工作汇报

程硕

Adaptive Server Parameter Tuning

- Applications should manage the agreed QoS metrics:
  - Response time
  - Throughput

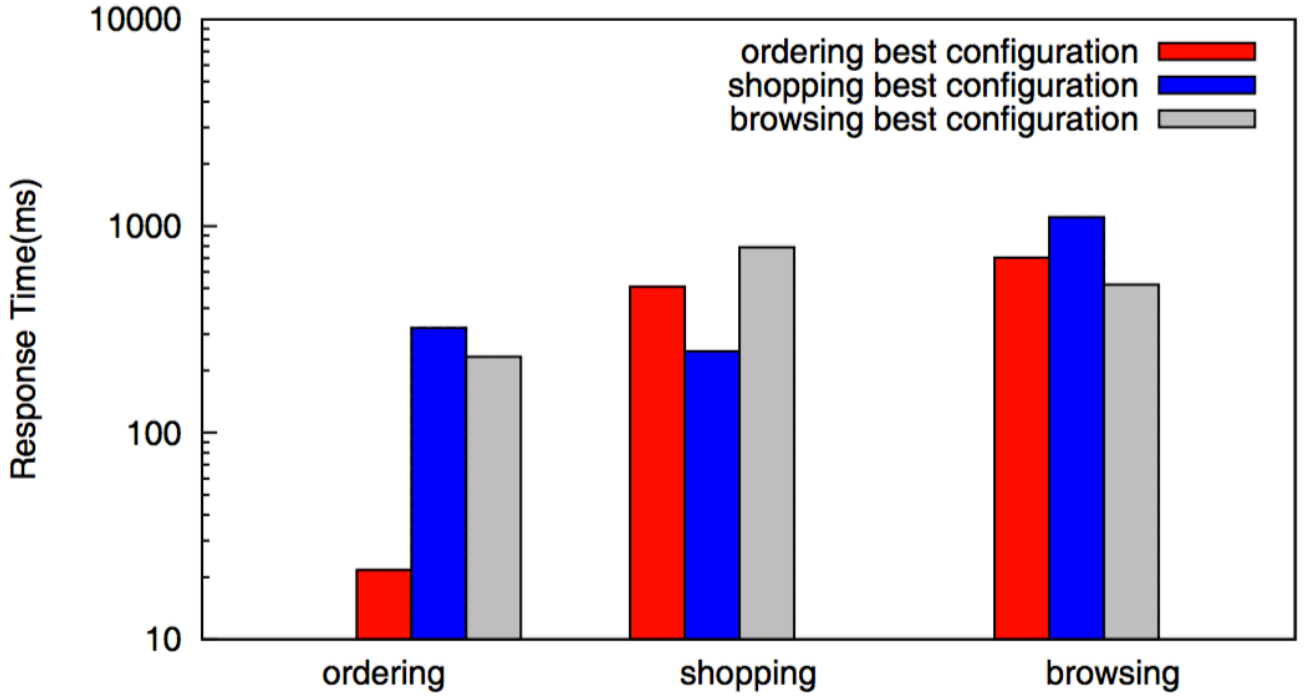
- Servers have many configurable parameters:
  - Apache web server: MaxClients, KeepAliveTimeout …
  - Database: max_connections, thread_concurrency …
  - Application: thread pool size, timeout, retry values …

Adaptive Server Parameter Tuning

- Conventional Method:
  - rules-of-thumb
  - trial-and-error
  - intuition
  - …

Adaptive Server Parameter Tuning

No single universal configuration is good for all workloads [1].

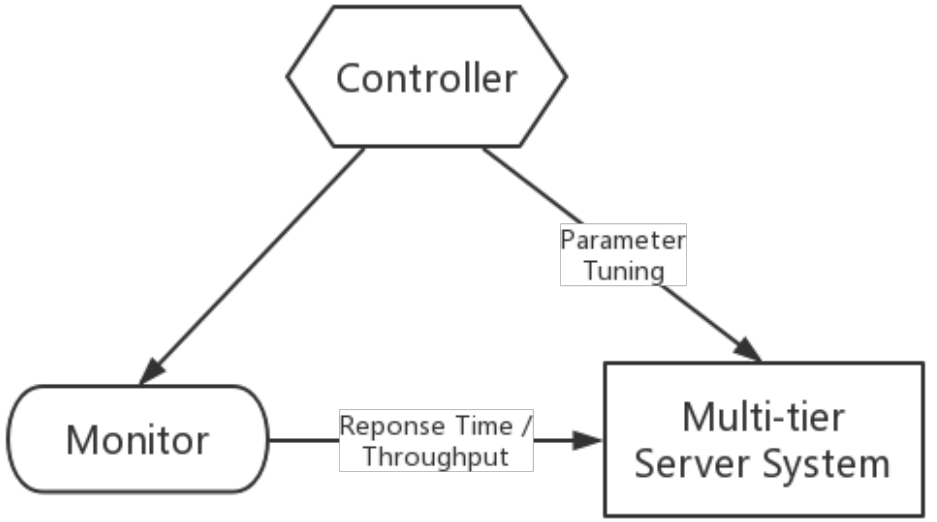


Performance under configurations tuned for different workloads

Adaptive Server Parameter Tuning

- Complexities of Adaptive Tuning
  - high workload dynamics
  - burstiness
  - multi-tier service architecture
  - virtualized server infrastructure
  - …

Adaptive Server Parameter Tuning



General architecture of Adaptive Tuning System

Adaptive Server Parameter Tuning

Main Concern -- How to generate optimal configuration?

- Ferrari et al. [2]:
  - Offline testing with benchmark applications;
  - Best configurations at different loads are generated;
  - Online applications select one from the candidate configurations

Adaptive Server Parameter Tuning

- Bu et al. [4]: Reinforcement Learning
  - State space: group of parameters
  - Action set: *increase*, *decrease* and *keep the parameter value*
  - Reward: *SLA* - $perf_t$

- Guo et al. [3]:
  - Neural networks and Fuzzy control

参考资料

[1] I.-H. Chung and J. K. Hollingsworth. Automated cluster- based web service performance tuning. In *HPDC*, pages 36– 44, 2004.

[2] Ferrari, Giovanna, Santosh Shrivastava, and Paul Ezhilchelvan. "An approach to adaptive performance tuning of application servers." *IEEE International Workshop on QoS in Application Servers*. 2004.

[3] Guo, Yanfei, et al. "Automated and agile server parametertuning by coordinated learning and control." *IEEE Transactions on Parallel and Distributed Systems* 25.4 (2014): 876-886.

[4] Bu, Xiangping, Jia Rao, and Cheng-Zhong Xu. "A reinforcement learning approach to online web systems auto-configuration." *Distributed Computing Systems, 2009. ICDCS'09. 29th IEEE International Conference on*. IEEE, 2009.

PostCopy VS PreCopy in QEMU

- 实验环境
  - Dell Mini Server * 2 (4G Mem; i5-4590T@2.00GHz)
  - Ubuntu 14.04.1 Server (4.3 kernel)
  - Qemu 2.6
  - Migrat speed limit: 10 Gbps

PostCopy VS PreCopy In QEMU

PreCopy

PostCopy