Debugging your debugging tools;

What to do when your service mesh goes down in production?

Neeraj Poddar
Co-founder & Chief Architect, Aspen Mesh

John Howard
Software Engineer, Google

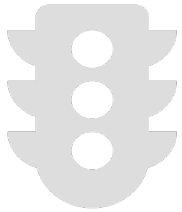
What Is a Service Mesh

A service mesh is…

- a transparent **infrastructure layer** that manages communication between microservices
- so that **developers** can focus on business logic
- while **operators** work independent of dev cycles to provide a more resilient environment

CLOUD NATIVE COMPUTING FOUNDATION ASPEN MESH

Key Benefits of a Service Mesh



Traffic Management

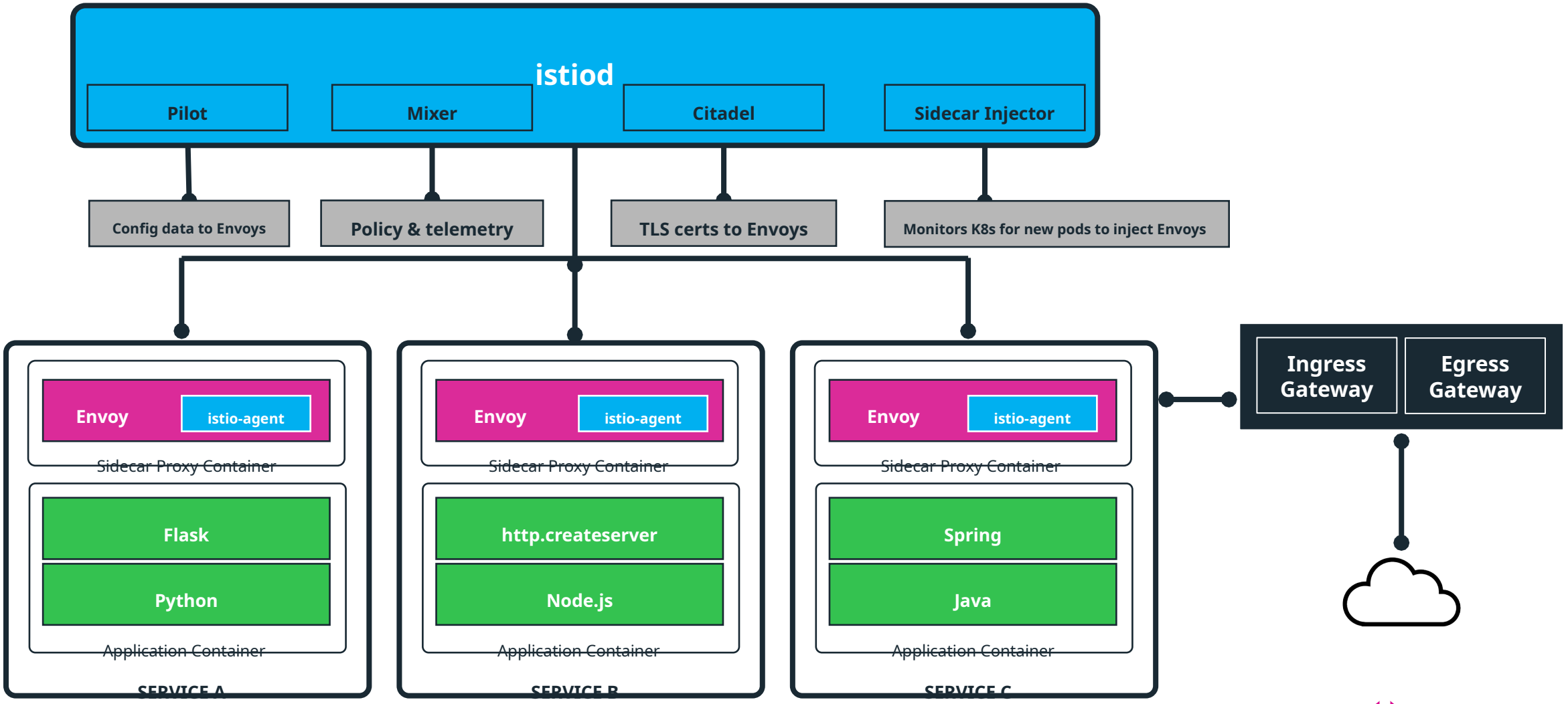
Security

Observability

CLOUD NATIVE COMPUTING FOUNDATION

ASPEN MESH

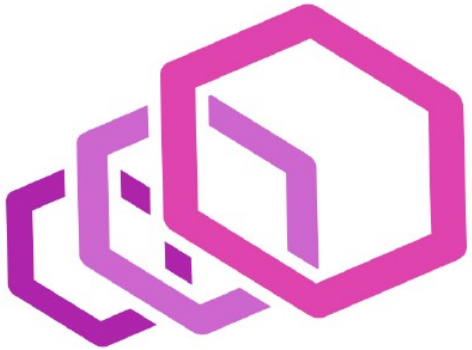
Istio Architecture Overview



Debugging Istio

Data plane

- Envoy sidecar
- Istio Agent



Control plane

- istiod

Debugging Envoy

- Connectivity to istiod
- Enabling & understanding access logs
- Debug logging
- Configuration dump

Connectivity to Istiod

```
kubectl -n foo exec -ti sleep-f8cbf5b76-htrpd -c sleep -- curl -sS istiod.istio-system:15014/debug/endpointz
```


```
{"svc": "details.default.svc.cluster.local:http", "ep": [
{
    "service": {
      "Attributes": {
        "ServiceRegistry": "Kubernetes",
        "Name": "details",
        "Namespace": "default",
        "UID": "istio://default/services/details",
        "ExportTo": null,
        "ClusterExternalAddresses": null,
        "ClusterExternalPorts": null
      },
      "ports": [
        {
          "name": "http",
          "port": 9080,
          "protocol": "HTTP"
        }
      ],
      "creationTime": "2020-06-15T23:16:59Z",
      "hostname": "details.default.svc.cluster.local",
      "address": "100.71.194.74",
      "Mutex": {},
      "cluster-vips": {
        "Kubernetes": "100.71.194.74"
      },
      "Resolution": 0,
      "MeshExternal": false
    },
    "servicePort": {
      "name": "http",
      "port": 9080,
      "protocol": "HTTP"
    },
```

Application Access Logs

- Useful for diagnosing traffic flows and failures
- End-to-end request routing across all microservices
- Default “turned off” in Istio
  - Enabled only in “demo” profile
- First tool in your debugging toolkit

Globally Enabling Access Logs

- Globally enabling using “istioctl install”

```
istioctl install --set profile=demo --set meshConfig.accessLogFile="/dev/stdout"
```

- Globally customizing encoding which defaults to “TEXT”

```
istioctl install --set profile=demo --set meshConfig.accessLogEncoding="JSON"
```

- Reverting back to no access log

```
istioctl install --set profile=demo --set meshConfig.accessLogFile="" --set meshConfig.accessLogEncoding="TEXT"
```

- This configuration is stored in “istio” configmap as default “mesh” configuration

```
kubectl -n istio-system get cm istio -o yaml | grep accessLog
```

```
accessLogEncoding: TEXT
accessLogFile: ""
accessLogFormat: ""
```

Enabling Access Logs for a Namespace

Use Namespace scope EnvoyFilter resource (Shouldn't be used if enabled globally)

```yaml
apiVersion: networking.istio.io/v1alpha3
kind: EnvoyFilter
metadata:
  name: access-log
  namespace: default
spec:
  configPatches:
  - applyTo: NETWORK_FILTER
    match:
      context: ANY
      listener:
        filterChain:
          filter:
            name: "envoy.http_connection_manager"
    patch:
      operation: MERGE
      value:
        typed_config:
          "@type": "type.googleapis.com/envoy.config.filter.network.http_connection_manager.v2.HttpConnectionManager"
          access_log:
          - name: envoy.file_access_log
            config:
              path: /dev/stdout
```

Understanding Access Logs


```
{
  "downstream_remote_address": "100.96.3.12:45004",
  "authority": "details:9080",
  "path": "/details/0",
  "protocol": "HTTP/1.1",
  "upstream_service_time": "0",
  "upstream_local_address": "127.0.0.1:33120",
  "duration": "1",
  "upstream_transport_failure_reason": "-",
  "route_name": "default",
  "downstream_local_address": "100.96.4.11:9080",
  "user_agent": "curl/7.35.0",
  "response_code": "200",
  "response_flags": "-",
  "start_time": "2020-07-21T00:45:48.439Z",
  "method": "GET",
  "request_id": "800604a6-613f-9946-b7a9-9be1520f65ee",
  "upstream_host": "127.0.0.1:9080",
  "x_forwarded_for": "-",
  "requested_server_name": "outbound_.9080_._.details.default.svc.cluster.local",
  "bytes_received": "0",
  "istio_policy_status": "-",
  "bytes_sent": "178",
  "upstream_cluster": "inbound|9080|http|details.default.svc.cluster.local"
}
```

Understanding Response Flags

`%RESPONSE_FLAGS%`

Additional details about the response or connection, if any. For TCP connections, the response codes mentioned in the descriptions do not apply. Possible values are:

HTTP and TCP

- **UH**: No healthy upstream hosts in upstream cluster in addition to 503 response code.
- **UF**: Upstream connection failure in addition to 503 response code.
- **UO**: Upstream overflow (circuit breaking) in addition to 503 response code.
- **NR**: No route configured for a given request in addition to 404 response code, or no matching filter chain for a downstream connection.
- **URX**: The request was rejected because the upstream retry limit (HTTP) or maximum connect attempts (TCP) was reached.

HTTP only

- **DC**: Downstream connection termination.
- **LH**: Local service failed health check request in addition to 503 response code.
- **UT**: Upstream request timeout in addition to 504 response code.
- **LR**: Connection local reset in addition to 503 response code.
- **UR**: Upstream remote reset in addition to 503 response code.
- **UC**: Upstream connection termination in addition to 503 response code.
- **DI**: The request processing was delayed for a period specified via fault injection.
- **FI**: The request was aborted with a response code specified via fault injection.
- **RL**: The request was ratelimited locally by the HTTP rate limit filter in addition to 429 response code.
- **UAEX**: The request was denied by the external authorization service.
- **RLSE**: The request was rejected because there was an error in rate limit service.
- **IH**: The request was rejected because it set an invalid value for a strictly-checked header in addition to 400 response code.
- **SI**: Stream idle timeout in addition to 408 response code.
- **DPE**: The downstream request had an HTTP protocol error.
- **UMSDR**: The upstream request reached to max stream duration.

Envoy Debug Logging

- Debug logging is verbose and expensive
  - Not recommended for permanent production usage
  - Defaults to “warning”
- Enable debug logging for a workload using istioctl
  - **Doesn’t require restarting the pod**

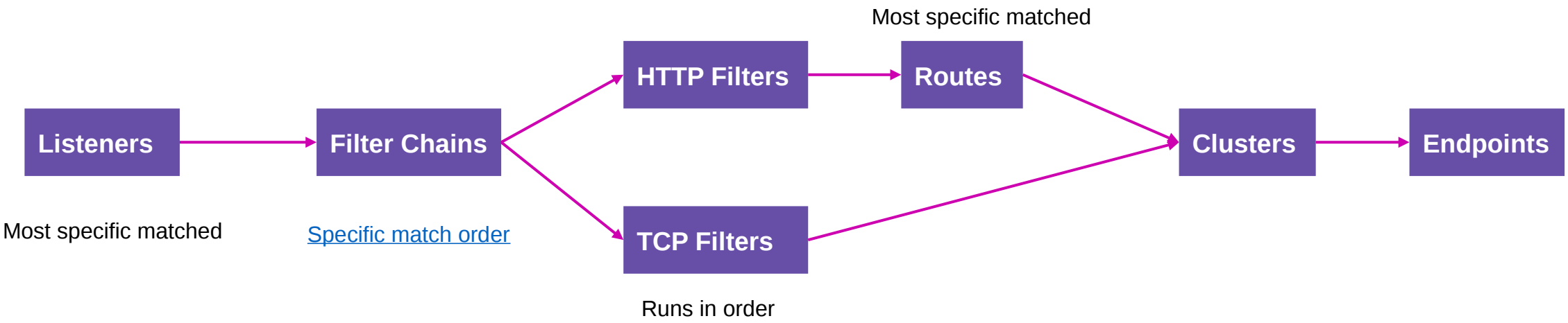
```
istioctl proxy-config log details-v1-78db589446-5q6cz.default --level debug
```

- Enable debug logging for a workload using annotations
  - Add pod annotation “sidecar.istio.io/logLevel: debug”
  - Requires pod restarts
- Enable debug logging globally

```
istioctl install --set profile=demo --set values.global.proxy.logLevel=debug
```

  - Requires pod restarts

Envoy Configuration Overview



Envoy Listener Dump

- Virtual inbound listener

```
istioctl proxy-config listeners details-v1-78db589446-5q6cz.default --address 0.0.0.0 --port 15006 -o json
```

- Virtual outbound listener

```
istioctl proxy-config listeners details-v1-78db589446-5q6cz.default --address 0.0.0.0 --port 15001 -o json
```

- Outbound listeners

```
istioctl proxy-config listeners details-v1-78db589446-5q6cz.default --address 0.0.0.0 --port 80 -o json
```

- All listeners

```
istioctl proxy-config listeners details-v1-78db589446-5q6cz.default -o json
```

Envoy HTTP Listener Configuration


```
{
    "name": "0.0.0.0_80",
    "address": {
        "socketAddress": {
            "address": "0.0.0.0",
            "portValue": 80
        }
    },
    "filterChains": [
        {
            "filterChainMatch": {
                "applicationProtocols": [
                    "http/1.0",
                    "http/1.1",
                    "h2c"
                ]
            },
            "filters": [
                {
                    "name": "envoy.http_connection_manager",
                    "typedConfig": {
                        "@type": "type.googleapis.com/envoy.config.filter.network.http_connection_manager.v2.HttpConnectionManager",
                        "statPrefix": "outbound_0.0.0.0_80",
                        "rds": {
                            "configSource": {
                                "ads": {}
                            },
                            "routeConfigName": "80"
                        },
                        "httpFilters": [
                            {
                                "name": "istio.metadata_exchange",
                                "typedConfig": {
                                    "@type": "type.googleapis.com/udpa.type.v1.TypedStruct",
                                    "typeUrl": "type.googleapis.com/envoy.extensions.filters.http.wasm.v3.Wasm",
                                    "value": {
                                        "config": {
                                            "configuration": "{}\n",
                                            "vm_config": {
                                                "code": {
                                                    "local": {
                                                        "inline_string": "envoy.wasm.metadata_exchange"
                                                    }
                                                },
                                                "runtime": "envoy.wasm.runtime.null"
                                            }
                                        }
                                    }
                                }
                            },
```

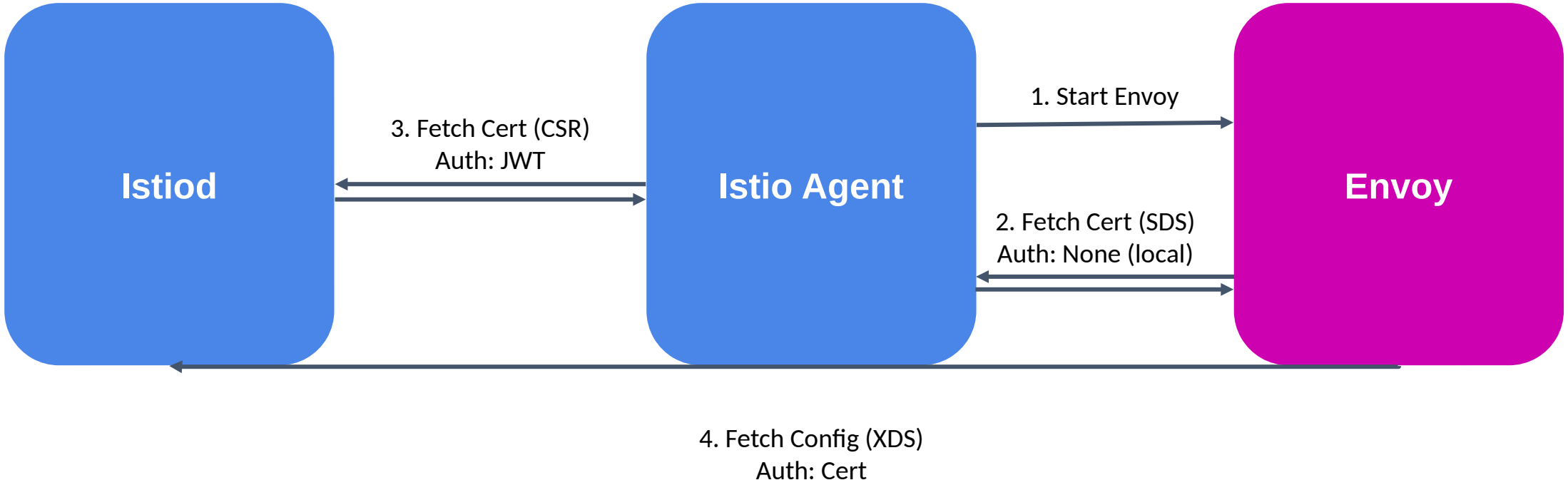
Envoy TCP Listener Configuration

```
,,
{
    "name": "envoy.tcp_proxy",
    "typedConfig": {
        "@type": "type.googleapis.com/envoy.config.filter.network.tcp_proxy.v2.TcpProxy",
        "statPrefix": "outbound|15012||istiod.istio-system.svc.cluster.local",
        "cluster": "outbound|15012||istiod.istio-system.svc.cluster.local",
        "accessLog": [
            {
                "name": "envoy.file_access_log",
                "typedConfig": {
                    "@type": "type.googleapis.com/envoy.config.accesslog.v2.FileAccessLog",
                    "path": "/dev/stdout",
                    "format": "[%START_TIME%] \"%REQ(:METHOD)% %REQ(X-ENVOY-ORIGINAL-PATH?:PATH)% %
DATA(istio.mixer:status)%\" \"%UPSTREAM_TRANSPORT_FAILURE_REASON%\" %BYTES_RECEIVED% %BYTES_SENT% %DU
DED-FOR)%\" \"%REQ(USER-AGENT)%\" \"%REQ(X-REQUEST-ID)%\" \"%REQ(:AUTHORITY)%\" \"%UPSTREAM_HOST%\" %
ADDRESS% %DOWNSTREAM_REMOTE_ADDRESS% %REQUESTED_SERVER_NAME% %ROUTE_NAME%\n"
                }
            }
        ]
    }
}
```

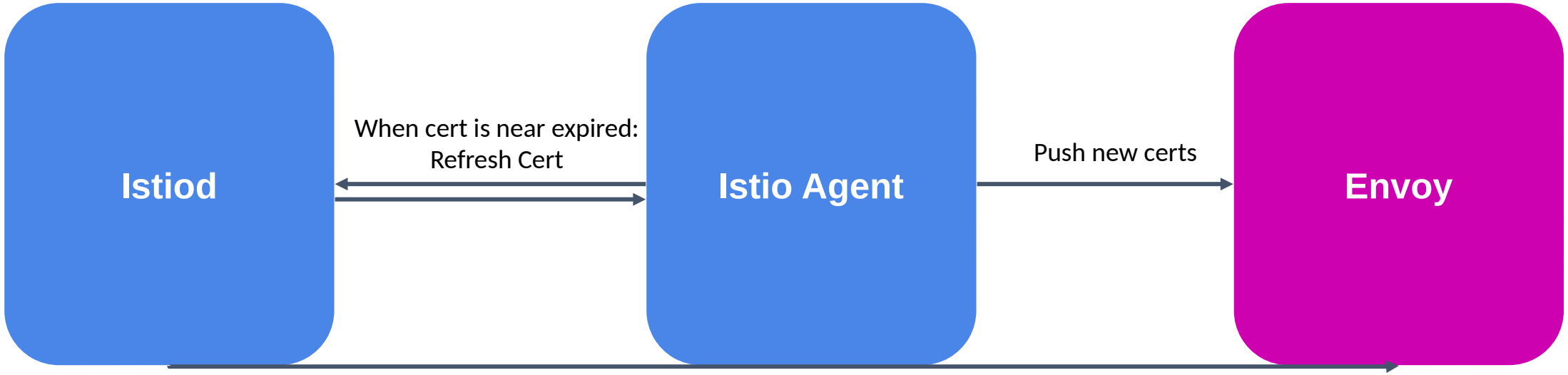
Mapping Resources to Envoy Configuration

Resource Type	Envoy Configuration	Notes
Kubernetes Services	Listeners Routes Clusters	New listeners if port/protocol combo is unique Add virtual hosts for existing routes One cluster per Service/Port/Subset
Kubernetes Endpoints	Endpoints	
Istio Gateway	Listeners	Apply to Ingress/Egress Gateways
Istio VirtualService	Listeners Routes	Client side proxies TLS/TCP affect listeners HTTP match blocks affect routes
Istio DestinationRule	Clusters Endpoints	Client side proxies Connection/HTTP/TLS settings
Istio ServiceEntry	Clusters Endpoints	Client side proxies
Istio PeerAuthentication	Listeners Clusters	Server side proxies
Istio RequestAuthentication	Listeners	Server side proxies
Istio Authorization Policies	Listeners	Server side proxies
Istio EnvoyFilter	All	Break glass API to directly manipulate Envoy
Istio Sidecar	All	Client or server side proxies Sidecar scope sets config visibility

Istio Agent Bootstrapping



Istio Agent Lifecycle



Periodic reconnect based on
max conn age/keep alives

When config/endpoints change:
Push updated config

Istio Agent Debugging

- Viewing secrets/certs using istioctl
  - `istioctl proxy-config secret <POD>`
  - `default` secret is the workload certificate
  - `ROOTCA` is the workload root certificate
  - Other secrets are certificates from DestinationRule/Gateway
  - Describe the full certificate information with openssl
    - ```
      istioctl proxy-config secret <POD> -o json | jq
      '.dynamicActiveSecrets[0].secret.tlsCertificate.certificateChain.inlineBytes'
      -r | base64 -d | openssl x509 -noout -text -in -
      ```

- "Warming" secrets
  - If a secret isn't found it will be stuck warming. Traffic that relies on that secret will fail.
  - For gateways, may indicate the referenced Secret is not present. Check logs or `istioctl analyze`.
  - For workload certificate, may indicate issues connecting to the CA (Istiod).
      ```

Istio Agent Readiness

- Probe: every 2s, check readiness status at localhost:15021/healthz/ready
  - On success, mark container ready
  - On 30 consecutive failures, mark container not ready
  - Runs for entire lifetime of the pod
- Envoy proxy is NOT ready: config not received from Pilot
  - Indicates we don't have the correct XDS configuration
  - Can be from rejected config
    - Envoy log: gRPC config for type.googleapis.com/... rejected
    - istiod log: ADS: ACK ERROR ...
    - Metrics: pilot_xds_{cds,lds,eds,rds}_reject
    - Generally a bug cause by non-standard configuration
  - Can be from connectivity issues to istiod
    - StreamAggregatedResources gRPC config stream closed: 14, no healthy upstream
    - Check certificates are ready
    - Check istiod is ready
    - Check Network Policies if applicable
- Envoy proxy is ready - Output on first successful probe

Istiod Debugging

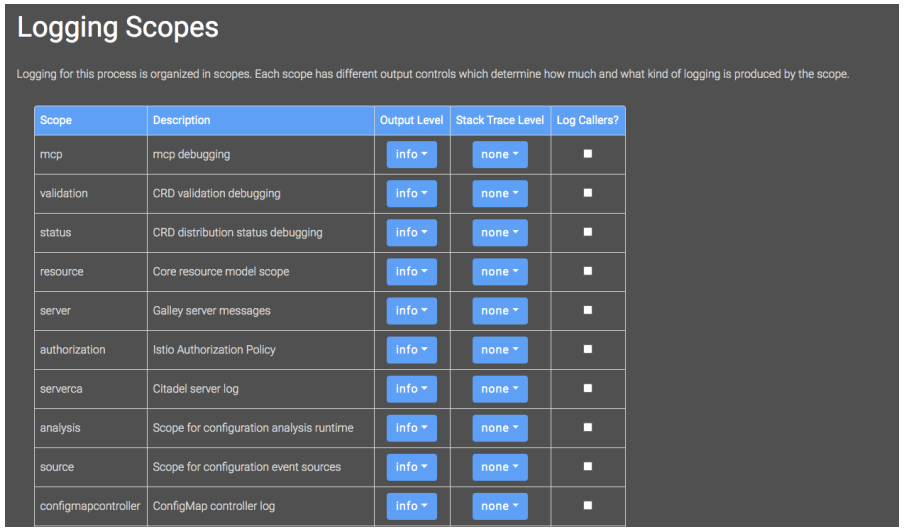
- Enabling debug logging
- ControlZ dashboard
- Istiod metrics for config synchronization/pushes

Istiod Debug Logging

- Default level is “info”.Available options: “debug”, “info”, “warn”, “error”, “fatal”, “none”
- Enable debug logging without restarting “istiod”

```
istioctl dashboard controlz istiod-7968744c5b-4bjkr.istio-system
```

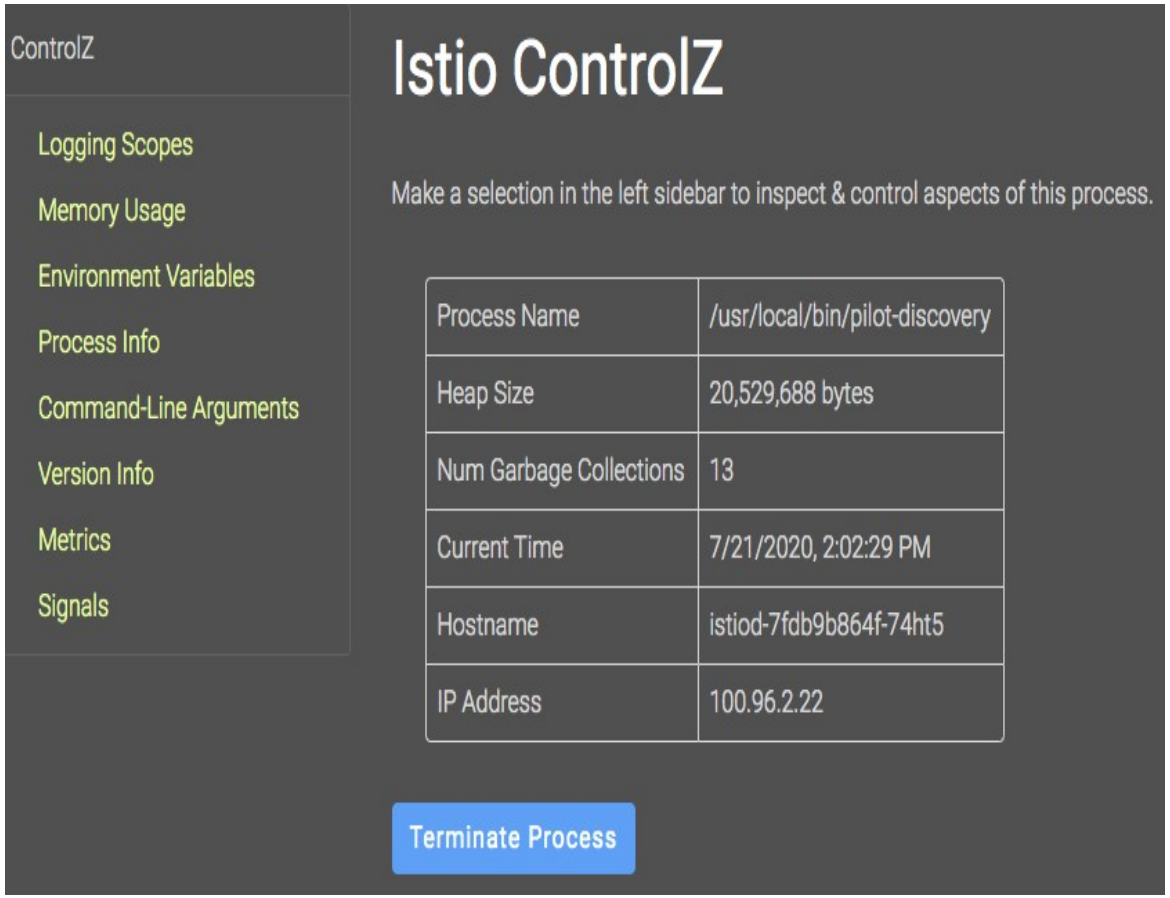
  - Toggle logging for a scope or for all scopes

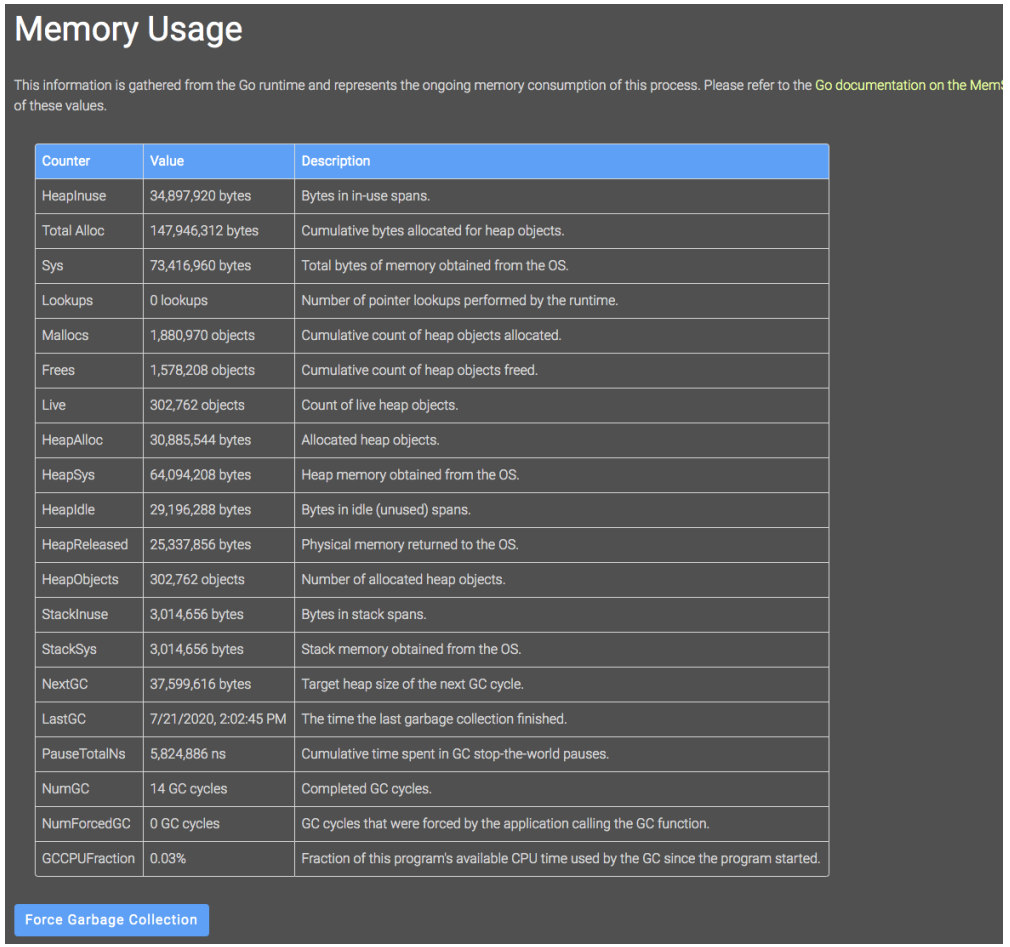


- Enable debug logging via “istioctl”
  - `istioctl install --set profile=demo --set values.global.logging.level=debug`

Istiod ControlZ Dashboard

```
istioctl dashboard controlz istiod-7968744c5b-4bjkr.istio-system
```



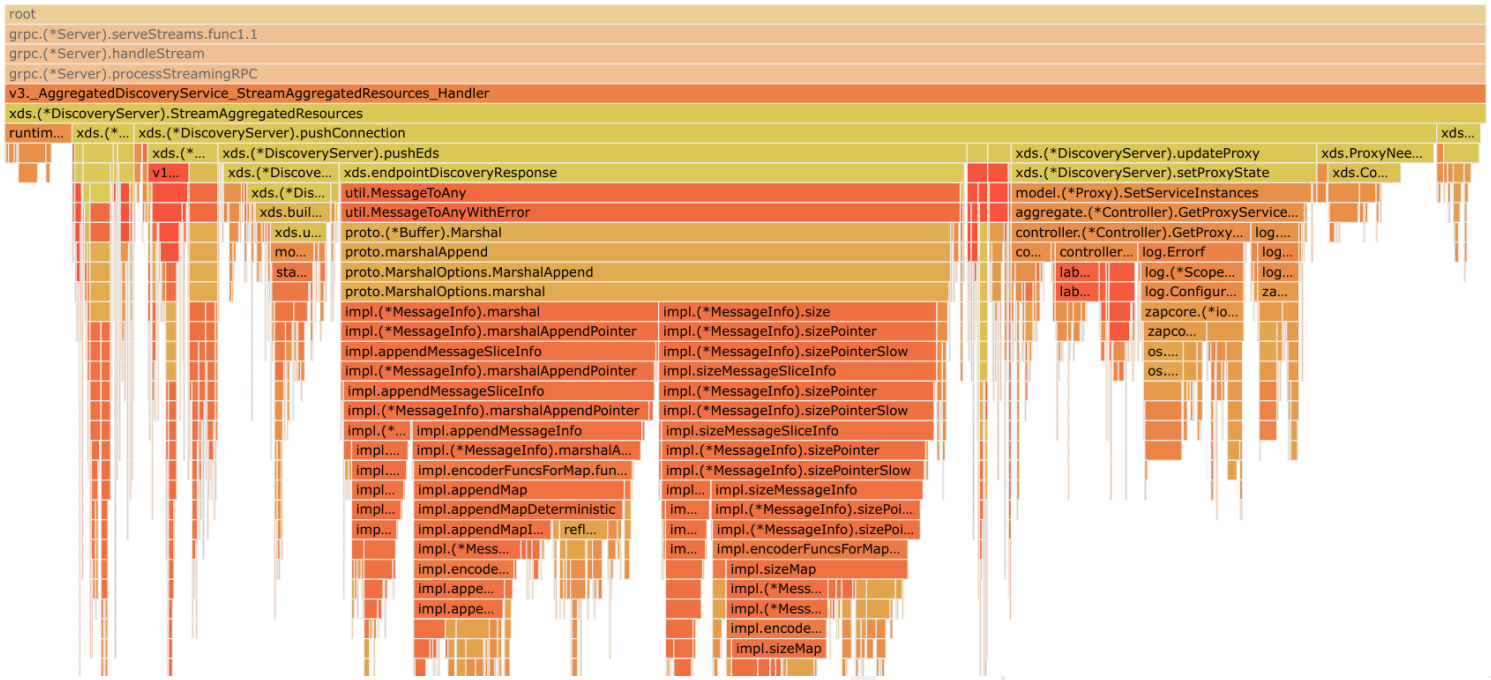


Istiod Metrics

- Grafana Dashboard comes prepopulated with useful metrics to track.
- Datadog has an excellent deep dive on Istio metrics.
- Useful metrics
  - Total number of invalid config: `pilot_total_xds_rejects`
    - Rejections generally indicate misconfiguration or bugs
  - Time to push config update to proxy: `pilot_proxy_convergence_time`
    - Slow times may lead to delays in config or endpoints updates. May indicate under-scaled control plane.
  - Number of XDS clients: `pilot_xds`
    - Useful to spot unbalanced load distribution

Profiling

- Both control plane and data plane support profiling which can help identify CPU and memory usages.
- `go tool pprof -http=: localhost:8080/debug/pprof/<ENDPOINT>`
  - profile: Captures 30s of CPU usage
  - heap: Captures in use and lifetime memory allocations
  - goroutine: Capture all goroutines in use



CLOUD NATIVE COMPUTING FOUNDATION

ASPEN MESH

Istioctl Debugging Capabilities

- Proxy status cluster-wide (Per-proxy view gives detailed config “diffs”)

```
istioctl proxy-status
```

NAME	CDS	LDS	EDS	RDS	PILOT	VERSION
details-v1-78db589446-5q6cz.default	SYNCED	SYNCED	SYNCED	SYNCED	istiod-7968744c5b-hlvgm	1.6.5
httpbin-779c54bf49-5wnmg.foo	SYNCED	SYNCED	SYNCED	SYNCED	istiod-7968744c5b-hlvgm	1.6.5
istio-egressgateway-57b9c7d65b-8xh22.istio-system	SYNCED	SYNCED	SYNCED	NOT SENT	istiod-7968744c5b-hlvgm	1.6.5
istio-ingressgateway-5c8b785c88-b64db.istio-system	SYNCED	SYNCED	SYNCED	NOT SENT	istiod-7968744c5b-hlvgm	1.6.5
productpage-v1-7f4cc988c6-44tlj.default	SYNCED	SYNCED	SYNCED	SYNCED	istiod-7968744c5b-hlvgm	1.6.5
prometheus-7c4b6d955-65t6s.istio-system	SYNCED	SYNCED	SYNCED	SYNCED	istiod-7968744c5b-hlvgm	1.6.5
ratings-v1-756b788d54-f9k4g.default	SYNCED	SYNCED	SYNCED	SYNCED	istiod-7968744c5b-hlvgm	1.6.5
reviews-v1-849fcdfd8b-rs5s7.default	SYNCED	SYNCED	SYNCED	SYNCED	istiod-7968744c5b-hlvgm	1.6.5
reviews-v2-5b6fb6c4fb-pb9hr.default	SYNCED	SYNCED	SYNCED	SYNCED	istiod-7968744c5b-hlvgm	1.6.5
reviews-v3-7d94d58566-v2qdc.default	SYNCED	SYNCED	SYNCED	SYNCED	istiod-7968744c5b-hlvgm	1.6.5
sleep-f8cbf5b76-htrpd.foo	SYNCED	SYNCED	SYNCED	SYNCED	istiod-7968744c5b-hlvgm	1.6.5
traffic-generator-productpage-fc97f5595-pmrmd.default	SYNCED	SYNCED	SYNCED	SYNCED	istiod-7968744c5b-hlvgm	1.6.5

- Wait for config distribution (experimental)

```
istioctl experimental wait --for=distribution gateway bookinfo-gateway.default
```

Istioctl Configuration Analysis

- Provides multi resource validation e.g. “Is a VirtualService tied to a Gateway that doesn’t

```
istioctl analyze -A

Error [IST0101] (VirtualService bookinfo.default) Referenced gateway not found: "bookinfos-gateway"
Error: Analyzers found issues when analyzing namespace: default.
See https://istio.io/docs/reference/config/analysis for more information about causes and resolutions.
```

- Complete list of analyzers

- Enable analyzer reporting via CRD status field

```
istioctl install --set profile=demo --set values.global.istiod.enableAnalysis=true

status:
  validationMessages:
  - code: IST0101
    documentation_url: https://istio.io/docs/reference/config/analysis/IST0101?ref=status-controller
    level: Error
    message: 'Referenced gateway not found: "bookinfos-gateway"'
```

Common Problems

Scaling Istiod

- Istiod can scale horizontally and vertically
- Ensure CPU limit is disabled or give sufficient CPU
  - Throttling may lead to slow configuration updates
- Use the latest Istio version!
  - Each release has substantial performance improvements

Factors for Istiod Scaling

- Size of config to generate
  - Impacted by total number of services/pods & Istio resources
  - For large scale, Sidecar scoping is recommended to bring down size of config
- Rate of change of environment
  - Every time a new Service is created or Istio configuration is changed full updates are sent to proxies
  - Adding new endpoints are cheap as only incremental updates are sent
- Number of proxies for which configuration needs to be generated
  - Impacted by number of pods with a sidecar, and gateways

Unbalanced Istiod Load

- xDS connection is a long lived gRPC stream
  - Makes load balancing challenging
- Connection will close every 30m to slowly rebalance load
- Quick scaling events may cause new instances to have small load
  - This can trigger HPA to scale up and down repeatedly
  - Running at least 2 replicas can help alleviate this issue

Understanding mTLS

DestinationRule: what type of traffic is sent

- mode: DISABLE sends plain text. Common for services outside of the mesh
- mode: ISTIO_MUTUAL sends mTLS
- mode: SIMPLE/MUTUAL can be used to originate TLS

PeerAuthentication: what type of traffic is accepted

- mode: DISABLE accept only plain text
- mode: STRICT accept only mTLS
- mode: PERMISSIVE accept mTLS *or* plain text

Understanding Auto mTLS

Auto mTLS: what type of traffic is sent

- If DestinationRule is configured, use DestinationRule settings
- If server has a sidecar and PeerAuthentication allows mTLS, send mTLS
- Otherwise, send plain text

PeerAuthentication: what type of traffic is accepted

- mode: DISABLE accept only plain text
- mode: STRICT accept only mTLS.
- mode: PERMISSIVE accept mTLS *or* plain text.

Troubleshooting Guide

“istioctl analyze”	“istioctl proxy-status”	“istiod” logs	Access logs Resp flags	Envoy config dump	“istiod” metrics	“debug” logging	Profiling

Production Istio Installation

- Metrics & logs from control & data plane
  - Setup alerts
- Enable Access logs
- Outbound traffic control
- Strict mTLS instead of “auto”
- Scale out control plane
  - Configure HPA
  - Configure pod anti-affinity
- Non self signed CA certificates
- Locking down Ingress GW ports
- Auto sidecar injection
- Production grade Prometheus & Jaeger

Questions?

Neeraj Poddar
@nrjpoddar
neeraj@aspenmesh.io

CLOUD NATIVE COMPUTING FOUNDATION
ASPEN MESH