



System Resiliency and Overload Throttling

MediaSense keeps track of a number of metrics and statistics about its own performance and raises alarms when certain thresholds are exceeded. The system also protects itself by rejecting requests that would cause it to exceed its critical capacity limits. When the node is at capacity, new recordings are redirected to other nodes (if available) or rejected and lost.

Because recording is always considered to be the highest priority operation, MediaSense reserves a certain amount of capacity specifically for that purpose, electing to reject media output requests while still continuing to accept new recording requests. Media output requests (such as live monitoring, playback, raw download and .mp4 or .wav conversion) result in 503 responses when the node is at capacity.

The relative weight of various media is also considered for overload throttling. For example, video takes significantly more capacity than audio.

However, these are overload-protection methods only; they are not intended to enforce licensed or rated capacity. They reflect the levels at which the product has been tested, and they exist so that MediaSense nodes can protect themselves and offer graceful service degradation in case of severe overuse.



Note

It is still the customer's responsibility to engineer his or her deployment such that the overall rated node and cluster capacities are not exceeded.

MediaSense also protects itself with respect to media storage capacity. It raises alarms, redirects new calls to other nodes (if available), prunes older recordings to recover space (if permitted), and even drops existing calls (as a last resort) in order to maintain the integrity of existing recordings.

The Real Time Monitoring Tool (RTMT) provides a great deal of statistical information about use levels and throttling activities for each node.

