



B E T S Y

Deliverable D6b
Evaluation Plan
Version 1.0

Public



Project Number : IST-004042

Project Title : BETSY
Deliverable Type : Report

Deliverable Number : D6b
Title of Deliverable : Evaluation Plan
Nature of Deliverable : Report, Public
Internal Document Number : betsy-phr-d6b-v01.0.doc
Contractual Delivery Date : 10 June 2005
Actual Delivery Date : 10 July 2005
Contributing WPs : WP6
Author(s) : P. van der Stok (Philips) H. de Groot (Philips) M. Joosten (C-lab)

Abstract

This document describes the evaluation plan for the work-package results developed during the BETSY project.

Keyword list

Evaluation criteria

Table of Contents

Table of Contents.....	2
1 Introduction.....	3
2 Proposed assessment approach.....	4
3 Assessment of the BETSY objectives	5
3.1 Task 2.1 Identification of end-to-end delivery chain parameters	6
3.2 Task 2.2 Stream description methods.....	7
3.3 Task 2.3 Decomposition of end-to-end constraints.....	8
3.4 Task 3.1 Parameter and resource requirements.....	9
3.5 Task 3.2 Device functional component modeling	10
3.6 Task 3.3 Network functional component modeling.....	11
3.7 Task 3.4 Model composition	12
3.8 Task 4.1 Specification of the BETSY Framework.....	13
3.9 Task 4.2 Design of the test bed and the demonstrator	14
3.10 Task 4.3 Reference implementation of the BETSY framework on the test-bed.....	15
3.11 Task 4.4 Porting of reference implementation on the demonstrator set-up	16
3.12 Task 4.5 Validation and evaluation of BETSY models	17
4 Conclusions	18
References	19

1 Introduction

This document describes how the results of the BETSY project [1] can be assessed and measured at the end of the project. The goal is to measure the quantity and quality of the results of the results of the BETSY project at the task level against the initial objectives.

The deliverable D1d described the evaluation criteria on the appropriateness of the chosen and developed BETSY techniques. It provides a first insight in the actual technical objectives of the BETSY project. As remarked by the reviewers, these technical objectives should be visible in the demonstrator. This document will also define the criteria of success for the final demonstrators as the consortium sees them today. Together with this document a *Risk Management Plan* (D6c) is written that highlights the risks of the BETSY project and their fall-back scenarios.

Rather than providing a set of seemingly unrelated numbers by quantifying absolute metrics for the BETSY objectives, a staged assessment approach is proposed that is based on the individual tasks of the BETSY project. In the next section the proposed assessment approach is presented after which the results of the approach are described.

The described tasks are directly related with the technical objectives. That means that work-packages wp5 and wp6 are not considered in this document. Work-package 1 has prepared the terrain for wp2-wp5 and does not contribute directly to the BETSY objectives. Therefore, wp1 will not be considered either.

2 Proposed assessment approach

The *Risk management plan* (D6c) describes the risks that each task may run. These risks need to be taken in consideration when evaluating the success of the BETSY objectives. For each task a table is inserted in which figures:

- The name of the task
- The objectives of the task taken over from the *Description of Work (DOW)*
- The success criteria which determine whether the objective is successfully met
- The impact of failing to meet/ partially meeting the objectives.

Accordingly, the success of the BETSY tasks can be evaluated. The success of a given BETSY task can be expressed as follows:

- When the task has reached its objectives, the task is completely successful
- The objectives have not been reached but the fall-back scenario has been implemented. The task is considered partly successful.
- The objectives have not been reached and none of the fall-back scenarios were realized. In that case the task is considered failed.

The evaluation process is targeted to the individual tasks. The overall project success depends on the sum of all successes and failures. Contributions of the success of the results of wp2 and wp3 are as important as wp4. The success of the test-bed and prototype should not be equated to the technical success of the BETSY project. A rough indication might be that wp2 and wp3 cover 70% of the technical success of BETSY and wp4 the remaining 30%.

This evaluation only tells something about the technical success criteria according to the DOW. For the final evaluation other factors come into play like the timeliness of the delivery of the deliverables, the dissemination success, the contribution to standards and the impact on the scientific and industrial communities. All these aspects are not covered by this document.

This document complements the deliverable D1d *Evaluation Criteria*. This document looks at the individual objectives of the tasks without taking into account their mutual relations. The deliverable D1d describes the higher-level success criteria of the technology developed during the BETSY project.

3 Assessment of the BETSY objectives

This section lists the final objectives and their success criteria for the BETSY tasks. All information is stored in tables for readability and to assure a uniform presentation.

At the end of the BETSY project, these tables will be used for the self-assessment together with the evaluation criteria described in Deliverable D1d.

3.1 Task 2.1 Identification of end-to-end delivery chain parameters

BETSY Task	Task 2.1 Identification of end-to-end delivery chain parameters	
Task objective(s)	<ul style="list-style-type: none"> • Identification of components that contribute to end-to-end delivery chain. • Identification of attributes of these components, which describe the end-to-end delivery chain fully such that all relevant timing aspects can be quantified. • Relation between resource model and identified components is expressed 	
Success criteria	<ul style="list-style-type: none"> • All end-to-end delivery chain timing aspects can be described with the components and their attributes. • There is an unambiguous relation between the identified components and the resource models of wp3. 	
Task objective reached?	Yes: Complete success	
	No: fallback solution(s) implemented?	Yes: Impact depends on the contribution of the missing components and how limited the number of variable parameters has to become to make the relationship between the components and resource models unambiguous.
		No: Failure

3.2 Task 2.2 Stream description methods

BETSY Task	Task 2.2 Stream description methods	
Task objective(s)	<ul style="list-style-type: none"> • Identification of relevant stream characteristics to put constraints on function parameters • Identification of relevant stream characteristics to efficiently allocate resources • Identify relevant semantic dependencies between video frames 	
Success criteria	<ul style="list-style-type: none"> • All stream characteristics that influence resource allocation and the parameter settings of the functions have been identified. 	
Task objective reached?	Yes: Complete success	
	No: fallback solution(s) implemented?	Yes: Impact depends on the contribution of the missing stream characteristics.
	No: Failure	

3.3 Task 2.3 Decomposition of end-to-end constraints

BETSY Task	Task 2.3 Decomposition of end-to-end constraints	
Task objective(s)	<ul style="list-style-type: none"> • Construction of complete end-to-end chains composed of analyzable components • Identification of composable attributes of components such that attribute value of all components yield attribute values for whole chain • Method to calculate end-to-end timing characteristics over whole chain. 	
Success criteria	<ul style="list-style-type: none"> • End-to-end constraint calculation can be decomposed into tractable constraint calculations on components. • From calculation of constraints of components the end-to end constraints can be calculated. 	
Task objective reached?	Yes: Complete success	
	No: fallback solution(s) implemented?	Yes: This constitutes a major problem, and should be accompanied by an impossibility proof.
	No: Failure	

3.4 Task 3.1 Parameter and resource requirements

BETSY Task	Task 3.1 Parameter and resource requirements	
Task objective(s)	<ul style="list-style-type: none"> • definition of equivalence classes of resources defined in WP1 • mapping of abstract resources to concrete resources 	
Success criteria	<ul style="list-style-type: none"> • mapping and definition is complete and validated 	
Task objective reached?	Yes: Complete success	
	No: fallback solution(s) implemented?	Yes: Impact of not being able to define equivalence classes depends on the importance of the missing classes. Impact of not being able to map would be huge and need impossibility proof, otherwise the success depends on how close the mapping can come to optimal use of concrete resources
	No: Failure	

3.5 Task 3.2 Device functional component modeling

BETSY Task	Task 3.2 Device functional component modeling	
Task objective(s)	<ul style="list-style-type: none"> • Determine abstract resource requirements of components in chain • Mapping of abstract resource requirements to physical resources 	
Success criteria	<ul style="list-style-type: none"> • All abstract and concrete resources are identified and validated • A complete and validated mapping exists 	
Task objective reached?	Yes: Complete success	
	No: fallback solution(s) implemented?	Yes: Impact depends on the contribution of the missing resources (abstract and concrete)
	No: Failure	

3.6 Task 3.3 Network functional component modeling

BETSY Task	Task 3.3 Network functional component modeling	
Task objective(s)	<ul style="list-style-type: none"> • identification of devices in the network • mapping of physical devices to physical resources 	
Success criteria	<ul style="list-style-type: none"> • Representative devices are identified • A complete and validated mapping exists 	
Task objective reached?	Yes: Complete success	
	No: fallback solution(s) implemented?	Yes: Number of expressible scenarios determines measure of success.
	No: Failure	

3.7 Task 3.4 Model composition

BETSY Task	Task 3.4 Model composition	
Task objective(s)	<ul style="list-style-type: none"> Compose the resource requirements of the components, as done in task 3.3 and 3.2, to the resource requirements of complete devices and of a complete end-to-en video streaming chain. 	
Success criteria	<ul style="list-style-type: none"> A complete and validated set of composition rules 	
Task objective reached?	Yes: Complete success	
	No: fallback solution(s) implemented?	Yes: Number of realizable scenarios is a measure of success.
	No: Failure	

3.8 Task 4.1 Specification of the BETSY Framework

BETSY Task	Task 4.1 Specification of the BETSY Framework	
Task objective(s)	<ul style="list-style-type: none"> • Architecture of WP1 and model of WP2 are combined to a software framework • Interfaces of the components in the framework are specified • Interfaces are refined to support the timeliness and resource parameters of the components 	
Success criteria	<ul style="list-style-type: none"> • The required optimized settings defined in wp2 and wp3 can be realized with software framework. • Timeliness settings can be set and evaluated. 	
Task objective reached?	Yes: Complete success	
	No: fallback solution(s) implemented?	Yes: Success is quantifiable by the number of resources parameters that can be supported and the percentage of execution time and resource overhead introduced (goal could be less than 5%).
	No: failure	

3.9 Task 4.2 Design of the test bed and the demonstrator

BETSY Task	Task 4.2 Design of the test bed and the demonstrator	
Task objective(s)	<ul style="list-style-type: none"> • A set of components is identified to be realized on selected hardware and software platforms • The components realizable on an existing platform and the components realized by simulation are defined • Definition of quantifiable metrics for both simulation and selected platform 	
Success criteria	<ul style="list-style-type: none"> • The design must allow for real implementations as well as simulations of components in the test bed. • Provision of interfaces that promote the integration of components in a system and access to real components from application. • At least one complete processing chain (sender, receiver) of one scenario can be implemented on the demonstrator. • The design of the test bed provides for the implementation of additional scenarios, by using simulated components (e.g. WLAN network simulator). 	
Task objective reached?	Yes: Complete success	
	No: fallback solution(s) implemented?	Yes: A demonstrator design can be restricted to a single scenario with single sender /single receiver and measurable energy and bandwidth consumption The test bed can only measure the power consumption of selected components or their simulation.
	No: failure	

3.10 Task 4.3 Reference implementation of the BETSY framework on the test-bed

BETSY Task	Task 4.3 Reference implementation of the BETSY framework on the test-bed	
Task objective(s)	<ul style="list-style-type: none"> • implementation of task 4.2 design on test-bed 	
Success criteria	<ul style="list-style-type: none"> • Possibility to measure and set timeliness aspects and resource aspects for a single sender and single receiver interconnected by a network connection. • Possibility to simulate one scenario. 	
Task objective reached?	Yes: Complete success	
	No: fallback solution(s) implemented?	Yes: single sender and single receiver demonstration constitutes 80% success. Some resources might not be measurable, like power consumption of functional components on test-bed. Therefore, estimations and simulations must be used. Simulation of part of scenario constitutes 15% success.
	No: failure	

3.11 Task 4.4 Porting of reference implementation on the demonstrator set-up

BETSY Task	Task 4.4 Porting of reference implementation on the demonstrator set-up	
Task objective(s)	<ul style="list-style-type: none"> • execution of task 4.2 and task 4.3 implementation 	
Success criteria	<ul style="list-style-type: none"> • Possibility to measure and set timeliness aspects and resource aspects for a single sender and single receiver interconnected by a network connection. • Possibility to simulate one scenario. 	
Task objective reached?	Yes: Complete success	
	No: fallback solution(s) implemented?	Yes: single sender and single receiver demonstration constitutes 80% success. Simulation of part of scenario constitutes 15% success.
		No: failure

3.12 Task 4.5 Validation and evaluation of BETSY models

BETSY Task	Task 4.5 Validation and evaluation of BETSY models	
Task objective(s)	<ul style="list-style-type: none"> • Show that design can handle fluctuations in network conditions • Show that design can accommodate the sharing of resources by applications • Show that design can effectively trade battery consumption against quality 	
Success criteria	<ul style="list-style-type: none"> • All three objectives are met 	
Task objective reached?	Yes: Complete success	
	No: fallback solution(s) implemented?	Yes: 33% per met objective
	No: failure	

4 Conclusions

A set of criteria has been defined for the acceptance of the technology developed during the BETSY project.

References

- [1] Harmke de Groot et al., *Deliverable D1a, Scenario and use cases*, BETSY project, November 2004
- [2] BETSY project, *Description of Work*, May 2004