Generic Environment for Full Automation of Benchmarking

Tomáš Kalibera, Lubomír Bulej, Petr Tůma

DISTRIBUTED SYSTEMS RESEARCH GROUP
http://nenya.ms.mff.cuni.cz

CHARLES UNIVERSITY, PRAGUE
Faculty of Mathematics and Physics

History: Middleware Benchmarking Projects

- Vendor testing
  - Borland, IONA, MLC Systeme
- Open source testing
  - omniORB, TAO, OpenORB, …
- Open CORBA Benchmarking
  - anyone can upload their own results

Motivation: Regression Testing for Performance

- Regression testing
  - integrated into development environment
  - tests performed regularly

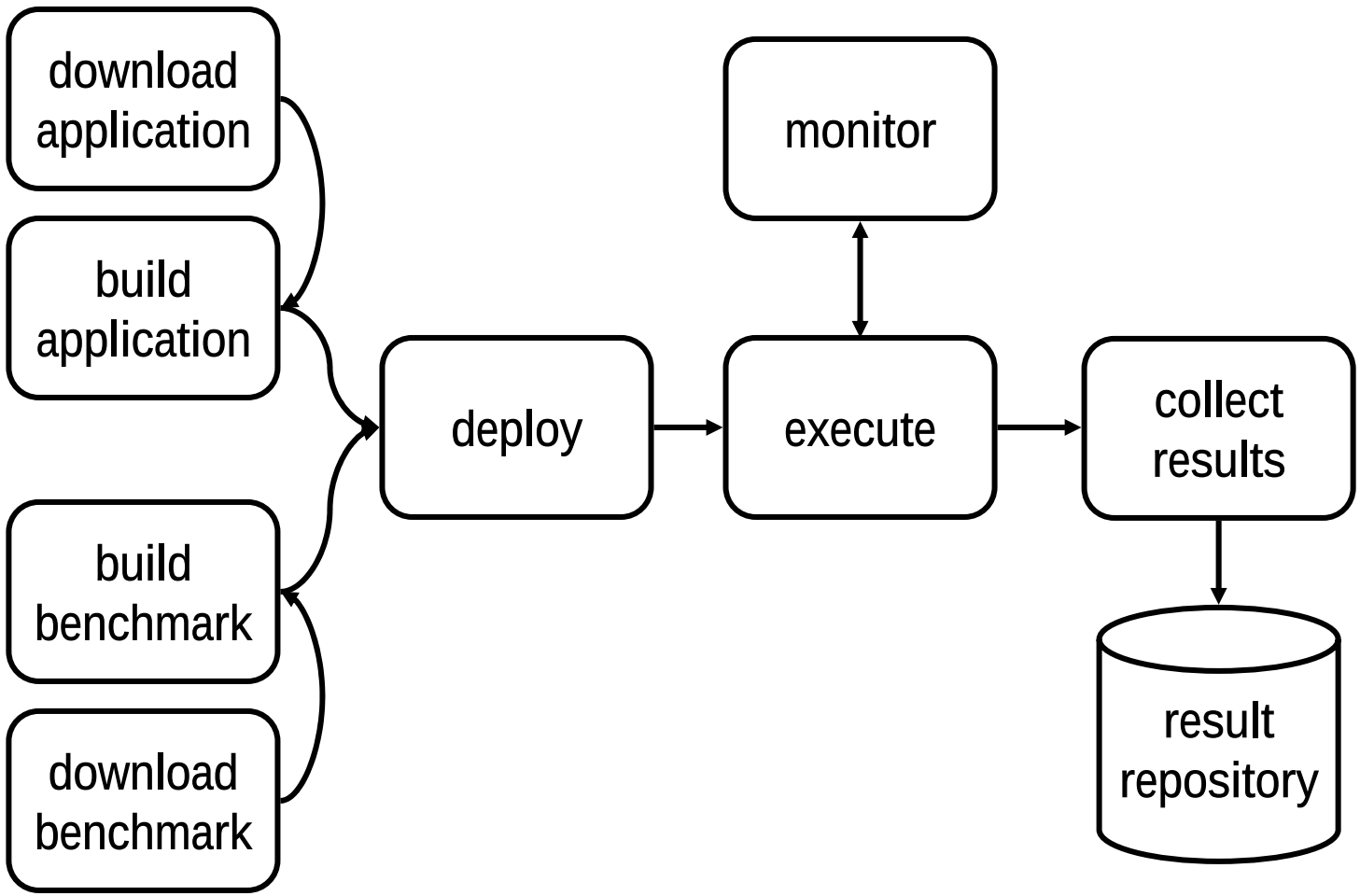
- Correctness tests
  - commonly used
  - detect bugs

- Performance tests
  - in research stage
  - detect performance regressions

Regression Benchmarking

- Detection of performance regressions
  - benchmarking performance of consecutive versions of software
  - automatic comparison of results
- Issues
  - automatic comparison of results
    - fluctuations in results
    - results format, different level of detail
  - automatic running of benchmarks
    - monitoring, failure resolution

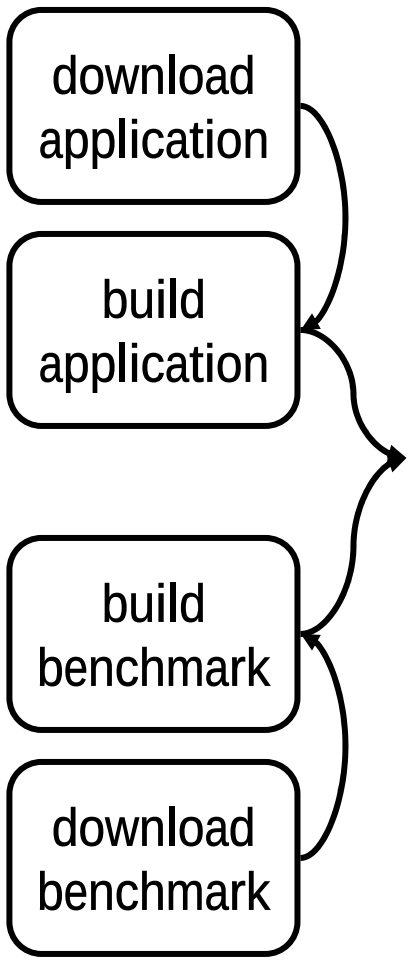
Steps of Running a Benchmark



Generic Benchmarking Environment

- Automated processing
  - monitoring, handling of failures
  - management tools
- Common form of results
  - allow benchmark independent analysis
  - raw data, system configuration
- Flexibility
  - benchmark and analysis independence

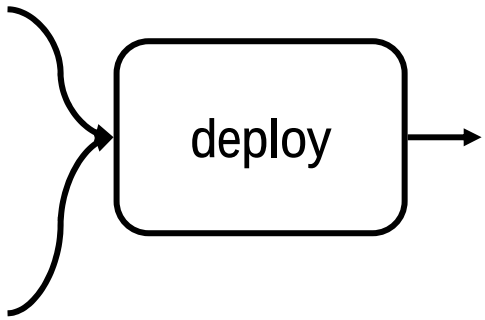
Automatic Downloading and Building



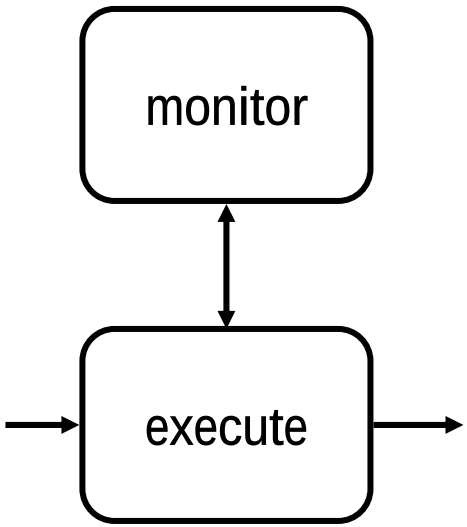
- Download methods
  - cvs checkout, http, ftp
- Build methods
  - Ant, make, scripts
  - support different platforms
- Software repository
  - storage for sources, binaries
  - annotated for future reference

Automatic Deployment

- Reproducibility
- Platform dependencies
  - CPU type
  - operating system
- Resource requirements
  - CPU frequency
  - RAM
- Software requirements
  - database server
  - web server



Automatic Execution

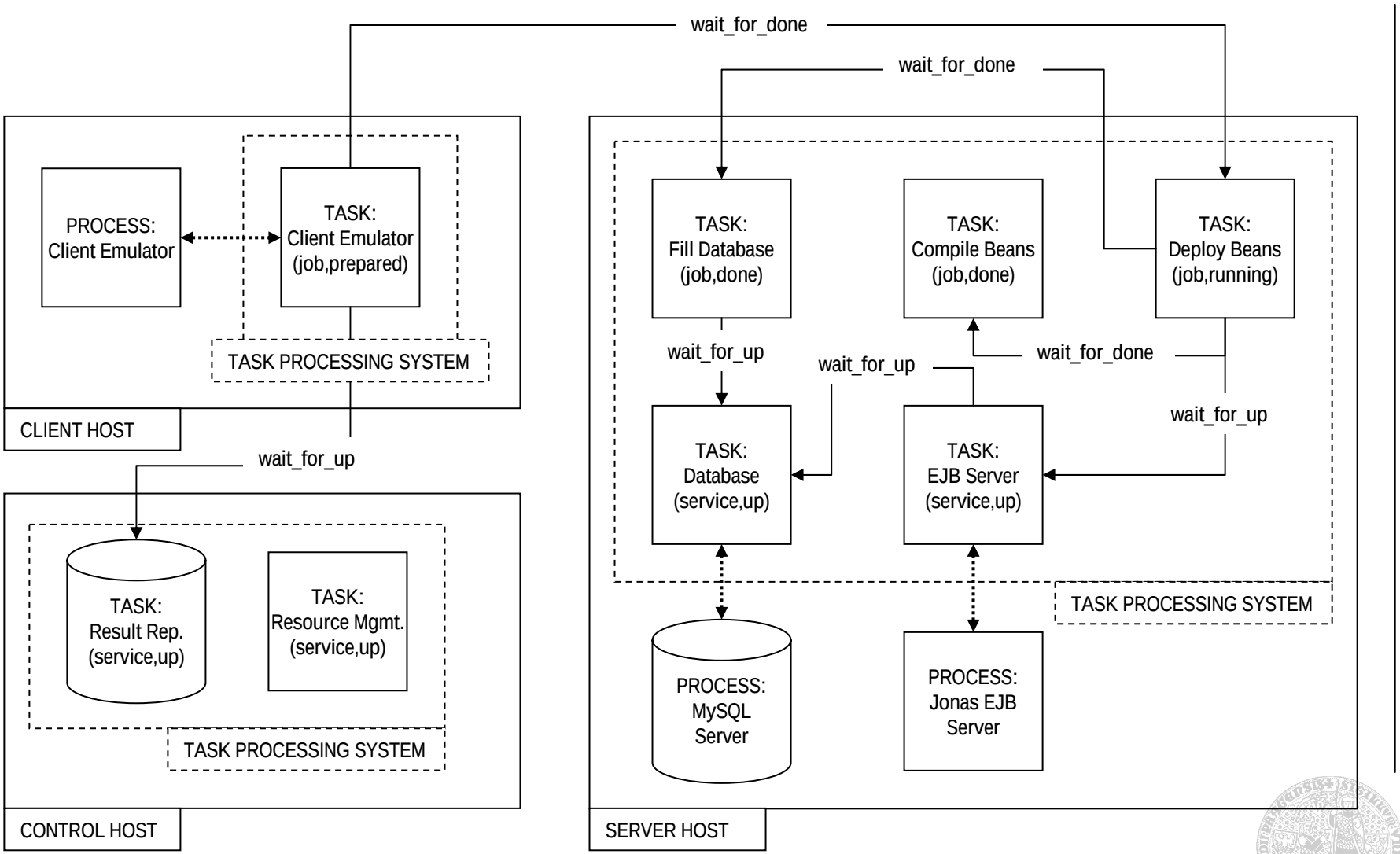


- Multiple applications
  - run in correct order
  - wait for initialization
- Monitoring
  - detect crashes
  - detect deadlocks
  - but do not distort the results !

Architecture of Benchmarking Environment

- Task processing system
  - deployment, execution, monitoring of tasks
  - task scheduler – dependencies on other tasks, checkpoints
  - jobs, services
- Environment tasks
  - result repository, software repository
- Benchmarking tasks
  - benchmarks, compilations, required apps

Example: RUBiS Benchmark



Conclusion & Future Work

- Generic benchmarking environment
  - automatic running of (existing) benchmarks
  - common form of results, result repository
- Current status
  - early implementation phase
- Future work
  - support for Xampler, RUBiS benchmarks
  - automatic detection of regressions
  - regression benchmarking of CORBA, EJB

Publications

- Bulej L., Kalibera T., Tůma P.: Repeated Results Analysis for Middleware Regression Benchmarking, accepted for publication in Special Issue on Performance Modeling and Evaluation of High-Performance Parallel and Distributed Systems, in Performance Evaluation: An International Journal, Elsevier

- Bulej, L., Kalibera, T., Tůma, P.: Regression Benchmarking with Simple Middleware Benchmarks, in proceedings of IPCCC 2004, International Workshop on Middleware Performance, Phoenix, AZ, USA

- Buble, A., Bulej, L., Tůma, P.: CORBA Benchmarking: A Course With Hidden Obstacles, in proceedings of the IPDPS Workshop on Performance Modeling, Evaluation and Optimization of Parallel and Distributed Systems (PMEOPDS 2003), Nice, France

- Tůma, P., Buble, A.: Open CORBA Benchmarking, in proceedings of the 2001 International Symposium on Performance Evaluation of Computer and Telecommunication Systems (SPECTS 2001), published by SCS, Orlando, FL, USA