YAHOO!

Performance Engineering Group

Rack Scale Computing

Hugo Gunnarsen | IDF 2015 - San Francisco

Bio – Hugo Eide Gunnarsen

Hugo Gunnarsen has been leading Yahoo's server hardware strategy for the last 10+ years. He is the director of Yahoo's Performance Engineering Team, which is in charge of implementing new technologies, model total cost of ownership, and creating the server offerings available to Yahoo's engineers. His daily work consists of working with Yahoo's business units and optimizing software and hardware to improve latency, throughput, and cost. Prior to joining Yahoo he was in charge of designing and operating data centers for Fast Search & Transfer that hosted the search engine AllTheWeb (the largest web index at the time). He also worked for Silicon Graphics, specializing in High Performance Computing and Visualization markets. Hugo studied computer science at the Norwegian Institute of Technology (NTNU).

Overview

- Yahoo / Performance Engineering Group
  - Hardware strategy
  - Application optimization
  - TCO optimization

- Today's challenges
  - Time to market - fast turnaround on HW deployment
  - Dedicated systems leaves some resources underutilized
  - Rapid development cycles results in outgrowing initial environment

- How can Intel® Rack Scale Architecture help addressing these challenges?

Intel® Rack Scale Architecture - Value add

- Incorporate hyper scale best practices to improve data center efficiencies.

- Incorporate pooling of hardware to achieve savings in shared infrastructure.

- Incorporate pooling of storage to achieve savings via shared boot.

- Future concepts like pooled NVMe provide functionalities for higher performance Hadoop workloads.

YAHOO!

Intel® Rack Scale Architecture

Incorporate hyper scale best practices to improve data center efficiencies:

- Estimated 5-15% reduction in HW cost due to shared power and cooling.

- Improved redundancy, and decreased overhead through sharing and large scale.

- Power shelf accepts multiple voltage inputs, eliminates AC/DC transformers at server level.

- Seamless upgrades of storage and compute.

- Improved cooling design by disaggregation.

- Allows for optimization through CPU technology selection.

YAHOO!

Intel® Rack Scale Architecture

Incorporate pooling of hardware to achieve savings in shared infrastructure:

- Traditional rack servers non-compute infrastructure can be as much as 25% of the total cost (chassis, power supply, fans, boot drive, NIC, BMC, etc.)

- Cable reduction and simplification through aggregation of ports in pooled systems.

- Dynamic hardware allocation allow for greater flexibility and utilization of available resources by matching workload requirements.

- Gain similar benefits traditionally found when virtualizing and combining workloads.

- Common management framework provide wider ecosystem choice.

- Future proofing of management interface – Intel Rack Scale Architecture APIs are open standards.

- Support of OpenStack allows for seamless integration into existing environments.

Intel® Rack Scale Architecture

Incorporate pooling of storage to achieve savings via shared boot:

- As much as 42% of the available drive space goes unused.

- For NVMe as much as 90% of throughput is not being utilized.

- Most cases it's the lower latency (10's of μs) that's sought after.

- Disaggregation of compute and storage allow for deployment flexibility:
  - Moving workload to another node require just a reboot.
  - No more locked up storage resources on individual nodes.
  - Allows for multiple dormant systems to share single compute node.

- Extending memory through use of NVMe as swap or shared memory.

- Opens up for more fine grained checkpoint-restart functionality.

Intel® Rack Scale Architecture

Future concepts like pooled NVMe provide functionalities for higher performance Hadoop/Hbase workloads:

- Using shared NVMe for temporary files.

- Avoid overprovision of fast storage through sharing.

- Ability to bypass network for local rack data exchange.

- Move temporary data to fast/low latency storage instead of main memory.

- Simplifies data sharing for scale-out applications.

- Yahoo looking at areas where Intel® Rack Scale Architecture will have significant impact to existing applications. The introduction of pooled NVMe allow for optimizations with minimal SW modifications.

YAHOO!