Your results

View how your workload aligns to best practices and recommendations to help you improve.

Guidance | Answer Summary

Shown below are the assessment's questions and how they were answered.

☑ Show all original responses available for each question.

Reliability

What reliability targets and metrics have you defined for your application?

☑ **Recovery targets to identify how long the workload can be unavailable (Recovery Time Objective) and how much data is acceptable to lose during a disaster (Recovery Point Objective).**

☑ **Availability targets such as Service Level Agreements (SLAs) and Service Level Objectives (SLOs).**

Availability metrics to measure and monitor availability such as Mean Time To Recover (MTTR) and Mean Time Between Failure (MTBF).

Composite SLA for the workload derived using the Azure SLAs for all relevant resources.

SLAs for all internal and external dependencies.

Independent availability and recovery targets for critical application subsystems and scenarios.

None of the above.

How have you ensured that your application architecture is resilient to failures?

Deployed the application across multiple regions.

Removed all single points of failure by running multiple instances of application components.

Deployed the application across Availability Zones within a region.

Performed Failure Mode Analysis (FMA) to identify fault-points and fault-modes.

Planned for component level faults to minimize application downtime.

Planned for dependency failures to minimize application downtime.

☑ **None of the above.**

How have you ensured required capacity and services are available in targeted regions?

Built a capacity model for the application

Planned for expected usage patterns.

☑ **Confirmed Azure service availability in required regions.**

Confirmed Availability Zones are available in required regions.

☑ **Validated required capacity is within Azure service scale limits and quotas.**

Validated all APIs/SDKs against target run-times and languages for required functionality.

Aligned with Azure roadmaps for required preview services and capabilities.

None of the above.

How are you handling disaster recovery for this workload?

Application is available across multiple regions in an active-active configuration.

Application is deployed across multiple regions in an active-passive configuration in alignment with recovery targets.

Traffic is routable to the application in the case of a regional failure.

- [x] **Defined a backup strategy in alignment with recovery targets.**

Defined a disaster recovery strategy to capture recovery steps for failover and failback.

Failover and failback steps and processes are automated.

Successfully tested and validated the failover and failback approach at least once.

Decomposed the application into distinct subsystems with independent disaster recovery strategies.

Network connectivity redundancy for on premise data/application sources.

None of the above.

What decisions have been taken to ensure the application platform meets your reliability requirements?

Application processes are stateless.

Session state is non-sticky and externalized to a data store.

Application configuration is treated as code and deployed with the application.

- [x] **Application platform services are running in a highly available configuration/SKU.**
- [x] **Application platform components are deployed across Availability Zones or Availability Sets.**

Leveraged platform services are Availability Zone aware.

Application platform components are deployed across multiple active regions.

- [x] **Load balancing is implemented to distribute traffic across multiple nodes.**
- [x] **Health probes are implemented to check the health of application components and compound application health.**

Queuing and reliable messaging patterns are used to integrate application tiers.

Client traffic can be routed to the application in the case of region/zone/network outages.

Procedures to scale out application platform components are automated.

None of the above.

What decisions have been taken to ensure the data platform meets your reliability requirements?

Data types are categorized by data consistency requirements.

Data platform services are running in a highly available configuration/SKU.

Data is replicated across multiple regions.

Data is replicated across Availability Zones.

- [x] **Data is backed-up on zone/geo-redundant storage.**

Active geo-replication is used for data platform components such as storage and databases.

Application traffic can be routed to data stores in the case of region/zone/network outages.

Read operations are segregated from update operations.

Load balancer health probes assess data platform components.

- [x] **Data restore processes have been defined to ensure consistent application state when data is corrupted or deleted.**

Data restore processes have been validated and tested to ensure consistent application state when data is corrupted or deleted.

None of the above.

How does your application logic handle exceptions and errors?

Have a method to handle faults that might take a variable amount of time to recover from.

Request timeouts are configured to manage inter-component calls.

- [x] **Retry logic is implemented to handle transient failures, with appropriate back-off strategies to avoid cascading failures.**
- [x] **The application is instrumented with semantic logs and metrics.**

None of the above.

What decisions have been taken to ensure networking and connectivity meets your reliability requirements?

All single points of failure have been eliminated from application communication flows.

- [x] **Health probes are configured for Azure Load Balancer(s) to assess application traffic flows and compound health.**

Azure Load Balancer Standard or Zone redundant application gateways are used to load balance traffic across Availability Zones.

Redundant connections from different locations are used for cross-premises connectivity (ExpressRoute or VPN).

A failure path has been simulated for cross-premises connectivity.

Zone redundant gateways are used for cross-premises connectivity (ExpressRoute or VPN).

Network traffic is monitored, and a response plan is in place to address network outages.

None of the above.

What reliability allowances for scalability and performance have you made?

- [x] **The application has dedicated cross-premises bandwidth.**

Components with sensitive latency requirements are collocated.

- [x] **Gateways (ExpressRoute or VPN) have been sized according to expected cross-premises network throughput.**

Expected throughput passing through security/network appliances has been tested and autoscaling is configured based on throughput requirements.

Autoscaling is enabled for application components and integrated with Azure Monitor.

Autoscaling has been tested and the time to scale in/out has been measured.

Tested and validated defined latency and defined throughput targets per scenario and component.

Calculated target data sizes and associated growth rates per scenario and component.

Operational procedures are defined in case data sizes exceed limits.

- [x] **Validated that long-running TCP connections are not required for the workload.**

Throttling is implemented to govern inbound application calls and inter-component calls.

None of the above.

What reliability allowances for security have you made?

The identity provider (AAD/ADFS/AD/Other) is highly available and aligns with application availability and recovery targets.

All external application endpoints are secured? i.e. Firewall, WAF, DDoS Protection Standard Plan, etc.

- [x] **Communication to Azure PaaS services secured using Virtual Network Service Endpoints or Private Link.**

Keys and secrets are backed-up to geo-redundant storage.

The process for key rotation is automated and tested

Emergency access break glass accounts have been tested and secured for recovering from Identity provider failure scenarios.

None of the above.

What reliability allowances for operations have you made?

Application can be automatically deployed to a new region without any manual operations to recover from disaster scenarios.

Application deployments can be rolled-back and rolled-forward through automated deployment pipelines.

The lifecycle of the application is decoupled from its dependencies.

The time it takes to deploy an entire production environment is tested and validated.

- [x] **None of the above.**

How do you test the application to ensure it is fault tolerant?

The application is tested against critical Non-Functional requirements for performance.

- [x] **Load Testing is conducted with expected peak volumes to test scalability and performance under load.**

Chaos Testing is performed by injecting faults.

Tests are automated and carried out periodically or on-demand.

- [x] **Critical test environments have 1:1 parity with the production environment.**

None of the above.

How do you monitor and measure application health?

- [x] **The application is instrumented with semantic logs and metrics.**
- [x] **Application logs are correlated across components.**

All components are monitored and correlated with application telemetry.

Key metrics, thresholds, and indicators are defined and captured.

A health model has been defined based on performance, availability, and recovery targets and is represented through monitoring dashboard and alerts.

Azure Service Health events are used to alert on applicable Service level events.

- [x] **Azure Resource Health events are used to alert on resource health events.**

Monitor long-running workflows for failures.

None of the above.

Security

Have you done a threat analysis of your workload?

Threat modeling processes are adopted, identified threats are ranked based on organizational impact, mapped to mitigations and communicated to stakeholders.

There's a process to track, triage and address security threats in the application development cycle.

Timelines and processess are established to deploy mitigations (security fixes) for identified threats.

Security requirements are defined for this workload.

Threat protection was addressed for this workload.

Security posture was evaluated with standard benchmarks (CIS Control Framework, MITRE framework etc.).

Business critical workloads, which may adversely affect operations if they are compromised or become unavailable, were identified and classified.

- [x] **None of the above.**

What considerations for compliance and governance did you make in this workload?

Regulatory and governance requirements of this workload are known and well understood.

Landing Zone concept is implemented for this workload using Azure Blueprints and/or Azure Policies.

Azure Policies are used to enforce and control security and organizational standards.

Root management group is used and any changes that are applied using this group are carefully considered.

Compliance for this workload is systematically monitored and maintained. Regular compliance attestations are performed.

External or internal audits of this workload are performed periodically.

Security plan for this workload was developed and is maintained.

Best practices and guidelines, based on industry recommendations, are reviewed and applied proactively.

Attacker vs. defender costs are considered when implementing defenses. Easy and cheap attack methods are always prevented.

Attacker access containment is considered when making investments into security solutions.

- [x] **None of the above.**

What practices and tools have you implemented as part of the development cycle?

A list of dependencies, frameworks and libraries used by this workload is maintained and updated regularly.

Framework and library updates are included into the workload lifecycle.

- [x] **Technologies and frameworks used in this workload are fully understood, including their vulnerabilities.**
- [x] **Security updates to VMs are applied in a timely manner, and strong passwords exist on those VMs for any local administrative accounts that may be in use.**
- [x] **All cloud services used by this workload are identified and it is understood how to configure them securely.**

Personally identifiable information (PII) is detected and removed/obfuscated automatically for this workload, including application logs.

Azure Tags are used to enrich Azure resources with operational metadata.

Elevated security capabilities such as dedicated Hardware Security Modules (HSMs) or the use of Confidential Computing was implemented or considered implementing?

None of the above.

Have you adopted a formal secure DevOps approach to building and maintaining software?

Formal DevOps approach to building and maintaining software in this workload was adopted.

DevOps security guidance based on industry lessons-learned, and available automation tools (OWASP guidance, Microsoft toolkit for Secure DevOps etc.) is leveraged.

Gates and approvals are configured in DevOps release process of this workload.

Security team is involved in planning, design and the rest of DevOps process of this workload.

Deployments are automated and it's possible to deploy N+1 and N-1 version (where N is the current production).

Code scanning tools are integrated as part of the continuous integration (CI) process for this workload and cover also 3rd party dependencies.

Credentials, certificates and other secrets are managed in a secure manner inside of CI/CD pipelines.

Branch policies are used in source control management, main branch is protected and code reviews are required.

Security controls are applied to all self-hosted build agents used by this workload (if any).

CI/CD roles and permissions are clearly defined for this workload.

☑ **None of the above.**

Is the workload developed and configured in a secure way?

Cloud services are used for well-established functions instead of building custom service implementations.

Detailed error messages and verbose information are hidden from the end user/client applications. Exceptions in code are handled gracefully and logged.

Platform specific information (e.g. web server version) is removed from server-client communication channels.

CDN (content delivery network) is used to separate the hosting platform and end-users/clients.

Application configuration is stored using a dedicated configuration management system (Azure App Configuration, Azure Key Vault etc.)

Access to data storage is identity-based, whenever possible.

Authentication tokens are cached securely and encrypted when sharing across web servers.

There are controls in place for this workload to detect and protect from data exfiltration.

☑ **None of the above.**

How are you monitoring security-related events in this workload?

Tools like Azure Security Center are used to discover and remediate common risks within Azure tenants.

A central SecOps team monitors security related telemetry data for this workload.

☑ **The security team has read-only access into all cloud environment resources for this workload.**

☑ **The security team has access to and monitor all subscriptions and tenants that are connected to the existing cloud environment, relative to this workload.**

Identity related risk events related to potentially compromised identities are actively monitored.

Communication, investigation and hunting activities are aligned with the workload team.

Periodic & automated access reviews of the workload are conducted to ensure that only authorized people have access?

Cloud application security broker (CASB) is leveraged in this workload.

A designated point of contact was assigned for this workload to receive Azure incident notifications from Microsoft.

None of the above.

How is security validated and how do you handle incident response when breach happens?

For containerized workloads, Azure Defender (Azure Security Center) or other third-party solution is used to scan for vulnerabilities.

Penetration testing is performed in-house or a third-party entity performs penetration testing of this workload to validate the current security defenses.

Simulated attacks on users of this workload, such as phishing campaigns, are carried out regularly.

Operational processes for incident response are defined and tested for this workload.

Playbooks are built to help incident responders quickly understand the workload and components, to mitigate an attack and do an investigation.

There's a security operations center (SOC) that leverages a modern security approach.

A security training program is developed and maintained to ensure security staff of this workload are well-informed and equipped with the appropriate skills.

- [x] **None of the above.**

How is connectivity secured for this workload?

- [x] **Services used by this workload, which should not be accessible from public IP addresses, are protected with network restrictions / IP firewall rules.**
- [x] **Service Endpoints or Private Links are used for accessing Azure PaaS services.**

Azure Firewall or any 3rd party next generation firewall is used for this workload to control outgoing traffic of Azure PaaS services (data exfiltration protection) where Private Link is not available.

- [x] **Network security groups (NSG) are used to isolate and protect traffic within the workloads VNet.**

NSG flow logs are configured to get insights about incoming and outgoing traffic of this workload.

- [x] **Access to the workload backend infrastructure (APIs, databases, etc.) is restricted to only a minimal set of public IP addresses - only those who really need it.**

Identified groups of resources are isolated from other parts of the organization to aid in detecting and containing adversary movement within the enterprise.

All public endpoints of this workload are protected/secured with appropriate solution (i.e. Azure Front Door, Azure Firewall...).

Publishing methods for this workload (e.g FTP, Web Deploy) are protected.

Code is published to this workload using CI/CD process instead of manually.

- [x] **Workload virtual machines running on premises or in the cloud don't have direct internet connectivity for users that may perform interactive logins, or by applications running on virtual machines.**

There's a capability and plans in place to mitigate DDoS attacks for this workload.

None of the above.

How have you secured the network of your workload?

- [x] **There's a designated group within the organization, which is responsible for centralized network management security of this workload.**
- [x] **There are controls in place to ensure that security extends past the network boundaries of the workload in order to effectively prevent, detect, and respond to threats.**

Enhanced network visibility is enabled by integrating network logs into a Security information and event management (SIEM) solution or similar technology.

Cloud virtual networks are designed for growth based on an intentional subnet security strategy.

This workload has a security containment strategy that blends existing on-premises security controls and practices with native security controls available in Azure, and uses a zero-trust approach.

Legacy network security controls for data loss prevention were deprecated.

- [x] **Traffic between subnets, Azure components and tiers of the workload is managed and protected.**

None of the above.

How are you managing encryption for this workload?

The workload uses industry standard encryption algorithms instead of creating own.

- [x] **The workload communicates over encrypted (TLS / HTTPS) network channels only.**

TLS 1.2 or 1.3 is used by default across this workload.

Secure modern hashing algorithms (SHA-2 family) are used.

Data at rest is protected with encryption.

- [x] **Data in transit is encrypted.**

Virtual disk files for virtual machines which are associated with this workload are encrypted.

None of the above.

Are keys, secrets and certificates managed in a secure way?

There's a clear guidance or requirement on what type of keys (PMK - Platform Managed Keys vs. CMK - Customer Managed Keys) should be used for this workload.

Passwords and secrets are managed outside of application artifacts, using tools like Azure Key Vault.

Access model for keys and secrets is defined for this workload.

A clear responsibility / role concept for managing keys and secrets is defined for this workload.

Secret/key rotation procedures are in place.

- [x] **Expiry dates of SSL/TLS certificates are monitored and there are renewal processes in place.**

None of the above.

What security controls do you have in place for access to Azure infrastructure?

- [x] **There are tools and processes in place to grant just-in-time access.**
- [x] **No user accounts have long-standing write access to production environments.**

Appropriate emergency access accounts are configured for this workload in case of an emergency.

Lines of responsibility and designated responsible parties were clearly defined for specific functions in Azure.

The application team has a clear view on responsibilities and individual/group access levels for this workload.

Workload infrastructure is protected with role-based access control (RBAC).

Resource locks are leveraged to protect critical infrastructure of this workload.

Direct access to the infrastructure through Azure Portal, command-line Interface (CLI) or REST API is limited and CI/CD is preferred.

Permissions to Azure workloads are rarely based on individual resources and custom permissions are rarely used.

There are processes and tools being used to manage privileged activities. Long standing administrative access is avoided whenever possible.

There is a lifecycle management policy for critical accounts in this workload and privileged accounts are reviewed regularly.

None of the above.

How are you managing identity for this workload?

When communicating with Azure platform services managed identities are preferred over API keys and connection strings.

- [x] **All APIs in this workload require clients to authenticate.**
- [x] **Modern authentication protocols (OAuth 2.0, OpenID) are used by this workload.**

- [x] **Azure Active Directory or other managed identity provider (Microsoft Account, Azure B2C etc.) is used for user authentication.**
- [x] **Authentication via identity services is prioritized for this workload vs. cryptographic keys.**
- [ ] Conditional access policies are implemented for users of this workload.
- [x] **Password-less or multi-factor authentication (MFA) is enforced for users of this workload.**
- [x] **Current on-premises Active Directory is synchronized with Azure AD or other cloud identity system.**
- [ ] None of the above.

Cost Optimization

How are you modeling cloud costs of this workload?

- [x] **Cloud costs are being modelled for this workload.**
- [ ] The price model of the workload is clear.
- [ ] Critical system flows through the application have been defined for all key business scenarios.
- [ ] There is a well-understood capacity model for the workload.
- [ ] Internal and external dependencies are identified and cost implications understood.
- [ ] Cost implications of each Azure service used by the application are understood.
- [ ] The right operational capabilities are used for Azure services.
- [ ] Special discounts given to services or licenses are factored in when calculating new cost models for services being moved to the cloud.
- [ ] Azure Hybrid Use Benefit is used to drive down cost in the cloud.
- [ ] None of the above.

How do you govern budgets and application lifespan for this workload?

- [ ] Budgets are assigned to all services in this workload.
- [ ] There is a cost owner for every service used by this workload.
- [x] **Cost forecasting is done to ensure it aligns with the budget.**
- [ ] There is a monthly or yearly meeting where the budget is reviewed.
- [ ] Every environment has a target end-date.
- [x] **Every environment has a plan for migrating to PaaS or serverless to lower the all up cost and transfer risk.**
- [ ] There is a clear understanding of how budget is defined.
- [ ] Budget is factored into the building phase.
- [x] **There is an ongoing conversation between the app owner and the business.**
- [x] **There is a plan to modernize the workload.**
- [ ] Azure Tags are used to enrich Azure resources with operational metadata.
- [ ] The application has a well-defined naming standard for Azure resources.
- [x] **Role Based Access Control (RBAC) is used to control access to operational and financial dashboards and underlying data.**
- [ ] None of the above.

How are you monitoring costs of this workload?

Alerts are set for cost thresholds and limits.

Specific owners and processes are defined for each alert type.

Application Performance Management (APM) tools and log aggregation technologies are used to collect logs and metrics from Azure resources.

Cost Management Tools (such as Azure Cost Management) are being used to track spending in this workload.

- [x] **None of the above.**

How do you optimize the design of this workload?

The application was built natively for the cloud.

There is an availability strategy defined and cost implications of it are understood.

This workload benefits from higher density.

Data is being transferred between regions.

Multi-region deployment is supported and cost implications understood.

- [x] **The workload is designed to use Availability Zones within a region.**

None of the above.

How do you ensure that cloud services are appropriately provisioned?

- [x] **Performance requirements are well-defined.**

Targets for the time it takes to perform scale operations are defined and monitored.

The workload is designed to scale independently.

- [x] **The application has been designed to scale both in and out.**

Application components and data are split into groups as part of your disaster recovery strategy.

Tools (such as Azure Advisor) are being used to optimise SKUs discovered in this workload.

Resources are reviewed weekly or bi-weekly for optimization.

Cost-effective regions are considered as part of the deployment selection.

- [x] **Dev/Test offerings are used correctly.**

Shared hosting platforms are used correctly.

None of the above.

What considerations for DevOps practices are you making in this workload?

There is an automated process to deploy application releases to production.

There is a difference in configuration for production and non-production environments.

Test-environments are deployed automatically and deleted after use.

- [x] **There is awareness around how the application has been built and is being maintained (in house or via an external partner).**

There is awareness regarding the ratio of cost of production and non-production environments for this workload.

None of the above.

How do you manage compute costs for this workload?

- [ ] Appropriate SKUs are used for workload servers.
- [ ] Appropriate operating systems are used in the workload.
- [ ] A recent review of SKUs that could benefit from Reserved Instances for 1 or 3 years or more has been performed.
- [ ] Burstable (B) series VM sizes are used for VMs that are idle most of the time and have high usage only in certain periods.
- [ ] VM instances which are not used are shut down.
- [ ] Spot virtual machines are used.
- [ ] PaaS is used as an alternative to buying virtual machines.
- [ ] Costs are optimized by using the App Service Premium (v3) plan over the Premium (Pv2) plan.
- [ ] Zone to Zone disaster recovery is used for virtual machines.
- [ ] The Start/Stop feature in Azure Kubernetes Services (AKS) is used.
- [x] **None of the above.**

How do you manage networking costs for this workload?

- [ ] Service Endpoints or Private Link are used for accessing Azure PaaS services.
- [ ] Hub and spoke design pricing is understood.
- [ ] Microsoft backbone network is preferred.
- [ ] DDoS attack mitigation plans and capabilities are in place.
- [ ] Azure Front Door, Azure App Gateway or Web Application Firewall is used.
- [ ] The workload is connected between regions (using network peering or gateways).
- [ ] Azure resources are connecting to the internet via on-premises.
- [ ] Public IPs and orphaned NICs are regularly cleaned up.
- [x] **None of the above.**

How do you manage storage and data costs for this workload?

- [ ] Reserved capacity is used for data in block blob storage.
- [ ] Data is organized into access tiers.
- [ ] Life-cycle policy is used to move data between access tiers.
- [ ] Shared disks are leveraged for suitable workloads.
- [ ] Reserved premium disks (P30 & above) are used.
- [ ] Bursting for P20 and below disks is utilized for suitable workloads.
- [x] **For database workloads, data and log files are stored on separate disks.**
- [ ] Unused storage resources (e.g. unattached disks, old snapshots) are periodically cleaned up.
- [x] **Selective disk backup and restore for Azure VMs is used.**
- [ ] None of the above.

Operational Excellence

Have you identified and planned out the roles and responsibilities to ensure your workload follows operational excellence best practices?

Development and operations processes are connected to a Service Management framework like ISO or ITIL

- [x] **There is no separation between development and operations teams.**

You have identified all broader teams responsible for operational aspects of the application and have established remediation plans with them for any issues that occur.

- [x] **Features and development tasks for the application are prioritized and executed on in a consistent fashion.**

You understand how the choices and desired configuration of Azure services are managed.

None of the above.

What design considerations for operations have you made?

- [x] **You have documented any components that are on-premises or in another cloud.**

Deployed the application across multiple regions.

Application is deployed across multiple regions in an active-passive configuration in alignment with recovery targets.

Application platform components are deployed across multiple active regions.

The workload is implemented with strategies for resiliency and self-healing.

- [x] **All platform-level dependencies are identified and understood.**

None of the above.

Have you defined key scenarios for your workload and how they relate to operational targets and non-functional requirements?

- [x] **Availability targets such as SLAs, SLIs and SLOs are defined for the application and key scenarios and monitored**
- [x] **Availability targets such as SLAs, SLIs and SLOs for all leveraged dependencies are understood and monitored**
- [x] **Recovery targets such as Recovery Time Objective (RTO) and Recovery Point Objective (RPO) are defined for the application and key scenarios**
- [x] **The consequences if availability and recovery targets are not satisfied are well understood**

There are targets defined for the time it takes to perform scale operations

Critical system flows through the application have been defined for all key business scenarios and have distinct availability, performance and recovery targets

There are well defined performance requirements for the application and key scenarios

- [x] **Any application components which are less critical and have lower availability or performance requirements are well understood**

None of the above.

How are you monitoring your resources?

An Application Performance Management (APM) tool like Azure Application Insights is used to collect application level logs

- [x] **Application logs are collected from different application environments**
- [x] **Log messages are captured in a structured format and can be indexed and searched**

Application events are correlated across all application components

- [ ] It is possible to evaluate critical application performance targets and non-functional requirements based on application logs and metrics
- [ ] End-to-end performance of critical system flows is monitored
- [ ] Black-box monitoring is used to measure platform services and the resulting customer experience.
- [ ] None of the above.

How do you interpret the collected data to inform about application health?

- [ ] A log aggregation technology, such as Azure Log Analytics or Splunk, is used to collect logs and metrics from Azure resources
- [ ] Azure Activity Logs are collected within the log aggregation tool
- [ ] Resource-level monitoring is enforced throughout the application
- [x] **Logs and metrics are available for critical internal dependencies**
- [x] **Log levels are used to capture different types of application events.**
- [ ] Critical external dependencies are monitored
- [ ] There are no known gaps in application observability that led to missed incidents and/or false positives.
- [ ] The workload is instrumented to measure customer experience.
- [ ] None of the above.

How do you visualize workload data and then alert relevant teams when issues occur?

- [ ] Application and resource level logs are either aggregated in a single data sink, or it is possible to cross-query events at both levels
- [ ] Application level events are automatically correlated with resource-level metrics to quantify the current application state
- [ ] A health model is used to qualify what 'healthy' and 'unhealthy' states represent for the workload
- [ ] Critical system flows are used to inform the health model
- [ ] The health model can distinguish between transient and non-transient faults
- [ ] Long-term trends are analysed to predict operational issues before they occur
- [x] **Retention times for logs and metrics have been defined and with housekeeping mechanisms configured**
- [ ] None of the above.

How are you using Azure platform notifications and updates?

- [ ] A tool such as Azure Monitor or Grafana is used to visualize the application health model and encompassed logs and metrics
- [ ] Dashboards are tailored to a specific audience such as developers, security or networking teams
- [ ] A tool such as Azure Monitor or Splunk is used for alerting
- [ ] Specific owners and processes are defined for each alert type
- [ ] Operational events are prioritized based on business impact
- [ ] Push notifications are used to inform responsible parties of alerts in real time
- [ ] Alerting is integrated with an IT Service Management (ITSM) system such as ServiceNow
- [x] **Azure Service Health alerts been created to respond to Service-level events.**
- [x] **Azure Resource Health alerts been created to respond to Resource-level events.**
- [ ] None of the above.

What is your approach to recovery and failover?

- [ ] Recovery steps are defined and well understood for failover and failback
- [ ] The failover and failback approach has been tested/validated at least once
- [ ] The health model is being used to classify failover situations
- [ ] Automated recovery procedures are in place for common failure events
- [ ] Automated recovery procedures are tested and validated on a regular basis
- [x] **Critical manual processes are defined and documented for failure responses.**
- [ ] Manual operational runbooks are tested and validated on a regular basis
- [ ] None of the above.

How are scale operations performed?

- [x] **There is a capacity model for the workload**
- [ ] Auto-scaling is enabled for supporting PaaS and IaaS services
- [ ] The process to provision and de-provision capacity is codified
- [ ] The impact of changes in application health on capacity is fully understood
- [x] **It has been validated that the required capacity (initial and future growth) is within Azure service scale limits and quotas**
- [ ] It has been validated that the required capacity (initial and future growth) is available within targeted regions
- [ ] Capacity utilization is monitored and used to forecast future growth
- [ ] None of the above.

How are you managing the configuration of your workload?

- [ ] You monitor and take advantage of new features and capabilities of underlying services used in your workload.
- [ ] Application configuration information is stored using a dedicated management system such as Azure App Configuration or Azure Key Vault
- [ ] Soft-Delete is enabled for your keys and credentials such as things stored in Key Vaults and Key Vault objects.
- [ ] Configuration settings can be changed or modified without rebuilding or redeploying the application
- [ ] Passwords and other secrets are managed in a secure store like Azure Key Vault or HashiCorp Vault
- [ ] Procedures are in place for key/secret rotation
- [ ] The application uses Azure Managed Identities
- [x] **The expiry dates of SSL certificates are monitored and there are processes in place to renew them**
- [ ] Components are hosted on shared application or data platforms as appropriate.
- [ ] Your workload takes advantage of multiple Azure subscriptions.
- [ ] The workload is designed to leverage managed services.
- [ ] None of the above.

What operational considerations are you making regarding the deployment of your workload?

- [ ] There is a systematic approach to the development and release process.
- [ ] The application can be deployed automatically from scratch without any manual operations
- [x] **There is a documented process for any portions of the deployment that require manual intervention**

- [ ] N-1 or N+1 versions can be deployed via automated pipelines where N is current deployment version in production
- [ ] There is a defined hotfix process which bypasses normal deployment procedures
- [ ] The application deployment process leverages blue-green deployments and/or canary releases
- [ ] Releases to production are gated by having it successfully deployed and tested in other environments
- [ ] Feature flags are used to test features before rolling them out to everyone
- [ ] None of the above.

What operational considerations are you making regarding the deployment of your infrastructure?

- [ ] The entire application infrastructure is defined as code
- [ ] No operational changes are performed outside of infrastructure as code
- [ ] Configuration drift is tracked and addressed
- [ ] The process to deploy infrastructure is automated
- [x] **Critical test environments have 1:1 parity with the production environment**
- [ ] Direct write access to infrastructure is not possible and all resources are provisioned or configured through IaC processes.
- [ ] None of the above.

How are you managing and distributing your patches

- [x] **You have a defined process to patch and update all relevant workload components.**
- [ ] You have a defined rollback strategy for patches.
- [ ] There is an playbook to deploy emergency patches as needed.
- [ ] None of the above.

How are you testing and validating your workload?

- [x] **The application is tested for performance, scalability, and resiliency**
- [ ] Tests for performance, scalability, and resiliency are performed as part of each major change
- [ ] At least a subset of tests is also performed in the production environment
- [ ] Fault injection tests are being utilized
- [ ] Smoke tests are performed during application deployments
- [ ] Unit and integration testing is performed as part of the application deployment process
- [ ] All these tests are automated and carried out periodically
- [ ] Failing tests at least temporarily block a deployment and lead to a deeper analysis of what has happened
- [ ] Business Continuity 'fire drills' are performed to test regional failover scenarios
- [ ] Security and penetration testing is performed regularly
- [ ] You regularly validate and update your tests to reflect any necessary changes.
- [ ] Operational procedures are reviewed and refined regularly.
- [ ] Mocks and stubs are used to test external dependencies in non-production environments.
- [ ] None of the above.

What processes and procedures have you adopted to optimize workload operability?

Specific methodologies, like DevOps, are used to structure the development and operations process

Collaboration between development and operations team to resolve production issue is clearly defined and well understood

Operational shortcomings and failures are analyzed and used to improve and refine operational procedures

There are tools or processes in place, such as Azure AD Privileged Identity Management, to grant access to critical systems on a just in-time basis

No users have long-standing write-access to production environments

Azure Resource Tags are used to enrich resources with operational meta-data

There are tools and processes, like Azure Policy, in place to govern available services, enforce mandatory operational functionality and ensure compliance

Standards, policies, restrictions and best practices are defined as code, for example by using solutions like Azure Policy or HashiCorp Sentinel

- [x] **None of the above.**

What operational excellence allowances for reliability have you made?

Error budgets used to track service reliability.

There is a policy that governs what happens when the error budget is exhausted.

- [x] **Availability targets such as Service Level Agreements (SLAs) and Service Level Objectives (SLOs) have been set.**

The life-cycle of the application is decoupled from its dependencies.

Application logs are correlated across components.

- [x] **The application is instrumented with semantic logs and metrics.**
- [x] **Validated required capacity is within Azure service scale limits and quotas.**

Calculated target data sizes and associated growth rates per scenario and component.

Operational procedures are defined in case data sizes exceed limits.

Tested and validated defined latency and throughput targets per scenario and component.

Autoscaling is enabled for application components and integrated with Azure Monitor.

Defined a disaster recovery strategy to capture recovery steps for failover and failback.

Keys and secrets are backed-up to geo-redundant storage.

Application can be automatically deployed to a new region without any manual operations to recover from disaster scenarios.

None of the above.

What operational excellence allowances for cost have you made?

The application was built natively for the cloud.

- [x] **The workload is designed to use Availability Zones within a region.**
- [x] **The application has been designed to scale both in and out.**

Performance requirements are well-defined.

Critical system flows through the application have been defined for all key business scenarios.

Application Performance Management (APM) tools and log aggregation technologies are used to collect logs and metrics from Azure resources.

- [x] **Role Based Access Control (RBAC) is used to control access to operational and financial dashboards and underlying data.**

Specific owners and processes are defined for each alert type.

There is an automated process to deploy application releases to production.

- [x] **There is awareness around how the application has been built and is being maintained (in house or via an external partner).**

The application has a well-defined naming standard for Azure resources.

Targets for the time it takes to perform scale operations are defined and monitored.

All internal and external dependencies identified and categorized as either weak or strong.

None of the above.

What operational excellence allowances for security have you made?

- [x] **Regulatory and governance requirements of this workload are known and well understood.**

There are tools and processes in place to grant just-in-time access.

Appropriate emergency access accounts are configured for this workload.

None of the above.

Performance Efficiency

What design considerations have you made for performance efficiency in your workload?

The workload is deployed across multiple regions.

- [x] **Regions were chosen based on location, proximity to users, and resource type availability.**

Paired regions are used appropriately.

You have ensured that both (all) regions in use have the same performance and scale SKUs that are currently leveraged in the primary region.

- [x] **Within a region the application architecture is designed to use Availability Zones.**

The application is implemented with strategies for resiliency and self-healing.

- [x] **Component proximity is considered for application performance reasons.**

The application can operate with reduced functionality or degraded performance in the case of an outage.

- [x] **You choose appropriate datastores for the workload during the application design.**

Your application is using a micro-service architecture.

- [x] **You understand where state will be stored for the workload.**

None of the above.

Have you identified the performance targets and non-functional requirements for your workload?

You are able to predict general application usage.

There are well-defined performance requirements for the workload and its key scenarios.

Targets for scale operations are defined.

You understand and have documented the expected maximum traffic volume before performance degradation occurs.

- [x] **None of the above.**

How are you ensuring that your workload is elastic and responsive to changes?

- [x] **The workload can scale horizontally in response to changing load.**

Have policies to scale in and scale down when the load decreases.

Configured scaling policies to use the appropriate metrics.

Automatically schedule autoscaling to add resources based on time of day trends.

Autoscaling has been tested under sustained load.

You have measured the time it takes to scale in and out.

None of the above.

How have you accounted for capacity and scaling requirements of your workload?

You have a capacity model for the workload.

Capacity utilization is monitored and used to forecast future growth.

A process for provisioning and de-provisioning capacity has been established.

You have enabled auto-scaling for all PaaS and IaaS services that support it.

- [x] **You are aware of relevant Azure service limits and quotas.**
- [x] **You have validated the SKU and configuration choices are appropriate for your anticipated loads.**

There is a strategy in place to manage events that may cause a spike in load.

None of the above.

What considerations for performance efficiency have you made in your networking stack?

You are using a Content Delivery Network.

You are offloading SSL.

You are using authentication/token verification offloading.

You have defined, tested, and validated latency targets for key scenarios.

You have defined, tested, and validated throughput targets for key scenarios.

- [x] **You have identified all components that are sensitive to network latency.**
- [x] **Dedicated bandwidth has been acquired where needed.**

None of the above.

How are you managing your data to handle scale?

- [x] **You know the growth rate of your data.**

You have documented plans for data growth and retention.

Design for eventual consistency.

You are using database replicas and data partitioning (sharding) as appropriate.

Minimize the load on the data store.

- [x] **Normalize the data appropriately.**
- [x] **Optimize database queries and indexes.**

None of the above.

How are you testing to ensure that you workload can appropriately handle user load?

- [ ] There is a defined testing strategy.
- [ ] Performance tests are performed regularly.
- [x] **You have identified the human and environmental resources needed to create performance tests.**
- [ ] You are using appropriate tools to conduct performance tests on your workload.
- [ ] You are testing all appropriate components for performance.
- [x] **You have identified all services being utilized in Azure (and on-premise) that need to be measured.**
- [ ] Some tests are performed in production.
- [ ] The testing plan includes occasionally injecting faults.
- [ ] None of the above.

How are you benchmarking your workload?

- [x] **You have identified goals or a baseline for workload performance.**
- [ ] Performance goals are based on device and/or connectivity type as appropriate.
- [x] **You have defined an initial connection goal for your workload.**
- [ ] There is a goal defined for complete page load times.
- [ ] You have defined goals for an API (service) endpoint complete response.
- [ ] There are goals defined for server response time.
- [ ] You have goals for latency between the systems & microservices of your workload.
- [x] **There are goals on database query efficiency.**
- [x] **You have a methodology to determine what acceptable performance is.**
- [ ] None of the above.

How have you modeled the health of your workload?

- [ ] Application and resource level logs are aggregated in a single data sink or able to be cross-queried.
- [ ] A health model is used to qualify what 'healthy' and 'unhealthy' states represent for the application.
- [ ] Critical system flows are used to inform the health model.
- [ ] The health model can distinguish between transient and non-transient faults.
- [ ] The health model can determine if the workload is performing at the expected targets.
- [x] **Retention times for logs and metrics been defined and housekeeping mechanisms are configured.**
- [ ] Long-term trends are analyzed to predict performance issues before they occur.
- [ ] None of the above.

How are you monitoring to ensure the workload is scaling appropriately?

- [ ] Track when resources scale in and out.
- [ ] Have an overall monitoring strategy for scalability and performance.
- [x] **Application logs are collected from different application environments.**

- [x] **Logs are captured in a structured format.**

- [ ] You monitor how much of an application is involved in serving a single request.
- [ ] Application events are correlated across all application components.
- [ ] You have determined an acceptable operational margin between your peak utilization and the application's maximum load, and monitor for this.
- [ ] You are aware of the appropriate metrics to monitor for performance tests under standard load.
- [ ] You monitor critical external dependencies for performance.
- [ ] None of the above.

What common problems do you have steps to troubleshoot in your operations playbook?

- [x] **You have steps to troubleshoot database issues.**

- [ ] You know how to handle high CPU or memory situations.
- [ ] You know what to do when the application response times increase while not using all the CPU or memory allocated to the system.

- [x] **You use profiling tools to profile your application code.**

- [ ] You have a response plan for network performance problems that includes traffic capturing tools.
- [ ] None of the above.
