Introduction to HP Serviceguard Clusters
Harish Kamath
HP Master Architect
March 2014

Clusters of computers can provide both high availability and high performance. While IT infrastructures are susceptible to planned and unplanned downtime, a key requirement of IT is to minimize or eliminate these disruptions related to business. Proactively protecting IT infrastructures from failures and service interruptions can be solved via HP's clustering technology, HP Serviceguard, to provide high availability of IT services.

In addition to high availability, science and engineering functions depend upon a high performance infrastructure to solve large problems in their work area. Weather forecasting, DNA analysis, market predictions, where to drill for oil and gas, and advanced research in the field of science are a few examples belonging to this sector. The requirement for high performance computing, which is not solvable through typical stand-alone computers, is addressed by aggregating computing power in a way that delivers higher computing capacity. Aggregation of computers for high performance computing is also achieved through clusters of computers managed by Serviceguard.

In failover clusters, we refer to how quickly a workload fails over, thus providing high availability. Compute clusters, however, refer to the rate at which a job gets processed.

Clustering technology allows individual nodes to operate in a coordinated manner to meet the needs of high availability and scalability. Clusters can further be classified as Shared-Nothing, Shared-Storage and Shared-Everything, a classification essentially indicating the extent of resource sharing among nodes of a cluster. An application running on any node of a cluster that must have simultaneous access with applications running on other nodes to the same file or data typically requires a distributed lock manager; this is a function provided by the Shared-Everything model. HP's Serviceguard adopts the approach of Shared-Nothing clusters, where no simultaneous access to data is permitted. HP addresses the Shared-Everything requirement through its OpenVMS Clustering technology.

While HP Serviceguard has taken the approach of Shared-Nothing, it is extensible to Shared-Storage since it is tightly integrated with select cluster file systems. This facilitates simultaneous access to data from cluster members. Yet another capability of Serviceguard is to integrate multi-node packages, thus bringing in the notion of active-active processing to a cluster.

Traditional high availability clusters are formed through a concept of cluster membership. Serviceguard nodes within a cluster maintain a consistent view of all of the other member nodes in the cluster. The nodes within a cluster continuously monitor the health and status of each other through a heartbeat mechanism.

System or applications may malfunction due to hardware errors. Failures could also be the result of certain dependent operations exceeding acceptable threshold limits. Network congestion, for instance, is

1

a form of threshold violation in which systems or applications appear to respond incorrectly or initiate fallback actions. Since the cluster heartbeat is a core functionality of a cluster and has a direct impact on uptime, Serviceguard delivers multiple configuration options such as separating data from heartbeat traffic by utilizing separate network channels and by applying network card level redundancy for heartbeat requirements, thus introducing tolerance to network card failures.

When a Serviceguard cluster is partitioned, conditions such as split-brain can be easily recovered with the help of HP's advanced arbitration mechanisms, lock-LUN or quorum services. This helps maintain data integrity and reliability.

System facilities also contribute to redundancy at various levels. HP Serviceguard leverages system facilities such as network card-binding or floating IP addresses in its environment in order to provide an additional layer of high availability. In short, the integration of a complex set of technologies and mechanisms by Serviceguard over the years prevents application failures, which leads to high application uptime.

The clustered platform enables workloads to fail over across nodes of the cluster. The conditions under which failover may take place may be due to error conditions observed in one of the components or by an act of policy enforcement related to load-balancing. HP Serviceguard delivers a platform to integrate a variety of applications for addressing high availability with a high degree of modular automation capabilities.

Virtualization products such as VMware and KVM attempt to deliver high availability to virtual machines on their hypervisor platforms. This technology is often referred to as a pool manager. A pool of virtual machines is typically managed via a centralized console, so systems are not clustered through a membership mechanism. Application failover needs in such an environment are handled with the help of HP Serviceguard through a set of supported configurations.

Organizations utilizing the HP Serviceguard Solutions Suite are able to leverage highly available services with a wide range of applications to meet the needs of businesses. This solution maximizes uptime, provides high performance, and meets the high availability requirements at and across sites to best serve an organization.

> *B Harish Kamath carries a diverse profile, with over two decades of software industry experience. He has contributed to infrastructure elements, be it in enterprise or cloud related technologies. As a Master Architect, he currently leads the effort in developing artifacts to meet the needs of business continuity and security at large.*