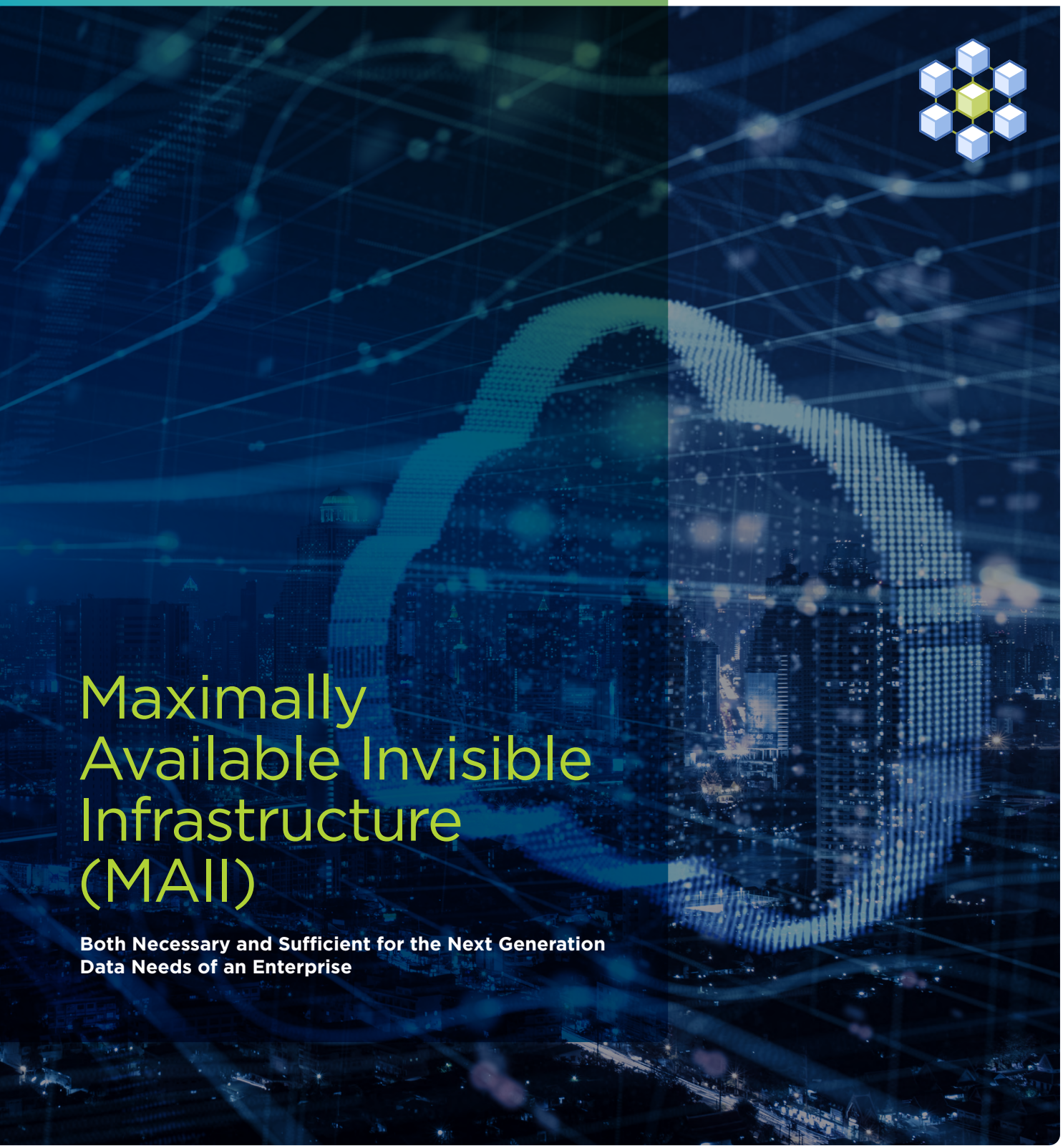
JULY 2020



Maximally Available Invisible Infrastructure (MAII)

Both Necessary and Sufficient for the Next Generation Data Needs of an Enterprise

NUTANIX
YOUR ENTERPRISE CLOUD

Abstract

Business needs are rapidly changing. Digitization, the Internet of Things (IoT), and social media create a data-driven culture where the amount and variety of information produced grows exponentially. New customer demands and the current prevalence of digital disruptors also contribute toward this trend.

Against this background, traditional data management infrastructures are coming under fire as rigid and unable to combine diverse data types. Enterprises today grapple with the problems of storing, managing and securing growing volumes of data as well as extracting insights (information) out of it. And data complexity rapidly increases with multi-cloud requirements such as moving data closer to both applications and users.

This paper is geared towards the CDO and explains both the problems involved in current data infrastructure and why a modern data infrastructure solves these issues with much greater flexibility, agility and reliability.

Introduction

In the journey to digitalization, enterprises aspire to develop data driven culture in order to lead or more effectively compete with digital disruptors. They collect both structured and unstructured data from various sources such as social media, IoT, and traditional systems of records. They deploy specialized skilled resources such as business analysts and data scientists to bring real time insights to bear in providing in-the-moment effectiveness of every stakeholder. They run analytical tools to gain information on customer service, customer engagement, market share, business optimization, financial performance, risk control, etc.

Effectively delivering a data driven culture requires exceptionally capable IT which in turn relies upon a robust, simplified, integrated and secure modern data infrastructure. The reality today is that executives from organizations of all shapes and sizes – both public and private, complain about a lack of access to accurate, timely, relevant, and reliable information. This impedes decision-making and either blocks the path from data capture to actionable insight, or leaves it littered with obstacles.

Most of the executives complaining run a legacy datacenter infrastructure – one that is based around proprietary SAN arrays and storage fabric networks. These "hardware-defined" datacenters are inefficient at scaling, are not natively resilient, and require specialized, expensive, and error-prone administration. This leaves them hopelessly unable to keep up with rapidly changing information demands. They make it extremely difficult to bring together diverse (structured and unstructured) data types. They suffer from time-consuming reporting, data availability issues, and most importantly – an inability to answer forward-looking, predictive questions.

IT staff dependent upon legacy datacenters are typically unable to keep up with the structural changes to data sources required to meet rapidly changing business relationships, customer demands, and organizational structures. They have difficulty adapting to entirely new data sets from newly implemented systems or as a result of mergers and acquisitions. The dilemma is that they need to address these changes while trying to meet the enterprise's evolving, and often challenging, business needs.

Companies that bring both an effective **architecture** and **strategy** to information management are finding that data becomes a core competitive advantage in driving digital transformation. At Nutanix, we believe a data services platform must support **Open Architecture** as well as be both **Maximally Available** to ensure continuous access and insights from the data, and effectively **Invisible** to operators and users allowing them to focus on higher order functions.

Open Architecture

Making data a competitive advantage requires supporting infrastructure that caters to all **data types**. Data in modern organizations tends to be organized primarily in three ways: Relational databases, unstructured data, and high-velocity data.

1) Relational databases: These include popular databases such as Oracle, SQL Server, DB2, MongoDB, MarkLogic, and PostgreSQL as well as key-value stores such as Cassandra and Couchbase.

2) Unstructured data: Unstructured data is typically stored in file servers and object storage repositories. It includes large files such as streaming media, healthcare imaging and indexers, along with billions of small files such as machine data, home directories and group shares.

3) High-Velocity data: High-velocity data represents a mix of structured and unstructured data and is typically stored in NoSQL databases. Examples include data lakes, data warehouses, ELK, Hadoop, and Kafka along with analytics workloads such as Tableau and Qlik.

This diversity of data and associated applications have resulted in the formation of many individual repositories or silos in which organizations store and organize their data. These silos share little to nothing with each other because of the lack of standardization of storage formats and data types. As a result, organizations have to deal with complex and poorly utilized storage infrastructure and poor utilization of their data.

An organization wanting to eliminate these silos and maximize the value derived from its data has several best practices to consider. These include adherence to open standards, adoption of open paradigms, offering choice and avoiding lock-in to users, customers, and operators alike. The ability to to bring in data in any shape or form and subsequently leave with that data is paramount.

In today's API-centric IT economy, adopting an API-first development methodology ensures software interoperability via generic endpoints and enables automation for rapid deployment and remediation of multiple production-ready environments. Using an API-first best practice also facilitates fast feedback loops from best in class alerting/monitoring stacks that consume our APIs.

Making Infrastructure Maximally Available

Traditional data repositories and operations result in a significant probability that data will be unavailable or even lost putting critical application services and entire organizations at risk. And the cumulative cost of maintaining acceptable uptime across the board can be significant. In order for data to become a competitive advantage, it first must be maximally available.

Making infrastructure maximally available mandates an ability to self-heal, to dynamically adjust the write path to suit the workload, and to dynamically apply data transforms based on heuristics. This requires control of initial placement and layout of applications across all possible hybrid cloud locations dependant upon:

- Data sovereignty
- Workload affinity
- Workload availability
- Workload redundancy

Maximally available data should be elegantly simple to manage with a management plane driven by the principles of consumer-grade design. This management plane should provide end-to-end capabilities around alerting, events and monitoring. Having access to performance indicators and self-build analytics and dashboards for all data types at once simplifies the day-to-day management processes.

MAXIMALLY AVAILABLE MEANS CONTINUOUS

In traditional database environments, an application or database update may create issues. As a preventative measure, many DBAs take a snapshot of their environment before any update to ensure they can roll back to the ideal state.

This manual step is unnecessary with a Maximally Available data architecture. The ability to upgrade the underlying infrastructure supporting your data should allow for continuous updates without interruption to services. This means that growing adding capacity for storage or performance, or performing a software update, should be seamless and non-disruptive to your data services.

Data protection mandates keeping the good state of the database environment. This includes secondary storage,storage for both legacy physical applications as well as virtual disks for virtual machines and persistent volumes for containers. Archiving, backups, and cold tiers (snapshopst, clones, and older versioned data) all should be readily accessible.

An ever growing concern with legacy data environments is the ability to keep up with, and maintain, database patches and updates. Database vulnerabilities with known patches are commonly highlighted in high profile cases where an organization's data is compromised or exposed, predominantly through lack of a process or appetite to maintain appropriate versioning control. Preventing these type of PR nightmares demands the ability to execute on two fronts:

- Easily identify what versions of databases are installed across the entire environment and apply patches (e.g. security patches), non-disruptively, in a timely fashion.

- Standardize the patching policy and process by reducing multiple different software configurations across dev, test, and production environments. This provides a unified management, producing less troubleshooting of bugs, and more uptime.

Making Data Infrastructure Invisible

A data infrastructure as a competitive advantage is effectively invisible. This requires strong architecture and design underpinnings geared toward optimally managing data. It must be both trusted and autonomous.

TRUSTED

An Invisible infrastructure for data delivers "Trust" by offering key capabilities around data and application security, data protection, and data governance and compliance:

Data and Application Security

Data platforms need to ensure that data stays secure, protected from compromise, corruption and is provided only on a need-to-access basis. This requires capabilities across the data, network and application plane. The data should be secured through at-rest encryption and enabled by effective key management solutions.

At a network level, the principle of zero-trust should provide a layer of security using micro-segmentation solutions that enable application micro-firewalling of east-west traffic. Finally, user or application level access to data needs to be authenticated and authorized using IAM solutions that support multi-factor authentication and fine grained role-based access controls to deliver a least privilege access.

Data Protection

In today's dynamic business environment with the volume, variety and velocity of data, it's critical that data platforms enable a seamless experience for ensuring data protection. The complexity of legacy infrastructure and software in use for backup and DR does not support adequate data protection.

Data protection entails safeguarding important information with an ability to ensure quick restoration in event of data corruption, compromise or loss. It includes systems and processes for, checksum, backup of data and business continuity/disaster recovery.

Nutanix provides a uniform set of tools enabling easily protecting all data assets to the appropriate level including options for in-place with space efficient Snapshots, replication (asynchronous, near synchronous, and synchronous), and cloning. It enables Hyper-Converged Backups or Disaggregated Backups with secondary storage as targets. These capabilities encompass all data types, providing a single set of tools to address your data protection needs, greatly reducing complexity.

This Data Governance and Compliance

While data security, data protection and data privacy capabilities need to be part of the invisible data infrastructure, in today's world there is an additional need to have a comprehensive lifecycle approach to data governance and compliance where there is real-time monitoring and detection of anomalous, potentially malicious activity, and hardening against common risks. Also required is a related capability to drive timely response and remediation to ensure that the data availability and integrity is not compromised at any point in time. To enable such an adaptive system, a host of native capabilities are required in the data platform.

Data governance also entails building safeguards to monitor data at ingress and egress hubs and based on the sensitivity of data, determine risk and remediation actions. Of course, regulation and standards compliance is imperative.

AUTONOMOUS

Organizations are looking for the same cloudlike simplicity & ease of use experienced by their infrastructure counterparts and apply that simplicity to their provisioning, management, and patching. Managing Databases tend to be very siloed with separate distinct groups running various database engines (Oracle, SQL Server Postgres, etc), database level groups, and compliance groups. There are a lot of steps involved from the point that a database request comes in and when a DB gets deployed.

The ability to automate a lot of the manual error-prone steps involved in provisioning or cloning of databases not only saves deployment time, but it also removes the potential for time consuming troubleshooting down the road from having multiple non-standard deployments. Time spent on the mundane day to day operational tasks of managing data is time not spent on projects supporting business initiatives or addressing technology growth.

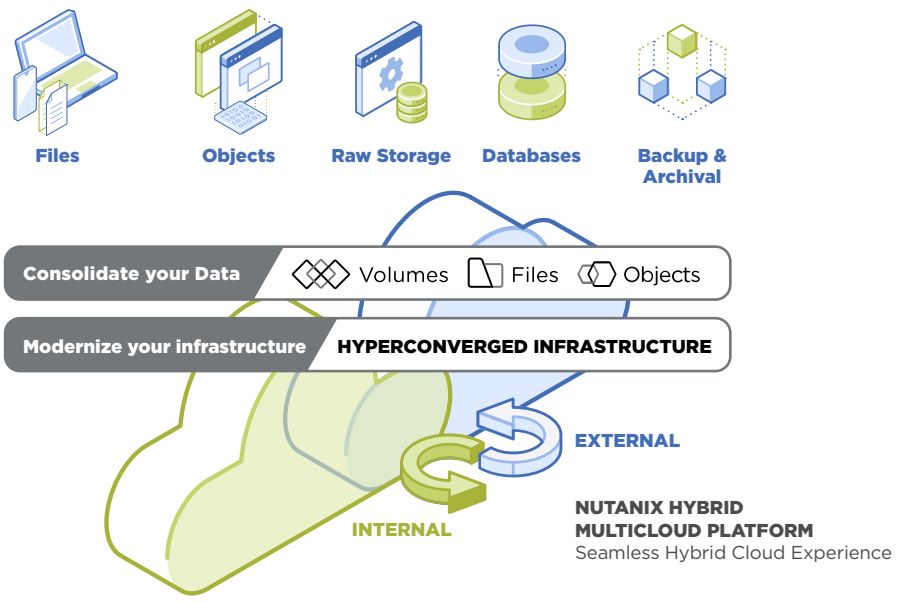
Integrating with infrastructure automation tools or self-service platforms gains additional efficiencies in the lifecycle management of data. This requires a data platform that easily integrates with existing automation tools by providing the ability to create a self-service catalog and delegation that controls who gets access, what they get access to (provision, clone, patch etc), how much (storage, RAM, vCPU), and for how long (spin down of resources).

Data growth and the accompanying requirements to manage ever more data can tax an IT organization. Most companies lack the required headcount to manage the resulting sprawl of their database estate. But even databases that are terabytes in size should be managed with the ease of use in the ability to copy-and-paste operations on your computer. For enterprise databases, for example, the requirement is not just to copy/paste a particular instance, but the ability to create clones from any point-in-time.

Data-Driven Digital Transformation

The software-defined Nutanix data services platform enables organizations to unify how their data is stored and consumed with a data services platform that is Open, Operative, and Opaque. This platform is supported both on-premises or in the public cloud. It delivers:

- An end-to-end environment for structured and unstructured data with raw storage.
- A cloud-like experience with full API-first automation and self-service—without sacrificing the security and control of an on-premises solution.
- The agility to embrace DevOps while maintaining a high level of service for traditional enterprise apps. Developers and QA teams leverage self-service to obtain resources without delay whenever they are needed.
- Fast provisioning of new file servers, object repositories, and databases with simplified recovery, empowering end-users including DBAs and reducing their dependence on infrastructure and ops teams.



Nutanix spans across HW platform, SPs and AWS -- serving files, objects, volumes, databases, and big data workloads.

- Nutanix Enterprise Cloud. Industry leading hyperconverged infrastructure platform delivering compute, storage, and networking services on a wide range hardware platforms and cloud infrastructure services. Integrated management functionality from Prism and Prism Pro eliminate the need for separate silos and deliver APIs for all management tasks.

- Nutanix Era. Integrated software that automates and simplifies database management, bringing one-click simplicity and invisible operations to database provisioning, lifecycle and copy data management. Era delivers a simple 1-Click path to delivering Databases as a Service (DBaaS).

- Nutanix Files. A software-defined scale-out file storage solution that supports both SMB and NFS, eliminating the need for NAS systems or file servers.

- Nutanix Volumes. Integrated SAN provides native scale-out block storage with direct block level access that replaces the cost and complexity of managing legacy SAN architectures.

- Nutanix Objects. A software-defined object storage solution with an S3- compatible REST API interface -- enabling a portfolio of cloud-native applications.

- Nutanix Mine. Integrated appliance to the integrated data protection to include secondary data backup and archiving all in a turnkey backup solution powered by our platform partners.

Data as Competitive Advantage

Digital transformation promises industry-shifting new products, services, and customer experiences, but all of these depend upon data-driven analytics and business intelligence. Database applications require an agile, scalable, and secure data infrastructure that frees organizations to stay focused on innovation, rather than remediation. Data provides a powerful, differentiating competitive advantage, but only if you can get to it quickly and reliably. For that, you'll need an effective data architecture and strategy.

We've painstakingly written down how we built this architecture, from design principles and product portfolio point of view. For those readers who care, here is our Nutanix Bible. And to learn more about how Nutanix can help transform your database operations, including DBaaS, visit Nutanix.com.

Nutanix makes infrastructure invisible, elevating IT to focus on the applications and services that power their business. The Nutanix Enterprise Cloud OS leverages web-scale engineering and consumer-grade design to natively converge compute, virtualization, and storage into a resilient, software-defined solution with rich machine intelligence. The result is predictable performance, cloud-like infrastructure consumption, robust security, and seamless application mobility for a broad range of enterprise applications. Learn more at www.nutanix.com or follow us on Twitter @nutanix.

NUTANIX™
Your Enterprise Cloud Platform

info@nutanix.com | www.nutanix.com | @nutanix