vTMM

Paper summary:

The paper proposes a smart design to balance memory pages between fast and slow memory sub-systems in multi-VM co-running virtualized environments. Because stat-of-the-art techniques suffer from severe overheads due to software-based balloning and page table scanning techniques, the paper leverages hardware features for virtualization, notably Intel PML, to overcome these limits.

Strengths:

- The paper focuses on virtualized environment and cloud is nowadays the prevalent execution environment for user applications
- The paper leverages existing hardware features for virtualization

Weaknesses:

- The paper needs to be better motivated, especially the virtualization focus: it is clear to me in which mean does vTMM specifically differ from the state-of-the-art solutions in the virtualization virtualization context (appart from the PML utilization)

Comments to the authors:

++ Motivation

1. Is hybrid memory system a common practice in the cloud? Have you investigated the predominance of such environment with cloud providers?

++ Design

1. Consider increasing the size font in Figure 1 for the smallest strings like "multi-level queue", it is very difficult to read
2. It will be good to number steps in the design's figure for the reader to have a better overview. In [3.2.1 second paragraph], you talk about an "initialization" step which we don't know where it comes from: is it the VM boot? The vTMM process initialization? etc.
3. In the description, you talk about GPT pages which are, according to my understanding, pages of the (guest) page table. If so, it is not clear to me if you are tracking guest processes' pages or only pages of the PT. Or is it just a mistake?
4. [3.2.3 Intermittent monitoring]: not clear what is the notion of *counter* here, it has not been defined
5. [3.3.1 Page-degree]: *read_count* and *write_count* not defined relatively to the page whose degree is been calculated
6. [3.4 PML-based migration]: doesn't KVM already implement live migration with PML like it is the case with the Xen hypervisor? Or are you the one to introduce it in KVM?
7. [3.5 vTMM memory pool]: how does vTMM behaves when the free remaining memory in the pool is lower than current requested memory for migration round?

++ Evals

1. How is the NUMA CXL simulated?
2. You often invoke pages limitation while I consider they are many redoundant text throughout the paper; for example, each time you refer to the related work you re-explain what is already presented and explained in section 2.2.
3. [4.3.2 Page tracking overhead]: if I understand Figure 2.b and the text that explains it, does this mean that we a sufficient MWS the stat-of-the-art solution is able to provide same results (or even approximatively closed to) as vTMM? -this is even the case for redis from MWS=600ms in Figure 2.b-. If so, I am not sure to perceive what is the importance of a small MWS: scanning the VM memory more frequently does not incur more degradation?
4. [4.3.3 Multi-level queue]:
   - In Figure 3 is the y-axis the frequency or the cumulated one (CDF)?
   - *We observe that the frequency of hot pages is higher in the distribution when enabling multi-level queue*: you stated in the previous sentences that the two distributions are basically the same, and this is what is even observable in the figure; so why this affirmation?
5. [4.4.2 page migration]: in Table 6, after a migration round, are A/D bits reset? If so, how can the number of VMTraps generated by PML be less than 1? Am I missing something in the exeperiment explanation?
6. [4.5 Ablation study]: why are the other benchs ommitted here?
7. [4.8 THP support]: I would suggest to put vlines in Table 7 for readability. Similarly in Figure 7, or increase the gap between clusters

++ General Notes

- I would suggest to review the text redoundancy in the document, and save space for more important results in the evaluation or the motivation
- Figures need to be refined: adjust the fonts, the placement in the document (e.g., In Figure 6 the title of bottom figures are attached to the x-axis of top one and it makes it diffucul to detect and read), etc.
- There are many small typos, like duplicate words in sentences or verbs conjugation (e.g., Introduction 7th paragraph: *[...] a shared **a** memory pool.*), but they don't make the text non-understandable anyway.
- It may be only on my side, but in all graphs, the gray pattern with bubbles inside does not appear when printing the document on paper