

ABSTRACT

An apparatus and method for closed loop dynamic resource allocation. For example, one embodiment of a method comprises: collecting data related to usage of a plurality of resources by a plurality of workloads over one or more time periods, the workloads including priority workloads associated with one or more guaranteed performance levels and best effort workloads not associated with guaranteed performance levels; analyzing the data to identify resource reallocations from one or more of the priority workloads to one or more of the best effort workloads in one or more subsequent time periods while still maintaining the guaranteed performance levels; reallocating the resources from the priority workloads to the best effort workloads for the subsequent time periods; monitoring execution of the priority workloads with respect to the guaranteed performance level during the subsequent time periods; and preemptively reallocating resources from the best effort workloads to the priority workloads during the subsequent time periods to ensure compliance with the guaranteed performance level and responsive to detecting that the guaranteed performance level is in danger of being breached.

Fig. 1A