Vantage provides the capability to manage CPU load through Session Based load balancing. When greater control over workflow processing is desired to maximize server utilization and improve workflow throughput, Vantage provides an Advanced Task Scheduling option.

**Introduction**
Within the Advanced Task Scheduling option there are two types of additional load balancing capabilities. *Cost Based load balancing* provides greater control for Action execution within a session and *Task Based load balancing* maximizes server efficiency across all services, on all servers, and helps avoid the situation where some Vantage servers are over-utilized and others are under-utilized. Task Based load balancing provides a generational improvement over Cost Based load balancing by balancing load across services of all types (e.g. Transcode, Analysis, Conform, etc.)—not just services of like types (e.g. only Transcode). It is therefore recommended that Task Based load balancing by utilized over Cost Based load balancing.

**Session Based Load Balancing**
Session Based load balancing is a default Vantage Array feature that distributes service action processing across multiple servers supporting the same service, up to the configured session limit for each server (or a maximum of four concurrent Lightspeed jobs on a Lightspeed server). Vantage takes CPU utilization of each machine into account, and also ensures that any *Run On* rules (routing tasks to specific servers) are observed. If all servers reach the configured limit, Vantage queues actions until the workload drops below the limit.
Session Based load balancing offers good, basic load balancing as a starting point for Vantage Array systems, and it comes built-in and operating with the Array license. You can make adjustments to this automatic system by setting the session limits for each service. If your Vantage system utilizes a combination of multiple transcoding services, then using Task Based load balancing will allow you to balance between these services providing optimal machine saturation.

The diagram at the top of the page shows how an IPTV Flip Action is scheduled on Server 1 when Server 2 and Server 3 have no more current processing capacity.

In this example, Server 1 has a more powerful CPU and is deemed to have twice the horsepower capability of Server 2 and Server 3. The IPTV Service therefore has a Session Limit of four for Server 1, and two for Server 2 and Server 3. The Vantage Scheduler determines that an IPTV Flip processing slot is available on Server 1.

**Cost Based Load Balancing**

Cost Based load balancing allows you to assign a default cost for each Action type. You have the option to override the default Action cost in workflows as needed. You can also define a target resource usage level for each service on a server.

To determine its available resources, each service monitors the total cost of all Actions it is processing and compares that cost to the configured target resource usage level for that service. A service can accept an Action for processing if the available resource usage is at least half the cost of the Action.

Actions are assigned to services in the order they are processed, which prevents low-cost Actions from consuming all service resources and starving higher-cost actions. Each Action is forwarded to the service with the most available resources. If the Action is forwarded to a service that does not have resources to immediately process the Action, the Action is then queued for later processing.

Cost Based load balancing offers considerable control and flexibility in balancing a Vantage Array for optimum, efficient operation. Dynamic task scheduling helps ensure very large jobs or many small jobs do not overwhelm the service. You can also assign processing “costs” independently to services and Actions, allowing you to finely tune processing resource usage. There are many benefits to this option, and it is easy to set up, but it requires some time to monitor and adjust.
The diagram above shows how an IPTV Flip Action is scheduled on Server 2 when Server 1 has no more processing capacity.

In this example, Server 1 and Server 2 are identical. The IPTV Flip Action has been defined to have a Cost of 4, and the IPTV Service has been defined to have Target Resource Usage of 8. Hence, two IPTV Flip jobs can run on a single IPTV service and in this example, Server 2 has capacity to execute the IPTV Flip Action.

**Task Based Load Balancing**

Task Based load balancing is the most efficient and capable form of Vantage system load balancing. It not only provides the ability to balance workload across servers in a Vantage Array, but also across services (as opposed to Session Based and Cost Based load balancing which only balance workload within a single service).

For example, Task Based load balancing allows you to balance the workload between CPU intensive Actions such as a Flip Action, a Multiscreen Action, an IPTV Action and an Analyze Action. Each of these Actions is managed by a separate Vantage service yet Task Based load balancing allows for tuning of the execution to maximize server utilization and minimize processing time.

Task Based load balancing uses the concept of **Capacity** for the servers (or entire Vantage Array) and **Cost** for the Actions that need to be processed. **Rules** are created to define Capacity or Cost and are assigned at a default level for all servers and services respectively. Additionally, the default Cost assigned to an Action can be overridden within the Workflow that utilizes the Action.
In the diagram above, Server 1 has a capacity of 100 and for the three Actions; each has varying costs associated with their execution.

Note that each of these Actions is executed by a separate Service. Task Based load balancing spans Service boundaries as opposed to Cost Based or Session Based load balancing which only works within the confines of a single service.

Not only can Rules be created to control the balancing of Actions across services, but also Rules can be created that control access to physical characteristics of a Vantage Array.

Access to a storage sub-system can be controlled through Task Based load balancing to ensure that the storage sub-system is not oversaturated with too many processes. Likewise, if you need to send files via FTP to a remote server, and the remote server only allows a single FTP connection, then Rules defined in Task Based load balancing allow this to be accomplished.
The diagram above illustrates the constraint of a single FTP session available on a remote server.

The Domain Rule of one FTP connection is what is used to control the access to the physical resource.

Requirements
Session Based load balancing comes standard with any Vantage system. The Vantage Advanced Task Scheduling license is required to take advantage of Cost Based or Task Based load balancing. The one license enables both capabilities.