Pipeline
Technology for Service Providers.
www.pipelinepub.com Volume 9, Issue 2

Big Data, Big Brother, and You

By Jesse Cryderman

Right now you are being monitored.

Your data is being gathered. A record of your visit to this page has been created. Your phone and text records sit on a hard drive next to your location data, and transaction logs and tollbooth tallies are available for access. If you use any "free" online services or social networking such as Gmail, YouTube or Facebook, you have generated great volumes of personal profile information that is bought and sold without your knowledge. If you live in many metropolitan cities, a photographic history of your red light violations has been stored. Your medial records have been digitized just like your cable bill. Your IP address can be easily gleaned, your GPS signal tracked, and from that, a satellite image of your location, accurate within feet, can be created.

This much is indisputable. The question is, how did we get here and what happens next?

What Warrant?

In the United States, conducting electronic surveillance has historically faced two hurdles: legal protections afforded by the Fourth Amendment to the U.S. constitution, and the cost and complexity required to monitor and collect data on a person of interest. As a result, any wiretapping or electronic tracking activities required a warrant

Over the past two decades, though, four factors have completely changed this framework lowering the hurdles to mere bumps in the road:

- The Communications Assistance for Law Enforcement Act (CALEA)
- The national tragedy of September 11, 2011 and subsequent executive orders (Patriot Act)
- The Foreign Intelligence Surveillance ACT (FISA) Amendments Act of 2008 (FAA)
- A massive increase in the amount of voluntarily shared personal data

Additional intersection points have arisen as well that could be cause for concern. The development

of the next-generation 9-1-1 (NG9-1-1) system includes provisions for the government to shut down communications and utilities; burgeoning smart grid initiatives making the proxy control of these facilities even easier. But the single biggest factor that has altered the data collection and privacy landscape is the U.S. Government's response to the September 11 attacks. In the interest of national security and the prevention of future attacks, executive orders have been issued by the last and current presidents that authorize warrant-less electronic data collection. With the stroke of a pen, the Patriot Act authorized the National Security Administration (NSA) to monitor, without search warrants, phone calls, internet activity, text messaging, and all other communications.

This surveillance, however, wouldn't be possible without an infrastructure to support it and the participation of communications service providers (CSPs). This groundwork was laid back in 1994 when then-President Bill Clinton signed the Communications Assistance for Law Enforcement Act (CALEA) act into law. The language of CALEA outlines its intent:

"To amend title 18, United States Code, to make clear a telecommunications carrier's duty to cooperate in the interception of communications for Law Enforcement purposes, and for other purposes."

CALEA sets forth mandates that require CSPs and their vendor partners to design their infrastructure with federal monitoring in mind, and permits the surveillance of telephone, broadband internet, and VoIP traffic. This means if a warrant is issued, CSPs must be able to provide investigators with a variety

The NSA's crown jewel is being built in Utah: a $2 billion data center.

of customer records. There are several solutions that CSPs turn to for CALEA compliance. There are on-premise hardware solutions like forensics appliances from Solera Networks and Cisco, application-based solutions like NetSentry, and trusted third party solutions like Com Net. There is even an open source project designed to help smaller ISPs meet CALEA compliance without incurring large costs.

What about Skype, Facebook Chat, or WhatsApp? As it's currently worded, CALEA doesn't apply to over-the-top (OTT) communications service, and there hasn't been a formal proposal made to Congress to include such data collection. Nevertheless,"the Department of Justice is complaining that they are going dark, and losing the ability to intercept communication," says Mark Rumold, Legal Fellow with the Electronic Frontier Foundation (EFF). "The Administration is pushing to expand CALEA; what they would have to do is build a backdoor in the code. When you build a backdoor for the FBI, it's not just for the FBI, it's for everybody. Laws that mandate a backdoor threaten everyone's security."

Compared to collecting and storing data for future analysis, monitoring all communications data in real-time is a much larger undertaking that requires a massive amount of horsepower. The monitoring

systems are barely ten years old and the data storage facilities are still being built. Real-time monitoring of all communications would require the ability to break encrypted messaging from third-party services, such as Skype, and/or backdoors in the code of various messaging services. Under current law, CSPs are only required to provide plain-text content of encrypted messages for which they (the CSP) provide the encryption key. In other words, Verizon will not be found in violation of CALEA if it cannot comply with a warrant request because certain traffic was Skype-based encrypted communication (encryption that is currently uncrackable). Similarly, APIs that enable various forms of real-time communication are not built with a backdoor for government investigators. On this topic, both privacy and business advocates have offered strong opposition; it's bad for both innovation and privacy if app developers were mandated to build surveillance portals for the Feds.

Currently, the FBI operates the Digital Collection System Network (DCSNet), a point-and-click wiretapping platform that can perform instant surveillance on any device in the U.S. It's like TiVO for spooks. At inception in 2003, the platform ran on Sprint's Peerless IP fiber network, which is immune to attack as it is not connected to the public internet. In 2007, news broke that the FBI had contracts with AT&T, MCI, and Verizon. Several large-scale telecom hardware vendors provide the monitoring apparatus used for the FBI's data collection efforts. Disturbingly, the EFF has discovered that some of these same vendors provide hardware to repressive regimes (Egypt, China, Oman), and their use has a decidedly negative impact on human rights.

Interestingly, newer service providers trend stronger on the transparency side, with Sonic.net at the top.

The NCS, on the other hand, collects data all over the world, and operates several data centers across the United States where the aggregate data is stored. Today, the crown jewel is being built in Utah: a $2 billion data center in Bluffdale, due to be completed in September of 2013. So what kind of tech does $2 billion buy in the desert?

One Data Center to Rule Them All

Few adjectives capture the magnitude of the NSA's latest data center build in Utah. Situated on a 240-acre site, the Bluffdale facility's footprint will stretch 1 million square feet, twice the size of the US Capitol Building and 18 times larger than the White House. The goal of the facility is to collect, "all forms of communication, including the complete contents of private emails, cell phone calls, and Google searches, as well as all sorts of personal data trails—parking receipts, travel itineraries, bookstore purchases, and other digital "pocket litter," says NSA expert James Bamford.

Basically, the Bulffdale facility is big enough and powerful enough to capture all things digital. To do this, it must have as-yet-unseen technological capabilities in the world of computing. To achieve its goals, the Utah data center will have to capture data on the order of yottabytes. A yottabyte is equal to

one quadrillion gigabytes; at the beginning of 2012, storage networks could only handle one-thousandth of a yottabyte. Gobbling up yottabytes takes a lot of energy—65 megawatts to be exact. This will push the annual electricity bill for the data center upwards of $40 million.

While the NSA claims it is not targeting average citizens, the technical ability to do so is clearly in place. Unlike traditional law enforcement, which has a level of transparency in terms of operation, the NSA operates outside the boundaries of most inquiry; this makes it impossible to know exactly what data is being collected and stored. According to undisputed evidence in Jewel v. NSA, however, AT&T telecommunications technician Mark Klein revealed that AT&T routed copies of internet traffic to a secret room in San Francisco controlled by the NSA. What is most likely is that all digital data will be stored for future reference. In other words, in the event of an investigation, the government will have retroactive access to all data on a person of interest, even if they are a U.S. citizen.

The Dilemma for CSPs

CSPs are in an unenviable position, forced to balance their customers' privacy concerns and trusted customer relationships against federal data collection mandates and "strong suggestions" from the NSA. If a CSP is working with the NSA, that fact will be classified, meaning by law, the company cannot comment on its involvement with the NSA. Currently, CSPs retain all types of personal data on subscribers that may or may not be shared. How much is shared with the government is anyone's guess; these are questions that service providers, naturally, are unwilling to discuss. Still, CSPs advocate for their customers' privacy in several ways; the Electronic Frontier Foundation pored over available data to create an overview of the companies' efforts in 2012 called, "Who Has Your Back." In the chart below, you can see how some select service providers fare in terms of advocacy and transparency.

Interestingly, newer service providers trend stronger on the transparency side. One standout was Sonic.net. This ultra-high-speed ISP out of California offers its customers a very high level of both transparency and advocacy. What makes them unique in a sea of similar companies isn't quite clear, but it proves that a balance can be struck between surveillance pressures and customer privacy The achievement of this balance can also be effectively marketed and monetized via increased customer loyalty and reduced churn.

Wireless providers could grant their subscribers access to their data records, perhaps as a premium service.

	Informs Users of Data Demands?	Transparency re: Government requests?	Fights for user privacy in courts?	Fights for user privacy in Congress?
AT&T	No	No	No	Yes
Verizon	No	No	No	No
Comcast	No	No	Yes	No
Google	Partial	Partial	Yes	Yes
Apple	No	No	No	Yes
Microsoft	No	No	No	Yes
Skype	No	No	No	No
Facebook	No	Partial	No	No
Dropbox	Yes	Yes	No	Yes
Twitter	Yes	Partial	Yes	Yes

Free Services—May Come With a Cost

Beyond all of these factors, the numerous methods that enable digital consumers to share vast amounts of personal information have exploded. In order to enjoy the benefits of no-cost social networking, email, storage, VoIP, and internet services, consumers have unveiled an unprecedented amount of private data to their CSP, social networks, and the public at-large.

On the positive side, this has created more possibilities with the sharing of ideas and information. This data can also be leveraged by CSPs to improve marketing efforts and create personalized experiences that enhance a customer relationship rather than diminish it. On the other hand, the amount of publicly available data on individuals is downright scary. You don't need a $2 billion data warehouse in Utah and a Patriot Act to snag personal details and a home photo of anyone with a phone number. Where traditional research falls short, Facebook has proven a treasure trove for news reporters and journalists. As long as consumers use no-cost services that are subsidized by data

collection and monitoring, little will ever change.

A Dark Future?

Does all of this mean we are headed for a Orwellian future? And if so, can or should CSPs help us avoid it? It's important to note that while some of the tools may be in place for the theoretical real-time monitoring of U.S. citizens, there are many groups fighting hard to ensure that we never reach this reality, including the EFF, the ACLU, and numerous CSPs. The EFF has initiated numerous lawsuits against the NSA; one of the most recent focused on GPS location data. "The court came to the right conclusion," said Mark Rumold."They decided that the installation of the GPS device was a search under the fourth amendment, but held off on the more difficult question of whether tracking someone's movement is a search within the meaning of the fourth amendment." Likewise, the ACLU has challenged the NSA in court, business advocates have argued ardently against mandates that would undermine innovation and privacy , security experts have sounded the alarm regarding the potential abuse of surveillance backdoors, and CSPs have joined the Digital Due Process coalition to fight for privacy rights in Congress.

The widespread popularity of non-traditional communication is a blessing in disguise for consumer privacy as well. OTT messaging and communication is currently outside the bounds of CALEA, encrypted messaging can't be cracked (unless it originates from a CSP, i.e. text messaging), and there aren't backdoors built into the code that provide for surveillance. The safest way to communicate right now would be over an encrypted connection provided by a third-party service (not an ISP or CSP), or over an encrypted OTT service. Building real-time surveillance (not simply collection) abilities into network hardware and communication software would create such pervasive security vulnerability--the same kind of vulnerabilities that have prompted investigations into ZTE and Huawei--that is unlikely we will see this type of development anytime soon.

Data collection and monitoring is big business, but as the EFF opined in a white paper in April of 2012, "The issue is complicated because most of these technologies are 'dual use.' This means that along with the ability to facilitate human rights abuses, nearly all of these technologies can be used for legitimate purposes." The same technology that can be used to launch un-warranted surveillance can also be used to create better network and computer security, enable valuable research, and better protect citizens. It's fundamentally different if Google has personal communications data stored, than if the NSA has communications data stored. This is why it's crucial that consumers and their service providers (and their vendors) understand who has access to personal and communications data and understand exactly how that data is being used.

In light of these developments, some interesting business cases arise to manage privacy, deliver enhanced business services, automate transparency for subscribers, and ensure greater levels of privacy and online anonymity. Perhaps the increased level of data collection will drive a market boom in private networking and private Cloud. Wireless providers could grant their subscribers access to their data records, perhaps as a premium service. CSPs or technology providers could offer "privacy-as-a-service" solutions. And transparency of business practices should be the cornerstone for all CSPs who wish to improve trust and customer loyalty. This transparency could extend not only to public tracking of government requests (similar to what Google does), but also transparency in vendor selection. Huawei, for instance, is being investigated by the U.S. government for building backdoors into its telecom gear to enable Chinese espionage. Ironically, some of the same companies that refuse to buy network gear from Huwaei currently purchase gear from vendors who allegedly supply surveillance equipment to the Chinese--hardware that is used to repress the citizenry and maintain the "Great Firewall of China." Vendor companies that maintain best practices in terms of human rights and privacy could create a coalition to leverage this position. The telecommunications industry as a whole could create a certification, similar to "green" certifications for environmentally friendly business practices, that would indicate positioning on customer privacy. There are many ways for the industry to advocate for security and a better network through data collection without undermining privacy that is guaranteed by the U.S. Constitution.