# Azure VMSS: Upgrade Policies, Health Monitoring, and Repair Options

## 1. Upgrade Policies in VMSS

Azure VMSS supports multiple upgrade policies to control how instances are updated when the model changes or when a new image version is available.

|  |  |  |
| --- | --- | --- |
| Mode | Description | Use Case |
| Manual | No automatic upgrades. You must explicitly update instances using az vmss update-instances. | Controlled, staged upgrades |
| Automatic | Azure automatically upgrades instances when the model changes. | Seamless updates without manual effort |
| Rolling | Upgrade is applied in batches with health checks between batches. | Production environments where minimal disruption is required |
| Automatic OS Upgrade (AOUG) | Automatically upgrades OS images when a new image version is published. | Automatically apply security or version updates |

## 2. Health Monitoring in VMSS

Azure VMSS can monitor application health using custom probes or load balancer probes. Health monitoring helps detect and respond to unresponsive or degraded VM instances.

|  |  |
| --- | --- |
| Feature | Description |
| Application Health Probes | Use custom HTTP probes (e.g., /healthz) to check app health. |
| Load Balancer Health Probes | Detect unresponsive instances using Azure Load Balancer probes. |
| Instance Health Status | Track instance states like Healthy, Unhealthy, or Unknown. |

## 3. Repair and Auto-Heal Options in VMSS

Azure VMSS supports automatic repair of unhealthy VM instances through health probes or platform metrics. This helps maintain high availability by replacing or restarting faulty VMs.

|  |  |
| --- | --- |
| Feature | Description |
| Automatic Instance Repair | Automatically detects and repairs unhealthy VMs. |
| Custom Repair Actions | Use Azure Monitor or Logic Apps to define custom repair workflows. |

## 4. VMSS Upgrade Policy and Health Management Comparison

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Feature | Manual | Automatic | Rolling | Automatic OS Upgrade | Health Probes | Auto-Heal / Repair |
| Trigger | Manual CLI/API | Auto on model change | Manual or pipeline | New image version | HTTP/TCP probes | Platform/probe-based |
| Granular Control | ✅ | ❌ | ✅ | ❌ | ✅ | ✅ |
| Downtime Control | ✅ High | ❌ Low | ✅ Very High | ✅ High | ✅ | ✅ |
| Use Case | Staged updates | Dev/Test | Prod rollout | Security patches | App health check | Auto-fix faults |
| CLI Command | az vmss update-instances | N/A | az vmss rolling-upgrade start | --enableAutomaticOSUpgrade | Configured in LB | az vmss update --set automaticRepairsPolicy.enabled=true |
| Supports Custom Image | ✅ | ✅ | ✅ | ✅ | ✅ | ✅ |

## 5. In-Place Upgrade vs Max Surge in Rolling Upgrades

|  |  |  |
| --- | --- | --- |
| Aspect | In-Place Upgrade | Max Surge |
| Definition | Reuses existing VM instances and upgrades in-place. | Creates new VM instances during upgrade (additional capacity). |
| Capacity | No extra capacity required. | Requires extra capacity temporarily. |
| Downtime | Possible short unavailability per VM. | Near-zero downtime (new VM becomes ready before old is removed). |
| Speed | Slower, sequential upgrade. | Faster due to parallel new VM provisioning. |
| Cost | No extra cost. | May incur cost for temporary surge capacity. |
| Use Case | Resource-constrained environments. | Mission-critical systems requiring zero downtime. |