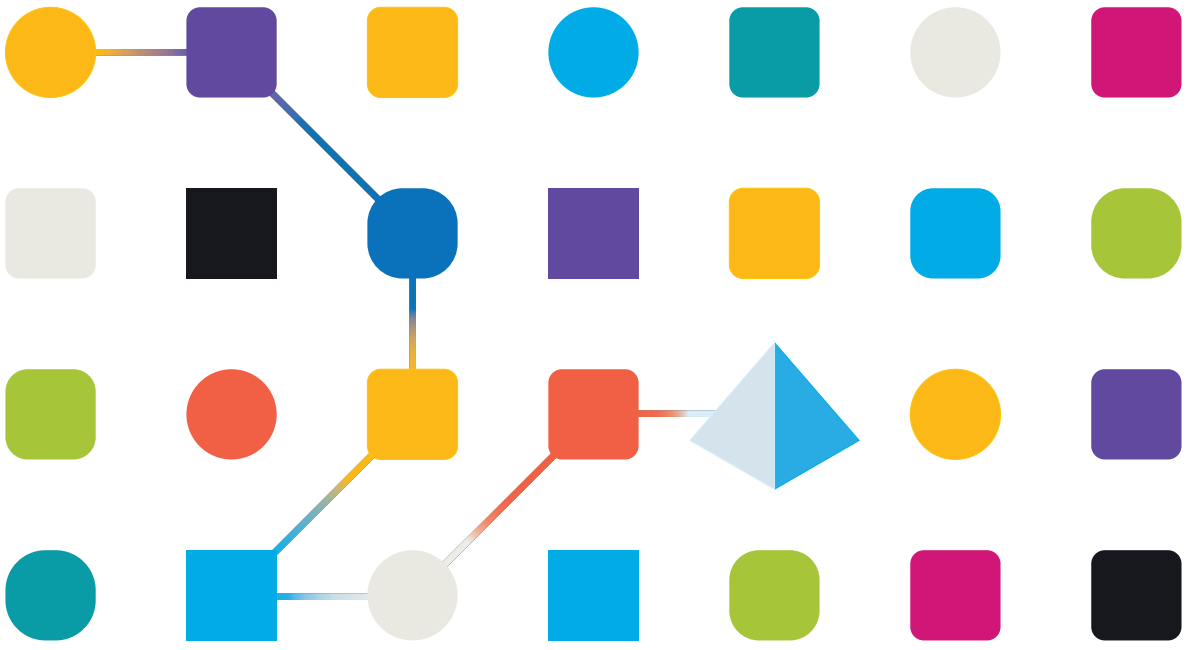
blueprism®

Blue Prism 6

Google Cloud Platform Reference Architecture

Document Revision: 1.0



Trademarks and Copyright

Blue Prism Limited, 2 Cinnamon Park, Crab Lane, Warrington, WA2 0XP, United Kingdom.
Registered in England: Reg. No. 4260035. Tel: +44 370 879 3000. Web: www.blueprism.com

Contents

Introduction

Intended audience

This reference guide is intended for use by system architects and designers who are seeking to gain an understanding of the options and considerations for deploying a Blue Prism environment in the Google Cloud Platform (GCP).

About this guide

The document provides an overview of the key considerations for a GCP based deployment of Blue Prism, along with reference architectures for the commonly requested patterns for a Cloud based deployment. The objective of this document is to explain the key considerations for deployment on GCP. A basic understanding of the GCP architecture is expected. The reference architectures contained within this document are based upon generalized assumptions and GCP design best practices and Reference Architectures. The architecture may need to be modified to suit a client deployment.

Google Cloud Services and key concepts

The following sections outline some of the relevant services and concepts that are key in designing a GCP based Blue Prism deployment.

Google Cloud Platform Services

The Blue Prism architecture can utilise several key GCP services and concepts. These are outlined below. Refer to the GCP documentation in the links for further information.

- Virtual Private Cloud (VPC) – The Google Cloud Platform Virtual Private Cloud (VPC) allows you to securely connect devices as well as isolate others. A VPC can be thought of as your network in the cloud, it allows the same fine-grained controls as you get on premise. It is possible to extend a VPC to your on-premise network using it's built in VPN connectivity.
- Firewall Rules – Firewall Rules are virtual firewalls that and are mechanism used to allow resources to access resources in other subnets. By default, all traffic is denied unless there is a specified Firewall Rule to allow it, this technique can be used to isolate certain subnets within a VPC.
- Global, Regional and Zonal Resources – In GCP it is important to understand the relationship of infrastructure resources. Global is the top level, there are a few resources that operate at the Global level – VPC is one. The globe is split into several Regions, these refer directly to geographic locations of Google data centres. Each Region is made up of two or three Zones, zones are completely isolated areas of each data centre for fault tolerance. Inter-zone communication benefits from petabit per second communication across Google's internal network. Many of the resources available in GCP are deployed into Zones, for Blue Prism GCP Compute Engine (Virtual Machines) are deployed at this level.
- Managed Instance Group – Managed Instance Groups (MIGs) use an instance template to deploy many Virtual Machines with the same machine configuration but manage them as a single entity. This allows for easy scaling back and out based on current usage in addition to providing load balancing capabilities. MIGs can identify when resources are unresponsive and recreate the instance automatically to ensure the target number of instances is maintained, they follow a similar process for unhealthy resources to ensure all instances are running optimally. Managed Instance Groups operate at two levels – zonal and regional. Zonal MIGs ensure that all instances in the group are in the same zone whereas regional MIGs ensure equal numbers of instances are deployed to available zones in the region allowing for higher availability in the unlikely event of a total zone failure.
- Google Organisations – An organisation is the root node of the resource hierarchy; all resources will be linked back through a structure to the organisation at the top level which provides a view of everything contained inside. GCPs Cloud Identity and Access Management (IAM) allows access to be granted across the entire organisation or just a specific, smaller part. Organisations are at the very highest level so operate as a global level resource. All GCP resources would be found under a single organisation, as an organisation also acts as a billing mechanism.

GCP SLAs and Impact on Target Architecture

The impact of the Google Cloud Service Level Agreements (SLAs) on each component needs to be assessed against the non-functional requirements of the target architecture.

Google are continually improving the levels of service offered and the areas in which they are providing them. To match target architecture requirements, it is recommended to thoroughly consult the GCP SLAs and geographic availability documentation to ensure that all requirements are fulfilled. If resources are not available in the desired region they should be placed in the nearest possible appropriate geographic region.

GCP SQL PaaS

Currently there is no option in Google Cloud Platform for a managed instance of Microsoft SQL Server. GCP does provide a method for using SQL Server on the platform by building a Virtual Machine from a provided, preconfigured image. Virtual Machines in GCP are very configurable so CPU and memory can be adjusted to suit the needs of a VM as required. The provided images for SQL Server are quite wide ranging in terms of OS and SQL versions with licensing included in the price.

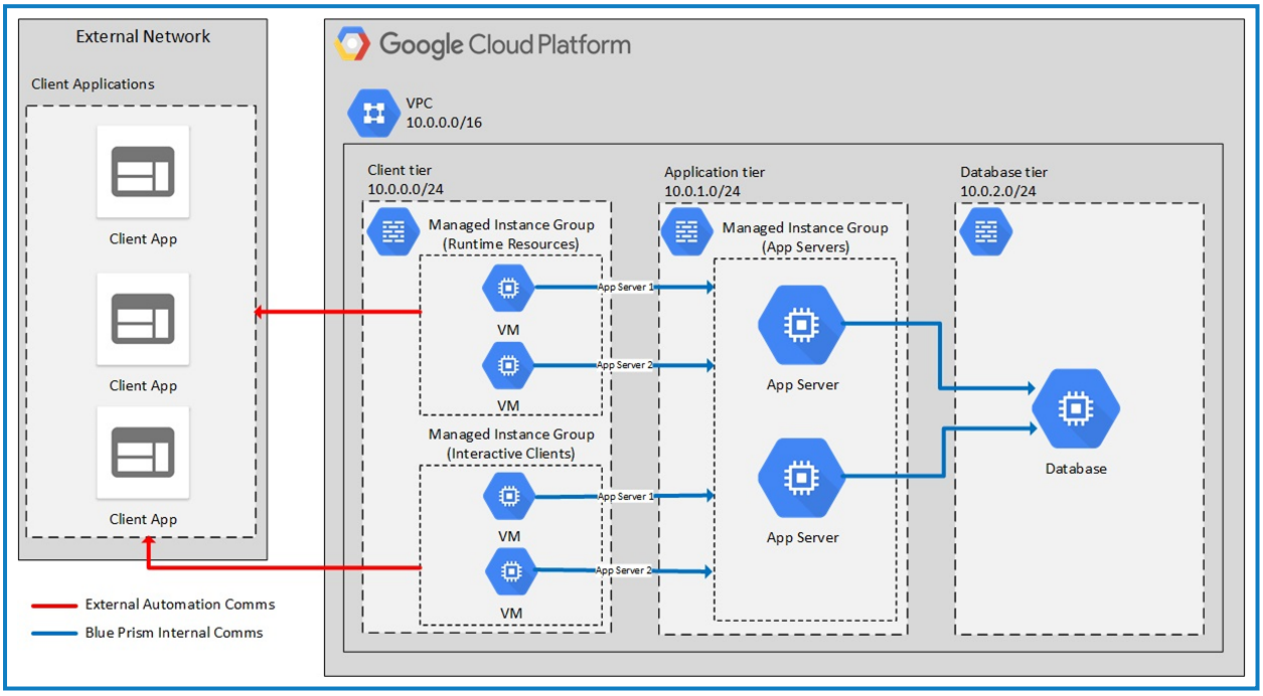
This approach brings the cost of extra management overhead for the SQL Server and the disk space used on the SQL Server. However, it is more familiar for traditional database administrators whilst providing an extra level of control compared to a fully managed solution.

Blue Prism Google Cloud platform reference architectures

The following sections outline the core reference architectures for the expected deployment models for Blue Prism on the Google Cloud Platform. These are supplemented by peripheral design considerations and scenarios, such as integration with Single Sign On (SSO).

Small scale deployment using SQL Express

This pattern makes use of the GCP SQL Express image for a small, lightweight deployment.



Assumptions

- A connection is in place between the network hosting Blue Prism and any external networks where the automated applications reside. This may be via a VPN or internal VPC Peering (for cloud hosted applications).
- Blue Prism clients and runtimes will need to have all necessary software installed to facilitate the automation and/or monitoring of runtimes.

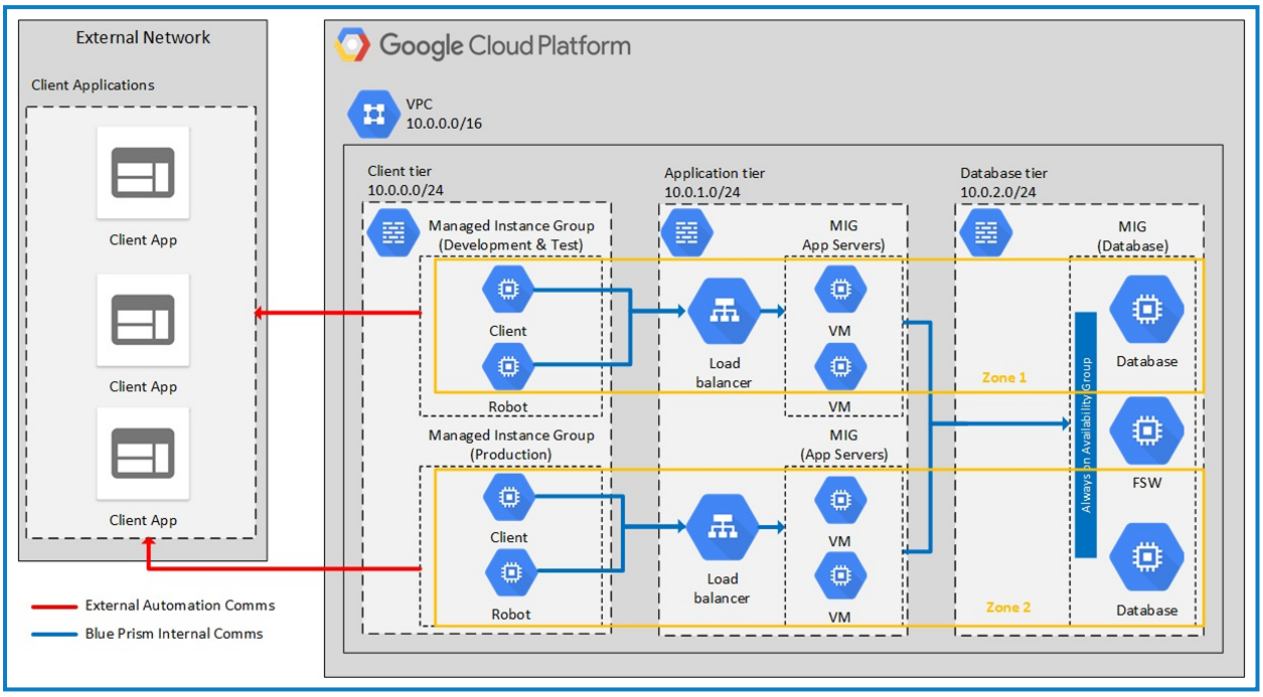
The Authentication components are not shown here. See section 4 for considerations.

Key design considerations

- The database sizing recommendations for Blue Prism in the Infrastructure Reference Guide should be used as a baseline for selecting the appropriate specification. The guide is accessible from the Blue Prism Portal with Privilged access.
- Multiple application servers are included in this design, to provide some degree of high availability in event of a failure. This may be reduced to 1 for small POC environments or removed from the Managed Instance Group to remove complexity.
- Consider splitting the connections between application servers and distributing workload between Runtimes on both, to account for failures or maintenance of one Application Server within the Availability set.
- The use of SQL Express has several restrictive performance restrictions, as such, this pattern should only be used for very small scale deployments.

Medium to large scale deployment using SQL clustering

This pattern involves the deployment of a SQL cluster with Always on Availability Groups (AAG) spread across two zones for High Availability and Disaster Recovery (HADR).



Assumptions

- A connection is in place between the Blue Prism network and any external networks where the automated applications reside. This may be via a VPN or internal VPC Peering (for cloud hosted applications).
- Blue Prism clients and runtimes will need to have all necessary software installed to facilitate the automation of remote applications
- The Authentication components are not shown here. See section 4 for considerations.

Key design considerations

- The sizing of the SQL environment (compute and database) should be based on the recommendations in the Infrastructure Reference Guide, accessible from the Blue Prism Portal with Privileged access. Google recommendations can be found here.

  Compute VMs in Google Cloud Platform are preloaded with SQL Server and an appropriate license for use in GCP, which is included in the price. For customers with Software Assurance from Microsoft, already purchased licenses are directly transferrable to GCP.
- This design pattern separates the Application servers into 2 separate Availability groups for the Interactive Clients and Runtimes across two different zones. This is optional and the application servers may be combined for both if desired. To provide very high availability spread application servers across regions and split them across zones.
- The number of application servers may be scaled up according to the size of environment. For more information about sizing, see the Infrastructure Reference Guide, accessible from the Blue Prism Portal with Privileged access.

- If Disaster Recovery is required, it may be necessary to deploy a cross-region scenario. As this is an unmanaged SQL Server instance regular best practices should be followed regarding backup procedures, with these being stored in a separate location either in GCP, on premise or another cloud provider. This is normal practice and should not be out of the ordinary for Database Administrators.

Supporting architecture patterns

The following section outlines the supporting architecture patterns for peripheral services, such as Authentication and remote connectivity. These are likely to be highly variable, depending on the client's existing GCP usage and strategy.

Authentication

Google does not offer a managed Microsoft Active Directory (AD) as part of their cloud services. Blue Prism relies on a Microsoft AD for Single Sign-On (SSO), so an AD infrastructure must be provisioned and managed as part of the deployed environment for scenarios where SSO for Blue Prism is intended.

There are multiple methods for integration of the Blue Prism environment into an Active Directory. This section will outline the two most common scenarios:

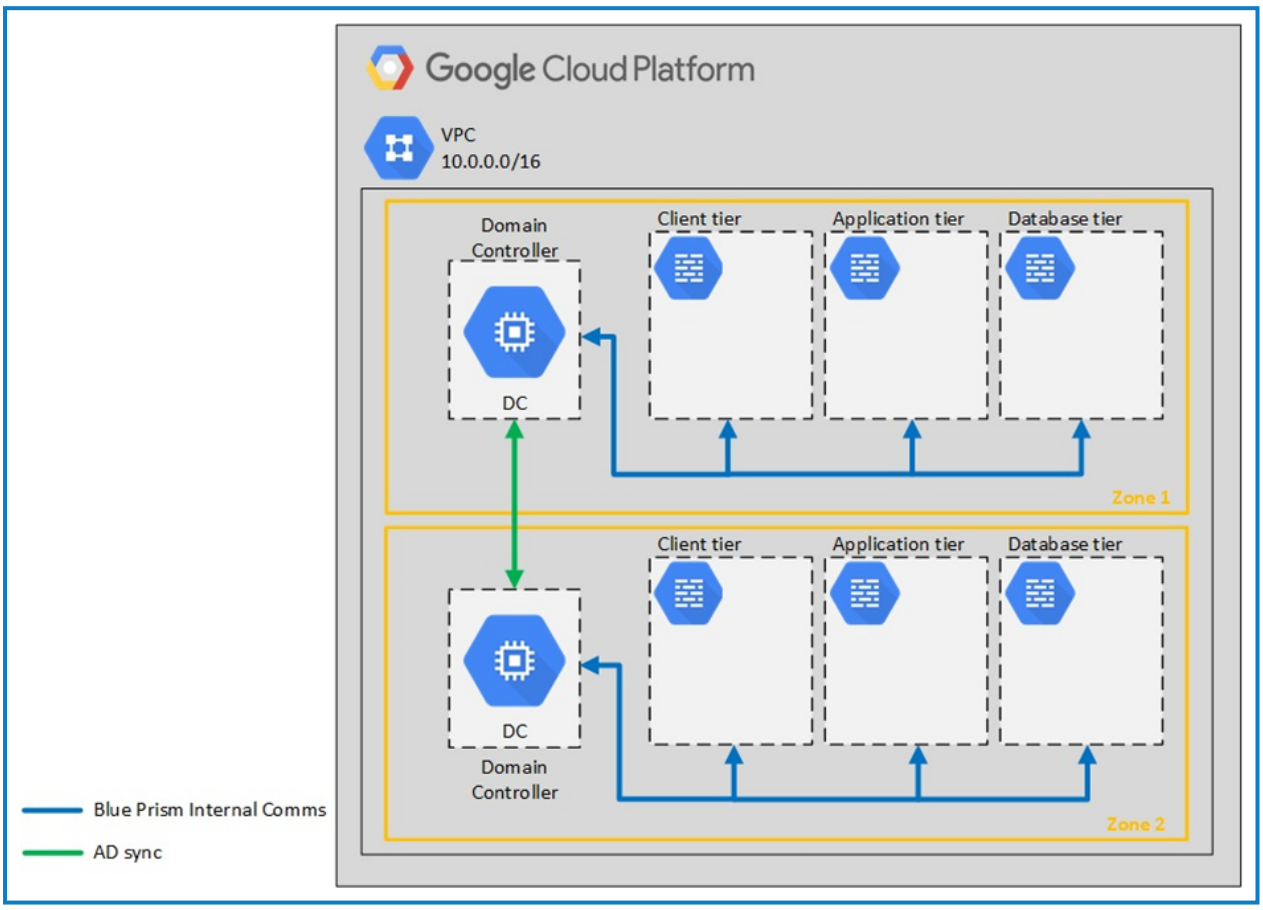
- Cloud only authentication via dedicated domain controller/s deployed on Google Cloud Platform.
- Hybrid authentication via dedicated domain controller/s in GCP replicating back to On-Premise network.

The approaches outlined below are documented to indicate the options in which Blue Prism may be integrated with a domain. The overall assessment and selection of an authentication approach is likely to be dependent on many factors and other unrelated client decisions.

Microsoft's introduction to Active Directory Replication can be found here.

Cloud only authentication using dedicated domain controllers

This pattern is applicable in a scenario where authentication for the Blue Prism application will be provided solely from the Google Cloud environment. This domain controller and directory structure is unique to this Blue Prism environment and is only accessible from within the Google Cloud VPC. As such, the logins are separate from any pre-existing corporate logins.



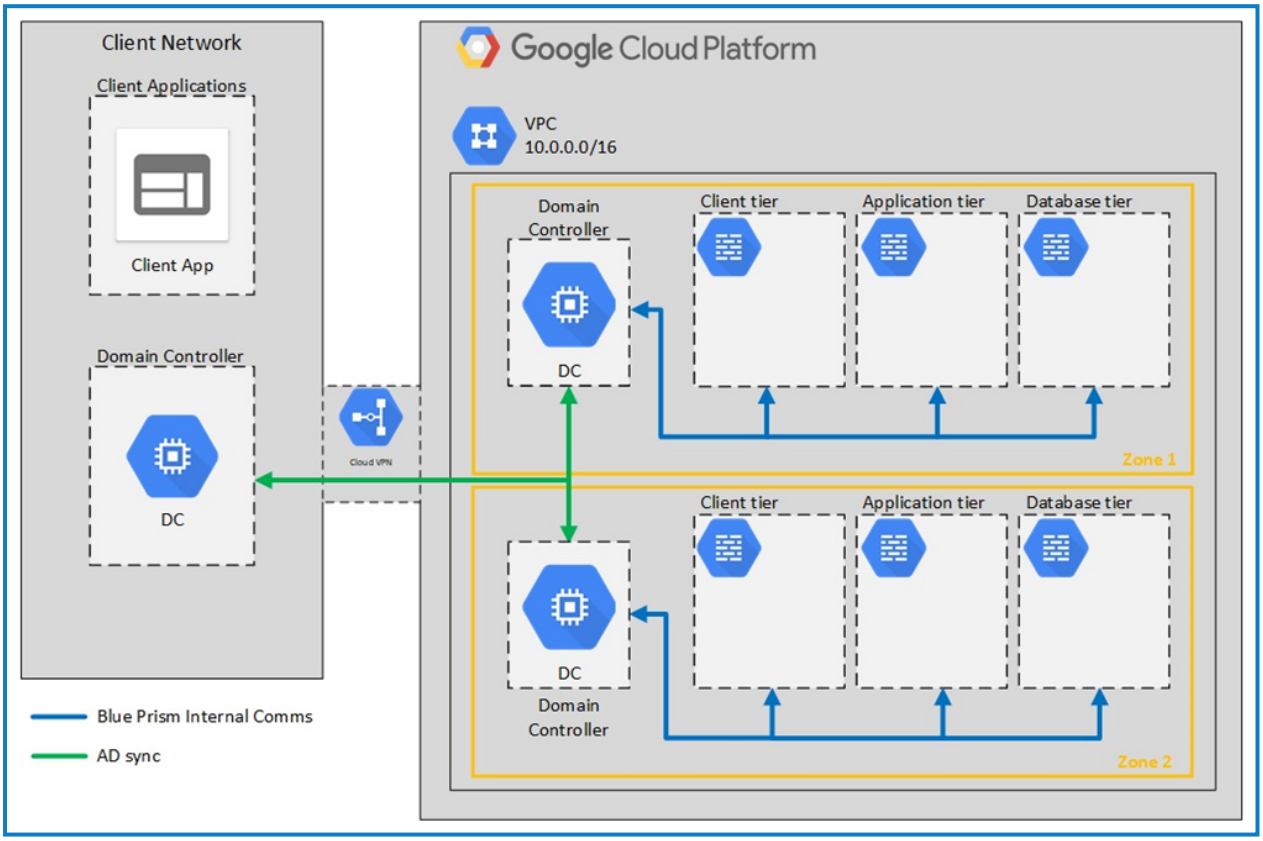
Key considerations

- Active Directory replication is setup across zones for High Availability.
- Active Directory is configured to use a single forest as outlined in the Blue Prism Active Directory Integration Guide. Blue Prism will not provide support for Active Directory administration, accounts and configuration are the responsibility of the customer.

Cloud only authentication using dedicated domain controllers

The following pattern is applicable where authentication for Blue Prism application will be controlled from a remote Active Directory on the client network. Domain Controllers are deployed into the cloud environment to reduce network latency of authentication calls.

An approach of this nature could still use on-premise applications and infrastructure whilst utilising the scalability of a cloud-based environment. This is the commonly preferred method of authentication for a Blue Prism environment on the Google Cloud Platform.



Key design considerations

The computer groups and OUs will be synced from the on-premise domain (i.e. the computer accounts for the cloud VMs will be visible on premise and vice versa for the shadow domain accounts.

> This is a sync relationship, not a trust-based setup. This setup is based on an extension of the current domain by deploying a domain controller in the cloud. It is possible to add an additional domain to the exisiting corporate forest, maintaining a single forest enables the use of Single Sign-On (SSO) funcionality.

Google Cloud platform resource configuration recommendations

The following section aims to outline some helpful hints and tips to configure common capabilities to successfully deploy Blue Prism in Google Cloud. The contents of this section may not apply to all GCP deployments of Blue Prism as each customer will have differing cloud strategies.

VPC and firewall rules configuration

A Virtual Private Cloud (VPC), as outlined in section 2.1 Google Cloud Platform Services, allows resources to be connected as though they were in the same physical network. The VPC is the highest level of configurable networking in GCP, if required VPCs can be peered to allow communication between each other.

All traffic in Google Cloud is denied by default unless it is explicitly allowed by a Firewall Rule associated with the VPC. Firewall Rules can have varying levels of granularity from single IP addresses to entire subnets or specific service accounts or tags. Some organisations may mandate that network connections be limited to only specific, associated machines or subnets. Large organisations or those who are particularly security conscious with very security conscious requirements may want to configure fine grained Firewall Rules as otherwise this is exposes a large attack plain for potential rogue resources on the network with access to every server on any port. Organisations with such requirements may want to allow inbound or outbound requests based on source or target IP and restrict communication ports to just Blue Prism configured ports, guidance on this area can be found in the Blue Prism Infrastructure Reference Guide, accessible from the Blue Prism Portal with Privileged access. This is configurable through the Google Cloud Console or the GCP REST APIs.

Google Intelligent Services integrations

The following section outlines the integrations that Blue Prism have produced for some of Google's Intelligence services. These are delivered as part of version 6.2 of the Blue Prism product, accessing these powerful services via Blue Prism only require an API key and the Blue Prism VBOs. The provided VBOs are available in the 6.2 release section of the Blue Prism portal.

All of Google's Intelligence services are being constantly evolved using pioneering deep learning techniques by Google data scientists and will continue to improve with use.

Vision

The Google Vision API exposes Google's class leading computer vision capabilities which can be accessed using VBOs provided by Blue Prism. The Vision API exposes several functions, which are listed and explained below.

- **Label Detection** – Identifies a broad range of categories from within the general image.
- **Face Detection** – Detects multiple faces within an image as well as key facial attributes, this is not facial recognition.
- **Explicit Content Detection** – Identifies the likelihood that an image has adult, spoofed, medical, violent or racy content.
- **Image Attributes** – Returns analysis of the image such as dominant colours and cropping suggestions.
- **Logo Detection** – Identifies the presence of any popular logos in the image.
- **Landmark Detection** – Detects popular natural or man-made landmarks from within the image.
- **Web Detection** – Searches the internet for similar images, identifying identical images if appropriate.
- **Optical Character Recognition (OCR)** – Detect and extract text from within an image, supports a range of languages.

The Vision API is accessible via a REST API, it is possible to request that more than one of the above services are carried out on the submitted image.

Natural Language Processing

Google's Natural Language Processing (NLP) service allows users to access powerful text analysis functions allowing the extraction of actionable insights. There are several functions, which are listed and explained below.

- **Syntax Analysis** – Extracts tokens and sentences to create dependency parse trees for each sentence.
- **Content Classification** – Documents are classified based on over 700 categories.
- **Entity Recognition** – Identifies pertinent entities and labels them by various types.
- **Sentiment Analysis** – Returns an overall sentiment value based on the block of text.

The Natural Language API is accessible via REST API and supports numerous varying languages including English, French, German, Chinese (Simplified and Traditional) and Japanese.

Translation

The Google Translation service only offers a few specialised features, which are the cornerstone of Google's successful Google Translate product. The key services are listed and explained below.

- **Text Translation** – Translation is provided from over 100 supported languages and thousands of supported language pairs.
- **Language Detection** – Text is analysed, and the source language is returned.

The Google Translation API is accessible via REST API and a small selection of programming languages.