

AEP新硬件测试 -- Redis

本文记录了使用新硬件环境下 Redis 的性能测试对比结果。

环境说明

OS:

CentOS Linux release 7.4.1708 (Core)

Linux version 4.18.8 (gcc version 4.8.5 20150623 (Red Hat 4.8.5-16))

Memory: 256G

AEP: 256G*2*2

Redis:

Base: 4.0.0

Pmem-Redis: base on redis-4.0.0

机器:

机器 A:

机器 B: (内存写入性能优于机器 A)

测试情况: 单机单实例测试

测试使用 redis-benchmark 和 memtier-benchmark 工具分别进行测试。

注意: 测试过程中均关闭了 aof&rdb 持久化, 避免持久化的影响。

测试结果

1.redis-benchmark -n 2000000 -c 30 -d 128 -r 100000000

以机器 B 作为 base, 使用 redis 自带的 redis-benchmark 为测试工具, 分别选取 Pmem-Redis 和 Redis-4.0.0 作为测试对象。

value 长度 128Bytes		
	原生 Redis	Pmem-Redis
PING_INLINE	122122.48 (<1ms 100 分位)	121758.19 (<1ms 100 分位)
PING_BULK	119381.60 (<1ms 100 分位)	120019.20 (<1ms 100 分位)
SET	129895.43 (<2ms	109517.03 (<3ms 100 分位)

value 长度 128Bytes		
	100 分位)	
GET	128949.06 (<1ms 100 分位)	131882.62 (<1ms 100 分位)
INCR	132828.58 (<5ms 100 分位)	132178.97 (<6ms 100 分位)
LPUSH	124649.42 (<1ms 100 分位)	123708.79 (<1ms 100 分位)
RPUSH	120307.98 (<1ms 100 分位)	125352.55 (<1ms 99.8 分位)
LPOP	119846.60 (<1ms 100 分位)	124828.37 (<1ms 100 分位)
RPOP	124984.37 (<1ms 100 分位)	127820.02 (<1ms 100 分位)
SADD	134662.00 (<3ms 100 分位)	136360.53 (<3ms 100 分位)
SPOP	138773.25 (<1ms 100 分位)	137306.06 (<1ms 100 分位)
HSET	133564.84 (<2ms 100 分位)	110235.35 (<7ms 100 分位)

value 长度 128Bytes		
MSET(10keys)	72629.55 (<38ms 100 分位)	21932.95 (<5ms 99.99 分位, <42ms 100 分位)
LRANGE_100	56216.09 (<2ms 100 分位)	55532.42 (<1ms 100 分位)
value 长度 1k Bytes		
	原生 Redis	Pmem-Redis
PING_INLINE	116767.87 (<1ms 100 分位)	165494.41 (<1ms 100 分位)
PING_BULK	115794.35 (<1ms 100 分位)	142257.62 (<1ms 100 分位)
SET	121477.16 (<3ms 100 分位)	95643.45 (<2ms 100 分位)
GET	128849.38 (<2ms 100 分位)	140864.91 (<1ms 100 分位)
INCR	130429.12 (<5ms 100 分位)	140449.44 (<4ms 99.93 分位, <12ms 100 分位)
LPUSH	115834.59 (<1ms 100 分位)	132424.03 (<1ms 100 分位)
RPUSH	112599.94 (<1ms 100 分位)	134961.88 (<2ms 100 分位)
LPOP	120999.46 (<1ms 100 分位)	137684.16 (<1ms 100 分位)
RPOP	124820.56 (<1ms 100 分位)	135034.77 (<1ms 100 分位)
SADD	133877.77 (<1ms 100 分位)	137343.77 (<3ms 100 分位)

value 长度 128Bytes		
SPOP	137164.80 (<1ms 100 分位)	138869.59 (<1ms 100 分位)
HSET	124766.05 (<8ms 100 分位)	97575.26 (<7ms 100 分位)
MSET(10keys)	54869.68 (<8ms 99.99 分位, <40ms 100 分位)	14677.50 (<9ms 99.99 分位, <43ms 100 分位)
LRANGE_100	13659.52 (<6ms 99.99 分位)	13698.82 (<5ms 100 分位)
value 长度 4k Bytes		
	原生 Redis	Pmem-Redis
PING_INLINE	121580.54 (<1ms 100 分位)	126095.45 (<1ms 100 分位)
PING_BULK	116022.73 (<1ms 100 分位)	120221.21 (<1ms 100 分位)
SET	99472.79 (<3ms 100 分位)	92148.91 (<3ms 100 分位)
GET	128832.77 (<1ms 100 分位)	132213.92 (<1ms 100 分位)
INCR	133005.25 (<2ms 100 分位)	133031.80 (<5ms 100 分位)
LPUSH	100462.12 (<1ms 100 分位)	111900.64 (<1ms 100 分位)
RPUSH	98338.09 (<1ms 99.99 分位)	113720.36 (<1ms 100 分位)
LPOP	121182.74 (<2ms 100 分位)	104047.45 (<1ms 100 分位)
RPOP	121036.07 (<2ms 100 分位)	106496.27 (<1ms 100 分位)

value 长度 4k Bytes		
SADD	132459.11 (<3ms 100 分位)	137485.39 (<3ms 100 分位)
SPOP	135804.98 (<1ms 100 分位)	139382.53 (<1ms 100 分位)
HSET	101579.57 (<4ms 100 分位)	86718.98 (<7ms 100 分位)
MSET (10keys)	26868.72 (<4ms 99.99 分位, <41ms 100 分位)	11758.90 (<5ms 99.90 分位, <45ms 100 分位)
LRANGE_100	3141.18 (<6ms 99.99 分位, <13ms 100 分位)	3288.34 (<10ms 100 分位)
value 长度 10k Bytes		
	原生 Redis	Pmem-Redis
PING_INLINE	115874.85 (<1ms 100 分位)	120736.49 (<1ms 100 分位)
PING_BULK	112815.88 (<1ms 100 分位)	117000.12 (<1ms 100 分位)
SET	82481.03 (<3ms 100 分位)	74468.48 (<3ms 100 分位)
GET	128377.94 (<2ms 100 分位)	129794.27 (<1ms 100 分位)
INCR	134662.00 (<5ms 100 分位)	131518.38 (<5ms 100 分位)
LPUSH	82274.05 (<1ms 100 分位)	73967.23 (<1ms 100 分位)
RPUSH	82362.14 (<2ms 100 分位)	74109.76 (<1ms 100 分位)

value 长度 4k Bytes		
LPOP	96195.47 (<1ms 100 分位)	74421.38 (<12ms 100 分位)
RPOP	95790.03 (<2ms 100 分位)	74618.52 (<23 100 分位)
SADD	82552.52 (<1ms 100 分位)	134399.58 (<3ms 100 分位)
SPOP	139411.69 (<1ms 100 分位)	134834.48 (<1ms 100 分位)
HSET	82552.52 (<8ms 100 分位)	63021.90 (<7ms 100 分位)
MSET (10keys)	10730.72 (<13ms 99.99 分位, <23ms 100 分位)	7413.56 (<8ms 99.99 分位, <45ms 100 分位)
LRANGE_100	1130.18 (<26ms 100 分位)	1082.12 (<8ms 99.89 分位, <25ms 100 分位)

value 长度 30k Bytes		
	原生 Redis	Pmem-Redis
PING_INLINE	121160.72 (<1ms 100 分位)	140479.03 (<1ms 100 分位)
PING_BULK	115068.17 (<1ms 100 分位)	165289.25 (<1ms 100 分位)
SET	44594.08 (<3ms 100 分位)	19240.02 (<4ms 100 分位)
GET	127429.12 (<3ms 100 分位)	137598.91 (<1ms 100 分位)
INCR	132819.77 (<5ms 100 分位)	130625.04 (<6ms 100 分位)

value 长度 30k Bytes		
LPUSH	39000.04 (<2ms 100 分位)	20372.20 (<4ms 100 分位)
RPUSH	30472.01 (<3ms 100 分位)	20257.06 (<3ms 100 分位)
LPOP	40268.19 (<44ms 99.99 分位, <590ms 100 分位)	21363.18 (<3ms 99.99 分位, <208ms 100 分位)
RPOP	37863.73 (<7ms 99.98 分位, <390ms 100 分位)	21374.37 (<3ms 99.99 分位, <141ms 100 分位)
SADD	137296.62 (<6ms 100 分位)	133395.58 (<3ms 100 分位)
SPOP	134879.95 (<1ms 100 分位)	118941.41 (<2ms 100 分位)
HSET	33698.97 (<3ms 100 分位)	13172.37 (<9ms 100 分位)
MSET (10keys)	1906.24 (<77ms 100 分位)	1887.05 (<66ms 100 分位)
LRANGE_100	399.57 (<31ms 99.99 分位, <88ms 100 分位)	317.58 (<60ms 100 分位)

总结：（测试结果存在一定的性能波动，如 SADD 可能受重复 key 影响等）

以 redis-benchmark 的结果来看，Pmem-Redis 的性能受 value 大小的影响比较明显，随着 val 的不断增大，在 **set&hset&lpush&rpsh** 等写入操作上与原生基于内存的 redis 存在较大的性能差异，在读性能上与内存相差无几（与 Intel 宣称的 AEP 的读性能接近内存，写性能较差相符合）。

例如，在<1k 的场景下，内存态 redis 的写性能要优于 Pmem-redis 约在 16%~25%之间；在 1K~10k 的场景下，内存态 Redis 的写性能约超过 Pmem-Redis 在 20%~25%之间；在 30k 的场景下，内存态 redis 的 set&hset 性能要超过 Pmem-redis2 倍以上。

2.memtier_benchmark

memtier_benchmark 是 Redis-Lab 推的一种比 Redis-Benchmark 更灵活的测试工具，可配置读写比等特殊选项。

参考命令：

```
memtier_benchmark -s 127.0.0.1 -p 6379 -n 100000 -c 30 -R --data-size-range=1-1024 --data-size-pattern=S --key-minimum=200 --key-maximum=400
```

说明：以机器 A 作为 base 环境，分别以 memtier_benchmark 工具进行测试。

参数说明：

- 并发数：30
- 线程数：4
- 请求数：50w
- 读写比：1:1 4:1 10:1 1:4 1:10
- data-size：128bytes, 2kb, 5kb

set/get -- 1 : 1

data_size		128bytes		2kb		5kb	
		pmem-redis	redis-4.0	pmem-redis	redis-4.0	pmem-redis	redis-4.0
avg qps		112759	132059	103326	127297	84936	128660
avg latency		1.06	0.91	1.16 ms	0.94 ms	1.41 ms	0.93 ms
IOPS (MB/s)		13.09	15.32	133.49	161.38	263.76	399.85
set	Ops/sec	56380.09	66026.34	51664.24	63649.22	42463.16	64329.11
	Hits/sec	--	--	--	--	--	--
	Miss/sec	--	--	--	--	--	--
	Latency	1.064 ms < 1.5ms 99.99 分位 < 28ms 100 分位	0.957ms < 1.3ms 99.99 分位 < 13ms 100 分位	1.163ms < 1.9ms 99.99 分位 < 26ms 100 分位	0.942 ms < 1.4 ms 99.98 分位 < 13ms 100 分位	1.417 ms < 2.3ms 99.99 分位 < 26ms 100 分位	0.932 ms < 1.5ms 99.99 分位 < 13ms 100 分位
KB/sec		9620.00	11265.99	105736.29	130264.89	214294.82	324643.65
get	Ops/sec	56380.09	66026.34	51664.24	63649.22	42463.16	64329.11
	Hits/sec	14490.40	16934.00	14566.01	16372.16	10854.44	16506.10

	Miss/sec	41889.69	49092.34	37098.23	47277.05	31608.71	47823.01
	Latency	1.062 ms < 1.5ms 99.99 分位	0.907 ms < 1.3ms 99.99 分位	1.157ms < 1.8ms 99.98 分位	0.941 ms < 1.4 ms 99.98 分位	1.405 ms < 2.2ms 99.99 分位	0.930 ms < 1.5ms 99.99 分位
		< 3.4ms 100 分位	< 2.6ms 100 分位	< 3.6ms 100 分位	< 3.2ms 100 分位	< 4.1ms 100 分位	< 3.2ms 100 分位
	KB/sec	3781.28	4423.56	30961.50	34984.32	55766.10	84794.09
set/get -- 1 : 4							
data_size		128bytes		2kb		5kb	
		pmem-redis	redis-4.0	pmem-redis	redis-4.0	pmem-redis	redis-4.0
avg qps		123427	127651	113151	134290	105300	130553
avg latency		0.97 ms	0.94 ms	1.06 ms	0.89 ms	1.14 ms	0.92 ms
IOPS (MB/s)		8.76	9.07	68.61	81.35	154.03	190.68
set	Ops/sec	24685.09	25530.10	22630.29	26858.02	21060.16	26110.76
	Hits/sec	--	--	--	--	--	--
	Miss/sec	--	--	--	--	--	--

	Latency	0.973 ms < 1.4ms 99.99 分位 < 23ms 100 分位	0.939 ms < 1.3ms 99.99 分位 < 16ms 100 分位	1.064 ms < 1.6ms 99.99 分位 < 26ms 100 分位	0.893 ms < 1.3ms 99.98 分位 < 13ms 100 分位	1.148 ms < 1.8ms 99.99 分位 < 24ms 100 分位	0.919 ms < 1.4ms 99.99 分位 < 16ms 100 分位
	KB/sec	4211.97	4356.22	46315.32	54967.75	106282.31	131770.61
get	Ops/sec	98740.36	102120.42	90521.17	107432.10	84240.66	104443.03
	Hits/sec	11335.87	11842.58	10434.49	12345.61	9716.86	11988.47
	Miss/sec	87404.49	90277.84	80086.69	95086.49	74523.80	92454.56
	Latency	0.97 ms < 1.3ms 99.97 分位 < 3.3ms 100 分位	0.938 ms < 1.3ms 99.99 分位 < 2.6ms 100 分位	1.058 ms < 1.5ms 99.98 分位 < 3.2ms 100 分位	0.892 ms < 1.3ms 99.98 分位 < 2.8ms 100 分位	1.136 ms < 1.7ms 99.98 分位 < 3.7ms 100 分位	0.917 ms < 1.4ms 99.99 分位 < 3ms 100 分位
	KB/sec	4757.37	4936.04	23941.89	28337.89	51444.16	63487.29
set/get -- 1 : 10							

data_size		128bytes		2kb		5kb	
		pmem-redis	redis-4.0	pmem-redis	redis-4.0	pmem-redis	redis-4.0
avg qps		131783	135869	117818	122958	120378	129185
avg latency		0.91 ms	0.88 ms	1.02 ms	0.98 ms	1.00 ms	0.93 ms
IOPS (MB/s)		6.71	6.91	36.33	37.87	86.80	93.21
set	Ops/sec	11954.03	12352.70	10710.89	11178.02	10943.58	11719.28
	Hits/sec	--	--	--	--	--	--
	Miss/sec	--	--	--	--	--	--
	Latency	0.912 ms	0.882 ms	1.025 ms	0.976 ms	1.009 ms	0.930 ms
		< 1.3ms 99.99 分位 < 21ms 100 分位	< 1.3ms 99.99 分位 < 11ms 100 分位	< 1.5ms 99.99 分位 < 25ms 100 分位	< 1.4ms 99.99 分位 < 12ms 100 分位	< 1.6ms 99.99 分位 < 25ms 100 分位	< 1.4ms 99.99 分位 < 16ms 100 分位
KB/sec	2039.70	2107.72	21920.97	22876.98	55227.93	59142.58	
get	Ops/sec	119538.94	123525.62	107107.68	111778.94	109434.59	117191.56
	Hits/sec	6597.55	6805.29	5852.10	6085.28	5998.83	6436.72
	Miss/sec	112941.39	116720.33	101255.58	105693.66	103435.76	110754.84

		0.909 ms	0.882 ms	1.016 ms	0.974 ms	0.994 ms	0.927 ms
	Latency	< 1.3ms 99.99 分位 < 2.5ms 100 分位	< 1.3ms 99.99 分位 < 2.3ms 100 分位	< 1.4ms 99.99 分位 < 3.4ms 100 分位	< 1.3ms 99.99 分位 < 2.9ms 100 分位	< 1.4ms 99.96 分位 < 3ms 100 分位	< 1.3ms 99.98 分位 < 2.4ms 100 分位
	KB/sec	4813.15	4972.09	15283.07	15905.36	33650.97	36099.70
set/get -- 4 : 1							
data_size		128bytes		2kb		5kb	
		pmem-redis	redis-4.0	pmem-redis	redis-4.0	pmem-redis	redis-4.0
avg qps		107123	141377	86783	131784	74509	104850
avg latency		1.12 ms	0.85	1.38 ms	0.91 ms	1.61 ms	1.14 ms
IOPS (MB/s)		15.99	21.90	151.88	230.55	321.29	451.55
set	Ops/sec	85700.41	113370.60	69430.36	105428.30	59679.28	83879.41
	Hits/sec	--	--	--	--	--	--
	Miss/sec	--	--	--	--	--	--
	Latency	1.119 ms < 1.6ms 99.99 分位	0.847 ms < 1.3ms 99.99 分位	1.382 ms < 2.2ms 99.99 分位	0.909 ms < 1.4ms 99.99 分位	1.611 ms < 3.4ms 99.98 分位	1.143 ms < 1.8ms 99.99 分位

		< 20ms 100 分位	< 15ms 100 分位	< 24ms 100 分位	< 15ms 100 分位	< 22ms 100 分位	< 10ms 100 分位
	KB/sec	14622.95	19344.34	142096.58	215770.09	301177.85	423306.26
get	Ops/sec	21425.10	28342.65	17357.59	26357.08	14919.82	20969.85
	Hits/sec	7868.13	10368.50	6401.66	9679.15	5536.19	7662.54
	Miss/sec	13556.98	17974.15	10955.93	16677.92	9383.63	13307.32
	Latency	1.117 ms < 1.6ms 99.99 分位 < 3.2ms 100 分位	0.847 ms < 2.7ms 99.99 分位 < 3.5ms 100 分位	1.377 ms < 2.2ms 99.99 分位 < 4.1ms 100 分位	0.909 ms < 1.4ms 99.99 分位 < 2.7ms 100 分位	1.6 ms < 3.4ms 99.98 分位 < 4.8ms 100 分位	1.142 ms < 1.8ms 99.99 分位 < 3.6ms 100 分位
	KB/sec	1750.60	2310.53	13431.22	20311.42	28220.97	39070.68
set/get -- 10 : 1							
data_size		128bytes		2kb		5kb	
		pmem-redis	redis-4.0	pmem-redis	redis-4.0	pmem-redis	redis-4.0
avg qps		108472	120747	84464	118171	65336	104806
avg latency		1.11 ms	0.99 ms	1.42 ms	1.01 ms	1.84 ms	1.14 ms
IOPS (MB/s)		17.26	19.22	159.77	223.59	304.63	488.51
set	Ops/sec	98609.58	109770.09	76789.88	107432.01	59422.75	95279.44

	Hits/sec	--	--	--	--	--	--
	Miss/sec	--	--	--	--	--	--
	Latency	1.105 ms < 1.7ms 99.99 分位 < 27ms 100 分位	0.992 ms < 1.4ms 99.99 分位 < 13ms 100 分位	1.42 ms < 2.4ms 99.99 分位 < 21ms 100 分位	1.014 ms < 1.6ms 99.99 分位 < 10ms 100 分位	1.836 ms < 3.5ms 99.99 分位 < 25ms 100 分位	1.144 ms < 1.8 ms 99.99 分位 < 16ms 100 分位
	KB/sec	16825.66	18729.86	157158.66	219870.77	299883.17	480837.60
get	Ops/sec	9860.83	10976.86	7678.89	10743.06	5942.20	9527.82
	Hits/sec	3929.54	4408.28	3085.57	4352.00	2395.70	3811.27
	Miss/sec	5931.29	6568.58	4593.32	6391.06	3546.50	5716.54
	Latency	1.102 ms < 1.7ms 99.99 分位 < 3.6ms 100 分位	0.992 ms < 1.3ms 99.99 分位 < 3.3ms 100 分位	1.413 ms < 2.3ms 99.99 分位 < 4.6ms 100 分位	1.013 ms < 1.6ms 99.99 分位 < 3.3ms 100 分位	1.825 ms < 3.4ms 99.99 分位 < 5.3ms 100 分位	1.142 ms < 1.8ms 99.99 分位 < 3.3ms 100 分位
	KB/sec	846.67	946.99	6451.13	9096.05	12195.23	19403.66

总结：（测试结果存在一定的性能波动）

以 memtier-benchmark 的结果来看，Pmem-Redis 的性能受 value 大小的影响比较明显，随着 val 的不断增加，与原生基于内存的 redis 存在较大的性能差异。（性能差在 20%~35%之间）

同时，以不同读写比的测试场景来看，在读性能上，Pmem-Redis 的性能更稳定，写性能波动性较大，与 Intel 宣称的 AEP 的读性稍弱内存，写性能较差相符合。

就长尾问题来看，在写操作上，Pmem-Redis 受 key-val 的大小的影响比内存更为严重，对于是否能够满足线上实际运行要求，需结合业务场景进行进一步测试。

测试总结

结合 redis-benchmark 和 memtier-benchmark 的测试结果来看，intel 推出的基于 AEP 的 Pmem-Redis 的性能要弱于纯内存的 Redis-4.0 的性能。

- 性能上而言，整体性能按照读写比 1:1 的场景来看，pmem-redis 的性能随着 key-val 的大小的不断增加而急剧下降，性能差异在 17.8%~52.4%之间；
- 长尾问题来看，Pmem-redis 的长尾问题在写场景下比内存态 redis 波动性更大，更为明显。

建议：可以结合业务的实际使用场景来做进一步测试，以便考虑结合 AEP 来提升资源利用率的问题。