

# OVS-DPDK Shared mempool improvements

Kevin Traynor

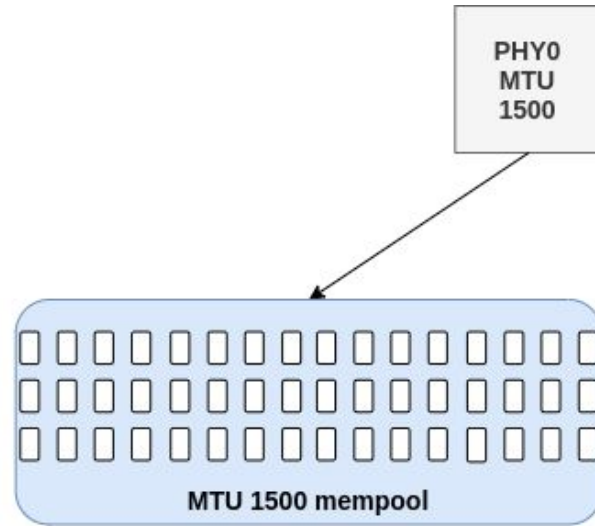
[ktraynor@redhat.com](mailto:ktraynor@redhat.com)

# OVS-DPDK mempools

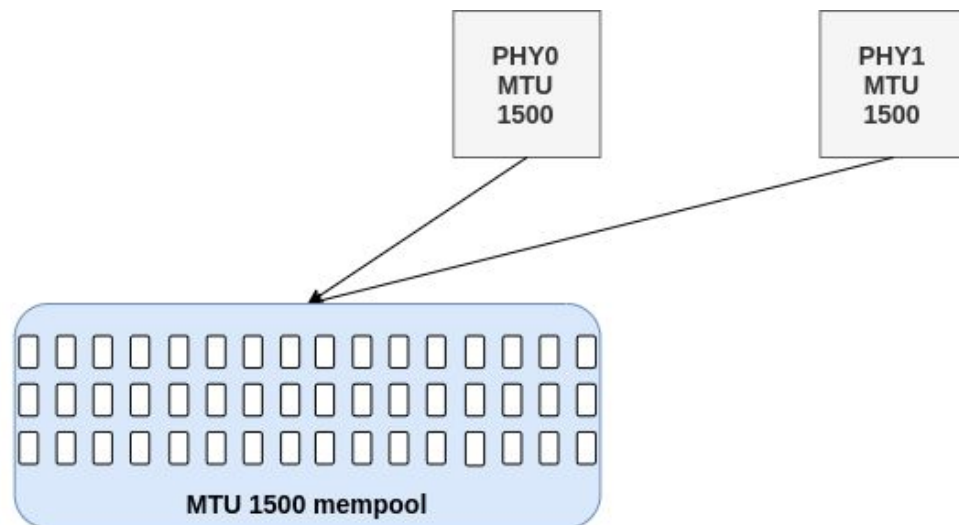
- ▶ OVS-DPDK preallocates a number of buffers (mbufs) into a group (mempool)
- ▶ Avoids allocation on datapath
  
- ▶ Two operating modes: Shared mempools and Per-port mempools
- ▶ Shared mempools
  - Can be used by multiple ports based on NUMA and MTU
  - Larger mempools
  - Default
- ▶ Per-port mempools
  - Can only be used by one port
  - Smaller mempools
  - May be useful for small number of ports (e.g. <10) but do not scale well
  
- ▶ Two recent improvements in Shared mempools
  - User configured shared mempools
  - Delayed vhost mempool creation

User configured shared mempools

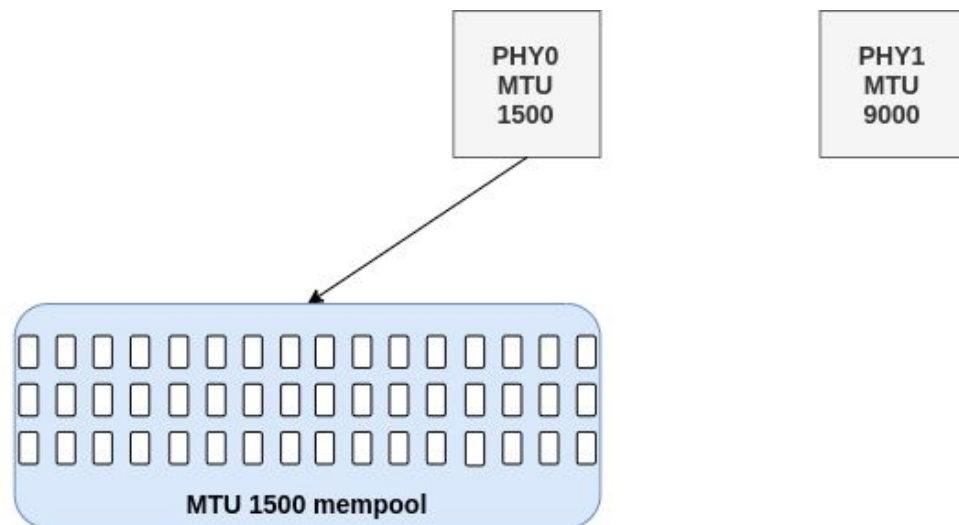
# Add port for a physical NIC with DPDK - create mempool



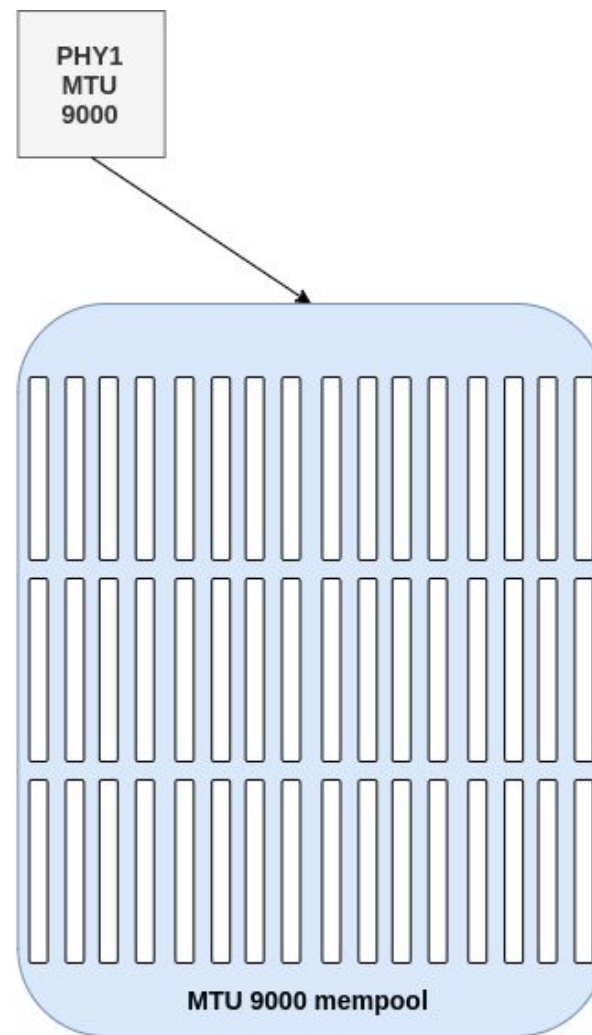
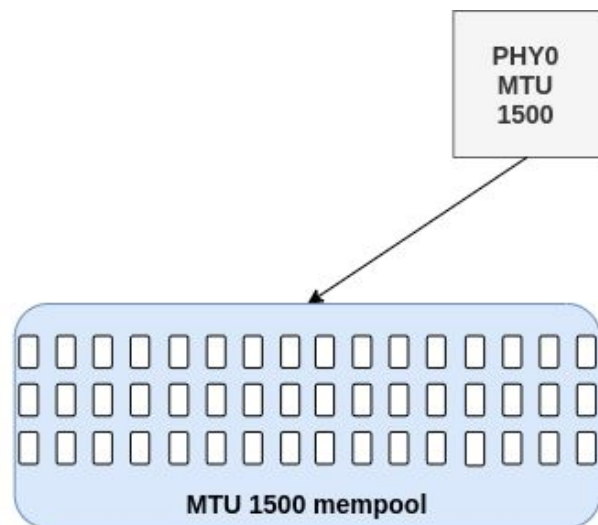
## Add another port - share mempool



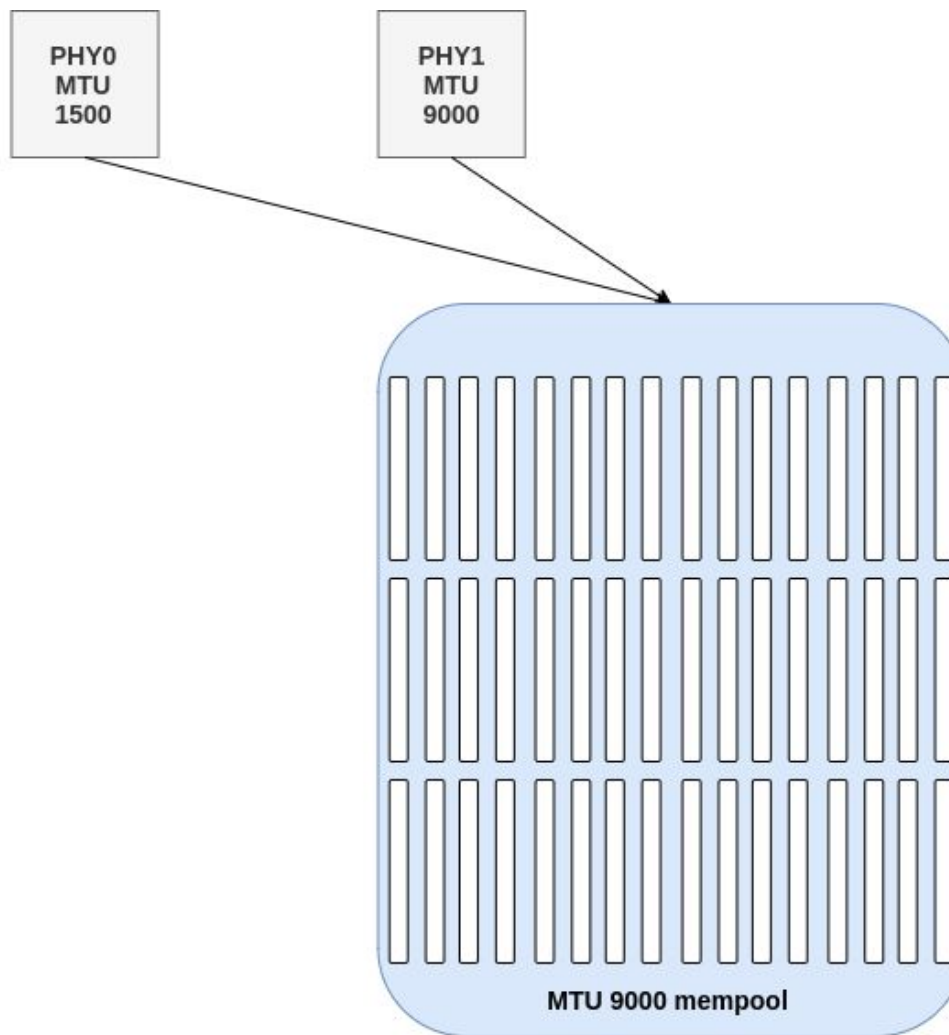
Change MTU of one port to 9K - cannot share mempool anymore



# Create new mempool

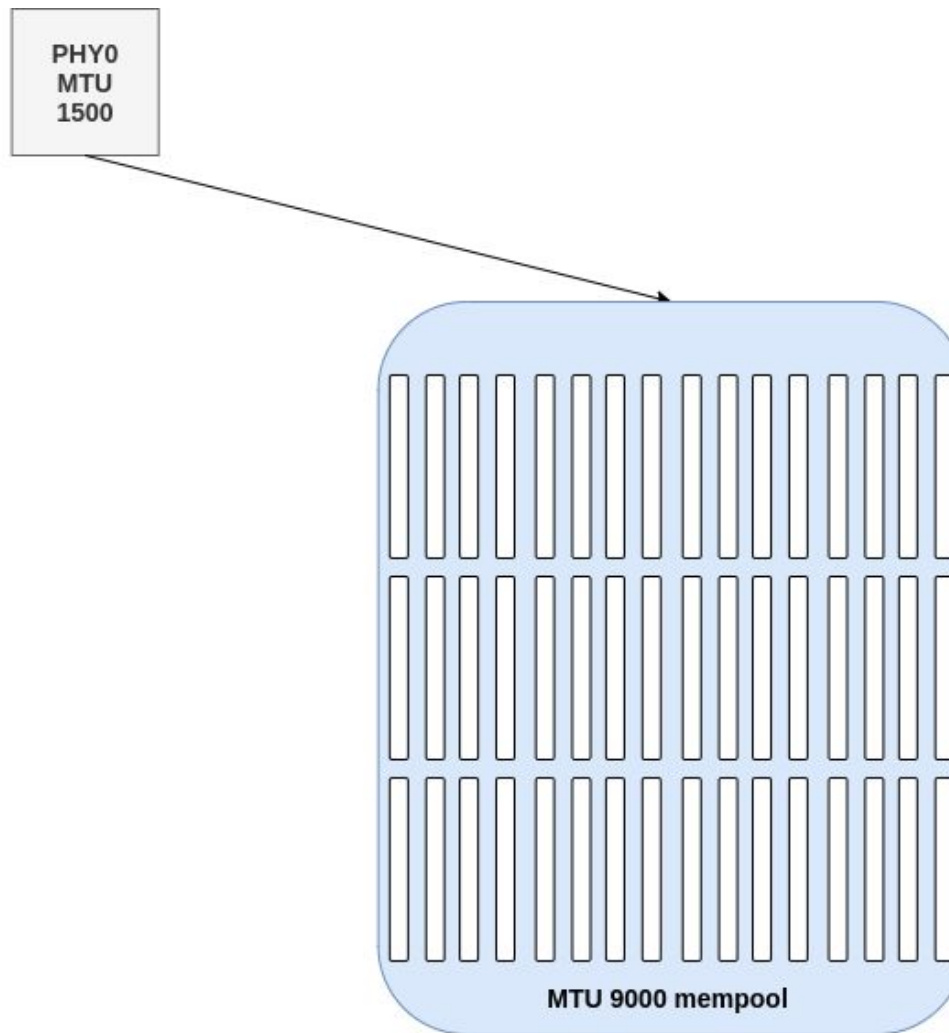


Could we now do this instead?



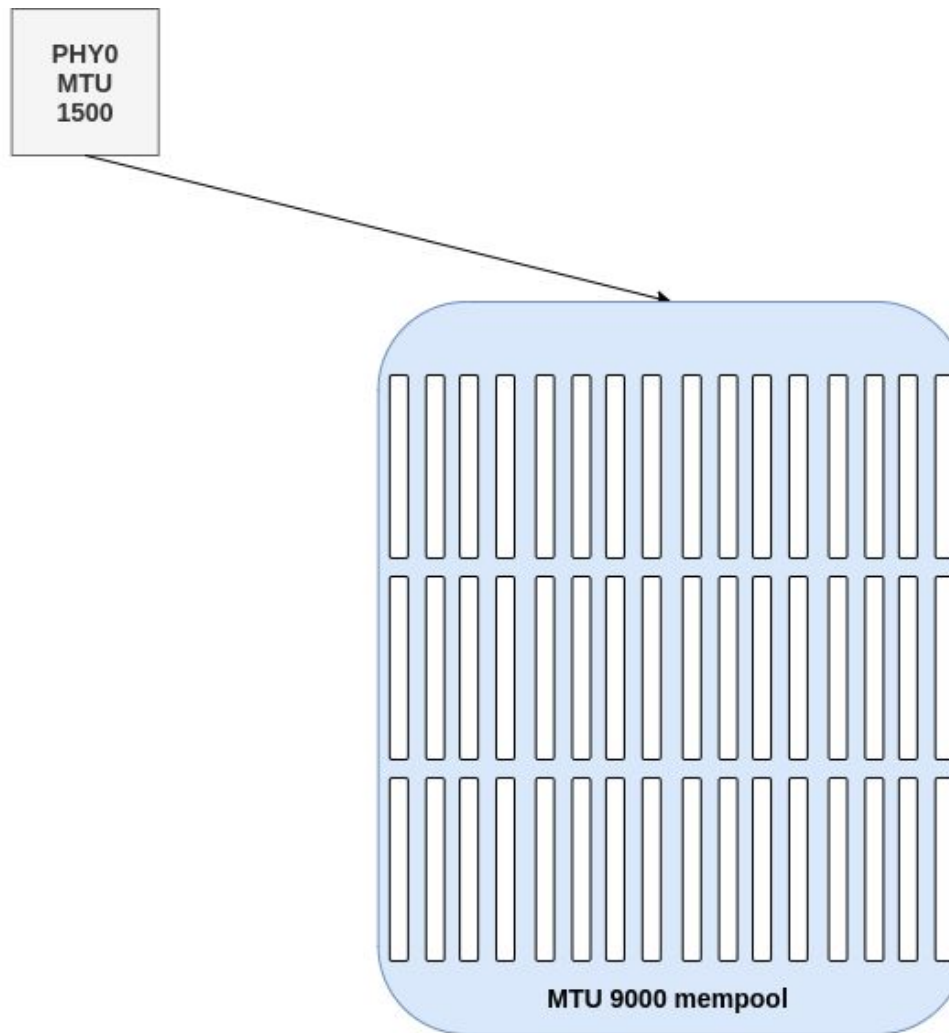


Could we have done this instead?

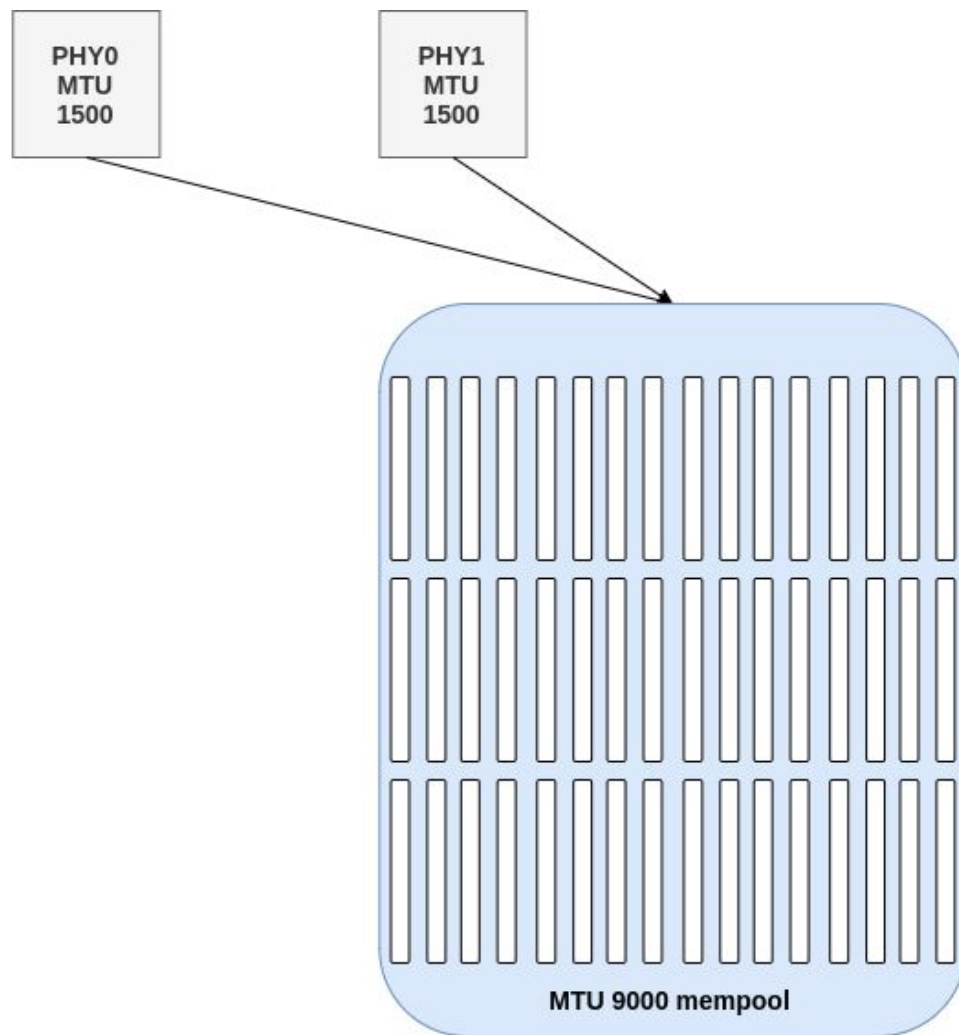


What if the user gave a hint?

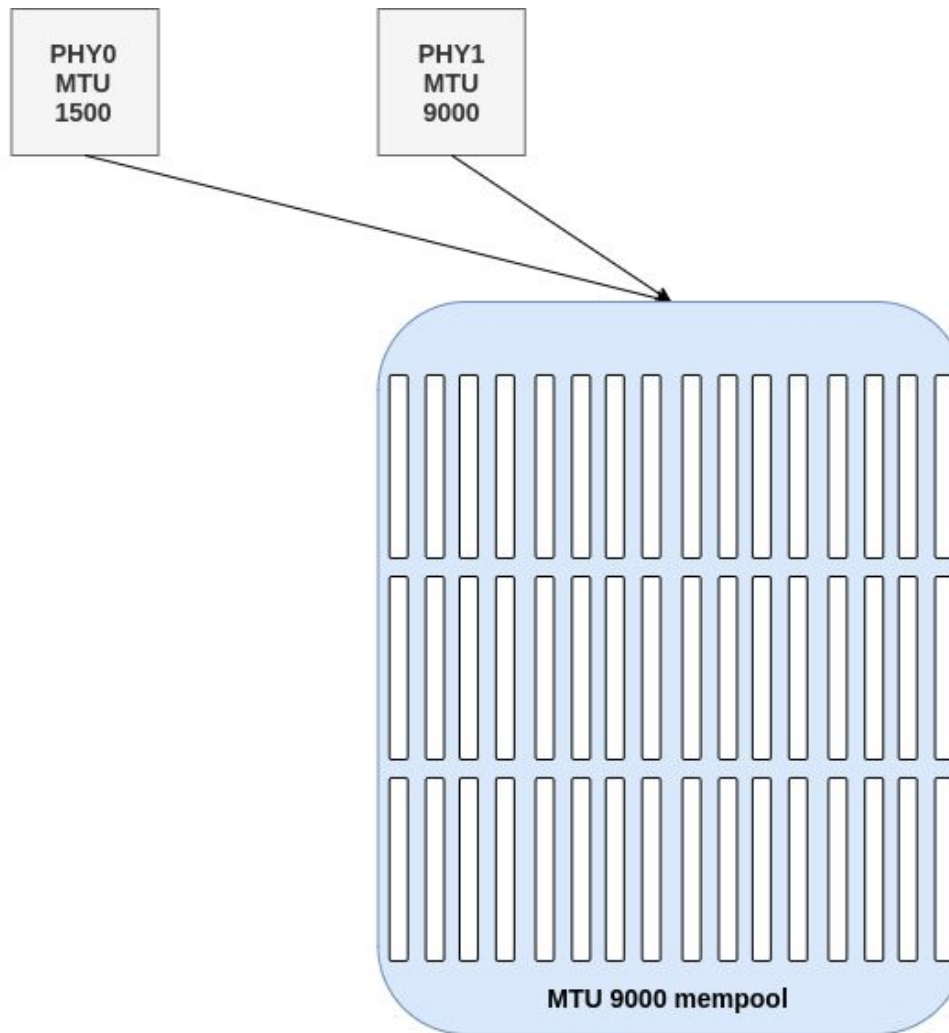
```
ovs-vsctl --no-wait set Open_vSwitch . other_config:shared-mempool-config=9000
```



## Add another port - share mempool



# Change MTU of one port to 9K

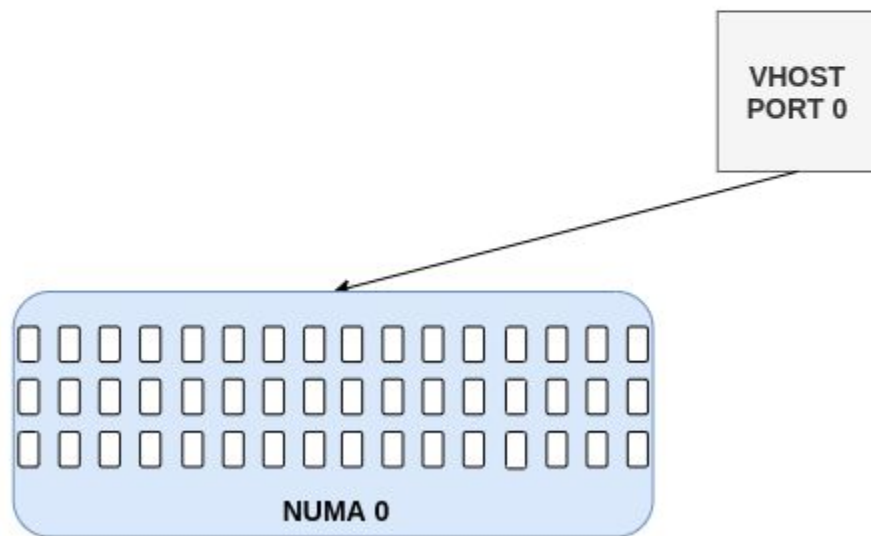


# More details

- ▶ Allows sharing of mempools between ports on the same NUMA with different MTUs
  - Avoids the need for multiple mempools based on MTU
- ▶ `ovs-vsctl --no-wait set Open_vSwitch .`  
`other_config:shared-mempool-config=<MTU>[:NUMA],...`
- ▶ `other_config:shared-mempool-config=9000`
  - User configured shared mempool hint of 9000 on any NUMA
- ▶ `other_config:shared-mempool-config=1500:0, 9000:1`
  - User configured shared mempool hint of 1500 on NUMA 0 and 9000 on NUMA 1
- ▶ Available in OVS 3.0
- ▶ Defaults are unchanged, requires user hint
- ▶ MTU outside of range will fallback to existing behaviour
- ▶ Hint is only used for when mempools need to be created

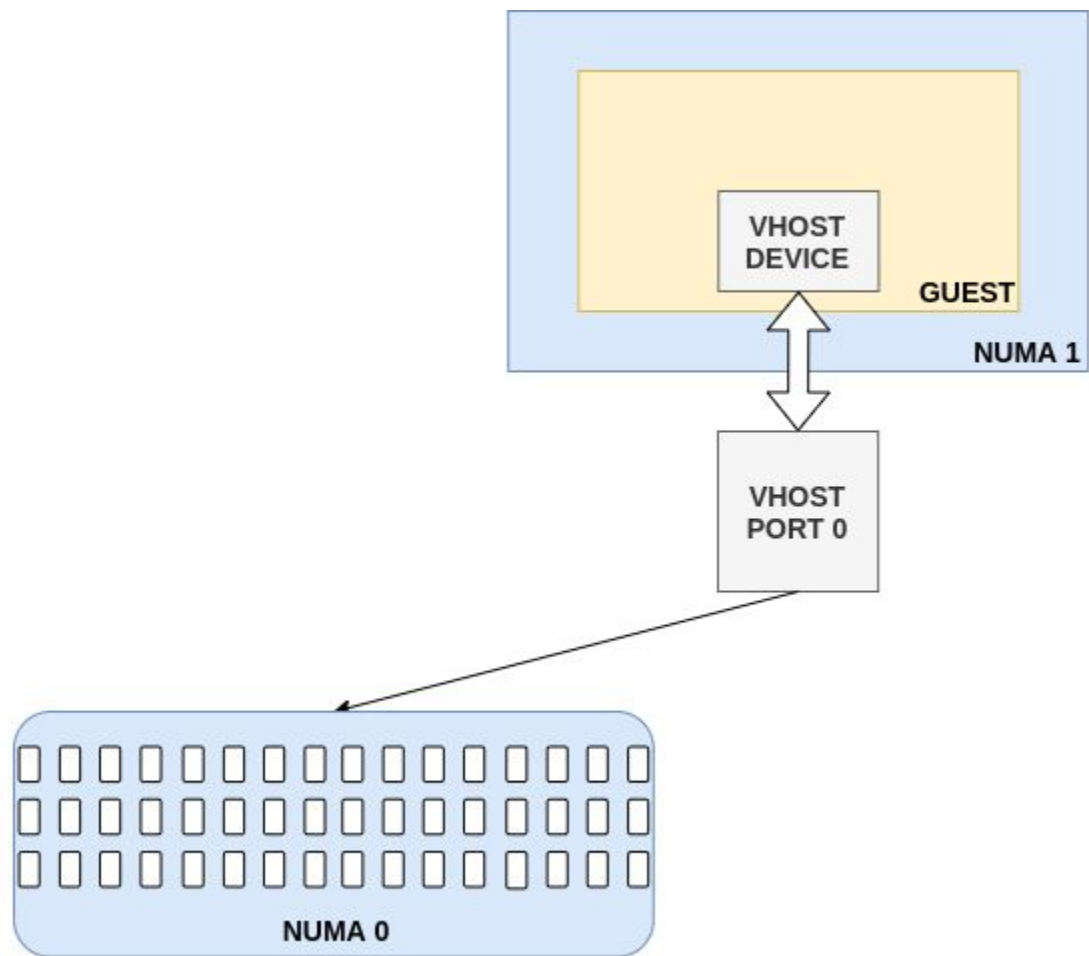
# Delayed vhost mempool creation

# Add port for vhost-user - create mempool

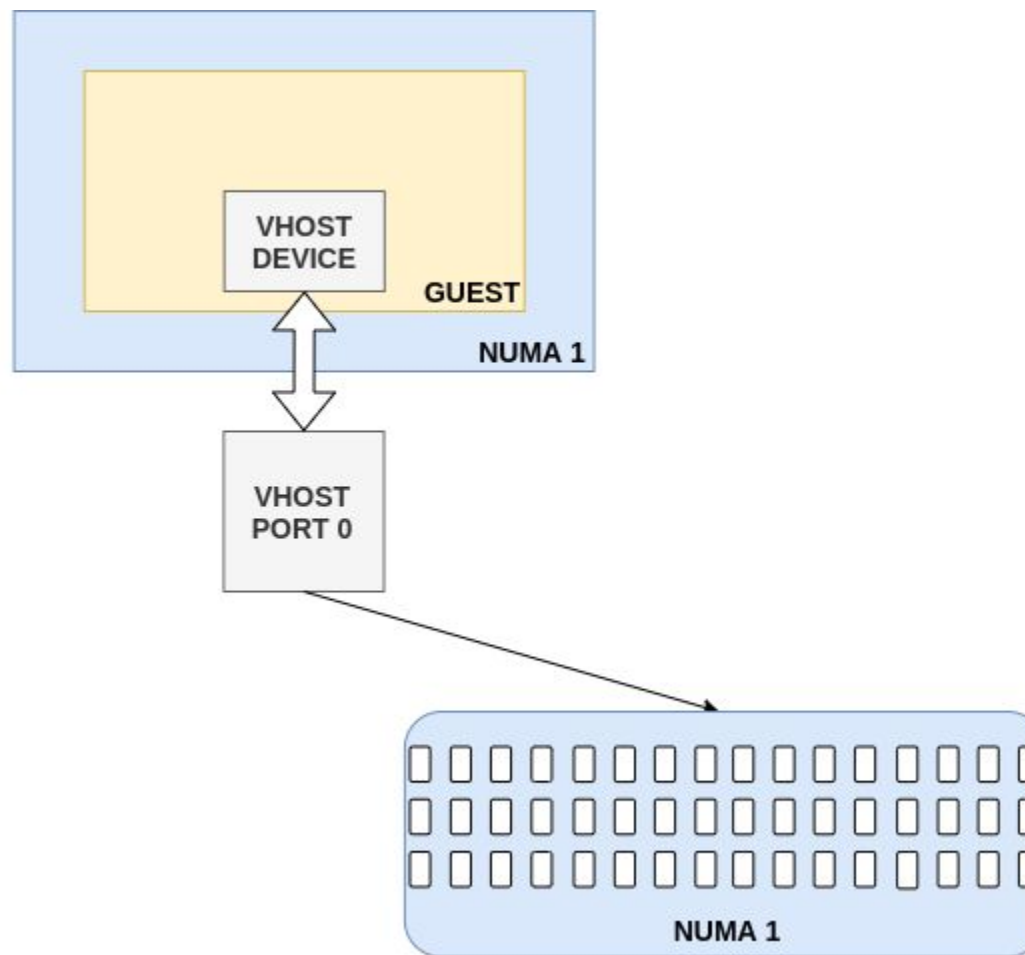




# Boot VM - mempool is on wrong NUMA



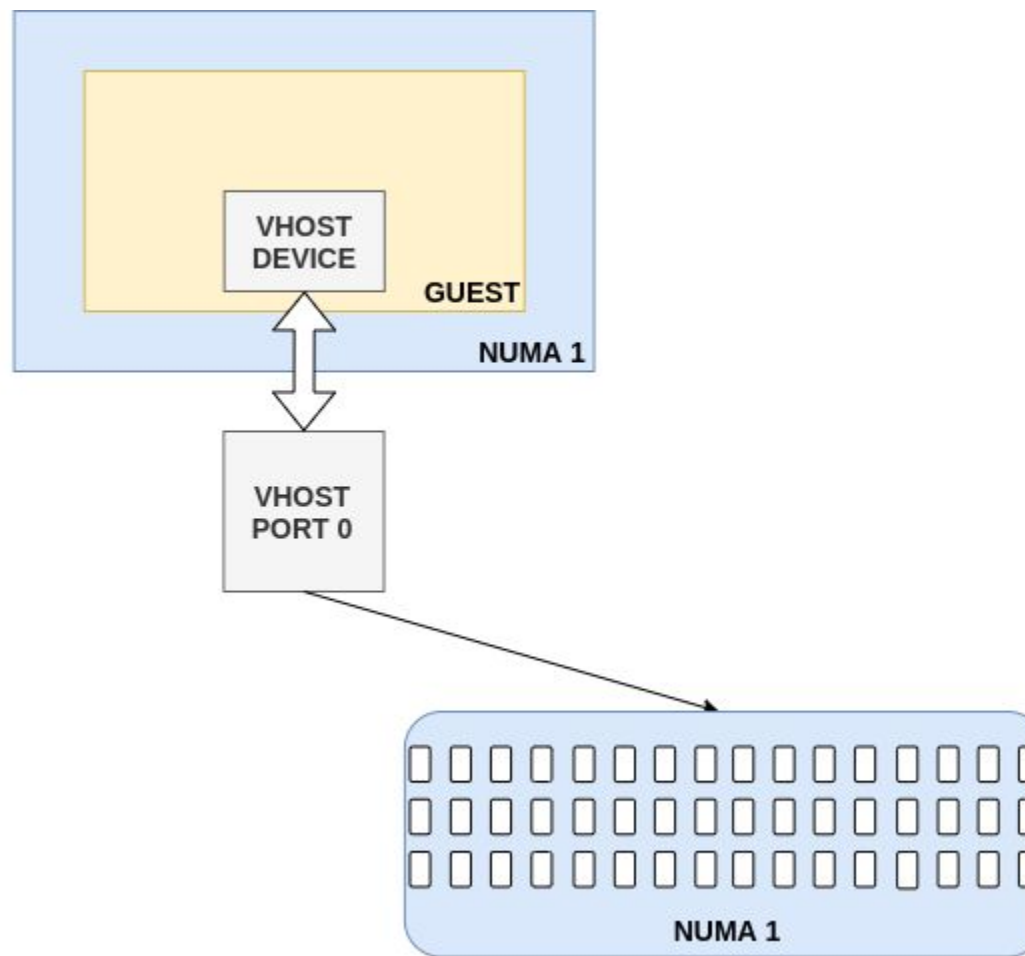
# Create mempool on correct NUMA



Add port for vhost-user - do not create mempool



# Boot VM - create mempool once NUMA info is available



# Delayed vhost mempool creation

- ▶ Avoids creating a mempool before it is needed
- ▶ Avoids creating a mempool on what ultimately ends up being the wrong NUMA
- ▶ Avoids user having to provision for mempool on a NUMA where it might not be needed
- ▶ Default in OVS 3.0

# Thank you

Red Hat is the world's leading provider of enterprise open source software solutions. Award-winning support, training, and consulting services make Red Hat a trusted adviser to the Fortune 500.



[linkedin.com/company/red-hat](https://www.linkedin.com/company/red-hat)



[youtube.com/user/RedHatVideos](https://www.youtube.com/user/RedHatVideos)



[facebook.com/redhatinc](https://www.facebook.com/redhatinc)



[twitter.com/RedHat](https://twitter.com/RedHat)