Power Aware Packet Processing

Chris MacNamara, Dave Hunt

DPDK Summit - San Jose – 2017
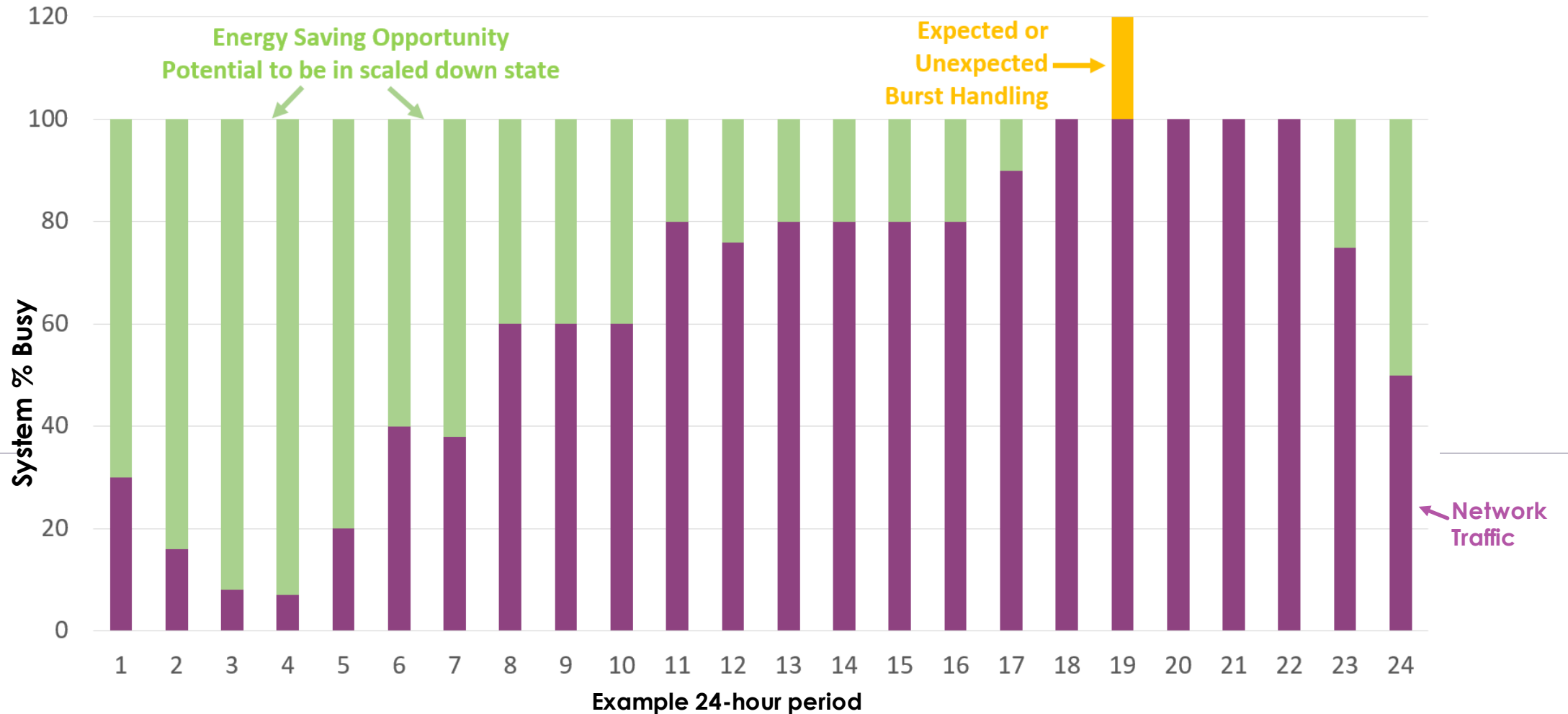
#DPDKSummit

Why We Are Here

- ▶ Drive for data and always on networks

- ▶ Opportunity for green DPDK
    - ▶ Based on continued polling and varying traffic rates

- ▶ Achieve electricity cost saving & increase performance

- ▶ Updates to the power management scheme in 17.11
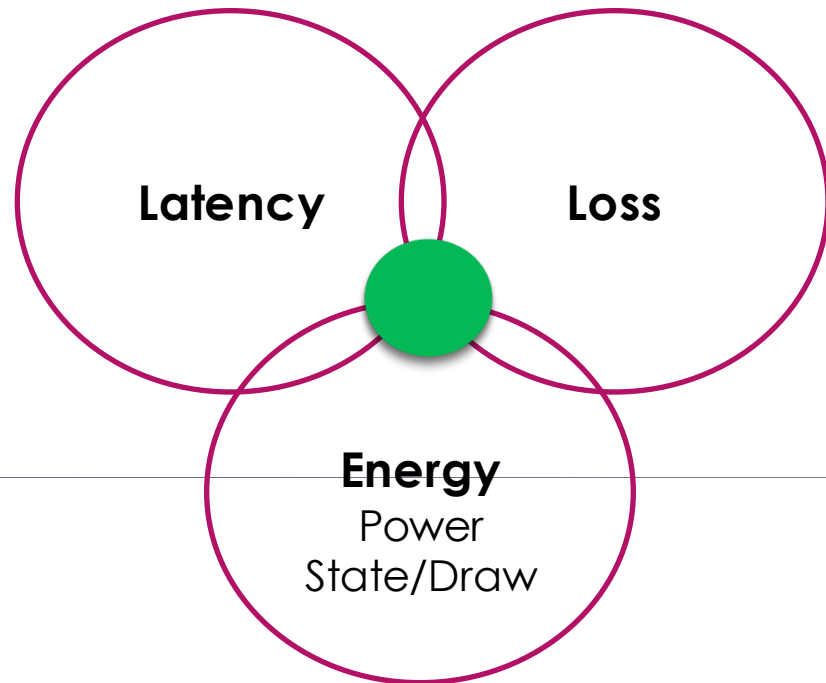
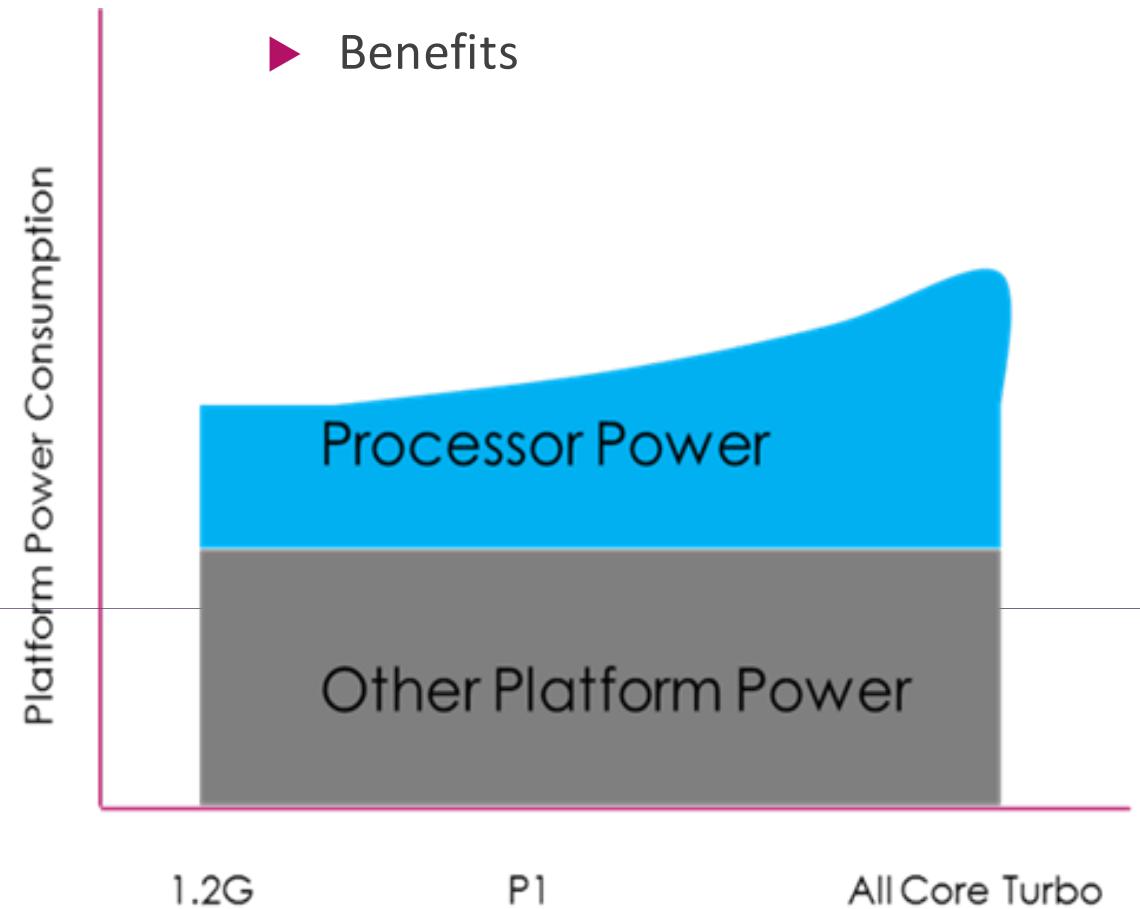# Mapping Power Usage To Network Traffic

DPDK

For Illustration Purposes Only

Energy Saving Opportunity
Potential to be in scaled down state

Expected or Unexpected Burst Handling

Network Traffic

System % Busy

Example 24-hour period

Moving To Green DPDK

DPDK

▶ Evaluation considerations

Latency

Loss

Energy
Power
State/Draw

▶ Benefits

Platform Power Consumption

Processor Power

Other Platform Power

1.2G   P1   All Core Turbo

Matching Platform Energy to Network Load

▶ Out of the box frequency management

  ▶ Freq Up / Freq Down

  ▶ Freq Min / Freq Max

▶ Enable Turbo: Enable Intel® Turbo Boost Technology on the specific lcore

  ▶ Core frequency will go to whatever frequency is allowed for that core based on number of active cores on the packet, thermal limits, etc.

▶ Disable Turbo: Disable Intel® Turbo Boost Technology on the specific lcore

  ▶ Core frequency will return to the maximum non-turbo frequency, if lower freq required, a further library call is required to scale down, go to minimum, etc.
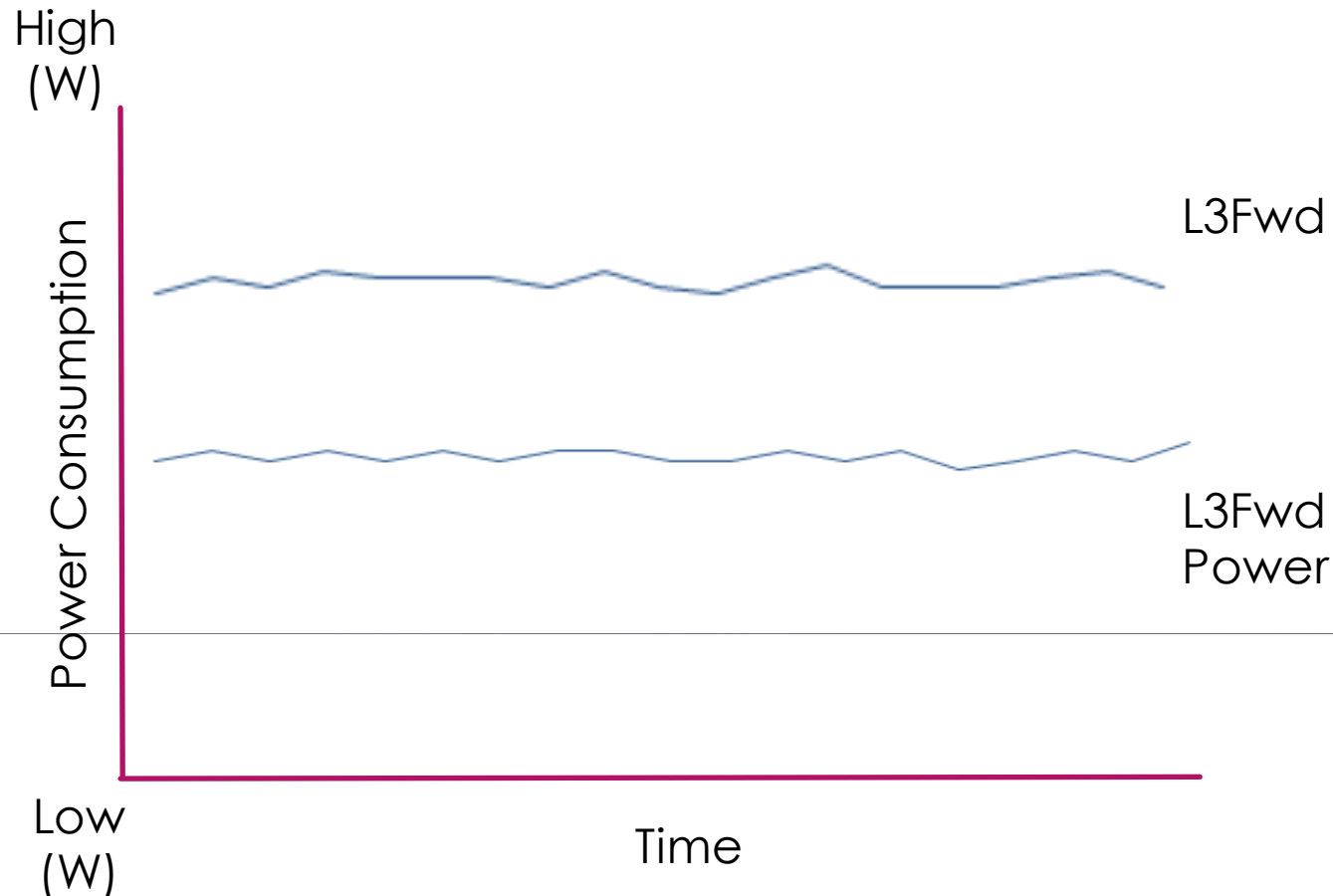
- Scale always on DPDK performance with the network demand
- Common challenges
  - Always On
    - Adjust PMD cores frequency to adjust to packet demand
    - Potential to save power drawn per core using frequency scaling
      - Additional savings from sleeping
- Speed of (re)action
  - Challenge: Fast Scale Up to react to increases in n/w traffic
  - Time = queueing/buffering
- Challenge: fast monitor & reaction time
  - Closer to hardware gives faster reaction time
- Move to policy-based control

(Re)Action

Monitor

Measure
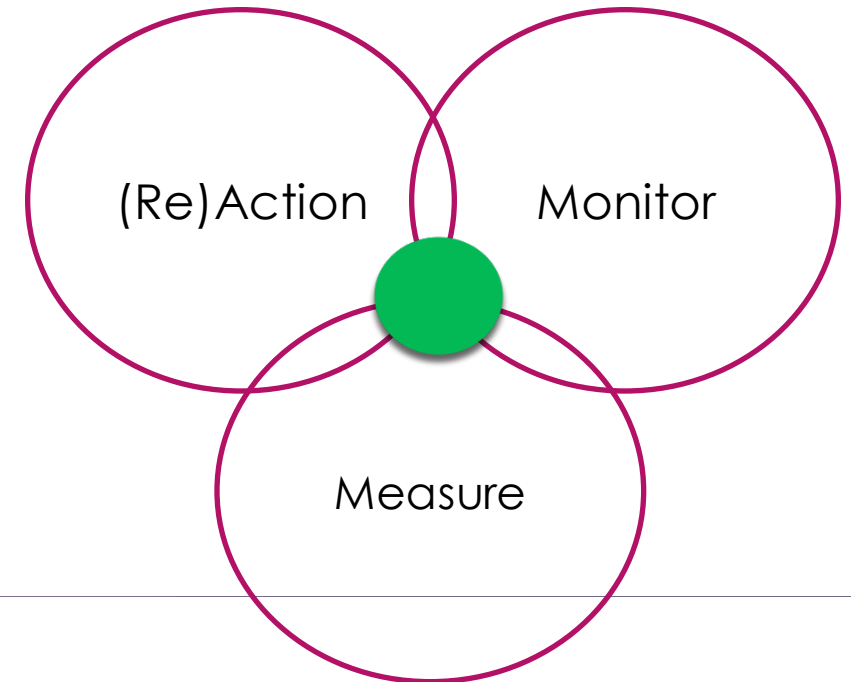
Apply Power Where and When it's needed

Elements Of An Ideal Scheme

▶ A system with a penalty-free reconfiguration capability

▶ DPDK: Be deliberate & control change, only change when you need to change

▶ Fast detection at the micro-burst level

▶ DPDK: Decide on key performance indicators (KPIs)

▶ Mechanism to determine the compute configuration

▶ DPDK: Use DPDK APIs to manage

▶ Power/energy savings with minimum impact to network performance

▶ DPDK: Toolbox allows fine grained control, network load

In-band Policy Control For Power Mgmt

- Patch Set for 17.11
- Power governor on host
- Takes profiles from Guest
- Scale up/down based on:
  - Traffic Rates
  - Time of Day
  - Workload (next)
- **Match compute to network/CPU load**
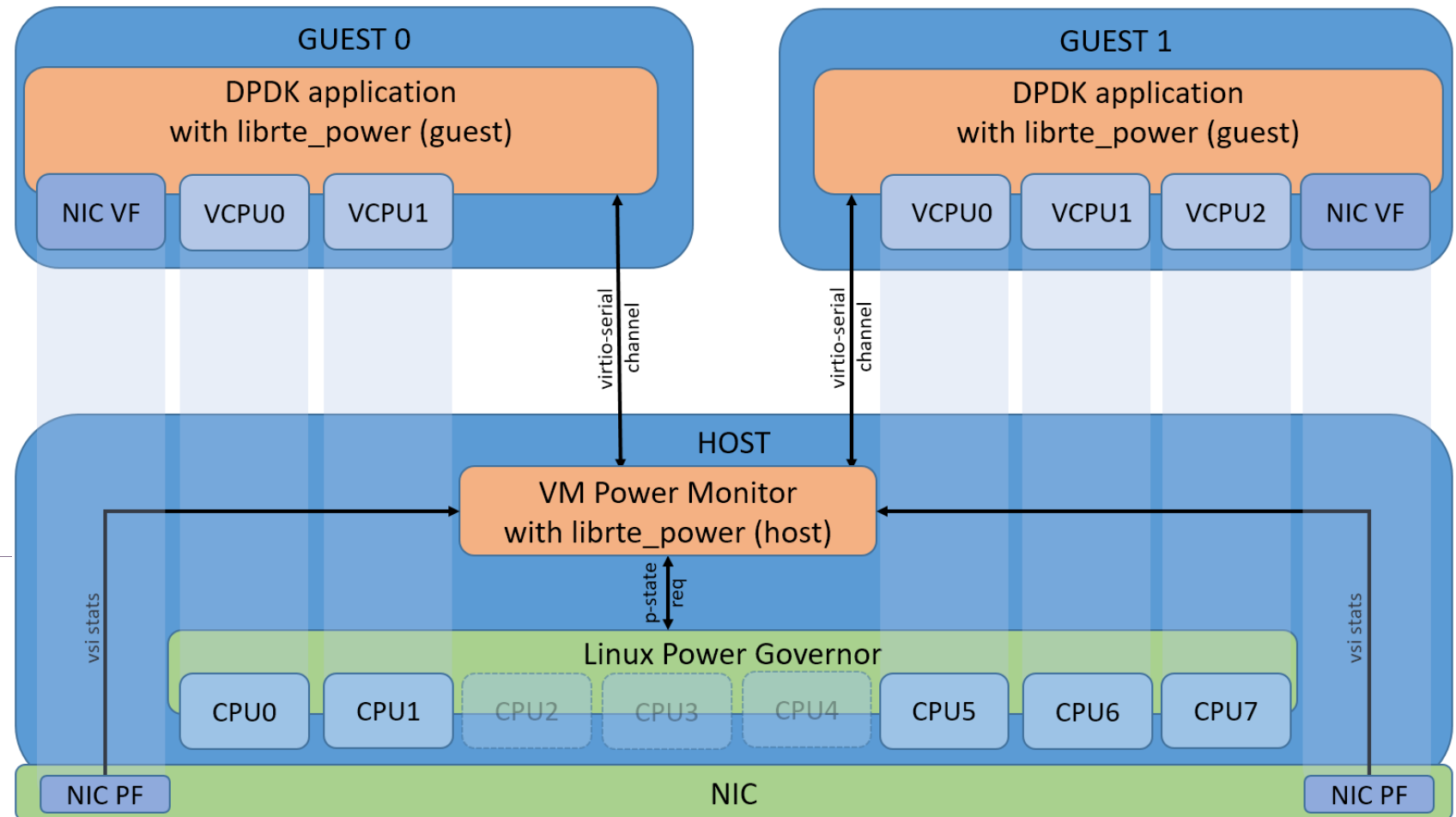
Acknowledgement

- Nemanja Marjanovic

- Rory Sexton

- Konstantin Ananyev

- John Browne

Questions?

Chris MacNamara

chris.macnamara@intel.com

David Hunt

david.hunt@intel.com