Co-funded by the European Commission within the Seventh Framework Programme

Project no. 318521

HARNESS

Specific Targeted Research Project
HARDWARE- AND NETWORK-ENHANCED SOFTWARE SYSTEMS FOR CLOUD COMPUTING

Setup of Integration, Testing and Demonstration Testbed

D7.1

Due date: 30 September 2013
Submission date: 30 October 2013

Start date of project: 1 October 2012

Document type: Deliverable
Activity: RTD
Work package: WP7

Editor: John McGlone (SAP)

Contributing partners: All

Reviewers: Gabriel Figueiredo, Alexandros Koliousis

Dissemination Level

PU	Public	√
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Revision history:

Version	Date	Authors	Institution	Description
0.1	2013/06/14	John McGlone	SAP	Outline
0.2	2013/07/26	John McGlone, Gabriel Figueiredo, Katerina Argyraki, Guillaume Pierre, Thorsten Scheutt	SAP, IMP, EPFL, UR1, ZIB	Added details on inventory and testbed usage
0.3	2013/08/21	John McGlone	SAP	Review and edit
0.4	2013/09/3	John McGlone, Gabriel Figueiredo, Katerina Argyraki, Guillaume Pierre, Thorsten Scheutt	SAP, IMP, EPFL, UR1, ZIB	Updates to testbed usage
0.5	2013/09/02	John McGlone	SAP	Add acronyms, review and edit
0.6	2013/09/02	Gabriel Figueiredo	IMP	Review
0.7	2013/09/02	John McGlone	SAP	Edits
0.8	2013/09/06	Alexandros Koliousis	IMP	Review
0.9	2013/09/02	John McGlone	SAP	Final Edits
1.0	2013/09/30	Alexander Wolf	IMP	Final review and edits by Coordinator

Tasks related to this deliverable:

Task No.	Task description	Partners involved°
T7.1	Establish and maintain integrated source-code management.	IMP*, All
T7.2	Establish and maintain testbed and demonstration environment.	SAP*, ZIB, MAX

°This task list may not be equivalent to the list of partners contributing as authors to the deliverable
*Task leader

Executive Summary

This deliverable report presents an overview of the integration, testing and demonstration environment currently supporting the *Hardware- and Network-Enhanced Software Systems for Cloud Computing* (HARNESS) technical work packages (WP3–WP6).

The report covers Tasks T7.1 and T7.2, providing details on the establishment and maintenance of the integrated source-code management used within HARNESS. It is organised into two chapters. Chapter 1 provides an inventory of hardware resources utilised by the project on a per partner basis. Chapter 2 gives a high-level overview of how current hardware is being used by each partner, also reporting on future plans for the current testbeds as the project moves beyond Year 1 and towards an integrated HARNESS platform, addressing the initial requirements of Task T7.2.

This deliverable will be followed by Deliverable D7.2.1, which will be submitted in M20 with details of the initial integrated HARNESS platform and validation use cases.

Contents

Acronyms

ASIC *application-specific integrated circuit*. 3

ConPaaS *Contrail platform-as-a-service*. 10

CPU *central processing unit*. 1, 3, 4, 7, 9

DFE *dataflow engine*. 9

EPL *École Polytechnique Fédérale de Lausanne*. 3

FPGA *field-programmable gate array*. 1, 8, 9

Gbps *gigabits per seconds*. 1

GPGPU *general-purpose graphics processing unit*. 1, 8, 9

HARNESS *Hardware- and Network-Enhanced Software Systems for Cloud Computing*. i, 1, 3, 6–10

IMP *Imperial College London*. 1, 2, 9

MAX *Maxeler Techologies*. 7

NaaS *network-as-a-service*. 1

PCIe *peripheral component interconnect express*. 8

RTM *reverse time migration*. 9, 10

SATA *serial advanced technology attachment*. 10

SDN *software-defined networking*. 9

SSD *solid-state drive*. 10

UR1 *Université de Rennes 1*. 4

ZIB *Konrad-Zuse-Zentrum für Informationstechnik Berlin*. 6, 10

1 Hardware Inventory

This chapter presents an inventory of all hardware used in the *Hardware- and Network-Enhanced Software Systems for Cloud Computing* (HARNESS) project on a per-partner basis.

1.1 Imperial College London

Imperial College London (IMP) is currently using two hardware setups for HARNESS, the Custom Computing Group testbed and the Large-Scale Distributed Systems Group testbed. The former is used to investigate how applications can make effective use of heterogeneous computational resources, while the latter supports the development of the *network-as-a-service* (NaaS) framework, a programming model for the HARNESS communication layer.

The Large-Scale Distributed System testbed, shown in Table 1.1, is composed of 32 nodes and 164 *central processing unit* (CPU) cores. Each node supports two to six Gigabit Ethernet ports with bandwidth ranging from 1 *gigabits per seconds* (Gbps) to 10 Gbps.

Table 1.2 presents the available hardware resources at the Custom Computing Group testbed. It currently harbours 10 nodes, 96 CPU cores and over 10TB of disk space. On top of these general-purpose resources, the testbed also has 4 *general-purpose graphics processing unit*s (GPGPUs), including the NVIDIA Fermi architecture, and 26 *field-programmable gate array* (FPGA) accelerators, supporting the Xilinx Virtex-6 FPGA architecture. All nodes are connected through a dedicated Gigabit Ethernet switch.

Node	**CPU**	**# of Cores**	**Clock rate**	**RAM**	**Ports**
wombat01–10	Opteron 2346 HE	4	1.8Ghz	4GB	4 × 1Gbps
wombat11–15	Xeon X3430	4	2.4Ghz	4GB	2 × 1GB
wombat16–29	Xeon E3-1240	4	3.3Ghz	8GB	2 × 1GB
koala1–2	Xeon E5-2690	16	2.9Ghz	32GB	4 × 1GB + 1 × 10GB
possum	Opteron 6136	16	2.4Ghz	16GB	4 × 1GB + 2 × 10GB

Table 1.1: The Large-Scale Distributed Systems Group testbed at IMP.

Node	CPU	RAM	HDD/SSD	GPU	FPGA
cccad1	Dual Intel Xeon X5650 @2.67GHz (6-core per CPU)	120 DDR3@1333MHz	1.1TB LVM	-	-
cccad2	Dual Intel Xeon X5650 @2.67GHz (6-core per CPU)	96GB DDR3@1333MHz	1.6TB RAID10	-	-
cccad3	Dual Intel Xeon E5-2650 @2.00GHz (8-core per CPU)	256GB DDR3@1600MHz	2.7TB RAID6	-	-
axel01	Dual Intel Xeon E5420 @2.50GHz (4-core per CPU)	16GB DDR2@1066MHz	250GB	NVIDIA GeForce GTX 480	6 × Alpha-Data ADM-XRC-5T2
axel04	Intel Core i7 950 @3.07GHz (4-core)	12GB DDR3@1600MHz	1TB	NVIDIA Tesla C2070	-
maxnode1–2	Dual Intel Core Xeon X5650 @2.67GHz (6-core per CPU)	48GB DDR3@1600MHz	1TB LVM	-	4 × Maxeler MAX3 (V6-SXT475)
maxstation1–2	Intel Core i7 870 @ 2.93GHz (4-core)	16GB DDR3@1600MHz	500GB LVM	-	Maxeler MAX3 (V6-SXT475)
pictor	Intel Core i7 950 @3.07GHz (4-core)	12GB DDR3@1600MHz	1TB	NVIDIA GeForce GTX 580	4 × Alpha-Data ADM-XRC-6T1

Table 1.2: The Custom Computing Research Group testbed at IMP.

1.2 École Polytechnique Fédérale de Lausanne

Table 1.3 presents the hardware resources used by *École Polytechnique Fédérale de Lausanne* (EPL). These resources are used for developing the network layer of the HARNESS platform, hence we report the relevant properties: their hardware (*application-specific integrated circuit* (ASIC)) and/or software (CPUs) processing power, as well as the number and type of network ports.

Node	**CPUs or ASICs**	**RAM**	**Ports**
1 × Broadcom switch	1 × Trident+ ASIC	-	16 × 40Gbps
4 × 16-core Intel servers	4 × 4-core Xeon CPUs @ 2.9GHz	32GB	8 × 10Gbps
1 × 24-core Intel servers	4 × 6-core Xeon CPUs @ 2.9GHz	768GB	-

Table 1.3: EPL hardware inventory.

1.3 Université de Rennes 1

The main experimental facility at *Université de Rennes 1* (UR1) is the Grid'5000 testbed.[1] As of August 2013, it consists of 1171 nodes totalling 2218 CPUs (7896 CPU cores), distributed across 10 locations in France. Its current hardware specification is listed in Table 1.4.

[1] https://www.grid5000.fr/

Location	Node	CPU	RAM	HDD	GPU
Bordeaux	51x Dell PowerEdge 1855	Dual Intel Xeon EM64T	2GB	70GB	
Bordeaux	93x IBM System x3455	Dual AMD Opteron 2218	4GB	80GB	
Bordeaux	10x IBM System x3755	Quad AMD Opteron 2218	32GB	600GB	
Grenoble	10x Bullx R422-E2	Dual Intel Xeon E5520	24GB	250GB	Nvidia Tesla S1070
Grenoble	72x Bullx Blade B500	Dual Intel Xeon E5520	24GB	60GB	
Grenoble	34x Bull R422-E1	Dual Intel Xeon E5420 QC	8GB	160GB	
Lille	20x Dell PowerEdge R410	Dual Intel Xeon E5620	16GB	2x150GB	
Lille	8x Dell PowerEdge C6100	Dual Intel Xeon E5620	8GB	300GB	4x M2050
Lille	46x SGI Altix Xe 310	Dual Intel Xeon E5440 QC	8GB	250GB	
Lyon	79x Sun Fire V20z	Dual AMD Opteron 250	2GB	73GB	
Lyon	16x Dell PowerEdge R720	Dual Intel Xeon E5-2630	32GB	598GB	
Lyon	4x Dell PowerEdge R720	Dual Intel Xeon E5-2630	32GB	598GB	Nvidia Tesla M2075
Lyon	4x Dell PowerEdge C6220	Dual Intel Xeon E5-2620	32GB	6TB	
Nancy	144x Carri System CS-5393B	Intel Xeon X3440	16GB	320GB	
Nancy	92x Carri System CS-5393B	Dual Intel Xeon L5420	16GB	320GB	
Rennes	40x HP Proliant DL165 G7	Dual AMD Opteron(tm) 6164 HE	48GB	250GB	
Rennes	25x SUN FIRE X2270	Dual Intel Xeon X5570	24GB	500GB	
Rennes	64x Carri System CS-5393B	Dual Intel Xeon L5420	32GB	320GB	
Sophia	56x Sun Fire X4100	Dual AMD Opteron 275	4GB	2x73GB	
Sophia	50x Sun Fire X2200 M2	Dual AMD Opteron 2218	4GB	250GB	
Sophia	45x Dell R410	Dual Intel Xeon E5520	32GB	2x300GB	
Toulouse	140x Sun Fire X2200 M2	Dual AMD Opteron 2218	8GB	250GB	

Table 1.4: The Grid'5000 hardware inventory.

1.4 Konrad-Zuse-Zentrum für Informationstechnik Berlin

Konrad-Zuse-Zentrum für Informationstechnik Berlin (ZIB) uses the `cumulus` cluster for HARNESS. It is the main resource for cloud research at ZIB and it currently consists of three different types of nodes, listed in Table 1.5.

Node	CPU	RAM	HDD/SSD	Other
3 × Dell	Dual Intel E5-2430	32GB	2xSSD, 6xHDD	10 GigE
1 × SUN			12x1TB	10 GigE
32 × SUN	Dual AMD Opteron 2376	8GB	1TB HDD	1 GigE

Table 1.5: ZIB's hardware inventory.

1.5 Maxeler Technologies

Maxeler Techologies (MAX) utilises the MPC-X test cluster for HARNESS. This cluster currently consists of 5 MPC-X Series Maxeler data-flow nodes and 3 Intel x86 CPU nodes. The two types of nodes are connected by an Infiniband network. Table 1.6 details their specifications.

Node	**CPU**	**RAM**	**HDD/SSD**	**DFE**	**Other**
MPC-X Series node	-	192-384GB (24-48GB per DFE)	-	8	Infiniband
CPU node	Dual Intel X5650	192GB	500GB	-	Infiniband

Table 1.6: MAX's MPC-X test cluster.

1.6 SAP

SAP Belfast have two main systems that are being used for the HARNESS project: the first one is a development machine that has a *peripheral component interconnect express* (PCIe)-based FPGA card, while the second has two PCIe-based Intel Phi accelerator cards and an NVIDIA K20x GPGPU. The specifications of the nodes are detailed in Table 1.7.

Node	**CPU**	**RAM**	**HDD/SSD**	**GPU**	**FPGA**	**Other**
D20	Dual Intel Xeon X5650 (6-core per CPU) @ 2.67GHz	96GB DDR3 @ 1333MHz (ECC)	700GB	n/a	Alpha-Data ADM XRC-6T1 Virtex 6 SX475T	n/a
Awaiting arrival	2 Xeon E5-2680	256 GB	1TB	Nvidia K20x	n/a	2 × Xeon Phi 5110P

Table 1.7: SAP's test cluster.

2 Testbed Usage and Future Plans

To date, the main technical work packages use distinct testbeds. This is in line with current work plans but, as the project moves towards the integrated HARNESS platform, an integrated testbed will be established to demonstrate it.

To investigate how to make effective use of heterogeneous computational resources available in a cloud data centre, and to understand how to exploit and coordinate a wide range of processing elements, including *dataflow engine*s (DFEs), FPGAs, GPGPUs and multi-core CPUs, there are two testbeds employed: the custom computing cluster (Table 1.2) and the SAP Belfast development system machines (Table 1.7).

Two of the HARNESS validation use cases, *reverse time migration* (RTM) and AdPredictor, have been evaluated with individual Maxeler nodes in the custom computing cluster, whereas the delta-merge use case has been developed and evaluated on the SAP development machines. In addition, dedicated nodes from IMP's Custom Computing Group cluster, such as the *cccad1* node (Table 1.2), have been used for hardware synthesis and design-space exploration experiments. In this context, the plan for Year 2 is: (i) to support vertical and horizontal scaling across multiple heterogeneous nodes in the Custom Computing Group cluster using the three HARNESS validation use cases [1] and (ii) to investigate new types of accelerators such as the Intel Phi currently available in the SAP testbed. For the third year, we plan to evaluate the full HARNESS platform on top of IMP's Custom Computing Group and Large-Scale Distributed System Group testbeds.

To address the problem of heterogeneous communication within HARNESS, we are currently using two testbeds:

- A 24-core Intel server is used to run a verification tool, which takes as input traffic-processing software and proves that it satisfies key properties (e.g., it does not crash, or it introduces bounded per-packet latency). This tool is being designed and built in order to ensure that heterogeneous network devices performing traffic processing in software do not destabilise network operation. This is necessary as, like all software, traffic-processing software is prone to behaviour and performance bugs.

- A Broadcom switch and a 16-core Intel server are used as components of a hybrid traffic-processing platform combining tightly coupled hardware and software pipelines. This is a new, heterogeneous network device, that can perform sophisticated traffic processing in software, while offloading simple traffic-processing tasks to specialised hardware (i.e., it aims to combine the flexibility of a software pipeline with the performance of a hardware pipeline). This as a prototypical example of a heterogeneous network device expected in future data centres. Moreover, two 16-core Intel servers are used to send traffic through the hybrid (one as a generator, the other as a sink), while another 16-core Intel server is used as an OpenFlow controller. These are necessary for testing the hybrid and evaluating its performance in a *software-defined networking* (SDN) setting.

For the time being, these two testbeds are developed and used separately, i.e., there are different groups of researchers working on building the verification tool and the hybrid traffic-processing platform. Our ultimate goal, however, is to make these two elements work together, i.e. use the verification tool to prove properties about the software running in the hybrid traffic-processing platform.

Development of the HARNESS cloud platform, which will use *Contrail platform-as-a-service* (ConPaaS) as an initial baseline, will primarily use Grid'5000 as its main development testbed (Table 1.3). This experimental facility offers extensive support for cloud software deployment and testing, while its hardware heterogeneity will also be useful for designing and validating the performance-model generation techniques required for the cloud platform. In a second phase, we will ensure that the HARNESS platform can be easily ported to the other testbeds mentioned in this deliverable.

The *cumulus* cluster at ZIB will be the main testbed for the developments on heterogeneous storage resource management. The cluster provides three nodes with fast *solid-state drive*s (SSDs) and 10 GigE that are especially suited for random access workloads. One node has 12 *serial advanced technology attachment* (SATA) disks and 10 GigE, which can provide high streaming performance. The remaining nodes have older disks and can be either used as clients or as slower storage devices. Primary integration of the cloud platform and storage work will make used of the *cumulus* cluster. ZIB is already running the AdPredictor validation use case on *cumulus*, and the RTM use case is expected to follow soon. Because of the lack of heterogeneous network devices, we expect to use the cluster for compute- and storage-focused use cases.

Work within WP7 will have a focus on producing a properly integrated set of technologies. WP7 also has responsibility for setting up a test and demonstration environment for the HARNESS platform, and creating and managing the validation use cases. For managing the software created within the HARNESS project and to facilitate the natural transition towards an integrated platform, a software repository, open source bug-tracking database, ticketing system and other necessary features tools, including an internal wiki and forums facility, have been established early in Year 1. Initial implementations of validation use cases have been completed with multiple implementations across multiple technologies, and are available to the consortium via this software repository.

Bibliography

[1] FP7 HARNESS Consortium. Industrial requirements. Project Deliverable D2.2, 2013.