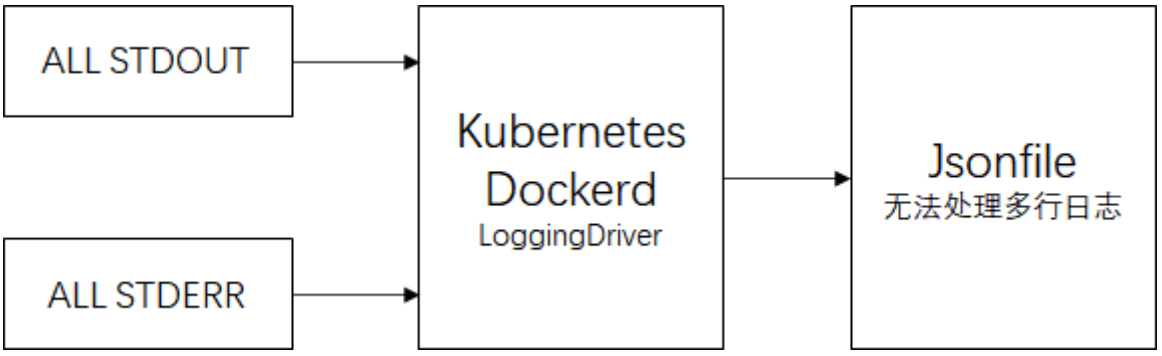
EFK生态

大数据时代背景下，应用数据是业务分析的核心，同时日志是应用Bug分析，监控集群状态的数据基础。

Kubernetes不提供远程日志存储解决方案

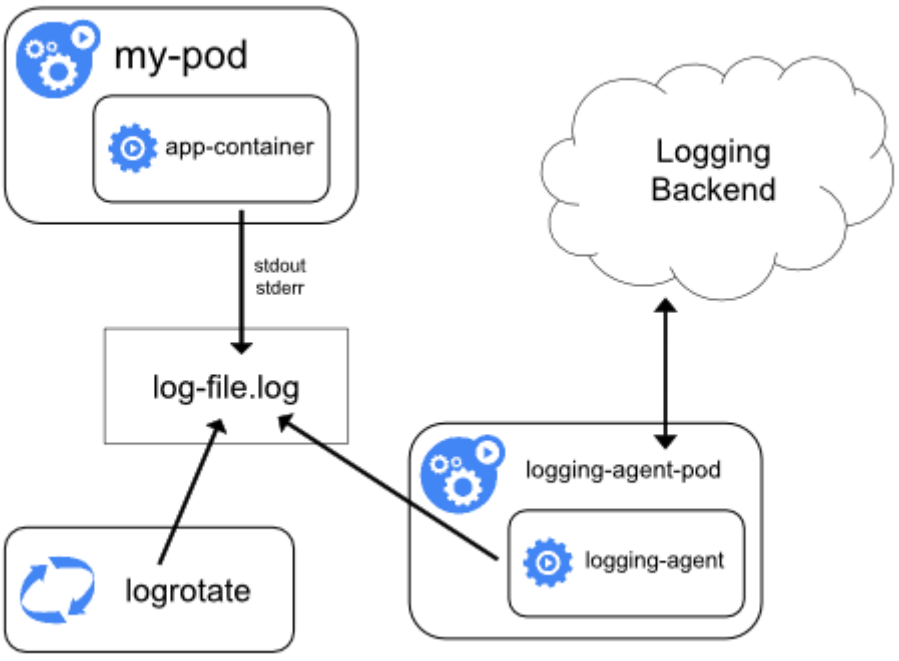


默认情况下container重启，保留原有container日志，POD被驱逐，则原POD相关日志目录同样被删除

集群内日志采集架构方案：

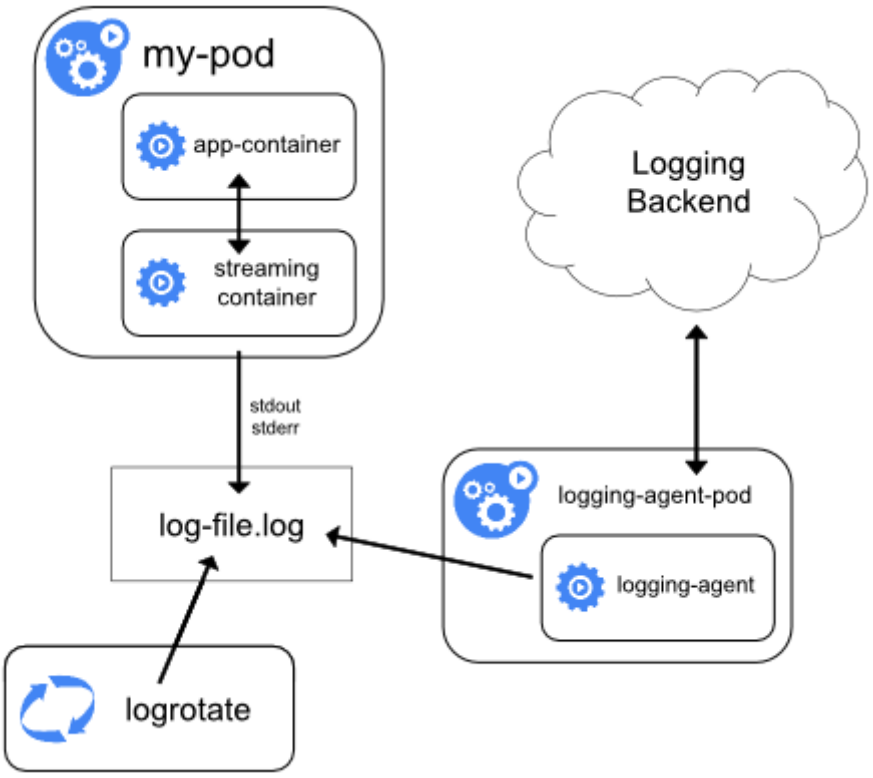
- 节点级日志代理
- Pod Sidecar容器
- 直接投递

节点级代理

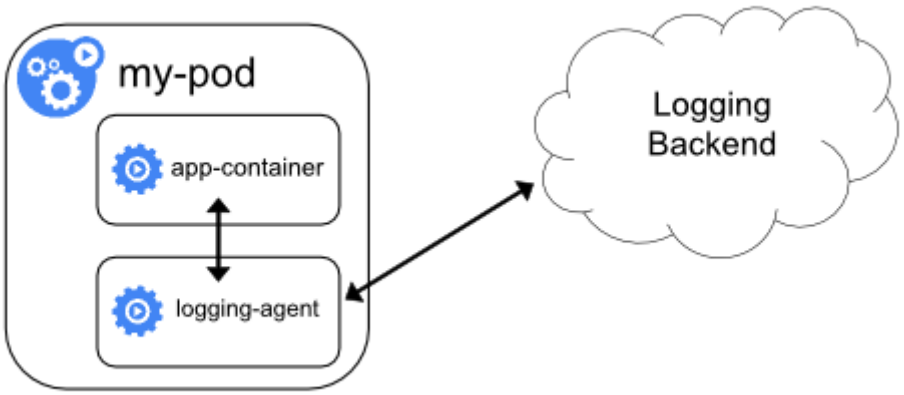


Sidecar容器

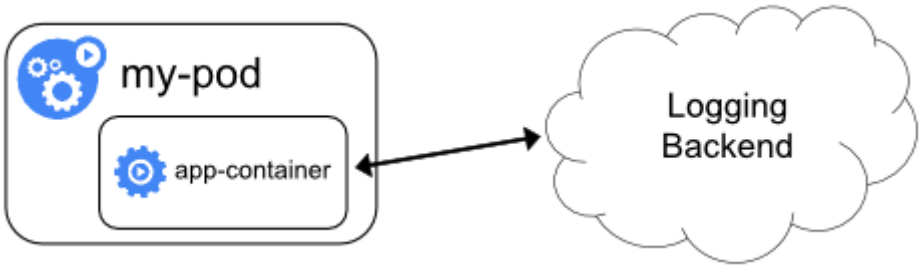
日志流处理



日志代理Sidecar



直接投递



EFK stack

- Elasticsearch分布式存储，检索系统
- Fluentd（日志采集组件，多种实现）
- Kibana（日志分析、展示平台）

ElasticSearch StatefulSet

servcie 用以固定地址

```yaml
kind: Service
apiVersion: v1
metadata:
  name: elasticsearch
  namespace: efk
  labels:
    app: elasticsearch
spec:
  selector:
    app: elasticsearch
  ports:
    - port: 9200
      name: rest
    - port: 9300
      name: inter-node
  type: ClusterIP
```

statefulset工作负载

```yaml
apiVersion: apps/v1              # 三要素
kind: StatefulSet                # 三要素
metadata:                        # 三要素
  name: es-cluster
  namespace: efk
spec:                            # 资源结构
  serviceName: elasticsearch     # 指定用以固定地址的service name
  replicas: 3                    # 有状态节点数量
  selector:
    matchLabels:
```

```
      app: elasticsearch
  template:
    metadata:
      labels:
        app: elasticsearch
    spec:
      containers:
      - name: elasticsearch
        image: elasticsearch:7.11.2
        resources:
            limits:
              cpu: 1000m
            requests:
              cpu: 100m
        ports:
        - containerPort: 9200
          name: rest
          protocol: TCP
        - containerPort: 9300
          name: inter-node
          protocol: TCP
        volumeMounts:
        - name: data
          mountPath: /usr/share/elasticsearch/data
        env:
          - name: cluster.name
            value: kubelogs
          - name: node.name
            valueFrom:
              fieldRef:
                fieldPath: metadata.name
          - name: discovery.seed_hosts
            value: "es-cluster-0.elasticsearch,es-cluster-1.elasticsearch,es-cluster-
2.elasticsearch"
          - name: cluster.initial_master_nodes
            value: "es-cluster-0,es-cluster-1,es-cluster-2"
          - name: ES_JAVA_OPTS
            value: "-Xms4096m -Xmx4096m"
      initContainers:
      - name: fix-permissions
        image: busybox
        command: ["sh", "-c", "chown -R 1000:1000 /usr/share/elasticsearch/data"]
        securityContext:
          privileged: true
        volumeMounts:
        - name: data
          mountPath: /usr/share/elasticsearch/data
      - name: increase-vm-max-map
        image: busybox
        command: ["sysctl", "-w", "vm.max_map_count=262144"]
        securityContext:
          privileged: true
      - name: increase-fd-ulimit
        image: busybox
        command: ["sh", "-c", "ulimit -n 65536"]
        securityContext:
          privileged: true
  volumeClaimTemplates:                    # 持久化存储生命
```

```
  - metadata:
      name: data
      labels:
        app: elasticsearch
    spec:
      accessModes: [ "ReadWriteOnce" ]
      storageClassName: rook-es
      resources:
        requests:
          storage: 500Gi
```

生成的PVC

```
NAME                STATUS   VOLUME                                     CAPACITY   ACCESS
MODES   STORAGECLASS   AGE
data-es-cluster-0   Bound    pvc-6599e2f4-53cd-427a-8213-1e82303f3af9   500Gi      RWO
rook-es        120d
data-es-cluster-1   Bound    pvc-35695491-1361-494f-9f6d-ef806755bd0d   500Gi      RWO
rook-es        120d
data-es-cluster-2   Bound    pvc-866583f6-ceea-46b1-9a79-d220836a5533   500Gi      RWO
rook-es        120d
```

Statefulset的固定地址

```
[root@centos-tools-68c5975c76-7j4rj /]# nslookup es-cluster-0.elasticsearch.efk.svc
Server:     10.96.0.10
Address:    10.96.0.10#53

Name:   es-cluster-0.elasticsearch.efk.svc.cluster.local    #
PODNAME.SERVICE.NAMESPACE.svc.cluster.local
Address: 10.244.100.228
```

Fluentd

节点级代理配置

```
---
apiVersion: v1
kind: ServiceAccount
metadata:
  name: fluentd
  namespace: efk

---
apiVersion: rbac.authorization.k8s.io/v1
kind: ClusterRole
metadata:
```

```
  name: fluentd
  namespace: efk
rules:
- apiGroups:
  - ""
  resources:
  - pods
  - namespaces
  verbs:
  - get
  - list
  - watch

---
kind: ClusterRoleBinding
apiVersion: rbac.authorization.k8s.io/v1
metadata:
  name: fluentd
roleRef:
  kind: ClusterRole
  name: fluentd
  apiGroup: rbac.authorization.k8s.io
subjects:
- kind: ServiceAccount
  name: fluentd
  namespace: efk
---
apiVersion: apps/v1
kind: DaemonSet
metadata:
  name: fluentd
  namespace: efk
  labels:
    k8s-app: fluentd-logging
    version: v1
spec:
  selector:
    matchLabels:
      k8s-app: fluentd-logging
      version: v1
  template:
    metadata:
      labels:
        k8s-app: fluentd-logging
        version: v1
    spec:
      serviceAccount: fluentd
      serviceAccountName: fluentd
      tolerations:
      - key: node-role.kubernetes.io/master
        effect: NoSchedule
      containers:
      - name: fluentd
        image: fluent/fluentd-kubernetes-daemonset:v1-debian-elasticsearch
        env:
          - name:  FLUENT_ELASTICSEARCH_HOST
            value: "elasticsearch"
          - name:  FLUENT_ELASTICSEARCH_PORT
```

```
            value: "9200"
          - name: FLUENT_ELASTICSEARCH_SCHEME
            value: "http"
          # Option to configure elasticsearch plugin with self signed certs
          # ================================================================
          - name: FLUENT_ELASTICSEARCH_SSL_VERIFY
            value: "false"
          # Option to configure elasticsearch plugin with tls
          # ================================================================
          - name: FLUENT_ELASTICSEARCH_SSL_VERSION
            value: "TLSv1_2"
          - name: FLUENT_ELASTICSEARCH_BUFFER_FLUSH_THREAD_COUNT
            value: "2"
          # X-Pack Authentication
          # =====================
          #- name: FLUENT_ELASTICSEARCH_USER
          #  value: "elastic"
          #- name: FLUENT_ELASTICSEARCH_PASSWORD
          #  value: "changeme"
        resources:
          limits:
            memory: 200Mi
          requests:
            cpu: 100m
            memory: 200Mi
        volumeMounts:
        - name: varlog
          mountPath: /var/log
        - name: varlibdockercontainers
          mountPath: /var/lib/docker/containers
          readOnly: true
        - name: config
          mountPath: /fluentd/etc
          readOnly: true
        - name: localtime
          mountPath: /etc/localtime
      terminationGracePeriodSeconds: 30
      volumes:
      - name: varlog
        hostPath:
          path: /var/log
      - name: varlibdockercontainers
        hostPath:
          path: /var/log
      - name: varlibdockercontainers
        hostPath:
          path: /var/lib/docker/containers
      - name: config
        configMap:
          name: fluentd-conf
      - name: localtime
        hostPath:
          path: /etc/localtime
```

相关配置文件

```
[root@controller01 flutetd]# tree conf/
conf/
├── disable.conf
├── fluent.conf
├── kubernetes.conf
├── prometheus.conf
└── systemd.conf

0 directories, 5 files

[root@controller01 flutetd]# kubectl get configmaps -n efk
NAME                 DATA   AGE
fluentd-conf         5      111d
```

sidecar配置文件

```
apiVersion: v1
kind: ConfigMap
metadata:
  name: fluentd-config
data:
  fluentd.conf: |
    <source>
      type tail
      format none
      path /var/log/1.log
      pos_file /var/log/1.log.pos
      tag count.format1
    </source>

    <source>
      type tail
      format none
      path /var/log/2.log
      pos_file /var/log/2.log.pos
      tag count.format2
    </source>
```

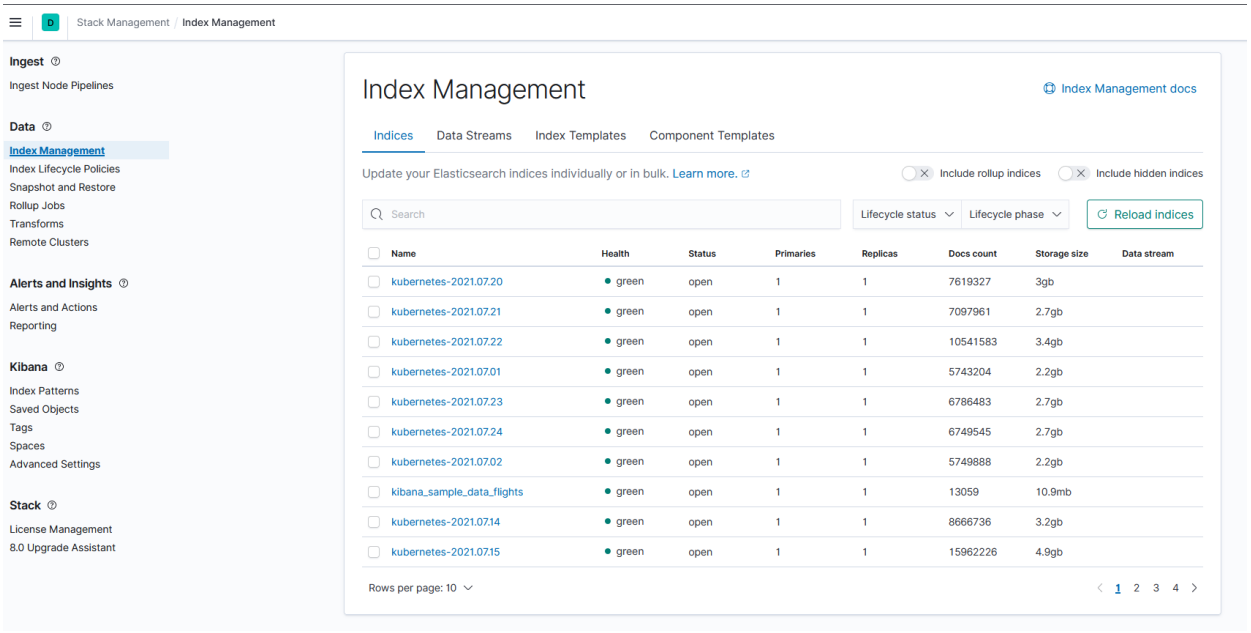
Kibana

```
apiVersion: v1
kind: Service
metadata:
  name: kibana
  namespace: efk
  labels:
    app: kibana
spec:
  ports:
  - port: 5601
```

```
  selector:
    app: kibana
---
apiVersion: apps/v1
kind: Deployment
metadata:
  name: kibana
  namespace: efk
  labels:
    app: kibana
spec:
  replicas: 1
  selector:
    matchLabels:
      app: kibana
  template:
    metadata:
      labels:
        app: kibana
    spec:
      containers:
      - name: kibana
        image: kibana:7.11.2
        resources:
          limits:
            cpu: 1000m
          requests:
            cpu: 100m
        env:
          - name: ELASTICSEARCH_URL
            value: http://elasticsearch:9200
        ports:
        - containerPort: 5601
        volumeMounts:
        - name: localtime
          mountPath: /etc/localtime
      volumes:
      - name: localtime
        hostPath:
          path: /etc/localtime
```

DashBoard

管理ES Index



Prometheus生态

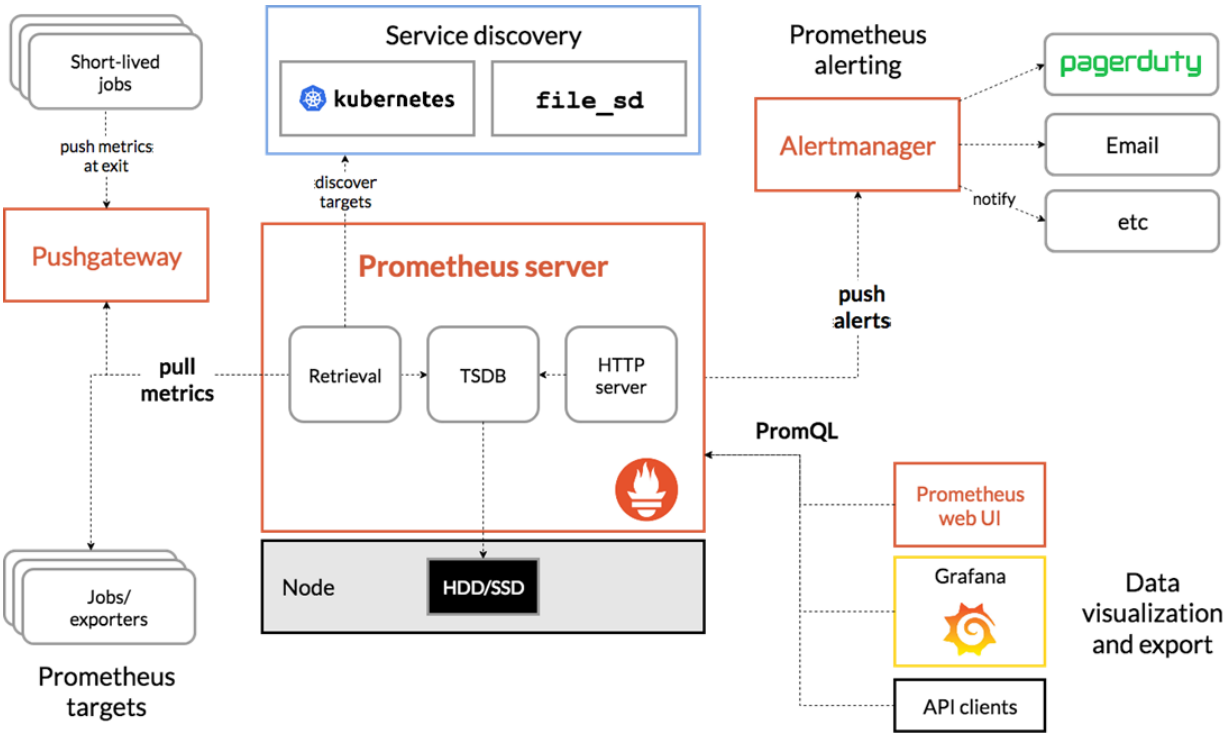
监控系统意义

监控系统提供了系统资源依据时序排列的度量，与日志想类似都是时序数据，但日志旨在描述发生了什么，监控旨在显示某一指标在某一时刻的状态。

组件

- Prometheus Server
- client libraries
- push gateway
- exporters
- alertmanager
- other tools

架构



Helm 部署 kube-prometheus-stack

```
[root@controller01 c_cluster]# tree kube-prometheus-stack -L 2
kube-prometheus-stack
├── Chart.lock
├── charts
│   ├── grafana
│   ├── kube-state-metrics
│   └── prometheus-node-exporter
├── Chart.yaml
├── CONTRIBUTING.md
├── crds
│   ├── crd-alertmanagerconfigs.yaml
│   ├── crd-alertmanagers.yaml
│   ├── crd-podmonitors.yaml
│   ├── crd-probes.yaml
│   ├── crd-prometheuses.yaml
│   ├── crd-prometheusrules.yaml
│   ├── crd-servicemonitors.yaml
│   └── crd-thanosrulers.yaml
├── README.md
├── templates
│   ├── alertmanager
│   ├── exporters
│   ├── grafana
│   ├── _helpers.tpl
│   ├── NOTES.txt
│   ├── prometheus
│   └── prometheus-operator
└── values.yaml
```

从示例系统中，分析values.yaml文件所包含的配置信息