Cloud resource management poses new and extremely challenging problems, so it should be no surprise

that it is a very active area of research. A fair number of papers, including [22,65,75,77,113,114,115,120,

138,139,151,155,162,184,238,275,307,313] are dedicated to various resource management policies.

Several papers are concerned with SLA and QoS; e.g., [4] covers SLA-driven capacity management

and [23] covers SLA-based resource allocation policies. Dynamic request scheduling of applications

subject to SLA requirements is presented in [54]. The QoS in clouds is analyzed in [121].

Autonomic computing [130] is the subject of papers such as [24], which covers energy-aware resource

allocation in autonomic computing; [188], which analyzes policies for autonomic computing based on

utility functions; [187], which discusses coordination of multiple autonomic managers and power performance

tradeoffs; and [9], which presents autonomic management of cloud services subject to

availability guarantees.

Auctions in which participants can bid on combinations of items or packages are called combinatorial

auctions [93]. Such auctions provide a relatively simple, scalable, and tractable solution to cloud resource

allocation. Two recent combinatorial auction algorithms are the simultaneous clock auction [29] and

the clock proxy auction [30]; the algorithm discussed in this chapter and introduced in [333] is called

the ascending clock auction (ASCA).

An authoritative reference on fault tolerance is [31]; applications of control theory to resource allocation

discussed in [70,109] cover resource multiplexing in data centers. Admission control policies are

discussed in [150]. Power and performance management are the subject of [202], and performance management

for cluster-basedWeb services is covered in [280]. Autonomic management of heterogeneous

workloads is discussed in [344], and application placement controllers are the topic of [346].

Scheduling and resource allocation are also covered by numerous papers: a batch queuing system

on clouds with Hadoop and HBase is presented in [387]; data flow-driven scheduling for business

applications is covered in [106]. Scalable thread scheduling is the topic of [374]. Scheduling of

real-time services in cloud computing is presented in [220]. The Open Grid Forum (OGF) Open Cloud Computing Interface (OCCI) is involved in the definition of virtualization formats and APIs for IaaS

[175] presents a performance analyzer.