Service Placement in Ad Hoc Networks

(Extended Abstract)

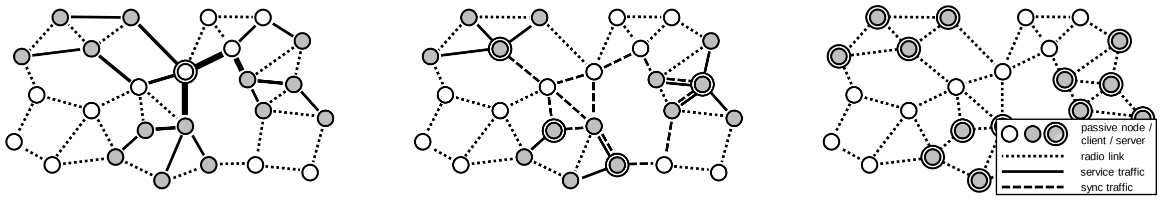
Georg Wittenburg and Jochen Schiller
Department of Mathematics and Computer Science
Freie Universität Berlin
Takustr. 9, 14195 Berlin, Germany
Email: {wittenbu,schiller}@inf.fu-berlin.de

Background Service provisioning in ad hoc networks is challenging given the difficulties of communicating over a wireless channel and the heterogeneity and mobility of devices. In order to optimize the performance of the network as a whole, it is necessary to continuously adapt the logical network topology to both external (e.g., wireless connectivity, mobility, churn) and internal (e.g., communication patterns, service demand) factors. Recent proposals [1, 2] advocate that nodes should dynamically choose which nodes in the network should provide application-level services to other nodes. The process of selecting an optimal service configuration, i.e., a set of nodes to host a service in light of a given service demand and network topology, is referred to as *service placement*.

Service placement addresses the questions of *how many* instances of the same service should be available in the network and cooperate to process clients' service requests; *where* these service instances should be placed, i.e., which nodes are best suited for hosting them; and *when* to adapt the current service configuration. The main benefit of this approach is that the service configuration, i.e., the set of nodes to host a service, is adapted automatically at run-time. A good service configuration reduces overall network traffic and latency, and it can also be used to optimize the network performance according to service-specific metrics.

In our recent work [3], we discussed the general applicability of service placement and gave a survey of current approaches. We found that the service placement problem is either tackled as a byproduct of middleware research or as an application of facility location theory. The interactions between service placement and existing service discovery and routing protocols are widely unexplored, which gives raise to questions regarding optimality, stability, and overhead of current solutions.

Contribution We propose a novel approach to service placement in ad hoc networks that optimizes the number and the location of service instances based on usage statistics and a partial network topology derived from routing information. Our system takes advantage of the interdependencies between service placement, service discovery and the routing of service requests to reduce network overhead. The system only requires minimal knowledge about the service it is tasked with placing in the network, and it is unique in that it explicitly considers the communication between service instances that is required to synchronize shared data.



(a) High sync traffic; one single service instance in the entire network

(b) Medium sync traffic; several instances

(c) No sync traffic; one service instance per client

Figure 1: Service configurations for different levels of synchronization traffic

In our architecture, usage statistics are transmitted periodically to a dynamically-assigned control node. This node calculates the optimal service configuration using a cost metric that is based on service demand in terms of used bandwidth and the network topology in terms of hop count or link quality. With this service configuration as input, the control node establishes a set of actions required to migrate from the current to the optimal configuration. If the combined cost of these actions in terms of network traffic less than the difference in cost between the current and the optimal configuration, the commands for adapting the configuration are issued to the nodes that currently host service instances. These nodes are then in charge of replicating, migrating, or shutting down individual service instances.

A crucial aspect in the calculation of the optimal service configuration is the traffic between service instances. Service instances need to communicate among themselves to keep global state and data synchronized, e.g., in a DNS-like service, updates to a record received by one service instance have to be propagated to all other instances. Figure 1 illustrates the interdependency between synchronization traffic and optimal service configuration. If a high volume of synchronization traffic, as it may be the case for distributed databases that provide transactional semantics, the optimal service configuration is to have only a single instance of the service (cf. Fig. 1a). On the other extreme, if no synchronization traffic is required, e.g., for a spell checking service, each client hosts its own service instance (cf. Fig. 1c). For the more interesting cases of medium to low synchronization traffic (cf. Fig. 1b), the service configuration needs to be calculated using a cost model based on the (partially known) network topology and the service demand.

Results from a preliminary evaluation of our architecture have shown that our approach is particularly effective for services that require mild synchronization traffic between service instances. Hence, it is applicable to wide-spread services such as DNS or WWW and can reduce their respective cost of service provisioning in ad hoc networks.

References

[1] K. Herrmann. *Self-Organizing Infrastructures for Ambient Services.* PhD thesis, Technische Universität Berlin, Berlin, Germany, July 2006.

[2] M. Lipphardt, J. Neumann, S. Groppe, and C. Werner. DySSCo - A Protocol for Dynamic Self-organizing Service Coverage. In *Proc. of IWSOS '08*, Vienna, Austria, Dec. 2008.

[3] G. Wittenburg and J. Schiller. A Survey of Current Directions in Service Placement in Mobile Ad-hoc Networks. In *Proc. of PerCom '08*, Hong Kong, Mar. 2008.