» spotinst_elastigroup_aws

Provides a Spotinst AWS group resource.

» Example Usage

```
# Create an Elastigroup
resource "spotinst_elastigroup_aws" "default-elastigroup" {

  name        = "default-elastigroup"
  description = "created by Terraform"
  product     = "Linux/UNIX"

  max_size          = 0
  min_size          = 0
  desired_capacity  = 0
  capacity_unit     = "weight"

  region      = "us-west-2"
  subnet_ids  = ["sb-123456", "sb-456789"]

  image_id              = "ami-a27d8fda"
  iam_instance_profile  = "iam-profile"
  key_name              = "my-key.ssh"
  security_groups       = ["sg-123456"]
  user_data             = "echo hello world"
  enable_monitoring     = false
  ebs_optimized         = false
  placement_tenancy     = "default"

  instance_types_ondemand       = "m3.2xlarge"
  instance_types_spot           = ["m3.xlarge", "m3.2xlarge"]
  instance_types_preferred_spot = ["m3.xlarge"]

  instance_types_weights = [
  {
    instance_type = "c3.large"
    weight        = 10
  },
  {
    instance_type = "c4.xlarge"
    weight        = 16
  }]
```

```
orientation          = "balanced"
fallback_to_ondemand = false
cpu_credits          = "unlimited"

wait_for_capacity         = 5
wait_for_capacity_timeout = 300

scaling_strategy = {
  terminate_at_end_of_billing_hour = true
  termination_policy = "default"
}

scaling_up_policy = {
  policy_name       = "Default Scaling Up Policy"
  metric_name       = "DefaultQueuesDepth"
  statistic         = "average"
  unit              = "none"
  adjustment        = 1
  namespace         = "custom"
  threshold         = 100
  period            = 60
  evaluation_periods = 5
  cooldown          = 300
}

scaling_down_policy = {
  policy_name       = "Default Scaling Down Policy"
  metric_name       = "DefaultQueuesDepth"
  statistic         = "average"
  unit              = "none"
  adjustment        = 1
  namespace         = "custom"
  threshold         = 10
  period            = 60
  evaluation_periods = 10
  cooldown          = 300
}

tags = [
{
   key   = "Env"
   value = "production"
},
{
   key   = "Name"
   value = "default-production"
```

```
  },
  {
     key   = "Project"
     value = "app_v2"
  }
 ]

  lifecycle {
    ignore_changes = [
      "desired_capacity",
    ]
  }
}
```

» Argument Reference

The following arguments are supported:

- `name` - (Required) The group name.

- `description` - (Optional) The group description.

- `product` - (Required) Operation system type. Valid values: `"Linux/UNIX"`, `"SUSE Linux"`, `"Windows"`. For EC2 Classic instances: `"Linux/UNIX (Amazon VPC)"`, `"SUSE Linux (Amazon VPC)"`, `"Windows (Amazon VPC)"`.

- `availability_zones` - (Optional) List of Strings of availability zones. Note: When this parameter is set, `subnet_ids` should be left unused.

- `subnet_ids` - (Optional) List of Strings of subnet identifiers. Note: When this parameter is set, `availability_zones` should be left unused.

- `region` - (Optional) The AWS region your group will be created in. Note: This parameter is required if you specify subnets (through subnet_ids). This parameter is optional if you specify Availability Zones (through availability_zones).

- `preferred_availability_zones` - The AZs to prioritize when launching Spot instances. If no markets are available in the Preferred AZs, Spot instances are launched in the non-preferred AZs. Note: Must be a sublist of `availability_zones` and `orientation` value must not be `"equalAzDistribution"`.

- `max_size` - (Optional; Required if using scaling policies) The maximum number of instances the group should have at any time.

- `min_size` - (Optional; Required if using scaling policies) The minimum number of instances the group should have at any time.

- `desired_capacity` - (Optional) The desired number of instances the group should have at any time.

- `capacity_unit` - (Optional, Default: `"instance"`) The capacity unit to launch instances by. If not specified, when choosing the weight unit, each instance will weight as the number of its vCPUs.

- `security_groups` - (Required) A list of associated security group IDS.

- `image_id` - (Optional) The ID of the AMI used to launch the instance.

- `iam_instance_profile` - (Optional) The ARN or name of an IAM instance profile to associate with launched instances.

- `key_name` - (Optional) The key name that should be used for the instance.

- `enable_monitoring` - (Optional) Indicates whether monitoring is enabled for the instance.

- `user_data` - (Optional) The user data to provide when launching the instance.

- `shutdown_script` - (Optional) The Base64-encoded shutdown script that executes prior to instance termination, for more information please see: Shutdown Script

- `ebs_optimized` - (Optional) Enable high bandwidth connectivity between instances and AWS's Elastic Block Store (EBS). For instance types that are EBS-optimized by default this parameter will be ignored.

- `placement_tenancy` - (Optional) Enable dedicated tenancy. Note: There is a flat hourly fee for each region in which dedicated tenancy is used.

- `instance_types_ondemand` - (Required) The type of instance determines your instance's CPU capacity, memory and storage (e.g., m1.small, c1.xlarge).

- `instance_types_spot` - (Required) One or more instance types.

- `instance_types_preferred_spot` - (Optional) Prioritize a subset of spot instance types. Must be a subset of the selected spot instance types.

- `instance_types_weights` - (Optional) List of weights per instance type for weighted groups. Each object in the list should have the following attributes:

    - `weight` - (Required) Weight per instance type (Integer).
    - `instance_type` - (Required) Name of instance type (String).

- `cpu_credits` - (Optional) Controls how T3 instances are launched. Valid values: `standard`, `unlimited`.

- `fallback_to_ondemand` - (Required) In a case of no Spot instances available, Elastigroup will launch on-demand instances instead.

- `wait_for_capacity` - (Optional) Minimum number of instances in a 'HEALTHY' status that is required before continuing. This is ignored when updating with blue/green deployment. Cannot exceed `desired_capacity`.

- `wait_for_capacity_timeout` - (Optional) Time (seconds) to wait for instances to report a 'HEALTHY' status. Useful for plans with multiple dependencies that take some time to initialize. Leave undefined or set to `0` to indicate no wait. This is ignored when updating with blue/green deployment.

- `orientation` - (Required, Default: `"balanced"`) Select a prediction strategy. Valid values: `"balanced"`, `"costOriented"`, `"equalAzDistribution"`, `"availabilityOriented"`.

- `spot_percentage` - (Optional; Required if not using `ondemand_count`) The percentage of Spot instances that would spin up from the `desired_capacity` number.

- `ondemand_count` - (Optional; Required if not using `spot_percentage`) Number of on demand instances to launch in the group. All other instances will be spot instances. When this parameter is set the `spot_percentage` parameter is being ignored.

- `draining_timeout` - (Optional) The time in seconds, the instance is allowed to run while detached from the ELB. This is to allow the instance time to be drained from incoming TCP connections before terminating it, during a scale down operation.

- `utilize_reserved_instances` - (Optional) In a case of any available reserved instances, Elastigroup will utilize them first before purchasing Spot instances.

- `scaling_strategy` - (Optional) Set termination policy.

  - `terminate_at_end_of_billing_hour` - (Optional) Specify whether to terminate instances at the end of each billing hour.
  - `termination_policy` - (Optional) - Determines whether to terminate the newest instances when performing a scaling action. Valid values: `"default"`, `"newestInstance"`.

- `health_check_type` - (Optional) The service that will perform health checks for the instance. Valid values: `"ELB"`, `"HCS"`, `"TARGET_GROUP"`, `"MLB"`, `"EC2"`, `"MULTAI_TARGET_SET"`, `"MLB_RUNTIME"`, `"K8S_NODE"`, `"NOMAD_NODE"`, `"ECS_CLUSTER_INSTANCE"`.

- `health_check_grace_period` - (Optional) The amount of time, in seconds, after the instance has launched to starts and check its health.

- `health_check_unhealthy_duration_before_replacement` - (Optional) The amount of time, in seconds, that we will wait before replacing an

5

instance that is running and became unhealthy (this is only applicable for instances that were once healthy).

- `tags` - (Optional) A key/value mapping of tags to assign to the resource.
- `elastic_ips` - (Optional) A list of AWS Elastic IP allocation IDs to associate to the group instances.
- `revert_to_spot` - (Optional) Hold settings for strategy correction – replacing On-Demand for Spot instances. Supported Values: `"never"`, `"always"`, `"timeWindow"`
  - `perform_at` - (Required) In the event of a fallback to On-Demand instances, select the time period to revert back to Spot. Supported Arguments – always (default), timeWindow, never. For timeWindow or never to be valid the group must have availabilityOriented OR persistence defined.
  - `time_windows` - (Optional) Specify a list of time windows for to execute revertToSpot strategy. Time window format: `ddd:hh:mm-ddd:hh:mm`. Example: `Mon:03:00-Wed:02:30`

» Load Balancers

- `elastic_load_balancers` - (Optional) List of Elastic Load Balancers names (ELB).
- `target_group_arns` - (Optional) List of Target Group ARNs to register the instances to.
- `multai_target_sets` - (Optional) Set of targets to register.
  - `target_set_id` - (Required) ID of Multai target set.
  - `balancer_id` - (Required) ID of Multai Load Balancer.

Usage:

```
elastic_load_balancers = ["bal5", "bal2"]
target_group_arns = ["tg-arn"]
multai_target_sets = [{
  target_set_id = "ts-123",
  balancer_id   = "bal-123"
},
{
  target_set_id = "ts-234",
  balancer_id   = "bal-234"
}]
```

» Signals

Each `signal` supports the following:

- `name` - (Required) The name of the signal defined for the group. Valid Values: `"INSTANCE_READY"`, `"INSTANCE_READY_TO_SHUTDOWN"`
- `timeout` - (Optional) The signals defined timeout- default is 40 minutes (1800 seconds).

Usage:

```
signal = {
  name    = "INSTANCE_READY_TO_SHUTDOWN"
  timeout = 100
}
```

» Scheduled Tasks

Each `scheduled_task` supports the following:

- `task_type` - (Required) The task type to run. Supported task types are: `"scale"`, `"backup_ami"`, `"roll"`, `"scaleUp"`, `"percentageScaleUp"`, `"scaleDown"`, `"percentageScaleDown"`, `"statefulUpdateCapacity"`.
- `cron_expression` - (Optional; Required if not using `frequency`) A valid cron expression. The cron is running in UTC time zone and is in Unix cron format.
- `start_time` - (Optional; Format: ISO 8601) Set a start time for one time tasks.
- `frequency` - (Optional; Required if not using `cron_expression`) The recurrence frequency to run this task. Supported values are `"hourly"`, `"daily"`, `"weekly"` and `"continuous"`.
- `scale_target_capacity` - (Optional) The desired number of instances the group should have.
- `scale_min_capacity` - (Optional) The minimum number of instances the group should have.
- `scale_max_capacity` - (Optional) The maximum number of instances the group should have.
- `is_enabled` - (Optional, Default: `true`) Setting the task to being enabled or disabled.
- `target_capacity` - (Optional; Only valid for statefulUpdateCapacity) The desired number of instances the group should have.
- `min_capacity` - (Optional; Only valid for statefulUpdateCapacity) The minimum number of instances the group should have.
- `max_capacity` - (Optional; Only valid for statefulUpdateCapacity) The maximum number of instances the group should have.
- `batch_size_percentage` - (Optional; Required when the `task_type` is `"roll"`.) The percentage size of each batch in the scheduled deployment roll.
- `grace_period` - (Optional) The period of time (seconds) to wait before checking a batch's health after it's deployment.

7

- `adjustment` - (Optional; Min 1) The number of instances to add or remove.
- `adjustment_percentage` - (Optional; Min 1) The percentage of instances to add or remove.

Usage:

```
scheduled_task = [{
  task_type             = "backup_ami"
  cron_expression       = ""
  start_time            = "1970-01-01T01:00:00Z"
  frequency             = "hourly"
  scale_target_capacity = 5
  scale_min_capacity    = 0
  scale_max_capacity    = 10
  is_enabled            = false
  target_capacity       = 5
  min_capacity          = 0
  max_capacity          = 10
  batch_size_percentage = 33
  grace_period          = 300
}]
```

» Scaling Policies

Each `scaling_*_policy` supports the following:

- `namespace` - (Required) The namespace for the alarm's associated metric.

- `metric_name` - (Required) The name of the metric, with or without spaces.

- `threshold` - (Required) The value against which the specified statistic is compared.

- `policy_name` - (Required) The name of the policy.

- `statistic` - (Optional, Default: `"average"`) The metric statistics to return. For information about specific statistics go to Statistics in the Amazon CloudWatch Developer Guide.

- `unit` - (Required) The unit for the alarm's associated metric. Valid values: `"percent, "seconds", "microseconds", "milliseconds", "bytes", "kilobytes", "megabytes", "gigabytes", "terabytes", "bits", "kilobits", "megabits", "gigabits", "terabits", "count", "bytes/second", "kilobytes/second", "megabytes/second", "gigabytes/second", "terabytes/second", "bits/second", "kilobits/second", "megabits/second", "gigabits/second", "terabits/second", "count/second", "none"`.

- `is_enabled` - (Optional, Default: `true`) Specifies whether the scaling policy described in this block is enabled.

- `period` - (Optional, Default: `300`) The granularity, in seconds, of the returned datapoints. Period must be at least 60 seconds and must be a multiple of 60.

- `evaluation_periods` - (Optional, Default: `1`) The number of periods over which data is compared to the specified threshold.

- `cooldown` - (Optional, Default: `300`) The amount of time, in seconds, after a scaling activity completes and before the next scaling activity can start. If this parameter is not specified, the default cooldown period for the group applies.

- `dimensions` - (Optional) A list of dimensions describing qualities of the metric.

  - `name` - (Required) The dimension name.
  - `value` - (Required) The dimension value.

- `operator` - (Optional, Scale Up Default: `gte`, Scale Down Default: `lte`) The operator to use in order to determine if the scaling policy is applicable. Valid values: `"gt"`, `"gte"`, `"lt"`, `"lte"`.

- `source` - (Optional) The source of the metric. Valid values: `"cloudWatch"`, `"spectrum"`.

- `action_type` - (Optional; if not using `min_target_capacity` or `max_target_capacity`) The type of action to perform for scaling. Valid values: `"adjustment"`, `"percentageAdjustment"`, `"setMaxTarget"`, `"setMinTarget"`, `"updateCapacity"`.

If you do not specify an action type, you can only use – `adjustment`, `minTargetCapacity`, `maxTargetCapacity`. While using action_type, please also set the following:

When using `adjustment` – set the field `adjustment` When using `percentageAdjustment` - set the field `adjustment` When using `setMaxTarget` – set the field `max_target_capacity` When using `setMinTarget` – set the field `min_target_capacity` When using `updateCapacity` – set the fields `minimum`, `maximum`, and `target`

- `adjustment` - (Optional; if not using `min_target_capacity` or `max_target_capacity`;) The number of instances to add/remove to/from the target capacity when scale is needed. Can be used as advanced expression for scaling of instances to add/remove to/from the target capacity when scale is needed. You can see more information here: Advanced expression. Example value: `"MAX(currCapacity / 5, value * 10)"`

- `min_target_capacity` - (Optional; if not using `adjustment`; available only for scale up). The number of the desired target (and minimum)

capacity

- `max_target_capacity` - (Optional; if not using `adjustment`; available only for scale down). The number of the desired target (and maximum) capacity

- `minimum` - (Optional; if using `updateCapacity`) The minimal number of instances to have in the group.

- `maximum` - (Optional; if using `updateCapacity`) The maximal number of instances to have in the group.

- `target` - (Optional; if using `updateCapacity`) The target number of instances to have in the group.

Usage:

```
scaling_up_policy = [{
  policy_name = "policy-name"
  metric_name = "CPUUtilization"
  namespace   = "AWS/EC2"
  source      = ""
  statistic   = "average"
  unit        = ""
  cooldown    = 60
  is_enabled  = false

  dimensions = {
      name  = "name-1"
      value = "value-1"
  }

  threshold         = 10
  operator          = "gt"
  evaluation_periods = 10
  period            = 60

  // === MIN TARGET ==================
  action_type        = "setMinTarget"
  min_target_capacity = 1
  // =================================

  // === ADJUSTMENT ==================
  # action_type = "adjustment"
  # action_type = "percentageAdjustment"
  # adjustment  = "MAX(5,10)"
  // =================================

  // === UPDATE CAPACITY =============
```

10

```
    # action_type = "updateCapacity"
    # minimum     = 0
    # maximum     = 10
    # target      = 5
    // =================================

}]
scaling_target_policy = [{
    policy_name = ""
    metric_name = ""
    namespace   = ""
    source      = ""
    statistic   = ""
    unit        = ""
    cooldown    = 10
    target      = 1

    dimensions = [{
      name  = ""
      value = ""
    }]
}]
```

» Network Interfaces

Each of the `network_interface` attributes controls a portion of the AWS Instance's "Elastic Network Interfaces". It's a good idea to familiarize yourself with AWS's Elastic Network Interfaces docs to understand the implications of using these attributes.

- `network_interface_id` - (Optional) The ID of the network interface.
- `device_index` - (Required) The index of the device on the instance for the network interface attachment.
- `description` - (Required) The description of the network interface.
- `private_ip_address` - (Optional) The private IP address of the network interface.
- `delete_on_termination` - (Optional) If set to true, the interface is deleted when the instance is terminated.
- `secondary_private_ip_address_count` - (Optional) The number of secondary private IP addresses.
- `associate_public_ip_address` - (Optional) Indicates whether to assign a public IP address to an instance you launch in a VPC. The public IP address can only be assigned to a network interface for eth0, and can only be assigned to a new network interface, not an existing one.

- `associate_ipv6_address` - (Optional) Indicates whether to assign IPV6 addresses to your instance. Requires a subnet with IPV6 CIDR block ranges.

Usage:

```
network_interface = [{
  network_interface_id                = ""
  device_index                        = 1
  description                         = "nic description in here"
  private_ip_address                  = "1.1.1.1"
  delete_on_termination               = false
  secondary_private_ip_address_count  = 1
  associate_public_ip_address         = true
}]
```

» Block Devices

Each of the `*_block_device` attributes controls a portion of the AWS Instance's "Block Device Mapping". It's a good idea to familiarize yourself with AWS's Block Device Mapping docs to understand the implications of using these attributes.

Each `ebs_block_device` supports the following:

- `device_name` - (Required) The name of the device to mount.
- `snapshot_id` - (Optional) The Snapshot ID to mount.
- `volume_type` - (Optional, Default: `"standard"`) The type of volume. Can be `"standard"`, `"gp2"`, `"io1"`, `"st1"` or `"sc1"`.
- `volume_size` - (Optional) The size of the volume in gigabytes.
- `iops` - (Optional) The amount of provisioned IOPS. This must be set with a `volume_type` of `"io1"`.
- `delete_on_termination` - (Optional) Whether the volume should be destroyed on instance termination.
- `encrypted` - (Optional) Enables EBS encryption on the volume.
- `kms_key_id` - (Optional) ID for a user managed CMK under which the EBS Volume is encrypted

Modifying any `ebs_block_device` currently requires resource replacement.

Usage:

```
ebs_block_device = [{
  device_name       = "/dev/sdb"
  snapshot_id       = ""
  volume_type       = "gp2"
  volume_size       = 8
  iops              = 1
```

```
      delete_on_termination = true
      encrypted             = false
      kms_key_id            = "kms-key-01"
  },
  {
      device_name           = "/dev/sdc"
      snapshot_id           = ""
      volume_type           = "gp2"
      volume_size           = 8
      iops                  = 1
      delete_on_termination = true
      encrypted             = true
      kms_key_id            = "kms-key-02"
  }]
```

Each `ephemeral_block_device` supports the following:

- `device_name` - (Required) The name of the block device to mount on the instance.
- `virtual_name` - (Required) The Instance Store Device Name (e.g. `"ephemeral0"`).

Usage:

```
  ephemeral_block_device = [{
    device_name  = "/dev/xvdc"
    virtual_name = "ephemeral0"
  }]
```

» Stateful

We support instance persistence via the following configurations. all values are boolean. For more information on instance persistence please see: Stateful configuration

- `persist_root_device` - (Optional) Boolean, should the instance maintain its root device volumes.
- `persist_block_devices` - (Optional) Boolean, should the instance maintain its Data volumes.
- `persist_private_ip` - (Optional) Boolean, should the instance maintain its private IP.
- `block_devices_mode` - (Optional) String, determine the way we attach the data volumes to the data devices, possible values: `"reattach"` and `"onLaunch"` (default is onLaunch).
- `private_ips` - (Optional) List of Private IPs to associate to the group instances.(e.g. "172.1.1.0"). Please note: This setting will only apply if persistence.persist_private_ip is set to true.

13

Usage:

```
persist_root_device   = false
persist_block_devices = false
persist_private_ip    = true
block_devices_mode    = "onLaunch"
private_ips           = ["1.1.1.1", "2.2.2.2"]
```

» Stateful Deallocation

- `stateful_deallocation` - (Optional)
  - `should_delete_images` - (Optional) For stateful groups: remove persistent images.
  - `should_delete_network_interfaces` - (Optional) For stateful groups: remove network interfaces.
  - `should_delete_volumes` - (Optional) For stateful groups: remove persistent volumes.
  - `should_delete_snapshots` - (Optional) For stateful groups: remove snapshots.

Usage:

```
stateful_deallocation = {
   should_delete_images              = false
   should_delete_network_interfaces  = false
   should_delete_volumes             = false
   should_delete_snapshots           = false
 }
```

» Health Check

- `health_check_type` - (Optional) The service that will perform health checks for the instance. Supported values : `"ELB"`, `"HCS"`, `"TARGET_GROUP"`, `"CUSTOM"`, `"K8S_NODE"`, `"MLB"`, `"EC2"`, `"MULTAI_TARGET_SET"`, `"MLB_RUNTIME"`, `"K8S_NODE"`, `"NOMAD_NODE"`, `"ECS_CLUSTER_INSTANCE"`.
- `health_check_grace_period` - (Optional) The amount of time, in seconds, after the instance has launched to starts and check its health
- `health_check_unhealthy_duration_before_replacement` - (Optional) The amount of time, in seconds, that we will wait before replacing an instance that is running and became unhealthy (this is only applicable for instances that were once healthy)

Usage:

```
health_check_type                       = "ELB"
health_check_grace_period               = 100
```

```
health_check_unhealthy_duration_before_replacement = 120
```

» Third-Party Integrations

- `integration_rancher` - (Optional) Describes the Rancher integration.
  - `master_host` - (Required) The URL of the Rancher Master host.
  - `access_key` - (Required) The access key of the Rancher API.
  - `secret_key` - (Required) The secret key of the Rancher API.
  - `version` - (Optional) The Rancher version. Must be `"1"` or `"2"`. If this field is omitted, it's assumed that the Rancher cluster is version 1. Note that Kubernetes is required when using Rancher version 2. Usage:

```
integration_rancher = {
  master_host = "master_host"
  access_key  = "access_key"
  secret_key  = "secret_key"
  version     = "2"
}
```

- `integration_ecs` - (Optional) Describes the EC2 Container Service integration.
  - `cluster_name` - (Required) The name of the EC2 Container Service cluster.
  - `autoscale_is_enabled` - (Optional, Default: `false`) Specifies whether the auto scaling feature is enabled.
  - `autoscale_cooldown` - (Optional, Default: `300`) The amount of time, in seconds, after a scaling activity completes before any further trigger-related scaling activities can start.
  - `autoscale_is_auto_config` - (Optional, Default: `false`) Enabling the automatic auto-scaler functionality. For more information please see: ECS auto scaler.
  - `autoscale_scale_down_non_service_tasks` - (Optional) Determines whether to scale down non-service tasks.
  - `autoscale_headroom` - (Optional) Headroom for the cluster.
    * `cpu_per_unit` - (Optional, Default: `0`) Cpu units for compute.
    * `memory_per_unit` - (Optional, Default: `0`) RAM units for compute.
    * `num_of_units` - (Optional, Default: `0`) Amount of units for compute.
  - `autoscale_down` - (Optional) Enabling scale down.
    * `evaluation_periods` - (Optional, Default: `5`) Amount of cooldown evaluation periods for scale down.
    * `max_scale_down_percentage` - (Optional) Represents the maximum percent to scale-down. Number between 1-100.

- **autoscale_attributes** - (Optional) A key/value mapping of tags to assign to the resource.

Usage:

```
integration_ecs = {
  cluster_name       = "ecs-cluster"
  autoscale_is_enabled = false
  autoscale_cooldown   = 300
  autoscale_scale_down_non_service_tasks = false

  autoscale_headroom = {
    cpu_per_unit    = 1024
    memory_per_unit = 512
    num_of_units    = 2
  }

  autoscale_down = {
    evaluation_periods      = 300
    max_scale_down_percentage = 70
  }

  autoscale_attributes = [{
    key   = "test.ecs.key"
    value = "test.ecs.value"
  }]
}
```

- **integration_codedeploy** - (Optional) Describes the Code Deploy integration.

  - **cleanup_on_failure** - (Optional) Cleanup automatically after a failed deploy.
  - **terminate_instance_on_failure** - (Optional) Terminate the instance automatically after a failed deploy.
  - **deployment_groups** - (Optional) Specify the deployment groups details.
    * **application_name** - (Optional) The application name.
    * **deployment_group_name** - (Optional) The deployment group name.

Usage:

```
integration_codedeploy = {
  cleanup_on_failure          = false
  terminate_instance_on_failure = false

  deployment_groups = {
    application_name      = "my-app"
```

```
      deployment_group_name = "my-group"
    }
  }
```

- `integration_route53` - (Optional) Describes the Route53 integration.

    - `domains` - (Required) Collection of one or more domains to register.
        * `hosted_zone_id` - (Required) The id associated with a hosted zone.
        * `spotinst_acct_id` - (Optional) The Spotinst account ID that is linked to the AWS account that holds the Route 53 hosted Zone ID. The default is the user Spotinst account provided as a URL parameter.
        * `record_sets` - (Required) Collection of records containing authoritative DNS information for the specified domain name.
            · `name` - (Required) The record set name.
            · `use_public_ip` - (Optional, Default: `false`) - Designates if the IP address should be exposed to connections outside the VPC.

Usage:

```
integration_route53 = {
  domains = {
    hosted_zone_id   = "zone-id"
    spotinst_acct_id = "act-123456"

    record_sets = {
      name          = "foo.example.com"
      use_public_ip = true
    }
  }
}
```

- `integration_docker_swarm` - (Optional) Describes the Docker Swarm integration.

    - `master_host` - (Required) IP or FQDN of one of your swarm managers.
    - `master_port` - (Required) Network port used by your swarm.
    - `autoscale_is_enabled` - (Optional, Default: `false`) Specifies whether the auto scaling feature is enabled.
    - `autoscale_cooldown` - (Optional, Default: `300`) The amount of time, in seconds, after a scaling activity completes before any further trigger-related scaling activities can start. Minimum 180, must be a multiple of 60.
    - `autoscale_headroom` - (Optional) An option to set compute reserve for the cluster.

* `cpu_per_unit` - (Optional, Default: 0) How much CPU to allocate for headroom unit.
* `memory_per_unit` - (Optional, Default: 0) The amount of memory in each headroom unit. Measured in MiB.
* `num_of_units` - (Optional, Default: 0) How many units to allocate for headroom unit.
  - `autoscale_down` - (Optional) Setting for scale down actions.
    * `evaluation_periods` - (Optional, Default: 5) Number of periods over which data is compared. Minimum 3, Measured in consecutive minutes.

Usage:

```
integration_docker_swarm = {
    master_host         = "10.10.10.10"
    master_port         = 2376
    autoscale_is_enabled = true
    autoscale_cooldown   = 180

    autoscale_headroom = {
        cpu_per_unit    = 2048
        memory_per_unit = 2048
        num_of_units    = 1
    }

    autoscale_down = {
        evaluation_periods = 3
    }
}
```

- `integration_kubernetes` - (Optional) Describes the Kubernetes integration.

  - `integration_mode` - (Required) Valid values: `"saas"`, `"pod"`.
  - `cluster_identifier` - (Required; if using integration_mode as pod)
  - `api_server` - (Required; if using integration_mode as saas)
  - `token` - (Required; if using integration_mode as saas) Kubernetes Token
  - `autoscale_is_enabled` - (Optional, Default: `false`) Specifies whether the auto scaling feature is enabled.
  - `autoscale_is_auto_config` - (Optional, Default: `false`) Enabling the automatic k8s auto-scaler functionality. For more information please see: Kubernetes auto scaler.
  - `autoscale_cooldown` - (Optional, Default: `300`) The amount of time, in seconds, after a scaling activity completes before any further trigger-related scaling activities can start.
  - `autoscale_headroom` - (Optional) An option to set compute reserve for the cluster.

* `cpu_per_unit` - (Optional, Default: `0`) How much CPU to allocate for headroom unit.
* `memory_per_unit` - (Optional, Default: `0`) How much Memory allocate for headroom unit.
* `num_of_units` - (Optional, Default: `0`) How many units to allocate for headroom unit.
- `autoscale_down` - (Optional) Setting for scale down actions.
  * `evaluation_periods` - (Optional, Default: `5`) How many evaluation periods should accumulate before a scale down action takes place.
- `autoscale_labels` - (Optional) A key/value mapping of tags to assign to the resource.

Usage:

```
integration_kubernetes = {
  integration_mode  = "pod"
  cluster_identifier = "my-identifier.ek8s.com"

  // === SAAS ===================
  # integration_mode = "saas"
  # api_server       = "https://api.my-identifier.ek8s.com/api/v1/namespaces/kube-system/s
  # token            = "top-secret"
  // ===========================

  autoscale_is_enabled    = false
  autoscale_is_auto_config = false
  autoscale_cooldown      = 300

  autoscale_headroom = {
    cpu_per_unit    = 1024
    memory_per_unit = 512
    num_of_units    = 1
  }

  autoscale_down = {
    evaluation_periods = 300
  }

  autoscale_labels = [{
    key   = "test.k8s.key"
    value = "test.k8s.value"
  }]
}
```

- `integration_nomad` - (Optional) Describes the Nomad integration.

  - `master_host` - (Required) The URL for the Nomad master host.

- **master_port** - (Required) The network port for the master host.
- **acl_token** - (Required) Nomad ACL Token
- **autoscale_is_enabled** - (Optional, Default: **false**) Specifies whether the auto scaling feature is enabled.
- **autoscale_cooldown** - (Optional, Default: **300**) The amount of time, in seconds, after a scaling activity completes before any further trigger-related scaling activities can start.
- **autoscale_headroom** - (Optional) An option to set compute reserve for the cluster.
  - **cpu_per_unit** - (Optional, Default: **0**) How much CPU (MHz) to allocate for headroom unit.
  - **memory_per_unit** - (Optional, Default: **0**) How much Memory allocate for headroom unit.
  - **num_of_units** - (Optional, Default: **0**) How many units of headroom to allocate.
- **autoscale_down** - (Optional) Settings for scale down actions.
  - **evaluation_periods** - (Optional, Default: **5**) How many evaluation periods should accumulate before a scale down action takes place.
- **autoscale_constraints** - (Optional) A key/value mapping of tags to assign to the resource.

Usage:

```
integration_nomad = {
  master_host         = "my-nomad-host"
  master_port         = 9000
  acl_token           = "top-secret"
  autoscale_is_enabled = false
  autoscale_cooldown   = 300

  autoscale_headroom = {
    cpu_per_unit    = 1024
    memory_per_unit = 512
    num_of_units    = 2
  }

  autoscale_down = {
    evaluation_periods = 300
  }

  autoscale_constraints = [{
    key   = "test.nomad.key"
    value = "test.nomad.value"
  }]
}
```

- `integration_mesosphere` - (Optional) Describes the Mesosphere integration.

    - `api_server` - (Optional) The public IP of the DC/OS Master.

Usage:

```
integration_mesosphere = {
  api_server = ""
}
```

- `integration_multai_runtime` - (Optional) Describes the Multai Runtime integration.

    - `deployment_id` - (Optional) The deployment id you want to get

Usage:

```
integration_multai_runtime = {
  deployment_id = ""
}
```

- `integration_gitlab` - (Optional) Describes the Gitlab integration.

    - `runner` - (Optional) Settings for Gitlab runner.
        * `is_enabled` - (Optional, Default: `false`) Specifies whether the integration is enabled.

Usage:

```
integration_gitlab = {
  runner = {
    is_enabled = true
  }
}
```

- `integration_beanstalk` - (Optional) Describes the Beanstalk integration.

    - `deployment_preferences` - (Optional) Preferences when performing a roll
        * `automatic_roll` - (Required) Should roll perform automatically
        * `batch_size_percentage` - (Required) Percent size of each batch
        * `grace_period` - (Required) Amount of time to wait between batches
        * `strategy` - (Optional) Strategy parameters
            · `action` - (Required) Action to take
            · `should_drain_instances` - (Required) Bool value if to wait to drain instance
    - `managed_actions` - (Optional) Managed Actions parameters
        * `platform_update` - (Optional) Platform Update parameters

- · `perform_at` - (Required) Actions to perform (options: timeWindow, never)
- · `time_window` - (Required) Time Window for when action occurs ex. Mon:23:50-Tue:00:20
- · `update_level` - (Required) - Level to update

Usage:

```
integration_beanstalk = {
  environment_id        = "e-3tkmbj7hzc"

  deployment_preferences = {
    automatic_roll       = true
    batch_size_percentage = 100
    grace_period          = 90
    strategy              = {
      action                = "REPLACE_SERVER"
      should_drain_instance = true
    }
  }

  managed_actions        = {
    platform_update = {
      perform_at   = "timeWindow"
      field_name   = "Mon:23:50-Tue:00:20"
      update_level = "minorAndPatch"
    }
  }
}
```

» Update Policy

- `update_policy` - (Optional)
  - `should_resume_stateful` - (Required) This will apply resuming action for Stateful instances in the Elastigroup upon scale up or capacity changes. Example usage will be for Elastigroups that will have scheduling rules to set a target capacity of 0 instances in the night and automatically restore the same state of the instances in the morning.
  - `auto_apply_tags` - (Optional) Enables updates to tags without rolling the group when set to `true`.
  - `should_roll` - (Required) Sets the enablement of the roll option.
  - `roll_config` - (Required) While used, you can control whether the group should perform a deployment after an update to the configuration.

* `batch_size_percentage` - (Required) Sets the percentage of the instances to deploy in each batch.
* `health_check_type` - (Optional) Sets the health check type to use. Valid values: `"EC2"`, `"ECS_CLUSTER_INSTANCE"`, `"ELB"`, `"HCS"`, `"MLB"`, `"TARGET_GROUP"`, `"MULTAI_TARGET_SET"`, `"NONE"`.
* `grace_period` - (Optional) Sets the grace period for new instances to become healthy.
* `wait_for_roll_percentage` - (Optional) For use with `should_roll`. Sets minimum % of roll required to complete before continuing the plan. Required if `wait_for_roll_timeout` is set.
* `wait_for_roll_timeout` - (Optional) For use with `should_roll`. Sets how long to wait for the deployed % of a roll to exceed `wait_for_roll_percentage` before continuing the plan. Required if `wait_for_roll_percentage` is set.
* `strategy` - (Optional) Strategy parameters
    · `action` - (Required) Action to take. Valid values: `REPLACE_SERVER`, `RESTART_SERVER`.
    · `should_drain_instances` - (Optional) Specify whether to drain incoming TCP connections before terminating a server.
    · `batch_min_healthy_percentage` - (Optional, Default 50) Indicates the threshold of minimum healthy instances in single batch. If the amount of healthy instances in single batch is under the threshold, the deployment will fail. Range `1` - `100`.

```
update_policy = {
  should_resume_stateful = false
  should_roll            = false
  auto_apply_tags        = false

  roll_config = {
    batch_size_percentage = 33
    health_check_type     = "ELB"
    grace_period          = 300
    wait_for_roll_percentage = 10
    wait_for_roll_timeout    = 1500

    strategy = {
      action = "REPLACE_SERVER"
      should_drain_instances = false
      batch_min_healthy_percentage = 10
    }
  }
}
```

» Attributes Reference

The following attributes are exported:

- `id` - The group ID.

» spotinst_elastigroup_aws_beanstalk

Provides a Spotinst AWS group resource using Elastic Beanstalk.

» Example Usage

```
resource "spotinst_elastigroup_aws_beanstalk" "elastigoup-aws=beanstalk" {

 name    = "example-elastigroup-beanstalk"
 region  = "us-west-2"
 product = "Linux/UNIX"

 min_size         = 0
 max_size         = 1
 desired_capacity = 0

 beanstalk_environment_name = "example-env"
 beanstalk_environment_id   = "e-example"
 instance_types_spot        = ["t2.micro", "t2.medium", "t2.large"]
}

deployment_preferences = {
    automatic_roll       = true
    batch_size_percentage = 100
    grace_period         = 90
    strategy = {
        action               = "REPLACE_SERVER"
        should_drain_instances = true
    }
}

managed_actions = {
    platform_update = {
        perform_at   = "timeWindow"
        time_window  = "Mon:23:50-Tue:00:20"
        update_level = "minorAndPatch"
    }
}
```

» **Argument Reference**

The following arguments are supported:

- `name` - (Required) The group name.

- `region` - (Required) The AWS region your group will be created in. Cannot be changed after the group has been created.

- `description` - (Optional) The group description.

- `product` - (Required) Operation system type. Valid values: `"Linux/UNIX"`, `"SUSE Linux"`, `"Windows"`. For EC2 Classic instances: `"Linux/UNIX (Amazon VPC)"`, `"SUSE Linux (Amazon VPC)"`, `"Windows (Amazon VPC)"`.

- `max_size` - (Required) The maximum number of instances the group should have at any time.

- `min_size` - (Required) The minimum number of instances the group should have at any time.

- `desired_capacity` - (Required) The desired number of instances the group should have at any time.

- `beanstalk_environment_name` - (Optional) The name of an existing Beanstalk environment.

- `beanstalk_environment_id` - (Optional) The id of an existing Beanstalk environment.

- `instance_types_spot` - (Required) One or more instance types. To maximize the availability of Spot instances, select as many instance types as possible.

- `deployment_preferences` - (Optional) Preferences when performing a roll

    - `automatic_roll` - (Required) Should roll perform automatically
    - `batch_size_percentage` - (Required) Percent size of each batch
    - `grace_period` - (Required) Amount of time to wait between batches
    - `strategy` - (Optional) Strategy parameters
        * `action` - (Required) Action to take
        * `should_drain_instances` - (Required) Bool value if to wait to drain instance

- `managed_actions` - (Optional) Managed Actions parameters

    - `platform_update` - (Optional) Platform Update parameters
        * `perform_at` - (Required) Actions to perform (options: timeWindow, never)

* `time_window` - (Required) Time Window for when action occurs ex. Mon:23:50-Tue:00:20
* `update_level` - (Required) - Level to update

» spotinst_elastigroup_azure

Provides a Spotinst elastigroup Azure resource.

» Example Usage

```
resource "spotinst_elastigroup_azure" "test_azure_group" {
  name                = "example_elastigroup_azure"
  resource_group_name = "spotinst-azure"
  region              = "eastus"
  product             = "Linux"

  user_data       = ""
  shutdown_script = ""

  managed_service_identities = {
    resource_group_name = "spotinst-azure"
    name = "example-identity"
  }

  // --- CAPACITY --------------------------------------------------------
  min_size         = 0
  max_size         = 1
  desired_capacity = 1
  // ---------------------------------------------------------------------

  // --- INSTANCE TYPES --------------------------------------------------
  od_sizes          = ["standard_a1_v1", "standard_a1_v2"]
  low_priority_sizes = ["standard_a1_v1", "standard_a1_v2"]
  // ---------------------------------------------------------------------

  // --- IMAGE -----------------------------------------------------------
  image = {
    marketplace = {
      publisher = "Canonical"
      offer     = "UbuntuServer"
      sku       = "16.04-LTS"
    }
  }
```

```
// --------------------------------------------------------------------

// --- STRATEGY -------------------------------------------------------
strategy = {
  od_count          = 1
  draining_timeout  = 300
}
// --------------------------------------------------------------------

// --- LOAD BALANCERS -------------------------------------------------
load_balancers = [{
  type          = "MULTAI_TARGET_SET"
  balancer_id   = "lb-1ee2e3q"
  target_set_id = "ts-3eq"
  auto_weight   = true
}]
// --------------------------------------------------------------------

// --- HEALTH-CHECKS --------------------------------------------------
health_check = {
  health_check_type = "INSTANCE_STATE"
  grace_period      = 120
  auto_healing      = true
}
// --------------------------------------------------------------------

// --- NETWORK --------------------------------------------------------
network = {
  virtual_network_name = "vname"
  subnet_name          = "my-subnet-name"
  resource_group_name  = "subnetResourceGroup"
  assign_public_ip     = true
}
// --------------------------------------------------------------------

// --- LOGIN ----------------------------------------------------------
login = {
  user_name      = "admin"
  ssh_public_key = "33a2s1f3g5a1df5g1ad3f2g1adfg56dfg=="
}
// --------------------------------------------------------------------

// --- SCHEDULED TASK -------------------------------------------------
scheduled_task = [{
  is_enabled      = true
  cron_expression = "* * * * *"
```

```
      task_type       = "scale"

      scale_min_capacity = 5
      scale_max_capacity = 8
      adjustment         = 2

      adjustment_percentage = 50
      scale_target_capacity = 6
      batch_size_percentage = 33
      grace_period          = 300
 }]
 // ----------------------------------------------------------------

 // --- SCALING POLICIES -------------------------------------------
   scaling_up_policy = [{
       policy_name = "policy-name"
       metric_name = "CPUUtilization"
       namespace   = "Microsoft.Compute"
       statistic   = "average"
       threshold   = 10
       unit        = "percent"
       cooldown    = 60

       dimensions = [
         {
           name  = "resourceName"
           value = "resource-name"
         },
         {
           name  = "resourceGroupName"
           value = "resource-group-name"
         },
       ]

       operator            = "gt"
       evaluation_periods  = "10"
       period              = "60"
       action_type         = "setMinTarget"
       min_target_capacity = 1
     }]

   scaling_down_policy = [{
       policy_name = "policy-name"
       metric_name = "CPUUtilization"
       namespace   = "Microsoft.Compute"
       statistic   = "average"
```

```
        threshold   = 10
        unit        = "percent"
        cooldown    = 60

        dimensions = {
            name  = "name-1"
            value = "value-1"
        }

        operator           = "gt"
        evaluation_periods = "10"
        period             = "60"
        action_type        = "adjustment"
        adjustment         = "MIN(5,10)"
    }]
}
 // -----------------------------------------------------------------
```

» Argument Reference

The following arguments are supported:

- `name` - (Required) The group name.

- `region` - (Required) The region your Azure group will be created in.

- `resource_group_name` - (Required) Name of the Resource Group for Elastigroup.

- `product` - (Required) Operation system type. Valid values: `"Linux"`, `"Windows"`.

- `max_size` - (Required) The maximum number of instances the group should have at any time.

- `min_size` - (Required) The minimum number of instances the group should have at any time.

- `desired_capacity` - (Required) The desired number of instances the group should have at any time.

- `od_sizes` - (Required) Available On-Demand sizes

- `low_priority_sizes` - (Required) Available Low-Priority sizes.

- `user_data` - (Optional) Base64-encoded MIME user data to make available to the instances.

- `shutdown_script` - (Optional) Shutdown script for the group. Value should be passed as a string encoded at Base64 only.

- **managed_service_identity** - (Optional) Add a user-assigned managed identity to the VMs in the cluster.

  - **resource_group_name** - (Required) The Resource Group that the user-assigned managed identity resides in.
  - **name** - (Required) The name of the managed identity.

- **strategy** - (Required) Describes the deployment strategy.

- **low_priority_percentage** - (Optional, Default 100) Percentage of Low Priority instances to maintain. Required if **od_count** is not specified.

- **od_count** - (Optional) Number of On-Demand instances to maintain. Required if low_priority_percentage is not specified.

- **draining_timeout** - (Optional, Default 120) Time (seconds) to allow the instance to be drained from incoming TCP connections and detached from MLB before terminating it during a scale-down operation.

» Load Balancers

- **load_balancers** - (Required) Describes a set of one or more classic load balancer target groups and/or Multai load balancer target sets.
- **type** - (Required) The resource type. Valid values: CLASSIC, TARGET_GROUP, MULTAI_TARGET_SET.
- **balancer_id** - (Required) The balancer ID.
- **target_set_id** - (Required) The scale set ID associated with the load balancer.
- **auto_weight** - (Optional, Default: **false**)

```
load_balancers = [{
  type         = "MULTAI_TARGET_SET"
  balancer_id  = "lb-1ee2e3q"
  target_set_id = "ts-3eq"
  auto_weight  = true
}]
```

» Image

- **image** - (Required) Image of a VM. An image is a template for creating new VMs. Choose from Azure image catalogue (marketplace) or use a custom image.
- **publisher** - (Optional) Image publisher. Required if resource_group_name is not specified.
- **offer** - (Optional) Name of the image to use. Required if publisher is specified.

- **sku** - (Optional) Image's Stock Keeping Unit, which is the specific version of the image. Required if publisher is specified.
- **resource_group_name** - (Optional) Name of Resource Group for custom image. Required if publisher not specified.
- **image_name** - (Optional) Name of the custom image. Required if resource_group_name is specified.

```
// market image
image = {
  marketplace = {
    publisher = "Canonical"
    offer     = "UbuntuServer"
    sku       = "16.04-LTS"
  }
}


// custom image
image = {
  custom = {
    image_name          = "customImage"
    resource_group_name = "resourceGroup"
  }
}
```

» Health Check

- **health_check** - (Optional) Describes the health check configuration.
- **health_check_type** - (Optional) Health check used to validate VM health. Valid values: "INSTANCE_STATE".
- **grace_period** - (Optional) Period of time (seconds) to wait for VM to reach healthiness before monitoring for unhealthiness.
- **auto_healing** - (Optional) Enable auto-healing of unhealthy VMs.

```
health_check = {
  health_check_type = "INSTANCE_STATE"
  grace_period      = 120
  auto_healing      = true
}
```

» Network

- **network** - (Required) Defines the Virtual Network and Subnet for your Elastigroup.
- **virtual_network_name** - (Required) Name of Vnet.
- **subnet_name** - (Required) ID of subnet.

- `resource_group_name` - (Required) Vnet Resource Group Name.
- `assign_public_up` - (Optional, Default: `false`) Assign a public IP to each VM in the Elastigroup.
- `additional_ip_configs` - (Optional) Array of additional IP configuration objects.
- `name` - (Required) The IP configuration name.
- `private_ip_version` - (Optional) Available from Azure Api-Version 2017-03-30 onwards, it represents whether the specific ipconfiguration is IPv4 or IPv6. Valid values: IPv4, IPv6.

```
network = {
  virtual_network_name = "vname"
  subnet_name          = "my-subnet-name"
  resource_group_name  = "subnetResourceGroup"
  assign_public_ip     = true

  additional_ip_configs = [{
    name = "test"
    private_ip_version = "IPv4"
  }]
}
```

» Login

```
network = {
  virtual_network_name = "vname"
  subnet_name          = "my-subnet-name"
  resource_group_name  = "subnetResourceGroup"
  assign_public_ip     = true
}
```

» Login

- `login` - (Required) Describes the login configuration.
- `user_name` - (Required) Set admin access for accessing your VMs.
- `ssh_public_key` - (Optional) SSH for admin access to Linux VMs. Required for Linux product types.
- `password` - (Optional) Password for admin access to Windows VMs. Required for Windows product types.

```
login = {
  user_name      = "admin"
  ssh_public_key = "33a2s1f3g5a1df5g1ad21651sag56dfg=="
}
```

» Scaling Policies

Each `scaling_*_policy` supports the following:

- `policy_name` - (Optional) The name of the policy.
- `metric_name` - (Required) Metric to monitor by Azure metric display name.
- `namespace` - (Optional, Default: `"Microsoft.Compute"`) The namespace for the alarm's associated metric. Valid values:

```
Microsoft.AnalysisServices/servers
Microsoft.ApiManagement/service
Microsoft.Automation/automationAccounts
Microsoft.Batch/batchAccounts
Microsoft.Cache/redis
Microsoft.CognitiveServices/accounts
Microsoft.Compute
Microsoft.ContainerInstance/containerGroups
Microsoft.ContainerService/managedClusters
Microsoft.CustomerInsights/hubs
Microsoft.DataFactory/datafactories
Microsoft.DataFactory/factories
Microsoft.DataLakeAnalytics/accounts
Microsoft.DataLakeStore/accounts
Microsoft.DBforMariaDB/servers
Microsoft.DBforMySQL/servers
Microsoft.DBforPostgreSQL/servers
Microsoft.Devices/IotHubs
Microsoft.Devices/provisioningServices
Microsoft.DocumentDB/databaseAccounts
Microsoft.EventGrid/eventSubscriptions
Microsoft.EventGrid/extensionTopics
Microsoft.EventGrid/topics
Microsoft.EventHub/clusters
Microsoft.EventHub/namespaces
Microsoft.HDInsight/clusters
Microsoft.Insights/AutoscaleSettings
Microsoft.Insights/Components
Microsoft.KeyVault/vaults
Microsoft.Kusto/Clusters
Microsoft.LocationBasedServices/accounts
Microsoft.Logic/workflows
Microsoft.NetApp/netAppAccounts/capacityPools/Volumes
Microsoft.NetApp/netAppAccounts/capacityPools
Microsoft.Network/applicationGateways
Microsoft.Network/dnszones
```

```
Microsoft.Network/connections
Microsoft.Network/expressRouteCircuits
Microsoft.Network/expressRouteCircuits/peerings
Microsoft.Network/frontdoors
Microsoft.Network/loadBalancers
Microsoft.Network/networkInterfaces
Microsoft.Network/networkWatchers/connectionMonitors
Microsoft.Network/publicIPAddresses
Microsoft.Network/trafficManagerProfiles
Microsoft.Network/virtualNetworkGateways
Microsoft.NotificationHubs/Namespaces/NotificationHubs
Microsoft.OperationalInsights/workspaces
Microsoft.PowerBIDedicated/capacities
Microsoft.Relay/namespaces
Microsoft.Search/searchServices
Microsoft.ServiceBus/namespaces
Microsoft.SignalRService/SignalR
Microsoft.Sql/managedInstances
Microsoft.Sql/servers/databases
Microsoft.Sql/servers/elasticPools
Microsoft.Storage/storageAccounts
Microsoft.Storage/storageAccounts/blobServices
Microsoft.Storage/storageAccounts/fileServices
Microsoft.Storage/storageAccounts/queueServices
Microsoft.Storage/storageAccounts/tableServices
Microsoft.StreamAnalytics/streamingjobs
Microsoft.TimeSeriesInsights/environments
Microsoft.TimeSeriesInsights/environments/eventsources
Microsoft.Web/hostingEnvironments/multiRolePools
Microsoft.Web/hostingEnvironments/workerPools
Microsoft.Web/serverfarms
Microsoft.Web/sites (excluding functions)
Microsoft.Web/sites (functions)
Microsoft.Web/sites/slots
```

- `statistic` - (Optional) The metric statistics to return. Valid values: `average`.

- `threshold` - (Required) The value against which the specified statistic is compared.

- `unit` - (Required) The unit for the alarm's associated metric. Valid values: `"percent, "seconds", "microseconds", "milliseconds", "bytes", "kilobytes", "megabytes", "gigabytes", "terabytes", "bits", "kilobits", "megabits", "gigabits", "terabits", "count", "bytes/second", "kilobytes/second", "megabytes/second", "gigabytes/second", "terabytes/second", "bits/second", "kilobits/second", "megabits/second",`

"gigabits/second", "terabits/second", "count/second", "none".

- `cooldown` - (Optional, Default: `300`) The amount of time, in seconds, after a scaling activity completes and before the next scaling activity can start. If this parameter is not specified, the default cooldown period for the group applies.

- `operator` - (Optional, Scale Up Default: `gte`, Scale Down Default: `lte`) The operator to use in order to determine if the scaling policy is applicable. Valid values: "gt", "gte", "lt", "lte".

- `evaluation_periods` - (Optional, Default: `1`) The number of periods over which data is compared to the specified threshold.

- `period` - (Optional, Default: `300`) The granularity, in seconds, of the returned datapoints. Period must be at least 60 seconds and must be a multiple of 60.

- `dimensions` - (Optional) A list of dimensions describing qualities of the metric. Required when `namespace` is defined AND not "Microsoft.Compute".

  - `name` - (Required) The dimension name.
  - `value` - (Optional) The dimension value.

When `namespace` is defined and is not "Microsoft.Compute" the list of dimensions must contain the following:

```
dimensions = [
  {
    name  = "resourceName"
    value = "example-resource-name"
  },
  {
    name  = "resourceGroupName"
    value = "example-resource-group-name"
  },
]
```

- `action_type` - (Optional; if not using `min_target_capacity` or `max_target_capacity`) The type of action to perform for scaling. Valid values: "adjustment", "percentageAdjustment", "setMaxTarget", "setMinTarget", "updateCapacity".

If you do not specify an action type, you can only use – `adjustment`, `min_target_capacity`, `max_target_capacity`. While using action_type, please also set the following:

When using `adjustment` – set the field `adjustment` When using `percentageAdjustment` - set the field `adjustment` When using `setMaxTarget` – set the field

max_target_capacity When using **setMinTarget** – set the field **min_target_capacity** When using **updateCapacity** – set the fields **minimum**, **maximum**, and **target**

- **adjustment** - (Optional) Value to which the action type will be adjusted. Required if using **numeric** or **percentage_adjustment** action types.

- **min_target_capacity** - (Optional; if not using **adjustment**; available only for scale up). The number of the desired target (and minimum) capacity

- **max_target_capacity** - (Optional; if not using **adjustment**; available only for scale down). The number of the desired target (and maximum) capacity

- **minimum** - (Optional; if using **updateCapacity**) The minimal number of instances to have in the group.

- **maximum** - (Optional; if using **updateCapacity**) The maximal number of instances to have in the group.

- **target** - (Optional; if using **updateCapacity**) The target number of instances to have in the group.

Usage:

```
// --- SCALE DOWN POLICY ------------------
  scaling_down_policy = [{
    policy_name = "policy-name"
    metric_name = "CPUUtilization"
    namespace   = "Microsoft.Compute"
    statistic   = "average"
    threshold   = 10
    unit        = "percent"
    cooldown    = 60

    dimensions = {
      name  = "name-1"
      value = "value-1"
    }

    operator           = "gt"
    evaluation_periods = "10"
    period             = "60"

    // === MIN TARGET ==================
    # action_type       = "setMinTarget"
    # min_target_capacity = 1
    // ================================

    // === ADJUSTMENT ==================
```

```
    action_type   = "adjustment"
    # action_type = "percentageAdjustment"
    adjustment    = "MIN(5,10)"
    // ================================

    // === UPDATE CAPACITY =============
    # action_type = "updateCapacity"
    # minimum     = 0
    # maximum     = 10
    # target      = 5
    // ================================

  }]
// ------------------------------------

// --- SCALE DOWN POLICY -----------------
  scaling_down_policy = [{
    policy_name = "policy-name-update"
    metric_name = "CPUUtilization"
    namespace   = "Microsoft.Compute"
    statistic   = "sum"
    threshold   = 5
    unit        = "bytes"
    cooldown    = 120

    dimensions = {
        name  = "name-1-update"
        value = "value-1-update"
    }

    operator          = "lt"
    evaluation_periods = 5
    period            = 120

    //// === MIN TARGET ==================
    # action_type         = "setMinTarget"
    # min_target_capacity = 1
    //// ===============================

    // === ADJUSTMENT ==================
    # action_type = "percentageAdjustment"
    # action_type = "adjustment"
    # adjustment  = "MAX(5,10)"
    // ================================

    // === UPDATE CAPACITY =============
```

```
    action_type = "updateCapacity"
    minimum     = 0
    maximum     = 10
    target      = 5
    // =================================

  }]
// -------------------------------------
```

» Scheduling

- `scheduled_task` - (Optional) Describes the configuration of one or more scheduled tasks.
- `is_enabled` - (Optional, Default: `true`) Describes whether the task is enabled. When true the task should run when false it should not run.
- `cron_expression` - (Required) A valid cron expression (`* * * * *`). The cron is running in UTC time zone and is in Unix cron format Cron Expression Validator Script.
- `task_type` - (Required) The task type to run. Valid Values: `backup_ami`, `scale`, `scaleUp`, `roll`, `statefulUpdateCapacity`, `statefulRecycle`.
- `scale_min_capacity` - (Optional) The min capacity of the group. Should be used when choosing 'task_type' of 'scale'.
- `scale_max_capacity` - (Optional) The max capacity of the group. Required when 'task_type' is 'scale'.
- `scale_target_capacity` - (Optional) The target capacity of the group. Should be used when choosing 'task_type' of 'scale'.
- `adjustment` - (Optional) The number of instances to add/remove to/from the target capacity when scale is needed.
- `adjustment_percentage` - (Optional) The percent of instances to add/remove to/from the target capacity when scale is needed.
- `batch_size_percentage` - (Optional) The percentage size of each batch in the scheduled deployment roll. Required when the 'task_type' is 'roll'.
- `grace_period` - (Optional) The time to allow instances to become healthy.

```
scheduled_task = [{
  is_enabled      = true
  cron_expression = "* * * * *"
  task_type       = "scale"

  scale_min_capacity = 5
  scale_max_capacity = 8
  adjustment         = 2

  adjustment_percentage = 50
  scale_target_capacity = 6
```

```
    batch_size_percentage = 33
    grace_period          = 300
}]
```

» Update Policy

- `update_policy` - (Optional)

    - `should_roll` - (Required) Sets the enablement of the roll option.
    - `roll_config` - (Required) While used, you can control whether the group should perform a deployment after an update to the configuration.
        * `batch_size_percentage` - (Required) Sets the percentage of the instances to deploy in each batch.
        * `health_check_type` - (Optional) Sets the health check type to use. Valid values: `"INSTANCE_STATE"`, `"NONE"`.
        * `grace_period` - (Optional) Sets the grace period for new instances to become healthy.

```
update_policy = {
  should_roll = false

  roll_config = {
    batch_size_percentage = 33
    health_check_type     = "INSTANCE_STATE"
    grace_period          = 300
  }
}
```

» Third-Party Integrations

- `integration_kubernetes` - (Optional) Describes the Kubernetes integration.
    - `cluster_identifier` - (Required) The cluster ID.

Usage:

```
integration_kubernetes = {
  cluster_identifier = "k8s-cluster-id"
}
```

- `integration_multai_runtime` - (Optional) Describes the Multai Runtime integration.
    - `deployment_id` - (Optional) The deployment id you want to get

Usage:

```
  integration_multai_runtime = {
    deployment_id = ""
  }
```

» spotinst_elastigroup_gcp

Provides a Spotinst elastigroup GCP resource.

» Example Usage

```
resource "spotinst_elastigroup_gcp" "example" {
  name            = "example-gcp"
  description     = "spotinst gcp group"
  service_account = "example@myProject.iam.gservicecct.com"
  startup_script  = ""

  min_size         = 0
  max_size         = 1
  desired_capacity = 1

  availability_zones = ["asia-east1-c", "us-central1-a"]

  preemptible_percentage = 50
  # on_demand_count        = 2
  fallback_to_ondemand        = true
  draining_timeout       = 180

  labels = [
    {
      key = "test_key"
      value = "test_value"
    }
  ]

  tags             = ["http", "https"]

  backend_services_config = [
    {
      service_name = "spotinst-elb-backend-service"
      ports = {
        port_name = "port-name"
        ports = [8000, 6000]
      }
```

```
    },
]

disks = [
  {
    device_name = "device"
    mode        = "READ_WRITE"
    type        = "PERSISTENT"
    auto_delete = true
    boot        = true
    interface   = "SCSI"

    initialize_params = {
      disk_size_gb = 10
      disk_type    = "pd-standard"
      source_image = ""
    }
  }
]

network_interface = [
  {
    network = "spot-network"
  }
]

instance_types_ondemand    = ["n1-standard-1"]
instance_types_preemptible = ["n1-standard-1", "n1-standard-2"]

instance_types_custom = [
  {
    vCPU      = 2
    memoryGiB = 7.5
  }
]

subnets = [
  {
    region       = "asia-east1"
    subnet_names = ""
  }
]

scaling = {
  up = {
    policy_name = "scale_up_1"
```

```
    source      = "stackdriver"
    metric_name = "instance/disk/read_ops_count"
    namespace   = "compute"
    statistic   = "average"
    unit        = "percent"
    threshold   = 10000
    period      = 300
    cooldown    = 300
    operator    = "gte"

    evaluation_periods = 1

    action = {
      type       = "adjustment"
      adjustment = 1
    }

    dimensions = [
      {
        name  = "storage_type"
        value = "pd-ssd"
      }
    ]
  }
 }
}
```

» Argument Reference

The following arguments are supported:

- `name` - (Required) The group name.
- `description` - (Optional) The region your GCP group will be created in.
- `startup_script` - (Optional) Create and run your own startup scripts on your virtual machines to perform automated tasks every time your instance boots up.
- `shutdown_script` - (Optional) The Base64-encoded shutdown script that executes prior to instance termination, for more information please see: Shutdown Script
- `service_account` - (Optional) The email of the service account in which the group instances will be launched.
- `max_size` - (Required) The maximum number of instances the group should have at any time.
- `min_size` - (Required) The minimum number of instances the group should have at any time.

- `desired_capacity` - (Required) The desired number of instances the group should have at any time.
- `availability_zones` - (Required) List of availability zones for the group.
- `subnets` - (Optional) A list of regions and subnets.
    - `region` - (Required) The region for the group of subnets.
    - `subnet_names` - (Required) The names of the subnets in the region.
- `instance_types_preemptible` - (Required) The preemptible VMs instance type. To maximize cost savings and market availability, select as many types as possible. Required if instance_types_ondemand is not set.
- `instance_types_ondemand` - (Required) The regular VM instance type to use for mixed-type groups and when falling back to on-demand. Required if instance_types_preemptible is not set.
- `instance_types_custom` - (Required) Defines a set of custom instance types. Required if instance_types_preemptible and instance_types_ondemand are not set.
    - `vCPU` - (Optional) The number of vCPUs in the custom instance type. GCP has a number of limitations on accepted vCPU values. For more information, see the GCP documentation (here.)[https://cloud.google.com/compute/docs/instances/creating-instance-with-custom-machine-type#specifications]
    - `memory_gib` - (Optional) The memory (in GiB) in the custom instance types. GCP has a number of limitations on accepted memory values.For more information, see the GCP documentation (here.)[https://cloud.google.com/compute/docs/instances/creating-instance-with-custom-machine-type#specifications]
- `preemptible_percentage` - (Optional) Percentage of Preemptible VMs to spin up from the "desired_capacity".
- `on_demand_count` - (Optional) Number of regular VMs to launch in the group. The rest will be Preemptible VMs. When this parameter is specified, the preemptible_percentage parameter is being ignored.
- `fallback_to_ondemand` - (Optional) Activate fallback-to-on-demand. When provisioning an instance, if no Preemptible market is available, fallback-to-on-demand will provision an On-Demand instance to maintain the group capacity.
- `draining_timeout` - (Optional) Time (seconds) the instance is allowed to run after it is detached from the group. This is to allow the instance time to drain all the current TCP connections before terminating it.
- `metadata` - (Optional) Array of objects with key-value pairs.
    - `key` - (Optional) Metadata key.
    - `value` - (Optional) Metadata value.
- `labels` - (Optional) Array of objects with key-value pairs.
    - `key` - (Optional) Labels key.
    - `value` - (Optional) Labels value.
- `tags` - (Optional) Tags to mark created instances.

» GPU

- `gpu` - (Optional) Defines the GPU configuration.
  - `type` - (Required) The type of GPU instance. Valid values:
    `nvidia-tesla-v100`, `nvidia-tesla-p100`, `nvidia-tesla-k80`.
  - `count` - (Required) The number of GPUs. Must be 0, 2, 4, 6, 8.

Usage:

```
gpu = {
  count = 2
  type = "nvidia-tesla-p100"
}
```

» Health Check

- `auto_healing` - (Optional) Enable auto-replacement of unhealthy
  instances.
- `health_check_grace_period` - (Optional) Period of time (seconds) to
  wait for VM to reach healthiness before monitoring for unhealthiness.
- `health_check_type` - (Optional) The kind of health check to perform
  when monitoring for unhealthiness.
- `unhealthy_duration` - (Optional) Period of time (seconds) to remain in
  an unhealthy status before a replacement is triggered.

```
auto_health              = true
health_check_grace_period = 100
health_check_type        = ""
unhealthy_duration       = ""
```

» Backend Services

- `backend_services` - (Optional) Describes the backend service configura-
  tions.
  - `service_name` - (Required) The name of the backend service.
  - `location_type` - (Optional) Sets which location the backend services
    will be active. Valid values: `regional`, `global`.
  - `scheme` - (Optional) Use when `location_type` is "regional". Set the
    traffic for the backend service to either between the instances in the
    vpc or to traffic from the internet. Valid values: `INTERNAL`, `EXTERNAL`.
  - `named_port` - (Optional) Describes a named port and a list of ports.
    * `port_name` - (Required) The name of the port.
    * `ports` - (Required) A list of ports.

Usage:

```
backend_services_config = [
  {
    service_name = "spotinst-elb-backend-service"
    locationType = "regional"
    scheme       = "INTERNAL"
    ports = {
      port_name = "port-name"
      ports = [8000, 6000]
    }
  },
]
```

» Disks

- `disks` - (Optional) Array of disks associated with this instance. Persistent disks must be created before you can assign them.
  - `auto_delete` - (Optional) Specifies whether the disk will be auto-deleted when the instance is deleted.
  - `boot` - (Optional) Indicates that this is a boot disk. The virtual machine will use the first partition of the disk for its root filesystem.
  - `device_name` - (Optional) Specifies a unique device name of your choice.
  - `interface` - (Optional, Default: `SCSI`) Specifies the disk interface to use for attaching this disk, which is either SCSI or NVME.
  - `mode` - (Optional, Default: `READ_WRITE`) The mode in which to attach this disk, either READ_WRITE or READ_ONLY.
  - `source` - (Optional) Specifies a valid partial or full URL to an existing Persistent Disk resource. This field is only applicable for persistent disks.
  - `type` - (Optional, Default: `PERSISTENT`) Specifies the type of disk, either SCRATCH or PERSISTENT.
  - `initialize_params` - (Optional) Specifies the parameters for a new disk that will be created alongside the new instance. Use initialization parameters to create boot disks or local SSDs attached to the new instance.
    * `disk_size_gb` - (Optional) Specifies disk size in gigabytes. Must be in increments of 2.
    * `disk_type` - (Optional, Default" `pd-standard`) Specifies the disk type to use to create the instance. Valid values: pd-ssd, local-ssd.
    * `source_image` - (Optional) A source image used to create the disk. You can provide a private (custom) image, and Compute Engine will use the corresponding image from your project.

Usage:

```
disks = [
```

```
  {
    device_name = "device"
    mode        = "READ_WRITE"
    type        = "PERSISTENT"
    auto_delete = true
    boot        = true
    interface   = "SCSI"

    initialize_params = {
      disk_size_gb = 10
      disk_type    = "pd-standard"
      source_image = ""
    }
  }
]
```

» Network Interfaces

Each of the `network_interface` attributes controls a portion of the GCP Instance's "Network Interfaces". It's a good idea to familiarize yourself with GCP's Network Interfaces docs to understand the implications of using these attributes.

- `network_interface` - (Required, minimum 1) Array of objects representing the network configuration for the elastigroup.
    - `network` - (Required) Network resource for this group.
    - `access_configs` - (Optional) Array of configurations.
        * `name` - (Optional) Name of this access configuration.
        * `type` - (Optional) Array of configurations for this interface. Currently, only ONE_TO_ONE_NAT is supported.

```
network_interface = [{
  network = "default"

  access_configs = {
    name = "config1"
    type = "ONE_TO_ONE_NAT"
  }

  alias_ip_ranges = {
   subnetwork_range_name = "range-name-1"
   ip_cidr_range = "10.128.0.0/20"
  }
}]
```

» Scaling Policies

- `scaling_up_policy` - (Optional) Contains scaling policies for scaling the Elastigroup up.
- `scaling_down_policy` - (Optional) Contains scaling policies for scaling the Elastigroup down.

Each `scaling_*_policy` supports the following:

- `policy_name` - (Optional) Name of scaling policy.
- `metric_name` - (Optional) Metric to monitor. Valid values: "Percentage CPU", "Network In", "Network Out", "Disk Read Bytes", "Disk Write Bytes", "Disk Write Operations/Sec", "Disk Read Operations/Sec".
- `statistic` - (Optional) Statistic by which to evaluate the selected metric. Valid values: "AVERAGE", "SAMPLE_COUNT", "SUM", "MINIMUM", "MAXIMUM", "PERCENTILE", "COUNT".
- `threshold` - (Optional) The value at which the scaling action is triggered.
- `period` - (Optional) Amount of time (seconds) for which the threshold must be met in order to trigger the scaling action.
- `evaluation_periods` - (Optional) Number of consecutive periods in which the threshold must be met in order to trigger a scaling action.
- `cooldown` - (Optional) Time (seconds) to wait after a scaling action before resuming monitoring.
- `operator` - (Optional) The operator used to evaluate the threshold against the current metric value. Valid values: "gt" (greater than), "get" (greater-than or equal), "lt" (less than), "lte" (less than or equal).
- `action` - (Optional) Scaling action to take when the policy is triggered.
  - `type` - (Optional) Type of scaling action to take when the scaling policy is triggered. Valid values: "adjustment", "setMinTarget", "updateCapacity", "percentageAdjustment"
  - `adjustment` - (Optional) Value to which the action type will be adjusted. Required if using "numeric" or "percentageAdjustment" action types.
- `dimensions` - (Optional) A list of dimensions describing qualities of the metric.
  - `name` - (Required) The dimension name.
  - `value` - (Required) The dimension value.

Usage:

```
scaling = {
    up = {
      policy_name = "scale_up_1"
      source      = "stackdriver"
      metric_name = "instance/disk/read_ops_count"
      namespace   = "compute"
      statistic   = "average"
```

```
        unit        = "percent"
        threshold   = 10000
        period      = 300
        cooldown    = 300
        operator    = "gte"

        evaluation_periods = 1

        action = {
          type        = "adjustment"
          adjustment = 1
        }

        dimensions = [
          {
            name  = "storage_type"
            value = "pd-ssd"
          }
        ]
    }
  }
```

» Third-Party Integrations

- `integration_docker_swarm` - (Optional) Describes the Docker Swarm integration.
    - `master_host` - (Required) IP or FQDN of one of your swarm managers.
    - `master_port` - (Required) Network port used by your swarm.

Usage:

```
integration_docker_swarm = {
    master_host = "10.10.10.10"
    master_port = 2376
}
```

» spotinst_elastigroup_gke

Provides a Spotinst Elastigroup GKE resource. Please see Importing a GKE cluster for detailed information.

» Example Usage

A spotinst_elastigroup_gke supports all of the fields defined in spotinst_elastigroup_gcp.

There are two main differences:

- you must include `cluster_zone_name` and `cluster_id`
- a handful of parameters are created remotely and will not appear in the diff. A complete list can be found below.

```
resource "spotinst_elastigroup_gke" "example-gke-elastigroup" {
 name              = "example-gke"
 // cluster_id        = "terraform-acc-test-cluster" // deprecated
 cluster_zone_name = "us-central1-a"
 node_image        = "COS"

 // --- CAPACITY ------------
 max_size         = 5
 min_size         = 1
 desired_capacity = 3
 // ------------------------

 // --- INSTANCE TYPES --------------------------------
 instance_types_ondemand    = "n1-standard-1"
 instance_types_preemptible = ["n1-standard-1", "n1-standard-2"]
 // --------------------------------------------------

 // --- STRATEGY -------------------
 preemptible_percentage = 100
 // -------------------------------

 integration_gke = {
  location                 = "us-central1-a"
  cluster_id               = "example-cluster-id"
  autoscale_is_enabled     = true
  autoscale_is_auto_config = false
  autoscale_cooldown       = 300

  autoscale_headroom = {
    cpu_per_unit    = 1024
    memory_per_unit = 512
    num_of_units    = 2
  }

  autoscale_down = {
    evaluation_periods = 300
  }
```

```
  autoscale_labels = {
    key   = "label_key"
    value = "label_value"
  }
}

  backend_services = [{
    service_name  = "backend-service"
    location_type = "global"
    named_ports = {
      name  = "http"
      ports = [80, 8080]
    }
  }]

}
```

» Argument Reference

All `spotisnt_elastigroup_gcp` arguments are supported. Please be sure to
include the following parameters in your `spotinst_elastigroup_gke` template:

- `cluster_zone_name` - (Required) The zone where the cluster is hosted.
- `cluster_id` - (Required) The name of the GKE cluster you wish to import.
- `node_image` - (Optional, Default: `COS`) The image that will be used for
  the node VMs. Possible values: COS, UBUNTU.

» Third-Party Integrations

- `integration_gke` - (Required) Describes the GKE integration.

  - `location` - (Optional) The location of your GKE cluster.
  - `cluster_id` - (Optional) The GKE cluster ID you wish to import.
  - `autoscale_is_enabled` - (Optional, Default: `false`) Specifies
    whether the auto scaling feature is enabled.
  - `autoscale_is_autoconfig` - (Optional, Default: `false`) Enabling
    the automatic auto-scaler functionality. For more information please
    see: .
  - `autoscale_cooldown` - (Optional, Default: `300`) The amount of
    time, in seconds, after a scaling activity completes before any further
    trigger-related scaling activities can start.
  - `autoscale_headroom` - (Optional) Headroom for the cluster.
    * `cpu_per_unit` - (Optional, Default: `0`) Cpu units for compute.

* `memory_per_unit` - (Optional, Default: `0`) RAM units for compute.
  * `num_of_units` - (Optional, Default: `0`) Amount of units for compute.
  – `autoscale_down` - (Optional) Enabling scale down.
    * `evaluation_periods` - (Optional, Default: `5`) Amount of cooldown evaluation periods for scale down.
  – `autoscale_labels` - (Optional) Labels to assign to the resource.
    * `key` - (Optional) The label name.
    * `value` - (Optional) The label value.

Usage:

```
integration_gke = {
 location = "us-central1-a"
 cluster_id = "terraform-acc-test-cluster"
 autoscale_is_enabled    = true
 autoscale_is_auto_config = false
 autoscale_cooldown      = 300

 autoscale_headroom = {
   cpu_per_unit    = 1024
   memory_per_unit = 512
   num_of_units    = 2
 }

 autoscale_down = {
   evaluation_periods = 300
 }

 autoscale_labels = {
   key   = "label_key"
   value = "label_value"
 }
}
```

» Diff-suppressed Parameters

The following parameters are created remotely and imported. The diffs have been suppressed in order to maintain plan legibility. You may update the values of these imported parameters by defining them in your template with your desired new value (including null values).

- `backend_services`
  – `service_name`
  – `location_type`

- scheme
- named_port
  - * port_name
  - * ports
- labels
  - key
  - value
- metadata
  - key
  - value
- tags
  - key
  - value
- service_account
- ip_forwarding
- fallback_to_od
- subnets
  - region
  - subnet_name

» spotinst_mrscaler

Provides a Spotinst AWS MrScaler resource.

» Example Usage - New Strategy

```
# Create a Mr Scaler with New strategy

resource "spotinst_mrscaler_aws" "Terraform-MrScaler-01" {
  name        = "Terraform-MrScaler-01"
  description = "Testing MrScaler creation via Terraform"
  region      = "us-west-2"
  strategy    = "new"
  release_label = "emr-5.17.0"

  availability_zones = ["us-west-2a:subnet-123456"]

  provisioning_timeout = {
    timeout        = 15
    timeout_action = "terminate"
  }

// --- CLUSTER ------------
```

```
  log_uri         = "s3://example-logs"
  additional_info = "{'test':'more information'}"
  job_flow_role   = "EMR_EC2_ExampleRole"
  security_config = "example-config"
  service_role    = "example-role"

  termination_protected = false
  keep_job_flow_alive   = true
// -----------------------

// --- OPTONAL COMPUTE -----
  custom_ami_id        = "ami-123456"
  repo_upgrade_on_boot = "NONE"
  ec2_key_name         = "test-key"

  managed_primary_security_group = "sg-123456"
  managed_replica_security_group = "sg-987654"
  service_access_security_group  = "access-example"

  additional_primary_security_groups = ["sg-456321"]
  additional_replica_security_groups = ["sg-123654"]

  applications = [
    {
      name = "Ganglia"
      version = "1.0"
    },
    {
      name = "Hadoop"
    },
    {
      name = "Pig"
      args = ["fake", "args"]
    }
  ]

  instance_weights = [
    {
      instance_type     = "t2.small"
      weighted_capacity = 10
    },
    {
      instance_type     = "t2.medium"
      weighted_capacity = 90
    }
  ]
```

```
  steps_file = {
    bucket = "example-bucket"
    key = "steps.json"
  }

  configurations_file = {
    bucket = "example-bucket"
    key = "configurations.json"
  }

  bootstrap_actions_file = {
    bucket = "terraform-emr-test"
    key = "bootstrap-actions.json"
  }
// ------------------------

// --- MASTER GROUP -------------
  master_instance_types = ["c3.xlarge"]
  master_lifecycle      = "SPOT"
  master_ebs_optimized  = true

  master_ebs_block_device = {
    volumes_per_instance = 1
    volume_type          = "gp2"
    size_in_gb           = 30
  }
// -----------------------------

// --- CORE GROUP -------------
  core_instance_types = ["c3.xlarge", "c4.xlarge"]
  core_min_size         = 1
  core_max_size         = 1
  core_desired_capacity = 1
  core_lifecycle        = "ON_DEMAND"
  core_ebs_optimized    = false

  core_ebs_block_device = {
    volumes_per_instance = 2
    volume_type          = "gp2"
    size_in_gb           = 40
  }
// --------------------------

// --- TASK GROUP -------------
  task_instance_types = ["c3.xlarge", "c4.xlarge"]
```

```
  task_min_size       = 0
  task_max_size       = 30
  task_desired_capacity = 1
  task_lifecycle      = "SPOT"
  task_ebs_optimized  = false

  task_ebs_block_device = {
    volumes_per_instance = 2
    volume_type          = "gp2"
    size_in_gb           = 40
  }
// --------------------------

// --- TAGS ------------------
  tags = [{
      key   = "Creator"
      value = "Terraform"
  }]
// --------------------------
```

» Example Usage - Clone Strategy

```
# Create a Mr Scaler with Clone strategy and Task scaling

output "mrscaler-name" {
  value = "${spotinst_mrscaler_aws.Terraform-MrScaler-01.name}"
}

output "mrscaler-created-cluster-id" {
  value = "${spotinst_mrscaler_aws.Terraform-MrScaler-01.output_cluster_id}"
}

resource "spotinst_mrscaler_aws" "Terraform-MrScaler-01" {
  name        = "Terraform-MrScaler-01"
  description = "Testing MrScaler creation via Terraform"
  region      = "us-west-2"
  strategy    = "clone"

  cluster_id       = "j-123456789"
  expose_cluster_id = true

  availability_zones = ["us-west-2a:subnet-12345678"]

// --- MASTER GROUP -------------
  master_instance_types = ["c3.xlarge"]
```

```
  master_lifecycle      = "SPOT"
  master_ebs_optimized  = true

  master_ebs_block_device = {
    volumes_per_instance = 1
    volume_type          = "gp2"
    size_in_gb           = 30
  }
// ----------------------------

// --- CORE GROUP -------------
  core_instance_types   = ["c3.xlarge", "c4.xlarge"]
  core_min_size         = 1
  core_max_size         = 1
  core_desired_capacity = 1
  core_lifecycle        = "ON_DEMAND"
  core_ebs_optimized    = false

  core_ebs_block_device = {
    volumes_per_instance = 2
    volume_type          = "gp2"
    size_in_gb           = 40
  }
// --------------------------

// --- TASK GROUP -------------
  task_instance_types   = ["c3.xlarge", "c4.xlarge"]
  task_min_size         = 0
  task_max_size         = 30
  task_desired_capacity = 1
  task_lifecycle        = "SPOT"
  task_ebs_optimized    = false

  task_ebs_block_device = {
    volumes_per_instance = 2
    volume_type          = "gp2"
    size_in_gb           = 40
  }
// --------------------------

// --- TAGS ------------------
  tags = [{
      key   = "Creator"
      value = "Terraform"
  }]
// --------------------------
```

```
// --- TASK SCALING POLICY ------
  task_scaling_down_policy = [{
    policy_name = "policy-name"
    metric_name = "CPUUtilization"
    namespace   = "AWS/EC2"
    statistic   = "average"
    unit        = ""
    threshold   = 10
    adjustment  = "1"
    cooldown    = 60
    dimensions = {
      name  = "name-1"
      value = "value-1"
    }

    operator           = "gt"
    evaluation_periods = 10
    period             = 60

    action_type = ""
    minimum     = 0
    maximum     = 10
    target      = 5
    max_target_capacity = 1
  }]
// ---------------------------
```

» Example Usage - Wrap Strategy

```
# Create a Mr Scaler with Wrap strategy

resource "spotinst_mrscaler" "example-scaler-2" {
  name        = "spotinst-mr-scaler-2"
  description = "created by Terraform"
  region      = "us-west-2"
  strategy    = "wrap"
  cluster_id  = "j-27UVDEHXL4OQM"

// --- TASK GROUP -------------
  task_instance_types = ["c3.xlarge","c4.xlarge"]

  task_target    = 2
  task_minimum   = 0
  task_maximum   = 4
```

```
  task_lifecycle = "SPOT"

  task_ebs_block_device = {
    volumes_per_instance = 1
    volume_type          = "gp2"
    size_in_gb           = 20
  }
// --------------------------
}
```

» Argument Reference

The following arguments are supported:

- `name` - (Required) The MrScaler name.
- `description` - (Optional) The MrScaler description.
- `region` - (Required) The MrScaler region.
- `strategy` - (Required) The MrScaler strategy. Allowed values are `new` `clone` and `wrap`.
- `cluster_id` - (Optional) The MrScaler cluster id.
- `expose_cluster_id` - (Optional) Allow the `cluster_id` to set a Terraform output variable.

» Provisioning Timeout (Clone, New strategies)

- `timeout` - (Optional) The amount of time (minutes) after which the cluster is automatically terminated if it's still in provisioning status. Minimum: '15'.
- `timeout_action` - (Optional) The action to take if the timeout is exceeded. Valid values: `terminate`, `terminateAndRetry`.

» Cluster Configuration (New strategy only)

- `log_uri` - (Optional) The path to the Amazon S3 location where logs for this cluster are stored.
- `additional_info` - (Optional) This is meta information about third-party applications that third-party vendors use for testing purposes.
- `security_config` - (Optional) The name of the security configuration applied to the cluster.
- `service_role` - (Optional) The IAM role that will be assumed by the Amazon EMR service to access AWS resources on your behalf.
- `job_flow_role` - (Optional) The IAM role that was specified when the job flow was launched. The EC2 instances of the job flow assume this role.

- `termination_protected` - (Optional) Specifies whether the Amazon EC2 instances in the cluster are protected from termination by API calls, user intervention, or in the event of a job-flow error.
- `keep_job_flow_alive` - (Optional) Specifies whether the cluster should remain available after completing all steps.

» Task Group (Wrap, Clone, and New strategies)

- `task_instance_types` - (Required) The MrScaler instance types for the task nodes.
- `task_target` - (Required) amount of instances in task group.
- `task_maximum` - (Optional) maximal amount of instances in task group.
- `task_minimum` - (Optional) The minimal amount of instances in task group.
- `task_lifecycle` - (Required) The MrScaler lifecycle for instances in task group. Allowed values are 'SPOT' and 'ON_DEMAND'.
- `task_ebs_optimized` - (Optional) EBS Optimization setting for instances in group.
- `task_ebs_block_device` - (Required) This determines the ebs configuration for your task group instances. Only a single block is allowed.
    - `volumes_per_instance` - (Optional; Default 1) Amount of volumes per instance in the task group.
    - `volume_type` - (Required) volume type. Allowed values are 'gp2', 'io1' and others.
    - `size_in_gb` - (Required) Size of the volume, in GBs.
    - `iops` - (Optional) IOPS for the volume. Required in some volume types, such as io1.

» Core Group (Clone, New strategies)

- `core_instance_types` - (Required) The MrScaler instance types for the core nodes.
- `core_target` - (Required) amount of instances in core group.
- `core_maximum` - (Optional) maximal amount of instances in core group.
- `core_minimum` - (Optional) The minimal amount of instances in core group.
- `core_lifecycle` - (Required) The MrScaler lifecycle for instances in core group. Allowed values are 'SPOT' and 'ON_DEMAND'.
- `core_ebs_optimized` - (Optional) EBS Optimization setting for instances in group.
- `core_ebs_block_device` - (Required) This determines the ebs configuration for your core group instances. Only a single block is allowed.
    - `volumes_per_instance` - (Optional; Default 1) Amount of volumes per instance in the core group.

- **volume_type** - (Required) volume type. Allowed values are 'gp2', 'io1' and others.
- **size_in_gb** - (Required) Size of the volume, in GBs.
- **iops** - (Optional) IOPS for the volume. Required in some volume types, such as io1.

» Master Group (Clone, New strategies)

- **master_instance_types** - (Required) The MrScaler instance types for the master nodes.
- **master_lifecycle** - (Required) The MrScaler lifecycle for instances in master group. Allowed values are 'SPOT' and 'ON_DEMAND'.
- **master_ebs_optimized** - (Optional) EBS Optimization setting for instances in group.
- **master_ebs_block_device** - (Required) This determines the ebs configuration for your master group instances. Only a single block is allowed.
  - **volumes_per_instance** - (Optional; Default 1) Amount of volumes per instance in the master group.
  - **volume_type** - (Required) volume type. Allowed values are 'gp2', 'io1' and others.
  - **size_in_gb** - (Required) Size of the volume, in GBs.
  - **iops** - (Optional) IOPS for the volume. Required in some volume types, such as io1.

» Tags (Clone, New strategies)

- **tags** - (Optional) A list of tags to assign to the resource. You may define multiple tags.
  - **key** - (Required) Tag key.
  - **value** - (Required) Tag value.

» Optional Compute Parameters (New strategy)

- **managed_primary_security_group** - (Optional) EMR Managed Security group that will be set to the primary instance group.
- **managed_replica_security_group** - (Optional) EMR Managed Security group that will be set to the replica instance group.
- **service_access_security_group** - (Optional) The identifier of the Amazon EC2 security group for the Amazon EMR service to access clusters in VPC private subnets.
- **additional_primary_security_groups** - (Optional) A list of additional Amazon EC2 security group IDs for the master node.

- **additional_replica_security_groups** - (Optional) A list of additional Amazon EC2 security group IDs for the core and task nodes.
- **custom_ami_id** - (Optional) The ID of a custom Amazon EBS-backed Linux AMI if the cluster uses a custom AMI.
- **repo_upgrade_on_boot** - (Optional) Applies only when **custom_ami_id** is used. Specifies the type of updates that are applied from the Amazon Linux AMI package repositories when an instance boots using the AMI. Possible values include: **SECURITY, NONE**.
- **ec2_key_name** - (Optional) The name of an Amazon EC2 key pair that can be used to ssh to the master node.
- **applications** - (Optional) A case-insensitive list of applications for Amazon EMR to install and configure when launching the cluster
    - **args** - (Optional) Arguments for EMR to pass to the application.
    - **name** - (Required) The application name.
    - **version**- (Optional)T he version of the application.
- **instance_weights** - (Optional) Describes the instance and weights. Check out Elastigroup Weighted Instances for more info.
    - **instance_type** - (Required) The type of the instance.
    - **weighted_capacity** - (Required) The weight given to the associated instance type.

» Availability Zones (Clone, New strategies)

- **availability_zones** - (Required in Clone) List of AZs and their subnet Ids. See example above for usage.

» Configurations (Clone, New strategies)

- **configurations_file** - (Optional) Describes path to S3 file containing description of configurations. More Information
    - **bucket** - (Required) S3 Bucket name for configurations.
    - **key**- (Required) S3 key for configurations.

» Steps (Clone, New strategies)

- **steps_file** - (Optional) Steps from S3.
    - **bucket** - (Required) S3 Bucket name for steps.
    - **key**- (Required) S3 key for steps.

» Bootstrap Actions (Clone, New strategies)

- `bootstrap_actions_file` - (Optional) Describes path to S3 file containing description of bootstrap actions. More Information
  - `bucket` - (Required) S3 Bucket name for bootstrap actions.
  - `key`- (Required) S3 key for bootstrap actions.

» Scaling Policies

Possible task group scaling policies (Wrap, Clone, and New strategies): * `task_scaling_up_policy` * `task_scaling_down_policy`

Possible core group scaling policies (Clone, New strategies): * `core_scaling_up_policy` * `core_scaling_down_policy`

Each `*_scaling_*_policy` supports the following:

- `policy_name` - (Required) The name of the policy.
- `metric_name` - (Required) The name of the metric, with or without spaces.
- `statistic` - (Required) The metric statistics to return. For information about specific statistics go to Statistics in the Amazon CloudWatch Developer Guide.
- `unit` - (Required) The unit for the metric.
- `threshold` - (Required) The value against which the specified statistic is compared.
- `adjustment` - (Optional) The number of instances to add/remove to/from the target capacity when scale is needed.
- `min_target_capacity` - (Optional) Min target capacity for scale up.
- `max_target_capacity` - (Optional) Max target capacity for scale down.
- `namespace` - (Required) The namespace for the metric.
- `operator` - (Required) The operator to use. Allowed values are : 'gt', 'gte', 'lt' , 'lte'.
- `evaluation_periods` - (Required) The number of periods over which data is compared to the specified threshold.
- `period` - (Required) The granularity, in seconds, of the returned datapoints. Period must be at least 60 seconds and must be a multiple of 60.
- `cooldown` - (Required) The amount of time, in seconds, after a scaling activity completes and before the next scaling activity can start.
- `dimensions` - (Optional) A mapping of dimensions describing qualities of the metric.
- `minimum` - (Optional) The minimum to set when scale is needed.
- `maximum` - (Optional) The maximum to set when scale is needed.
- `target` - (Optional) The number of instances to set when scale is needed.
- `action_type` - (Required) The type of action to perform. Allowed values are : 'adjustment', 'setMinTarget', 'setMaxTarget', 'updateCapacity',

'percentageAdjustment'

» Scheduled Tasks

- `scheduled_task` - (Optional) An array of scheduled tasks.
- `is_enabled` - (Optional) Enable/Disable the specified scheduling task.
- `task_type` - (Required) The type of task to be scheduled. Valid values: `setCapacity`.
- `instance_group_type` - (Required) Select the EMR instance groups to execute the scheduled task on. Valid values: `task`.
- `cron` - (Required) A cron expression representing the schedule for the task.
- `desired_capacity` - (Optional) New desired capacity for the elastigroup.
- `min_capacity` - (Optional) New min capacity for the elastigroup.
- `max_capacity` - (Optional) New max capacity for the elastigroup.

» Attributes Reference

The following attributes are exported:

- `id` - The scaler ID.

» spotinst_multai_balancer

Provides a Spotinst Multai Balancer.

» Example Usage

```
resource "spotinst_multai_balancer" "my_balancer" {
  name   = "foo"
  scheme = "internal"

  connection_timeouts = {
    idle     = 10
    draining = 10
  }

  tags = [{
    key   = "env"
    value = "prod"
  }]
}
```

» Argument Reference

The following arguments are supported:

- `name` - (Required) The balancer name. May contain only alphanumeric characters or hyphens, and must not begin or end with a hyphen.
- `scheme` - (Optional)
- `dns_cname_aliases` - (Optional)
- `connection_timeouts` - (Optional)
- `idle` - (Optional) The idle timeout value, in seconds. (range: 1 - 3600).
- `draining` - (Optional) The time for the load balancer to keep connections alive before reporting the target as de-registered, in seconds (range: 1 - 3600).
- `tags` - (Optional) A list of key:value paired tags.
- `key` - (Required) The tag's key.
- `value` - (Required) The tag's value.

» spotinst_multai_deployment

Provides a Spotinst Multai Deployment.

» Example Usage

```
resource "spotinst_multai_deployment" "my_deployment" {
  name = "foo"
}
```

» Argument Reference

The following arguments are supported:

- `name` - (Required) The deployment name.

» spotinst_multai_listener

Provides a Spotinst Multai Listener.

» Example Usage

```
resource "spotinst_multai_listener" "my_listener" {
  name        = "foo"
  balancer_id = "b-12345"
  protocol    = "http"
  port        = 1337

  tls_config = {
    certificate_ids            = ["ce-12345"]
    min_version                = "TLS10"
    max_version                = "TLS12"
    cipher_suites              = [""]
    prefer_server_cipher_suites = true
    session_tickets_disabled   = false
  }

  tags = [{
    key   = "env"
    value = "prod"
  }]
}
```

» Argument Reference

The following arguments are supported:

- `balancer_id` - (Required) The ID of the balancer.

- `protocol` - (Required) The protocol to allow connections to the load balancer.

- `port` - (Required) The port on which the load balancer is listening.

- `tls_config` - (Optional) Describes the TLSConfig configuration.

  - `min_version` - (Required) MinVersion contains the minimum SSL/TLS version that is acceptable (1.0 is the minimum).
  - `max_version` - (Required) MaxVersion contains the maximum SSL/TLS version that is acceptable.
  - `certificate_ids` - (Optional) Contains one or more certificate chains to present to the other side of the connection.
  - `cipher_suites` - (Optional) List of supported cipher suites. If cipherSuites is nil, TLS uses a list of suites supported by the implementation.
  - `prefer_server_cipher_suites` - (Optional) Controls whether the server selects the client's most preferred ciphersuite, or the server's

most preferred ciphersuite.

- **session_tickets_disabled** - (Optional) May be set to true to disable session ticket (resumption) support.

- **tags** - (Optional) A list of key:value paired tags.

  - **key** - (Required) The tag's key.
  - **value** - (Required) The tag's value.

» spotinst_multai_routing_rule

Provides a Spotinst Multai Routing Rule.

» Example Usage

```
resource "spotinst_multai_routing_rule" "my_routing_rule" {
  balancer_id = "b-12345"
  listener_id = "l-98765"
  route       = "Path(\x60/bar\x60)"
  strategy    = "LEASTCONN"

  tags = [{
    key   = "env"
    value = "prod"
  }]
}
```

» Argument Reference

The following arguments are supported:

- **balancer_id** - (Required) The ID of the balancer.

- **listener_id** - (Required) The ID of the listener.

- **route** - (Required) Route defines a simple language for matching HTTP requests and route the traffic accordingly. Route provides series of matchers that follow the syntax: Path matcher: — Path("/foo/bar") // trie-based PathRegexp("/foo/.") // regexp-based Method matcher: — Method("GET") // trie-based MethodRegexp("POST|PUT") // regexp based Header matcher: — Header("Content-Type", "application/json") // trie-based HeaderRegexp("Content-Type", "application/.") // regexp based Matchers can be combined using && operator: — Method("POST") && Path("/v1")

66

- `strategy` - (Optional) Balancing strategy. Valid values: `ROUNDROBIN`, `RANDOM`, `LEASTCONN`, `IPHASH`.

- `tags` - (Optional) A list of key:value paired tags.

- `key` - (Required) The tag's key.

- `value` - (Required) The tag's value.

» spotinst_multai_target

Provides a Spotinst Multai Target.

» Example Usage

```
resource "spotinst_multai_target" "my_target" {
  balancer_id   = "b-12345"
  target_set_id = "l-98765"

  name   = "foo"
  port   = 1338
  host   = "host"
  weight = 1

  tags = [{
    key   = "env"
    value = "prod"
  }]
}
```

» Argument Reference

The following arguments are supported:

- `balancer_id` - (Required) The ID of the balancer.

- `target_set_id` - (Required) The ID of the target set.

- `name` - (Required) The name of the Target . Must contain only alphanumeric characters or hyphens, and must not begin or end with a hyphen.

- `port` - (Required) The port the target will register to.

- `host` - (Required) The address (IP or URL) of the targets to register

- `weight` - (Required) Defines how traffic is distributed between targets.

- `tags` - (Optional) A list of key:value paired tags.
- `key` - (Required) The tag's key.
- `value` - (Required) The tag's value.

» spotinst_multai_target_set

Provides a Spotinst Multai Target Set.

» Example Usage

```
resource "spotinst_multai_target_set" "my_target_set" {
  balancer_id   = "b-12345"
  deployment_id = "dp-12345"
  name          = "foo"
  protocol      = "http"
  port          = 1338
  weight        = 2

  health_check = {
    protocol = "http"
    path     = "/"
    port     = 3001
    interval = 20
    timeout  = 5

    healthy_threshold   = 3
    unhealthy_threshold = 3
  }

  tags = [{
   key   = "env"
   value = "prod"
  }]
}
```

» Argument Reference

The following arguments are supported:

- `name` - (Required) The name of the Target Set. Must contain only alphanumeric characters or hyphens, and must not begin or end with a hyphen.

- `balancer_id` - (Required) The id of the balancer.

- `deployment_id` - (Required) The id of the deployment.

- `protocol` - (Required) The protocol to allow connections to the target.

- `port`

- `weight` - (Required) Defines how traffic is distributed between the Target Set.

- `health_check`

- `protocol` - (Required) The protocol to allow connections to the target for the health check.

- `path` - (Required) The path to perform the health check.

- `port` - (Required) The port on which the load balancer is listening.

- `interval` - (Required) The interval for the health check.

- `timeout` - (Required) The time out for the health check.

- `healthy_threshold` - (Required) Total number of allowed healthy Targets.

- `unhealthy_threshold` - (Required) Total number of allowed unhealthy Targets.

- `tags` - (Optional) A list of key:value paired tags.

- `key` - (Required) The tag's key.

- `value` - (Required) The tag's value.

» spotinst_ocean_aws

Provides a Spotinst Ocean AWS resource.

» Example Usage

```
resource "spotinst_ocean_aws" "example" {
  name = "demo"
  controller_id = "fakeClusterId"
  region = "us-west-2"

  max_size         = 2
  min_size         = 1
  desired_capacity = 2
```

```
subnet_ids = ["subnet-123456789"]
whitelist  = ["t1.micro", "m1.small"]

// --- LAUNCH CONFIGURATION --------------
image_id             = "ami-123456"
security_groups      = ["sg-987654321"]
key_name             = "fake key"
user_data            = "echo hello world"
iam_instance_profile = "iam-profile"
root_volume_size     = 20

associate_public_ip_address = true

load_balancers = [
  {
    arn = "arn:aws:elasticloadbalancing:us-west-2:fake-arn"
    type = "TARGET_GROUP"
  },
  {
    name = "AntonK"
    type = "CLASSIC"
  }
]
// -------------------------------------

// --- STRATEGY --------------------
fallback_to_ondemand       = true
spot_percentage            = 100
utilize_reserved_instances = false
// -------------------------------

// --- AUTOSCALER -----------------
autoscaler = {
  autoscale_is_enabled     = false
  autoscale_is_auto_config = false
  autoscale_cooldown       = 300

  autoscale_headroom = {
    cpu_per_unit    = 1024
    gpu_per_unit    = 1
    memory_per_unit = 512
    num_of_units    = 2
  }

  autoscale_down = {
    evaluation_periods = 300
```

```
    }

    resource_limits = {
      max_vcpu       = 1024
      max_memory_gib = 20
    }
  }
  // ------------------------------
}
```

» Argument Reference

The following arguments are supported:

- `name` - (Required) The cluster name.
- `controller_id` - (Required) The ocean cluster identifier. Example: `ocean.k8s`
- `region` - (Required) The region the cluster will run in.
- `max_size` - (Optional, Default: `1000`) The upper limit of instances the cluster can scale up to.
- `min_size` - (Optional) The lower limit of instances the cluster can scale down to.
- `desired_capacity` - (Optional) The number of instances to launch and maintain in the cluster.
- `subnet_ids` - (Required) A comma-separated list of subnet identifiers for the Ocean cluster. Subnet IDs should be configured with auto assign public ip.

```
name = "demo"
controller_id = "fakeClusterId"
region = "us-west-2"

max_size         = 2
min_size         = 1
desired_capacity = 2
```

- `whitelist` - (Optional) Instance types allowed in the Ocean cluster. Cannot be configured if `blacklist` is configured.
- `blacklist` - (Optional) Instance types not allowed in the Ocean cluster. Cannot be configured if `whitelist` is configured.

```
whitelist = ["t1.micro", "m1.small"]
// blacklist = ["t1.micro", "m1.small"]
```

- `user_data` - (Optional) Base64-encoded MIME user data to make available to the instances.

- `image_id` - (Required) ID of the image used to launch the instances.

- `security_groups` - (Required) One or more security group ids.

- `key_name` - (Optional) The key pair to attach the instances.

- `iam_instance_profile` - (Optional) The instance profile iam role.

- `associate_public_ip_address` - (Optional, Default: `false`) Configure public IP address allocation.

- `root_volume_size` - (Optional) The size (in Gb) to allocate for the root volume. Minimum 20.

- `load_balancers` - (Optional) - Array of load balancer objects to add to ocean cluster

  - `arn` - (Optional) Required if type is set to TARGET_GROUP
  - `name` - (Optional) Required if type is set to CLASSIC
  - `type` - (Required) Can be set to CLASSIC or TARGET_GROUP

```
image_id             = "ami-79826301"
security_groups      = ["sg-042d658b3ee907848"]
key_name             = "fake key"
user_data            = "echo hello world"
iam_instance_profile = "iam-profile"
root_volume_size     = 20

associate_public_ip_address = true

load_balancers = [
  {
    arn = "arn:aws:elasticloadbalancing:us-west-2:fake-arn"
    type = "TARGET_GROUP"
  },
  {
    name = "AntonK"
    type = "CLASSIC"
  }
]
```

- `fallback_to_ondemand` - (Optional, Default: `true`) If not Spot instance markets are available, enable Ocean to launch On-Demand instances instead.
- `spot_percentage` - (Optional, Default: `100`) The percentage of Spot instances the cluster should maintain. Min 0, max 100.
- `utilize_reserved_instances` - (Optional, Default `false`) If Reserved instances exist, OCean will utilize them before launching Spot instances.

```
fallback_to_ondemand       = true
spot_percentage            = 100
utilize_reserved_instances = false
```

- `autoscaler` - (Optional) Describes the Ocean Kubernetes autoscaler.
- `autoscale_is_enabled` - (Optional, Default: `true`) Enable the Ocean Kubernetes autoscaler.
- `autoscale_is_auto_config` - (Optional, Default: `true`) Automatically configure and optimize headroom resources.
- `autoscale_cooldown` - (Optional, Default: `null`) Cooldown period between scaling actions.
- `autoscale_headroom` - (Optional) Spare resource capacity management enabling fast assignment of Pods without waiting for new resources to launch.
- `cpu_per_unit` - (Optional) Optionally configure the number of CPUs to allocate the headroom. CPUs are denoted in millicores, where 1000 millicores = 1 vCPU.
- `gpu_per_unit` - (Optional) Optionally configure the number of GPUS to allocate the headroom.
- `memory_per_unit` - (Optional) Optionally configure the amount of memory (MB) to allocate the headroom.
- `num_of_units` - (Optional) The number of units to retain as headroom, where each unit has the defined headroom CPU and memory.
- `autoscale_down` - (Optional) Auto Scaling scale down operations.
- `evaluation_periods` - (Optional, Default: `null`) The number of evaluation periods that should accumulate before a scale down action takes place.
- `resource_limits` - (Optional) Optionally set upper and lower bounds on the resource usage of the cluster.
- `max_vcpu` - (Optional) The maximum cpu in vCPU units that can be allocated to the cluster.
- `max_memory_gib` - (Optional) The maximum memory in GiB units that can be allocated to the cluster.

```
autoscaler = {
  autoscale_is_enabled    = false
  autoscale_is_auto_config = false
  autoscale_cooldown      = 300

  autoscale_headroom = {
    cpu_per_unit    = 1024
    gpu_per_unit    = 1
    memory_per_unit = 512
    num_of_units    = 2
  }

  autoscale_down = {
    evaluation_periods = 300
  }
```

```
    resource_limits = {
      max_vcpu      = 1024
      max_memory_gib = 20
    }
  }
```

- `tags` - (Optional) Optionally adds tags to instances launched in an Ocean cluster.
- `key` - (Optional) The tag key.
- `value` - (Optional) The tag value.

```
tags = [{
  key   = "fakeKey"
  value = "fakeValue"
}]
```

» spotinst_ocean_aws_launch_spec

Provides a custom Spotinst Ocean AWS Launch Spec resource.

» Example Usage

```
resource "spotinst_ocean_aws_launch_spec" "example" {
  ocean_id  = "o-123456"
  image_id  = "ami-123456"
  user_data = "echo hello world"

  labels = [{
    key   = "fakeKey"
    value = "fakeValue"
  }]

  taints = [{
    key    = "taint key updated"
    value  = "taint value updated"
    effect = "NoExecute"
  }]

}
```

» Argument Reference

The following arguments are supported:

- `ocean_id` - (Required) The ocean cluster you wish to

- `user_data` - (Optional) Base64-encoded MIME user data to make available to the instances.

- `image_id` - (Optional) ID of the image used to launch the instances.

- `labels` - (Optional) Optionally adds labels to instances launched in an Ocean cluster.

  - `key` - (Required) The tag key.
  - `value` - (Required) The tag value.

- `taints` - (Optional) Optionally adds labels to instances launched in an Ocean cluster.

  - `key` - (Required) The tag key.
  - `value` - (Required) The tag value.
  - `effect` - (Required) The effect of the taint. Valid values: `"NoSchedule"`, `"PreferNoSchedule"`, `"NoExecute"`.

» spotinst_ocean_gke

Provides a Spotinst Ocean GKE resource.

» Example Usage

```
resource "spotinst_ocean_gke" "example" {
  name               = "example-ocean-cluster-name"
  controller_id      = "example-cluster-id"
  cluster_name       = "example-cluster-name"
  master_location    = "us-central1-a"
  subnet_name        = "example-subnet-1"
  availability_zones = ["us-central1-a"]
  whitelist          = ["n1-standard-1", "n1-standard-2"]

  max_size         = 1000
  min_size         = 0
  desired_capacity = 500

  // --- LAUNCH CONFIGURATION --------------
  source_image          = "https://www.googleapis.com/compute/v1/projects/my-project/global
  service_account       = "example-account@my-account.iam.gserviceaccount.com"
  root_volume_size_in_gb = 100
  ip_forwarding         = true
```

```
  labels = [{
    key   = "spotinst-gke-original-node-pool",
    value = "example-cluster-name__default-pool"
  }]

  metadata = [{
    key   = "cluster-name"
    value = "example-cluster"
  }]

  tags = ["gke-example-vpc-1234567-node"]

  backend_services = [{
    service_name  = "example-backend-service"
    location_type = "global"

    named_ports = {
      name  = "http"
      ports = [80, 8080]
    }
   }]
  // --------------------------------------

  // --- NETWORK INTERFACE -----------------
   network_interface = [{
     network = "example-vpc-network"

    access_configs = {
      name = "config1"
      type = "ONE_TO_ONE_NAT"
    }

    alias_ip_ranges = {
      subnetwork_range_name = "range-1"
      ip_cidr_range         = "10.8.0.0/20"
    }
   }]
  // --------------------------------------
}
```

» Argument Reference

The following arguments are supported:

- **name** - (Required) The cluster name.

- `controller_id` - (Required) The ocean cluster identifier. Example: `ocean.k8s`
- `cluster_name` - (Required) The GKE cluster name.
- `master_location` - (Required) The zone the master cluster is located in.
- `subnet_name` - (Required) Subnet identifier for the Ocean cluster.
- `availability_zones` - (Required) List of availability zones available to the cluster.
- `whitelist` - (Optional) Instance types allowed in the Ocean cluster.
- `max_size` - (Optional, Default: `1000`) The upper limit of instances the cluster can scale up to.
- `min_size` - (Optional) The lower limit of instances the cluster can scale down to.
- `desired_capacity` - (Optional) The number of instances to launch and maintain in the cluster.

Usage:

```
name               = "example-ocean-cluster-name"
controller_id      = "example-cluster-id"
cluster_name       = "example-cluster-name"
master_location    = "us-central1-a"
subnet_name        = "example-subnet-1"
availability_zones = ["us-central1-a"]
whitelist          = ["n1-standard-1", "n1-standard-2"]


max_size         = 1000
min_size         = 0
desired_capacity = 500
```

» Launch Configuration

Note: label, metadata, and tag keys are required, and depend on your GKE cluster. Please modify the values to match your configuration. You may also add additional key/value pairs. This resource is intended to be used as part of a Module.

- `source_image` - (Optional) A source image used to create the disk. You can provide a private (custom) image, and Compute Engine will use the corresponding image from your project.
- `service_account` - (Optional) The email of the service account in which the group instances will be launched.
- `root_volume_size_in_gb` - (Optional) The size (in Gb) to allocate for the root volume. Minimum `100`.
- `ip_forwarding` - (Optional) Enables the transfer IP packets from one network to another.
- `labels` - (Optional) Array of objects with key-value pairs.

77

- – `key` - (Optional) Labels key.
- – `value` - (Optional) Labels value.
- `metadata` - (Optional) Array of objects with key-value pairs.
  - – `key` - (Optional) Metadata key.
  - – `value` - (Optional) Metadata value.
- `tags` - (Optional) Tags to mark created instances. Minimum 1.

Usage:

```
source_image         = "https://www.googleapis.com/compute/v1/projects/my-project/global
service_account      = "example-account@my-account.iam.gserviceaccount.com"
root_volume_size_in_gb = 100
ip_forwarding        = true

labels = [{
  key   = "spotinst-gke-original-node-pool",
  value = "example-cluster-name__default-pool"
}]

metadata = [{
  key   = "cluster-name"
  value = "example-cluster"
}]

tags = ["gke-example-vpc-1234567-node"]
```

» Backend Services

- `backend_services` - (Optional) Describes the backend service configurations.
  - – `service_name` - (Required) The name of the backend service.
  - – `location_type` - (Optional) Sets which location the backend services will be active. Valid values: `regional`, `global`.
  - – `scheme` - (Optional) Use when `location_type` is "regional". Set the traffic for the backend service to either between the instances in the vpc or to traffic from the internet. Valid values: `INTERNAL`, `EXTERNAL`.
  - – `named_port` - (Optional) Describes a named port and a list of ports.
    - * `port_name` - (Required) The name of the port.
    - * `ports` - (Required) A list of ports.

Usage:

```
backend_services = [{
  service_name  = "example-backend-service"
  location_type = "global"
  scheme        = "INTERNAL"
  named_ports = {
```

```
    name  = "http"
    ports = [80, 8080]
  }
}]
```

» **Autoscaler**

- `autoscaler` - (Optional) Describes the Ocean Kubernetes autoscaler.
  - `autoscale_is_enabled` - (Optional, Default: `true`) Enable the Ocean Kubernetes autoscaler.
  - `autoscale_is_auto_config` - (Optional, Default: `true`) Automatically configure and optimize headroom resources.
  - `autoscale_cooldown` - (Optional, Default: `null`) Cooldown period between scaling actions.
  - `autoscale_headroom` - (Optional) Spare resource capacity management enabling fast assignment of Pods without waiting for new resources to launch.
    * `cpu_per_unit` - (Optional) Optionally configure the number of CPUs to allocate the headroom. CPUs are denoted in millicores, where 1000 millicores = 1 vCPU.
    * `gpu_per_unit` - (Optional) Optionally configure the number of GPUS to allocate the headroom.
    * `memory_per_unit` - (Optional) Optionally configure the amount of memory (MB) to allocate the headroom.
    * `num_of_units` - (Optional) The number of units to retain as headroom, where each unit has the defined headroom CPU and memory.
  - `autoscale_down` - (Optional) Auto Scaling scale down operations.
    * `evaluation_periods` - (Optional, Default: `null`) The number of evaluation periods that should accumulate before a scale down action takes place.
  - `resource_limits` - (Optional) Optionally set upper and lower bounds on the resource usage of the cluster.
    * `max_vcpu` - (Optional) The maximum cpu in vCPU units that can be allocated to the cluster.
    * `max_memory_gib` - (Optional) The maximum memory in GiB units that can be allocated to the cluster.

Usage:

```
autoscaler = {
  autoscale_is_enabled     = false
  autoscale_is_auto_config = false
  autoscale_cooldown       = 300

  autoscale_headroom = {
```

```
      cpu_per_unit    = 1024
      gpu_per_unit    = 1
      memory_per_unit = 512
      num_of_units    = 2
    }

    autoscale_down = {
      evaluation_periods = 300
    }

    resource_limits = {
      max_vcpu       = 1024
      max_memory_gib = 20
    }
  }
```

» spotinst_subscription

Provides a Spotinst subscription resource.

» Example Usage

```
# Create a Subscription
resource "spotinst_subscription" "default-subscription" {

  resource_id = "${spotinst_elastigroup_aws.my-eg.id}"
  event_type  = "AWS_EC2_INSTANCE_LAUNCH"
  protocol    = "http"
  endpoint    = "http://endpoint.com"

  format = {
    event         = "%event%"
    instance_id   = "%instance-id%"
    resource_id   = "%resource-id%"
    resource_name = "%resource-name%"
    tags          = "foo,baz,baz"
  }
}
```

» Argument Reference

The following arguments are supported:

- `resource_id` - (Required) Spotinst Resource ID (Elastigroup ID).
- `event_type` - (Required) The event to send the notification when triggered.
  Valid values: `"AWS_EC2_INSTANCE_TERMINATE"`, `"AWS_EC2_INSTANCE_TERMINATED"`,
  `"AWS_EC2_INSTANCE_LAUNCH"`, `"AWS_EC2_INSTANCE_UNHEALTHY_IN_ELB"`,
  `"GROUP_ROLL_FAILED"`, `"GROUP_ROLL_FINISHED"`, `"CANT_SCALE_UP_GROUP_MAX_CAPACITY"`,
  `"GROUP_UPDATED"`, `"AWS_EC2_CANT_SPIN_OD"`, `"AWS_EMR_PROVISION_TIMEOUT"`,
  `"AWS_EC2_INSTANCE_READY_SIGNAL_TIMEOUT"`.
- `protocol` - (Required) The protocol to send the notification. Valid values:
  `"http"`, `"https"`, `"email"`, `"email-json"`, `"aws-sns"`, `"web"`.
- `endpoint` - (Required) The endpoint the notification will be sent to: url in
  case of `"http"`/`"https"`, email address in case of `"email"`/`"email-json"`,
  sns-topic-arn in case of `"aws-sns"`.
- `format` - (Optional) The format of the notification content (JSON
  Format - Key+Value). Valid values: `"%instance-id%"`, `"%event%"`,
  `"%resource-id%"`, `"%resource-name%"`.

» Attributes Reference

The following attributes are exported:

- `id` - The subscription ID.