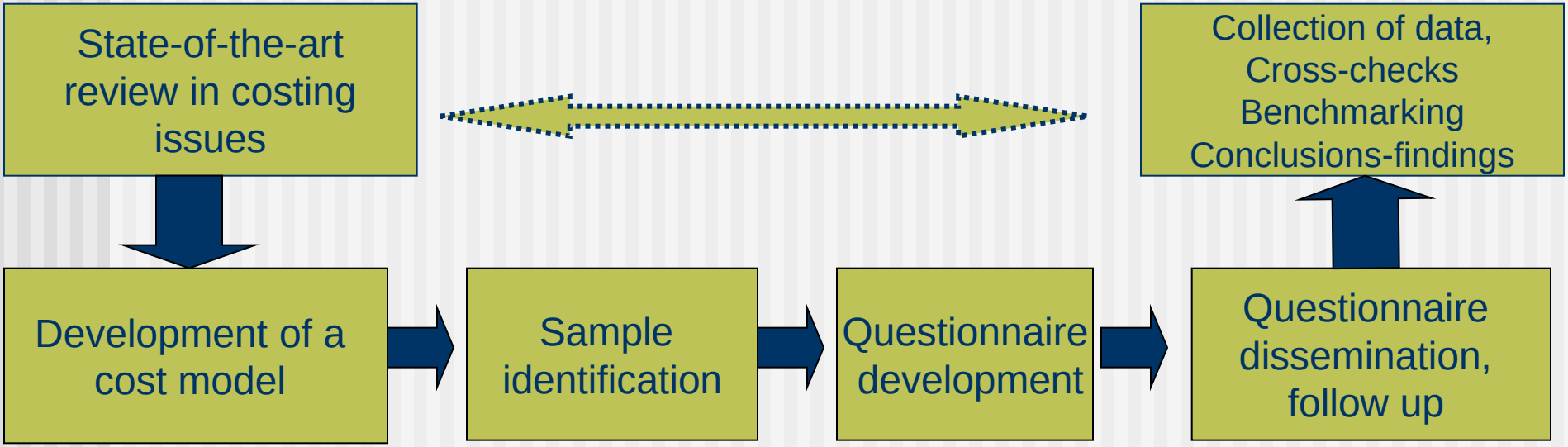
e-FISCAL

e-FISCAL project Workshop
21 September 2011

Setting the scene for e-Infrastructure Cost analysis
e-FISCAL methodology
Draft Questionnaire

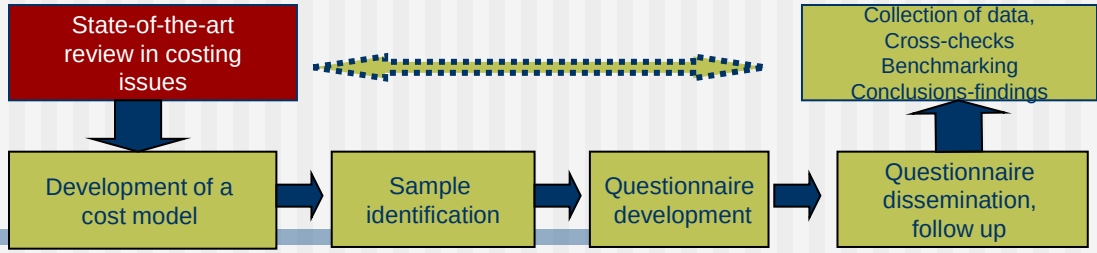
Sandra Cohen and Fotis Karagiannis
AUEB

Overview



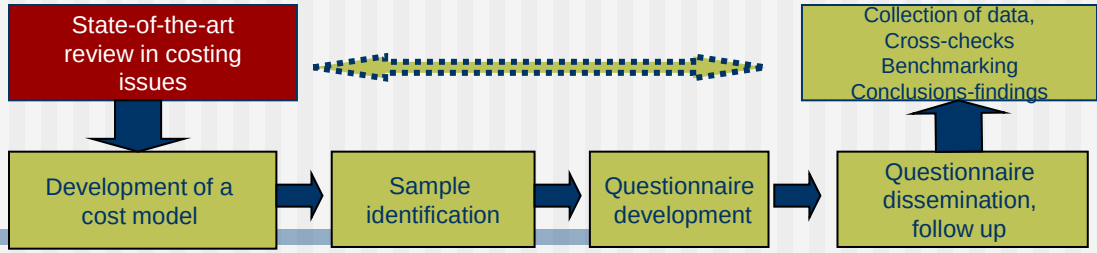
- Calculation of the total yearly cost of ownership
  - Calculation of the cost per logical CPU/hour (core/hour) under different settings
  - Calculation of several metrics
  - Publication of a generic cost model with benchmarks

State of the art



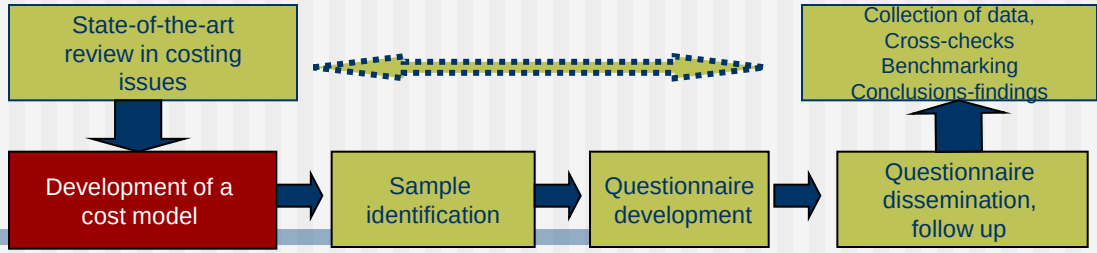
- Prepare a repository
  - Academic papers, Industry project results, EU studies' results, Vendor analyses, etc.
- Multi scope orientation:
  - Costing issues in general
  - Business Models
  - Cloud vs Grid papers
  - Migration to the Cloud papers
  - Industry benchmarks
- On-going procedure throughout the project
- Use of the state of the art input to:
  - Develop the cost model to be applied
  - Get input to business model, cost and pricing analysis
  - Contribute to relevant discussions
  - Highlight publications and results from the participating organisations

State of the art



- Basic conclusions from literature review :
  - Rather typical break down of HTC/HPC costs:
    - Computing and storage hardware costs including interconnection costs, auxiliary equipment costs (cooling, UPSs, power generators), software costs, personnel costs, site operating costs, network connectivity costs
  - Same cost structure, but very different results
    - Cost per (logical) CPU/hour
  - Comparisons with commercial providers under different hypotheses
    - Service quality characteristics pertaining the numerical results
  - The initial e-IRGSP2 financial exercise identified several pitfalls in cost calculations
    - Reason for variation in €/CPU hour results
    - Addressed in the e-Fiscal survey design

The model



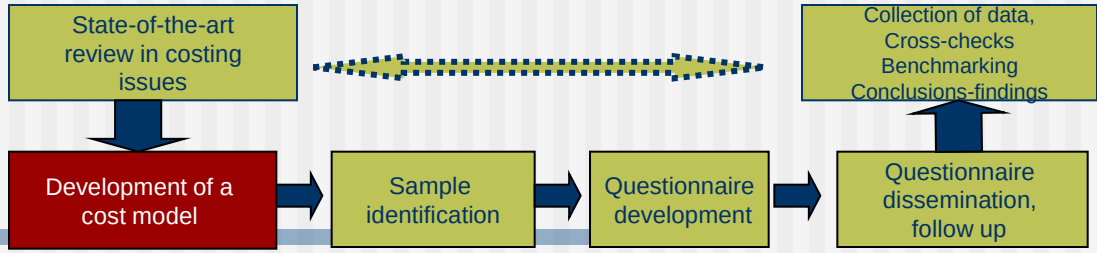
- Two step methodology to approximate the total yearly cost of ownership
- Simulation of the physical infrastructure:
  - Cross-checks with available accounting data will be performed.

Estimation of the site/centre investment in terms of logical CPUs, storage devices, auxiliary equipment, connectivity devices
X
Prices per logical CPU, for storage, percentages, etc.

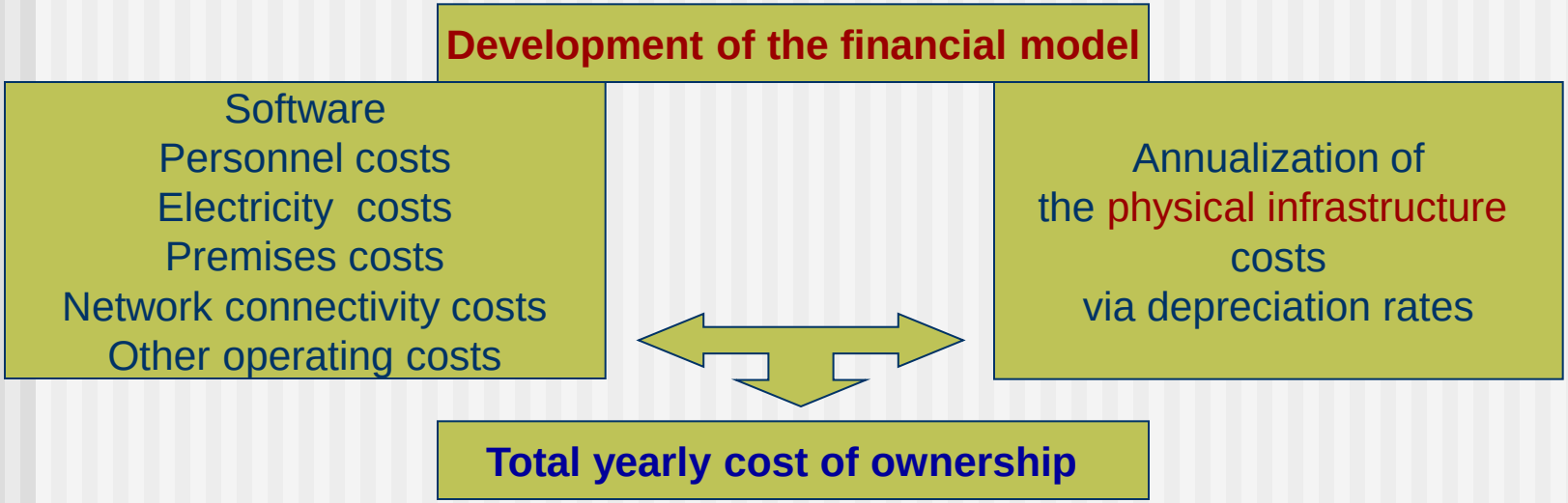
→

Simulation of the physical infrastructure
Approximation of the current physical infrastructure investment cost

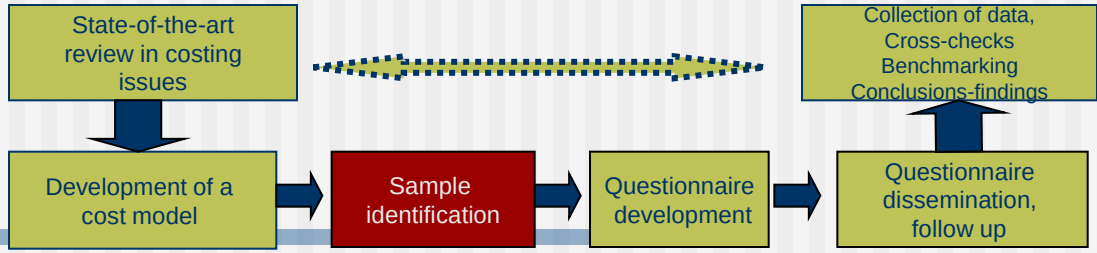
The model



- Development of the financial model:
  - Annualized cost of the simulated physical infrastructure
    - Application of depreciation rates
  - Operating cost of the physical infrastructure
    - Fine-grained breakdown of the operating costs is important: model validation, economies of scale, identifying "outliers" for further analysis

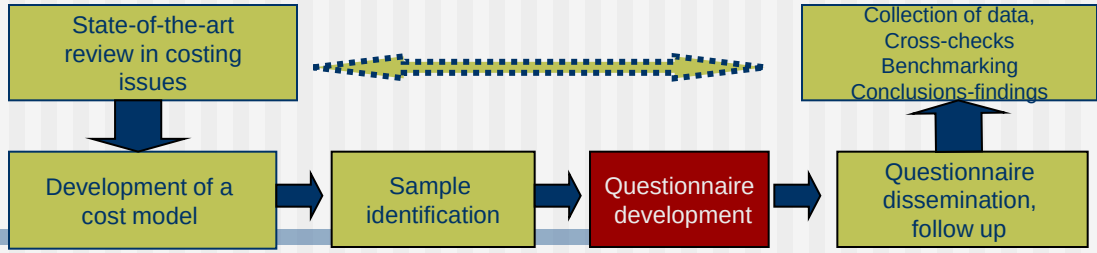


Sample



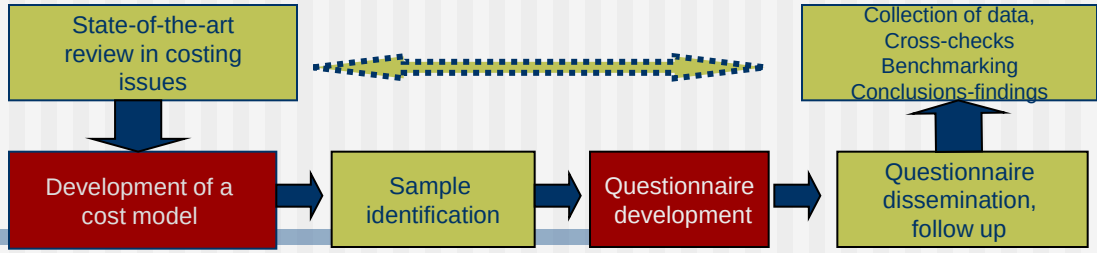
- Ideally our sample (respondents) is the total population of:
  - EGI/HTC sites
  - PRACE/HPC centers
- Practically the sample should be adequate to ensure representation of different
  - HTC/HPC site/centre sizes
  - Countries
  - e-infrastructure types
- We aim at having a response rate above 50%
- Prerequisites for:
  - Extrapolating the costs of the sample to the "population"
  - Assessing any economies of scale effects

Questionnaire



- Questionnaire characteristics (so far):
  - Focus on 2010-2011 costs (plus forward looking considerations)
  - Thematic development (12 sections – 16 pages)
  - Monetary values (Euros) and metric volumes questions (e.g. number of cores, number of FTEs)
  - Questionnaire is administered on special software
    - On line version (http://www.surveymonkey.com/s/7N5QDCT)
    - Editable pdf. version
- Cost questionnaires require considerable effort
  - Especially ones designed to avoid "apples and oranges" comparisons
- Your efforts are very much appreciated!
- Reliability of the results depend on the level of participation
  - Reliable, robust financial models and cost estimates are prerequisites for **sustainability**
  - Governments need to know how much we're asking and for what - and they need to believe in value for money

Model and Questionnaire



Reference to 2010 and 2011

Estimation of the site/centre investment in terms of logical CPUs, storage devices, auxiliary equipment, connectivity devices
X
Prices per logical CPU, for storage, percentages, etc.
retrieved by questionnaires

Simulation of the physical infrastructure
Approximation of the current physical infrastructure investment cost

Development of the financial model

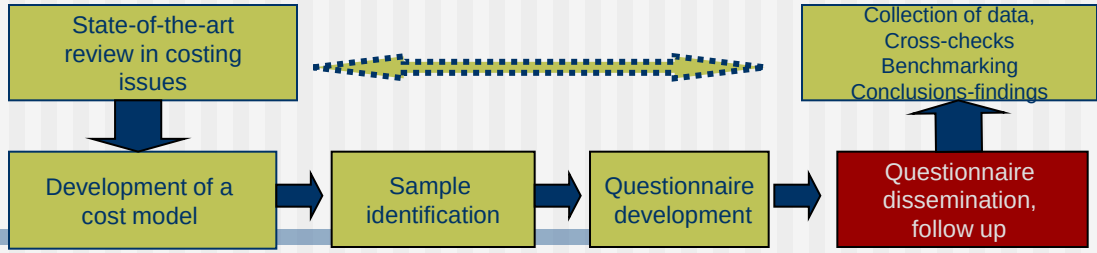
Software
Personnel costs
Electricity costs
Premises costs
Network connectivity costs
Other operating costs
(questionnaire)

Annualization of the physical infrastructure costs via depreciation rates
(questionnaire)

- Cross-checks with available data (from EGI, PRACE, market or other literature sources) will be consistently performed

Total yearly cost of ownership

Survey flow chart



Survey execution

Basic Questionnaire (common)

Feedback (from different sources)

Check and refine

Infrastructure Type?

EGI (HTC)

PRACE (HPC)

Other
Not neatly fit to cost HPC/HTC division

NGI HQ quest.

PRACE Partner Quest.

Personalized process

Select sites

Site questionnaire

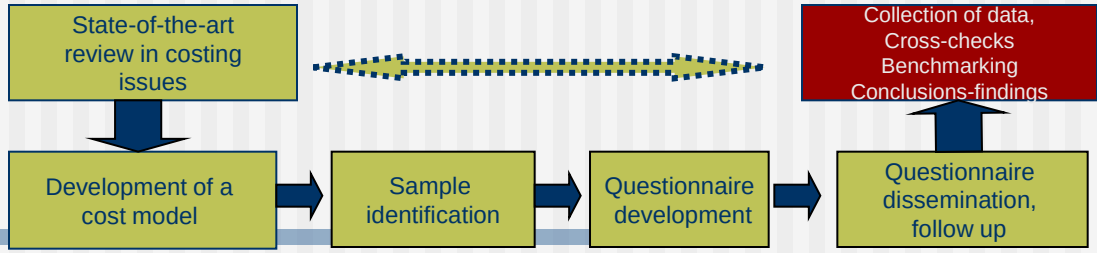
OR

Consistency, Completeness and Cross checking

Model re-verification Questionnaire review

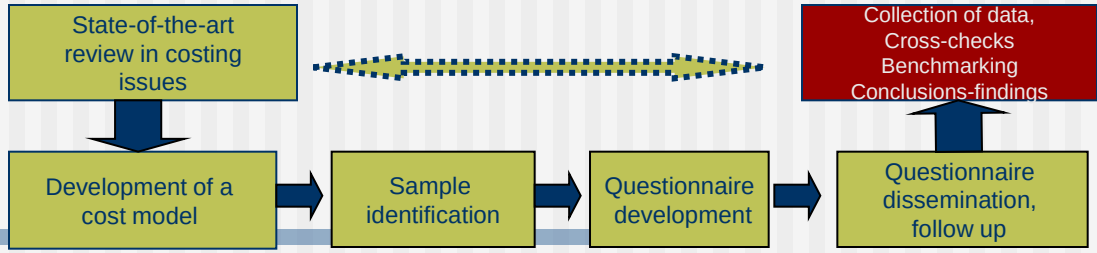
Follow –up processn

Analysis



- Cross checks of our finding with literature
  - Comparisons with published data
- Use questionnaire information (and industry data) to calculate costs and produce metrics
  - Cost per CPU (or CPU core), Opex/Capex ratio, Personnel number (FTEs)/CPU, etc.
- Pay attention to
  - Confidentially of data/ Anonymity
  - Validation of results/ Cross – checks
    - Inform our findings with input from interviews, case studies
- Development of a generic cost model
  - Development of an excel calculation spreadsheet (public)

Analysis (II)



- Performance of comparisons with commercial leased and on-demand offerings
- Selection of services with which our cost estimates would be compared
  - Important to find services for which meaningful comparisons can be performed
    - e.g Infrastructure as a Service (IaaS) offerings such as Amazon EC2, and S3
    - Commercial HPC offerings such as those offered by SGI's Cyclone, Penguin's On Demand computing or Amazon's EC2 Cluster HPC
- Caution: Comparing prices with costs
- Inclusion of other than cost considerations
  - Trust, reliability, customer satisfaction…

Any questions?
A detailed presentation of the questionnaire follows