io-fire for AWS EC2

Fast Block Storage on AWS from WildFire Storage

io-fire is a new storage option for EC2 instances that leverages local SSDs and EBS block storage volumes. It delivers higher performance at lower cost than other storage options without leaving AWS or modifying your applications.

- **Performance of a local NVMe SSD**
- **Cross-boot durability**
- **Snapshots and cloning capable**
- **Much lower cost than io2**

Limitations:

EC2 – Linux Only: ***io-fire*** is a hybrid solution that combines existing AWS services with block software that runs on Linux. ***io-fire*** supports many, but not all Linux kernels offered by EC2 on both x86_64 and ARM.

Ephemeral or Durable:

io-fire can be deployed to deliver either ephemeral or durable block storage. For ephemeral storage, a local ephemeral SSD is used as the backing store. For durable storage, this local SSD is paired with one or more EBS gp3 volumes creating storage that has high IOPS, very low latency, and durability.

For gp3 Users:

More IOPS: > 20X the IOPS
Super low latency: > 10X lower latency

For Local SSD Users:

More IOPS: Burst 4K random writes to 500K IOPS or more depending on your SSD configuration.
Durability: When paired with an EBS gp3 volume or volumes, you get full volume durability across instance shutdowns.
Lower SSD Wear: Not that you really care, ***io-fire*** actually lowers SSD wear.

For io2 Users:

More IOPS: 2X or more IOPS.
Lower latency: Similar read latency and much lower write latency
Much lower cost: Save thousands per month by avoiding the io2 “provisioning premium”.

For Spot Instance Users:

If you use spot instances, the lack of durability of SSDs can be very problematic. ***io-fire*** lets you use local SSDs and spot instances without having to schedule copies of the ephemeral storage.

In that spot instances can save you 70% or more, this is a great solution.

How it Works:

io-fire is built around WildFire-Storage’s patented Enterprise Storage Stack (ESS) block translation layer. This software driver converts the random write workload from your application into a perfect linear workload that is ideally suited to optimize the bandwidth of storage devices. Usually, ESS is deployed in ultra performance local storage engines utilizing large arrays of SSDs. These arrays can reach over 20 Gbytes/second of write bandwidth and over 5 million 4K random writes. But this is AWS and we work with what we have (which is still quite a bit).

AWS offers local NVMe SSDs with above 2 GB/sec of write bandwidth from a single drive. This storage is even affordable. The downside is that this storage is totally "ephemeral". If you shut the system down, the data is wiped. If AWS does server maintenance, your data is "often" wiped. If you need fast storage that is durable, you are supposed to use Elastic Block Storage or EBS. The SSD variants are gp2, gp3, and io1/2. gp2 is general purpose, and generally slow (in SSD terms). gp3 lets you tune the storage somewhat, but maxes out at 16K IOPS. io1/2 starts to get interesting, supporting up to 260K IOPS for io2/block-express and decent latency, but the cost for a maxed-out io2 "IOPS only" volume is nearly $10,000 per month before you pay $0.125/GB.

io-fire combines the performance of one or more local NVMe SSDS with the persistent storage of EBS. By mirroring a local SSD to one or more gp3 volumes, you can create a durable store that has the performance of the local SSD without the ephemeral downside. Normally, such a mirror would be a disaster as gp3 will not handle a write intensive workload that is not linear. Because ESS is in place, both the local SSD and the gp3 mirror will see only a perfect workload of exactly 256K linear writes. Configuring an i4i or c6id instance with 1250 MB/sec or more of EBS storage using gp3 quite easy and not horribly expensive. At a cost less than 1:20th that of a fully configured io2 volume, you can max out the IOPS to over 250K read and write while maintaining cross-boot durability.

Performance Limits with EC2 SSDs and EBS

For ***io-fire*** in durable mode, we use a combination of local SSDs and EBS volumes. In that all reads come from the local SSD, read performance is completely determined by the local SSD performance, both in terms of transfer rates, IOPS, and latency.

For writes, ***io-fire*** linearizes writes into long linear block operations. These operations are mirrored to both the local SSD and to one or more EBS volumes attached to the instance. In that all writes are long and linear, only the write bandwidth, and not the IOPS, of the mirror volumes matter. This linear write pattern lends itself to gp3 volumes. gp3 volumes let you provision bandwidth at up to 1000 MB/sec at attractive costs. Even with the bandwidth provision costs, gp3 volumes are lower cost than gp2 and much lower than provisioned io2 volumes.

The limit for gp3 bandwidth is often not the volume, but the EBS connectivity to the instance. AWS documents this well for newer instances. Be careful with instances that are specified as "up to 10Gbit/sec". These instances will run at 10 Gbits for about an hour and then slow to half that speed. If your instance is spec'd at 10 Gbit/sec for EBS, then you should be able to sustain about 1100 MB/sec of linear writes which translates to 275K 4K random write IOPS.

Some Use Case Example Instances:

2 TB instance – 16 vCPUs / 128 GB RAM / Intel Xeon 8975C Ice Lake								
			io-fire		**gp3**		**io2**	
Storage Performance								
	4K random reads		350K	IOPS	16K	IOPS	64K	IOPS
	4K random reads latency		0.1	ms	1	ms	0.2	ms
	4K random writes burst		220K	IOPS	16K	IOPS	64K	IOPS
	4K random writes sustained		120K	IOPS	16K	IOPS	64K	IOPS
	4K random writes latency		0.01	ms	3	ms	0.4	ms
Monthly Cost								
	EC2 instance		i4i.4xlarge	$1,002	r6i.4xlarge	$736	r6i.4xlarge	$736
	gp3 volume space	GB	3578	286	2000	160		
	performance	IOPS			16K	64		
	performance	BW	625 x 2	40	1000	35		
	io2 volume space	GB					2000	250
	performance	IOPS					64K	3,520
	io-fire license	GB	2000	200				
	Total Monthly Cost			$1,528		$995		$4,506

This is 2TB of usable storage on an i4i.4xlarge instance. The total cost is about 50% more than the same instance with a 2TB gp3 volume, but it has about 20X the IOPS. Compared to the io2 option, ***io-fire*** is 4X the performance for 1/3rd the cost.

1 TB instance – 32 vCPUs / 64 GB RAM / Intel Xeon 8375C Ice Lake								
			io-fire		**gp3**		**io2**	
Storage Performance								
	4K random reads		350K	IOPS	16K	IOPS	256K	IOPS
	4K random reads latency		0.1	ms	1	ms	0.1	ms
	4K random writes burst		220K	IOPS	16K	IOPS	256K	IOPS
	4K random writes sustained		120K	IOPS	16K	IOPS	256K	IOPS
	4K random writes latency		0.01	ms	3	ms	0.2	ms
Monthly Cost								
	EC2 instance		c6id.8xlarge	$1,177	c6i.8xlarge	$993	c6i.8xlarge	$993
	gp3 volume space	GB	1812	145	1000	80		
	performance	IOPS			16K	64		
	performance	BW	625 x 2	40	1000	35		
	io2 volume space	GB					1000K	125
	performance	IOPS					64K	3,520
	io-fire license	GB	1000	100				
	Total Monthly Cost			$1,462		$1,172		$4,638

This example is 1TB of local storage on a c6id.8xlarge instance. At 256K IOPS, the ***io-fire*** implementation is somewhat faster than io2/block-express. The io2/block-express solution is just over 7X the cost.

Benchmark Examples:

i4i.4xlarge SSD only in ephemeral mode:

This is an instance with only the local SSD. ***io-fire*** in this mode does not mirror to EBS, but just manages the local drive. Because the local drive is written to 100% linearly, you will note that write performance is constant regardless of block size.

Read Tests:

Block Size	1 thread IOPS	1 thread BW	10 threads IOPS	10 threads BW	100 threads IOPS	100 threads BW
4K	8201	32.0M	76712	299.6M	366590	1431.9M
8K	7216	56.3M	66665	520.8M	312284	2439.7M
16K	6179	96.5M	55031	859.8M	202116	3158.0M
32K	5439	169.9M	44837	1401.1M	101463	3170.7M
64K	4905	306.5M	34465	2154.0M	50648	3165.5M
128K	4417	552.1M	23897	2987.1M	24320	3040.0M
256K	3520	880.0M	12721	3180.2M	11299	2824.7M

Write Tests:

Block Size	1 thread IOPS	1 thread BW	10 threads IOPS	10 threads BW	100 threads IOPS	100 threads BW
4K	341796	1335.1M	610503	2384.7M	604135	2359.9M
8K	275706	2153.9M	275234	2150.2M	275295	2150.7M
16K	137391	2146.7M	137291	2145.1M	137972	2155.8M
32K	68772	2149.1M	68693	2146.6M	68823	2150.7M
64K	34391	2149.4M	34429	2151.8M	34392	2149.5M
128K	17183	2147.8M	17207	2150.8M	17206	2150.7M
256K	8586	2146.5M	8604	2151.0M	8626	2156.5M

In general, ESS at these bandwidth's is nowhere near saturation. 8 SSDs striped should achieve over 3M IOPS and over 12 GB/sec of write bandwidth. WildFire-Storage has tested this size of array on local systems, but not on AWS as of this flyer's preparation.

c6id.8xlarge with gp3 EBS mirror active:

This is a fully resilient instance with ***io-fire*** mirroring gp3 volumes and a local SSD. The read performance is the SSD. The write performance is the 10 Gbit EBS bandwith that this instance has access to.

Read Tests:

Block Size	1 thread IOPS	1 thread BW	10 threads IOPS	10 threads BW	100 threads IOPS	100 threads BW
4K	10918	42.6M	103906	405.8M	388645	1518.1M
8K	10147	79.2M	92894	725.7M	337748	2638.6M
16K	9068	141.6M	79578	1243.4M	196846	3075.7M
32K	8158	254.9M	66053	2064.1M	97438	3044.9M
64K	7003	437.6M	48584	3036.5M	39980	2498.7M
128K	5989	748.6M	24621	3077.6M	20013	2501.6M
256K	4776	1194.0M	12387	3096.7M	10007	2501.7M

Write Tests:

Block Size	1 thread IOPS	1 thread BW	10 threads IOPS	10 threads BW	100 threads IOPS	100 threads BW
4K	373323	1458.2M	268245	1047.8M	267856	1046.3M
8K	128833	1006.5M	136475	1066.2M	133216	1040.7M
16K	67537	1055.2M	67911	1061.1M	66934	1045.8M
32K	34082	1065.0M	33790	1055.9M	33917	1059.9M
64K	16422	1026.3M	16837	1052.3M	17026	1064.1M
128K	8555	1069.3M	8376	1047.0M	8280	1035.0M
256K	4233	1058.2M	4219	1054.7M	4116	1029.0M

I4i.16xlarge with gp3 EBS mirror active:

This is a fully resilient instance with ***io-fire*** mirroring gp3 volumes and 4 striped local SSDs. The read performance is the SSD. The write performance is limited by the EBS bandwidth that this instance has access to. This level of performance, in terms of IOPS is well above io2 performance levels for both reads and writes.

Read Tests:

Block Size	1 thread IOPS	1 thread BW	10 threads IOPS	10 threads BW	100 threads IOPS	100 threads BW
4K	5094	19.8M	29068	113.5M	295664	1154.9M
8K	4431	34.6M	29385	229.5M	280217	2189.1M
16K	3809	59.5M	27348	427.3M	264541	4133.4M
32K	3004	93.8M	25254	789.1M	232908	7278.3M
64K	2807	175.4M	23861	1491.3M	171803	10.4G
128K	2494	311.7M	21246	2655.7M	90309	11.0G
256K	2324	581.0M	18522	4630.5M	44935	10.9G

Write Tests:

Block Size	1 thread IOPS	1 thread BW	10 threads IOPS	10 threads BW	100 threads IOPS	100 threads BW
4K	368263	1438.5M	572248	2235.3M	397152	1551.3M
8K	220738	1724.5M	201955	1577.7M	198149	1548.0M
16K	100111	1564.2M	100470	1569.8M	98000	1531.2M
32K	49142	1535.6M	49621	1550.6M	49706	1553.3M
64K	24427	1526.6M	24907	1556.6M	24640	1540.0M
128K	12310	1538.7M	12289	1536.1M	12309	1538.6M
256K	6225	1556.2M	6079	1519.7M	6220	1555.0M

Availability

io-fire is based on ESS 5.3.0 which is a stable build for 5.4 and later Linux kernels. The installer for ***io-fire*** is still a work in progress, so early users are invited to contact us. Once installed, ***io-fire*** is pretty much "zero touch" as it just presents a block device to your application.

Eventually, the ***io-fire*** installer will auto sense your instance, load the required kernel support files, and auto configure the local SSD and mirror raid devices.

Demos

io-fire is billed by the GB by the hour, so a demo is not really a thing. You get a $5 credit on your account, so your initial use is free anyway. You can start or stop instances at any time and are only billed when ***io-fire*** is active. You can "start" ***io-fire*** in read only mode against the EBS volume without a local SSD to access snapshots for backups or other purposes. Mounts in read-only mode are not billed. Read-only mounts can be done on low performance instances provided the instance has sufficient memory to map the ***io-fire*** volume's LBA capacity.

Legal

Io-fire is a proprietary software product of EasyCo LLC dba WildFire Storage. It is protected by US patents.

AWS, Amazon Web Services, EBS, Elastic Block Storage, EC2, Elastic Compute Cloud are all tradenames and trademarks of Amazon Web Services. The AWS logo is used per the AWS logo usage guidelines and does not imply endorsement of any products or services by AWS.

WildFire Storage

https://wildfire-storage.com +1 (610) 237-2000
sales@wildfire-storage.com +1 (888) 473-7866