Fluentbit with Opensearch Ingestion

前提条件

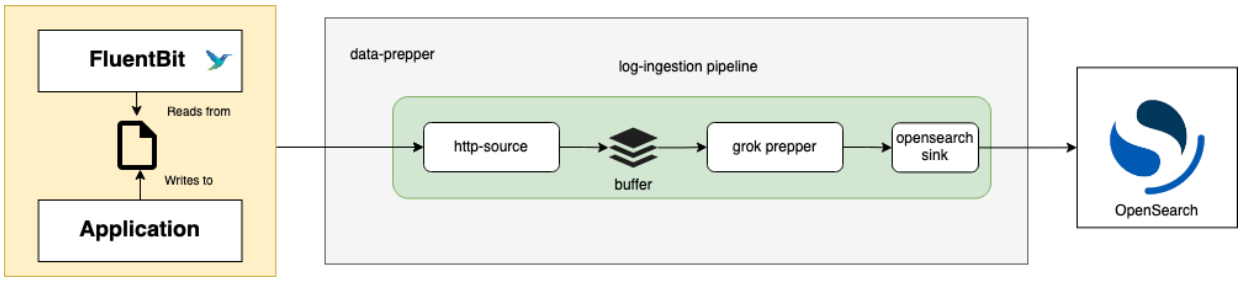
1. 创建好Provisioned Opensearch Domain(VPC Accessed)
    a. 记得修改**Access policy**中的Deny为Allow.
2. 创建一台EC2 作为 Opensearch Dashboard Proxy. (记得安全组放开443端口)

```
sudo yum update
sudo yum install nginx
sudo service nginx start
```

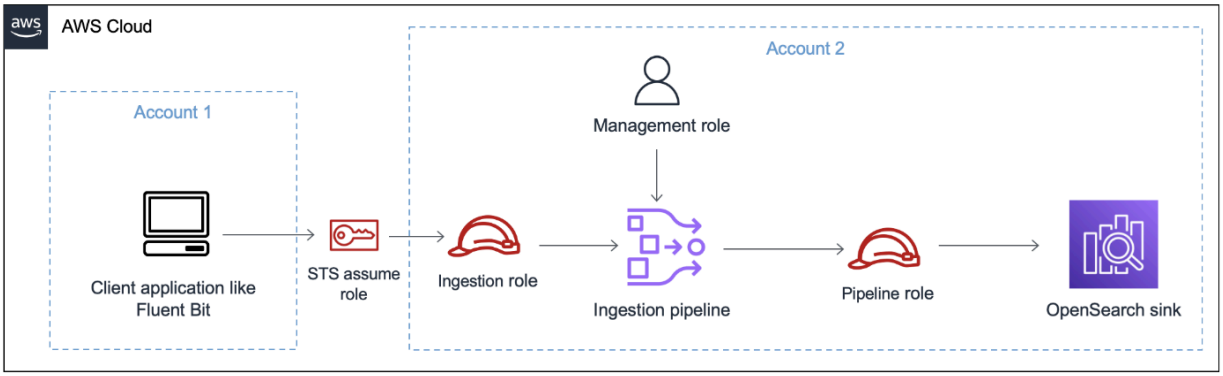
参考`https://github.com/huahua0601/OpenSearch_Dashboard_Nginx_Proxy`来进行配置Proxy, conf.d 下配置

```
server {
    listen 443 ssl;
    server_name localhost;

    ssl_certificate           /etc/nginx/self_signed_certificate.crt;
    ssl_certificate_key       /etc/nginx/public.key;

    location / {
        proxy_pass https://{host};
        proxy_set_header Host $host;
        proxy_set_header X-Real-IP $remote_addr;
        proxy_set_header X-Forwarded-For $proxy_add_x_forwarded_for;
        proxy_set_header X-Forwarded-Proto $scheme;

        # Ensure cookies are rewritten for correct domain
        proxy_cookie_domain {host} $host;

        # Handle redirects properly
        proxy_redirect https://{host} https://$host;
    }
}
```

部署过程

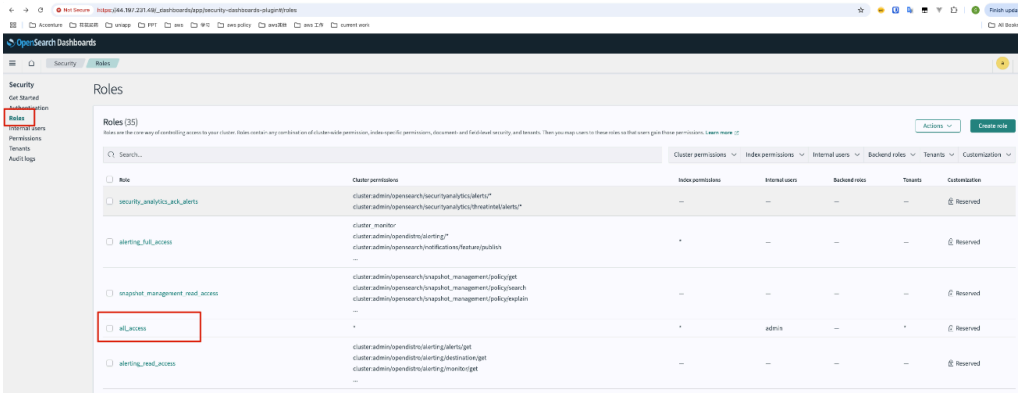


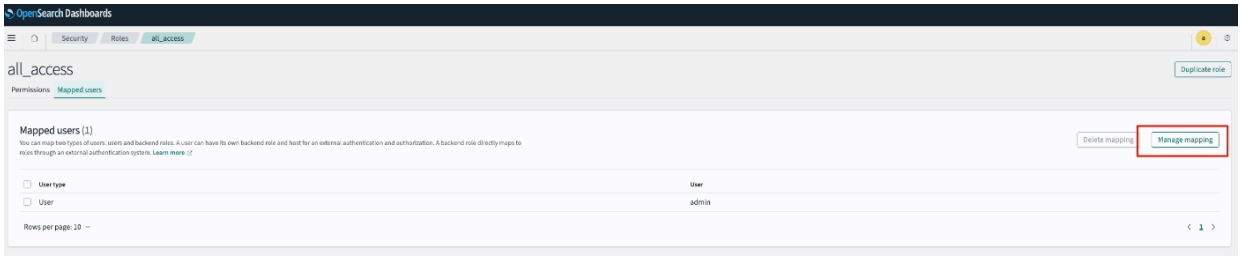
1. 参考文章创建Pipeline role和policy
2. 参考文章创建Ingestion Role



2. 参考文章: https://docs.aws.amazon.com/opensearch-service/latest/developerguide/configure-client-fluentbit.html 创建 Pipeline

3. Opensearch Dashboards 配置

4. 



5. Opensearch Ingestion 配置文件

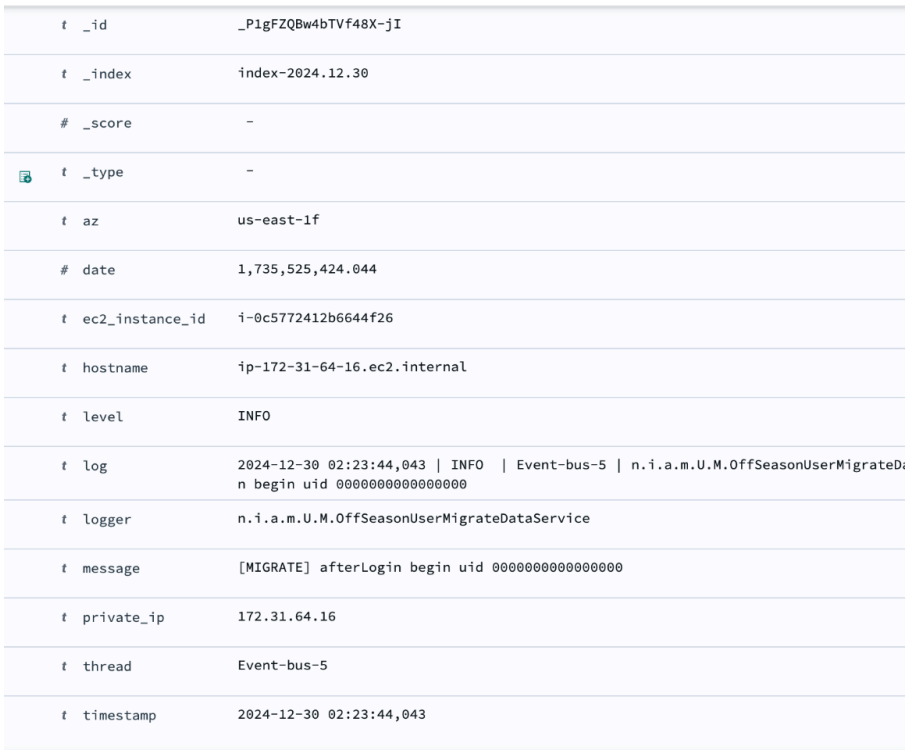
```
{
  "version": "2",
  "unaggregated-log-pipeline": {
    "source": {
      "http": {
        "path": "/log/ingest"
      }
    },
    "processor": [
      {
```

```
        "grok": {
          "match": {
            "log": ["%{TIMESTAMP_ISO8601:timestamp}\\s+\\|\\s+%{LOGLEVEL:level}\\s+\\|
          }
        }
      },
      {
        "date": {
            "from_time_received": true,
            "destination": "@timestamp"
        }
      }
    ],
    "sink": [
      {
        "opensearch": {
          "hosts": [
            "https://vpc-ingestion-test-ar3quly6s3yrh7lkvyqjqqki3q.us-east-1.es.amazon
          ],
          "index": "index_name",
          "index_type": "custom",
          "bulk_size": 20,
          "aws": {
            "sts_role_arn": "arn:aws:iam::715371302281:role/osis-pipeline-role",
            "region": "us-east-1"
          }
        }
      }
    ]
  }
}
```

Fluentbit 配置文件:

```
[SERVICE]
    Flush          5
    Daemon         Off
    Log_Level      info
    log_file       /var/log/fluent-bit.log

[INPUT]
    Name           tail
    Path           /home/ec2-user/test.log
    Tag            app.logs
    read_from_head        true

[FILTER]
    Name           aws
    Match          *
    private_ip     true
    imds_version   v2
    hostname       true
```

```
    ec2_instance_id    true

[OUTPUT]
    Name http
    Match *
    Host fluentbit-pipeline-3xrdun24eq5ewj42z255gy37i4.us-east-1.osis.amazonaws.com
    Port 443
    URI /log/ingest
    aws_auth true
    aws_region us-east-1
    aws_service osis
    Log_Level info
    tls On
```



t	_id	_P1gFZQBw4bTVf48X-jI
t	_index	index-2024.12.30
#	_score	-
t	_type	-
t	az	us-east-1f
#	date	1,735,525,424.044
t	ec2_instance_id	i-0c5772412b6644f26
t	hostname	ip-172-31-64-16.ec2.internal
t	level	INFO
t	log	2024-12-30 02:23:44,043 | INFO | Event-bus-5 | n.i.a.m.U.M.OffSeasonUserMigrateDa n begin uid 0000000000000000
t	logger	n.i.a.m.U.M.OffSeasonUserMigrateDataService
t	message	[MIGRATE] afterLogin begin uid 0000000000000000
t	private_ip	172.31.64.16
t	thread	Event-bus-5
t	timestamp	2024-12-30 02:23:44,043

常用工具:

https://grokconstructor.appspot.com/do/match#result