Persistent disk performance

Send feedback

Contents

Understand your workload
Choose a storage option
Calculate the price per IOP for a persistent disk
Configure your disks and instances
...

To configure a storage option for apps that run on your instances, use the following process:

Understand your workload

For workloads that primarily involve small (4 kB to 16 kB) random I/Os, the limiting factor is random input/output operations per second (IOPS) [1].

For workloads that primarily involve sequential or large (256 kB to 1 MB) random I/Os, the limiting factor is throughput [2].

Choose a storage option

You can provide several different types of block storage for your instances to use. Each type has different price, performance, and durability characteristics. See Storage options for a full comparison.

When you configure a zonal or regional persistent disk, you must select one of the following disk types:

- **Standard persistent disks** (`pd-standard`) are suited for large data processing workloads that primarily use sequential I/Os.
- **Balanced persistent disks** (`pd-balanced`) are an alternative to SSD persistent disks that balance performance and cost. With the same maximum IOPS as SSD persistent disks and lower IOPS per GB, balanced PD offers performance levels suitable for most general purpose applications at a price point between that of standard and SSD persistent disks.
- **SSD persistent disks** (`pd-ssd`) are suited for enterprise applications and high-performance database needs that require lower latency and more IOPS than standard persistent disks provide. SSD persistent disks are designed for single-digit millisecond latencies; the observed latency is sub-millisecond.

	Zonal standard PD	Regional standard PD	Zonal balanced PD	Regional balanced PD	Zonal SSD PD	Regional SSD PD	Local SSD (NVMe)	Local SSD (SCSI)
Maximum sustained IOPS								
Read IOPS per GB	0.75	0.75	6	6	30	30	–	–
Write IOPS per GB	1.5	1.5	6	6	30	30	–	–
Read IOPS per instance	7,500†	3,000†	15,000†	9,000†	15,000† 60,000‡	15,000† 60,000‡	900,000† 100,000‡	450,000† 100,000‡
Write IOPS per instance	15,000†	15,000†	15,000† 60,000‡	15,000† 60,000‡	15,000† 60,000‡	15,000† 60,000‡	800,000†	1,200,000†
Maximum sustained throughput (MB/s)								
Read throughput per GB	0.12	0.12	0.28	0.28	0.48	0.48	–	–
Write throughput per GB	0.12	0.12	0.28	0.28	0.48	0.48	–	–
Read throughput per instance	240 / 1,200†	240 / 240†	240 / 1,200†	240 / 240†	240 / 1,200†	240 / 1,200†	9,360	9,360
Write throughput per instance	76 / 400†	38 / 240†	240 / 400†	240 / 240†	120 / 400†	240 / 1,200†	120 / 400†	4,680

† Persistent disk IOPS and throughput performance depends on disk size, instance vCPU count, and I/O block size, among other factors.
‡ Persistent disks can achieve greater throughput performance on instances with more vCPUs. Read fsync ops cap on write throughput.

Attaching a disk to multiple virtual machines does not affect aggregate performance or cost. Each machine gets a share of the per-disk performance limit.

Note that SSD read bandwidth and IOPS consistency near the maximum performance limits depend largely on network ingress utilization. Some variability in the performance limits is to be expected, especially when operating near the maximum IOPS limits with an I/O size of 16 kB.

Calculate the price per IOP for a persistent disk

Persistent disk has no per-I/O costs, so there is no need to estimate monthly I/O to calculate budget for what you will spend on disks. However, for IOPS-oriented workloads, it is possible to break down the per-month cost to look at price per IOPS, for comparison purposes.

The following calculation examples use U.S. persistent disk pricing. Consider the relative costs of standard persistent disks compared to SSD persistent disks. For example, in the zone `us-central1`, standard persistent disks are priced at $0.040 per GB and SSD persistent disks are priced at $0.170 per GB. When you increase the size of a volume, you also increase the performance caps automatically, at no additional cost.

To determine the cost per IOPS of a persistent disk, divide the price per GB per month with the number of IOPS per GB. The following table calculates the price per random read IOPS per GB. You can use the same calculations to calculate the price per write IOPS as well.

	Price per GB per month	Read IOPS per GB	Price per IOPS per GB
Standard persistent disk	$0.040	0.75	$0.040 / 0.75 = $0.0533
Balanced persistent disk	$0.100	6	$0.100 / 6 = $0.0167
SSD persistent disk	$0.170	30	$0.170 / 30 = $0.0057

Learn more about persistent disk pricing and local SSD pricing.

Configure your disks and instances

To maximize performance, configure the correct disk size, vCPU count, and machine type.

Compare persistent disk to a physical hard drive

When you specify the size of your persistent disks, consider how these disks compare to traditional physical hard drives. The following tables compare standard persistent disks and SSD persistent disks to the typical performance that you would expect from a 7200 RPM SATA drive, which typically achieves 75 IOPS or 120 MB per second.

I/O type	I/O pattern	Size required to match a 7200 RPM SATA drive (GB)		
		Standard persistent disk	Balanced persistent disk	SSD persistent disk
Small random reads	75 small random reads	100	12	3
Small random writes	75 small random writes	50	12	3
Streaming large reads	120 MB/s streaming reads	1,200	428	250
Streaming large writes	120 MB/s streaming writes	1,200	428	250

Disk size and vCPU count

Persistent disk performance scales with the size of the disk. Performance also depends on the number of vCPUs on your VM instance due to network egress caps on write throughput.

Performance scales linearly until it reaches either the limits of the disk or the limits of the instance to which the disk is attached. Limits of the instance to which the disk is attached are based on the machine type and the number of vCPUs on the instance. To learn how persistent disk performance depends on instance limits, see Machine types and vCPU counts.

For example, consider a 1,000 GB standard persistent disk attached to an instance with an N2 machine type and 4 vCPUs. The read limit based solely on the size of the disk is 0.75(1,000) = 750 IOPS. However, because the instance has 4 vCPUs, the read limit is restricted to 15,000 IOPS.

The following tables show performance by disk size in increments where performance changes significantly. However, you can specify a disk size in any 1 GB increment.

Standard persistent disk performance by disk size

Disk size (GB)	IOPS				Sustained throughput (MB/s)	
	Read (<=16 kB per I/O)	Write (<=8 kB per I/O)	Write (>8 kB per I/O)		Read	Write
10	*	*	*	*	*	*
32	24	24	48	24	0.37	3
64	48	48	96	96	7	7
128	96	192	192	192	15	15
256	192	384	384	384	30	30
512	384	768	768	768	61	61
1,000	750	1,500	1,500	1,500	120	120
1,500	1,125	2,250	2,250	2,250	180	180
2,048	1,536	3,072	3,072	3,072	245	245
4,000	3,000	6,000	6,000	6,000	480	400
5,000	3,750	7,500	7,500	7,500	600	400
8,192	6,144	12,288	7,500	7,500	983	400
10,000–65,536	7,500	15,000	7,500	7,500	1,200	400

* Use the disk size only for boot disks. I/O bursting provides higher performance for boot volumes than the linear scaling described here.

SSD persistent disk performance by disk size

Disk size (GB)	IOPS				Sustained throughput (MB/s)	
	Read (<=4 kB per I/O)	Read (>=16 kB per I/O)	Write (<=8 kB per I/O)	Write (16 kB per I/O)	Read	Write
10	300	300	300	300	4.8	4.8
32	960	960	960	960	15	15
64	1,920	1,920	1,920	1,920	30	30
128	3,840	3,840	3,840	3,840	61	61
256	7,680	7,680	7,680	7,680	122	122
500	15,000	15,000	15,000	15,000	240	240
834	25,000	25,000	25,000	25,000	400	400
1,000	30,000	30,000	30,000	25,000	480	400
1,334	40,000	40,000	30,000	25,000	640	400
1,667	50,000	50,000	30,000	25,000	800	400
2,048	60,000	60,000	30,000	25,000	983	400
3,000–65,536	100,000	100,000	30,000	25,000	1,200	400

Balanced persistent disk performance by disk size

Balanced PD scales linearly with disk size, at a rate of 6 IOPS per GB and 0.28 MB per second for both reads and writes.

Machine type and vCPU count

In Compute Engine, machine types are grouped and curated for different workloads. Each machine type is subject to specific persistent disk limits per vCPU. Performance by disk size is the same across machine types.

- **See performance by vCPU count for E2 VMs**
- **See performance by vCPU count for N1, N2, and N2D VMs**
- **See performance by vCPU count for C2 VMs**

Consider factors that affect performance

Network egress caps on write throughput

Your virtual machine (VM) instance has a network egress cap that depends on the machine type of the VM.

Compute Engine stores data on persistent disks with multiple parallel writes to ensure high durability and redundancy. Additionally, each write request has some overhead that uses additional write bandwidth.

The maximum write traffic that a VM instance can send to the network egress is divided by a bandwidth multiplier that accounts for the write bandwidth used by this redundancy and overhead. The network egress cap depends on the machine type of the VM instance. The network egress caps for each machine type are listed in the Machine type table in the **Network egress bandwidth (Gbps)** column.

In a situation where persistent disk is competing with IP traffic for network egress bandwidth, 60% of the **maximum write bandwidth** goes to persistent disk traffic, leaving 40% for IP traffic.

Click below to see an example of how to calculate the maximum persistent disk write traffic that a VM instance can issue.

- **See example expected persistent disk write bandwidth with and without additional IP traffic**

Simultaneous reads and writes

For standard persistent disks, simultaneous reads and writes share the same resources. While your instance is using more read throughput or IOPS, it is able to perform fewer writes. Conversely, instances that use more write throughput or IOPS are able to perform fewer reads.

SSD and balanced persistent disk are capable of achieving maximum throughput limits for both reads and writes simultaneously. However, it is not possible for SSD and balanced persistent disks to reach their maximum IOPS limits for reads and writes simultaneously.

Note that **throughput + IOPS * I/O size**. To take advantage of maximum throughput limits for simultaneous reads and writes on SSD persistent disks, use an I/O size such that read and write IOPS combined don't exceed the IOPS limit.

- **See instance limits for simultaneous reads and writes**

Logical volume size

Persistent disks can be up to 64 TB in size, and you can create single logical volumes of up to 257 TB using logical volume management inside your VM. A larger volume size impacts performance in the following ways:

- Not all local file systems work well at this scale. Common operations, such as mounting and file system checking might take longer than expected.
- Maximum persistent disk performance is achieved at smaller sizes. Disks take longer to fully read or write with this much storage on one VM. If your application requires this much storage, consider using multiple VMs for greater total system throughput.
- Snapshotting large amounts of persistent disk might take longer than expected to complete and might provide an inconsistent view of your logical volume without careful coordination with your applications.

Multiple disks on a single VM instance

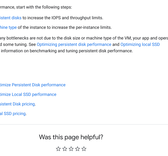
Multiple disks of the same type

Suppose you have multiple disks of the same type (standard or SSD) attached to the same inode (for example, read/write). If you use only one disk, then that single disk can reach the performance limit corresponding to the combined size of the disks. If you split all of the disks at 100%, the aggregate performance limit is split evenly among the disks regardless of relative disk size.

For example, suppose you have a 200 GB standard disk and a 1,000 GB standard disk.

If you do not use the 1,000 GB disk, then the 200 GB disk can reach the performance limit of a 200 GB standard disk. If you use both disks at 100%, then each will have the performance limit of a 200 GB standard disk (1,200 GB / 2 disks = 600 GB disk).

To demonstrate this, consider the following instances `instance-1` has one 200 GB standard persistent disk and one 1,000 GB standard persistent disk. `instance-1` has one CPU and 3.75 GB of memory. Both of the two instances have the same configuration.



The following commands test read IOPS for the 200 GB disk on `instance-1` and for the 200 GB disk on `instance-2`, separately.

```
sudo fioman read-iops-random --device=/dev/sdb --rwmixpct=0 --numjobs=8 --iodepth=16 --ioengine=libaio
Run read_iops: jobs: percentage: 0=0,10=0,20=0,30=0,40=0,50=0,60=0,70=0,80=0,90=0,100=0
read_iops: (g=0): rw=randrw, bs=(R) 4KiB-4KiB, (W) 4KiB-4KiB, ioengine=libaio, iodepth=16
...
fio-3.1-Starting 8 threads
read_iops: (groupid=0, jobs=8): err= 0: pid=2470: Wed Jun 12 21:17:15 2019
  read: IOPS=1500, BW=6000KiB/s (6144kB/s)(1000MiB/170600msec)
    slat (usec): min=2, max=1000, avg=5.33, stdev=3.00
    clat (msec): min=1, max=1000, avg=84.00, stdev=0.00
     lat (msec): min=1, max=1000, avg=84.00, stdev=0.00
   bw (  KiB/s): min=  4, max=  8, per=100.00%, avg=  6.00, stdev=  0.00, samples=340000
   iops        : min=  1, max=  2, per=100.00%, avg=  1.50, stdev=  0.00, samples=340000
  cpu          : usr=0.00%, sys=0.00%, ctx=1000000, majf=0, minf=0
  IO depths    : 1=100.0%, 2=0.0%, 4=0.0%, 8=0.0%, 16=0.0%, 32=0.0%, >=64=0.0%
     submit    : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, >=64=0.0%
     complete  : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, >=64=0.0%
     issued rwt: total=1000000,0,0, short=0,0,0, dropped=0,0,0
     latency   : target=0, window=0, percentile=100.00%, depth=16

Run status group 0 (all jobs):
   READ: bw=6000KiB/s (6144kB/s), 6000KiB/s-6000KiB/s (6144kB/s-6144kB/s), io=1000MiB (1048576MB), run=170600-170600msec
```

```
sudo fioman read-iops-random --device=/dev/sdb --rwmixpct=0 --numjobs=8 --iodepth=16 --ioengine=libaio
Run read_iops: jobs: percentage: 0=0,10=0,20=0,30=0,40=0,50=0,60=0,70=0,80=0,90=0,100=0
read_iops: (g=0): rw=randrw, bs=(R) 4KiB-4KiB, (W) 4KiB-4KiB, ioengine=libaio, iodepth=16
...
fio-3.1-Starting 8 threads
read_iops: (groupid=0, jobs=8): err= 0: pid=2470: Wed Jun 12 21:17:15 2019
  read: IOPS=1500, BW=6000KiB/s (6144kB/s)(1000MiB/170600msec)
    slat (usec): min=2, max=1000, avg=5.33, stdev=3.00
    clat (msec): min=1, max=1000, avg=84.00, stdev=0.00
     lat (msec): min=1, max=1000, avg=84.00, stdev=0.00
   bw (  KiB/s): min=  4, max=  8, per=100.00%, avg=  6.00, stdev=  0.00, samples=340000
   iops        : min=  1, max=  2, per=100.00%, avg=  1.50, stdev=  0.00, samples=340000
  cpu          : usr=0.00%, sys=0.00%, ctx=1000000, majf=0, minf=0
  IO depths    : 1=100.0%, 2=0.0%, 4=0.0%, 8=0.0%, 16=0.0%, 32=0.0%, >=64=0.0%
     submit    : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, >=64=0.0%
     complete  : 0=0.0%, 4=100.0%, 8=0.0%, 16=0.0%, 32=0.0%, >=64=0.0%
     issued rwt: total=1000000,0,0, short=0,0,0, dropped=0,0,0
     latency   : target=0, window=0, percentile=100.00%, depth=16

Run status group 0 (all jobs):
   READ: bw=6000KiB/s (6144kB/s), 6000KiB/s-6000KiB/s (6144kB/s-6144kB/s), io=1000MiB (1048576MB), run=170600-170600msec
```

The observed read IOPS for the 200 GB disk on `instance-1` is 902, which matches the expected read IOPS level for a combined disk size of 200 GB + 1,000 GB = 1,200 GB.

Multiple disks of different types

If you have multiple disks of different types attached to a single VM, then the total performance limit for the VM is determined by the SSD disk type. The total performance limit is shared between all disks attached to the VM.

For example, suppose you have one 5,000 GB Standard disk and one 1,000 GB SSD disk attached to an N2 VM with one vCPU. The read IOPS limit for the Standard disk is 3,000 and the read IOPS limit for the SSD disk is 15,000. Because the limit of the SSD disk determines the overall limit, the total read IOPS limit for your VM is 15,000. This limit is shared between all attached disks.

Review performance metrics

You can review persistent disk performance metrics in Cloud Monitoring, Google Cloud's integrated monitoring solution.

Several of these metrics are useful for understanding if and when your disks are being throttled. Throttling is intended to smooth out bursty I/Os. With throttling, bursty I/Os can be spread out over a period of time, so that the performance limits of your disk can be met but not exceeded at any given instant.

To learn more, see Reviewing persistent disk performance metrics.

Optimize disk performance

To increase disk performance, start with the following steps:

- Resize your persistent disks to increase the IOPS and throughput limits.
- Change the machine type of the instance to increase the per-instance limits.

After you ensure that any bottlenecks are not due to the disk size or machine type of the VM, your app and operating system might still need some tuning. See Optimizing persistent disk performance and Optimizing local SSD performance for more information on benchmarking and tuning persistent disk performance.

What's next

- Learn how to optimize Persistent Disk performance.
- Learn how to optimize Local SSD performance.
- Learn about Persistent Disk pricing.
- Learn about Local SSD pricing.

Was this page helpful?

Send feedback

Except as otherwise noted, the content of this page is licensed under the Creative Commons Attribution 4.0 License, and code samples are licensed under the Apache 2.0 License. For details, see the Google Developers Site Policies. Java is a registered trademark of Oracle and/or its affiliates.

Last updated 2020-08-04 UTC.