

块存储数据迁移—降低成本+提升性能

AWS EBS 通用型卷 GP3 于 2021 年 1 月即在亚马逊云科技中国北京区和宁夏区推出，距今已有一年多时间。作为下一代通用型 SSD 卷，GP3 使客户不仅能够独立于存储容量来预置性能，并且每 GB 价格点比现有 GP2 优惠多达 20%。

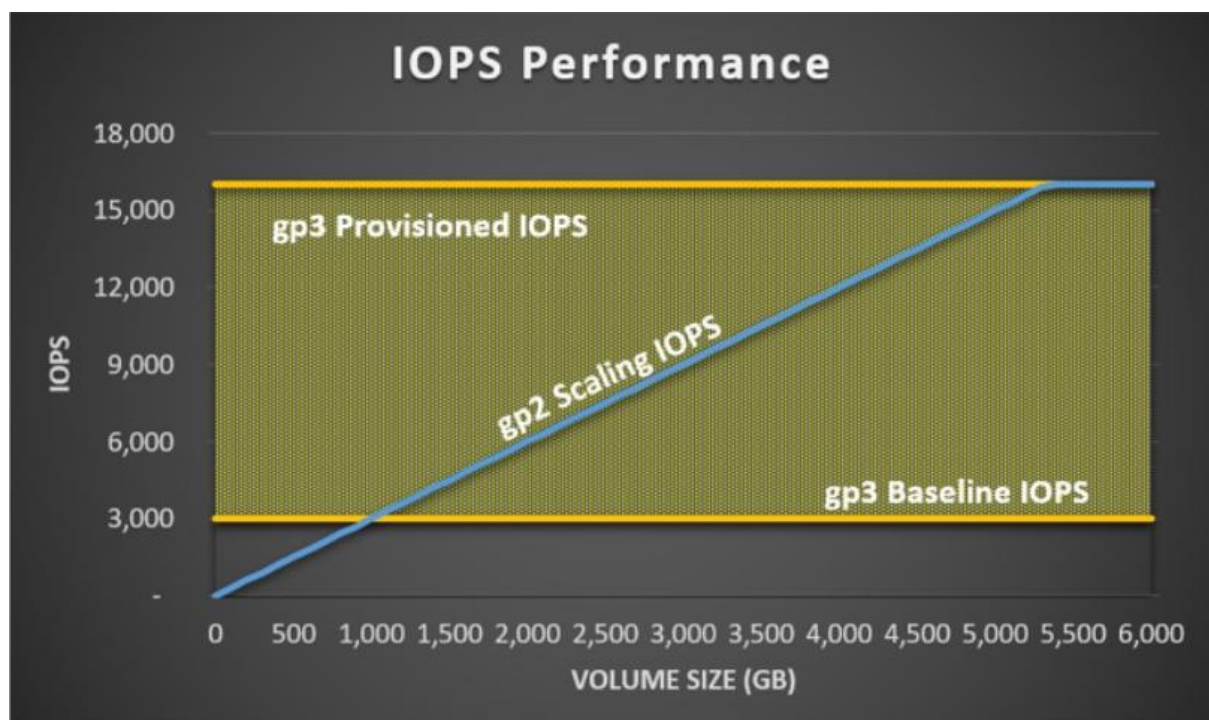
GP2 vs GP3

首先，通过 GP2、GP3 在性能、价格上做详细对比，可以看到 GP3 作为下一代通用型 SSD 卷能够为客户提供更高的性价比：

性能

GP2 能够提供的 IOPS 在 100 ~ 16000 之间，基准性能以每 GiB 卷大小 3 IOPS 的速度线性扩展，对于 1TB 以下的卷提供通过积分累计实现一段时间可以突增至 3000IOPS 的能力。因此，GP2 卷的性能与卷的大小相关联，卷越大，基准性能级别就越高，I/O 积分积累速度也越快。基于这一关联，客户在使用 GP2 时为获得较高的 IOPS，往往需要预制较大的磁盘容量，导致成本偏高。

而 GP3 的推出正是解耦了 IO 性能与卷大小的关联关系。GP3 提供了一致的基准速率，即 3,000 IOPS 和 125 MiB/s，基准速率包含在存储价格中。在此基础上，客户还可以独立设置 IOPS（高达 16,000 个）和吞吐量（高达 1,000 MiB/s），并根据实际的预置容量付费。在 GP3 中，预置 IOPS 与预置卷大小的最大比率为 500 IOPS / GiB。也就是说，如果需要 16000IOPS，在 GP3 只需要 32GB 存储就可达到，而对 GP2 则需要 5.3TB。



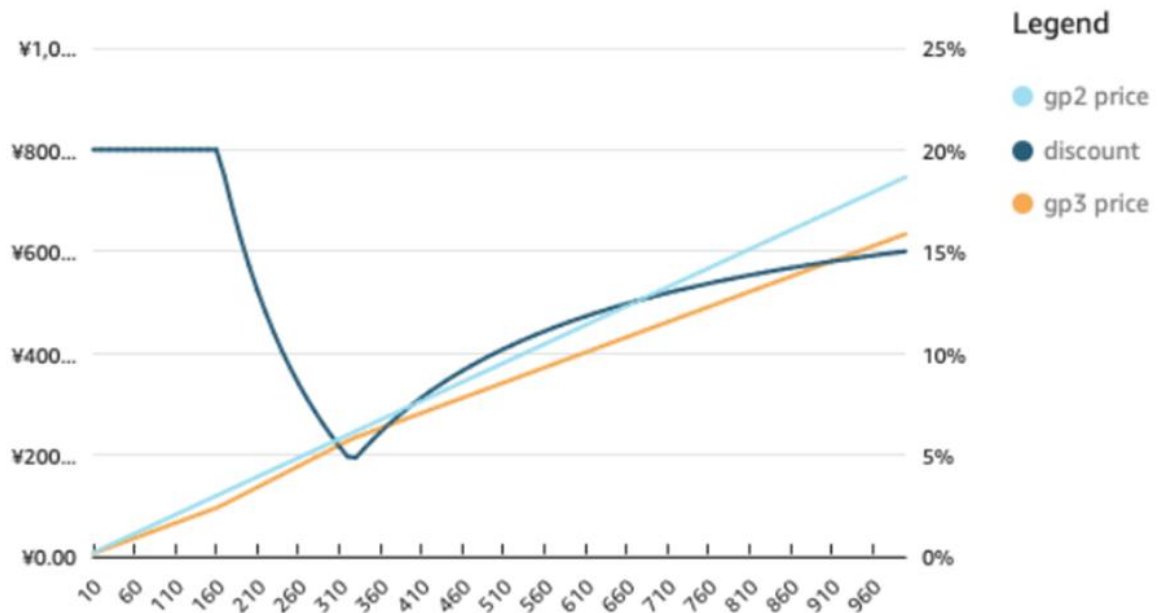
价格

目前, GP3 每 GB 的基础价格较 GP2 降低 20%, 同时能提供 4 倍于 GP2 的最大 MPS。下面以提供相同性能的情况下, 对 GP2 和 GP3 的价格进行对比, 可以看到 GP3 总可以提供较优的价格。

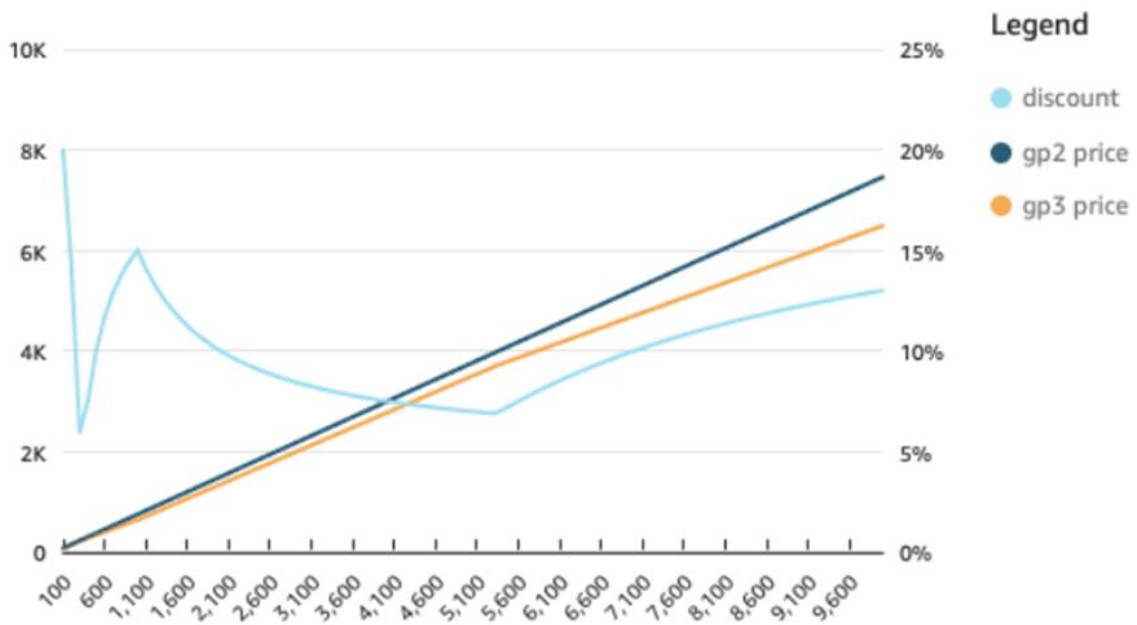
IOPS	3IOPS*卷大小(GB)	3000+预置 IOPS
MPS	卷大小(GB) * 256KB*3IOPS	125MPS+预置 MPS
最大 MPS	250MB/s	1000MB/s
	GP2	GP3

180GB 之内, GP3 比 GP2 提供 20% 的优惠。之后优惠比率开始降低, 是因为 GP3 需要为超出基准的 MPS 付费。在超过 320GB 之后, 优惠比率从 5% 开始回升至 1TB 的 15%, 是因为 GP2 在 1TB 可达到 3000IOPS, 而 GP3 在 IOPS 为 3000 之内都是免费。在超过 1TB 后, GP3 将为超过 3000 的 IOPS 付费。

1TB内gp2 vs gp3 价格对比



10TB gp2 vs gp3 价格对比



应用的当前配置

磁盘用途	当前大小	需要的最大 IOPS
data	2TB	<6000
log	2TB	<6000
tmp	2TB	<6000

在我们实施迁移的实际案例中，客户的生产环境拥有上百套数据库，为 OLTP 交易型负载。为保证磁盘读写性能，每套库起初统一配置了 3 块 2TB 大小存储，分别用于数据卷、日志卷和临时表空间，用于提供最大 6000IOPS/卷的读写性能。在推出 GP3 后，从成本优化的角度开始规划存储类型的迁移。

迁移到 GP3

迁移需要考虑的问题

考虑到本次迁移涉及的存储量较大，生产环境中为保证迁移的顺利进行，在迁移之前我们做了如下准备工作：

1. 确认每账户的存储限额，必要时开启案例申请提限，包括：Total gp3 Storage Needed (in TiB/account) — 每账户所需的 GP3 存储容量

Max Modifying General Purpose SSD (gp2) Provisioned Storage in TiB — 修改过程中处于修改状态的 GP2 存储容量

Max Modifying General Purpose SSD (gp3) Provisioned Storage in TiB — 修改过程中处于修改状态的 GP3 存储容量

比如我们本次迁移涉及超过 1PB 的数据，计划通过 3 个周期分别做迁移，则每账户所需的 GP3 存储容量为 1PB，修改过程中处于修改状态的 GP2、GP3 的存储容量则分别为 350TB。

2. 控制单次执行磁盘修改的 API 调用次数，以保证不被 API 限流。

本次迁移，除了上面提到的数据库涉及的磁盘，还包括诸多应用、操作系统使用的 EBS 盘，迁移盘的数量多达 5000 多块。我们在迁移过程中，单次执行修改控制在 600-800 块盘，避免造成 API 限流。

3. 选择业务低峰期执行磁盘修改

GP2 到 GP3 的修改支持在线进行，不影响业务的正常运行。磁盘类型在修改过程中磁盘可以对外提供的性能介于修改前和修改后磁盘指定的性能。通常建议选择业务低峰期执行大批量的磁盘修改。

4. 对同一个卷的持续修改不能少于 6 小时。

修改卷后，须等待至少六个小时并确保卷处于 in-use 或 available 状态，然后再对同一个卷进行其他修改。因此在修改卷时，需要避免对同一卷做过于频繁连续修改。

成本优化

将磁盘类型从 GP2 修改为 GP3 时，我们在迁移的第一阶段首先选择保持磁盘大小不变，只将磁盘类型修改为 GP3，并设定 IOPS 为 6000。这一迁移操作可以在不影响上层应用的情况下在线进行，且能够提供与原配置完全一致的性能。

通过这一简单修改, 以当前北京区域为例, 每块 2TB 的磁盘每月成本节省近 190 元, 上百套数据库共涉及 500 多块盘超过 1PB 的数据, 一年仅转换 GP3 就可以实现成本节省超过 100 万人民币。