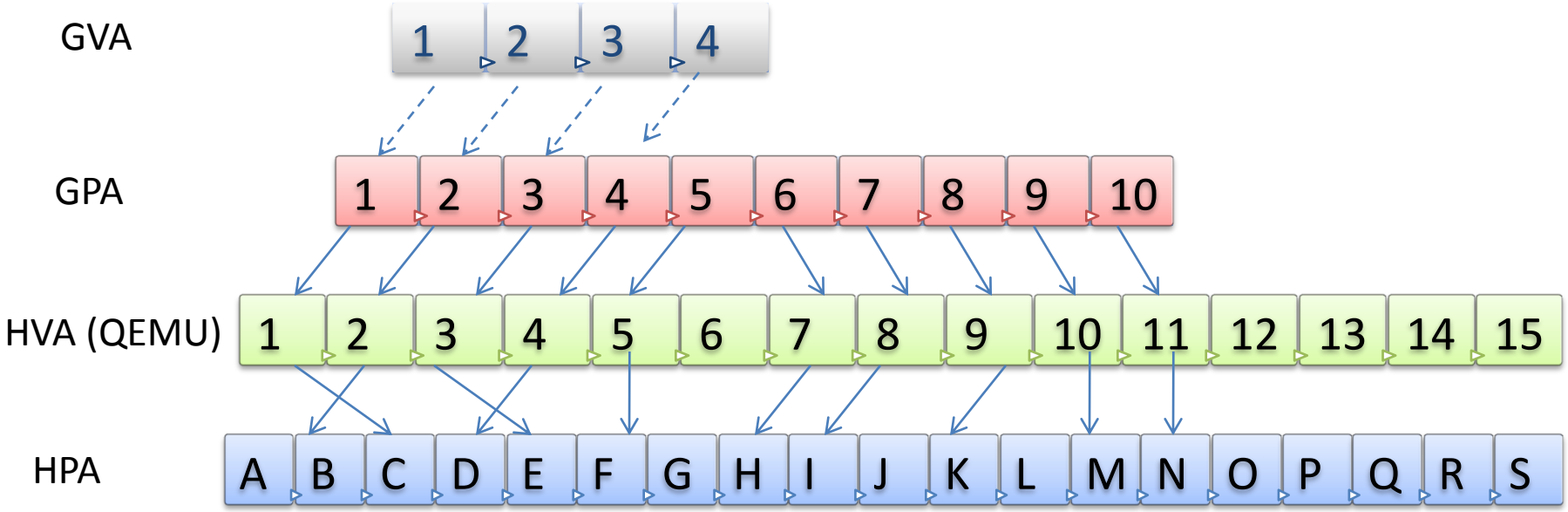
# Shadow page table

- Problems in memory virtualization
  - 3 levels of indirection, MMU can translate 1 level
  - GVA -> GPA -> HVA -> HPA must be achieved

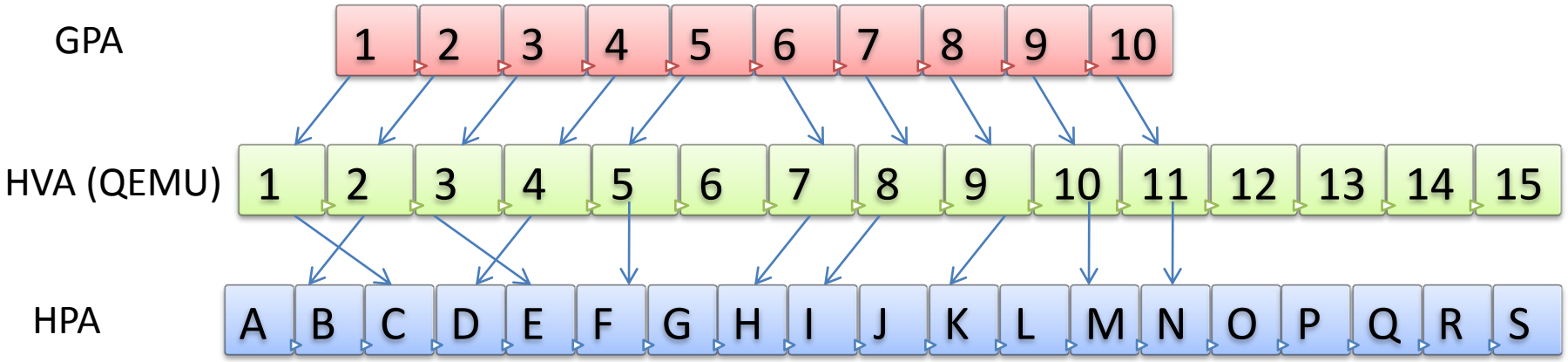
- Solution1 - Shadow page table
  - Contains  GVA -> HPA. MMU will use this instead of guest page table
  - One shadow table for each guest page table
  - Incrementally build

# Shadow page table building

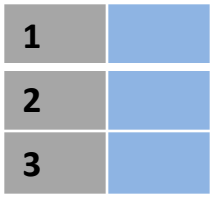



- Guest wants to create a linear mapping for a process
- Guest does pure demand
- QEMU knows GPA-> HVA mapping ( malloc())

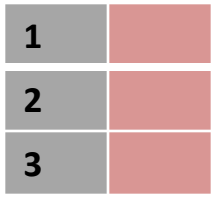
# Shadow page table building




Shadow page table
GVA -> HPA



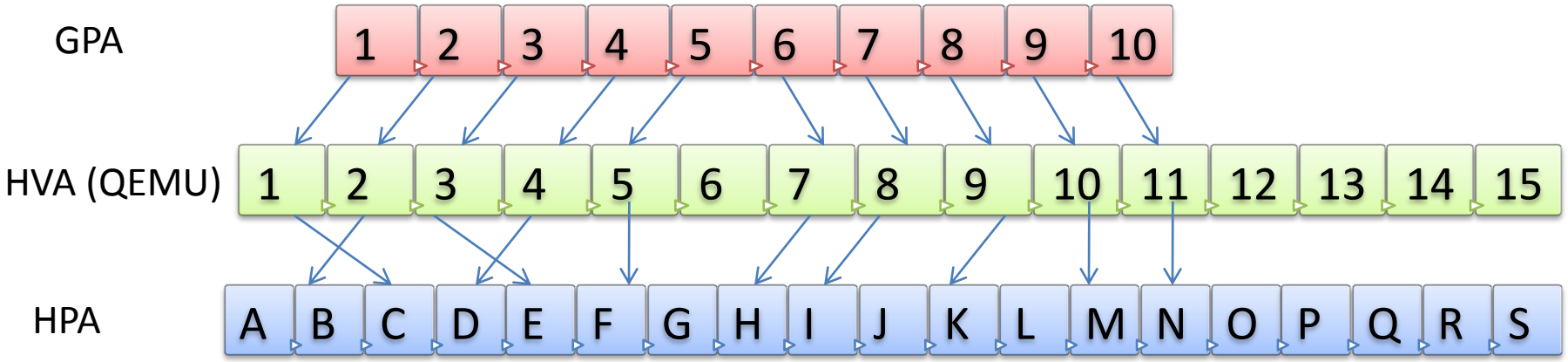

Guest process page table
GVA -> GPA ( Read only)



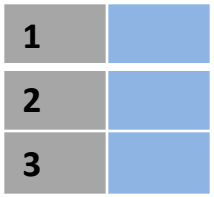

Step 1:

- Guest tries to map GVA 1 -> GPA 1
- Page fault (because of RO) causes VM exit
- KVM sees GPA as 1 by instruction emulation /using register contents

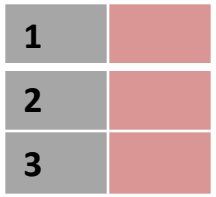
# Shadow page table building




Shadow page table
GVA -> HPA




Guest process page table
GVA -> GPA ( Read only)

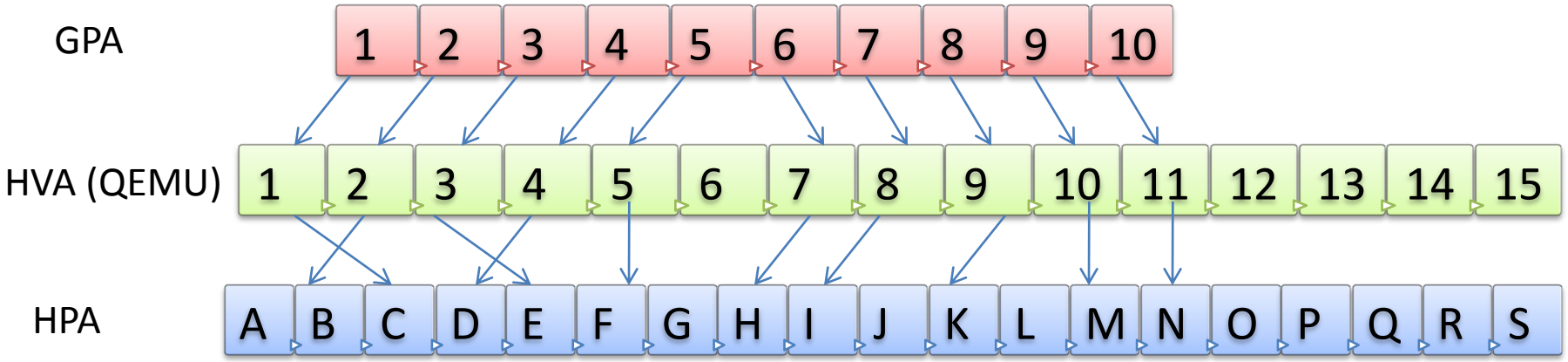



Step 2:

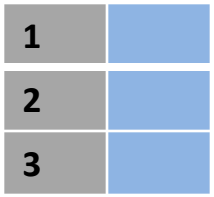
•GPA 1 -> HVA 1 is obtained

• This possible because GPA -> HVA mapping is known to QEMU/KMV

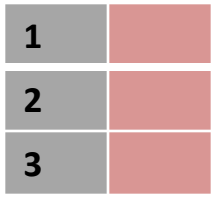
# Shadow page table building




Shadow page table
GVA -> HPA



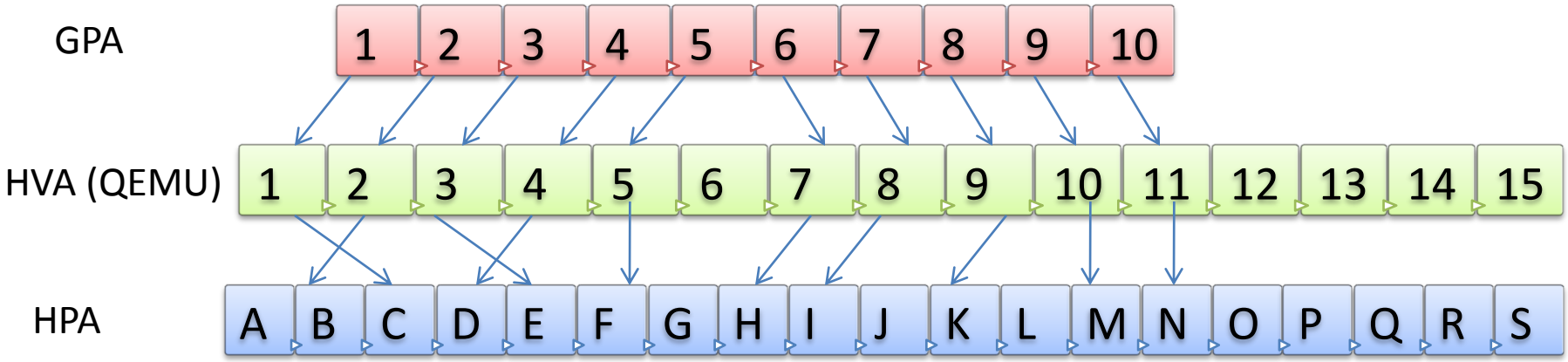

Guest process page table
GVA -> GPA ( Read only)




Step 3:

- KVM does lookup on QEMU's page table to find out HVA->HPA
- KVM finds out HVA 1 -> HPA C

# Shadow page table building




Shadow page table
GVA -> HPA

| | |
|---|---|
| 1 | C |
| 2 | |
| 3 | |

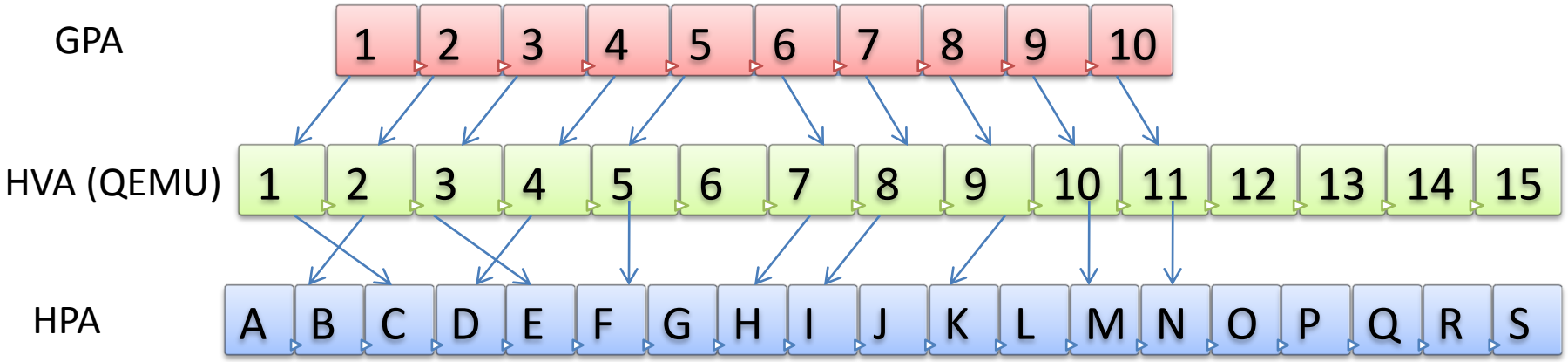
Guest process page table
GVA -> GPA ( Read only)

| | |
|---|---|
| 1 | 1 |
| 2 | |
| 3 | |

Step 4:

- KVM updates shadow page table with GVA 1 -> HPA C
- KVM also updates guest page table – by emulating the instruction which tried to map GVA 1 -> GPA 1
- GVA -> GPA -> HVA -> HPA is done

# Shadow page table building




Shadow page table
GVA -> HPA

| | |
|---|---|
| 1 | C |
| 2 | B |
| 3 | E |

Guest process page table
GVA -> GPA ( Read only)

| | |
|---|---|
| 1 | 1 |
| 2 | 2 |
| 3 | 3 |

Step 5:

- Similarly other entries are update as and when page fault happens
- GVA 2 -> GPA 2 -> HVA 2 -> HPA B
- GVA 3 -> GPA 3 -> HVA 3 -> HPA E