**CHAPTER 5**

**APPLICATION**

Cloud based applications may be scaled by either or both of two methods: scaling up by procuring a more powerful computational resource, and scaling out by procuring more instances of a computational resource, each of which offer some distinct advantages.

Scaling up is the most common method to improve performance, but is restricted by the capabilities of the most powerful hardware the evolution of hardware performance should also be considered. Migrating existing solutions to more powerful hardware is a well understood problem and is particularly applicable where the task cannot easily be decomposed into smaller units of work. In the SSA example, each worker performs a unit of work that would be difficult to decompose, and satisfactory performance is within the capability of existing hardware. In order to benefit by scaling out, an understanding of the computation is required as the algorithm has to be decomposed to take advantage of parallel operations. Scale-out often requires more development effort than migrating to a scale-up method. Dividing a task across multiple computational resources incurs an overhead, thus limiting the theoretical improvement in performance as more resources are added. In the SSA example, complex units of work consume tasks from a queue, which makes scaling-out easier since the number of workers consuming tasks from a queue can be varied with the length of the queue.

Using Microsoft Windows Azure was advantageous in this example as this is a PaaS, negating the need to maintain, patch and update the underlying OS. The environment also supports queues, various types of storage, including an SQL server and even includes a data market place to monetise datasets. One key advance which emerged during this work is a cloud-based high performance compute (HPC) cluster. Although not incorporated into this example architecture, HPC is a very powerful asset that ensures legacy MPI applications can seamlessly migrate into a cloud-based architecture.

Throughout this example it became obvious that cloud providers are offering powerful, cost effective infrastructures, but harnessing the power and migrating existing applications is often painful and just out of reach of most application scientists. Cloud providers are still evolving their offerings, and as migration scenarios and remote debugging capabilities. Can expect to see scientists consuming more cloud resources.

The cloud-based solution provides additional advantages, including the ability to share data with trusted partners simply, rapidly and securely. The partners, at their option, could then fund additional compute resources located close to the data to perform further analysis. The data marketplace provided by Windows Azure is also potentially advantageous, in that it extends the concept of readily and securely sharing data to include the option for the data owner to monetise the data set, the income from which could fund additional analysis.