ALIBABA CLOUD

Alibaba Cloud

Apsara File Storage NAS
Performance

Document Version: 20210922

Alibaba Cloud

Legal disclaimer

Alibaba Cloud reminds you to carefully read and fully understand the terms and conditions of this legal disclaimer before you read or use this document. If you have read or used this document, it shall be deemed as your total acceptance of this legal disclaimer.

1. You shall download and obtain this document from the Alibaba Cloud website or other Alibaba Cloud-authorized channels, and use this document for your own legal business activities only. The content of this document is considered confidential information of Alibaba Cloud. You shall strictly abide by the confidentiality obligations. No part of this document shall be disclosed or provided to any third party for use without the prior written consent of Alibaba Cloud.

2. No part of this document shall be excerpted, translated, reproduced, transmitted, or disseminated by any organization, company or individual in any form or by any means without the prior written consent of Alibaba Cloud.

3. The content of this document may be changed because of product version upgrade, adjustment, or other reasons. Alibaba Cloud reserves the right to modify the content of this document without notice and an updated version of this document will be released through Alibaba Cloud-authorized channels from time to time. You should pay attention to the version changes of this document as they occur and download and obtain the most up-to-date version of this document from Alibaba Cloud-authorized channels.

4. This document serves only as a reference guide for your use of Alibaba Cloud products and services. Alibaba Cloud provides this document based on the "status quo", "being defective", and "existing functions" of its products and services. Alibaba Cloud makes every effort to provide relevant operational guidance based on existing technologies. However, Alibaba Cloud hereby makes a clear statement that it in no way guarantees the accuracy, integrity, applicability, and reliability of the content of this document, either explicitly or implicitly. Alibaba Cloud shall not take legal responsibility for any errors or lost profits incurred by any organization, company, or individual arising from download, use, or trust in this document. Alibaba Cloud shall not, under any circumstances, take responsibility for any indirect, consequential, punitive, contingent, special, or punitive damages, including lost profits arising from the use or trust in this document (even if Alibaba Cloud has been notified of the possibility of such a loss).

5. By law, all the contents in Alibaba Cloud documents, including but not limited to pictures, architecture design, page layout, and text description, are intellectual property of Alibaba Cloud and/or its affiliates. This intellectual property includes, but is not limited to, trademark rights, patent rights, copyrights, and trade secrets. No part of this document shall be used, modified, reproduced, publicly transmitted, changed, disseminated, distributed, or published without the prior written consent of Alibaba Cloud and/or its affiliates. The names owned by Alibaba Cloud shall not be used, published, or reproduced for marketing, advertising, promotion, or other purposes without the prior written consent of Alibaba Cloud. The names owned by Alibaba Cloud include, but are not limited to, "Alibaba Cloud", "Aliyun", "HiChina", and other brands of Alibaba Cloud and/or its affiliates, which appear separately or in combination, as well as the auxiliary signs and patterns of the preceding brands, or anything similar to the company names, trade names, trademarks, product or service names, domain names, patterns, logos, marks, signs, or special descriptions that third parties identify as Alibaba Cloud and/or its affiliates.

6. Please directly contact Alibaba Cloud for any errors of this document.

Document conventions

Style	Description	Example
Danger	A danger notice indicates a situation that will cause major system changes, faults, physical injuries, and other adverse results.	**Danger:** Resetting will result in the loss of user configuration data.
Warning	A warning notice indicates a situation that may cause major system changes, faults, physical injuries, and other adverse results.	**Warning:** Restarting will cause business interruption. About 10 minutes are required to restart an instance.
Notice	A caution notice indicates warning information, supplementary instructions, and other content that the user must understand.	**Notice:** If the weight is set to 0, the server no longer receives new requests.
Note	A note indicates supplemental instructions, best practices, tips, and other content.	**Note:** You can use Ctrl + A to select all files.
>	Closing angle brackets are used to indicate a multi-level menu cascade.	Click **Settings**> **Network**> **Set network type**.
Bold	Bold formatting is used for buttons , menus, page names, and other UI elements.	Click **OK**.
`Courier font`	Courier font is used for commands	Run the `cd /d C:/window` command to enter the Windows system folder.
Italic	Italic formatting is used for parameters and variables.	`bae log list --instanceid` *Instance_ID*
[] or [a|b]	This format is used for an optional value, where only one item can be selected.	`ipconfig [-all|-t]`
{} or {a|b}	This format is used for a required value, where only one item can be selected.	`switch {active|stand}`

Table of Contents

1.Performance testing for Apsara File Storage NAS

This topic describes how to use the fio tool to measure the throughput and input/output operations per second (IOPS) of Apsara File Storage NAS.

Performance testing in Linux

Before you conduct performance tests, note that:

- The sunrpc_slot parameter must be set to a proper value. For more information, see Why is the performance of an NFS client poor on Linux?.
- The maximum throughput of a file system cannot exceed the bandwidth of an Elastic Compute Service (ECS) instance. If the bandwidth of the ECS instance is 1 Gbit/s, the throughput of the file system can reach a maximum of 125 MB/s.
- The fio tool is pre-installed on your ECS instance. If the fio tool is unavailable, run the following command to install the tool:
  - If CentOS, Red Hat Enterprise Linux (RHEL), or Aliyun Linux is running on the ECS instance, run the following command to install the tool:

    ```
    sudo yum install fio
    ```

  - If Ubuntu or Debian is running on the ECS instance, run the following commands to install the tool:

    ```
    sudo apt-get  update
    ```

    ```
    sudo apt-get install fio
    ```

The following examples show several performance testing cases:

> **Note** The following values are all estimated based on the test results of a single ECS instance. To meet the performance metrics that are provided on the product page of NAS, we recommend that you use multiple ECS instances to conduct performance tests. For more information about the performance metrics, visit Apsara File Storage NAS.

- Random read IOPS

  ```
  fio -numjobs=1 -iodepth=128 -direct=1 -ioengine=libaio -sync=1 -rw=randread -bs=4K -size=1G -time_based -runtime=60 -name=Fio -directory=/mnt
  ```

  Estimated value: 14,000

```
Fio: (g=0): rw=randread, bs=(R) 4096B-4096B, (W) 4096B-4096B, (T) 4096B-4096B, ioengine=libaio, iodepth=128
fio-3.1
Starting 1 process
Fio: Laying out IO file (1 file / 1024MiB)
fio: native_fallocate call failed: Operation not supported
Jobs: 1 (f=1): [r(1)][100.0%][r=58.0MiB/s,w=0KiB/s][r=15.1k,w=0 IOPS][eta 00m:00s]
Fio: (groupid=0, jobs=1): err= 0: pid=5196: Mon Mar 11 11:37:36 2019
   read: IOPS=15.3k, BW=59.6MiB/s (62.5MB/s)(1793MiB/30061msec)
    slat (usec): min=2, max=185, avg= 8.00, stdev= 4.00
    clat (usec): min=909, max=684981, avg=8370.96, stdev=13265.50
     lat (usec): min=916, max=684986, avg=8379.54, stdev=13265.48
    clat percentiles (usec):
     |  1.00th=[  1139],  5.00th=[  1336], 10.00th=[  1598], 20.00th=[  2212],
     | 30.00th=[  2835], 40.00th=[  3458], 50.00th=[  4113], 60.00th=[  4817],
     | 70.00th=[  5800], 80.00th=[  7898], 90.00th=[ 32900], 95.00th=[ 35390],
     | 99.00th=[ 40109], 99.50th=[ 85459], 99.90th=[175113], 99.95th=[177210],
     | 99.99th=[181404]
   bw (  KiB/s): min=47176, max=96280, per=100.00%, avg=61176.13, stdev=7812.35, samples=60
   iops        : min=11794, max=24070, avg=15294.02, stdev=1953.09, samples=60
  lat (usec)   : 1000=0.04%
  lat (msec)   : 2=16.63%, 4=31.67%, 10=35.17%, 20=5.03%, 50=10.69%
  lat (msec)   : 100=0.67%, 250=0.10%, 500=0.01%, 750=0.01%
  cpu          : usr=4.42%, sys=23.50%, ctx=254636, majf=0, minf=155
  IO depths    : 1=0.1%, 2=0.1%, 4=0.1%, 8=0.1%, 16=0.1%, 32=0.1%, >=64=100.0%
     submit    : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
     complete  : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.1%
     issued rwt: total=459008,0,0, short=0,0,0, dropped=0,0,0
     latency   : target=0, window=0, percentile=100.00%, depth=128

Run status group 0 (all jobs):
   READ: bw=59.6MiB/s (62.5MB/s), 59.6MiB/s-59.6MiB/s (62.5MB/s-62.5MB/s), io=1793MiB (1880MB), run=30061-30061msec
```

- Random write IOPS

  ```
  fio -numjobs=1 -iodepth=128 -direct=1 -ioengine=libaio -sync=1 -rw=randwrite -bs=4K -size=1G -time_based -runtime=60 -name=Fio -directory=/mnt
  ```

  Estimated value: 10,000

```
Fio: (g=0): rw=randwrite, bs=(R) 4096B-4096B, (W) 4096B-4096B, (T) 4096B-4096B, ioengine=libaio, iodepth=128
fio-3.1
Starting 1 process
Jobs: 1 (f=1): [w(1)][100.0%][r=0KiB/s,w=59.4MiB/s][r=0,w=15.2k IOPS][eta 00m:00s]
Fio: (groupid=0, jobs=1): err= 0: pid=5205: Mon Mar 11 11:38:53 2019
  write: IOPS=15.5k, BW=60.6MiB/s (63.6MB/s)(1820MiB/30016msec)
    slat (usec): min=2, max=7876, avg= 7.64, stdev=11.81
    clat (usec): min=1928, max=247922, avg=8233.26, stdev=4356.30
     lat (usec): min=1933, max=247926, avg=8241.46, stdev=4356.26
    clat percentiles (msec):
     |  1.00th=[    4],  5.00th=[    5], 10.00th=[    5], 20.00th=[    6],
     | 30.00th=[    7], 40.00th=[    7], 50.00th=[    8], 60.00th=[    9],
     | 70.00th=[    9], 80.00th=[   11], 90.00th=[   13], 95.00th=[   16],
     | 99.00th=[   21], 99.50th=[   23], 99.90th=[   40], 99.95th=[   42],
     | 99.99th=[  178]
   bw (  KiB/s): min=54968, max=70856, per=100.00%, avg=62109.13, stdev=3242.17, samples=60
   iops        : min=13742, max=17714, avg=15527.25, stdev=810.54, samples=60
  lat (msec)   : 2=0.01%, 4=2.44%, 10=76.87%, 20=19.71%, 50=0.95%
  lat (msec)   : 100=0.01%, 250=0.02%
  cpu          : usr=5.25%, sys=24.21%, ctx=304678, majf=0, minf=24
  IO depths    : 1=0.1%, 2=0.1%, 4=0.1%, 8=0.1%, 16=0.1%, 32=0.1%, >=64=100.0%
     submit    : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
     complete  : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.1%
     issued rwt: total=0,466009,0, short=0,0,0, dropped=0,0,0
     latency   : target=0, window=0, percentile=100.00%, depth=128

Run status group 0 (all jobs):
  WRITE: bw=60.6MiB/s (63.6MB/s), 60.6MiB/s-60.6MiB/s (63.6MB/s-63.6MB/s), io=1820MiB (1909MB), run=30016-30016msec
```

- Random read throughput

  ```
  fio -numjobs=1 -iodepth=128 -direct=1 -ioengine=libaio -sync=1 -rw=randread -bs=1M -size=1G -time_based -runtime=60 -name=Fio -directory=/mnt
  ```

- Estimated value for NAS Capacity: 150 MB/s

```
Fio: (g=0): rw=randread, bs=(R) 1024KiB-1024KiB, (W) 1024KiB-1024KiB, (T) 1024KiB-1024KiB, ioengine=libaio, iodepth=128
fio-3.1
Starting 1 process
Jobs: 1 (f=1): [r(1)][100.0%][r=0KiB/s,w=0KiB/s][r=0,w=0 IOPS][eta 00m:00s]
Fio: (groupid=0, jobs=1): err= 0:  pid=5208: Mon Mar 11 11:40:35 2019
   read: IOPS=148,  BW=149MiB/s (156MB/s)(9214MiB/62045msec)
    slat (usec): min=74, max=2201.5k, avg=361.66, stdev=22934.09
    clat (msec): min=4, max=42052, avg=860.88, stdev=2559.16
     lat (msec): min=4, max=42052, avg=861.25, stdev=2559.19
    clat percentiles (msec):
     |  1.00th=[    8],  5.00th=[    9], 10.00th=[   10], 20.00th=[   13],
     | 30.00th=[   24], 40.00th=[   42], 50.00th=[   56], 60.00th=[  100],
     | 70.00th=[  197], 80.00th=[  693], 90.00th=[ 2534], 95.00th=[ 4665],
     | 99.00th=[12684], 99.50th=[16711], 99.90th=[17113], 99.95th=[17113],
     | 99.99th=[17113]
   bw (  KiB/s): min= 2048, max=382211, per=100.00%, avg=157685.99, stdev=45144.76, samples=118
   iops        : min=    2, max=  373, avg=153.94, stdev=44.08, samples=118
  lat (msec)   : 10=11.38%, 20=17.58%, 50=19.70%, 100=11.97%, 250=10.48%
  lat (msec)   : 500=6.14%, 750=5.37%, 1000=0.80%, 2000=5.84%, >=2000=10.72%
  cpu          : usr=0.12%, sys=2.07%, ctx=9087, majf=0, minf=604
  IO depths    : 1=0.1%, 2=0.1%, 4=0.1%, 8=0.1%, 16=0.2%, 32=0.3%, >=64=99.3%
     submit    : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
     complete  : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.1%
     issued rwt: total=9214,0,0, short=0,0,0, dropped=0,0,0
     latency   : target=0, window=0, percentile=100.00%, depth=128

Run status group 0 (all jobs):
   READ: bw=149MiB/s (156MB/s), 149MiB/s-149MiB/s (156MB/s-156MB/s), io=9214MiB (9662MB), run=62045-62045msec
```

- Estimated value for NAS Performance: 300 MB/s

- Random write throughput

```
fio -numjobs=1 -iodepth=128 -direct=1 -ioengine=libaio -sync=1 -rw=randwrite -bs=1M -size=1G -time_based -runtime=60 -name=Fio -directory=/mnt
```

- Estimated value for NAS Capacity: 150 MB/s

```
Fio: (g=0): rw=randwrite, bs=(R) 1024KiB-1024KiB, (W) 1024KiB-1024KiB, (T) 1024KiB-1024KiB, ioengine=libaio, iodepth=128
fio-3.1
Starting 1 process
Fio: Laying out IO file (1 file / 1024MiB)
fio: native_fallocate call failed: Operation not supported
Jobs: 1 (f=1): [w(1)][100.0%][r=0KiB/s,w=30.0MiB/s][r=0,w=30 IOPS][eta 00m:00s]
Fio: (groupid=0, jobs=1): err= 0: pid=4966: Mon Mar 11 12:27:51 2019
  write: IOPS=152,  BW=152MiB/s (160MB/s)(9262MiB/60849msec)
    slat (usec): min=125, max=861634, avg=303.96, stdev=8950.88
    clat (msec): min=26, max=1728, avg=840.31, stdev=143.40
     lat (msec): min=28, max=1729, avg=840.61, stdev=143.15
    clat percentiles (msec):
     |  1.00th=[  268],  5.00th=[  726], 10.00th=[  793], 20.00th=[  818],
     | 30.00th=[  835], 40.00th=[  844], 50.00th=[  852], 60.00th=[  860],
     | 70.00th=[  869], 80.00th=[  885], 90.00th=[  911], 95.00th=[  927],
     | 99.00th=[ 1401], 99.50th=[ 1552], 99.90th=[ 1687], 99.95th=[ 1703],
     | 99.99th=[ 1737]
   bw (  KiB/s): min= 2048, max=354304, per=100.00%, avg=155880.98, stdev=28138.64, samples=120
   iops        : min=    2, max=  346, avg=152.17, stdev=27.48, samples=120
  lat (msec)   : 50=0.12%, 100=0.08%, 250=0.51%, 500=3.21%, 750=1.37%
  lat (msec)   : 1000=92.38%, 2000=2.34%
  cpu          : usr=1.56%, sys=1.99%, ctx=8453, majf=0, minf=23
  IO depths    : 1=0.1%, 2=0.1%, 4=0.1%, 8=0.1%, 16=0.2%, 32=0.3%, >=64=99.3%
     submit    : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
     complete  : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.1%
     issued rwt: total=0,9262,0, short=0,0,0, dropped=0,0,0
     latency   : target=0, window=0, percentile=100.00%, depth=128

Run status group 0 (all jobs):
  WRITE: bw=152MiB/s (160MB/s), 152MiB/s-152MiB/s (160MB/s-160MB/s), io=9262MiB (9712MB), run=60849-60849msec
```

- Estimated value for NAS Performance: 300 MB/s

Performance testing in Windows

Before you conduct performance tests, note that:

- The requests to access a file system from a Windows server are processed in serial. Therefore,

multiple files are used in the following performance tests.

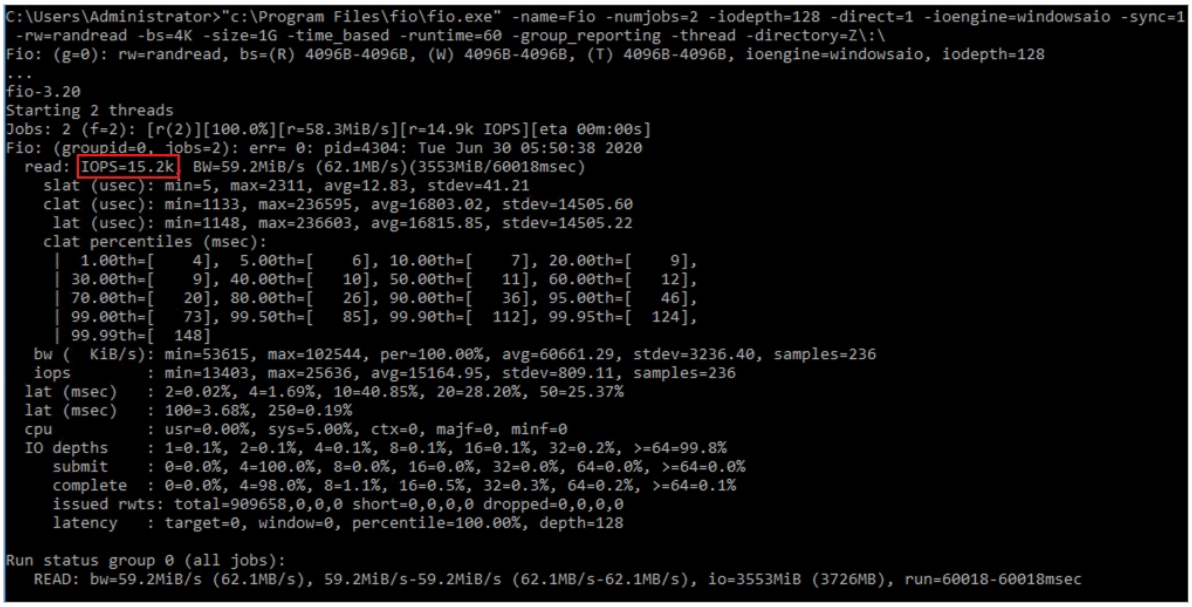
You can download and install a Microsoft Installer (MSI) package of the fio tool on a Windows server. To download the MSI package, visit the fio official website.

In this example, a NAS file system is mounted on Disk Z and the fio tool is installed in the *C:\Program Files\fio\fio.exe* directory. You can run the following commands to conduct performance tests.

- Random read IOPS

  ```
  "c:\Program Files\fio\fio.exe" -name=Fio -numjobs=2 -iodepth=128 -direct=1 -ioengine=windowsaio -sync=1 -rw=randread -bs=4K -size=1G -time_based -runtime=60 -group_reporting -thread -directory=Z\:\
  ```

  Estimated value: 14,000

- Random write IOPS

  ```
  "c:\Program Files\fio\fio.exe" -name=Fio -numjobs=2 -iodepth=128 -direct=1 -ioengine=windowsaio -sync=1 -rw=randwrite -bs=4K -size=1G -time_based -runtime=60 -group_reporting -thread -directory=Z\:\
  ```

  Estimated value: 10,000

```
C:\Users\Administrator>"c:\Program Files\fio\fio.exe" -name=Fio -numjobs=2 -iodepth=128 -direct=1 -ioengine=windowsaio -sync=1
 -rw=randwrite -bs=4K -size=1G -time_based -runtime=60 -group_reporting -thread -directory=Z\:\
Fio: (g=0): rw=randwrite, bs=(R) 4096B-4096B, (W) 4096B-4096B, (T) 4096B-4096B, ioengine=windowsaio, iodepth=128
...
fio-3.20
Starting 2 threads
Jobs: 2 (f=2): [w(2)][100.0%][w=53.7MiB/s][w=13.7k IOPS][eta 00m:00s]
Fio: (groupid=0, jobs=2): err= 0: pid=4740: Tue Jun 30 05:49:01 2020
  write: IOPS=13.5k, BW=52.6MiB/s (55.1MB/s)(3154MiB/60018msec); 0 zone resets
    slat (usec): min=3, max=1365, avg=11.34, stdev= 8.34
    clat (usec): min=2005, max=63791, avg=18979.27, stdev=3108.76
     lat (usec): min=2044, max=63795, avg=18990.61, stdev=3108.62
    clat percentiles (usec):
     |  1.00th=[12387],  5.00th=[15270], 10.00th=[16057], 20.00th=[16909],
     | 30.00th=[17695], 40.00th=[18220], 50.00th=[18744], 60.00th=[19268],
     | 70.00th=[19792], 80.00th=[20579], 90.00th=[21890], 95.00th=[23462],
     | 99.00th=[30540], 99.50th=[34866], 99.90th=[40109], 99.95th=[41681],
     | 99.99th=[53740]
   bw (  KiB/s): min=45616, max=60022, per=100.00%, avg=53851.28, stdev=1308.69, samples=232
   iops        : min=11404, max=15005, avg=13462.21, stdev=327.15, samples=232
  lat (msec)   : 4=0.01%, 10=0.59%, 20=70.94%, 50=28.46%, 100=0.02%
  cpu          : usr=0.00%, sys=4.17%, ctx=0, majf=0, minf=0
  IO depths    : 1=0.1%, 2=0.1%, 4=0.1%, 8=0.1%, 16=0.1%, 32=0.1%, >=64=100.0%
     submit    : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
     complete  : 0=0.0%, 4=95.2%, 8=3.7%, 16=0.9%, 32=0.2%, 64=0.1%, >=64=0.1%
     issued rwts: total=0,807502,0,0 short=0,0,0,0 dropped=0,0,0,0
     latency   : target=0, window=0, percentile=100.00%, depth=128

Run status group 0 (all jobs):
  WRITE: bw=52.6MiB/s (55.1MB/s), 52.6MiB/s-52.6MiB/s (55.1MB/s-55.1MB/s), io=3154MiB (3308MB), run=60018-60018msec
```

- Random read throughput

  ```
  "c:\Program Files\fio\fio.exe" -name=Fio -numjobs=2 -iodepth=128 -direct=1 -ioengine=windowsaio -sync
  =1 -rw=randread -bs=1M -size=1G -time_based -runtime=60 -group_reporting -thread -directory=Z\:\
  ```

  - Estimated value for NAS Capacity: 150 MB/s

    ```
    C:\Users\Administrator>"c:\Program Files\fio\fio.exe" -name=Fio -numjobs=2 -iodepth=128 -direct=1 -ioengine=windowsaio -sync=1
     -rw=randread -bs=1M -size=1G -time_based -runtime=60 -group_reporting -thread -directory=Z\:\
    Fio: (g=0): rw=randread, bs=(R) 1024KiB-1024KiB, (W) 1024KiB-1024KiB, (T) 1024KiB-1024KiB, ioengine=windowsaio, iodepth=128
    ...
    fio-3.20
    Starting 2 threads
    Jobs: 2 (f=0): [f(2)][100.0%][r=254MiB/s][r=254 IOPS][eta 00m:00s]
    Fio: (groupid=0, jobs=2): err= 0: pid=1264: Tue Jun 30 05:43:14 2020
      read: IOPS=152, BW=153MiB/s (160MB/s)(9439MiB/61697msec)
        slat (usec): min=33, max=71264, avg=102.71, stdev=754.88
        clat (msec): min=5, max=14473, avg=1670.15, stdev=1711.22
         lat (msec): min=5, max=14473, avg=1670.25, stdev=1711.21
        clat percentiles (msec):
         |  1.00th=[   10],  5.00th=[   83], 10.00th=[  153], 20.00th=[  347],
         | 30.00th=[  567], 40.00th=[  835], 50.00th=[ 1150], 60.00th=[ 1536],
         | 70.00th=[ 2005], 80.00th=[ 2668], 90.00th=[ 3876], 95.00th=[ 5067],
         | 99.00th=[ 8020], 99.50th=[ 9463], 99.90th=[12013], 99.95th=[13221],
         | 99.99th=[14429]
       bw (  KiB/s): min=108326, max=457835, per=99.89%, avg=156494.95, stdev=19940.82, samples=240
       iops        : min=  104, max=  446, avg=151.55, stdev=19.50, samples=240
      lat (msec)   : 10=1.14%, 20=0.56%, 50=1.11%, 100=3.39%, 250=9.37%
      lat (msec)   : 500=11.47%, 750=9.92%, 1000=8.62%, 2000=24.27%, >=2000=30.14%
      cpu          : usr=0.00%, sys=0.00%, ctx=0, majf=0, minf=0
      IO depths    : 1=0.1%, 2=0.1%, 4=0.1%, 8=0.2%, 16=0.3%, 32=0.7%, >=64=98.7%
         submit    : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
         complete  : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.1%, 32=0.0%, 64=0.1%, >=64=0.1%
         issued rwts: total=9439,0,0,0 short=0,0,0,0 dropped=0,0,0,0
         latency   : target=0, window=0, percentile=100.00%, depth=128

    Run status group 0 (all jobs):
       READ: bw=153MiB/s (160MB/s), 153MiB/s-153MiB/s (160MB/s-160MB/s), io=9439MiB (9898MB), run=61697-61697msec
    ```

  - Estimated value for NAS Performance: 300 MB/s

- Random write throughput

  ```
  "c:\Program Files\fio\fio.exe" -name=Fio -numjobs=2 -iodepth=128 -direct=1 -ioengine=windowsaio -sync
  =1 -rw=randwrite -bs=1M -size=1G -time_based -runtime=60 -group_reporting -thread -directory=Z\:\
  ```

- Estimated value for NAS Capacity: 150 MB/s

```
C:\Users\Administrator>"c:\Program Files\fio\fio.exe" -name=Fio -numjobs=2 -iodepth=128 -direct=1 -ioengine=windowsaio -sync=1
 -rw=randwrite -bs=1M -size=1G -time_based -runtime=60 -group_reporting -thread -directory=Z\:\
Fio: (g=0): rw=randwrite, bs=(R) 1024KiB-1024KiB, (W) 1024KiB-1024KiB, (T) 1024KiB-1024KiB, ioengine=windowsaio, iodepth=128
...
fio-3.20
Starting 2 threads
Jobs: 2 (f=0): [f(2)][100.0%][w=254MiB/s][w=254 IOPS][eta 00m:00s]
Fio: (groupid=0, jobs=2): err= 0: pid=1020: Tue Jun 30 05:46:20 2020
  write: IOPS=152, BW=153MiB/s (160MB/s)(9439MiB/61702msec); 0 zone resets
    slat (usec): min=50, max=1337, avg=93.71, stdev=22.73
    clat (msec): min=10, max=4072, avg=1672.27, stdev=334.35
     lat (msec): min=10, max=4072, avg=1672.36, stdev=334.34
    clat percentiles (msec):
     |  1.00th=[  239],  5.00th=[ 1284], 10.00th=[ 1469], 20.00th=[ 1586],
     | 30.00th=[ 1620], 40.00th=[ 1670], 50.00th=[ 1687], 60.00th=[ 1720],
     | 70.00th=[ 1754], 80.00th=[ 1821], 90.00th=[ 1905], 95.00th=[ 2005],
     | 99.00th=[ 2769], 99.50th=[ 3071], 99.90th=[ 3675], 99.95th=[ 3809],
     | 99.99th=[ 4077]
   bw (  KiB/s): min=21871, max=455791, per=99.37%, avg=155661.83, stdev=17383.30, samples=240
   iops        : min=   20, max=  444, avg=151.70, stdev=16.97, samples=240
  lat (msec)   : 20=0.21%, 50=0.23%, 100=0.11%, 250=0.56%, 500=1.28%
  lat (msec)   : 750=0.49%, 1000=0.53%, 2000=91.61%, >=2000=4.98%
  cpu          : usr=0.00%, sys=0.00%, ctx=0, majf=0, minf=0
  IO depths    : 1=0.1%, 2=0.1%, 4=0.1%, 8=0.2%, 16=0.3%, 32=0.7%, >=64=98.7%
     submit    : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.0%
     complete  : 0=0.0%, 4=99.9%, 8=0.1%, 16=0.0%, 32=0.0%, 64=0.0%, >=64=0.1%
     issued rwts: total=0,9439,0,0 short=0,0,0,0 dropped=0,0,0,0
     latency   : target=0, window=0, percentile=100.00%, depth=128

Run status group 0 (all jobs):
  WRITE: bw=153MiB/s (160MB/s), 153MiB/s-153MiB/s (160MB/s-160MB/s), io=9439MiB (9898MB), run=61702-61702msec
```

- Estimated value for NAS Performance: 600 MB/s

2.FAQ about performance of Apsara File Storage NAS

What is the read and write performance of a file system related to?

The peak performance of read and write operations is linearly proportional to the used capacity of the file system. A higher capacity indicates a higher throughput.

What is IOPS? What are the relationships between IOPS and throughput, read and write block size, and latency?

IOPS refers to the number of input/output operations per second.

The following formulas indicate the relationships between IOPS and read and write block size, throughput, number of reads and writes, and latency:

Throughput = IOPS × Read and write block size, where, IOPS = Number of reads and writes/Overall latency

For example, a Capacity NAS file system has a write latency of 100 ms per 1 MiB, 15 ms per 8 KiB, and 10 ms per 4 KiB. You can send a maximum of 128 concurrent requests to a Capacity NAS file system. If you want to write 1 MiB of data per second to a file system, you can use the solutions in the following table.

No.	Read and write block size	Concurrency	Number of write operations	Overall latency	IOPS	Throughput	Description
Solution 1	4 KiB	1	250	10 ms*250=2.5s	250/2.5s=100	4 KiB*100=400 KiB/s	Small read and write block size and low concurrency result in poor throughput and high latency. In this case, the throughput cannot reach 1 MiB/s.
Solution 2	1 MiB	1	1	100 ms	1/0.1s=10	1 MiB*10=10 MiB/s	Compared with Solution 1, the read and write block size of Solution 2 is increased, and the throughput and latency performance are improved. The throughput can reach 1 MiB/s. However, the overall latency is high.

No.	Read and write block size	Concurrency	Number of write operations	Overall latency	IOPS	Throughput	Description
Solution 3	4 KiB	125	250	10 ms × (250/125) = 20 ms	250/0.02s =12500	4 KiB × 12500 ≈ 49 MiB/s	Compared with Solution 1, the concurrency of Solution 3 is increased, and the throughput and latency performance are improved. The throughput can reach 1 MiB/s. The overall latency is low, but the IOPS reaches the upper limit of the file system.
Solution 4	8 KiB	125	125	15 ms × (125/125) = 15 ms	125/0.015 s ≈ 8333	8 KiB × 8333 ≈ 65 MiB/s	Compared with Solution 1, the read and write block size and concurrency of Solution 4 are increased, and the throughput and latency performance are improved. The throughput can reach 1 MiB/s. The overall latency is the lowest among the four solutions and the IOPS is low.

What happens if the read and write throughput of a request exceeds the bandwidth threshold?

If the read and write throughput of a request sent by you or your application exceeds the bandwidth threshold, NAS throttle the request. In this case, the latency is increased. For more information, see Performance metrics.

Why does NGINX take a long time to write logs to a file system?

- Background information

  You can use the following two directives to specify NGINX logs: The **log_format** directive specifies the format of the logs. The **access_log** directive specifies the log storage path, format name, and cache size.

- Issue

  NGINX requires a long time to write logs to the file system and the performance of the file system is compromised.

- Cause

  The path that is specified in the **access_log** directive contains variables. Each time NGINX attempts to write logs to the file system, the destination files are opened. After the logs are written, the files are closed. To ensure data visibility, NAS writes the data back to the NAS server when the files are closed. This compromises the performance of the file system.

- Solution
  - Solution 1: Delete the variables in the **access_log** directive and store the logs in a fixed path.
  - Solution 2: Use the **open_log_file_cache** directive to cache the file descriptors of frequently used logs. This improves the performance of log storage to the path that contains the variables. For more information, see open_log_file_cache.

    Recommended configurations:

    ```
    open_log_file_cache max=1000 inactive=1m valid=3m min_uses=2;
    ```

Why does an SMB file system have I/O latency?

- Issue

  If you access a Server Message Block (SMB) file system by using a mount target, you must wait for several minutes before you can perform I/O operations on the file system.

- Cause
  - You must wait for several minutes because a Network File System (NFS) client is installed but not used.
  - The file server cannot access the SMB file system because the WebClient service is enabled.
  - The files in the file system cannot be opened because `Nfsnp` is included in the value of the ProviderOrder key.

- Solution
  i. When you access an SMB file system for the first time, **ping** the domain name of the mount target to check the network connectivity between the compute node and the file system. You can also check whether the latency is within the allowed range.
     - If the **ping** command fails, check your network settings and make sure that the network is connected.
     - If the latency is high, run the **ping** command to ping the IP address of the mount target. If the latency of accessing the IP address is lower than the latency of accessing the domain name, check the configurations of the Domain Name System (DNS) server.

  ii. If an NFS client is installed but not used, we recommend that you delete the NFS client.

  iii. Disable the WebClient service.

  iv. Check the Registry key in the following path: *HKEY_LOCAL_MACHINE\System\CurrentControlSet\Control\NetworkProvider\Order\ProviderOrder*. If the value of the ProviderOrder key contains `Nfsnp`, remove `Nfsnp` and restart the ECS instance on which the file system is mounted.

Note

- You can use the fio tool to check the performance of the file system.

```
fio.exe --name=./iotest1 --direct=1 --rwmixread=0 --rw=write --bs=4K --numjobs=1 --thread --iodepth=128 --runtime=300 --group_reporting --size=5G --verify=md5 --randrepeat=0 --norandommap --refill_buffers --filename=\\<mount point dns>\myshare\testfio1
```

- We recommend that you perform read and write operations based on large data blocks. Small data blocks consume more network resources. If you cannot modify the block size, you can construct the BufferedOutputStream class to write data to a specified output stream with a specified buffer size.

Why am I unable to improve the I/O performance of Windows SMB clients?

- Cause

  By default, the `large MTU` feature is disabled on Windows SMB clients. This limits the I/O performance of Windows SMB clients.

- Solution

  You can enable the `large MTU` feature by modifying the Windows Registry. The Registry key is stored in the following path: *HKEY_LOCAL_MACHINE\System\CurrentControlSet\Services\LanmanWorkstation\Parameters*.

  Create a key of the `DWORD` data type and name the key as DisableLargeMtu. Set the value of the key to `0`. Restart the ECS instance on which the file system is mounted to validate the key.

How can I improve the performance for access from IIS to NAS?

- Cause

  When Internet Information Service (IIS) accesses a file in the shared directory of a NAS file system, the backend of IIS may access the shared directory multiple times. When you access the NAS file system, you must interact with the network at least once. This is different from the scenario when you access a local disk. Although each access request does not take a long time, the client may take a long time to respond if multiple access requests are sent.

- Solution

  i. Use the SMB Redirector component to optimize the performance of SMB file systems. For more information, see SMB2 Client Redirector Caches Explained.

  Modify the Registry keys in the following path: *HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\services\LanmanWorkstation\Parameters*. Modify the values of the following three keys to 600 or higher:

  - FileInfoCacheLifetime
  - FileNotFoundCacheLifetime
  - DirectoryCacheLifetime

> **Note** If none of the preceding keys exists, perform the following operations:
>
> a. Make sure that the file system uses the SMB protocol.
>
> b. Check whether the Windows version supports the keys. If the Windows version supports the keys but the keys do not exist, create the keys. For more information, see Performance tuning for file servers.

ii. We recommend that you store the web-related files such as JS and CSS files to local disks if IIS frequently accesses these files.

If the read and write performance of IIS cannot meet your business requirements, submit a ticket.

Why is the performance of an NFS client poor on Linux?

- Issue

  The read and write speed of an NFS client on Linux is only several MB/s.

- Cause

  By default, a maximum of 2 concurrent NFS requests can be sent from a Linux-based NFS client. This limits the performance of the NFS client.

- Solution

  After an NFS client is installed, modify the maximum number of concurrent NFS requests to improve the performance of the NFS client. For more information, see How can I modify the maximum number of concurrent NFS requests?.

How do I increase the read and write bandwidth of a General-purpose NAS file system?

The read and write bandwidth of a General-purpose NAS file system linearly increases with the capacity of the file system. For more information, see General-purpose NAS file systems.

You can increase the capacity of the file system by writing hole files to the file system. Then, the read and write bandwidth of the file system is increased. File holes occupy storage space in NAS. Therefore, you are charged for the file holes in your NAS file system. For more information, see Billing overview of General-purpose NAS.

For example, if you write a 1,000 GiB hole file to a Capacity NAS file system, you can increase the read and write bandwidth of the file system by 150 MB/s. If you write a 1,000 GiB hole file to a Performance NAS file system, you can increase the read and write bandwidth of the file system by 600 MB/s.

- Linux

  ```
  dd if=/dev/zero of=/mnt/sparse_file.txt bs=1073741824000 count=1
  ```

  In the preceding command, */mnt* is the mount path of the file system on the compute node.

- Windows

  ```
  fsutil file createnew Z:\sparse_file.txt 1073741824000
  ```

  In the preceding command, *Z:* is the mount path of the file system on the compute node.

Why does a file system give a slow response or does not respond when I run the ls command?

- Issue

  When you traverse a directory of a file system, the file system gives a slow response or does not respond. This may occur when you run the **ls** command or commands that contain wildcards such as `*` and `?`.It may also occur when you run the `rm -rf` command, or when the getdents system call is invoked.

- Cause
  - The directory is being modified. For example, a directory is being created or deleted, or a file in the directory is being renamed. This leads to slow responses due to frequent cache invalidations.
  - The data size of the directory is too large. This leads to slow responses due to cache eviction.

- Solution
  - Limit the number of files stored in the directory. Make sure that you can store less than 10,000 files in a single directory.
  - Do not frequently modify the directory when you traverse the directory.
  - If you store more than 10,000 files in the directory and the directory does not need to be frequently modified, you can accelerate the traversal process to some extent. However, you must make sure that you mount the file system by using the NFSv3 protocol and use the *nordirplus* parameter. For more information, see Mount parameters.