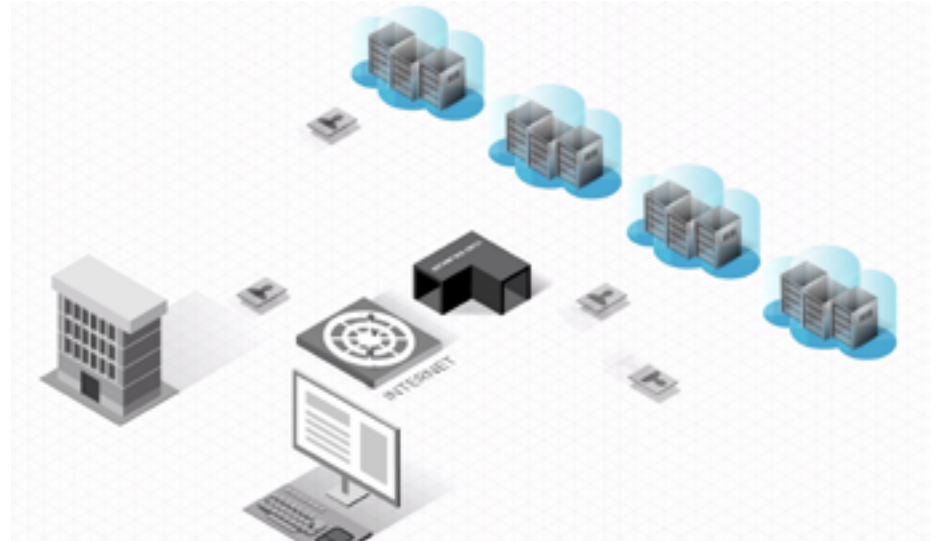## AUTO-SCALING

Once a balanced instance has been configured, an auto-scaling system can be configured. Auto-scaling is a set of instructions whose verification involves the creation (or deletion) of a cloud server (using unassigned resources within the Cloud Datacenter itself). You set the instructions, create a template, and link them to a specific server.

The aim is to maximize resource efficiency and provide maximum flexibility to your Gigas solution.

Gigas' auto-scaling system allows you to define the automatic creation and deletion of servers in a simple and intuitive way, either from Gigas' API or from the Control Panel itself.

Auto-scaling uses

For example, you can auto-scale your resources:

- Depending on the number of user sessions of the Gigas balancer. The auto-scaling system is fully integrated with our balancer making it possible to create a group of servers that automatically incorporate themselves into the balancer, thus increasing the capacity of service when a specific number of sessions is reached.

- If the load of one of your servers goes above the limit of CPU usage that your have set and for the minimum time you have established.

- For certain pre-set time periods: You can create instructions to increase and/or shut down servers at specific times.

REQUIREMENTS

- A customized template for a given cloud server is needed

- A configured and working balanced instance and a configured backend are needed

- Sufficient (unassigned) resources must be available in the cloud to cover the number of configurations for the auto-scaled cloud servers (cloned from the template of the first server) that need to be working concurrently.