Advanced Parallel Memory Virtualization

Zhang Yulei <yuleixzhang@tencent.com>
Tencent Cloud can serve globally with large scale of resources
Agenda

• Background
• Design
• Future works
Background

- Significant performance drop after live migration
  - Multiple VCPUS
  - Numerous memory
  - with huge page table
Overhead Analysis

• Numerous page fault happens after migration
• setup the page mapping for guest try to access memory
• enlarge the memory will cause performance decline
Current EPT setup

EPT Violation

- Address is mmio
  - Yes: mmio emulation
  - NO: Fast page fault

  - Can fast page fault
    - NO: Ping guest page
      - acquire mmu_lock
        - update page table
          - release mmu_lock
Our Proposal

• Improve concurrency of the vcpus with pre-setup page table
• lockless update the R/W status
Overview

- KVM_SET_USER_MEMORY_REGION: create page table
- KVM_GET_DIRTY_LOG: enable dirty log
- kvm->slots_lock
- vMMU Page Table
- Page table update
Fast pin guest memory and setup page table

- KVM_SET_USER_MEMORY_REGION
- Ping guest memory
- setup pgtable
- global_root_hpa
- root point
- mmu_load
- entering guest

VCPU...VCPU
Live migration support

- KVM_GET_DIRTY_LOG
- Break down to 4K granularity if necessary
- Support PML
  - Set WP on PTE
- Clear the D-bit in PTE
Update page entry with mmu_lock
Lockless access

vMMU page table

VCPU VCPU ...... VCPU
Performance data

Test VM with 32 vCPUs and 64G memories, force each vCPU to dirty 2G memory.

<table>
<thead>
<tr>
<th>Page size</th>
<th>Normal (second)</th>
<th>Pre-population(second)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4K</td>
<td>18~21</td>
<td>2~2.5</td>
</tr>
<tr>
<td>2M</td>
<td>3.2~3.6</td>
<td>2~2.5</td>
</tr>
</tbody>
</table>
**Benefit for Parallel Memory Virtualization**

- lockless access
- Eliminate the page fault latency
- Save system resource
  - mmu notification
  - Shadow page caches
  - Parent reverse mapping
Limitation

1. TDP enabled mode
2. SMM is not supported
3. memory overcommit is not supported
Future works
Future works

- Post copy in live migration
Link to the source code

https://lkml.org/lkml/2020/9/1/425