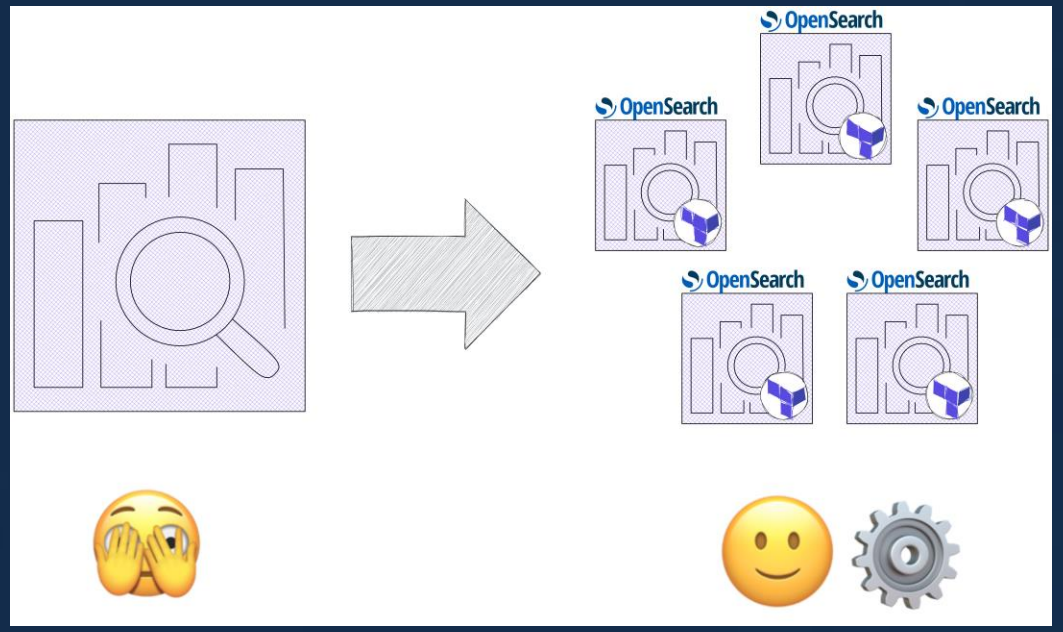
CHASE | J.P.Morgan

Delicacies of Observability: AWS OpenSearch Cluster from 'rare' to 'well-done'

Eugene Tolbakov

WHAT'S ON MENU



Starter:
Introduction

Main Course:
From uber – cluster to 5 log-specific clusters
IaC approach & overview of OpenSearch settings

Dessert:
Key learnings, insights & tips

Nutmeg. a J.P.Morgan company

CHASE

DISCLAIMER

Nutmeg. a J.P.Morgan company

CHASE

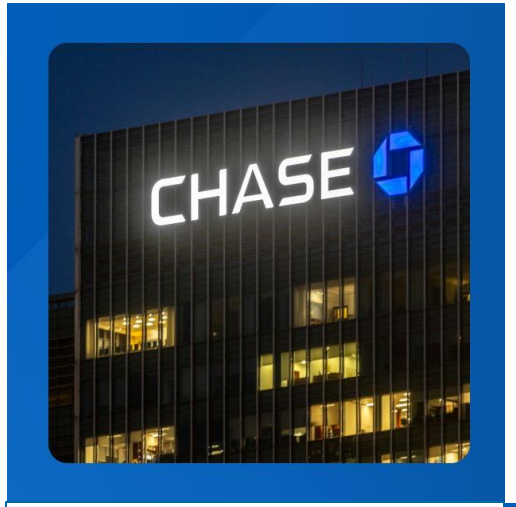
ABOUT ME

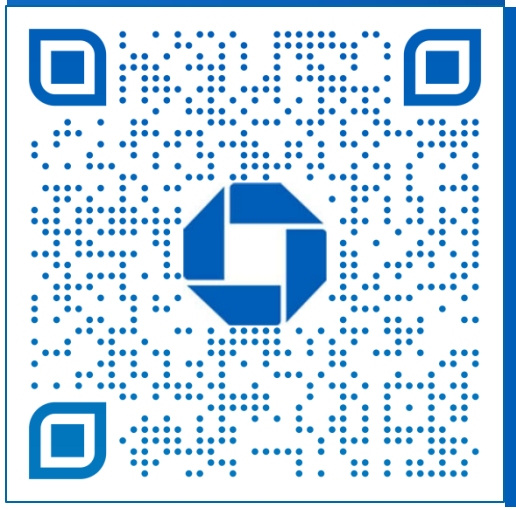

```
GET london-o11y-engineering-meetup-speakers/_doc/4
```


```
{
  "_index": "london-o11y-engineering-meetup-speakers",
  "_id": "4",
  "_version": 1,
  "_seq_no": 0,
  "_primary_term": 1,
  "found": true,
  "_source": {
    "name": "Eugene Tolbakov",
    "organization": "Chase UK",
    "position": "Platform Engineer",
    "skills": [
      "JVM-based languages (Clojure, Kotlin, Java)",
      "Python",
      "Rust"
    ],
    "interests": [
      "Observability",
      "Distributed Systems",
      "Databases"
    ],
    "contacts": {
      "LinkedIn": "Eugene Tolbakov",
      "Twitter/X": "@evtolbakov",
      "Telegram": "@evtolbakov"
    }
  }
}
```


Nutmeg. a J.P.Morgan company CHASE

CHASE | AWARD-WINNING BANKING





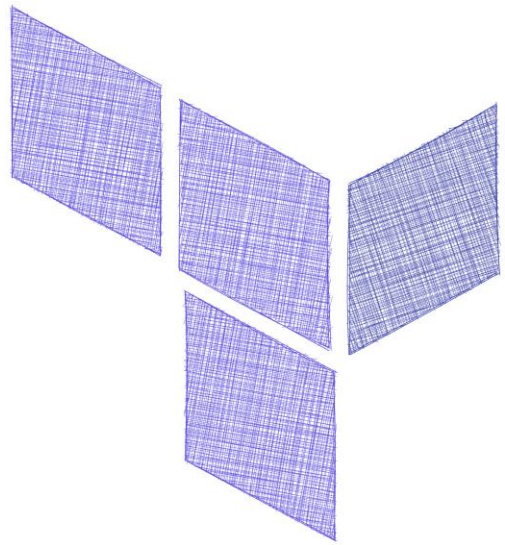

```
GET  british_banks/_doc/1


```


```
{
  "_index": "british_banks",
  "_id": "1",
  "_version": 1,
  "_seq_no": 0,
  "_primary_term": 1,
  "found": true,
  "_source": {
    "name": "Chase UK",
    "launch_date": "09/2021",
    "number_of_customers": "2M",
    "balance": "£20B",
    "products": [
      "current account",
      "savings account",
      "digital wealth & investments"
    ],
    "awards": {
      "2023": [
        "Best savings provider",
        "Best current account provider"
      ],
      "2024": [
        "Best British Bank",
        "Best Current Account Provider"
      ]
    },
    "ratings": {
      "Trustpilot": 4.1,
      "Google Play": 4.8,
      "iOS App Store": 4.9
    }
  }
}
```


Nutmeg. a J.P.Morgan company

CHASE

TERRAFORM 101



- ***What?***
  🛠️ to define(declare) what your infrastructure looks like,
  instead of manually setting everything up (a.k.a. IaC)
- ***Why?***
  consistency, repeatability, collaboration
- ***How?***
  configuration files(HCL) that describes your infrastructure (VMs, storage, networking).
  Terraform figures out how to make those changes.
    - "**modules**" - reuse/avoid repetitive declarations
    - "**providers**" - interact with various APIs
      - ☁️ (AWS, etc)
      - ▯ (Mysql, Postrges)
      - 😳 Spotify / Dominos
      - your_next_best_tool
      - **Elasticsearch | OpenSearch**

Nutmeg. a J.P.Morgan company CHASE

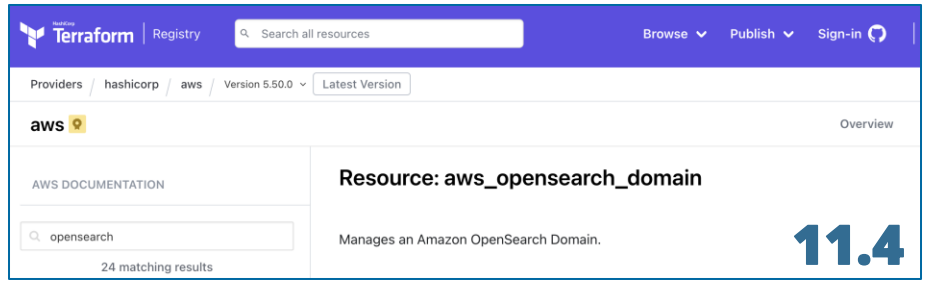
TERRAFORM 101 | AWS_OPENSEARCH_DOMAIN

```
locals {
  cluster_params = {
    cluster-x = {
      engine_version       = "OpenSearch_2.13"
      master_node_type     = "m6g.large.search"
      master_node_number   = 3
      data_node_type       = "r6g.2xlarge.search"
      data_node_number     = 3
    },
    cluster-y = {

    },
    cluster-z = {

    }
  }
}
```

11.1

```
.
├── locals.tf
├── main.tf
└── module
    ├── iam.tf
    ├── logs.tf
    └── opensearch.tf
```

11.3



11.4

```
module "opensearch" {
  source  = "./module/"
  for_each            = local.cluster_params

  cluster_name        = each.key
  engine_version      = each.value["engine_version"]
  master_node_type    = each.value["master_node_type"]
  master_node_number  = each.value["master_node_number"]
  data_node_type      = each.value["data_node_type"]
  data_node_number    = each.value["data_node_number"]
}
```

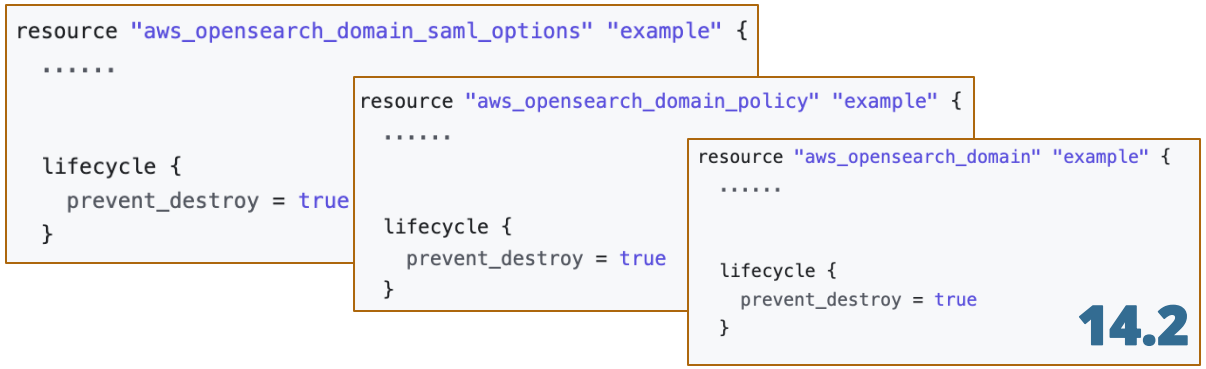
11.2

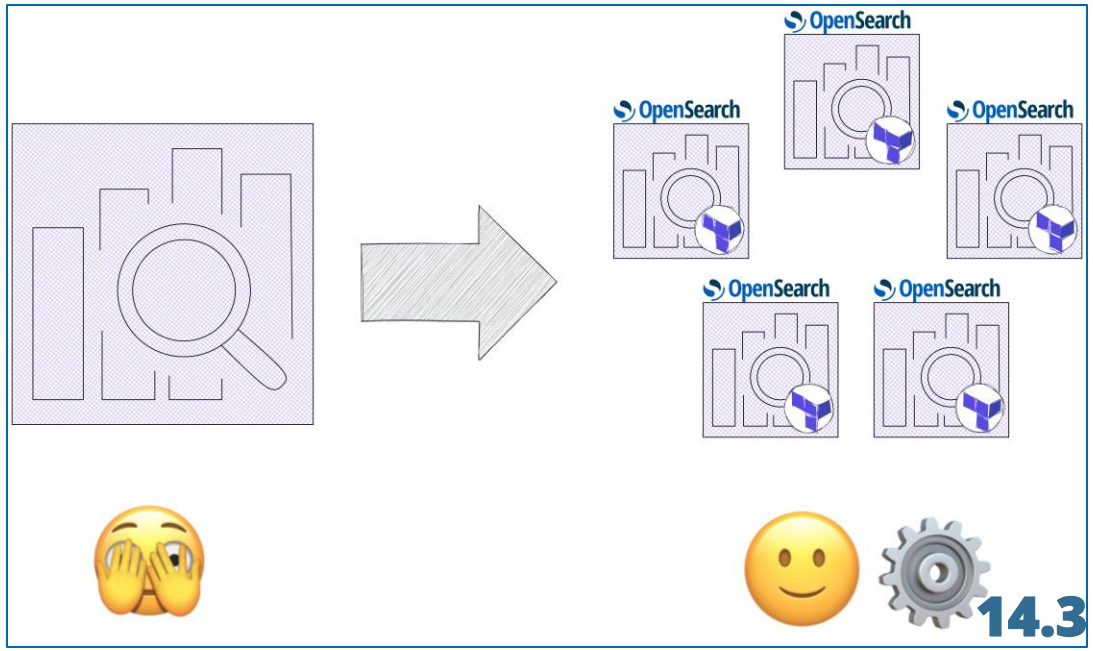
Nutmeg. a J.P.Morgan company

CHASE

SMALLER CLUSTERS

- ❗Bigger overheard
- ✅ Smaller cluster state
- ✅ More stability
- ✅ Smaller blast radius
  - ✓ Migrations 😰 (2.13 improved 👏, Workaround)
  - ✓ Migrations are not "backward" 🦊





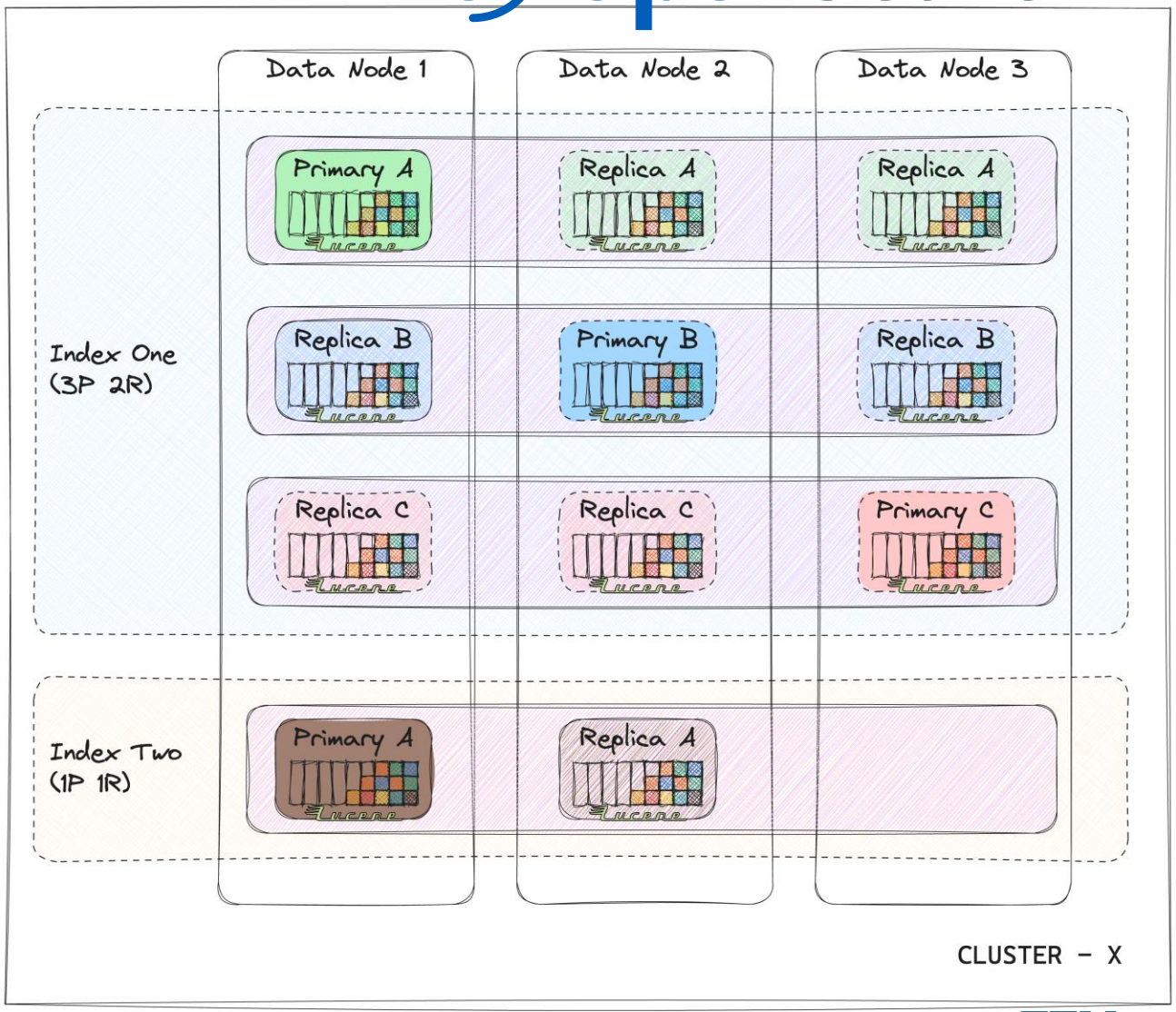
OpenSearch cluster		
Cluster configuration	Data (_cat/allocation?v)	Environment
3 master nodes (r6g.xlarge.search) 24 data nodes (r6g.4xlarge.search)	120TB 70B docs 14k shards 300 indicies / 10 index templates 2.3 - 2.4 billion entries/day	Managed cluster Multi-AZ Engine version 2.11 Indexing Data Rate : 4k/12k (avg/peak) docs/second Average Search Latency: 10ms/500ms (avg/peak) 14.4

Nutmeg. a J.P.Morgan company CHASE

OPENSEARCH 101

- ✓ Apache 2.0;
- ✓ 2021 Elasticsearch (version 7.10.2)
- ✓ Workloads
  - ✓ search engine
  - ✓ logging

- ✓ Cluster
- ✓ Data node
- ✓ Index
- ✓ Shard
- ✓ Segment
- ✓ Document



TERRAFORM 101 | OPENSEARCH RESOURCE

- ✓ 2 providers
- ✓ .tftpl files
- ✓ terraform plan

```yaml
- _id: "index-pattern:my-logs"
  _source:
    index-pattern:
      timeFieldName: "@timestamp"
      title: "${pass_from_terraform}my-logs*"
      fieldFormatMap: "{\"payload_size\":{\"id\":\"bytes\"}}"
    references: []
    type: "index-pattern"
```

19.1

```
module "opensearch_resources" {
  source = "./module"

  index_pattern_files   = fileset(path.module, "../cluster-x-config/index-patterns/*.yml.tftpl")
  index_template_files  = fileset(path.module, "../cluster-x-config/index-templates/*.yml.tftpl")
  ism_policy_files      = fileset(path.module, "../cluster-x-config/ism-policies/*.yml.tftpl")
  ingest_pipeline_files = fileset(path.module, "../cluster-x-config/ingest-pipelines/*.yml.tftpl")
  role_files            = fileset(path.module, "../cluster-x-config/roles/*.yml.tftpl")
  role_mapping_files    = fileset(path.module, "../cluster-x-config/role-mappings/*.yml.tftpl")
}
```

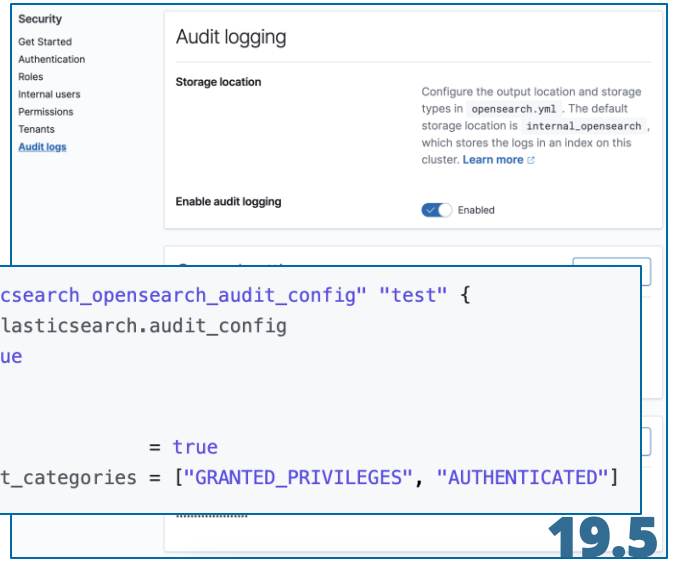
19.2

```
locals {
  index_templates = merge({
    for filename in var.index_template_files :
    replace(basename(filename), "/\\.yaml.tftpl$/", "")
        => yamldecode(file(filename))
  }, var.index_templates)
```

19.3

```
├── cluster-x-config
│   ├── index-patterns
│   │   └── example-pattern.yml.tftpl
│   ├── index-templates
│   │   └── example-template.yml.tftpl
│   ├── ingest-pipelines
│   │   └── example-pipeline.yaml.tftpl
│   ├── ism-policies
│   │   └── example-policy.yml.tftpl
│   ├── role-mappings
│   │   └── all_access.yml.tftpl
│   └── roles
│       └── write_access.yml.tftpl
├── cluster-y-config
├── cluster-z-config
├── main.tf
└── module
    ├── audit_config.tf
    ├── index_patterns.tf
    ├── index_templates.tf
    ├── ingest_pipeline.tf
    ├── ism_policy.tf
    ├── locals.tf
    ├── provider.tf
    ├── role_mappings.tf
    ├── roles.tf
    ├── variables.tf
    └── versions.tf
```

19.4




```
resource "elasticsearch_opensearch_audit_config" "test" {
    provider = elasticsearch.audit_config
    enabled = true

  audit {
    enable_rest              = true
    disabled_rest_categories = ["GRANTED_PRIVILEGES", "AUTHENTICATED"]
```

19.5

```
provider "elasticsearch" {
  url                  = local.cluster_endpoint
  kibana_url           = "${local.cluster_endpoint}/_dashboards"
  #  ...
  elasticsearch_version = "OpenSearch_2.11"
}

provider "elasticsearch" {
  alias                = "audit_config"
  url                  = local.cluster_endpoint
  kibana_url           = "${local.cluster_endpoint}/_dashboards"
  #  ...
  elasticsearch_version = "8.5"
}
```

19.6

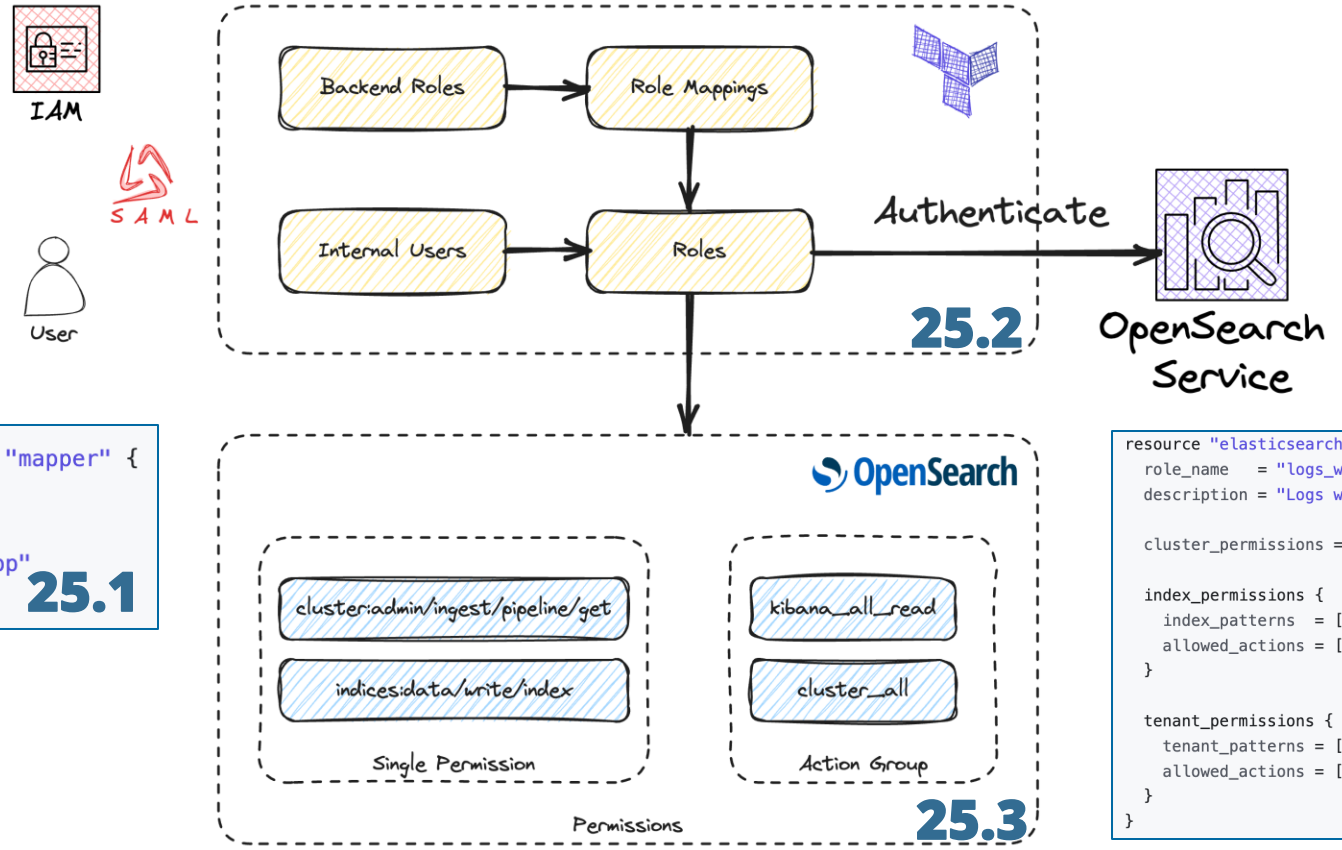
Nutmeg. a J.P.Morgan company

CHASE

ROLES / ROLE-MAPPINGS

```
# Create a role mapping
resource "elasticsearch_opensearch_roles_mapping" "mapper" {
  role_name     = "logs_writer"
  description   = "Mapping AWS IAM roles to ES role"
  backend_roles = [
    "arn:aws:iam::123456789012:role/lambda-call-elasticsearch",
    "arn:aws:iam::123456789012:role/run-containers",
  ]
}
```

25.4




```
resource "elasticsearch_opensearch_user" "mapper" {
  username    = "app-reader"
  password    = "SuperSekret123!"
  description = "a reader role for our app"
}
```

25.1

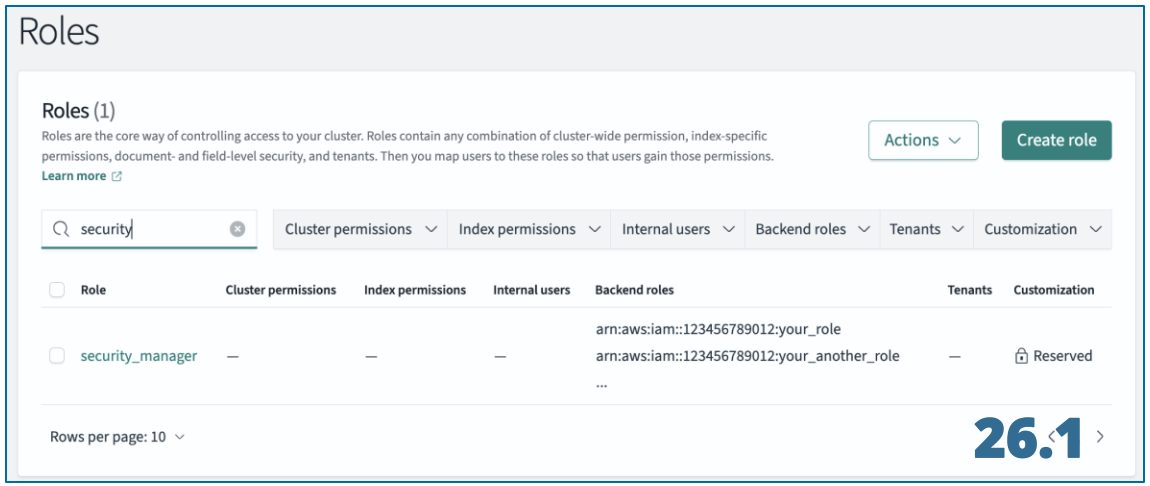
```
resource "elasticsearch_opensearch_role" "writer" {
  role_name   = "logs_writer"
  description = "Logs writer role"

  cluster_permissions = ["*"]

  index_permissions {
    index_patterns  = ["logstash-*"]
    allowed_actions = ["write"]
  }

  tenant_permissions {
    tenant_patterns = ["logstash-*"]
    allowed_actions = ["kibana_all_write"]
  }
}
```

25.5

Nutmeg. a J.P.Morgan company CHASE

ROLES / ROLE-MAPPINGS Cont'd

- ✓ field-level security 🤓
- ✓ advanced settings granularity 😌
- ✓ 'security_manager' in terraform 😵‍💫
- ✓ debugging 🔎



Nutmeg. a J.P.Morgan company CHASE

INDEX TEMPLATES

- ***What?***
  Blueprints for creating with predefined settings/mappings
- ***How?***
  Index matches a template
- ***Why?***
  Homogenous setting

composable / component

```
PUT _index_template/my_logs_template
{
  "index_patterns": [ ...
  ],
  "template": {
    "aliases": { ...
    },
    "settings": { ...
    },
    "mappings": {
      "properties": { ...
      }
    }
  }
}
```

```
"index_patterns": [
  "my_logs-*"
],
```

```
"aliases": {
    "my_logs": {}
},
```

```
"settings": {
    "number_of_shards": 2,
    "number_of_replicas": 1
},
```

29.2

```
"mappings": {
    "properties": {
      "timestamp": {
        "type": "date",
        "format": "epoch_millis"
      },
      "value": {
        "type": "double"
      }
    }
  }
```

```
resource "elasticsearch_index_template" "template_1" {
  name = "template_1"
  body = <<EOF
{
  "template": "te*",
  "settings": {
    "number_of_shards": 1
  },
```

29.1

- ✓ link indices with UI
- ✓ reference to multiple indices
- ✓ refresh interval
- ✓ pipelines
- ✓ number of shards

- ✓ auto—generated IDs
- ✓ @timestamp and 'date_nanos'
- ✓ 'dynamic: false | true | runtime'
- ✓ 'type: flat_object '
- ✓ keyword 'doc_values'
- ✓ 'type: match_only_text' (stacktrace)
- ✓ '_size: enable=true' / painless

Nutmeg. a J.P.Morgan company

CHASE

INDEX PATTERNS

- ✓ '.kibana' index alias
- ✓ refresh
- ✓ Type: number / Format: Bytes

```
resource "elasticsearch_kibana_object" "index_pattern" {
  index = ".kibana"
  body = <<EOF
[
  {
```

31.3

Name	Type	Format	Searchable	Aggregatable
payload_size	number	Bytes	•	•

31.1

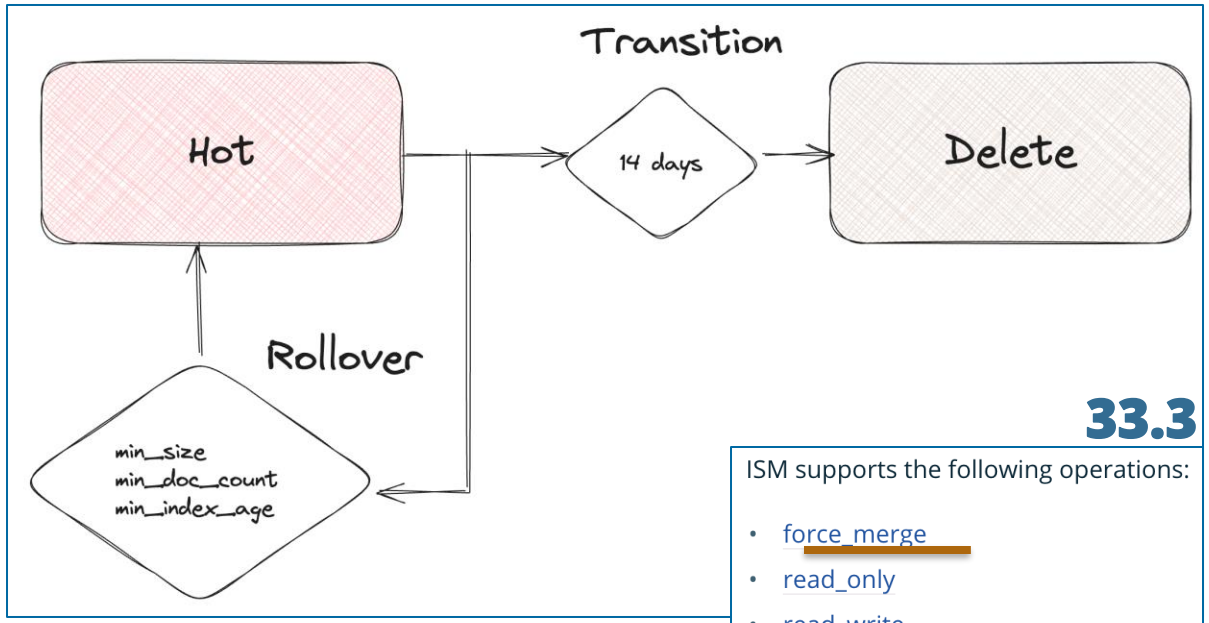
```
{
  "_id": "index-pattern:my-logs",
  "_source": {
    "index-pattern": {
      "timeFieldName": "@timestamp",
      "title": "my-logs*",
      "fieldFormatMap": "{\"payload_size\":{\"id\":\"bytes\"}}"
    },
    "references": [],
    "type": "index-pattern"
  }
}
```

31.2

Nutmeg. a J.P.Morgan company CHASE

ISM / ILM POLICIES

- ✓ Hot – fastest/costly (*im4g/MRC-6g* instances)
- ✓ UltraWarm (*HDD*)
- ✓ Cold (*S3*)
- ✓ Mixed Hot/Warm (*OR1* instances*)



33.3

- ✓ Look after the rollover

```
PUT /my-index1
{
  "settings": {
    "index": {
      "replication.type": "SEGMENT"
    }
  }
}
```

33.1

Parameter
`min_size`
`min_primary_shard_size`
`min_doc_count`
`min_index_age`

33.2

ISM supports the following operations:

- force_merge
- read_only
- read_write
- replica_count
- shrink
- close
- open
- delete
- rollover
- notification
- snapshot
- index_priority
- allocation
- rollup

33.4

```
{
  "default_state": "hot",
  "states": [
    {
      "name": "hot",
      "actions": [
        {
          "retry": {"count": 3...},
          "rollover": {
            "min_size": "100gb",
            "min_doc_count": 100000000,
            "min_index_age": "10d"
          }
        }
      ],
      "transitions": [
        {
          "state_name": "delete",
          "conditions": {
            "min_index_age": "14d"
          }
        }
      ]
    },
    {
      "name": "delete",
      "actions": [
        {
          "retry": {"count": 3...},
          "delete": {}
        }
      ],
      "transitions": []
    }
  ],
  "ism_template": [
    {
      "index_patterns": ["my-logs*"],
      "priority": 100
    }
  ]
```

33.5

INGEST PIPELINES

✓ Enrichment 💰

```
# Create a simple ingest pipeline
resource "elasticsearch_ingest_pipeline" "test" {
  name = "terraform-test"
  body = <<EOF
{
  "description" : "describe pipeline",
```

35.1

```
PUT _ingest/pipeline/my-pipeline
{
  "processors": [
    {
      "set": {
        "description": "Index the ingest timestamp as 'log_ingested'",
        "field": "log_ingested",
        "value": "{{{_ingest.timestamp}}}"
      }
    }
  ]
}
```

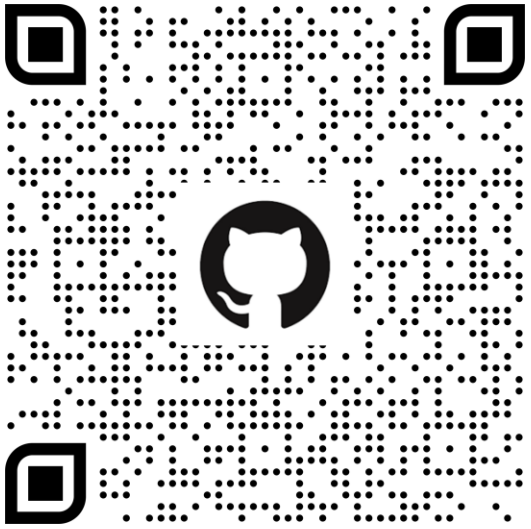
35.2

```
PUT _ingest/pipeline/my-second-pipeline
{
  "description": "Set a human-understandable names for mobile device ",
  "processors": [
    {
      "script": {
        "lang": "painless",
        "params" : {
          "MAR-LX1A" : "HUAWEI P30 lite",
          "SM-A137F" : "Galaxy A13"
        },
        "source": "ctx['mobile_device_name'] = params[ctx['mobile_device_code']]"
      }
    }
  ]
}
```

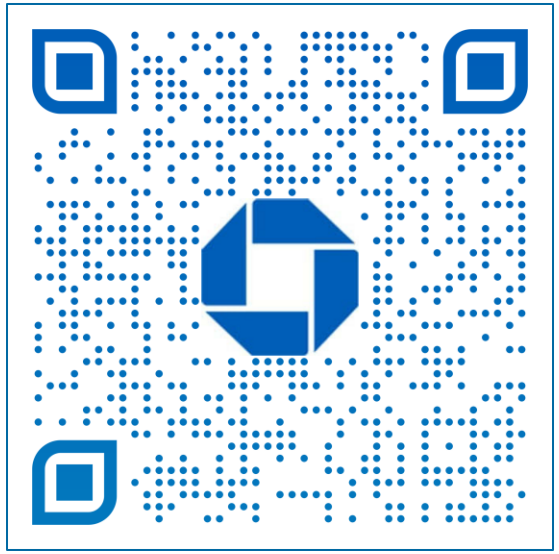
35.3

Nutmeg. a J.P.Morgan company CHASE

DESSERT | CHECKLIST | BEST PRACTICES



- ✓ prefer smaller clusters (<30k shards)
- ✓ lifecycle {prevent_destroy = true }
- ✓ use _*bulk* API for ingestion
- ✓ extra care with 'security_manager' role
- ✓ use index templates
  - ✓ define settings (number of shards)
  - ✓ define mappings
    - ✓ auto—generated IDs
    - ✓ @timestamp and 'date_nanos'
    - ✓ 'dynamic: false'
    - ✓ 'type: flat_object '
    - ✓ keyword 'doc_values'
    - ✓ 'type: match_only_text'
- ✓ extra care with ILM (avoid oversharding)



Nutmeg. a J.P.Morgan company CHASE