14-2985-cv

United States Court of Appeals

for the

Second Circuit

In the Matter of a Warrant to Search a Certain E-mail Account
Controlled and Maintained by Microsoft Corporation

MICROSOFT CORPORATION,

Appellant,

– v. –

UNITED STATES OF AMERICA,

Appellee.

ON APPEAL FROM THE UNITED STATES DISTRICT COURT
FOR THE SOUTHERN DISTRICT OF NEW YORK

BRIEF FOR *AMICI CURIAE* COMPUTER AND DATA SCIENCE EXPERTS IN SUPPORT OF APPELLANT MICROSOFT CORPORATION

PHILIP WARRICK
KLARQUIST SPARKMAN, LLP
One World Trade Center
121 S.W. Salmon Street, Suite 1600
Portland, Oregon 97204
(503) 595-5300
Counsel for Amici Curiae

TABLE OF CONTENTS

Page

TABLE OF AUTHORITIES

Page(s)

IDENTITY AND INTEREST OF *AMICI CURIAE*

The thirty-five *amici* identified below are computer and data science experts.[1] *Amici* have an interest in ensuring that the intersection between law and technology reflects an accurate awareness of the technology at issue and its real-life implementations. As professors who routinely research and teach computer science concepts, *amici* are well-positioned to provide a firm technological foundation for the resolution of the important legal disputes of this case regarding the storing and accessing of electronic data.

Amici are leading researchers in fields that include computer systems, networking, distributed systems, computer security, cryptography, and computer architecture—the foundations of cloud computing. They include members of the National Academy of Engineering and the National Academy of Sciences; winners of the Turing Award (the "Nobel Prize" of computer science); and Fellows of the American Academy of Arts & Sciences, the Association for Computing Machinery, the Institute of Electrical and Electronics Engineers, and the American Association for the Advancement of Science. While many have industry experience, all are now

[1] Pursuant to Fed. R. App. P. 29(c)(5), *amici* certify that no party's counsel authored this brief in whole or in part; no party or party's counsel contributed money intended to fund the preparation or submission of the brief; and no person other than *amici* or their counsel contributed money intended to fund the preparation or submission of the brief. Pursuant to Rule 29(a), all parties have consented to the filing of this brief.

faculty members at the leading computer science programs, including MIT, Stanford, Berkeley, Carnegie Mellon, Cornell, the University of Washington, Princeton, Georgia Tech, and Harvard, among others.

A list of *amici* appears below. *Amici* are signing this brief on their own individual behalf and not on behalf of any company, university, or other organization with whom they are affiliated.

Harold Abelson, Department of Electrical Engineering and Computer Science, *Massachusetts Institute of Technology*

Andrew W. Appel, Department of Computer Science, *Princeton University*

Steven M. Bellovin, Department of Computer Science, *Columbia University*

Matthew Bishop, Department of Computer Science, *University of California at Davis*

Avrim L. Blum, School of Computer Science, *Carnegie Mellon University*

Dan Boneh, Department of Computer Science, *Stanford University*

Douglas E. Comer, Computer Science Department, *Purdue University*

David L. Dill, Department of Computer Science, *Stanford University*

Edward W. Felten, Department of Computer Science and Public Affairs, *Princeton University*

Lance Fortnow, School of Computer Science, College of Computing, *Georgia Institute of Technology*

Shafrira Goldwasser, Department of Electrical Engineering and Computer Science, *Massachusetts Institute of Technology*

Allan Gottlieb, Computer Science Department, Courant Institute, *New York University*

J. Alex Halderman, Electrical Engineering and Computer Science Department, *University of Michigan*

Nadia Heninger, Computer and Information Science Department, *University of Pennsylvania*

Haym B. Hirsh, Department of Computer Science and Information Science, *Cornell University*

Daniel Peter Huttenlocher, Computer Science Department and Johnson Graduate School of Management, *Cornell University*

Brian Kernighan, Computer Science Department, *Princeton University*

Edward D. Lazowska, Department of Computer Science & Engineering, *University of Washington*

Henry M. Levy, Department of Computer Science & Engineering, *University of Washington*

Kathleen R. McKeown, Department of Computer Science, *Columbia University*

Nick W. McKeown, School of Engineering and Department of Computer Science, *Stanford University*

John Gregory Morrisett, School of Engineering and Applied Sciences, *Harvard University*

Asu Ozdaglar, Department of Electrical Engineering and Computer Science, *Massachusetts Institute of Technology*

David A. Patterson, Electrical Engineering and Computer Sciences Department, *University of California at Berkeley*

Vern Paxson, Electrical Engineering and Computer Sciences Department, *University of California at Berkeley*

William Worthington Pugh Jr., Department of Computer Science, *University of Maryland*

Jennifer Rexford, Computer Science Department, *Princeton University*

Ronald L. Rivest, Department of Electrical Engineering and Computer Science, *Massachusetts Institute of Technology*

Aviel D. Rubin, Department of Computer Science, *Johns Hopkins University*

Fred Schneider, Samuel B. Eckert Professor of Computer Science, *Cornell University*

Scott Shenker, Electrical Engineering and Computer Sciences Department, *University of California at Berkeley*

Eugene H. Spafford, Department of Computer Sciences, *Purdue University*

Philip Wadler, Laboratory for Foundations of Computer Science, School of Informatics, *University of Edinburgh*

James H. Waldo, School of Engineering and Applied Sciences, *Harvard University*

Dan S. Wallach, Department of Computer Science and Baker Institute for Public Policy, *Rice University*

INTRODUCTION AND SUMMARY OF ARGUMENT

Amici write to clarify certain technology underpinning "the cloud" as it pertains to this appeal. Cloud computing has revolutionized how we interact with data in the form of emails, photos, music, and other information now available at the touch of a button on a computer or smartphone located anywhere in the world. But, while "the cloud" has become a widely-used buzzword in recent years, many people have little idea what it is or how it works. Indeed, this confusion was aptly captured by comedian Amy Poehler in a recent national television advertisement in which she frantically asks an electronics store employee: "What's the cloud? Where *is* the cloud? Are we in the cloud now!?"[2] At least the first two questions posed by Ms. Poehler directly relate to the email technology at issue in this appeal.

Amici respectfully submit that the proper resolution of this appeal requires an understanding of certain fundamental points regarding the infrastructure and practices underlying cloud-based email. Thus, *amici* submit this brief to explain the significance of cloud computing and to clarify at least the following three points: (1) emails accessible "in the cloud" are stored in at least one identifiable *physical location*; (2) the "cloud" enables easier *access* to data, *not* new storage techniques;

[2] The commercial may be viewed at: https://www.youtube.com/watch?v=y9g5zaJ4bIM (last visited Dec. 10, 2014); *see also* Tim Robbins, *Are We in the Cloud Now?* (Feb. 6, 2013), *available at* http://creativemms.com/are-we-in-the-cloud-now/.

and (3) customer emails are *secured* as the confidential property of the account holder. These facts have important implications for cases like this involving customer emails accessible "in the cloud" but stored across international boundaries.

ARGUMENT

I. CLOUD COMPUTING HAS REVOLUTIONIZED HOW WE INTERACT WITH DATA.

Computing revolves around data—the vast quantities of ones and zeros (called bits) that collectively represent the photos, letters, spreadsheets, emails, and everything else we use computers to store, view, edit, and share. The advent of cloud computing has transformed how we interact with that data. No longer must we be tied to specific computers and physical storage devices. Instead, "the cloud" enables us to retrieve data from—and share it with—any device with an Internet connection. Thus, while cloud computing has been defined, by the government for example, as "a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction,"[3] the basic concept is actually quite simple—"cloud computing means storing and accessing data and programs

[3] Peter Mell & Timothy Grance, *The NIST Definition of Cloud Computing*, NIST Special Publication No. 800-145, at 2 (Sept. 2011), *available at* http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf.

over the Internet instead of your computer's hard drive."[4]

For many years, data storage was primarily local. Whether located inside a personal computer (e.g., a hard drive) or on portable devices that can be read by the computer (such as a USB "thumb drive" or a disc like a CD-ROM), people have stored (and in many cases continue to store) their data on local, physical media. For example, someone might draft a document using word processing software installed on her laptop computer and then save the document to the internal hard drive of the computer or a shared company server, and perhaps create a copy on a portable drive. The devices that store data have evolved over time from large sets of unwieldy "floppy disks" that were prone to decay and slow to store or retrieve data to fast, modern, solid-state storage devices that can fit on a keychain, have no moving parts, cost less than a fast-food meal, and have the capacity of thousands of archaic floppy disks. Corporate data storage has likewise evolved, with modern data servers consistently becoming smaller, faster, more efficient, and capable of storing immense amounts of data. These data servers traditionally housed the important files and communications of an entire company in an on-site server room or data center. Yet all of these devices relate to an increasingly outmoded paradigm—local data storage and retrieval.

[4] Eric Griffith, *What Is Cloud Computing?*, PC Mag. (Mar. 13, 2013), *available at* http://www.pcmag.com/article2/0,2817,2372163,00.asp.

Local data storage has significant drawbacks. Prior to the advent of the Internet, users could only access their documents, photos, and other files if they had access to the specific computer or physical media on which they had been stored. Thus, ensuring access to important information away from one's desk required physical access to the computer, storage device, or corporate network on which the data had been saved. And if two users wanted to share a file, they would have to exchange a physical disk or device or share access to a corporate network server computer. Perhaps more importantly, trusting one's data to physical media is only as secure and reliable as the media itself. Hard drives can and do fail, computers can and do crash, portable drives are easily misplaced, and critical data can be lost forever. Computers and storage media are also vulnerable to theft, which not only deprives the user of her data, but also exposes it to a stranger.

Storing data "in the cloud" mitigates many of these issues and presents an array of new opportunities.[5] While "the cloud" is merely an abstraction, it represents complex software and hardware that enable convenient access to remote data servers that house the data that formerly resided primarily on local storage devices.[6] Put

---

[5] *See* Antonio Regalado, *Who Coined "Cloud Computing"?*, MIT Tech. Rev. (Oct. 31, 2011), *available at* http://www.technologyreview.com/news/425970/who-coined-cloud-computing/ (describing the origins of "cloud computing").

[6] *See* Mell & Grance, *supra* n.3, at 2 n.2 ("The cloud infrastructure can be viewed as containing both a physical layer and an abstraction layer. The physical layer consists of the hardware resources that are necessary to support the cloud services

simply, the cloud permits access to data and applications from *any* location with an Internet connection. Web-based applications "in the cloud" allow users to access and manipulate their data essentially anywhere, no matter where it is physically stored by the service provider. Thus, with today's widespread broadband and wireless Internet access, computers (including not only traditional desktop and laptop computers, but also tablets and smartphones) are increasingly becoming mere terminals that serve as access points for remotely-stored data and remotely-executed programs.[7] Web-based email services like Microsoft's Outlook.com, for example, do not require any email software to be installed on one's local machine, and the emails themselves are stored on Microsoft servers and accessible through the Internet using a standard web browser like Internet Explorer.[8]

---

being provided, and typically includes server, storage and network components. The abstraction layer consists of the software deployed across the physical layer . . . .").

[7] *See* Jonathan Strickland, *How Cloud Computing Works*, HowStuffWorks.com (Apr. 8, 2008), http://computer.howstuffworks.com/cloud-computing/cloud-computing.htm (predicting that "[r]emote machines owned by another company would run everything from e-mail to word processing to complex data analysis programs. It's called cloud computing, and it could change the entire computer industry," and noting that "[i]f you have an e-mail account with a Web-based e-mail service like Hotmail, Yahoo! Mail or Gmail, then you've had some experience with cloud computing. Instead of running an e-mail program on your computer, you log in to a Web e-mail account remotely. The software and storage for your account doesn't exist on your computer – it's on the service's computer cloud.").

[8] *See* Christopher Barnatt, *A Brief Guide to Cloud Computing: An Essential Guide to the Next Computing Revolution* (2010) (noting that before "online e-mail service, such as Google's Gmail, Yahoo! Mail or Windows Live Hotmail [a predecessor to

This remotely-stored data is typically secured with advanced physical and electronic safeguards, access controls, and other technological security measures to prevent unauthorized access.[9] Storing data "in the cloud" also averts the risk of data loss, as the data is no longer stored on one's personal devices, but rather in physically-secure datacenters, which employ built-in redundancies (e.g., copies on multiple hard drives, possibly in multiple locations) to ensure against hardware failures and physical damage.[10] Cloud storage also permits concurrent access from

---

Outlook.com]" existed, "all e-mails were written in an e-mail application, such as Outlook Express, that was installed on the sender's computer. The message was then sent over the Internet and downloaded to the e-mail application installed on the recipient's computer. However, when e-mailing takes place between to people who use services like Gmail, Yahoo! Mail or Hotmail, their messages never leave the cloud. The e-mail software used to write and read the message is also never installed on either user's PC").

[9] *See, e.g.*, Microsoft, *Information Security Management System for Microsoft's Cloud Infrastructure* 1 (updated Feb. 2014), *available at* http://download.microsoft.com/download/A/0/3/A03FD8F0-6106-4E64-BB26-13C87203A763/Information_Security_Management_System_for_Microsofts_Cloud_Infrastructure.pdf (noting that "[h]osting such familiar consumer-oriented services as Outlook.com and Bing . . . means the company must adhere to numerous regulatory, statutory, and industry standards for securing personal and financial data").

[10] *See, e.g.*, Microsoft, *Cloud Operations Excellence & Reliability* 5 (2014), *available at* http://download.microsoft.com/download/E/3/0/E30B17E4-E70D-41E3-83E1-C22B767A76BC/Cloud_Operations_Excellence_Reliability_Strategy_Brief.pdf (noting that "[i]n the event of a natural disaster or service outage, we have programs, procedures, engineers, and operations experts in place to contain issues or rapidly recover with minimal impact on your organization"); *id.* at 2 ("[W]e are investing in developing greater application resiliency in our software so it will instantly recognize a disruption and gracefully fail over to a different set of servers or even a different datacenter, without interrupting the availability of the service.").

a variety of different devices and applications, thereby streamlining the distribution and sharing of data and facilitating collaboration. Moreover, applications hosted "in the cloud" may be updated by vendors with minimal user inconvenience. While beyond the scope of this brief, cloud computing also opens the door to many applications beyond email and data storage, including for example, on-demand streaming of music and video content.[11]

II. DATA HAS AN IDENTIFIABLE PHYSICAL LOCATION— "THE CLOUD" RELATES TO REMOTE DATA ACCESS, NOT A NEW WAY TO STORE DATA.

While "the cloud" is not a physical thing, data stored "in the cloud" does have at least one identifiable physical location. The cloud is merely an *abstraction* related to data *access*. The underlying data, however, is *stored using traditional physical media*, typically on hard drives in servers within large data centers like Microsoft's facility in Dublin, Ireland pictured below.[12]

---

[11] For additional examples of the broad range of services possible within the cloud computing environment, see Amazon.com, Inc.'s description of its various web services, available at: http://aws.amazon.com/products/ (last visited Dec. 10, 2014).

[12] Image from Microsoft, *Microsoft's Quest for Greater Efficiency in the Cloud* (Apr. 19, 2011), http://news.microsoft.com/2011/04/19/microsofts-quest-for-greater-efficiency-in-the-cloud/; *see also* Microsoft, *Cloud-Scale Datacenters* 2 (2014), *available at* http://download.microsoft.com/download/B/9/3/B93FCE14-50A2-40F6-86EE-8C1E1F0D3A95/Cloud_Scale_Datacenters_Strategy_Brief.pdf (noting that Microsoft "has invested over $15 billion in building a highly scalable, reliable, secure, and efficient globally distributed data center infrastructure"). Additional information regarding Microsoft's data centers, including a short video, is also available at: http://www.microsoft.com/en-us/server-cloud/cloud-os/global-

These data centers house thousands of server computers, all linked together as shown in the picture below to provide reliable and efficient data access.[13]

datacenters.aspx (last visited Dec. 10, 2014) (noting that Microsoft's cloud platform "infrastructure is comprised of a large global portfolio of more than 100 datacenters, 1 million servers, content distribution networks, edge computing nodes, and fiber optic networks").

[13] Image from Microsoft, *Microsoft's Cloud Infrastructure, Datacenters and Network Fact Sheet* (Nov. 2014), *available at* http://download.microsoft.com/download/8/2/9/8297F7C7-AE81-4E99-B1DB-D65A01F7A8EF/Microsoft_Cloud_Infrastructure_Datacenter_and_Network_Fact_Sheet.pdf.

These servers organize data in a structured database, which contains information permitting the server to store and retrieve the underlying data from the server computer's file management system, which in turn ultimately stores the data as ones and zeroes on magnetic or solid-state storage drives within the data-center servers.[14]

For example, assume you have an email account with Microsoft's Outlook.com service and you want to view an email message. You will likely either use a web browser like Microsoft's Internet Explorer or perhaps an email application on your smart phone to access one of the Outlook.com servers. Software on this server will validate that you are authorized to access the account—typically through a password or similar access control—and then reference a database to determine where your emails are stored. In some cases this database may also store certain information (called metadata) regarding individual email messages such as the sender, receiver, date, etc. Once the relevant data center location has been identified, the Outlook.com server requests that the data servers at that location retrieve the email content (i.e., the text of the email message). The data servers utilize their own structured database software, which has recorded where the relevant data was previously stored in the file-management system, to determine which file or files

[14] *See, e.g.*, Microsoft, *How Microsoft Designs Its Cloud-Scale Servers* 4 (2014) (describing the hard drives utilized in Microsoft's servers), *available at* http://download.microsoft.com/download/5/7/6/576F498A-2031-4F35-A156-BF8DB1ED3452/How_MS_designs_its_cloud_scale_servers_strategy_paper.pdf.

contain the requested content. The data servers then utilize the file management system software to determine where the individual ones and zeroes that constitute the email message are saved on the server's physical storage media (e.g., hard drives). The data is then read from the disk or other computer-readable storage mechanism and copied to the data center's external communication system and ultimately through the Internet to the Outlook.com server, which then forwards the complete email to your device for display on its screen.

Therefore, while web-based email content may pass through numerous communication channels and pieces of data infrastructure making up the Internet in the course of being retrieved and delivered to a user, *accessing email entails retrieving data from a physical location*. As explained above, email, like any other data, is stored using traditional, physical storage devices. To access email, the underlying data is retrieved from these devices and transmitted through the Internet to the user.

Thus, every email, photograph, or document stored "in the cloud" is in fact stored as a series of bits on at least one discrete physical storage device not unlike the hard drive in a personal computer. The difference is not that the data itself is virtual, but rather that one needn't be near the physical storage device to access the stored data. That is not to say that the data center's physical location is irrelevant, however. Although the Internet is very fast, it has not obviated geographic

considerations. To the contrary, choices about where to store data must take network latency into account. Latency is the delay between the time data is requested and the time it is delivered. While network latency is often measured in fractions of a second, these seemingly infinitesimal delays have dramatic effects. One study found, for example, "that a half-second delay causes a 20 percent drop in traffic on Google, and a one tenth of a second delay can lower Amazon's sales by 1 percent."[15] Retrieving data from multiple data centers generally involves greater latency (and thus more delay) than from a single data center, and requests for data from a nearby data center will generally result in significantly less latency (i.e., delay) than requests for data stored on the other side of the globe. *See In re Warrant to Search a Certain E-Mail Account Controlled and Maintained by Microsoft Corp.*, 15 F. Supp. 3d 466, 467 (S.D.N.Y. 2014) (Francis, Mag. J.) (citing Microsoft affidavits explaining that "because the quality of service decreases the farther a user is from the datacenter where his account is hosted, efforts are made to assign each account to the closest datacenter").

In sum, the cloud is merely the latest iteration of an information access revolution, which started with couriers, followed by telegraph, facsimile, and most recently the Internet. However, the cloud does not require any new storage

[15] Jelle Frank Van Der Zwet, *Layers of Latency: Cloud Complexity and Performance*, Wired (Sept. 18, 2012), *available at* http://www.wired.com/2012/09/layers-of-latency/.

mechanisms. Although data such as emails are more easily *accessed* by authorized parties, they still need to be *stored* on physical storage media in one or more servers in a datacenter as described above.

A. The Practice Of Load Balancing Does Not Make Data More Difficult To Locate.

In some advanced systems, multiple data centers cooperate to provide data access. For example, some data centers mirror data (i.e., store exact copies) across multiple locations to which user requests are distributed for efficiency purposes. This practice, called "load balancing," prevents single server computers from being overwhelmed by user requests and also permits redirection to alternate resources when a server computer suffers a failure (or "crash").[16] In another common scenario, secondary copies of data are saved to remote servers or data centers for disaster-recovery purposes (e.g., to preserve backups in the event of a natural disaster or terrorist attack).

In those scenarios, data has *multiple* physical storage locations. But this does not render the data more difficult to locate. To the contrary, to the extent the data

[16] *See, e.g.*, Citrix, *What is Load Balancing*, http://www.citrix.com/glossary/load-balancing.html (last visited Dec. 10, 2014) ("Load balancing is a core networking solution responsible for distributing incoming traffic among servers hosting the same application content. By balancing application requests across multiple servers, a load balancer prevents any application server from becoming a single point of failure, thus improving overall application availability and responsiveness. For example, when one application server becomes unavailable, the load balancer simply directs all new application requests to other available servers in the pool.").

has been replicated (copied) to multiple data centers, this serves only to facilitate the identification of *at least one* such location. Nor do these practices imply that email accounts, for example, are frequently transferred between various servers around the globe due to load balancing or other maintenance concerns. This is highly unlikely for at least two reasons. First, simply copying data to a new location does not remove the data from its initial physical location. And second, actually moving data between data centers at any frequent interval, particularly those in different parts of the world, would be inefficient and expensive and require bandwidth that could otherwise be used to satisfy customer requests.[17] As such, the data still has an identifiable physical location, even if that location occasionally shifts if and when the data is moved or when a secondary backup copy becomes a primary data resource following an equipment crash or natural disaster. After all, the service provider needs to be able to determine from where to retrieve data when requested by its customers.

B. Nor Do The Practices Of Sharding Or Partitioning Make Data More Difficult To Locate.

The terms "sharding" and "partitioning" refer to techniques for splitting a large database (but not necessarily individual database records) across several

[17] *See* David Patterson et al., *A View of Cloud Computing*, 53 Comm. ACM (Issue 4) 50-58 (2010).

computers (e.g., data servers).[18] Service providers typically utilize these techniques with very large sets of data, where splitting the database enables more computing power to be used to respond to queries and commands, and to distribute the data among multiple storage resources. Both Google and Microsoft, for example, distribute their massive indexes of the Internet (used for Google and Bing searches) over hundreds or even thousands of computers.[19]

These practices of sharding or partitioning data do *not* make it difficult or impossible to identify the physical location of a user's web-based email. As an initial matter, splitting a database into "shards" or "partitions" does not change the fact that each piece of underlying data has a physical storage location; it is the database as a whole that is partitioned.[20] Much like dividing a set of hard-copy encyclopedias into

---

[18] *See, e.g.*, Tony Morales, *Oracle Database VLDB and Partitioning Guide, 11g Release 1 (11.1)*, at 2-1 (2007), *available at* http://docs.oracle.com/cd/B28359_01/server.111/b32024.pdf ("Partitioning allows tables, indexes, and index-organized tables to be subdivided into smaller pieces, enabling these database objects to be managed and accessed at a finer level of granularity.").

[19] *See* Google, *How Search Works Handout*, http://static.googleusercontent.com/media/www.google.com/en/us/insidesearch/howsearchworks/assets/searchInfographic.pdf (last visited Dec. 10, 2014) (explaining that Google's Internet search index "is well over 100,000,000 gigabytes" and is distributed across "different data centers around the world").

[20] *See, e.g.*, Microsoft Developer Network, *Sharding Pattern*, *available at* http://msdn.microsoft.com/en-us/library/dn589797.aspx (last visited Dec. 10, 2014) (noting that a database administrator "can shard data based on the location of [users]. It may be possible to take the data for [users] in a specific geographic region offline

three subsets and storing each group of volumes in a different room of your house, the information as a whole has been distributed, but each individual encyclopedia entry has a discrete physical location within a particular volume in a specific room of the house. Moreover, the smaller and more geographically-dispersed the shards, the more the disadvantages outweigh the benefits of splitting the database.[21] Due to the network latency problems discussed above, it would be inefficient and highly abnormal for individual database records to be split between distant data centers.[22] Files for which sharding offers benefits are extremely large—unlike individual email messages, which represent very small pieces of data in modern computing. It would be very inefficient to break these small pieces of data into even smaller pieces, distribute them around the world, and then retrieve them to reconstitute the email message each time a user requests access. In sum, email providers have no incentive

---

for backup and maintenance during off-peak hours in that region, while the data for [users] in other regions remains online and accessible during their business hours.").

[21] *See, e.g.*, *id.* ("It can be difficult to maintain referential integrity and consistency between shards, so you should minimize operations that affect data in multiple shards.").

[22] *See, e.g.*, Jason Baker et al., *Megastore: Providing Scalable, Highly Available Storage for Interactive Services*, *in* 2011 Proc. Conf. on Innovative Data Sys. Res. (CIDR) 223, 225, *available at* http://research.google.com/pubs/pub36971.html ("To minimize latency, applications try to keep data near users and replicas near each other. They assign each entity group to the region or continent from which it is accessed most. Within that region they assign a triplet or quintuplet of replicas to datacenters . . . .").

to fracture and distribute email messages, which would reflect a very inefficient use of these techniques.

For that reason, both of the major web-mail vendors, Google and Microsoft, take similar approaches to sharding their web-mail databases. Both vendors divide their email databases by mailbox and then assign each mailbox to a particular datacenter or region. Microsoft employees have testified to this fact in affidavits submitted to the district court, and public Google documents indicate a similar approach. Specifically, Google's cloud-based Gmail data is stored using its Megastore storage model, in which "[t]o minimize latency, applications try to keep data near users," and more specifically, "[e]ach email account forms a natural entity group," which is assigned "to the region or continent from which it is accessed most."[23] In Microsoft's case, by way of further example, the precise server in the datacenter that hosts the active copy of the mailbox can be identified as can the precise server that hosts the replica of the mailbox which is used when there is a service outage impacting the active copy. *See In re Warrant*, 15 F. Supp. 3d at 468. As such, any specific user's mailbox has at least one specific physical location that

[23] *Id.* at 225 (describing Google's Megastore storage model); *see also* James C. Corbett et al., *Spanner: Google's Globally-Distributed Database*, *in* 2012 USENIX Symp. on Operating Systems Design & Implementation (OSDI '12) 251, 254, *available at* http://static.googleusercontent.com/media/research.google.com/en/us/archive/spanner-osdi2012.pdf (noting that Gmail is an example of "well-known Google applications that use Megastore").

can be readily identified. And the impracticalities of sharding or partitioning very small segments of data across geographically dispersed data centers mean that a given individual's email will generally be isolated to a particular region, if not a particular datacenter and server, regardless of the vendor. Thus, a web-mail vendor should likely have no difficulty identifying at least one discrete region in which it physically stores a customer's email data. Indeed, in this case Microsoft identified a data center in Dublin, Ireland as the physical storage location of the email content sought by the government. *See id.*

III. CUSTOMER EMAILS ARE TYPICALLY STORED SECURELY AS THE CONFIDENTIAL PROPERTY OF THE ACCOUNT HOLDER.

Unlike a company's own data records, an individual customer's email content is considered the private property of the email account holder.[24] As such, email providers typically employ significant security measures to ensure that only those authorized to access the email account may read, edit, or delete the contents of stored

[24] *See, e.g.*, Microsoft, *Protecting Data and Privacy in the Cloud* 8 (2014), http://download.microsoft.com/download/2/0/A/20A1529E-65CB-4266-8651-1B57B0E42DAA/Protecting-Data-and-Privacy-in-the-Cloud.pdf ("Microsoft believes that its customers should control their own information whether stored on their premises or in a cloud service. Accordingly, we will not disclose Customer Data to a third party . . . except as customers direct or required by law."); Google, *Your Security and Privacy*, https://support.google.com/a/answer/60762 (last visited Dec. 10, 2014) ("To put it simply, Google does not own your data. . . . Google does not share or reveal private user content such as email or personal information with third parties except as required by law . . . on request by a user or system administrator, or to protect our systems.").

messages. For example, login systems ensure that only authorized parties may access an electronic mailbox. Furthermore, email data may be encrypted or otherwise protected during transmission over the Internet, which prevents unauthorized parties from utilizing the data, even if it is intercepted in transit.[25] And data centers typically employ dozens of additional security measures including facility security, firewalls, intrusion detection systems, and many others.[26] These extensive security efforts comport with email providers' stated conviction that,

[25] *See, e.g.*, Matt Thomlinson, *Advancing Our Encryption and Transparency Efforts*, Microsoft on the Issues (July 1, 2014), http://blogs.microsoft.com/on-the-issues/2014/07/01/advancing-our-encryption-and-transparency-efforts/ (noting that "when you send an email to someone, your email is encrypted and thus better protected as it travels between Microsoft and other email providers"); Stephan Somogyi, *Making End-to-End Encryption Easier to Use*, Google Online Security Blog (June 3, 2014), http://googleonlinesecurity.blogspot.com/2014/06/making-end-to-end-encryption-easier-to.html (noting that Gmail "now always uses an encrypted connection when you check or send email in your browser" and discussing a new tool providing "end-to-end encryption," which allows "data leaving your browser [to] be encrypted until the message's intended recipient decrypts it").

[26] *See, e.g.*, Microsoft, *Securing the Microsoft Cloud Strategy Brief* 5 (2014), *available at* http://download.microsoft.com/download/D/5/E/D5E0E59E-B8BC-4D08-B222-8BE36B233508/Securing_the_Microsoft_Cloud_Strategy_Brief.pdf ("When we deploy a service to our datacenters, we assess and address every part of the service stack – from the physical controls, to encrypting data moving over the network, to locking down the host servers and keeping malware protection up-to-date, to ensuring applications themselves have appropriate safeguards in place.").

"[w]e believe your email belongs to you, not us, and that it should receive the same privacy protection as paper letters sent by mail–no matter where it is stored."[27]

CONCLUSION

Amici respectfully submit that the resolution of this appeal should take into account the fact that web-based email and other data stored "in the cloud" has at least one identifiable, physical location, and that the content of customer emails is securely stored as the confidential property of the account holder.

DATED: December 15, 2014

By: */s/ Philip Warrick*

Philip Warrick
KLARQUIST SPARKMAN, LLP
One World Trade Center
121 S.W. Salmon Street, Suite 1600
Portland, Oregon 97204
Telephone: 503-595-5300
Facsimile: 503-595-5301
philip.warrick@klarquist.com

Counsel for Amici Curiae

[27] Brendon Lynch, *Your Email Belongs to You, Not Us*, Microsoft Cyber Trust Blog (Aug. 27, 2014), http://blogs.microsoft.com/cybertrust/2014/08/27/your-email-belongs-to-you-not-us/.

CERTIFICATE OF COMPLIANCE

Pursuant to Rule 32(a)(7)(C) of the Federal Rules of Appellate Procedure, the foregoing brief is in 14-Point Times New Roman proportional font and contains 5,233 words and thus is in compliance with the type-volume limitation set forth in Rule 32(a)(7)(B) of the Federal Rules of Appellate Procedure.

DATED: December 15, 2014

By: */s/ Philip Warrick*

Philip Warrick
KLARQUIST SPARKMAN, LLP
One World Trade Center
121 S.W. Salmon Street, Suite 1600
Portland, Oregon 97204
Telephone: 503-595-5300
Facsimile: 503-595-5301
philip.warrick@klarquist.com

Counsel for Amici Curiae