

Alert Name: OCI Reserved Instance Usage Below Threshold

Description: This alert is triggered when the percentage of used to reserved OCI instances drops below 98%. This typically indicates that some reserved compute capacity is not being utilized, potentially leading to unnecessary costs.

Trigger Condition: $\text{Used OCI Instances} / \text{Reserved OCI Instances} < 98\%$

Initial Triage When the alert is received:

1. Acknowledge the alert in your monitoring or alerting system.
2. Collect relevant context:
 - OCI Region
 - Compartment OCID
 - Instance shape(s) affected
 - Time of alert trigger
3. Check recent change history or deployment events that may have affected compute usage.

Preliminary Checks: Use the OCI CLI to gather current usage and reservation status:

1. Check reserved capacity: `bash CopyEdit oci compute capacity-reservation list \ --compartment-id <compartment_ocid> \ --region <region> \ --query "data[?lifecycle-state=='ACTIVE'].[display-name,reserved-instance-count,used-instance-count,availability-domain]" \ --output table`
2. List currently running compute instances: `bash CopyEdit oci compute instance list \ --compartment-id <compartment_ocid> \ --region <region> \ --query "data[?lifecycle-state=='RUNNING'].[display-name,shape,availability-domain]" \ --output table`
3. Check general capacity availability for compute: `bash CopyEdit oci limits resource-availability get \ --service-name compute \ --limit-name standard-e2-core-count \ --compartment-id <compartment_ocid> \ --region <region>`

Investigation Steps:

- Compare the number of used instances to the reserved capacity.
- Identify whether any instances were recently terminated or stopped.
- Check audit logs for instance terminations: `bash CopyEdit oci audit event list \ --compartment-id <compartment_ocid> \ --start-time <timestamp_1h_before_alert> \ --end-time <timestamp_of_alert>`
- Confirm if any scaling policies, manual terminations, or autoscaling events occurred.
- Determine whether the reservation is being underutilized temporarily (e.g., during maintenance) or is consistently underutilized.

Resolution Steps:

Option A: Increase Usage to Match Reserved Instances

If the reserved capacity is still required but underutilized:

- Identify workloads that can be deployed using the existing reservation.
- Launch new instances using the same shape and in the same availability domain as the reservation: `bash CopyEdit oci compute instance launch \ --compartment-id <compartment_ocid> \ --availability-domain <AD> \ --shape <reserved_shape> \ --subnet-id <subnet_ocid> \ --image-id <image_ocid> \ --display-name <new_instance_name>`

Option B: Adjust or Cancel Reservation

If the reserved capacity is no longer required:

- Update the reserved instance count: `bash CopyEdit oci compute capacity-reservation update \ --capacity-reservation-id <reservation_id> \ --instance-reservation-configurations ' [{"availabilityDomain": "<AD>", "instanceShape": "<shape>", "reservedCount": <new_count> }] '`
- Or delete the reservation if no longer needed: `bash CopyEdit oci compute capacity-reservation delete \ --capacity-reservation-id <reservation_id>`

Post-Resolution Steps:

- Monitor the alert for 24-48 hours to confirm it doesn't re-trigger.
- Document any actions taken in your change log or ticketing system.
- Notify relevant stakeholders (e.g., FinOps or infrastructure teams) if ongoing cost inefficiencies are identified.
- Consider automating or scheduling reviews of reservation usage monthly or quarterly.