Results from Running an Experiment as a Service Platform for Mobile Networks

Miguel Peón-Quirós
IMDEA Networks, Spain
mikepeon@imdea.org

Vincenzo Mancuso
IMDEA Networks, Spain
vincenzo.mancuso@imdea.org

Vincenzo Comite
La Sapienza University, Italy
vincenzo.comite@gmail.com

Andra Lutu
Simula Research Lab, Norway
andra@simula.no

Özgü Alay
Simula Research Lab, Norway
ozgu@simula.no

Stefan Alfredsson
Karlstad University, Sweden
stefan.alfredsson@kau.se

Jonas Karlsson
Karlstad University, Sweden
jonas.karlsson@kau.se

Anna Brunström
Karlstad University, Sweden
anna.brunstrom@kau.se

Marco Mellia
Politecnico di Torino, Italy
marco.mellia@polito.it

Ali Safari Khatouni
Politecnico di Torino, Italy
ali.safari@polito.it

Thomas Hirsch
Celerway, Norway
thomas.hirsch@celerway.com

ABSTRACT

In this paper we present a selection from a vast range of experiments run with MONROE, our experiment as a service platform open to external researchers. We show that the platform can be used to benchmark network performance in a repeatable and controlled manner thanks to the collection of a rich set of geotagged metadata and the execution of discretionary user experiments. Indeed, with the sheer amount of data collected from 12 commercial mobile operators across Europe, MONROE offers an unprecedented opportunity to monitor, analyze and ultimately improve the status of current and future mobile broadband networks. We also show how flexibly the platform allows combining metadata and experimental data series during the experiments or by means of post-processing, and show results produced by our own experiments as well as by external research groups and developers that have been granted access to the platform.

1 INTRODUCTION

The field of networking offers the possibility of gathering large volumes of information from network elements and end hosts. Analyzing these data is crucial to understand how networks perform under different usage patterns and adapt them to future requirements. This is particularly important for mobile broadband networks (MBBs), which are the segment with the strongest growth forecast and higher variability in operating conditions. Two main challenges arise when trying to analyze the performance and reliability of MBBs: The difficulty of obtaining systematic data from reliable repetition of experiments on commercial operational MBB networks, and sifting through the big amount of variables that can be monitored and measured.

MONROE[1] is a Europewide experiment oriented network of around 150 custom measurement devices (or *nodes*), designed to enable collection and analysis of the characteristics of commercial mobile broadband networks and execution of discretionary experiments from external researchers. The platform nodes operate under a wide variety of conditions (e.g., aboard trains, buses and delivery trucks, or inside residential homes and laboratories). In the final design, nodes are co-located in pairs, where one node connects to two mobile providers using customer-grade commercial subscriptions, and the other connects to a third operator and potentially to a WiFi network. Both nodes can connect to Ethernet where available.

The testbed performs periodic passive and active measurements and continuously monitors the status of the MBB networks through metadata collection. The collected metadata are centrally stored in a NoSQL database to ensure scalability past billions of records. We offer to the community the unique possibility of accessing our curated dataset through periodic data dumps, which enable data analysis across all the nodes and lifespan of the platform. Additionally, we encourage external experimenters to devise novel experiments and add to the diversity of MONROE open data.

The following is a list of the main characteristics and innovations of MONROE.

Large-scale deployment in diversified scenarios: MONROE nodes are being deployed across Norway, Sweden, Italy and Spain, with external partners deploying additional nodes in Greece, Portugal, Slovenia and the UK. Some nodes have stationary locations in dense urban areas, while a significant number (114 at the time of writing) operate aboard public inter-city trains, buses and delivery

trucks. Whereas trains traverse large distances, sometimes at high speeds, buses cover urban areas – both settings enable us to collect a unique dataset under mobility scenarios along the fix routes of those vehicles. Nodes aboard delivery trucks, which traverse both urban and rural areas without fixed routes, complement the previous dataset.

Open experimentation platform on commercial cellular operators: MONROE is an open platform that allows authenticated researchers to run their own custom experiments on commercial MBB networks. Researchers can then opt to add their data to the MONROE open dataset, increasing its diversity and allowing us to look past performance metrics and metadata. Notable examples are a web performance experiment and video QoE measurements [23], which are being evaluated for inclusion in the set of periodic measurements run on the nodes. In addition to the actual data, experiment source code[2] and supporting material[3] for those wanting to create new experiments on MONROE are also openly available.

Consistency and repeatability: MONROE provides a uniform hardware and software environment to measure and monitor MBB networks at fixed locations and times. Furthermore, the public transportation vehicles that host MONROE nodes ensure fairly repeatable routes for mobility experiments. Even more, they repeat the same itineraries several times a day at different hours (i.e., mixing peak and normal hours) and on different days (i.e., weekdays and weekends). This provides the dataset rich spatio-temporal dimension, which is key to enable the comparison of different measurements over different operators, places and times of day.

Metadata-rich dataset: Each MONROE node is instrumented to periodically measure the performance of its MBB providers. They continuously gather metadata, including, for example, location, signal strength and link technology for each network provider. Additionally, ICMP ping and HTTP download experiments are executed periodically to asses network performance. Since MONROE does not involve real users (which usually entail privacy protection restrictions), rich metadata collection, including geo-temporal tagging, is possible, which enables the evaluation of mobile services under mobility. In particular, MONROE collection of data enables purely off-line experiments for analysis of MBB network performance.

In the rest of this paper, we first review other approaches that aim to study mobile networks (Section 2). Then, we briefly describe the design of the MONROE platform (Section 3). Finally, we present a selection of experiments from ourselves and from several external groups that showcase the possibilities that MONROE opens (Section 4). Those examples aim to entice other researchers to exploit the data gathered by our platform in innovative ways or to design their own experiments and so contribute to improve our overall knowledge on the behavior of MBB networks.

2 RELATED WORK

Due to growing interest by regulators, policy makers and networking community, several nationwide efforts to measure the performance of home and mobile broadband networks (e.g., the US FCC's
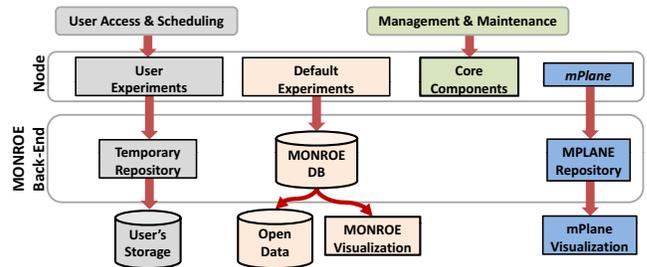

Figure 1: MONROE system design. Researchers access the system through the WEB user interface and scheduler, or directly through the various repositories and data bases. Several passive (metadata, mPlane) and active (RTT, HTTP bandwidth) probes monitor continuously network usage and performance.

Measuring Broadband America initiative [7]) have been initiated. MONROE goes beyond proposing a trans-national platform.

In contrast with operator-driven measurement campaigns [10, 11, 18], or existing small-case drive-by tests [20], MONROE offers open access to cross-operator collected data, including device-level metadata, which is key to interpret measurement results, across a wide variety of locations.

Moreover, there have been several crowdsourcing projects devoted to measure MBBs using tools such as Mobiperf,[4] Netalyzer [12] and Haystack [22]. Such projects allow crawling through mobile network performance factors to identify the causes of experienced performance figures. In general, such approaches lack rich metadata due to the privacy concerns created by the involvement of real users, hindering the analysis of their datasets. Also, reliance on users can provide high coverage, but at the cost of repeatability regarding location, route or equipment. However, in combination with a platform like MONROE, they could be used in a more systematic and controllable way [8].

3 PLATFORM DESIGN AND IMPLEMENTATION

The MONROE platform was designed with the purpose of collecting, storing and offering open access to large amounts of diverse mobile network data, and providing an EaaS platform for the execution of discretionary experiments by external researchers. Therefore, enriching measurement data with abundant context information (metadata), and enabling a wide variety of experiments, are the two key aspects that have steered the platform design since its inception. Figure 1 offers a high-level overview of the complete MONROE platform design.

An introduction to the platform was presented in [2], while its detailed description and the experience of operating it are presented in [3]. Therefore, here we briefly present the platform components and focus on the processes of collection and storage of measurement results and the concrete implementation choices made during the platform design.

---

[2]https://github.com/MONROE-PROJECT/Experiments
[3]https://github.com/MONROE-PROJECT/UserManual

[4]https://sites.google.com/site/mobiperfdev/

MONROE nodes execute a series of default experiments including, for example, periodic ping measurements for connectivity survey, HTTP downloads from a series of targets under our control, or web performance measurements. The results of these experiments and the collected metadata are transferred as JSON files to the main MONROE server via rsync over SSH channels. Once at the server, the JSON files are stored in a NoSQL database. In addition to the default experiments, each node runs an mPlane [5] compatible passive traffic analysis tool called Tstat [9]. Tstat generates a series of logs that the nodes send to the mPlane repository, from where users can consume the data using the mPlane visualization solution. Tstat data is also stored in the MONROE DB.

Off-line data analysis can happen both at the server side in the form of database queries or at the experimenter's side (with custom applications) if further processing is required. Since data are the main asset of the platform, we implement several backup and duplication mechanisms to provide data safety and access redundancy.

MONROE allows authenticated external researchers to access the platform via the WEB user interface and deploy their own custom experiments. Separate storage for the results of user experiments is offered in a temporary repository accessible through the platform web user interface. We encourage users to make their results public and include them in the MONROE open dataset. Finally, a visualization solution facilitates the surveillance of the platform health and its available resources in near real-time.

3.1 Node instrumentation

MONROE nodes collect four types of information:

1. Metadata: Network parameters (RSSI, cell identifiers, link technology, etc.), node location and speed (GPS), node working parameters (CPU temperature, processing load, etc.) and node events (watchdogs).

2. Connectivity and latency measurements: RTT measurements for ICMP (UDP and TCP as future extensions).

3. MONROE and user experiments: Bandwidth estimation, traceroute results to popular targets, and the results of user experiments (additionally, we are in the process of integrating web performance results and video quality of experience in the MONROE open data).

4. Traffic monitoring as captured by Tstat.

The differentiation between the aforementioned types of data responds to their differing natures and purposes. In that way, passive metadata can be gathered at the nodes with minimal impact on any experiments; thus, they are recorded on a continuous basis. Similarly, the passive mPlane Tstat probe, which produces low processing load, runs continuously. Background experiments such as RTT measurement may create a moderate (controlled) interference with other experiments; however, the value obtained by gathering these data are worth their cost. Experimenters are made aware of those background experiments; furthermore, we are evaluating a mechanism to allow them to pause their execution. Finally, some MONROE experiments such as bandwidth measurements might produce a higher impact on user experiments. Therefore, those experiments are not scheduled concurrently with any user ones: Each
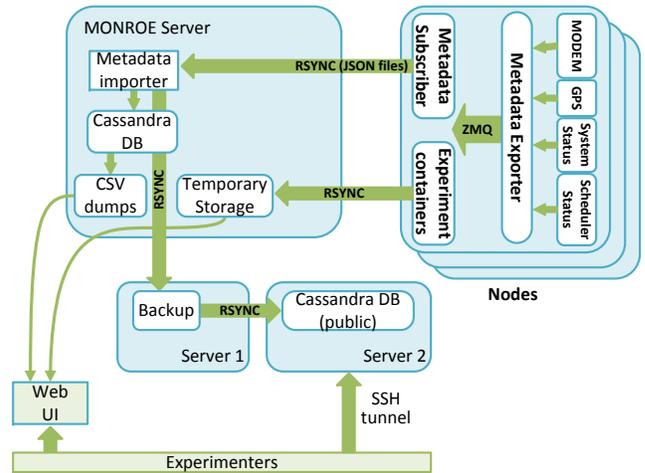

Figure 2: Flow of information in the MONROE platform.

user experiment runs in exclusivity with respect to experiments from any other users.

3.2 Data flows

Figure 2 shows the different flows of information through the platform, since it is generated in a node until it is collected and stored in our databases for later analysis. MONROE nodes implement a metadata distribution mechanism based on a publish/subscribe model. Experiments running in the nodes can subscribe to different information "topics" to monitor system status and events such as network interface (dis)connections, link technology changes or GPS location variations. This flexible design eases the implementation of each platform component as data producers do not need to keep track of their clients, and new data consumers can choose the information topics they are interested on without caring about the details of the producers.

Independently of their origin, all data items are transferred to the MONROE servers via rsync over SSH. Once at the server, each item is processed and stored according to its nature: Metadata, the results of the MONROE experiments and Tstat measurements, which arrived as JSON files, are stored in a NoSQL database, whereas the results of user experiments are temporarily kept at a repository for easy access through a web user interface.

In the server, several scripts create backups of the database contents and a dump of the database in CSV format is produced daily; experimenters may use those 24 h feeds if their experiments are focused on small periods of time. Furthermore, a secondary copy of the database is updated every day for direct access by external researchers. That secondary copy is not a normal database "replica" to avoid the risk that accidental (or malicious) modifications to the (open) database spread to the primary one. The daily CSV dumps are available for direct download to registered users through the web user interface; access to the (secondary) database is provided to external researchers via SSH tunnels.

3.3 At the node side

At the node side, metadata distribution is implemented in a publish/subscribe pattern using ZeroMQ.[5] The metadata stream is available for experiments during their execution using the ZMQ subscription mechanisms. Metadata entries are generated in a single-line JSON format, which eases human analysis. Every data entry is labeled with a "topic" field; consumers may subscribe to the whole stream of metadata or just to some topics. The metadata subscriber module runs in the nodes and subscribes to all the topics, writing JSON entries to files in a special file system location. A synchronization process transfers those files at certain intervals (tens of seconds) to the MONROE server.

Regarding node stability, several monitoring and recovery methods ensure that they remain online and capable of executing experiments. Node stability is ensured via lightweight virtualization (Docker containers), thus guaranteeing a clean environment for each experiment. Several surveillance mechanisms (watchdogs) in the nodes can force a complete reinstallation of the operating system and environments if they detect system malfunctions such as filesystem corruption.

3.4 At the server side

Information received from the nodes in JSON format is stored at the server in a NoSQL database. The choice of a NoSQL solution was based on the need to permanently store a potentially very large dataset consisting of billions of entries. As a quick calculation to illustrate the scale of the dataset, RTT measurements are executed for each of the three MBB interfaces of each node every second. Therefore, $3 \times 3\,600\,\mathrm{s} \times 24\,\mathrm{h} \times 365\,\mathrm{days} \times 150\,\mathrm{nodes} = 14\,191 \times 10^6$ entries are stored in the database every year, only for RTT measurements. Based on the concrete storage and access needs of MONROE, Apache Cassandra[6] was chosen as the system NoSQL database for its scaling abilities, both in performance and storage capacity: If the space available in a machine is exhausted, new space may be added simply by configuring a new replica. Additionally, Cassandra is a mature technology that offers access drivers for multiple programming languages and production-grade tools for data analytics, widening access options for researchers.

Several Python scripts produce a backup of the JSON files received at the server and a daily CSV dump of the database. Those results are transferred to a backup server that provides off-site backups. The secondary copy of the database that is accessible to external researchers is also hosted in an independent server, thus avoiding performance interferences with the main database.

3.5 Access to data

The metadata produced by the nodes can be accessed in several ways. First, experiments may access the metadata stream during execution using the ZMQ subscription mechanisms. In this way, they can monitor and react to events such as interface reconnections or link technology and signal strength changes for each MBB at run-time. Second, researchers may access the database (or the CSV dumps) to correlate their results with the metadata matching by the corresponding timestamps. As an example, the results of an
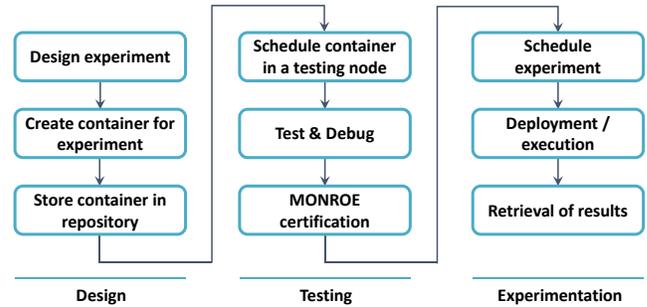
Figure 3: Experiment workflow covering the design, test and experimentation phases.

experiment may be related to the network conditions during its execution, even if at that time not all the metadata was checked online. Researchers may also import the CSV dumps into their own tools for more specific data analyses. Finally, several open access options such as Zenodo are under evaluation for permanent storage beyond the lifetime of the MONROE consortium.

3.6 Experimentation workflow

Figure 3 shows the general workflow of the experiments executed on MONROE nodes. The first step is to design the experiment selecting the appropriate tools. The required files have to be collected in a Docker container, which is submitted to a repository. MONROE offers a set of dedicated nodes for testing that can execute containers from any public repository. Once the experiment is ready, it undergoes a certification process in which MONROE administrators check that it is generally safe for execution and move the container to a private repository. Deployed (i.e., real experimentation nodes) can download containers only from the MONROE private repository. Container execution can be scheduled as many times and on as many nodes as required, always subject to quota availability. Using the platform web interface, users can monitor the progress of all their experiments, including repetitions on multiple nodes. Finally, the results can be downloaded directly from the platform web page.

4 EXPERIMENTS

In this section we present several of the most interesting experiments that have already been conducted on MONROE or using its dataset, which, at the time of this writing, contained more than 975 M metadata entries, 1882 M RTT and 107 K bandwidth measurements, 38 M Tstat entries and more than 18 K experimenter results.

As an example of our own work, in [3] we show preliminary results from our experiments on web page load time (PLT) and proxy identification over mobile broadband networks. There, we use a headless browser to fetch two popular websites from 37 nodes operating in four countries and using 11 operators. As an example, we observe large variations of PLT for the same website between Sweden and Norway. In that work we also report results on identification of Performance Enhancement Proxies (PEP) in MBBs.
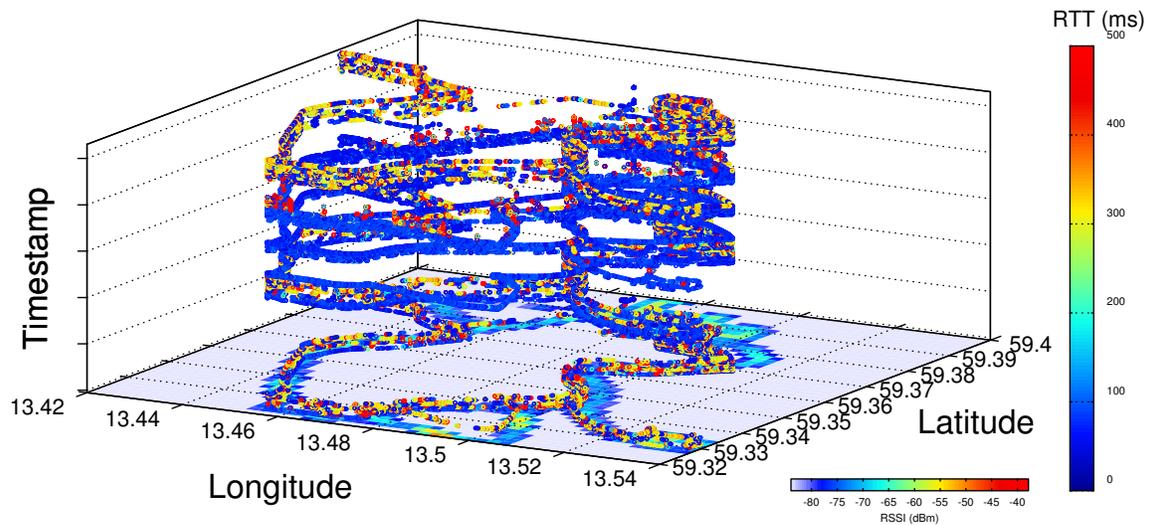
Figure 4: RTT vs. RSSI measured in a bus at Karlstad from Aug. 5th, 2016 12:59:20 to Aug. 11th, 2016 17:51:53 (GMT+2). Average RSSI values are shown on the XY plane. Individual RTT measures are plotted on the Z-axis using their relative timestamps as height to visualize successive laps.
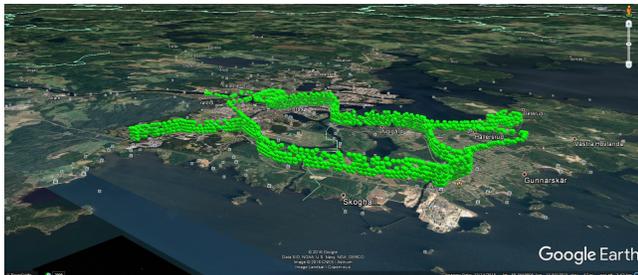

Figure 5: This representation of link technology for the bus at Karlstad reveals that 4G coverage is consistently available for the complete route during the analyzed period.

Also as a result of our own experiments, in [16] we present our experience estimating the download speed offered by actual 3G/4G networks. For that experiment, we analyzed data from 50 nodes in 4 countries over 11 operators during more than two months. The conclusion of that study is that measuring the performance of MBB networks is quite complex as different network configurations such as the presence of NATs or Performance Enhancing Proxies (PEP), which do vary over time, have a significant impact on measurements.

In general, MONROE enables two broad types of experiments: Passive experiments that analyze the curated dataset, and active experiments that use the platform nodes to perform new measurements.

4.1 Passive experiments

Passive experiments use the dataset collected by MONROE, which contains metadata, the results of the default experiments and the results of experiments shared by their owners with the broader community. They can perform queries directly on our NoSQL database or process the CSV files that are generated daily (e.g., for more complex analyses on smaller amounts of data). Those experiments can use the whole range of MONROE data, since the moment it started to collect information, and for all the nodes in all the countries, and can be repeated at any point in time.

Passive metadata/QoS analysis to build coverage and latency maps. MONROE deployment in public transportation vehicles enables evaluation of MBBs on wide urban mobility environments. Route predictability provides high confidence, whereas measurements taken at similar positions on different hours allow comparing the behavior of the MBBs at different times (e.g., rush versus normal hours).

Figure 4 follows the typical route of a bus around Karlstad (Sweden), showing the measured RSSI (signal strength) and RTT (ICMP ping). The different laps along several days are represented vertically ascending to ease the visualization of the dense information obtained. Figure 5 shows the negotiated link technology for the same route. The analysis of the collected data (signal strength, link technology and measured delay) gives insights into the performance perceived by users during their bus trips. Such information might then be used by network operators to improve the service offered to users during commuting.

Based on the same dataset and on previous theory and observations that show that line-of-sight (LOS) fading follows a Rice distribution [15], while NoLOS fading follows a Rayleigh one, we are currently developing a method to infer which distribution yields a better fit for experimental data, potentially providing information to operators to optimize the location of their base stations.

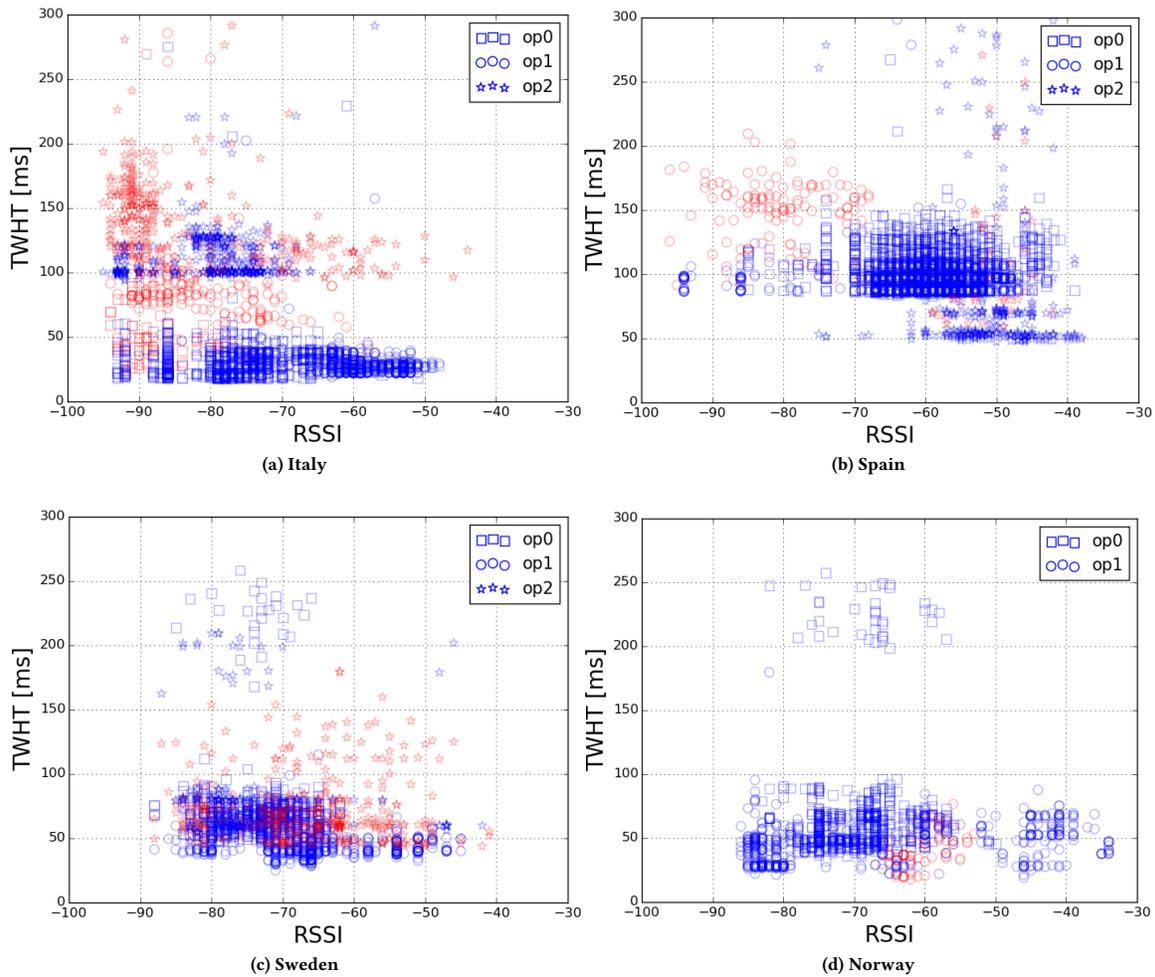*Passive traffic analysis and network monitoring with Tstat.*

Figure 6: TCP three-way handshake times (TWHT) obtained using the HTTP download experiment for bandwidth measurement versus the RSSI reported in MONROE metadata. Blue and red correspond to 4G and 3G, respectively.

Tstat [9] is a powerful passive monitoring tool that rebuilds TCP flows reporting more than 100 flow descriptors (e.g., client and server IP and port, RTT, number of retransmissions) and more than a thousand packet level metrics.[7] Experimenters can use Graphite[8] to easily navigate through the logs and store a dashboard showing relevant data within an adjustable time window.

As an example, Figure 6 shows a correlation between 3-way handshake time (TWHT), as measured by Tstat, and RSSI from the metadata, illustrating the many possibilities that MONROE creates for cross-domain data analysis.

4.2 Active experiments

Active experiments are executed on MONROE nodes under schedule. They are built as Docker containers and scheduled on real nodes through the platform WEB user interface. Those experiments can
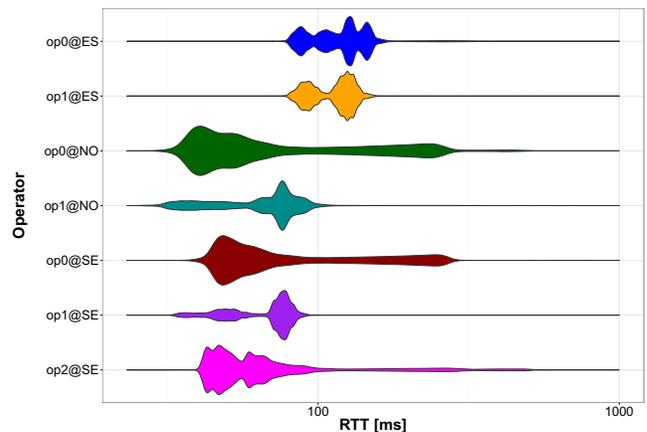

Figure 7: Violin plots of the RTT measurements for different operators in Spain (ES), Norway (NO) and Sweden (SE).

---

[7]http://tstat.polito.it/

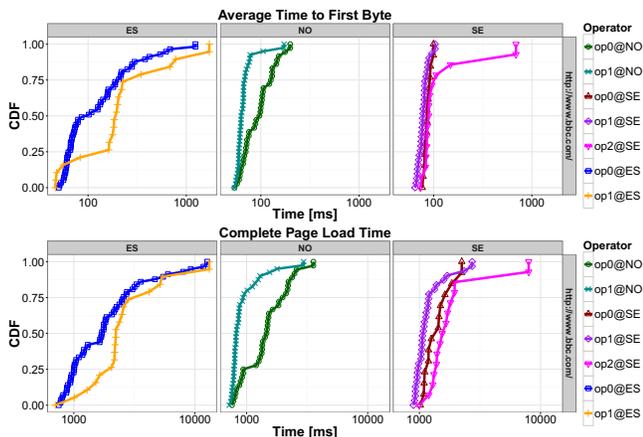[8]http://graphite.readthedocs.org/en/latest/index.html

Figure 8: Average Time to First Byte and Complete Page Load Time for operators in Spain (ES), Norway (NO) and Sweden (SE) for *www.bbc.com.*

consist of any software compatible with the container architecture and use all networking resources available in the nodes at the moment of execution, subject to user quotas. Experiments can be repeated as desired to verify the consistency of the results or to analyze changes on network behavior along time. The new data generated by active experiments may become part of the dataset available for passive experiments.

Cross operator/country performance. MONROE enables comparison of different operators (in terms of network characteristics and user-perceived application performance) in and among countries.

Figure 7 shows a violin plot for the RTT samples collected (using ICMP ping) during one week (from July, 8th until July, 15th, 2016) with 30 stationary nodes for 7 different operators in 3 countries. Each "violin" shows the probability density of the RTT at different values; the higher the area, the higher the probability of observing a measurement in that range. Nodes in Norway and Sweden exhibit lower delays than nodes in Spain because they are closer to the target measurement server, which is hosted in the MONROE backend in Sweden. Interestingly, measurement variance is much higher than in fixed networks, showing that MBBs introduce complexity even for such basic tests as RTT monitoring. For example, RTT measurements exhibit typically a multimodal distribution that corresponds to the different access delays faced by different radio access technologies (e.g., 3G vs. 4G). MONROE's repetitive measurements enable correlation with time, location and context conditions such as variations in signal strength.

In Figure 8, we present the CDFs of the complete page load time and average time to first byte for *www.bbc.com* broken down per country. Interestingly, for the Spanish operators we detected multiple DNS iterations, which partially account for their higher TTFB values.

4.3 Other experiments

Apart from the previously described experiments, MONROE is in the process of expanding its range of supported measurements and

thus enrich the dataset it collects and offers to the community. In the past months, MONROE incorporated, among others, traceroute measurements, path transparency measurements with *PATHspider* [14], web performance measurements using a headless browser and video QoE measurements with YoMoApp [23]. The MONROE traceroute experiment aims to compare routing from nodes in different countries and, inside a country, different operators (some of our measurements are performed with SIMs in roaming that show home-routing patterns). Traceroute experiments also allow identifying middleboxes and their differences between operators and countries. Building on this idea, *PATHspider* [14] is a tool developed for A/B testing of path transparency. It allows testing the feasibility of deploying new protocols in the Internet and quantifying the impact of path impairments. The headless browser experiment measures web performance against multiple popular targets, enabling, for example, the tracking of page load time and its correlation with metadata information. Similarly, YoMoApp (YouTube Performance Monitoring Application) allows analyzing mobile network performance with respect to YouTube traffic, and developing optimization solutions and QoE models for mobile HTTP adaptive streaming.

4.4 Experiments from external projects

In addition to the MONROE experiments presented before, the platform is currently used for research and experimentation by 25 external groups from academy and industry who run both passive and active experiments. In the following paragraphs, we describe the most relevant or promising results obtained by those groups.

Software radio extensions. The SOPHIA project has developed an extension to enhance MONROE nodes with software radio capabilities. In [19], its members present detailed performance measurements of LTE networks to illustrate the potential benefits and new possible measurements.

Forecasting LTE cell congestion. In [21], the authors try to forecast the average downlink throughput for LTE cells using data collected from multiple probes and to apply that knowledge to Self-Organizing Network (SON) strategies to shift coverage and capacity according to predicted demand. This group updated some MONROE nodes to address the benchmarking of voice calls, showing the flexibility of the platform nodes.

Available Bandwidth measurement on SDN deployments. In [1], the authors employ MONROE as a testbed to study the complexity of available bandwidth (ABw) estimation using SDN-based active measurements. They conduct their experiments using one node in each of the four main countries of the project. Their ongoing work tries to improve the accuracy and reliability of existing tools, using the MONROE testbed to isolate and better understand different aspects of the measurement process.

Designing application performance with MBB analytics. The authors of [4] use the radio parameters measured by MONROE nodes to determine the best application protocol for a service, identifying the most suitable Key Performance Indicators (KPIs) to characterize the network state. These type of works are very relevant to close the gap between network performance measurements and user experience. Interestingly, the authors see an opportunity on the data generated by other experiments running in the platform (and

made openly available by the respective researchers) as a means to obtain additional data points for their own investigation.

Surveying DSCP modifications in mobile networks. MONROE is used by a group of researchers in [6] to conduct a survey on path-level treatment of DiffServ packets in MBB networks and identify behaviors that potentially violate the IETF specifications. Differentiated Services (DiffServ) enable classification of traffic into Quality of Service classes via usage of the Differentiated Services Code Point (DSCP) field in the IP header. Using MONROE to analyze the behavior at the edge mobile network, they find that there is a high probability that the corresponding fields are remarked in the first two hops.

Path protocol transparency. The multiple middleboxes present between Internet users pose serious issues to deploy new protocol features. Understanding point-to-point connectivity for particular protocols is helpful to identify and solve those problems. In [13], the authors, in collaboration with part of the MONROE consortium, present the results of adapting PATHspider, an extensible framework for A/B testing of protocols, to the realm of commercial mobile networks using MONROE nodes deployed by themselves in the UK. Among their conclusions, the most relevant is that MBB networks provide a considerably different environment to the one provided by the cloud access points that PATHspider was using in the past.

Streaming video QoE measurements. In [17], MONROE is extended into the domain of QoE with the design and implementation of a measurement tool for YouTube video streaming. That work emulates and monitors at the network and application layers an end-user device requesting video streams. The experiments used MONROE nodes for three and a half weeks, measuring 165 streaming sessions. Among other factors relevant for HTTP adaptive video streaming, they successfully gathered data on stalling, initial delay and adaptation, and HTTP request and response information.

5 CONCLUSIONS

In this paper, we have described the unique dataset that MONROE collects, curates and makes available to the community. We have focused on how MONROE helps to improve the knowledge on the usage and behavior of current and future commercial mobile broadband networks. We have also explained the main design characteristics of the platform that make it unique and how, from the generation of data at the nodes to their storage in a NoSQL database that can scale past billions of records, MONROE offers the unprecedented possibility of data analysis across all the nodes and lifespan of the platform. To illustrate the potential and flexibility of the platform, we have presented results from our own experiments and from several other groups that have access to our platform.

ACKNOWLEDGMENTS

REFERENCES

[1] G. Aceto, V. Persico, A. Pescapé, and G. Ventre. 2017. SOMETIME: SOftware defined network-based Available Bandwidth MEasuremenT In MONROE. In *Proc. of the IEEE/IFIP Workshop on Mobile Network Measurement.*

[2] Ö. Alay, A. Lutu, R. García, M. Peón-Quirós, V. Mancuso, T. Hirsch, T. Dély, J. Werme, K. Evensen, A. Hansen, S. Alfredsson, J. Karlsson, A. Brunström, A. Safari Khatouni, M. Mellia, M. Ajmone Marsan, R. Monno, and H. Lønsethagen. 2016. MONROE, a Distributed Platform to Measure and Assess Mobile Broadband Networks: Demo. In *Proceedings of the Tenth ACM International Workshop on Wireless Network Testbeds, Experimental evaluation, and Charaterization (WiNTECH '16).* ACM, New York, NY, USA, 85–86.

[3] Ö. Alay, A. Lutu, M. Peón-Quirós, V. Mancuso, T. Hirsch, K. Evensen, A. Hansen, S. Alfredsson, J. Karlsson, A. Brunström, A. Safari Khatouni, M. Mellia, and M. Ajmone Marsan. 2017. Experience: An Open Platform for Experimentation with Commercial Mobile Broadband Networks. In *Proc. of ACM Mobicom.*

[4] I. Alepuz, J. Cabrejas, J.-F. Monserrat, A.-G. Perez, G. Pajares, and R. Gimenez. 2017. Use of Mobile Network Analytics for Application Performance Design. In *Proc. of the IEEE/IFIP Workshop on Mobile Network Measurement.*

[5] P. Casas, P. Fiadino, S. Wassermann, S. Traverso, A. D'Alconzo, E. Tego, F. Matera, and M. Mellia. 2016. Unveiling network and service performance degradation in the wild with mplane. *IEEE Communications Magazine* 54, 3 (March 2016), 71–79.

[6] A. Custura, A. Venne, and G. Fairhurst. 2017. Exploring DSCP modification pathologies in mobile edge networks. In *Proc. of the IEEE/IFIP Workshop on Mobile Network Measurement.*

[7] FCC. 2013. *2013 Measuring Broadband America February Report.* Technical Report. FCC's Office of Engineering and Technology and Consumer and Governmental Affairs Bureau.

[8] M.-R. Fida, A. Lutu, M. Marina, and Ö. Alay. 2017. ZipWeave: Towards Efficient and Reliable Measurement based Mobile Coverage Maps. *Proc. IEEE INFOCOM* (May 2017).

[9] A. Finamore, M. Mellia, M. Meo, M. M. Munafo, P. D. Torino, and D. Rossi. 2011. Experiences of Internet traffic monitoring with tstat. *IEEE Network* 25, 3 (May 2011), 8–14.

[10] E. Halepovic, J. Pang, and O. Spatscheck. 2012. Can you GET me now?: Estimating the time-to-first-byte of HTTP transactions with passive measurements.. In *Proc. of IMC.*

[11] J. Huang, F. Qian, Y. Guo, Y. Zhou, Q. Xu, Z.-M. Mao, S. Sen, and O. Spatscheck. 2013. An In-depth Study of LTE: Effect of Network Protocol and Application Behavior on Performance. In *Proc. of SIGCOMM.*

[12] C. Kreibich, N. Weaver, B. Nechaev, and V. Paxson. 2010. Netalyzr: Illuminating the edge network. In *Proc. of the 10th ACM SIGCOMM conference on Internet measurement.* 246–259.

[13] I.-R. Learmonth, A. Lutu, G. Fairhurst, D. Ros, and Ö. Alay. 2017. Path Transparency Measurements from the Mobile Edge with PATHspider. In *Proc. of the IEEE/IFIP Workshop on Mobile Network Measurement.*

[14] I.-R. Learmonth, B. Trammell, M. Kühlewind, and G. Fairhurst. 2016. PATHspider: A tool for active measurement of path transparency. In *First ACM/IRTF Applied Networking Research Workshop.* Berlin, Germany.

[15] M.A. Richards. 2006. *Rice Distribution for RCS.* Georgia Institute of Technology.

[16] A. Safari Khatouni, M. Mellia, M. Ajmone Marsan, S. Alfredsson, J. Karlsson, A. Brunström, Ö. Alay, C. Midoglu A. Lutu, and V. Mancuso. 2017. Speedtest-like Measurements in 3G/4G Networks: The MONROE Experience. In *Proc. of ITC29.*

[17] A. Schwind, M. Seufert, Ö. Alay, P. Casas, P. Tran-Gia, and F. Wamser. 2017. Concept and Implementation of Video QoE Measurements in a Mobile Broadband Testbed. In *Proc. of the IEEE/IFIP Workshop on Mobile Network Measurement.*

[18] M. Z. Shafiq, L. Ji, A. X. Liu, J. Pang, S. Venkataraman, and J. Wang. 2013. A first look at cellular network performance during crowded events. In *Proc. of SIGMETRICS.*

[19] P. Sutton and I. Gomez. 2017. MONROE-SOPHIA - A Software Radio Platform for Mobile Network Measurement. In *Proc. of the IEEE/IFIP Workshop on Mobile Network Measurement.*

[20] Tektronix. 2009. *Reduce Drive Test Costs and Increase Effectiveness of 3G Network Optimization.* Technical Report. Tektronix Comm.

[21] P. Torres, P. Marques, H. Marques, R. Dionísio, T. Alves, L. Pereira, and J. Ribeiro. 2017. Data Analytics for Forecasting Cell Congestion on LTE Networks. In *Proc. of the IEEE/IFIP Workshop on Mobile Network Measurement.*

[22] N. Vallina-Rodriguez. 2017. Illuminating the Third Party Mobile Ecosystem with the Lumen Privacy Monitor. In *FTC PrivacyCon 2017.*

[23] F. Wamser, M. Seufert, P. Casas, R. Irmer, P. Tran-Gia, and R. Schatz. 2015. YoMoApp: A Tool for Analyzing QoE of YouTube HTTP Adaptive Streaming in Mobile Networks. In *European Conference on Networks and Communications (EuCNC).*