

Scalable Automated Provisioning Toolkit

Helen Cunningham
R&D User Experience
hcunningham@vmware.com

Deanna McCusker
R&D User Experience
deanna@vmware.com

Abstract. The complexity of provisioning new datacenters and virtualizing existing ones remains a barrier to the adoption of VMware technologies. This paper envisions a Scalable Automated Provisioning Toolkit that operates at three levels of customer sophistication, providing automation tailored to their needs. At the entry level, a customer with no IT experience needs BUSINESS INTELLIGENCE provisioning that takes advantage of industry best practices and expresses IT needs in terms of business metrics – how many employees do I have and what are they doing? At the intermediate level, a customer who understands traditional IT may need help virtualizing the datacenter. An ASSISTIVE service allows him to tweak basic settings, look at tradeoffs, and compare his choices with best practices. Finally, senior IT administrators with mastery of datacenter operations will appreciate a SELF-REVEALING view of objects and metrics in the inventory. Our paper presents plausible models for using VMware best-practices knowledge to create engaging customer experiences at all three levels. While this is not a software architecture paper, we have written prototypes in Java™ and Matlab™ that embody the proposed models.

Introduction

A key advantage of the virtualized datacenter is the ability to automate various tasks, especially provisioning. Provisioning requires a user to make decisions that will impact performance and reliability of the system, and the best decisions are based on a full understanding of virtualization technologies. VMware has a variety of offerings in this space, ranging from vCenter Orchestrator (for implementing the nuts and bolts of ITaaS) all the way up to App Director (which provides rapid authoring and deployment of vApps.)

We seek to provide automation even earlier in the customer acquisition process, at the point of sale. A speedy datacenter deployment makes for happy customers and a well-respected “VMware” brand.

In talking with customers and customer-facing technology experts, it quickly becomes clear that one size of automation does not fit all. Customers are heterogeneous: some know little about IT or virtualization, and want someone to just “make it so”. At the other extreme, veteran IT administrators want to control everything. Because small customers can grow into big ones, and because Service Providers will adopt our technologies if it helps them sell more services, we can make everyone happy by providing automation levels for different

customer needs, getting ALL new customers up and running as quickly as possible. Furthermore, we can encourage IT administrators to adopt a more automated approach when they can see that VMware recommendations are safe and can be trusted.

Customer Story 1. Business Intelligence

At the low end of sophistication, we seek a simple, fun, VMware-branded service provider interface. As a typical customer, consider Jason, the owner of a car dealership. His business is rapidly growing and he’s on top of it, with precise projections for demand growth, number of employees, how many deliveries per week he will receive, how much he expects to spend on advertising, web promotions, and so on. Jason has no background in IT, doesn’t have time to learn it, and doesn’t have the budget to hire an IT person. The Business Intelligence system is a front end that requires only business information and budget. It outputs a recommended IT system scaled to his needs within his cost constraints. The output can be a recommended buy list, or it could go directly into a service provider system. This is like Amazon’s cloud service, but will use VMware expertise to provide a better user experience.

What Does Business Intelligence Automation Look Like?

The business intelligence system must be a lightweight, web- or mobile-based experience. It asks for the type of business, number of employees, security needs, and budget. Additional buttons in the UI provide optional advanced actions.

How is Business Intelligence implemented?

The back end uses VMware best practices and provisioning knowledge to create a fixed set of configuration templates for small datacenters. Sales Engineer knowledge comes from sources such as the “Journey Assessment Tool”, here:

http://www.vmware.com/files/include/journey_assessment/

and the “Why VMware? Calculator” (referenced later in this paper). For a fully automated solution, VMware technologies such as Service Catalog, vService Manager, vCenter Orchestrator, and Autodeploy come into play. We limit the number of templates visible to the user to 20 or fewer, because large IT catalogs can be overwhelming for new users. We want something more detailed than “Gold-Silver-Bronze”, but simple

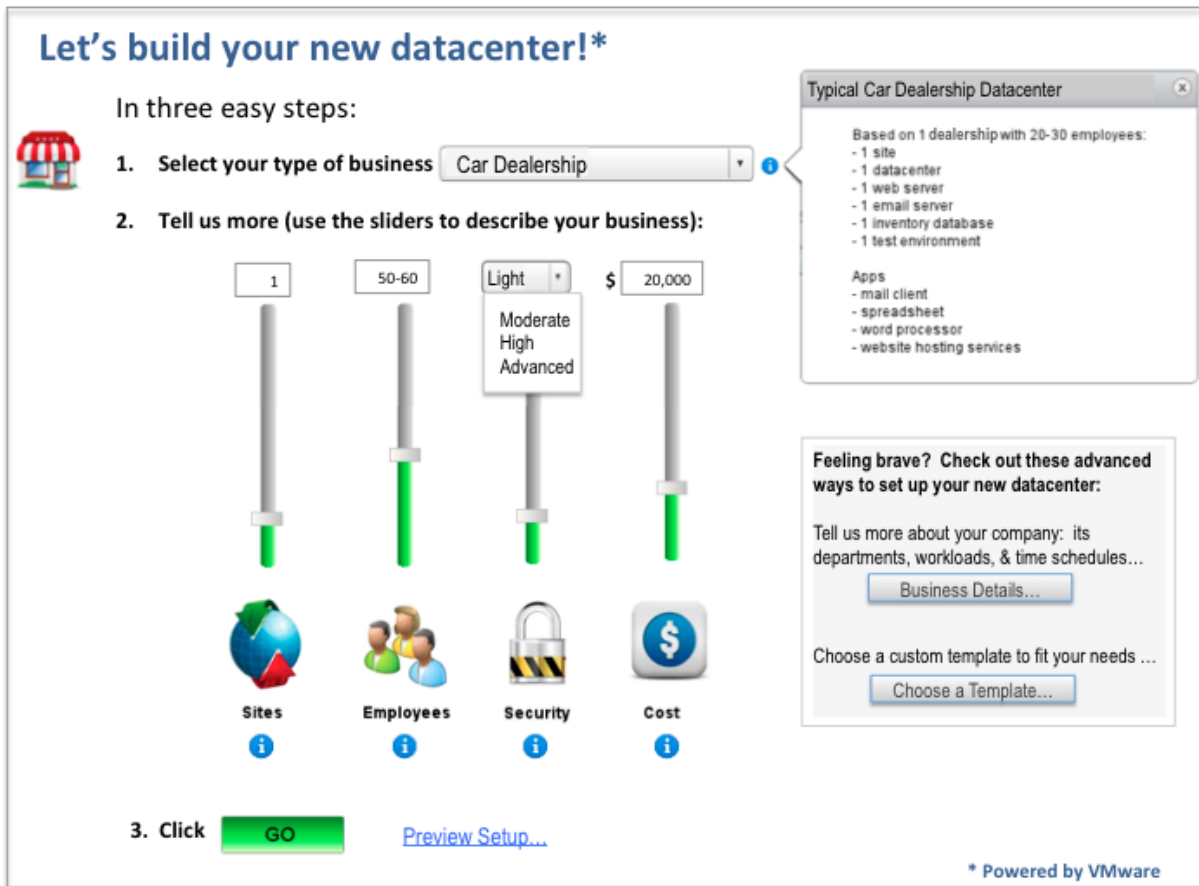


Figure 1 - Business Intelligence Automation

enough to fit into one web page (a rule of thumb that user research has shown to work well). At the Assistive and Self-Revealing levels, customers will have the ability to create and save new templates.

Figure 1 illustrates the simplicity of the Business Intelligence UI – a friendly style with simple elements for describing the business – popup contextual help, advanced options, a setup preview, and a GO button.

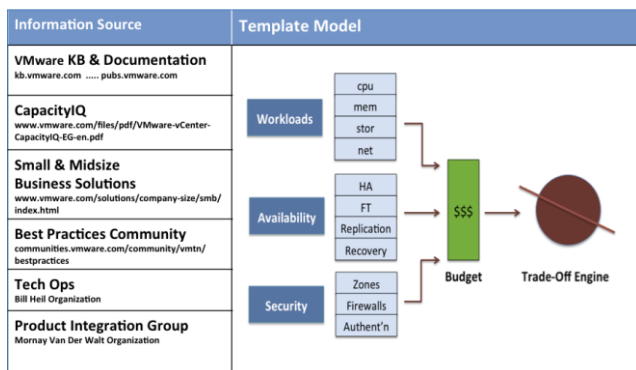


Figure 2: Proposed Automation Model

Figure 2 shows a sketch of a possible automation architecture. Technical knowledge is combined with resource needs and budget, and a trade-off engine determines the best template for a particular customer.

Customer Story 2. Assistive Automation

At an intermediate level of IT sophistication, consider the manager of a small datacenter, whose skills include masters work in system administration and 5 years of datacenter management experience. Let's call this person Stephanie. She person understands the impact of different kinds of resources on operational efficiency and reliability, but has never provisioned a virtual datacenter. Stephanie's task is to virtualize an entire corporate division. She's excited about the idea of consolidation, and hopes to save money in some areas so she can afford high-end computing in others.

What Does Assistive Automation Look Like?

The assistive interface starts with the templates of the Business Intelligence interface, then adds the ability to modify resources for various workloads, security, and availability requirements. It is basically a sandbox where any IT admin (junior or seasoned) can try out various scenarios.

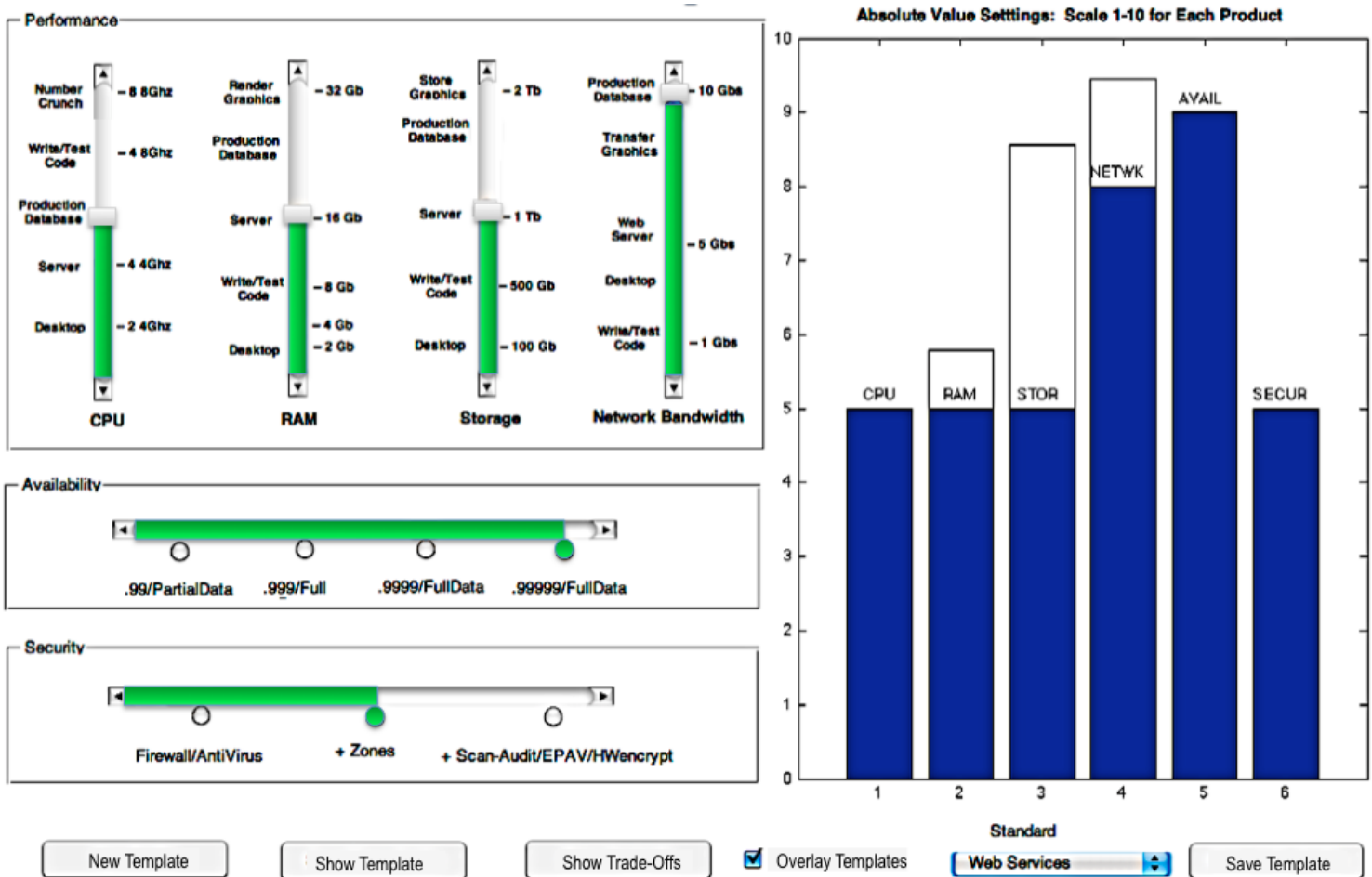


Figure 3 - Assistive Automation

Operating a datacenter requires trade-offs which control the relationship between resources and performance. Let's say, for example, that a basic template assigns 10 accounting employees one cluster with a total CPU capacity of 12 GHz. Stephanie knows that the end of the month payroll period generates very high loads, and she is willing to pay more money for an extra 8 GHz to cover peak loads. This interface allows her to run that scenario in a virtual environment assuming vMotion, to evaluate its effectiveness and cost savings. The Assistive UI with a mobile implementation would also be a great customer-facing tool for VMware Sales Engineers and service provider partners.

The user experience advantages of the Assistive UI include:

- A strong visual representation of best practices
- A strong visual representation of tradeoffs
- Easy to justify expenditures to management
- Teaches virtualization concepts
- Scalability as the company or organization grows

Figure 3 is a screenshot from the current version of the Assistive interface implemented in Matlab™.

Compute, Memory, Storage, and Network resources have vertical sliders labeled on the right in units of quantity (GHz, TB, etc.) and on the left in workload categories (Desktop, Web Server, etc.) The quantity units may be considered continuous metrics whereas the workload units are discrete. The correspondence between quantity units and workload units is determined by the parameters of the underlying model.

The horizontal sliders show Availability and Security settings. They are horizontal to distinguish these parameters from resources, which are shown vertically. The slider values are levels of service in categories provided by Sales Engineers.

The overlay view on the right shows all six elements (Compute, Memory, Storage, Networking, Availability and Security) together in a single view. Here the user can tweak, play, and run scenarios. The prototype offers two main ways to play with the overlay view:

- **Fixed-budget tradeoffs:** Changing the amount of one element redistributes financial resources across the remaining elements.
- **Industry benchmark comparisons:** Select an industry-standard workload to overlay those values on the user's choices.

In this figure, Stephanie has chosen the Web Services standard (shown in dark blue in the graph). Her settings (shown in white) are not aligned with industry benchmarks. Additional elements not shown here will include the ability to enter budget for computing cost trade-offs.

How is Assistive Automation implemented?

The automation architecture would be identical to the one shown in Figure 2 for Business Intelligence, with the output at this level being a continuous function of resources, security and availability instead of a discrete choice of a template. A complete software architecture is beyond the scope of this RADIO paper, but here is what we envision. VMware products, documentation and sales professionals embody algorithmic knowledge that can be built into an Assistive interface. The following represent a small sample of what is available internally:

Capacity IQ - An important part of the vCOPS monitoring product. CapIQ builds models of capacity usage and provides the ability to extrapolate provisioning needs into the future. <https://www.vmware.com/files/pdf/VMware-vCenter-CapacityIQ-EG-en.pdf>

Cost-per-Application Calculator - An existing service provided in the "Why VMware? Calculator" tool, here: <http://www.vmware.com/technology/whyvmware/calculator/>

Figure 4 shows output of the calculator ...

vCloud Director Placement Engine - Go here for details: <http://download3.vmware.com/vcat/documentationcenter/index.html>

"5.4.4 vApp Storage Placement

1. For fast provisioning-enabled virtual datacenters, find a datastore containing a base disk. If a base disk for the virtual machine exists, place a virtual machine on that datastore. The following conditions apply if the target datastore is reaching yellow or red disk thresholds.
 - If base disk exists but target datastore exceeds red threshold, look for a normal or yellow-threshold datastore. If no suitable datastores are available, the operation will fail.
 - If base disk exists but target datastore exceeds yellow threshold, look for a datastore that has not reached its yellow threshold. If none exists, deploy on the target datastore if capacity is sufficient.
2. If no base disk exists, place the virtual machine on the datastore with the most available capacity that does not exceed yellow threshold."

Figure 5 shows the vApp placement algorithm used by the vCloud Director Placement Engine.

Figure 13. VCD placement Engine vApp Placement Algorithm

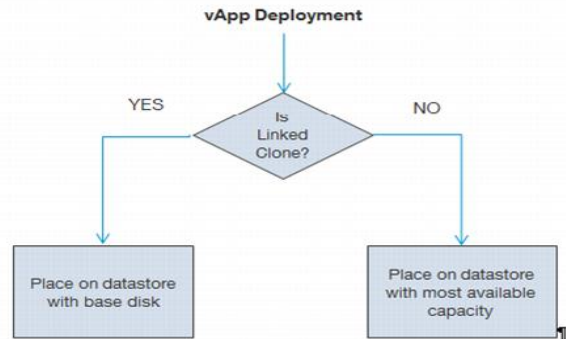


Figure 5 - vApp Storage Provisioning

Table 1 – Detailed Cost-per-Application Comparison

As VMware vSphere 5.1 enables greater application density, it requires fewer servers to host your applications, resulting in a lower cost per application.

	VMware vSphere 5.1			Citrix
	Enterprise Plus Edition	Enterprise Edition	Standard Edition	XenServer 5.6 Advanced Edition
Number of applications virtualized*	100	100	100	100
Number of VMs per host	29	29	29	24
Number of hosts	4	4	4	5
Infrastructure Costs	\$113,570	\$113,570	\$113,570	\$128,204
Software Costs	\$80,613	\$73,347	\$52,414	\$51,195
Total Costs	\$194,183	\$186,917	\$165,984	\$179,399
Cost-per-application	\$1,942	\$1,869	\$1,660	\$1,794
Cost-per-application Savings	-8%	-4%	7%	

* If management software is deployed in virtual machines, includes management VMs.
 Note: A minimum of 2 virtualization hosts is required for high availability or live migration.

Figure 4 - Output of "Why VMware?" Cost Calculator

Customer Story 3. Self-Revealing Automation

The senior IT administrator or architect – let’s call him Dave – is responsible for making sure that the IT services of an entire enterprise or large service provider are implemented and managed correctly, from environmental control through hardware, all the way up to the cloud. Over the last two years we have interviewed many “Daves”, both within the VMware Tech Ops organization and from large customers. Here is what we have learned about the typical “Dave”:

- Resists automation when it comes to decision-making and monkeying with the equipment, but wants to automate simple tasks, and is eager to use the output of automated data analytics to optimize his decision-making.
- Delegates much work but needs a high-level understanding and a holistic view of the operations. Wants junior admins to learn this holistic thinking as well. Feels that too often management tools make it difficult to see the forest among the trees.
- For initial VI provisioning, needs a clear sense of where the hardware is and what it can do.
- Wants a nicely packaged collection of scripts developed by VMware professionals.
- Wants a visual interface for batch-assigning host profiles and getting hosts into clusters.
- For VM placement, Dave wants to see visual metrics showing which clusters can handle additional load, in real time, and which cannot.

What Does Self-Revealing Automation Look Like?

This interface offers a detailed dynamic view of instantaneous state of Dave’s inventory, including hardware. Special emphasis in our prototype is on the *provisioning* use case where cluster metrics inform the correct placement of new VMs and vApps.

Aggregate View

Figure 6 shows a “Clusterviewer” plug-in to the vSphere user interface. In the center panel are roll-ups by vCenter and datacenter, where a graphical element represents an average metric for vClusters within the relevant scope. Each roll-up icon has a metric associated with it (cluster CPU, VMs On, Network Throughput, and so on.)

For greater information density, and to see correlations and relationships between related data metrics, Dave has chosen to show related data in icon pairs (e.g., VMsOn & VMSOff next to each other).

Detailed View

In the far right panel, *individual* clusters are shown. Notice that even when shown in a small size, the clusters with anomalous values visually “jump out” of the background. This facilitates both monitoring and provisioning decisions.

Customizing the View

Figure 7 shows that Dave can use right-click to see a list of available cluster metrics and select a new one to map onto each icon.

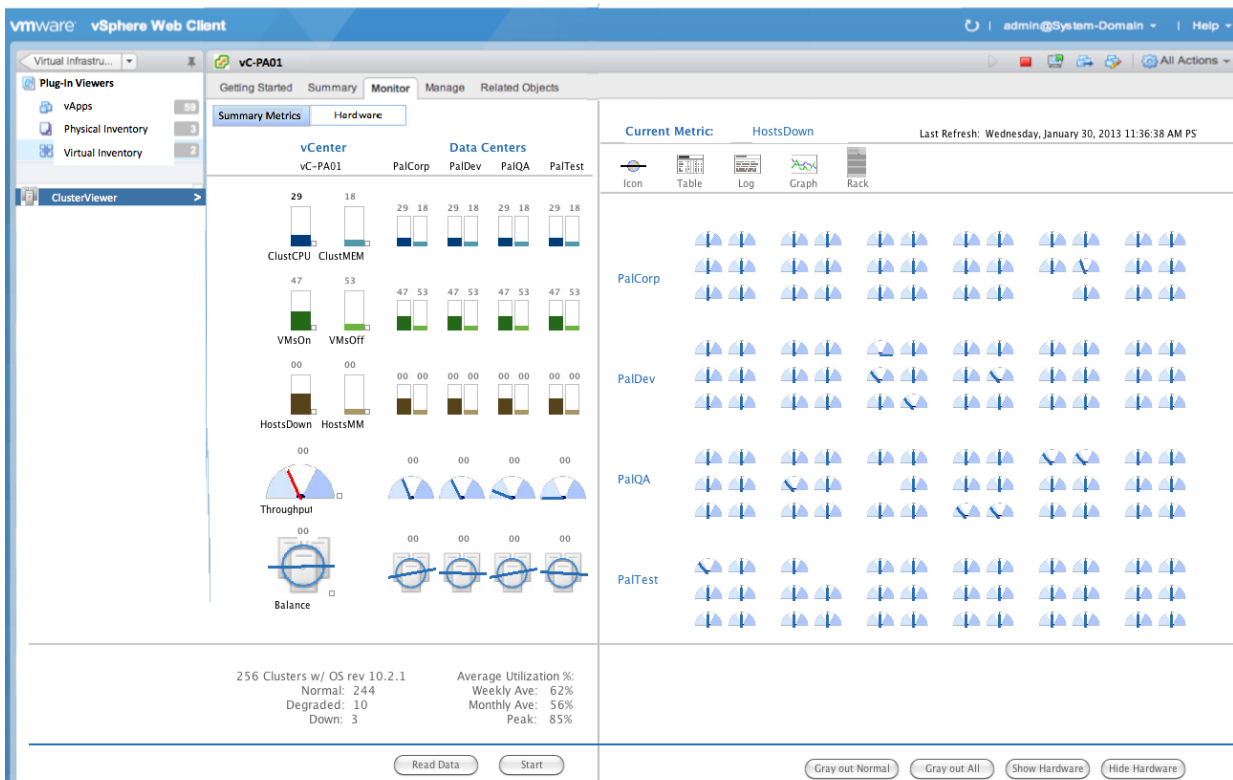


Figure 6: Self-Revealing Automation

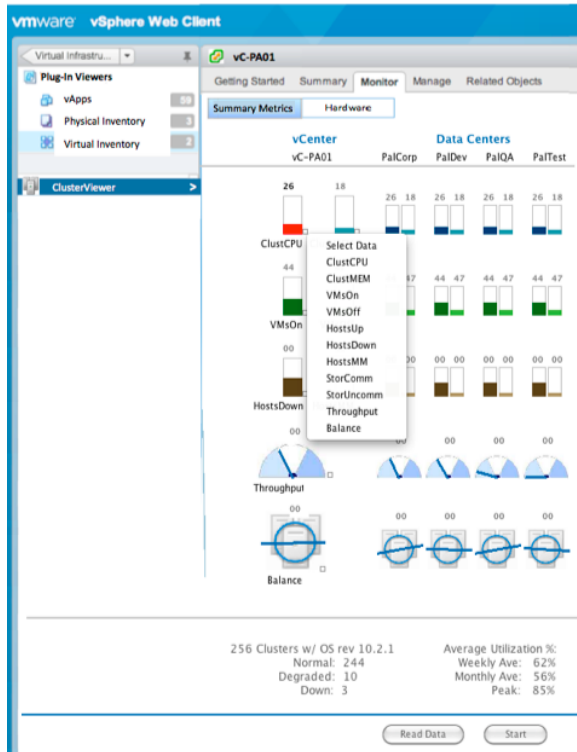


Figure 7: Drop-down Menu

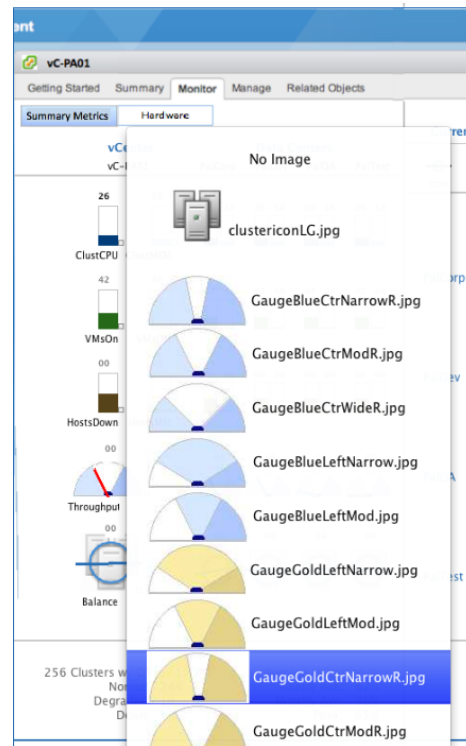


Figure 8: Choosing Metric Icons

Figure 8 shows that Dave can also select the metric icons themselves. It is important for an icon to match the metric it represents. Quantities map well into histograms. Rates map well into speedometer-like gauges. Metric icons are drawn so as to show normal-versus-abnormal ranges. Dave can choose an icon with the appropriate range for easier visual monitoring.

How is Self-Revealing Automation implemented?

This prototype is implemented in Java. A full product implementation would be a straightforward addition to vCenter Operations (vCOPs). Because it is animated, performance will be critical and this view would probably work best as a stand-alone separate from vSphere. A powerful implementation would integrate the assistive and self-revealing views very closely, such that a change in the assistive view is reflected immediately in a simulation of the self-revealing view. This way users can see the impact of trade-offs in the underlying operational reality.

Viewing the Prototypes

Prototypes will be shown at RADIO 2013, and are also available by contacting Helen Cunningham (hcunningham@vmware.com).

Conclusion

Automation in the datacenter is a requirement if VMware wants to be competitive in the cloud. As we have seen here, automation can be complex because one size doesn't fit all. But VMware has a wealth of best-practices knowledge, and an opportunity to put this knowledge to good use in templates with reasonable defaults to entry-level users and small businesses.

Entry level customers will appreciate automated provisioning based on business needs rather than technical knowledge. They can also be lured into a deeper relationship with VMware by playing with scenarios in a safe sandbox environment. Conversely, expert customers who often distrust automation will appreciate the sandbox environment for testing VMware's automation technology before committing to it.

Acknowledgements

We thank Harry Vertelney (hvertelney@vmware.com) for the original idea behind this work ("outcome-based UI"), and Eric Hulteen & Karen Wong who worked on an early version.

Thanks also to Kevin Johnson (former Corp IT Architect) and Randy Keener (former Tech Ops / R&D Liaison) for their valuable contributions to the ideas described here.