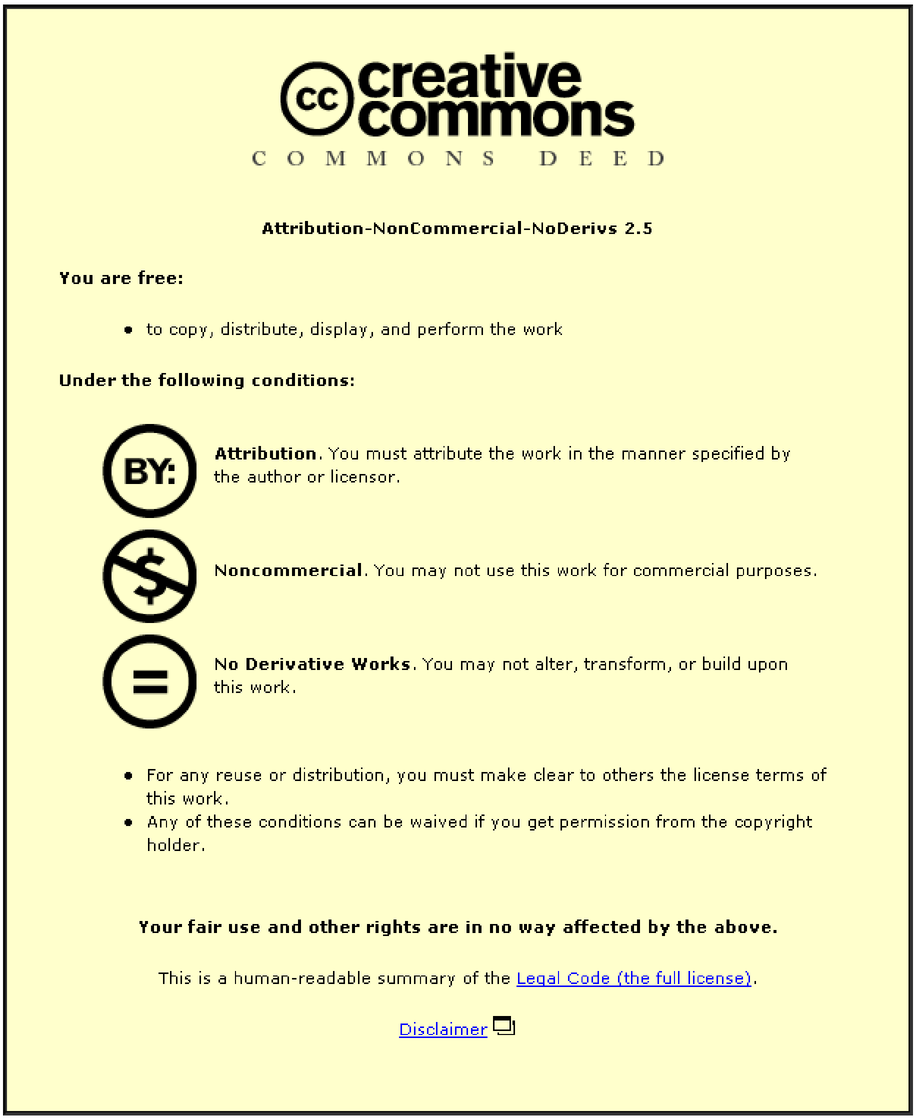

Challenges in the capture and dissemination of measurements from high-speed networks

R. G. Clegg[1] (richard@richardclegg.org)
M. S. Withall[2] (M.S.Withall@lboro.ac.uk)
A. W. Moore[3] (andrew.moore@cl.cam.ac.uk)
I. W. Phillips[4] (I.W.Phillips@lboro.ac.uk)
D. J. Parish[2] (D.J.Parish@lboro.ac.uk)
M. Rio[1] (m.rio@ee.ucl.ac.uk)
R. Landa[1] (rlanda@ee.ucl.ac.uk)
H. Haddadi[1] (hamed@ee.ucl.ac.uk)
K. Kyriakopoulos[2] (K.Kyriakopoulos@lboro.ac.uk)
J. Augé[3] (jordan.auge@cl.cam.ac.uk)
R. Clayton[3] (richard@highwayman.com)
D. Salmon[5] (D.Salmon@ukerna.ac.uk)
[1] Electronic & Electrical Engineering, University College London
[2] Electronic & Electrical Engineering, Loughborough University
[3] Computer Laboratory, University of Cambridge
[4] Computer Science, Loughborough University
[5] JANET(UK)

November 6, 2008

Abstract

The production of a large-scale monitoring system for a high-speed network leads to a number of challenges. These challenges are not purely techinical but also socio-political and legal. The number of stakeholders in a such a monitoring activity is large including the network operators, the users, the equipment manufacturers and of course the monitoring researchers. The MASTS project (Measurement at All Scales in Time and

Space) was created to instrument the high-speed JANET Lightpath network, and has been extended to incorporate other paths supported by JANET(UK).

Challenges the project has faced have included: simple access to the network; legal issues involved in the storage and dissemination of the captured information, which may be personal; the volume of data captured and the rate at which this data appears at store. To this end the MASTS system will have established four monitoring points each capturing packets on a high speed link. Traffic header data will be continuously collected, anonymised, indexed, stored and made available to the research community. A legal framework for the capture and storage of network measurement data has been developed which allows the anonymised IP traces to be used for research purposes.

1 Introduction

The common availability of quality monitoring hardware, high-performance computers and a ready supply of interesting network-use has led the research community to somewhat become blasé about network monitoring. However, a short discussion with any practitioners of network-monitoring reveals that the topic is both complex and fraught. We intend this paper to serve two purposes. Firstly it provides a roadmap, a commentary and insight for future contributors in the monitoring field and secondly it describes a data resource which will be of great use to the network modelling and analysis community.

The MASTS project (Monitoring at All Scales in Time and Space) [1] is an Engineering and Physical Sciences Research Council (EPSRC) funded collaborative research project between three universities: Loughborough, Cambridge and University College London (UCL). The project aim was to create and operate a monitoring system for various links on JANET (Joint Academic NETwork and JANET Lightpath. JANET is the National Research and Educational Network (NREN) for the United Kingdom, it provides Internet connectivity among all UK Universities and institutions of education along with most research organizations (for example, Welcome Trust and Research Council UK facilitie). JANET Lightpath is a 10Gb/s network, previously known as UKLIGHT, which is operated in the UK by JANET(UK), previously known as UKERNA. It supports a range of research activities and carries traffic from a number of Grid research projects. In the MASTS project the JANET Lightpath network provides both a system to monitor and a backhaul for the collected data.

MASTS aims to provide information to network operators, network users and network researchers. The challenges of monitoring high performance networks as addressed by MASTS offer a systems level set of solutions to communication network monitoring and the monitoring interface, storage and legal aspects are presented in this paper. The project has also investigated the visualisation, compression and analysis of monitored network data, but these aspects are reported elsewhere [2, 3].

The ultimate aim is to provide to researchers a database of layer two, three and four header information for four monitoring points on the network, three of which are carrying scientific/technical data on the JANET Lightpath network and one of which is carrying data on the main JANET network. An anonymised version of this data is made available to all researchers who sign an Acceptable Usage Policy. The data sets are catalogued in a searchable database and enhanced with metadata.

The internet, once a mere research-vehicle, now forms the background for substantial parts of the economy and is fundamental to much social intercourse. Improving performance of broadband IP networks have been central to this with IP networks able to carry any data type. The heterogeneity of IP networks, their ability to carry a *triple-play* of services (Television, telephone and data-services) to every broadband consumer has lead many ISPs to transition to IP-based national backbones (*e.g.* British Telecom's 21st CN [4]) and will motivate movement to an IP-based network at the foundation of all communication services. Our system, aimed at 10Gb/s, is ideal for monitoring the current-generation backbones and next generation distribution-networks of such new broadband networks. Understanding drawn from MASTS will permit both a better understanding and more sophisticated optimisation of an IP-based world.

1.1 Background and Motivation

Many researchers approaching network monitoring with a need for network data (perhaps to validate a theory or provide input to a simulation or study) quickly find themselves overwhelmed by the complexities of monitoring. Performing meaningful monitoring operations on high performance networks is a complex challenge, which embraces not only the technical issues of connecting to a network and storage of the information collected, but also the procedural and legal issues of allowing this information to be disseminated to interested users world wide.

Network monitoring covers a vast spectrum of activities from using *tcpdump* [5] or *wireshark* [6] on a personal computer and understanding why your browser is misbehaving, through to the wide-area monitoring of data flows across an entire ISP as input to auditing, accounting or network intrusion detection. However, in all but the most trivial network-monitoring, the researcher will need to interact with the operators of the relevant network. When such networks are in-house this can make the process easier. However, there is no guarantee. Network operations staff are focused on the day-to-day and longer-term operational needs of a network; researchers wishing to monitor networks — often focused on their own research-deliverables — may only distract from the day-to-day operations and are commonly seen as a *tax* upon operators time and resources. Such diverging interests are not the only trap in network-monitoring; there is a vast array of different legal and technical challenges to be faced [7].

A number of research groups and organisations have developed network monitoring systems including CAIDA [8], RIPE [9] and NLANR (which is no longer operating). Many of these projects use active measurement whereby specific packets are generated and added to the existing network traffic. In such scenarios, a degree of control exists with respect to the rate at which measurements are made. Many of the probe designs such as those produced by the RIPE Test Traffic Measurement project [10] are intended for use on multiple internet paths and provide results for many different paths at a low rate. However, the lack of performance monitoring and diagnostic mechanisms has been highlighted in several places [11, 12].

The MASTS project therefore has designed passive probes for use on 10Gb/s link (note that throughout this article we will use B for bytes and b for bits, thus 10Gb/s represents 10Gbits/s). As such, node deployment is sparse compared with many other projects but more data is generated from each probe. It is generally recognised that network monitoring is a complex, multidisciplinary activity requiring the optimisation of many parameters. Some of these issues have been addressed by CAIDA [13]; whereas this paper presents the main issues and solutions as seen by the MASTS project.

From the outset, the project took to heart the adage *good data outlives bad theory* [14, 11], hence a long-term archive of activity in the JANET Lightpath network was planned. The monitoring system was designed to cope with a growing network and the database system is intended to provide a long-lived resource to the community.

Many worldwide projects exist that collected and/or disseminated packet-level traces, *e.g.* the previously mentioned CAIDA and NLANR [15] projects, the CRAWDAD repository [16] and the Bellcore project [17] are all good examples. The aim of MASTS is to complement these data sets with data from faster links, provided online soon after it is generated. The availability of this data and extra-derived data (we keep both packet traces and aggregated flow information for longer periods) is crucial for several communities. Traffic analysis, long range dependency, fault analysis, denial-of-service detection, *etc.* can all profit from a large data set representative of a significant Autonomous System.

1.2 Legal Issues

The very process of passive network monitoring involves the capture (and usually) storage and analysis of information generated by users other than those involved in the monitoring process itself. Potentially, this can lead to serious legal issues associated with privacy and data protection. This situation is generally influenced by some or all of the following characteristics:

- the purpose of the monitoring operation;
- the ownership of the data so collected and its location;
- the anonymisation approach adopted;
- the nature of the data to be collected, including the protocol layers;
- the sources of the data and
- the form of the data to be stored and disseminated.

In order to manage the legal status of the monitoring activities, the particular combination of these characteristics determines the legal status of the monitoring activity and the liable parties for any abuse. The approaches developed by the MASTS project are discussed in Section 3.2.

2 Architecture

The architecture of any network monitor is largely informed by the link to be monitored, the constraints of cost and the objectives the project may seek to optimise. In the case of the MASTS project the intention from the outset was to design capture systems that perform full line-rate capture. This is not to imply capture every octet of every packet will always be captured. However, a system was desired that was engineered to allow as close to this as technically and legally permissible.

The first hurdle was to design a monitoring system to the physical interface of the network-link to be monitored. The opportunity of the JANET Lightpath project, a new network infrastructure, provided a unique chance to build a monitoring system in concert with a specific physical infrastructure. Network practitioners will recognise that there are numerous ways a particular link/capacity may be provisioned. A range of physical options (copper, fibre, wireless) along with a range of data-link-layers (SDH, packet over SONET, raw (LAN) Ethernet) and a wide range of speed options led to a huge number of alternatives, each with its own cost and benefit.

By being involved with the operational-deployment from the outset the project allowed for fibre (interception)-needs and space-needs be accommodated, while keeping the monitoring team appraised of the operational-network's deployment. The physical links of the infrastructure are capable of 10Gb/s, however, the majority of the installation was based upon a specific vendor's proprietary SDH frame format. We could not monitor these links using a splitter alone. This led to two different solutions: one for parts of the JANET Lightpath network with an alternative approach for other monitoring installations.

The physical interconnections dictated two different approaches to present data from the three different links being monitored:

- JANET Lightpath: dedicated line card;
- JANET Lightpath RAL–CERN: 10Gb/s splitter and
- JANET internet interconnect: 10Gb/s splitter.

Aside from physical interconnections the specific project goals led to an architecture optimised to minimise uncontrolled loss while allowing best control over the long-term archiving of data.

2.1 Physical Architecture

Traffic interception is subject to constraints in both the political and engineering fields. While the political considerations are discussed elsewhere, we describe the two physical solutions employed in our implementations. Clearly the design of any capture system is tightly coupled to the physical media. In the case of MASTS, some physical media installations did not lend themselves to interception. To be economically intercepted (without the need to construct special purpose equipment running to millions of pounds) the physical line representation needs to be able to be interpreted by monitoring hardware (best thought-of as enhanced interface boards.) This is entirely practical when the physical line-encoding is one of a number of standards: for 10Gb/s the relevant IEEE Ethernet agreed standards are:

- the 10Gb/s LAN PHY: a physical layer for use in short-haul networks and
- the 10Gb/s WAN PHY: a physical layer for use in longer-haul networks and compatible with common telecommunications (SONET/SDH) equipment.

Figure 1 **A** provides a diagrammatic representation of a splitter internal: for each direction of flow a percentage of light is redirected to a second output. Splitting the light-flow in each direction provides two flows of data from the intercepted physical interface. (While 50:50 splitters may be used, more usually 80:20 or 90:10 where the majority of the photons are not intercepted, is common practice.)

Figure 1 **B** provides illustration of a trivial intercept: collecting data flowing between *host1* and *host2*. The splitter provides intercepted data for the *Capture System*, the hardware of the capture system may range from simply a pair of unused network interface adapters through to dedicated capture hardware. The differences between a simple solution and a more sophisticated approach, such as that described here, relate to the accuracy of time-stamping within the capture system. Standard network interface cards may not provide an accurate timestamp or sufficient card capture facilities to minimize loss. Buffer memory is a critical resource to overcome bandwidth limitations in a computer architecture, (often orders of magnitude more than that provisioned on a regular network interface card). Alongside this, the network interface card needs to provide appropriate hardware and software support for the most efficient mechanisms to move data into the capture system. Given our approach is, at first approximation, to capture all data on the physical interface we do not need the ability to selectively filter and discard irrelevant data (a feature often present on network hardware).

While standards such as the LAN PHY 10Gb/s Ethernet and the WAN PHY 10Gb/s Ethernet are common, the physical presentation may not follow such an open standard; such is the case for some of the links within JANET Lightpath. This led to a rather different solution for one of the MASTS monitoring systems; a dedicated monitoring port in the network infrastructure is used to mirror traffic from particular ports. Illustrated in Figure 1 **C**, this approach may be recognizable to readers as similar to the *Switched Port Analyzer* (SPAN) on Cisco switch equipment [18].

The project's use of port-mirroring differs in several important ways. Firstly, in a switching infrastructure the use of port-mirroring may lead to high levels of packet jitter and packet loss [19]. Secondly, it is important to over-provision the monitoring port. Clearly monitoring a 1Gb/s connection will require 2Gb/s of monitoring capacity (1Gb/s for each direction). For architectural reasons these two problems have limited impact on our use of port-mirroring in JANET Lightpath. The port-mirroring activity is done by a TDM (time-division multiplex) switch at the TDM level, this means that the timing relationship between packets in a single direction is undisturbed and the timing-error between packets of each direction within a multiplex is a small bounded number of the order of a TDM slot-length (*e.g.* 15.625μs); thus it may be easily corrected in the capture system. The second issue of over-provisioning is addressed by this approach being limiting to the monitoring of at most 5 full-duplex circuits. Within the JANET Lightpath service each circuit is typically provisioned at 1Gb/s and most services are based upon these 1Gb/s circuits (although finer grained provisioning is possible).

In this particular configuration a port loopback is used and a splitter is employed to extract the intercepted data-stream. This is because the intercept board does not provide any input data. The capture board has no reason to transmit data and thus has no 10Gb/s laser. However, in-common with much telecommunications equipment, without a valid input the monitoring port will not initialise and send any data. One solution is to use the loopback, sending the monitor-port data back into the monitor-port. The switch will not actually process this data as no paths are configured from the monitor-port to any destination. This eliminates the risk of (unintentionally) injecting replica junk traffic back into the switch.

Once intercepted, packets need to be stored, processed and passed to the database back-end without loss (or at least loss-limited) continuously. Capture systems in the past have often operated in a *capture to local disk* for a period and then off-line they would move or process relevant data. As noted above, continuous capture was a driving imperative for this architecture. Figure 2 illustrates the capture system we employ. The physical architecture is optimised for lossless capture of all packets on a particular physical link. This means adequate provisioning of intermediate storage is needed throughout the capture system. Obviously data will require buffering at every point where throughput may be discontinuous. These discontinuities are the interfaces between parts of the capture system as well as the parts of the capture system where data-processing may, for short periods, exceed available resources. Data is intercepted using an Endace DAG6.2SE (a purpose-built network monitoring card) in a dedicated Dell PowerEdge 2850 (dual 3GHz Xeon processors, 4GB memory). While the capture card is capable of receiving 10Gb/s, legal constraints restrict the capture to only the transport/network headers; the data in the packet body itself is not monitored. As noted in Section 3.3, this significantly reduces the required bandwidth. Even in the worst case (a continuous stream of the smallest packets) the throughput requirements are well within the specification of the host machine. The host machine must move the data (captured packets along with timestamps) from the capture card to intermediate storage. Along with the captured data, the host machine also logs metadata related to the health of the capture hardware, the host machine and so on.

In our architecture a SAN (System Area Network) is employed that allows tight coupling to the capture-card host. The SAN permits low-overhead/high-performance sharing of manipulated files and is based upon the Global File System (GFS) cluster filesystem (see [20] or [21]) over the ATA-over-Ethernet interface [22]. The storage disks of the SAN provide access to the captured data (and associated meta-data) through ancillary machines. The use of a SAN provides coarse grained control of priorities which in-turn allows the capture system writing new data to always have priority writing new data to the SAN over any unduly heavy data-read operation. The current (over)-specification of hardware can accommodate the 10Gb/s stream, further, with the use of the intermediate disk, the system has significant local storage capacity allowing buffering of captured data if the down-stream nodes, (capture access-node) require rebooting or have become CPU-bound in tasks such as the anonymization of headers. Like any SAN, there is no reason why multiple access-nodes cannot read data from the SAN storage if required; this may prove particularly useful if intermediate process tasks such as the anonymization of headers (Section 2.2.1) required multiple machines.

Captured data and log data formats consist of regular data-files with a strict, pre-agreed naming convention incorporating the time of capture. The capture system employs a fixed upper size, however, the capture system also has a maximum period of time to wait before capture and log files are rotated (closed, renamed and re-opened). In this way a steady upper and lower-bounded stream of information can be guaranteed to be made available from the capture system to the database back-end.

2.2 Database Architecture

An overview of the physical architecture is shown in Figure 3. This shows the systems currently in place at UCL and (using dotted lines) those planned additions later in the project. In the current deployment the webserver and database are on the same physical machine.

Once the capture system has finished writing a trace file and its associated metadata to the archive, the IP addresses are anonymised (see Section 2.2.1). The accompanying metadata contains information including the time window covered by the trace file, the monitoring point and several basic statistics such as number of packets and bytes captured. In addition to the per trace file metadata, probe configuration and monitoring point information is provided out-of-band with the packet capture process, in the form of an XML file. This includes information about the hardware and software used, which link is being monitored and the bandwidth of the link. Both the trace file metadata and capture system information are inserted into a PostgreSQL database [23]. This database is suitably indexed to allow trace files to be found and simple statistics to be derived. In conjunction with the database importing system is an archive disk management system, which handles removal of expired trace files (although some metadata for the removed files is maintained).

External users can search though and access the trace file archive via a web-based interface (written in Python TurboGears [24]). Before accessing the archive, the user must first register and accept the terms and conditions of use (see Section 3.2). Only registered users may download the trace files. Once the user has registered they are issued a unique username and password for accessing the web-based interface. Within the interface users can search for files by link, probe, time or other combinations of the metadata. The resulting trace files can then be downloaded by end users. In addition to searching for and downloading, metadata visualisations can also be created (such as graphs of throughput for a particular link and time period). Further preprocessing and visualisation capabilities are planned (see Section 4).

2.2.1 Anonymisation

A network-researcher may ideally wish to access a high-fidelity network trace where the payloads of the data indicate clearly the activity of the users and IP addresses easily identify end-points in the real world. However, implications of the legal constraints control the data accessible: requiring payload data be removed and, in the United Kingdom, the end-users not be identifiable. These needs lead to the anonymization process. Payloads are stripped at the capture system. The removal of payloads improves the performance of the capture system; Section 3.3 illustrates the significant difference in the raw data-rate of captured data that discarding payloads can provide. When engineering a capture system it is thus advantageous to discard payloads at the capture point reducing the quantity of data to be managed within the capture architecture.

Aside from the removal of payloads, the industry standard Crypto-Pan [25] is employed to provision a prefix-preserved, anonymised IP address. Preserving address prefixes maintains the structure of the IP address allowing for studies of routing and identifying groups of end-systems but removes information permitting the specific identification of a user, thereby satisfying the legal constraints. Users are required to sign an acceptable use policy forbidding attempts to reverse engineer the anonymisation before downloading the data (see Section 3.2).

3 Practical Implications

The results of a project such as MASTS are varied and not limited to purely measurements. In reality, as the project has had to interact with real operators on real networks, results include documentation of these interactions. This section details the issues we have had operationally and legally. Some initial results follow.

3.1 Operational Issues

As with any monitoring and measurement project a significant problem is issues arising when working on real networks and with their operators. As those responsible for the running of the network, operators need to ensure that the user service is always supported. In this section the common practical problems in network monitoring are examined and the solutions for the MASTS project enunciated.

Availability – The primary purpose of a network is to provide a connectivity service, and thus the primary purpose of the operator is to ensure that the connectivity remains. A common monitoring method is to insert an optical splitter into the fibre to take a copy of the traffic. Such an operation has two consequences: firstly that the fibre will need to be broken, with subsequent loss of service; and secondly a fear that the reduction in signal level will affect traffic. It is quickly apparent that *at-risk* maintenance periods need to be scheduled for such installation and testing — this requires a close and ongoing relationship with the network-operations staff.

Standards – Although a number of common standards exist for interoperability between different suppliers, it is most common for a network to be constructed from a single supplier's equipment. This usually allows for the use of specific non-standard, proprietary extensions and this caused problems in getting data from parts of the JANET Lightpath network as discussed in Section 2.1. To overcome this a novel hardware solution was necessary and the difficulties of obtaining, installing and configuring cutting-edge monitoring hardware (and consequent delays to the project) should never be under-estimated.

Operator Cooperation – Placing a new card into an operational switch requires a number of considerations. When the researchers are not operators of the network to be monitored the problem becomes far more complex than simply purchasing monitoring hardware and plugging it into a rack. Often purchase, installation and configuration will need the active cooperation of the network operator and this can lead to delays at each stage. One solution to be considered for future projects is the *embedding* of a project member within the network operator to act as a liaison.

Data Storage Requirements – The project proposed monitoring several bidirectional 10Gb/s links. Obviously this produces an enormous amount of data. The issues involved with storing this amount of data are discussed in Section 3.3.

3.2 Legal Framework

In Europe there are two relevant legal frameworks – limitations on interception of communications (wiretapping), and data protection legislation. Each member country of the EU is obliged to transpose EU Directives into their own national law, to provide a consistent legal framework, although there may be further national obligations to meet as well. In the UK, interception must be done in such a manner as to be made legal by the statutory exemptions in the Regulation of Investigatory Powers Act 2000 (RIPA) [26], with the equivalent EU level provision being enacted two years later as Article 5 of the Privacy and Electronic Communications Directive (2002/58/EC) [27]. The EU Data Protection Directive (95/46/EC) dates from 1995 [28] and was transposed into UK law as the Data Protection Act 1998 [29]. This latter legislation imposes strict rules on data processing and storage whenever it is possible to identify individuals within the data.

In order to manage the legal status of the monitoring activities, the MASTS project recognises 5 different categories of user or organisation for its monitoring operations:

1. The Network Operator (JANET(UK) in this case).
2. The organisation holding the monitored data (UCL for this work).
3. Other MASTS project members using the data (researchers at Cambridge and Loughborough).
4. External users using packet level data.
5. External users using summary data.

It was necessary to establish different agreements for each of the above groups due to the differing legal nature of the relationships. Agreement A is signed between 1 and 2 and covers their relationship. Users in 3 are covered by A as well by the mechanism specific in B; Agreement C covers users in category 4; finally, users in category 5 are not covered by an explicit legal agreement because they only have access to summary data.

A. A legal agreement between the Network Operator and UCL as the site holding the data. The resulting document available at `http://www.mastsproject.org/legal.html` establishes a practical example of a monitoring agreement between a UK operator and a UK University group. The agreement defines what data may be collected; what uses it may be put to; how privacy of the data originators is to be protected and that any machines storing data must be protected to the standard of best practice for their operating environments. The detailed text of these issues has been based on a framework document previously generated for this purpose by JANET(UK). The document is available at `http://www.ja.net/documents/development/legal-and-regulatory/regulated-activities/traffic-data-for-research.doc`. The legal aspects of this agreement were made considerably simpler by the aims of the MASTS project to record only protocol information from the Transport Layer and below. As such, no Application Layer data are stored and hence no information produced directly by a user (such as email text) is collected. Privacy is however still potentially compromised by the presence of the Network Layer (*i.e.* IP) Address. The agreement therefore requires such addresses to be anonymised in such a way that end user privacy is maintained. Technical solutions to this issue are discussed in Section 2.2.1.

B. A legal agreement which allows MASTS users at other institutions to be registered as visitors to the UCL network. This allows the cover provided by agreement A to extend to them if they are named in this agreement.

C. A legal agreement between the site holding the data and non-MASTS users. This is similar to the other agreements and ensures that data are not disclosed to third parties or used for purposes other than those agreed. In addition, users must agree to acknowledge the source of the data in any work published and not attempt to reverse the anonymisation.

The project has also established a dissemination approach which does not require data users to sign a legal agreement. This approach makes available summary data in which individual packet-level data is not available. For example, total data rate, or the number of different source IP addresses (but not the anonymised IP addresses themselves) within a given period of time can be provided world wide via the web interface without the establishment of a formal agreement. The only requirement made to a user is that of acknowledging the source of the data and using this for approved purposes only.

3.3 Data Available

Figure 4 shows the path of the data and the various transformations which occur between the monitoring point and the database. The initial traffic streams are expected to have a maximum rate of 10Gb/s. The traffic is split as described in Section 2.1 and only the headers retained. The data is anonymised as described in Section 2.2.1 (because of limite rack space and processing power, this anonymisation cannot be performed on site). Metadata about the capture process and extracted summary data is placed in a searchable database as described in Section 2.2. The data is provided in the Extensible Record Format (ERF) which has high timestamp fidelity and includes loss information [30]. Tools are provided to convert the data to the *pcap* format [31].

Obviously with such a high data arrival rate the data store would fill quickly. Tests have been performed on several large trace files to estimate this. The data sets considered here include a 500GB data set covering 24 hours from the site-connection of a medium size research institute and some typical data sets (each approximately 10GB) collected in 2002 and downloaded from the CAIDA website. In the first data set, stripping headers reduced the data to 14% of its original volume. Compression techniques on the headers (gzip [32] and lzo [33] were both tried in `--best` and `--fast` modes) reduced the headers further to between 4.5% and 6.1% of the original volume depending on the technique used. Using standard parameters, taking netflow style summary data without sampling reduced the data to 1.2% of the original volume and taking 1/512 packet samples reduced that data to 0.0071% of the original volume.

Table 1 shows how quickly various summary methods would fill 10TB of storage which represents the amount of storage that this project could reasonable devote to storing a single type of data from one monitoring point. The table shows the full data, the headers only, the headers compressed using gzip (the differences between the various compression algorithms tried were quite small), Netflow data without sampling and netflow data using 1/512 packet sampling. The figures are based on the assumption that, on average, the data arrives in the system at 10% of the maximum system capacity (that is, the data is arriving at a mean rate of 1Gb/s rather than the maximum rate 10Gb/s) – naturally if this assumption is changed the results would be scaled appropriately. The figures are given to only a single figure of accuracy and are based upon the results of the previous paragraph. It is obvious that for all but the most extremely compressed data storage formats storage can only be for a limited time period. Those extremely compressed formats, however, carry much less information. For example, one of the options is to store the number of bytes of data seen in every millisecond interval as a time-series for the lifetime of the project. However, the research value of this data is much less than the research value of full header data.

Data Format	Max rate	Mean rate	Time to collect 10TB of data
Full data	10Gb/s	1Gb/s	1 day
Headers	1Gb/s	100Mb/s	1 week
Comp. headers	500Mb/s	50Mb/s	2 weeks
Full netflow	100Mb/s	10Mb/s	3 months
1/512 netflow	700Kb/s	70Kb/s	30 years

Table 1: Types of data which might be stored with approximate data rates and estimated time to fill 10TB of storage.

The final solution which is used for the MASTS project is to have several levels of data kept. Extremely summarised data (for example bytes seen in a given time unit) can be stored for the lifetime of the project. Complete header information is stored for a short period for those researchers who wish to look at the current day of traces or who might want to examine traces to investigate a particular special event which has recently occurred on the network. A small repository of complete header files is kept for a longer time period. This repository will be useful for researchers who want representative traces to test data analysis schemes or hypotheses about, for example, creation of synthetic traffic traces. Finally, representative metadata (such as sampled netflow) may be stored for a longer time period, which will be determined by the amount of storage space taken up by that data.

4 Conclusions and Future Work

While the MASTS project does not finish until 2009, considerable progress has been made. Obviously, the laws applying to such data collection vary considerably across jurisdictions, the legal framework given here would be directly useful to those considering monitoring in the UK and could be a model to adapt for those in other countries. This legal framework is an important outcome of the project which could be useful to other monitoring researchers.

The difficulty of a monitoring project of this type should not be underestimated. The experiences described in this paper should provide a useful guide for those considering attempting such a project. As described in Section 3.1, there are several concerns which may cause problems and delay in monitoring projects. In specific, monitoring equipment must be deployed with minimum harm to network availability, equipment may use protocols which differ from those established and delays in scheduling the installation of equipment can cause difficulties. The insertion of a passive (optical) tap into a network link requires a small period of downtime during which the network link is not available. While such a downtime may be planned and, with alternative routing, scheduled-around it will require specific actions on the part of the operations staff. Commonly, alternative pathways are not used but the downtime is scheduled into a well-advertised at-risk period.

Data sets are available from the project website: `http://www.mastsproject.org/`. These data sets are a valuable resource for networking research. The ability to monitor recent traces from the JANET network will allow researchers to save data sets of particular value and when network events of interest occur. The utility of a monitoring project is best judged by the research it stimulates and it is hoped that the data provided here will be of considerable use both in the understanding it will bring and in the new research opportunities it will provide.

The MASTS project has provided a combination of both legal and engineering tools, as well as encouraging the operational relationships to ease future monitoring, particularly at the large scale. It is clear that the monitoring systems in place within MASTS may be easily extended to cover larger aspects both of the JANET interconnect to the internet and across the regions of the JANET infrastructure. There is no reason to be limited to the JANET networks and with the great interconnection diversity in the UK, provided by many broadband providers and peering locations such as LINX [34] (the London Internet Exchange), this will lead to a rich and diverse set of monitoring opportunities.

As an extension to the basic search functions and visualisation of the meta-data, more flexible preprocessing and advanced visualisation [2] of the data will be developed. The extensions to the data processing will partly be based on the idea of storing intermediate information [35] and also incorporate ideas from other network data processing work *e.g.* [36]. In addition, caching of downloaded trace files may be incorporated as part of the web server to minimise the load on the archive.

Thanks

The JANET Lightpath network was funded by JISC and HEFCE. MASTS was funded under the EPSRC grants GR/T10503/01, GR/T10510/03, and GR/T10527/01.

We would also like to thank Andrew Cormack who contributed invaluable expertise in the preparation of the legal documents and David Miller for contributing his technical expertise.

References

[1] MASTS Consortium, “MASTS website.” `http://www.mastsproject.org/`.

[2] M. Withall, I. Phillips, and D. Parish, “Network visualization: a review,” *IET Communications*, vol. 1, pp. 365–372, June 2007.

[3] K. Kyriakopoulos and D. Parish, “A live system for wavelet compression of high speed computer network measurements,” in *Proceedings of Passive and Active Measurement PAM 2007*, pp. 241–244, April 2007.

[4] BT, “Delivering the future, bt’s 21st century network.” `http://www.btplc.com/21CN/`.

[5] “TCPDUMP website.” `http://www.tcpdump.org/`.

[6] “Wireshark website.” `http://www.wireshark.org/`.

[7] V. Paxson, “Strategies for sound internet measurement,” in *IMC ’04: Proceedings of the 4th ACM/SIGCOMM conference on internet measurement*, (New York, NY, USA), pp. 263–271, ACM, 2004.

[8] K. Claffy, “Measuring the internet,” *IEEE Internet Computing*, vol. 4, no. 1, pp. 73–75, 2000.

[9] X. Zhou and P. V. Mieghem, “Hopcount and E2E delay: IPv6 versus IPv4,” in *Proceedings of Passive and Active Measurement PAM 2005*, pp. 345–348, March–April 2005.

[10] “RIPE website.” `http://www.ripe.net/ttm/`.

[11] D. A. Patterson, D. D. Clark, A. Karlin, J. Kurose, E. D. Lazowska, D. Liddle, D. McAuley, V. Paxson, S. Savage, and E. W. Zegura, *Looking over the fence at networks: A neighbor’s view of networking research.* Washington, DC: Computer Science and Telecommunications Board, National Academy of Sciences, 2001.

[12] G. Fox and D. Walker, "e-science gap analysis." Technical report, Indiana University, June 2003. `http://grids.ucs.indiana.edu/ptliupages/publications/GapAnalysis30June03v2.pdf`.

[13] M. Murray and K. Claffy, "Measuring the immeasurable: Global internet measurement infrastructure," in *Proceedings of Passive and Active Measurement PAM 2001*, p. 9, April 2001.

[14] C. D. Keeling, "Rewards and penalties of monitoring the earth," *Annual Review of Energy and the Environment*, vol. 23, pp. 25–82, 1998.

[15] "NLANR website." `http://pma.nlanr.net/`.

[16] "CRAWDAD website." `http://crawdad.cs.dartmouth.edu/`.

[17] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson, "On the self-similar nature of Ethernet traffic (extended version)," *IEEE/ACM Trans. Netw.*, vol. 2, pp. 1–15, 1994.

[18] CISCO, "Catalyst switched port analyzer (SPAN) configuration example." `http://www.cisco.com/warp/public/473/41.html`.

[19] J. Zhang and A. Moore, "Traffic trace artifacts due to monitoring via port mirroring," in *Workshop on End-to-End Monitoring Techniques and Services, 2007. E2EMON '07*, (Munich, Germany), pp. 1–8, May 2007.

[20] "GFS project page." `http://sources.redhat.com/cluster/gfs/`.

[21] S. R. Soltis, T. M. Ruwart, and M. T. O'Keefe, "The Global File System," in *Fifth NASA Goddard Space Flight Center Conference on Mass Storage Systems and Technologies*, (College Park, Maryland), Sept. 1996.

[22] E. L. Cashin, "Kernel korner: ATA over Ethernet: putting hard drives on the LAN," *Linux J.*, vol. 2005, no. 134, pp. 24–31, 2005.

[23] "Postgresql website." `http://www.postgresql.org/`.

[24] "Turbogears website." `http://www.turbogears.org/`.

[25] J. Fan, J. Xu, M. H. Ammar, and S. B. Moon, "Prefix-preserving IP address anonymization: measurement-based security evaluation and a new cryptography-based scheme," *Comput. Networks*, vol. 46, no. 2, pp. 253–272, 2004.

[26] Her Majesty's Stationery Office, "Regulation of investigatory powers act. chapter 23." http://www.opsi.gov.uk/acts/acts2000/ukpga_20000023_en_1, 2000.

[27] European Union, "Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications). Official Journal L 201, pp. 0037–0047." http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32002L0058:EN:HTML, July 2002.

[28] European Union, "Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data. Official Journal L 281, pp. 0031–0050." http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:EN:HTML, November 1995.

[29] Her Majesty's Stationery Office, "Data Protection Act. Chapter 29." http://www.opsi.gov.uk/Acts/Acts1998/ukpga_19980029_en_1, 1998.

[30] S. Donnelly, *High precision timing in passive measurements of data networks*. PhD thesis, University of Waikato, June 2002.

[31] "WAND website." http://research.wand.net.nz/software/libtrace.php.

[32] "gzip website." http://www.gzip.org.

[33] "LZO website." http://www.oberhumer.com/opensource/lzo/.

[34] "LINX website." https://www.linx.net/.

[35] O. Bashir, I. Phillips, D. Parish, J. Adams, and T. Spencer, "The management and processing of network performance information," *BT Technology Journal*, vol. 16, pp. 203–212, October 1998.

[36] M. Fisk and G. Varghese, "Agile and scalable analysis of network events," in *Proceedings of the SIGCOMM Internet Measurement Workshop*, ACM, November 2002.

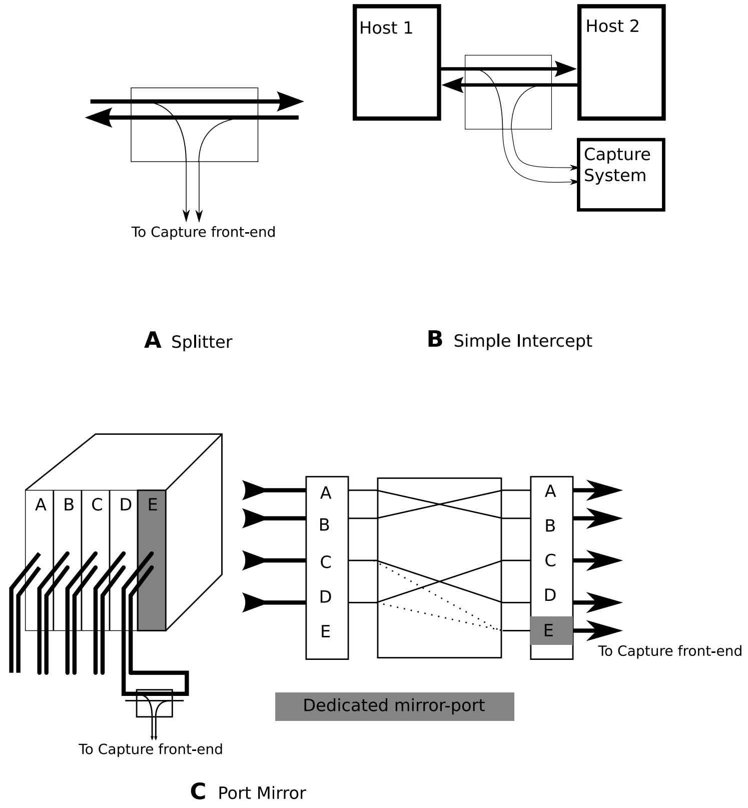

Figure 1: Intercepting data for capture.

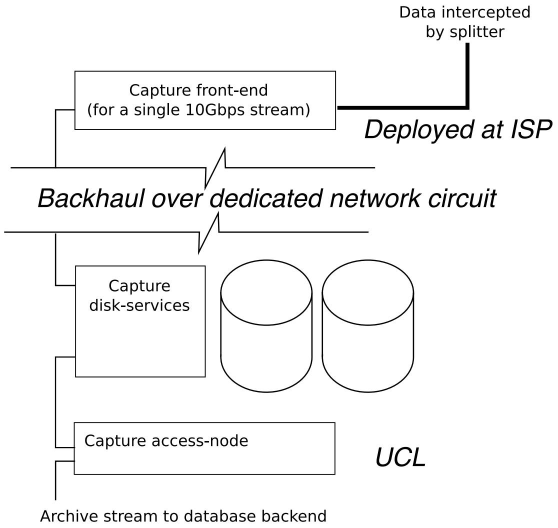
Data intercepted
by splitter
Capture front-end
(for a single 10Gbps stream)
Deployed at ISP
Backhaul over dedicated network circuit
Capture
disk-services
Capture access-node
UCL
Archive stream to database backend

Figure 2: The structure of the 10Gb/s monitor elements used within the MASTS project.

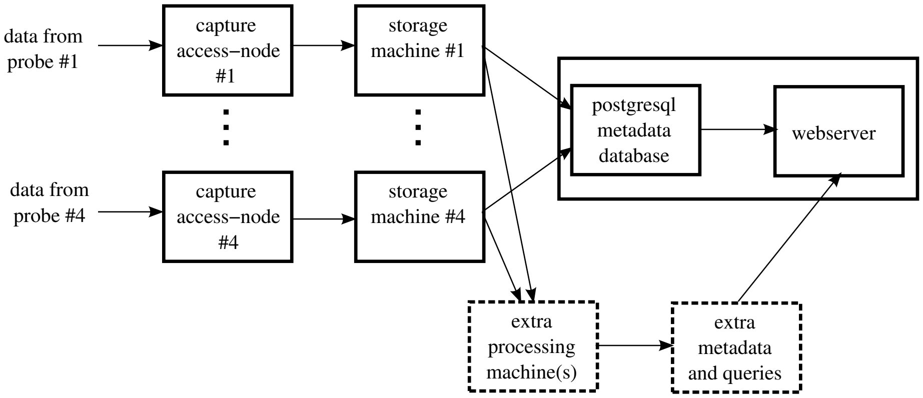

Figure 3: The machines which are deployed to collect and analyse the data.

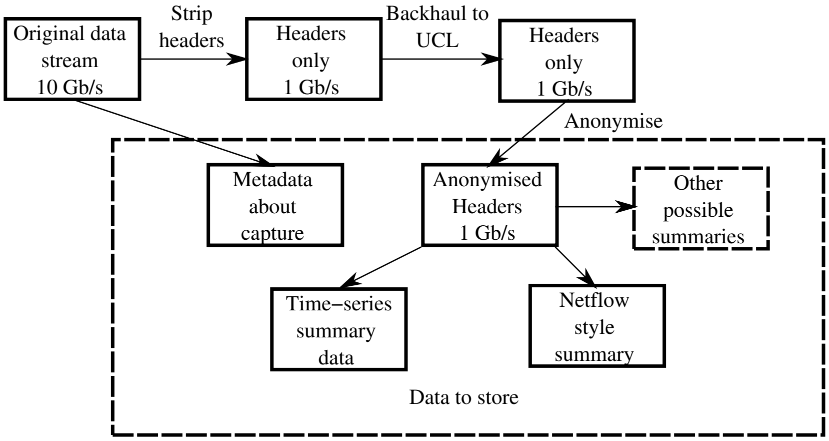

Figure 4: The path of the data from monitoring point to database.