Comparison of Availability Between Local and Cloud Storage

Leviathan Security Group

Introduction

The U.S. National Institute for Standards and Technology defines cloud computing as having five essential characteristics: on-demand self-service for provisioning, network access, the use of virtual machines to provide servers, rapid elastic provisioning and release of servers, and metering capability on the service provided.[1] This paper seeks to answer whether, from the perspective of availability of data, especially in the face of natural or manmade threats to data and access, cloud computing is able to provide a greater benefit than local data storage. For the purposes of this discussion, local data storage will include arrangements where data is kept on business premises or in a nearby datacenter, such as a colocation facility or other rigidly-provisioned resource. We will first examine the significant differentiator that cloud computing provides in this space, and then examine different scales of threats to availability to see whether local data storage or cloud data storage fares better against known historical threats.

The Cloud Differentiator: Geographic Redundancy

At its heart, cloud computing and cloud storage are exactly the same as computing and storage provided through a datacenter or SOHO server; the cloud is made of computers. The difference between cloud-based storage and local storage is that using cloud-based storage allows replication between separate geographic regions. Indeed, the key feature of cloud computing, from the perspective of maximizing availability and survivability of data, is replication for geographic redundancy. When incidents curtail availability from one datacenter (whether due to an incident occurring at a datacenter itself, such as a power issue, localized disaster, or interruption to the network connections that serve the datacenter), a correctly-used cloud provider should allow businesses to continue to access their data from other sources. As will be discussed below, using a cloud provider as though they were a local datacenter provides no protection from incidents that impact availability.

Using a cloud provider as though they were a local datacenter provides no protection from incidents that impact availability.

Classification and Examples of Availability Failures

Availability of datacenters can be impacted by events occurring at many scales. This section will endeavor to categorize and provide information on data impacts from several different types of incidents:

- Local and sub-local events: area of effect less than 2mi/3km radius (for instance, the City of London), marked as the red circle in Figure 1.

- Area events: area of effect less than 50mi/80km radius (for instance, Wales, or Beijing), marked as the green circle in Figure 1.

- Regional events: area of effect less than 300mi/500km radius (for instance, France, or an area equivalent to all of Central America), marked as the purple circle in Figure 1.

- National and supranational events: area of effect more than 300mi/500km radius, marked as the blue circle in Figure 1.

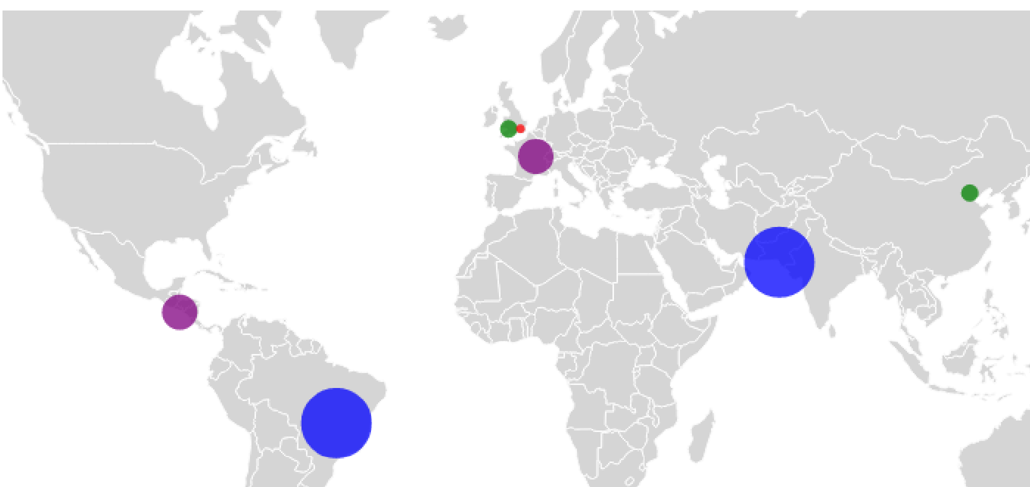

Figure 1: Relative Incident Size Visualization

Local and Sub-local Events

Events at this level include highly-localized weather events (for instance, flooding in a downtown area), as well as events occurring within a datacenter, such as power failures, failures in management systems, or other incidents not related to an external cause. Some examples include:

Shaw Building Explosion On July 12, 2012, a "small explosion" in a datacenter owned by Shaw Communications Inc. in downtown Calgary, Alberta caused widespread communications failures. The resulting datacenter failure caused landline phone service to 30,000 customers to fail (including 911 calls), intermittent cell phone outages in the area, shut down the Repac system (the system that coordinates emergency medical services in the city of Calgary, including ambulance-based computers and communication of hospital diversions), took an IBM-owned datacenter (that hosted services for both government and non-government customers) and office offline, and shut down the city's 311 information line service (which is ordinarily used, among other activities, to inquire as to when power and phone service will be restored in the event of an outage). It also shut down three radio stations, the local parking and fire authorities, and a taxi service, as well as some province-level government functions. Restoration of service took multiple days.[2,3]

Amazon Single-AZ Outages Amazon Web Services' US-East-1 "Availability Zone" (AWS' term for a group of datacenters in a very small area), located in northern Virginia, is its oldest datacenter, and is the cheapest for customers to use. As such, it is no surprise that many businesses that use AWS place instances in this group. As a side effect of this clustering, each time US-East-1 suffers an outage, a large number of services are rendered inaccessible. In September 2013, GitHub, Heroku, and other major Internet services were taken offline due to a networking issue.[4] Even more customers were taken offline in December 2012 due to a developer mistake related to Amazon's Elastic Load Balancing system,[5] in October 2012 due to a memory leak in a monitoring server,[6] and in April 2011 due to a networking mistake made during a migration.[7] These events are just a sample of those that this particular group of datacenters has suffered.

Amazon's Irish datacenter has also suffered critical failures, and like US-East-1, it is also a common choice for companies who do not wish to use multiple datacenters. In 2011, an alleged lightning strike to a major transformer close to Dublin-based datacenters for both Amazon and Microsoft caused damage to both datacenters. In Amazon's case, the outage corrupted Elastic Block Stores, requiring a very time-consuming recovery process.[8]

While any datacenter can suffer a failure, it is of particular note that none of these critical functions (for both public and private entities) had any resilience against a single point of failure; while in these cases a "small explosion" or "mistake" was the cause, a multitude of causes can shut down a datacenter temporarily. While there are some small benefits to ease of access and consumption of data through using a single datacenter of a cloud provider, as discussed below, the greatest benefits of cloud storage and computing only occur when customers take advantage of the geographic spread that providers offer, and ensure that there are no single points of failure in their applications (whether they be single computers, single network connections, or single buildings).

Area Events

Events at this level include large-scale incidents at the size of counties or smaller states in the US; these events cause impacts to facilities over a relatively widespread area (for instance, the size of a large metropolitan area. These will often be weather-related events, such as tornadoes, wildfires, or large-scale flooding. While there have not been recent major datacenter outages attributed to this type of event at these scales, this may change; as more datacenters are constructed in areas that are more prone to wildfires and tornadoes (and as those events increase in scope and move to new areas, as a result of global climate change), it is likely that datacenters might be in the path of the next major tornado event (such as that which struck Joplin, MO, and its surrounding communities in 2011) or wildfire.

Regional Events

Events at this level include major regional disasters, such as large-scale earthquakes or hurricanes; these events will sap recovery resources from an entire nation (or possibly surrounding nations).

Typhoon Haiyan, Tōhoku Tsunami, and Others Large-scale disasters happen throughout the world. The 2011 Tōhoku earthquake, and ensuing tsunami, damaged or destroyed Internet and telecom infrastructure throughout Japan, even in areas that still had power after the Fukushima meltdown; while many areas were able to continue some global Internet access, this was only possible due to redundant fiber-optic links, as many major fiber-optic lines were damaged or destroyed.[9] Typhoon Haiyan (also known as Typhoon Yolanda) critically damaged large areas of the Phillipines, destroying infrastructure of every type. Many other events of equivalent size and destruction occur throughout the world, but at the moment, the number of major datacenters in non-Global North countries is fairly low. As the the number of datacenters distributed around the world continues to increase, these types of events will take major Internet-scale services offline if those services are not designed to fail over to backups out of the region.

Superstorm Sandy During the October 2012 near-hurricane “superstorm” Sandy, massive flooding destroyed or took offline many datacenters throughout the US East Coast. Most of the datacenters taken down failed due to being unable to supply fuel to their generators when electricity failed, either due to running out of fuel (and being unable to obtain more), or due to the fuel pumps failing (often due to being under water).[10] [11] No one expects most commercial-grade datacenters to be able to survive disasters on this scale without any service interruption; that said, applications that rely on datacenters within a single region—a single US state, or even a single small country—will be unable to maintain functionality or access to data when natural disasters arrive. Since regional emergency services often use regional datacenters (being entities of governments that have an interest in investing in their local economies), it is particularly critical for these services also to put replicated, redundant computing power at a great distance from the areas they serve, and to invest in redundant data connections to connect the emergency services to the data needed in an emergency.

Hurricane Katrina The widespread damage of Hurricane Katrina hardly needs to be restated, including the cost of lives, the years of recovery, and the economic damage to the region. Somewhat less known is the number of datacenters that were taken offline during the disaster (and for weeks afterward); indeed, the only datacenter that stayed online, jointly operated by I-55, Data Protection Services, and Intercosmos Media Group, did so due to a combination of the services of one former US Special Forces operator, Michael Barnett, and the National Guard, which trucked fuel to the datacenter after they began using it as a communications link. Barnett was not an ordinary employee at the datacenter; in fact, he was hired for the specific event, and had no IT or computer security experience.[12]

The reporting on what capabilities this datacenter preserved is also illustrative as to services that were not designed with geographic redundancy; communications with the local Emergency Operations Centers, the National Guard, and area medical centers were destroyed, and only restored through the efforts of this one particular group (which had no obligation to help). This is particularly troubling, as these are services specifically needed in the event of a large-scale disaster. While ham radio support staff were able to support some critical operations through their ARES and RACES support agreements,[13] this type of assistance definitionally cannot aid commercial organizations, nor, even if it could, would it be anywhere near sufficient to preserve access to business data.[14]

It seems fairly clear that in the next large hurricane, businesses cannot rely on Special Forces-trained contract employees for one firm to volunteer their time to help others in the middle of a wide-spread natural disaster.

National/Supranational Events

These are events that curtail data access from an entire country or continent. While large-scale wars would fall into this category, so too would incidents that cut off Internet access to entire countries, such as the failure of submarine cables, or technical or political failures that cause whole-country routing to the outside world to be shut down. Crucially, due to the way in which many countries' infrastructure is designed, losing external connectivity may also destroy internal connectivity, causing even domestic network services to become unavailable.

The government of Egypt removed access to the global Internet as a tactic to quell protests against the government in 2011.[15] While volunteer groups like the Telecomix Crypto Munitions Bureau were able to pass some information into and out of Egypt during the blackouts through a combination of amateur radio and dialup modems,[16] this approach does not scale effectively to meet demand, nor is it designed to meet commercial data availability needs. There were reports that the government of Syria also shut down its Internet connections during protests in 2013,[17] but since that time, there have been allegations that the Internet blackout was not caused by the Syrian government.

While these Internet cuts may have been intentional, others have not; for instance, a shipping accident in the Suez Canal caused Internet outages to the majority of people and businesses in Pakistan, Egypt, India, Kuwait, Maldives, Lebanon, and Algeria in 2008.[18] Similar fiber cuts have occurred due to earthquakes in Taiwan and Southeast Asia.

Domestic-only traffic, and domestic-only clouds, are not necessarily a defense against outages caused by international traffic routes being lost. While in theory, an interruption to international Internet access would not affect domestic or regional routing, the truth is that many countries' domestic Internet traffic is routed outside the country. This affects countries large and small; for instance, many African and South American countries route all domestic traffic (going between separate Internet Service Providers) across North American or European nodes. Routing between neighboring countries uses long-distance routes even more often. One example is traffic between Brazil and Peru, which uses a submarine cable to reach Miami, FL, where it is routed across the United States before being routed down the West Coast back to Peru. For communications between Brazil and Argentina, data runs to Europe and back. Cutting external fiber-optic cables, then, destroys nearly all domestic and regional traffic, not just intercontinental traffic.[19] While plans now exist to add regional fiber lines within South America to fix this particular issue, it is a common situation for non-Global North countries—and maps of global fiber capacity, such as Cablemap Info[20], show that most countries have only one or two fiber-optic cables (mostly at single landing spots) connecting their traffic to the outside world. Backup links for critical traffic, such as satellite connections, are nearly always international in nature, with multi-corporation, multi-country consortia owning most geosynchronous communications satellites. In sum, the possibility of external connectivity outages does not counsel against using globalized cloud data storage.

Losing a country's Internet connectivity will often impact its domestic Internet, so the loss of external fiber will degrade or destroy internetworking within the country.

Cloud vs. Local: The Essential Qualities

Modern entities do not simply store data locally and use it locally; they communicate, process, and transfer data continuously. This applies to all sectors; as examples:

- Students in schools access learning resources, communicate with other institutions, and reach out to parents,
- Law firms exchange data with courts and governments, and
- Businesses in all lines of work exchange shipping manifests, customer information, medical records, and indeed almost any conceivable type of data as part of their business-critical processes.

A question that arises, then, is this: is local data storage, such as that provided by a local "normal" datacenter, something that is easier to access or less subject to availability concerns than a cloud-based storage solution? To find an answer to this question, we used ourselves as an example of a widely-distributed workforce. Leviathan has employees spread across the United States and Canada, which gives us a useful perspective on accessing both Leviathan's own datacenter resources (stored in a local colocation facility in the Seattle, WA area) and those of typical cloud vendors. To get a wider range of possibilities, we examined three major points of presence, from three different cloud vendors:

- Rackspace Inc. facility in Dulles, VA
- Linode, LLC facility in Fremont, CA
- Amazon Web Services US-West-2 facility in Oregon

We asked a sampling of our employees to run traceroutes[21] from their work locations (offices, homes, and coworking spaces) to computers in each datacenter; we then mapped the results, as shown in Figure 2.

Modern entities do not simply store data locally and use it locally; they communicate, process, and transfer data continuously.

The results were illuminating. From each Leviathan location, each of the major cloud platforms was just a few hops away, whether the

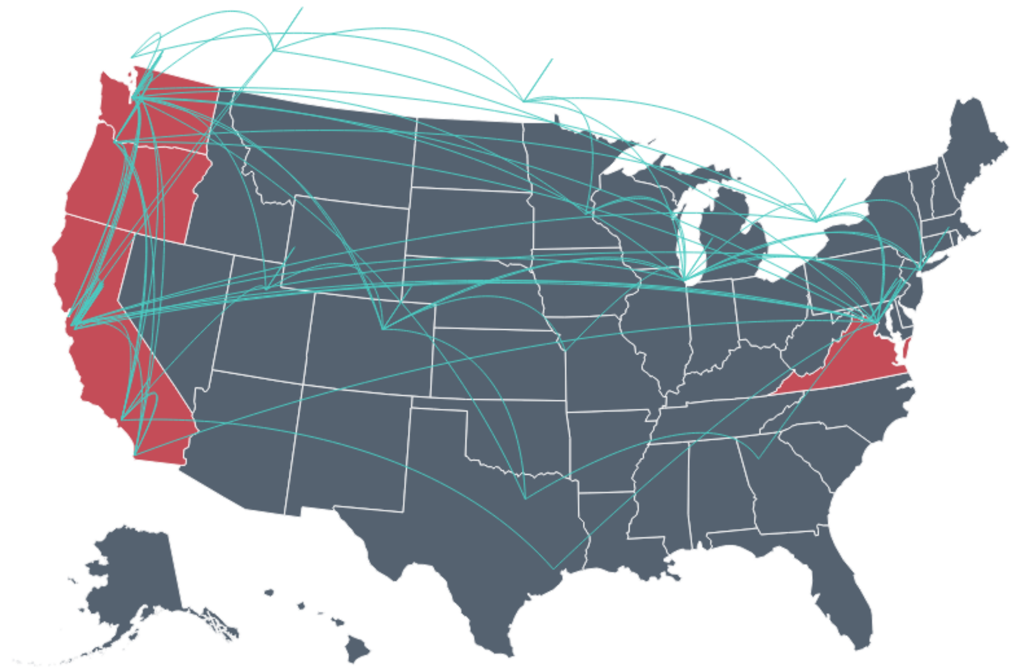

Figure 2: Leviathan National Paths in the US and Canada

facility was in the same area or not; this is likely due to the major cloud platforms' systems having nearly-direct connections to major peering points, where many networks come together. In Figure 2, there are four datacenters marked, including Leviathan's Washington-based colocation facility.

By contrast, a connection to the colocation facility, even from the Seattle office, required many more hops; communications ran from the office, to the ISP, to a small peering point, to a large regional peering point, back to a different small peering point, to the datacenter's ISP, and finally to the datacenter itself. This leads to the tangled web around Seattle shown in Figure 3. Ultimately, using a local datacenter provides no lower latency than using a cloud datacenter—even one across the country.

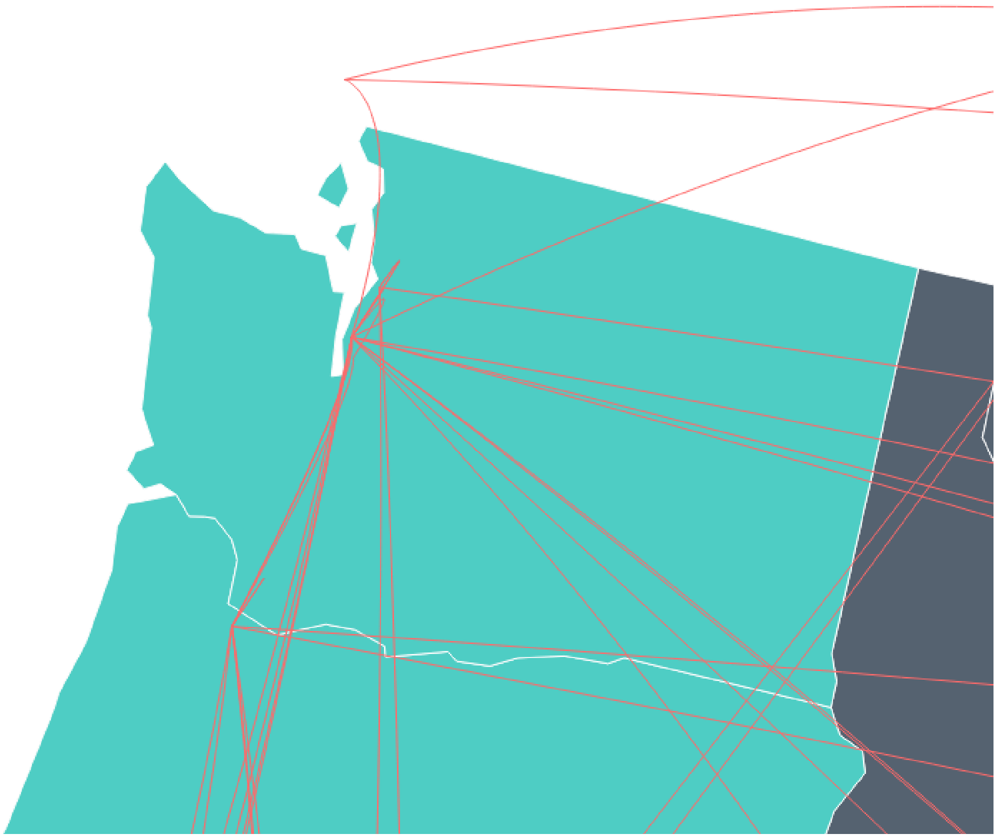

Figure 3: Leviathan Paths in Washington and Northern Oregon

Conclusion

Companies with large data storage needs have two options: use a local datacenter, or use cloud storage. When utilized properly, cloud storage gives companies the ability to use resources in different geographic regions to ensure high availability even in the face of local/area/regional incidents. Achieving this, however, requires taking advantage of geographical redundancy—ensuring that data is replicated not just across a city, but across a continent or an ocean. Many companies treat cloud providers like colocation facilities, storing all their data in a single region of a single cloud provider and relying on that facility to provide continuous access. This produces predictable results: this paper has discussed many situations where single-datacenter storage, despite "cloud" branding, caused substantial failures in availability.

This has important implications for restricting data storage to particular nations. Even when companies spread their data among multiple locations in a region, most countries—or continents—are simply not big enough to ensure data availability in the event of large-scale disasters. World-wide cloud storage allows data continuity and survivability when local datacenters, or single-nation clouds, would fail. As an additional benefit, data stored in these large-scale clouds will be closer in network terms to the point of data consumption, whether that's information transfer between entities, cloud computing, or simply ensuring that workers can access data quickly wherever they are.

The Internet was originally created to be able to route around damage to ensure continuity of information flow; it is critical that companies and governments not artificially restrict themselves from taking advantage of its possibilities.

The Internet was originally created to be able to route around damage to ensure the continuity of information. The capability exists to make data storage, not just communication, resilient in the face of large-scale threats; it requires only that companies and governments not restrict communications on the basis of geographic boundaries.

Notes

1. NIST Publication SP800-145, http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf

2. "Calgary 911 service affected by small downtown fire," CBCNews, July 11, 2012, http://www.cbc.ca/news/canada/calgary/calgary-911-service-affected-by-small-downtown-fire-1.1178503

3. "Explosion, Fire Cause Data Meltdown in Calgary," Wired, Caleb Garling, July 13, 2012, http://www.wired.com/2012/07/explosion-shaw-ibm/

4. "Amazon Web Services Stability and the September 13th US East 1 Outage", InfoQ, Chris Swan, September 20, 2013, http://www.infoq.com/news/2013/09/aws-east-outage

5. "Summary of the December 24, 2012 Amazon ELB Service Event in the US-East Region," AWS, https://aws.amazon.com/message/680587/

6. "Summary of the October 22, 2012 AWS Service Event in the US-East Region," AWS, https://aws.amazon.com/message/680342

7. "Summary of the Amazon EC2 and Amazon RDS Service Disruption in the US East Region," AWS, https://aws.amazon.com/message/65648

8. "Outage in Dublin Knocks Amazon, Microsoft Data Centers Offline," Rich Miller, Data Center Knowledge, August 7, 2011, http://www.datacenterknowledge.com/archives/2011/08/07/lightning-in-dublin-knocks-amazon-microsoft-data-centers-offline/

9. "In Japan, Many Undersea Cables Are Damaged," Om Malik, GigaOm, March 14, 2011, https://gigaom.com/2011/03/14/in-japan-many-under-sea-cables-are-damaged/

10. "Hurricane Sandy takes data centers offline with flooding, power outages," Jon Brodkin, Wired, October 30, 2012, http://arstechnica.com/information-technology/2012/10/hurricane-sandy-takes-data-centers-offline-with-flooding-power-outages/

11. "How one NYC data center survived Hurricane Sandy," Jon Brodkin, Wired, November 15, 2012, http://arstechnica.com/information-technology/2012/11/how-one-nyc-data-center-survived-hurricane-sandy/

12. "Hurricane Heroics," Wired, http://archive.wired.com/wired/archive/13.11/posts.html?pg=2

13. "Hurricane Katrina Amateur Radio Emergency Communications Relief Effort Operations Review Summary," Testimony before the Federal Communications Commission,

Gregory Sarratt, W4OZK, http://transition.fcc.gov/pshs/docs/advisory/hkip/GSpeakers060306/ACT1045.pdf

14. High-speed ham radio digital connections transmit at 19200 symbols per second; more typically, amateur radio communications transmit at 1200 symbols per second. These rates do not compare favorably to the 0.5-3 trillion bits (or more) per second sent over modern fiber-optic lines.

15. “Egyptians were Unplugged, and Uncowed,” Noam Cohen, The New York Times, February 20, 2011, http://www.nytimes.com/2011/02/21/business/media/21link.html

16. “Telecomix: tech support for the Arab spring,” Chavala Madiena, The Guardian, July 7, 2011, http://www.theguardian.com/technology/2011/jul/07/telecomix-arab-spring

17. “Syrian internet back after 19-hour blackout,” BBC, May 8, 2013, http://www.bbc.com/news/world-middle-east-22447247

18. “Fiber Optic Cable Cuts Isolate Millions From Internet, Future Cuts Likely,” Ryan Singel, Wired, January 31, 2008, http://www.wired.com/2008/01/fiber-optic-cab/

19. “South American Fiber Optic Ring,” Raul Zibechi, CIP Americas, April 12, 2012, http://www.cipamericas.org/archives/6734

20. http://www.cablemap.info

21. Traceroute is a standard networking tool that uses small packets to map communications pathways between hosts, such as between a client computer and an Internet server.

Project Team

- **James Arlen** - Director of Risk and Advisory Services - james.arlen@leviathansecurity.com
- **Lee Brotherston** - Security Advisor - lee.brotherston@leviathansecurity.com
- **Steve Manzuik** - Security Advisor - steve.manzuik@leviathansecurity.com
- **Brendan O'Connor** - Senior Security Consultant - brendan@leviathansecurity.com
- **Chad Thunberg** - Chief Operating Officer - chad.thunberg@leviathansecurity.com

Leviathan Security Group, Inc.
3220 1st Ave S, Suite 100
Seattle, WA 98134

p: 866.452.6997
f: 206.225.2004
e: contact@leviathansecurity.com

www.leviathansecurity.com

This work was commissioned by Google Inc., and created by Leviathan Security Group, Inc.
Google had no editorial control over this document.