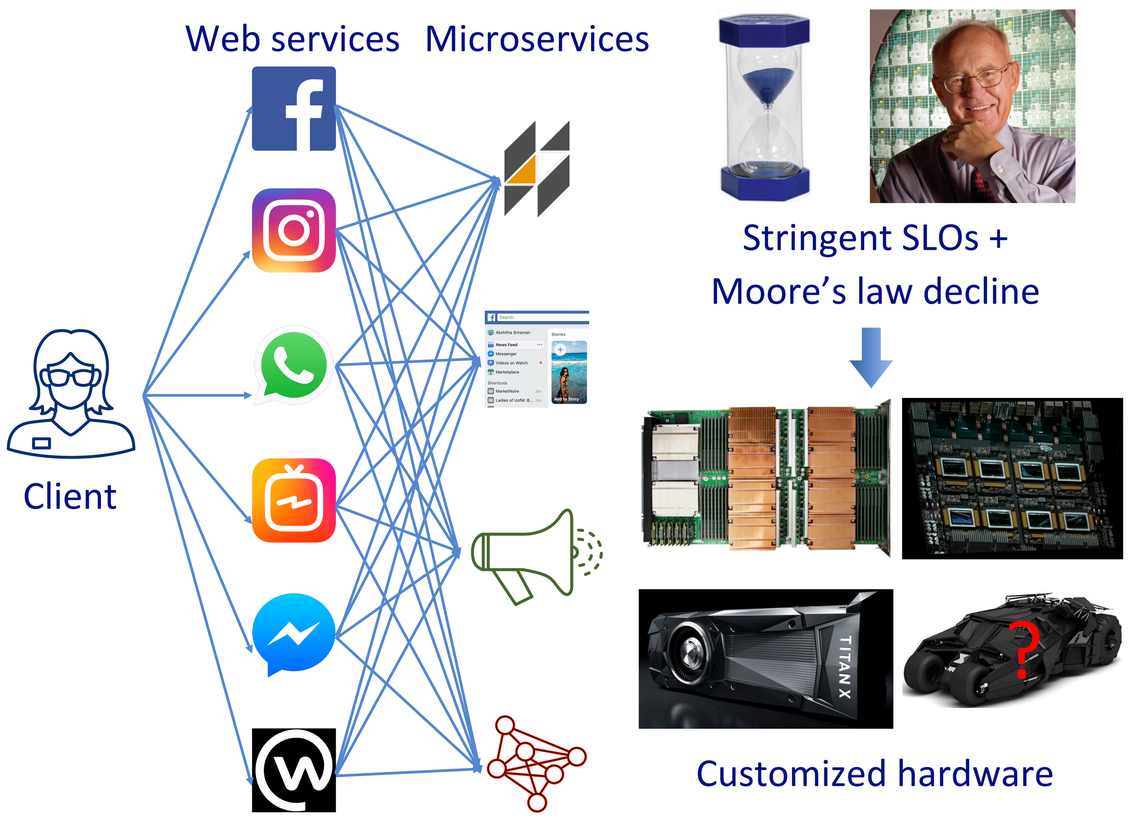
SoftSKU: Optimizing Server Architectures for Microservice Diversity @Scale

Akshitha Sriraman[✣▲], Abhishek Dhanotia[▲], Thomas F. Wenisch[✣]
University of Michigan[✣], Facebook[▲]

Rapid Increase in Modern Web Services



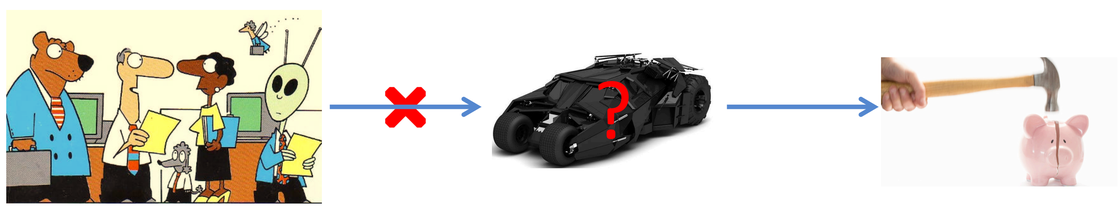
Stringent SLOs +
Moore's law decline

Customized hardware

Rapid increase in µservices -> greater need for custom hardware

Facebook µServices' Characterization

µService	Throughput (QPS)	Response latency	Pathlength
Web	O(100)	O(ms)	$O(10^6)$
Feed1	O(1000)	O(ms)	$O(10^9)$
Feed2	O(10)	O(s)	$O(10^9)$
Ads1	O(10)	O(ms)	$O(10^9)$
Ads2	O(100)	O(ms)	$O(10^9)$
Cache1	O(100K)	O(µs)	$O(10^3)$
Cache2	O(100K)	O(µs)	$O(10^3)$



Great diversity in bottlenecks — Use custom SKUs? — Prohibitively expensive

Can we achieve perf. efficiency without building custom SKUs?

"Soft" SKUs: Best of Both Worlds

Tune coarse HW & OS knobs on commodity HW

Performance efficiency

Procurement efficiency & scalability

- Core freq.
- Uncore freq.
- Core count
- CDP
- Prefetcher
- THP
- SHP

Are Custom Platforms Always Needed?

Customized platforms -> expensive

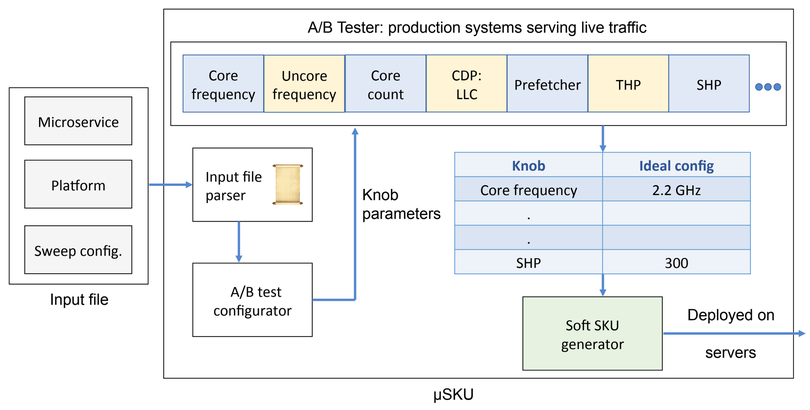
Procurement @scale

Data centers prefer hardware resource fungibility

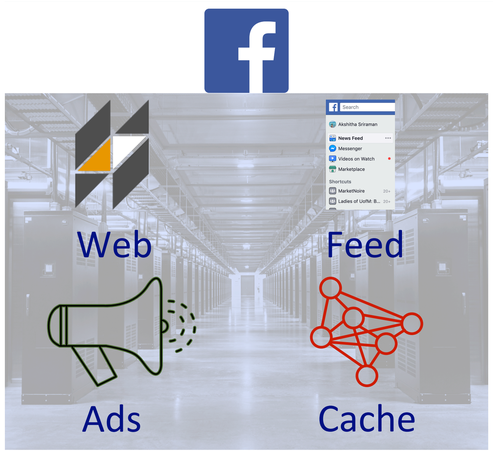
Low testing overhead

Dire need for limited CPU SKUs that support a variety of µservices

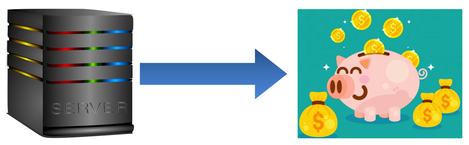
µSKU: Soft SKU Design & Deployment



Performance of Commodity Servers



Key FB µservices occupy a large portion of the data center

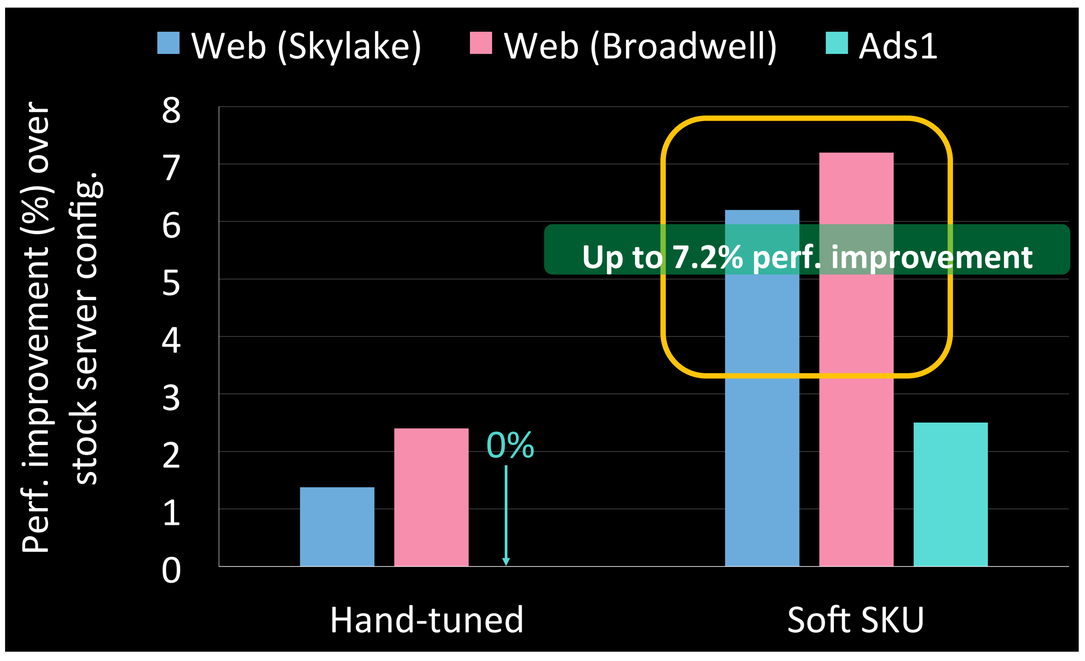


Use commodity hardware for procurement efficiency & scalability?

How performant is commodity hardware for these µservices?

Are there common bottlenecks that can inspire future SKUs?

Soft SKU Performance



*Soft SKU can achieve **~7.2% throughput improvement** on **production systems** with **no extra hardware requirement***

Contributions

- Comprehensive characterization of Facebook's microservices
  - ❑ System-level & architectural bottlenecks
  - Reveals enormous bottleneck diversity across microservices
- Concept of "soft" server SKUs
  - ❑ Tuning coarse-grained OS & hardware configuration knobs
- µSKU
  - ❑ Automates soft-SKU search & configuration via production A/B tests
  - ❑ Deploys soft SKUs on production microservices

~7.2% perf. boost on **production** µservices + **no extra hardware**