



PLAYBOOK

Commercializing Telco AI Infrastructure

A phased playbook for telcos to
successfully launch a revenue-ready
AI Factory



INTRODUCTION: FROM VISION TO EXECUTION

Telcos need a way to stand up a commercially viable AI Factory fast, without over-engineering or drifting through endless pilots.

This document highlights a roadmap that can bridge the gap between vision and operational reality: a plan that takes you from ready hardware to delivering production-grade AI services in roughly twelve weeks. The approach is anchored in two principles:

- **Scope for first revenue.** Start with a credible minimum catalog (fractional GPUs, inference endpoints) that customers can consume immediately.
- **Execute in parallel tracks.** While the core platform comes online, integrations, training, and support preparation run in parallel to avoid bottlenecks.

Each phase ends with clear, verifiable outcomes: a validated design, a non-production mirror, a live control plane with AI Service SKUs, and an expanded catalog supported by enterprise-ready operations.

By sequencing this way, telcos reduce risk while accelerating time-to-revenue from AI factory investments and strengthening their position as cloud-grade AI service providers.

PHASE 0: PLANNING (WEEK 1)

Phase 0 aligns business objectives with technical design. The goal is to create an approved architecture, define the first set of SKUs, and identify integration needs across the AI factory stack.



KEY ACTIVITIES

Workstream	Activities
Architecture Review & Design	Define the initial GPU Cloud blueprint, covering NVIDIA-certified compute clusters, networking, storage, monitoring, and billing systems
SKU Design	Identify the first monetizable SKUs
3rd-Party Tools Review & Design	Review integration needs across identity management, security, observability, CI/CD, and storage platforms.

READINESS CRITERIA

Criteria	Status
Architecture blueprint validated and approved	<input type="checkbox"/>
Draft SKU catalog	<input type="checkbox"/>
3rd-party integration requirements documented	<input type="checkbox"/>
Ownership for integration workstreams assigned	<input type="checkbox"/>

CHECKLIST: PHASE 0

Item	Covered
Architecture design approved	<input type="checkbox"/>
Initial monetizable SKUs defined (with draft pricing and quotas)	<input type="checkbox"/>
Integration scope reviewed and documented	<input type="checkbox"/>
Responsibilities and owners assigned	<input type="checkbox"/>

Telcos that start with a validated blueprint and a draft SKU catalog are able to reach revenue faster. Phase 0 avoids downstream rework and ensures all teams are aligned before execution begins.

PHASE 1: NON-PRODUCTION ENVIRONMENT SETUP (WEEKS 2-7)

With the design and SKUs defined, telcos establish a non-production environment that mirrors production. This provides a safe space to test services, validate integrations, and fine-tune tenant experiences.

KEY ACTIVITIES

Workstream	Activities	Assigned To
Platform Setup & Configuration	Deploy AI factory control plane, enable secure multi-tenancy, configure workflows for operators and tenants.	Rafay & Telco Operator
SKU Catalog Setup	Publish initial SKUs in a non-prod catalog.	Rafay & Telco Operator
Validation & Certification	Test SKUs for functionality, enforce quotas and SLAs, validate compliance and security controls including data residency and tenant isolation requirements.	Rafay & Telco Operator
UI/UX White-Labeling	Apply branding and tenant-facing customization for enterprise and B2B portals.	Rafay & Telco Operator
Systems Integration	Connect identity (SSO/IdP), OSS/BSS (billing, usage, dashboards), monitoring, logging, and storage systems.	Rafay & Telco Operator

READINESS CRITERIA

Criteria	Status
Control plane operational with validated multi-tenant workflows	■
Initial SKUs published and validated in non-prod	■
White-labeled tenant portal available for demo	■
OSS/BSS, IdP, monitoring integrations tested for at least one end-to-end usage and billing flow	■
SLA, quota, and compliance policies verified	■

CHECKLIST: PHASE 1

Item	Covered
AI factory control plane set up	
Initial SKUs live in non-prod	
Tenant experience branded and customized	
OSS/BSS & monitoring wired into platform	
SLA/compliance validated in test runs	

Why This Matters

Phase 1 de-risks the launch. By validating SKUs, integrations, and portals in a controlled environment, telcos gain confidence and can demonstrate progress to stakeholders before moving to live services.

PHASE 2: GO-LIVE (WEEK 8)

This is the **zero-to-one moment**: moving from test to live production. The focus is on activating the control plane, launching a minimal but credible set of SKUs, and validating billing and support workflows.

KEY ACTIVITIES

Workstream	Activities
Production Control Plane Activation	Deploy and configure the production environment, mirroring validated non-prod setup "including multi-tenancy, RBAC, and baseline policies.
SKU Offering Launch	Launch two initial revenue generating SKUs
Integration Validation	Test OSS/BSS workflows for billing, usage reporting, and SLA enforcement.
Go-Live Readiness	Complete operator training, finalize Day-2 runbooks, secure approvals for launch.

READINESS CRITERIA

Criteria	Status
Production control plane activated and healthy	■
First two SKUs live, priced, and accessible to tenants	■
OSS/BSS billing and usage workflows validated end-to-end	■
Operator training completed	■
Day-2 runbooks reviewed and approved	■

CHECKLIST: PHASE 2

Item	Covered
Production control plane active and aligned to validated non-prod configuration	
2 SKUs launched in production and ready for customer onboarding	
Billing & SLA workflows confirmed via test tenants	
Operators trained for Day-2 support (incidents, changes, capacity)	
Runbooks approved for readiness	

Why This Matters

Phase 2 turns infrastructure into a live service. By starting small, two SKUs with full billing and SLA validation, telcos prove commercial viability while limiting scope. This creates a foundation to scale with confidence and gives sales and partner teams something concrete to sell and demonstrate.



PHASE 3: EXPANSION & SCALE (WEEKS 9–12)

With production live, focus shifts to **broadening the catalog and hardening operations**. This phase adds new services, brings partners into the ecosystem, and strengthens monitoring and support processes.

KEY ACTIVITIES

Workstream	Activities
SKU Expansion	Add new compute and AI/GenAI SKUs such as distributed training clusters, fine-tuning pipelines, and GenAI assistants.
Marketplace Onboarding	Integrate ISV and partner applications into the catalog.
Operational Maturity	Enhance monitoring, SLA dashboards, and per-tenant and per-SKU cost attribution; implement FinOps practices.

READINESS CRITERIA

Criteria	Status
Expanded SKUs available in production catalog with documented pricing and SLA tiers	■
ISV/partner marketplace applications onboarded	■
Advanced monitoring and SLA dashboards operational	■
Cost attribution and FinOps controls in place with reports available to product and finance teams.	■
Enterprise support workflows tested and validated with at least one simulated major incident.	■

CHECKLIST: PHASE 3

Item	Covered
New compute & AI/GenAI SKUs launched and marketed to target customers	
ISV/partner marketplace onboarded with curated offers	
SLA & monitoring dashboards live and integrated with NOC/Operations workflows	
FinOps practices operationalized	
Enterprise support processes validated	

Phase 3 evolves the AI factory from an MVP into an enterprise-ready platform. A broader catalog enables new revenue streams, while operational maturity ensures the telco can scale reliably as demand grows.

PARALLEL WORKSTREAMS (WEEKS 8–12)

In parallel with the core rollout, supporting workstreams prepare teams and processes for enterprise-grade service delivery so that go-live doesn't stall on documentation or support readiness.

KEY ACTIVITIES

Workstream	Activities
Ops Documentation	Create detailed runbooks for provisioning, Day-2 operations, upgrades, and incident response including severity levels, escalation paths, rollback steps, and RFO templates.
Operator Training	Train telco ops teams on AI factory management, tenant workflows, SLA enforcement, and troubleshooting.
Enterprise Support Readiness	Establish enterprise-grade support, including escalation paths, ticketing workflows, and 24/7 monitoring coverage.
Continuous Lifecycle Management	Define patching, upgrade cadence, and governance processes including maintenance windows and approval workflow .

CHECKLIST: PARALLEL WORKSTREAMS

Item	Covered
Ops runbooks in place	
Operators trained and certified	
Enterprise support workflows live and tested via at least one simulated incident	
Continuous lifecycle management established and scheduled	

Why This Matters

Parallel workstreams ensure operators are trained, runbooks are in place, and enterprise support workflows are live without extending the 12-week timeline.

ENTERPRISE SUPPORT (WEEK 9 - N)

In parallel with the core rollout, supporting workstreams prepare teams and processes for enterprise-grade service delivery so that go-live doesn't stall on documentation or support readiness.

KEY ACTIVITIES

Workstream	Activities	Timeline
Ongoing Enterprise Support	Provide post go-live enterprise-grade support, including monitoring, escalation paths, ticketing workflows, and SLA enforcement.	Week 9+



LAUNCHING ENTERPRISE-READY TELCO AI FACTORY SERVICES

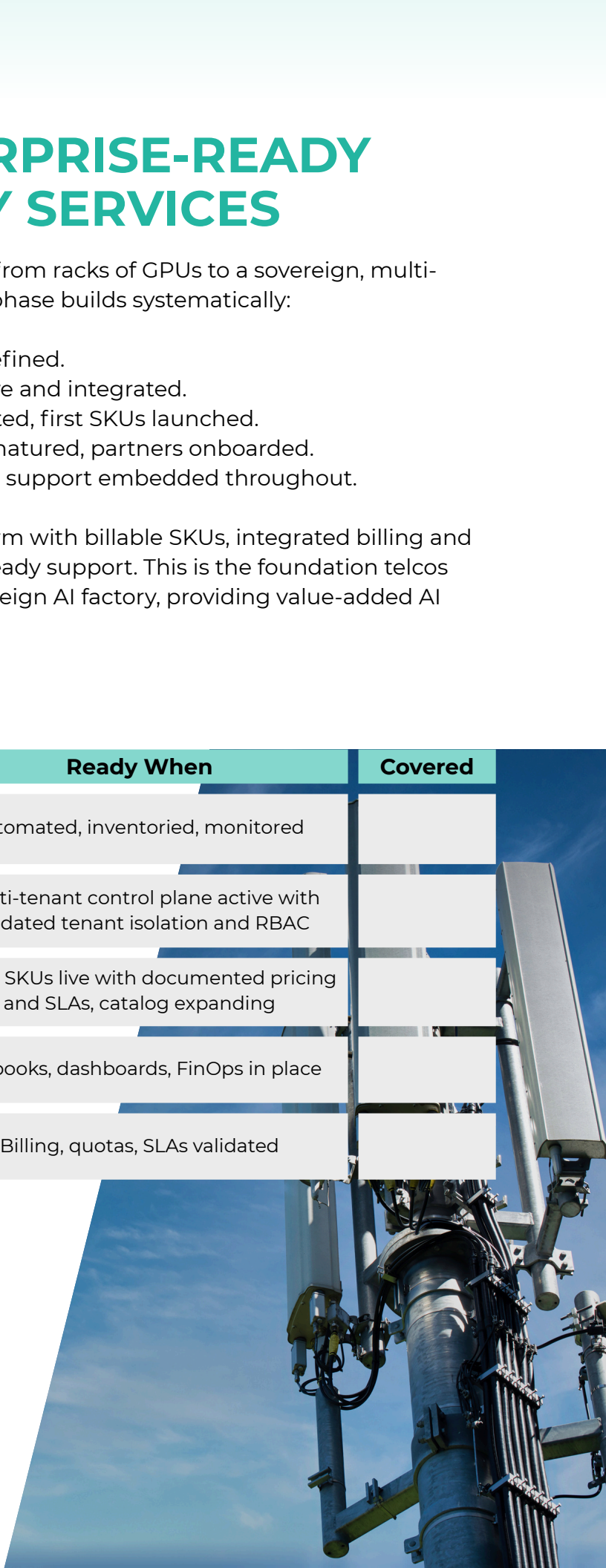
By following this roadmap, telcos can move from racks of GPUs to a sovereign, multi-tenant AI Factory in roughly 12 weeks. Each phase builds systematically:

- **Phase 0:** Design validated, initial SKUs defined.
- **Phase 1:** Non-production environment live and integrated.
- **Phase 2:** Production control plane activated, first SKUs launched.
- **Phase 3:** Catalog expanded, operations matured, partners onboarded.
- **Parallel Workstreams:** Ops, training, and support embedded throughout.

The outcome is a commercially viable platform with billable SKUs, integrated billing and monitoring, multi-tenancy, and enterprise-ready support. This is the foundation telcos need to position themselves as trusted sovereign AI factory, providing value-added AI infrastructure services.

FINAL READINESS CHECKLIST

Area	Ready When	Covered
Infrastructure	Automated, inventoried, monitored	<input type="checkbox"/>
Platform	Multi-tenant control plane active with validated tenant isolation and RBAC	<input type="checkbox"/>
Services	Initial SKUs live with documented pricing and SLAs, catalog expanding	<input type="checkbox"/>
Operations	Runbooks, dashboards, FinOps in place	<input type="checkbox"/>
Monetization	Billing, quotas, SLAs validated	<input type="checkbox"/>



SCALING TO MATURITY

By following this roadmap, telcos move from hardware to a commercially viable AI factory in twelve weeks. Each phase builds on the last, validated architecture, a live control plane, billable SKUs, and enterprise-ready operations, without over-engineering or extended pilots.

The outcome is a production-grade platform with the technical depth to serve regulated enterprises and the operational maturity to scale reliably as demand grows. Multi-tenancy, integrated billing, SLA enforcement, and FinOps controls are built in from Phase 0, ensuring the platform is commercially credible from day one.

About Rafay Systems

Turn complex compute environments into self-service consumption engines for enterprises, service providers, neoclouds and sovereign AI clouds. With Rafay's orchestration and workflow automation, your accelerated hardware gets transformed into composable infrastructure—ready for critical AI and cloud-native (Kubernetes-based) workloads.

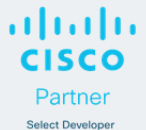


partner
network

Advanced
Technology
Partner



Inception
Program



Rafay Systems, Inc. | 530 Lakeside Dr, Ste 210 Sunnyvale, CA 94085 | 669.247.2551 | rafay.co

© 2026 Rafay Systems, Inc. All rights reserved. All other trademarks and service marks are property of their respective owners.