

Report from the Clean Slate Network Research post-SIGCOMM 2006 Workshop

Jon Crowcroft
University of Cambridge
15 JJ Thomson Avenue
Cambridge CB3 0FD, UK
jon.crowcroft@cl.cam.ac.uk

Peter Key
Microsoft Research Cambridge
7 JJ Thomson Avenue
Cambridge CB3 0FB, UK
peterkey@microsoft.com

ABSTRACT

Europe has often followed in the footsteps of US research, but here we are trying to lead in Clean Slate networking research, rather than Cleans late networking. This is a report from a recent workshop on this topic.

Categories and Subject Descriptors

A [.]: 2—*Reference*, C.1.1[Single Datastream Architectures], C.2.1 [Packet-switching networks], C.2.4 [Distributed applications], D1.3 [Distributed Programming], D4.1 [Scheduling], D4.4 [Network Communication], D.4.8 [Stochastic Analysis], E.1 [Data Structures], E.2 [Hash Tables], E.4 [Data Compaction and compression], G.1.6 [Constrained Optimization], H.2.8 [Spatial Databases]

General Terms

General Terms: Algorithms, Performance, Design

Keywords

Data Communications, Review

1. INTRODUCTION

It cannot have escaped many peoples' notice that research in systems is burgeoning in Europe now. Partly this is because of the negative effect from US funding failure, but also it is from positive trends in Europe in long-term research funding and thinking, as well as inevitable political impact from different climates in Europe and the US (e.g. attitudes to foreign researchers, and simple ease of obtaining visas and work permits).

In the last 12 months, not only have ACM's flagship conferences SOSP, and SIGCOMM, but also IEEE Infocom and ACM Mobihoc *and* Mobisys, have all been held within Europe. After the SOSP conference in Brighton, England, in autumn of 2005, we held a small Systems Colloquium in Cambridge, taking advantage of the many people from overseas who were nearby at the time. For similar reasons, we called a small (invitation only, purely for reasons of size) workshop on Clean Slate Network Research, in Cambridge, right after SIGCOMM in Pisa, Italy. (Of course, going from Brighton to Cambridge is perhaps more of a bonus trip, whereas going from Pisa to Cambridge, during semester might be seen as somewhat of a penalty, though of course, it certainly selects for people that are really keen to discuss hot topics even more than they already had done in the 5 marvellous days of SIGCOMM in Pisa already!).

We decided to hold the meeting in Cambridge because of the local mix and support from Intel Research Cambridge and Microsoft Research Cambridge, as well as the ease of low cost flights from Europe to nearby Stansted Airport (a mistake, as it turned out, due to the lamentable under-provisioning of immigration and security staff in UK airports for the current workloads).

Rather than specify talks, we invited people to think outside the box, although in the event, the talks did roughly group into the areas: architecture; routing; measurement and wireless.

The goal was to have a modest sized event where attendees can present work-in-progress in a friendly non-competitive context. Some of you may recall a similar event last year after the European SOSP.

Jon Crowcroft called the meeting, and doesn't believe in "top-down driven innovation" (an oxymoron, if ever there was one), so instead asked for a collection of bottom-up talks and discussions, stemming from people's work and ideas, to stimulate imagination and cross-fertilisation. Any resemblance to a structure like FIND/GENI (see <http://find.isi.edu/> and <http://www.geni.net>) or FP7 (see <http://cordis.europa.eu/fp7/>) is, therefore, purely coincidence.

There were several novel areas discussed, especially in role based architecture, metarouting, privacy preserving network monitoring, the split between software and hardware support for network and router virtualization, and vehicular and other novel wireless network applications.

The talks and these notes are linked from: <http://www.cl.cam.ac.uk/Research/SRG/netos/cleanslate.html>

2. ARCHITECTURES

Mark Handley talked about his work on Role Based Architectures [4], and the role of protocol heaps, and especially, of supporting multiple control planes. There was an extensive discussion on binding, and on performance, using some worked examples of control plane activities (admission control, NAT traversal, etc), as well as a discussion on trying to introduce expressiveness constraints (remembering the success-disaster that was SIP, since the Session Description Protocol language was so overly expressive). There was also discussion on the lessons that might be imported from older (Telecom/GSM/cellular) architectures.

We also talked about the lessons from cellular nets where roaming/mobile/handover and QoS work well and scalably, and yet so badly for mobile IP. Why not wholesale adopt these techniques in next generation data nets? [16]

There was a brief discussion on Denial-of-service proof ar-

architectures, and whether there were lessons from telecommunications networks there although largely these were ones of innovation versus openness, and micro-call accounting versus lower transaction-entry-cost communications systems for commercial providers - the baby/bathwater argument was mentioned¹.

Kevin Fall started from the radical design space of DTNs[9] and talked about novel name spaces, the problems of provenance, and tags. He also briefly discussed his routing work on stretch-1 algorithms and how well they fit the problems in Manet, DTN, and mobile IP. There was further discussion on declarative networking.

Christophe Diot gave a brief outline of the initiatives in Europe including the Arcadia Cost action (see <http://www.cost.esf.org/>). He noted that the EU can fund non US things (outside EU), and can fund infrastructure before research agenda (extreme e.g. underwater net testbed!) - see also <http://cordis.europa.eu/fp7/> for more EU funding information.

Nick Feamster extended the talk on Vini[2] given by Andy Bavier at SIGCOMM, to talk about technology for enabling comparative architecture bake-off research. The question was asked whether the platforms described would let us replace the “sub-IP” layer - i.e. is the Vini initiative mainly about routing research? What about circuit versus packet, MPLS, lambda versus an IP tunnel? Emulab lets you model that: does Vini? The conclusion of the discussion was that Vini probably did not address sub-IP without a lot more work, and it has much the same limitation for wireless (except maybe edge) really, mainly for *packet switched* network research.

3. ROUTING

Tim Griffin presented a sneak preview of what to expect from a (soon to be released) prototype implementation of metarouting[10]. The tool takes high level specifications of routing protocols and network and outputs a set of (steady state) routing tables. It represents a “first step” that allows users to learn about metarouting and experiment with protocol specifications. Future steps will involve integrating this with real routing codes such as XORP (see <http://www.xorp.org/>). Discussion centred on how to build systems now in practice, and what is the hard bit (state machine for policy?).

We also discussed the Intel/Berkeley ideas on declarative routing, and whether metarouting could be implemented on this or vice versa or both.

Steve Uhlig outlined the challenges for interdomain routing research

- Measurements are not enough to improve our knowledge of what matters in interdomain routing behaviour. [22]
- Playing with models of the Internet might give some significant insight into what matters for interdomain routing behaviour. This started in their 2006 SIGCOMM paper[18].

¹It is conjectured that the flat rate system of billing in the Internet has contributed to innovation, and that introducing a more complex usage based accounting system might decrease misuse, but might also impact the rate of innovation. Thus one might throw out the dirty bathwater of DDoS, but lose the Baby of creativity in doing so.

- Leverage (how to do it is still open, pure data analysis for sure not, coupling data with model probably better) all the available BGP data we have, might still improve this poor knowledge of routing dynamics.

4. MONITORING AND MEASUREMENT

Matthew Roughan gave one of the most thought-provoking talks of the day about how to design routing and network systems that prevent privacy invasions, *by design*[20][19].

There was an extensive discussion also on privacy preserving data mining&trust, and on how to carry these ideas over to the wireless domain (e.g. geo routing). [1].

This was followed by a talk by Lukas Kencl, on content anonymization[15] (<http://www.ieee-infocom.org/2006/>) (i.e. avoiding the “AOL debacle”). A key benefit was in the change to the log-file ownership model. The tradeoff in privacy and accuracy of analysis was key here.

Patrick Crowley talked about the Washington University work on diversified networking[21] in their open network laboratory[8], with emphasis on hardware support for network Virtualization and Planetlab. A key question is how to do isolation right. [7]

Richard Mortier discussed the work on Anemone[6], which is a mix of all end-systems monitoring with access to the network devices (e.g. for routing tables). A key problem here is (e.g. bot-nets) colluding liars. Mort also talked about their Seaweed[17] system, which is a network event distribution system supporting the end-system monitoring, to illustrate how the connections behind various projects could evolve into different areas: Seaweed is uses ideas from databases, which have shown up in sensor network research (e.g. sketches[5]) as well.

5. WIRELESS/SENSORS

Dina Papagiannaki talked about Intel work on “energy-star” power saving mode for wireless APs. While the absolute amount of power in wireless LAN APs isn’t a whole lot, the percentage that can be saved is significant - in developing areas of the world, this might be crucial. The (decentralised) algorithms may well carry over to fixed networks where (esp. in co-lo data center environments) we are rapidly approaching a brick wall in terms of ability to dissipate enough heat, let alone costs![11]

David Hadaller from Waterloo presented work that he has been doing with Keshav on fairness v. throughput in vehicular drive-by net access scenarios. He showed conclusively that he can build system with more capacity, but less fairness, but still more capacity for “poorer” users as well as more-more for “richer”: there was a nice use of a Lorenz curve to illustrate wealth gap/variance. Again, the generality of the argument (re-visiting fairness versus capacity) is interesting in other contexts. [12]

Alastair Beresford also talked about automotive network applications, looking at parasitic sensor nets on car for pollution monitoring, with uses in transport efficiency, automatic map updates from road usage pattern. He also has relevant work on privacy preserving monitoring.[3]

Milan Vojnovic re-visited the source of the heavy tail/power law in mobility models (c.f. Huggle project)[13], and He also alluded to Jon’s own conjecture about the heavy tail of contact time and duration being merely the result of overlaying individuals’ mobility from a set of cyclic trips with a small

set of different frequencies and phases. Milan showed that in fact, the power law isn't at odds with current mobility models (e.g. the Manhattan street grid mobile model), and also described many such in nature. This is important as there may be a horizon-effect on the distribution of inter-contact time and contact durations, which might mean that feasible (finite mean delivery delay) stateless DTN routing algorithms can be found.

Wenjun Hu talked about the tussle space between research and products, in the hardware versus software for wireless network research, partly based on the experience in the COPE (core algorithm in XORS in the Air paper) and other work, emphasising the difficulty of doing novel work within current constraints of commodity hardware (i.e. optimising for current hardware is suboptimal). [14]

6. TAKE HOME MESSAGES

Jon Crowcroft summarised the talks by trying to capture a cognitive model of networks where adjacency is probabilistic, and there are fuzzy clouds perhaps not always with silver linings!

Virtualization hardware/software split - not just like page table, device, processor user/supervisor mode, but also, amount of address recognition hardware, network processors state, VPN state/identifier space/processing, queue/buffer isolation, etc, etc

Routing we are about ready to try out metarouting and we need to think hard about deploying it - some of the same platforms (Xen/Xorp) might serve well. There may be some very nice computer science to be done between declarative networking and metarouting.

Monitoring architect for privacy, at routing, logging and database/storage stages.

Wireless policy v. mechanisms in scheduling MAC can still be surprising. We are still at a very early stage on having good models of mobility. Commodity hardware can be inflexible in terms of trying to do more radical radio level research.

Other power management is interesting in many, many places.

Discussion continued over dinner at Jesus College Cambridge, sponsored by Intel Research.

6.1 Attendees

- Jon Crowcroft, University of Cambridge
- Peter Key, Microsoft Research Cambridge
- Kevin Fall, Intel Research Berkeley
- Nick Feamster, Georgia Tech
- Mark Handley, UCL
- Graham Holmes, Cisco
- Tim Griffin, Cambridge
- Alex Gurney, Cambridge
- Steve Uhlig, UCL

- Dina Papagiannaki, Intel Research Cambridge
- Wenjun Hu, University of Cambridge
- David Hadaller, Waterloo
- Milan Vojnovic, Microsoft Research Cambridge
- Alastair Beresford, University of Cambridge
- Matthew Roughan, Adelaide
- Patrick Crowley, Washington University,
- Christophe Diot, Thomson Research Paris
- Christos Gkantsidis, Microsoft Research Cambridge
- Thomas Karagiannis, Riverside
- Richard Mortier, Microsoft Research Cambridge
- Lukas Kencl Intel Research Cambridge
- Andrew Moore, QMUL

7. REFERENCES

- [1] AGRAWAL, R., AND SRIKANT, R. Privacy-preserving data mining. *ACM SIGMOD Record* 29, 2 (2000), 439–450.
- [2] BAVIER, A., FEAMSTER, N., HUANG, M., PETERSON, L., AND REXFORD, J. In VINI Veritas: Realistic and Controlled Network Experimentation. *Proceedings of ACM SIGCOMM 2006* (2006).
- [3] BERESFORD, A., AND STAJANO, F. Location privacy in pervasive computing. *Pervasive Computing, IEEE* 2, 1 (2003), 46–55.
- [4] BRADEN, R., FABER, T., AND HANDLEY, M. From protocol stack to protocol heap: role-based architecture. *ACM SIGCOMM Computer Communication Review* 33, 1 (2003), 17–22.
- [5] CONSIDINE, J., LI, F., KOLLIOS, G., AND BYERS, J. Approximate aggregation techniques for sensor databases. *IEEE ICDE* (2004), 449–460.
- [6] COOKE, E., MORTIER, R., DONNELLY, A., BARHAM, P., AND ISAACS, R. Reclaiming network-wide visibility using ubiquitous end system monitors. In *Proceedings of USENIX 2006 Annual Technical Conference* (June 2006).
- [7] CROWLEY, P., FLUCZYNSKI, M., BAER, J., AND BERSHAD, B. Characterizing processor architectures for programmable network interfaces. *Proceedings of the 14th international conference on Supercomputing* (2000), 54–65.
- [8] DEHART, J., KUHN, F., PARWATIKAR, TURNER, J., WISEMAN, C., AND WONG, K. The open network laboratory. In *Proceedings of the 37th SIGCSE Technical Symposium on Computer Science Education* (2006), 107–111.
- [9] FALL, K. A delay-tolerant network architecture for challenged internets. *Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications* (2003), 27–34.
- [10] GRIFFIN, T., AND SOBRINHO, J. Metarouting. *Proc. ACM SIGCOMM* (2005).

- [11] GUPTA, M., AND SINGH, S. Greening of the internet. *Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications* (2003), 19–26.
- [12] HADALLER, D., KESHAV, S., AND BRECHT, T. MV-MAX: Improving Wireless Infrastructure Access for Multi-Vehicular Communication. *Proceedings of ACM SIGCOMM CHANTS Workshop* (2006).
- [13] HUI, P., CHAINTREAU, A., SCOTT, J., GASS, R., CROWCROFT, J., AND DIOT, C. Pocket switched networks and human mobility in conference environments. *Applications, Technologies, Architectures, and Protocols for Computer Communication* (2005), 244–251.
- [14] KATTI, S., RAHUL, H., HU, W., KATABI, D., AND CROWCROFT, J. Network Coding Made Practical. *Proc ACM SIGCOMM 2006* (2006).
- [15] KENCL, L., LOEBL, M., AND ZAMORA, J. Packet content anonymization by hiding words. In *Demo at IEEE INFOCOM 2006, is on website* (2006).
- [16] LOO, B., CONDIE, T., GAROFALAKIS, M., GAY, D., HELLERSTEIN, J., MANIATIS, P., RAMAKRISHNAN, R., ROSCOE, T., AND STOICA, I. Declarative networking: language, execution and optimization. *Proceedings of the 2006 ACM SIGMOD international conference on Management of data* (2006), 97–108.
- [17] MORTIER, R., NARAYANAN, D., DONNELLY, A., AND ROWSTRON, A. Seaweed: Distributed scalable ad hoc querying. In *32nd Intl. Conference on Very Large Databases (VLDB 2006)* (2006).
- [18] MÜHLBAUER, W., FELDMANN, A., MAENNEL, O., ROUGHAN, M., , AND UHLIG, S. Building an as-topology model that captures route diversity. In *Proc. of ACM SIGCOMM 2006* (2006), ACM Press New York, NY, USA.
- [19] ROUGHAN, M., AND ZHANG, Y. Privacy-preserving performance measurements. In *ACM SIGCOMM MineNet-2006* (2006).
- [20] ROUGHAN, M., AND ZHANG, Y. Secure distributed data-mining and its application to large-scale network measurements. In *ACM SIGCOMM Computer Communication Review* (January 2006), vol. 36, pp. 7–14.
- [21] TURNER, J., AND TAYLOR, D. Diversifying the internet. *Proceedings of the Global Telecommunications Conference* (2005), 755–760.
- [22] UHLIG, S., BONAVENTURE, O., AND QUOITIN, B. Interdomain traffic engineering with minimal BGP configurations. *18th International Teletraffic Congress (ITC)* (2003).