PETE HERZOG

the protocol historian

Pete is the Managing Director and co-founder of ISECOM, the Institute for Security and Open Methodologies, and is directly involved in all ISECOM projects. His main objective is to make security make sense.

pete@isecom.org

THE MANY PROTOCOLS CREATED, thriving, dying, and dead are quietly being documented in detail beyond that of the RFCs that introduced them. Accidental protocol curator and historian Dru Lavigne has been going beyond the technical details of Internet protocols since early 2001 to get the human side of invention.

Dru found that what started as a project to list common port numbers mapped to their associated applications for the appendix of the OSSTMM (www.osstmm.org), the standard methodology for security testing, would quickly evolve into more. Now as the OPRP (Open Protocol Resource Project; www.isecom.org/oprp/) it is one of the main projects for the open, nonprofit, security research community, ISECOM (www.isecom.org). When she volunteered in 2001 to assist in reviewing the OSSTMM, she couldn't help but notice that the mappings were woefully incomplete and, in her opinion, "not much of a help to anyone who would be interested in knowing which application was most likely associated with a port." Keep in mind, this was years before Fyodor introduced nmap -sV. She had already experienced her own frustrations in scouring the Internet looking for information on various ports. This seemed like the perfect opportunity to organize her previous research forays and make them publicly available so others could benefit as well. And since no one had previously shown any interest in this section of the OSSTMM, it became hers to do with as she could.

The actual goal of the OPRP is to provide a quick reference for those who are wondering what application may be running on a particular port. This has actually become easier since LAS (www.localareasecurity.com) created a Firefox plug-in to allow quick searches of the OPRP. The OPRP is meant to augment, not supplant, the official repository of registered port numbers (www.iana.org/assignments/port-numbers). This is the reason why Dru tries to contact the original protocol registrants for a description to include within the OPRP. The IANA has been registering port numbers for over two decades, and much has changed during that period: Products have come and gone and been EOL'd, and companies have been merged and purchased and perhaps swallowed by the dot-com bubble. The OPRP tries to determine whether each particular port is still in use today, and if so, in what products one can expect to find its usage.

The OPRP isn't meant to be a definitive source or a guarantee of what is running on a particular port. That is impossible, seeing that it is trivial to change the default port for almost any TCP/IP application. However, for the security tester or sysadmin reviewing firewall logs, it gives a starting point to see what is supposed to be there and if that is a likely application for the given environment.

The number of her protocol descriptions has now surpassed 1500. Currently there are approximately 5000 IANA registered protocols. That means she has curated descriptions for roughly one-third of the registered protocols. To put this in perspective, approximately 165 protocols have been registered thus far in 2006. Dru waits six months after those protocols are registered before contacting the registrants to give them time to get their protocols in use. "I quickly learned that it wasn't productive to contact registrants immediately, as protocols are often registered in the early stages of product development," Dru states. "I've found that a window of six to eight months after registration is most effective; by that time, the protocol is often actively in use and 'out in the wild.'"

Dru started with a simple guiding principle: who better to know whether a protocol is still in use and who better suited to provide a useful description than the person who registered the protocol? In her first round of contacts, she simply emailed all of the email addresses found in the IANA official list of registered ports. Since many of those addresses were long extinct, she saved all of the nondelivery messages for the next stage. It also did not help that only recently has IANA began dating registrations. However, a surprising number of email addresses did still work, and several hundred descriptions were received and input into the OPRP.

In the next stage she used her Google skills to see if she could find the remaining registrants. That garnered another 500 or so descriptions. Now she has two folders she works with: the nondelivery messages for newer but extinct email addresses and a folder for email that was successfully delivered but to which she didn't receive a response.

Stage three involved finding contact information for the companies that had registered protocols but for which the original registrant was unresponsive or could not be found. Although some may be unresponsive for trade-secret or corporate confidentiality reasons, another problem is sometimes that the protocol seems to have disappeared completely. "At this stage I'm still working out a plan for how to get descriptions for the protocols which perhaps didn't survive company mergers," she says. "For example, how many DEC and Compaq protocols are still being used in HP products? Or what of the protocols that were registered by companies since swallowed by IBM, Nortel, or Cisco?"

While some might think that a hobby or job as protocol historian may be dull, Dru finds it fascinating. She says she just naturally likes to organize information. She has a particular fondness for protocols, which is a natural extension of her need to know how things work. She is also fascinated by history, including the history of the Internet and TCP/IP. This is apparent in how the OPRP has started to become a repository of descriptions of historical protocols. IANA simply puts a "de-registered on date" note on file. "I would hate to see the name and history behind a de-registered protocol lost forever," Dru says. "I currently have 1185 delivered emails which I haven't received a response to and 118 nondeliverable emails." Her goal is to catalog them all.

When asked about herself, Dru says, "For those that are curious about my age: Neil Armstrong said, 'That's one small step for man but one giant leap for mankind' on my fourth birthday. At the time this was memorable simply because I was irritated that a bunch of boring grownups had preempted my favorite TV shows in order to talk endlessly about the same news clip. Since then, I've come to appreciate that seemingly small actions have ripple effects. This is part of what attracts me to open source. It is also a prime motivator for the many projects I am involved with, including the OPRP."

When she returned to school to study networking, she was bemused that most classmates found protocols to be so much boring theory. "I'm fascinated by anything that gives insight into how things work. I'm also fascinated by the stories behind how things came to be, so I was naturally drawn to RFCs and Internet history," she says.

She's just like any other busy person in IT who somewhere along the way became "known" within various open source communities. As to where she works (being a protocol historian doesn't pay a salary), Dru says it's not a short answer. She says every day is a bit different, with the threads of several ongoing works intertwining. She's been teaching IT certifications, most recently in Ottawa, since 1998 and is the acting chair of the BSD Certification Group (www.bsdcertification.org), a registered nonprofit with a goal of providing an IT certification for assessing the skills of BSD system administrators. She's also been a system administrator since 1996, starting with Novell and Microsoft systems and later integrating these with Linux and BSD systems. Since 2000, she has been writing technical documentation for various products, courseware and labs for various curricula, a column for O'Reilly (www.onlamp.com/pub/ct/15), and, most recently, another for IT Toolbox (blogs.ittoolbox.com/unix/bsd). She also attends and/or speaks at various technical conferences as well as meeting regularly with my local BUG (BSD User Group) and GOSLING (Get Open Source Logic INto Government). However, the OPRP project is something that she cares about deeply, and it puts her in touch with the movers and shakers of the information age.

"I've received everything from very terse replies indicating that the protocol is still in use but covered under an NDA to long essays on the details of the protocol," she explains. "Some responses could be considered a marketing slick, but that's fine as it still answers the fundamental questions, 'Is this protocol still in use, and what company/application(s) are using it?'"

Dru says she's been pleasantly surprised at the overwhelming positive response by registrants to the OPRP and has had only two belligerent responses since 2001. "I think this speaks to the professionalism shown by the registrants and the respect in the IT community for the ISECOM organization," she says. "I've also been humbled by receiving responses from very big names in the IT industry, the type of names that networking geeks such as myself considered to be demigods when it comes to the Internet and TCP/IP."

Some of these legendary responders include Bob Braden, whose research interests include end-to-end network protocols, especially in the transport and Internetwork layers (en.wikipedia.org/wiki/Bob_Braden); Joe Touch, whose interests include Internet protocols, network architecture, high-speed and low-latency nets, network device design, and experimental network analysis (www.isi.edu/touch/bio.html); Joe Pato, whose current research focus is on the security needs of collaborative communities, addressing both large-scale inter-enterprise models and the challenges of ubiquitous devices (www.hpl.hp.com/personal/Joe_Pato); and Linus Torvalds, who, she says, responded within fifteen minutes on an Easter Sunday.

For Dru to say which protocol has the best story is like asking a kid in a candy store what her favorite candy is. It really is hard for her to say. Since every protocol has a story, oftentimes a story of genius and hopes and dreams of real people trying to push forward the information age to an even greater age of ubiquity and enlightenment. For instance, there is the port number that represents the birthdate of a developer's daughter. Another port number represents the date of a wedding anniversary. Then there are protocols that were as ubiquitous as HTTP is today but that have since become extinct, including protocols that represent the excitement of the dot-com era but that never saw a single shipped product. TCP 1456 was registered for use by OpenMind, a groupware application published by DCA and then Attachmate. Even though it won Product of the Year in 1995, it is no longer in production or commercially available. TCP/UDP 1305 was originally registered as pe-mike, but the company was bought out and no products were ever released to a customer that used this protocol.

This is exactly what Dru likes best about being a protocol curator: the human drama behind the invention. Those she finds most notable are as follows:

"MilliCent used to use ports 1180, 2180 and 3180," says one email response Dru received. "When it existed, [it used] TCP. Now it doesn't exist and it uses neither. This MilliCent protocol was originally created for DEC, then Compaq and now HP. The project is now defunct, but it was great."

The response regarding port 1989 says, "Originally developed by the University of Sydney and Message Handling Systems Py Ltd, Australia, and first sold in 1989. MHSnet has been used to build message networks where the links range from poor quality up to Internet quality. It was used by the Australian Govt Department of Foreign Affairs and Trade to build a message network between embassies and posts. It has been used to build many private networks but was also the backbone of an academic network (ACSnet) in the early days of networking in Australia."

Another responder wrote, "I think 585 was an administrative error. I believe it had been originally ear-marked for IMAP-SSL, but that turned out to be 993. I either never knew or have long forgotten what caused that situation, and alas, we can no longer ask Jon (Postel). In any case, if 585 is alive, I don't know anything about it."

The response regarding ports 309, 709, and 710 says, "When I was employed at Entrust, Inc. (1994–2001), I registered those ports (which are all based on my daughter's birth statistics, time, date, and weight, respectively). . . . Note that 709 is deprecated in favour of 829 (PKIX CA/RA; 829 is the wedding anniversary of the fellow who registered that one, Carlisle Adams, the CA in CAST)."

The protocol Gopher, which was the precursor to the Web on port 70, had been hugely popular until it got eclipsed by the Web. Dru received the following response in regards to this behemoth that has nearly shrunk to nothing: "Internet Gopher popularized the notion of distributed information systems before the World Wide Web. Client and server software is available for most popular platforms. Although the original Gopher developers at the University of Minnesota are no longer actively working on this project, other groups are. For instance see http://gofish.sourceforge.net/

and http://gopher.quux.org:70/devel/gopher/pygopherd and the usenet newsgroup comp.infosystems.gopher."

Going back to the backbone of ARPANET, the precursor to the Internet, Dru received: "51 was implemented on the BBN IMPs, which formed the backbone of the original ARPANET and later MILNET (a.k.a. Defense Data Network, DDN). It was used to add a layer of indirection to ARPANET addresses, which were originally tied to the physical ports on each particular IMP (like IMP 18, port 4). Logical addresses made it possible to keep the same ARPANET address without being tied to one particular physical port. However, the use of IP made this moot, since once IP was used, packets were sent to a particular IP address, rather than a particular ARPANET address, and an IP address resolution protocol was used to do the mapping. And of course, the shutdown of the ARPANET made it REALLY moot. However, that's not to say that there's not an old military network somewhere still running IMPs (although I REALLY doubt it)."

Some protocols are more specific and more rare. One such protocol is the one registered for port 91. The responder writes, "The port assignment was for a protocol peculiar to equipment and arrangements of equipment use in the MIT Lab for Computer Science over 20 years ago. All the equipment is now long gone. As far as I know, the protocol has not been reassigned, but I have not tracked such things. As I recall, we used it for TCP, not UDP, but it was a long time ago. The tool in question also handled the Chaos net protocol, a completely different network that I think never propagated elsewhere."

The final response Dru provides is of a dot-com invention that, although perhaps superior to what is now commonly implemented, never got released publicly. The registrant of port 1228 writes, "Florence was a proof-of-concept remote method invocation facility with application-hinted client-side caching designed to improve latency in hierarchical arrangement of nodes. It was hastily rushed into production by a dot-com whose time was running out. I was one of two engineers in charge of its design and implementation. There was one application, a business-to-business exchange running atop Florence, that made it to the demo stage, but as iventurelab.com is now thoroughly and completely out of business and—to my knowledge—nobody purchased the IP, I doubt if any of the source code implementing it actually still exists. The idea was sound, and I've been meaning to re-implement the concept atop a more portable software infrastructure and release it as open source software, using this assigned port, but frankly, this is #3 on my list even of open source priorities, and implementing yet another remote method invocation facility gratuitously incompatible with SOAP, while worth it for the performance gain of protocol-supported response caching, will probably not get anyone too excited about using it."

Other protocols of note from the OPRP are shown in the table on the next page.

Number	*Transport*	*Application*	*RFC/Vendor's URL/MS KB Article*	*Description*
47	TCP	deprecated		Originally registered as NI FTP, the Network Independent File Transfer Protocol, known as "Blue Book." It operated over many years in the UK academic community, primarily over x.25.
51	NPC	deprecated	www.ietf.org/rfc/rfc851.txt	Was used by IMP Logical Address Maintenance on the original BBN ARPANET. It was used to map ARPANET addresses to physical ports on an IMP. This functionality was superseded by TCP/IP.
61	TCP	deprecated		Originally registered as NI Mail and also known as "Grey Book." It was a mail protocol based on RFC 822, operating over NIFTP (see port 47).
81	TCP UDP	deprecated		Originally registered as HOSTS2 Name Server; its registered use seems to have been long deprecated.
96	TCP UDP	DIXIE	http://www.ietf.org/rfc/rfc1249.txt	Was used by DIXIE, which has since been replaced by LDAP on port 389.
105	TCP	deprecated	www.ietf.org/rfc/ rfc2378.txt	Was the CCSO Name Server, the backend of the Ph function of Eudora. It has since been replaced by LDAP.
402	TCP UDP	deprecated	http://web.archive.org/web/19991009142042/www-genie.mrrl.lut.ac.uk/interfaces.html	Registered for the genie protocol, but unused since 1998.
692	TCP	deprecated	http://www.hyperwave.com	Was used for the Distributed Interactive Services (DIS) protocol for core-level access to Hyperwave's backend server architecture.
1228	TCP	deprecated		Originally registered for Florence, a proof-of-concept remote method invocation facility with application-hinted client-side caching designed to improve latency in hierarchical arrangement of nodes. It never shipped, as a result of a business failure.
1427	UDP	deprecated		Was used by a private, experimental protocol developed as part of DARPA-funded research.

Like any other open project, though, the OPRP does have its detractors. Some registrants disagree that the OPRP should include nonregistered usage. Dru's philosophy is that there should be an entry for what is likely to run on a port: for example, well-known worms or Trojans as well as usage by common, though unregistered, applications. She feels this is what will be most useful to an administrator.

However, since worms and Trojans are discovered more quickly than she has time to research, anyone is welcome to add an entry to the OPRP. It is an open database, after all. She reviews these entries before inclusion and her deciding factor for permanent entry is based on the reliability of the information. Something as simple as an URL pointing to supporting documentation, however, can be considered reliable information.

However, Dru does keep tight control over what is entered and she does review it all personally. For future would-be researchers and curators, she advises, "Some registrants have changed the name of the protocol or the company since the original IANA registration and have not updated their info with IANA. If you see a description for a registered protocol in the OPRP, which isn't a Trojan or marked as for unregistered use, that description and name change is from the registrant. Please don't try to add an entry with the outdated IANA information, as it won't be included in the OPRP."

It's an ambitious project. When asked when she thinks it will be finished, she says, "Probably never." As long as IANA continues to register protocols, entries will need to be updated. The OPRP needs to have at least a description for every registered protocol. With that, she comments, "I think any article on protocols should make a reference to Postel (http://www.livinginternet.com/i/iw_mgmt_iana.htm). Postel's contributions to the IANA and RFCs are deeply appreciated by the networking community and a conversation on ports isn't complete without paying respect to him."

As far as anyone can tell, Dru's ambition and busy schedule have already tagged her as a remarkable person, especially within the open source and Internet communities. Her dedication and contribution as a protocol historian are nothing short of amazing.

Dru's final comment to those out there is this: "If you have registered a protocol but haven't received the OPRP questionnaire, email me to request a copy. If your registered protocol needs an updated description, email me the details. If you have contacts for a large corporation's intellectual property department and want to sort out what registered protocols are or aren't still in use, just email me."

You can contact Dru easily at dru@isecom.org. The OPRP is available at www.isecom.org/oprp/.