Benchmarking in WP 2.1

Oscar Corcho, Raúl García-Castro, Asunción Gómez-Pérez
<ocorcho,rgarcia,asun@fi.upm.es>

March 4th, 2004

Table of Contents

Benchmark and benchmarking definitions

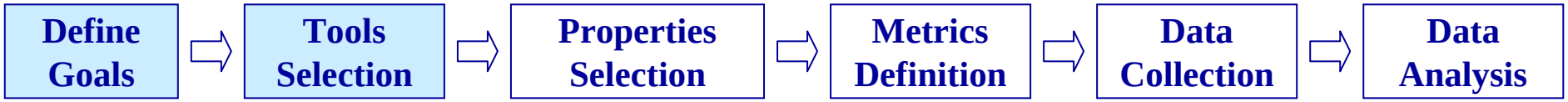
Benchmarks are experiments to produce data (primarily) for:

- Hardware and software system evaluations.
- Load testing. To analyze the effect of varying the load of a system.
- Performance measurement.

Benchmarking is a systematic and continuous process of measuring and evaluating performance, practices, and processes, within and between systems, to obtain information for improvement.

Smith, Connie U. "Performance Engineering of Software Systems". Addison-Wesley, 1990
Alstete, J. W. "Competitive Benchmarking Course". Technical Report, 1992
http://www.iona.edu/faculty/jalstete/MNG992/documents.htm

Benchmarking. State of the Art



Goals

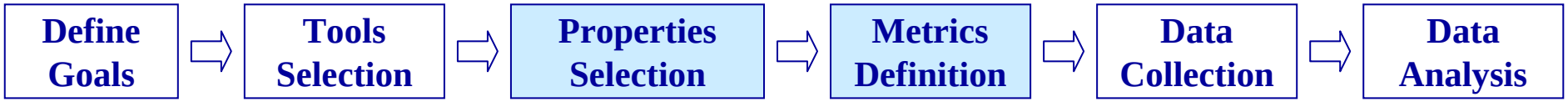
- Evaluate technical suitability
- Evaluate economic suitability
- Obtain improvement recommendations

Tools

- Internal benchmarking
  - All through the system's life cycle.
- Competitive benchmarking
  - At the operation phase of the life cycle.

Alstete, J. W. "Competitive Benchmarking Course". Technical Report, 1992
http://www.iona.edu/faculty/jalstete/MNG992/documents.htm
Ontoweb Deliverable 2.1Successful Scenarios for Ontology-based Applications

Benchmarking. State of the Art



Properties

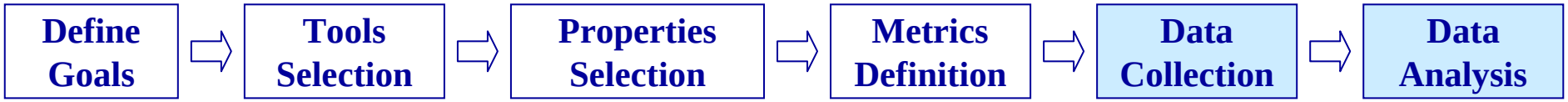
- Tool properties:
  - Performance
  - Correctness
  - …
- Service properties:
  - Stability
  - Usability
  - …
- Business properties:
  - Viability
  - …
- Support:
  - Documentation
  - Tutorials
  - …

Measured by

Metrics

- Quantitative:
  - Results per second
  - Max. number of users
  - Memory allocation
  - Return of investment
  - …
- Qualitative:
  - Degree of interoperability
  - Automated functionality
  - Provides support
  - …

Benchmarking. State of the Art



Data collection

- Benchmarks: •Automatic •Interactive
- Measurements
- Surveys
- …

Analysis

- Summarize and interpret data
- Analyse

We need test beds to help in the data collection and analysis.

Test bed. An environment containing the hardware, instrumentation, simulators, software tools, and other support elements needed to conduct a test.

Although there are general test beds for testing software, they don't provide:

- Specific domain data
- Granularity
- Control over data

IEEE Std 610.12-1990, IEEE Standard Glossary of Software Engineering Terminology
Gray, J. *The Benchmark Handbook for Database and Transaction Systems* (2nd Edition). Morgan Kaufmann 1993

Table of Contents

WP 2.1. Benchmarking of ontology based tools

Goals (4 years):

1. Methodology and general criteria for different types of ontology based tools benchmarking
2. Construction of prototypes of tools for benchmarking ontology tools.
3. Benchmarking of ontology tools according to the criteria and test beds produced.

WP 2.1. Benchmarking of ontology development tools

Goals (18 months):

1. State of the Art (month 6)
2. First draft of a methodology (month 12)
3. Identification of criteria (month 12)
4. Identification of metrics (month 12)
5. Definition of test beds for benchmarking (month 12)
6. Development of first versions of prototypes of tools (month 18)
7. Benchmarking of ontology development tools according to the criteria and test beds produced (month 18)

Types of ontology-based tools

- **Ontology development (WP2.1, first 18 months)**
  - **Editors**
  - **Translators**
- **Annotation (WP2.1, after 18m)**
- **Storage, querying and reasoning (WP2.1, after 18m)**
- **Others (WP2.1, after 18m)**
- Merging and alignment (WP2.2)

Ontology development tools included

- OntoEdit
- Protégé 2000
- WebODE
- ...

Benchmarking in WP 2.1

- Scope
  - Internal/Competitive benchmarking?
  - Tools/Ontology development tools?
  - What properties?
- Summary/index
- Prototypes of tools/test beds
  - •Data collection
  - •Analysis
  - Requirements?
- Benchmarking?
- Contributors?

Partners involved

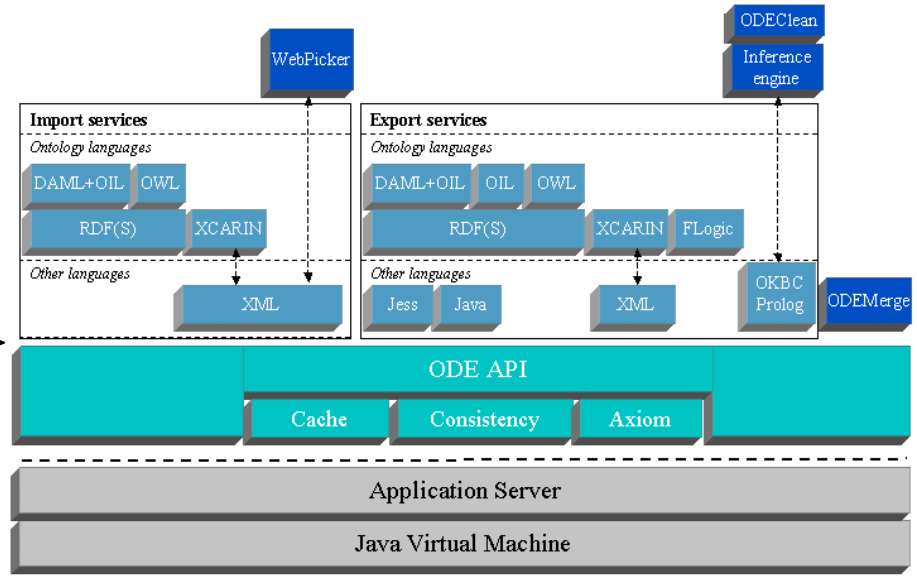
Partner	Person	State of the Art	Methodology	Test Beds	Prototypes of tools	Benchmarking
UPM	Raul Garcia-Castro rgarcia@fi.upm.es	Yes	Yes	Yes	Yes	Yes (WebODE)

An example

ESPERONTO

WebODE

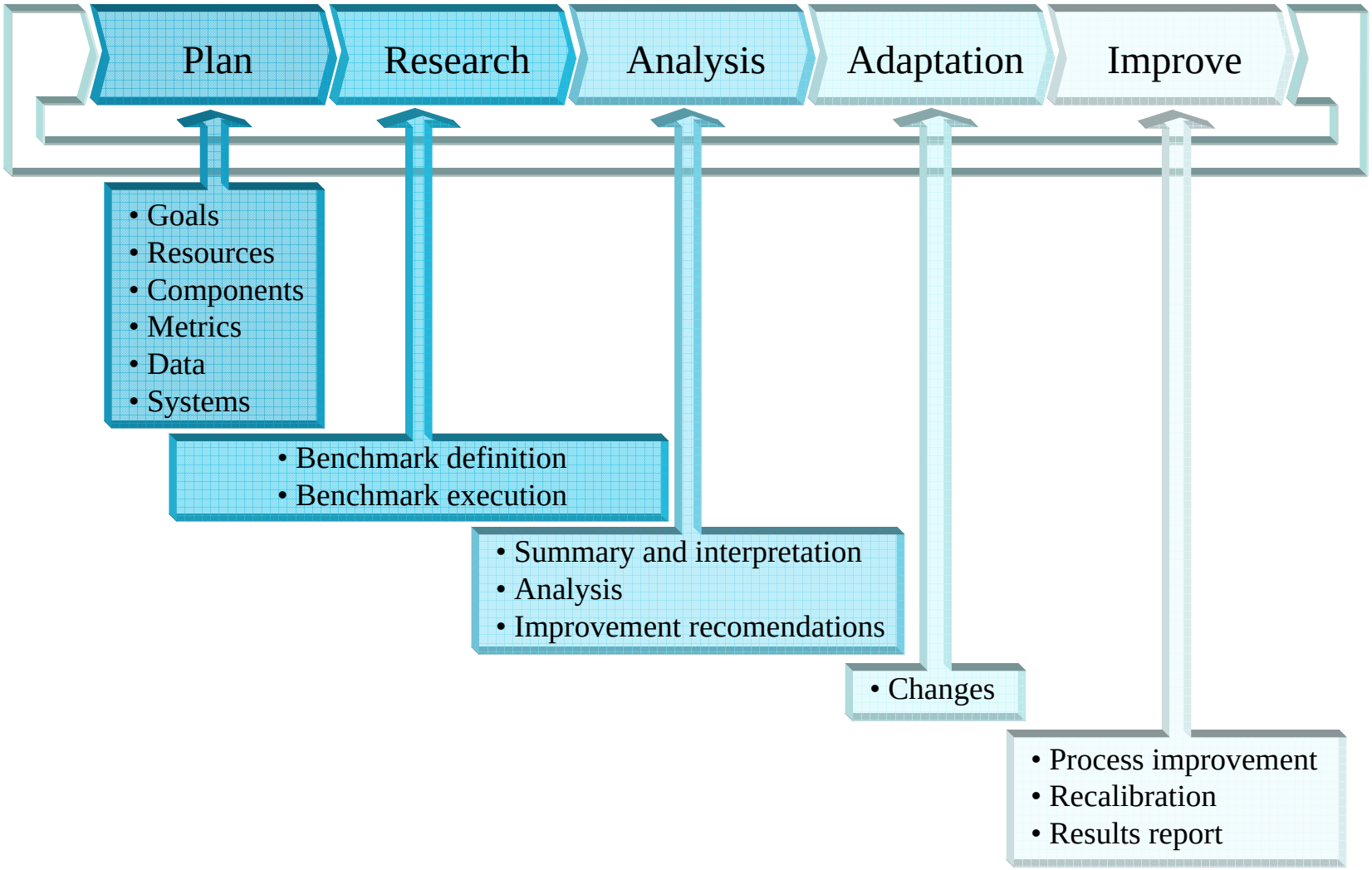
Assess the performance of WebODE measuring the execution time of its API methods.



Scenarios:

- **Over a high load state** (ANOMALIES)
- **Repeatedly over the same load state** (STABILITY)
- **Over incremental load states** (LOAD/PERFORMANCE)
- **With different input parameters** (DIFFERENT BEHAVIOR)

Methodology



Benchmarking in Knowledge Web

Benchmarking present in:

- WP 1.2
- WP 1.3
- WP 2.1
- WP 2.2

Unify effort in benchmarking?

Benchmarking in WP 2.1

Oscar Corcho, Raúl García-Castro, Asunción Gómez-Pérez
<ocorcho,rgarcia,asun@fi.upm.es>

March 4th, 2004