connections, rebooted machines and lost work. It was decided to reimplement the Filestore using a 160Mb Winchester disc attached to a Fred machine. This took place from 1983 onwards, in two phases:

1. as the Interdata Filestore's ether connection was too unreliable for the intensive traffic involved alongside its normal load, the filestore protocol interpreter was initially implemented on ECSVAX, calling out to VMS's native filesystem

2. having uploaded an image of the necessary file structure to the ECSVAX "filestore," the APM-based Filestore was developed using the relatively more reliable facilities it provided.

Once the APM-based Filestore was brought into service, the original Interdata Filestore was retired, and run as a printer despooler driving the original lineprinter over its Link interface. In addition to the 160Mb Filestore, two 300Mb Filestores were eventually deployed, one mainly for staff use and the other mainly for student use, along with four based on 70Mb drives for development and small projects.

10Mbps Ethernet

Economies of scale kicked in and researchers started to buy machines from Sun Microsystems rather than designing and building them in-house. These machines came with standard ethernet, as had lfcs's VAX-11/750 and Pyramid machines ("cstvax" and "cheops"). A LANCE-based ethernet card was built for the APMs and installed in a couple of the Filestores. Initially 10base5 ("thick ether") was used, but as clusters of Suns appeared in an *ad hoc* manner they were linked with segments of 10base2 ("thinnnet") and routed by adding second interfaces to one or more machines in each cluster.

The Department's machine halls also housed a Gould PowerNode ("itspna"), shared with the Artificial Intelligence ("AI") and Electrical Engineering ("EE") Departments, and a couple of GEC machines for ERCC's commercial arm. AI accessed the Gould by serial lines connected to their South Bridge site using a statmux pair and a BT kilostream link, while EE had their own ethernet bridged to ours over a fibre pair run through one of the KB site's heating ducts (figure 2). This link in fact remained in service until building work in 2000 meant the removal of the fibre.

Although there was this common infrastructure, there were actually at first a few islands of connectivity: all the BSD-based machines spoke TCP/IP; the GECs spoke Coloured Book protocols, for political reasons; and the Filestores and ECSVAX initially spoke neither. A TCP/IP stack was therefore written for these latter machines, including RIP-based routing, rlogin and rcp clients and servers, and lpr-based printing.
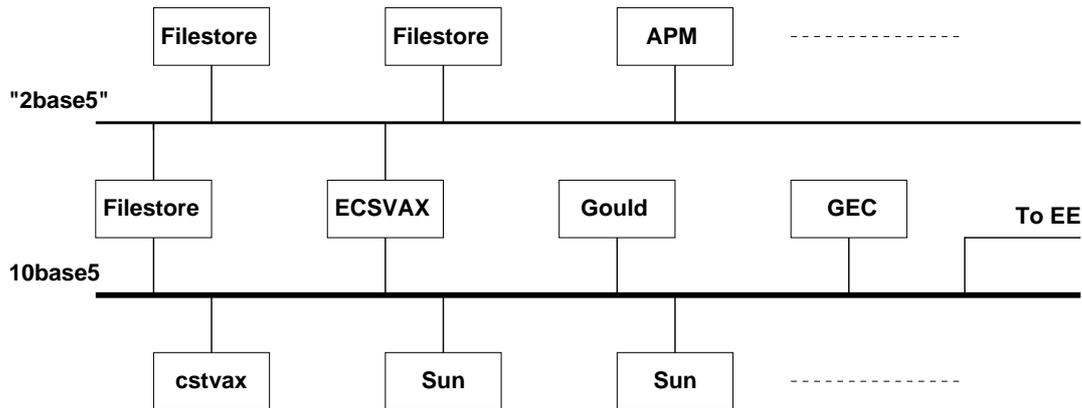
Figure 2: Network diagram from around 1986

Though perhaps not strictly part of the data network, it's worth mentioning that at this time the Department also ran serial lines to offices and a terminal room in 2501 from `ECSVAX`, `cstvax`, `cheops` and `itspna` using a Gandalf multiplexor.

Gradually as the APMs reached the end of their service life they too were replaced by Sun workstations, and the thinnet structure was necessarily rationalised. Discless Suns were heavy network loads, and so the network was explicitly subnetted into several staff segments and several student segments, with the Sun fileservers doubling up as routers. The "CS1 lab" in Appleton Tower was initially separate, before it too was linked to the rest of the Department at JCMB, first using UUCP and then SLIP over a 4800bps EPS8 BT line (fortunately both sites were within the same 667 exchange area). All subnets were /24, and routing information was distributed using Routing Information Protocol ("RIP").

Wider: RCOnet and EdNet

The wider University network is described elsewhere[3], so only those aspects which were CS-related are mentioned here.

For the first period of time covered by this note (mid-1970s to roughly 1980) the Department used serial terminals attached to the University's multi-access EMAS systems for its teaching lab in Appleton Tower, until that was replaced by Sun workstations. At JCMB there were also EMAS terminals in staff offices and Block O/P, but the Departmental and ERCC networks were initially almost completely separate. The only connection, such as it was, was the ERTE system, which could be used for small-scale file transfer between EMAS and the Filestore when it wasn't required for its research project work.

When the ERTE project finished the hardware was mostly disposed of. Its HDLC interface (a DQS11) was, however, redeployed in `ECSVAX`, with a locally-written driver and protocol modules (cite Gordon? PMM?). This provided a full set of remote login, file transfer, printing and email applications. `cstvax` also had a connection to EdNet, using an LSI-11 as a front-end.

TP and the "First Great Rewiring"

As the number of Suns in the Department increased, it was necessary to control the allocation of machines to subnets ever more carefully. Discless Suns paged heavily, and the applications used by some groups of users were relatively more network-intensive. This was done by having numerous 10base2 segments strung around the various parts of the Department's accommodation.

---

[3]https://www.wiki.ed.ac.uk/display/iticis/University+of+Edinburgh+Networking+Overview
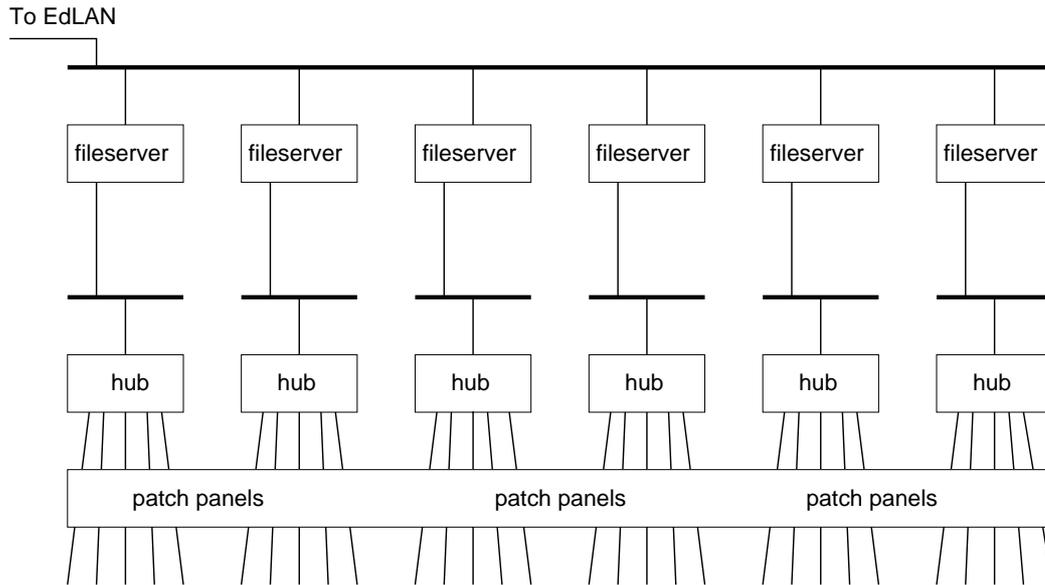
To EdLAN

Figure 3: Outline diagram after the First Great Rewiring

Physically, 10base2 isn't very robust. It's disruptive to add new connections, as the cable has to be cut and connectors carefully fitted. Connections are then subject to strains as machines are moved and feet tangle in the cables. Corrosion sets in (not helped by drips from wet bicycles!). Hum-loops can be created as connectors come in contact with building metalwork. Faults might affect everyone on the wire, or just the odd machine. Debugging is tedious and time-consuming, and can often "fix" problems for no obvious reason. Finally, we had too many subnets in use to be able to run them physically through all the offices, simply because of the size of the cable bundle, which put a severe constraint on the allocation of people to offices!

By 1991 we had concluded that we couldn't carry on with 10base2. We decided to have the "First Great Rewiring," introducing 10baseT on cat-3 cables. Physically this is much more robust than 10base2, and gives better, but not perfect, fault isolation. Crucially it made it possible to patch any office to any subnet (figure 3), considerably easing the room-allocation problems of the admin staff. It was also now even easier to split subnets to add bandwidth, and eventually we had around 25 subnets in use for various purposes. We were back to point-to-point connections, which had proved problematic with the Links and the Filestore, but we weren't now limited by backplane space. We had active electronics instead of a passive piece of wire, but the overall robustness more than made up for the added complexity. Wiring was installed by the Department's technicians, with one hub site in an alcove in room 2501, covering the offices, and one in the South Machine Hall covering the servers. A few hubs were also spread around under the Machine Hall floor for the students' client machines. At its peak this network had over 30 hubs in 2501 and another 15 or so in the Machine Halls.

This was a considerable improvement over what we had had before, but there were still issues. Although we had individual connections from machines and offices to the hubs, those hubs were interlinked using 10base2 (figure 4). We still had a shared medium — hubs are simple repeaters, and copy everything they receive out onto all their ports — and bandwidth was still under pressure. The hubs were connected together using the 10base2 cabling, which was rather old and fragile. And although faults were generally much more contained, hub faults could affect everyone on an entire subnet, often in strange ways, and were hard to diagnose.

And we now had the issue of how to manage the large number of connections at the 2501 hub site, particularly as people moved offices or had new machines to connect. Churn resulted in the 10baseT patch areas getting completely out of hand over time, so that on a couple of occasions we just ripped the whole lot out and repatched again from scratch.
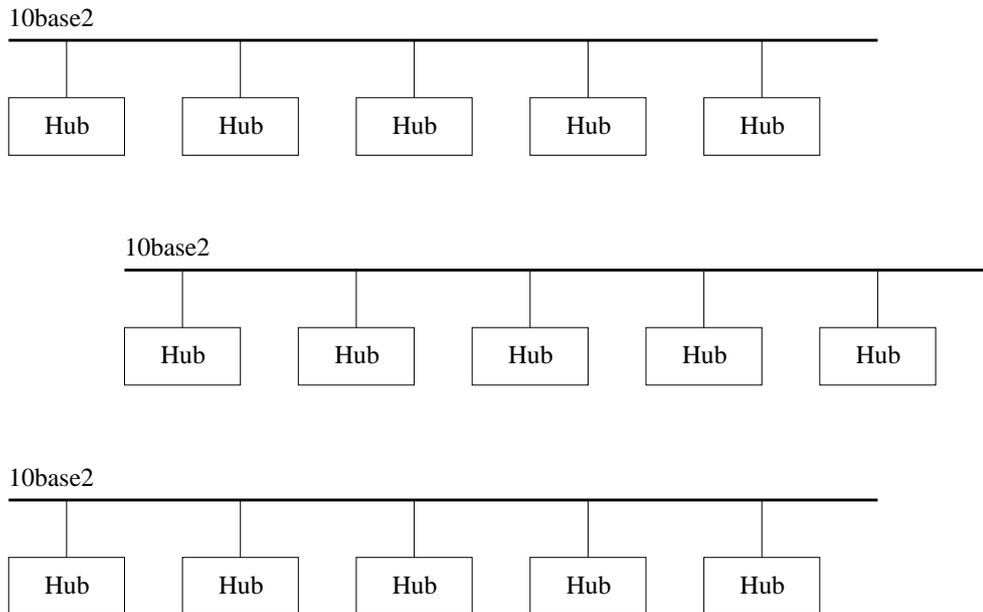
Figure 4: Hubs connected by 10base2

EdLAN

When EdNet was replaced by EdLAN the Department connected both its JCMB and Appleton Tower sites, and dropped the SLIP line as fast as we could. Up until then we had been using addresses chosen somewhat arbitrarily, but renumbered both sites onto the officially allocated addresses as they were connected to the new EdLAN.

Switches and the "Second Great Rewiring"

The arrival of switches (marketing-speak for multi-port bridges) opened up a couple of avenues for improving the network. These were intelligent devices (well, slightly...) with much better fault diagnosis and reporting than was available from the hubs. They also implemented VLANs, allowing for several subnets to be carried over the same physical wiring, and could group several links together into one "fatter pipe" to increase aggregate bandwidth.

We introduced a stacked pair of 3Com switches in the 2501 hub area and another pair in the machine halls server area. Hubs were now connected as individual leaf nodes off the switches, immediately reducing contention for bandwidth, and inter-switch traffic was carried over dual 100baseTX links (figure 5). Faults now had a much more localised effect, and the switches' diagnostic tools allowed for faster resolution and better forward planning. Once again we had complex electronics replacing simple wires, but the overall result was generally much more reliable. However, the patch areas were still completely out of hand!

Although the 3Com switches inproved things, there were still a few rough edges. The firmware didn't initially support VLANs; and the switches would sometimes crash and lose their configurations, though unstacking them did help with this. The biggest issue, though, was that we still had servers connected to shared segments at 10Mbps, albeit now at least single-hub segments.

The introduction by HP of their Procurve 4000M modular switches, at an affordable price and with a lifetime warranty, allowed us to rethink our network connectivity. Each machine could be given its own private 100Mbps link to one of the switches, and the switches themselves could be connected together using 1Gbps fibre and copper links. Ports could be assigned to VLANs based on configuration files, so there would be no more need for physical repatching and the mess that
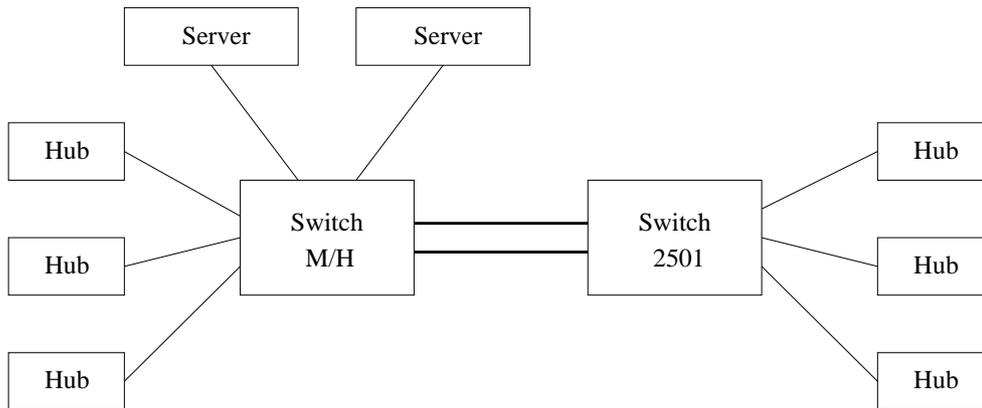
Figure 5: Hubs interconnected using switches

resulted. Redundant links could be added, under Spanning Tree Protocol control.

Unfortunately, in order to make use of these opportunities we had to undertake a Second Great Rewiring (in 1998)! Our initial TP cable installation had too many over-length lines, which passed the 10baseT electrical requirements but not the 100baseTX timing requirements. We had also doubled up many connections using the "spare" cable pairs, which 10baseT could tolerate but which caused unacceptable crosstalk for 100baseTX. Finally, most of our initial cable installation was cat-3. Once again the Department's technicians installed the new cat-5 cable, now terminated in two sites (rooms 1415 and 1619), and a pair of 4000M switches was installed at each. Further switches covered the teaching machines in the Machine Halls, and three more switches provided "core" connectivity as well as links to the servers. Most users had two outlets installed for them, with mini-hubs being used where more machines were needed, the realisation that the panels above the office doors were asbestos-based having meant that only existing holes could be reused for wiring! Though expanded later by the addition of the IPAB robot labs and some student workrooms, this design remained essentially unchanged until we moved from JCMB to the Informatics Forum in 2008 (figure 6).

Routing was still initially performed by the fileservers. In order to handle increasing traffic levels, however, we added a ProCurve 5308xl at the core, able to route traffic at wirespeed.

Configuration and Monitoring

While we still had only a small handful of switches connected as a backbone for the hubs, it was reasonable to configure them using their CLI or web interfaces. Once we had decided to roll out "switch ports for all," however, it was immediately apparent that some kind of configuration tool would be needed. The requirements were that it should support both the 3Com and Procurve switches with a common user interface; and that interface should expose only those items which were necessary, and in such a way that a simple subset could be made available for the technicians and support staff to allow them to make routine changes. The tool should run on a commodity machine, with the configuration stored in its usual filesystem and backed up in the usual way.

At the time there wasn't a suitable tool commercially available, so we developed our own (described elsewhere), based around Tcl, `expect`, and the Tnm SNMP package. A Procurve or 3Com switch is queried using SNMP for its current state, which is then compared against the switch's configuration file and any updates applied as required. For the Procurve switches those updates could also be done using SNMP. The 3Com switches, however, had such a baroque VLAN implementation that it was simply not safe to try to use SNMP there too, and so updates for those were applied by sending commands to the CLI using an `expect` script. The tool was deployed using the "`rfe`" remote file editing tool and `make`. CogSci also used the tool, with a front-end to
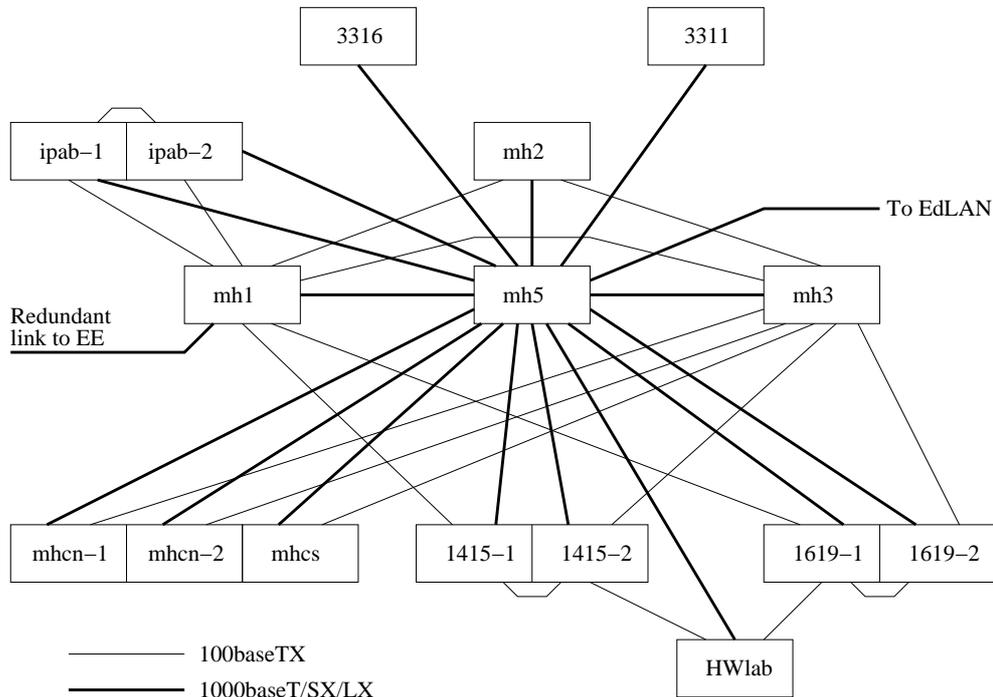
Figure 6: JCMB network using switches throughout

generate the configuration files. (A 2009 development project took another look and still couldn't find anything more suitable for our use than this in-house tool.)

We also needed better routine monitoring, for error and basic statistical trend reporting. Again, this was written in house, using Tcl and Tnm. All the switches are polled every 5 minutes, and their traffic counters extracted. These are stored using `rrdtool` for future use. Index files are generated several times per day, and the graphs are available as necessary using cgi-bin script wrappers for rrdtool's facilities.

Traffic filtering

While we had an isolated data network, the issue of allowing or blocking traffic from outside didn't, of course, arise, though not all our internal users were as benign as might have been desired. When JANET's "shoestring" project came along and we connected to the Internet, we initially relied on the underlying strength of the systems we were running to keep outsiders out. As the Internet grew and connectivity improved, this position became clearly unsustainable.

We were using Suns as routers at the time, so the natural tool to use was `ipfilter`. We reorganised our subnets slightly so as to have one "outside" subnet, linked onwards through EdLAN, with all the rest being "inside." We did consider setting up a DMZ in the usual way, but concluded that it didn't really fit into the way our network was used. We configured ipfilter initially by hand, and later under lcfg control, so as to block inbound traffic by default but permit outbound traffic by default. The more restrictive position of blocking everything by default would have resulted in an excessive management load then, as it would now. Later as we moved to Linux perimeter routers we ported our rules to `iptables`, effectively unchanged in the first instance, as this was the native filtering system available. We did consider using the ProCurves' ACLs, but these were, and remain, less flexible.

Blocking inbound and permitting outbound by default has remained the configuration to date. It's not ideal, but it is still the best fit for us from the well-known paradigm of "choose any two

from 'secure,' 'cheap' and 'easy to use'."

Formation of Informatics

With the formation of the Division (later School) of Informatics in 1998 came the requirement to integrate the constituent Departments' networks.

- The situation in CS was as described above: we had our main site at JCMB with the "CS1 lab" in Appleton Tower as a satellite operation. All machines were connected to a 100Mbps switch port, either directly or through a mini-hub. Hard patching had been replaced by soft configuration of the switch ports under control of the configuration tools.

- Artificial Intelligence ("AI") was split across two sites: Forrest Hill and South Bridge. The Forrest Hill site had been rewired with cat-5. South Bridge was still mostly 10base2.

- Cognitive Science ("CogSci") was sited at the East end of Buccleuch Place. The buildings had been rewired with cat-5 and a switch-based installation was in use, very similar to CS's.

The requirements were as follows: Informatics machines should be treated as part of one unified network which happened to be spread across the five (initially) sites; and the network should be robust, there should be transparent fast paths between sites, and all external traffic should be filtered. We decided that it would be best to keep subnets and their associated VLANs within sites as much as possible, to minimise inter-site and EdLAN dependencies, a position we have maintained to date; we introduced a "transit" subnet for inter-site traffic, with only routers and a small number of network infrastructure machines on it; and we gradually introduced switch/routers throughout, to improve intra- and inter-site routing throughput. Inter-site traffic was carried by tagged VLANs across EdLAN, for the most part. We did have some dark-fibre links between AT and FH and between AT and BP for a while, but instabilities introduced by being part of the EdLAN spanning tree meant that we had to keep them mostly disabled.

Routing was still controlled using RIP, initially using `gated` and latterly `quagga`. Traffic was directed through preferred primary and secondary routers by adjusting the metrics advertised. We realised quite soon that the use of RIP at our privately-interconnected sites did not interact well with the OSPF that EdLAN was running — in particular, metrics were lost in the conversion, and so taking routes from EdLAN at one site, propagating them internally to another site, and then forwarding them from that site to a different EdLAN router was not safe — and so we had to introduce heavy filtering of the RIP routes we sent and received. The realisation that a large amount of Physics's traffic was transiting across our networks as a result of RIP/OSPF interaction helped to increase the priority of this work.

South Bridge fire and relocation to Appleton Tower

The South Bridge fire in 2002 did not actually have much of an immediate impact on the network, other than getting rid of all remaining 10base2. Our Appleton Tower core was promptly upgraded by the addition of a switch/router, and temporary edge switches in level-3 and level-5 were connected to it. The relocation of the core from level-5 to the basement server room was done essentially transparently to the users by relying on the redundant inter-switch links to keep the traffic flowing while relocating the main links to new fibre. As floors were refurbished, a pair of edge switches was installed in each and linked back to the core. (When floors were later upgraded we moved to the "Forum" model of a hierarchy of switches, as a more cost-effective solution.)

The Appleton Tower network was designed in the light of experience at our other sites (figure 7 on page 10). Core and edge switches were clearly distinguished by rôle, and models were chosen for their particular rôle, not as a compromise for mixed use. The initial design had one switch/router at the core, alongside one non-routing switch. The edge switches, two per floor, had a "main" 1000baseSX link to atba and a backup 100baseFX link to atbb. A PoE switch was added, with
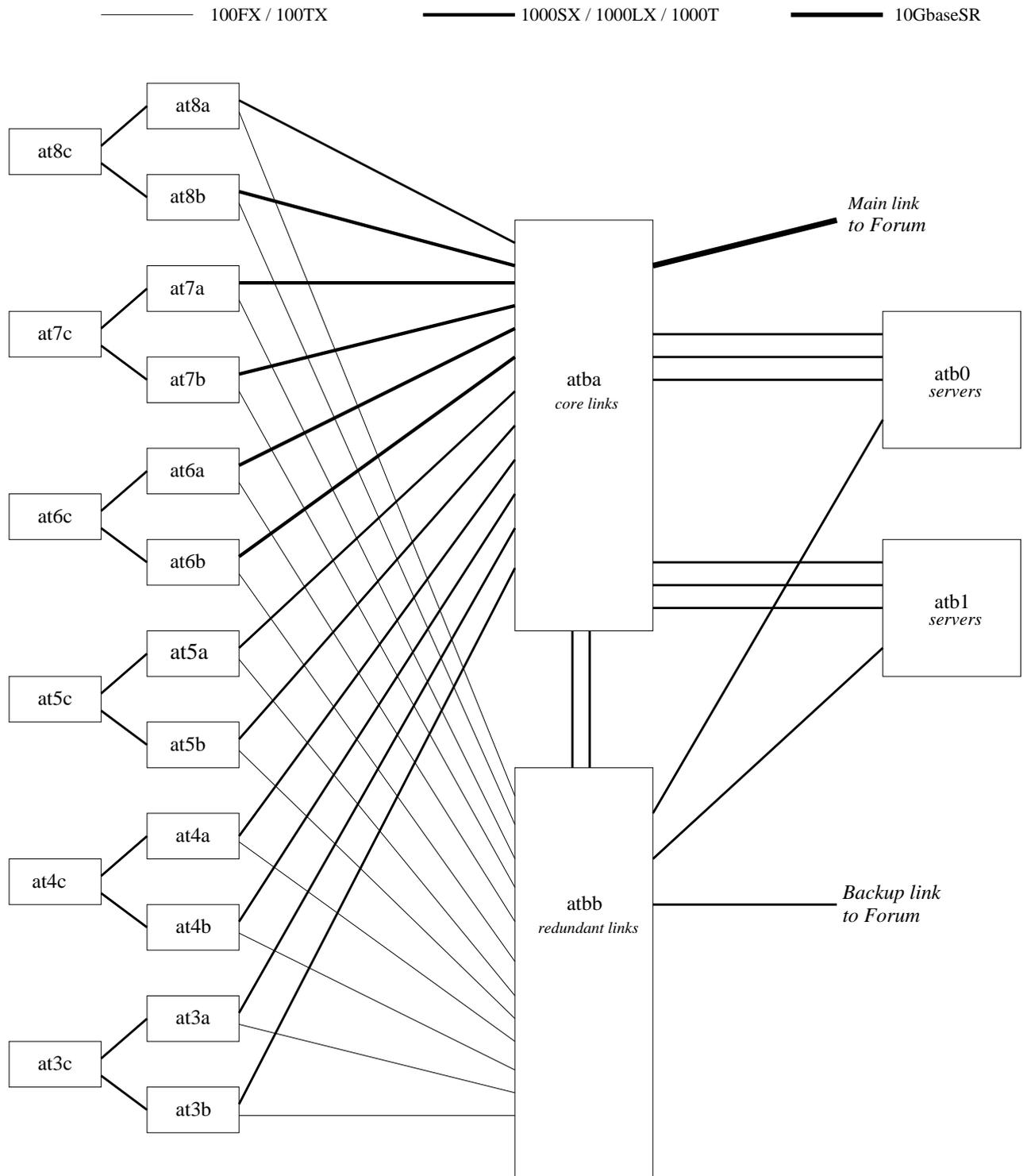
9

Figure 7: Appleton Tower network

one link to each of the floor's edge switches. Those original core switches have since been replaced by a pair of more capable switch/routers, with a 20Gbps trunk between them and with one edge switch from each floor connected to each.

Informatics Forum

The design of the network in the Informatics Forum once again drew on experience at our other sites. As the building was being constructed from scratch we were able to get a network design which was reasonably close to the way we wanted it, rather than having to be retro-fitted to an existing building. The funding included provision of new switches, as the ones we had installed in our previous sites would have been due for replacement anyway.

At the core of the building network is a cluster of three ProCurve 5412zl switch/routers, each connected to the other two by 20Gbps trunked links (later increased to 40Gbps). Each of these switches also has a link to one of the Appleton Tower core switches, the whole forming one tree under Rapid Spanning Tree Protocol control with priorities and link costs chosen so that the tree is stable and traffic is contained within the two sites and passed preferentially over the desired links. These switches have four power supplies, connected to three different power circuits and three local UPSes, for robustness, but the intention of the design is such that one or even two of the core switches can fail without the service for the rest of the building being much affected.

The physical cable installation was designed in detail by the building's engineers, to our specifications. Each floor has from one to three "IT closets," located to fit in with the building's other services. In the upper floors the closets are sited in three columns, vertically one above the other, with a direct cable path between. Unfortunately the ground floor and basement closets are offset slightly from this arrangement. Fixed copper wiring radiates out from each closet to a number of distribution boxes under the raised floors. Floorboxes, with six data and four power sockets, are then connected to these under-floor distribution points. The intention was to allow flexibility for floorboxes to be added or moved later without the disruption of having to wire them all the way back to the closets. There are also a number of fixed data outlets spread around the walls of the building, for printers, wireless acecss points and the like.

Of the six data points in each floorbox, four were initially patched in. Each user was provided with one 1000baseT link, two 100baseTX links and one VoIP phone link. As with our other sites, these could be assigned to VLANs under the control of the configuration system. This arrangement was chosen as being a reasonable compromise between cost and speed. Each closet then has fixed-format ProCurve switches installed, sufficient to service the required number of each type of port. At the time of writing there are more than 150 switches in the Forum, providing for rather more than 3000 data connections and almost 400 phone connections, patched in 17 closets and two server rooms.

Unlike Appleton Tower, where each main edge switch had its own direct link back to the core, not all the Forum switches are directly connected, due to the cost of doing so being prohibitive. Instead (see figure 8 on page 12) all the Gbps switches in each column of closets are interconnected vertically using 10GbaseT links (initially 10GbaseCX4). Roughly every third switch in each column then has its own direct 10GbaseSR link back to the core, with the remainder being indirectly connected. 100Mbps switches are then connected indirectly via one or other of the Gbps switches. In all cases, connections are arranged so that each switch has at least two independent paths back to the core, all still under RSTP control, so that the failure of one switch does not result in a loss of service for any other. There are also some cross-links between the columns, partly for robustness and partly because it turned out in some cases that this was the best way to provide dual paths back to the core. Each closet also has one PoE switch, to power wireless access points. The phone network effectively parallels the Gbps data network, to minimise the chances of either impacting on the service of the other. In all cases "close" switches are linked back to different core switches for added robustness. This arrangement has shown itself to be easy to extend: additional Gbps switches can be added into the existing columns, with additional uplinks to the core as necessary, with new 100Mbps switches connected to the existing Gbps and 100Mbps switches.
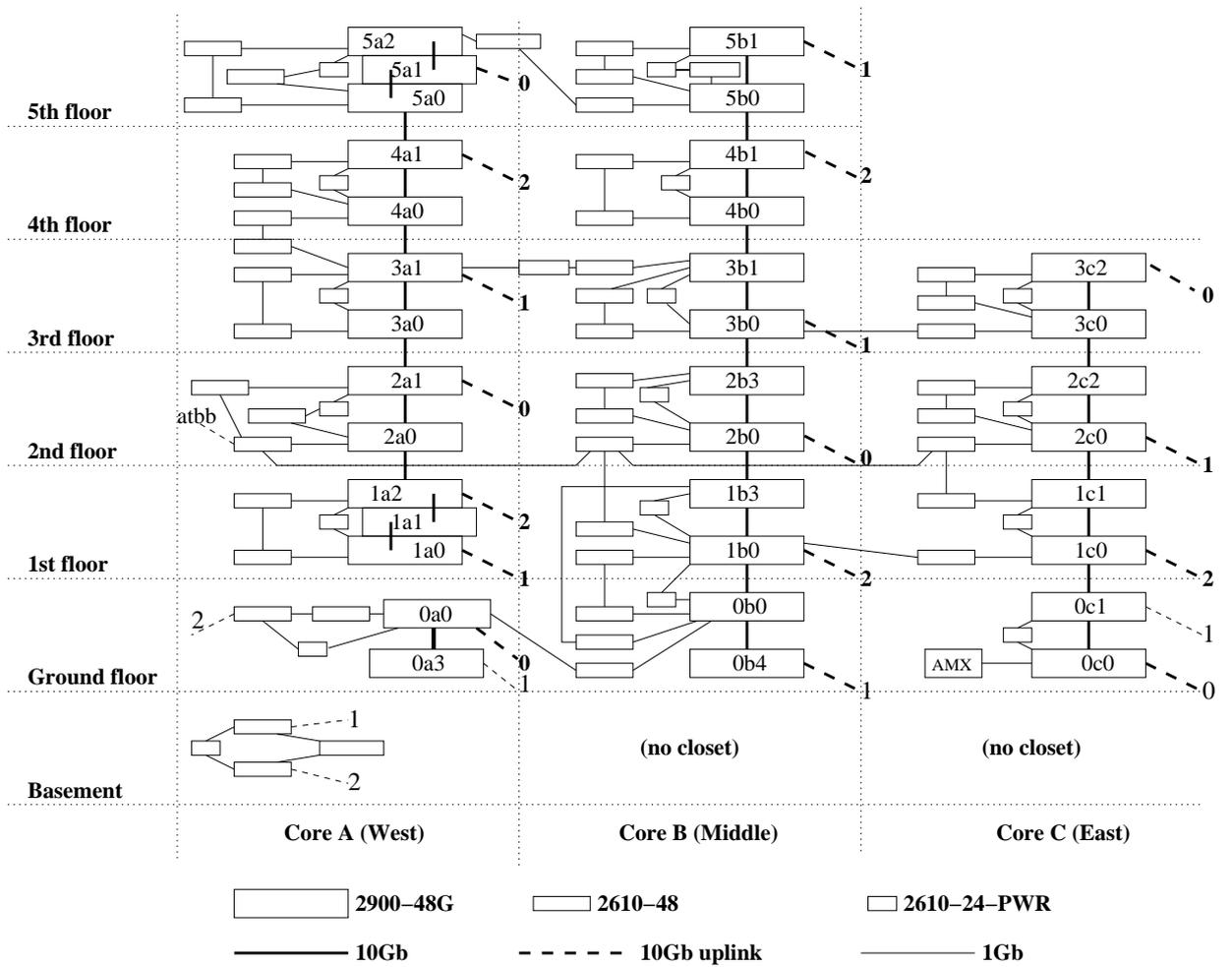
Figure 8: Forum edge date network

The main server room in the Forum is fairly straightforward. Each group of three racks has a pair of Gbps switches installed along with a redundant power supply. The pair is connected together using 10GbaseT, and each switch is connected back to a different core switch, again using 10GbaseT. Finally, each switch has a 1000baseT link to a switch somewhere else in the room, chosen so that it links back to the third core switch. The self-managed server room has four switches, linked together in pairs and linked to different core switches.

Routing using OSPF

We had been using RIP as our routing protocol ever since we had started running TCP/IP, in spite of its awkward interface with EdLAN. However the move to the Forum prompted us to re-evaluate. A larger building suggested that sticking with /24-sized subnets throughout would make client management rather more complicated than it really needed to be. We did consider RIPv2, but as EdLAN was using OSPF as its native routing protocol it made much more sense for us to adopt that too.

In conjunction with Information Services ("IS," formerly UCS, formerly ERCC), it was decided that all of Informatics should be in one OSPF area (area 64). The Area Boundary Routers are the EdLAN routers at JCMB (`kb6` at the time) and Appleton Tower (`at5` at the time; both have since been upgraded). The two "external" subnets (42 and 160) each have one EdLAN router and all our Linux perimeter routers on them, to provide for better control of external traffic as well as robustness against router failures. The transit subnet remains, but now has only switch/routers (and a handful of infrastructure machines) but no Linux routers. Each site has one subnet where Linux routers and switch/routers speak OSPF, and routes for all other subnets are injected by the switch/routers acting as ASBRs. Traffic is directed preferentially by setting the routers' path costs — this is necessary to ensure that external traffic is routed symmetrically, so that `iptables`'s connection-tracking works properly. Figure 9 shows this in outline form: indicative switch/routers are labelled "Sxx" and Linux routers "Lxx." As well as allowing for non-/24 subnets (we now have a variety of /22–/26-sized subnets in use) this arrangement provides improved robustness, by allowing any site to act as a standby perimeter router for any other site.

The transition from RIP to OSPF was handled in several stages. The first step was to enable OSPF on all the core switch/routers, but with the administrative distance set so that RIP continued to be preferred. Once that was in place, OSPF was introduced on all the Linux routers. Finally, RIP was disabled everywhere, and we had a pure OSPF site.

Moving to OSPF also meant that we had to change the way routes are distributed to non-routers. While we had RIP this was done by those non-routers snooping on the RIP traffic. We didn't want non-routers participating in the OSPF exchanges, as this would have loaded the DRs and BDRs unnecessarily as well as weakening security. We considered RIPv2, but concluded that as we only really needed to distribute default routes it would have no advantage over router-discovery. We have, however, tuned the router-discovery parameters to provide a faster failover time. Those machines which can not listen to router-discovery are expected to honour the values sent to them by DHCP.

AT Basement Refurbishment

The Appleton Tower basement server rooms were refurbished in 2009, and we had to decant all our kit out, mostly to Forrest Hill, for the duration. The comms rack was able to stay *in situ*, but behind a temporary wall and with no power or ventilation. Although spread across three locations (Appleton Tower upper floors, Forrest Hill server room and the Forum server room), we decided to maintain the "AT network" identity throughout, to make management of clients, servers and self-managed systems easier.

We installed a pair of temporary switch/routers in the Forum server room, initially running in parallel with the AT core. They were connected to the Forum core switches, which had had the

EdLAN

kb6   at5

.160

.42

Informatics

KB
1210   1210   1210
1210
LK1
1210
800
SK1
800

AT
1110   1110   1120   1120
LA1   LA2
1110   1120
800   800
SA1   SA2
800   800

Forum
1010   1010   1020   1020
LF1   LF2
1010   1020
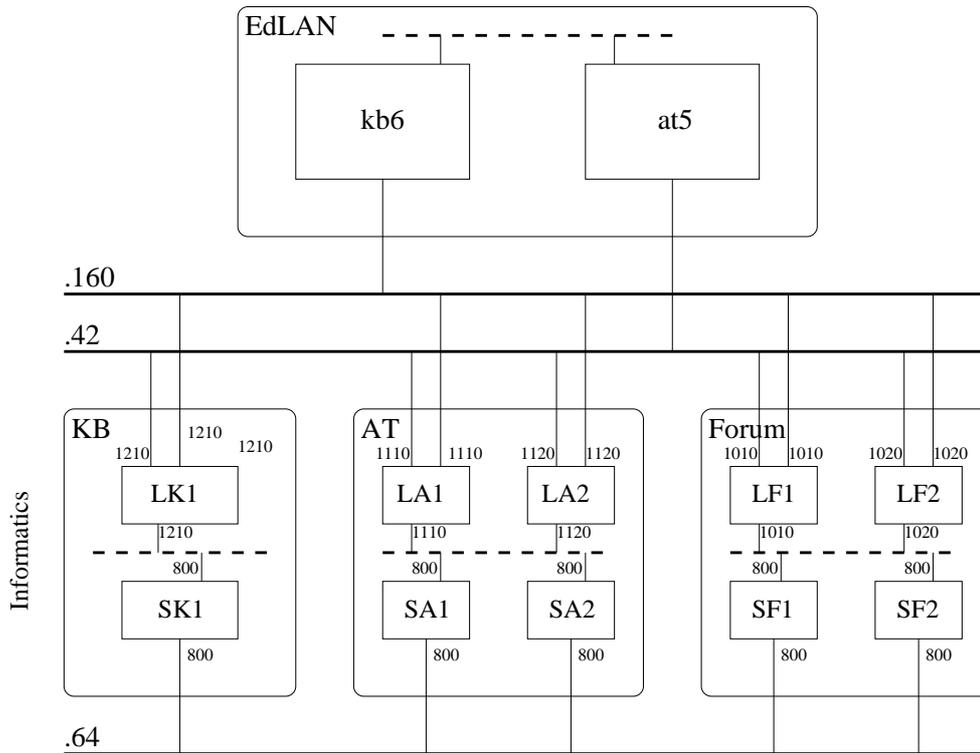800   800
SF1   SF2
800   800

.64

Figure 9: Outline OSPF routing arrangements

AT VLANs added. Once those switches were up and running we were able to re-patch the upstairs switches to them using the fibre we had had installed between the buildings when the Forum was built. The decanted servers in Forrest Hill were catered for by installing a pair of Gbps switches, connected back to the Forum over single-mode dark-fibre. Figure 10 shows the arrangement, with the four switches above the dotted line being in the Forum and the two below in Forrest Hill.

Once the refurbishment was complete, the AT core switches were reinstalled and the upstairs floors re-patched. A pair of new switches was installed in the server racks, connected back to the core using three 1Gbps lines trunked together as the distance was too far for CX4 and the expense of fibre wasn't really justified for the switches involved.

When these server room switches were later upgraded we did also upgrade their uplinks to use 10GbaseSR. We decided on fibre rather than using 10GbaseT for a couple of reasons:

1. the links were run on cable baskets shared with power and other unidentified data lines, and we weren't confident about 10GbaseT's noise immunity under this arrangement; and

2. we needed SFP+ cards in the switches anyway, for links to the Forum and to EdLAN, and this allowed us to share resources.

Moving out of Buccleuch Place and Forrest Hill

After we had moved out of the Buccleuch Place and Forrest Hill offices into the Forum we still had machines in the server rooms there for quite a while afterwards. In conjunction with IS, therefore, we decided to continue to manage the edge networks for the new occupants of the offices in the buildings using the existing switch infrastructure. This was reasonably straightforward on the whole, as there was not nearly as much VLAN churn as would have been produced by our own users. Once all our machines were out of Buccleuch Place we reset all the switch passwords and SNMP communities to temporary transitional values and handed the entire network installation
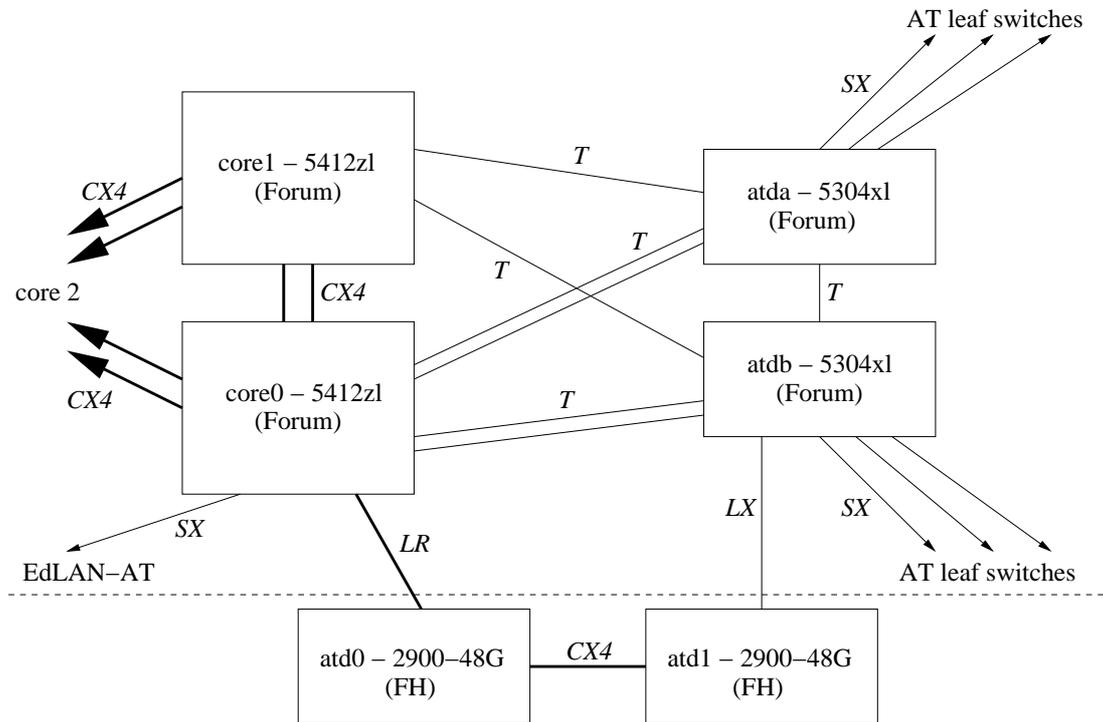
Figure 10: Appleton Tower decant arrangements

over to IS to manage from then on. The same procedure was followed when we moved out of Forrest Hill, though this was delayed by the Appleton Tower basement server room decant.

Decanting out of Appleton Tower again!

In 2015 the Appleton Tower recladding project[4] required us to decant out of the upper floors of the building while the work was in progress. Teaching went back to Forrest Hill, which had been extensively refurbished to suit our needs as much as possible, while non-teaching moved to the Wilkie Building. We moved out in Summer 2015, with the intention of moving back in Summer 2016. Additional building work may, however, result in this being delayed by a year.

For simplicity it was decided to run both Forrest Hill and Wilkie as "virtual floors" of Appleton Tower, connecting back to the core using 10GbaseLR links over dark fibre. There were a couple of advantages to this:

1. machines could be taken off their AT desks, installed and connected in at their new locations, and would then continue to work with no further reconfiguration required; and

2. there was a cost saving, as core infrastructure was not required at either site.

The disadvantage, of course, is that it introduces outside dependencies, but given the expected duration of the decant this was felt to be an acceptable compromise.

IPv6

Recognising that IPv6 support was becoming more and more necessary, a development project was started in 2015 to investigate how this might be provided on the Informatics network, and to

---

[4] http://www.ed.ac.uk/projects-estates/current-projects/appleton-tower-recladding

provide an initial service for managed machines. The expectation is that this initial service will be available from mid-2016. Extending it to self-managed machines will then require some more development, primarily for host configuration and audit, and the replacement of some more old, less capable switches.

People

The following were involved in the events described above:

- Peter Lindsay designed the Link boards

- Rainer Thönnes and Paul McLellan wrote the original Interdata Filestore system, with input from Hamish Dewar

- Rainer Thönnes designed the 2Mbps ether stations and wrote the Fred machine system

- Fred King designed the Fred machines (APMs)

- Gordon Brebner wrote the WAN software for `ECSVAX`; Paul McLellan wrote the DQS11 driver

- the Author wrote the APM-Filestore system, based heavily on the "1976" Interdata version and with contributions by Rainer; the TCP/IP stack for APM and VMS; and the switch configuration and monitoring tools; and designed the Department's and School's data networks from the First Great Rewiring to date.

The following made some very useful comments on an early draft: John Davis, Hamish Dewar, Gordon Hughes, Mike Hurst, Nick Rothwell, Frank Stacey, Rainer Thönnes, Graham Toal. Thanks to Alan Boyd for the EdLAN documents.