

Warehouse-grade at the Speed of Web-scale IT

A Brief on Automating Infrastructure for Hadoop and OpenStack

By Don MacVittie and John Oh
StackIQ, Inc.

Table of Contents

New Realities

Don't Bring a Knife to a Gunfight

Hadoop Is No Panacea (Yet) But Does Enlighten

Warehouse-grade: Much Ado About Something

Automation: Same But Not Equal

What's on the Warehouse Floor: The Parts

Requirements of Warehouse-grade Automation

New Realities

Today's highly integrated systems and customer-centric business environment generate more data than ever before. This massive influx of data is forcing IT to consider new infrastructure solutions for storing, querying, and managing all of this data. This results in today's enterprise and service provider datacenters experiencing rapid change at an unprecedented velocity.

Whether addressing machine generated data, user-collected data, or purchased data, the volume and level of analysis being demanded of IT systems has led to this re-evaluation of the architecture. There is a mountain of data to be created, processed, stored, and analyzed. It is clear to virtually everybody that Big Data is deeply valuable to any organization seeking new insight to drive growth. The endeavor to gain this competitive insight has been elusive for the vast majority of organizations. Until now.

Don't Bring a Knife to a Gunfight

New datacenter technologies have emerged as a result of the additions of "big data systems." In this context, one can view the Big Data trend as a forcing function for datacenter architects and operators to rethink how they envision, construct, deploy, and manage this new infrastructure.

But, wait. Do we even need a new type of infrastructure? To address that fundamental inquiry, let us peer into the vast differences between a general purpose server farm and a large-scale distributed computing system.

The two words most frequently mentioned when facing a mountain of data to process and analyze are: speed and scale. The infrastructure must rapidly make necessary changes - with near zero downtime - while scaling easily. Some common traits of an infrastructure that supports Big Data systems are found below, in no particular order:

- **Require optimal workload management**
- **Aggregation of pooled resources (for compute and data intensive apps)**
- **Support for heterogeneous hardware**
- **Shared cluster capability for multiple user groups and sites**
- **Easy integration into automation workflow**
- **Adherence to compliance and corporate governance**
- **Workload-driven dynamic infrastructure (at node and cluster levels)**

Big Data compute has more similarities to traditional technical computing such, as HPC (high performance computing) than it does to datacenter server farms - whether physical or virtual. Many reasons cause this similarity but, namely, the requirements for high performance metadata, rolling upgrades, “on the fly” node repairs, and extreme scalability are just some top level drivers. Analytics also push the boundaries for I/O while demanding longer running workloads.

In contrast, traditional infrastructure - and the servers that reside in it - is designed to be optimal within a constrained setting. That is, general purpose server farms are built to grow slow. Some servers will be highly utilized while others sit idle. In cases of spikes, there is a load balancer playing the role of traffic cop to even out the requests to those back end servers. For Big Data services, the underlying cluster is optimized for performance at all times. Moreover, as long as it is working properly, the cluster is rapidly scalable since that is a primary benefit of a clustered system.

Technical computing is no longer just the fancy of large, scientific problems. Those large problems have now been introduced to businesses. We are in a new era of data science in business. And, the explosion of data has to be harnessed for business advantage. That much is certain. Enter Hadoop and private cloud.

Hadoop Is No Panacea (Yet) But Does Enlighten

In a groundbreaking paper called *Datacenter as a Computer*, two authors from Google painstakingly detailed their perspectives on building a “Warehouse-scale computer” out of datacenters. The notion of pooling resources, making those assets portable, and striping smaller, or lighter, units of workload across a distributed design impressed many IT operators. But, only if they had the same magnitude of challenges that Google faced with regard to data - in both volume and speed. Fast forward half a decade and the same challenges early Internet companies faced are now increasingly common in the enterprise market.

Since the Google paper, distributed data platforms, such as Hadoop, have had a force multiplier effect on the rate of scale-out infrastructure buildouts. The idea of loading commodity servers with commercial Linux to efficiently support such a demanding Big Data platform as Hadoop is now commonplace. The only meaningful differences across companies is in the size and nature of the production Hadoop project itself (i.e., internal IT app vs. customer-facing Web app).

What follows in our paper is a continuation of that seminal Google work, albeit with a focus around cluster level automation. Thus, this paper places emphasis around the complexities of provisioning and configuration tasks as datacenters increasingly house “Warehouse-grade everything.” Second, whereas the Google paper was penned at a time when Big Data and hyperscale were largely the purview of the largest Web companies on the planet, Warehouse-grade reflects today’s reality of the same needs and ambitions having widely penetrated both enterprise and service provider segments.

Warehouse-grade: Much Ado About Something

We describe the use cases of highly distributed systems running atop large-scale clusters as Warehouse-grade in scope and scale. Similar to physical warehouses and manufacturing factories that are automated, we believe that datacenter operations teams are increasingly viewing their own environments as compute warehouses, big data warehouses, cloud warehouses, media warehouses, and the like. In short, these IT ops teams are building Warehouse-grade infrastructure inside datacenter factories.

In stark contrast to developer-grade configuration tools, Warehouse-grade automation tools are built from the ground-up for IT operators (Figure 1). It reduces the reliance on time-consuming coding skills (which itself goes countergrain to automation) and places greater emphasis on an “off the shelf” solution to get a cluster up-and-running quickly, easily, and reliably.

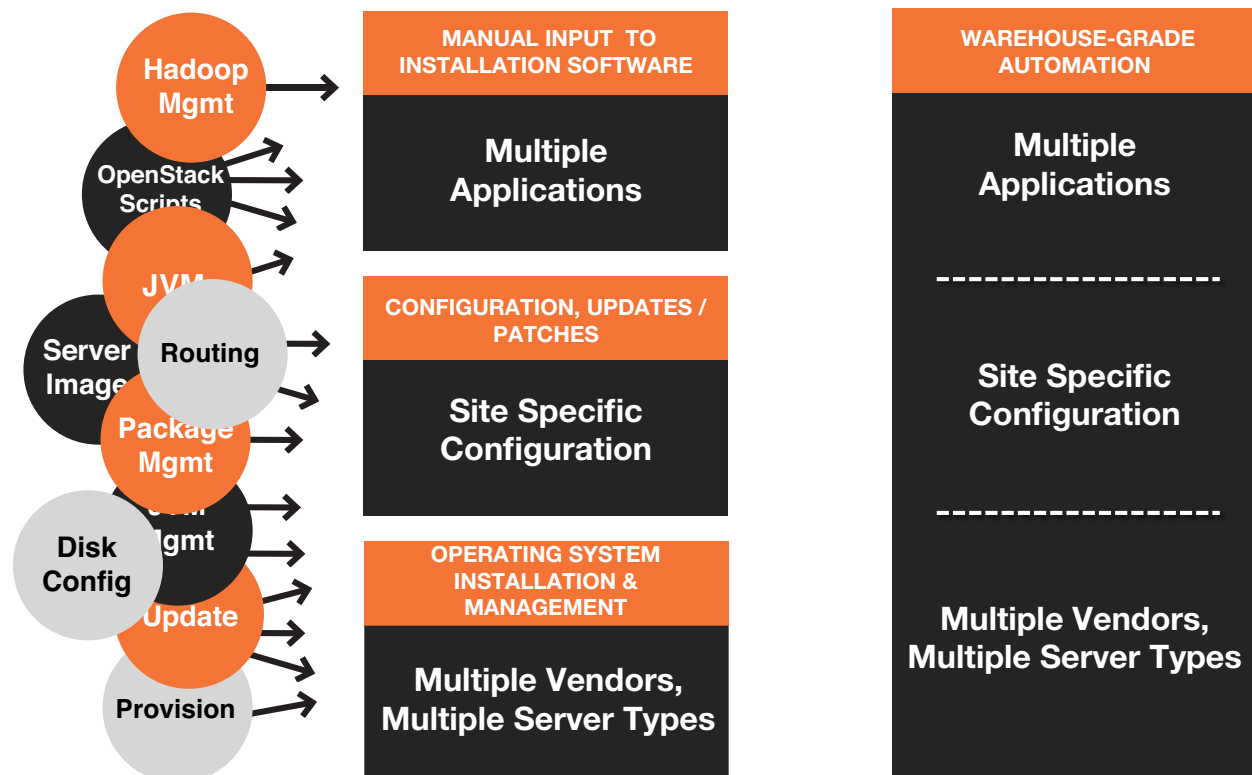


Figure 1: Clustered servers vs. general purpose servers

Automation: Same But Not Equal

Automating the installation, deployment, and management of general-purpose servers is time intensive. Achieving the same for clustered servers that run today's most compute-intensive, distributed systems is orders of magnitude more complex and time consuming.

This reality is attributable to three major factors. First, clusters are designed out-the-gate to scale to many hundreds or thousands of servers (or nodes) in a shorter span of time than their general purpose counterparts. Manual entry and working with a slew of partial automation tools combined with time-to-deploy demands exerting pressure on actual automation speed frequently cause errors.

Second, the system is an entire cluster of servers – more complex to automate than either a single application or a single server. Whole cluster performance hinges greatly on consistency and availability of each inter-dependent node. Errors that manifest as cluster performance degradation are harder to troubleshoot than errors on a single server, and thus, consistent configuration of the infrastructure becomes that much more of a critical imperative. Without this consistency, the cluster experiences severe performance degradation and cascading compute failures.

And, third, commercially available versions of Big Data platforms (e.g., Hadoop, NoSQL) and private cloud frameworks (e.g., OpenStack) are rapidly maturing but still saddled with traits found in most emerging technologies. That is, these next generation platforms come with powerful capabilities but have a heavy site-specific dependency. Further, these new scale-out servers must also adhere to the processes and policies that a typical large enterprise or service provider requires. There are hundreds of steps involved in setting up a bare metal, virtual, or containerized infrastructure for all of these platforms.

The sheer magnitude of the tasks involved in standing up this type of scale-out infrastructure very quickly becomes daunting – and unprecedented (Figure 2). Hence, there is a need for a new kind of “warehouse-grade” automation to solve the challenge of stamping out clusters reliably, easily, and quickly – that is, to do so at the speed of Web-scale IT.

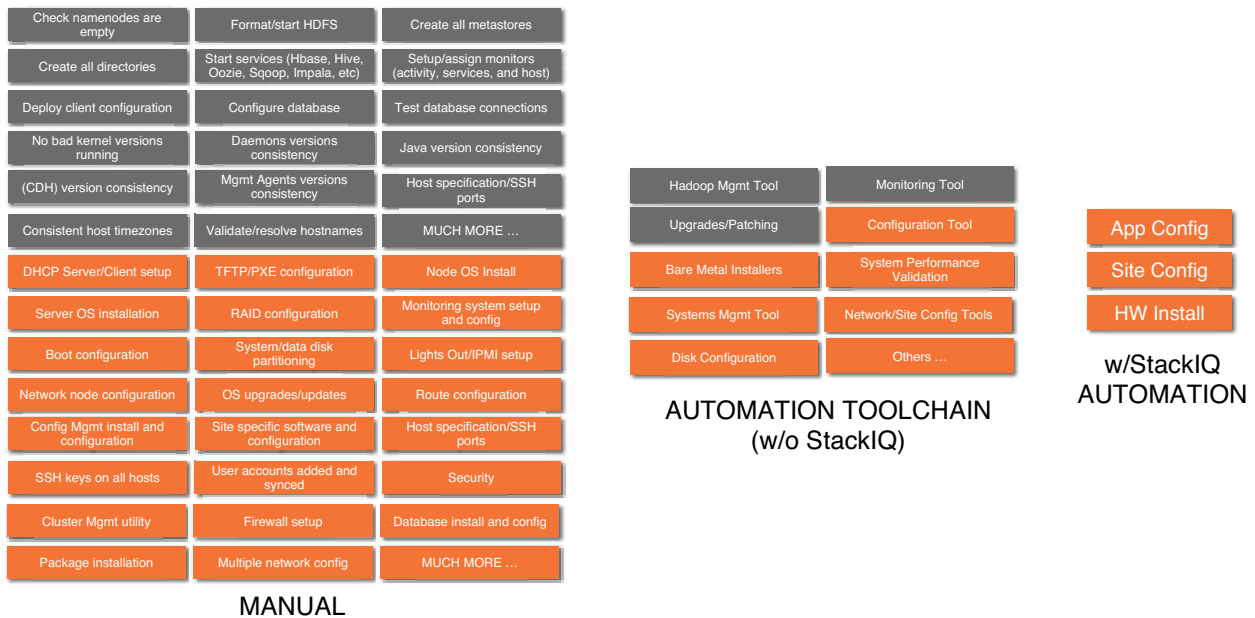


Figure 2: Steps involved in automating Hadoop infrastructure)

StackIQ Boss is a complete automation platform built specifically for any Big Data and private cloud infrastructure. The *Boss* product (formerly called Cluster Manager) is comprised of three parts that – together – deliver an easy, complete and agile infrastructure (Figure 3).

What's on the Warehouse Floor: The Parts

Before we get into the requirements that make up the warehouse-grade automation platform, let us spend a moment on the enemy of automation - sprawls.

If you are spinning up a high value cluster, you want to be on high alert for all types of sprawl. There are three types of sprawl to avoid as if your life depended on it.

Image-sprawl:

Golden images or VMs. While scaling your cluster, it will be difficult to manage the proliferating image libraries that come from images for each type of server. Instead, ask your solution provider if they have the flexibility of offering both package- and image-based installers. Package-based methods allow for dynamic automation and let you avoid image-sprawl.

Toolchain-sprawl:

If you are like most IT shops, you have many types of automation tools. This toolchain has not been perfect but it is still adequate in terms of letting you provision general purpose servers over the years. You are not in love with it, but through familiarity and time, you would never characterize your toolchain as an enemy.

For Big Data clusters in bare metal or cloud, you cannot have multiple tools managed by multiple teams trying to keep things running. The cluster is a tightly coupled interconnected family of servers that either (a) scales fast or (b) degrades fast. It acts as a single unit dependent upon each of its massive subsystems and servers. Thus, you must automate the entire stack as a single step - from bare metal installation all the way up to (and through) application and site configuration. Do not get trapped in the toolchain-sprawl because it feels familiar. Preserving your team's workflow for server provisioning is smart but you should incorporate a complete cluster automation platform into the mix. Do not keep generating new scripts for every install, upgrade, or configuration change. The term to look out for is: FULL STACK automation.

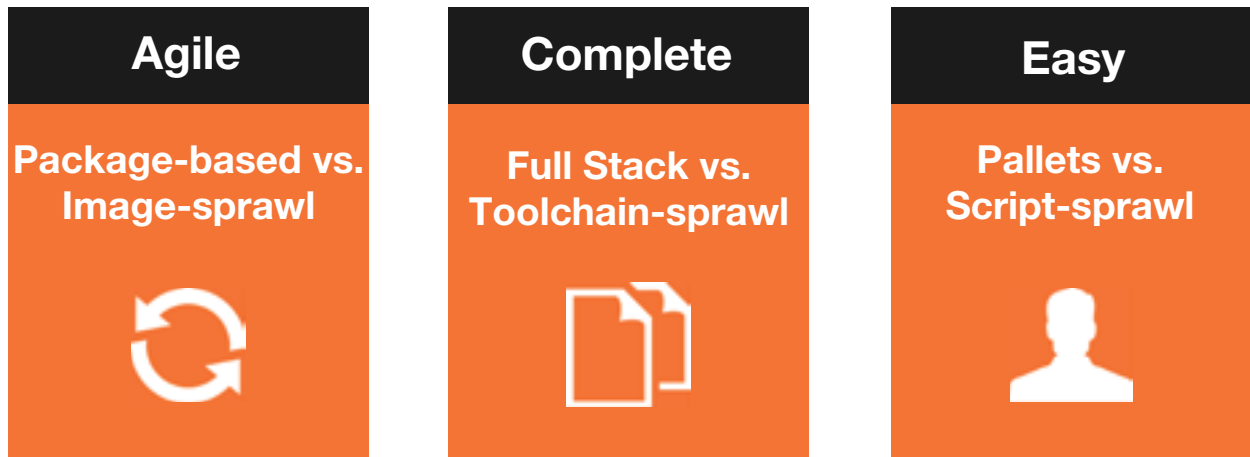


Figure 3: Benefits of Warehouse-grade Automation

Script-sprawl:

Replacing manual scripting with automated scripts makes zero sense if you have to write lots and lots of automation scripts! And yet, we see vendors and customers alike get caught in this treadmill. “Write it once, run it everywhere” sounds catchy if you are into marketing gimmicks and cute one-liners. But, script-sprawl for cluster automation is a highly dangerous habit to get into. Your cluster will simply reach production outages at an alarming rate as we have seen with one of the largest banks in the world. Ask your vendor if they have some way of rolling out pre-built configuration packs that come ready-made with settings baked in. Our own versions of these pre-built configuration packs are called *Pallets* - in keeping with the warehouse-grade theme - for their ability to move application and platform-specific auto-configuration details into our base product called *Boss* (the automated warehouse).

Requirements of Warehouse-grade Automation

1. Complete system installation and management.

A Warehouse-grade automation system must take cluster installation and management together as a single set of actions. Installation should be complete from beginning to end with management providing a view of all that is occurring on the cluster, no matter the machine or service that needs attention.

2. Cluster growth aware.

Clusters grow. From pilot to production, from production to additional capacity; they continuously evolve. Warehouse-grade automation must enable these changes at the same rate that it enables initial installation. Additionally, due to the large number of servers that are brought into most clusters at once, the system must have both automated and parallel installation. All systems to be configured or updated must be delivered at nearly the same time with no user interaction required to qualify for Warehouse-grade Automation. Any other solution stretches out the time window for initial implementation, planned growth, and planned upgrades.

3. Adaptable to the environment.

Warehouse automation requires fulfilling customized site requirements due to the policies, procedures, and existing environment, so must Warehouse-grade automation adapt to the needs of the organization. Security policies, minimum software versions, existing scripting toolsets, and other automation tools that exist in the environment. All must be taken into account. A Warehouse-grade automation system must work in a vacuum – assuming nothing – but must also be able to take advantage of existing automation tools, accept custom security requirements, and ensure that policy is conformed to. The best Warehouse-grade automation tools make this process as painless as possible.

4. Real-time monitoring of overall cluster performance.

The monitoring software included in Big Data toolsets can inform admins about the performance of queries and the stability of subsystems, but a cluster is more than that. A good hardware management platform can inform admins about the state of the hardware for any given machine, but again, a cluster is more than a collection of machines. Warehouse-grade automation must inform admins about the overall performance of the pooled hardware, and allow drill down to discover the source of problems. This functionality, when partnered with the monitoring provided by Big Data vendors, can then offer a complete view of what is happening on the cluster at any given moment. If the Big Data monitoring tool tells you that subsystems deployed to server X are having degraded response times, the Warehouse-grade automation tool needs to inform admins of the status of the physical architecture of server X and its performance relative to the rest of the cluster. Together this presents a complete picture of what is happening.

5. Heterogeneous hardware support.

Modern complex systems (Big Data, OpenStack, or HPC) require a large volume of servers and normally require frequent introduction of new hardware – from repairs, new servers added, or system upgrades. A Warehouse-grade automation tool must take these differences in systems in stride and provide tools for configuring complex subsystems, such as RAID arrays and high-end networking cards. The Warehouse-grade automation tool must be able to account for hardware differences without a significant amount of development or implementation changes.

6. Automated upgrade of all layers of the software stack.

A Warehouse-grade automation tool must enable simple and effective upgrades without a large number of scripts to modify, or manual intervention on each machine. A Warehouse-grade tool will allow for bulk or rolling upgrades on demand, with a minimum of development or administration time investment. Due to the volume of potential upgrades across a system, this last part – a minimum of time investment – is critical.

7. Automated reconfiguration at all layers.

When the cluster changes – and it will – Warehouse-grade automation must be able to adapt the installation to accommodate those changes. It is not enough for a Warehouse-grade automation tool to have the ability to add 200 new servers to a cluster without a hitch or a line of code. The tool must also be able to direct the cluster to distribute sub-processes onto those new servers. The distribution of sub-processes is one of the key performance factors of a highly clustered environment, so any tool that does not allow for customization and automation of reconfiguring sub-process distribution is only doing half of the job.

Conclusion

Clusters and clouds are highly complex systems that are increasingly finding a critical role in the enterprise - providing agility and information that previously were not possible. Yet, configuration and management of these systems are more difficult than traditional IT datacenter systems due to their highly interconnected and complex nature.

Effective management of these systems requires more than most IT shops can or want to invest in putting together a customized tool chain and writing sufficient scripts to both deploy and manage the cluster. Warehouse-grade Automation provides the complete solution to installing, upgrading, managing, and expanding these clusters - reducing both golden image repositories and scripting needs, while automating the full-stack deployment of complex systems.

A top-tier Warehouse Grade Automation system can reduce time spent on deployment by as much as 90%, and also reduces time investments for monitoring and maintenance. This is time that IT can spend focused on tuning the cluster or serving other business needs. And full-stack automation means less manual errors, reducing stress and improving cluster reliability and uptime.

Note to the Reader

We very much appreciate any thoughts, suggestions, or corrections you might have on our white paper. We plan to revise this document relatively often and will make sure to acknowledge explicitly any input that can help us improve its usefulness and accuracy. Thanks in advance for taking the time to contribute.



StackIQ.com | info@StackIQ.com | @StackIQ |
StackIQ, 420 Stevens Ave., Suite 100, Solana Beach, CA 92075