

VISA Workshop Public Cloud Provider Demo



Disclaimer

This work uses Google Cloud Platform (GCP)
with absolutely no intention advertise it
for open research, my PhD project with European XFEL / University of Hamburg

This talk is not about service development or deployment by European XFEL or DESY.

VISA in knowledge transfer
photon/neutron science <-> university
virtual environments <-> empirical methods in research software engineering research

Visa deployment

Helm Chart for VISA (3.3.1) -> github.com/SchuhMichael/visa-helm

Migrated to Github, CI/CD workflow, release in the OCI package registry

Repository private; collaborators, reviewers welcome

Load balancing enabled

No default cloud provider installed

No dependency on Openstack

Helm Chart for cloud provider for GCE (Compute) -> github.com/SchuhMichael/visa-helm-gce-provider

Flux control repo on Github, deployed to GKE (Kubernetes) -> github.com/SchuhMichael/flux

- visa-helm
 - Postgres
 - Redis
- visa-helm-gce-provider
- Cert Manager
- Sealed Secrets
- Traefik Ingress Controller (moved away from nginx)

Not covered: dbmate (!), ETL, VISA-[FS, Patch, Print, Security Groups]

Visa deployment

GKE cluster on pre-emptible nodes (!)

API replicated

- But not the scheduler, load balancing effective for vdi sessions
- Redis standalone pub/sub, no persistence
 - When Redis restarts, but not VISA API, is the subscription renewed?

Quarkus Framework

- Stopped (paused) work on dev features. Dev-ui, live coding, debugger
- Helm Chart template: `application.yaml`

Spam filters (DFN, UHH) block my emails from GCE via gmxd@dfn.de

Todo: end-to-end tests (Playwright)

Artefacts in visa namespace, complete list

Before and outside of the Helm Charts:

- SealedSecret/visa/api-vdi-key-private
- SealedSecret/visa/api-vdi-key-public
- SealedSecret/visa/dockerhub-pull-secret
- SealedSecret/visa/gc-secret
- SealedSecret/visa/visa-cloud-openstack-secret
- SealedSecret/visa/visa-db-secret
- SealedSecret/visa/visa-openstack-secret
- SealedSecret/visa/visa-redis
- SealedSecret/visa/visa-smtp-credentials

Stateless Helm Chart for Cloud Provider

- HelmRelease/visa/visa-gce-provider
 - ConfigMap/visa/visa-gce-provider-cloud-gce-cm
 - Service/visa/visa-gce-provider-cloud-openstack-service
 - Deployment/visa/visa-gce-provider-cloud-openstack

- HelmRelease/visa/visa
 - NetworkPolicy/visa/visa-postgresql
 - NetworkPolicy/visa/visa-redis
 - PodDisruptionBudget/visa/visa-postgresql
 - PodDisruptionBudget/visa/visa-redis-master
 - ServiceAccount/visa/visa-postgresql
 - ServiceAccount/visa/visa-redis-master
 - ConfigMap/visa/visa-redis-configuration
 - ConfigMap/visa/visa-redis-health
 - ConfigMap/visa/visa-redis-scripts
 - ConfigMap/visa/visa-attribute-provider
 - ConfigMap/visa/visa-account-config
 - ConfigMap/visa/visa-api-config
 - ConfigMap/visa/visa-quarkus-application
 - ConfigMap/visa/visa-jupyter-proxy-config
 - Service/visa/visa-postgresql-hl
 - Service/visa/visa-postgresql
 - Service/visa/visa-redis-headless
 - Service/visa/visa-redis-master
 - Service/visa/visa-panosc-account-service
 - Service/visa/visa-quarkus-service
 - Service/visa/visa-panosc-api-service
 - Service/visa/visa-external-git
 - Service/visa/visa-panosc-jupyter-proxy-service
 - Service/visa/visa-panosc-web-service
 - Deployment/visa/visa-account-service
 - Deployment/visa/visa-api-service-primary
 - Deployment/visa/visa-api-service-secondary
 - Deployment/visa/visa-jupyter-proxy-service
 - Deployment/visa/visa-web-service
 - StatefulSet/visa/visa-postgresql
 - StatefulSet/visa/visa-redis-master
 - Ingress/visa/visa-api-ingress
 - Ingress/visa/visa-jupyter-proxy-ingress
 - Ingress/visa/visa-web-ingress
 - IngressRoute/visa/visa-docs-proxy
 - Middleware/visa/visa-rewrite-docs

Visa image

Packer GCE builder

Ubuntu 24 base image

<https://documentation.ubuntu.com/gcp/google-explanation/guest-agents/>

google-guest-agent

gce-compute-image-packages

google-compute-engine-oslogin

google-osconfig-agent <- **PAM Modules** (!)

Ubuntu 22 deprecated and deleted images

Images “hosted somewhere close to the zone/region”

startup delay due to image transfer, compared to image caching on Nova hosts

2cpu/4GB: stable; 4cpu/16GB: VDI better

Instance creation

```
createInstance(name: string,  
              imageId: string,  
              flavourId: string,  
              securityGroups: string[],  
              metadata: Map<string, string>,  
              bootCommand: string): Promise<string>;
```

```
const instanceResource: protos.google.cloud.compute.v1.IInstance = {  
  name: name,  
  machineType: `zones/${this._zone}/machineTypes/${flavourId}`,  
  networkInterfaces: [{  
    network: `projects/${this._projectId}/global/networks/default`,  
  }],  
  disks: [{  
    boot: true,  
    initializeParams: {  
      sourceImage: image.selfLink,  
      diskSizeGb: "50",  
    },  
  }],  
  tags: {  
    items: securityGroups,  
  },  
  metadata: {  
    items: Array.from(metadata.entries()).map(([key, value]) => ({ key, value })),  
  },  
};
```

Openstack / Visa couple disk size to flavour!
Visa does not manage or store disk sizes!
.. hardcoding disk size in cloud provider
.. to be added: `autoDelete: true`

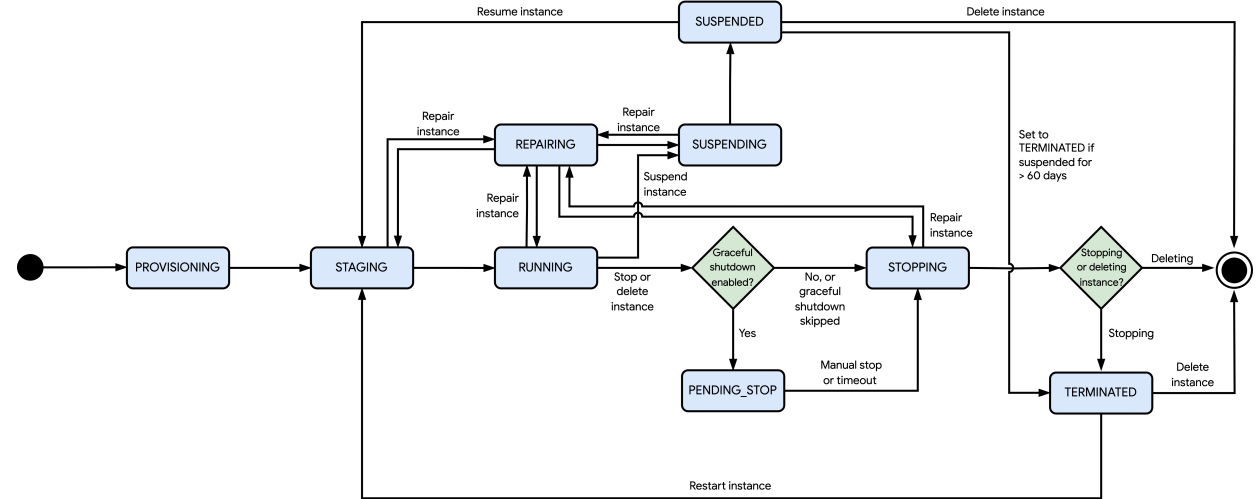
GCE instances can be tagged
GCE firewall-rules for source / target tags
Using tags as security group ids

Instance states and commands

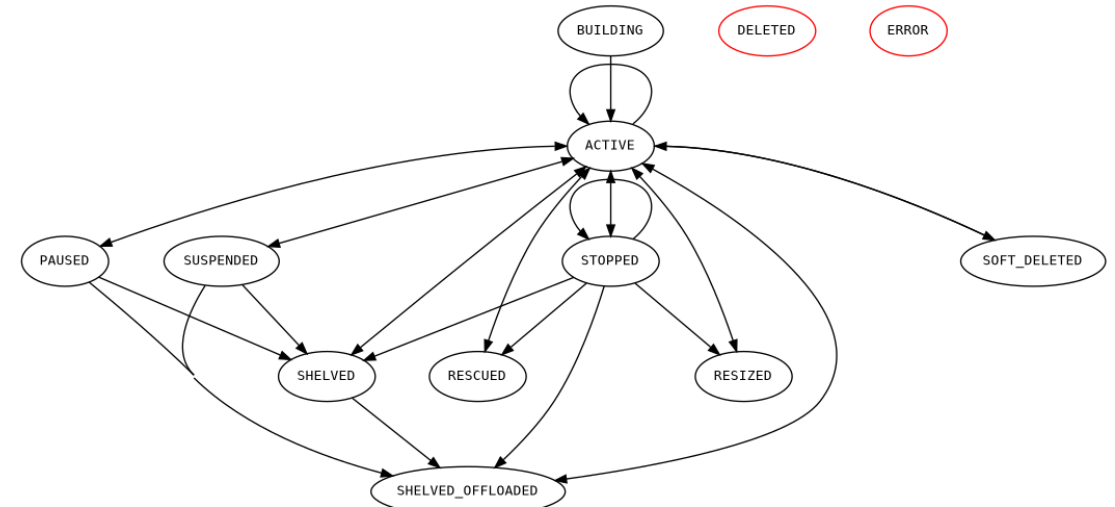
```
public enum InstanceState
{
    ACTIVE,
    PARTIALLY_ACTIVE,
    BUILDING,
    DELETED,
    DELETING,
    ERROR,
    STARTING,
    STOPPED,
    STOPPING,
    REBOOTING,
    UNAVAILABLE,
    UNKNOWN
}
```

post	/instances	create
delete	/instances/:id	remove
post	/instances/:id/start	start
post	/instances/:id/shutdown	shutdown
post	/instances/:id/reboot	reboot

<https://cloud.google.com/compute/docs/instances/instance-lifecycle>



<https://docs.openstack.org/nova/latest/reference/vm-states.html>



All states are allowed to transition to **DELETED** and **ERROR**.

Quotas and metrics

Openstack resource quotas appear in multiple APIs

OpenStackClient presents them as a single object with multiple properties.

Block Storage v1, v2, Compute v2, Network v2

In GCP resource quotas amount to 34,821 system limits (looked up today)

Metrics Client, Quotas Web UI in Google Cloud Console, Google Analytics

<https://cloud.google.com/compute/docs/api/compute-api-quota-metrics>

Outline: use project scopes

- Cloud Providers for dedicated projects
- Structure for billing, metrics, access
- Auxiliary services, specialised images, privileged deployments
- Tailscale (or similar) interesting for portable, platform agnostic networking