A Hybrid Storage Page Management Method Based on Thermal Perception

Feng-jun Shang, Wen-kai Wang and Xuan-ling Chen

ABSTRACT

In recent years, with the rapid development of computer technology, because the phase change memory has many advantages, like non-volatile, high storage density and so on. It becomes the most likely replacement for DRAM in the future. But it also has disadvantages, like reading and writing asymmetric. It make PCM cannot completely replace the traditional DRAM memory. Therefore, the hybrid storage system based on PCM&DRAM becomes the main direction of storage technology. In this paper, we proposed that the page partition management of heat perception is based on PCM&DRAM hybrid storage system. It predicts write-hot page. The introduction of the concept of writing distance. It effectively reduces the number of writing operations and increases the PCM life. It improves the system data storage capability.

INTRODUCTION

With the rapid development of the computer, the requirements of the program are continuously improved. DRAM is the most widely used main memory, but the research of it has almost reached the bottleneck. In order to make up for the deficiencies of DRAM, PCM was born. Because of the non-volatile and other characters, it is expected to become the main equipment, but there are many problems to be solved. Like the number of PCM write limited. In general, it uses the hybrid memory system to improve the overall performance and has become more feasible. How to design a hybrid frame and design an algorithm of page replacement based on heat perception is the focus of this paper.

RELATED WORK

This paper proposes a hybrid main memory page management method based on hybrid main memory architecture, which combines IR and IRR, and combines the
recent write information of the page to define the writing heat of the page. When a page is replaced by main memory, it is not necessary to immediately discard its write information, save it for a period of time, so that when it is replaced again into the main memory To determine its write heat, so as to allocate the appropriate main memory medium.

Migration mechanism: When a page in PCM because of a new write request to change their write heat, the system will immediately migrate it from the PCM in the DRAM; the contrary, if the DRAM in the page write heat cooling , Will be moved to the PCM. In the migration, the migration between different media in different ways, when the page from the PCM into the DRAM, we use the immediate migration mechanism; from the DRAM to the PCM, we use delayed migration strategy. In addition, we prefer to migrate less frequently written but frequently read pages in DRAM to PCM.

PAGE MANAGEMENT ALGORITHM BASED ON THERMAL PERCEPTION

System Design

Formatting the title

PCM&DRAM page memory management algorithm based on hybrid memory architecture is to ensure memory hit rate and reduce the number of writing on PCM. The key is to predict writing heat page accurately and store on DRAM in time.

Definition1. RRD: The number of pages have been written during the last two writing operations of a page. After a page into memory, if there is only once or no writing, then the RRD is infinite.

Definition2. WR: The other number of pages after the page was last written.

If RRD value is smaller, the heat of the corresponding page is higher. But we need both RRD and WR to define it. It prevent a RRD not to be write-hot page which is not writing for a long time and low value. We try to put the write-hot page to DRAM and put the write-cold page to PCM.

The basic structure

The algorithm maintains three different clock lists: 1. Common clock: manages all main memory pages, and use the traditional clock algorithm and replacement strategy to ensure its accuracy. 2. Write clock: write the least page of information, the purpose is used to determine the page write heat; 3. DRAM exchange clock: write clock from the management to be replaced out of the DRAM Page read-only page in DRAM. When a DRAM page in the DRAM swap clock list is hit write request, we move it to the write clock page. However, when all of the DRAM pages have been updated recently or are write-hot pages, they are stored in the clock, so the DRAM exchange clock list is empty.
Each list has one or more pointers. The Hg pointer in an ordinary CLOCK points to the least accessed page recently. It is equal to HAND point of CLOCK. Hhot points to write-hot page that has the most value of CLOCK. It shows a threshold of changing during write-hot page and write-cold page. The pointed page shows tail of CLOCK. If the pointer is moved, the write-hot page will turn into a write-cold page, or write-cold page is out of CLOCK. It shows the least accessed write-cold page recently when Hcold points nearest the write-cold page clockwise. By limiting the CLOCK to reduce historical space pay. When page number in CLOCK exceeds the upper limit, move Hcold to remove write-cold page from clock. When Hcold points the nearest write-cold page from Hhot clockwise in DRAM, trigger the move from DRAM to PCM. And we can find a write-cold page in DRAM quickly, otherwise, it is empty. Finally, Hswap points the tail of CLOCK on DRAM. And the DRAM page that is removed from CLOCK will be added to anti-clockwise point. It shows the header is placed into the corresponding linked list..

**Data Detection And Migration Strategy**

When a page is written on PCM and it is written in clock during detection period, the page is considered to be a write-hot page. And it need to be remove to DRAM. If DRAM memory have no free space, we should get space by move a DRAM page to PCM. First, we should check the exchanging clock on DRAM. If it is not empty, we select the page from the exchanging clock on DRAM to exchange prior. Because of the Hswap pointing to tail of clock, we begin with the anti-clockwise. Priority search page that read access bit is set to "1" that has reading tendency. If there is no page set to “1”, so we select the head of list to be the moving page. If the DRAM exchanging clock is empty, it shows the DRAM pages are all in clock. And we should select write-cold page from writing clock. First, ensure the DRAM write-cold page is exist by Hswap. If writing access bit and detection flag of the moving DRAM page are set, update them to “0” and add them into the head of DRAM. If HDcold is empty, it shows the DRAM page are all write-hot page or to be. So we need to start writing heat transfer mechanism to get write-cold page.
The purpose of the process is to find a write-cold page in DRAM to migrate to PCM. Figure 2 shows the changing and migrating of page writing heat. In figure 2(a), all pages are write-hot page or to be. So, HDcold pointer is empty. When it access page A, set its writing access to “0” and consider it is write-hot page. So it need to be moved to DRAM. But there is no space to store page A on DRAM, so move it to PCM. Writing heat exchange machine change the c to write-hot page and move it to head of linked list. Now write-hot page number exceeds a specified threshold and then trigger Hhot pointer to move clockwise. It also changes to write-cold page that pointed to B. Because of the B storage in DRAM, the HDcold pointer is updated. Then, Hhot and Hcold pointers are pointing to write-hot F and write-cold A in figure 2(b). When Hhot pointer scans over the page E, it shows detection period of the page is over. So remove it. Finally, by exchanging page B on PCM, we get a free DRAM storage space A, in figure 2(c). And the HDcold pointer is empty again.

In this algorithm, the page replacement decisions by the general clock list, and Hhot pointer will change write-hot page to write-cold page that is not in memory. In figure 3(c-d), if pointer points to page G of PCM, it will trigger writing heat exchange machine. Page A will change to write-hot page and it must arrive at linked head. In figure 3.2(c), write-cold page is pointed by Hhot pointer. But page F is not exist. The migration will be late from PCM to DRAM in figure 3.2(d) until it is accessed again. Figure 3.2(d) shows when Hhot scan to page H, we don’t remove it immediately. Because the page is accessed during detection period. But it has not exchanged to write-hot page yet. So we can give it a chance again to be in.

EXPERIMENT

Through an effective simulation system based on the architecture of the PCM&DRAM hybrid memory test algorithm, the PCM and DRAM unified addressing, each page size is set to 4KB. And compared with the traditional CLOCK algorithm, CLOCK-DWF algorithm and D-CLOCK. In this experiment, the artificial Zipf data set is used to test the validity of the experiment. For the Zipf data set with a total of N pages, the probability that the I page is accessed can satisfy the following formula:

\[ p_i = \frac{1}{H_N^{1-\theta} \times i^{1-\theta}} \]  

(1)
Among them, $\theta=\log a / \log B$, $a$ and $B$ represent the access locality of the data set. For example, $a=0.6$, $b=0.4$ represents 60% of the data access focused on the page of 40%. All of the four generated Zipf data sets are available for only 10000 different pages, but with different read / write ratios and locality.

<table>
<thead>
<tr>
<th>Traces</th>
<th>Footprints</th>
<th>Read/write ratio</th>
<th>Locality</th>
<th>Total Request</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zipf1982</td>
<td>10000</td>
<td>10%/90%</td>
<td>80%/20%</td>
<td>400000</td>
</tr>
<tr>
<td>Zipf1955</td>
<td>10000</td>
<td>10%/90%</td>
<td>50%/50%</td>
<td>400000</td>
</tr>
<tr>
<td>Zipf2873</td>
<td>10000</td>
<td>20%/80%</td>
<td>70%/30%</td>
<td>400000</td>
</tr>
<tr>
<td>Zipf4682</td>
<td>10000</td>
<td>40%/60%</td>
<td>80%/20%</td>
<td>400000</td>
</tr>
</tbody>
</table>

One of the goals of the algorithm is to reduce the number of write on PCM, so the Zipf series of data sets, we respectively for different memory capacity, different DRAM and PCM capacity ratio experiment.

Figure 3 shows the effect of different capacity of PCM and DRAM in same memory capacity for PCM writing number. The Y axis represents the number of writes caused by each algorithm to PCM. X axis represents the capacity ratio of PCM and DRAM. In figure, we know when PCM capacity is increasing, the PCM writing number that run different data sets are increasing. And this algorithm is better than others. And with the depression of DRAM capacity, it is lower than D-CLOCK algorithm in Zipf1982. Because we set the upper limit of write-hot page that is DRAM capacity, but the upper limit value is lower than the write-hot page number of data set. So a part of write-hot pages have to be defined to write-cold page and move between PCM and DRAM.

CONCLUSION

In this paper, we proposed a method of page partition management based on heat perception and introduce the idea of IRR. And we combined the recent page with writing information to define the writing heat of the page. By a writing information with history of writing page ring, we can judge writing heat of a page and store the write-hot page to DRAM in time to reduce the number of operations on PCM. The simulation results show that this method can effectively reduce the number of pages written on PCM. For the hybrid storage architecture, we will conduct a more in-depth study in the future.

ACKNOWLEDGMENTS

The author would like to thank the Chongqing Basic and Frontier Research Project under Grant NO. cstc2016jcyjA0590. The work is partly funded by the National Nature Science Foundation of China (No. 61672004).

REFERENCES


