Isolation & QoS In PaaS Cloud CLSF Open Discussion

Oliver Yang

http://oliveryang.net

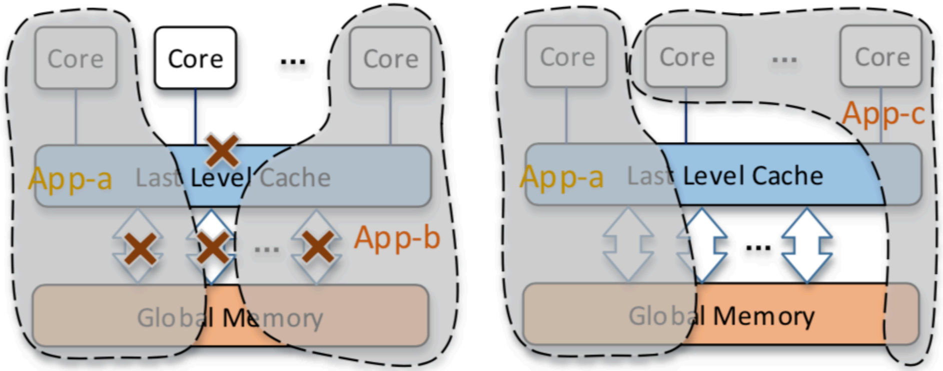
2018 10.11

Background

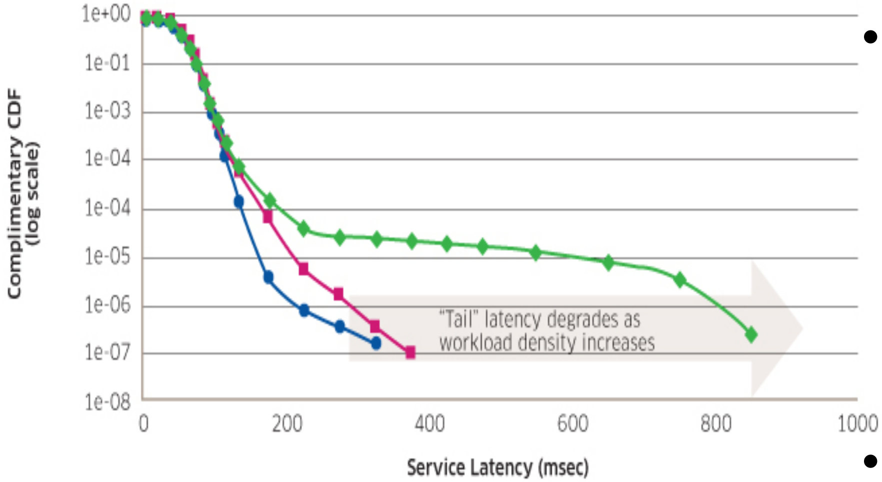
- Machine
  - Virtualization (Xen, KVM, Container)
  - Service consolidation
- Application
  - Docker
  - Colocation

Colocation

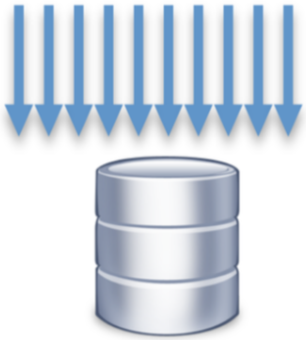
- Increase data center hardware utilization
- Consolidate multiple workloads with acceptable QoS loss
  - Borg/Kubernetes, VMware DRS etc.
- Challenge: Avoid contentions among multiple workloads
  - Workload characterizations
    - Resource model: Compute bound + IO bound
    - Time sharing: Peak time + Off-peak time
  - OS resource control & isolation features



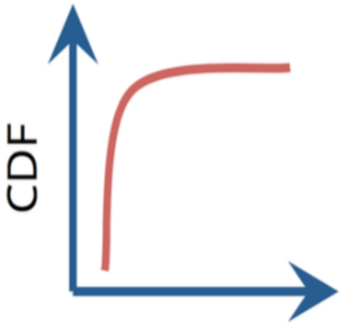
Performance Jitters & Causes



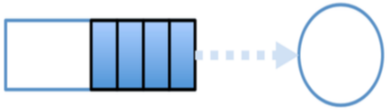
- System performance jitters RCA (Root Cause Analysis) can be very challenge…
  - Burst randomly
  - In a very short period
  - Reported/found very late
  - Always postmortem analysis
  - Without enough debug data
- Tradeoff between utilization and QoS
  - Avoid jitters by resource isolations



HW Resource Contentions



Skewed Access Pattern

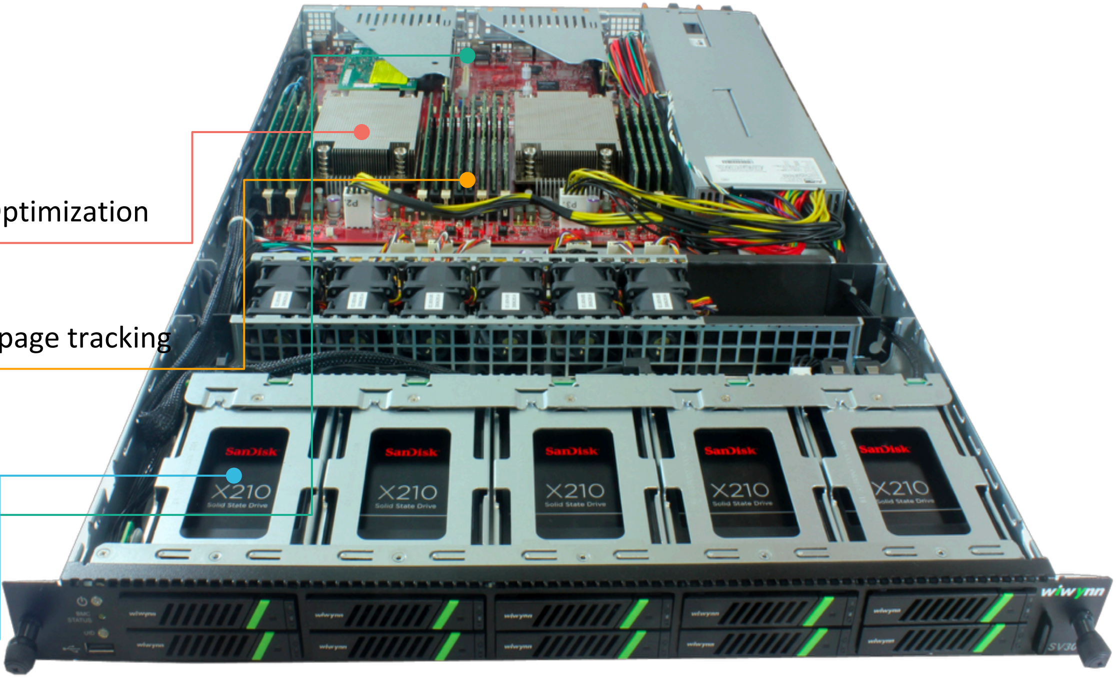


Queuing Delays

Background Activities

System wide optimizations

- Hardware & Software cooperate together
- Full stack optimizations



CPU: CAT, Cgroup, Noise Clean, Scheduler Optimization

Memory: MBM, MBA, Cgroup, Idle page tracking

Network: TC, HTB, Lockless Qdisc

IO: Write back Cgroup, Low/Upper limit，
Weight, File Level-Based IO Throttle

Challenges from web-scale data center

- Scale-out optimization at large scale
  - Resource QoS classifications & scheduling
  - Workload placements(Bin-packing) algorithm
  - Dynamic distributed load balance & scheduling
- Constantly varying load challenges
  - The synthetic test environment has poor coverage
  - Many issues can be only found on real online environment
  - Using statistics to determine the impact of a change
  - Performance or resource characters could be totally changed after a time period
  - Paper: Performance Analysis of Cloud Applications – NSDI 2018

Application-centric Architecture

- 4 key aspects need to be re-considered…
  - Security Isolation
  - Fault Isolation
  - Performance Isolation
  - Resource Utilization

Security Isolation

- Virtualization spectrum: Kata, Gvisor, LinuxD, Cgroup
    - Quantitative security measurement?
    - Performance overhead？
    - What are the key performance metrics?
        - Boot time
        - Run time
            - Compute bound?
            - IO bound?

Fault Isolation

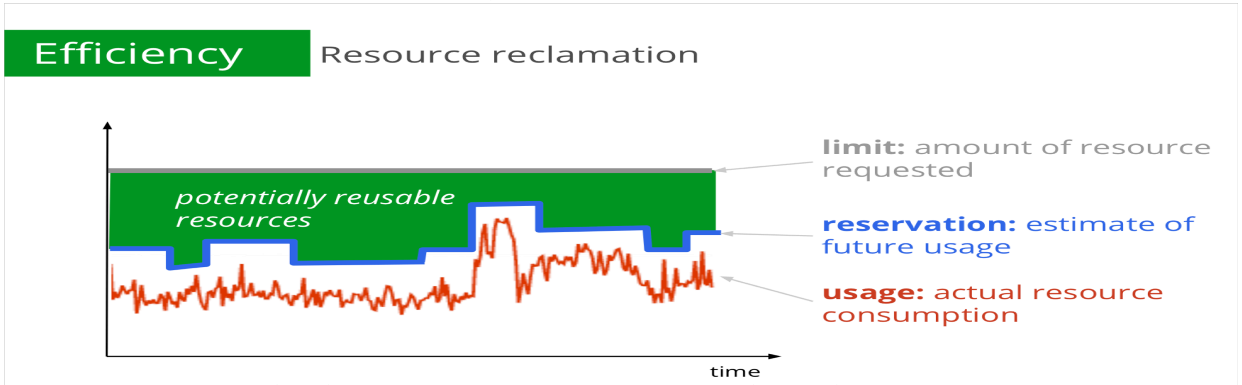
- Hardware Isolation
  - CE, UE, TDP issues
  - Device retire: offline and fence
  - Disk Hotplug?
- Software Isolation
  - Split-lock usage
  - OOM Handling
  - Segfault from application
  - Priority Reversion in FS
- Error Resilience
  - Checkpoint/Restore
    - Live Migration
    - CRIU
  - Backup schedule
    - VM/Container/Job reschedule

Performance Isolation

- Hardware Resource Contention
  - Intel RDT, SRIOV, NVME WRR
  - Flexibility (Static vs. Dynamic)
  - Fine granularity
- Lack of latency bound support
  - Share, Weight, Quota caused latency problems
  - Scheduling: preempt latency
  - Memory: slow path
  - IO: sync io

Resource Utilization

- Deployment density & scalability
  - Partitioning vs. Sharing
  - Small resource footprint
- QoS Classes & SLA
  - Kubernetes: Guaranteed, Burstable, Besteffort
- Resource reclamation & QoS



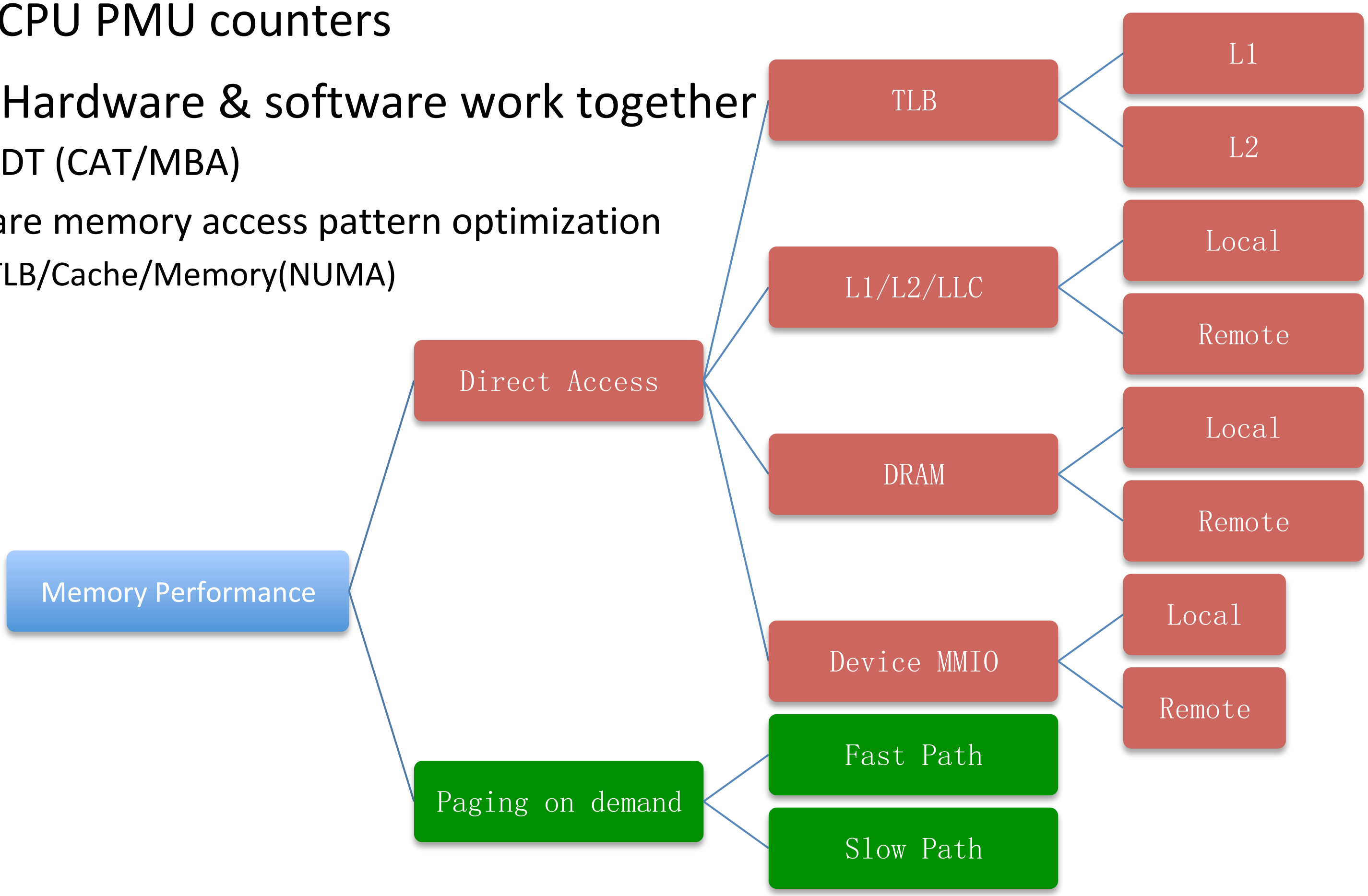
Paper: Large-scale cluster management at Google with Borg - EuroSys 2015

References

- System memory QoS Challenges
- Colocation in Alibaba Private Cloud

Performance Isolation - Direct Memory Access

- Problem: Hardware resource contention
- Analysis: CPU PMU counters
- Solution: Hardware & software work together
  - Intel RDT (CAT/MBA)
  - Software memory access pattern optimization
    - TLB/Cache/Memory(NUMA)



Perf Metrics For Direct Memory Access

CPI/IPC/MIPS

- Perf stat - IPC
- Pmu-tools – toplev, ocperf

TLB Hit/Miss/ Flush

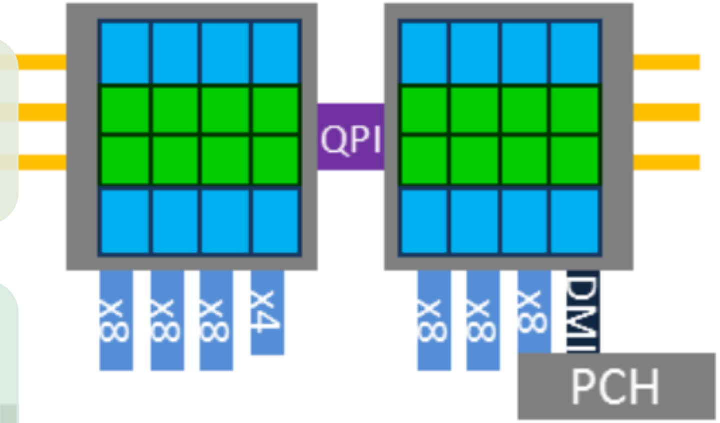
- Perf stat – tlb miss, flush events
- perf mem – access ratio

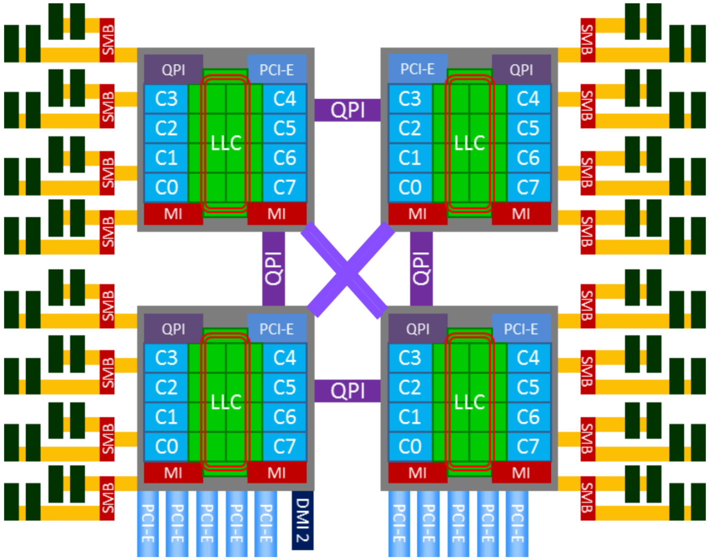
Cache Hit/Miss

- Perf stat – cache miss events
- Perf mem – access ratio
- Perf c2c – access ratio, latency

Memory Latency/BPS/ IOPS

- Resctrl - CQM and MBM
- Numatop - Local/Rmote
- Perf c2c – access ratio, latency

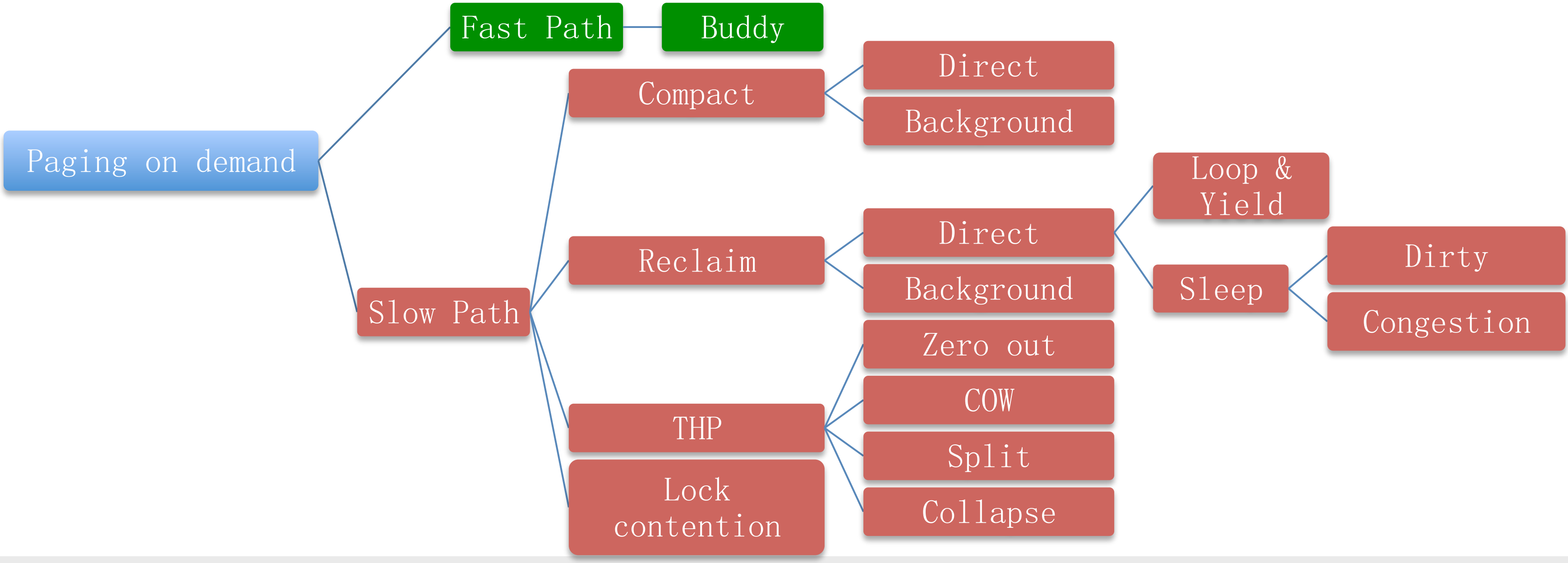




Notes: Some metrics have performance overheads

Performance Isolation - On Demand Paging

- Problem: Overheads from kernel paging mechanism
- Analysis: Kernel counters and dynamic tracing
- Solutions: Avoid running into slow path
  - Warm up and lock the pages
  - Increase min_free_kbytes vs. memory utilization
  - Google patch: per-memcg kswapd (incomplete & need more enhancements)


Paging on demand
Fast Path
Buddy
Slow Path
Compact
Direct
Background
Reclaim
Direct
Background
Loop & Yield
Sleep
Dirty
Congestion
THP
Zero out
COW
Split
Collapse
Lock contention