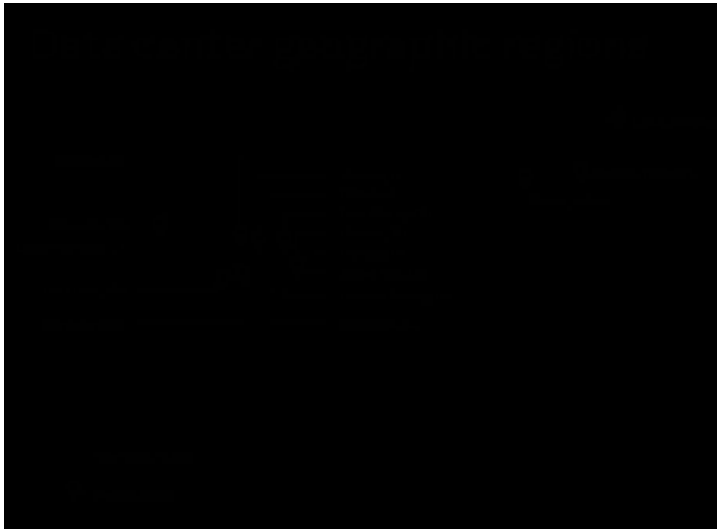
DCPerf: An Open-Source, Battle-Tested Performance Benchmark Suite for Datacenter Workloads

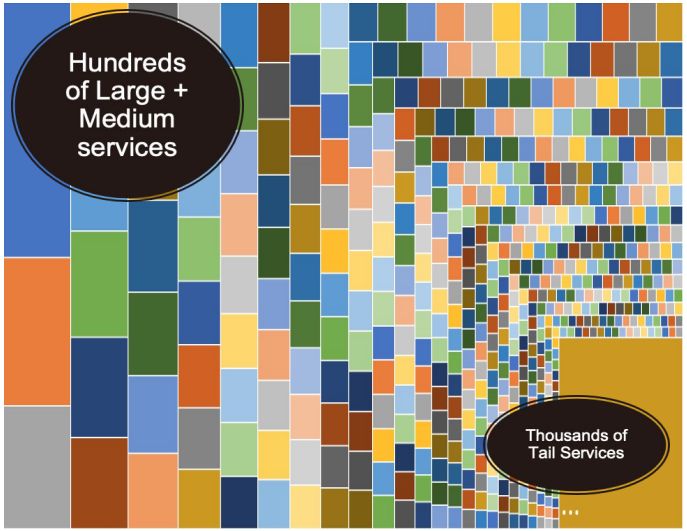
Wei Su, **Abhishek Dhanotia**, Carlos Torres, Jayneel Gandhi, Neha Gholkar, Shobhit Kanaujia, Maxim Naumov, Kalyan Subramanian, Valentin Andrei, Yifan Yuan, Chunqiang Tang

Meta

We work on hyperscale, fast-growing fleet



- 30+ regions with millions of servers
- Billions of users

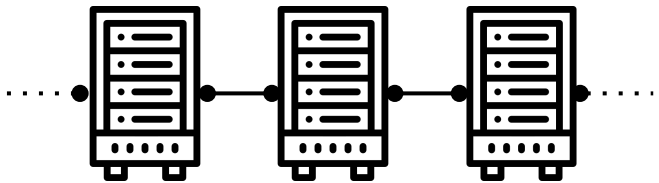


- Multiple server types - Compute, Storage & AI
- 1000s of workloads
- Service oriented architecture

Every compute roadmap decision needs to be carefully evaluated

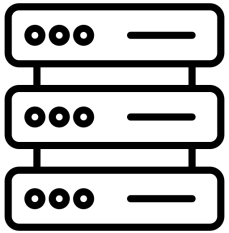
Perf evaluation of new HW in production is challenging

What a production service requires



- Racks in production datacenters, 100s of servers
- Meta-managed secure environments for live traffic
- Proprietary tools and software stack

What's available at early-stage…



- Simulators, emulators, a few engineering samples
- Reference boards, generic OS and software stack
- No access to production code and data

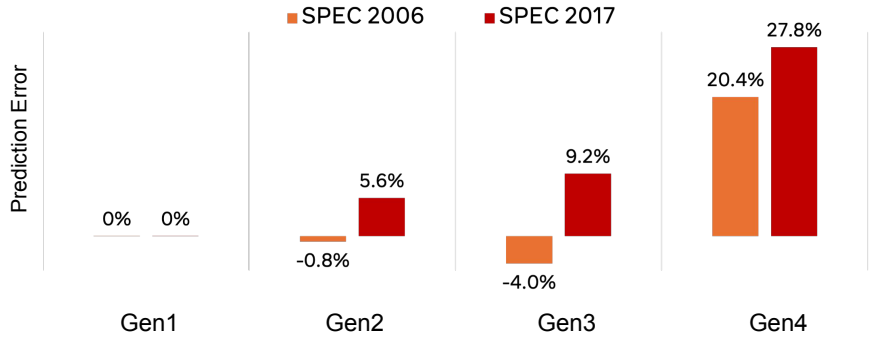
Highly representative benchmarks are a must!

Standardized benchmarks are not sufficient

- Accuracy is the primary concern

Further, they do not

- Support scalable application architecture
- Use diverse software libraries
- Reflect system arch and uArch behaviors
- Evolve over time



Prediction error in gen-over-gen performance between “SPEC vs. Prod”

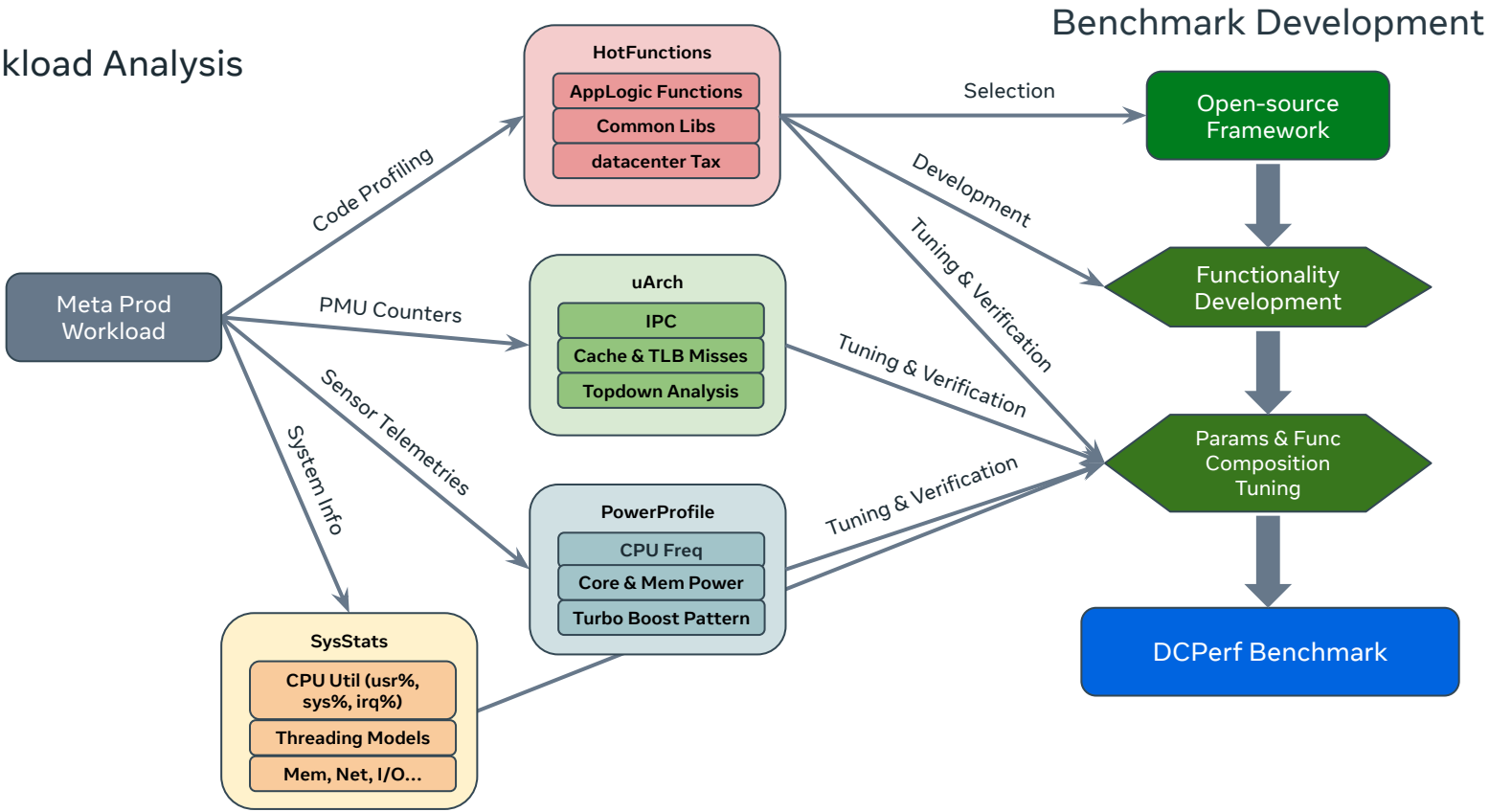
DCPerf is designed to address these problems!

DCPerf benchmark suite

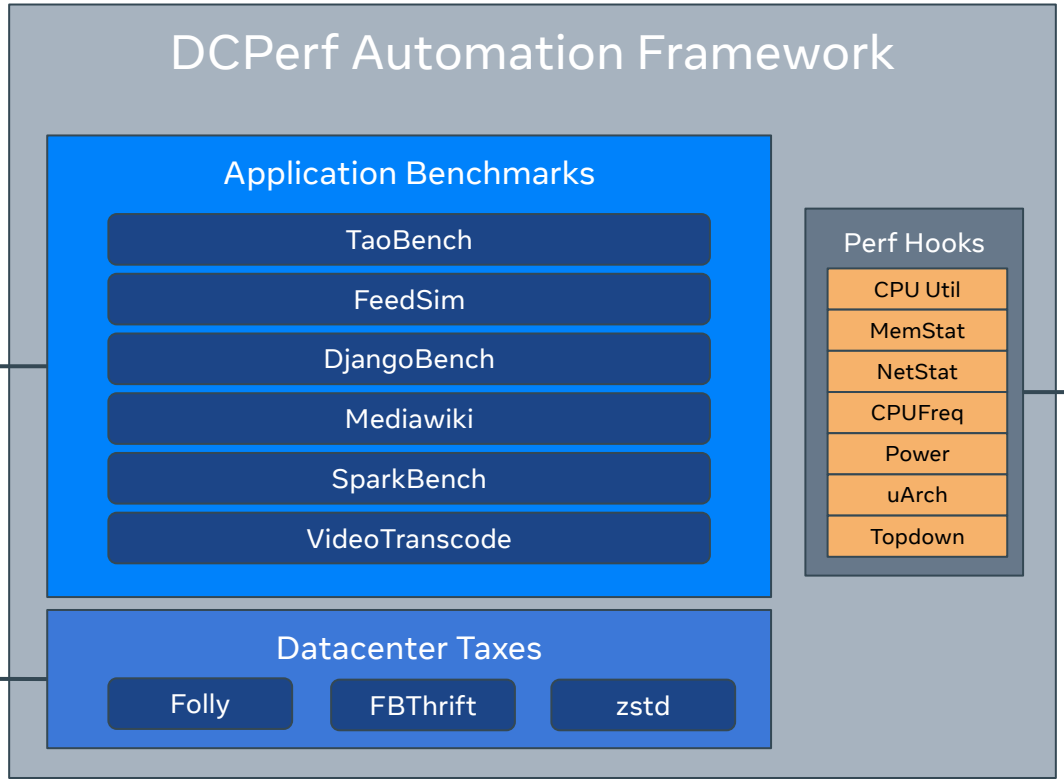
- Designed to representing major hyperscale workloads at Meta datacenters
  - Leverage open-source software for ease of use outside Meta
  - Modular, parameterized and scalable - similar to distributed hyperscalar applications
  - Intended to be run on modern processors with high core counts, latest OS and kernels.

- Designed for
  - Hardware performance evaluations
  - Early-stage software optimizations and hardware pathfinding
  - Hardware-software co-design
  - Open source to enable collaboration with industry and academic partners

DCPerf design process



DCPerf components



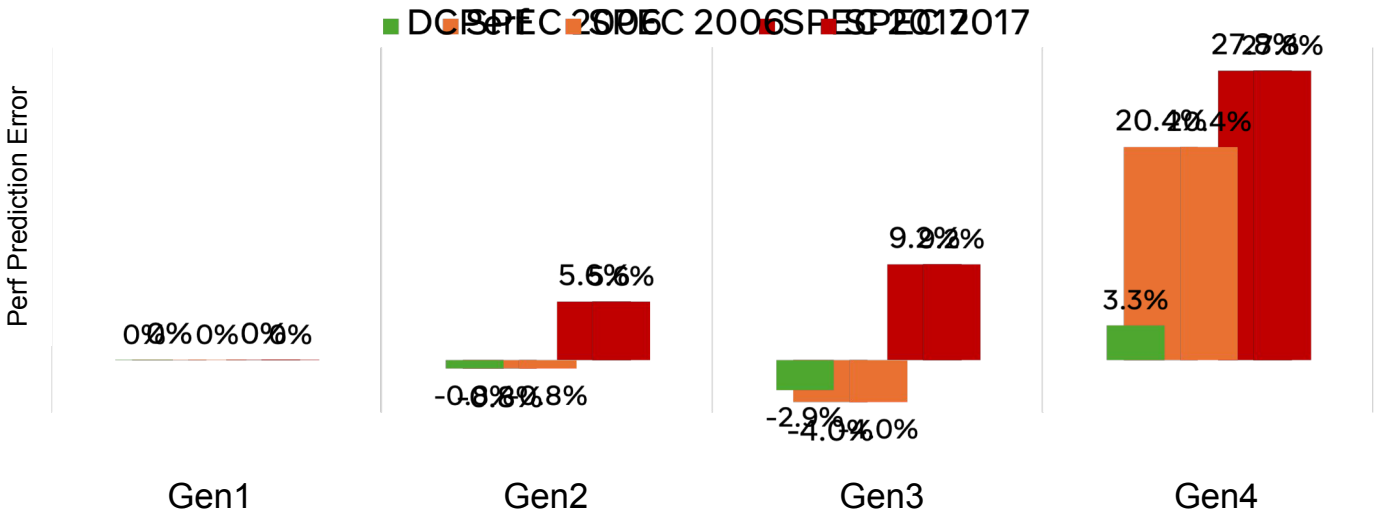
Wide coverage of major Meta workloads: Web, Ranking, caching, big data, video processing

Key libs of common routines/operations

Monitors important system & perf telemetries during benchmark runs

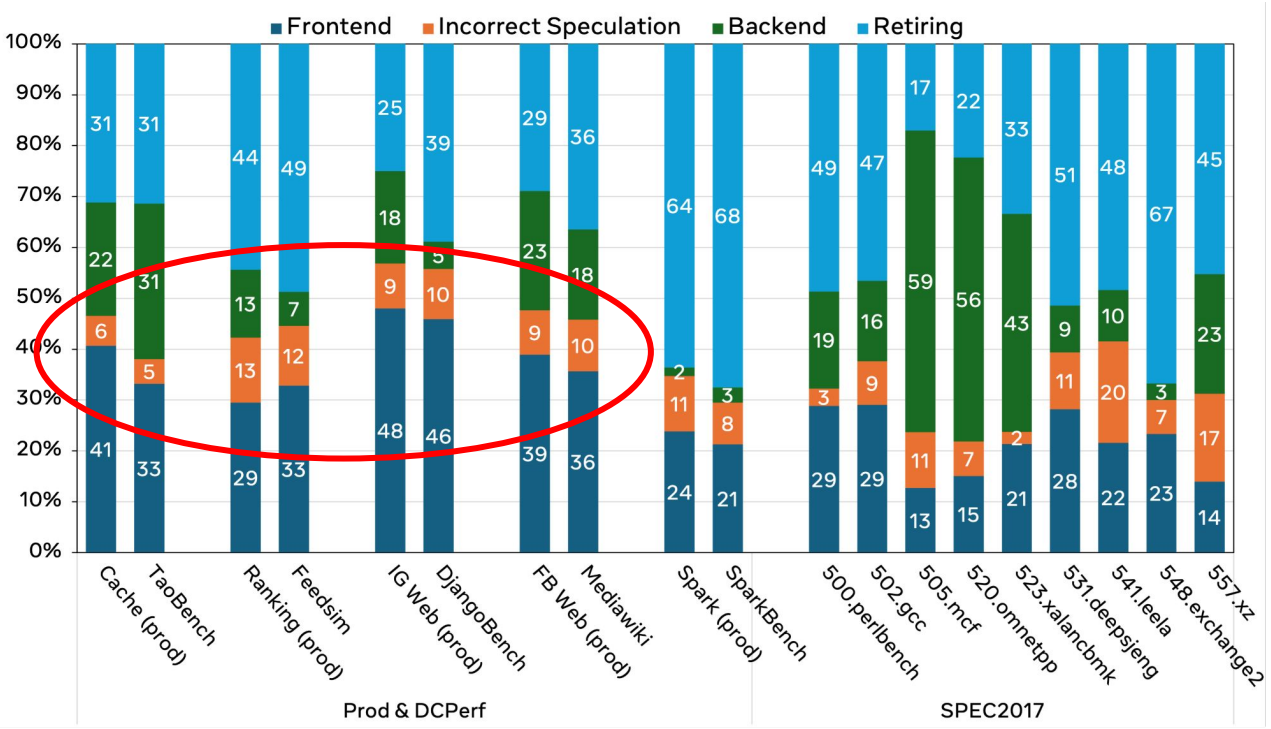
DCPerf supports x86/ARM, CentOS/Ubuntu and latest Linux versions

Validating DCPerf: Predicting workload performance



DCPerf can project Prod Services' performance more accurately

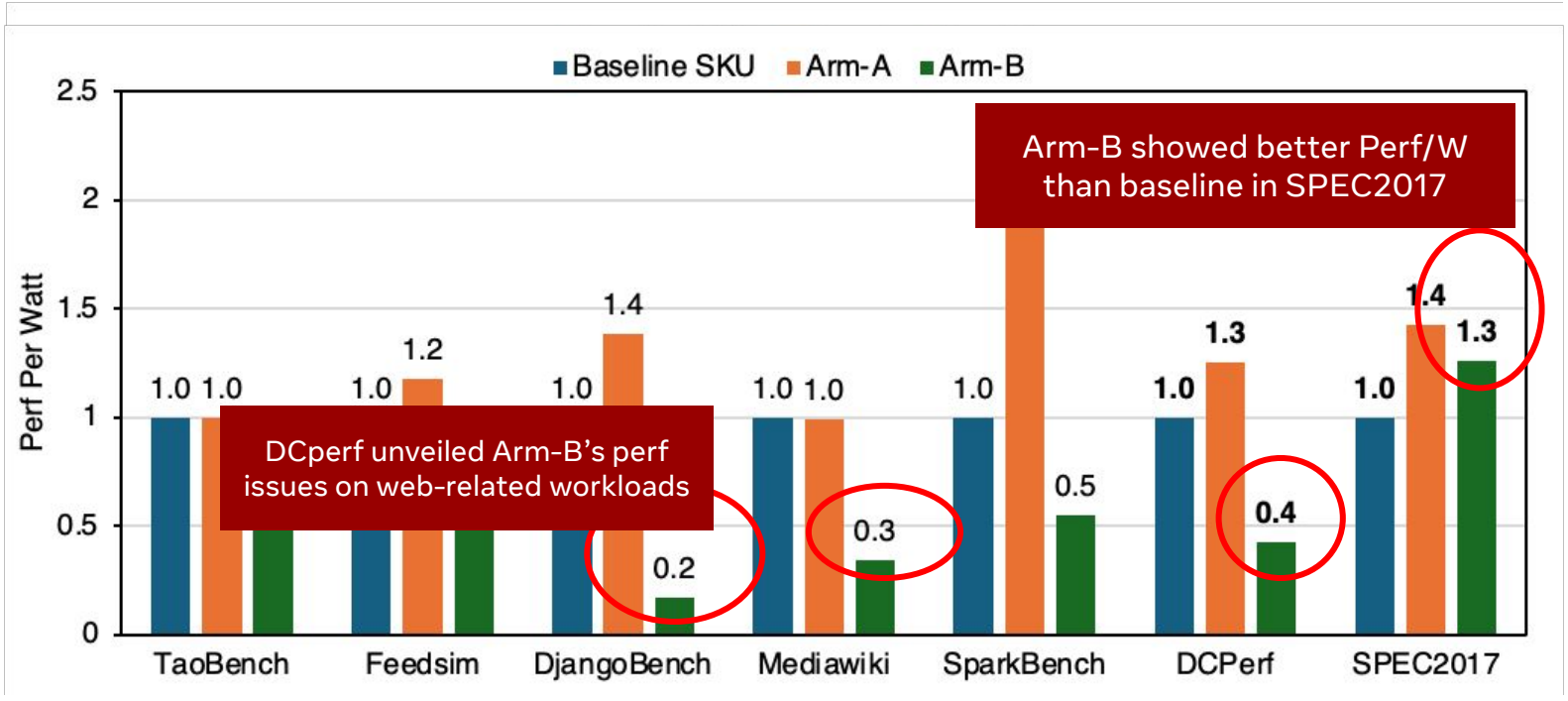
Validating DCPerf: Representing topdown metrics



More metrics (uArch, hot functions, power etc) in the paper!

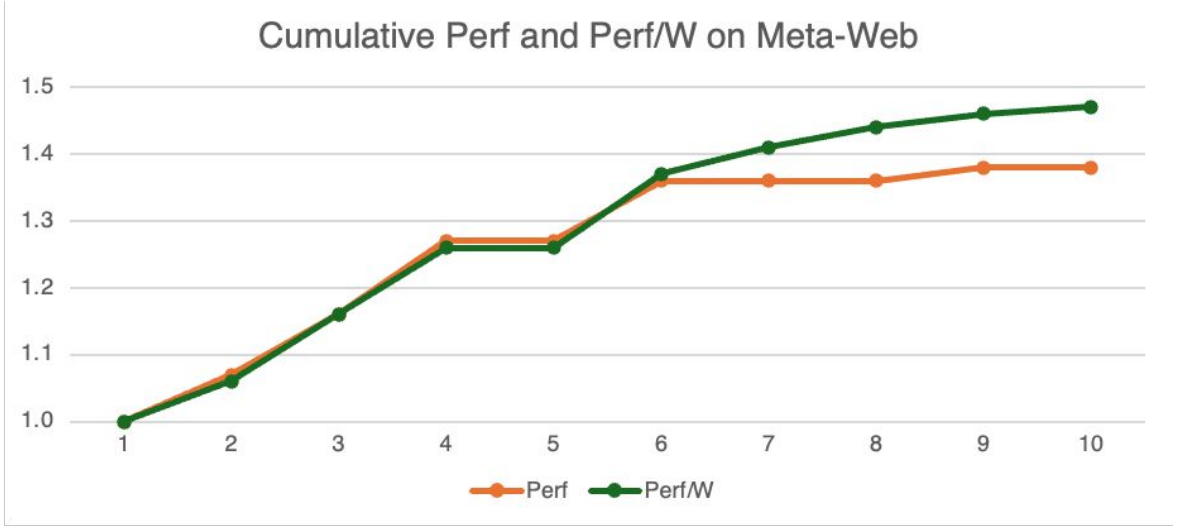
Case 1: ARM CPU decision

- First-gen Arm-based server at Meta, we needed to select between two options.
- DCPerf data helped make the decision deploy ARM-A @scale in the fleet



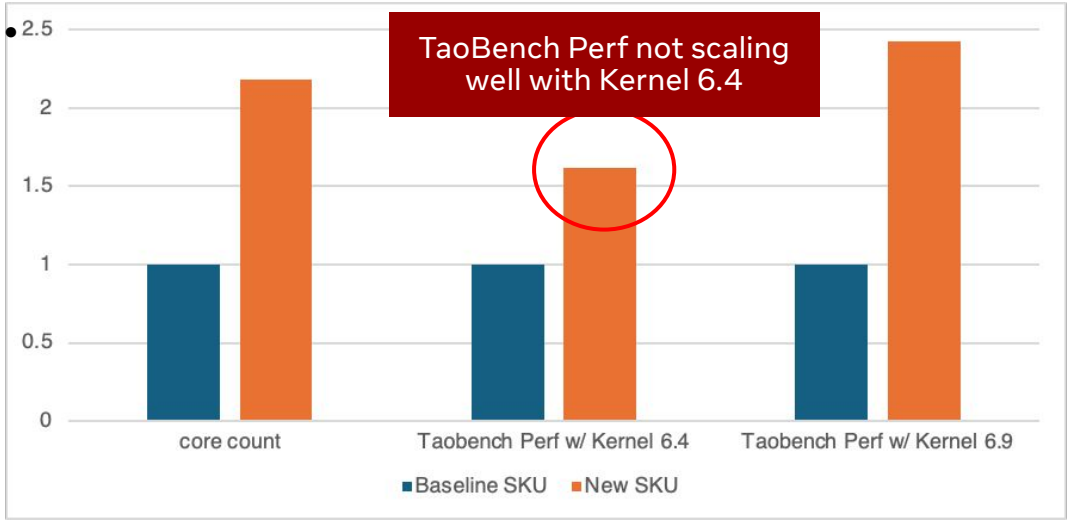
Case 2: Guiding uarch optimizations

- Worked with a vendor to tune one of Meta's large workloads on a new CPU
  - Used Mediawiki as a proxy to improve performance with vendor
  - After iterating over few weeks, we delivered 48% Perf/W improvement
  - Insights and improvements did NOT show up on SPEC CPU

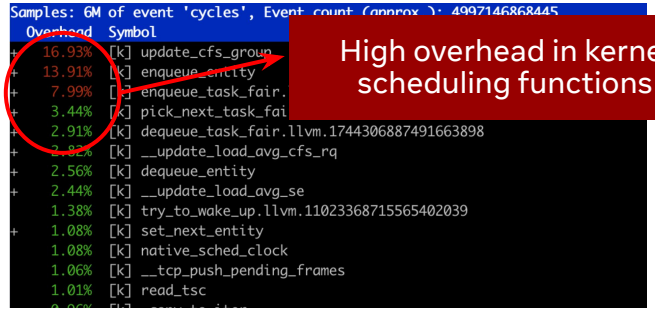


Case 3: Linux kernel improvement

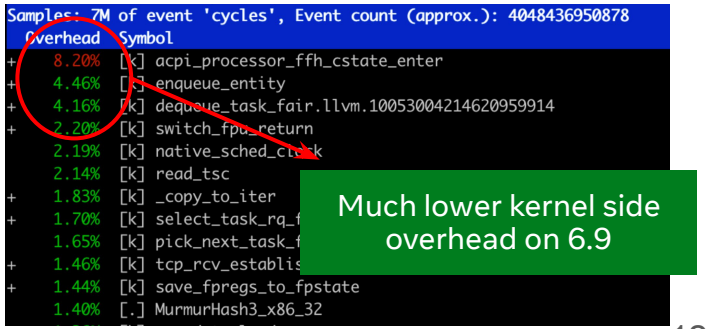
- Tested a new CPU SKU which very high core count
- Ran TaoBench on Linux 6.4 => scheduler issues
- Issue was not seen on targeted microbenchmarks
- Fast kernel perf issue detection and optimization.



New SKU on Kernel 6.4



New SKU on Kernel 6.9



DCPerf keeps evolving

Workload coverage

New benchmarks for emerging AI applications

DCPerf Mini

Shrunk version with short execution time for silicon exploration (simulation/emulation)

Timely refresh

Keeping up with the emerging HW & SW trends

Conclusion

- DCPerf is a representative benchmark suite for our datacenter applications
- It is validated using production workloads and guides procurement decisions and codesign for millions of servers
- We made it open-source to foster community innovation
  - Contributions from other hyperscalars and the broader community are welcome!