Power Aware Packet Processing

Chris MacNamara, Dave Hunt

DPDK Summit Userspace - Dublin- 2017

Legal Notices and Disclaimers

Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Learn more at intel.com, or from the OEM or retailer.

No computer system can be absolutely secure.

Tests document performance of components on a particular test, in specific systems. Differences in hardware, software, or configuration will affect actual performance. Consult other sources of information to evaluate performance as you consider your purchase. For more complete information about performance and benchmark results, visit http://www.intel.com/performance.

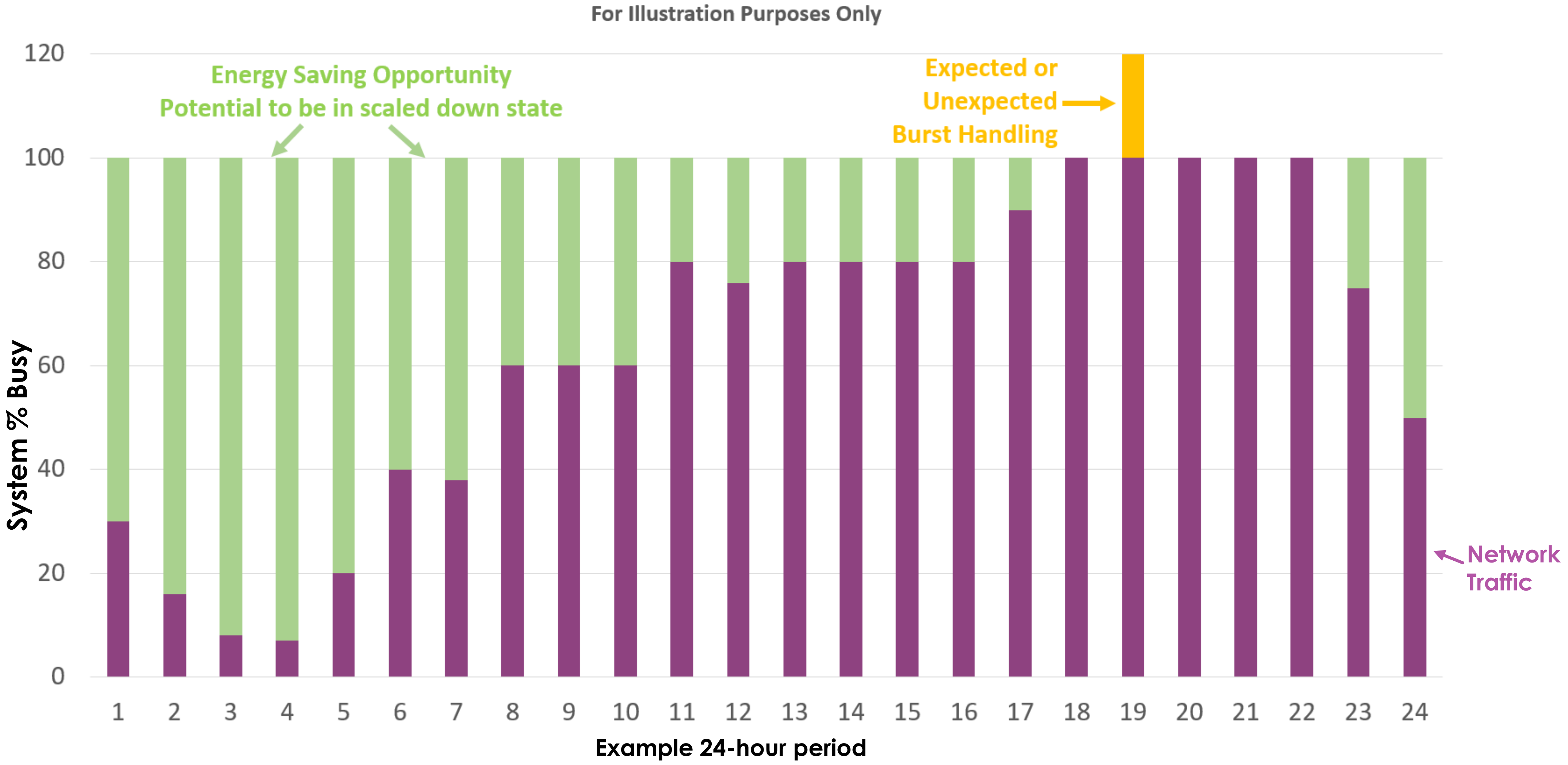
Intel, the Intel logo and others are trademarks of Intel Corporation in the U.S. and/or other countries. *Other names and brands may be claimed as the property of others.

© 2017 Intel Corporation.

Why we are here

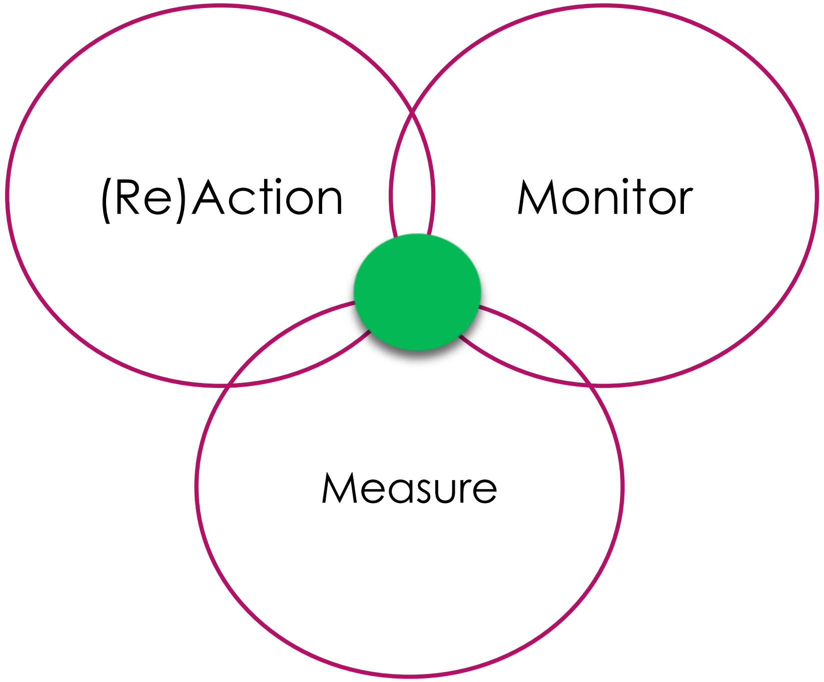
- Drive for data and always on networks
- Opportunity for Green DPDK
  - Based on continued polling and varying traffic rates
- Achieve electricity cost saving & Increase Performance
- Proposed Updates to the existing Power Management scheme

Mapping Power Usage to Network Traffic



Meeting the needs of an on demand network

- 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
      - ++ 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



Apply Power Where and When it’s needed

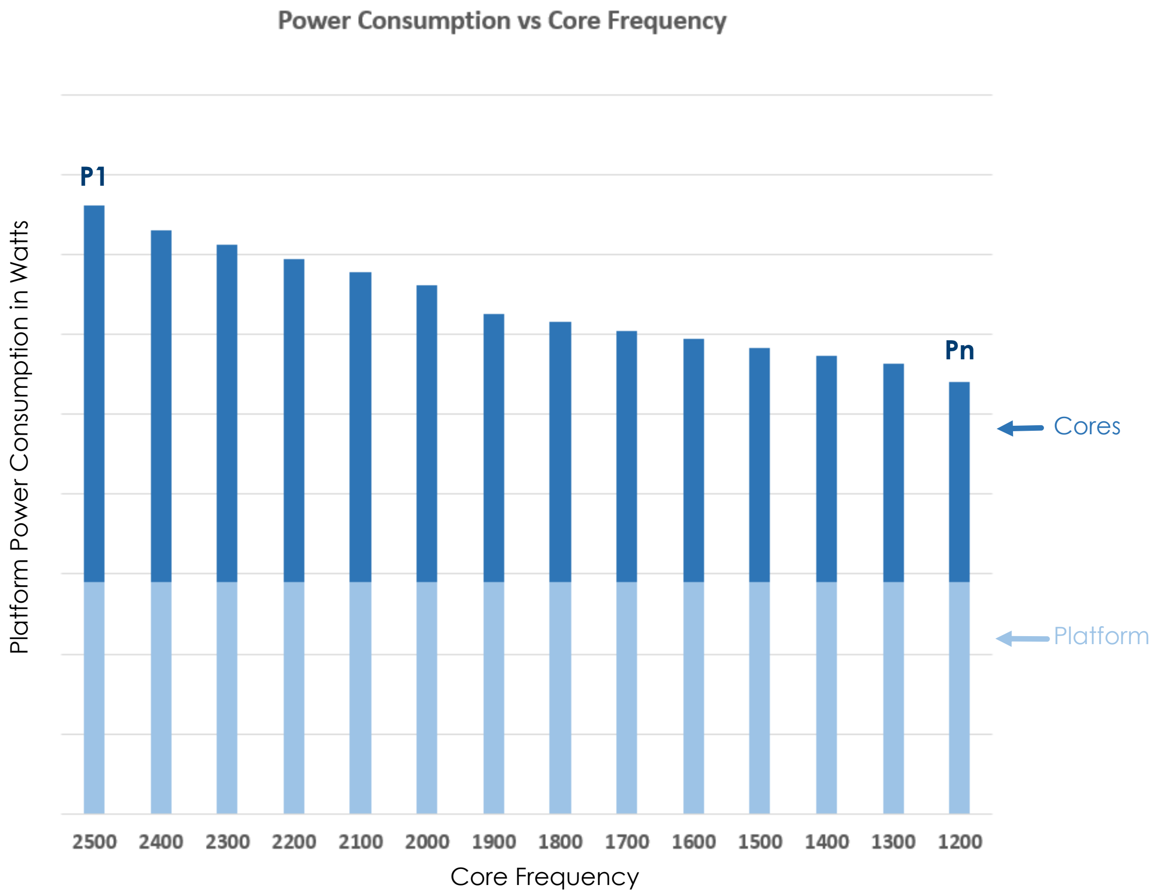
Design Considerations

DPDK

- Ability to scale up quickly e.g. burst detection
- Dimension queue depth and monitoring for worst case or maximum latency
- Dimension receive descriptors in the DPDK poll mode driver, to handle bursts
- Example: 128 default size/depth holds ~8us
- Depth of queues related to energy efficient & performance technology state changes
  - Identifying modules with shallow queueing, to avoid packet loss

Opportunity for Energy Efficiency & Performance

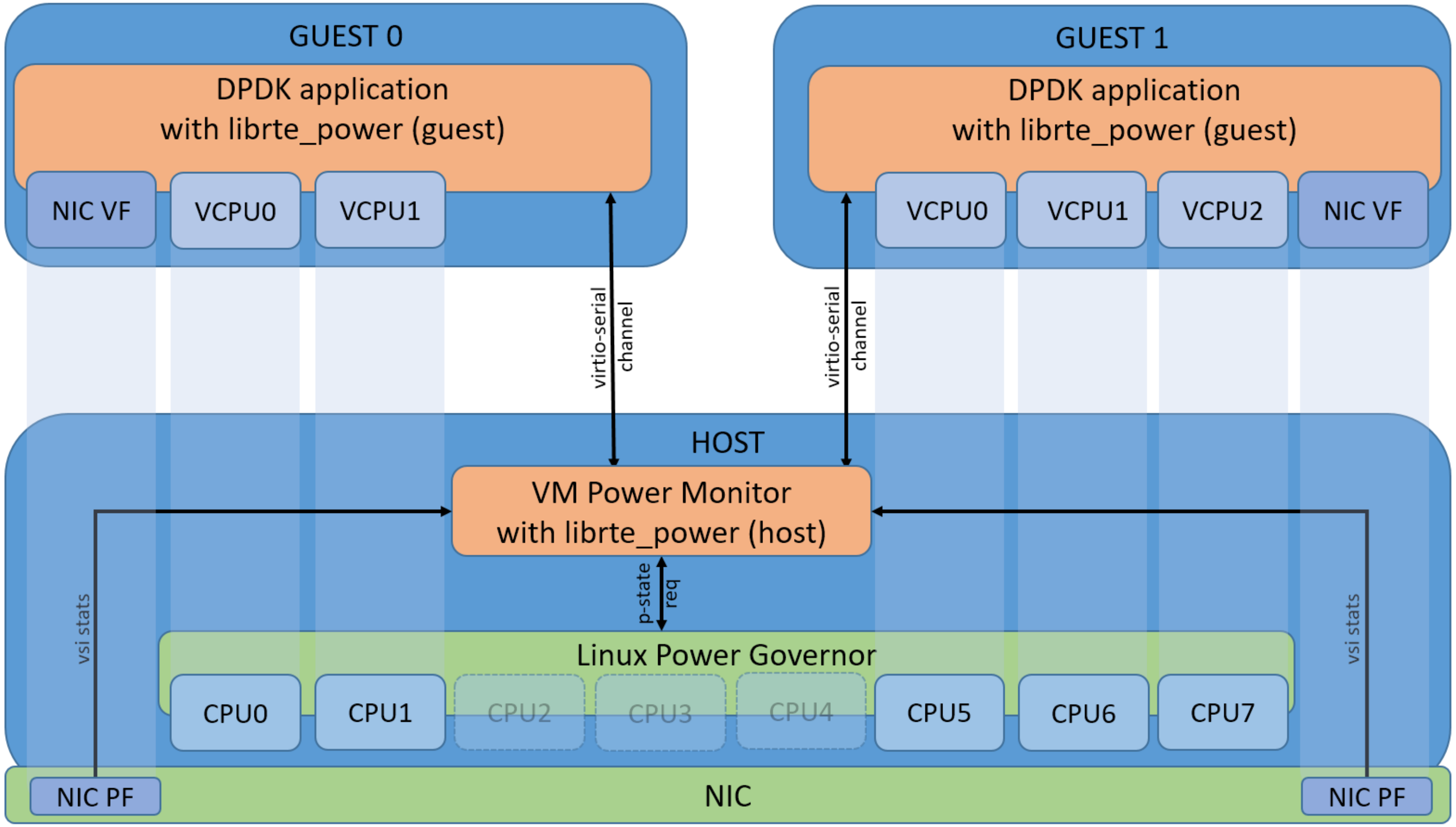
- Expect energy headroom to be available in most scenarios
- Varying frequency can save energy
- For example Intel® Xeon™ Scalable Processor 8180 frequency variation per core allowed from 1200MHz to 2500MHz (without Turbo)
- Impactful power saving on 1200MHz vs 2500MHz
- Further savings possible with additional technology tuning
  - For example, varying core activity / sleep states



Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more complete information visit http://www.intel.com/performance.

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**



Additional DPDK Performance APIs

- 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.

Acknowledgement

- Nemanja Marjanovic
- Rory Sexton
- Konstantin Ananyev

Questions?

Chris MacNamara

chris.macnamara@intel.com

David Hunt

david.hunt@intel.com