Chapter 9
Decentralized Overlay for Federation of Enterprise Clouds

Rajiv Ranjan
The University of Melbourne, Australia

Rajkumar Buyya
The University of Melbourne, Australia

ABSTRACT
This chapter describes Aneka-Federation, a decentralized and distributed system that combines enterprise Clouds, overlay networking, and structured peer-to-peer techniques to create scalable wide-area networking of compute nodes for high-throughput computing. The Aneka-Federation integrates numerous small scale Aneka Enterprise Cloud services and nodes that are distributed over multiple control and enterprise domains as parts of a single coordinated resource leasing abstraction. The system is designed with the aim of making distributed enterprise Cloud resource integration and application programming flexible, efficient, and scalable. The system is engineered such that it: enables seamless integration of existing Aneka Enterprise Clouds as part of single wide-area resource leasing federation; self-organizes the system components based on a structured peer-to-peer routing methodology; and presents end-users with a distributed application composition environment that can support variety of programming and execution models. This chapter describes the design and implementation of a novel, extensible and decentralized peer-to-peer technique that helps to discover, connect and provision the services of Aneka Enterprise Clouds among the users who can use different programming models to compose their applications. Evaluations of the system with applications that are programmed using the Task and Thread execution models on top of an overlay of Aneka Enterprise Clouds have been described here.

INTRODUCTION
Wide-area overlays of enterprise Grids (Luther, Buyya, Ranjan, & Venugopal, 2005; Andrade, Cirne, Brasileiro, & Roissenberg, 2003; Butt, Zhang, & Hu, 2003; Mason & Kelly, 2005) and Clouds (Amazon
Elastic Compute Cloud, 2008; Google App Engine, 2008; Microsoft Live Mesh, 2008; Buyya, Yeo, Venugopal, 2008) are an appealing platform for the creation of high-throughput computing resource pools and cross-domain virtual organizations. An enterprise Cloud is a type of computing infrastructure that consists of a collection of inter-connected computing nodes, virtualized computers, and software services that are dynamically provisioned among the competing end-user’s applications based on their availability, performance, capability, and Quality of Service (QoS) requirements. Various enterprise Clouds can be pooled together to form a federated infrastructure of resource pools (nodes, services, virtual computers). In a federated organisation: (i) every participant gets access to much larger pools of resources; (ii) the peak-load handling capacity of every enterprise Cloud increases without having the need to maintain or administer any additional computing nodes, services, and storage devices; and (iii) the reliability of a enterprise Cloud is enhanced as a result of multiple redundant clouds that can efficiently tackle disaster condition and ensure business continuity.

Emerging enterprise Cloud applications and the underlying federated hardware infrastructure (Data Centers) are inherently large, with heterogeneous resource types that may exhibit temporal resource conditions. The unique challenges in efficiently managing a federated Cloud computing environment include:

- **Large scale**: composed of distributed components (services, nodes, applications, users, virtualized computers) that combine together to form a massive environment. These days enterprise Clouds consisting of hundreds of thousands of computing nodes are common (Amazon Elastic Compute Cloud, 2008; Google App Engine, 2008; Microsoft Live Mesh, 2008) and hence federating them together leads to a massive scale environment;

- **Resource contention**: driven by the resource demand pattern and a lack of cooperation among end-user’s applications, particular set of resources can get swamped with excessive workload, which significantly undermines the overall utility delivered by the system; and

- **Dynamic**: the components can leave and join the system at will.

The aforementioned characteristics of the infrastructure accounts to significant development, system integration, configuration, and resource management challenges. Further, the end-users follow a variety of programming models to compose their applications. In other words, in order to efficiently harness the computing power of enterprise Cloud infrastructures (Chu, Nandiminti, Jin, Venugopal, & Buyya, 2007; Amazon Elastic Compute Cloud, 2008; Google App Engine, 2008; Microsoft Live Mesh, 2008), software services that can support high level of scalability, robustness, self-organization, and application composition flexibility are required.

This chapter has two objectives. The first is to investigate the challenges as regards to design and development of decentralized, scalable, self-organizing, and federated Cloud computing system. The second is to introduce the Aneka-Federation software system that includes various software services, peer-to-peer resource discovery protocols and resource provisioning methods (Ranjan, 2007; Ranjan, Harwood, & Buyya, 2008) to deal with the challenges in designing decentralized resource management system in a complex, dynamic and heterogeneous enterprise Cloud computing environment. The components of the Aneka-Federation including computing nodes, services, providers and end-users self-organize themselves based on a structured peer-to-peer routing methodology to create a scalable wide-area overlay of enterprise Clouds. In rest of this chapter, the terms Aneka Cloud(s) and Aneka Enterprise Cloud(s) are used interchangeably.