Ismail Oukid, Daniel Booss, Adrien Lespinasse

On testing persistent-memory-based software

Erstveröffentlichung in / First published in:
DOI: https://doi.org/10.1145/2933349.2933354

Diese Version ist verfügbar / This version is available on:
https://nbn-resolving.org/urn:nbn:de:bsz:14-qucosa2-806385
On Testing Persistent-Memory-Based Software

Ismail Oukid  
TU Dresden & SAP SE  
i.oukid@sap.com

Daniel Booss  
SAP SE  
daniel.booss@sap.com

Adrien Lespinasse*  
Grenoble INP-Ensimag  
adrien.lespinasse@ensimag.fr

Wolfgang Lehner  
TU Dresden  
wolfgang.lehner@tu-dresden.de

ABSTRACT

Leveraging Storage Class Memory (SCM) as a universal memory—i.e. as memory and storage at the same time—has deep implications on database architectures. It becomes possible to store a single copy of the data in SCM and directly operate on it at a fine granularity. However, exposing the whole database with direct access to the application dramatically increases the risk of data corruption. In this paper we propose a lightweight on-line testing framework that helps find a nd d ebug S CM-related e rrors that a n o ccur u pon software or power failures. Our testing framework simulates failures in critical code paths and achieves fast code coverage by leveraging call stack information to limit duplicate testing. It also partially covers the errors that might arise as a result of reordered memory operations. We show through an experimental evaluation that our testing framework is fast enough to be used with large software systems and discuss its use during the development of our in-house persistent SCM allocator.

1. INTRODUCTION

The advent of Storage Class Memory (SCM) is disrupting the database landscape and driving novel database architectures that store data, access it, and modify it directly from SCM at a cache-line granularity [3, 7, 9]. However, the no free lunch folklore conjecture holds more than ever as SCM brings unprecedented challenges. Consistency failure scenarios and recovery strategies of software that persists data depend on the underlying storage technology. In the traditional case of block-based devices, software has full control over when data is made persistent. Basically, software schedules I/O to persist modified data at a page granularity. The user level has no direct access to the primary copy of the data and can only access copies of the data that are buffered in main memory. Hence, software errors can corrupt data only in main memory which can be reverted as long as the corruption was not explicitly propagated to storage. In fact, crash-safety for block-based software highly depends on the correctness of the underlying file system. In contrast, SCM is byte-addressable and is accessed via a long volatility chain that includes store buffers, CPU caches, and the memory controller buffers, over all of which software has little control. The SNIA [2] recommends to manage SCM using an SCM-aware file system that grants the application layer direct access to SCM with mmap, enabling load/store semantics. As a side effect, changes can be speculatively propagated from the CPU cache to SCM at any time, and compilers and out-of-order CPU execution can jeopardize consistency by reordering memory operations. Moreover, changes are made persistent at a cache line granularity which necessitates the use of CPU persistence primitives. This adds another level of complexity as enforcing the kind of changes which are made persistent cannot be delayed like with block-based devices, and must be synchronous. In addition to data consistency, memory leaks in SCM have a deeper impact than in DRAM. This is because SCM allocations are persistent, hence, a memory leak would also be persistent.

Several proposals tackled these challenges following two main approaches. The first one focuses on providing global software-based solutions, mainly transactional-memory-like libraries, to make it easier for developers to write SCM-based software. Examples of these solutions include Mnemosyne [12], NVHeap [6], and REWIND [4]. The second and more mainstream approach is to rely solely on existing hardware persistence primitives, such as cache line flushing instructions and memory barriers to achieve consistency. Several persistent data structures were proposed following this approach, such as the CDDS B-Tree [11], the wBTree [5] and the NV-Tree [14]. Nevertheless, all approaches have in common that SCM-related errors may result in data corruption. In contrast to volatile RAM where data corruption can be cured with a restart of the program, data corruption in SCM might be irreversible as it is persistent. Therefore, we argue for the need of testing the correctness of SCM-based software against software crashes and power failures—which result in the loss of the content of the CPU cache. We remark that testing is orthogonal to devising recovery strategies that solve the challenges introduced by SCM.

In this paper we propose a lightweight automated on-line testing framework that helps detect and debug a wide range of SCM-related bugs that can arise upon software or power failures. We particularly focus on detecting missing cache line flushing instructions. Our testing framework is based on a suspend-test-resume approach and is able to simulate different crash scenarios including the loss of the content of the CPU cache. An important feature of our testing framework is its ability to avoid excessive duplicate testing by tracking the call stack information of already tested code paths, which leads to achieving fast code coverage. Additionally, our testing framework is able to detect errors that might arise due to the compiler or the CPU speculatively reordering memory operations. An additional capability of our testing framework is simulating crashes in the recovery procedure of the tested program, which we argue is very
important since hidden SCM-related errors in the recovery procedure may compromise the integrity of the data upon every restart.

We show with an experimental evaluation on a persistent B-Tree and a persistent SCM allocator that our testing framework exhibits good performance, even in the case of nested crash simulations, and is fast enough to be used on fairly large software systems. In particular, we demonstrate that taking into account the call stack can improve the testing time by several orders of magnitude.

This paper is organized as follows: Section 2 discusses related work and elaborates on our system architecture assumptions. Then, Section 3 presents our testing framework and its different optimizations. Thereafter, we evaluate our testing framework in Section 4. Finally, Section 5 concludes this paper and outlines future directions for a field, we believe, much in need of research.

2. BACKGROUND AND RELATED WORK

Crash-safety for disk-based software has been extensively researched and several tools that combine experimental testing and model checking have been proposed [10, 13]. Although they share the same goals, crash-safety testing for disk-based and SCM-based software are different in that they have to address radically different failure modes. Consistency and recovery testing of SCM-based software did not get much attention so far. Lantz et al. [8] proposed Yat, a hypervisor-based off-line testing framework for SCM-based software. Yat is based on a record-and-replay approach. First, it records all SCM write operations by logging VMM exits that are caused by writes to SCM. Persistence primitive instructions need to be replaced in tested software by illegal instructions to make them traceable by causing a VMM exit. Yat then divides the memory trace into segments, each of which is delimited by two persistence barriers. It considers that SCM write operations can be arbitrarily reordered within a segment, with the exception of writes to the same cache line which are considered to be of fixed order. Yat replays the trace until a non-tested segment is encountered, then runs the recovery procedure of the tested software for every possible reordering combination inside that segment. Since the number of combinations can grow exponentially, the authors propose to limit the number of combinations per segment to a certain threshold.

Given a program that covers the whole code base of the tested software, and given a sanity check program that detects any present data corruption–both of which are challenging to produce–, Yat may theoretically achieve comprehensive testing for single-threaded SCM-based software. In practice however, that may still require prohibitive testing time. In the case of a multi-threaded program, Yat records the sequence of operations executed by the various threads, which can differ between two runs due to the non-determinism of multi-threading. Hence, comprehensive testing of a recorded sequence does not imply comprehensive testing of the software.

In contrast to Yat, our testing framework performs on-line testing and is non-invasive as it does not require software changes in most cases. While it covers only partially memory-reordering-related errors, our testing framework achieves fast code coverage by limiting duplicate testing, and is able to automate crash testing inside the recovery procedure of a program, both of which Yat does not provide. We explain in Section 3.7 how our testing framework can be combined with Yat to make up for the limitations of both tools.

We conjecture that providing comprehensive consistency and recovery testing for large SCM-based software is practically not feasible. Instead, our focus is on improving the quality of such software by covering a wide range of SCM-related errors in a reasonable amount of time. We argue that similarly to concurrent software, providing theoretical correctness guarantees should be a prerequisite for any SCM-based software, and that experimental testing should not (and cannot) make up for the lack of such theoretical guarantees.

In this work we assume, without loss of generality, an Intel x86 architecture that provides the following persistence primitives:

- mfence: It is a memory barrier that guarantees that all load and store instructions finish executing before proceeding further.
- clflush: It is an instruction that evicts a cache line from the CPU cache and writes it back to memory. clflush is asynchronous and is not guaranteed to take place until an mfence is issued [1].

clflush needs to be wrapped by two memory barriers to be fully ordered and serialized. In the following we refer to a serialized clflush by the persist function. Other persistence instructions, such as clflushopt, clwb, and pcommit, have been announced by Intel. However, since there is currently no available hardware that supports them, we do not consider them in this work. Nevertheless, the flexibility of our testing framework makes it easy to take them into account in the modeling of SCM persistence in the future.

3. TESTING OF SCM-BASED SOFTWARE

We propose a lightweight on-line testing framework that is able to simulate software crashes and power failures that cover a wide range of consistency and recovery bugs. As an implementation example, we integrate the framework with our own persistent SCM allocator. The persistent allocator creates large files, referred to as segments, and then logically divides them into smaller blocks for allocation.

3.1 Crash simulation

In this section we explain how our testing framework is able to simulate software and power failures. The main challenge in simulating a power failure is to simulate the loss of the content of the CPU caches. To achieve this, we devise the following strategy that is based on a suspend-test-resume approach:

- For every segment created by the persistent allocator, a corresponding mirror segment is created.
- When a cache line is explicitly flushed, its content is copied from the original segment to the same offset in its corresponding mirror segment. Therefore, only explicitly flushed data is present in the mirror segments.
- Malfunctions that simulate a crash are randomly triggered when calling persistent primitive functions. These simulated crashes proceed as follows, as illustrated in Figure 1: (1) Suspend (pause) the main process by forking a test process and waiting on it; (2) the test process creates a copy of each mirror segment; and (3) the test process executes the binary of the test program which recovers...
using the created copies of the mirror segments. The test process starts then the recovery procedure the same as would have done the main process if it crashed where it was suspended. After recovery, the test process executes a user-defined consistency check procedure (step (4) in Figure 1).

- Once the test process exits, the copies of the mirror segments are deleted and the main program resumes execution (step (5) in Figure 1) until the next crash simulation.

This strategy is fully automated and the state of the main process is never changed during the crash simulation phase. More specifically, the state of the original segments and the mirror segments is never changed during the crash simulation phase. Capturing explicit flushes and replicating them in the mirror segments allows to simulate the loss of the CPU cache. In fact, this corresponds to the case where no cache lines are evicted from the CPU cache without being explicitly flushed. In a real scenario, some of the content of the CPU cache might have been speculatively written back. Nevertheless, we argue that by capturing only explicitly flushed data, we can more reliably detect missing flushes in the tested code.

Crash simulations are triggered inside the Flush function, as shows in Algorithm 1, because it is the only function that modifies the state of the mirror segments. In the case of a detected error or a crash in the recovery procedure or in the user-defined checking procedure, the test process is suspended and a debugger can be attached to both the main process and the test process. This greatly helps in finding the reason of the crash, since both the simulated crash scenario in the main process and the recovery crash in the test process can be fully traced and the content of their corresponding data structures examined. To help reproduce errors, a seeded random number generator is used to generate crash probabilities. Hence, in single-threaded execution, a bug is always reproducible as long as the same seed is used. Besides, it is possible to allow crash simulations for only a specific code region.

### Algorithm 1 Random crash simulation

1. procedure FLUSH(PPtr PAddr, Char Content[CacheLineSize])
2.   crashProb ← rand(0, 1)  ⬤ Get a random crash probability
3.   if crashProb < 0.5 then  ⬤ Crash with a probability of 0.5
4.     SimulateCrash()
5.   end if
6.   copyToMirror(PAddr, Content)
7.   crashProb ← rand(0, 1)
8.   if crashProb < 0.5 then
9.     SimulateCrash()
10. end if

To illustrate different classes of errors that our testing framework can detect, we consider in the following the case of a simplified array append operation. The correct code is:

```
array[size] = val;
persist(&array[size]);
size++;
persist(&size);
```

The newly appended value must be persisted before size is incremented and persisted. Consider the following code:

```
array[size] = val;
size++;
persist(&size);
```

In this case there is no guarantee that size will refer to only valid entries in the array after a failure, because appended values are never explicitly persisted. Our testing framework successfully detects such errors since array will never be updated in the mirror segments and a simple check of the content of array during a simulated crash will detect this issue.

### 3.2 Faster testing with copy-on-write

In practice, the cost of making copies of the mirror segments for every simulated crash is proportional to the size and number of segments. Therefore, this step can be prohibitively expensive for programs with a large memory footprint. To remedy this issue, we use copy-on-write memory mapping. Copy-on-write enables to read data directly from the mirror segments while copying only the memory pages that are modified by the test process. When the test program terminates, the memory pages that hold the changes that were made to the mirror segments are discarded. Hence, both the original segments and the mirror segments remain unchanged during the crash simulation phase, but at a much lower cost than that of making copies of the mirror segments.

### 3.3 Towards faster code coverage

Triggering crash simulations purely randomly gives no code coverage guarantees as some critical paths might not be tested while others might be tested multiple times. An alternative to the purely random approach is systematic crash simulation, similar to the one followed by Yat. This approach has the disadvantage of testing the same critical path as many times as it is executed. For instance, if a program appends one thousand values to a persistent array, systematic crash simulation will test the append function of the persistent array one thousand times, leading to prohibitive testing times.

### Algorithm 2 Trigger crash based on the probability of a call stack

```
1: procedure GETCRASHPROBABILITY(CallStack S)
2:   (prob, found) ← CallStackMap.find(S)
3:   if found then
4:     return prob
5:   else  ⬤ First time visiting this call stack
6:     CallStackMap.insert(S, 1)
7:     return 1  ⬤ Simulate crash
8: procedure RUSSIANROULETTE(CallStack S)
9:   prob ← GetCrashProbability(S)
10:  crashProb ← rand(0,1)
11:  if crashProb < prob then
12:    SimulateCrash()
13:  CallStackMap.update(S, prob/2)
```

To solve this issue, we propose to capture the program call stack when a crash is simulated as a means to limit duplicate testing. Basically, we cache the call stacks of the scenarios that were already tested and avoid triggering a crash simulation when the same call stack is visited again. However, the same call stack does not mean the exact same scenario. We argue that testing the same caller stack several times is beneficial in the sense that corner cases are more likely to be covered. Consider the following code for the previous example of an array append operation:

```
array[size] = val;
persist(array);
size++;
persist(size);
```

The error is that it is always the first cache-line-sized piece of array that is flushed instead of the cache line that holds the newly appended value. If a crash is simulated only once in the append operation, this error might remain unnoticed since the first appended value will be located in the first cache-line-sized piece of array. Therefore, we propose to exponentially decrease the probability of

---

1See the MAP_PRIVATE flag of mmap: [http://man7.org/linux/man-pages/man2/mmap.2.html](http://man7.org/linux/man-pages/man2/mmap.2.html)
simulating a crash with the same call stack. Algorithm 2 illustrates the different steps of this process. Basically, we map tested call stacks to a corresponding probability. If the call stack has never been visited before, we insert the new call stack in the map together with a corresponding probability of 1. This probability is divided by two whenever a crash is simulated at the same call stack. We use our own call stack unwinder that fetches the whole call stack to avoid any ambiguity.

3.4 Memory reordering

Although it is advised to always protect a flushing instruction with two memory barriers, it can sometimes be useful to group several flushing instructions together for optimization purposes—e.g., when the order of writes is not important. However, such optimizations may stem from wrong ordering assumptions, hence leading to errors. Consider again the previous example of an array append operation:

```c
void persist(size);  
size++;  
/* Missing memory barrier */  

1: array[size] = val;  
2: barrier();  
3: flush(&array[size]);  
4: /* Missing memory barrier */  
5: size++;  
6: persist(&size);
```

The code above tries to optimize by skipping one memory barrier that should be at line 4. As a result, lines 3 and 5 might be reordered by the CPU and the new value of `size` might be made persistent before the newly appended value.
duplicate testing by caching the call stack of the thread that triggered the crash simulation becomes less effective in a multi-threaded environment. This is because the crash scenario depends on the state of all threads and not only the thread that triggered the crash simulation. Nevertheless, as discussed in Section 2, exhaustive testing for multi-threaded programs is infeasible because thread scheduling is non-deterministic, hence, two runs of the same multi-threaded program may produce different scenarios.

### 3.7 Limitations and complementarity with Yat

While our testing framework is able to detect a wide range of consistency-related errors, it is unable to find errors that stem from the reordering of consecutive SCM writes whose persistence is delayed but still enforced in the right order. Consider the following code snippet of an array append operation:

```c
array[size] = val;
size++;
persist(&array[size]);
persist(&size);
```

The issue in the code above is that the new value of `size` can be speculatively evicted from the CPU cache and become persistent before the newly appended value. A power failure at this same instant would leave the array in an inconsistent state. Our testing framework cannot detect this issue because it captures only persistence primitives and not individual SCM writes. If lines 3 and 4 are swapped, our testing framework would detect it as a wrong persistence order. In contrast, Yat is able to detect such errors since it captures all SCM writes.

We advocate coupling our testing framework with Yat. Our testing framework can quickly test several classes of SCM-related errors, leaving out only the class of errors described above, for which Yat can be used. Yat can eliminate the scenarios that were tested using our testing framework and focus instead on SCM write reordering between two cache line flushing instructions. This significantly decreases the amount of combinations that Yat needs to test.

Unfortunately, there is another class of errors that neither our testing framework nor Yat can detect. Consider the example of Section 3.3:

```c
array[size] = val;
persist(&array);
size++;
persist(&size);
```

The error in this case is in line 2 where we persist the first cache-line-sized piece of `array` instead of the appended value. If the test program appends only a few values which fit in a single cache line, line 2 will still persist the newly appended value. Hence, the error will remain unnoticed both in our testing framework and in Yat. If in the release version of the software the array spans more than one cache line, data consistency will not be guaranteed.

### 4. EVALUATION

We evaluated our testing framework on a persistent B-Tree and on a persistent SCM allocator. The allocator code base is around 7000 lines of code (excluding blank lines and comments), which is fairly large for an evaluation. We use `tempfs` to simulate SCM. We consider the following test scenarios:

- **PTree-N**: The main program consists of inserting `N` key-value pairs in the persistent B-Tree, then erasing them. The test program that is executed upon crash simulation consists in executing recovery and checking that the tree is in a consistent state.
- **PAlloc-N**: The main program consists of interleaving `N` allocation and `N` deallocation of blocks of sizes between 128 Bytes and 1 KB. The test program consists in executing recovery, deallocating all currently allocated blocks, then inspecting the allocator against memory leaks.

In the case of PTree-N, the persistent tree uses the persistent SCM allocator. Hence, some of the simulated crashes will be triggered inside the allocator code. This is useful because it will stress the scenarios where a memory leak might happen because the persistent tree did not properly track its own allocations or deallocations.

First, we execute the test scenarios without allowing crash simulation in the test process (i.e. no nested crash simulation). Table 1 illustrates the number of persistence primitives, the number of crash scenarios, and the total testing time with and without taking into account the call stack. We observe that when not taking into account the call stack, the number of crash scenarios grows linearly with the number of operations. When taking into account the call stack however, the number of crash scenarios grows very slowly, which allows speedups of 41.7x and 30.5x for PTree-10000 and PAlloc-10000, respectively, compared with not taking into account the call stack.

We depict in Figure 3 the frequency of simulated crashes in visited call stacks for PTree-10000 and PAlloc-10000. We observe that when taking into account the call stack, the highest number of simulated crashes with the same call stack is limited to 16. When not taking into the account the call stack, this number increases up to...
Table 1: Performance of the testing framework with nested crash simulation disabled. Call stack is abbreviated as CS.

<table>
<thead>
<tr>
<th>Test</th>
<th>Main program stats</th>
<th>#Crash simulation</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>#Flush</td>
<td>#Barrier</td>
<td>w/o CS</td>
</tr>
<tr>
<td>PTree-100</td>
<td>747</td>
<td>1238</td>
<td>966</td>
</tr>
<tr>
<td>PTree-1000</td>
<td>4488</td>
<td>8440</td>
<td>8524</td>
</tr>
<tr>
<td>PAlloc-100</td>
<td>2254</td>
<td>4346</td>
<td>4034</td>
</tr>
<tr>
<td>PAlloc-1000</td>
<td>17643</td>
<td>38730</td>
<td>34902</td>
</tr>
<tr>
<td>PAlloc-10000</td>
<td>173479</td>
<td>384046</td>
<td>346580</td>
</tr>
</tbody>
</table>

Table 2: Performance of the testing framework with nested crash simulation enabled.

<table>
<thead>
<tr>
<th>Test</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>w/o CS</td>
</tr>
<tr>
<td>PTree-100</td>
<td>218s</td>
</tr>
<tr>
<td>PTree-1000</td>
<td>710s</td>
</tr>
<tr>
<td>PTree-10000</td>
<td>6120s</td>
</tr>
<tr>
<td>PAlloc-100</td>
<td>DNF (≈ 1.3d)</td>
</tr>
<tr>
<td>PAlloc-1000</td>
<td>DNF (≈ 11.5d)</td>
</tr>
<tr>
<td>PAlloc-10000</td>
<td>DNF (≈ 114.8d)</td>
</tr>
</tbody>
</table>

Figure 3: Frequency distribution of the number of simulated crashes per call stack. The y axis represents the considered range groups of the number of crash simulations per call stack, while the x axis represents the cardinality of these groups.

As a second step, we allow nested crash simulation in order to test the crash-safety of recovery procedures. We set a limit of one day per run. We report the results in Table 2. As expected, the testing time is higher than without nested crash simulation. Still, taking into account the call stack keeps the testing cost reasonable and yields significant speedups: 22.4x and 7352.6x for PTree-10000 and PAlloc-10000, respectively, compared with not taking into account the call stack. The three call-stack-oblivious PAlloc test scenarios exceeded the one-day threshold and we report an estimation of DNF.

5. CONCLUSION

In this paper we presented a lightweight automated testing framework for software that uses SCM as a universal memory, especially modern database systems. Our testing framework can simulate software crashes and power failures following a suspend-test-resume approach. It makes efficient usage of copy-on-write memory mapping to speedup the testing process. Additionally, it achieves fast code coverage by caching the call stacks of the already tested crash scenarios. Our testing framework is also able to simulate nested crashes in a fully automated way, that is, crashes that occur during the recovery procedure of a test program executing in an ongoing crash simulation. Our experimental evaluation shows that our testing framework successfully finds a wide range of consistency and recovery errors, and is fast enough to be used continuously during development of fairly large software systems.

The focus of our testing framework is to improve the quality of SCM-based software rather than to achieve comprehensive crash-safety testing, which we argue is infeasible. For future work, we plan to extend our testing framework to take into account data that is speculatively evicted from the CPU cache. We also plan to refine our memory model to take into account new persistence primitives, such as clwb and pcommit. Moreover, we plan to investigate new development life cycles that are fit for SCM-based software and that involve model-based verification. Indeed, we argue that theoretical consistency and recovery guarantees must precede experimental verification. Finally, we believe that testing of SCM-based software will receive increasing attention in the near future as real SCM-based systems start to emerge.

Acknowledgements

We thank the anonymous reviewers for their constructive comments that helped us improve the paper. This work is partially supported by the German Research Foundation (DFG) within the Collaborative Research Center “SFB 912/HAEC” as well as the Cluster of Excellence “Center for Advancing Electronics Dresden (cfaed)”.

6. REFERENCES


