Transform Your Data Center: Five Steps to Optimization

The data center architect's guide

Table of contents

intel

Facing today's rapid technological evolution, ensuring your data center is up-to-date isn't just a strategic ideal—it's a business imperative.

Continuing to use ageing infrastructure to run increasingly complex and data-hungry workloads can hinder productivity and performance. As efficiency dwindles and maintenance costs increase in the data center, business users can be left unable to access and use emerging capabilities like generative AI that are fast becoming a competitive necessity. Meanwhile, evolving cyberthreats mean it's essential to stay up to date with the latest security capabilities to protect your data assets.

Re-working the foundations of your data center can seem daunting. Architects frequently face the challenge of modernizing core systems while meeting the demands of ever-expanding workloads. However, modernization doesn't have to be prohibitively complex. By adopting a thoughtful and structured approach, you can break down the process into manageable stages, creating a future-ready foundation without compromising day-to-day operations.

intel

This eBook outlines five actionable steps to optimize your data center for scalability, security, and future readiness:

1. **Align IT with Business Goals:** Prioritize initiatives that deliver the greatest business value and ROI.

2. **Audit Your Infrastructure:** Identify gaps, dependencies, and opportunities for improvement.

3. **Choose the Right Strategy:** From retiring outdated systems to leveraging SaaS, select the best paths for optimization.

4. **Address Barriers Early:** Mitigate risks like technical debt, budget constraints, and stakeholder resistance.

5. **Commit to Continuous Improvement:** Regularly assess, adapt, and document lessons learned.

STEP 1

Define Business Objectives and Initiatives

Any data center optimization efforts should support wider business goals and priorities, so start by understanding what these are and how they translate to data center and IT infrastructure requirements. This will also help you prioritize what and where to optimize first.

Mapping business priorities to IT infrastructure requirements

The workloads and applications that your business depends on will each have their own combination of requirements for your data center. These requirements may depend on a variety of factors, such as:

- **Performance targets:** For example, a financial services firm might prioritize reducing transaction processing times, while a retail company could focus on improving the speed and reliability of its e-commerce platform. Meanwhile, a healthcare provider may aim to optimize system uptime to ensure seamless patient care.

- **Scalability objectives:** If the business strategy is to grow significantly, such as by expanding into a new market or geography, then the ability to scale up without impacting the performance of core services will also be important. Or, it may be necessary to allow for seasonal scaling both up and down, if your business operates on a cyclical model with peaks and troughs.

- **Security requirements:** Companies face various security requirements, such as protecting customer payment information in retail, meeting HIPAA compliance in healthcare, or ensuring GDPR compliance in Europe across multiple industries. Prioritize confidential computing capabilities to safeguard sensitive data.

- **Sustainability goals:** Prioritize energy-efficient infrastructure and technologies that reduce your data center's carbon footprint. Align optimization efforts with corporate sustainability objectives, such as minimizing power consumption, adopting renewable energy sources, and achieving net-zero emissions targets.

While each of these considerations is an important priority, the weighting of each will vary depending on your organization's specific circumstances. Use a weighted scoring system to evaluate initiatives based on business impact, urgency, and cost. For instance, assign scores to each initiative based on its expected ROI, alignment with strategic goals, and resource requirements. Sum these scores to rank and prioritize projects objectively.

Managing competing business objectives

▪ **Stakeholder collaboration:** It's common to find that different stakeholders have different, even opposing, views. Foster regular communication and use collaborative platforms to ensure transparency and alignment across teams and business functions.

▪ **Scenario planning:** Develop contingency plans to handle unexpected shifts in priorities. For example, a company undergoing a major application migration could create fallback plans to revert to legacy systems temporarily if critical issues arise during deployment, ensuring minimal downtime.

Evaluate Your Current IT Infrastructure

Once you're clear on your priority workloads or platforms from a business perspective, you can begin to map out where and how you'll need to optimize your data center to best support those workloads. You can do this by auditing your current IT infrastructure.

Conducting an effective technology audit

1. **Inventory your assets:** Document your applications, workloads, and supporting infrastructure, including detailed metadata such as application owners, utilization metrics, and support lifecycles. This provides a clear picture of which assets are mission-critical and which might be candidates for decommissioning. Also, take advantage of any vendor-imposed milestones to drive positive improvements. For example, the end of support for Microsoft SQL Server 2012 and Microsoft Windows Server 2012 is being used by many enterprises as an opportunity to upgrade hardware as they migrate to the newer software versions.

2. **Evaluate dependencies:** Identify data flows, integrations, and interoperability requirements. This involves mapping not just direct connections but also indirect dependencies, such as shared databases or middleware services, to ensure any changes do not disrupt downstream operations.

intel

3. **Audit security measures:** Review your current strategy and policies, including endpoint protections and data encryption protocols. In addition to checking compliance with existing regulations, perform penetration testing to identify vulnerabilities, evaluate your incident response plan, and assess the efficacy of threat monitoring systems. Consider how and where the use of new technologies such as AI will impact your existing security strategy and where additional protections may need to be introduced.

4. **Decide workload placement strategy:** Different workloads perform best on different architectures. Consider whether you need high-performance and optimized compute cores for HPC applications, or for software which is licensed on a per-core basis (many of the most important enterprise applications are). Or you may need to scale many compute cores for applications like web and microservices. It's likely you'll need both, along with a common architecture so that your hypervisor can run them wherever is best.

5. **Consider data gravity:** Large datasets tend to attract applications and services, making data movement challenging and costly. In IT infrastructure planning, especially within multi-cloud environments, understanding your data gravity is crucial for optimizing performance, minimizing costs, and ensuring effective data management.

6. **Avoid over-provisioning:** Power management and provisioning are crucial in data center planning to ensure operational efficiency, scalability, and sustainability, especially if you're using GPUs, which are enormously power-hungry. Match application needs to compute resource so that applications which don't need such powerful cores run on lower-wattage options.

7. **Follow your hybrid cloud strategy:** With data and applications inevitably spread across multiple cloud and on-prem environments, make sure you have a clear strategy and policies in place that will help guide decision making and ensure consistency. This should cover critical areas such as: data placement, workload distribution, governance policies, and disaster recovery planning.

Decide on the Best Path(s) Forward

Strategic objectives and technology audit findings in-hand, you can now start making informed decisions about how best to optimize your data center. There are a number of optimization strategies to consider:

- **Retire:** Decommission outdated systems that no longer serve strategic objectives. For instance, legacy applications that are expensive to maintain and provide little business value can be retired to free up resources for more strategic investments.

- **Rehost:** Migrate existing applications to cloud infrastructure without significant changes. This lift-and-shift approach can quickly improve infrastructure flexibility while keeping application logic intact.

- **Replatform:** Make minor modifications to optimize performance on new platforms. For example, updating database software or tweaking application configurations to leverage modern hardware can provide immediate performance gains.

- **Refactor:** Re-architect applications to leverage modern frameworks and tools. This may involve breaking monolithic applications into microservices to improve scalability and agility.

- **Replace:** Transition to SaaS solutions for improved scalability and reduced management overhead. Replacing custom-built solutions with SaaS platforms often leads to reduced maintenance burdens and faster feature updates.

Criteria for decision-making

Choosing the right strategy involves evaluating several key factors. Use the criteria below to guide your decision-making process:

- **Business value:** Evaluate the ROI of each option. Consider how each strategy contributes to your goals, such as consolidating workloads and reducing operational costs, improving performance of key workloads, enhancing customer satisfaction, or supporting growth.

- **Technical complexity:** Consider the resources required for implementation. Assess the technical skills, time, and costs involved in adopting each option, and weigh these against the anticipated benefits. Seek opportunities to uplevel your capabilities without adding unnecessary complexity or cost, for example by running your first AI inference workloads on your existing x86 architecture where possible to keep down total cost of ownership (TCO).

- **Future proofing:** Align your choices with long-term strategic goals. Ensure the approach you choose positions your organization for scalability, compliance with future regulations, and the ability to adopt emerging technologies.

Identify and Resolve Potential Barriers

It's also useful to be aware from the outset of the obstacles that may hinder your journey. Addressing these upfront minimizes delays and ensures a smoother implementation process further down the line.

Common barriers to look out for

- **Technical debt:** Begin by cataloguing legacy assets and assessing their compatibility with your data center refresh initiatives. For example, ensure any mission-critical assets such as core banking platforms or patient record systems can integrate with innovative tools like AI co-pilots that your organization wants to use. If not, you may need to consider first migrating those legacy systems to the cloud, or conducting other preparatory work.

- **Budget constraints:** Misaligned financial planning can quickly derail projects. Develop a clear ROI model to align any planned investments with expected business outcomes. Start by identifying key performance metrics such as operational cost savings, increased system uptime, or improved workload efficiency. Next, forecast these benefits over a set period, compare them against implementation costs, and present the results in a clear, visual format such as a cost-benefit analysis chart.

- **Stakeholder resistance:** A lack of buy-in from key stakeholders can slow progress. Early and consistent communication is critical to align priorities. Identify your key stakeholders early by mapping all individuals and groups affected by your plans, such as IT managers, finance leads, and end users. Use stakeholder analysis tools to prioritize engagement based on their influence and interest in the project.

Unearthing hidden barriers

In addition to the universal challenges, you're likely to find a few less-predictable hurdles as well. Every organization has its own priorities, cultural considerations and idiosyncrasies, and it pays to spend some time making sure you're aware of these before you get started. Some ways to go about this include:

- Conduct **stakeholder interviews** to understand the nuances of how your proposed initiative will impact different groups within the business. Gathering input from across the organization early helps avoid surprises later while also helping to build trust and engagement.

- Perform **dependency mapping** to reveal interconnected systems that may be affected by changes. For instance, a retail organization looking to update its inventory management system will need to ensure that its POS system is also up-to-date so the two can continue to work seamlessly together.

- Use **proof-of-concept pilots** to identify technical or process issues before full-scale deployment.

intel

Learn and Adapt for Continuous Improvement

Optimizing your data center should be an ongoing initiative. As business needs change and technologies evolve, it's important to keep monitoring and evaluating what these developments mean for your technology stack and how you run it.
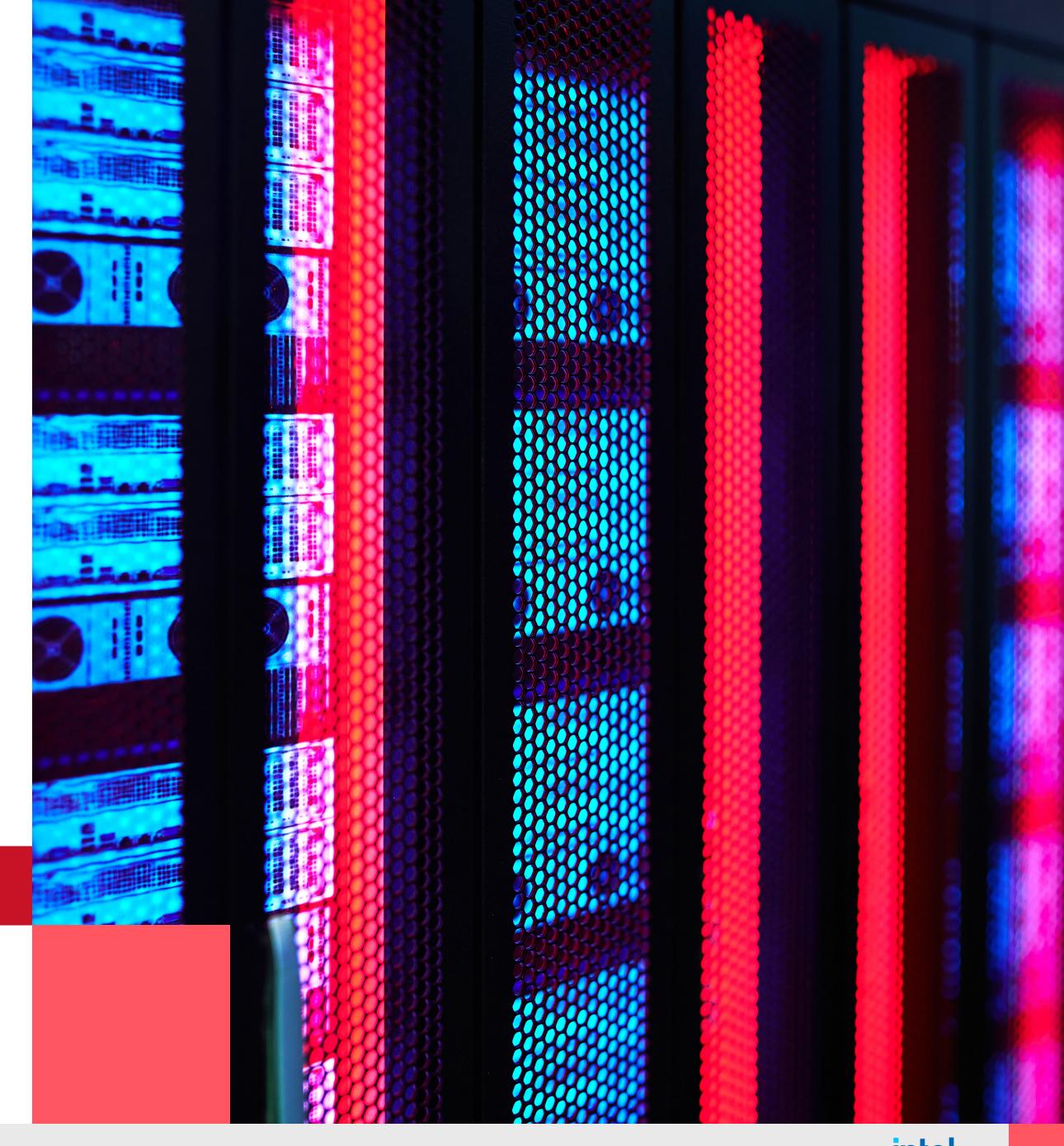
Building an effective feedback loop

- **Set clear metrics:** Track KPIs like workload performance, cost savings, and security compliance. Define specific thresholds for each metric to provide a clear benchmark for success and use tools like dashboards for real-time tracking.

- **Frequent reviews:** Conduct quarterly assessments to measure progress against your KPIs. These reviews should include cross-functional teams to evaluate all aspects of the modernization strategy and ensure it's aligned with business goals.

- **Iterative adjustments:** Use insights from your monitoring to fine-tune strategies. For example, if monitoring reveals a bottleneck in workload performance during peak usage, resources can be reallocated or scaled to optimize processing power, ensuring consistent performance and user satisfaction. Adjust project plans, timelines, or resource allocations based on performance data and evolving priorities.

Scaling successfully

- Start with **small, high-impact projects** to validate methodologies. Focus on workloads or applications that deliver quick wins, building momentum and stakeholder confidence for larger initiatives.

- Document **best practices** and create **reusable templates**. Capture detailed lessons learned and establish standard operating procedures that can be applied to future projects.

- Use **Intel's ecosystem of partners** to expand capabilities: Our broad ecosystem of technology partners offers expertise, tools, and resources to support you through every phase of your ongoing data center optimization journey.

Conclusion

Every day that passes with outdated infrastructure puts your organization at a disadvantage. Operational inefficiencies grow, costs increase, and the opportunity to leverage technologies like AI diminishes.

Now is the time to act—optimizing your data center is not just a choice; it's a strategic imperative. Intel® Xeon® processors and Intel® Gaudi® AI accelerators are purpose-built to meet the demands of today's data center architects. By taking the first step today, you position your business to thrive in tomorrow's digital-first economy.

For more information on Intel's latest technologies and how they can support your data center optimization, visit www.intel.com /content

Reasons to Choose Intel

Every journey needs the right traveling companions, so finding the right partners to support your optimization efforts for the long term is essential. Intel's solutions provide a robust foundation for ongoing innovation and scalability, supported by an ecosystem of experts to guide your journey. Consider Intel® Xeon® processors and Intel® Gaudi® AI accelerators as part of your data center optimization roadmap to:

1. Upgrade your infrastructure and save costs by consolidating up to 3:1 compared to 5-year-old systems[1]. Reduce server count and power consumption while meeting demanding workload requirements.

2. Achieve up to 69% lower total cost of ownership (TCO) on a range of workloads, including AI and high-performance computing[2].

3. Combine your software migration with new hardware to unlock the full value of Windows Server 2022 and SQL Server 2022.

4. For industries relying on sensitive and regulated data, enable Confidential AI on-prem or in the cloud, through Intel's innovative and comprehensive approach to confidential computing.

1. See 7T2 at intel.com/processorclaims: Intel® Xeon® 6. Your costs and results may vary. Test by Intel as of 04/16/24.
2. See 9T9, 9T8, 9T7 at intel.com/processorclaims: Intel Xeon 6. Results may vary. Test by Intel as of July 2024.