

Report from the NSF 'Large-scale Networking Platforms "Communities of Practice"' Workshop

Workshop organizers:

Ivan Seskar & **Morley Mao**
(Rutgers University) (University of Michigan)

Workshop Dates: October 24 – 25, 2016

Report Publication Date: February 1, 2017

Acknowledgments

This report was compiled from workshop notes and draft review comments from: Prasad Calyam, Tom Mitchell, Ken Baker, Thyaga Nandagopal, Abhimanyu Gosain, Mark Berman, Joe Mambretti, James Gigrich, Stefano Basagni, and Vuk Marojevic.

TABLE OF CONTENTS

EXECUTIVE SUMMARY	3
1.1. SUMMARY RECOMMENDATIONS – COMMUNITY OF PRACTICE	4
1.2. SUMMARY RECOMMENDATIONS – PAWR SPECIFIC	4
2. EXPERIENCES WITH EXISTING TESTBEDS	4
2.1. PLANNING ISSUES	7
2.2. DEVELOPMENT ISSUES	8
2.3. USER AND COMMUNITY ISSUES	8
2.4. LEGAL ISSUES	9
2.5. SOFTWARE ISSUES	9
2.6. OPERATIONAL ISSUES	10
2.7. GUIDANCE FROM NSF LARGE FACILITIES	10
3. BREAKOUT SESSIONS	11
3.1. PLANNING AND ORGANIZATION	11
3.2. USABILITY	13
4. VISIONS FOR PAWR	16
5. CONCLUSIONS AND RECOMMENDATIONS	19
5.1. COMMUNITY OF PRACTICE RECOMMENDATIONS	19
5.2. PAWR RECOMMENDATIONS	21
APPENDIX A: WORKSHOP ATTENDEES	22

Executive Summary

The NSF 'Large-scale Networking Platforms "Communities of Practice"' workshop was held on October 24-25, 2016 in Washington DC. The main aim of the workshop was to bring together researchers from both academia and industry involved in (data) networking and community infrastructure projects to distill lessons learned from starting/building/operating/sustaining research-focused networking testbeds. Another aim was to use past experiences and see how they could be applied to the Platforms for Advanced Wireless Research (PAWR) initiative and on the establishment of 'communities of practice' in order to help the PAWR Project Office (PPO) with design, development, deployment and management of a set of advanced wireless research platforms/testbeds.

The workshop had 44 attendees including 28 from academia, 8 from industry and 8 representatives from non-profit, NSF and other government agencies. Two of the attendees were representatives of large European testbeds.

The workshop was structured to address two goals. The first goal, addressed primarily on day one, was seeding a community of practice by building on experiences of current large-scale platform developers and operators. The second goal, and the focus of day two, was obtaining community input applicable to the PAWR program. The morning of the first day was reserved for short presentations on past experiences in building and running individual testbeds. The afternoon program included two breakout sessions on: a) planning and organization and b) usability. The second day opened with the keynote presentation by Walter Johnston of the FCC who spoke on future challenges that the regulatory agency is facing in managing the spectrum. This was followed by the morning plenary session that was reserved for individual position statements and discussion on recommendations for PAWR. The workshop closed with the recommendation development session in the afternoon.

This report summarizes the discussions at the workshop. Key recommendations emerging from the workshop are summarized below and are also detailed in Section 5. Although the recommendations are roughly organized into two groups in accordance with the workshop's goals, both sets of recommendations are applicable to PAWR planning activities. The workshop website contains links to the agenda, the various presentations made during the workshop and this report. The URL for this website is: <http://www.winlab.rutgers.edu/events/tbcopws/Index.html>.

1.1. Summary Recommendations – Community of Practice

1. Adequately staff engineering, deployment and maintenance roles with dedicated professionals to assure successful platform operations.
2. Prioritize service needs and set expectations for each group of experimenters. Drive a rigorous set of priorities for each testbed.
3. Introduce a common data collection standard of data collected on the infrastructure, and data collected by the experimenters. While different they are both significantly beneficial.
4. Support real world usability with a mix of emerging technologies and current state-of-the-art by using open programmable platforms, and avoid vertical stack silo-only architectures.
5. Connect with other testbeds, compute grids, identity management systems and data analytics platforms, utilizing proven services from other instances without re-inventing the wheel.
6. Encourage testbeds to have a user reputation system that makes an appropriate set of testbed resources available to the right kind of users.

1.2. Summary Recommendations – PAWR specific

7. Primacy of research is key. A researcher-driven group should be identified to serve as counterpart to the recently formed advanced wireless industry consortium. The PAWR platform technology should be selected to emphasize variety over commonality. At least 50% of PAWR airtime should be used for (NSF sanctioned) publishable, non-proprietary research programs.
8. The Community needs help in connecting with city stakeholders.

2. Experiences with existing testbeds

The set of presentations on experiences with existing networking testbeds covered a large array of testbeds as shown in Table 1.

<p>StarLight International/National Communications Exchange Facility (http://www.startap.net/starlight/) Joe Mambretti</p>	<p>StarLight is a multi-layer service exchange facility for high-performance access to participating networks, and a true optical switching facility for wavelengths, advanced networking, database, visualization and computing research using IP-over-lambda networks interconnecting services at all levels. StarLight also supports experimental protocol</p>
--	---

	and middleware research of high-performance application provisioning of lightpaths over optical networks.
<p>GENI (http://www.geni.net/) Tom Mitchell</p>	<p>GENI is an open infrastructure for at-scale networking and distributed systems research and education that spans the US. GENI engages a large community of testbed developers, operators, and research end users.</p>
<p>CORNET (https://cornet.wireless.vt.edu/) Vuk Marojevic</p>	<p>The Cognitive Radio Network Testbed (CORNET) is a collection of 48 software-defined radio nodes deployed within a four-story building.</p>
<p>EMULAB (https://www.emulab.net/) Eric Eide</p>	<p>Emulab is a network testbed, giving researchers a wide range of environments in which to develop, debug, and evaluate their systems. The name Emulab refers both to a facility and to a software system</p>
<p>CentMesh (http://centmesh.csc.ncsu.edu/) Rudra Dutta</p>	<p>Outdoor wireless mesh facility consists of a completely user-programmable mesh of 14 rooftop and pole top nodes, 8 pushcart-mounted nodes, and a small number of airborne nodes.</p>
<p>DETER (https://deter-project.org/) John Wroclawski</p>	<p>Cyber Defense Technology Experimental Research testbed is a computer facility to support experiments in a broad range of cyber-security research projects, including those experiments that involve "risky" code</p>
<p>Distributed Testbeds Rick Mcgeer</p>	<p>Experiences with distributed testbeds including: PlanetLab, GENICloud, InstaGENI and GENI Experiment Engine</p>
<p>CEER Cyber Physical Testbeds</p>	<p>Testbed that enables researchers to validate new technology and provide faster response and recovery following an attack on the electric grid.</p>
<p>Lessons from Houston (http://www.techforall.org/) Edward Knightly</p>	<p>"Technology for all" community deployed wireless network, created to provide free, secure wireless Internet to 19,000 residents in low-income neighborhood.</p>
<p>PHANTOMNET (https://www.phantomnet.org/) Jacobus Van der Merwe</p>	<p>PhantomNet is a mobile networking testbed that provides researchers with a set of hardware and software resources that they can use to develop,</p>

	debug, and evaluate their mobility ideas
<p>iMinds http://ilabt.iminds.be/iminds-wilabt-overview Brecht Vermeulen</p>	<p>The iMinds iLab.t hosts multiple state-of-the-art remotely accessible technical testbeds. The two most used generic testbeds in iLab.t today are the "Virtual Wall" as the primary test environment for experimentation related to wired networks and "w-iLab.t" as the test environment for wireless networks related experimentation.</p>
<p>PANACEA'S CLOUD https://panaceascloud.wordpress.com Prasad Calyam</p>	<p>Panacea's Cloud is a platform for public safety applications and communications that incorporates Internet of Things (IoT) with heads-up displays, virtual beacons, QR-code cards, and wireless mesh network and Incident Command System (ICS).</p>
<p>PLANETLAB https://www.planet-lab.org Andy Bavier</p>	<p>PlanetLab is a global research network that supports the development of new network services. Since the beginning of 2003, more than 1,000 researchers at top academic institutions and industrial research labs have used PlanetLab to develop new technologies for distributed storage, network mapping, peer-to-peer systems, distributed hash tables, and query processing.</p>
<p>WiTEST https://witestlab.poly.edu/ Thanasis Korakis</p>	<p>The Wireless Implementation Testbed Lab (WiTest) in the Department of Electrical and Computer Engineering at NYU Polytechnic School of Engineering conducts research, education, and outreach focused on implementation of, and experimentation with, wireless networking protocols, applications and services</p>
<p>FlexiCoN & WiMNET http://flexicon.ee.columbia.edu/ http://wimnet.ee.columbia.edu/ Harish Krishnaswamy</p>	<p>Testbeds with novel PHY layer (RFIC) implementations for emerging wireless communication paradigms.</p>
<p>SmartSANTANDER http://www.smartsantander.eu/ Luis Muñoz Verónica Gutiérrez Polidura</p>	<p>The Santander is a city-scale smart city testbed composed of around 3000 IEEE 802.15.4 devices, 200 GPRS modules and 2000 joint RFID tag/QR code labels deployed both at static locations (streetlights, facades, bus stops) as</p>

	well as onboard of mobile vehicles (buses, taxis).
CIAN (www.cian-erc.org) Daniel Kilper	The CIAN Testbed provides resources for both theoretical and applied research programs in all areas related to optics and the optical sciences.
NetSense / NetHealth (http://netsense.nd.edu/about.html http://netscale.cse.nd.edu/) Aaron Striegel	Two large scale mobile phone / Fitbit based social studies that collected information on communication events and proximity between the devices.
ORBIT (http://www.orbit-lab.org) Ivan Seskar	The ORBIT testbed is an indoor 400-node programmable radio grid and an outdoor field trial system of short- and long-range radios. Includes SDR platforms, LTE, MIMO and cloud RAN capabilities

Table 1: Past Experience Presentations

2.1. PLANNING ISSUES

The most common mistake identified in these past experience presentations related to project planning is in underestimating the level of effort required:

- (a) In a number of instances, it took much longer to build the testbed than originally planned and projects failed to account for contingencies.
- (b) Operational costs, which are generally quite high, are another undervalued element in project planning.
- (c) The majority of successful projects ended up running much longer than what was originally planned. This, among other things, causes problems with maintenance in the later stages. It was noted that a number of successful projects were able to engage other agencies in order to secure sustainable funding sources. In addition to benefits of additional funding, success of city-scale projects is quite dependent on engagement of a tech-savvy local community partner (i.e., city/regional government and/or the local entrepreneurial ecosystem). It is also possible to improve sustainability by charging service fees for use in production operations (i.e., use of data or infrastructure by independent service vendors).

2.2. DEVELOPMENT ISSUES

Another interesting observation was that quite often the effort of building the testbed ended up being research on the infrastructure itself and that, in order to get the best possible outcome, people building the testbed should also be the researchers using it. This is necessary because just building a testbed based on existing platforms (i.e., COTS) without much novelty is often not very successful because it mostly draws from existing experiences and thus limits research scope. A recommended approach is to allocate 50% of grant funding to experimental research related to the testbed being developed. This has worked well in earlier projects because it ensures that meaningful experiments are available to drive testbed design and evaluation.

In general, existing platforms and testbeds support two broad types of experiments resulting in different requirements:

1. Deployment/open-ended studies requiring flexibility and evolvability in the testbed
2. Rigorous hypothesis testing requiring stability and reproducibility

In a number of instances, the choice was made to support the diversity of experiments rather than a large number of experiments.

2.3. USER AND COMMUNITY ISSUES

A number of past experience presentations discussed user expectations and the fact that they significantly depend on the user's experience levels. Ease of use especially for the novice users should be the principal design criteria. Keeping a minimal core and adding advanced features over time was seen as an important way of getting buy-in from the user community.

Users in general expect high level of availability and reliability for the testbed, which also requires dedicated support personnel and 24/7 testbed operation (which is quite unusual for a typical university/research environment). The importance of varying documentation levels supporting novice, intermediate and advanced users, especially for radio-based experimentation was also highlighted. In the initial project phase, development roadmaps need to be clearly stated and revised periodically to make sure that user expectations are properly managed. Each project needs a person whose primary job is to enhance the testbed and who is committed to improving research outcomes as well as user experience without ad-hoc customizations.

The need for an identity management federation was identified as one of the important aspects of testbed usability. In particular, existing platforms like InCommon and OpenID, in addition to being a form of consolidated and secure identity management, can significantly simplify the new user sign-up process as well as minimize security risks. However, it was also noted that each project has to maintain local identity management in order to support users whose organizations are not participating in the existing identity management platforms.

The value of common platforms that can be shared between academia and industry was also recognized. An engaged research community and strong industry participation can significantly improve outcomes. It was pointed out that, in addition to helping with the build-out and development, the industry has dedicated resources that can significantly help with the maintenance as well as speed up the experimentation processes.

Finally, teaching was recognized as one of the testbed use cases that need special attention given that it can significantly improve the students' learning experience. Successful implementation of the testbed based "classroom-as-a-service" that provides packaged educational materials, exercises and experiments, was recognized as one of the main objectives in platform development.

2.4. LEGAL ISSUES

A number of legal issues were raised during the discussions. The issue most often mentioned was the intellectual property license for the code developed in the project and the need for it to be specified up front. Also, the project should allow for more restrictive code licenses, but that has to be limited and specified up front. Another licensing related issue discussed was regarding vendor supplied software or tools that might be required for experimentation and their right-to-use not just by developers and/or operators but also by experimenters (including addressing potential ITAR issues). Also, every large project had to address the consortium issues (rights and obligations) as well, and various existing projects were cited as examples that are readily available and should be used as a blueprint because they passed the test of time (in particular PlanetLab Consortium agreement was mentioned as a very successful industry/academia type of agreement).

2.5. SOFTWARE ISSUES

Many testbeds discussed at the workshop had problems with proprietary software (mainly with software maintenance) while on the other hand, reusing code from other (open source) projects in order to build a common code base proved quite successful. In addition to simplified maintenance, open source based platforms can facilitate integration of new and legacy infrastructures. Similarly, community wide repositories

(GitHub/Bitbucket/GitLab), development practices and tools for both code and documentation, are essential to attract a wide developer community and increase software adoption.

2.6. OPERATIONAL ISSUES

Testbed and platform maintenance was repeatedly raised as one of the main operational issues. Especially with wireless equipment, constant real-time monitoring of testbed/resources for misuse is required. Furthermore, most resources need periodic calibration and determining how to do this while the testbed is in full operation is a major challenge. Another observation was that even simple maintenance tasks become challenging at scale. Automation is key to success and having professional staff at hand for these tasks is essential. Most of the testbed resources need monitoring for failure detection and, in order to improve user support, this monitoring data has to be publicly available (e.g., global NOC status pages). Successful testbeds need skilled people, and training such staff takes a long period of time.

2.7. GUIDANCE FROM NSF LARGE FACILITIES

Robert M. Hengst, Rebecca Yasky and Jeffrey Zivick from the NSF Large Facilities Office (LFO) gave a presentation on planning processes and efforts related to running large facilities and how they can be applied to smaller projects without significant overheads that are associated with large scale NSF projects. Typically, there are science and program sides to each project with three stages: a.) planning/design/development, b.) operation and c.) divestment. For the duration of the project, the LFO holds monthly meetings with recipients in order to track the execution.

On the science side, the focus has to include the full lifecycle cost of the project and longevity of the proposed facility. The design stage is typically broken into three phases: conceptual, preliminary and final with both internal and external reviews before getting full approval (and funding). The Operations stage is not just about running the facility but also about improving technology and science instrumentation on an ongoing basis (i.e., in the development stage the instrumentation is built to a minimum requirement under the assumption that it will be improved later). Once up and running, the project needs to start working on divestment strategies. Long-term projects require risk-benefit analysis up front but NSF is also considering spiral development (that was given as an example for the GENI project) as a way of managing risks associated with projects of longer duration.

Workshop participants recognized some areas where these two project styles may differ, but also identified areas of commonality, such as risk assessment and management, which is applicable at essentially any project scale. The overall

conclusion was that the risks should be identified up front and be based on input from all participants and should be tracked on an ongoing basis. Dependencies between project elements should be minimized and linkages decoupled as much as possible. Partners in the project should be prepared to share risks on an ongoing basis. The following risks were identified as critical for any larger project in the data-networking domain:

- Not properly estimating labor requirements for running the testbed
- Relying on graduate students to run the testbed
- Not figuring out how to sustain the testbed
- Managing the “scope creep”

In this domain, builders are engineers and scientists are users. Another important observation was that new technology can force a mid-course reset and that any large project should be prepared for that.

3. Breakout Sessions

The first day afternoon session was split into two breakout groups:

- a) Planning and organization and
- b) Usability.

The bulk of the discussion was aimed at using the lessons from the earlier sessions to focus on how the PAWR initiative can benefit from this shared knowledge.

3.1. Planning and Organization

The objectives of the Planning and Organization breakout session were to discuss the issue of maintaining architectural viability and changing the scope of the technology over time. This includes understanding budgetary requirements and fiscal planning issues and in particular staffing needs (based on the past experiences where the community seems to be underestimating how much effort is needed to run testbeds - especially given city- scale, a la PAWR). Another important question for this breakout session was how do we sustain a large testbed beyond NSF support and what are the incentives, challenges and solutions for industrial and municipal involvement.

Given the scale of the planned PAWR testbed deployments, there are a number of participants that need to be involved in a planning process: researchers, city government, city infrastructure providers, private infrastructure providers and equipment vendors.

NSF has helped convene an industry consortium that so far consists of 24 companies (21 companies and 3 associations) committing 50 million dollars for the PAWR initiative. By virtue of putting money and in-kind support/services, the consortium members will be allowed to participate in any of the winning deployments. Facilities provided by the consortium members will be announced as part of the solicitation process so that they can be included in proposals. Other corporate partners will be able to participate only as a part of a proposal for a particular testbed/city deployment. A number of cities have already expressed interest in participating, suggesting that there will be a large number of potential deployments. At the same time, cities are already deploying various forms of IoT networks providing the opportunity to integrate these efforts with PAWR.

During the discussion, the need for the balance of long term research vs. near term industrial goals was identified. The challenge is to form a proper balance of advance wireless technology research with production network related research that is using deployed wireless platforms. The science plan and scoping statement are hard problems for wireless networking given the range of potential uses and variety of infrastructure elements. One of the conclusions was that we are not supposed to compete with commercial services and should focus on academic research (services) and find an intersection with industrial interest and at the same time allow use of the resulting infrastructure (possibly simultaneously through use of virtualization). The conclusion was that the focus should be on pre-competitive research looking 3 to 5 years ahead rather than the immediate future in order to avoid industry exclusivity requirements but still be attractive enough for their participation. Another observation was that cities (and not just cities but also states since there are a number of related state-wide a.k.a "SmartState" initiatives), included as the third major player in deployments, would have to find their own interest in supporting the testbed/infrastructure.

One observation is that in order to achieve sustainability of research we need to design a system that has a number of features that can be expanded over time. Modularity and evolvability are important aspects that should be built into the architecture. This is especially important given that wireless hardware becomes obsolete very quickly so the architecture has to demonstrate the ability to accept new wireless platforms over time and support generational change. Special care has to be taken to avoid interference with critical production services and we need to define a list of protected services that are always going to be guaranteed.

Each deployment will need to carefully define the actual scope since deployments can span support infrastructure, wireless technologies and applications. The variety of

possible technologies includes: novel physical layer devices, new wireless systems, scaled deployments of regular sensors, IoT devices, licensed and unlicensed infrastructure/production devices, services over production networks, etc. This diversity implies that it is quite possible that each deployment will have results with a distinct architecture as opposed to an instance of a single modular architecture.

The industry consortium in collaboration with the research community will have to define a technical and organizational strategy for continuous relevance and sustainability. A lengthy discussion was focused on possible evolutions towards commercial use but the preference is for continuous research focus, which should be supported at least partially with funding from federal agencies (other than NSF).

3.2. Usability

UI design and ease of use

A number of issues on UI design and ease of use were discussed including:

- Modern and standardized UI design with built-in contextual help
- REST APIs for users to leverage federated resources
- Potentially a specification language that can be ported to different hardware
- Experimentation-as-a-Service: tools to co-create experiments
- Hardware opt-in capabilities
- Users support and follow-up (e.g., knowledge base, common guidelines ticket systems, GitHub, mailing lists, questionnaires, etc.)

One of the main conclusions was that the driving principle of UI design has to be: “make it simple”. Another observation was that it could take a year to train a graduate student. As a result, there is a strong need for step-by-step user documentation, example experiments and tutorials. These tutorials should include steps for, along with complete experimental set-up examples, a progressive introduction to features and capabilities. A repository of successful experiments was recommended.

Providing documentation, training materials to provide an educational component and lower the slope of the learning curve will lead to wide participation from non-core networking researchers. Automation of help/user mailing lists also allows building a community that helps each other, as some users become experts.

We also need to be cognizant of smaller university participation so that we don't widen the gap between those that have resources and those who do not. This implies that we need to learn about who the users are and what their competency level is.

Service tasks such as reserving time on the testbed or renewing resources should be accessible from a mobile platform as well since students, our main user base, want this feature.

The perspective of tiered access to users is discussed below and has implications for UI design where a feature set for a user depends on their classification and, as they become experts more features are available for them to use.

Supporting a diverse and large user-base of researchers

The first question to answer is “define the term testbed user”. PAWR envisions a diverse set of users such as community residents, service providers, researchers and students. Generally, it is important to make clear in documentation that “users” can mean the community that is accessing services from a community centric testbed. This needs to be distinguished from “users” that are researchers that use the system as a testbed.

The various user type discussions raised an interesting point from the perspective of access control to resources. There will be different experiences and expertise levels amongst this diverse group. There needs to be a mechanism of being vetted before full access to resources is allowed.

We propose a Reputation management system where users are given tiered access to this high-end equipment based on their expertise level. There could be policies in place for different users. Since this is a wireless outdoor testbed the experimental license onus falls on the users. A research faculty advisor or a committee can also evaluate if the experiment proposed is valid and the user has the capability. We propose a concept of a RF firewall to notch standard frequencies, on which the equipment would not be able to transmit.

These concepts can be tightened and vetted by getting community feedback via workshops that have worked well for previous wireless testbeds.

Portability of experiments

Portability, as defined by the group, in the PAWR context would mean that platforms investigating similar research challenges would allow experimenters to export resources and experiment state in a common format across platforms e.g., virtual box, topology and shared within other testbeds. This would entail developing features such as common API(s), data model for accepting experiment state and defining models for what is common across platforms as not every feature is common across PAWR

platforms(s). This will entail trade-offs over the complexity of what is common vs. ease of portability.

There are also interesting questions around what are native API(s) for certain vendor platforms and how there will be a need for a common wrapper to port across different hardware platforms.

Long-term experiment curation/dataset storage

One of the questions that was raised during the discussion was: how long does an experimenter have access to old experiment results and what sort of sharing mechanisms should PAWR provide for others to access data collected on the platform? Prior experiences have ranged from testbeds not having the resources to keep experiments to testbeds who leave the management of data and archiving to experimenters. We acknowledge that storage is cheap and maybe outsourcing the data storage to an AWS or third party cloud provider is the easiest solution. We also need to define good or bad data as there is too much information to collect, parse, analyze and it may not all be relevant.

As was mentioned in a number of prior experience presentations, the community needs a common format for looking at data and organizing it using a recommended classification system with descriptors enabling efficient searching as well as curating GOLD standard experiments that reflect best practices (requiring "editing" or even "peer reviewing" for the 'important' artifacts in the datasets). Another related question that should be addressed was how to deal with experiments and/or data artifacts that may have 'expired' (and are thus not repeatable anymore).

Evaluation of testbed performance

Performance in this context is not bits/sec/Hz or number of real users. The performance here is comparing what was promised at the development phase of the testbed to what is actually deployed. NSF's prior role in evaluating performance of testbeds was largely cursory, which will not be the case with the PAWR initiative. Previous experiences have ranged from monitoring user lists but this ties into who will survey the lists and that depends on what state the testbed is in: build-out, deployed, in-use, saturated.

The ideas were to have third party reviews in the form of advisory boards and steering committees look at proposed vs. actual architecture, users, and deployed technology. We propose a close look at real research user feedback or organized surveys to gauge what went wrong and, going a step further, to monitor and follow up those users who tried but could not finish using the platform i.e., to follow up users throughout the process as part of support.

Securing testbed infrastructure against misuse, abuse and cyber-attacks

The platform will host a multitude of OS, kernel, and user file system versions. The security of these images, which will become obsolete and vulnerable to attacks within a couple of years, is a serious issue. However, monitoring and de-commissioning software will require active support to ensure known vulnerabilities are addressed and cannot be exploited.

The testbed community has plenty of reactive processes and procedures such as KILL and Emergency stop that are practiced regularly when security issues are discovered. There are well-documented processes to identify the custodian of a compromised resource or software, as resources, time or space shared are accountable to a single point of contact.

Another layer of security is physical security of outdoor deployed hardware, which is expensive and a target for theft and vandalism. Also, even if malicious actors do not compromise resources, monitoring infrastructure has to be able to detect and react to inadvertent misuse, which is quite common in experimental systems.

Lessons in staffing from pre-planning to sustenance stages

There needs to be adequate full time staff to operate, monitor, develop, deploy and an adequate supply pipeline for staffing, as the field is a hot area and staff turnover will be high. The emphasis on documentation is essential and requires a full time staff dedicated to its creation and maintenance.

4. Visions for PAWR

The second day started with the keynote talk by Walter Johnston outlining future challenges and top priorities from a city-scale testbed from the FCC perspective identifying new mm bands, flexible licensing rules and ensuring communications security as main priorities. It also posed two important questions for the community that could potentially be addressed by the research enabled with the (large-scale) experimental platforms:

1. Can you run in experimental spectrum without interfering?
2. Can you reliably identify and provide proof of emission for interferer?

The keynote talk was followed with the short presentations on individual visions for PAWR that are summarized in Table 2.

Mooi Choo Chuah	<p>Need for IoT testbed with secure automated software updates, incremental upgrade of nodes and continuous software verification while supporting:</p> <ul style="list-style-type: none"> • dynamic network configuration and bandwidth allocation • wireless charging of IoT devices, real-time data analytics (edge cloud analytic engine), privacy aware data mining
Mark Berman	<p>Distributed scaling through federation:</p> <ul style="list-style-type: none"> • identify (access credentials), control plane (common API), data plane (connectivity and traffic rules) • local ownership/management of resources • unique capabilities/goals of local testbeds <p>Building a user base with active outreach to both experimenters and community users (user needs change over time mandating shift from 'retail' to 'wholesale' support model)</p> <p>Community structures:</p> <ul style="list-style-type: none"> • experimenters, developers, infrastructure owner/operator, community users, industry partners • researchers with priority allocation and strong representation in decision-making
Vuk Marojevic	<p>Additional industry and state/federal agencies should be allowed to participate</p> <p>Allow matching funds and flexible cash flow rates</p> <p>De-conflict production and experimental environment</p> <p>Remote automation and management are critical</p>
ESNet	<p>Software-defined everything is a way to keep the architecture viable over the long term.</p>
Rudra Dutta	<p>Advanced Wireless testbed or a SmartCity testbed?</p> <p>Why talk about 'city/community users'? What about IRB issues for such a large user base?</p> <p>Can it be city street poles?</p> <p>Bring Your Own X - what is X? (device/ idea/ equipment/ code/ people?) Needs to be clearer</p> <p>Being able to plug in diverse X (mostly unforeseen) is paramount - without assuming large support staff, but enable users</p> <p>Orchestration is a primary function, not just slicing and composing, specifying clear interfaces and monitoring them</p>

	<p>Train users and staff on a continual basis</p> <p>Less important: Slicing, number of new users, any specific cyber infrastructure technology (prioritize above first)</p> <p>Come up with mom-and-apple-pie use cases to incentivize cities</p>
Dan Kilper	<p>Need to get into bed with the city: community support, matching funds, economic development, startups, digital divide, etc.</p> <p>Not a national backbone network, so can do things that were not feasible on national scale (e.g., Bristol is Open)</p> <ul style="list-style-type: none"> city experimentation as a service
Harish Krishnaswamy	<p>Be able to support emerging wireless technologies and custom hardware deployment (bring it and leave it).</p>
Aaron Striegel	<p>Debugging/management support should be first principle</p> <p>Calibration is a hard problem, need to figure how to calibrate on the fly, or in a quick turnaround (of order of a few hours)</p> <p>Curation/reproducibility is a key concern with a common data format (JSON or something equivalent)</p> <p>Reference stack/toolset - package support for a Docker wireless equivalent within a NSF Small proposal?</p>
Verónica Gutiérrez	<p>Pay attention to data and analytics that come out of it, not just wireless technologies.</p> <p>Interoperability of devices is critical - provide appropriate interfaces. Start small, and scale up over time, with a set of canonical use cases first.</p> <p>Use the right incentives (i.e. no need to go overboard with incentives).</p> <p>Periodic and appropriate communication with stakeholders is a key to success.</p> <p>Enable participation for smaller companies that will be eager to work with specific testbeds</p>
John Wroclawski	<p>Should be realistic (users, user behavior, economics, mgmt. skills)</p> <p>Need challenge problems that illustrate and motivate 'taxonomies' of realism.</p>
Manu Gosain	<p>Promote use-inspired research (Edison's quadrant or Pasteur's quadrant)</p> <p>Make the pioneers/early users happy; rest will follow.</p> <p>A successful platform mediates interaction and defines common APIs that facilitates widespread use</p> <ul style="list-style-type: none"> similar basic building blocks
Eric Eide	<p>Replication and reproducibility is critical - can multiple testbeds be used for this purpose?</p> <p>Can the experiments be portable across/within testbeds?</p> <p>Use existing technologies (RSpecs, Chef, Ansible, OMD, GEE</p>

	XPFlow) to specify configuration, execution and workflow
Stefano Basagni	<p>Can city be programmed? SDN technology needs to be brought to par with the distributed and performance requirements of heterogeneous wireless networking. We need to bring prevailing and new technologies together.</p> <p>New platforms will require experience in large testbed management, but also deep knowledge to appreciate the challenges solved by the platform and the relationships between platform and societal evolution.</p> <p>Academic institutions should have a central role.</p>

Table 2: PAWR Vision Presentations

5. Conclusions and Recommendations

This section presents the workshop’s conclusions and recommendations. These are broken into two categories: those that are generally applicable to a large-scale networking platform community of practice and those with more specific applicability to the PAWR program. It is worth noting that there is necessarily substantial overlap across these categories, and that the general recommendations are certainly also applicable to PAWR.

5.1. Community of Practice Recommendations

1. Dedicated roles should be staffed with key personnel: Platform plans must cover not only initial development but also deployment, maintenance, and operations. The key to successful deployment and operation of a large scale distributed testbed is the staff - both managerial as well as operational. Successful teams staff key roles with dedicated personnel and have an adequate staffing pipeline. A key recommendation is to provide enough resources for recruiting and maintaining sufficient staffing through the operational lifetime of a testbed to increase the likelihood of success.
2. ***ers** of the platform: Experimenters, Users, and Consumers (**ers*) of the platform will each have their service needs and expectations. Successful platforms cater to these experimenters, testbed builders and end users/consumers. Sometimes multiple of these **ers* in the same capacity. It will be important to have clear consensus of the classifications since the requirements of each group of **ers* will need to be met differently. While this recommendation is important as a general principle, it is particularly important for PAWR because it plans to include a very broad range of **ers*, including platform builders, researchers, industry, and the

general public. With cooperation from the PAWR community, the definition of a platform as it pertains to each constituency will help clarify the scope, scale and service requirements for these *ers. It is also imperative for PAWR to drive a much more rigorous set of priorities for each testbed.

3. Proper data management plan and data sharing: To facilitate data and result sharing, it is necessary for each project to establish the relevant (and common) data standards. This applies equally to data collected by the infrastructure, and data from experiments (that are typically quite different). Given the possible volume of data that large testbeds/experiments can collect, sufficient facilities and strategies for data processing (data analytics) should be provisioned as well. The benefits of having one common system across all the testbeds should be considered.
4. Testbed design needs to support real world usability with consumer devices, implying the need for a mix of emerging technologies and current state-of-the-art. In the PAWR case, it is critically important to avoid vertical stack silo architectures typically associated with cellular access, avoiding the temptation of easily acquired commercial equipment from vendors. It will need a research community investment in open programmable base stations and devices as well as open and fully programmable SDN and cloud platforms.
5. Reputation System: Large platforms often include different types of resources. These resources may have varying capabilities and performance characteristics and may require different levels of user expertise. Platforms should have a user-driven reputation system that makes an appropriate set of testbed resources available to the right kind of users.
6. Successful platforms take advantage of the work of others. A large-scale platform should implement from scratch only those core capabilities that are unique to that platform. Other testbeds, commercial infrastructure, and open standards can provide reliable implementations and encourage interoperation, while simultaneously reducing development, operations, and maintenance burden. Examples include compute grid, identity management systems, and data analytics platforms. (In the specific PAWR example, the PPO should encourage the use of existing research cyber-infrastructure to provide connectivity and encourage interoperation among different PAWR platforms and technology. For example, platforms should consider GENI for edge cloud capability and connectivity to wired core resources via Internet2, as well as NSF Future Cloud (CloudLab and Chameleon) and commercial cloud providers for experimental

and data center cloud reach-back. The PPO may wish to include as a requirement for PAWR platform proposals a specific cooperation plan, in which proposers should expect to highlight cooperation among entities.)

5.2. PAWR Recommendations

7. Primacy of research¹ is of highest priority: Maximizing PAWR's research utility will be best achieved by seeking the broadest feasible representation of participants and technologies. It is paramount that a large-scale testbed supports a wide variety of depth and breadth across various research topics. This principle should be observed in (a) community representation, (b) PAWR platform design, and (c) platform resource allocation. To address community representation, a researcher-driven group should be identified to serve as counterpart to the recently formed advanced wireless industry consortium. With respect to platform design, a wider variety of research trumps commonality of testbeds. Hence, the portfolio of testbeds in the PAWR program should be a greater sum than its parts and sacrificing some commonality to encompass diverse topics of research is important. In the area of resource allocation, at least 50% of airtime should be reserved for (NSF/community sponsored) research; the other prospective users can then make use of the other 50%. The PPO should be responsible to maintain this portfolio.

8. The Research Community needs help in connecting with city stakeholders: The success of PAWR platforms will require coordination and cooperation amongst diverse partners: city governments, non-profits, startups, and campus executives to name a few. The academic community needs to be educated on how to interact with these various actors. A matchmaking marketplace sponsored by NSF or a teaching workshop conducted by other successful research/ non-profit organizations that have formed such partnerships such as US Ignite and Mozilla should be used. An academic PI will need broader support from outside their institutions and NSF should provide a framework or know how to enable these interactions.

¹ In this context, the term "research" is intended to exclude proprietary or product-focused studies and refer exclusively to academic style research publishable in the open literature.

Appendix A: Workshop Attendees

Name	Organization
Harish Krishnaswamy	Columbia University
Padma Krishnaswamy	FCC
Walter Johnston	FCC-OET
Mark Berman	GENI Project Office
Tom Mitchell	GENI Project Office
Brecht Vermeulen	iMinds
Douglas Castor	InterDigital
John Kaewell	InterDigital
Byung Yi	InterDigital, Inc.
Joe Mambretti	International Center for Advanced Internet Research Northwestern University
James Gigrich	Keysight Technologies
Mooi Choo Chuah	Lehigh University
Sarah Yost	National Instruments
Ben Hilburn	National Instruments
Thyaga Nandagopal	National Science Foundation
Darleen Fisher	National Science Foundation
Jack Brassil	National Science Foundation
Thanasis Korakis	New York University
Rudra Dutta	North Carolina State University
Stefano Basagni	Northeastern University
Andy Bavier	Princeton University / ON.LAB
Abhimanyu Gosain	Raytheon BBN Technologies
Edward Knightly	Rice University
Anduo Wang	Temple University
Krishna Narayanan	Texas A&M University
Scott Miller	Texas A&M University, Dept of ECE
Dan Kilper	University of Arizona
Ramesh Rao	University of California, San Diego
Verónica Gutiérrez	University of Cantabria
Polidura	
Kenneth Baker	University of Colorado
Tim Yardley	University of Illinois Urbana-Champaign

Prasad Calyam	University of Missouri-Columbia
Aaron Striegel	University of Notre Dame
John Wroclawski	University of Southern California/ISI/DETER
Kobus Van Der Merwe	University of Utah
Eric Eide	University of Utah
Nick Maynard	US Ignite
Rick Mcgeer	US Ignite
Glenn Ricart	US Ignite and U. Utah
Jerry Park	Virginia Tech
Vuk Marojevic	Virginia Tech
Jeffrey Reed	Virginia Tech
Ivan Seskar	WINLAB, Rutgers University