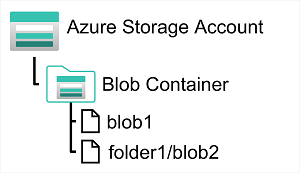
**Explore Azure blob storage**

Completed100 XP

* 4 minutes

Azure Blob Storage is a service that enables you to store massive amounts of unstructured data as binary large objects, or *blobs*, in the cloud. Blobs are an efficient way to store data files in a format that is optimized for cloud-based storage, and applications can read and write them by using the Azure blob storage API.



In an Azure storage account, you store blobs in *containers*. A container provides a convenient way of grouping related blobs together. You control who can read and write blobs inside a container at the container level.

Within a container, you can organize blobs in a hierarchy of virtual folders, similar to files in a file system on disk. However, by default, these folders are simply a way of using a "/" character in a blob name to organize the blobs into namespaces. The folders are purely virtual, and you can't perform folder-level operations to control access or perform bulk operations.

Azure Blob Storage supports three different types of blob:

* **Block blobs**. A block blob is handled as a set of blocks. Each block can vary in size, up to 4000 MiB. A block blob can contain up to 190.7 TiB (4000 MiB X 50,000 blocks), giving a maximum size of over 5000 MiB. The block is the smallest amount of data that can be read or written as an individual unit. Block blobs are best used to store discrete, large, binary objects that change infrequently.
* **Page blobs**. A page blob is organized as a collection of fixed size 512-byte pages. A page blob is optimized to support random read and write operations; you can fetch and store data for a single page if necessary. A page blob can hold up to 8 TB of data. Azure uses page blobs to implement virtual disk storage for virtual machines.
* **Append blobs**. An append blob is a block blob optimized to support append operations. You can only add blocks to the end of an append blob; updating or deleting existing blocks isn't supported. Each block can vary in size, up to 4 MB. The maximum size of an append blob is just over 195 GB.

Blob storage provides three access tiers, which help to balance access latency and storage cost:

* The *Hot* tier is the default. You use this tier for blobs that are accessed frequently. The blob data is stored on high-performance media.
* The *Cool* tier has lower performance and incurs reduced storage charges compared to the Hot tier. Use the Cool tier for data that is accessed infrequently. It's common for newly created blobs to be accessed frequently initially, but less so as time passes. In these situations, you can create the blob in the Hot tier, but migrate it to the Cool tier later. You can migrate a blob from the Cool tier back to the Hot tier.
* The *Archive* tier provides the lowest storage cost, but with increased latency. The Archive tier is intended for historical data that mustn't be lost, but is required only rarely. Blobs in the Archive tier are effectively stored in an offline state. Typical reading latency for the Hot and Cool tiers is a few milliseconds, but for the Archive tier, it can take hours for the data to become available. To retrieve a blob from the Archive tier, you must change the access tier to Hot or Cool. The blob will then be rehydrated. You can read the blob only when the rehydration process is complete.

**Lưu trữ Blob Azure: Lưu trữ dữ liệu phi cấu trúc hiệu quả trên đám mây**

**Lưu trữ Blob Azure** là dịch vụ cho phép bạn lưu trữ một lượng lớn dữ liệu phi cấu trúc dưới dạng các đối tượng nhị phân lớn (binary large objects) hoặc blob trên đám mây. Blob là một cách hiệu quả để lưu trữ các tập tin dữ liệu theo định dạng được tối ưu hóa cho lưu trữ đám mây. Các ứng dụng có thể đọc và ghi chúng bằng cách sử dụng API lưu trữ blob Azure.

**Trong tài khoản lưu trữ Azure**, bạn lưu trữ các blob trong **container** (thùng chứa). Container cung cấp một cách thuận tiện để nhóm các blob liên quan lại với nhau. Bạn kiểm soát ai có thể đọc và ghi blob bên trong một container ở cấp độ container.

**Bên trong một container**, bạn có thể tổ chức các blob theo một hệ thống thư mục ảo, tương tự như các tệp tin trong hệ thống tệp trên đĩa. Tuy nhiên, theo mặc định, các thư mục này chỉ đơn giản là cách sử dụng ký tự "/" trong tên blob để tổ chức các blob thành các không gian tên (namespaces). Các thư mục hoàn toàn ảo và bạn không thể thực hiện các thao tác cấp độ thư mục để kiểm soát quyền truy cập hoặc thực hiện các thao tác hàng loạt.

**Lưu trữ Blob Azure hỗ trợ ba loại blob khác nhau:**

* **Blob khối (Block blobs):** Là một tập hợp các khối. Mỗi khối có thể có kích thước khác nhau, tối đa 4000 MiB. Một blob khối có thể chứa tới 190,7 TiB (4000 MiB x 50.000 khối), cho kích thước tối đa vượt quá 5000 MiB. Khối là đơn vị dữ liệu nhỏ nhất có thể được đọc hoặc ghi riêng lẻ. Blob khối được sử dụng tốt nhất để lưu trữ các đối tượng nhị phân riêng biệt, lớn, thay đổi không thường xuyên.
* **Blob trang (Page blobs):** Được tổ chức thành một bộ các trang kích thước cố định 512 byte. Blob trang được tối ưu hóa để hỗ trợ các hoạt động đọc và ghi ngẫu nhiên; bạn có thể truy nạp và lưu trữ dữ liệu cho một trang riêng lẻ nếu cần. Một blob trang có thể chứa tới 8 TB dữ liệu. Azure sử dụng blob trang để triển khai lưu trữ đĩa ảo cho các máy ảo.
* **Blob nối thêm (Append blobs):** Là một blob khối được tối ưu hóa để hỗ trợ các hoạt động nối thêm. Bạn chỉ có thể thêm các khối vào cuối của một blob nối thêm; việc cập nhật hoặc xóa các khối hiện có không được hỗ trợ. Mỗi khối có thể có kích thước khác nhau, tối đa 4 MB. Kích thước tối đa của một blob nối thêm chỉ hơn 195 GB.

**Lưu trữ Blob cung cấp ba tầng truy cập, giúp cân bằng giữa độ trễ truy cập và chi phí lưu trữ:**

* **Tầng Nóng (Hot):** Là tầng mặc định. Bạn sử dụng tầng này cho các blob được truy cập thường xuyên. Dữ liệu blob được lưu trữ trên phương tiện truyền thông hiệu suất cao.
* **Tầng Mát (Cool):** Có hiệu suất thấp hơn và chi phí lưu trữ giảm so với tầng Nóng. Sử dụng tầng Mát cho dữ liệu được truy cập không thường xuyên. Thông thường các blob mới được tạo sẽ được truy cập thường xuyên ban đầu, nhưng theo thời gian thì ít hơn. Trong những tình huống này, bạn có thể tạo blob trong tầng Nóng, nhưng sau đó di chuyển nó sang tầng Mát. Bạn có thể di chuyển một blob từ tầng Mát trở lại tầng Nóng.
* **Tầng Lưu trữ (Archive):** Cung cấp chi phí lưu trữ thấp nhất, nhưng độ trễ truy cập tăng lên. Tầng Lưu trữ dành cho dữ liệu lịch sử không được phép bị mất, nhưng chỉ cần thiết hiếm khi. Các blob trong tầng Lưu trữ được lưu trữ ngoại tuyến. Độ trễ đọc điển hình cho các tầng Nóng và Mát là vài mili giây, nhưng đối với tầng Lưu trữ, có thể mất hàng giờ để dữ liệu sẵn sàng. Để truy xuất một blob từ tầng Lưu trữ, bạn phải thay đổi tầng truy cập thành Nóng hoặc Mát. Blob sau đó sẽ được "cấp nước" (rehydrated) trở lại. Bạn chỉ có thể đọc blob khi quá trình cấp nước hoàn tất.

**Ví dụ về việc sử dụng 3 loại blob trong Lưu trữ Blob Azure:**

**1. Blob khối (Block blobs):**

* **Lưu trữ ảnh:** Thích hợp để lưu trữ ảnh vì có thể lưu trữ các tệp lớn và thay đổi không thường xuyên.
* **Lưu trữ video:** Thích hợp để lưu trữ video vì có thể lưu trữ các tệp lớn và có thể truy cập các phần cụ thể của video một cách hiệu quả.
* **Lưu trữ các bản sao lưu:** Thích hợp để lưu trữ các bản sao lưu dữ liệu vì có thể lưu trữ lượng dữ liệu lớn và truy cập khi cần thiết.

**2. Blob trang (Page blobs):**

* **Lưu trữ cơ sở dữ liệu:** Thích hợp để lưu trữ cơ sở dữ liệu vì có thể truy cập dữ liệu ngẫu nhiên một cách hiệu quả.
* **Lưu trữ hệ thống tệp:** Thích hợp để lưu trữ hệ thống tệp vì có thể truy cập các tệp và thư mục theo cách tương tự như hệ thống tệp cục bộ.
* **Lưu trữ các ứng dụng web:** Thích hợp để lưu trữ các ứng dụng web vì có thể truy cập các tệp và thư mục theo cách hiệu quả.

**3. Blob nối thêm (Append blobs):**

* **Lưu trữ nhật ký:** Thích hợp để lưu trữ nhật ký vì có thể ghi dữ liệu liên tục vào blob.
* **Lưu trữ dữ liệu chuỗi thời gian:** Thích hợp để lưu trữ dữ liệu chuỗi thời gian vì có thể ghi dữ liệu liên tục vào blob.
* **Lưu trữ các bản ghi trò chuyện:** Thích hợp để lưu trữ các bản ghi trò chuyện vì có thể ghi dữ liệu liên tục vào blob.