

The Washington Post

THURSDAY, OCTOBER 31, 2013

NSA taps Yahoo, Google links

DATA CENTERS
AT ISSUE

Agency has access to
millions of accounts

BY **BARTON GELLMAN**
AND **ASHKAN SOLTANI**

The National Security Agency has secretly broken into the main communications links that connect Yahoo and Google data centers around the world, according to documents obtained from former NSA contractor Edward Snowden and interviews with knowledgeable officials.

By tapping those links, the agency has positioned itself to collect at will from hundreds of millions of user accounts, many of them belonging to Americans. The NSA does not keep everything it collects, but it keeps a lot.

According to a top-secret accounting dated Jan. 9, 2013, the NSA's acquisitions directorate sends millions of records every day from internal Yahoo and Google networks to data warehouses at the agency's headquarters at Fort Meade, Md. In the preceding 30 days, the report said, field collectors had processed and sent back 181,280,466 new records — including “metadata,” which would indicate who sent or received e-mails and when, as well as content such as text, audio and video.

The NSA's principal tool to exploit the data links is a project called MUSCULAR, operated jointly with the agency's British counterpart, the Government Communications Headquarters. From undisclosed interception points, the NSA and the GCHQ are copying entire data flows across fiber-optic cables that carry information among

the data centers of the Silicon Valley giants.

The infiltration is especially striking because the NSA, under a separate program known as PRISM, has front-door access to Google and Yahoo user accounts through a court-approved process.

The MUSCULAR project appears to be an unusually aggressive use of NSA tradecraft against flagship American companies. The agency is built for high-tech spying, with a wide range of digital tools, but it has not been known to use them routinely against U.S. companies.

In a statement, the NSA said it is “focused on discovering and developing intelligence about valid foreign intelligence targets only.”

“NSA applies Attorney General-approved processes to protect the privacy of U.S. persons — minimizing the likelihood of their information in our targeting, collection, processing, exploitation, retention, and dissemination,” it said.

In a statement, Google's chief legal officer, David Drummond, said the company has “long been concerned about the possibility of this kind of snooping” and has not provided the government with access to its systems.

“We are outraged at the lengths to which the government seems to have gone to intercept data from our private fiber networks, and it underscores the need for urgent reform,” he said.

A Yahoo spokeswoman said, “We have strict controls in place to protect the security of our data centers, and we have not given access to our data centers to the NSA or to any other government agency.”

Under PRISM, the NSA gathers huge volumes of online communications records by legally compelling U.S. technology com-

The National Security Agency, working with Britain's intelligence agency (GCHQ), secretly taps into the internal networks of Yahoo and Google. The operation intercepts information flowing between data centers that those companies maintain around the world. In general, Google and Yahoo use privately owned or leased lines to synchronize their data centers. How the NSA and GCHQ break into those internal networks, using Google's as an example:

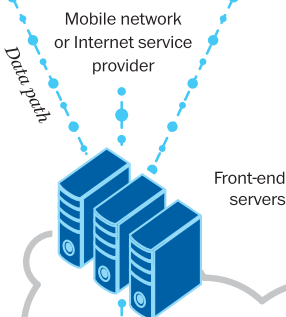
1 Public Internet

Internet and mobile users who send Gmail, create Google Drive documents or use other Google products over the public Internet typically do so via encrypted Web connections with Google.



2 Front-end servers

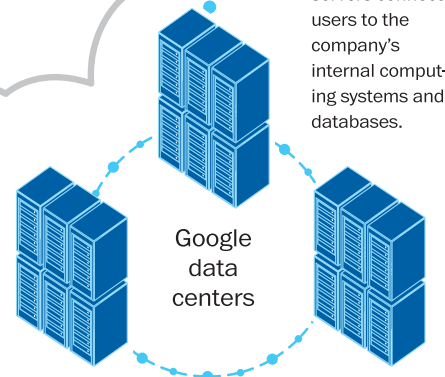
All Google requests are received by front-end servers that handle and process Web requests and return the data to the user.



3 Google's private cloud

Google's data centers, located around the world, are networks of computers linked by private fiber-optic cables.

Google's servers, in enormous buildings under 24-hour guard, synchronize their data and work together to balance traffic loads.



4 Infiltrating Google's private cloud

The NSA intercepts user account information as it flows among data centers. The precise collection points and methods are unknown. These are among the possibilities:

●●● Data flow ★ Point of data intercept

SCENARIO 1

Two data centers in separate geographic locations — wholly owned and operated by Google — are connected by Google-owned fiber, or by cables and network equipment managed and leased from a third party.

SCENARIO 2

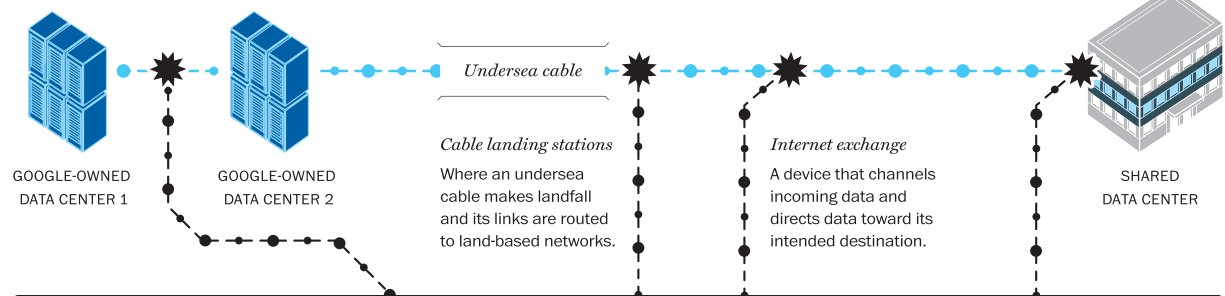
Google owns and operates major Internet connections, including some undersea cables, making them a primary Internet provider.

SCENARIO 3


Google also leases private links from global Internet providers that manage the network and its equipment, such as an exchange.

SCENARIO 4

Other times, Google's servers are housed in shared hosting facilities with other companies.



HOW THE NSA MAY SPY ON GOOGLE

 The NSA may have figured out ways to tap directly into Google's privately owned and managed internet links.

The NSA's British counterpart, the GCHQ, may have induced or compelled a third party — such as the operator of a cable landing station, a major Internet exchange or a data center that Google shares with other companies — to install surveillance equipment on Google's private cables.

panies, including Yahoo and Google, to turn over any data that match court-approved search terms. That program, which was first disclosed by The Washington Post and the Guardian newspaper in Britain, is authorized under Section 702 of the FISA Amendments Act and overseen by the Foreign Intelligence Surveillance Court (FISC).

Intercepting communications overseas has clear advantages for the NSA, with looser restrictions and less oversight. NSA documents about the effort refer directly to “full take,” “bulk access” and “high volume” operations on Yahoo and Google networks. Such large-scale collection of Internet content would be illegal in the United States, but the operations take place overseas, where the NSA is allowed to presume that anyone using a foreign data link is a foreigner.

Outside U.S. territory, statutory restrictions on surveillance seldom apply and the FISC has no jurisdiction. Senate Intelligence Committee Chairman Dianne Feinstein (D-Calif.) has acknowledged that Congress conducts little oversight of intelligence-gathering under the presidential authority of Executive Order 12333, which defines the basic powers and responsibilities of the intelligence agencies.

John Schindler, a former NSA chief analyst and frequent defender who teaches at the Naval War College, said it is obvious why the agency would prefer to avoid restrictions where it can.

“Look, NSA has platoons of lawyers, and their entire job is figuring out how to stay within the law and maximize collection by exploiting every loophole,” he said. “It’s fair to say the rules are less restrictive under Executive Order 12333 than they are under FISA,” the Foreign Intelligence Surveillance Act.

In a statement, the Office of the Director of National Intelligence denied that it was using executive authority to “get around the limitations” imposed by FISA.

The operation to infiltrate data links exploits a fundamental weakness in systems architecture. To guard against data loss and system slowdowns, Google and Yahoo maintain fortresslike data centers

across four continents and connect them with thousands of miles of fiber-optic cable. Data move seamlessly around these globe-spanning “cloud” networks, which represent billions of dollars of investment.

For the data centers to operate effectively, they synchronize large volumes of information about account holders. Yahoo’s internal network, for example, sometimes transmits entire e-mail archives — years of messages and attachments — from one data center to another.

Tapping the Google and Yahoo clouds allows the NSA to intercept communications in real time and to take “a retrospective look at target activity,” according to one internal NSA document.

To obtain free access to data-center traffic, the NSA had to circumvent gold-standard security measures. Google “goes to great lengths to protect the data and intellectual property in these centers,” according to one of the company’s blog posts, with tightly audited access controls, heat-sensitive cameras, round-the-clock guards and biometric verification of identities.

Google and Yahoo also pay for premium data links, designed to be faster, more reliable and more secure. In recent years, both of them are said to have bought or leased thousands of miles of fiber-optic cables for their own exclusive use. They had reason to think, insiders said, that their private, internal networks were safe from prying eyes.

In an NSA presentation slide on “Google Cloud Exploitation,” however, a sketch shows where the “Public Internet” meets the internal “Google Cloud” where their data reside. In hand-printed letters, the drawing notes that encryption is “added and removed here!” The artist adds a smiley face, a cheeky celebration of victory over Google security.

Two engineers with close ties to Google exploded in profanity when they saw the drawing. “I hope you publish this,” one of them said.

For the MUSCULAR project, the GCHQ directs all intake into a “buffer” that can hold three to five days of traffic before recycling storage space. From the buffer,

custom-built NSA tools unpack and decode the special data formats that the two companies use inside their clouds. Then the data are sent through a series of filters to “select” information the NSA wants and “defeat” what it does not.

PowerPoint slides about the Google cloud, for example, show that the NSA tries to filter out all data from the company’s “Web crawler,” which indexes Internet pages.

According to the briefing documents, prepared by participants in the MUSCULAR project, collection from inside Yahoo and Google has produced important intelligence leads against hostile foreign governments that are specified in the documents.

Last month, long before The Post approached Google to discuss the penetration of its cloud, Eric Grosse, vice president for security engineering, said the company is rushing to encrypt the links between its data centers. “It’s an arms race,” he said then. “We see these government agencies as among the most skilled players in this game.”

Yahoo has not announced plans to encrypt its data-center links.

Because digital communications and cloud storage do not usually adhere to national boundaries, MUSCULAR and a previously disclosed NSA operation to collect Internet address books have amassed content and metadata on a previously unknown scale from U.S. citizens and residents. Those operations have gone undebated in public or in Congress because their existence was classified.

The Google and Yahoo operations call attention to an asymmetry in U.S. surveillance law. Although Congress has lifted

some restrictions on NSA domestic surveillance on grounds that purely foreign communications sometimes pass over U.S. switches and cables, it has not added restrictions overseas, where American communications or data stores now cross over foreign switches.

“Thirty-five years ago, different countries had their own telecommunications infrastructure, so the division between foreign and domestic collection was clear,” Sen. Ron Wyden (D-Ore.), a member of the intelligence panel, said in an interview. “Today there’s a global communications infrastructure, so there’s a greater risk of collecting on Americans when the NSA collects overseas.”

It is not clear how much data from Americans is collected and how much of that is retained. One weekly report on MUSCULAR says the British operators of the site allow the NSA to contribute 100,000 “selectors,” or search terms. That is more than twice the number in use in the PRISM program, but even 100,000 cannot easily account for the millions of records that are said to be sent to Fort Meade each day.

In 2011, when the FISC learned that the NSA was using similar methods to collect and analyze data streams — on a much smaller scale — from cables on U.S. territory, Judge John D. Bates ruled that the program was illegal under FISA and inconsistent with the requirements of the Fourth Amendment.

bart.gellman@washpost.com

Soltani is an independent security researcher and consultant.