Service Deployment in Programmable Networks

A dissertation submitted to the
SWISS FEDERAL INSTITUTE OF TECHNOLOGY
ZURICH

for the degree of
DOCTOR OF TECHNICAL SCIENCES

presented by
ROBERT HAAS
Ing. dipl. en Systèmes de Communication, EPF Lausanne
born May 29, 1973
citizen of Montreux, Vaud

accepted on the recommendation of
Prof. Dr. B. Stiller, examiner
Prof. Dr. D. Hutchison, co-examiner

2003
Abstract

Networks have experienced a dramatic growth in size and functionality in the past years: network nodes not only forward datagrams using longest-prefix matching of the destination address, but also execute functions based on dynamic policies such as proxy-caching, encryption, tunnelling, and firewalling. More recently, programmable behaviors have begun to appear in network elements, allowing experimentation with even more sophisticated services. The need to accelerate the deployment of new services in networks while taking advantage of the sophisticated capabilities already installed poses a significant challenge to existing network management tools. Complexity of services, number of network nodes, programmability of such nodes, and fast deployment call for an novel approach to deploying services.

This thesis presents an automated approach to network service deployment that scales to large heterogeneous programmable networks. It focuses on three aspects. First, we present a framework that encompasses both network and node levels with hierarchically distributed computations. Second, a number of algorithms used in the vertical dimension of the hierarchy (to disseminate deployment messages and aggregate resulting information) and algorithms used in the horizontal dimension (to select appropriate nodes during deployment) are developed and evaluated. Third, a method to perform the adaptation between programming models used for services and for networking nodes is developed and validated. These three aspects lead to a global view and a top-down approach to our investigation of service deployment. When installing new services, automated service deployment allows the distributed and complex capabilities present in network elements to be leveraged more efficiently than is possible in traditional centralized network management-based approaches: this results in faster installation and optimized usage of available functions.
This thesis first introduces the topological categories of service deployment, namely, continuous and sparse path-based, node-based, fence-based, and combinations thereof. A five-step deployment mechanism of hierarchically distributed computations in the network and node levels is presented that comprises solicitation, summarization, dissemination, installation, and advertisement steps. These five steps perform the evaluation of capabilities of network nodes and the installation of a service in the suitable nodes. These steps take place in the vertical dimension of the service-deployment hierarchy and are formalized using the HIGCS (Hierarchical Iterative Gather-Compute-Scatter) model and simulated over large topologies for three typical categories of services. These categories include services such as quality of service, distributed proxy-caching, and virtual private networks. Novel information aggregation methods such as augmented transition-matrices and function summarization for cost-aware deployment are introduced. An algorithm for the deployment of spare path-based services over hierarchical network is presented. Numerical simulation results demonstrate the scalability of the approach.

In the horizontal dimension, s-straight paths are introduced as a generalization of simple, chordless, and shortest paths. Algorithms are introduced to compute such paths, in particular chordless paths and approximation of nodes that belong to such paths, as a means to restrict the scope of the search when querying nodes in a network. The problem of CHORDLESS PATH VIA NODE is introduced and shown to be NP-complete. Algorithms for fence-based service deployment are introduced and shown to reduce to max-flow min-cut problems through graph transformations. Strategy trees are introduced to optimize the number of queried nodes until a continuous path can be found. A construction for such trees is presented in the case of a tree network, namely a tree from a receiver to replicated content servers. Graph-based simplifications are shown to reduce the number of potential strategies that has to be to be examined.

Finally, at the node-level, a method to perform the on-line evaluation and activation of configurations is presented that performs the adaptation of programming models. A prototype is developed for the adaptation between the Linux traffic control and the IBM PowerNP network processor programming models, and validated using typical Differentiated Services configurations of edge and core routers.
Les réseaux de communication ont récemment évolué de façon significative: au lieu de simplement effectuer le routage de paquets sur la base de l'adresse de destination, les noeuds des réseaux effectuent également des opérations telles que proxy-caches, codage, encapsulation, ou filtrage/pare-feux. Depuis peu, certains éléments de réseau sont même programmables, ce qui permet l'introduction et l'expérimentation de services encore plus variés et sophistiqués. Le besoin d'accélérer le déploiement de nouveaux services dans de tels réseaux tout en tirant parti de toutes les fonctions présentes est un sérieux problème pour les outils actuels de gestion des réseaux. La complexité des services, le nombre grandissant de noeuds programmables dans les réseaux, et la rapidité demandée dans le déploiement exigent une approche nouvelle.

Cette thèse présente une approche automatisée du déploiement de services qui est adaptée aux grands réseaux hétérogènes. Trois aspects sont traités plus particulièrement. Premièrement, une architecture est présentée qui traite le problème au niveau du réseau ainsi que de chacun des noeuds d'une façon distribuée hiérarchique. Deuxièmement, des méthodes et algorithmes sont développés et évalués pour la dimension verticale de la hiérarchie (afin de disséminer les messages servant au déploiement et d'agréger les résultats obtenus) et pour la dimension horizontale (afin de sélectionner les noeuds appropriés durant le déploiement). Troisièmement, une méthode est développée et validée servant à adapter des modèles de programmation utilisés par les services d'un côté et les noeuds du réseau de l'autre. Ensemble, ces trois aspects donnent à notre approche pyramidale une vue d'ensemble des relations entre les différents problèmes. Lors de l'introduction de nouveaux services, le déploiement automatisé permet de mieux profiter des fonctions complexes distribuées dans les noeuds du réseau que ne le permettent les approches traditionnelles centralisées de gestion de réseau. Cela résulte en un déploiement
plus rapide et optimisant l'utilisation des fonctions disponibles.

Différentes catégories de déploiement de services sont introduites : des services continus requis sur tous les noeuds de chemins traversant le réseau, ou discontinus, c.à.d. seulement sur certains noeuds de ces chemins; des services requis sur certains noeuds indépendamment des chemins; des services requis sur des noeuds formant une clôture enfermant certains noeuds; et des combinaisons de toutes ces catégories. Un mécanisme est présenté composé de cinq étapes de calculs hiérarchiquement distribués à travers le réseau et dans chaque noeud. Ces étapes sont la sollicitation, l’agrégation, la dissémination, l’installation, et l’annonce. Ces cinq étapes permettent d’évaluer les capacités des noeuds du réseau et d’installer les services dans les noeuds les mieux adaptés. Ces étapes ont lieu à travers la dimension verticale de la hiérarchie et sont formalisées à l’aide du modèle HIGCS (collecte/calcul/distribution hiérarchiques itératifs). Ces étapes sont simulées sur de grands réseaux pour trois des catégories de services, qui incluent des services tels que la garantie de la qualité de service, les proxy-caches distribués, et les réseaux virtuels privés. Des méthodes originales d’agrégation sont présentées, telles que matrices de transition augmentées et agrégation de fonctions de coût. Un algorithme pour le déploiement de services discontinus dans un réseau hiérarchique est également présenté. La simulation à événements discrets démontre que l’approche proposée fonctionne à large échelle.

Dans la dimension horizontale de la hiérarchie, les chemins s-droits sont introduits comme généralisation des chemins simples, des chemins sans corde, et des plus courts chemins. Des algorithmes sont proposés pour calculer ces chemins, en particulier les chemins sans corde, ce qui permet de diminuer le nombre de noeuds interrogés lors de l’étape de sollicitation. Le problème de CHEMIN SANS CORDE VIA NOEUD est introduit et sa NP-complexité est démontrée. Le problème du déploiement de service en clôture est présenté avec les transformations nécessaires du graphe pour aboutir à un problème connu de flux-maximum coupe-minimum. Les arbres de stratégies sont introduits afin d’optimiser le nombre de noeuds sollicités avant qu’un chemin continu soit trouvé. La construction de ces arbres de stratégies est présentée dans le cas d’un réseau en arbre représentant par exemple l’arbre des chemins d’un client vers des serveurs de contenu répliqué. Les simplifications proposées du graphe de l’arbre de stratégies diminuent sensiblement le nombre de stratégies à examiner.

Finalement, au niveau des noeuds, une méthode est présentée pour évaluer directement et activer des configurations en appliquant l’adaptation
Résumé

des modèles de programmation. Un prototype est développé pour l’adaptation entre les modèles de programmation du contrôle du trafic sous Linux et du processeur réseau IBM PowerNP. Le prototype est validé à l’aide de configurations types de routeurs d’accès et de routeurs internes supportant la différenciation de la qualité de service.