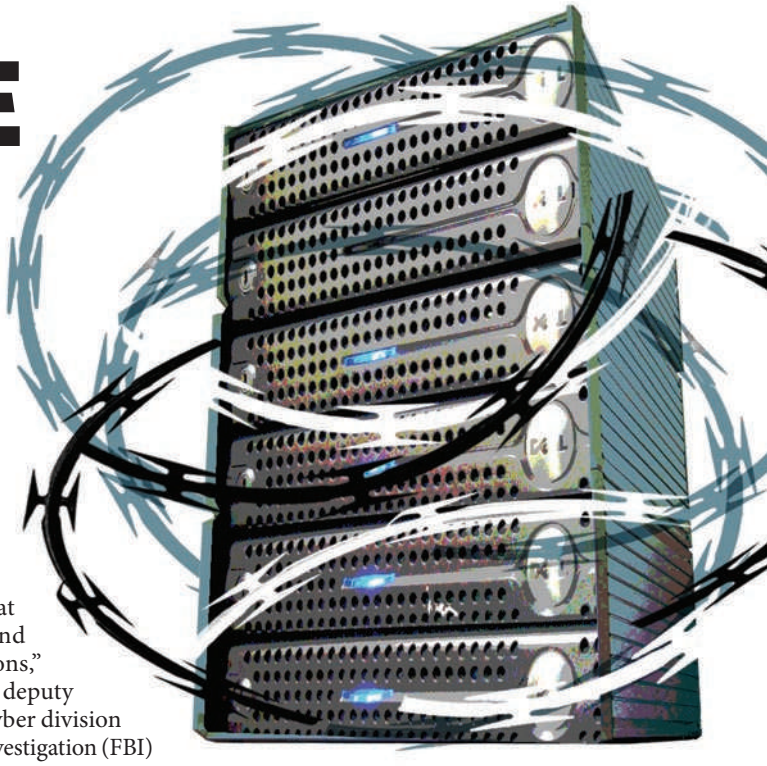
HOW SAFE ARE YOUR DATA?

Many scientists want to keep their data and resources free; cybersecurity specialists want them under lock and key. **Jeffrey Perkel** reports.



A. MARTIN

In 2002, bioinformatician Mark Gerstein and his colleagues set up a server to host some commonly used genomics databases. Operating within this was a free software application called Snort, which actively monitors anomalies in web traffic, surreptitious scans and server nudges that might be attempts to compromise a system's integrity. Then the researchers watched and waited.

Seven months later, the picture that emerged was one of a network under siege. "If you put up a server, it just continuously gets barraged," says Gerstein of Yale University in New Haven, Connecticut. Most days, the server was hit ten times or fewer, but on occasion the hit count spiked into the thousands[1]. Not all of those hits were attacks, Gerstein notes, but many were. On two high-hit days, for example, more than 90% of events attempted to induce a 'buffer overflow', whereby superfluous amounts of data are written into a system's memory in an attempt to make it fail, opening it up to exploitation.

In the face of such a relentless onslaught, most ventures onto the World Wide Web should be taken with caution. Having a graduate student with little experience set up a website could be disastrous, Gerstein says. Entering either through poorly written code or open ports — digital entry points into computer hardware and operating systems — hackers could deface the content on the site or, worse, install malicious software on the machine running it. Attacks can run the gamut from the installation of programs intended to co-opt system resources to keystroke loggers and scanning software designed to purloin user information and passwords. Some hackers target university systems to steal computing resources or even intellectual property — proprietary compounds, instrument designs, patient data and personal communications. Hacking is suspected in the November 2009 release of e-mails from the Climatic Research Unit at the University of East Anglia, UK, which resulted in a global crisis of confidence in climate science.

"There is no sector that has been able to withstand this onslaught of intrusions," warns Steven Chabinsky, deputy assistant director in the cyber division of the Federal Bureau of Investigation (FBI) in Washington DC.

Protecting research data presents particular challenges. Most information-technology (IT) professionals suggest ensuring that large or sensitive data stores are managed by a centralized IT team that can monitor and administer systems, keeping a close watch over traffic and limiting access. But this can conflict with the ethos of researchers who need such systems to be accessed by a wide variety of students, postdocs and collaborators. "Universities tend to be fairly open kinds of environments," says John McCanny, lead scientist at the Centre for Secure Information Technologies at Queen's University Belfast, UK. And some researchers bristle at the idea of losing control. "I'd rather be insecure and free," says Phillip Zamore, a biologist at the University of Massachusetts Medical School in Worcester.

> **"Researchers are the best in the world at basically ignoring the administration."**

Drive-by hacking

It is a rough world in which to take such a stance. Michael Corn, chief privacy and security officer at the University of Illinois at Urbana-Champaign, estimates that each day his university's firewalls block some two million to three million scans — essentially Internet drive-bys looking for open communications ports. A "significant percentage" of these, he says, are likely to stem from "professionals in the employ of organized crime or possibly state actors". The network at the San Diego Supercomputer Center of the University of California, San Diego, logged an average of 27,000 intrusion attempts per day during the first three months of 2010, according to Dallas Thornton, division director of the university's cyberinfrastructure services. Thornton says that this number actually underestimates the problem, because many common attacks are simply ignored.

Successful intrusions are rarer. Until recently, the Educational Security Incidents website (www.adamdodge.com/esi) reported breaches in cybersecurity for higher education, recording nearly 500 incidents of data theft, loss or exposure between July 2005 and December 2009. Most of these occurred in the United States; just 10 occurred in the United Kingdom, 11 in Canada and 8 in Japan. But that may be just the tip of the iceberg; most incidents go unreported, says Chabinsky. McCanny reports that cybercrime's growth in the United Kingdom has greatly outpaced expectation. "The train is coming down the track much faster than most people would have anticipated even 15 months ago," he says.

EDUCAUSE, a non-profit organization of higher-education IT professionals, undertakes annual surveys of academic chief information officers. The 2009 survey ranked security third overall among IT problems facing higher-education institutions today[2]. Security officers in such institutions face two crucial problems. One is funding, which ranked first in the survey. The chemistry department at the University of Wisconsin-Madison, for instance, has one IT person overseeing 1,300 computers used by 500 staff and several thousand students. In

such a climate, says departmental chairman Robert Hamers, "it's just very limiting in terms of what you can actually expect one person to do". Last year, Hamers' department suffered a breach. In the absence of a firewall (now installed), foreign hackers infiltrated 40 computers, installing software that turned those systems into file-sharing servers for copyrighted music, movies and television programmes.

The other problem, says Alan Paller, director of research at the cybersecurity-focused SANS Institute in Bethesda, Maryland, is researcher independence. Academic researchers need to be able to install software on demand, to collaborate with colleagues, to develop new tools for public or private consumption, and to cater for an ever-changing and heterogeneous user base. Besides that, "researchers are the best in the world at basically ignoring the administration", says Paller. In 2006, a researcher at Georgetown University in Washington DC independently decided to migrate a server handling confidential patient data from an IT-monitored UNIX system to an unmonitored Windows one. The server was hacked, and the names, birth dates and social-security numbers of some 41,000 individuals may have been accessed. The university was legally obliged to notify each of them.

Such incidents, says Brian Voss, chief information officer at Louisiana State University in Baton Rouge, mostly occur when researchers strike out on their own. "It's like herding cats," he says. "The challenge is, how do you get them all under the umbrella?"

Generally, universities do that by promoting the benefits of their centralized services. Many academic institutions offer a battery of common but effective defences — everything from pushing operating system and antivirus patches out to users, to remotely monitoring network traffic, to the establishment of secure virtual private networks for encrypted communication and virtual machines. Virtual machines are hardware surrogates: although users can interface as if they are working on physical computers or web servers, they're actually just working through windows running on a centralized host computer.

Gerstein says that most of his lab is set up on a framework of virtual machines. This allows him and his colleagues to maintain back-ups for all the resources they put up on the web. "If something gets hacked into, we are not as worried any more," he says, "because we can roll back really easily and then put the thing back up."

The American University in Washington DC also offers virtual-machine services to its faculty members, says chief information officer David Swartz, who (with Voss) co-chairs the EDUCAUSE Higher Education Information Security Council. "If they need a server I say, 'Come to me, I'll set you up'," Swartz says. The resulting sites are professionally secured, continuously monitored, and easily backed up and restored. And in the event of a breach, that intrusion is often effectively contained by virtue of the virtual machines' isolated architecture.

Cybercrime is evolving rapidly. At the University of Nevada in Las Vegas, Lori Temple, vice-provost for information technology, says she is seeing increased intrusion through social-networking sites. Facebook games and images, for example, are particularly problematic, she says. "As much as you try to be ahead, you are almost always one step behind."

How much individual researchers need to worry depends on their own cost–benefit analysis, says Chabinsky — a reflection of the value of the data's confidentiality and integrity both to the researcher and his or her competitors. Many will calculate that they have little cause for concern, save the loss or corruption of their own data, which can be mitigated by routine back-ups. But for those who must store personal, confidential or economically valuable information, he says, "they should ask some strong questions of their system security providers". They should enquire, for example, about the confidence they have in the systems in place and what additional measures they can take to mitigate risks.

"The best way to secure a computer is to remove the keyboard, then the human cannot make a mistake."

This is especially important as US federal granting agencies are beginning to require certification that sensitive information such as health records will be secured in compliance with the Federal Information Security Management Act of 2002. "An incident at the university could cause more than public embarrassment or a lawsuit," says Swartz, "it might actually cause the loss of federal grants."

Ultimately, preemptive training might be the most cost-effective weapon IT departments have. Knowledge, after all, is power (see 'Ten tips for cybersecure science'). Especially given that no amount of infrastructure can overcome user carelessness. In 2007, the University of Illinois at Urbana-Champaign suffered a breach in which a spreadsheet containing the personal information of 5,247 students was accidentally e-mailed to 700 unauthorized individuals.

As Marty Lindner, a principal engineer in the computer emergency response team (CERT) at the Carnegie Mellon Software Engineering Institute programme in Pittsburgh, Pennsylvania, says: "The best way to secure a computer is to remove the keyboard, because then the human cannot make a mistake." ■

Jeffrey Perkel is a freelance writer in Pocatello, Idaho.

Ten tips for cybersecure science

DO enable automatic operating-system updates.

DO install and update your antivirus and anti-malware software, most of which is available for little or no cost from universities.

DON'T run your computer with administrator privileges, but as a non-privileged user. Then, if somebody does hack into your computer, they cannot install anything.

DO consider purging sensitive data from connected computers and confining them to offline machines.

DO encrypt your hard drive, for instance with FileVault (Mac), TrueCrypt (Windows/Mac/Linux) or PGP Whole Disk Encryption (Windows/Mac/Linux).

DON'T send sensitive data by standard e-mail. If you're not using encrypted e-mail, encrypt the material itself, for instance in a password-protected PDF.

DO ensure all your applications are patched to the current level, for instance with Secunia's free Personal Software Inspector (Windows).

DO password-protect your computer and smartphone.

DON'T let your web browser remember your passwords; instead, use password vaults, such as KeePass and LastPass, which store them in encrypted databases.

DO use strong passwords — or better, passphrases — that include both upper- and lower-case letters, numbers and symbols. Change passwords regularly, and don't use the same one for everything.

1. Smith, A., Greenbaum, D., Douglas, S. M., Long, M. & Gerstein, M. *Genome Biol.* **6,** 119 (2005).
2. Agee, A. S. *et al. EDUCASE Rev.* **44,** 44–59 (2009).

See Editorial, page 1246.