Contents

- Billing

Resources

- Azure limits and quotas
- Azure regions
- Azure Roadmap
- Pricing calculator
- Samples
- Videos

Get started guide for Azure developers

2/25/2020 • 20 minutes to read • Edit Online

What is Azure?

Azure is a complete cloud platform that can host your existing applications and streamline new application development. Azure can even enhance on-premises applications. Azure integrates the cloud services that you need to develop, test, deploy, and manage your applications, all while taking advantage of the efficiencies of cloud computing.

By hosting your applications in Azure, you can start small and easily scale your application as your customer demand grows. Azure also offers the reliability that's needed for high-availability applications, even including failover between different regions. The Azure portal lets you easily manage all your Azure services. You can also manage your services programmatically by using service-specific APIs and templates.

This guide is an introduction to the Azure platform for application developers. It provides guidance and direction that you need to start building new applications in Azure or migrating existing applications to Azure.

Where do I start?

With all the services that Azure offers, it can be an intimidating task to figure out which services you need to support your solution architecture. This section highlights the Azure services that developers commonly use. For a list of all Azure services, see the Azure documentation.

First, you must decide on how to host your application in Azure. Do you need to manage your entire infrastructure as a virtual machine (VM). Can you use the platform management facilities that Azure provides? Maybe you need a serverless framework to host code execution only?

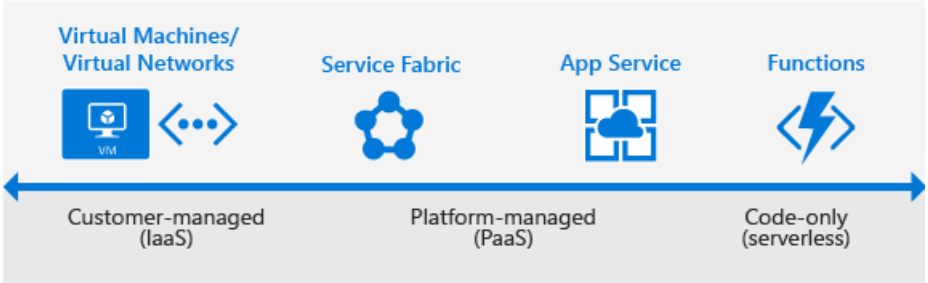
Your application needs cloud storage, which Azure provides several options for. You can take advantage of Azure's enterprise authentication. There are also tools for cloud-based development and monitoring, and most hosting services offer DevOps integration.

Now, let's look at some of the specific services that we recommend investigating for your applications.

Application hosting

Azure provides several cloud-based compute offerings to run your application so that you don't have to worry about the infrastructure details. You can easily scale up or scale out your resources as your application usage grows.

Azure offers services that support your application development and hosting needs. Azure provides Infrastructure as a Service (IaaS) to give you full control over your application hosting. Azure's Platform as a Service (PaaS) offerings provide the fully managed services needed to power your apps. There's even true serverless hosting in Azure where all you need to do is write your code.



Azure App Service

When you want the quickest path to publish your web-based projects, consider Azure App Service. App Service makes it easy to extend your web apps to support your mobile clients and publish easily consumed REST APIs. This platform provides authentication by using social providers, traffic-based autoscaling, testing in production, and continuous and container-based deployments.

You can create web apps, mobile app back ends, and API apps.

Because all three app types share the App Service runtime, you can host a website, support mobile clients, and expose your APIs in Azure, all from the same project or solution. To learn more about App Service, see What is Azure Web Apps.

App Service has been designed with DevOps in mind. It supports various tools for publishing and continuous integration deployments. These tools include GitHub webhooks, Jenkins, Azure DevOps, TeamCity, and others.

You can migrate your existing applications to App Service by using the online migration tool.

> **When to use**: Use App Service when you're migrating existing web applications to Azure, and when you need a fully-managed hosting platform for your web apps. You can also use App Service when you need to support mobile clients or expose REST APIs with your app.
>
> **Get started**: App Service makes it easy to create and deploy your first web app, mobile app, or API app.
>
> **Try it now**: App Service lets you provision a short-lived app to try the platform without having to sign up for an Azure account. Try the platform and create your Azure App Service app.

Azure Virtual Machines

As an Infrastructure as a Service (IaaS) provider, Azure lets you deploy to or migrate your application to either Windows or Linux VMs. Together with Azure Virtual Network, Azure Virtual Machines supports the deployment of Windows or Linux VMs to Azure. With VMs, you have total control over the configuration of the machine. When using VMs, you're responsible for all server software installation, configuration, maintenance, and operating system patches.

Because of the level of control that you have with VMs, you can run a wide range of server workloads on Azure that don't fit into a PaaS model. These workloads include database servers, Windows Server Active Directory, and Microsoft SharePoint. For more information, see the Virtual Machines documentation for either Linux or Windows.

> **When to use**: Use Virtual Machines when you want full control over your application infrastructure or to migrate on-premises application workloads to Azure without having to make changes.
>
> **Get started**: Create a Linux VM or Windows VM from the Azure portal.

Azure Functions (serverless)

Rather than worrying about building out and managing a whole application or the infrastructure to run your code, what if you could just write your code and have it run in response to events or on a schedule? Azure Functions is a "serverless"-style offering that lets you write just the code you need. With Functions, you can trigger code execution with HTTP requests, webhooks, cloud service events, or on a schedule. You can code in your development language of choice, such as C#, F#, Node.js, Python, or PHP. With consumption-based billing, you pay only for the time that your code executes, and Azure scales as needed.

> **When to use**: Use Azure Functions when you have code that is triggered by other Azure services, by web-based events, or on a schedule. You can also use Functions when you don't need the overhead of a complete hosted project or when you only want to pay for the time that your code runs. To learn more, see Azure Functions Overview.
>
> **Get started**: Follow the Functions quickstart tutorial to create your first function from the portal.

> **Try it now**: Azure Functions lets you run your code without having to sign up for an Azure account. Try it now at and create your first Azure Function.

Azure Service Fabric

Azure Service Fabric is a distributed systems platform. This platform makes it easy to build, package, deploy, and manage scalable and reliable microservices. It also provides comprehensive application management capabilities such as:

- Provisioning
- Deploying
- Monitoring
- Upgrading/patching
- Deleting

Apps, which run on a shared pool of machines, can start small and scale to hundreds or thousands of machines as needed.

Service Fabric supports WebAPI with Open Web Interface for .NET (OWIN) and ASP.NET Core. It provides SDKs for building services on Linux in both .NET Core and Java. To learn more about Service Fabric, see the Service Fabric documentation.

> **When to use:** Service Fabric is a good choice when you're creating an application or rewriting an existing application to use a microservice architecture. Use Service Fabric when you need more control over, or direct access to, the underlying infrastructure.
>
> **Get started:** Create your first Azure Service Fabric application.

Enhance your applications with Azure services

Along with application hosting, Azure provides service offerings that can enhance the functionality. Azure can also improve the development and maintenance of your applications, both in the cloud and on-premises.

Hosted storage and data access

Most applications must store data, so however you decide to host your application in Azure, consider one or more of the following storage and data services.

- **Azure Cosmos DB**: A globally distributed, multi-model database service. This database enables you to elastically scale throughput and storage across any number of geographical regions with a comprehensive SLA.

  > **When to use:** When your application needs document, table, or graph databases, including MongoDB databases, with multiple well-defined consistency models.
  >
  > **Get started**: Build an Azure Cosmos DB web app. If you're a MongoDB developer, see Build a MongoDB web app with Azure Cosmos DB.

- **Azure Storage**: Offers durable, highly available storage for blobs, queues, files, and other kinds of nonrelational data. Storage provides the storage foundation for VMs.

  > **When to use**: When your app stores nonrelational data, such as key-value pairs (tables), blobs, files shares, or messages (queues).
  >
  > **Get started**: Choose from one of these types storage: blobs, tables, queues, or files.

- **Azure SQL Database**: An Azure-based version of the Microsoft SQL Server engine for storing relational tabular data in the cloud. SQL Database provides predictable performance, scalability with no downtime,

business continuity, and data protection.

> **When to use**: When your application requires data storage with referential integrity, transactional support, and support for TSQL queries.
>
> **Get started**: Create a SQL database in minutes by using the Azure portal.

You can use Azure Data Factory to move existing on-premises data to Azure. If you aren't ready to move data to the cloud, Hybrid Connections in Azure App Service lets you connect your App Service hosted app to on-premises resources. You can also connect to Azure data and storage services from your on-premises applications.

Docker support

Docker containers, a form of OS virtualization, let you deploy applications in a more efficient and predictable way. A containerized application works in production the same way as on your development and test systems. You can manage containers by using standard Docker tools. You can use your existing skills and popular open-source tools to deploy and manage container-based applications on Azure.

Azure provides several ways to use containers in your applications.

- **Azure Docker VM extension**: Lets you configure your VM with Docker tools to act as a Docker host.

  > **When to use**: When you want to generate consistent container deployments for your applications on a VM, or when you want to use Docker Compose.
  >
  > **Get started**: Create a Docker environment in Azure by using the Docker VM extension.

- **Azure Kubernetes Service**: Lets you create, configure, and manage a cluster of virtual machines that are preconfigured to run containerized applications. To learn more about Azure Kubernetes Service, see Azure Kubernetes Service introduction.

  > **When to use**: When you need to build production-ready, scalable environments that provide additional scheduling and management tools, or when you're deploying a Docker Swarm cluster.
  >
  > **Get started**: Deploy a Kubernetes Service cluster.

- **Docker Machine**: Lets you install and manage a Docker Engine on virtual hosts by using docker-machine commands.

  > **When to use**: When you need to quickly prototype an app by creating a single Docker host.

- **Custom Docker image for App Service**: Lets you use Docker containers from a container registry or a customer container when you deploy a web app on Linux.

  > **When to use**: When deploying a web app on Linux to a Docker image.
  >
  > **Get started**: Use a custom Docker image for App Service on Linux.

Authentication

It's crucial to not only know who is using your applications, but also to prevent unauthorized access to your resources. Azure provides several ways to authenticate your app clients.

- **Azure Active Directory (Azure AD)**: The Microsoft multitenant, cloud-based identity and access management service. You can add single-sign on (SSO) to your applications by integrating with Azure AD. You can access directory properties by using the Azure AD Graph API directly or the Microsoft Graph API. You can integrate with Azure AD support for the OAuth2.0 authorization framework and Open ID Connect

by using native HTTP/REST endpoints and the multiplatform Azure AD authentication libraries.

> **When to use**: When you want to provide an SSO experience, work with Graph-based data, or authenticate domain-based users.
>
> **Get started**: To learn more, see the Azure Active Directory developer's guide.

- **App Service Authentication**: When you choose App Service to host your app, you also get built-in authentication support for Azure AD, along with social identity providers—including Facebook, Google, Microsoft, and Twitter.

> **When to use**: When you want to enable authentication in an App Service app by using Azure AD, social identity providers, or both.
>
> **Get started**: To learn more about authentication in App Service, see Authentication and authorization in Azure App Service.

To learn more about security best practices in Azure, see Azure security best practices and patterns.

Monitoring

With your application up and running in Azure, you need to monitor performance, watch for issues, and see how customers are using your app. Azure provides several monitoring options.

- **Application Insights**: An Azure-hosted extensible analytics service that integrates with Visual Studio to monitor your live web applications. It gives you the data that you need to improve the performance and usability of your apps continuously. This improvement occurs whether you host your applications on Azure or not.

> **Get started**: Follow the Application Insights tutorial.

- **Azure Monitor**: A service that helps you to visualize, query, route, archive, and act on the metrics and logs that you generate with your Azure infrastructure and resources. Monitor is a single source for monitoring Azure resources and provides the data views that you see in the Azure portal.

> **Get started**: Get started with Azure Monitor.

DevOps integration

Whether it's provisioning VMs or publishing your web apps with continuous integration, Azure integrates with most of the popular DevOps tools. You can work with the tools that you already have and maximize your existing experience with support for tools like:

- Jenkins
- GitHub
- Puppet
- Chef
- TeamCity
- Ansible
- Azure DevOps

> **Get started**: To see DevOps options for an App Service app, see Continuous Deployment to Azure App Service.
>
> **Try it now:** Try out several of the DevOps integrations.

Azure regions

Azure is a global cloud platform that is generally available in many regions around the world. When you provision a service, application, or VM in Azure, you're asked to select a region. This region represents a specific datacenter where your application runs or where your data is stored. These regions correspond to specific locations, which are published on the Azure regions page.

Choose the best region for your application and data

One of the benefits of using Azure is that you can deploy your applications to various datacenters around the globe. The region that you choose can affect the performance of your application. For example, it's better to choose a region that's closer to most of your customers to reduce latency in network requests. You might also want to select your region to meet the legal requirements for distributing your app in certain countries/regions. It's always a best practice to store application data in the same datacenter or in a datacenter as near as possible to the datacenter that is hosting your application.

Multi-region apps

Although unlikely, it's not impossible for an entire datacenter to go offline because of an event such as a natural disaster or Internet failure. It's a best practice to host vital business applications in more than one datacenter to provide maximum availability. Using multiple regions can also reduce latency for global users and provide additional opportunities for flexibility when updating applications.

Some services, such as Virtual Machine and App Services, use Azure Traffic Manager to enable multi-region support with failover between regions to support high-availability enterprise applications. For an example, see Azure reference architecture: Run a web application in multiple regions.

> **When to use**: When you have enterprise and high-availability applications that benefit from failover and replication.

How do I manage my applications and projects?

Azure provides a rich set of experiences for you to create and manage your Azure resources, applications, and projects—both programmatically and in the Azure portal.

Command-line interfaces and PowerShell

Azure provides two ways to manage your applications and services from the command line. You can use tools like Bash, Terminal, the command prompt, or your command-line tool of choice. Usually, you can do the same tasks from the command line as in the Azure portal—such as creating and configuring virtual machines, virtual networks, web apps, and other services.

- Azure Command-Line Interface (CLI): Lets you connect to an Azure subscription and program various tasks against Azure resources from the command line.
- Azure PowerShell: Provides a set of modules with cmdlets that enable you to manage Azure resources by using Windows PowerShell.

Azure portal

The Azure portal is a web-based application. You can use the Azure portal to create, manage, and remove Azure resources and services. It includes:

- A configurable dashboard
- Azure resource management tools
- Access to subscription settings and billing information. For more information, see the Azure portal overview.

REST APIs

Azure is built on a set of REST APIs that support the Azure portal UI. Most of these REST APIs are also supported

to let you programmatically provision and manage your Azure resources and applications from any Internet-enabled device. For the complete set of REST API documentation, see the Azure REST SDK reference.

APIs

Along with REST APIs, many Azure services also let you programmatically manage resources from your applications by using platform-specific Azure SDKs, including SDKs for the following development platforms:

- .NET
- Node.js
- Java
- PHP
- Python
- Ruby
- Go

Services such as Mobile Apps and Azure Media Services provide client-side SDKs to let you access services from web and mobile client apps.

Azure Resource Manager

Running your app on Azure likely involves working with multiple Azure services. These services follow the same life cycle and can be thought of as a logical unit. For example, a web app might use Web Apps, SQL Database, Storage, Azure Cache for Redis, and Azure Content Delivery Network services. Azure Resource Manager lets you work with the resources in your application as a group. You can deploy, update, or delete all the resources in a single, coordinated operation.

Along with logically grouping and managing related resources, Azure Resource Manager includes deployment capabilities that let you customize the deployment and configuration of related resources. For example, you can use Resource Manager deploy and configure an application. This application can consist of multiple virtual machines, a load balancer, and an Azure SQL database as a single unit.

You develop these deployments by using an Azure Resource Manager template, which is a JSON-formatted document. Templates let you define a deployment and manage your applications by using declarative templates, rather than scripts. Your templates can work for different environments, such as testing, staging, and production. For example, you can use templates to add a button to a GitHub repo that deploys the code in the repo to a set of Azure services with a single click.

> **When to use**: Use Resource Manager templates when you want a template-based deployment for your app that you can manage programmatically by using REST APIs, the Azure CLI, and Azure PowerShell.
>
> **Get started**: To get started using templates, see Authoring Azure Resource Manager templates.

Understanding accounts, subscriptions, and billing

As developers, we like to dive right into the code and try to get started as fast as possible with making our applications run. We certainly want to encourage you to start working in Azure as easily as possible. To help make it easy, Azure offers a free trial. Some services even have a "Try it for free" functionality, like Azure App Service, which doesn't require you to even create an account. As fun as it is to dive into coding and deploying your application to Azure, it's also important to take some time to understand how Azure works. Specifically, you should understand how it works from a standpoint of user accounts, subscriptions, and billing.

What is an Azure account?

To create or work with an Azure subscription, you must have an Azure account. An Azure account is simply an identity in Azure AD or in a directory, such as a work or school organization, that Azure AD trusts. If you don't belong to such an organization, you can always create a subscription by using your Microsoft Account, which is

trusted by Azure AD. To learn more about integrating on-premises Windows Server Active Directory with Azure AD, see Integrating your on-premises identities with Azure Active Directory.

Every Azure subscription has a trust relationship with an Azure AD instance. This means that it trusts that directory to authenticate users, services, and devices. Multiple subscriptions can trust the same directory, but a subscription trusts only one directory. To learn more, see How Azure subscriptions are associated with Azure Active Directory.

As well as defining individual Azure account identities, also called *users*, you can define *groups* in Azure AD. Creating user groups is a good way to manage access to resources in a subscription by using role-based access control (RBAC). To learn how to create groups, see Create a group in Azure Active Directory preview. You can also create and manage groups by using PowerShell.

Manage your subscriptions

A subscription is a logical grouping of Azure services that is linked to an Azure account. A single Azure account can contain multiple subscriptions. Billing for Azure services is done on a per-subscription basis. For a list of the available subscription offers by type, see Microsoft Azure Offer Details. Azure subscriptions have an Account Administrator who has full control over the subscription. They also have a Service Administrator who has control over all services in the subscription. For information about classic subscription administrators, see Add or change Azure subscription administrators. Individual accounts can be granted detailed control of Azure resources using role-based access control (RBAC).

Resource groups

When you provision new Azure services, you do so in a given subscription. Individual Azure services, which are also called resources, are created in the context of a resource group. Resource groups make it easier to deploy and manage your application's resources. A resource group should contain all the resources for your application that you want to work with as a unit. You can move resources between resource groups and even to different subscriptions. To learn about moving resources, see Move resources to new resource group or subscription.

The Azure Resource Explorer is a great tool for visualizing the resources that you've already created in your subscription. To learn more, see Use Azure Resource Explorer to view and modify resources.

Grant access to resources

When you allow access to Azure resources, it's always a best practice to provide users with the least privilege that's required to do a given task.

- **Role-based access control (RBAC)**: In Azure, you can grant access to user accounts (principals) at a specified scope: subscription, resource group, or individual resources. RBAC lets you deploy resources into a resource group and grant permissions to a specific user or group. It also lets you limit access to only the resources that belong to the target resource group. You can also grant access to a single resource, such as a virtual machine or virtual network. To grant access, you assign a role to the user, group, or service principal. There are many predefined roles, and you can also define your own custom roles. To learn more, see What is role-based access control (RBAC)?.

  > **When to use**: When you need fine-grained access management for users and groups or when you need to make a user an owner of a subscription.
  >
  > **Get started**: To learn more, see Manage access using RBAC and the Azure portal.

- **Service principal objects**: Along with providing access to user principals and groups, you can grant the same access to a service principal.

  > **When to use**: When you're programmatically managing Azure resources or granting access for applications. For more information, see Create Active Directory application and service principal.

Tags

Azure Resource Manager lets you assign custom tags to individual resources. Tags, which are key-value pairs, can

be helpful when you need to organize resources for billing or monitoring. Tags provide you a way to track resources across multiple resource groups. You can assign tags the following ways:

- In the portal
- In the Azure Resource Manager template
- Using the REST API
- Using the Azure CLI
- Using PowerShell

You can assign multiple tags to each resource. To learn more, see Using tags to organize your Azure resources.

Billing

In the move from on-premises computing to cloud-hosted services, tracking and estimating service usage and related costs are significant concerns. It's important to estimate what new resources cost to run on a monthly basis. You can also project how the billing looks for a given month based on the current spending.

Get resource usage data

Azure provides a set of Billing REST APIs that give access to resource consumption and metadata information for Azure subscriptions. These Billing APIs give you the ability to better predict and manage Azure costs. You can track and analyze spending in hourly increments and create spending alerts. You can also predict future billing based on current usage trends.

> **Get started**: To learn more about using the Billing APIs, see Azure Billing Usage and RateCard APIs overview.

Predict future costs

Although it's challenging to estimate costs ahead of time, Azure has tools that can help. It has a pricing calculator to help estimate the cost of deployed resources. You can also use the Billing resources in the portal and the Billing REST APIs to estimate future costs, based on current consumption.

> **Get started**: See Azure Billing Usage and RateCard APIs overview.

Azure subscription and service limits, quotas, and constraints

2/25/2020 • 85 minutes to read • Edit Online

This document lists some of the most common Microsoft Azure limits, which are also sometimes called quotas.

To learn more about Azure pricing, see Azure pricing overview. There, you can estimate your costs by using the pricing calculator. You also can go to the pricing details page for a particular service, for example, Windows VMs. For tips to help manage your costs, see Prevent unexpected costs with Azure billing and cost management.

Managing limits

If you want to raise the limit or quota above the default limit, open an online customer support request at no charge. The limits can't be raised above the maximum limit value shown in the following tables. If there's no maximum limit column, the resource doesn't have adjustable limits.

Free Trial subscriptions aren't eligible for limit or quota increases. If you have a Free Trial subscription, you can upgrade to a Pay-As-You-Go subscription. For more information, see Upgrade your Azure Free Trial subscription to a Pay-As-You-Go subscription and the Free Trial subscription FAQ.

Some limits are managed at a regional level.

Let's use vCPU quotas as an example. To request a quota increase with support for vCPUs, you must decide how many vCPUs you want to use in which regions. You then make a specific request for Azure resource group vCPU quotas for the amounts and regions that you want. If you need to use 30 vCPUs in West Europe to run your application there, you specifically request 30 vCPUs in West Europe. Your vCPU quota isn't increased in any other region--only West Europe has the 30-vCPU quota.

As a result, decide what your Azure resource group quotas must be for your workload in any one region. Then request that amount in each region into which you want to deploy. For help in how to determine your current quotas for specific regions, see Resolve errors for resource quotas.

General limits

For limits on resource names, see Naming rules and restrictions for Azure resources.

For information about Resource Manager API read and write limits, see Throttling Resource Manager requests.

Subscription limits

The following limits apply when you use Azure Resource Manager and Azure resource groups.

RESOURCE	DEFAULT LIMIT	MAXIMUM LIMIT
Subscriptions per Azure Active Directory tenant	Unlimited.	Unlimited.
Coadministrators per subscription	Unlimited.	Unlimited.
Resource groups per subscription	980	980

RESOURCE	DEFAULT LIMIT	MAXIMUM LIMIT
Azure Resource Manager API request size	4,194,304 bytes.	4,194,304 bytes.
Tags per subscription[1]	Unlimited.	Unlimited.
Unique tag calculations per subscription[1]	10,000	10,000
Subscription-level deployments per location	800[2]	800

[1]You can apply an unlimited number of tags per subscription. The number of tags per resource or resource group is limited to 50. Resource Manager returns a list of unique tag name and values in the subscription only when the number of tags is 10,000 or less. You still can find a resource by tag when the number exceeds 10,000.

[2]If you reach the limit of 800 deployments, delete deployments from the history that are no longer needed. To delete subscription level deployments, use Remove-AzDeployment or az deployment delete.

Resource group limits

RESOURCE	DEFAULT LIMIT	MAXIMUM LIMIT
Resources per resource group	N/A	Resources aren't limited by resource group. Instead, they're limited by resource type in a resource group. See next row.
Resources per resource group, per resource type	800	Some resource types can exceed the 800 limit. See Resources not limited to 800 instances per resource group.
Deployments per resource group in the deployment history	800[1]	800
Resources per deployment	800	800
Management locks per unique scope	20	20
Number of tags per resource or resource group	50	50
Tag key length	512	512
Tag value length	256	256

[1]If you reach the limit of 800 deployments per resource group, delete deployments from the history that are no longer needed. Deleting an entry from the deployment history doesn't affect the deployed resources. For more information, see Resolve error when deployment count exceeds 800.

Template limits

VALUE	DEFAULT LIMIT	MAXIMUM LIMIT
Parameters	256	256

VALUE	DEFAULT LIMIT	MAXIMUM LIMIT
Variables	256	256
Resources (including copy count)	800	800
Outputs	64	64
Template expression	24,576 chars	24,576 chars
Resources in exported templates	200	200
Template size	4 MB	4 MB
Parameter file size	64 KB	64 KB

You can exceed some template limits by using a nested template. For more information, see Use linked templates when you deploy Azure resources. To reduce the number of parameters, variables, or outputs, you can combine several values into an object. For more information, see Objects as parameters.

Active Directory limits

Here are the usage constraints and other service limits for the Azure Active Directory (Azure AD) service.

CATEGORY	LIMITS
Directories	A single user can belong to a maximum of 500 Azure AD directories as a member or a guest. A single user can create a maximum of 20 directories.
Domains	You can add no more than 900 managed domain names. If you set up all of your domains for federation with on-premises Active Directory, you can add no more than 450 domain names in each directory.
Resources	• A maximum of 50,000 Azure AD resources can be created in a single directory by users of the Free edition of Azure Active Directory by default. If you have at least one verified domain, the default directory service quota in Azure AD is extended to 300,000 Azure AD resources. • A non-admin user can create no more than 250 Azure AD resources. Both active resources and deleted resources that are available to restore count toward this quota. Only deleted Azure AD resources that were deleted fewer than 30 days ago are available to restore. Deleted Azure AD resources that are no longer available to restore count toward this quota at a value of one-quarter for 30 days. If you have developers who are likely to repeatedly exceed this quota in the course of their regular duties, you can create and assign a custom role with permission to create a limitless number of app registrations.

CATEGORY	LIMITS
Schema extensions	• String-type extensions can have a maximum of 256 characters. • Binary-type extensions are limited to 256 bytes. • Only 100 extension values, across *all* types and *all* applications, can be written to any single Azure AD resource. • Only User, Group, TenantDetail, Device, Application, and ServicePrincipal entities can be extended with string-type or binary-type single-valued attributes. • Schema extensions are available only in the Graph API version 1.21 preview. The application must be granted write access to register an extension.
Applications	A maximum of 100 users can be owners of a single application.
Application Manifest	A maximum of 1200 entries can be added in the Application Manifest.

CATEGORY	LIMITS
Groups	• A user can create a maximum of 250 groups in an Azure AD organization. • An Azure AD organization can have a maximum of 5000 dynamic groups. • A maximum of 100 users can be owners of a single group. • Any number of Azure AD resources can be members of a single group. • A user can be a member of any number of groups. • The number of members in a group that you can synchronize from your on-premises Active Directory to Azure Active Directory by using Azure AD Connect is limited to 50,000 members. • Nested Groups in Azure AD are not supported within all scenarios At this time the following are the supported scenarios with nested groups. • One group can be added as a member of another group and you can achieve group nesting. • Group membership claims (when an app is configured to receive group membership claims in the token, nested groups the signed-in user is a member of are included) • Conditional access (when scoping a conditional access policy to a group) • Restricting access to self-serve password reset • Restricting which users can do Azure AD Join and device registration The following scenarios DO NOT supported nested groups: • App role assignment (assigning groups to an app is supported, but groups nested within the directly assigned group will not have access), both for access and for provisioning • Group-based licensing (assigning a license automatically to all members of a group) • Office 365 Groups.
Application Proxy	• A maximum of 500 transactions per second per App Proxy application • A maximum of 750 transactions per second for the Azure AD organization A transaction is defined as a single http request and response for a unique resource. When throttled, clients will receive a 429 response (too many requests).

CATEGORY	LIMITS
Access Panel	• There's no limit to the number of applications that can be seen in the Access Panel per user. This applies to users assigned licenses for Azure AD Premium or the Enterprise Mobility Suite. • A maximum of 10 app tiles can be seen in the Access Panel for each user. This limit applies to users who are assigned licenses for Azure AD Free license plan. Examples of app tiles include Box, Salesforce, or Dropbox. This limit doesn't apply to administrator accounts.
Reports	A maximum of 1,000 rows can be viewed or downloaded in any report. Any additional data is truncated.
Administrative units	An Azure AD resource can be a member of no more than 30 administrative units.
Admin roles and permissions	• A group cannot be added as an owner. • A group cannot be assigned to a role. • Users' ability to read other users' directory information cannot be restricted outside of the Azure AD organization-wide switch to disable all non-admin users' access to all directory information (not recommended). More information on default permissions here. • It may take up to 15 minutes or signing out/signing in before admin role membership additions and revocations take effect.

API Management limits

RESOURCE	LIMIT
Maximum number of scale units	10 per region[1]
Cache size	5 GiB per unit[2]
Concurrent back-end connections[3] per HTTP authority	2,048 per unit[4]
Maximum cached response size	2 MiB
Maximum policy document size	256 KiB[5]
Maximum custom gateway domains per service instance[6]	20
Maximum number of CA certificates per service instance	10
Maximum number of service instances per subscription[7]	20
Maximum number of subscriptions per service instance[7]	500
Maximum number of client certificates per service instance[7]	50

RESOURCE	LIMIT
Maximum number of APIs per service instance[7]	50
Maximum number of API operations per service instance[7]	1,000
Maximum total request duration[7]	30 seconds
Maximum buffered payload size[7]	2 MiB
Maximum request URL size[8]	4096 bytes

[1]Scaling limits depend on the pricing tier. To see the pricing tiers and their scaling limits, see API Management pricing.
[2]Per unit cache size depends on the pricing tier. To see the pricing tiers and their scaling limits, see API Management pricing.
[3]Connections are pooled and reused unless explicitly closed by the back end.
[4]This limit is per unit of the Basic, Standard, and Premium tiers. The Developer tier is limited to 1,024. This limit doesn't apply to the Consumption tier.
[5]This limit applies to the Basic, Standard, and Premium tiers. In the Consumption tier, policy document size is limited to 4 KiB.
[6]This resource is available in the Premium tier only.
[7]This resource applies to the Consumption tier only.
[8]Applies to the Consumption tier only. Includes an up to 2048 bytes long query string.

App Service limits

The following App Service limits include limits for Web Apps, Mobile Apps, and API Apps.

RESOURCE	FREE	SHARED	BASIC	STANDARD	PREMIUM (V2)	ISOLATED
Web, mobile, or API apps per Azure App Service plan[1]	10	100	Unlimited[2]	Unlimited[2]	Unlimited[2]	Unlimited[2]
App Service plan	10 per region	10 per resource group	100 per resource group	100 per resource group	100 per resource group	100 per resource group
Compute instance type	Shared	Shared	Dedicated[3]	Dedicated[3]	Dedicated[3]	Dedicated[3]
Scale out (maximum instances)	1 shared	1 shared	3 dedicated[3]	10 dedicated[3]	30 dedicated[3]	100 dedicated[4]
Storage[5]	1 GB[5]	1 GB[5]	10 GB[5]	50 GB[5]	250 GB[5]	1 TB[5]
CPU time (5 minutes)[6]	3 minutes	3 minutes	Unlimited, pay at standard rates	Unlimited, pay at standard rates	Unlimited, pay at standard rates	Unlimited, pay at standard rates

RESOURCE	FREE	SHARED	BASIC	STANDARD	PREMIUM (V2)	ISOLATED
CPU time (day)[6]	60 minutes	240 minutes	Unlimited, pay at standard rates	Unlimited, pay at standard rates	Unlimited, pay at standard rates	Unlimited, pay at standard rates
Memory (1 hour)	1,024 MB per App Service plan	1,024 MB per app	N/A	N/A	N/A	N/A
Bandwidth	165 MB	Unlimited, data transfer rates apply	Unlimited, data transfer rates apply	Unlimited, data transfer rates apply	Unlimited, data transfer rates apply	Unlimited, data transfer rates apply
Application architecture	32-bit	32-bit	32-bit/64-bit	32-bit/64-bit	32-bit/64-bit	32-bit/64-bit
Web sockets per instance[7]	5	35	350	Unlimited	Unlimited	Unlimited
IP connections	600	600	Depends on instance size[8]	Depends on instance size[8]	Depends on instance size[8]	16,000
Concurrent debugger connections per application	1	1	1	5	5	5
App Service Certificates per subscription[9]	Not supported	Not supported	10	10	10	10
Custom domains per app	0 (azurewebsites.net subdomain only)	500	500	500	500	500
Custom domain SSL support	Not supported, wildcard certificate for *.azurewebsites.net available by default	Not supported, wildcard certificate for *.azurewebsites.net available by default	Unlimited SNI SSL connections	Unlimited SNI SSL and 1 IP SSL connections included	Unlimited SNI SSL and 1 IP SSL connections included	Unlimited SNI SSL and 1 IP SSL connections included
Hybrid connections per plan			5	25	200	200
Integrated load balancer		X	X	X	X	X[10]
Always On			X	X	X	X

RESOURCE	FREE	SHARED	BASIC	STANDARD	PREMIUM (V2)	ISOLATED
Scheduled backups				Scheduled backups every 2 hours, a maximum of 12 backups per day (manual + scheduled)	Scheduled backups every hour, a maximum of 50 backups per day (manual + scheduled)	Scheduled backups every hour, a maximum of 50 backups per day (manual + scheduled)
Autoscale				X	X	X
WebJobs[11]	X	X	X	X	X	X
Endpoint monitoring			X	X	X	X
Staging slots				5	20	20
SLA			99.95%	99.95%	99.95%	99.95%

[1]Apps and storage quotas are per App Service plan unless noted otherwise.
[2]The actual number of apps that you can host on these machines depends on the activity of the apps, the size of the machine instances, and the corresponding resource utilization.
[3]Dedicated instances can be of different sizes. For more information, see App Service pricing.
[4]More are allowed upon request.
[5]The storage limit is the total content size across all apps in the same App service plan. The total content size of all apps across all App service plans in a single resource group and region cannot exceed 500GB.
[6]These resources are constrained by physical resources on the dedicated instances (the instance size and the number of instances).
[7]If you scale an app in the Basic tier to two instances, you have 350 concurrent connections for each of the two instances. For Standard tier and above, there are no theoretical limits to web sockets, but other factors can limit the number of web sockets. For example, maximum concurrent requests allowed (defined by `maxConcurrentRequestsPerCpu`) are: 7,500 per small VM, 15,000 per medium VM (7,500 x 2 cores), and 75,000 per large VM (18,750 x 4 cores).
[8]The maximum IP connections are per instance and depend on the instance size: 1,920 per B1/S1/P1V2 instance, 3,968 per B2/S2/P2V2 instance, 8,064 per B3/S3/P3V2 instance.
[9]The App Service Certificate quota limit per subscription can be increased via a support request to a maximum limit of 200.
[10]App Service Isolated SKUs can be internally load balanced (ILB) with Azure Load Balancer, so there's no public connectivity from the internet. As a result, some features of an ILB Isolated App Service must be used from machines that have direct access to the ILB network endpoint.
[11]Run custom executables and/or scripts on demand, on a schedule, or continuously as a background task within your App Service instance. Always On is required for continuous WebJobs execution. There's no predefined limit on the number of WebJobs that can run in an App Service instance. There are practical limits that depend on what the application code is trying to do.

Automation limits

Process automation

RESOURCE	MAXIMUM LIMIT	NOTES
Maximum number of new jobs that can be submitted every 30 seconds per Azure Automation account (nonscheduled jobs)	100	When this limit is reached, the subsequent requests to create a job fail. The client receives an error response.
Maximum number of concurrent running jobs at the same instance of time per Automation account (nonscheduled jobs)	200	When this limit is reached, the subsequent requests to create a job fail. The client receives an error response.
Maximum storage size of job metadata for a 30-day rolling period	10 GB (approximately 4 million jobs)	When this limit is reached, the subsequent requests to create a job fail.
Maximum job stream limit	1MB	A single stream cannot be larger than 1 MB.
Maximum number of modules that can be imported every 30 seconds per Automation account	5	
Maximum size of a module	100 MB	
Job run time, Free tier	500 minutes per subscription per calendar month	
Maximum amount of disk space allowed per sandbox[1]	1 GB	Applies to Azure sandboxes only.
Maximum amount of memory given to a sandbox[1]	400 MB	Applies to Azure sandboxes only.
Maximum number of network sockets allowed per sandbox[1]	1,000	Applies to Azure sandboxes only.
Maximum runtime allowed per runbook[1]	3 hours	Applies to Azure sandboxes only.
Maximum number of Automation accounts in a subscription	No limit	
Maximum number of Hybrid Worker Groups per Automation Account	4,000	
Maximum number of concurrent jobs that can be run on a single Hybrid Runbook Worker	50	
Maximum runbook job parameter size	512 kilobits	
Maximum runbook parameters	50	If you reach the 50-parameter limit, you can pass a JSON or XML string to a parameter and parse it with the runbook.

RESOURCE	MAXIMUM LIMIT	NOTES
Maximum webhook payload size	512 kilobits	
Maximum days that job data is retained	30 days	
Maximum PowerShell workflow state size	5 MB	Applies to PowerShell workflow runbooks when checkpointing workflow.

[1]A sandbox is a shared environment that can be used by multiple jobs. Jobs that use the same sandbox are bound by the resource limitations of the sandbox.

Change Tracking and Inventory

The following table shows the tracked item limits per machine for change tracking.

RESOURCE	LIMIT	NOTES
File	500	
Registry	250	
Windows software	250	Doesn't include software updates.
Linux packages	1,250	
Services	250	
Daemon	250	

Update Management

The following table shows the limits for Update Management.

RESOURCE	LIMIT	NOTES
Number of machines per update deployment	1000	

Azure Cache for Redis limits

RESOURCE	LIMIT
Cache size	1.2 TB
Databases	64
Maximum connected clients	40,000
Azure Cache for Redis replicas, for high availability	1
Shards in a premium cache with clustering	10

Azure Cache for Redis limits and sizes are different for each pricing tier. To see the pricing tiers and their associated sizes, see Azure Cache for Redis pricing.

For more information on Azure Cache for Redis configuration limits, see Default Redis server configuration.

Because configuration and management of Azure Cache for Redis instances is done by Microsoft, not all Redis commands are supported in Azure Cache for Redis. For more information, see Redis commands not supported in Azure Cache for Redis.

Azure Cloud Services limits

RESOURCE	DEFAULT LIMIT	MAXIMUM LIMIT
Web or worker roles per deployment[1]	25	25
Instance input endpoints per deployment	25	25
Input endpoints per deployment	25	25
Internal endpoints per deployment	25	25
Hosted service certificates per deployment	199	199

[1]Each Azure Cloud Service with web or worker roles can have two deployments, one for production and one for staging. This limit refers to the number of distinct roles, that is, configuration. This limit doesn't refer to the number of instances per role, that is, scaling.

Azure Cognitive Search limits

Pricing tiers determine the capacity and limits of your search service. Tiers include:

- **Free** multi-tenant service, shared with other Azure subscribers, is intended for evaluation and small development projects.
- **Basic** provides dedicated computing resources for production workloads at a smaller scale, with up to three replicas for highly available query workloads.
- **Standard**, which includes S1, S2, S3, and S3 High Density, is for larger production workloads. Multiple levels exist within the Standard tier so that you can choose a resource configuration that best matches your workload profile.

Limits per subscription

You can create multiple services within a subscription. Each one can be provisioned at a specific tier. You're limited only by the number of services allowed at each tier. For example, you could create up to 12 services at the Basic tier and another 12 services at the S1 tier within the same subscription. For more information about tiers, see Choose an SKU or tier for Azure Cognitive Search.

Maximum service limits can be raised upon request. If you need more services within the same subscription, contact Azure Support.

RESOURCE	FREE[1]	BASIC	S1	S2	S3	S3 HD	L1	L2
Maximum services	1	16	16	8	6	6	6	6

RESOURCE	FREE	BASIC	S1	S2	S3	S3 HD	L1	L2
Maximum scale in search units (SU)[2]	N/A	3 SU	36 SU	36 SU	36 SU	36 SU	36 SU	36 SU

[1] Free is based on shared, not dedicated, resources. Scale-up is not supported on shared resources.

[2] Search units are billing units, allocated as either a *replica* or a *partition*. You need both resources for storage, indexing, and query operations. To learn more about SU computations, see Scale resource levels for query and index workloads.

Limits per search service

Storage is constrained by disk space or by a hard limit on the *maximum number* of indexes, document, or other high-level resources, whichever comes first. The following table documents storage limits. For maximum limits on indexes, documents, and other objects, see Limits by resource.

RESOURCE	FREE	BASIC[1]	S1	S2	S3	S3 HD[2]	L1	L2
Service level agreement (SLA)[3]	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Storage per partition	50 MB	2 GB	25 GB	100 GB	200 GB	200 GB	1 TB	2 TB
Partitions per service	N/A	1	12	12	12	3	12	12
Partition size	N/A	2 GB	25 GB	100 GB	200 GB	200 GB	1 TB	2 TB
Replicas	N/A	3	12	12	12	12	12	12

[1] Basic has one fixed partition. At this tier, additional search units are used for allocating more replicas for increased query workloads.

[2] S3 HD has a hard limit of three partitions, which is lower than the partition limit for S3. The lower partition limit is imposed because the index count for S3 HD is substantially higher. Given that service limits exist for both computing resources (storage and processing) and content (indexes and documents), the content limit is reached first.

[3] Service level agreements are offered for billable services on dedicated resources. Free services and preview features have no SLA. For billable services, SLAs take effect when you provision sufficient redundancy for your service. Two or more replicas are required for query (read) SLAs. Three or more replicas are required for query and indexing (read-write) SLAs. The number of partitions isn't an SLA consideration.

To learn more about limits on a more granular level, such as document size, queries per second, keys, requests, and responses, see Service limits in Azure Cognitive Search.

Azure Cognitive Services limits

The following limits are for the number of Cognitive Services resources per Azure subscription. Each of the Cognitive Services may have additional limitations, for more information see Azure Cognitive Services.

TYPE	LIMIT	EXAMPLE
A mixture of Cognitive Services resources	Maximum of 200 total Cognitive Services resources.	100 Computer Vision resources in West US 2, 50 Speech Service resources in West US, and 50 Text Analytics resources in East US.
A single type of Cognitive Services resources.	Maximum of 100 resources per region, with a maximum of 200 total Cognitive Services resources.	100 Computer Vision resources in West US 2, and 100 Computer Vision resources in East US.

Azure Cosmos DB limits

For Azure Cosmos DB limits, see Limits in Azure Cosmos DB.

Azure Data Explorer limits

The following table describes the maximum limits for Azure Data Explorer clusters.

RESOURCE	LIMIT
Clusters per region per subscription	20
Instances per cluster	1000
Number of databases in a cluster	10,000
Number of attached database configurations in a cluster	70

The following table describes the limits on management operations performed on Azure Data Explorer clusters.

SCOPE	OPERATION	LIMIT
Cluster	read (for example, get a cluster)	500 per 5 minutes
Cluster	write (for example, create a database)	1000 per hour

Azure Database for MySQL

For Azure Database for MySQL limits, see Limitations in Azure Database for MySQL.

Azure Database for PostgreSQL

For Azure Database for PostgreSQL limits, see Limitations in Azure Database for PostgreSQL.

Azure Functions limits

RESOURCE	CONSUMPTION PLAN	PREMIUM PLAN	APP SERVICE PLAN[1]
Scale out	Event driven	Event driven	Manual/autoscale

RESOURCE	CONSUMPTION PLAN	PREMIUM PLAN	APP SERVICE PLAN
Max instances	200	100	10-20
Default timeout duration (min)	5	30	30[2]
Max timeout duration (min)	10	unbounded[8]	unbounded[3]
Max outbound connections (per instance)	600 active (1200 total)	unbounded	unbounded
Max request size (MB)[4]	100	100	100
Max query string length[4]	4096	4096	4096
Max request URL length[4]	8192	8192	8192
ACU per instance	100	210-840	100-840
Max memory (GB per instance)	1.5	3.5-14	1.75-14
Function apps per plan	100	100	unbounded[5]
App Service plans	100 per region	100 per resource group	100 per resource group
Storage[6]	1 GB	250 GB	50-1000 GB
Custom domains per app	500[7]	500	500
Custom domain SSL support	unbounded SNI SSL connection included	unbounded SNI SSL and 1 IP SSL connections included	unbounded SNI SSL and 1 IP SSL connections included

[1] For specific limits for the various App Service plan options, see the App Service plan limits.
[2] By default, the timeout for the Functions 1.x runtime in an App Service plan is unbounded.
[3] Requires the App Service plan be set to Always On. Pay at standard rates.
[4] These limits are set in the host.
[5] The actual number of function apps that you can host depends on the activity of the apps, the size of the machine instances, and the corresponding resource utilization.
[6] The storage limit is the total content size in temporary storage across all apps in the same App Service plan. Consumption plan uses Azure Files for temporary storage.
[7] When your function app is hosted in a Consumption plan, only the CNAME option is supported. For function apps in a Premium plan or an App Service plan, you can map a custom domain using either a CNAME or an A record.
[8] Guaranteed for up to 60 minutes.

Azure Kubernetes Service limits

RESOURCE	DEFAULT LIMIT
Maximum clusters per subscription	100

RESOURCE	DEFAULT LIMIT
Maximum nodes per cluster with Virtual Machine Availability Sets and Basic Load Balancer SKU	100
Maximum nodes per cluster with Virtual Machine Scale Sets and Standard Load Balancer SKU	1000 (100 nodes per node pool)
Maximum pods per node: Basic networking with Kubenet	110
Maximum pods per node: Advanced networking with Azure Container Networking Interface	Azure CLI deployment: 30[1] Azure Resource Manager template: 30[1] Portal deployment: 30

[1]When you deploy an Azure Kubernetes Service (AKS) cluster with the Azure CLI or a Resource Manager template, this value is configurable up to 250 pods per node. You can't configure maximum pods per node after you've already deployed an AKS cluster, or if you deploy a cluster by using the Azure portal.

Azure Machine Learning limits

The latest values for Azure Machine Learning Compute quotas can be found in the Azure Machine Learning quota page

Azure Maps limits

The following table shows the usage limit for the Azure Maps S0 pricing tier. Usage limit depends on the pricing tier.

RESOURCE	S0 PRICING TIER LIMIT
Maximum request rate per subscription	50 requests per second

The following table shows the data size limit for Azure Maps. The Azure Maps data service is available only at the S1 pricing tier.

RESOURCE	LIMIT
Maximum size of data	50 MB

For more information on the Azure Maps pricing tiers, see Azure Maps pricing.

Azure Monitor limits

Alerts

RESOURCE	DEFAULT LIMIT	MAXIMUM LIMIT
Metric alerts (classic)	100 active alert rules per subscription.	Call support.
Metric alerts	1000 active alert rules per subscription in Azure public, Azure China 21Vianet and Azure Government clouds.	Call support.
Activity log alerts	100 active alert rules per subscription.	Same as default.

RESOURCE	DEFAULT LIMIT	MAXIMUM LIMIT
Log alerts	512	Call support.
Action groups	2,000 action groups per subscription.	Call support.
Autoscale settings	100 per region per subscription.	Same as default.

Action groups

RESOURCE	DEFAULT LIMIT	MAXIMUM LIMIT
Azure app push	10 Azure app actions per action group.	Call support.
Email	1,000 email actions in an action group. No more than 100 emails in an hour. Also see the rate limiting information.	Call support.
ITSM	10 ITSM actions in an action group.	Call support.
Logic app	10 logic app actions in an action group.	Call support.
Runbook	10 runbook actions in an action group.	Call support.
SMS	10 SMS actions in an action group. No more than 1 SMS message every 5 minutes. Also see the rate limiting information.	Call support.
Voice	10 voice actions in an action group. No more than 1 voice call every 5 minutes. Also see the rate limiting information.	Call support.
Webhook	10 webhook actions in an action group. Maximum number of webhook calls is 1500 per minute per subscription. Other limits are available at action-specific information.	Call support.

Log queries and language

LIMIT	DESCRIPTION
Query language	Azure Monitor uses the same Kusto query language as Azure Data Explorer. See Azure Monitor log query language differences for KQL language elements not supported in Azure Monitor.
Azure regions	Log queries can experience excessive overhead when data spans Log Analytics workspaces in multiple Azure regions. See Query limits for details.

LIMIT	DESCRIPTION
Cross resource queries	Maximum number of Application Insights resources and Log Analytics workspaces in a single query limited to 100. Cross-resource query is not supported in View Designer. Cross-resource query in log alerts is supported in the new scheduledQueryRules API. See Cross-resource query limits for details.
Query throttling	A user is limited to 200 queries per 30 seconds on any number of workspaces. This limit applies to programmatic queries or to queries initiated by visualization parts such as Azure dashboards and the Log Analytics workspace summary page.

Log Analytics workspaces

Data collection volume and retention

TIER	LIMIT PER DAY	DATA RETENTION	COMMENT
Current Per GB pricing tier (introduced April 2018)	No limit	30 - 730 days	Data retention beyond 31 days is available for additional charges. Learn more about Azure Monitor pricing.
Legacy Free tiers (introduced April 2016)	500 MB	7 days	When your workspace reaches the 500 MB per day limit, data ingestion stops and resumes at the start of the next day. A day is based on UTC. Note that data collected by Azure Security Center is not included in this 500 MB per day limit and will continue to be collected above this limit.
Legacy Standalone Per GB tier (introduced April 2016)	No limit	30 to 730 days	Data retention beyond 31 days is available for additional charges. Learn more about Azure Monitor pricing.
Legacy Per Node (OMS) (introduced April 2016)	No limit	30 to 730 days	Data retention beyond 31 days is available for additional charges. Learn more about Azure Monitor pricing.
Legacy Standard tier	No limit	30 days	Retention can't be adjusted
Legacy Premium tier	No limit	365 days	Retention can't be adjusted

Number of workspaces per subscription.

PRICING TIER	WORKSPACE LIMIT	COMMENTS
Free tier	10	This limit can't be increased.
All other tiers	No limit	You're limited by the number of resources within a resource group and the number of resource groups per subscription.

Azure portal

CATEGORY	LIMITS	COMMENTS
Maximum records returned by a log query	10,000	Reduce results using query scope, time range, and filters in the query.

Data Collector API

CATEGORY	LIMITS	COMMENTS
Maximum size for a single post	30 MB	Split larger volumes into multiple posts.
Maximum size for field values	32 KB	Fields longer than 32 KB are truncated.

Search API

CATEGORY	LIMITS	COMMENTS
Maximum records returned in a single query	500,000	
Maximum size of data returned	64,000,000 bytes (~61 MiB)	
Maximum query running time	10 minutes	See Timeouts for details.
Maximum request rate	200 requests per 30 seconds per AAD user or client IP address	See Rate limits for details.

General workspace limits

CATEGORY	LIMITS	COMMENTS
Maximum columns in a table	500	
Maximum characters for column name	500	
Data export	Not currently available	Use Azure Function or Logic App to aggregate and export data.

Data ingestion volume rate

Azure Monitor is a high scale data service that serves thousands of customers sending terabytes of data each month at a growing pace. The default ingestion volume rate limit for data sent from Azure resources using diagnostic settings is approximately **6 GB/min** per workspace. This is an approximate value since the actual size can vary between data types depending on the log length and its compression ratio. This limit does not apply to

data that is sent from agents or Data Collector API.

If you send data at a higher rate to a single workspace, some data is dropped, and an event is sent to the *Operation* table in your workspace every 6 hours while the threshold continues to be exceeded. If your ingestion volume continues to exceed the rate limit or you are expecting to reach it sometime soon, you can request an increase to your workspace by opening a support request.

To be notified on such an event in your workspace, create a log alert rule using the following query with alert logic base on number of results grater than zero.

```
Operation
|where OperationCategory == "Ingestion"
|where Detail startswith "The rate of data crossed the threshold"
```

> **NOTE**
> Depending on how long you've been using Log Analytics, you might have access to legacy pricing tiers. Learn more about Log Analytics legacy pricing tiers.

Application Insights

There are some limits on the number of metrics and events per application, that is, per instrumentation key. Limits depend on the pricing plan that you choose.

RESOURCE	DEFAULT LIMIT	NOTE
Total data per day	100 GB	You can reduce data by setting a cap. If you need more data, you can increase the limit in the portal, up to 1,000 GB. For capacities greater than 1,000 GB, send email to AIDataCap@microsoft.com.
Throttling	32,000 events/second	The limit is measured over a minute.
Data retention	90 days	This resource is for Search, Analytics, and Metrics Explorer.
Availability multi-step test detailed results retention	90 days	This resource provides detailed results of each step.
Maximum event size	64,000,000 bytes	
Property and metric name length	150	See type schemas.
Property value string length	8,192	See type schemas.
Trace and exception message length	32,768	See type schemas.
Availability tests count per app	100	
Profiler data retention	5 days	
Profiler data sent per day	10 GB	

For more information, see About pricing and quotas in Application Insights.

Azure Policy limits

There's a maximum count for each object type for Azure Policy. An entry of *Scope* means either the subscription or the management group.

WHERE	WHAT	MAXIMUM COUNT
Scope	Policy definitions	500
Scope	Initiative definitions	100
Tenant	Initiative definitions	1,000
Scope	Policy or initiative assignments	100
Policy definition	Parameters	20
Initiative definition	Policies	100
Initiative definition	Parameters	100
Policy or initiative assignments	Exclusions (notScopes)	400
Policy rule	Nested conditionals	512

Azure SignalR Service limits

RESOURCE	DEFAULT LIMIT	MAXIMUM LIMIT
Azure SignalR Service units per instance for Free tier	1	1
Azure SignalR Service units per instance for Standard tier	100	100
Azure SignalR Service units per subscription per region for Free tier	5	5
Total Azure SignalR Service unit counts per subscription per region	150	Unlimited
Connections per unit per day for Free tier	20	20
Connections per unit per day for Standard tier	1,000	1,000
Included messages per unit per day for Free tier	20,000	20,000

RESOURCE	DEFAULT LIMIT	MAXIMUM LIMIT
Included messages per unit per day for Standard tier	1,000,000	1,000,000

To request an update to your subscription's default limits, open a support ticket.

Backup limits

For a summary of Azure Backup support settings and limitations, see Azure Backup Support Matrices.

Batch limits

RESOURCE	DEFAULT LIMIT	MAXIMUM LIMIT
Azure Batch accounts per region per subscription	1-3	50
Dedicated cores per Batch account	90-900	Contact support
Low-priority cores per Batch account	10-100	Contact support
Active jobs and job schedules per Batch account (**completed** jobs have no limit)	100-300	1,000[1]
Pools per Batch account	20-100	500[1]

> **NOTE**
> Default limits vary depending on the type of subscription you use to create a Batch account. Cores quotas shown are for Batch accounts in Batch service mode. View the quotas in your Batch account.

[1]To request an increase beyond this limit, contact Azure Support.

Classic deployment model limits

If you use classic deployment model instead of the Azure Resource Manager deployment model, the following limits apply.

RESOURCE	DEFAULT LIMIT	MAXIMUM LIMIT
vCPUs per subscription[1]	20	10,000
Coadministrators per subscription	200	200
Storage accounts per subscription[2]	100	100
Cloud services per subscription	20	200
Local networks per subscription	10	500
DNS servers per subscription	9	100

RESOURCE	DEFAULT LIMIT	MAXIMUM LIMIT
Reserved IPs per subscription	20	100
Affinity groups per subscription	256	256
Subscription name length (characters)	64	64

[1]Extra small instances count as one vCPU toward the vCPU limit despite using a partial CPU core.

[2]The storage account limit includes both Standard and Premium storage accounts.

Container Instances limits

RESOURCE	DEFAULT LIMIT
Standard sku container groups per region per subscription	100[1]
Dedicated sku container groups per region per subscription	0[1]
Number of containers per container group	60
Number of volumes per container group	20
Ports per IP	5
Container instance log size - running instance	4 MB
Container instance log size - stopped instance	16 KB or 1,000 lines
Container creates per hour	300[1]
Container creates per 5 minutes	100[1]
Container deletes per hour	300[1]
Container deletes per 5 minutes	100[1]

[1]To request a limit increase, create an Azure Support request.

Container Registry limits

The following table details the features and limits of the Basic, Standard, and Premium service tiers.

RESOURCE	BASIC	STANDARD	PREMIUM
Storage[1]	10 GiB	100 GiB	500 GiB
Maximum image layer size	200 GiB	200 GiB	200 GiB
ReadOps per minute[2, 3]	1,000	3,000	10,000
WriteOps per minute[2, 4]	100	500	2,000

RESOURCE	BASIC	STANDARD	PREMIUM
Download bandwidth MBps[2]	30	60	100
Upload bandwidth MBps[2]	10	20	50
Webhooks	2	10	500
Geo-replication	N/A	N/A	Supported
Content trust	N/A	N/A	Supported
Virtual network access	N/A	N/A	Preview
Repository-scoped permissions	N/A	N/A	Preview
• Tokens	N/A	N/A	20,000
• Scope maps	N/A	N/A	20,000
• Repositories per scope map	N/A	N/A	500

[1]The specified storage limits are the amount of *included* storage for each tier. You're charged an additional daily rate per GiB for image storage above these limits. For rate information, see Azure Container Registry pricing.

[2]*ReadOps*, *WriteOps*, and *Bandwidth* are minimum estimates. Azure Container Registry strives to improve performance as usage requires.

[3]A docker pull translates to multiple read operations based on the number of layers in the image, plus the manifest retrieval.

[4]A docker push translates to multiple write operations, based on the number of layers that must be pushed. A `docker push` includes *ReadOps* to retrieve a manifest for an existing image.

Content Delivery Network limits

RESOURCE	DEFAULT LIMIT
Azure Content Delivery Network profiles	25
Content Delivery Network endpoints per profile	25
Custom domains per endpoint	25

A Content Delivery Network subscription can contain one or more Content Delivery Network profiles. A Content Delivery Network profile can contain one or more Content Delivery Network endpoints. You might want to use multiple profiles to organize your Content Delivery Network endpoints by internet domain, web application, or some other criteria.

Data Factory limits

Azure Data Factory is a multitenant service that has the following default limits in place to make sure customer

subscriptions are protected from each other's workloads. To raise the limits up to the maximum for your subscription, contact support.

Version 2

RESOURCE	DEFAULT LIMIT	MAXIMUM LIMIT
Data factories in an Azure subscription	50	Contact support.
Total number of entities, such as pipelines, data sets, triggers, linked services, and integration runtimes, within a data factory	5,000	Contact support.
Total CPU cores for Azure-SSIS Integration Runtimes under one subscription	256	Contact support.
Concurrent pipeline runs per data factory that's shared among all pipelines in the factory	10,000	Contact support.
Concurrent External activity runs per subscription per Azure Integration Runtime region External activities are managed on integration runtime but execute on linked services, including Databricks, stored procedure, HDInsights, Web, and others.	3000	Contact support.
Concurrent Pipeline activity runs per subscription per Azure Integration Runtime region Pipeline activities execute on integration runtime, including Lookup, GetMetadata, and Delete.	1000	Contact support.
Concurrent authoring operations per subscription per Azure Integration Runtime region Including test connection, browse folder list and table list, preview data.	200	Contact support.
Concurrent Data Integration Units[1] consumption per subscription per Azure Integration Runtime region	Region group 1[2]: 6000 Region group 2[2]: 3000 Region group 3[2]: 1500	Contact support.
Maximum activities per pipeline, which includes inner activities for containers	40	40
Maximum number of linked integration runtimes that can be created against a single self-hosted integration runtime	100	Contact support.
Maximum parameters per pipeline	50	50
ForEach items	100,000	100,000

RESOURCE	DEFAULT LIMIT	MAXIMUM LIMIT
ForEach parallelism	20	50
Maximum queued runs per pipeline	100	100
Characters per expression	8,192	8,192
Minimum tumbling window trigger interval	15 min	15 min
Maximum timeout for pipeline activity runs	7 days	7 days
Bytes per object for pipeline objects[3]	200 KB	200 KB
Bytes per object for dataset and linked service objects[3]	100 KB	2,000 KB
Data Integration Units[1] per copy activity run	256	Contact support.
Write API calls	1,200/h This limit is imposed by Azure Resource Manager, not Azure Data Factory.	Contact support.
Read API calls	12,500/h This limit is imposed by Azure Resource Manager, not Azure Data Factory.	Contact support.
Monitoring queries per minute	1,000	Contact support.
Entity CRUD operations per minute	50	Contact support.
Maximum time of data flow debug session	8 hrs	8 hrs
Concurrent number of data flows per factory	50	Contact support.
Concurrent number of data flow debug sessions per user per factory	3	3
Data Flow Azure IR TTL limit	4 hrs	Contact support.

[1] The data integration unit (DIU) is used in a cloud-to-cloud copy operation, learn more from Data integration units (version 2). For information on billing, see Azure Data Factory pricing.

[2] Azure Integration Runtime is globally available to ensure data compliance, efficiency, and reduced network egress costs.

REGION GROUP	REGIONS
Region group 1	Central US, East US, East US2, North Europe, West Europe, West US, West US 2
Region group 2	Australia East, Australia Southeast, Brazil South, Central India, Japan East, Northcentral US, Southcentral US, Southeast Asia, West Central US
Region group 3	Canada Central, East Asia, France Central, Korea Central, UK South

[3] Pipeline, data set, and linked service objects represent a logical grouping of your workload. Limits for these objects don't relate to the amount of data you can move and process with Azure Data Factory. Data Factory is designed to scale to handle petabytes of data.

Version 1

RESOURCE	DEFAULT LIMIT	MAXIMUM LIMIT
Pipelines within a data factory	2,500	Contact support.
Data sets within a data factory	5,000	Contact support.
Concurrent slices per data set	10	10
Bytes per object for pipeline objects[1]	200 KB	200 KB
Bytes per object for data set and linked service objects[1]	100 KB	2,000 KB
Azure HDInsight on-demand cluster cores within a subscription[2]	60	Contact support.
Cloud data movement units per copy activity run[3]	32	Contact support.
Retry count for pipeline activity runs	1,000	MaxInt (32 bit)

[1] Pipeline, data set, and linked service objects represent a logical grouping of your workload. Limits for these objects don't relate to the amount of data you can move and process with Azure Data Factory. Data Factory is designed to scale to handle petabytes of data.

[2] On-demand HDInsight cores are allocated out of the subscription that contains the data factory. As a result, the previous limit is the Data Factory-enforced core limit for on-demand HDInsight cores. It's different from the core limit that's associated with your Azure subscription.

[3] The cloud data movement unit (DMU) for version 1 is used in a cloud-to-cloud copy operation, learn more from Cloud data movement units (version 1). For information on billing, see Azure Data Factory pricing.

RESOURCE	DEFAULT LOWER LIMIT	MINIMUM LIMIT
Scheduling interval	15 minutes	15 minutes
Interval between retry attempts	1 second	1 second

RESOURCE	DEFAULT LOWER LIMIT	MINIMUM LIMIT
Retry timeout value	1 second	1 second

Web service call limits

Azure Resource Manager has limits for API calls. You can make API calls at a rate within the Azure Resource Manager API limits.

Data Lake Analytics limits

Azure Data Lake Analytics makes the complex task of managing distributed infrastructure and complex code easy. It dynamically provisions resources, and you can use it to do analytics on exabytes of data. When the job completes, it winds down resources automatically. You pay only for the processing power that was used. As you increase or decrease the size of data stored or the amount of compute used, you don't have to rewrite code. To raise the default limits for your subscription, contact support.

RESOURCE	DEFAULT LIMIT	COMMENTS
Maximum number of concurrent jobs	20	
Maximum number of analytics units (AUs) per account	250	Use any combination of up to a maximum of 250 AUs across 20 jobs. To increase this limit, contact Microsoft Support.
Maximum script size for job submission	3 MB	
Maximum number of Data Lake Analytics accounts per region per subscription	5	To increase this limit, contact Microsoft Support.

Data Lake Store limits

Azure Data Lake Storage Gen1 is an enterprise-wide hyper-scale repository for big data analytic workloads. You can use Data Lake Storage Gen1 to capture data of any size, type, and ingestion speed in one single place for operational and exploratory analytics. There's no limit to the amount of data you can store in a Data Lake Storage Gen1 account.

RESOURCE	DEFAULT LIMIT	COMMENTS
Maximum number of Data Lake Storage Gen1 accounts, per subscription, per region	10	To request an increase for this limit, contact support.
Maximum number of access ACLs, per file or folder	32	This is a hard limit. Use groups to manage access with fewer entries.
Maximum number of default ACLs, per file or folder	32	This is a hard limit. Use groups to manage access with fewer entries.

Data Share limits

Azure Data Share enables organizations to simply and securely share data with their customers and partners.

RESOURCE	LIMIT
Maximum number of Data Share resources per Azure subscription	50
Maximum number of sent shares per Data Share resource	100
Maximum number of received shares per Data Share resource	100
Maximum number of invitations per sent share	100
Maximum number of share subscriptions per sent share	100
Maximum number of datasets per share	100
Maximum number of snapshot schedules per share	1

Database Migration Service Limits

Azure Database Migration Service is a fully managed service designed to enable seamless migrations from multiple database sources to Azure data platforms with minimal downtime.

RESOURCE	DEFAULT LIMIT	COMMENTS
Maximum number of services per subscription, per region	2	To request an increase for this limit, contact support.

Event Grid limits

The following limits apply to Azure Event Grid system topics and custom topics, *not* event domains.

RESOURCE	LIMIT
Custom topics per Azure subscription	100
Event subscriptions per topic	500
Publish rate for a custom topic (ingress)	5,000 events per second per topic
Publish requests	250 per second
Event size	1 MB (charged in as multiple 64-KB events)

The following limits apply to event domains only.

RESOURCE	LIMIT
Topics per event domain	100,000
Event subscriptions per topic within a domain	500
Domain scope event subscriptions	50

RESOURCE	LIMIT
Publish rate for an event domain (ingress)	5,000 events per second
Publish requests	250 per second
Event Domains per Azure Subscription	100

Event Hubs limits

The following tables provide quotas and limits specific to Azure Event Hubs. For information about Event Hubs pricing, see Event Hubs pricing.

The following limits are common across basic, standard, and dedicated tiers.

LIMIT	SCOPE	NOTES	VALUE
Number of Event Hubs namespaces per subscription	Subscription	-	100
Number of event hubs per namespace	Namespace	Subsequent requests for creation of a new event hub are rejected.	10
Number of partitions per event hub	Entity	-	32
Maximum size of an event hub name	Entity	-	50 characters
Number of non-epoch receivers per consumer group	Entity	-	5
Maximum throughput units	Namespace	Exceeding the throughput unit limit causes your data to be throttled and generates a server busy exception. To request a larger number of throughput units for a Standard tier, file a support request. Additional throughput units are available in blocks of 20 on a committed purchase basis.	20
Number of authorization rules per namespace	Namespace	Subsequent requests for authorization rule creation are rejected.	12
Number of calls to the GetRuntimeInformation method	Entity	-	50 per second
Number of virtual network (VNet) and IP Config rules	Entity	-	128

Event Hubs Basic and Standard - quotas and limits

LIMIT	SCOPE	NOTES	BASIC	STANDARD
Maximum size of Event Hubs event	Entity		256 KB	1 MB
Number of consumer groups per event hub	Entity		1	20
Number of AMQP connections per namespace	Namespace	Subsequent requests for additional connections are rejected, and an exception is received by the calling code.	100	5,000
Maximum retention period of event data	Entity		1 day	1-7 days
Apache Kafka enabled namespace	Namespace	Event Hubs namespace streams applications using Kafka protocol	No	Yes
Capture	Entity	When enabled, micro-batches on the same stream	No	Yes

Event Hubs Dedicated - quotas and limits

The Event Hubs Dedicated offering is billed at a fixed monthly price, with a minimum of 4 hours of usage. The Dedicated tier offers all the features of the Standard plan, but with enterprise scale capacity and limits for customers with demanding workloads.

FEATURE	LIMITS
Bandwidth	20 CUs
Namespaces	50 per CU
Event Hubs	1000 per namespace
Ingress events	Included
Message Size	1 MB
Partitions	2000 per CU
Consumer groups	No limit per CU, 1000 per event hub
Brokered connections	100 K included
Message Retention	90 days, 10 TB included per CU
Capture	Included

Identity Manager limits

CATEGORY	LIMIT
User-assigned managed identities	• When you create user-assigned managed identities, only alphanumeric characters (0-9, a-z, and A-Z) and the hyphen (-) are supported. For the assignment to a virtual machine or virtual machine scale set to work properly, the name is limited to 24 characters. • If you use the managed identity virtual machine extension, the supported limit is 32 user-assigned managed identities. Without the managed identity virtual machine extension, the supported limit is 512 user-assigned identities.

IoT Central limits

IoT Central limits the number of applications you can deploy in a subscription to 10. If you need to increase this limit, contact Microsoft support.

IoT Hub limits

The following table lists the limits associated with the different service tiers S1, S2, S3, and F1. For information about the cost of each *unit* in each tier, see Azure IoT Hub pricing.

RESOURCE	S1 STANDARD	S2 STANDARD	S3 STANDARD	F1 FREE
Messages/day	400,000	6,000,000	300,000,000	8,000
Maximum units	200	200	10	1

> **NOTE**
> If you anticipate using more than 200 units with an S1 or S2 tier hub or 10 units with an S3 tier hub, contact Microsoft Support.

The following table lists the limits that apply to IoT Hub resources.

RESOURCE	LIMIT
Maximum paid IoT hubs per Azure subscription	100
Maximum free IoT hubs per Azure subscription	1
Maximum number of characters in a device ID	128
Maximum number of device identities returned in a single call	1,000
IoT Hub message maximum retention for device-to-cloud messages	7 days
Maximum size of device-to-cloud message	256 KB

RESOURCE	LIMIT
Maximum size of device-to-cloud batch	AMQP and HTTP: 256 KB for the entire batch MQTT: 256 KB for each message
Maximum messages in device-to-cloud batch	500
Maximum size of cloud-to-device message	64 KB
Maximum TTL for cloud-to-device messages	2 days
Maximum delivery count for cloud-to-device messages	100
Maximum cloud-to-device queue depth per device	50
Maximum delivery count for feedback messages in response to a cloud-to-device message	100
Maximum TTL for feedback messages in response to a cloud-to-device message	2 days
Maximum size of device twin	8 KB for tags section, and 32 KB for desired and reported properties sections each
Maximum length of device twin string key	1 KB
Maximum length of device twin string value	4 KB
Maximum depth of object in device twin	10
Maximum size of direct method payload	128 KB
Job history maximum retention	30 days
Maximum concurrent jobs	10 (for S3), 5 for (S2), 1 (for S1)
Maximum additional endpoints	10 (for S1, S2, and S3)
Maximum message routing rules	100 (for S1, S2, and S3)
Maximum number of concurrently connected device streams	50 (for S1, S2, S3, and F1 only)
Maximum device stream data transfer	300 MB per day (for S1, S2, S3, and F1 only)

NOTE

If you need more than 100 paid IoT hubs in an Azure subscription, contact Microsoft Support.

> **NOTE**
> Currently, the total number of devices plus modules that can be registered to a single IoT hub is capped at 1,000,000. If you want to increase this limit, contact Microsoft Support.

IoT Hub throttles requests when the following quotas are exceeded.

THROTTLE	PER-HUB VALUE
Identity registry operations (create, retrieve, list, update, and delete), individual or bulk import/export	83.33/sec/unit (5,000/min/unit) (for S3). 1.67/sec/unit (100/min/unit) (for S1 and S2).
Device connections	6,000/sec/unit (for S3), 120/sec/unit (for S2), 12/sec/unit (for S1). Minimum of 100/sec.
Device-to-cloud sends	6,000/sec/unit (for S3), 120/sec/unit (for S2), 12/sec/unit (for S1). Minimum of 100/sec.
Cloud-to-device sends	83.33/sec/unit (5,000/min/unit) (for S3), 1.67/sec/unit (100/min/unit) (for S1 and S2).
Cloud-to-device receives	833.33/sec/unit (50,000/min/unit) (for S3), 16.67/sec/unit (1,000/min/unit) (for S1 and S2).
File upload operations	83.33 file upload initiations/sec/unit (5,000/min/unit) (for S3), 1.67 file upload initiations/sec/unit (100/min/unit) (for S1 and S2). 10,000 SAS URIs can be out for an Azure Storage account at one time. 10 SAS URIs/device can be out at one time.
Direct methods	24 MB/sec/unit (for S3), 480 KB/sec/unit (for S2), 160 KB/sec/unit (for S1). Based on 8-KB throttling meter size.
Device twin reads	500/sec/unit (for S3), Maximum of 100/sec or 10/sec/unit (for S2), 100/sec (for S1)
Device twin updates	250/sec/unit (for S3), Maximum of 50/sec or 5/sec/unit (for S2), 50/sec (for S1)
Jobs operations (create, update, list, and delete)	83.33/sec/unit (5,000/min/unit) (for S3), 1.67/sec/unit (100/min/unit) (for S2), 1.67/sec/unit (100/min/unit) (for S1).
Jobs per-device operation throughput	50/sec/unit (for S3), maximum of 10/sec or 1/sec/unit (for S2), 10/sec (for S1).
Device stream initiation rate	5 new streams/sec (for S1, S2, S3, and F1 only).

IoT Hub Device Provisioning Service limits

The following table lists the limits that apply to Azure IoT Hub Device Provisioning Service resources.

RESOURCE	LIMIT
Maximum device provisioning services per Azure subscription	10
Maximum number of enrollments	1,000,000
Maximum number of registrations	1,000,000
Maximum number of enrollment groups	100
Maximum number of CAs	25
Maximum number of linked IoT hubs	50
Maximum size of message	96 KB

NOTE

To increase the number of enrollments and registrations on your provisioning service, contact Microsoft Support.

NOTE

Increasing the maximum number of CAs is not supported.

The Device Provisioning Service throttles requests when the following quotas are exceeded.

THROTTLE	PER-UNIT VALUE
Operations	200/min/service
Device registrations	200/min/service
Device polling operation	5/10 sec/device

Key Vault limits

Key transactions (maximum transactions allowed in 10 seconds, per vault per region[1]):

KEY TYPE	HSM KEY CREATE KEY	HSM KEY ALL OTHER TRANSACTIONS	SOFTWARE KEY CREATE KEY	SOFTWARE KEY ALL OTHER TRANSACTIONS
RSA 2,048-bit	5	1,000	10	2,000
RSA 3,072-bit	5	250	10	500
RSA 4,096-bit	5	125	10	250
ECC P-256	5	1,000	10	2,000
ECC P-384	5	1,000	10	2,000

KEY TYPE	HSM KEY CREATE KEY	HSM KEY ALL OTHER TRANSACTIONS	SOFTWARE KEY CREATE KEY	SOFTWARE KEY ALL OTHER TRANSACTIONS
ECC P-521	5	1,000	10	2,000
ECC SECP256K1	5	1,000	10	2,000

NOTE

In the previous table, we see that for RSA 2,048-bit software keys, 2,000 GET transactions per 10 seconds are allowed. For RSA 2,048-bit HSM-keys, 1,000 GET transactions per 10 seconds are allowed.

The throttling thresholds are weighted, and enforcement is on their sum. For example, as shown in the previous table, when you perform GET operations on RSA HSM-keys, it's eight times more expensive to use 4,096-bit keys compared to 2,048-bit keys. That's because 1,000/125 = 8.

In a given 10-second interval, an Azure Key Vault client can do *only one* of the following operations before it encounters a `429` throttling HTTP status code:

- 2,000 RSA 2,048-bit software-key GET transactions
- 1,000 RSA 2,048-bit HSM-key GET transactions
- 125 RSA 4,096-bit HSM-key GET transactions
- 124 RSA 4,096-bit HSM-key GET transactions and 8 RSA 2,048-bit HSM-key GET transactions

Secrets, managed storage account keys, and vault transactions:

TRANSACTIONS TYPE	MAXIMUM TRANSACTIONS ALLOWED IN 10 SECONDS, PER VAULT PER REGION[1]
All transactions	2,000

For information on how to handle throttling when these limits are exceeded, see Azure Key Vault throttling guidance.

[1] A subscription-wide limit for all transaction types is five times per key vault limit. For example, HSM-other transactions per subscription are limited to 5,000 transactions in 10 seconds per subscription.

Media Services limits

NOTE

For resources that aren't fixed, open a support ticket to ask for an increase in the quotas. Don't create additional Azure Media Services accounts in an attempt to obtain higher limits.

RESOURCE	DEFAULT LIMIT
Azure Media Services accounts in a single subscription	25 (fixed)
Media reserved units per Media Services account	25 (S1) 10 (S2, S3)[1]
Jobs per Media Services account	50,000[2]
Chained tasks per job	30 (fixed)

RESOURCE	DEFAULT LIMIT
Assets per Media Services account	1,000,000
Assets per task	50
Assets per job	100
Unique locators associated with an asset at one time	5[4]
Live channels per Media Services account	5
Programs in stopped state per channel	50
Programs in running state per channel	3
Streaming endpoints that are stopped or running per Media Services account	2
Streaming units per streaming endpoint	10
Storage accounts	1,000[5] (fixed)
Policies	1,000,000[6]
File size	In some scenarios, there's a limit on the maximum file size supported for processing in Media Services.[7]

[1]If you change the type, for example, from S2 to S1, the maximum reserved unit limits are reset.

[2]This number includes queued, finished, active, and canceled jobs. It doesn't include deleted jobs. You can delete old jobs by using **IJob.Delete** or the **DELETE** HTTP request.

As of April 1, 2017, any job record in your account older than 90 days is automatically deleted, along with its associated task records. Automatic deletion occurs even if the total number of records is below the maximum quota. To archive the job and task information, use the code described in Manage assets with the Media Services .NET SDK.

[3]When you make a request to list job entities, a maximum of 1,000 jobs is returned per request. To keep track of all submitted jobs, use the top or skip queries as described in OData system query options.

[4]Locators aren't designed for managing per-user access control. To give different access rights to individual users, use digital rights management (DRM) solutions. For more information, see Protect your content with Azure Media Services.

[5]The storage accounts must be from the same Azure subscription.

[6]There's a limit of 1,000,000 policies for different Media Services policies. An example is for the Locator policy or ContentKeyAuthorizationPolicy.

> **NOTE**
> If you always use the same days and access permissions, use the same policy ID. For information and an example, see Manage assets with the Media Services .NET SDK.

[7]The maximum size supported for a single blob is currently up to 5 TB in Azure Blob Storage. Additional limits apply in Media Services based on the VM sizes that are used by the service. The size limit applies to the files that you upload and also the files that get generated as a result of Media Services processing (encoding or analyzing). If your source file is larger than 260-GB, your Job will likely fail.

The following table shows the limits on the media reserved units S1, S2, and S3. If your source file is larger than the limits defined in the table, your encoding job fails. If you encode 4K resolution sources of long duration, you're required to use S3 media reserved units to achieve the performance needed. If you have 4K content that's larger than the 260-GB limit on the S3 media reserved units, open a support ticket.

MEDIA RESERVED UNIT TYPE	MAXIMUM INPUT SIZE (GB)
S1	26
S2	60
S3	260

Mobile Services limits

TIER	FREE	BASIC	STANDARD
API calls	500,000	1.5 million per unit	15 million per unit
Active devices	500	Unlimited	Unlimited
Scale	N/A	Up to 6 units	Unlimited units
Push notifications	Azure Notification Hubs Free tier included, up to 1 million pushes	Notification Hubs Basic tier included, up to 10 million pushes	Notification Hubs Standard tier included, up to 10 million pushes
Real-time messaging/ Web Sockets	Limited	350 per mobile service	Unlimited
Offline synchronizations	Limited	Included	Included
Scheduled jobs	Limited	Included	Included
Azure SQL Database (required) Standard rates apply for additional capacity	20 MB included	20 MB included	20 MB included
CPU capacity	60 minutes per day	Unlimited	Unlimited
Outbound data transfer	165 MB per day (daily rollover)	Included	Included

For more information on limits and pricing, see Azure Mobile Services pricing.

Multi-Factor Authentication limits

RESOURCE	DEFAULT LIMIT	MAXIMUM LIMIT
Maximum number of trusted IP addresses or ranges per subscription	0	50
Remember my devices, number of days	14	60
Maximum number of app passwords	0	No limit
Allow **X** attempts during MFA call	1	99
Two-way text message timeout seconds	60	600
Default one-time bypass seconds	300	1,800
Lock user account after **X** consecutive MFA denials	Not set	99
Reset account lockout counter after **X** minutes	Not set	9,999
Unlock account after **X** minutes	Not set	9,999

Networking limits

Networking limits - Azure Resource Manager The following limits apply only for networking resources managed through **Azure Resource Manager** per region per subscription. Learn how to view your current resource usage against your subscription limits.

> **NOTE**
> We recently increased all default limits to their maximum limits. If there's no maximum limit column, the resource doesn't have adjustable limits. If you had these limits increased by support in the past and don't see updated limits in the following tables, open an online customer support request at no charge

RESOURCE	DEFAULT/MAXIMUM LIMIT
Virtual networks	1,000
Subnets per virtual network	3,000
Virtual network peerings per virtual network	500
Virtual network gateways (VPN gateways) per virtual network	1
Virtual network gateways (ExpressRoute gateways) per virtual network	1
DNS servers per virtual network	20
Private IP addresses per virtual network	65,536

RESOURCE	DEFAULT/MAXIMUM LIMIT
Private IP addresses per network interface	256
Private IP addresses per virtual machine	256
Public IP addresses per network interface	256
Public IP addresses per virtual machine	256
Concurrent TCP or UDP flows per NIC of a virtual machine or role instance	500,000
Network interface cards	65,536
Network Security Groups	5,000
NSG rules per NSG	1,000
IP addresses and ranges specified for source or destination in a security group	4,000
Application security groups	3,000
Application security groups per IP configuration, per NIC	20
IP configurations per application security group	4,000
Application security groups that can be specified within all security rules of a network security group	100
User-defined route tables	200
User-defined routes per route table	400
Point-to-site root certificates per Azure VPN Gateway	20
Virtual network TAPs	100
Network interface TAP configurations per virtual network TAP	100

Public IP address limits

RESOURCE	DEFAULT LIMIT	MAXIMUM LIMIT
Public IP addresses[1]	10 for Basic.	Contact support.
Static Public IP addresses[1]	10 for Basic.	Contact support.
Standard Public IP addresses[1]	10	Contact support.
Public IP Prefixes	limited by number of Standard Public IPs in a subscription	Contact support.

RESOURCE	DEFAULT LIMIT	MAXIMUM LIMIT
Public IP prefix length	/28	Contact support.

[1]Default limits for Public IP addresses vary by offer category type, such as Free Trial, Pay-As-You-Go, CSP. For example, the default for Enterprise Agreement subscriptions is 1000.

Load balancer limits

The following limits apply only for networking resources managed through Azure Resource Manager per region per subscription. Learn how to view your current resource usage against your subscription limits.

Standard Load Balancer

RESOURCE	DEFAULT/MAXIMUM LIMIT
Load balancers	1,000
Rules per resource	1,500
Rules per NIC (across all IPs on a NIC)	300
Frontend IP configurations	600
Backend pool size	1,000 IP configurations, single virtual network
High-availability ports	1 per internal frontend
Outbound rules per Load Balancer	20

Basic Load Balancer

RESOURCE	DEFAULT/MAXIMUM LIMIT
Load balancers	1,000
Rules per resource	250
Rules per NIC (across all IPs on a NIC)	300
Frontend IP configurations	200
Backend pool size	300 IP configurations, single availability set
Availability sets per Load Balancer	150

The following limits apply only for networking resources managed through the classic deployment model per subscription. Learn how to view your current resource usage against your subscription limits.

RESOURCE	DEFAULT LIMIT	MAXIMUM LIMIT
Virtual networks	100	100
Local network sites	20	50

RESOURCE	DEFAULT LIMIT	MAXIMUM LIMIT
DNS servers per virtual network	20	20
Private IP addresses per virtual network	4,096	4,096
Concurrent TCP or UDP flows per NIC of a virtual machine or role instance	500,000, up to 1,000,000 for two or more NICs.	500,000, up to 1,000,000 for two or more NICs.
Network Security Groups (NSGs)	200	200
NSG rules per NSG	1,000	1,000
User-defined route tables	200	200
User-defined routes per route table	400	400
Public IP addresses (dynamic)	500	500
Reserved public IP addresses	500	500
Public VIP per deployment	5	Contact support
Private VIP (internal load balancing) per deployment	1	1
Endpoint access control lists (ACLs)	50	50

ExpressRoute limits

RESOURCE	DEFAULT/MAXIMUM LIMIT
ExpressRoute circuits per subscription	10
ExpressRoute circuits per region per subscription, with Azure Resource Manager	10
Maximum number of routes advertised to Azure private peering with ExpressRoute Standard	4,000
Maximum number of routes advertised to Azure private peering with ExpressRoute Premium add-on	10,000
Maximum number of routes advertised from Azure private peering from the VNet address space for an ExpressRoute connection	200
Maximum number of routes advertised to Microsoft peering with ExpressRoute Standard	200
Maximum number of routes advertised to Microsoft peering with ExpressRoute Premium add-on	200

RESOURCE	DEFAULT/MAXIMUM LIMIT
Maximum number of ExpressRoute circuits linked to the same virtual network in the same peering location	4
Maximum number of ExpressRoute circuits linked to the same virtual network in different peering locations	4
Number of virtual network links allowed per ExpressRoute circuit	See the Number of virtual networks per ExpressRoute circuit table.

Number of virtual networks per ExpressRoute circuit

CIRCUIT SIZE	NUMBER OF VIRTUAL NETWORK LINKS FOR STANDARD	NUMBER OF VIRTUAL NETWORK LINKS WITH PREMIUM ADD-ON
50 Mbps	10	20
100 Mbps	10	25
200 Mbps	10	25
500 Mbps	10	40
1 Gbps	10	50
2 Gbps	10	60
5 Gbps	10	75
10 Gbps	10	100
40 Gbps*	10	100
100 Gbps*	10	100

**100 Gbps ExpressRoute Direct Only*

> **NOTE**
> Global Reach connections count against the limit of virtual network connections per ExpressRoute Circuit. For example, a 10 Gbps Premium Circuit would allow for 5 Global Reach connections and 95 connections to the ExpressRoute Gateways or 95 Global Reach connections and 5 connections to the ExpressRoute Gateways or any other combination up to the limit of 100 connections for the circuit.

Virtual WAN limits

RESOURCE	LIMIT
Virtual WAN hubs per region	1
Virtual WAN hubs per virtual wan	Azure regions
VPN (branch) connections per hub	1,000

RESOURCE	LIMIT
VNet connections per hub	500
Point-to-Site users per hub	10,000
Aggregate throughput per Virtual WAN VPN gateway	20 Gbps
Throughput per Virtual WAN VPN connection (2 tunnels)	2 Gbps with 1 Gbps/IPsec tunnel
Aggregate throughput per Virtual WAN ExpressRoute gateway	20 Gbps

Application Gateway limits

The following table applies to v1, v2, Standard, and WAF SKUs unless otherwise stated.

RESOURCE	DEFAULT/MAXIMUM LIMIT	NOTE
Azure Application Gateway	1,000 per subscription	
Front-end IP configurations	2	1 public and 1 private
Front-end ports	100[1]	
Back-end address pools	100[1]	
Back-end servers per pool	1,200	
HTTP listeners	100[1]	
HTTP load-balancing rules	100[1]	
Back-end HTTP settings	100[1]	
Instances per gateway	V1 SKU - 32 V2 SKU - 125	
SSL certificates	100[1]	1 per HTTP listener
Maximum SSL certificate size	V1 SKU - 10 KB V2 SKU - 16 KB	
Authentication certificates	100	
Trusted root certificates	100	
Request timeout minimum	1 second	
Request timeout maximum	24 hours	
Number of sites	100[1]	1 per HTTP listener
URL maps per listener	1	

RESOURCE	DEFAULT/MAXIMUM LIMIT	NOTE
Maximum path-based rules per URL map	100	
Redirect configurations	100[1]	
Concurrent WebSocket connections	Medium gateways 20k Large gateways 50k	
Maximum URL length	32KB	
Maximum header size for HTTP/2	4KB	
Maximum file upload size, Standard	2 GB	
Maximum file upload size WAF	V1 Medium WAF gateways, 100 MB V1 Large WAF gateways, 500 MB V2 WAF, 750 MB	
WAF body size limit, without files	128 KB	
Maximum WAF custom rules	100	
Maximum WAF exclusions	100	

[1] In case of WAF-enabled SKUs, we recommend that you limit the number of resources to 40 for optimal performance.

Network Watcher limits

RESOURCE	DEFAULT LIMIT	MAXIMUM LIMIT	NOTE
Azure Network Watcher	1 per region	1 per region	Network Watcher is created to enable access to the service. Only one instance of Network Watcher is required per subscription per region.
Packet capture sessions	10,000 per region	10,000	Number of sessions only, not saved captures.

Private Link limits

The following limits apply to Azure private link:

RESOURCE	LIMIT
Number of private endpoints per virtual network	1000
Number of private endpoints per subscription	64000
Number of private link service per subscription	800
Number of IP Configurations on a private link service	8 (This number is for the NAT IP addresses used per PLS)

RESOURCE	LIMIT
Number of private endpoints on the same private link service	1000

Traffic Manager limits

RESOURCE	DEFAULT/MAXIMUM LIMIT
Profiles per subscription	200
Endpoints per profile	200

Azure Bastion limits

RESOURCE	DEFAULT LIMIT
Concurrent RDP connections	25*
Concurrent SSH connections	More than 50**

*May vary due to other on-going RDP sessions or other on-going SSH sessions.
**May vary if there are existing RDP connections or usage from other on-going SSH sessions.

Azure DNS limits

Public DNS zones

RESOURCE	DEFAULT LIMIT
Public DNS Zones per subscription	250 [1]
Record sets per public DNS zone	10,000 [1]
Records per record set in public DNS zone	20
Number of Alias records for a single Azure resource	20
Private DNS zones per subscription	1000
Record sets per private DNS zone	25000
Records per record set for private DNS zones	20
Virtual Network Links per private DNS zone	1000
Virtual Networks Links per private DNS zones with auto-registration enabled	100
Number of private DNS zones a virtual network can get linked to with auto-registration enabled	1
Number of private DNS zones a virtual network can get linked	1000

RESOURCE	DEFAULT LIMIT
Number of DNS queries a virtual machine can send to Azure DNS resolver, per second	500 [2]
Maximum number of DNS queries queued (pending response) per virtual machine	200 [2]

[1]If you need to increase these limits, contact Azure Support.

[2]These limits are applied to every individual virtual machine and not at the virtual network level. DNS queries exceeding these limits are dropped.

Azure Firewall limits

RESOURCE	DEFAULT LIMIT
Data throughput	30 Gbps[1]
Rules	10,000. All rule types combined.
Maximum DNAT rules	299
Minimum AzureFirewallSubnet size	/26
Port range in network and application rules	0-64,000. Work is in progress to relax this limitation.
Public IP addresses	100 maximum (Currently, SNAT ports are added only for the first five public IP addresses.)
Route table	By default, AzureFirewallSubnet has a 0.0.0.0/0 route with the NextHopType value set to **Internet**. Azure Firewall must have direct Internet connectivity. If your AzureFirewallSubnet learns a default route to your on-premises network via BGP, you must override that with a 0.0.0.0/0 UDR with the **NextHopType** value set as **Internet** to maintain direct Internet connectivity. By default, Azure Firewall doesn't support forced tunneling to an on-premises network. However, if your configuration requires forced tunneling to an on-premises network, Microsoft will support it on a case by case basis. Contact Support so that we can review your case. If accepted, we'll allow your subscription and ensure the required firewall Internet connectivity is maintained.

[1]If you need to increase these limits, contact Azure Support.

Azure Front Door Service limits

RESOURCE	DEFAULT/MAXIMUM LIMIT
Azure Front Door Service resources per subscription	100
Front-end hosts, which includes custom domains per resource	100

RESOURCE	DEFAULT/MAXIMUM LIMIT
Routing rules per resource	100
Back-end pools per resource	50
Back ends per back-end pool	100
Path patterns to match for a routing rule	25
Custom web application firewall rules per policy	10
Web application firewall policy per subscription	100
Web application firewall match conditions per custom rule	10
Web application firewall IP address ranges per match condition	600
Web application firewall string match values per match condition	10
Web application firewall string match value length	256
Web application firewall POST body parameter name length	256
Web application firewall HTTP header name length	256
Web application firewall cookie name length	256
Web application firewall HTTP request body size inspected	128 KB
Web application firewall custom response body length	2 KB

Timeout values

Client to Front Door

- Front Door has an idle TCP connection timeout of 61 seconds.

Front Door to application back-end

- If the response is a chunked response, a 200 is returned if or when the first chunk is received.
- After the HTTP request is forwarded to the back end, Front Door waits for 30 seconds for the first packet from the back end. Then it returns a 503 error to the client.
- After the first packet is received from the back end, Front Door waits for 30 seconds in an idle timeout. Then it returns a 503 error to the client.
- Front Door to the back-end TCP session timeout is 30 minutes.

Upload and download data limit

	WITH CHUNKED TRANSFER ENCODING (CTE)	WITHOUT HTTP CHUNKING
Download	There's no limit on the download size.	There's no limit on the download size.

	WITH CHUNKED TRANSFER ENCODING (CTE)	WITHOUT HTTP CHUNKING
Upload	There's no limit as long as each CTE upload is less than 2 GB.	The size can't be larger than 2 GB.

Other limits

- Maximum URL size - 8,192 bytes - Specifies maximum length of the raw URL (scheme + hostname + port + path + query string of the URL)
- Maximum Query String size - 4,096 bytes - Specifies the maximum length of the query string, in bytes.

Notification Hubs limits

TIER	FREE	BASIC	STANDARD
Included pushes	1 million	10 million	10 million
Active devices	500	200,000	10 million
Tag quota per installation or registration	60	60	60

For more information on limits and pricing, see Notification Hubs pricing.

Role-based access control limits

RESOURCE	LIMIT
Role assignments for Azure resources per Azure subscription	2,000
Role assignments for Azure resources per management group	500
Custom roles for Azure resources per tenant	5,000
Custom roles for Azure resources per tenant (specialized clouds, such as Azure Government, Azure Germany, and Azure China 21Vianet)	2,000

Service Bus limits

The following table lists quota information specific to Azure Service Bus messaging. For information about pricing and other quotas for Service Bus, see Service Bus pricing.

QUOTA NAME	SCOPE	NOTES	VALUE
Maximum number of Basic or Standard namespaces per Azure subscription	Namespace	Subsequent requests for additional Basic or Standard namespaces are rejected by the Azure portal.	100

QUOTA NAME	SCOPE	NOTES	VALUE
Maximum number of Premium namespaces per Azure subscription	Namespace	Subsequent requests for additional Premium namespaces are rejected by the portal.	100
Queue or topic size	Entity	Defined upon creation of the queue or topic. Subsequent incoming messages are rejected, and an exception is received by the calling code.	1, 2, 3, 4 GB or 5 GB. In the Premium SKU, and the Standard SKU with partitioning enabled, the maximum queue or topic size is 80 GB.
Number of concurrent connections on a namespace	Namespace	Subsequent requests for additional connections are rejected, and an exception is received by the calling code. REST operations don't count toward concurrent TCP connections.	NetMessaging: 1,000. AMQP: 5,000.
Number of concurrent receive requests on a queue, topic, or subscription entity	Entity	Subsequent receive requests are rejected, and an exception is received by the calling code. This quota applies to the combined number of concurrent receive operations across all subscriptions on a topic.	5,000
Number of topics or queues per namespace	Namespace	Subsequent requests for creation of a new topic or queue on the namespace are rejected. As a result, if configured through the Azure portal, an error message is generated. If called from the management API, an exception is received by the calling code.	10,000 for the Basic or Standard tier. The total number of topics and queues in a namespace must be less than or equal to 10,000. For the Premium tier, 1,000 per messaging unit (MU). Maximum limit is 4,000.
Number of partitioned topics or queues per namespace	Namespace	Subsequent requests for creation of a new partitioned topic or queue on the namespace are rejected. As a result, if configured through the Azure portal, an error message is generated. If called from the management API, the exception **QuotaExceededException** is received by the calling code.	Basic and Standard tiers: 100. Partitioned entities aren't supported in the Premium tier. Each partitioned queue or topic counts toward the quota of 1,000 entities per namespace.
Maximum size of any messaging entity path: queue or topic	Entity	-	260 characters.

QUOTA NAME	SCOPE	NOTES	VALUE
Maximum size of any messaging entity name: namespace, subscription, or subscription rule	Entity	-	50 characters.
Maximum size of a message ID	Entity	-	128
Maximum size of a message session ID	Entity	-	128
Message size for a queue, topic, or subscription entity	Entity	Incoming messages that exceed these quotas are rejected, and an exception is received by the calling code.	Maximum message size: 256 KB for Standard tier, 1 MB for Premium tier. Due to system overhead, this limit is less than these values. Maximum header size: 64 KB. Maximum number of header properties in property bag: **byte/int.MaxValue**. Maximum size of property in property bag: No explicit limit. Limited by maximum header size.
Message property size for a queue, topic, or subscription entity	Entity	The exception **SerializationException** is generated.	Maximum message property size for each property is 32,000. Cumulative size of all properties can't exceed 64,000. This limit applies to the entire header of the BrokeredMessage, which has both user properties and system properties, such as SequenceNumber, Label, and MessageId.
Number of subscriptions per topic	Entity	Subsequent requests for creating additional subscriptions for the topic are rejected. As a result, if configured through the portal, an error message is shown. If called from the management API, an exception is received by the calling code.	2,000 per-topic for the Standard tier.
Number of SQL filters per topic	Entity	Subsequent requests for creation of additional filters on the topic are rejected, and an exception is received by the calling code.	2,000

QUOTA NAME	SCOPE	NOTES	VALUE
Number of correlation filters per topic	Entity	Subsequent requests for creation of additional filters on the topic are rejected, and an exception is received by the calling code.	100,000
Size of SQL filters or actions	Namespace	Subsequent requests for creation of additional filters are rejected, and an exception is received by the calling code.	Maximum length of filter condition string: 1,024 (1 K). Maximum length of rule action string: 1,024 (1 K). Maximum number of expressions per rule action: 32.
Number of SharedAccessAuthorizationRule rules per namespace, queue, or topic	Entity, namespace	Subsequent requests for creation of additional rules are rejected, and an exception is received by the calling code.	Maximum number of rules per entity type: 12. Rules that are configured on a Service Bus namespace apply to all types: queues, topics.
Number of messages per transaction	Transaction	Additional incoming messages are rejected, and an exception stating "Cannot send more than 100 messages in a single transaction" is received by the calling code.	100 For both **Send()** and **SendAsync()** operations.
Number of virtual network and IP filter rules	Namespace		128

Site Recovery limits

The following limits apply to Azure Site Recovery.

LIMIT IDENTIFIER	DEFAULT LIMIT
Number of vaults per subscription	500
Number of servers per Azure vault	250
Number of protection groups per Azure vault	No limit
Number of recovery plans per Azure vault	No limit
Number of servers per protection group	No limit
Number of servers per recovery plan	50

SQL Database limits

For SQL Database limits, see SQL Database resource limits for single databases, SQL Database resource limits for elastic pools and pooled databases, and SQL Database resource limits for managed instances.

SQL Data Warehouse limits

For SQL Data Warehouse limits, see SQL Data Warehouse resource limits.

Storage limits

The following table describes default limits for Azure general-purpose v1, v2, and Blob storage accounts. The *ingress* limit refers to all data from requests that are sent to a storage account. The *egress* limit refers to all data from responses that are received from a storage account.

RESOURCE	DEFAULT LIMIT
Number of storage accounts per region per subscription, including both standard and premium accounts	250
Maximum storage account capacity	2 PiB for US and Europe, and 500 TiB for all other regions (including the UK)[1]
Maximum number of blob containers, blobs, file shares, tables, queues, entities, or messages per storage account	No limit
Maximum request rate[1] per storage account	20,000 requests per second
Maximum ingress[1] per storage account (US, Europe regions)	25 Gbps
Maximum ingress[1] per storage account (regions other than US and Europe)	5 Gbps if RA-GRS/GRS is enabled, 10 Gbps for LRS/ZRS[2]
Maximum egress for general-purpose v2 and Blob storage accounts (all regions)	50 Gbps
Maximum egress for general-purpose v1 storage accounts (US regions)	20 Gbps if RA-GRS/GRS is enabled, 30 Gbps for LRS/ZRS[2]
Maximum egress for general-purpose v1 storage accounts (non-US regions)	10 Gbps if RA-GRS/GRS is enabled, 15 Gbps for LRS/ZRS[2]
Maximum number of virtual network rules per storage account	200
Maximum number of IP address rules per storage account	200

[1]Azure Storage standard accounts support higher capacity limits and higher limits for ingress by request. To request an increase in account limits for ingress, contact Azure Support. For more information, see Announcing larger, higher scale storage accounts.

[2] If your storage account has read-access enabled with geo-redundant storage (RA-GRS) or geo-zone-redundant storage (RA-GZRS), then the egress targets for the secondary location are identical to those of the primary location. Azure Storage replication options include:

- Locally redundant storage (LRS)
- Zone-redundant storage (ZRS)

- Geo-redundant storage (GRS)
- Read-access geo-redundant storage (RA-GRS)
- Geo-zone-redundant storage (GZRS)
- Read-access geo-zone-redundant storage (RA-GZRS)

> **NOTE**
> Microsoft recommends that you use a general-purpose v2 storage account for most scenarios. You can easily upgrade a general-purpose v1 or an Azure Blob storage account to a general-purpose v2 account with no downtime and without the need to copy data. For more information, see Upgrade to a general-purpose v2 storage account.

If the needs of your application exceed the scalability targets of a single storage account, you can build your application to use multiple storage accounts. You can then partition your data objects across those storage accounts. For information on volume pricing, see Azure Storage pricing.

All storage accounts run on a flat network topology and support the scalability and performance targets outlined in this article, regardless of when they were created. For more information on the Azure Storage flat network architecture and on scalability, see Microsoft Azure Storage: A Highly Available Cloud Storage Service with Strong Consistency.

For more information on limits for standard storage accounts, see Scalability targets for standard storage accounts.

Storage resource provider limits

The following limits apply only when you perform management operations by using Azure Resource Manager with Azure Storage.

RESOURCE	DEFAULT LIMIT
Storage account management operations (read)	800 per 5 minutes
Storage account management operations (write)	1200 per hour
Storage account management operations (list)	100 per 5 minutes

Azure Blob storage limits

RESOURCE	TARGET
Maximum size of single blob container	Same as maximum storage account capacity
Maximum number of blocks in a block blob or append blob	50,000 blocks
Maximum size of a block in a block blob	100 MiB
Maximum size of a block blob	50,000 X 100 MiB (approximately 4.75 TiB)
Maximum size of a block in an append blob	4 MiB
Maximum size of an append blob	50,000 x 4 MiB (approximately 195 GiB)
Maximum size of a page blob	8 TiB
Maximum number of stored access policies per blob container	5

RESOURCE	TARGET
Target request rate for a single blob	Up to 500 requests per second
Target throughput for a single page blob	Up to 60 MiB per second
Target throughput for a single block blob	Up to storage account ingress/egress limits[1]

[1] Throughput for a single blob depends on several factors, including, but not limited to: concurrency, request size, performance tier, speed of source for uploads, and destination for downloads. To take advantage of the performance enhancements of high-throughput block blobs, upload larger blobs or blocks. Specifically, call the Put Blob or Put Block operation with a blob or block size that is greater than 4 MiB for standard storage accounts. For premium block blob or for Data Lake Storage Gen2 storage accounts, use a block or blob size that is greater than 256 KiB.

Azure Files limits

For more information on Azure Files limits, see Azure Files scalability and performance targets.

RESOURCE	STANDARD FILE SHARES	PREMIUM FILE SHARES
Minimum size of a file share	No minimum; pay as you go	100 GiB; provisioned
Maximum size of a file share	100 TiB*, 5 TiB	100 TiB
Maximum size of a file in a file share	1 TiB	1 TiB
Maximum number of files in a file share	No limit	No limit
Maximum IOPS per share	10,000 IOPS*, 1,000 IOPS	100,000 IOPS
Maximum number of stored access policies per file share	5	5
Target throughput for a single file share	up to 300 MiB/sec*, Up to 60 MiB/sec ,	See premium file share ingress and egress values
Maximum egress for a single file share	See standard file share target throughput	Up to 6,204 MiB/s
Maximum ingress for a single file share	See standard file share target throughput	Up to 4,136 MiB/s
Maximum open handles per file	2,000 open handles	2,000 open handles
Maximum number of share snapshots	200 share snapshots	200 share snapshots
Maximum object (directories and files) name length	2,048 characters	2,048 characters
Maximum pathname component (in the path \A\B\C\D, each letter is a component)	255 characters	255 characters

* Available in most regions, see Regional availability for the details on available regions.

Azure File Sync limits

RESOURCE	TARGET	HARD LIMIT
Storage Sync Services per region	20 Storage Sync Services	Yes
Sync groups per Storage Sync Service	100 sync groups	Yes
Registered servers per Storage Sync Service	99 servers	Yes
Cloud endpoints per sync group	1 cloud endpoint	Yes
Server endpoints per sync group	50 server endpoints	No
Server endpoints per server	30 server endpoints	Yes
File system objects (directories and files) per sync group	100 million objects	No
Maximum number of file system objects (directories and files) in a directory	5 million objects	Yes
Maximum object (directories and files) security descriptor size	64 KiB	Yes
File size	100 GiB	No
Minimum file size for a file to be tiered	V9: Based on file system cluster size (double file system cluster size). For example, if the file system cluster size is 4kb, the minimum file size will be 8kb. V8 and older: 64 KiB	Yes

NOTE

An Azure File Sync endpoint can scale up to the size of an Azure file share. If the Azure file share size limit is reached, sync will not be able to operate.

Azure Queue storage limits

RESOURCE	TARGET
Maximum size of a single queue	500 TiB
Maximum size of a message in a queue	64 KiB
Maximum number of stored access policies per queue	5
Maximum request rate per storage account	20,000 messages per second, which assumes a 1-KiB message size
Target throughput for a single queue (1-KiB messages)	Up to 2,000 messages per second

Azure Table storage limits

RESOURCE	TARGET
Maximum size of a single table	500 TiB
Maximum size of a table entity	1 MiB
Maximum number of properties in a table entity	255, which includes three system properties: PartitionKey, RowKey, and Timestamp
Maximum total size of property values in an entity	1 MiB
Maximum total size of an individual property in an entity	Varies by property type. For more information, see **Property Types** in Understanding the Table Service Data Model.
Maximum number of stored access policies per table	5
Maximum request rate per storage account	20,000 transactions per second, which assumes a 1-KiB entity size
Target throughput for a single table partition (1 KiB-entities)	Up to 2,000 entities per second

Virtual machine disk limits

You can attach a number of data disks to an Azure virtual machine. Based on the scalability and performance targets for a VM's data disks, you can determine the number and type of disk that you need to meet your performance and capacity requirements.

> **IMPORTANT**
> For optimal performance, limit the number of highly utilized disks attached to the virtual machine to avoid possible throttling. If all attached disks aren't highly utilized at the same time, the virtual machine can support a larger number of disks.

For Azure managed disks:

The following table illustrates the default and maximum limits of the number of resources per region per subscription. There is no limit for the number of Managed Disks, snapshots and images per resource group.

RESOURCE	DEFAULT LIMIT	MAXIMUM LIMIT
Standard managed disks	50,000	50,000
Standard SSD managed disks	50,000	50,000
Premium managed disks	50,000	50,000
Standard_LRS snapshots	50,000	50,000
Standard_ZRS snapshots	50,000	50,000
Managed image	50,000	50,000

- **For Standard storage accounts:** A Standard storage account has a maximum total request rate of 20,000 IOPS. The total IOPS across all of your virtual machine disks in a Standard storage account should not

exceed this limit.

You can roughly calculate the number of highly utilized disks supported by a single Standard storage account based on the request rate limit. For example, for a Basic tier VM, the maximum number of highly utilized disks is about 66, which is 20,000/300 IOPS per disk. The maximum number of highly utilized disks for a Standard tier VM is about 40, which is 20,000/500 IOPS per disk.

- **For Premium storage accounts:** A Premium storage account has a maximum total throughput rate of 50 Gbps. The total throughput across all of your VM disks should not exceed this limit.

For more information, see Virtual machine sizes.

Managed virtual machine disks

Standard HDD managed disks

STANDARD DISK TYPE	S4	S6	S10	S15	S20	S30	S40	S50	S60	S70	S80
Disk size in GiB	32	64	128	256	512	1,024	2,048	4,096	8,192	16,384	32,767
IOPS per disk	Up to 500	Up to 500	Up to 500	Up to 500	Up to 500	Up to 500	Up to 500	Up to 500	Up to 1,300	Up to 2,000	Up to 2,000
Throughput per disk	Up to 60 MiB/sec	Up to 60 MiB/sec	Up to 60 MiB/sec	Up to 60 MiB/sec	Up to 60 MiB/sec	Up to 60 MiB/sec	Up to 60 MiB/sec	Up to 60 MiB/sec	Up to 300 MiB/sec	Up to 500 MiB/sec	Up to 500 MiB/sec

Standard SSD managed disks

STANDARD SSD SIZES	E1*	E2*	E3*	E4	E6	E10	E15	E20	E30	E40	E50	E60	E70	E80
Disk size in GiB	4	8	16	32	64	128	256	512	1,024	2,048	4,096	8,192	16,384	32,767
IOPS per disk	Up to 120	Up to 120	Up to 120	Up to 120	Up to 240	Up to 500	Up to 500	Up to 500	Up to 500	Up to 500	Up to 500	Up to 2,000	Up to 4,000	Up to 6,000
Throughput per disk	Up to 25 MiB/sec	Up to 25 MiB/sec	Up to 25 MiB/sec	Up to 25 MiB/sec	Up to 50 MiB/sec	Up to 60 MiB/sec	Up to 60 MiB/sec	Up to 60 MiB/sec	Up to 60 MiB/sec	Up to 60 MiB/sec	Up to 60 MiB/sec	Up to 400 MiB/sec	Up to 600 MiB/sec	Up to 750 MiB/sec

*Denotes a disk size that is currently in preview, for regional availability information see New disk sizes: Managed and unmanaged.

Premium SSD managed disks: Per-disk limits

PREMIUM SSD SIZES	P1*	P2*	P3*	P4	P6	P10	P15	P20	P30	P40	P50	P60	P70	P80
Disk size in GiB	4	8	16	32	64	128	256	512	1,024	2,048	4,096	8,192	16,384	32,767
IOPS per disk	120	120	120	120	240	500	1,100	2,300	5,000	7,500	7,500	16,000	18,000	20,000
Throughput per disk	25 MiB/sec	25 MiB/sec	25 MiB/sec	25 MiB/sec	50 MiB/sec	100 MiB/sec	125 MiB/sec	150 MiB/sec	200 MiB/sec	250 MiB/sec	250 MiB/sec	500 MiB/sec	750 MiB/sec	900 MiB/sec
Max burst IOPS per disk**	3,500	3,500	3,500	3,500	3,500	3,500	3,500	3,500						
Max burst throughput per disk**	170 MiB/sec	170 MiB/sec	170 MiB/sec	170 MiB/sec	170 MiB/sec	170 MiB/sec	170 MiB/sec	170 MiB/sec						
Max burst duration**	30 min	30 min	30 min	30 min	30 min	30 min	30 min	30 min						
Eligible for reservation	No	No	No	No	No	No	No	No	Yes, up to one year	Yes, up to one year	Yes, up to one year	Yes, up to one year	Yes, up to one year	Yes, up to one year

*Denotes a disk size that is currently in preview, for regional availability information see New disk sizes: Managed and unmanaged.

**Denotes a feature that is currently in preview, see Disk bursting for more information.

Premium SSD managed disks: Per-VM limits

RESOURCE	DEFAULT LIMIT
Maximum IOPS Per VM	80,000 IOPS with GS5 VM
Maximum throughput per VM	2,000 MB/s with GS5 VM

Unmanaged virtual machine disks

Standard unmanaged virtual machine disks: Per-disk limits

VM TIER	BASIC TIER VM	STANDARD TIER VM
Disk size	4,095 GB	4,095 GB
Maximum 8-KB IOPS per persistent disk	300	500
Maximum number of disks that perform the maximum IOPS	66	40

Premium unmanaged virtual machine disks: Per-account limits

RESOURCE	DEFAULT LIMIT
Total disk capacity per account	35 TB
Total snapshot capacity per account	10 TB
Maximum bandwidth per account (ingress + egress)[1]	<=50 Gbps

[1]*Ingress* refers to all data from requests that are sent to a storage account. *Egress* refers to all data from responses that are received from a storage account.

Premium unmanaged virtual machine disks: Per-disk limits

PREMIUM STORAGE DISK TYPE	P10	P20	P30	P40	P50
Disk size	128 GiB	512 GiB	1,024 GiB (1 TB)	2,048 GiB (2 TB)	4,095 GiB (4 TB)
Maximum IOPS per disk	500	2,300	5,000	7,500	7,500
Maximum throughput per disk	100 MB/sec	150 MB/sec	200 MB/sec	250 MB/sec	250 MB/sec
Maximum number of disks per storage account	280	70	35	17	8

Premium unmanaged virtual machine disks: Per-VM limits

RESOURCE	DEFAULT LIMIT
Maximum IOPS per VM	80,000 IOPS with GS5 VM
Maximum throughput per VM	2,000 MB/sec with GS5 VM

StorSimple System limits

LIMIT IDENTIFIER	LIMIT	COMMENTS
Maximum number of storage account credentials	64	
Maximum number of volume containers	64	
Maximum number of volumes	255	
Maximum number of schedules per bandwidth template	168	A schedule for every hour, every day of the week.
Maximum size of a tiered volume on physical devices	64 TB for StorSimple 8100 and StorSimple 8600	StorSimple 8100 and StorSimple 8600 are physical devices.
Maximum size of a tiered volume on virtual devices in Azure	30 TB for StorSimple 8010 64 TB for StorSimple 8020	StorSimple 8010 and StorSimple 8020 are virtual devices in Azure that use Standard storage and Premium storage, respectively.
Maximum size of a locally pinned volume on physical devices	9 TB for StorSimple 8100 24 TB for StorSimple 8600	StorSimple 8100 and StorSimple 8600 are physical devices.
Maximum number of iSCSI connections	512	
Maximum number of iSCSI connections from initiators	512	
Maximum number of access control records per device	64	
Maximum number of volumes per backup policy	24	
Maximum number of backups retained per backup policy	64	
Maximum number of schedules per backup policy	10	
Maximum number of snapshots of any type that can be retained per volume	256	This amount includes local snapshots and cloud snapshots.
Maximum number of snapshots that can be present in any device	10,000	

LIMIT IDENTIFIER	LIMIT	COMMENTS
Maximum number of volumes that can be processed in parallel for backup, restore, or clone	16	• If there are more than 16 volumes, they're processed sequentially as processing slots become available. • New backups of a cloned or a restored tiered volume can't occur until the operation is finished. For a local volume, backups are allowed after the volume is online.
Restore and clone recover time for tiered volumes	<2 minutes	• The volume is made available within 2 minutes of a restore or clone operation, regardless of the volume size. • The volume performance might initially be slower than normal as most of the data and metadata still resides in the cloud. Performance might increase as data flows from the cloud to the StorSimple device. • The total time to download metadata depends on the allocated volume size. Metadata is automatically brought into the device in the background at the rate of 5 minutes per TB of allocated volume data. This rate might be affected by Internet bandwidth to the cloud. • The restore or clone operation is complete when all the metadata is on the device. • Backup operations can't be performed until the restore or clone operation is fully complete.

LIMIT IDENTIFIER	LIMIT	COMMENTS
Restore recover time for locally pinned volumes	<2 minutes	• The volume is made available within 2 minutes of the restore operation, regardless of the volume size. • The volume performance might initially be slower than normal as most of the data and metadata still resides in the cloud. Performance might increase as data flows from the cloud to the StorSimple device. • The total time to download metadata depends on the allocated volume size. Metadata is automatically brought into the device in the background at the rate of 5 minutes per TB of allocated volume data. This rate might be affected by Internet bandwidth to the cloud. • Unlike tiered volumes, if there are locally pinned volumes, the volume data is also downloaded locally on the device. The restore operation is complete when all the volume data has been brought to the device. • The restore operations might be long and the total time to complete the restore will depend on the size of the provisioned local volume, your Internet bandwidth, and the existing data on the device. Backup operations on the locally pinned volume are allowed while the restore operation is in progress.
Thin-restore availability	Last failover	
Maximum client read/write throughput, when served from the SSD tier*	920/720 MB/sec with a single 10-gigabit Ethernet network interface	Up to two times with MPIO and two network interfaces.
Maximum client read/write throughput, when served from the HDD tier*	120/250 MB/sec	
Maximum client read/write throughput, when served from the cloud tier*	11/41 MB/sec	Read throughput depends on clients generating and maintaining sufficient I/O queue depth.

*Maximum throughput per I/O type was measured with 100 percent read and 100 percent write scenarios. Actual throughput might be lower and depends on I/O mix and network conditions.

Stream Analytics limits

LIMIT IDENTIFIER	LIMIT	COMMENTS
Maximum number of streaming units per subscription per region	500	To request an increase in streaming units for your subscription beyond 500, contact Microsoft Support.
Maximum number of inputs per job	60	There's a hard limit of 60 inputs per Azure Stream Analytics job.
Maximum number of outputs per job	60	There's a hard limit of 60 outputs per Stream Analytics job.
Maximum number of functions per job	60	There's a hard limit of 60 functions per Stream Analytics job.
Maximum number of streaming units per job	192	There's a hard limit of 192 streaming units per Stream Analytics job.
Maximum number of jobs per region	1,500	Each subscription can have up to 1,500 jobs per geographical region.
Reference data blob MB	300	Reference data blobs can't be larger than 300 MB each.

Virtual Machines limits

Virtual Machines limits

RESOURCE	DEFAULT LIMIT	MAXIMUM LIMIT
Virtual machines per cloud service[1]	50	50
Input endpoints per cloud service[2]	150	150

[1]Virtual machines created by using the classic deployment model instead of Azure Resource Manager are automatically stored in a cloud service. You can add more virtual machines to that cloud service for load balancing and availability.

[2]Input endpoints allow communications to a virtual machine from outside the virtual machine's cloud service. Virtual machines in the same cloud service or virtual network can automatically communicate with each other. For more information, see How to set up endpoints to a virtual machine.

Virtual Machines limits - Azure Resource Manager

The following limits apply when you use Azure Resource Manager and Azure resource groups.

RESOURCE	DEFAULT LIMIT
VMs per subscription	25,000[1] per region.
VM total cores per subscription	20[1] per region. Contact support to increase limit.
Azure Spot VM total cores per subscription	20[1] per region. Contact support to increase limit.
VM per series, such as Dv2 and F, cores per subscription	20[1] per region. Contact support to increase limit.

RESOURCE	DEFAULT LIMIT
Availability sets per subscription	2,000 per region.
Virtual machines per availability set	200
Certificates per subscription	Unlimited[2]

[1]Default limits vary by offer category type, such as Free Trial and Pay-As-You-Go, and by series, such as Dv2, F, and G. For example, the default for Enterprise Agreement subscriptions is 350.

[2]With Azure Resource Manager, certificates are stored in the Azure Key Vault. The number of certificates is unlimited for a subscription. There's a 1-MB limit of certificates per deployment, which consists of either a single VM or an availability set.

> **NOTE**
> Virtual machine cores have a regional total limit. They also have a limit for regional per-size series, such as Dv2 and F. These limits are separately enforced. For example, consider a subscription with a US East total VM core limit of 30, an A series core limit of 30, and a D series core limit of 30. This subscription can deploy 30 A1 VMs, or 30 D1 VMs, or a combination of the two not to exceed a total of 30 cores. An example of a combination is 10 A1 VMs and 20 D1 VMs.

Shared Image Gallery limits

There are limits, per subscription, for deploying resources using Shared Image Galleries:

- 100 shared image galleries, per subscription, per region
- 1,000 image definitions, per subscription, per region
- 10,000 image versions, per subscription, per region

Virtual machine scale sets limits

RESOURCE	DEFAULT LIMIT	MAXIMUM LIMIT
Maximum number of VMs in a scale set	1,000	1,000
Maximum number of VMs based on a custom VM image in a scale set	600	600
Maximum number of scale sets in a region	2,000	2,000

See also

- Understand Azure limits and increases
- Virtual machine and cloud service sizes for Azure
- Sizes for Azure Cloud Services
- Naming rules and restrictions for Azure resources