



— CLOUD + AI · VIETNAM

SOLUTIONS AND PRODUCTS

Solid as onyx. Sharp as technology.





Outline

1. CLOUD INFRASTRUCTURE
2. AI SOLUTIONS
3. PROFESSIONAL SERVICES
4. INTERNAL PRODUCTS

A platform that replaces vSphere component for component

Independent codebase, an operating model VMware admins already know, and certification against 260+ software and hardware products.



5,000+

enterprise customers

30+

countries

1,000+

VMware replacements

60,000+

hours stable

Source: platform vendor materials, 2025

What replaces what

Compute virtualisation	vSphere (ESXi + vCenter)	Hypervisor with a central console
Network virtualisation	NSX	Built-in VPC, VxLAN, LB, firewall
Storage	vSAN	Software-defined storage
Operations & automation	Aria Suite	Built-in ops, billing, tenants
Containers	Tanzu Kubernetes	Built-in Kubernetes platform
Backup & disaster recovery	SRM, native backup	Replication and DR orchestration

One operating model across public, hybrid and on-premises

The same workloads, wherever they need to live — one team, one control plane, one support line.

PUBLIC CLOUD

Public cloud landing zone

- Landing zone, network, IAM, cost guardrails
- Managed migration from your provider
- Managed data and AI services on tap
- Pay-as-you-go through campaign peaks

Best for fast launch and global reach

HYBRID

Public and private under one control plane

- One console over public clouds and private clusters
- Shared identity, quota, billing, monitoring
- Placement by cost, latency, data residency
- One migration path, both directions

Best for bursting to public with regulated data at home

ON-PREMISES

Private cloud in your own data centre

- Hypervisor, private cloud or hyperconverged on your hardware
- SDS storage, VPC network, DR to a second site
- Perpetual licence, no forced renewal
- No data leaves your building

Best for data sovereignty and predictable cost

Licensing that stays fair, migration without agents

The two questions that decide the business case: what it costs over five years, and what breaks while we move.

Three differences vs Broadcom-VMware

Perpetual, not subscription

An owned licence is never force-converted.

Per CPU socket, not per core

Core counts climb every generation. Sockets do not.

À la carte, not bundled

Add each product only when you need it.

50%

TCO reduction

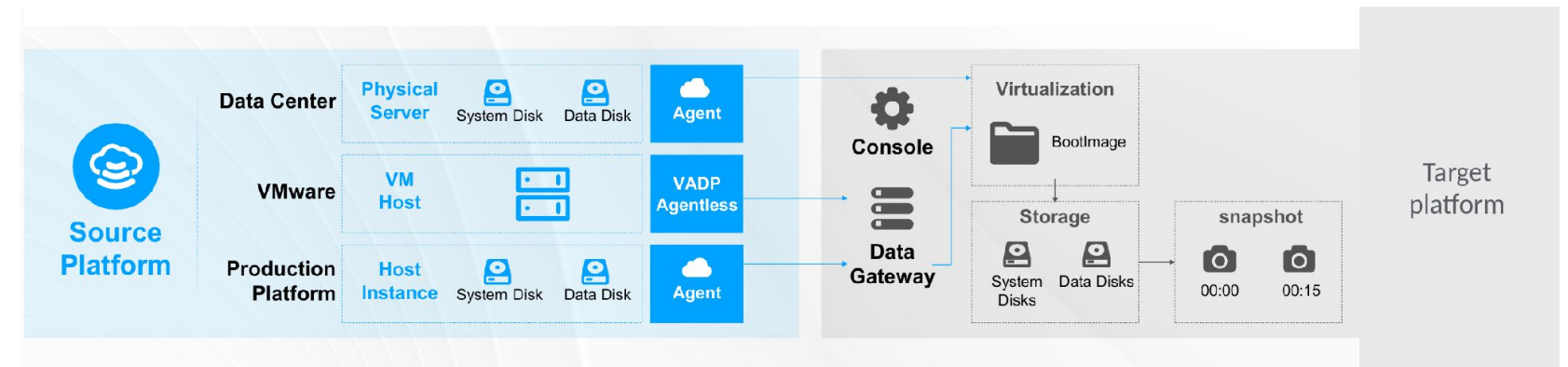
100%

scenario coverage

0

agents in the guest

Agentless from VMware, in five steps

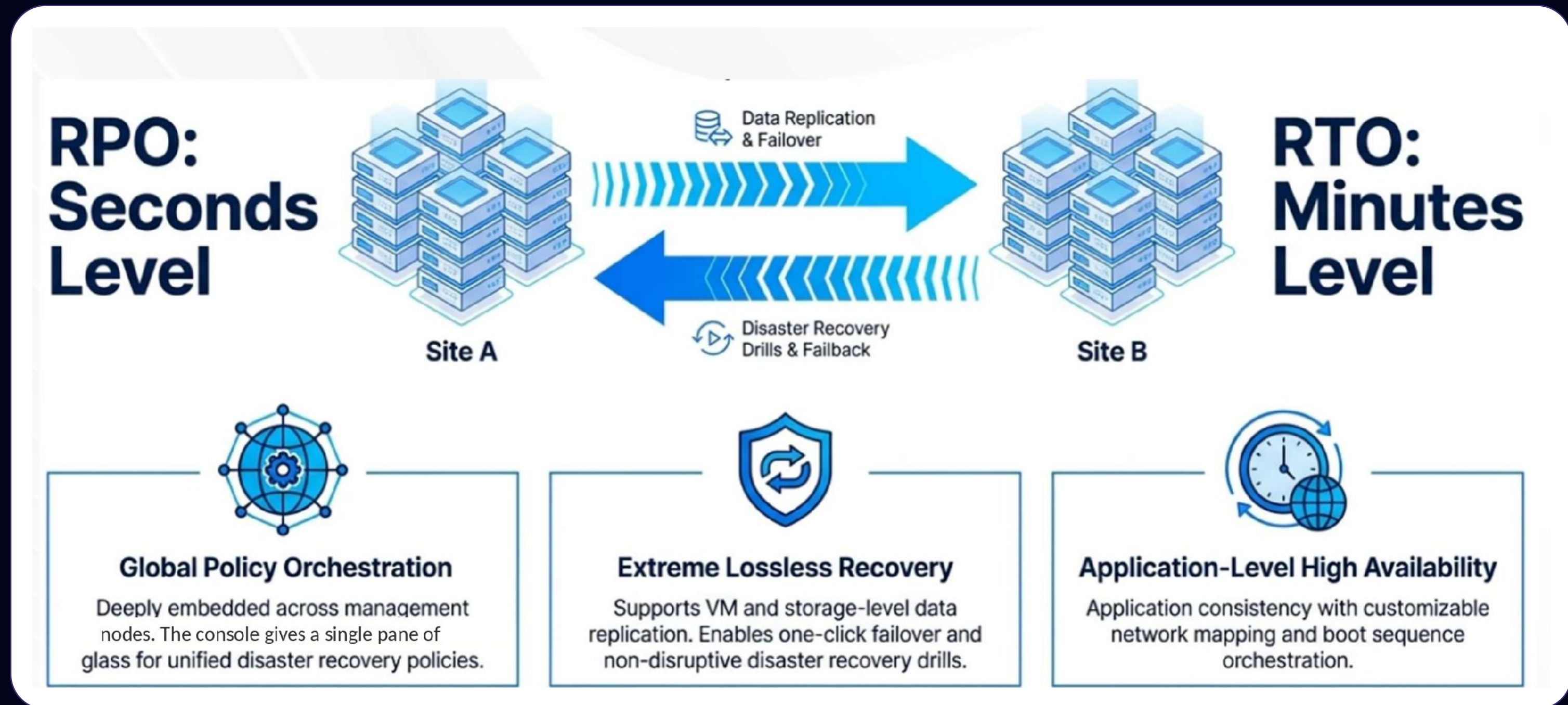


- 01 Discovery
- 02 Business assessment
- 03 Planning
- 04 Implementation
- 05 Post-deployment support

VMware sources move through VADP — nothing is installed inside the guest.

Disaster recovery that replaces VMware SRM

Replication between two sites with a recovery point measured in seconds and a recovery time measured in minutes. Licensed as an add-on module.



Bring AI into the systems you already run

No rip-and-replace. AI attaches to the applications and data you have today — grounded in your own content, and measured on a business metric.



01

Ground it

Answers with citations back to your own documents.

8OnyX AI Search

02

Personalise it

Behaviour trained into a ranking model on a managed AI platform.

8OnyX AI Recommendation

03

Automate it

Agents and workflows running on the tools your team uses.

Agents, workflows, MCP

04

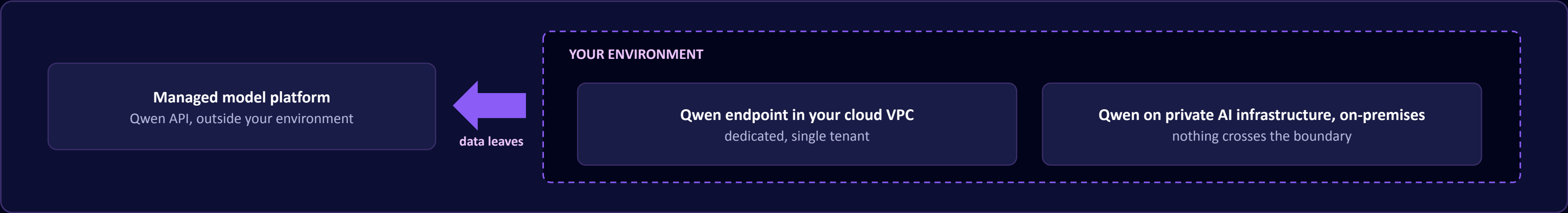
Govern it

Routing, quota, guardrails and tracing before anything scales.

AI gateway

Three places the AI can run — your data decides

The same application logic ships to all three. What changes is where the model sits and who can see the data. 8OnyX officially distributes the Qwen family.



PUBLIC API

Fastest to value

- Qwen called over a managed model platform
- No GPU investment, no infrastructure to run
- Usage-based cost, scales to zero

Fit: pilots, low-sensitivity content

PRIVATE MODEL

Isolated in your own tenancy

- Qwen open-weight, fine-tuned on your data
- No shared inference traffic with other tenants
- Reserved capacity, predictable monthly cost

Fit: customer data, internal knowledge

ON-PREMISES

Private AI platform in your data centre

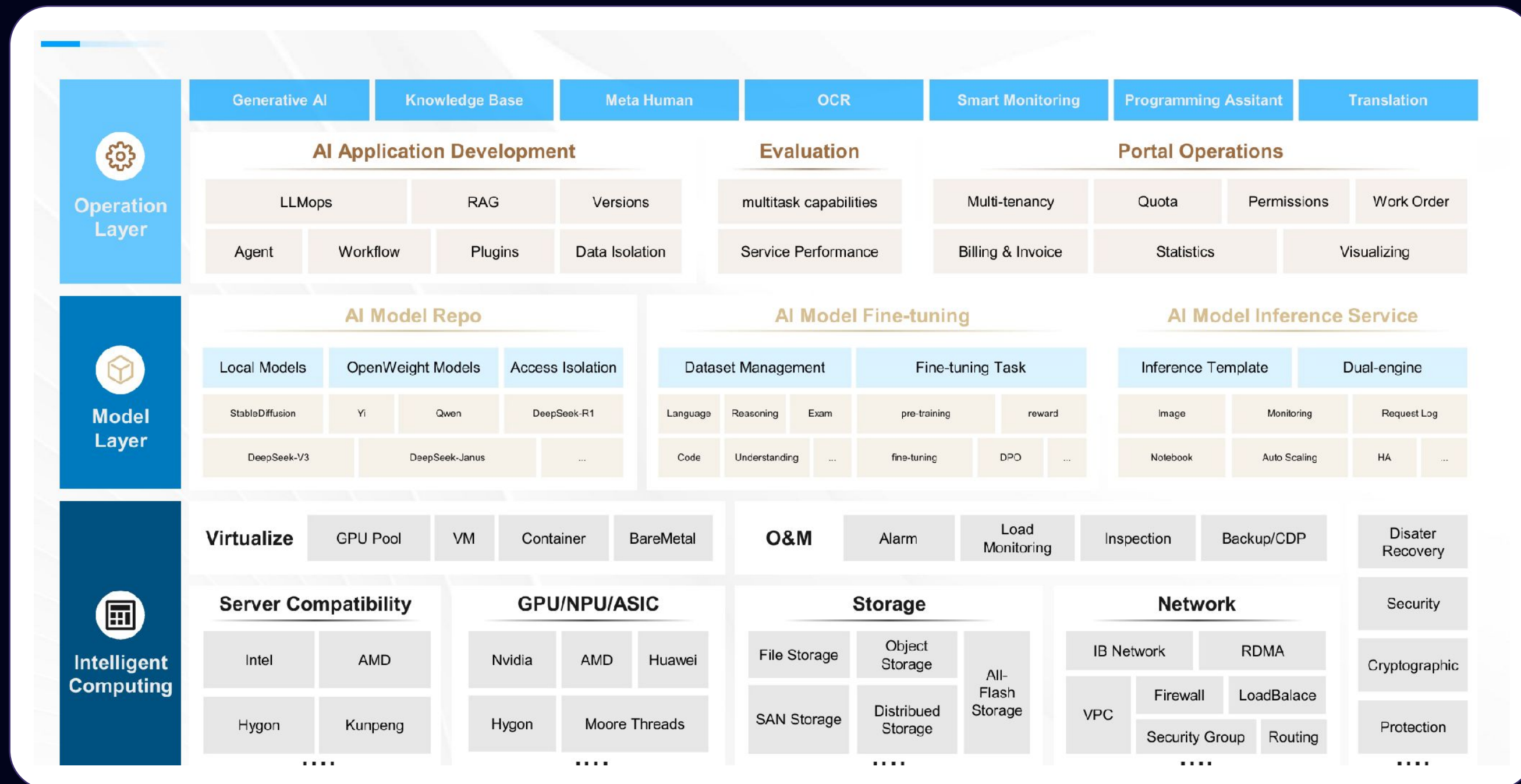
- GPU pooling, vGPU slicing, flexible scheduling
- Model repository, fine-tuning and inference
- Multi-tenant metering and billing built in

Fit: data sovereignty, steady heavy load

How an engagement usually runs Use-case selection → proof of concept on real data → evaluation against an agreed metric → rollout and handover

Private AI infrastructure — your GPUs, your data

Three layers: applications and portal on top, model repository and inference in the middle, pooled GPU, storage and network underneath.



The model repository can host many open-weight families; 8OnyX officially distributes the Qwen family.

AI Search

Typical Scenarios — Search Recommendation

03 Search Recommendation



Fine-grained Document Governance

Parse, clean and chunk raw docs (HTML/PDF/DOC) into unified format.



Pre-Retrieval & Retrieval

Query rewrite, term matching, embedding & vector search across the knowledge store.



Answer Generation

LLM answer generation with agent/workflow services on top of retrieved chunks.



Auto Evaluation

Retrieval and answer-correctness metrics validate quality continuously.

How AI Search works — 4 simple steps

01

Ingest

Parse & chunk your documents (PDF, Word, HTML) into a clean knowledge base



02

Retrieve

Rewrite the query, then semantic + vector search finds the most relevant chunks



03

Generate

The LLM writes a grounded answer, with product & MiniApp recommendations



04

Evaluate

Automatic quality checks keep answers accurate and relevant over time

Personalized recommendation engine

Turn user behaviour into higher click-through and conversion, on a managed AI platform.

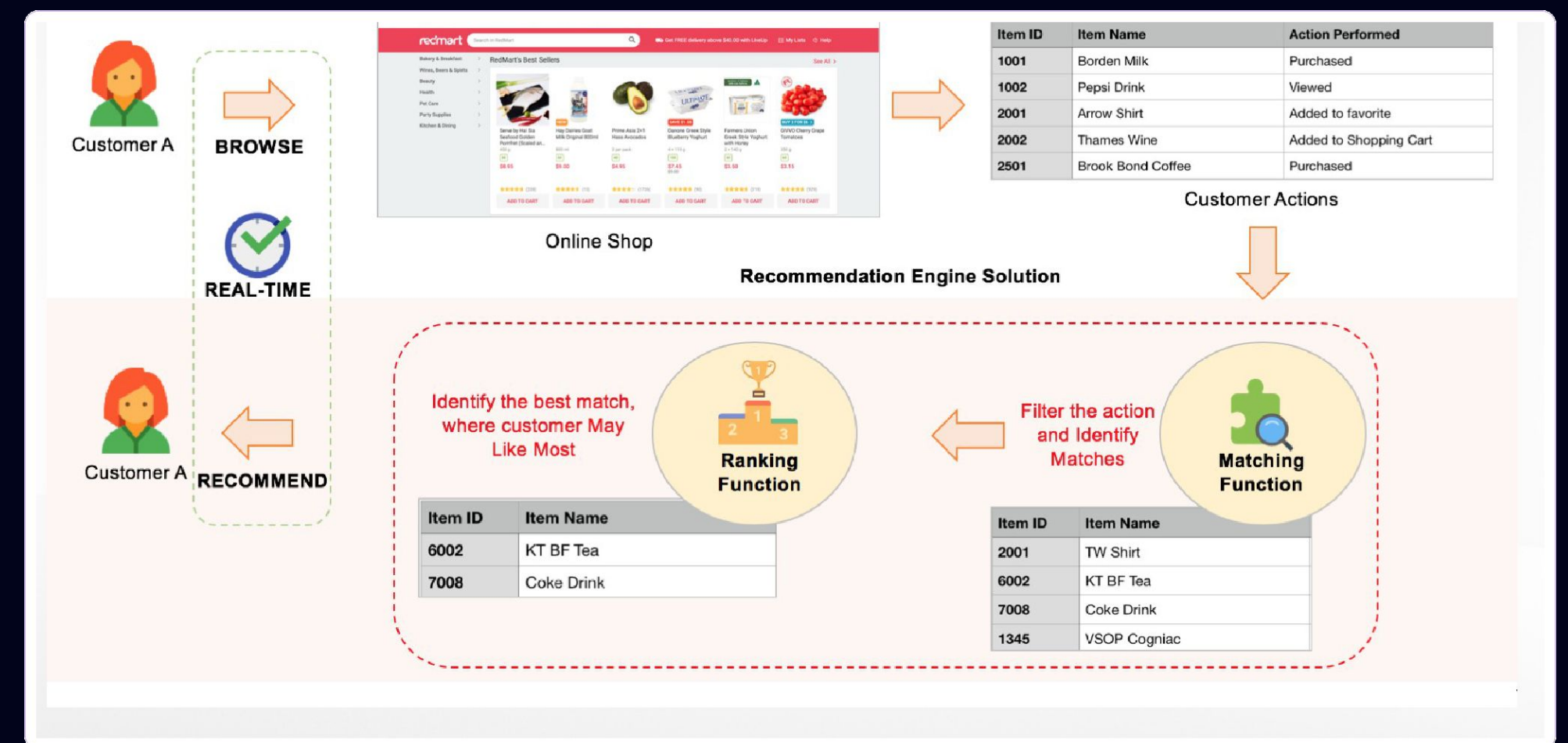
The challenge

- ✗ Lower click-through rate (CTR)
- ✗ Hard to identify customer preferences
- ✗ Struggling to maximize purchases

What the engine needs

- User — gender, age, education, income, area
- Item — size, price, category, area
- Action — buy, click, collect, add-to-cart

Real-time recommendation engine



Matching Function filters the catalog to candidates · **Ranking Function** orders by interest strength

— AI RECOMMENDATION

Proven at scale

Recommendation engines running in production across e-commerce and media.

Dolphin Home



Homepage

Shopping Cart

- **Cross-border beauty e-commerce**
Beauty, mother & baby, fruit, home
- **20M+ users · 380K+ SPU**
UV-CTR is the core success metric
- **Peak-ready**
Double-11 traffic 100× daily

Pumpkin Movie



You may also like

Related recommendations

- **Vertical membership film & TV**
Smart recommendation on app & TV
- **25M+ end users**
PV-CTR & UV-CTR are key indicators
- **Multi-dimension labels**
Film tags drive personalized recall

Chongqing Morning News



- **News & content platform**
Personalized content feed
- **Higher engagement**
Lifts click-through & retention
- **Real-time**
Fresh behavior-driven ranking

Four ways to put our team on your problem

Advisory when you need a decision, delivery when you need it built, engineers when you need capacity, and operations when you need it to keep running.

01

Advisory & assessment

- Architecture and infrastructure review
- Five-year TCO and licensing analysis
- Migration feasibility and risk
- AI opportunity mapping and scoring

Fixed fee · typically 2–4 weeks

02

Project delivery

- Design, build, migrate, hand over
- Landing zones and private cloud build-outs
- AI wired into existing applications
- Documentation and training at handover

Fixed scope · milestone-based

03

Dedicated engineering team

- Cloud, DevOps, backend, data, AI engineers
- In your team, your tools, your standup
- Ramp up or down with the roadmap
- You direct the work, we handle retention

Monthly rate per engineer · from 3 months

04

Managed operations

- Monitoring, patching, backup, DR drills
- Incident response with defined escalation
- Capacity and cost reviewed quarterly
- 24/7 where the service needs it

Monthly retainer · SLA-backed

We proudly introduce our products

Agency Technology Solutions

A suite of specialized platforms powering agencies, creators, advertisers, and AI-driven operations.



Data Studio
for advertiser and Buzz tracking



AI writing platform
for Generative Engine Optimization



Cloud and AI Consultancy
for innovation and sustainable technology



Operating System
for Agencies



Operating Platform
for Specialized Booking Operations

— LET'S BUILD TOGETHER · HỢP TÁC CÙNG CHÚNG TÔI



**Let's build
an unbreakable foundation!**