PTP in Media Virtualized Environment

Alex Vainman Sr. Staff Engineer
Mellanox

IP SHOWCASE THEATRE AT IBC2019 : 13–17 SEPT 2019

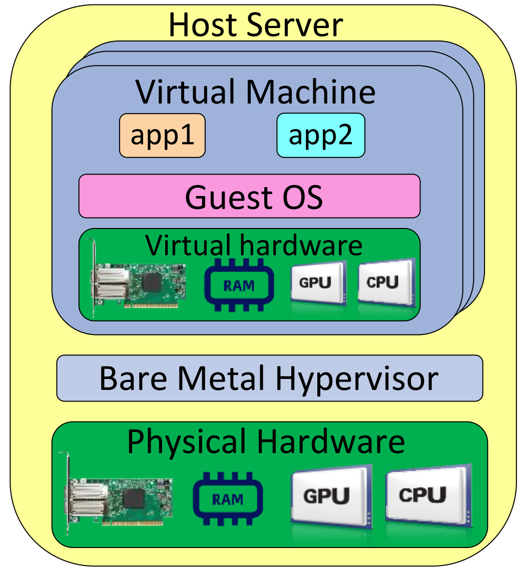
Intro & Agenda

- The M&E industry is moving to virtualized environment, cloud on premise and finally to cloud deployment
- This is the next trivial step of adopting SW IP solutions and its advantages – but is it trivial?
- Among the list of challenges, timing is probably high in the list
- In this presentation we will try to describe the virtualization environment, understand some of the key challenges, existing solution and looking forward

IP SHOWCASE THEATRE AT IBC2019 : 13–17 SEPT 2019

Virtualization Environment Intro

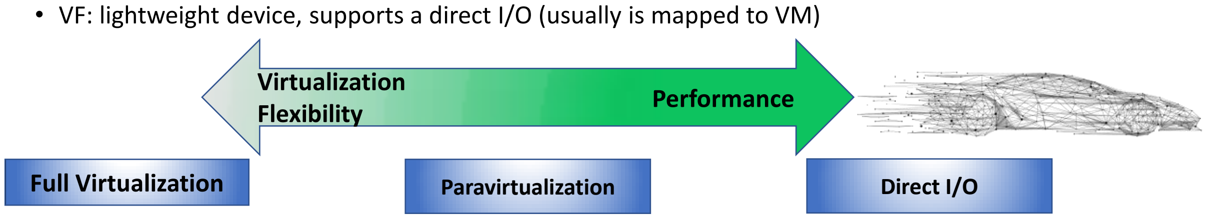
- **Hardware Virtualization:** Physical resources abstraction done by software
- **Host**: Physical server that is hosting the virtualization environment
- **Virtual Machine / Guest Machine**: Emulates a physical server in software
  - **Guest OS**: Operating System that is running on the VM
- **Hypervisor**:
  - Decouples the VMs from the physical server
  - Creates and runs VMs
  - Allocates and shares physical resources for VMs



Mellanox TECHNOLOGIES

I/O Virtualization Modes

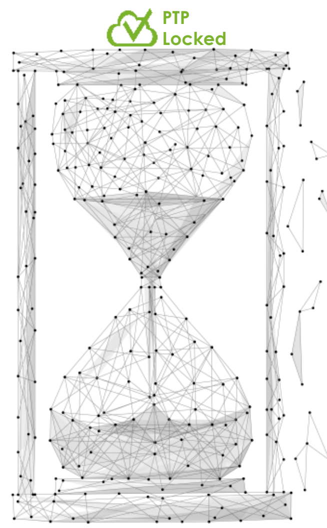
- **Full Virtualization**: I/O is trapped and emulated by HV, Guest is unchanged
- **Paravirtualization**: Modified Guest is doing I/O via hyper-calls to HV
- **Direct I/O:**
  - **Passthrough**: A single VM owns the device and directly accesses it
  - **SR-IOV**: PCIe device is partitioned to multiple logical PCIe devices
    - PF: is fully capably device (usually is mapped to trusted entity)
    - VF: lightweight device, supports a direct I/O (usually is mapped to VM)



Mellanox TECHNOLOGIES

SMPTE 2110 Timing Requirements

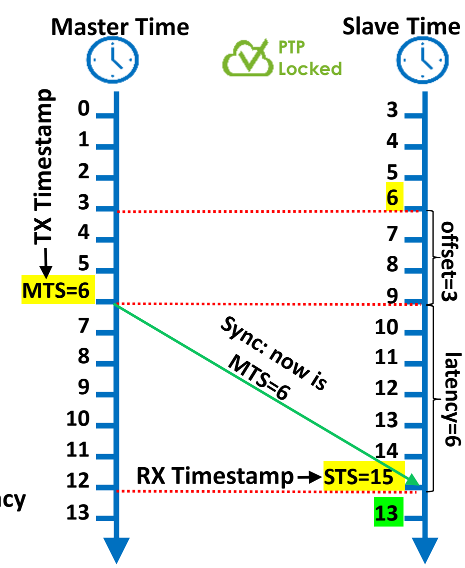
- Time synchronization shall be achieved using PTP
  - Protocol for clock synchronization over a network
  - PTP Master distributes time to Slaves via PTP messages
  - Achieves sub-microsecond range accuracy
  - ST2059-2 PTP profile must be supported
- Media signals timing must be according to ST2059-1
- ST2059 accuracy requirements:
  - Maximum +-500ns offset from the master clock
    - Slave shall be synchronized within 5 seconds
- For UHD streams sub-microsecond accuracy is critical



PTP Clock Offset Calculation: The Idea

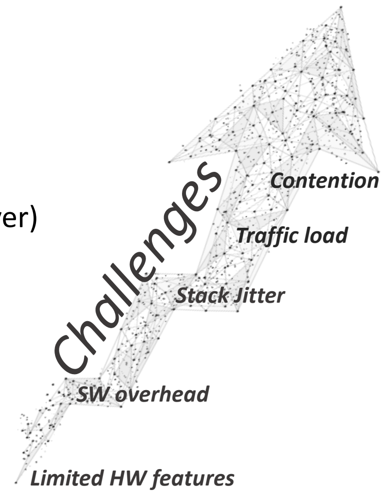
How slave synchs to Master Clock (**simplified**):

- Master sends to slave current time, **MTS** (TX Timestamp)
- Slave takes RX timestamp (**STS**) when it gets MTS
- *Offset_from_master = STS - MTS - latency*
  - *Latency = RTT / 2 (RTT = Master2Slave_Lat + Slave2Master_Lat)*
  - **Problem: if Master2Slave_Lat and Slave2Master_Lat are not symmetric the calculated offset is not accurate**
- PTP packets timestamping has a big impact on calculated latency and clock offset accuracies
  - Software based: not accurate due to OS jitter
  - HW based: very accurate (nanoseconds range)
  - **HW Timestamping is a must to get sub-microsecond PTP accuracy**



Virtualized PTP Performance Challenges

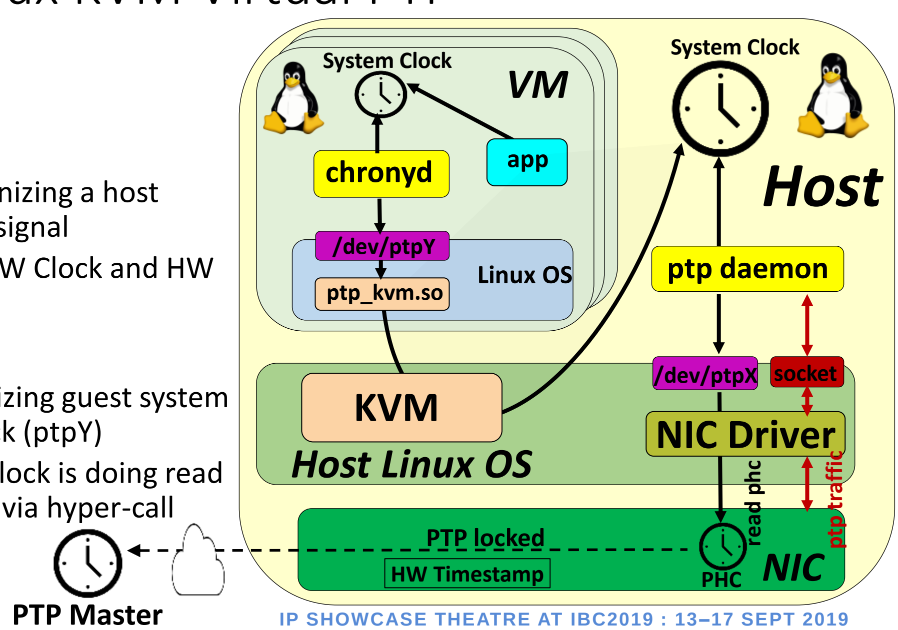
- Virtualization environment implications and costs:
  - Extra SW overhead
  - Increased resources contention (CPU, memory, network…)
  - Higher network stack jitter
  - Higher PTP traffic load (multiple PTP clients per physical server)
  - Limited ability to utilize HW features
- This hurts PTP accuracy for VMs
  - PTP SW timestamping is less accurate
  - Higher risk to exhaust Grand Master capacity
- In the next slides we will see how to overcome these challenges and get a near native performance for VMs



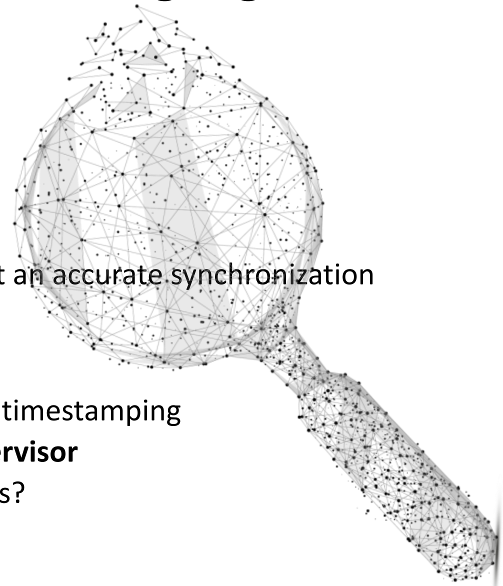
Linux KVM Virtual PTP

How it Works:

- Host
  - PTP client is synchronizing a host system clock to PTP signal
  - Utilizing Linux PTP HW Clock and HW timestamping
- Guest
  - Chronyd is synchronizing guest system clock to VM PTP clock (ptpY)
  - Read from VM PTP clock is doing read of host system clock via hyper-call



Linux KVM Virtual PTP Highlights



- Performance: **tens of nanoseconds for VMs**
- **Just a single PTP stack is running**
  - Minimal PTP traffic load
- NIC HW requirements:
  - **PTP packets HW timestamping is required** to get an accurate synchronization
  - Suitable for all I/O virtualizations modes
- SW requirements & restrictions
  - NIC Host driver must support Linux PHC and HW timestamping
  - **Supported for Linux guests and Linux KVM hypervisor**
  - But what about other guest OSes and hypervisors?

Mellanox TECHNOLOGIES

IP SHOWCASE THEATRE AT IBC2019 : 13–17 SEPT 2019 9

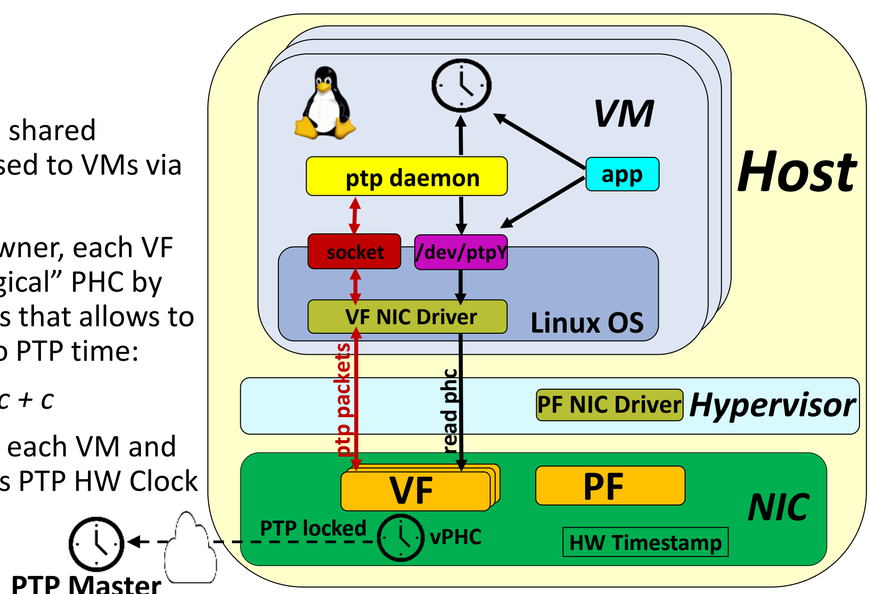
Virtualized Linux PTP HW Clock

How it Works:

- NIC PTP HW Clock (PHC) is shared between VMs and is exposed to VMs via PCIe VFs
- There is no a single PHC owner, each VF NIC driver managing a “logical” PHC by calculating PTP coefficients that allows to transform NIC HW Clock to PTP time:

*ptp_ns = m * hw_clock_cyc + c*

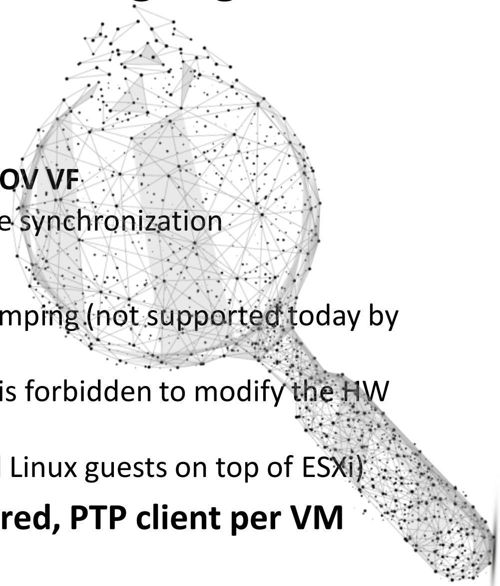
- PTP daemon is running on each VM and synchronizing Linux guest’s PTP HW Clock and system time



Mellanox TECHNOLOGIES

IP SHOWCASE THEATRE AT IBC2019 : 13–17 SEPT 2019 10

Virtualized Linux PHC Highlights



- **Performance: tens of nanoseconds for VMs**
- NIC HW requirements:
  - **SR-IOV support + exposure of HW clock via SR-IOV VF**
  - **PTP packets HW timestamping** to get an accurate synchronization
- SW requirements & restrictions:
  - NIC VF Driver must support PHC and HW timestamping (not supported today by Windows)
  - NIC VF Driver must manage PTP coefficients and is forbidden to modify the HW clock
  - Should be supported for different HVs (we tested Linux guests on top of ESXi)
- **Disadvantage: multiple PTP clients are required, PTP client per VM**

Mellanox TECHNOLOGIES

IP SHOWCASE THEATRE AT IBC2019 : 13–17 SEPT 2019 11

PTP on Windows VM

- Currently Windows doesn't support PTP hardware timestamping
  - Planned for future Windows release
- Due to lack of HW timestamping PTP accuracy that can be achieved today for Windows is limited:
  - For native environment the achievable accuracy is good enough for HD streams but not for UHD
  - For VMs the achievable accuracy is not good enough even for HD streams
- Virtualized PTP solution comes to overcome this, see the next slides…

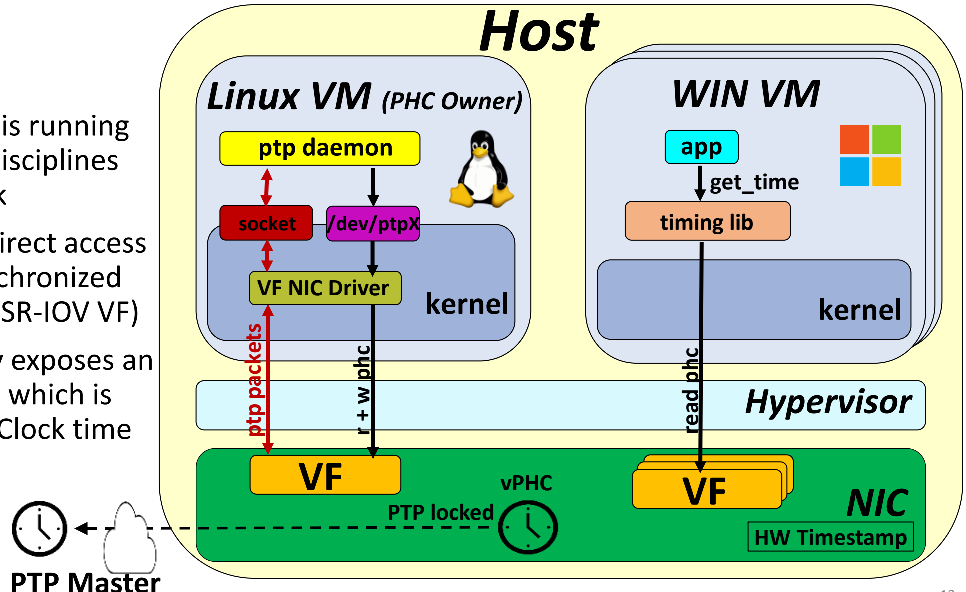
Mellanox TECHNOLOGIES

IP SHOWCASE THEATRE AT IBC2019 : 13–17 SEPT 2019 12

Virtualized PTP

How it Works:

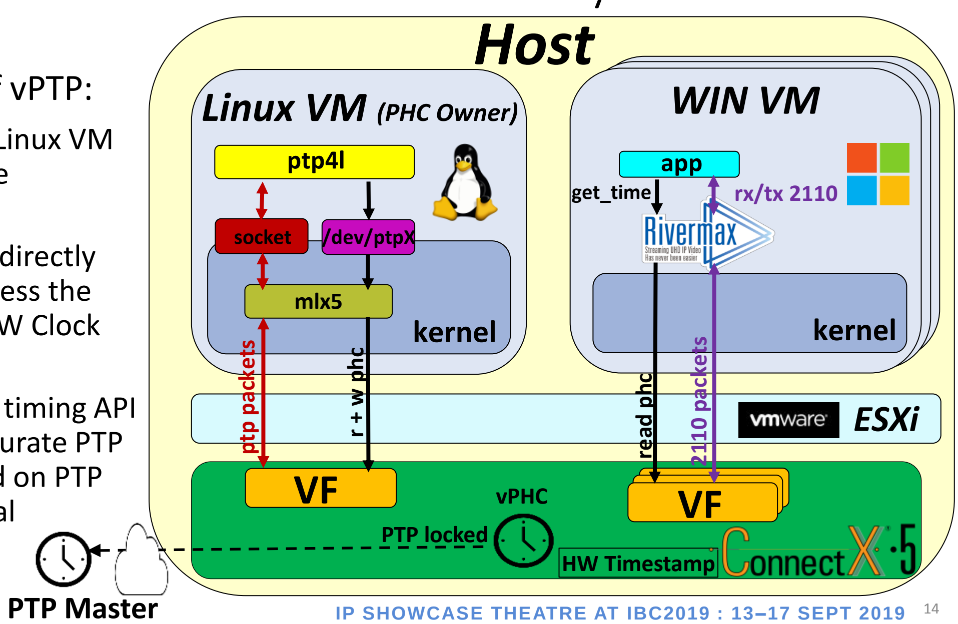
- A single PTP client is running on Linux VM and disciplines the hardware clock
- Other VMs has a direct access (Read Only) to synchronized PTP HW Clock (via SR-IOV VF)
- A dedicated library exposes an accurate PTP time, which is based on PTP HW Clock time signal



Virtualized PTP Case Study

Implementation of vPTP:

- ptp4l is running on Linux VM and "disciplines" the hardware clock
- Windows Rivermax directly from user space access the synchronized PTP HW Clock (via SR-IOV VF)
- Rivermax exposes a timing API and provides an accurate PTP time, which is based on PTP HW Clock time signal



Virtualized PTP Highlights

- Performance (using Rivermax): **tens of nanoseconds for Windows VMs**
- **Single VM (1 core) disciplines the PTP HW Clock**
  - PTP traffic is minimal
- NIC HW requirements:
  - **SR-IOV support + exposure of HW clock via SR-IOV VF to VM**
  - PTP packets HW timestamping to get an accurate synchronization
- SW requirements
  - **A dedicated library is needed to expose PTP HW Clock time**
  - Should be supported for different HVs (we tested Windows guests on top of ESXi)

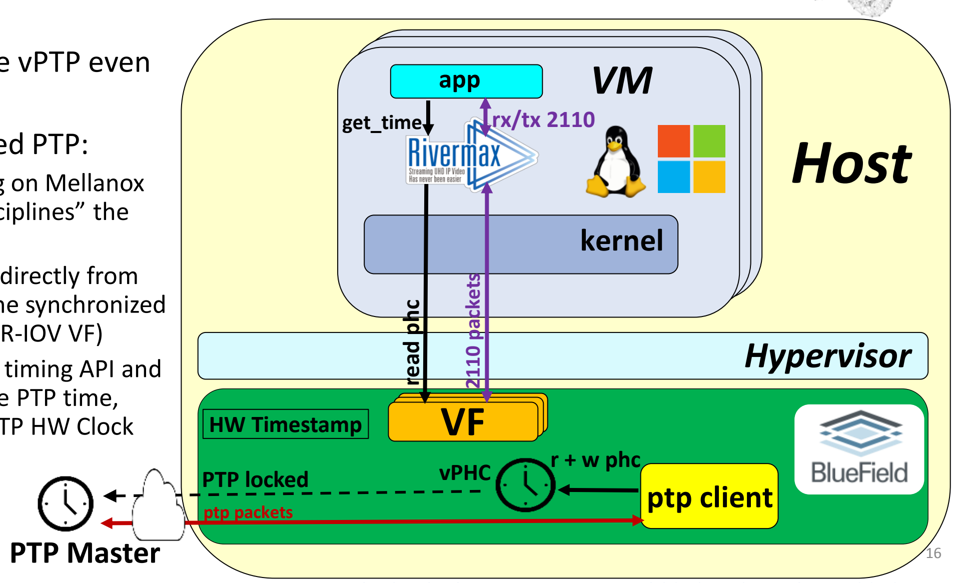
Rivermax SMPTE 2110-21

Mellanox TECHNOLOGIES

IP SHOWCASE THEATRE AT IBC2019 : 13–17 SEPT 2019 15

Looking Forward: Embedded vPTP

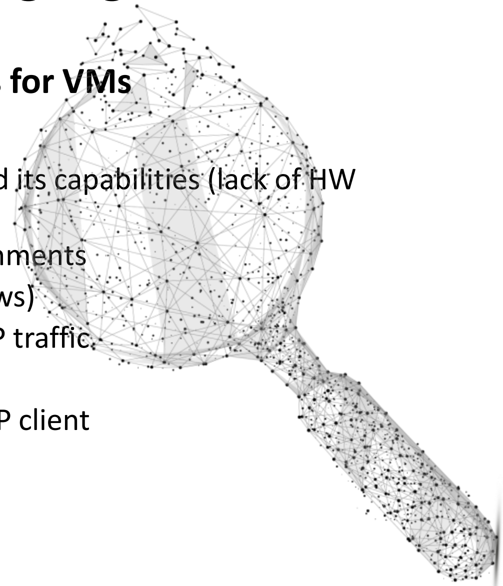
- Idea how to enhance vPTP even more
- Embedded Virtualized PTP:
  - PTP client is running on Mellanox Smart NIC and "disciplines" the hardware clock
  - Windows Rivermax directly from user space access the synchronized PTP HW Clock (via SR-IOV VF)
  - Rivermax exposes a timing API and provides an accurate PTP time, which is based on PTP HW Clock time signal



Mellanox TECHNOLOGIES

IP SHOWCASE THEATRE AT IBC2019 : 13–17 SEPT 2019 16

Embedded vPTP Highlights



- Expected performance: **tens of nanoseconds for VMs**
- Advantages:
  - **PTP performance is agnostic** to CPU load, OS and its capabilities (lack of HW timestamp, scheduling limitations, etc…)
  - Supports both virtualized and bare metal environments
  - **A single PTP daemon for all OSes** (Linux, Windows)
  - Just a single PTP client is running, minimizing PTP traffic.
  - Secured, PTP client is running on trusted domain
  - **Faster, easier integration and deployment** of PTP client

Mellanox TECHNOLOGIES

Thank you

Alex Vainman, Mellanox
alexv@mellanox.com

Come and see us at booth 8.E21!

Thank you to our Media Partners

SVG SPORTS VIDEO GROUP · SVG EUROPE SPORTS VIDEO GROUP · THE BROADCAST BRIDGE Connecting IT to Broadcast · TVB EUROPE Intelligence for the media & entertainment industry

Mellanox TECHNOLOGIES