

Turbonomic ARM

Role-based Demo Script

This script is intended as a guide to demonstrate how the Turbonomic Application Resource Management platform supports different personas in a client organization, specifically:

- Enterprise SRE Manager/ Application Operations Manager
- Site Reliability Engineer
- Infrastructure Manager

It is a good idea to have information about the customer's use case(s) and the problems they are trying to solve prior to doing the demo to show specifically the impact that Turbonomic will have on their environment.

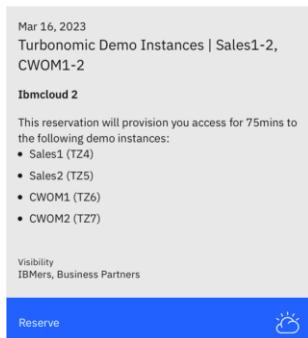
In the demo script:

- “**Action**” denotes a setup step for the presenter.
- “**Narration**” denotes what the presenter will say.
- “**Note**” denotes where the presenter may need to deviate from this demo script depending on the customer's environment setup.

To access the demo environment, open the following TechZone collection:

<https://techzone.ibm.com/collection/turbonomic-application-resource-management-demo-assets>

Request access by selecting the following tile titled “Turbonomic Demo Instances”.

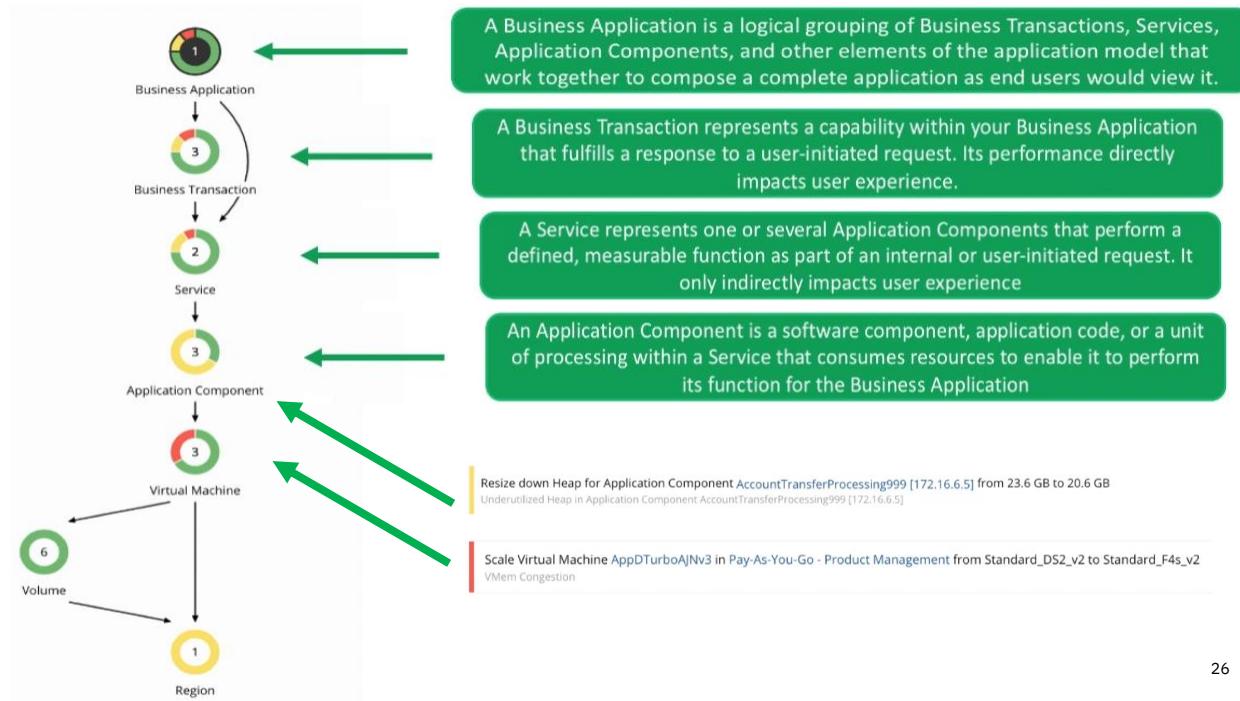


Reservation authorizes access to four demo instances:

- <https://tz4.demo.turbonomic.com>
- <https://tz5.demo.turbonomic.com>
- <https://tz6.demo.turbonomic.com> (Cisco OEM version)
- <https://tz7.demo.turbonomic.com> (Cisco OEM version)

Access is granted for 75 minutes. When reservation expires, you have to request access again.

KEY TERMINOLOGY



26

You should be familiar with the following terminology when discussing Turbonomic:

- **Business Application** - *a logical grouping of Business Transactions, Services, Application Components, and other elements of the application model that work together to compose a complete application as end-users would view it. (e.g. a travel booking application like SAP Concur)*
- **Business Transaction** - *represents a capability within your Business Application that fulfills a response to a user-initiated request. Its performance directly impacts user experience. (e.g. making a hotel booking)*
- **Service** - *represents one or several Application Components that perform a defined, measurable function as part of an internal or user-initiated request. It is consumed by internal parts of the application. (e.g. service to locate hotels around specified location)*
- **Application Component** - *a software component, application code, or a unit of processing within a Service that consumes resources to enable it to perform its function for the Business Application. (e.g. database)*

Actions are recommended for Application Components or underlying Infrastructure and propagated up the stack.

1. INTRODUCTION

Narration

Welcome to this demonstration of the Turbonomic platform. Turbonomic is an Application Resource Management (ARM) platform designed to put your hybrid environment into the optimal state by achieving three primary goals:

- assuring the performance of your applications by making sure that the applications get the resources they need when they need them,
- using limited on-premises and cloud resources as efficiently as possible at the lowest cost, and finally
- adhering to any compliance and business policies required by your organization.

Turbonomic can help multiple personas in a typical client organization in their daily duties. Let's start with Mike – the Enterprise Site Reliability Engineer (SRE) Manager. Mike manages a team of SREs and is responsible for the performance of key business applications, but also for optimizing costs of running them.

Let's see it in action!



Mark
Enterprise SRE Manager

Application performance prioritization
I want to identify the top applications where performance can be improved and assess the benefit to the end user.

Put performance assurance on autopilot
I want to delegate next best action decisions to a trusted and proven ML and AI engine to continuously optimize performance and reduce resource utilization with low/no touch.

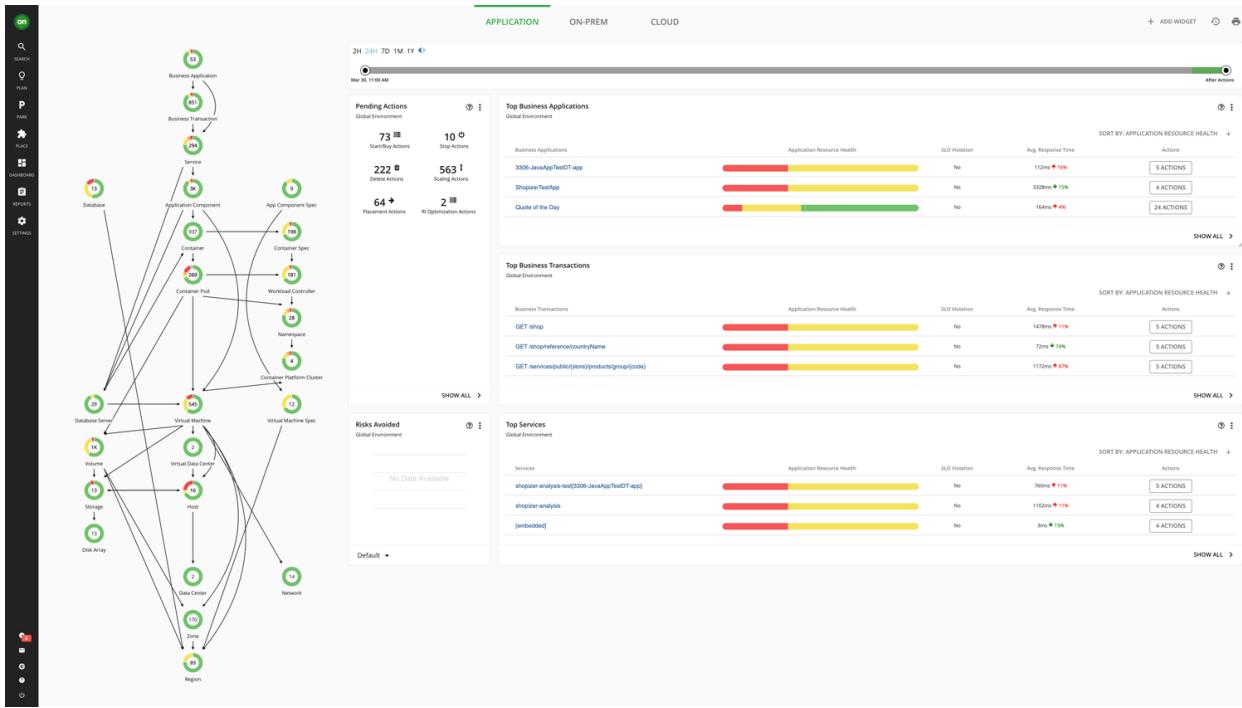
Cost savings across clouds
I want to assess and optimize overall resource utilization across my cloud and on-premises deployments to reduce my overall spend.



2. TURBONOMIC OVERVIEW

2.1 Action

Navigate to the Turbonomic home page (it opens by default). If you want to return there, click the green **ON** icon in the top left corner of the window.



Narration

When the Turbonomic ARM is deployed and selected environments are targeted, Turbonomic discovers all the entities in the targeted environments. It then builds out “the supply chain” (the graph on the left), stitching together all of the entities from the top-level business application down through the supporting infrastructure.

This shows the relationships among the various entities and gives you full-stack visibility into your entire hybrid environment, which is very powerful.

In this example, Turbonomic is pulling in:

- business applications from Application Performance Monitoring (APM) tools like Instana, AppDynamics, Dynatrace, or New Relic
- their associated transactions, services, and application components
- containers, pods, and nodes (represented as virtual machines) from any Kubernetes environment whether on-premises or in the cloud
- virtual machines and storage from on-prem environments like VMWare vCenter and public cloud environments like AWS and Azure
- physical hosts, storage, network with many more targets available.

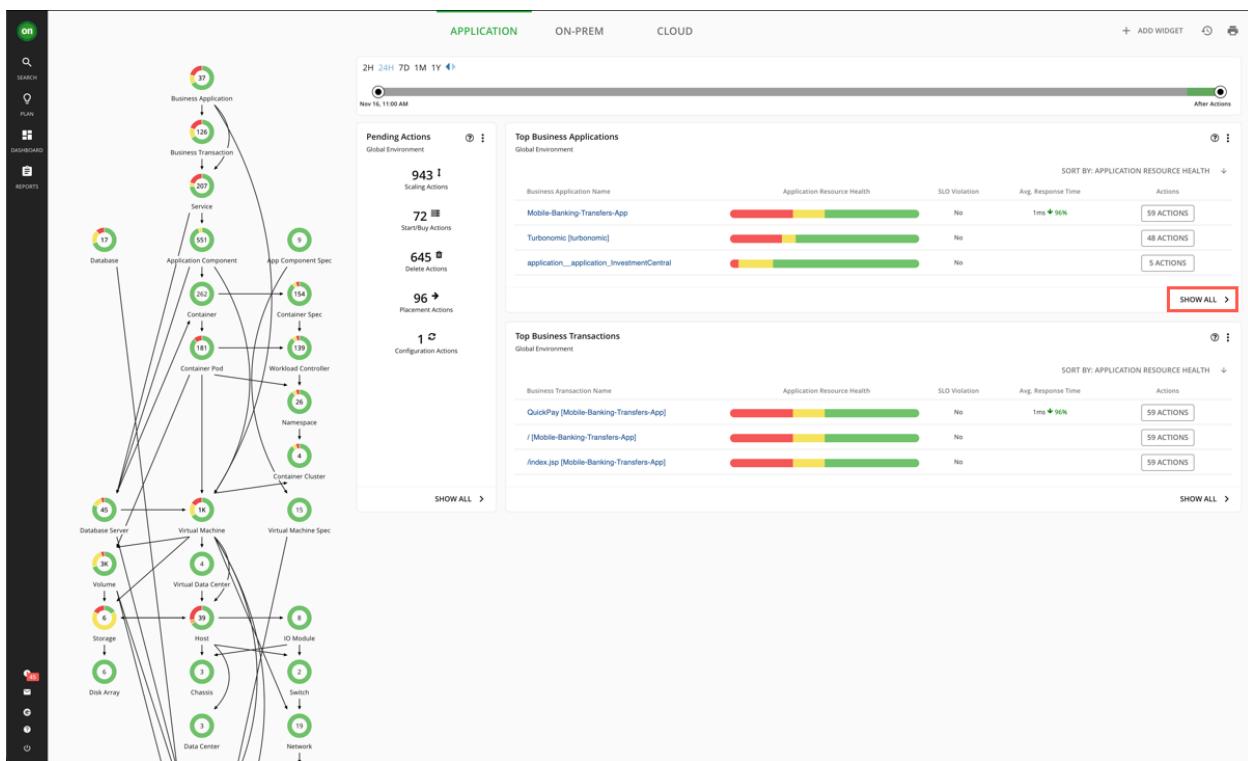
No agents are required to be installed as Turbonomic is agentless. It discovers all these entities automatically via the APIs of the targeted environments and pulls them into a common data model.

Within about 30 minutes of discovery, Turbonomic determines the actions required to put this hybrid environment into the optimal state that I mentioned previously. In this optimal state, the demand of the applications is met by the underlying supply of IT resources.

On this home page, there are widgets for Top Business Applications, Top Business Transactions, and Top Services that allow you to quickly assess the risk to the running applications. Not only can you quickly see the applications with the highest risk, but you can also click **Show All** to see the complete list of Business Applications and relevant metrics. This helps prioritize the SRE team effort.

2.2 Action

Click **Show All** on the “Top Business Applications” widget.



Note: The list of the applications as well as values for metrics may be different from the ones shown above. However, this should not affect the demo flow.

Global Environment

Total 26

Sort by APPLICATION RESOURCE HEALTH ▾

1	2	3	4
SLO Violation	Avg. Response Time	Peak Response Time	Avg. Transactions
No	2ms ↑ 94%	3ms	0TPS
No	63ms ↓ 52%	459ms ↓ 89%	0.33TPS ↓ 4%
Yes ↑ 51%	2089ms ↑ 33%	2089ms ↓ 39%	0TPS ↓ 100%
No	0ms	0ms	0TPS
No	11ms ↓ 12%	117ms ↓ 74%	0.05TPS
No	864ms ↑ 37%	3789ms ↓ 44%	0.02TPS ↓ 99%
No			0.02TPS
No			0TPS
Yes ↑ 72%	5504ms ↓ 7%	5504ms ↓ 7%	0.02TPS
			no actions

Narration

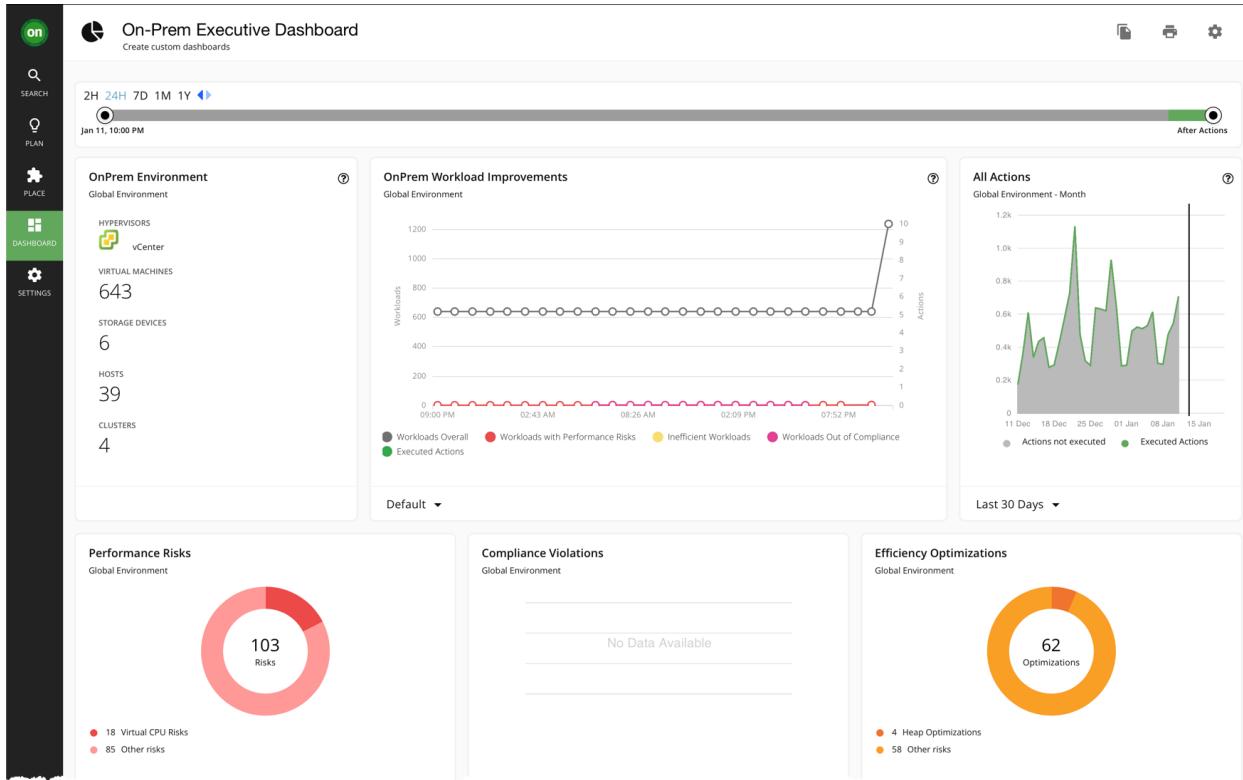
For example, if you look at the Top Business Applications widget, you can see the specific application, any resource risk from the underlying infrastructure, if there is a violation of a Service Level Objective (SLO) (1) like exceeding the desired transactions per second, the averages and peaks for response time (2) and transactions (3) and all of the actions (4) required to put this application into a healthy state. The underlying infrastructure actions are propagated up to the application so that when you click on “Actions”, you will see all the actions that impact the specific application.

2.3 Action

Click the **DASHBOARD** icon in the menu (1). Click **On-Prem Executive Dashboard** (2).



Note: Depending on your client's interest you may start with Cloud or Container dashboards. Adjust your narration accordingly.



Narration

An Enterprise SRE Manager usually wants to understand the overall status of the infrastructure and long-term utilization trends. Turbonomic provides several dashboards available out of the box and you can easily build new custom ones.

For example, **On-Prem Executive Dashboard** provides the summary of the status and risks related to on-premises virtualization resources. At the top of the screen, you can select the time for which you would like to see the data. Below, there are details for virtualization platforms, a summary of the performance, risks, and actions taken. At the bottom, you can find important metrics related to virtual machine density on hosts and storage, and how they are changing over time. Similarly, there is a dedicated dashboard for cloud infrastructure (Cloud Executive Dashboard) and Kubernetes infrastructure (Container Platform Dashboard).

Now, let's see how Turbonomic helps Site Reliability Engineers with their tasks. Compared to the SRE Manager role, an SRE is usually interested in a specific application status and actions that can be taken to assure its performance, mitigating any risks before they impact end-users.

Let's see it in action!



Eddie

Application SRE

Assure application performance & availability

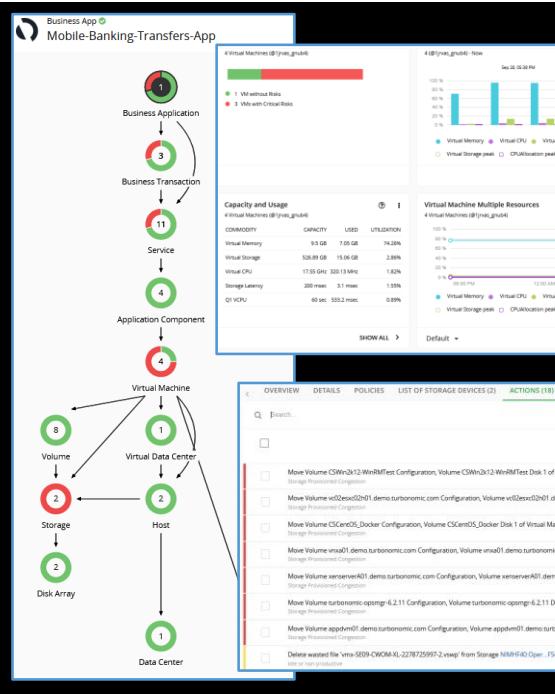
I want to define key SLOs for my application and rely on automated software to dynamically shift resources to meet them.

Manage business application resources

I want to monitor and optimize my business application's resource usage to reduce its total cost of ownership.

Incident prevention and assisted remediation

I want guidance on what actions to take to prevent incidents, reduce alerts, and remediate problems before they impact end users.

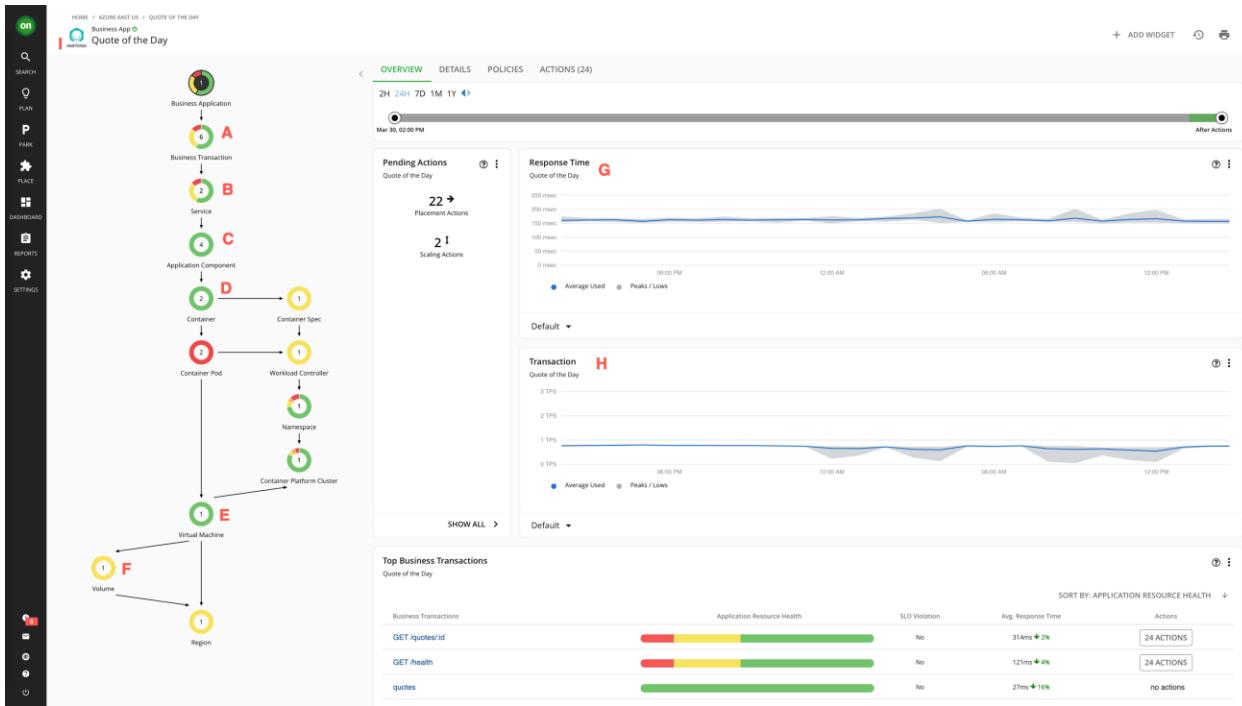


3. BUSINESS APPLICATION DRILL DOWN

Note: The business application that you choose may differ from what is presented in this script. Please adjust your story accordingly

3.1 Action

Click any of the applications listed in the Top Business Applications Widget, for example "Quote of the Day". If it is not listed on the main view, click **SHOW ALL** to open the full list.



Narration

When you scope to this business application, the supply chain has changed from showing all the entities in the global environment to only the entities that are relevant to the selected application. This business application consists of 6 business transactions (A), 2 services (B), and 4 application components (C) running in 2 containers (D) that sit on top 1 virtual machines (E). This virtual machine is connected to 1 storage volume (F). Notice that the entities may have different colors where:

- “red” indicates performance risk where entities need to be moved to a more optimal host or scaled up because of resource congestion or under-provisioning
- “yellow” indicates efficiency opportunities where resources can be reclaimed due to overprovisioning; and
- “green” indicates entities in their desired state, which is what we are after.

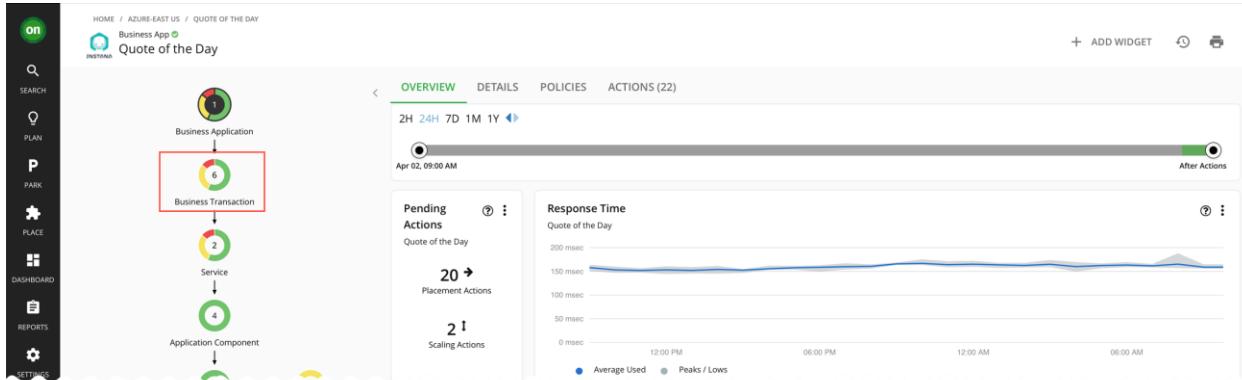
The result is that by scoping to the specific application, you can quickly see where there is any congestion risk or where the resources are potentially wasted.

You see the metrics for response times (G) and transactions per second (H) along with their service level objectives (SLO) that you can set in the policy section in Turbonomic. These information comes from the connected APM tool – in this case Instana, what is indicated in the upper left corner (I).

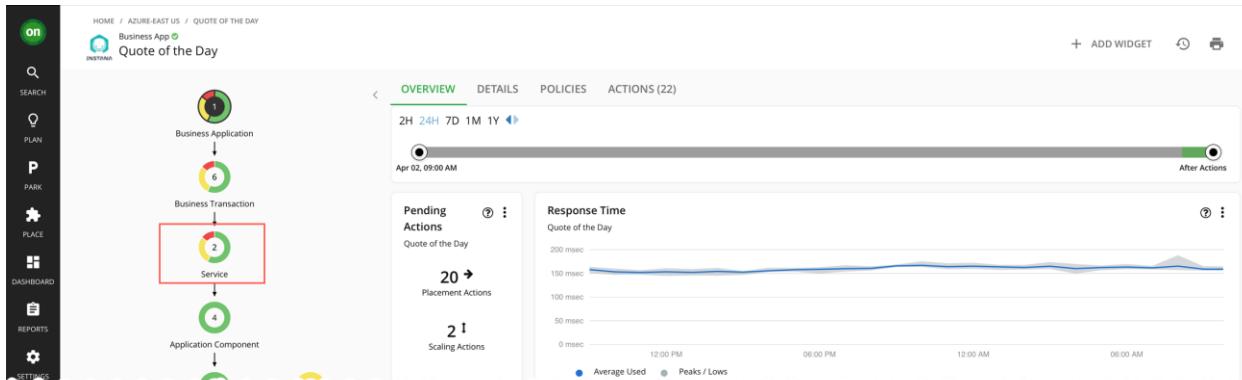
The structure of supply chain was also imported from Instana, where Business Application is defined as a logical grouping of Business Transactions, Services and Application Components.

3.2 Action

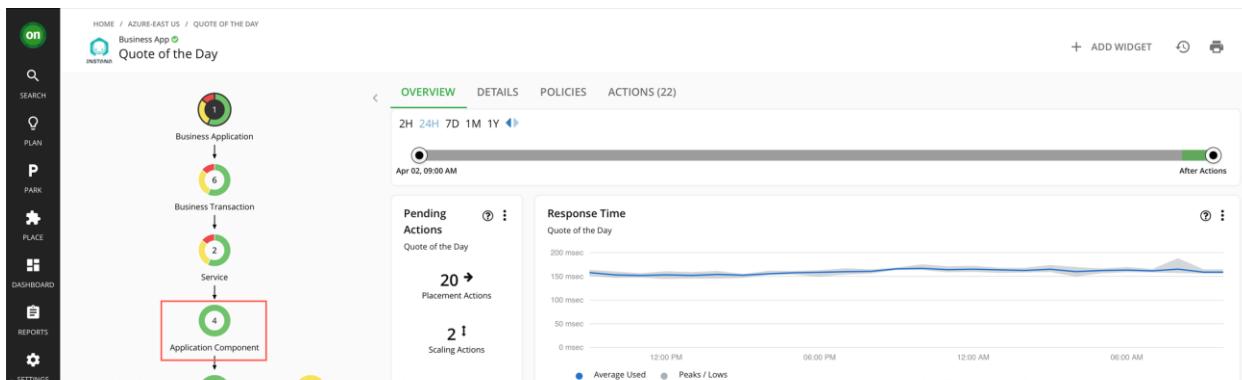
Click the **Business Transaction**, **Service**, and **Application Component** icons in sync with narration.



If you scope to the Business Transaction, it represents a capability within your Business Application that fulfills a response to a user-initiated request. Its performance directly impacts user experience.



The Service entity is made up of one or more Application Components that perform a defined, measurable function as part of an internal or user-initiated request. Its performance is key to understanding application performance, but only indirectly impacts user experience.



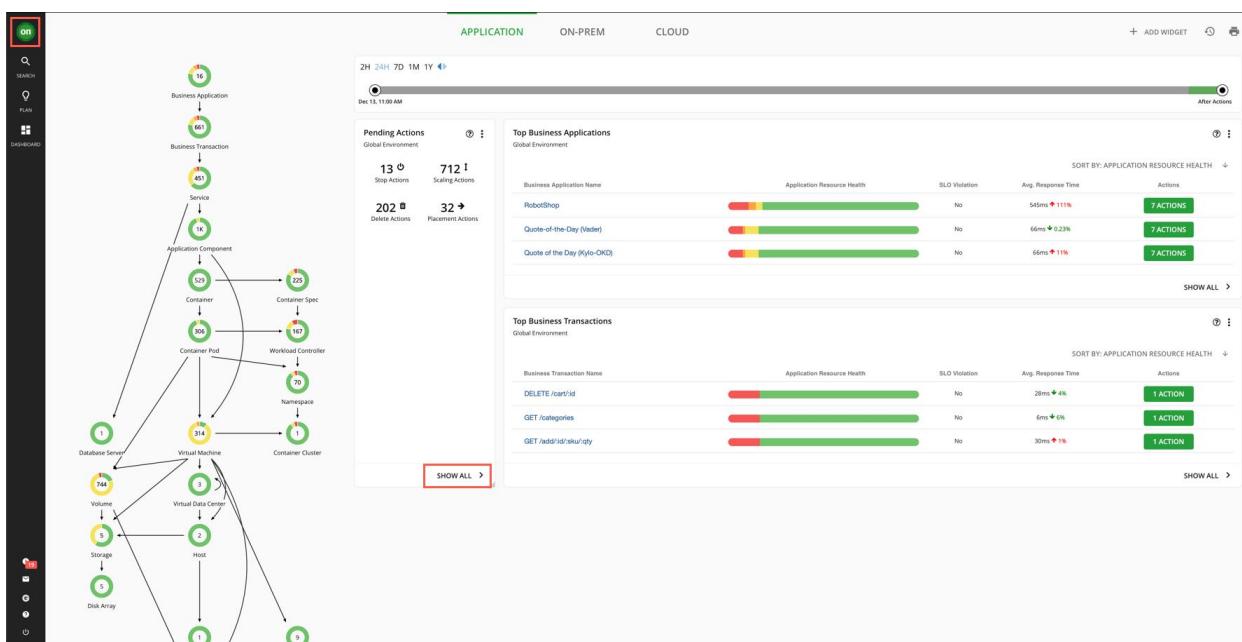
Finally, if you scope to the Application Component you will see a software component, application code, or a unit of processing within a Service that consumes resources to enable it to perform its function for the Business Application. Turbonomic can recommend actions to adjust the amount of resources available to Application Components.

Turbonomic does not recommend actions for a Business Application, Transactions, or Services, but it does recommend actions for the underlying Application Components and infrastructure entities, thus providing visibility into the risks that have a direct impact on the Business Application's performance. For example, you can see here the actions to resize down both the heap (memory allocation) for a Java Virtual Machine (JVM) and the virtual machine (VM) that hosts that JVM.

Let's explore the Turbonomic actions in more detail.

3.3 Action

Click the green **ON** button in the top-left corner to return to the Turbonomic home page. Click **SHOW ALL** at the bottom of the **Pending Actions** widget.



The screenshot shows the Turbonomic Action Center interface. The left sidebar contains navigation links for SEARCH, PLAN, PARK, PLACE, DASHBOARD, REPORTS, and SETTINGS. The main area displays a list of pending actions categorized by provider (AWS, AZURE, GCP) and type (DELETE, SCALE, RESIZE, MOVE, BUY, PROVISION). Each action row includes columns for Volume ID, Account, Tier, Size, State, Days Unattached, Disk Image, Last VM, Action Category, Cost Impact, and a Details button. A search bar and an 'ADD FILTER' button are located at the top right of the list.

Narration: You can see here in the Pending Actions widget the types of actions that are generated. These can be:

- “placement” actions to move a virtual machine to the most desirable host or storage from one tier to another
- “delete” actions to remove storage volumes that have been orphaned
- “scaling” actions to resize a virtual machine, container, or database
- “stop” or “suspend” actions to turn off idle resources to avoid unnecessary charges, and
- “start / buy” actions to turn on new VMs, provision a new host, or buy reserved instances (RIs) in the cloud.

In this example, there are 912 actions that can be taken to move this hybrid environment into the optimal or what we call the desired state. You can see there are actions to scale up the virtual CPU or virtual memory of virtual machines, provision actions to add a host in a cluster where resources are low and move actions to move a virtual machine from one host to another.

Because there are hundreds and sometimes thousands of actions in an environment, Turbonomic allows you to apply filters to the list of actions to help you quickly sort the list. Clicking **ADD FILTER** and opening the dropdown, you can filter the actions by action type, action mode, action category, action prerequisite, or entity type.

You may ask “How does an SRE know that these are the right actions to take?” For each action, Turbonomic not only shows you the current state of the environment, but it also shows the positive impact on the environment if the suggested action is taken. Let’s expand a couple of actions to showcase this starting with a move action.

3.4 Action

Expand **Move → Virtual Machines** (or **Move → Container Pods** if the other is not visible). Click the **DETAILS** button on any record.

Host	vc01esx08.na.cloud.techzon	vc01esx02.na.cloud.techzon	Improve overall performance	PREVENTION
vc01-centos-C04 (vm-79)	vc01esx08.na.cloud.techzon	vc01esx05.na.cloud.techzon	Q1 VCPU & CPU Congestion	PERFORMANCE
vc01-centos-C03	vc01esx04.na.cloud.techzon	vc01esx01.na.cloud.techzon	CPU & Mem Congestion	PERFORMANCE
vc01-centos-M30	vc01esx02.na.cloud.techzon	vc01esx05.na.cloud.techzon	Q1 VCPU & Mem Congestion	PERFORMANCE
vc01-centos-M19	vc01esx08.na.cloud.techzon	vc01esx04.na.cloud.techzon	Q1 VCPU Congestion	PERFORMANCE
vc01-centos-M08	vc01esx07.na.cloud.techzon	vc01esx04.na.cloud.techzon	Q1 VCPU Congestion	PERFORMANCE

Note: Depending on the action selected the risk detected (a reason for action to be generated) – PERFORMANCE in this case – may be different. Adjust your narration accordingly.

CPU	CPU PROVISIONED	MEMORY	MEMORY PROVISIONED
33.8 % → 13.2 GHz	32.5 % ↓ 13.2 GHz	94.3 % → 16 GB	76.1 % ↓ 16 GB
10.4 GHz	8.6 % ↑ 10.4 GHz	47.6 % → 16 GB	52.7 % ↑ 16 GB

Narration

If you look at the description of this move action, the VM (or Container Pod) is moving to a new host due to memory congestion on the current host. For each of the metrics shown, the value on the left is the current state while the value on the right is the projected state after the action is taken.

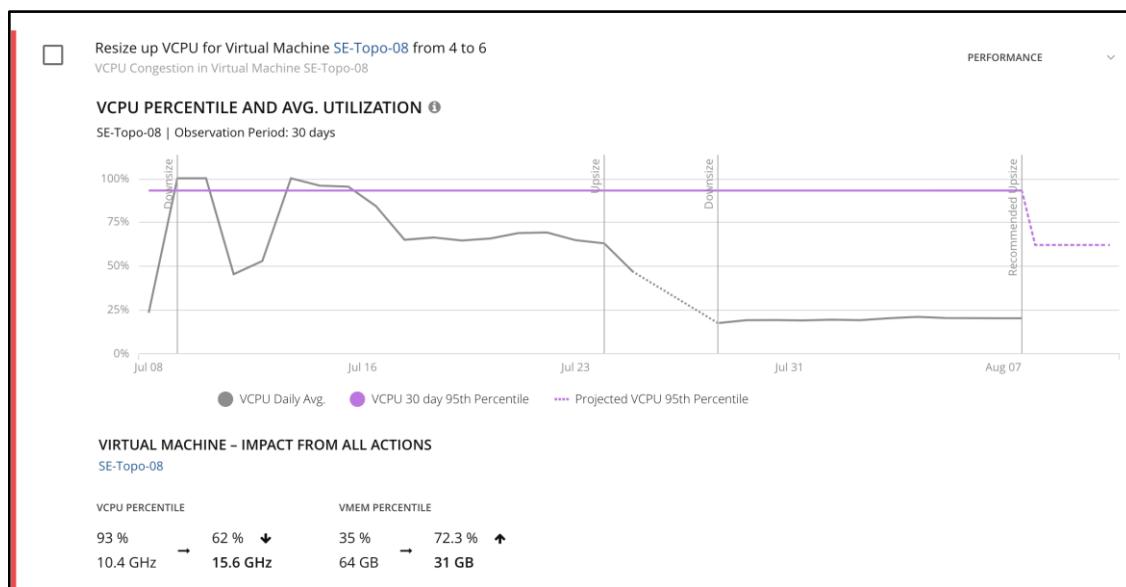
For example, on the current host's memory utilization is 94%. After the recommended actions are executed, the utilization will go down to 76.1%. For the destination host where the virtual machine will end up, you can see that the current memory utilization is 47.6% and will increase to 52.7% after the VM moves to this host.

Next, let's look at a VM scaling action.

3.5 Action

Expand **Scale → Virtual Machines**. Click **DETAILS** of any VM scale action.

Note: Depending on the action selected, the associated graphs and action details may be different. Adjust your narration accordingly.



Narration

For virtual machine scaling, Turbonomic uses percentile calculations to measure virtual CPU and virtual memory utilization more accurately, and drive scaling actions that improve overall utilization and reduce cost for cloud VMs. When you examine the details for a pending scaling action on a VM, you will see charts that highlight virtual CPU and/or virtual memory utilization percentiles for a given observation period, and the projected percentiles after you execute the action. Turbonomic's aggressiveness setting is 95th percentile with an observation period setting of 30 days. This means that in this example, Turbonomic shows that the vCPU utilization is below 93% for 95% of the time over the last 30 days. Because this is well above the current setting for the desired state of 70%, there is an action to resize up the vCPU. You can see that after this resize action is taken, the vCPU percentile utilization is projected to be below 62% for 95% of the time. The reason that you see the vMem percentile increasing is because there is a corresponding resize down action for vMem that is not shown here.

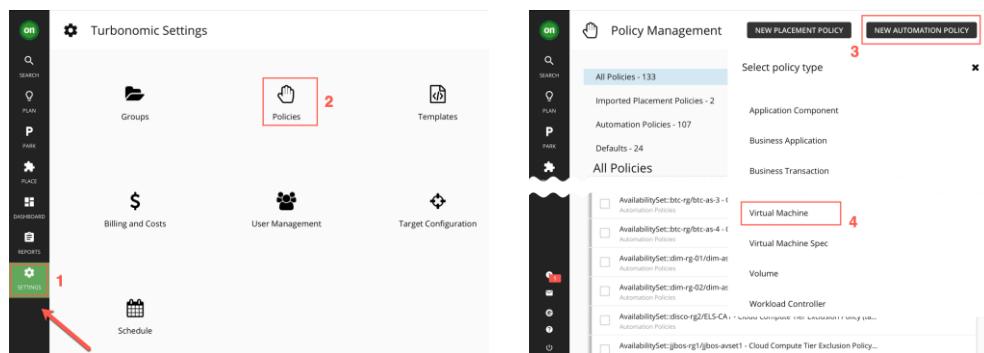
An important point to highlight here is that these actions are not simply recommendations on what a user should do; these are actions that can be taken within the Turbonomic platform. For example, if the action mode is set to manual, an SRE can click on this checkbox, scroll up, and click **EXECUTE ACTIONS** and this VM will move to the new host, or in the case of the resizing action, the VM will be resized. Taking this even further, you can create automation policies so that Turbonomic takes a preventative approach to automatically take these actions before issues occur.

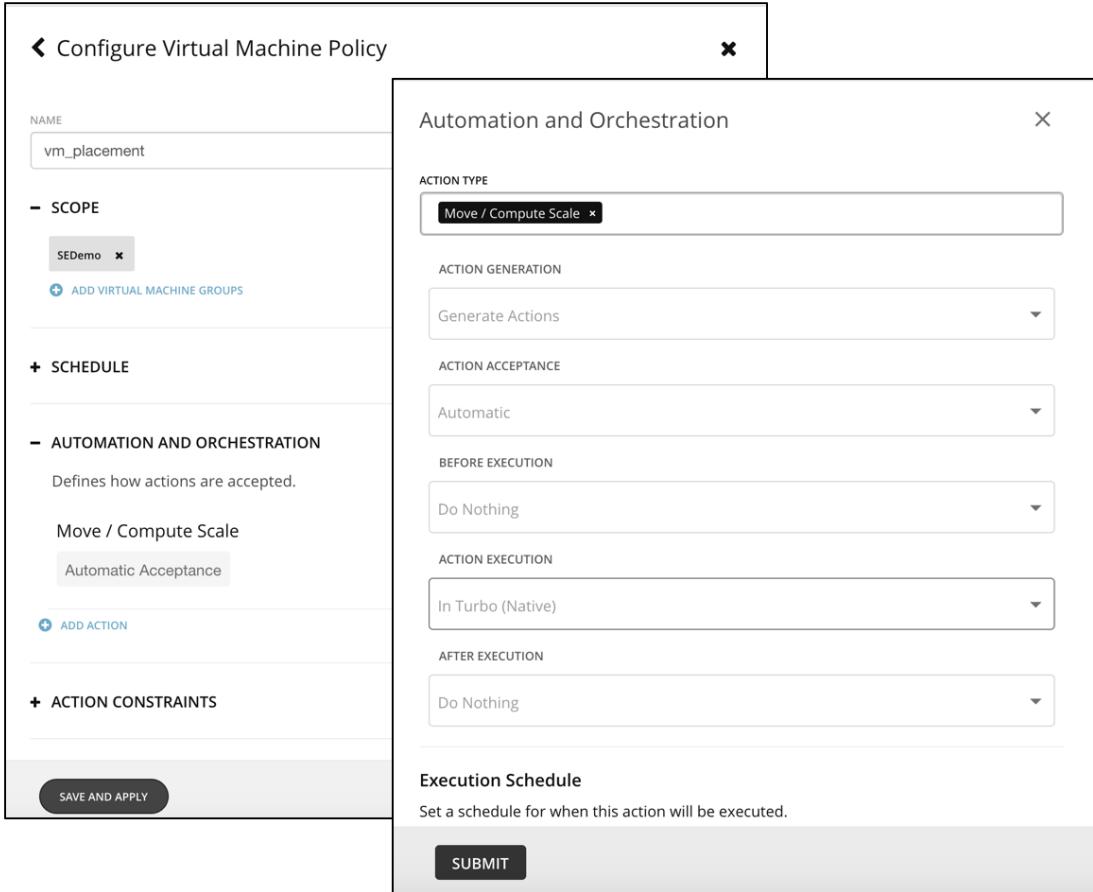
This keeps the environment in the desired state without human intervention and minimizes the amount of time that IT teams must spend debugging application resourcing issues after they occur and assures that the applications get the resources they need exactly when they need them.

In other words, Turbonomic assures that the demand of the applications is met in real-time by the supply of available IT resources. Finally, these actions can be scheduled to execute during a maintenance window and even get approvals via change requests through Turbonomic's integration with ServiceNow. Let's explore how SRE can set up automation policies.

3.6 Action

Navigate to **SETTINGS > Policies > NEW AUTOMATION POLICY >Virtual Machine**





Narration

Let's create a single automation policy for virtual machine placement and resize. In addition, let's assume that you want placement actions to execute anytime the action is generated and you want resize up and down actions for vCPU and vMem to occur nightly.

1. Navigate to **Settings > Policies > New Automation Policy > Virtual Machines**.
2. Give the policy a name: e.g. "**VM_placement_resize**".
3. Scope this policy to a group of VMs and in this case, let's automate within a single cluster.
4. Under **AUTOMATION AND ORCHESTRATION**, click **ADD ACTION**. You can choose the action type to **Move** to automate placement, action generation to **Generate Actions**, action acceptance to **Automatic**, then **Submit**.
5. To add resize automation, click **ADD ACTION** again. You can choose the four action types **vCPU Resize Up**, **vMem Resize Up**, **vCPU Resize Down**, and **vMem Resize Down** to automate VM resizes.
6. Set action generation to **Generate Actions**, action acceptance to **Automatic**.
7. For resizes, we said you wanted to do them on a schedule between midnight and 5 am, which can be done with an Action Execution Schedule. Let's create the schedule.

8. Save the new schedule, choose nightly, and click save, then click Submit to save the Automation.
9. **DO NOT** click Save and Apply. The demo environment is shared so you don't want to break it for other users. If you saved the policy, **delete it** after the demo is over.

You can see that there is a lot of flexibility in the setting of policies, and you have full control if those actions are taken manually or fully automated.

4. INFRASTRUCTURE CHANGES PLANNING

Finally, let's look how Turbonomic helps Infrastructure Managers with planning infrastructure changes, optimizations, or cloud migrations.

Let's see it in action!

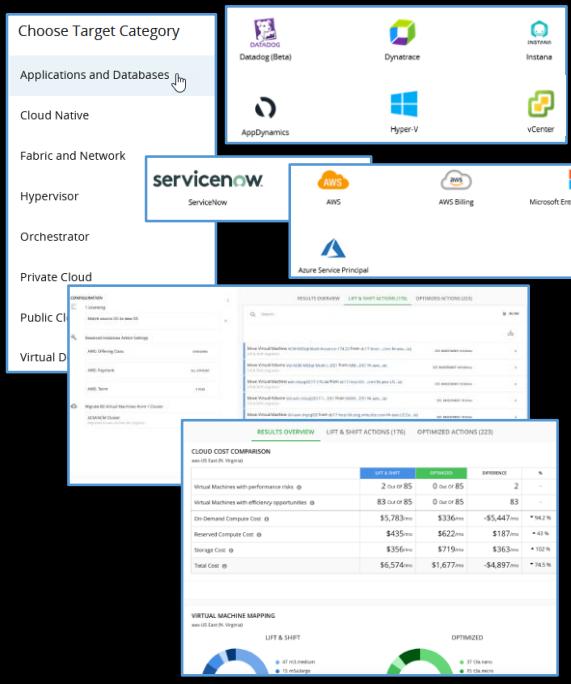


Jane
Infrastructure Manager

Limit tooling sprawl!
I want to integrate all my existing APM tools to visualize the end-to-end application dependencies and resource usage.

De-risk migrations to cloud
I want to plan the resources and costs for moving VMs to the cloud and migrate to the cloud with a full understanding of the cost implications.

Manage infrastructure changes
I want to control infrastructure changes, assess their associated risk, and simulate their cost impact before approving them.



4.1 Action

Click the **PLAN** icon in the left menu, then click the **NEW PLAN** button in the top-right corner.



Plan Management

USER PLAN **NIGHTLY PLAN CONFIGURATION**

Search... **FILTER**

NEW PLAN

Narration

Along with the dashboards that were explored before, Turbonomic provides a dedicated planning tool to help Infrastructure Managers decide infrastructure changes.

You can see there are multiple plan types available, for on-premises infrastructure, for Kubernetes clusters as well as for Public Clouds.

Let's see an example plan for optimizing the infrastructure.

4.2 Action

Select **Optimize On-prem** plan type. Select any available cluster. Click **RUN PLAN** at the bottom.

The screenshot shows a user interface for selecting a plan type. At the top, it says "Select a Plan Type" and has a close button "X". Below this, there are three main sections: "CONTAINER PLATFORM", "PUBLIC CLOUD", and "ON-PREM & PRIVATE CLOUD". In the "ON-PREM & PRIVATE CLOUD" section, the "Optimize On-prem" option is highlighted with a red box. This option is described as "Scale or move virtual machines and consolidate hardware". Other options in this section include "Add Virtual Machines", "Hardware Refresh", "Host Decommission", "Virtual Machine Migration", "Merge Clusters", "Alleviate Pressure", and "Custom Plan". Each option has a brief description and a "See how many nodes (Virtual Machines) are required" link.

Note: Depending on the client's interest you may choose other plan types. Adjust your narration accordingly.

Narration

When the plan is generated (it can take a while depending on the environment size) you can see the results. There is a table showing the summary of the current infrastructure and after the recommended actions are taken. Below, you can see the detailed list of proposed optimizations.

In each section, you can click **SHOW ALL** to see the list of proposed changes, and for each proposed change you can see the details as well as the rationale why this change should be made. These changes can be executed either manually or can be scheduled using the policies what was shown before.

5. DEMO SUMMARY

This demo has shown you how to use Turbonomic to optimize your hybrid estate. Turbonomic is an application resource management platform that takes a top-down approach to assure that your applications get the resources they need when they need them.

The Turbonomic platform automatically discovers the entities and metrics from the targeted environments, provides full-stack visibility from the application down to the infrastructure, and creates the placement, scale, start, and stop actions to optimize the environment, all in real-time with the goal of assuring application performance while adhering to business logic and compliance, and doing all of this at the lowest possible cost. All of these actions can be automated in order to prevent issues before they occur so that your IT team can focus on expanding your business, rather than reacting to alerts and trouble tickets.

6. CLEANUP

If you have created any artifacts such as a policy or a plan during the demo, please delete them after the demo is completed.