Why the operating system matters

Defining the foundation for an adaptable cloud strategy

Operating systems have been around since the beginning of modern computing, but in those early days, the operating system played a secondary role within the infrastructure. For decades, the hardware itself—massive mainframes in dedicated server rooms and laboratories—was the primary consideration. The operating system (OS) was just part of the framework, an efficient way to interact with the physical hardware, the peripherals, and the subsystems. With the shift to cloud-based infrastructure, it feels like we've circled back to that time. Today, the operating system is often considered incidental to the cloud provider and available services.

This approach to your cloud infrastructure is limiting and restrictive. One of the lessons learned from the old datacenter was the danger of brittle spaghetti architectures, and that lesson is being repeated in a lot of multicloud infrastructures now.

What most IT departments (and CTOs and business leaders responsible for strategic execution) really want is an environment that is flexible, adaptable, resilient, and—above all—manageable. And the ability to execute a flexible, adaptable, resilient, and manageable technology strategy is a reflection of your overall operating environment—your infrastructures, your services, and (yes) your operating system.

A brief history of operating systems

It can be helpful to describe what an operating system is.

Think of the entire infrastructure as a stack. The lowest level is the hardware and its firmware, then the operating system, then application services (middleware), and the applications themselves.

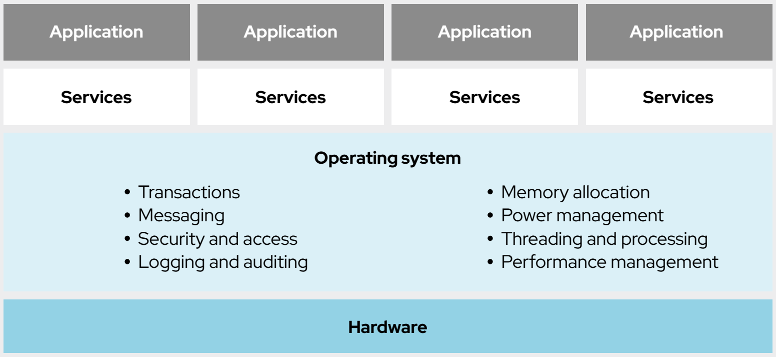

Figure 1. Illustration of an infrastructure stack

The operating system sits between the physical hardware and the layers of services and applications that users interact with directly. The operating system controls all of the housekeeping tasks that applications require—like power management, performance management, threading, logging, and security—not just to run individual software but for all of the applications and services to run together harmoniously.

At one time, that idea of multiple services running all together was a huge innovation. Originally, mainframes were devoted to a single application, and this is one reason the "operating system" wasn't as mature or meaningful as it is today. Brian Kernighan and Rob Pike—the original developers of the Unix operating system (and the entire concept of a portable, separate operating system)—wrote that "the power of a system comes more from *the relationships among programs* than from the programs themselves"(emphasis added).[1]

That idea of abstracting the hardware from the operating layer to the application was incredibly innovative and ultimately helped initiate incredible new developments in hardware technology. At first, the hardware and its operating system were still tightly linked (like Sun Solaris and SPARC or IBM and AIX). But over time, Windows and especially Linux® proved that you could choose an operating system completely independent from the hardware and start building an environment that was specifically designed for your technology needs.

One of the other pioneering ideas behind Unix was the idea of a community contributing ideas, requirements, and actual code to the larger project—the direct precursor of Linux communities.

Starting in the 1990s, there was a sudden growth in open communities around Linux-based operating systems: Debian (1993), FreeBSD (1993), the original Red Hat® Linux (1994), and Gentoo (1999). These Linux distributions had similar underlying technologies and capabilities, but each offered different approaches to development and different types of community interactions. You were choosing more than the technology—you were choosing the type of culture and design that best fit your IT department. As Corey Quinn wrote:

> *You used Gentoo or similar if you didn't value your time, you used Ubuntu (once it came out) if you valued community, you went with Debian if you enjoyed having the crap kicked out of you in IRC channels and mailing lists, and so on. But if you were a business[...], you used Red Hat Enterprise Linux.*[2]

Cloud is redefining infrastructure

As both users and applications have grown more sophisticated, infrastructures have had to evolve—and the interactions within the hardware and software stack have changed. Virtualization and containers abstracted away the hardware, so your applications only had to interact with the operating system or even with just a subset of libraries and services. Cloud has accelerated that shift even more, so that you make fluid choices in (or even ignore) the operating system and services you use and focus on the applications themselves.

[1] *Kernighan, Brian W., and Rob Pike. The UNIX Programming Environment. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1984.*

[2] *Quinn, Corey. "Nobody Cares about the Operating System Anymore." Last Week in AWS, April 2021.*

Public cloud services first launched in 2006, with Amazon Web Services, and that gave IT departments a chance to move beyond the limits of physical infrastructure—the waste of underutilization to manage demand, the difficulty in scaling or deploying applications, and limits on self-service. Public cloud was also supposed to simplify infrastructure provisioning, reducing the need to calculate storage and network requirements or manage server redundancy and failover.

The reality of public cloud is that the changes to IT infrastructure were just as fundamental as predicted, but in different ways.

> "Up until now companies were focused on bits and pieces of the modernization journey. Now every second or third CIO is talking about how their workforce won't ever go back to the same shape. It's become an existential issue; organizations are building for a completely new future. You no longer have to 'sell' anyone on the value of new technologies. It's more about the approach, the underlying principles and practices, that are needed to make the change."[3]
>
> **Ashok Subramanian**
> Head of Technology, UK,
> Thoughtworks

What cloud adoption looks like in reality

Public cloud has been steadily growing in prominence within IT infrastructures. According to Gartner, by 2024, 45% of all IT infrastructure spending will be on public cloud services.[4] In its State of the Cloud report for 2021, Flexera said that 50% of workloads are currently running in the public cloud, with another 7% shifting over within the next year.[5]

These changes are significant, but what is interesting is the dynamic between cloud and on-premise systems. The 2020 Linux Market Study, conducted by Red Hat and Management Insights, found that 62% of Linux servers are still deployed on physical systems, while the rest are deployed in either public or private clouds.[6] Part of the reason for the large physical deployment is the fact that workloads themselves are constantly shifting:

58% of respondents moved a workload from a physical system into the cloud in the past year.

30% moved a cloud workload onto a physical system.

27% "repatriated" a system that had gone from on-premise into the cloud, back onto on-premise systems.

The cloud promised scale, availability, and accessibility. However, these capabilities are not equally important for all workloads—security, management, and portability can be even more critical.

Among our customers, we see that they tend to use cloud environments for proofs-of-concept (POC), development, and scalability,[6] and for these use cases, the specific choice of an operating system is often less important. Over one-third of all cloud deployments use a community Linux distribution,[6] which works well for internal or temporary instances.

As workloads shift to on-premise systems or move from POC into production, other factors like automation, security, and life-cycle management become just as important as accessibility and availability.

[3] *Thoughtworks. "Making enterprise modernization a reality," Feb. 2021.*

[4] *Pettey, Christy. "Cloud Shift Impacts All IT Markets." Smarter with Gartner, Oct. 2020*

[5] *Flexera. "2021 State of the Cloud Report." 2021.*

[6] *Management Insights, sponsored by Red Hat. "2020 Linux Market Study." April 2021.*

When cloud projects fail

Just the act of moving applications and services to public cloud platforms does not guarantee that organizations actually see the intended benefits—like lower costs, simplicity, and speed. According to a study by Accenture, about two-thirds of companies did not see the results they were expecting—and fewer than half were "very satisfied" with the results of their cloud projects.[8]

Failure can mean a lot of things, and it doesn't necessarily mean that a project couldn't be completed. Most companies moved into cloud environments with specific goals for cost reduction, speed for service delivery, or better service levels. Almost everyone met at least some of their targets (96% according to Accenture[8]), but almost no one achieved all of what they intended to do.

The primary reasons for failure were not unexpected: security risks, the complexity around organizational change, and technical debt (or "legacy sprawl" within the survey).

Problems with prioritization, as noted by Thoughtworks, can be a key reason that IT projects fail. According to Thoughtworks research, 74% of all modernization projects fail, and it usually comes down to a disjointed approach, with IT teams focusing on technology and compatibility, while business teams focus on security and strategic initiatives.[3]

Neither technical nor business teams are wrong in this scenario, and their priorities are important. The problem is, as Subramanian put it, focusing on the "bits and pieces" rather than the holistic vision of what the infrastructure needs to be and needs to deliver to be successful.[3]

> "Hybrid cloud is about a capability. It's not about an end state. It's not about having this percentage in public cloud, and this percentage in a private cloud, and this percentage on bare metal. It's about the ability and the capability to be able to move and adapt and adjust as you see fit, and based upon your needs."[7]
>
> **Stefanie Chiras**
> Red Hat

The meaning of the OS is changing

The operating system is, in a sense, the proxy for an overall view of what your IT architecture needs to be. It is a way of defining your overall operating environment, beyond a single server or cloud instance.

Computerworld has periodically run articles questioning whether the operating system really matters, often based on the same premise: the OS matters from a technology implementation perspective, but it isn't a choice that matters at the strategic level.[9] This attitude sees the effect of a disciplined and intentional technology platform strategy, but it misses the cause. Your infrastructure is only as coherent and sustainable as the strategy underlying its foundation.

The primary motivating factor in infrastructure decisions—whether choosing a hardware vendor, datacenter, or cloud provider—should not be the immediate outcome. It should be the balance of current technology requirements, the team culture, and long-term IT strategy. That consideration may have been reflected in what mainframe hardware you chose or network and storage requirements for a given project, but those core cultural principles are what determine the success of the project.

And it is still why cloud projects can fail so easily. The lessons learned in datacenters in the early 2000s are still applicable, even if they can now be more easily obscured.

[7] Red Hat. *"Red Hat's approach to hybrid cloud," Sept. 10, 2020.*

[8] Accenture research report. *"Cloud Outcomes Survey: Expectation vs. Reality," Jan. 2020.*

[9] Hall, Mark. *"Does the OS Matter?" Computerworld, March 17, 2003.*

Your applications, services, and IT teams build on the features and capabilities of your operating system—even if you don't know what that operating system is. It is crucial to know and understand your operating environment, across clouds and services, because your IT infrastructure is relying on it.

As you approach a new project, assess your objectives and examine your operating system's capabilities:

- Do you have to maintain compliance with government or industry standards?
- Are you planning to operate across multiple cloud environments? How will those cloud environments interact with each other?
- How often do you migrate systems versus redeploying?
- Do you have existing applications, and will they be migrated? Or will you need to maintain existing systems in parallel with new projects?
- What are your IT teams' skills? What new skills do you need to bring on?
- What kind of interoperability do you need to maintain with service providers, customer applications, and Software-as-a-Service (SaaS) or third-party applications?
- Do you have a data management or data security strategy? How are cloud workloads storing or accessing data?
- Do you have a defined set of best practices or guidance for starting new IT projects?
- Are you planning to implement DevOps or other agile deployment methods? (Or have you already?)
- Do you have defined and well-understood workflows?
- Do you have centralized teams or is there an expectation for cross-collaboration and communication for new projects?

There is no right answer to any of these questions. These are just some of the factors that go into designing your IT foundation. Also, the operating system itself doesn't answer these questions. Rather, it creates a platform that allows you to develop, deploy, and manage systems and to change and pivot in a way that reflects your organization's priorities and culture.

The power of an IT foundation

When an operating environment is flexible, consistent, resilient, and focused on security, the overall application and user environments inherit that consistency and security—providing a reliable IT foundation for your software. There are three areas that work together to create that foundation:

- Automation
- Security
- Life-cycle management

The specifics of how an operating environment accomplishes those things will be different for every organization, reflecting teams' collaboration styles and culture, strategic priorities, technical debt, and other factors.

Consistency is powerful

Cloud environments are complex. As IT infrastructures start incorporating more distributed computing—from edge deployments to Internet of Things (IoT) to containers—that complexity compounds. Often, maintenance principles developed for datacenters prioritized the ability to isolate access or consolidate services and resources. That approach isn't possible in modern architectures. Hybrid cloud infrastructures are often incompatible with that kind of centralized control.

Defining some best practices can help reestablish control. That first best practice to implement: Simplify where you can.

As organizations have increasingly adopted hybrid cloud infrastructures, they have drifted away from the idea of a standard operating environment, but that approach offers some real benefits. Consistency is a powerful tool to increase productivity, efficiency, and visibility for both the people and the processes within your organization.

Standardization also introduces a key capability: IT automation. Automating routine tasks can boost efficiency by as much as 96% over manual scripting.[10] The efficiency gains are only the start—having a consistent platform across environments allows you to use monitoring and management tools for your entire infrastructure, rather than having a view into only the subset of systems in any given cloud.

Human error is the most common reason for unexpected downtime (causing 49% of known outages).[11] Service interruptions can result from a wide range of unintended oversights and mistakes, including simple errors like misconfiguration or overlooking systems when applying patches and updates.

Simplify and standardize. With a consistent baseline, a consistent operating system and profile, and a consistent set of management and monitoring tools, you can manage more systems more effectively with fewer errors.

Security is the top priority

Roughly one quarter of companies fire senior executives after a security breach, with some industries (like technology and finance) being much more likely to let someone go.[12] High-profile security breaches can lead to government probes[13] or cause financial problems for companies, such as short-term stock drops, lawsuits, or reduced revenue.[14] Cybercrime has shifted from targeting individuals to targeting businesses and infrastructure—something that accelerated during the pandemic, according to Interpol.[15]

Unsurprisingly, according to a Red Hat customer survey, 97% of customers rate security the most valued aspect of their subscription.[16]

[10] *Principled Technologies. "Save administrator time and effort by activating Red Hat Insights to automate monitoring." Sept. 2020.*

[11] *Cepero, Robert. "6 major causes of IT downtime." Bleuwire, Sept. 13, 2020.*

[12] *Swinhoe, Dan. "7 security incidents that cost CISOs their jobs." CSO, Jan. 2, 2020.*

[13] *Mello, John. "SEC reportedly probing SolarWinds breach." TechNewsWorld, June 23, 2021.*

[14] *Kvochko, Elena, and Rajiv Pant. "Why data breaches don't hurt stock prices." Harvard Business Review, March 31, 2015.*

[15] *Interpol. "Interpol report shows alarming rate of cyberattacks during COVID-19." August 4, 2020.*

[16] *Internal Red Hat customer survey, Oct. 2020.*

Security becomes more complex in hybrid cloud environments because of the inability to lock down services behind firewalls and demilitarized zones (DMZ). Data accessibility is critical, and data security (and user access) is decentralized and harder to maintain. This complexity is compounded with the need for rapid data processing to fulfill customer expectations.

The security of your applications and data depends on the security of the infrastructure they are running on. The environment (cloud, physical, or virtual) is less important for implementing security practices than the operating system itself. As a result, consistency and simplicity in your operating environment can have a significant effect—allowing you to implement security best practices across your entire infrastructure. Security is not a single state—it is the combined result of consistently applying different practices across your infrastructure, beginning with the basics:

- Automate routine processes.
- Set clear data management policies. Define how data is shared between applications and environments, who needs access, and where data is stored. Data is the most important part of your infrastructure—gain visibility and control to better manage your data, and you can mitigate and respond to threats more effectively.
- Perform frequent backups. Replacing a compromised system is much easier than trying to repair it. If you suffer a breach, act quickly—disable the compromised system and replace it.
- Be consistent and frequent in patching systems. Red Hat addressed more than 2000 security vulnerabilities in 2020, 58 of them concerning critical issues.[17] Updates are issued almost daily; keeping systems updated mitigates their exposure.
- Security starts with the operating system. Use features like process escalation, root controls, and user authentication to limit access of services and applications. Define your security policies consistently between settings. (Management tools that set baseline configuration and identify drift can be useful for maintaining your security profile.)
- Use integration to define key elements in your architecture. Integration can mean different things in different contexts. Services like Identity Management in Red Hat Enterprise Linux can integrate with Active Directory, allowing integration between user directories, service access, and access permissions. Application programming interfaces (APIs), Apache Kafka, and Apache Kafka Streams (among other technologies) can provide lightweight ways to connect applications and data. Every time data crosses an environmental boundary, there is a risk. Identify those integration points and the best-matching technology to manage that integration to maintain security at those connections.

Scalability is part of your life cycle

Scalability is the number one reason why people first choose to use Linux (49%).[19] When they choose to use Linux in the cloud, scalability is even more of a factor—57% of respondents said they chose Linux because of scalability.[18]

There are architectural models and considerations for scalability for infrastructure, but what scalability really means is the ability to change capacity as needed. Scalability is illustrated by how frequently organizations move workloads to the cloud, between clouds, or back on-premise.

[17] Red Hat. *"Red Hat Product Security risk report: 2020." Feb. 2021.*

[18] *Management Insights, sponsored by Red Hat. "2020 Linux Market Study." April 2021.*

Yet many organizations do not have a life-cycle management strategy for their infrastructure. According to a Red Hat customer survey, 69% of organizations do not have a migration strategy[19]—they only migrate a system when forced to by an end-of-life event for software or to retire hardware.

If you are considering a cloud infrastructure to improve scalability, you need a life-cycle strategy that supports scalability.

Forty percent of organizations, according to the Linux Market Study, are using a standard operating environment as part of their migration strategy.[18] This strategy is usually implemented in combination with provisioning and automation software, containers, or cloud provider services, but the basis still relies on a common operating structure to make management tasks easier to execute.

The key to scalability is having a plan for how you will manage your systems across their life cycle, from deployments to upgrades to retirement:

- Define an upgrade cadence for the operating system, whether that means migration or redeploying.
- Identify whether it's more efficient to upgrade or to redeploy and migrate systems for each workload.
- Have clear continuous integration/continuous delivery (CI/CD) pipelines, especially if development and production environments are in different cloud providers or infrastructures.
- Use golden images or configuration baselines to help with both updating and rolling back changes, as necessary.
- Know where workloads need to run—based on memory, bandwidth, or other requirements—to be most cost effective, and have a plan for when to migrate that workload if necessary.

A defined life-cycle strategy allows your IT teams to effectively scale your systems and maintain them. It's estimated that roughly half of cloud spending is wasted because of over-provisioning instances or running suboptimal workloads.[20] Make migration and workload management part of your cloud strategy to get the most out of your cloud projects and to keep your infrastructure flexible and scalable.

An OS is more than a technology choice

The best practices and challenges that existed in on-premise architectures still persist and are amplified in cloud environments. Infrastructure naturally becomes more complex, harder to manage, and harder to integrate as it grows.

The operating system can be the foundation that solidifies and unifies your IT infrastructure and whatever environments it encompasses.

Nearly half of all server deployments (47%) are Linux,[21] which is largely due to the technical capabilities of Linux, its flexibility, and its customizability. However, Linux is not a monolith. There are dozens of different distributions and communities. The differences are a reflection of the community and culture, not necessarily a line-item comparison of features.

[19] *Internal Red Hat customer survey, June 2020.*

[20] *Fadilpasic, Sead. "Majority of cloud spending is going to waste." ITProPortal. April 14, 2021.*

[21] *Management Insights, sponsored by Red Hat. "2020 Linux Market Study." April 2021.*

Research conducted by Red Hat revealed that customers are looking for a collaborative, knowledge-based relationship with their Linux vendors.[22] Red Hat provides that high level of partnership, guidance, and expertise through its subscription. A Red Hat subscription helps you:

- **Gain support.** Common support services include online and phone support, a knowledgebase for technical issues, and product documentation—but you can expand your definition (and expectations) for support. You should be proactive in how you collaborate with your Linux vendor, working with customer support managers to establish best practices, plan projects, and identify potential issues before they arise.
- **Make your voice heard.** Your experience is invaluable. With Red Hat Enterprise Linux, you can submit feature requests, provide feedback, and check out product roadmaps, so you have direct input into the product development process. Red Hat provides multiple ways to provide feedback directly to engineering teams—everything from user experience design and documentation requests to new tooling and help with using Red Hat products.[23] And changes and enhancements are contributed back to the community, so your experience can create better technology for everyone.
- **Build on available services.** You need to be able to use your technology in a way that is most effective and relevant for your organization. Red Hat provides additional services like customer dashboards,[24] Red Hat Insights[25] for system management and analytics, and configuration tools and labs. These integrated services build on our engineering expertise and allow you to manage your infrastructure with data-driven intelligence.
- **Rely on a larger portfolio and ecosystem.** As you launch new projects or modernize existing systems, you will need other services. Your Red Hat subscription includes a variety of critical applications, from development tools like OpenJDK to security tools like Identity Management. When you need more specialized applications, hardware, or cloud infrastructure, we have an ecosystem of thousands of certified partner solutions that are guaranteed to work with your Red Hat Enterprise Linux systems and are cross-supported by the partner and Red Hat. The strength of Red Hat Enterprise Linux is its community.

The core pillars of IT architecture are processes, people, and technology. With the support of a strong vendor relationship, your cloud projects can achieve more of your intended goals—and then pivot and establish new initiatives as your organization succeeds and evolves.

[22] *Internal messaging research with Management Insights, 2020.*

[23] *Red Hat Customer Portal. "You asked. We acted." Accessed Aug. 2021.*

[24] *Customer dashboards can be found at the Red Hat Customer Portal. Customer log-in required.*

[25] *Red Hat Insights can be found at the Red Hat Customer Portal. Customer log-in required.*

Do more with Red Hat

Modern IT starts with Red Hat Enterprise Linux. Learn how Red Hat Enterprise Linux is used in public clouds, containers, and edge deployments.

As you start planning your next cloud project, talk to a customer success manager or a Technical Account Manager, or schedule a Discovery Session with Red Hat Consulting.

About Red Hat

Red Hat is the world's leading provider of enterprise open source software solutions, using a community-powered approach to deliver reliable and high-performing Linux, hybrid cloud, container, and Kubernetes technologies. Red Hat helps customers develop cloud-native applications, integrate existing and new IT applications, and automate and manage complex environments. A trusted adviser to the Fortune 500, Red Hat provides award-winning support, training, and consulting services that bring the benefits of open innovation to any industry. Red Hat is a connective hub in a global network of enterprises, partners, and communities, helping organizations grow, transform, and prepare for the digital future.

facebook.com/redhatinc
@RedHat
linkedin.com/company/red-hat

North America
1 888 REDHAT1
www.redhat.com

Europe, Middle East, and Africa
00800 7334 2835
europe@redhat.com

Asia Pacific
+65 6490 4200
apac@redhat.com

Latin America
+54 11 4329 7300
info-latam@redhat.com