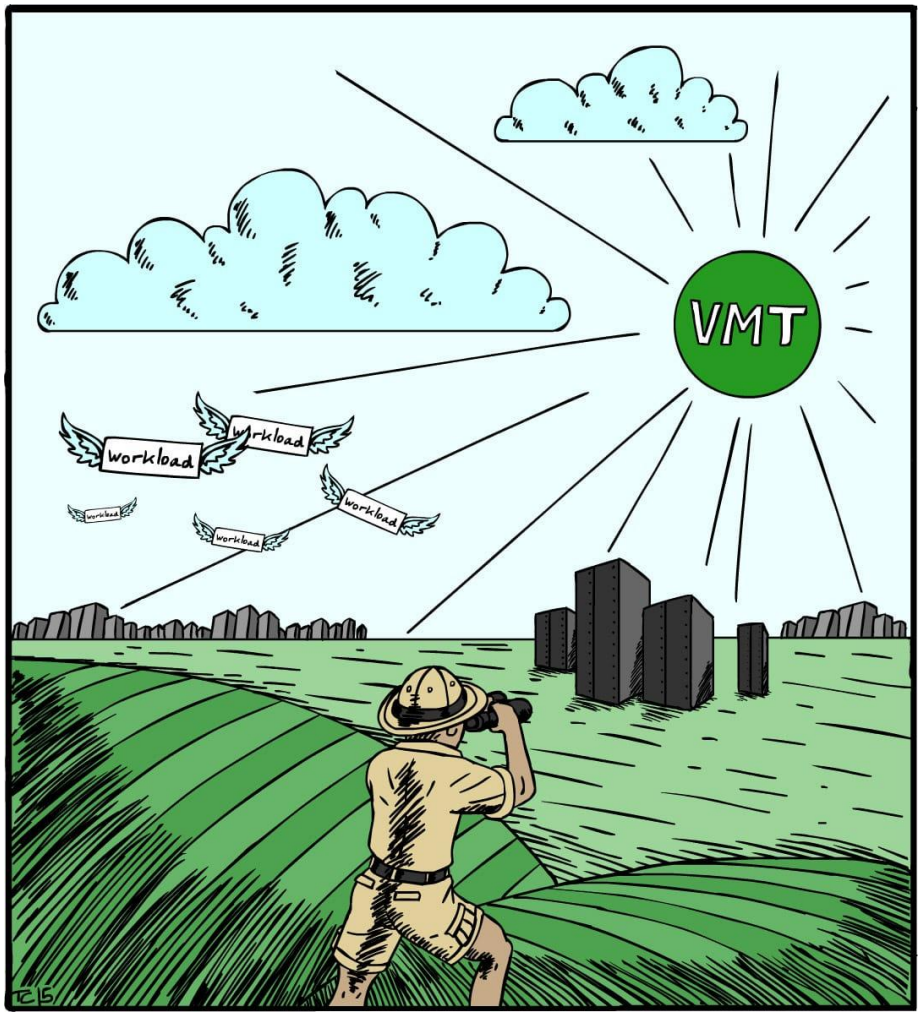
SOLVING THE RIGHT PROBLEM IN THE DATA CENTER



Executive Summary

Organizations are accelerating the adoption of virtualization and cloud deployments to meet the demands of their business and end-users. As they do so, the complexity of assuring application performance increases exponentially. Traditional approaches to virtualization and cloud management simply cannot keep up.

For decades IT has been trying to solve the wrong problem. Operating under the assumption that things must break and then IT will fix them, this break-fix loop precludes scalability, agility, and operational efficiency. When IT is bogged down with alerts, firefighting, and troubleshooting, there is no possibility of smarter, leaner growth.

This white paper discusses the limitations of traditional cloud and virtualization management. It outlines a new approach that enables IT to operate outside of the break-fix loop, preventing degradation in the environment before it starts.

Solving for the Wrong Problem

Since the dawn of enterprise applications IT management has focused on monitoring the data center. The standard mode of operations: detect anomalies, alert operators when the environment is in a “bad” state, and then force operations to drill down and analyze metrics, trends and reports to troubleshoot and identify the root cause. This series of events does not assure performance.

Enterprises have accelerated their adoption of virtualization and cloud deployments. The complexity of these environments and volume of data has grown exponentially. Today’s operational teams are continuously firefighting and finger pointing, generating unpredictable quality of service, if at all.

These labor-intensive, time-consuming activities are trying to solve the wrong problem. Because the focus is on figuring out what caused performance issues, performance issues by definition have already occurred. The environment must degrade into a “bad” state before an alert is sent. This way of thinking inherently *does not* assure application performance.

It’s time for a different way of thinking.

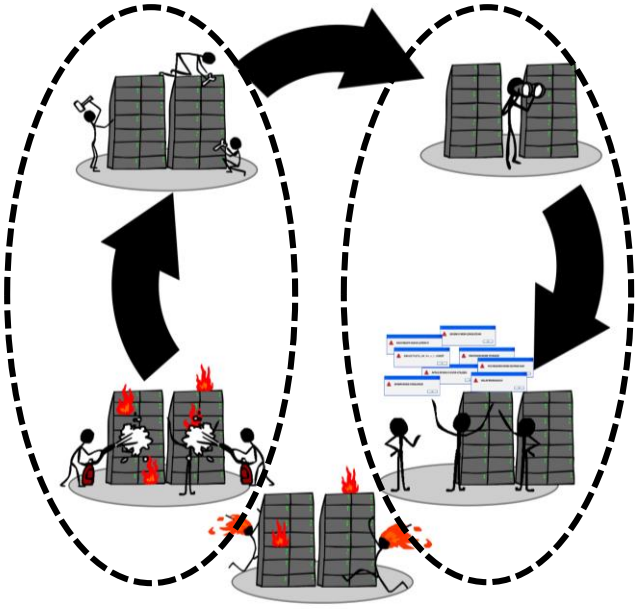
The Break-Fix Loop: Viewing and Doing

The sequence of monitoring, alerts, troubleshooting and root cause analysis is a perpetual break-fix loop. Figure 1 illustrates this loop, which can be divided into two separate sets of activities:

1. “Viewing”: Collecting data on the infrastructure, detecting anomalies and generating information for human consumption.
2. “Doing”: Manual/human troubleshooting to decide what to do and manually trigger the execution of automation scripts.

The break-fix loop requires humans to make sense of the data, “viewing,” and then fix what’s broken, “doing.” It is important to note that the troubleshooting step at the bottom of the loop, analyzing the root-cause based on the data collected is a combinatorial complex problem. It cannot be solved in software. Virtual and cloud environments are constantly changing with dynamic end-user demand, making the results of root cause analysis immediately irrelevant.

Figure 1: Viewing and Doing



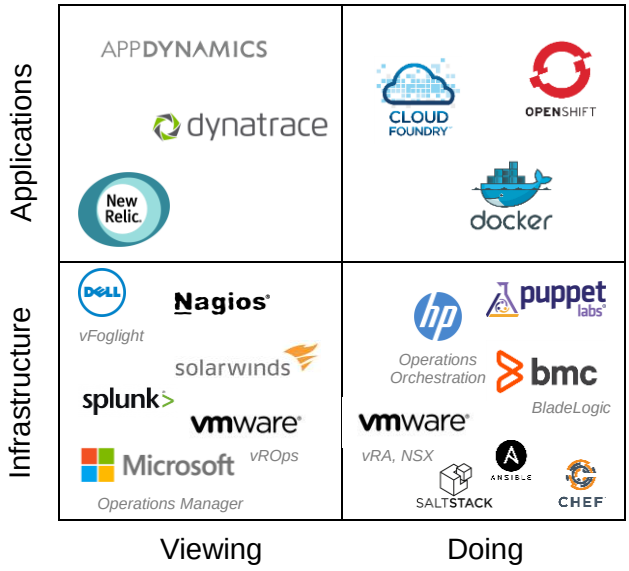
An Industry Rises Around Break-Fix IT

Over the past four decades, the IT management software industry evolved to support this "break-fix" loop where hundreds of vendors arm IT operators with tools for "viewing" and "doing." "Viewing" tools collect data and generate information for human consumption, most of them focusing on reducing the noise generated by the amount of data that shouldn't have been collected in the first place. "Doing" tools provide automated execution of scripts that would otherwise require multiple manual steps. These tools aim to reduce the time required for root cause analysis and mean time to incident resolution.

Some vendors have been attempting to connect their "Viewing" and "Doing" tools through a unified user interface. But connecting these two activities cannot be done in software. Given the complexity of root cause analysis, the heavy lifting and time-consuming troubleshooting are still left to humans.

Viewing and Doing Tools

Figure 2: Viewing and Doing tools across the Infrastructure and Application Layers



What these tools provide is visibility into your environment. Visibility is ***not*** control. They provide trends, reports and alerts for humans to consume, interpret, diagnose, and troubleshoot. It is a reactive approach that inherently fails to assure application quality of service.

Escaping the break-fix loop requires a new way to think about assuring application performance. Instead of waiting for performance degradation to occur, figure out how to prevent it altogether. Preventing problems before they start is a very difficult challenge. It is far too difficult for humans, but perfectly suited for software.

Assuring Application Performance: An N-Dimensional Challenge

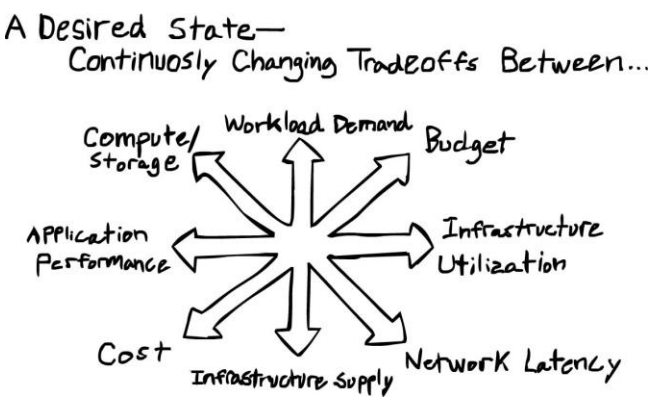
The inherent nature of all entities in a virtualized and shared environment is interdependent and exponentially complex with each additional entity. Applications and VMs demand memory, CPU, storage, and network. The placement of any and every workload to meet those demands changes the availability of resources for all other workloads.

Amid the complex interdependencies and trade-offs is a state in which application performance is assured *while* infrastructure utilization is maximized. This is the Desired State. Finding the Desired State is an N-dimensional challenge. It requires making tradeoffs between resource states. Ultimately an environment's state is an N-dimensional

universe of potential resource states (e.g. memory, CPU, storage, network), workload configuration states and constraints—each with infinite possibilities of states.

The Desired State is a N-dimensional subset of "good" states in which application performance is assured. It is an intersection of N spheres/subsets, each of which is a "good" state with respect to one dimension of the N dimensions.

Watch: [The n-Dimensional Application Performance Challenge]



Solving the N-Dimensional Challenge

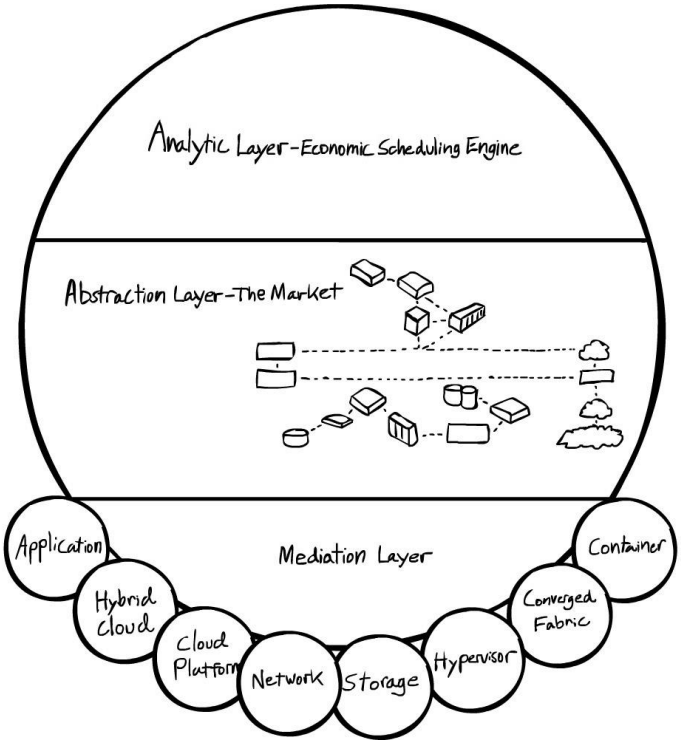
In 1988, Donald Ferguson and Cliristos Nikolaou at IBM T.J. Watson Research Center and Prof. Yechiam Yemini (one of VMTurbo's founders) at Columbia University, suggested in the paper *Microeconomic Algorithms for Load Balancing in Distributed Computer Systems*[1] that applying microeconomic principles of supply, demand and pricing can be used to solve the complexity of load balancing across distributed applications, enabling applications to get the resources they need to perform at scale.

In 2009 VMTurbo commercialized the ideas suggested by Prof. Yemini et al, delivering a platform that controls any type of workload on any type of infrastructure anywhere at anytime. VMTurbo's patented algorithm abstracts the data center and all its interdependent entities into a common data model, a marketplace of buyer and sellers, mapping the end-to-end relationships throughout the entire IT stack and across the entire data center. These abstractions enable application performance assurance that scales.

[1]Ferguson, D., Yemini, Y., and Nikolau, C. Microeconomic Algorithms for Load Balancing Systems in Distributed Computer Systems. June 1988.

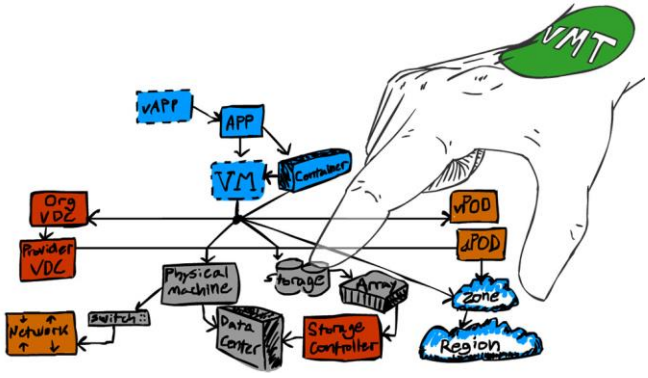
An “Invisible Hand” for the Data Center

VMTurbo understands the data center as a supply chain of consumers (“buyers”) and providers (“sellers”), from end-user service delivery to physical resource supply. The economic abstraction and common data model essentially create an “Invisible Hand” in the data center.



Everything in the data center is a buyer and a seller: hosts, data stores, VMs, applications, containers, zones, etc. The commodities they trade are compute resources, such as memory, CPU, IO, ready queues, IOPS, latency, transactions per second, etc. For example, a host sells memory, CPU, IO, network, CPU ready queues, ballooning, swapping, etc. A data store sells IOPS, latency, and storage amount. A VM buys these resources and sells vMem, vCPU, vStorage, etc. An Application buys these resources and sells Transactions per Second.

The entities independently make purely rational decisions based solely on the price of resources. The price of a resource reflects the state of the environment: the more constrained the resource, the higher the price. For example, if memory on a host is constrained (demand is high), the VMs on that host pay more for it. Price allows an entity to make a



decision based on a complete understanding of the entire stack and everything in the data center.

When you deploy VMTurbo in your data center, the software will understand the application workload demand and provide you with real-time decisions that match the demand with the available infrastructure supply. These decisions can be taken manually or, as most of our customers do, automated.

Watch: VMTurbo Provides Control

In short, VMTurbo enables the entities to independently determine where they should consume resources, if they need to size up or down, and if more or less should be provisioned.

VMTurbo's Application Performance Control System

VMTurbo is an Application Performance Control System for cloud and virtualized environments. It enables operations teams to control their data centers in a Desired State — a state in which application performance is assured at any scale.

The VMTurbo platform controls the application workload demands, as well as the compute, storage and network infrastructure supply in a private cloud, public cloud or a hybrid combination. It assures applications get all of the resources they need to perform—and only what they need to perform.

Control Any Workload on Any Infrastructure at Anytime

VMTurbo controls any type of workload on any type of infrastructure, anywhere, all the time. The workloads can run on any type of hypervisor, compute, storage and/or network and can run anywhere, on premises, in a private and/or public clouds. Whether traditional applications (such as Java-based apps, DBs, etc.) running in VMs or new cloud-ready apps running in containers, the software assures application performance while maximizing efficiency.

The VMTurbo Application Performance Control System frees IT operations from the traditional "break-fix" mode of operations. The software:

- Discovers the dependencies across the entire environment from the workload to the compute, storage and fabric resources.
- Monitors the workload, compute, storage and fabric performance, capacity and utilization.

- Controls compute, storage and fabric resources to assure that the workloads get the resources they need when they need them.
- Continuously and dynamically places workloads on the right compute and storage resources to assure performance, preventing bottlenecks and interferences.
- Continuously considers business constraints, such as D/R and planned outages, to assure workload performance given these constraints.

VMTurbo dramatically simplifies IT operations, eliminating the constant firefighting and finger pointing, while assuring quality of service and efficiency. It controls and deploys virtualized workloads, as well as plans for future fluctuations and changes in the environment.

Conclusion

The VMTurbo Application Performance Control System is the only product in the marketplace today with a patented control system that keeps virtual and cloud environments in a healthy state. It continuously controls the data center, intelligently placing and configuring workloads, allocating and configuring infrastructure resources to meet workload demand fluctuations, and accurately assessing when and how much of a resource is required. The system precludes resource congestion and performance bottlenecks and maximizes infrastructure utilization.

By trusting VMTurbo's Performance Control System in their infrastructure, thousands of engineers have put the burden of assuring application performance while maximizing infrastructure utilization on software. As IT plays a more prominent role in business, they have freed themselves from the break-fix loop to focus on more important projects, partnering with business stakeholders to shape the strategic direction of their organization.

As you think about your strategy for workload mobility, hybrid clouds, PaaS, IaaS, etc., and how to transform your current IT operations in support of this strategy, VMTurbo is the only unified control system that can get you there.

About VMTurbo

VMTurbo's Application Performance Control system enables customers to guarantee Quality of Service for any application while maximizing resource utilization of any infrastructure. VMTurbo's patented technology continuously matches any application workload demand to any infrastructure supply. With this unique, real-time capability, VMTurbo is the only technology capable of controlling and maintaining an environment in a healthy state.

The VMTurbo platform first launched in August 2010 and now has more than 40,000 users, including many of the world's leading banks, financial institutions, social and e-commerce sites, carriers and service providers.

vmturbo.com