The January 2017 Issue

This issue of Computer Communication Review marks an important change for our newsletter. As announced in the October 2016 issue, CCR now accepts longer papers provided that the authors release artefacts such as datasets or software that allow to replicate the main results of the paper. *Towards a Context-Aware Forwarding Plane in Named Data Networking Supporting QoS* is our first replicatable paper. D. Posch et al. propose and evaluate new techniques to improve Quality of Service in Information Centric Networking. More specifically, they investigate whether context information can be exploited by forwarding strategies to improve QoS. The proposed techniques are implemented in `ndnSim` and various simulation results are analysed. The authors provide all the simulation scripts and modifications to the simulator that are required to repeat the results. This has been checked by M. Tortelli who interacted with the authors to ensure that their software was working correctly and could be reused by others.

Our second technical paper, *A Database Approach to SDN Control Plane Design*, describes the lessons learned by B. Davie et al. when developing production-ready multi-vendor implementations of a network virtualization system. Multi-vendor implementations pose a specific challenge since different vendors need to agree on a common solution. In the networking industry, the classical way to solve such problems is to design a new protocol or extend an existing one and spend a lot of time in standardisation meetings. This paper used a completely different approach that turned out to be much better. After difficulties in extending Openflow to support their specific use case, they decided to apply database principles to solve their virtualization problem. Instead of designing a new protocol, they focused on modeling the information that had to be exchanged among the different components of the system. Once the information model was in place, they could leverage existing database synchronisation mechanisms to ensure the distribution of the required information. This contrasts with the traditional approach of putting in a single design both the information that needs to be exchanged and the communication mechanisms. They also explain the central role played by the open-source `OVSDB` software in ensuring interoperability. I encourage all protocol designers to read this paper in details and reread it again the next time they envisage to design a new protocol or extend an existing one.

Our third technical paper, *On the Potential Abuse of IGMP*, analyses security risks of the Internet Group Management Protocol (IGMP). This protocol was designed in the early days of IP Multicast and was intended to only be used by hosts to report their subscribed groups to their local routers. IGMP can also be used in unicast mode and some routers reply to unicast IGMP messages. This feature was initially designed to debug problems on distant routers. Unfortunately, when a router receives a unicast IGMP message, it may return a much larger IGMP response back, which opens the possibility of using IGMP to conduct amplification attacks. M. Sargent et al. analyse network scans to quantify the security risks posed by the 305k routers that reply to unicast IGMP messages. The reviewers discussed about the possibility of asking the authors to release their scanning software or the trace collected. They did not request these artefacts given that releasing them could have helped malicious exploitations of this vulnerability.

Our fourth technical paper, *Exploring Domain Name Based Features on the Effectiveness of DNS Caching*, is a measurement paper. The scalability of the Domain Name System (DNS) depends heavily on the util-

isation of caches in both resolvers and end-hosts. Caching assumes that a name will be reused in the future. However, there are domain names that are never or rarely reused. The paper analyses such domain names in more details and studies the impact of their usage on DNS caches. S. Hao and H. Wang have publicly released one of the DNS traces analysed in the paper.

Besides these four technical papers, this issue also contains three editorials. The first one, *Toward a Taxonomy and Attacker Model for Secure Routing Protocols*, is the summary of a recent Dagsthul seminar that focused on secure routing protocols. The second one, *Can We Make a Cake and Eat it Too? A Discussion of ICN Security and Privacy*, summarises the results of another Dagsthul seminar that focused on Information Centric Networking and Security. Our last editorial, *Workshop on Tracking Quality of Experience in the Internet: Summary and Outcomes* summarises a workshop sponsored by the NSF and the FCC on QoE.

Reproducible research

In early December 2016, we conducted an informal survey among the authors of papers published in CCR and the SIGCOMM sponsored conferences in 2016. We sent a short email to the authors of all accepted papers. 77 authors replied to our survey, which is a good subset of our community. Most of the responding authors published their paper at IMC (35.9%), Conext (23.1%), Hotnets (10.3%), SIGCOMM (7.7%) and CCR (16.3%). The vast majority of the responders were students or researchers working in university labs. Only a small portion of the responders indicated the presence of authors working for industry.

The first survey question was whether the authors had taken actions to ensure the repeatability of their published paper. 34% of the responders indicated that their paper was self-contained and did not require any software or dataset to repeat its results. 4% of the responders have released a longer technical report. Considering software, 42% of the responders have released the software used to perform the experiments described in the paper but only 25% have released the experimental data. 30% of the responders have created a website to distribute artefacts associated with the published paper. Creating a web site is usually a good idea because it allows to distribute information that can be easily updated. However, the persistence of such web sites is not always guaranteed and they may disappear after a few months or years.

Looking more specifically at software, more than 70% of the responders indicate that they have developed specific software for their paper. This software can range from small analysis scripts to complete protocol implementations. Slightly less than 50% of this software was available in open-source at publication time. 25% of the responders planned to release their software within a year after publication. 16% of the responders did not plan to release their software. The main reasons for not releasing software was that it was a prototype that was not ready for public use (18 responders). Only 6 responders indicated that commercial reasons blocked the software release.

Open-source software is important in our community. Indeed, 50 responders used open-source software to carry their research. A wide variety of open-source software was listed from compilers, network monitoring tools, simulators or protocol implementations. The next question evaluated how this usage of the open-source software was referenced in the paper. Surprisingly, the most popular method to reference the utilisation of open-source software is to simply list their names in the paper. Only 26% of the responders have referenced the main software that they used in their bibliography. We should probably require authors to correctly cite the software packages that they use in their bibliography so that the authors of these packages receive the credits that they desserve. Given the open-source nature of most software packages, one could expect that the authors who have modified open-

source packages have released their modifications to those packages. Unfortunately, less than 20% of the responders have released their modifications.

The last set of questions focussed on **experimental data**. Such data is key, notably in the measurement community. If data has been collected correctly, it can often be used for different papers. Unfortunately, the public release of experimental data is not the default in our community. Only 11 have publicly released their dataset. 10 responders replied that their dataset was available upon request, but this limited availability can block some utilisations of the collected data. 9 responders planned to release the data within a year after publication and 16 responders did not plan to release their experimental data. 40% the responders who released experimental data indicated they the data was released on a personal homepage or through a cloud provider such as `github` or `dropbox`.

The last question of the survey was an open question on whether our community should encourage the reproduction of networking papers. The responders were in favor of encouraging this and provided many interesting suggestions. Here are some representative ones. The entire survey is available as a supplement to this note on the Digital Library.

- *Force researchers to release their code*
- *Add visibility to papers releasing their code, and perhaps add a dedicated award for quality software. Have a public commenting system.*
- *Encourage PCs to look for reproducibility in evaluating papers submitted to conferences*
- *Make it easier to actually publish replication papers. They tend to get rejected due to 'lack of novelty'... But really, if it's not replicated it might as well not exist...*
- *Reproduction should always be the default. There should be a strong justification for accepting non-reproducible papers; Papers that are not reproducibles should be marked as so; Best paper award only to papers that are reproducible; being reproducible should be an important tie breaker when accepting papers.*
- *Yes, this should absolutely be encouraged. The main reason why we didn't do anything of this was "laziness" (or the fact that it wasn't required for publication of the paper).*
- *There exists an increasingly mature toolchain for replicating machine setups (puppt, vagrant, docker, etc), which could help readying infrastructure for repeating/replicating/reproducing research.*
- *We need to specifically encourage papers that attempt to reproduce or refute earlier results. Perhaps conferences could have a pre-allocated session or workshop for such papers and announce this as part of the CFP. In our community, we have a culture where novelty is prized above all else. In the past I have been on the receiving end of some harsh reviews because I dared to publish a paper that repeated someone else's experiment on a different dataset to see if the original findings held in other contexts. I will not attempt anything similar again unless I have a very good reason to believe that the paper will not be needlessly savaged by a review process that is overly oriented towards novelty rather than scientific value.*

This is clearly not the last word on the reproducibility of the research results published in our community. This discussion will continue at the **Reproducibility'17** workshop during the SIGCOMM'17 week.

Olivier Bonaventure
CCR Editor