



Vers une définition d'un système réparti multi-échelle

Sam Rottenberg, Sébastien Leriche, Claire Lecocq, Chantal Taconet

► To cite this version:

Sam Rottenberg, Sébastien Leriche, Claire Lecocq, Chantal Taconet. Vers une définition d'un système réparti multi-échelle. Ubimob '12: 8èmes journées francophones Mobilité et Ubiquité, Jun 2012, Anglet, France. pp.178-183. hal-00724426

HAL Id: hal-00724426

<https://hal.science/hal-00724426>

Submitted on 21 Aug 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Vers une définition d'un système réparti multi-échelle

Sam Rottenberg, Sebastien Leriche, Claire Lecocq, Chantal Taconet
Institut Télécom, Télécom SudParis, UMR CNRS 5157 SAMOVAR

9 rue Charles Fourier
Evry ; F-91011 ; France

Prenom.Nom@telecom-sudparis.eu

RÉSUMÉ

Les systèmes ambiants, d'une part, et l'informatique en nuage ou « cloud », d'autre part, sont au cœur de nombreux travaux de recherche récents. Les contributions dans chacun des domaines ont atteint un bon niveau de maturité. De nouveaux travaux identifient maintenant le besoin de faire coopérer ces deux types de systèmes pour créer des systèmes répartis sur plusieurs niveaux, aussi appelés multi-échelle. Mais le terme multi-échelle est employé pour décrire des situations de répartition variées et ne possède pas encore de définition stable et reconnue. Ces différentes situations laissent apparaître des problématiques spécifiques qui ne sont pas encore résolues.

Dans cet article, notre objectif est de proposer une définition du terme multi-échelle dans le cadre des systèmes répartis. Nous commençons par étudier différents scénarios dits multi-échelle afin d'identifier des caractéristiques communes et spécifiques à ce type de situations. À partir de ces observations nous proposons une définition homogène du concept de systèmes répartis multi-échelle. Enfin nous identifions les différentes problématiques liées à la conception et au développement de tels systèmes.

Ce travail est préliminaire à plusieurs travaux de recherches auxquels nous participons, dont les objectifs sont la conception de solutions intergicielles dans un contexte de répartition multi-échelle.

Mots-clés

Multi-échelle, Systèmes répartis, Intergiciels

1. INTRODUCTION

Les systèmes ambiants sont au cœur de nombreux travaux de recherche récents et leurs domaines d'application s'étendent de la domotique au tourisme en passant par la santé ou le commerce [5]. Ces systèmes ont pour but de transformer l'environnement proche de l'utilisateur en un environnement

intelligent et interactif qui l'informe, le guide, veille à sa santé, ou plus généralement améliore sa qualité de vie.

Parallèlement aux systèmes ambiants, l'informatique en nuage ou « cloud computing » [1] fait également l'objet de nombreux travaux de recherche récents. Ces infrastructures permettent notamment d'exposer des services complexes qui nécessitent une très forte puissance de calcul. De plus, ils offrent une très grande capacité de stockage et peuvent ainsi former des bases de connaissances très riches.

Nous remarquons que ces deux domaines de recherche ont atteint un niveau de maturité suffisant. Ainsi plusieurs travaux de recherches récents [7, 2, 10] envisagent de les faire coopérer pour permettre à chacun des domaines d'exploiter les capacités de l'autre. En effet, les systèmes ambiants étaient, jusqu'à présent, conçus pour être déployés dans un périmètre restreint (par exemple une maison, un musée, un magasin). Ils sont ainsi limités en terme de puissance de calcul et n'ont pas accès à la richesse des bases de connaissances du « cloud ». Inversement, les systèmes du « cloud » peuvent avoir besoin d'informations provenant de l'environnement ambiant de l'utilisateur pour améliorer leurs services et enrichir leurs bases de connaissances. Ces deux systèmes constituent les deux extrêmes dans les systèmes informatiques. Leur coopération passe parfois par des niveaux médians, comme les équipements mobiles ou les serveurs de proximité.

Nous avons constaté l'apparition de systèmes répartis sur plusieurs niveaux dans plusieurs travaux de recherche récents. Certains auteurs désignent ces systèmes comme multi-échelle. Pourtant le terme multi-échelle n'a jamais été clairement défini dans le cadre des systèmes répartis et peut-être utilisé pour désigner des systèmes très différents. De plus, nous pensons que certains systèmes sont en réalité multi-échelle sans avoir été explicitement désignés comme tels. L'objectif de cet article est de proposer une définition du terme multi-échelle dans le cadre des systèmes répartis.

La suite de cet article est organisé de la manière suivante. Dans la section 2, nous étudions plusieurs systèmes et cas d'utilisations ayant pour point commun de faire coopérer des systèmes de niveaux différents. En section 3, nous identifions les caractéristiques communes des différents travaux étudiés et proposons une définition concise du concept de système réparti multi-échelle. Puis, dans la section 4, nous identifions les différents verrous scientifiques liés à la conception et au

développement de solutions intergicielles pour de tels systèmes. Enfin, la section 5, conclue notre article et présente nos projets en cours sur le domaine du multi-échelle.

2. QUELQUES SYSTÈMES ET CAS D'UTILISATIONS MULTI-ÉCHELLE

Dans cette partie, nous étudions des travaux de recherche récents dans le domaine des systèmes répartis. Ces travaux ont pour point commun de faire coopérer des systèmes de niveaux différents. Certains systèmes ou cas d'utilisations présentés sont explicitement désignés comme multi-échelle. D'autres ne le sont pas, mais en possèdent les caractéristiques. Les sections suivantes présentent une sélection de travaux sur des systèmes multi-échelle : de surveillance de crue en 2.1, de gestion de ressources en 2.2, de gestion de contexte multi-échelle en 2.3 et enfin de serveurs de proximité en 2.4.

2.1 Surveillance de crue

Blair et Grace [2] présentent un scénario de surveillance de l'environnement et en particulier de surveillance de crue d'une rivière. En décrivant le scénario, les auteurs utilisent le terme d'« échelle » pour désigner des équipements de tailles différentes. En effet, des capteurs sont utilisés pour surveiller une rivière, les informations obtenues de ces capteurs sont envoyées vers le cloud car le traitement de ces informations nécessite une très grande puissance de calcul. Les informations scientifiques ainsi que les alertes d'urgence sont accessibles depuis les équipements mobiles de différents acteurs (scientifiques, services d'urgence, grand public).

Ce scénario est un exemple proposé par les auteurs pour illustrer les nouveaux défis liés aux systèmes multi-échelle. La principale difficulté technique identifiée par les auteurs est l'extrême hétérogénéité technique présente dans un tel système. Cette hétérogénéité en terme d'APIs, de systèmes d'exploitations (pour les smartphones), ou de paradigmes de programmation est telle que l'idée d'une solution standard est écartée par les auteurs. D'après eux, toute solution standard sera rapidement obsolète ou incompatible avec la complexité liée à une situation multi-échelle. Ils proposent donc le développement d'intergiciels « émergents » qui seraient générés à l'exécution pour former un pont entre les différentes parties de l'application.

2.2 Gestion de ressources en contexte multi-échelle

Kessiss et al. [4] proposent un intergiciel de gestion de ressources dans un contexte multi-échelle. Le but de ce système est de proposer une gestion flexible d'un ensemble de ressources réparties de façon multi-échelle. Dans cet article, le terme multi-échelle est employé pour décrire la répartition des ressources à travers des réseaux de tailles différentes : PAN (Personal Area Network), LAN (Local Area Network), WAN (Wide Area Network).

La solution proposée consiste à regrouper les ressources en domaines et fédérations de domaines afin d'implémenter différents modèles de gestion. Il est également possible de passer d'un modèle de gestion à un autre au cours de l'exécution.

Une représentation architecturale est extraite des ressources gérées afin de créer une couche d'abstraction homogène qui représente des ressources hétérogènes. Cette couche est composée d'éléments basiques qui peuvent être assemblés pour former des éléments composites. Les éléments basiques et composites sont tous considérés comme des éléments gérés de façon homogène du point de vue de l'application.

À partir de ces deux principes de gestion (domaines et éléments), l'intergiciel DASIMA a été développé. Il implémente une gestion de ressources flexible car il permet aux utilisateurs de changer le modèle de gestion mis en place et la granularité de l'application de gestion au cours de l'exécution.

2.3 Internet des objets et gestion de contexte

Le projet ANR INCOME [3] a pour objectif de proposer une infrastructure de gestion de contexte Multi-Échelle pour l'Internet des Objets. Le concept même d'Internet des objets possède certains aspects multi-échelle. Ce concept fait intervenir des équipements de très petites tailles (capteurs, effecteurs, etc.) qui communiquent entre eux tout en étant répartis sur l'ensemble d'Internet. L'entité logicielle responsable de la collecte, de la gestion (traitement et filtrage) et de la présentation des informations de contexte aux applications s'appelle communément un gestionnaire de contexte. C'est une entité essentielle pour le développement des applications sensibles au contexte grand public. La gestion de contexte est un problème traité généralement dans le cadre de réseaux ambiants. Dans le cadre de réseaux multi-échelles (ambiant, Internet, nuages) et au-dessus de l'Internet des objets, la gestion de contexte devient autrement plus complexe. Elle doit prendre en compte l'hétérogénéité des données, répartir les traitements et les flux d'informations, assurer le passage à l'échelle, gérer des informations de qualité de contexte pour permettre des prises de décision appropriées, respecter la vie privée lors de la transmission des informations de contexte, s'adapter à des environnements dynamiques, ou encore identifier des situations en mixant des données issues de l'Internet des objets et de bases de connaissances.

Le projet INCOME a pour ambition de réaliser de la gestion de contexte multi-échelle dans lequel des traitements seront à la fois réalisés sur des petits équipements au plus près de leur production, mais également sur des clouds pour bénéficier de véritables bases de connaissances et de puissance de calcul. Les consommateurs des informations de contexte produites seront quant à eux répartis sur l'ensemble de l'Internet.

2.4 Serveurs de proximité

Le concept de « cloudlet » est évoqué dans un article de Satyanarayanan [7] et développé plus en détails par Satyanarayanan et al. [9]. Les auteurs expliquent que, d'une part, les équipements mobiles seront toujours limités en terme de puissance de calcul et devront donc toujours déléguer une partie de leurs calculs, et que, d'autre part, le cloud computing n'est pas forcément la meilleure solution car cela pose des problèmes de latence non négligeables.

La solution proposée est la mise en place de serveurs de proximité sans état et connectés à Internet (cloudlets) auxquels les équipements mobiles peuvent se connecter directe-

ment via un réseau Wi-Fi. Ces équipements sont déployés comme des hotspots Wi-Fi dans des cafés, magasins etc. Ils possèdent une forte puissance de calcul et permettent aux équipements mobiles d'externaliser les calculs trop complexes sur une machine proche, ce qui résout les problèmes de latence du cloud computing.

Un scénario proposé est celui de l'aide aux personnes atteintes de la maladie d'Alzheimer grâce à des lunettes à réalité augmentée. Ces lunettes possèdent une caméra pour filmer ce que voit l'utilisateur et des écouteurs pour lui retourner des informations. À partir d'algorithmes comme la reconnaissance faciale ou le traitement d'informations de contexte, il serait possible de dire à l'utilisateur le nom de la personne qui est en face de lui, de lui rappeler de sortir son chien ou encore d'arroser ses plantes. Ce scénario fait intervenir des capteurs pour obtenir des informations de contexte, une base de connaissance distante et des cloudlets afin d'améliorer la performance et de pouvoir rapidement traiter les informations de contexte. Nous sommes donc bien en présence d'un scénario multi-échelle car il fait intervenir des équipements de très petites tailles (capteurs) dépourvus de puissance de calcul, des cloudlets accessibles via un réseau local, et une base de connaissance sur Internet.

Les différents travaux que nous venons de présenter décrivent tous des systèmes multi-échelle. L'étude de ces systèmes et de leurs caractéristiques nous permet de proposer une définition du multi-échelle dans la section suivante.

3. PROPOSITION DE DÉFINITION DU MULTI-ÉCHELLE

Dans cette partie nous effectuons une synthèse des différents systèmes et cas d'utilisation étudiés afin d'en identifier des caractéristiques communes. Puis, nous affinons notre raisonnement en comparant le concept de système multi-échelle avec celui de système grande échelle et nous proposons alors une définition du terme multi-échelle dans le cadre des systèmes répartis.

3.1 Les différentes dimensions du multi-échelle

À partir de ces différents cas d'utilisation, nous pouvons constater que le terme multi-échelle peut avoir plusieurs dimensions. Pour exemple, Kessiss et al. [4] proposent une définition orientée réseau alors que Blair et Grace [2] proposent plutôt une définition orientée équipements.

En plus des dimensions évoquées dans ces articles (réseau et équipement), d'autres dimensions de l'aspect multi-échelle d'un système peuvent être proposées : géographique, administratif, utilisateur, données. Ces différentes dimensions du multi-échelle sont détaillées dans le tableau ci-dessous.

[4] et [2] proposent des définitions du multi-échelle respectivement créées sur les dimensions réseaux et équipement. Le tableau ci-dessus montre que chacune des dimensions que nous avons identifiées fait intervenir des niveaux de *tailles* différentes et peut donc être considérée comme une dimension multi-échelle.

3.2 Systèmes grande échelle et systèmes multi-échelle

Dimension	Taille		
	Petite	Moyenne	Grande
Équipement	capteurs	smartphones, ordinateurs	cloud
Réseau	réseaux personnels	réseaux locaux	Internet
Géographique	interactions proches	interactions distantes	interactions très distantes
Administratif	individu	Intranet	Internet
Utilisateur	individu	groupes	communautés
Données	données de capteur	bases de données	bases de connaissances

Dans le but de proposer une définition d'un système multi-échelle, il est important d'identifier ce qui différencie le concept de répartition multi-échelle du concept de répartition grande échelle.

Le terme système réparti grande échelle est apparu dans les années 1990 [8, 6]. Dans le cadre des systèmes répartis à grande échelle, le terme d'échelle est utilisé dans un sens quantitatif, c'est-à-dire qu'un système grande échelle est fortement réparti en terme de nombre de tiers impliqués, d'éloignement géographique entre les tiers, de temps de communication entre les tiers. De même, le terme « passage à l'échelle » est utilisé pour désigner la capacité d'un système à passer d'une répartition faible (peu de tiers, tiers proches, réseau local) à une répartition plus large (beaucoup de tiers, tiers très éloignés géographiquement, temps de communication important entre les tiers).

Les défis liés aux systèmes grande échelle voire ultra grande échelle sont évoqués par Blair et Grace [2] mais les auteurs précisent que le défi du passage à l'échelle n'est pas le plus difficile à relever. D'après eux, le grand défi actuel réside dans l'hétérogénéité des technologies impliquées.

C'est à notre sens cette notion d'hétérogénéité qui différencie le concept de multi-échelle de celui de grande échelle. En effet, un système multi-échelle n'est pas nécessairement fortement réparti en terme d'éloignement géographique ou de nombre de tiers impliqués. Il est plutôt réparti « verticalement » sur des niveaux d'échelle différents et fait donc intervenir des technologies très hétérogènes entre les différents niveaux.

À partir de ces observations, nous proposons la définition suivante. « *Un système multi-échelle est un système réparti de façon verticale sur plusieurs niveaux de tailles différentes dans une ou plusieurs dimensions (équipement, réseau, géographie, etc.).* »

4. PROBLÉMATIQUES SPÉCIFIQUES AU MULTI-ÉCHELLE

À partir des différents cas d'utilisations étudiés, nous avons identifié plusieurs verrous scientifiques spécifiques aux systèmes multi-échelle.

Le premier verrou est l'extrême hétérogénéité des technologies impliquées dans un système multi-échelle. Ce point est en particulier développé par Blair et Grace [2].

Le deuxième verrou spécifique, décrit par Kessiss et al. [4], est la possibilité de changer d'architecture et de granularité au cours de l'exécution. Les systèmes multi-échelle ont souvent une architecture complexe et instable qui implique un système de gestion flexible et adaptable.

En outre, les situations multi-échelle héritent des verrous technologiques de chacun des niveaux de l'échelle qu'elles contiennent : gestion de la mobilité, des déconnexions, tolérance aux pannes, sécurité, problèmes de latences pour les interactions distantes, le passage à l'échelle. Ces verrous ne sont pas totalement résolus de façon indépendante et sont sans doute encore plus difficiles à résoudre dans un système qui les englobe tous.

Enfin, il nous semble important que les systèmes multi-échelle intègrent une gestion de la confidentialité adaptée. En effet, plus les systèmes sont complexes tout s'adaptant à l'utilisateur et à son contexte, plus ils nécessitent de rassembler des données sur l'utilisateur, ses préférences, ses activités. Ses données sont sensibles. L'utilisateur devrait pouvoir gérer ses données de vie privée de façon également multi-échelle, c'est-à-dire d'avoir la possibilité de rendre une information accessible à certains niveaux de l'échelle uniquement. Ainsi, un besoin de gestion fine de la confidentialité à chaque niveau de l'échelle vient s'ajouter au besoin d'une couche intergicielle qui permette de masquer l'hétérogénéité technique des systèmes multi-échelle.

5. CONCLUSION ET PROJETS EN COURS

Dans cet article, nous avons étudié plusieurs scénarios et cas d'utilisations de systèmes répartis multi-échelle. Nous avons pu en extraire une définition générique qui prend en compte les différentes dimensions de répartition de tels systèmes. Enfin, nous avons identifié les problématiques spécifiques soulevées par le multi-échelle. Ce travail a été réalisé en amont de plusieurs projets en cours dont le projet ANR INCOME. Ces projets ont pour objectif la conception de solutions intergicielles qui répondent aux problématiques du multi-échelle que nous avons identifiées.

Les propositions prendront en compte les particularités des systèmes répartis multi-échelle identifiés, à savoir la gestion de l'hétérogénéité des éléments tant logiciels que matériels, le choix de la bonne échelle pour chacun des traitements, le respect de la vie privée. Enfin, nous aurons un soin particulier concernant l'évaluation de nos propositions intergicielles dans un environnement multi-échelle. En effet, il sera difficile d'évaluer les propositions dans un environnement multi-échelle réel. Des émulations et simulations seront nécessaires dans un premier temps pour leurs évaluation. Ces projets participeront à la construction de systèmes répartis multi-échelle nécessaires aux applications mobiles et ubiquitaires de nouvelle génération.

Remerciements

Ce travail est financé en partie par le projet ANR INCOME (ANR-11-INFR-009, 2012-2015) auquel prennent part les partenaires français suivants : IRIT (Institut de Recherche en Informatique de Toulouse), Télécom SudParis, et ARTAL Technologies. Les auteurs remercient tous les membres qui participent à ce projet.

6. REFERENCES

- [1] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M. Zaharia. A view of cloud computing. *Commun. ACM*, 53(4) :50–58, 2010.
- [2] G. Blair and P. Grace. Emergent middleware : Tackling the interoperability problem. *Internet Computing, IEEE*, 16(1) :78–82, jan.-feb. 2012.
- [3] Infrastructure de gestion de COntexte Multi-Échelle pour l'Internet des Objets. Projet ANR INFRA 2011.
- [4] M. Kessiss, C. Roncancio, and A. Lefebvre. Dasima : A flexible management middleware in multi-scale contexts. In *Information Technology : New Generations, 2009. ITNG '09. Sixth International Conference on*, pages 1390–1396, april 2009.
- [5] F. Sadri. Ambient intelligence : A survey. *ACM Computing Surveys*, 43 :36 :1–36 :66, Oct. 2011.
- [6] H. Sandhu and S. Zhou. Cluster Based File Replication in Large-Scale Distributed Systems. In *Proc. ACM Sigmetrics Performance '92 Conference*, May 1992.
- [7] M. Satyanarayanan. Mobile computing : the next decade. *SIGMOBILE Mobile Computing and Communications Review*, 15 :2–10, August 2011.
- [8] M. Satyanarayanan. Scalable, Secure, and Highly Available Distributed File Access. *IEEE Computer*, May 1990.
- [9] M. Satyanarayanan, P. Bahl, R. Caceres, and N. Davies. The case for vm-based cloudlets in mobile computing. *IEEE Pervasive Computing*, 8 :14–23, October-December 2009.
- [10] M. van Steen, G. Pierre, and S. Voulgaris. Challenges in very large distributed systems. *Journal of Internet Services and Applications*, pages 1–8, 2011. 10.1007/s13174-011-0043-x.