Service Organization Controls (SOC) 3 Report

Report on the Google Cloud Platform System

Relevant to Security, Availability, Processing Integrity, and Confidentiality

For the Period 1 May 2016 to 30 April 2017
Management’s Assertion Regarding the Effectiveness of Its Controls Over the Google Cloud Platform System Based on the Trust Services Principles and Criteria for Security, Availability, Processing Integrity, and Confidentiality

30 June 2017

We, as management of, Google Inc. (“Google”) are responsible for designing, implementing and maintaining effective controls over the Google Cloud Platform System (System) to provide reasonable assurance that the commitments and system requirements related to the operation of the System are achieved.

There are inherent limitations in any system of internal control, including the possibility of human error and the circumvention of controls. Because of inherent limitations in Security controls, an entity may achieve reasonable, but not absolute, assurance that all security events are prevented and, for those that are not prevented, detected on a timely basis. Examples of inherent limitations in an entity’s Security’s controls include the following:

- Vulnerabilities in information technology components as a result of design by their manufacturer or developer
- Ineffective controls at a vendor or business partner
- Persistent attackers with the resources to use advanced technical means and sophisticated social engineering techniques specifically targeting the entity

We have performed an evaluation of the effectiveness of the controls over the system throughout the period 1 May 2016 to 30 April 2017, to achieve the commitments and system requirements related to the operation of the System using the criteria for the security, availability, processing integrity, and confidentiality (Control Criteria) set forth in the AICPA’s TSP section 100A, Trust Services Principles and Criteria for Security, Availability, Processing Integrity, Confidentiality, and Privacy. Based on this evaluation, we assert that the controls were effective throughout the period 1 May 2016 to 30 April 2017 to provide reasonable assurance that:

- the System was protected against unauthorized access, use, or modification to achieve Google’s commitments and system requirements
- the System was available for operation and use, to achieve Google’s commitments and system requirements
- the System processing is complete, valid, accurate, timely, and authorized to achieve Google’s commitments and system requirements
- the System information is collected, used, disclosed, and retained to achieve Google’s commitments and system requirements

based on the Control Criteria.
Our attached description of the boundaries of the Google Cloud Platform System identifies the aspects of the Google Cloud Platform System covered by our assertion.

Very truly yours,

GOOGLE Inc.
Report of Independent Accountants

To the Management of Google Inc.:

Approach:

We have examined management’s assertion that Google Inc. (“Google”) maintained effective controls to provide reasonable assurance that:

- the Google Cloud Platform System was protected against unauthorized access, use, or modification to achieve Google’s commitments and system requirements
- the Google Cloud Platform System was available for operation and use to achieve Google’s commitments and system requirements
- the Google Cloud Platform System processing is complete, valid, accurate, timely, and authorized to achieve Google’s commitments and system requirements
- the Google Cloud Platform System information is collected, used, disclosed, and retained to achieve Google’s commitments and system requirements

during the period 1 May 2016 through 30 April 2017 based on the criteria for security, availability, processing integrity, and confidentiality in the American Institute of Certified Public Accountants' TSP Section 100A, Trust Services Principles and Criteria, for Security, Availability, Processing Integrity, Confidentiality, and Privacy. This assertion is the responsibility of Google’s management. Our responsibility is to express an opinion based on our examination.

Our examination was conducted in accordance with attestation standards established by the American Institute of Certified Public Accountants. Those standards require that we plan and perform our examination to obtain reasonable assurance about whether management’s assertion is fairly stated, in all material respects. An examination involves performing procedures to obtain evidence about management’s assertion, which includes:

1. obtaining an understanding of Google’s relevant security, availability, processing integrity and confidentiality policies, processes and controls;
2. testing and evaluating the operating effectiveness of the controls; and
3. performing such other procedures as we considered necessary in the circumstances.

The nature, timing, and extent of the procedures selected depend on our judgment, including an assessment of the risk of material misstatement, whether due to fraud or error. We believe that the evidence obtained during our examination is sufficient and appropriate to provide a reasonable basis for our opinion.

Our examination was not conducted for the purpose of evaluating Google’s cybersecurity risk management program. Accordingly, we do not express an opinion or any other form of assurance on its cybersecurity risk management program.
Inherent limitations:

There are inherent limitations in the effectiveness of any system of internal control, including the possibility of human error and the circumvention of controls. Because of inherent limitations in its internal control, those controls may provide reasonable, but not absolute, assurance that its commitments and system requirements related to security, availability, processing integrity and confidentiality are achieved.

Examples of inherent limitations of internal controls related to security include (a) vulnerabilities in information technology components as a result of design by their manufacturer or developer; (b) breakdown of internal control at a vendor or business partner; and (c) persistent attackers with the resources to use advanced technical means and sophisticated social engineering techniques specifically targeting the entity. Furthermore, projections of any evaluation of effectiveness to future periods are subject to the risk that controls may become inadequate because of changes in conditions or that the degree of compliance with the policies or procedures may deteriorate.

Opinion:

In our opinion, Google’s management’s assertion referred to above is fairly stated, in all material respects, based on the aforementioned criteria for security, availability, processing integrity and confidentiality.

30 June 2017
System Description of the Google Cloud Platform System

Google Overview

Google Inc. ("Google") is a global technology service provider focused on improving the ways people connect with information. Google’s innovations in web search and advertising have made Google’s web site one of the most viewed Internet destinations and its brand among the most recognized in the world. Google maintains one of the world’s largest online index of web sites and other content, and makes this information freely available to anyone with an Internet connection. Google’s automated search technology helps people obtain nearly instant access to relevant information from their vast online index.

Google offers Internet-based services and tools that user entities can access to communicate, collaborate, and work more efficiently. The following Google product offerings automatically save all work performed by user entities in the cloud and enable user entities to work securely, regardless of where they are in the world and what device they are using.

Google Cloud Platform provides Infrastructure as a Service ("IaaS") and Platform as a Service ("PaaS"), allowing businesses and developers to build and run any or all of their applications on Google’s Cloud infrastructure. Users can benefit from performance, scale, reliability, ease-of-use, and a pay-as-you-go cost model. It includes the following services, hereafter described collectively as “Google Cloud Platform” or (“GCP”):

Big Data
- BigQuery
- Cloud Dataflow
- Cloud Dataproc
- Genomics

Compute
- App Engine
- App Engine Flexible Environment
- Compute Engine
- Container Engine
- Container Registry
- Cloud Functions

Machine Learning
- Machine Learning Engine
- Natural Language API
- Speech API
- Translation API
- Cloud Vision API

Networking
- Cloud CDN (Content Delivery Network)
- Cloud DNS (Domain Name System)
- Cloud Load Balancing
Google's product offerings for Google Cloud Platform provide the unique advantage of leveraging the resources of Google’s core engineering team while also having a dedicated team to develop solutions for the corporate market. As a result, these Google offerings are positioned to innovate at a rapid rate and provide the same level of service that users are familiar with on google.com.

Google Cloud Platform is targeted towards small and medium size business, and large corporates alike, as well as the development teams within those organizations. These products provide a comprehensive variety of technical services that organizations rely on:

- **Big Data** - tools to capture, process, store, and analyze data on a single platform
Google Inc. | Description of the System

- Computation - a scalable range of computing options tailored to match the size and needs of an organization
- Machine Learning - fast, scalable and easy to use modern machine learning services, with pre-trained models and the ability to generate tailored models
- Networking - a high quality private network using software-defined networking and distributed systems technologies to host and deliver services around the world
- Cloud Management - manage apps on GCP with a web-based console, mobile app, or Cloud Shell for real time monitoring, logging, diagnostics, and configuration
- Identity Management Services - manage the security and access to cloud assets, supported by Google’s own protection of its infrastructure
- Storage - scalable storage options and varieties for different needs and price points
- Developer Tools - a rich collection of tools and libraries that help development teams work quickly and effectively

The products are comprised of communication, productivity, collaboration and security tools that can be accessed from virtually any location with Internet connectivity. This means every employee and each user entity they work with can be productive from anywhere, using any device with an Internet connection.

The Google Cloud Platform products covered in this system description consist of the following services:

**BigQuery**

BigQuery Service is a fully managed data analysis service that enables businesses to analyze Big Data. It features highly scalable data storage that accommodates up to hundreds of terabytes, the ability to perform ad hoc queries on multi-terabyte datasets, and the ability to share data insights via the web.

**Cloud Dataflow**

Cloud Dataflow is a fully managed service for strongly consistent, parallel data-processing pipelines. It provides a Software Development Kit (“SDK”) for Java with composable primitives for building data-processing pipelines for batch or continuous processing. This service manages the lifecycle of Google Compute Engine resources of the processing pipeline(s). It also provides a monitoring user interface for understanding pipeline health.

**Cloud Dataproc**

Cloud Dataproc is a fast, easy to use, managed Spark and Hadoop service for distributed data processing. It provides management, integration, and development tools for unlocking the power of rich open source data processing tools. With Cloud Dataproc, user can create Spark/Hadoop clusters sized for their workloads precisely when they need them.
Genomics

Genomics provides an Application Programming Interface (“API”) to store, process, explore and share DNA sequence reads, reference-based alignments, and variant calls, using Google's cloud infrastructure.

App Engine

App Engine enables developers to build and host web apps on the same systems that power Google applications. App Engine offers fast development and deployment; simple administration with no need to worry about hardware, patches or backups; and effortless scalability. App Engine also provides the ability to create Managed VMs. In addition, developers can build client APIs for their App Engine applications using Google Cloud Endpoints.

App Engine Flexible Environment

App Engine Flexible makes it easy to build, deploy, and manage web applications and APIs on the same systems that power Google applications. App Engine flexible offers fast development and deployment; simple administration with no need to worry about hardware, patches or backups; and effortless scalability. For ultimate convenience and customizability, App Engine Flexible lets user run their applications using our built in runtimes, or they can bring their own Docker image.

Compute Engine

Compute Engine offers scalable and flexible virtual machine computing capabilities in the cloud. Developers can use Google Compute Engine to solve large-scale processing and analytic problems on Google’s computing, storage, and networking infrastructure.

Container Engine

Container Engine, powered by the open source container scheduler Kubernetes, enables users to run containers on Google Cloud Platform. Container Engine takes care of provisioning and maintaining the underlying virtual machine cluster, scaling their application, and operational logistics such as logging, monitoring, and cluster health management.

Container Registry

Container Registry is a private Docker image storage system on Google Cloud Platform. The registry can be accessed through an HTTPS endpoint, or your own hardware.

Cloud Functions

Cloud Functions is a lightweight, event-based, asynchronous compute solution that allows users to create small, single-purpose functions that respond to cloud events without the need to manage a server or a runtime environment. Events from Google Cloud Storage and Google Cloud Pub/Sub can trigger Cloud Functions asynchronously, or users can use HTTP invocation for synchronous execution.
Machine Learning Engine

Machine Learning Engine is a managed service that enables user to easily build machine learning models with the powerful TensorFlow framework. It provides scalable training and prediction services that work on large scale datasets.

Natural Language API

Natural Language API provides powerful natural language understanding as an easy to use API. This API enables application developers to answer the following questions: 1) What are the entities referred to in the block of text? 2) What is the sentiment (positive or negative) for this block of text? 3) What is the language of this block of text? and 4) What is the syntax for this block of text (including parts of speech and dependency trees)? Users can call this API by passing in a block of text or by referring to a document in Google Cloud Storage.

Speech API

Speech API allows developers to convert audio to text by applying powerful neural network models in an easy to use API. To support customer global user base, the API can recognize over 80 languages and variants. Users can transcribe the text of users dictating to an application’s microphone, enable command-and-control through voice, or transcribe audio files, among many other use cases. Users can stream audio directly to the API or pass URL to audio stored in Google Cloud Storage.

Translation API

Translation API automatically translates text from one language to another language (e.g., French to English). Users can use the API to programmatically translate text in their webpages or apps.

Cloud Vision API

Cloud Vision API enables developers to understand the content of an image by encapsulating powerful machine learning models in an easy to use API. It quickly classifies images into thousands of categories (e.g., "sailboat", "lion", "Eiffel Tower"), detects individual objects and faces within images, and finds and reads printed words contained within images. Users can build metadata on their image catalog, moderate offensive content, or enable new marketing scenarios through image sentiment analysis. Users can also analyze images uploaded in the request and integrate with their image storage on Google Cloud Storage.

Cloud CDN (Content Delivery Network)

Cloud CDN uses Google’s globally distributed edge points of presence to cache HTTP(S) load balanced content close to users.

Cloud DNS (Domain Name System)

Cloud DNS is a high performance, resilient, global, fully managed DNS service that provides a RESTful API to publish and manage DNS records for applications and services.
**Cloud Load Balancing**

Cloud Load Balancing is a distributed, software-defined, managed service for all traffic (HTTP(S), TCP/SSL, and UDP) to user's computing resources. It is not an instance or device based solution, so users aren't locked into a physical load balancing infrastructure or face the HA, scale and management challenges inherent in instance based load balancers. In contrast to DNS-based Global Load Balancing solutions, Cloud Load Balancing reacts instantaneously to changes in users, traffic, network, backend health and other related conditions.

**Cloud Pub/Sub**

Cloud Pub/Sub is designed to provide reliable, many-to-many, asynchronous messaging between applications. Publisher applications can send messages to a "topic" while other applications can subscribe to that topic to receive the messages. By decoupling senders and receivers, Google Cloud Pub/Sub allows developers to communicate between independently written applications.

**Cloud Router**

Cloud Router is a fully managed and resilient network service that enables GCP and non-GCP customer networks connected by Cloud VPN to auto-discover and auto-update each other. Cloud Router enables this by dynamically sending updates about changes in its GCP network topology and learning about changes in the other non-GCP network using BGP.

**Cloud Virtual Network**

Cloud Virtual Network is a comprehensive set of Google-managed networking capabilities including granular IP address range selection, routes and firewalls. The Virtual Network Core makes up the foundation for several Google Cloud networking services such as VPN and Cloud Router.

**Cloud VPN**

Cloud VPN securely connects customer's existing network to their Google Cloud Platform network through an IPsec VPN connection. Traffic traveling between the two networks is encrypted by one VPN gateway, then decrypted and authenticated by the other VPN gateway. This protects users' data as it travels over the Internet.

**Debugger**

Debugger enables developers to inspect the call-stack and variables of a running cloud application, in real-time, without stopping or slowing it down. It is safe to use in test, production or any other deployment environment. It can be used to debug applications written in any programming language.

**Error Reporting**

Error Reporting counts, analyzes and aggregates the crashes in applications. The crash data is extracted from application logs on Google Cloud or reported via the public API. Users can
inspect the collected data via the UI or the public API and may opt-in to receive notifications about the occurrence of errors.

*Stackdriver Logging*

Stackdriver Logging is a fully-managed service that allows users to store, search, analyze, monitor, and alert of log data and events from Google Cloud Platform and Amazon Web Services (“AWS”). The API also allows ingestion of any custom log data from any source. Stackdriver Logging can ingest application and system log data from thousands of VMs and analyze log data in real-time.

Software related to Stackdriver Logging (i.e. open-source logging agent) that customers are able to download and install on their own VMs is out of the scope of this report.

*Trace*

Trace collects latency data from customers’ applications and displays it in the Google Cloud Platform Console. It automatically analyzes trace data to generate in-depth performance reports that help surface performance bottlenecks.

*Google Service Control*

Google Service Control provides control plane functionality to managed services, such as logging, monitoring, and status checks.

*Cloud Billing API*

Cloud Billing API provides methods such that users can use to programmatically manage billing for their projects on the Google Cloud Platform.

*Cloud Console*

Cloud Console is used to manage and get insights into everything that powers the cloud application -- including web applications, data analysis, virtual machines, datastore, databases, networking, and developer services. Google Cloud Console helps users to deploy, scale and diagnose production issues in a simple web based interface. Search to quickly find resources and connect to instances via SSH in the browser. Handles devops workflows on the go with powerful native iOS and Android applications. Master the most complex development tasks with Google Cloud Shell, the admin machine in the cloud.

*Cloud IAM (Identity & Access Management)*

Cloud Identity & Access Management (“IAM”) lets administrators authorize who can take action on specific resources, giving them full control and visibility to manage access to cloud resources centrally. For established enterprises with complex organizational structures, hundreds of workgroups and potentially many more projects, Cloud IAM provides a unified view into security policy across users’ entire organization, built-in auditing to ease compliance processes.
Cloud Jobs API

Cloud Jobs API, our latest machine learning service, provides the necessary lingua franca between the job seeker and employer job postings in order to improve the hiring process. Much like how Google Cloud Translation API translates an arbitrary string into any supported language, Cloud Jobs API understands the nuances of job titles, descriptions, skills and preferences, and matches job seeker preferences with relevant job listings based on sophisticated classifications and relational models. The API is intended for job boards, career sites and applicant tracking systems.

Cloud Key Management Service

Cloud Key Management Service (“KMS”) is a cloud-hosted key management service that lets users manage encryption for their cloud services the same way it’s done on-premise. Users can generate, use, rotate and destroy AES256 encryption keys.

Cloud Launcher

Cloud Launcher is a marketplace that offers ready-to-go development stacks, solutions, and services to accelerate development. It enables users to deploy production-grade solutions in a few clicks, obtain direct access to partner support, and receive a single bill for all GCP and 3rd party services.

Cloud Resource Manager

Cloud Resource Manager API allows users to programmatically manage Google Cloud Platform container resources (such as Organizations and Projects), that allow users to group and hierarchically organize other Google Cloud Platform resources. This hierarchical organization lets users easily manage common aspects of resources such as access control and configuration settings.

Cloud Security Scanner

Cloud Security Scanner is a web application security scanner for Google App Engine. It enables developers to easily check for a subset of common web application vulnerabilities in websites built on App Engine.

Cloud Shell

Cloud Shell provides command-line access to Google Cloud Platform resources through an in-browser Linux shell backed by a temporary Linux VM in the cloud. It allows developers to manage their projects and resources without having to install additional tools on their system and comes equipped and configured with common developer tools such as text editors, a MySql client, and kubernetes.

Cloud Source Repositories

Cloud Source Repositories provides Git version control to support collaborative development of any application or service as well as a source browser that can be used to browse the contents of repositories and view individual files from within the Cloud Console. Cloud Source
Repositories and related tools (e.g., Stackdriver Debugger) can be used to view debugging information alongside users’ code during application runtime.

**Container Builder**

Container Builder lets developers create container images from application source code located in Google Cloud Storage or in a third party service (e.g., Github, Bitbucket). Created Container images can be stored in Container Registry and deployed on Container Engine, Compute Engine, App Engine Flexible Environment or other services where users can run applications from Docker containers.

**Deployment Manager**

Deployment Manager is an infrastructure management service that makes it simple to create, deploy, and manage Google Cloud Platform resources.

**Data Loss Prevention API**

Data Loss Prevention (“DLP”) API helps developers classify and redact content in text, images, and cloud assets using a set of predefined detectors for types of sensitive or personally identifiable information. It quickly classifies text into information types (e.g., phone_number, social_security_number, passport_number) along with metadata like offsets or bounding boxes for images. Users can use this API in their own workflows and applications to help better understand and manage their data.

**Cloud Bigtable**

Cloud Bigtable is a fast, fully managed, highly-scalable NoSQL database service. It is designed for the collection and retention of data from 1TB to hundreds of PB.

**Cloud Datastore**

Cloud Datastore is a fully managed, schemaless, non-relational datastore. It provides a rich set of query capabilities, supports atomic transactions, and automatically scales up and down in response to load. It can scale to support an application with 1,000 users or 10 million users with no code changes.

**Cloud SQL**

Cloud SQL is a web service that allows developers to create, configure, and use relational databases that live in Google’s cloud. It is a fully-managed service that maintains, manages, and administers databases, allowing users to focus on their applications and services.
Cloud Storage

Cloud Storage is a RESTful service for storing and accessing data on Google’s infrastructure. The service combines the performance and scalability of Google’s cloud with advanced security and sharing capabilities.

Cloud Mobile App

Cloud Mobile App is a native mobile app that enables customers to manage key Google Cloud services. It provides monitoring, alerting, and the ability to take actions on resources.

Cloud SDK

Cloud SDK is a set of command-line tools for the Cloud Platform that can be run interactively or in automated scripts. These tools can be used to manage Google Compute Engine virtual machines, network and firewall configurations, disk storage, and much more without having to use the Cloud Platform Console.

Infrastructure

Google Cloud Platform runs in a multi-tenant, distributed environment. Rather than segregating user entity data to one machine or set of machines, data from all user entities is distributed amongst a shared infrastructure. For Google Cloud Platform, this is achieved through a Google distributed file system designed to store extremely large amounts of data across many servers. Customer data is then stored in large distributed databases, built on top of this file system.

Data Centers and redundancy

Google maintains consistent policies and standards across all data centers for physical security to help protect production and corporate servers, network devices and network connections within Google data centers.

Redundant architecture exists such that data is replicated in real-time to at least two (2) geographically dispersed data centers. The data centers are connected through multiple encrypted network links and interfaces. This provides high availability by dynamically load balancing across those sites. Google uses a dashboard that provides details such as resource footprint, central processing unit capacity, and random-access memory availability to monitor resource availability across their data centers and to validate that data has been replicated to more than one location.

Authentication and access

Strong authentication and access controls are implemented to restrict administrative access to Google Cloud Platform production systems, internal support tools, and customer data. Machine-level access restriction relies on a Google-developed distributed authentication service based on Transport Layer Security ("TLS") certificates, which helps to positively identify the resource access requester. This service also offers transport encryption to enhance data confidentiality in transit. Data traffic is encrypted between Google production facilities.
Google follows a formal process to grant or revoke employee access to Google resources. Lightweight Directory Access Protocol ("LDAP"), Kerberos, and a Google proprietary system which utilizes Secure Shell ("SSH") and TLS certificates help provide secure and flexible access mechanisms. These mechanisms are designed to grant access rights to systems and data only to authorized users.

Both user and internal access to customer data is restricted through the use of unique user IDs. Access to sensitive systems and applications requires two-factor authentication in the form of a unique user ID, strong passwords, One-Time-Passwords ("OTP"), Security Keys and/or certificates. Periodic reviews of access lists are implemented to help ensure access to customer data is appropriate and authorized. Access to production machines, network devices and support tools is managed via an access group management system. Membership in these groups must be approved by respective group administrators. User groups are annually reviewed.

Change Management

Change Management policies, including security code reviews and emergency fixes, are in place, and procedures for tracking, testing approving, and validating changes are documented. Changes are developed utilizing the code versioning tool to manage source code, documentation, release labeling and other functions. Google requires all code changes to be reviewed and approved by a separate technical resource, other than the developer, to evaluate quality and accuracy of changes. Further, all application and configuration changes are tested prior to migration to production environment.

Data

Google provides controls at each level of data storage, access, and transfer. Google has established training programs for privacy and information security to support data confidentiality. All employees are required to complete these training programs annually. All product feature launches that include new collection, processing, or sharing of user data are required to go through an internal design review process. Google has also established incident response processes to report and handle events related to confidentiality. Google establishes agreements, including non-disclosure agreements, for preserving confidentiality of information and software exchange with external parties.

Network Architecture and Management

The Google Cloud Platform system architecture utilizes a fully redundant network infrastructure. Google has implemented perimeter devices to protect the Google network from external attacks. Network monitoring mechanisms are in place to prevent and disconnect unauthorized access to the Google network from unauthorized devices.

People

Google has implemented a process-based service quality environment designed to deliver the Google Cloud Platform products to customers. The fundamentals underlying the services provided are the adoption of standardized, repeatable processes; the hiring and development of highly skilled resources; and leading industry practices. Google’s repeatable process model includes key infrastructure and product related processes and controls over security, availability,
process integrity, and confidentiality.

Formal organizational structures exist and are available to Google employees on the Company’s intranet. The intranet provides drill-down functionality for identifying employees in the functional operations team. Google has developed and documented formal policies, procedures, and job descriptions for operational areas including data center operations, security administration, system and hardware change management, hiring, training, performance appraisals, terminations, and incident escalation. These policies and procedures have been designed to segregate duties and enforce responsibilities based on job functionality. Policies and procedures are reviewed and updated as necessary.