

Wireless Cloud Infrastructure for High-Performance Applications and Services

Xin Wang

Wireless Networking and Systems Lab, Stony Brook University
(xwang@ece.sunysb.edu)

Mobile Cloud Computing and Challenges

Cloud computing can potentially enable performance-hungry applications on almost any mobile devices by executing a resource-intensive application on a high-performance server or server cluster in a remote Cloud. Although promising, there are some potential barriers:

- Lack of mobile Internet access everywhere
- Limited wide-area wireless network bandwidth
- High delay involved in WAN-level Internet transmission
- Cost of commercial cloud

Potential Mobile Cloud Model

Instead of relying on a distant cloud, mobile devices can access nearby resource-rich computers or servers whenever possible while gracefully falling back to cloud resources when needed.

Issues

- Lack of support of concurrent use of local resources (including enterprise resources) and resources from the remote Cloud
- Lack of an efficient service management infrastructure to discover, recruit and coordinate the nodal resources.
- Difficulty in supporting wireless services over higher-speed but shorter-range wireless links, due to the unreliability of the wireless medium, the mobility of nodes, and the resource limitations of wireless devices.
- Need of reliable and high quality data collection and distribution schemes over widely distributed mobile nodes
- Lack of mechanisms to locate and track resources and operational states of a large number of nodes which are possibly mobile and unreliable, enable resource sharing and service collaborations, and maintain service reliability.

Objectives

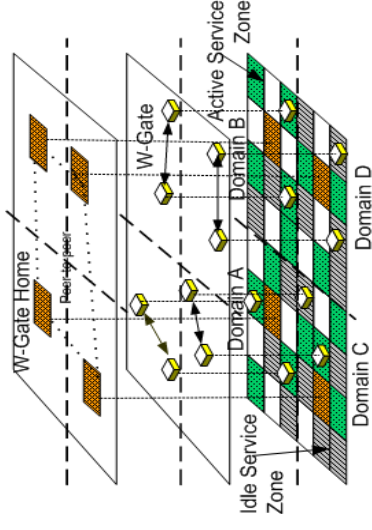
The objective of our work is to develop a scalable and robust service infrastructure to enable seamless access of resources from local cloudlet and remote clouds for high-performance wireless applications and services in presence of node mobility, network and resource dynamics, as well as uncertainty of environment and resource consumption.

Methodologies and Research Directions

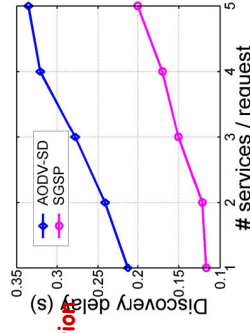
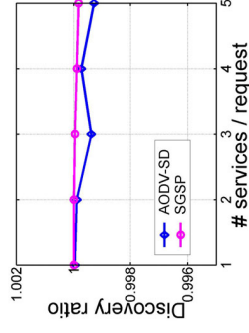
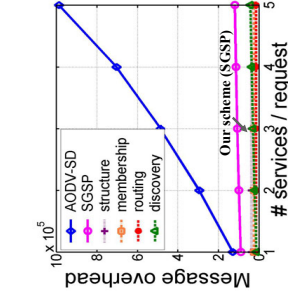
To better handle node mobility and radio channel dynamics, and consequent changes in network topology and resource availability without incurring excessive management overhead, we will integrate location information into the management infrastructure to significantly reduce control overhead and increase reliability through various *stateless virtual infrastructures*. Our research consists of five thrusts:

- (1) *Developing a robust and scalable monitoring framework* to track locations, resources and service availability, and liveness of resource providers and clients in a dynamic wireless environment;
- (2) *Designing an efficient service and resource discovery scheme* to allow the registration of distributed resources, the publication of services, the quick finding of resources from a large pool of servers, and the coordination of a group of servers to support services requested by a client;
- (3) *Developing a scalable, reliable, and flexible information and data management infrastructure* to meet various application needs in presence of wireless channel dynamics and mobility;
- (4) *Design of an efficient task allocation scheme* to approach optimal resource allocation, while ensuring scalability, reliability and stability of resource management in a dynamic system with inaccuracy and uncertainty in modeling of resource consumption and constant changes of resources;
- (5) *Testbed implementation of the proposed service infrastructure*, exploiting cloud tool such as Microsoft Windows Azure platform to enable wireless clients to seamless access resources from local pool and a remote Cloud.

Basic Service Architecture



Preliminary Results on Resource Discovery



Significant overhead reduction in supporting service collaborations