

Powering Retrieval Augmented Generation with

INFINIDAT

Transform private unstructured data into Intelligence with Security, Speed, Scalability



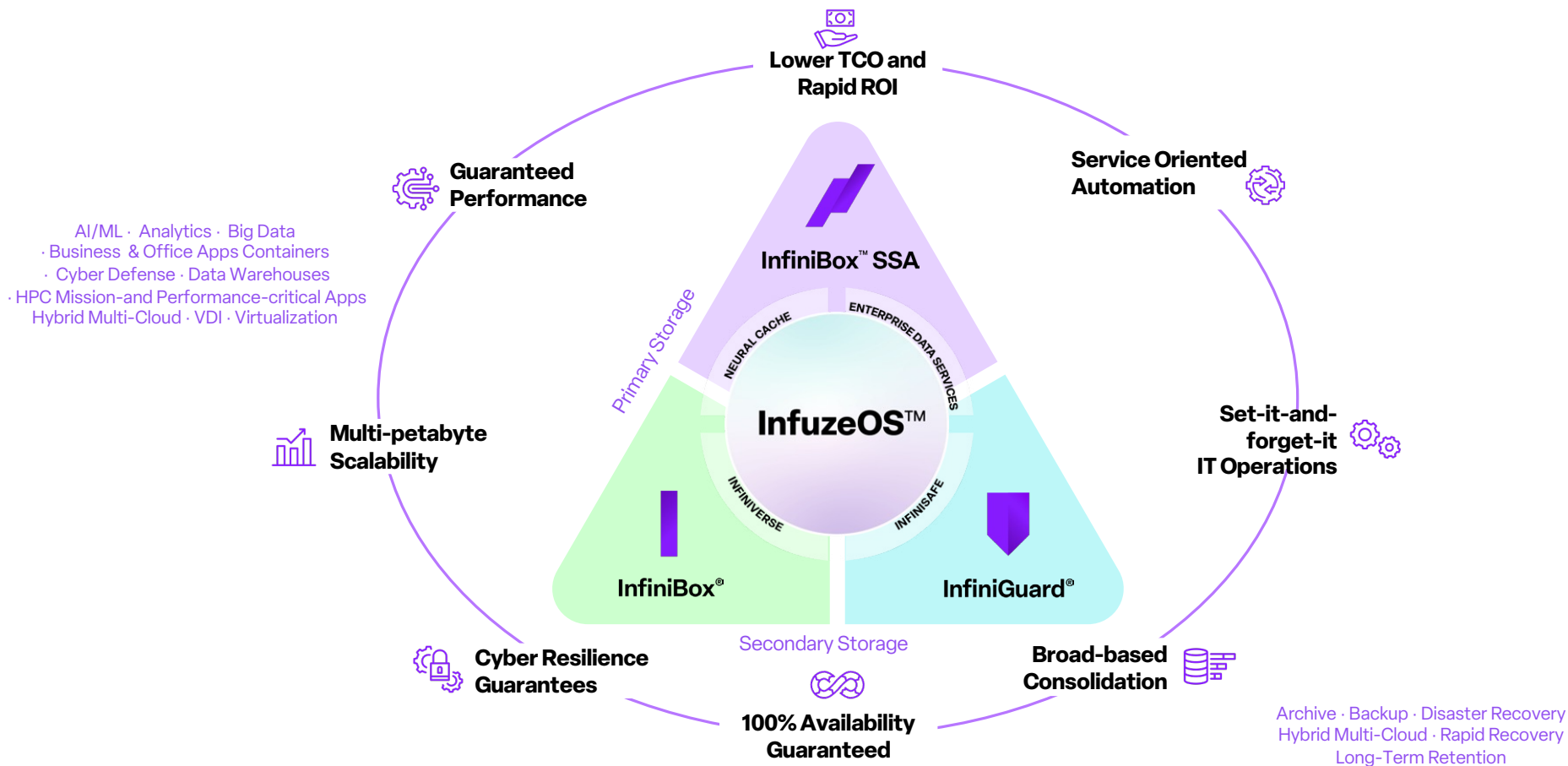
Eric Herzog – CMO
ehertzog@infinidat.com



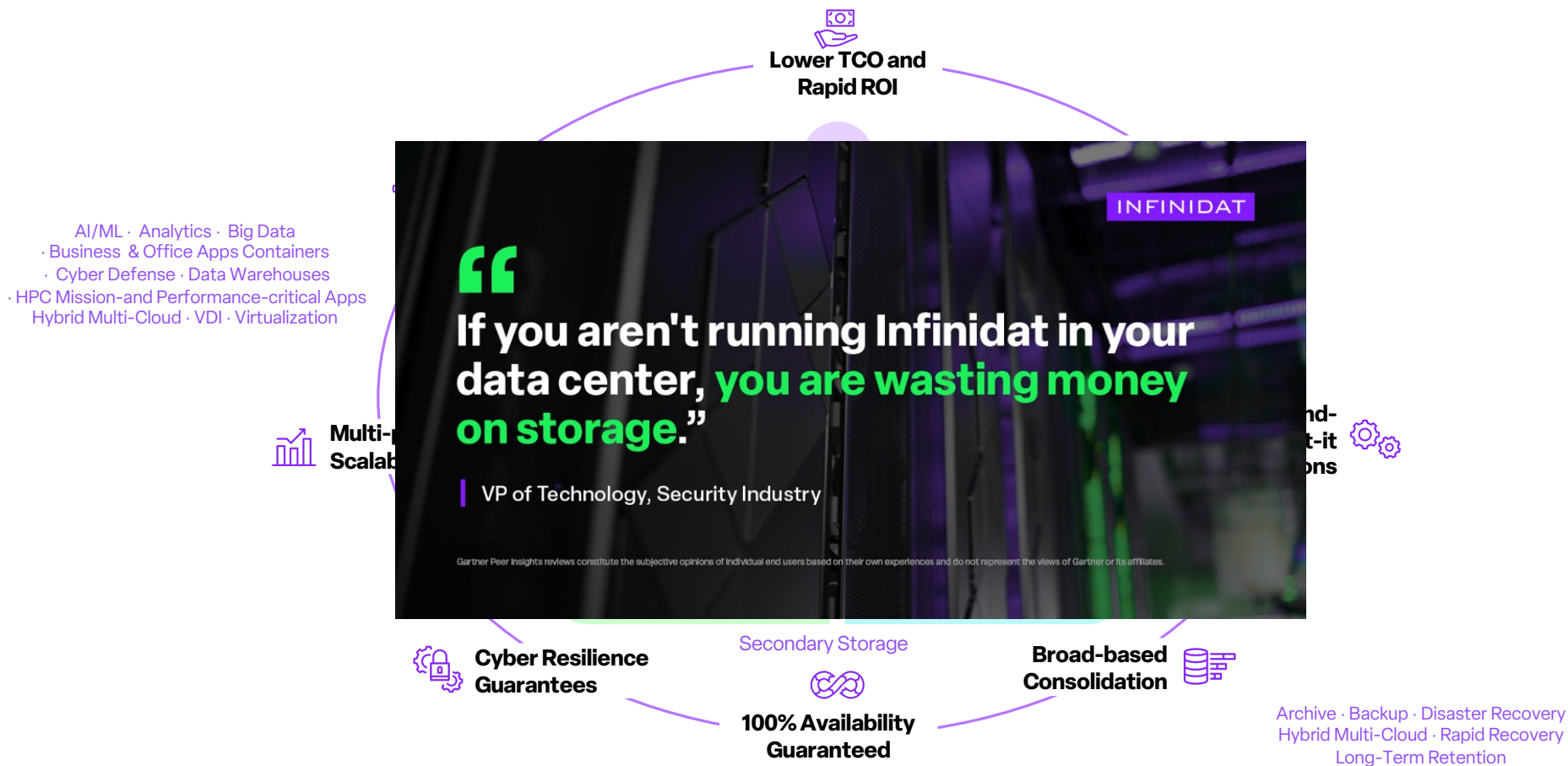
Bill Basinas – Sr. Dir. Product Marketing
bbasinas@infinidat.com

February 26, 2025

Review: InfuzeOS - One SDS Architecture, All Platforms



Review: InfuzeOS - One SDS Architecture, All Platforms



Industry Acclaim: 2024

GIGAOM



InfiniBox® and InfiniBox™ SSA II a Leader and Outperformer - 2024 Radar Report for Primary Storage For Large Enterprises



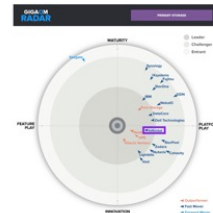
Storage Product of the Year: InfiniBox™ SSA II Gold Medal Award for “The best enterprise storage arrays of 2023”



InfiniBox™ SSA II Named One of Top 5 Cyber Secure High-End All-Flash Arrays for Enterprise Storage



Six Time Gartner® Peer Insights Customers' Choice Award Overall - 4 times North America - 2 times



InfiniBox® and InfiniBox™ SSA a Leader - GigaOm Radar Report for Primary Storage v.5



CRN® Storage Vendor 100 for the Fifth Year in a Row



Capacity Storage Vendor of the Year Winner - Storage Awards - Capacity Storage Vendor of the Year



5-Star Rating in CRN® Partner Program Guide for the 3rd Consecutive Year



Editor's Choice - Product Winner - Editor's Choice - Product - InfiniGuard® Runner-up - Storage Performance, Backup and Optimization



1 of the “Top 10 Leading Enterprise Storage Solution Providers of 2024”



Winner: Channel Storage Company of the Year and Cloud Security Innovation of the Year Runner up: Storage Hardware Innovation of the Year



InfiniBox® and InfiniGuard® Named One of Top 5 2PB+ Cyber Secure Backup Targets in the World



Infinidat Named a 2024 CRN® Tech Innovator Winner for InfuzeOS™ Cloud Edition, and a Finalist for InfiniSafe® Cyber Detection



Eric Herzog - CMO to Watch 2024 by the CMO Alliance



InfiniBox® G4 “Best of Show “for “Most Innovative Hyperscaler Implementation”



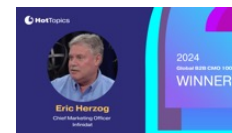
Eric Herzog – 2024 CRN® Inclusive Channel Leaders



Eric Herzog - Top 10 CMO/Marketer of the Year 2024



Infinidat – Finalist for Infrastructure and Storage Manufacturer of the Year



HotTopics - Eric Herzog – Global B2B CMO 100 Award 2024

GIGAOM



InfiniBox® and InfiniBox™ SSA II a Leader and Outperformer - 2024 Radar Report for Primary Storage Midsize Businesses



Steve Sullivan – 2024 CRN® Channel Chief



James ‘JT’ Lewis – 2024 CRN® Channel Chief - EMEA and APAC



Lynsy Marshall Top 100 Channel Marketing Leader - Channel Insider



Product of the Year Finalist - SSA Express



Infinidat named one of the World's Top 50 Marketing Teams



Lynsy Marshall CRN 2024 Women of the Year Finalist



One of 2024's 50 Coolest Software-Defined Storage Vendors

Industry Acclaim: 2024

CRN CLOUD 100 2025
Outperformer - 2024 Radar Report for Primary Storage For Large Enterprises

TechTarget STORAGE PRODUCTS OF THE YEAR 2023 GOLD

2024-25 DCIG TOP 5 GLOBAL EDITION

Gartner Peer Insights Customers' Choice 2024 North America
Six Time Gartner® Peer Insights Customers' Choice Award - 4 times in North America

CRN STORAGE 100 2024
CRN® Storage Vendor 100 for the Fifth Year in a Row

STORAGE AWARDS THE STORIES XXI WINNER
Capacity Storage Vendor of the Year
Winner - Storage Awards - Capacity Storage Vendor of the Year

THE CHANNEL CO. CRN TECH INNOVATORS WINNER 2024
Infinidat Named a 2024 CRN® Tech Innovator Winner for InfuzeOS™ Cloud Edition, and a Finalist for InfiniSafe® Cyber Detection

theCUBE TECHNOLOGY INNOVATION AWARDS 2025
Top Data Storage Innovation FINALIST

InfiniBox™ SSA II Named One of Top 5 Secure High-End Enterprise Storage

STORAGE PRODUCTS OF THE YEAR 2024 SILVER

Gartner Peer Insights Customers' Choice 2025
InfiniBox® and InfiniBox™ SSA a Leader - GigaOm Radar Report for Primary Storage v.5

CRN STORAGE 100 2024
InfiniBox® and InfiniGuard® Named One of Top 5 2PB+ Cyber Secure Backup Targets in 2024

theCUBE TECHNOLOGY INNOVATION AWARDS 2025
Top Data Protection Innovation FINALIST

CRN INCLUSIVE CHANNEL LEADERS 2024
Eric Herzog - CMO theCUBE
Eric Herzog - 2024 CRN® Inclusive Channel Leaders

Gartner Peer Insights Customers' Choice 2025
Infinitat - Finalist for Infrastructure and Storage Manufacturer of the Year

CRN PRODUCTS OF THE YEAR FINALIST 2024
Product of the Year Finalist - SSA Express

ONCON ICON AWARDS 2024
Infinidat named one of the World's Top 50 Marketing Teams

CRN WOMEN OF THE YEAR FINALIST 2024
Lynsy Marshall CRN 2024 Women of the Year Finalist

theCUBE TECHNOLOGY INNOVATION AWARDS 2025
One of 2024's 50 Coolest Software-Defined Storage Vendors

CRN CHANNEL CHIEFS REGIONAL 2024
James 'JT' Lewis - 2024 CRN® Channel Chief - EMEA and APAC

CRN CHANNEL CHIEFS 2024
Steve Sullivan - 2024 CRN® Channel Chief

100 TOP CHANNEL MARKETING LEADERS channelInsider 2024
Lynsy Marshall Top 100 Channel Marketing Leader - Channel Insider

FMS Best of Show *****
InfiniBox® G4 "Best of Show" for "Most Innovative Hyperscaler Implementation"

CMO to Watch IN 2024
Eric Herzog - CMO to Watch 2024 by the CMO Alliance

GIGAOM THE CHANNEL CO. CRN CHANNEL CHIEFS 2025
Eric Herzog

Gen AI has risk and reward!

Bad or inaccurate responses can lead to:

- Legal and Compliance Issues
- Brand Reputation/Perception Issue
- Customer Satisfaction
- Loss of business to Competitors

Accurate data produces in Gen AI

- Better Customer Experience
- Enhanced Productivity/Efficiency
- Higher Customer Satisfaction
- Innovate and Create Faster
- Cost Reductions and Higher Profits

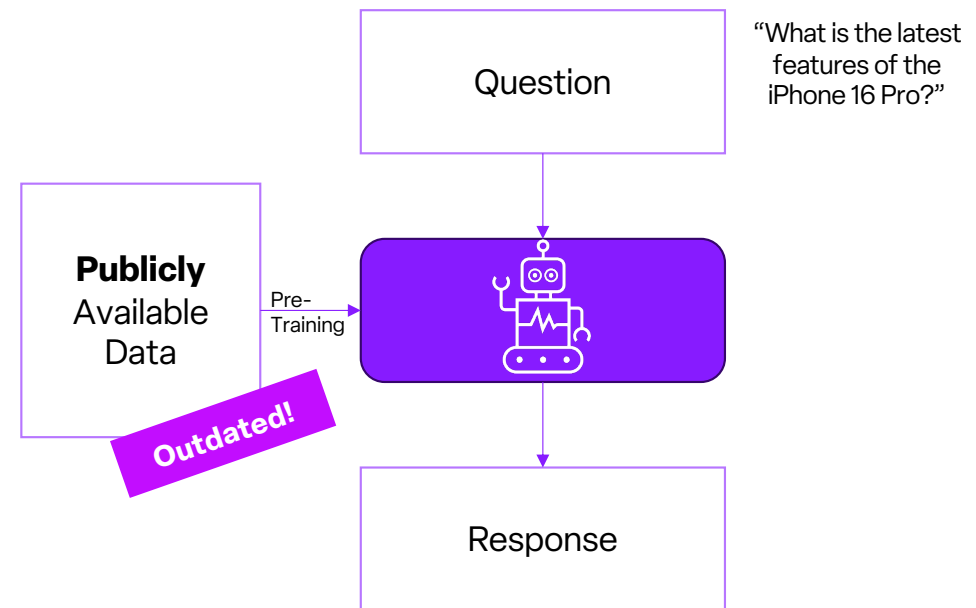
<https://markets.businessinsider.com/news/stocks/amazon-amzn-fails-to-provide-answers-in-ongoing-alexa-ai-delay-1034374823>



“Amazon’s (AMZN) AI upgrade of its Alexa device is expected to be delayed because of one rather inconvenient problem – it gives wrong answers to questions.”

Gen AI – Getting the right answer is a challenge!

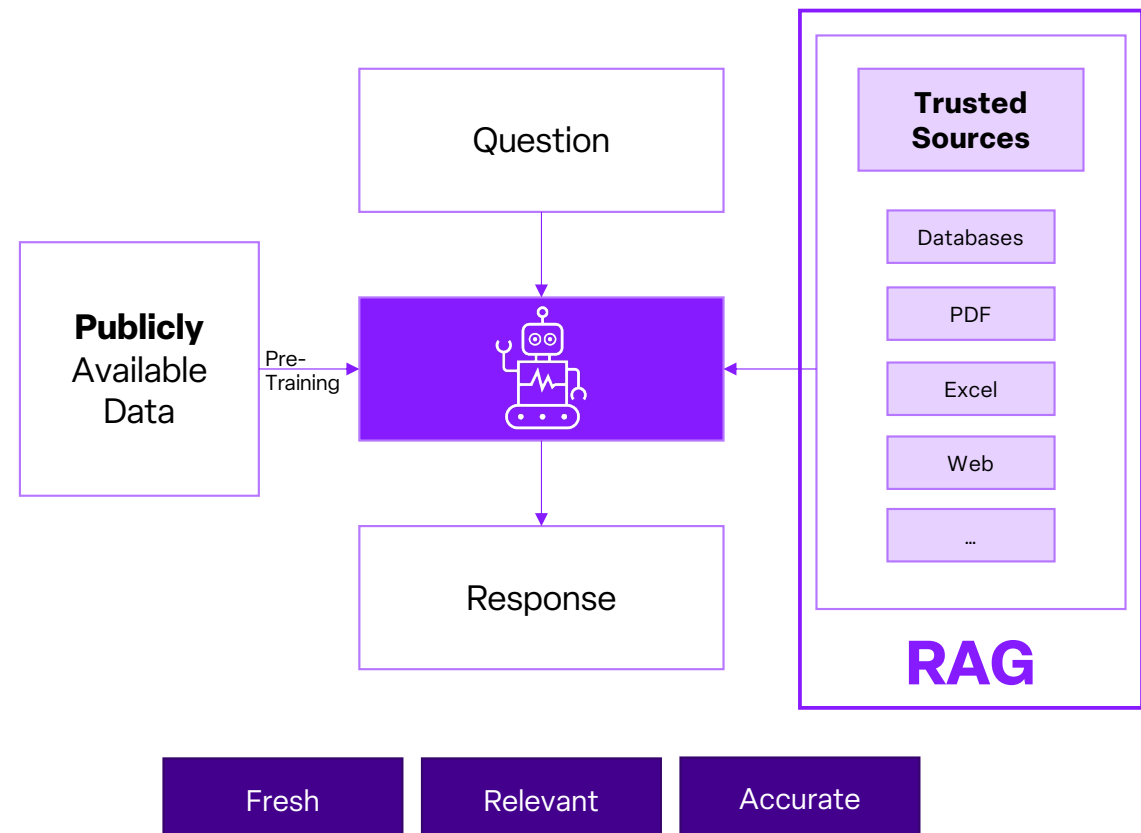
- LLMs and SLMs are typically trained on publicly available data, lack exposure to specific information and an enterprise's proprietary data
- Result - inaccurate and unreliable answers called "hallucinations"
- Appear to be plausible answers but are incorrect or outdated
- This leads to misinformed business decisions, providing customers with wrong information which can damage the enterprise's reputation
- Gen AI needs continuous augmentation with new, trusted, and accurate data to enhance and maintain its reliability



"As of my last update in January 2022, I don't have access to real-time information or updates beyond that date."

R(etrieval)A(ugmented)G(eneration) to the Rescue!

- Retrieval-Augmented Generation (RAG) combines the power of GenAI with an enterprise's private data to unlock fresh insights and streamline workflows
- Integrating information from external sources into LLMs and SLMs can achieve more accurate and specific results
- Agentic



The Infinidat Gen AI Solution

On-Premise



InfuzeOS on-prem



Replication



InfuzeOS
Cloud Edition

- Adding RAG to benefit Gen AI should work with existing infrastructure in the majority of cases
- Leveraging your existing storage eco-system is key
- Automate and orchestrate the deployment from data center to hybrid multi-cloud
- Easily integrate other storage environments as well (NFS/File)
- Agentic RAG adds new dimensions of flexibility and scalability





Retrieval Augmented Generation with InfuzeOS

Reference Architecture Demo

Wei Wang

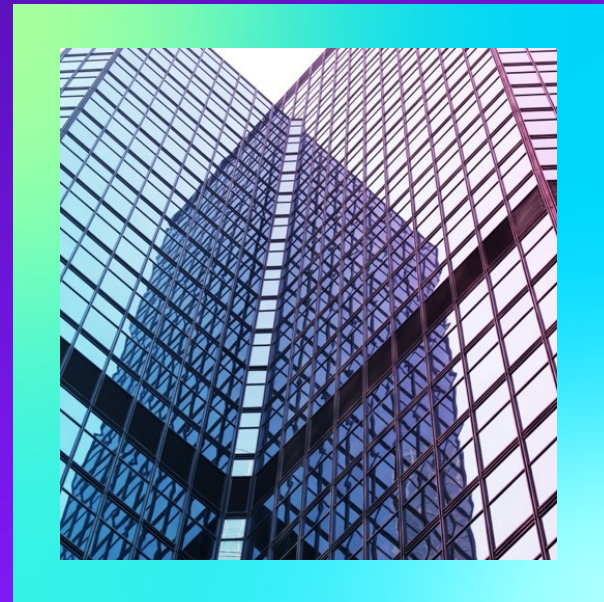
wwang@infinidat.com

Technology Integration Specialist

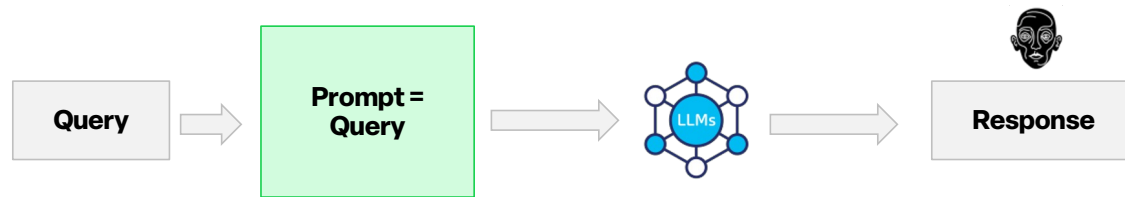
INFINIDAT

Demo Agenda

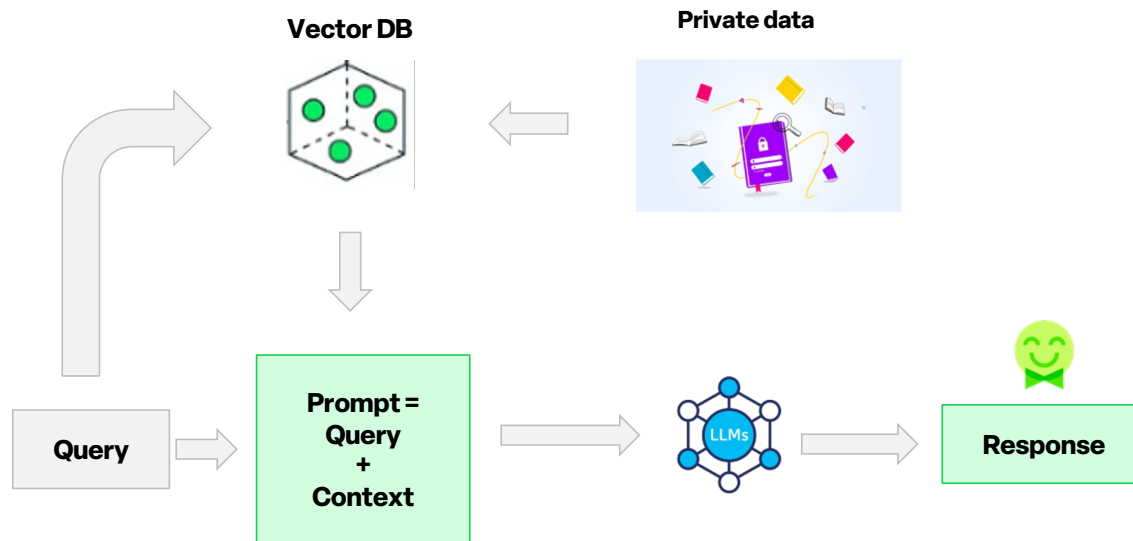
- Architecture walkthrough
- Use Case 1: Chat with your private data
 - Feed in private data
 - Key components with K8s dashboard
 - Vector database (neo4j) - Graph Knowledge
 - Chat with the dataset
- Use case 2: Agentic RAG with n8n
 - Mature open-source tools like n8n can be used to build powerful and secure AI solutions
 - Our framework is flexible and expandable



The Best Tool To Gain Intelligence From Private Data



- Outdated data
- Hallucinations

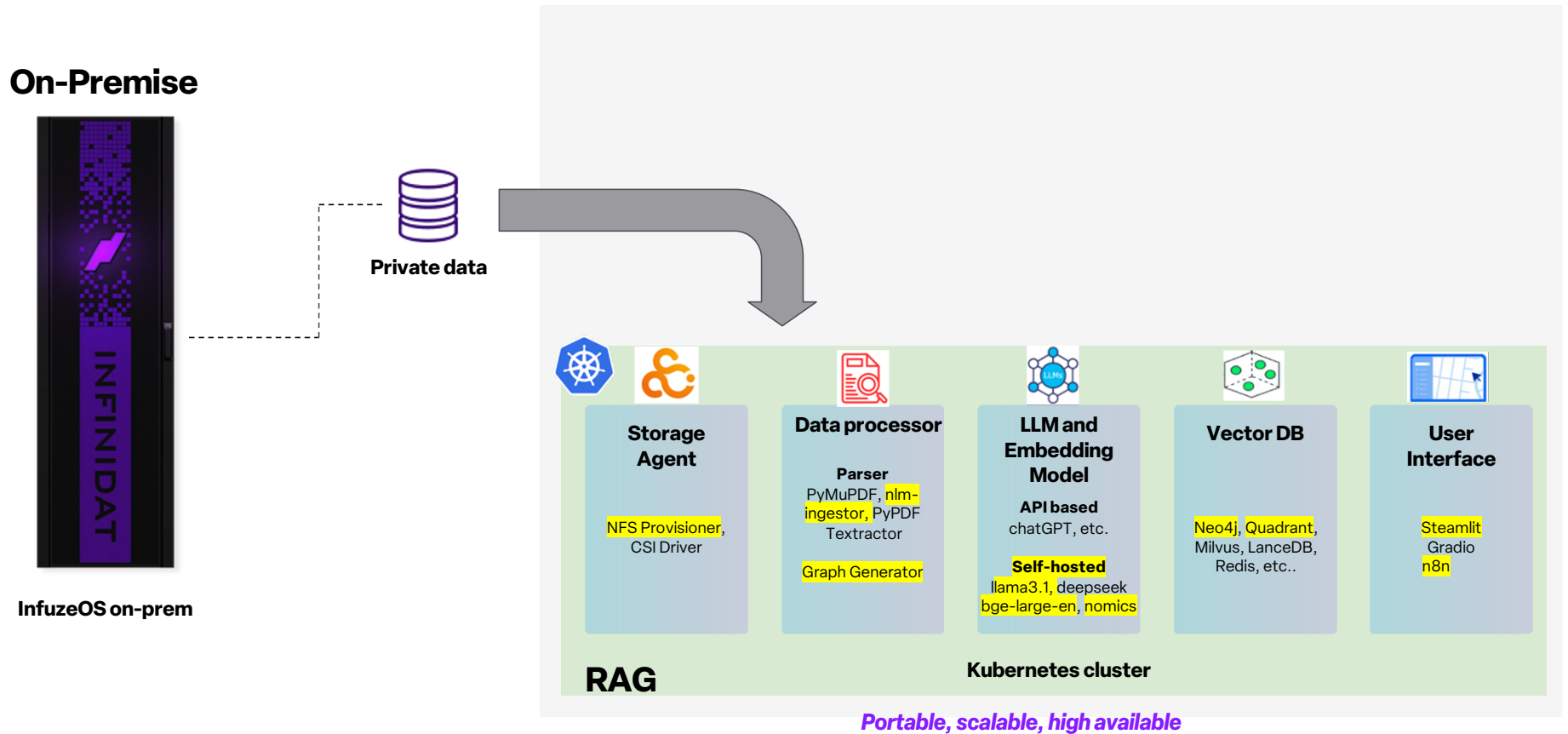


RAG

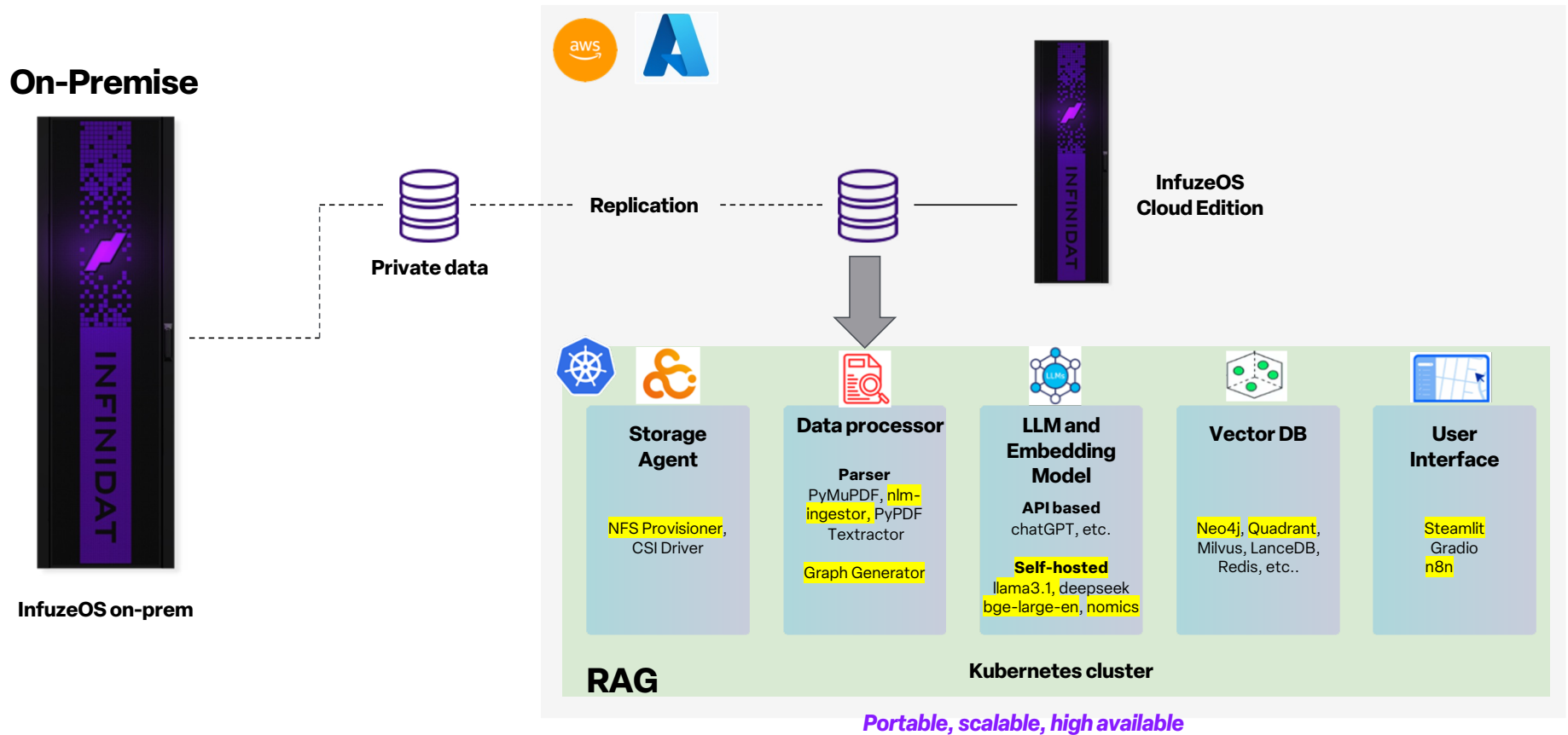
Major challenges:

- Security
- Scalability
- Operation complexity

A Framework for Secure and Limitless Intelligence (On-Prem)



A Framework for Secure and Limitless Intelligence (Cloud)



Automated Deployment with Terraform

info.infinidat.com/AI_Use_Case_RAG_with_Infinibox-EN_VideoLP.html



INFINIDAT

Video

AI Use Case: Retrieval-Augmented Generation (RAG) with InfiniBox

Deployment starts
at timestamp 2:22

Swiftly integrate your data with the latest GenAI technologies

Demo

- On-prem InfiniBox
- Dataset
- Replication to InfluxOS Cloud Edition
- **Deploy RAG to Cloud**
 - Dataset ingestion
 - Chat with the dataset

```
~/infinidat_RAG_Ref_Arch git:(Demo) # make demo
~/infinidat_RAG_Ref_Arch #
~/infinidat_RAG_Ref_Arch #
[18:58:25] INFO: InfluxRAG auto deployment
[18:58:25] INFO: Deploying an AWS EC2 and create a single node K3s Cluster, this will take about
7 min
make[1]: Entering directory '/home/wmang/workspace/infinidat_RAG_Ref_Arch/terraform'
terraform init
Initializing the backend...
```

InfiniBox Advantage for RAG


- Many of our current customers use their InfiniBox in mission-critical applications that utilize Databases which supports storing and retrieving vector data – RAG-ready!
- InfiniBox achieves lightning-fast performance with very low latency (Neural Cache), will not bog-down the RAG workload process during the Embedding stage.
- InfiniBox is a fast retrieval data storage system the Large Language Model (LLM) can use to generate coherent and contextually relevant responses.
- InfuzeOS Cloud Edition provides cloud-based storage operation similar to on-prem enabling storage admins to have the same powerful InfuzeOS in the cloud, to rapidly replicate datasets to AI resources.
- Flexibility – Leverage Agentic RAG architectures



The End Users Have Spoken



InfiniBox Reviews
by Infinidat in Primary Storage Platforms
4.9 ★★★★★ 491 Ratings



InfiniBox SSA Reviews
by Infinidat in Primary Storage Platforms
4.8 ★★★★★ 25 Ratings



InfiniGuard Reviews
by Infinidat in Data Center - Others
5.0 ★★★★★ 21 Ratings



As of Feb 25, 2025

The End Users Have Spoken

The image shows three screenshots of Gartner Peer Insights reviews for InfiniBox. Each screenshot includes the InfiniBox logo, the product name, the reviewer's name, the rating (stars), and the number of ratings. The reviews are for InfiniBox SSA, InfiniBox SSA, and InfiniBox SSA.

Review Title	Rating	Number of Ratings
InfiniBox SSA	4.9	21 Ratings
InfiniBox SSA	4.8	21 Ratings
InfiniBox SSA	5.0	21 Ratings

INFINIDAT

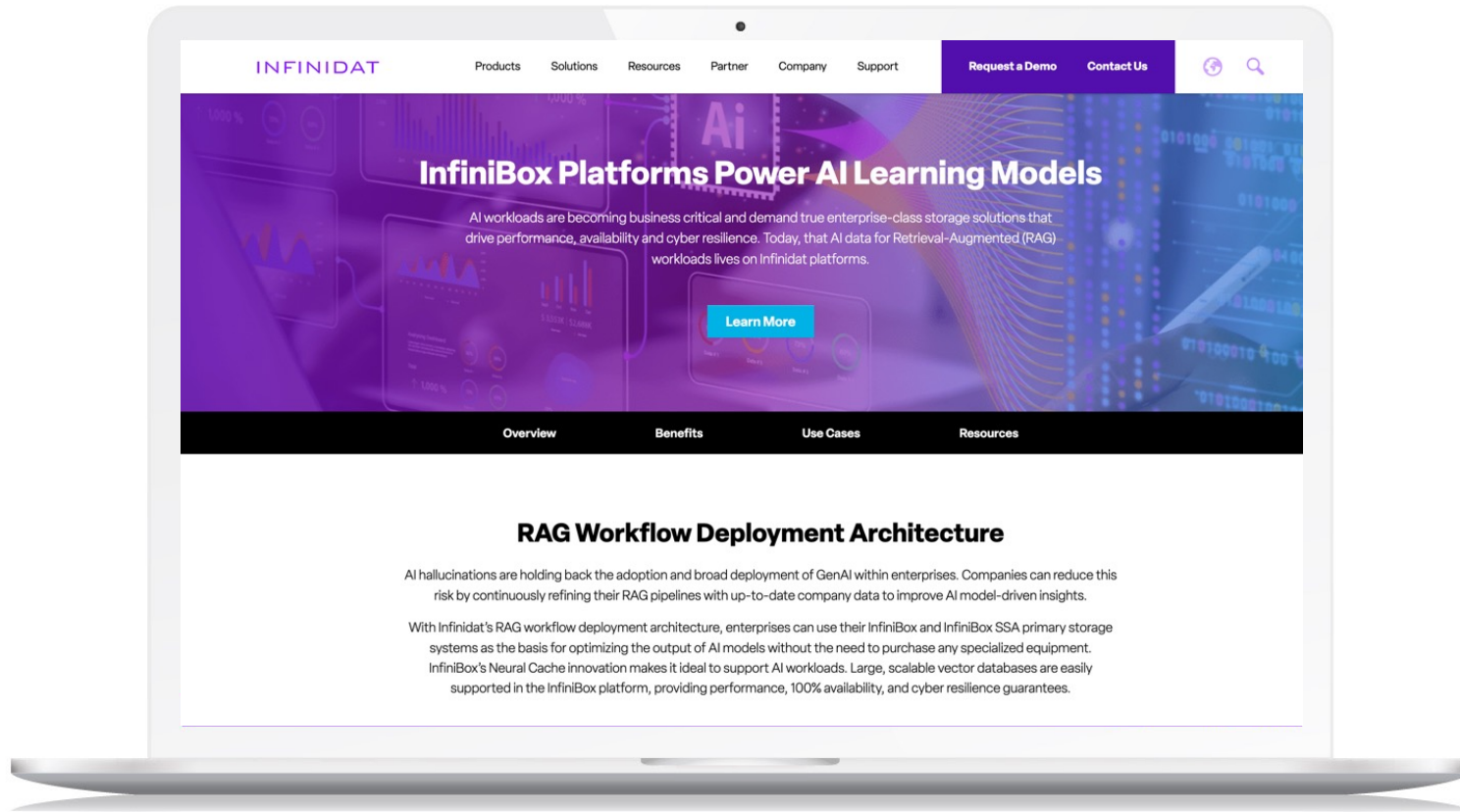
“**InfiniBox SSA is an outstanding storage array.**”

Account Executive, IT Services | InfiniBox™ SSA



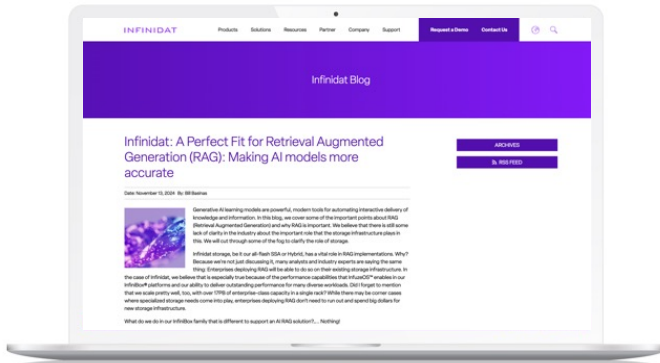
As of Feb 25, 2025

AI Solutions Web Page



<https://www.infinidat.com/en/artificial-intelligence>

AI Solutions Resources



Blog:

Infinidat: A Perfect Fit for Retrieval Augmented Generation (RAG): Making AI models more accurate

<https://www.infinidat.com/en/blog/infinidat-perfect-fit-retrieval-augmented-generation-rag-making-ai-models-more-accurate>

Analyst Whitepaper:

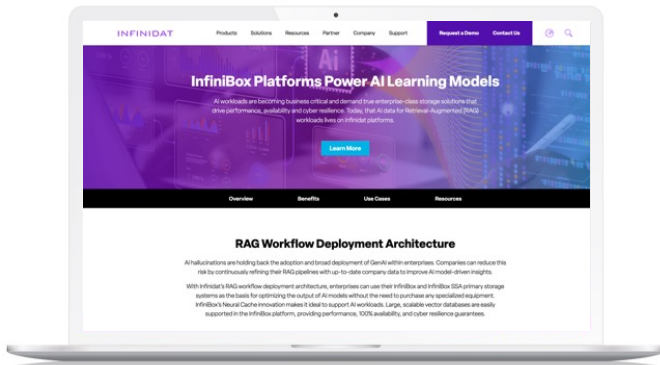
<https://www.infinidat.com/en/resource-pdfs/infinidat-cures-rag-response-time.pdf>

Analyst Video:

<https://blocksandfiles.com/2025/02/13/how-to-reduce-ai-hallucinations/>

Solution Webpage:

<https://www.infinidat.com/en/artificial-intelligence>



 **Thank You**