Units 9-10/12 Mars Rd
Lane Cove NSW 2066
Australia

Tel +61 2 9420 0690
Fax +61 2 9420 0692
www.cap-xx.com

CAP-XX (Australia) Pty Ltd
ABN 28 077 060 872
ACN 077 060 872

White Paper

CAP-XX Supercapacitors provide power backup for Solid State Drives

By Pierre Mars, V.P. Applications Engineering

1. SSD Introduction

Solid State Drives (SSDs) are electrically, mechanically and software compatible with a conventional Hard Disk Drive. The difference is instead of a using rotating magnetic media to retain data, an SSD uses semi conductor memory such as battery-backed RAM or other electrically erasable ICs such as Flash memory. NAND Flash has become the dominant storage media for SSDs and this paper will consider performance of NAND flash SSDs and how this can be enhanced using cache memory, in particular Synchronous Dynamic RAM (SDRAM) with backup power supplied by a supercapacitor.

Key to SSD performance is the SSD Controller which controls how read and write operations are performed to maximise both data transfer speed and disk life. Fig 1 shows a typical SSD controller.

2. SSD Performance

Many SDDs currently available do not use cache memory. We will first discuss performance without cache and then examine the improvements with cache and the power requirements for cache memory.

2.1 Endurance

Hard Disk Drives have no effective limit on the number of writes that can be performed. NAND Flash, on the other hand, can only be written to a finite number of times known as Endurance. There are two categories of NAND Flash:

- Single Level Cells, (SLC) where each cell has one of 2 states representing 1 bit. This type has typical endurance of 100,000 writes
- Multi Level Cells, (MLC), where each cell has one of 4 voltage levels representing 2 bits/cell. The penalty for this greater density is a lower endurance, typically 10,000 writes.

The endurance limit means the SSD controller must spread write operations evenly throughout the NAND Flash blocks, known as wear leveling, to prevent part of the memory prematurely failing. Wear leveling means files will not be written in sequential blocks but spread through the memory. This adds complexity to the File Allocation Table (FAT) and SSD management system, which must:

- Keep track of which physical blocks a file has been written to
- Keep track of which blocks to use next to optimize wear leveling

Therefore there are several writes to the FAT for each file write. As will be seen in the next section, these many small size writes can greatly impact SSD write performance.

2.2 Read & Write Speed

Conventional Hard Disk Drives are limited in their input/output (I/O) performance by the need to move the read/write heads. The time required to move the head to the correct position is given in HDD specification as Seek Times which is in the order of msecs. NAND flash, on the other hand, has the following limitations:

- The IC is organized in blocks and pages. As an example, a 2Gb device, Micron MT29F2G08[1], configured as 256Mbytes, is split into 2048 blocks of 64 pages, with each page 2048 Bytes, so each block is 128K Bytes.
- A location must be erased before it is written to.
- Erase operations are performed on entire blocks
- Read and Write operations are by page
- Block Erase takes ~2ms (typical)
- Time to program 1 page takes 300µsecs (typical)
- Time to read a random page is 25µsecs (max)
- Time to read sequential pages is 30nsecs (min)

Hence maximum write speed to an individual NAND Flash IC = time to write an entire block = (64 pages x 2048 bytes)/(0.002s block erase time + 64 x 0.0003s page write time) = 6.18MB/sec. This is very slow compared to HDD write speed which is in the order of 60MB/sec – 100MB/sec.

Fortunately, SSDs are not limited to individual NAND Flash IC performance. Fig 1 shows a typical block diagram of an SSD. The NAND flash ICs are partitioned in several banks and the controller can interleave operations between banks, so for example, one group can be read while another is being erased, or multiple blocks can be erased simultaneously. This greatly increases throughput.

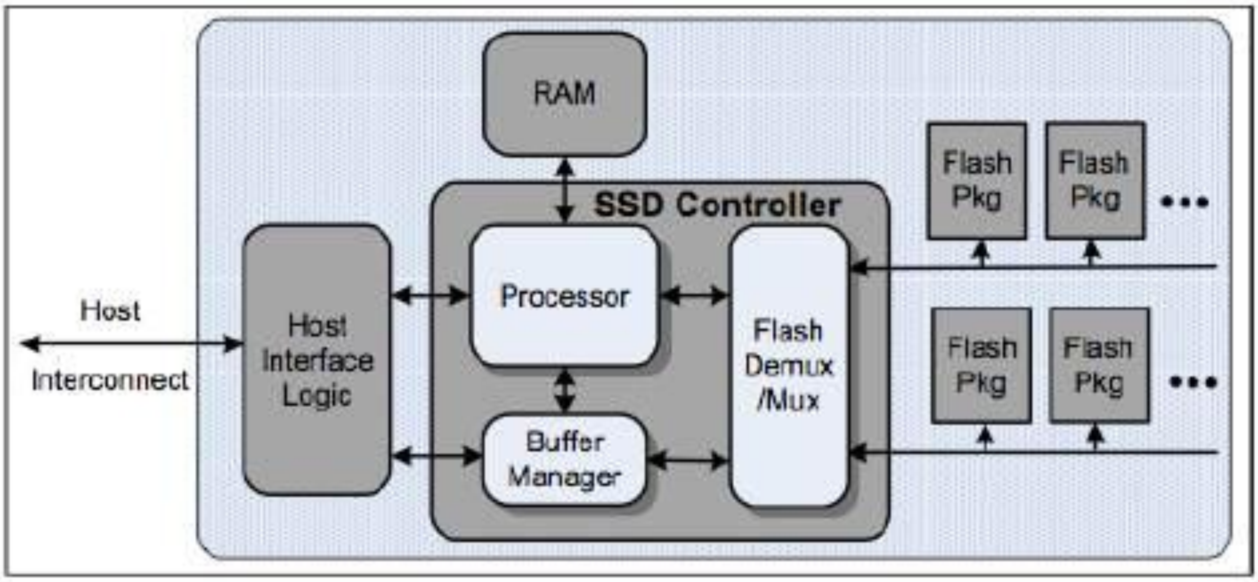


Fig 1: Typical SSD Block Diagram

---

[1] http://download.micron.com/pdf/datasheets/flash/nand/2_4_8gb_nand_m49a.pdf

Table 1 below lists typical drive performance figures for some HDDs and SSDs[2].

The SSDs all perform at speeds far in excess of the write speeds calculated for an individual NAND flash IC. However, HDD specs appear superior. This is misleading, since SSDs have much better access times. Read speed will depend on the transfer size – for large files of sequential blocks the HDD will perform well, while for small files the SDD performance will be far superior as shown in Fig 2[2].

Table 1: Sample performance comparisons between HDDs & SSDs

	Drive Model	Description	Seek Time			Latency	Read XFR Rate		Write XFR Rate	
			Track to Track	Average	Full Stroke		Outer Tracks	Inner Tracks	Outer Tracks	Inner Tracks
Hard Drives	Western Digital WD7500AYYS	7200 RPM 3.5" SATA	0.6 ms	8.9 ms	12.0 ms	4.2 ms	85 MB/sec	60 MB/sec*	85 MB/sec	60 MB/sec*
	Seagate ST936751SS	15K RPM 2.5" SAS	0.2 ms	2.9 ms	5.0 ms*	2.0 ms	112 MB/sec	79 MB/sec	112 MB/sec	79 MB/sec
Flash SSDs	Transcend TS8GCF266	8GB 266x CF Card	0.09ms				40 MB/sec		32 MB/sec	
	Samsung MCAQE32G5APP	32G 2.5" PATA	0.14ms				51 MB/sec		28 MB/sec	
	Sandisk SATA5000	32G 2.5" SATA	0.125ms				68 MB/sec		40 MB/sec	

* Figure is an estimate

Fig 2 shows that the SSDs will outperform HDDs unless the data transfer size is very large.

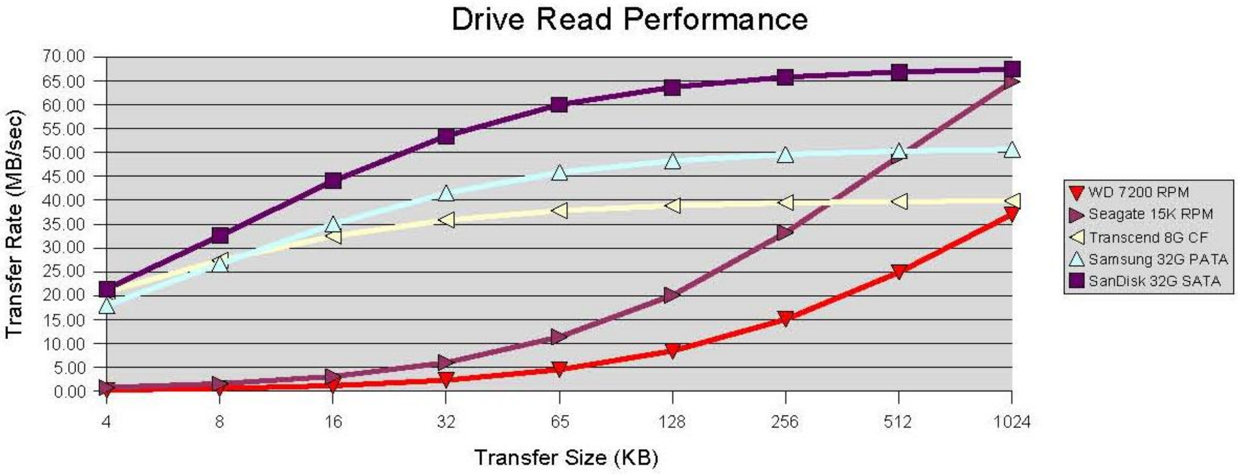


Fig 2: Read Performance as a function of data transfer size

Small data reads and writes are much more prevalent than most people realize. Firstly, many people work with relatively small Word files or similar. The FAT data will likely require 3 – 5 small reads to locate the file.

We will now look at how small writes affect SSD performance. Writing the file will similarly involve several small writes to the FAT. As foreshadowed with the description of NAND flash properties at

---

[2] "Understanding Flash SSD Performance" http://www.storagesearch.com/easyco-flashperformance-art.pdf

the introduction to this section, it is with writes to random blocks that SSD performance degrades. Blocks must be erased before data is written in them. If some data is to be overwritten in a block, then the remaining data must be temporarily stored elsewhere while the block is erased, then restored. The random write performance of the three SSD examples in Table 1 is shown in Table 2[3]:

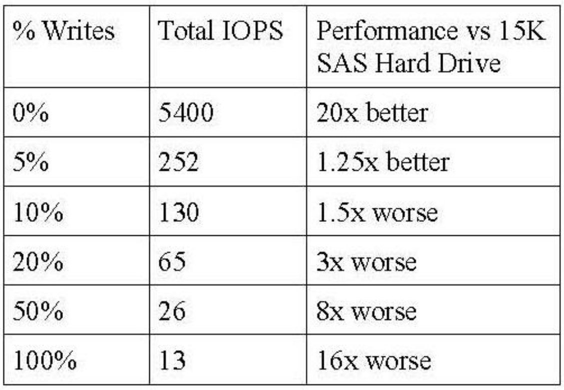
Table 2: SSD Random Write Performance

Drive	Random Writes/sec
Transcend 8G CF	47
Samsung 32G PATA	24
SanDisk 32G SATA	13

This performance is much worse than the quoted write speed. As the write block size approaches the Erase block size in the NAND Flash ICs, and blocks do not have data that needs temporary stored prior to erasure, then write speed will approach quoted speed.

Finally, consider the impact of writes on overall I/O performance. Fig 2 shows that for small random reads, a Flash SSD will be typically 20 x faster than a hard disk. What happens as the proportion of write operations increases? This is shown in Table 3 below which compares the overall I/O performance of the SanDisk SATA 500 SSD and Seagate SAS HDD. The specs for both these devices were included in Table 1.

Table 3: Effect of random writes on overall I/O performance



% Writes	Total IOPS	Performance vs 15K SAS Hard Drive
0%	5400	20x better
5%	252	1.25x better
10%	130	1.5x worse
20%	65	3x worse
50%	26	8x worse
100%	13	16x worse

Table 3 shows that even a small proportion of random writes severely impacts the overall I/O performance of a Flash SSD. This limits the effectiveness of SSDs in large server applications such as on line databases, mail queues, order entry and others which involve many small updates. This leads to the use of cache memory to overcome this limitation.

---

[3] "Understanding Flash SSD Performance" http://www.storagesearch.com/easyco-flashperformance-art.pdf

3. Use of Cache Memory to improve performance

Fig 3[4] shows an SSD Controller that uses cache memory, in this case as SDRAM buffer.

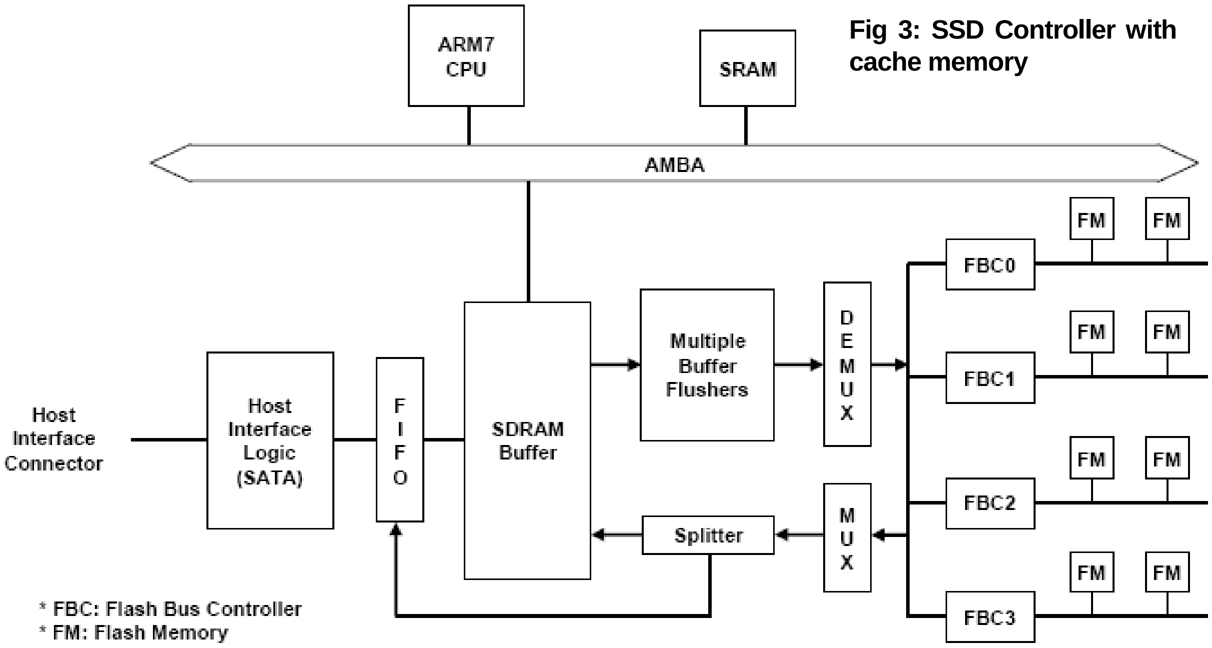


Fig 3: SSD Controller with cache memory

* FBC: Flash Bus Controller
* FM: Flash Memory

The cache holds the FAT and allows small random writes to the FAT and of data without NAND flash write characteristics affecting performance. As an example of performance improvement when using a cache compare the write speeds of two similar NAND Flash SSDs, one with cache, the other without as shown in Table 4:

Table 4: Performance comparison with/without cache.

	Quoted Read Speed	Benchmarked[5] Read Speed	Quoted Write Speed	Benchmarked Write Speed
MemoRight MR25.2-064S SSD 64GB SATA/150	120MB/s	115.6MB/s	120MB/s	124.0MB/s
Samsung SSD SATA 3.0Gbps 2.5" 64GB SATA/300	100MB/s	90.6MB/s	80MB/s	83.7MB/s

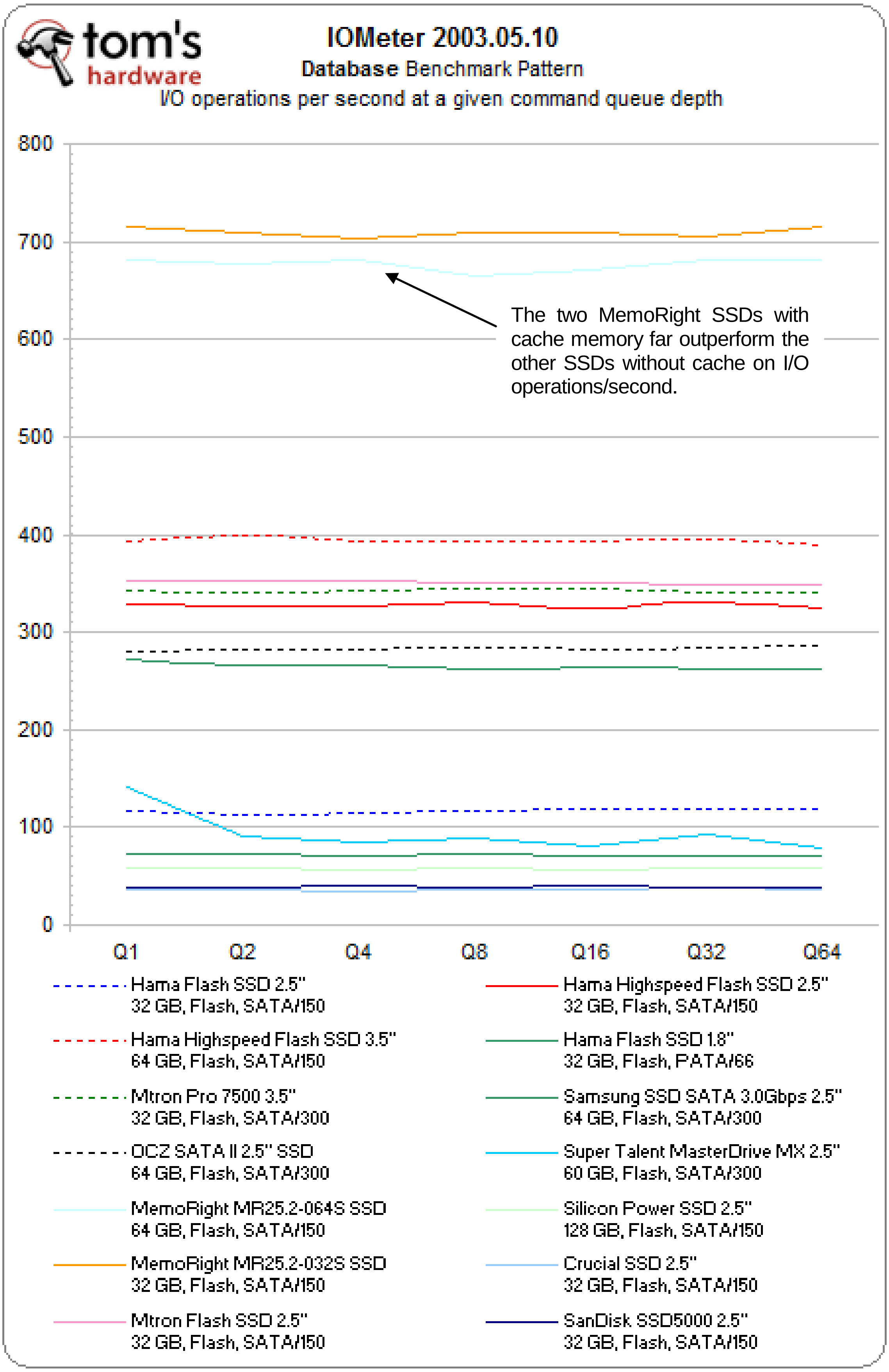
The MemoRight unit has 64MB of cache. Even though its serial interface is only 150MB/s compared to Samsung's 300MB/s, it outperforms the Samsung unit. Comparing I/O operations per second for a combination of small reads and writes shows a greater performance gap in Figs 4 and 5[5] for Database and File Server Benchmarks.

The conclusion from this section is that cache memory greatly enhances performance.

---

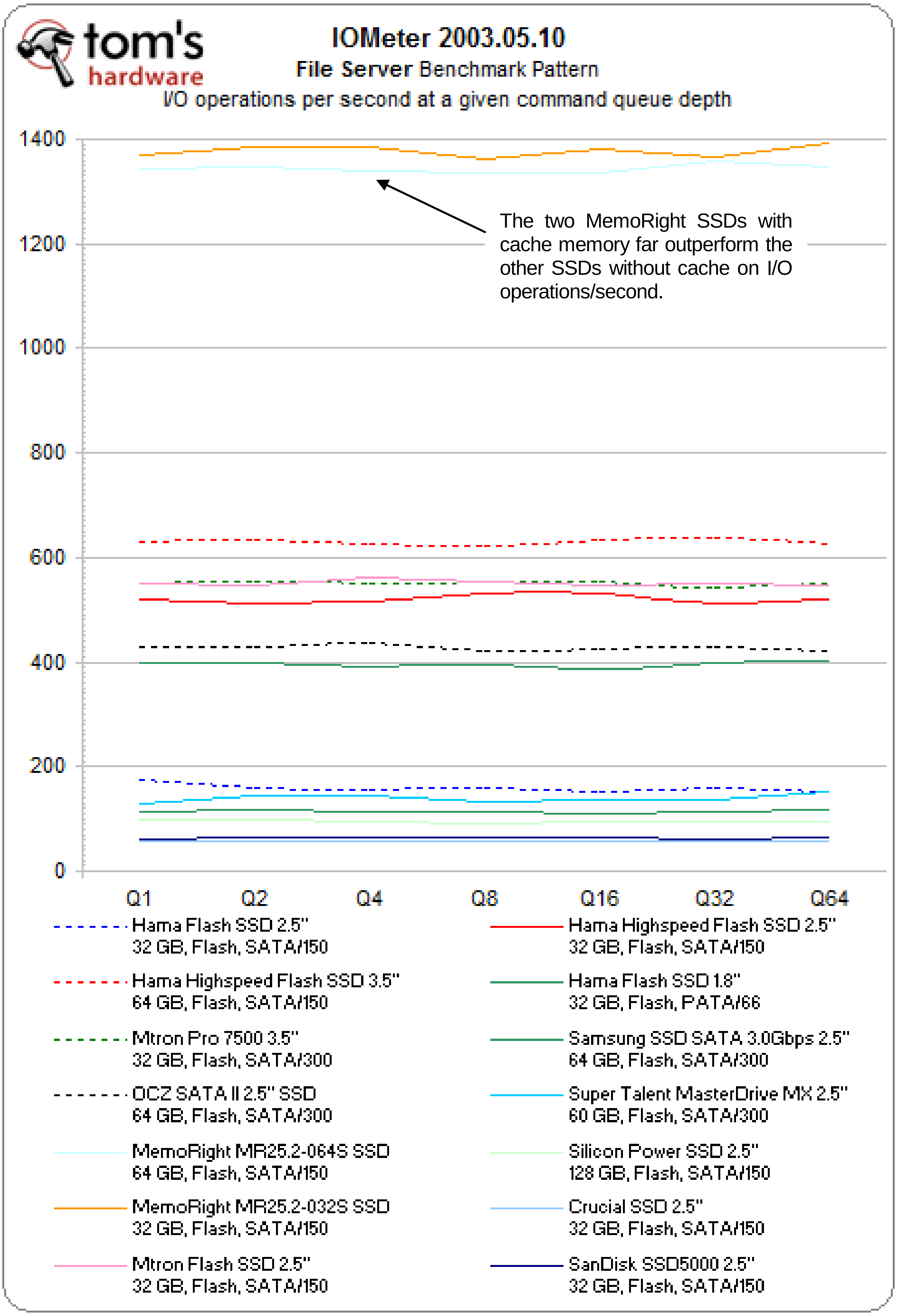
[4] http://download.micron.com/pdf/datasheets/flash/nand/2_4_8gb_nand_m49a.pdf

[5] http://www.tomshardware.com/reviews/flash-ssd-hard-drive,2000.html



The two MemoRight SSDs with cache memory far outperform the other SSDs without cache on I/O operations/second.

Fig 4: I/O Database Benchmark for a selection of Flash SSDs.



The two MemoRight SSDs with cache memory far outperform the other SSDs without cache on I/O operations/second.

Fig 5: I/O File Server Benchmark for a selection of Flash SSDs

4. Possible Solutions for Cache Memory

This section considers the different memory types that may be used for cache memory. The key attributes are:

- Random write and read speed
- Volatility – are contents preserved in case of power loss?
- Cost
- Density

Table 5 catalogues the alternatives for cache memory:

Cache Type	Read/Write Speed	Volatility	Density	Cost
None. Data + FAT written directly to Flash. Software (careful ordering of writing FAT & Data) ensures no data lost on power fail.	Limited by NAND Flash performance	Non Volatile, no data lost on power fail	No extra space required	$0
SDRAM to cache data only. No power backup. FAT written direct to NAND Flash	Data I/O at 133MB/s, but FAT writes limited by NAND Flash performance.	Data in last write lost, SSD reverts to last checkpoint, similar to HDD	Excellent. 64MB in 54 pin TSOP (22mm x 10mm)	Low. 64MB IC @ US$13.80[5]
SDRAM to cache FAT + Data. Needs power backup to flush to NAND	Full improvement as per previous section. I/O at 133MB/s	Non Volatile, no data lost.	Excellent. 64MB in 54 pin TSOP (22mm x 10mm)	Low. 64MB IC @ US$13.80
Battery Backed[6] RAM. Li-Ion battery integrated into the IC package.	~12MB/sec	Non Volatile, no data lost.	Poor, 2MB in 36 pin DIP	Expensive. US$121[7]
nVSRAM. Each SRAM cell is shadowed by an non volatile cell. There is a // write function to store all the contents of the SRAM to the non volatile memory. This happens	~66MB/sec	Non Volatile, no data lost.	Poor, 512KB in 48 ball FBGA, 10mm x 6mm, but each IC requires its own 82µF capacitor.	Expensive. US$46.33 per IC (single quantity)[7]

---

[5] Digikey 1000 off price

[6] ST Microelectronics M48Z2MiY

[7] http://app.arrownac.com/

automatically on power fail & supported by an 82µF cap per IC.				
FRAM[8], Ferroelectric RAM is a non volatile memory that uses a ferroelectric material to store an electric polarization as a non volatile data store. May eventually replace Flash.	~90MB/s	Non Volatile, no data lost.	Poor, 256KB in 48 pin TSOP, 14mm x 12mm.	Expensive, sampling at 2000Yen (~US$19)
MRAM, Magneto resistive Random Access Memory stores data by setting a magnetic field to one of two polarities for each cell.	~28MB/s (35ns read & write cycle times)	Non Volatile, no data lost.	Poor, 512KB in 44 pin TSOP, 10mm x 18mm	Expensive, samples @ US$25 ea
Phase Change RAM. Data stored by changing the state of call between an unstructured amorphous state or a highly structured crystalline state. The cell's phase is read by measuring electrical resistance through the cell.	Not yet in Production, prototypes announced.			

In the enterprise market, data integrity is paramount and high data transfer rates are extremely important. This leaves the first and third options of Table 5 as the only two practical solutions. SSD vendors are now moving towards the third option to be competitive on data transfer speed.

In the consumer market, cost is very important, so the first or second options tend to be used. Data transfer speed is not as noticeable to an individual user as it is for servers, so the lowest cost, no cache option is attractive until memory prices reduce. For the second option, losing the data in the cache and rolling back to the last checkpoint is no different to losing power during an HDD write.

---

[8] Fujitsu, MB85R2001. Price info from Fujitsu PR 18 Apr 2007.

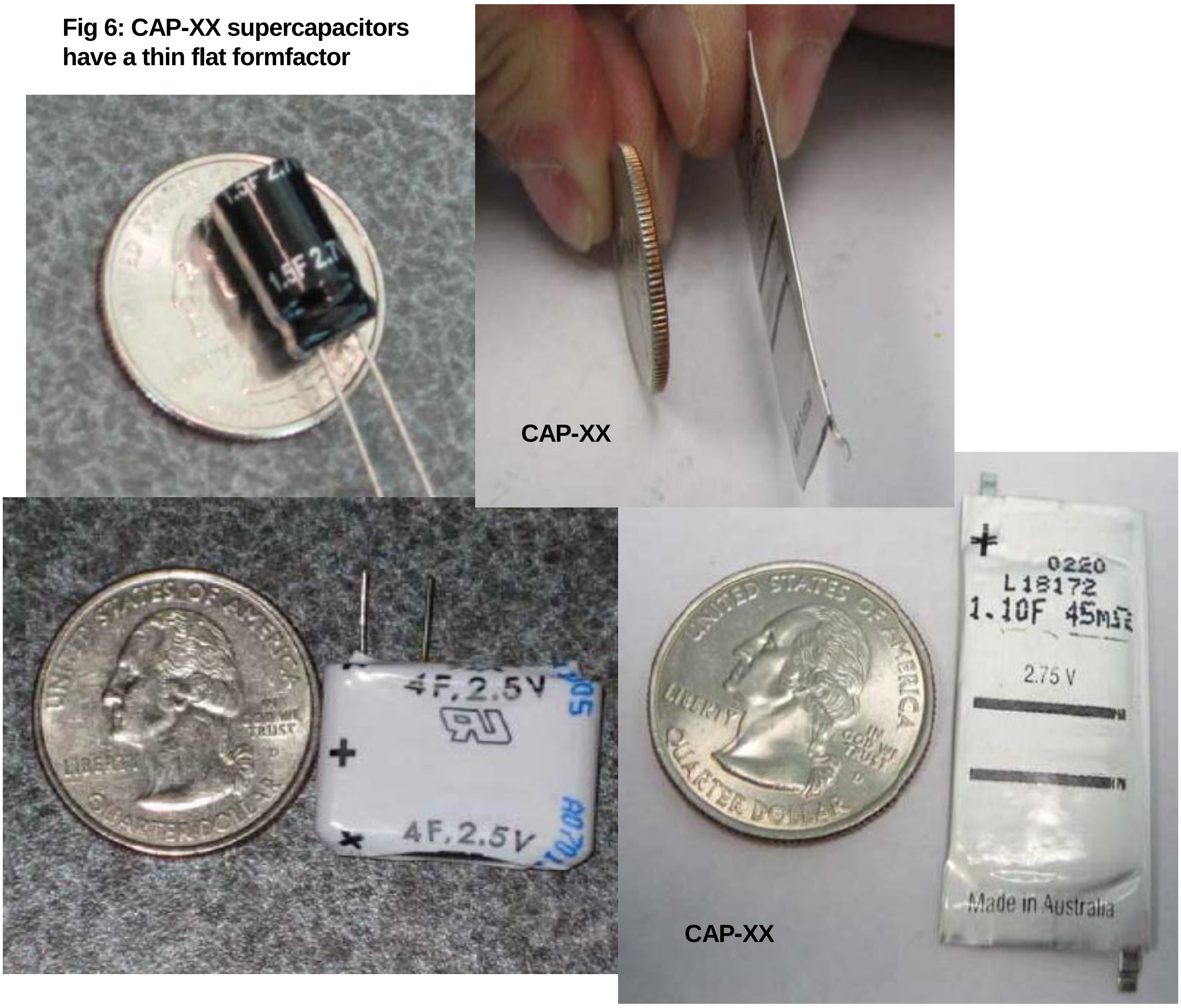
When cache is used, then the preferred solution, option 3 of the alternatives above, is to use SDRAM to cache data + FAT, as per the MemoRight SSD that had the fastest performance of the SSDs looked at in the previous section. This requires the whole system, Cache, Controller, NAND Flash to be backed up in case of power fail, so that the controller can write all the contents of the cache to the Flash. There are two alternatives for providing this backup power, shown in Table 6.

Table 6: Backup Power Options

	Pros	Cons
Secondary Battery		– Limited life, ~500 cycles – Size, in order to provide required current.
Supercapacitor	– Long Life – Wider temperature range (-30°C to +85°C)	
CAP-XX Supercapacitor	– Also small, thin	

The photos in Fig 6 compare the form factor of CAP-XX and some other supercapacitors. CAP-XX is uniquely suited for SSDs with thin form factors, for example in blade servers, or in thin, light high end notebooks.

Fig 6: CAP-XX supercapacitors have a thin flat formfactor



The next section details alternative methods of providing backup power using a CAP-XX supercapacitor, depending on the source voltage available.

5. SSD Backup Power Supply Design with CAP-XX Supercapacitors

Typically, in an SSD, the NAND Flash is powered at 1.8V, though it may also be powered at 3.3V, SDRAM is powered at 3.3V, and the microcontroller at ~1.2V, though this may also be at 3.3V or 1.8V. In the following cases, it will be assumed that the NAND is at 1.8V and the microcontroller is at 1.2V. If they are at one of the other possible rails, then the following cases simplify accordingly. The load currents and duration of backup depend on the size of the SSD.

The available power supply to an SSD is typically:

- Enterprise (rack mounted), 12V or 5V
- 2.5”, 5V
- 1.8” 3.3V

Each of these three cases will be considered separately.

5.1 Enterprise, 12V supply

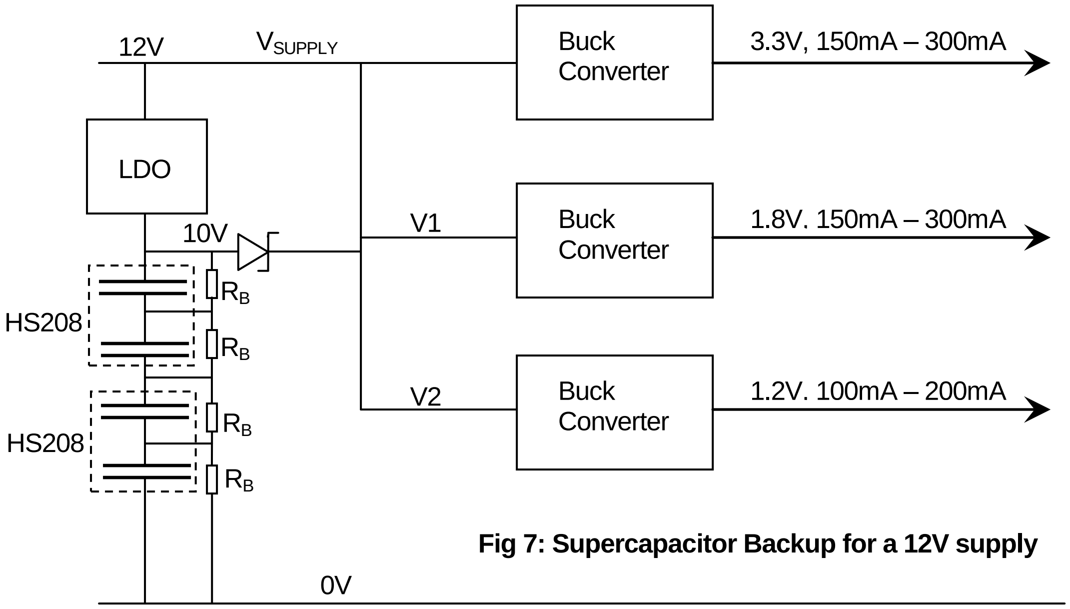


Fig 7: Supercapacitor Backup for a 12V supply

Referring to Fig 7, when power is available, V_{SUPPLY} = 12V and the 2 x HS208 in series are charged to 10V. The LDO prevents the supercapacitor from going over voltage and also protects the 12V source from inrush current to the supercapacitor, therefore an LDO with current limiting should be selected. A thermal current limit will do, since the current does not need to be controlled accurately, it just needs to protect the 12V source. The current limit value will also determine supercapacitor charge time = C x V_{CAP} / I_{LIM}. If the current limit is 1A, the supercapacitor = 0.45F (= 2 x HS208 in series) and V_{CAP} = 10V, then charge time = 0.45F*10V/1A = 4.5secs. The balancing resistors R_B balance the supercapacitor cells so each cell has 2.5V across it. A suitable value for R_B = 5.6KΩ. Assume each buck converter is 85% efficient.

Then Max Load Power as seen by the source

$= 3.3V \times 300mA/85\% + 1.8V \times 300mA/85\% + 1.2V \times 200mA/85\%$

$= 2.08W$

The power dissipated through the balancing resistors = $10^2V/(4 \times 5.6K\Omega) = 0.0044W << 2.08W$, so this can be ignored.

The minimum voltage at the input of the 3.3V buck converter in order to maintain a 3.3V output is typically 4V. A suitable example is the LT1506. It typically takes 2 – 3 secs to flush the cache and gracefully shutdown the SSD.

Therefore, we need to determine what size supercapacitor is required to deliver 2.1W for 3 secs, starting at 10V with a final voltage of 4V. The constant power sheet in the Voltage Decay simulator on the CAP-XX website solves this problem, http://www.cap-xx.com/resources/designaids/design_calc.htm. Fig 8 below shows the result for 2 x HS208 in series.

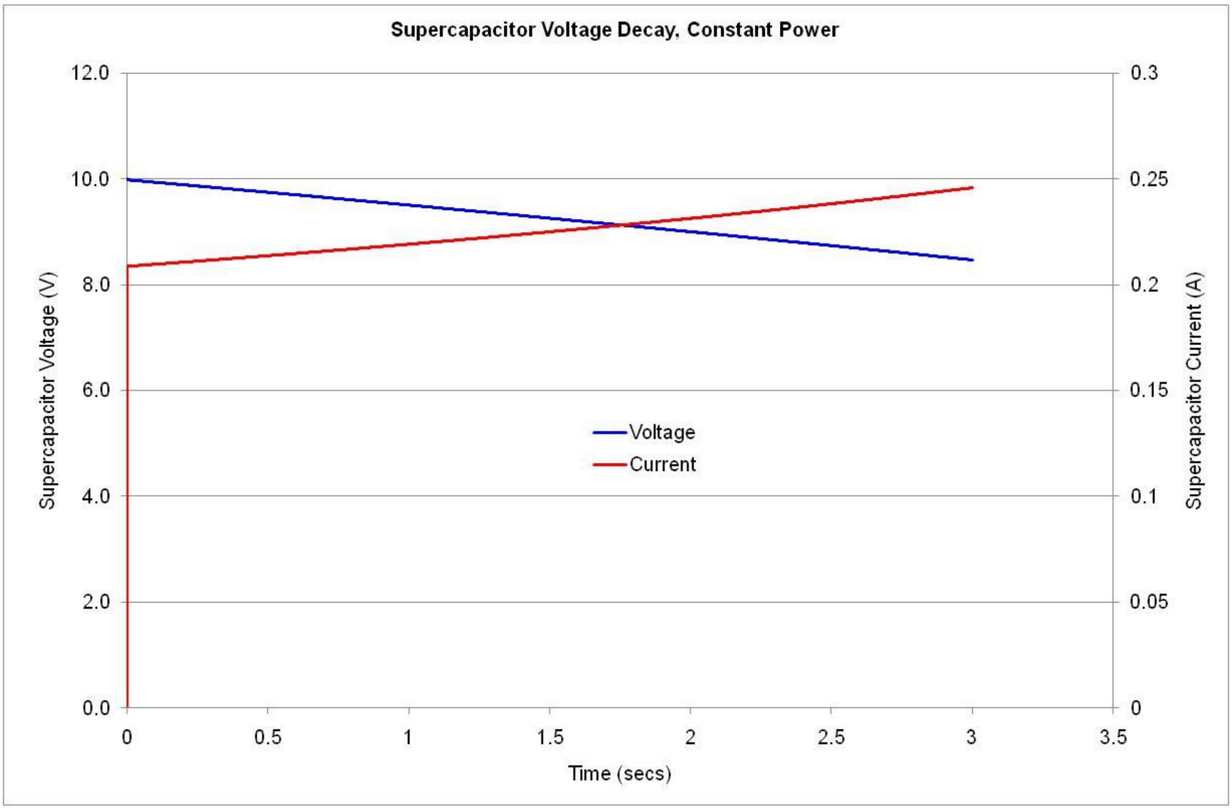
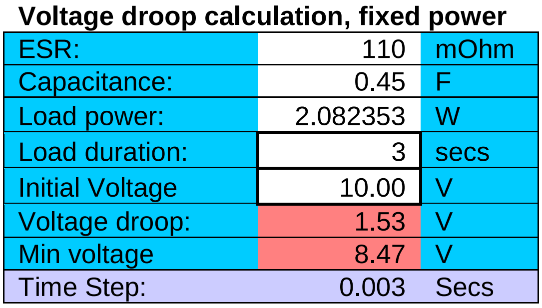


Fig 8: Voltage decay of 2 x HS208 supercapacitors. Final voltage is 8.47V >> min voltage of 4V which allows headroom for ageing.



Voltage droop calculation, fixed power

ESR:	110	mOhm
Capacitance:	0.45	F
Load power:	2.082353	W
Load duration:	3	secs
Initial Voltage	10.00	V
Voltage droop:	1.53	V
Min voltage	8.47	V
Time Step:	0.003	Secs

Table 7 shows the parameters entered in the CAP-XX Constant Power Voltage Decay simulator to produce Fig 7.

An alternative, simpler, but less efficient solution to Fig 7 is to feed the 3.3V output to V1 & V2 and use LDOs to derive the 1.8V and 1.2V rails as shown in Fig 9.

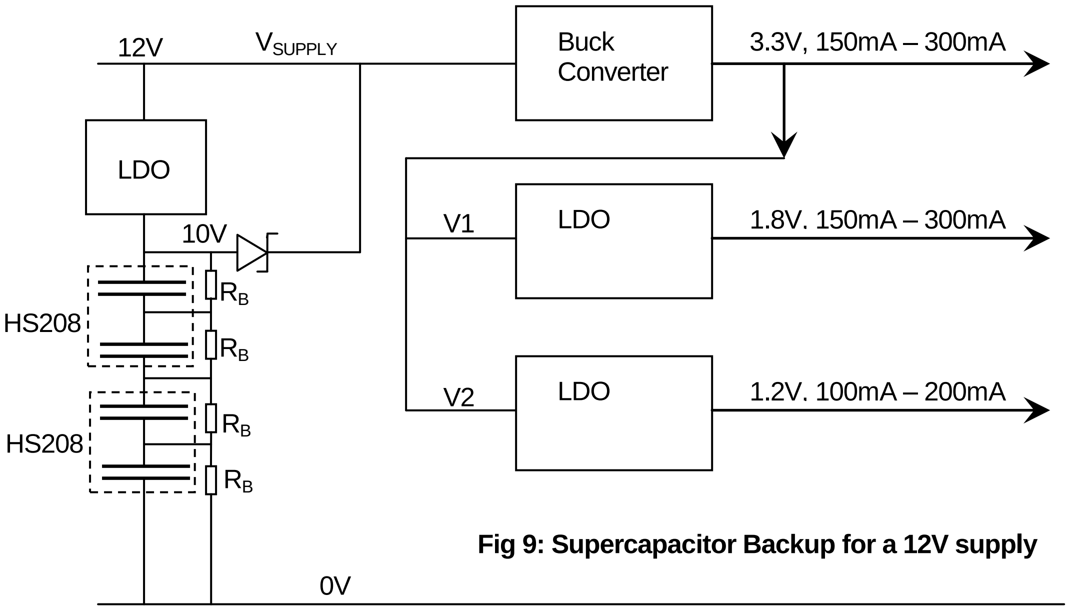


Fig 9: Supercapacitor Backup for a 12V supply

In this case the max load power

=3.3V x (300mA + 300mA + 200mA)/85% = 3.1W.

Note that this simplification is 50% less efficient! The voltage decay simulator in this case shows 2 x HS208 supercapacitors would only decay to 7.6V, so the reduced efficiency does not impact the effectiveness of the supercapacitor solution for power back up.

5.2 2.5” SSD, 5V supply

Fig 10 below shows a power block diagram for the 5V supply case.

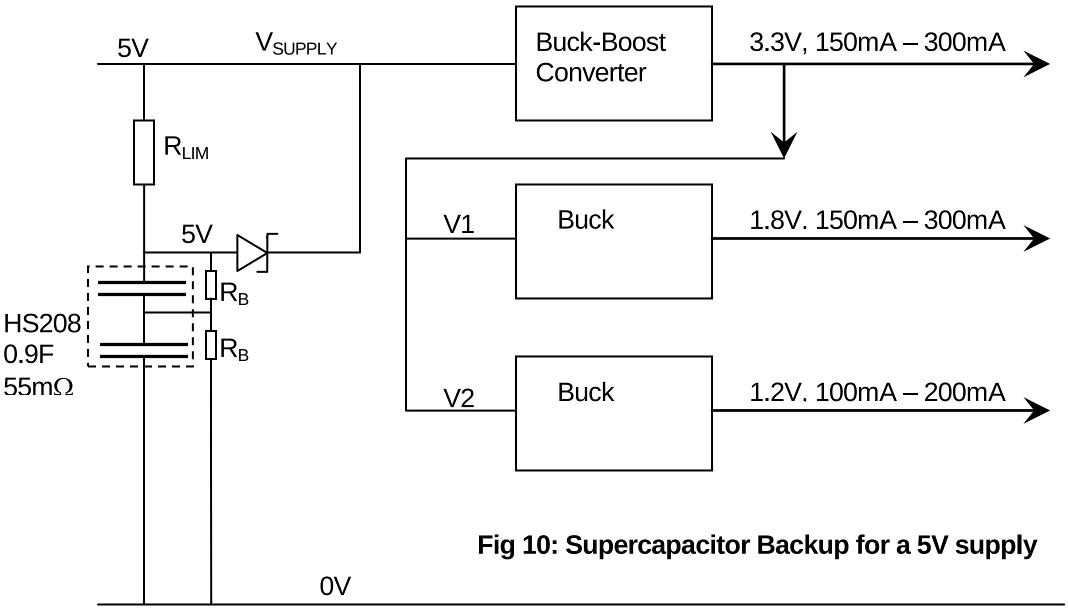


Fig 10: Supercapacitor Backup for a 5V supply

RLIM protects the 5V supply from the supercapacitor inrush current. Over voltage protection is not required since the HS208 can operate up to 5.5V. If RLIM = 5Ω, then the maximum initial inrush current = 1A, and it will take 4.5secs to charge the supercapacitor to 4.75V.

A buck-boost is used to generate the 3.3V rail so that the supercapacitor can more completely discharge which will support the SSD for longer. A suitable buck-boost is the TPS63001 which can operate with an input voltage down to 1.8V. Assuming 85% efficiency for the switching regulators, the maximum load power seen by the source

= 3.3V x 300mA/85% + 1.8V x 300mA /(85% x 85%) + 1.2V x 200mA /(85% x 85%)

= 2.25W.

Again, using the Constant Power sheet of the CAP-XX Voltage Decay Simulator to solve how long a given size supercapacitor supports the SSD load gives the results in Fig 11. Table 8 shows the simulator input parameters for this case.

ESR:	55	mOhm
Capacitance:	0.9	F
Load power:	2.25	W
Load duration:	3	secs
Initial Voltage	5.00	V
Voltage droop:	1.89	V
Min voltage	3.11	V
Time Step:	0.003	Secs

Table 8: Input Parameters for the Constant Power simulation for the 5V supply case.

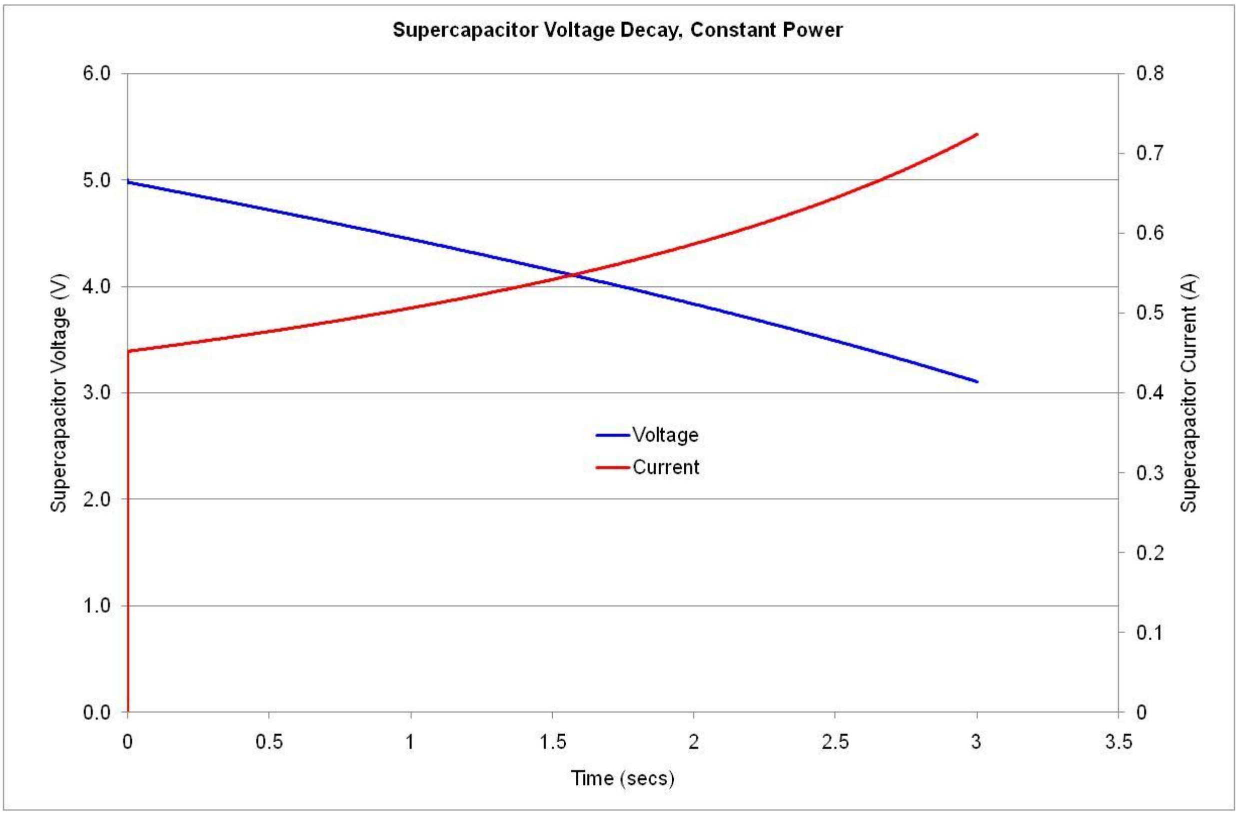


Fig 11: Voltage decay of 1 x HS208 supercapacitor with a 5V supply. Final voltage is 3V >> min voltage of 1.8V which allows headroom for ageing.

5.3 1.8” SSD, 3.3V Supply

Fig 12 shows a power supply block diagram for a 3.3V supply.

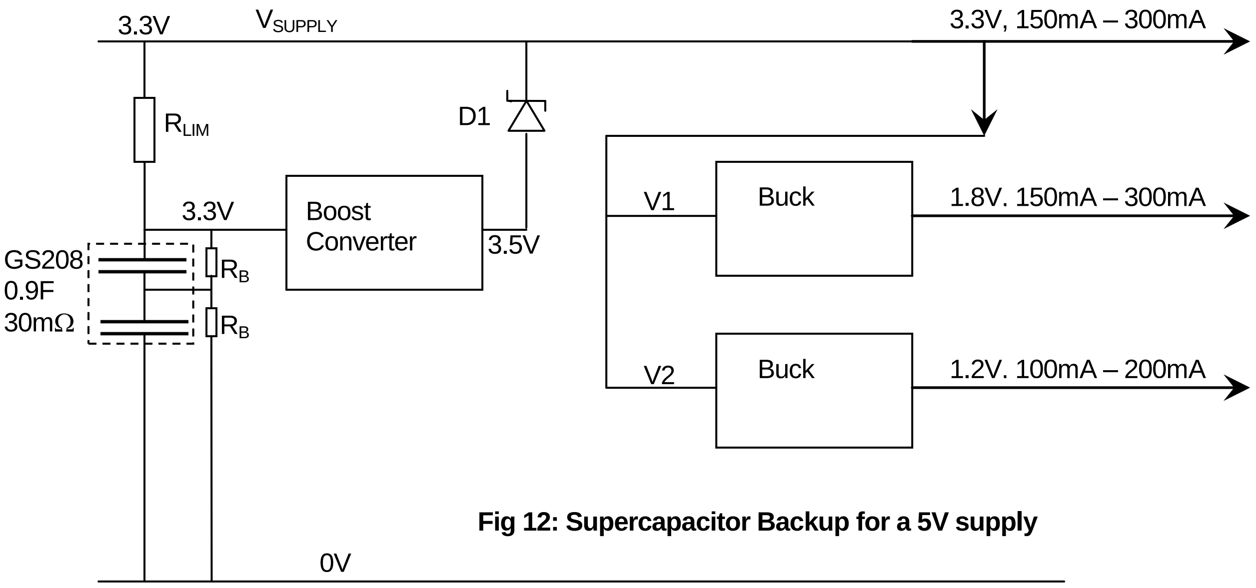


Fig 12: Supercapacitor Backup for a 5V supply

Referring to Fig 12, the supercapacitor is charged through R_{LIM} in exactly the same way as in the 5V case. RLIM would be set to 3.9Ω to limit initial inrush current to 0.85A. The supercapacitor would reach 3.13V after 3.5 secs. To maximize the backup time from the supercapacitor, the boost converter should be chosen with the lowest possible minimum input voltage. A suitable IC is the MAX1763 which has a minimum operating input voltage of 0.7V.

When V_{SUPPLY} is present at 3.3V, then D1 will not conduct. If V_{SUPPLY} is lost then D1 conducts and will drop ~0.3V forward voltage, so the 3.3V rail will be supplied at 3.2V. This assumes that VSUPPLY will not be a short circuit when lost, otherwise another diode is required in series with VSUPPLY. Assuming the DC:DC converters are 85% efficient, the load power as seen by the supercapacitor

$$= 3.5V \times 300mA/85\% + 1.8V \times 300mA /(85\% \times 85\%) + 1.2V \times 200mA /(85\% \times 85\%)$$

$$= 2.3W.$$

Fig 13 shows the results of providing power supply backup using one GS208.

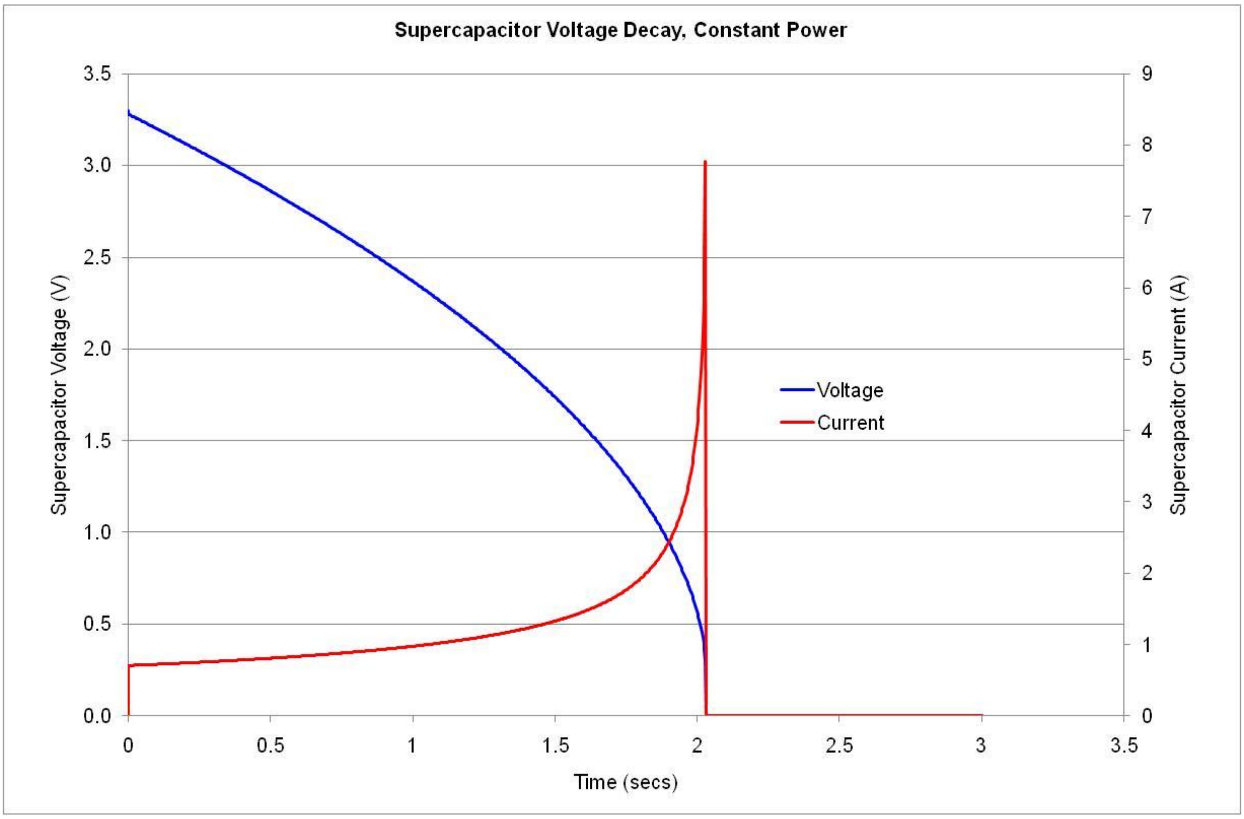


Fig 13: Supercapacitor Backup for a 3.3V supply maximum load using 1 x GS208

Fig 14 shows that 1 x GS208 supercapacitor can back up the maximum load for 1.8 secs. For them minimum load case, shown in Fig 14, the GS208 only decays to 1.8V after 3 secs, >> 1V minimum

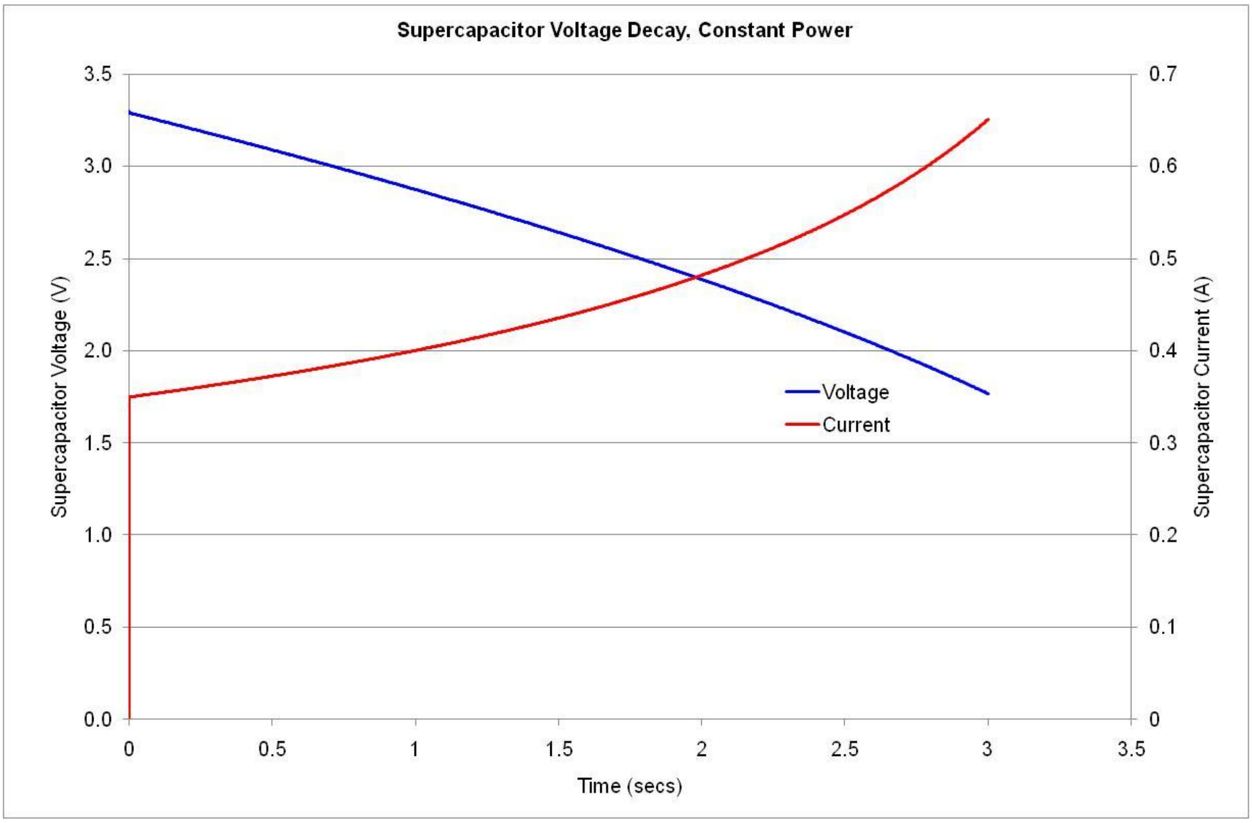


Fig 14: Supercapacitor Backup for a 3.3V supply minimum load using 1 x GS208

Fig 15 shows supercapacitor backup for a 3.3V supply using 2 x GS208 in parallel for the maximum 2.3W load.

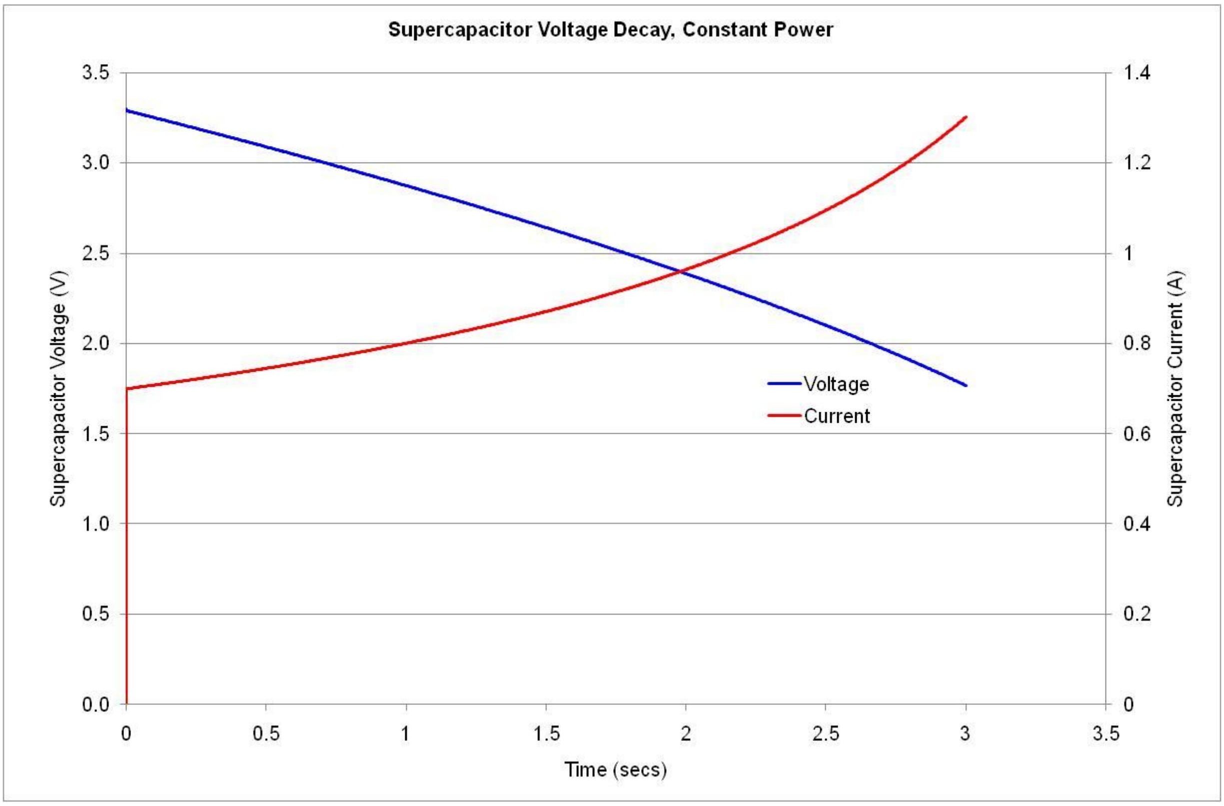


Fig 15: Supercapacitor Backup for a 3.3V supply minimum load using 2 x GS208 in parallel

Fig 15 shows that the 2 x GS208 support the maximum load for 3 seconds, with the supercapacitors decaying to 1.8V after 3 secs, >> 1V minimum supercapacitor voltage to supply 0.7V to the boost converter, allowing for the drop across D1.

Table 9 shows the input parameters to the Voltage Decay Simulator to obtain the results in Fig 15.

Voltage droop calculation, fixed power

ESR:	15	mOhm
Capacitance:	1.8	F
Load power:	2.3	W
Load duration:	3	secs
Initial Voltage	3.30	V
Voltage droop:	1.53	V
Min voltage	1.77	V
Time Step:	0.003	Secs

Table 9

6. Conclusions

This white paper has shown that:

- High speed cache memory greatly improves SSD performance.
- The best alternative for high speed cache memory in the medium term is SDRAM with power supply backup.
- Supercapacitors are preferable to batteries for power supply backup, and in particular CAP-XX supercapacitors with their thin flat form factor are an ideal solution.

Power supply topologies using CAP-XX supercapacitors have been provided with simulation results confirming they can support the SSD while it flushes the cache and gracefully shuts down with no loss of data in the event of power fail.