WOLFFISH RESEARCH

AGENT INFRASTRUCTURE · SEPTEMBER 2026

The agent stack just got a second centre of gravity.

Huawei opened a complete agent stack: an agent runtime inside the openEuler operating system, agentic tooling in CANN, shared Ascend compute, and a published MCP catalogue. Agent infrastructure is no longer a single-supplier story.

5,000+
General MCP assets published in the catalogue, plus 1,000+ industry-specific

10,000
Approximate Ascend NPUs in the shared clusters opened to developers

100
NPU-hours in the baseline developer programme — no hardware purchase needed

1st
OS-level agent runtime: ThinkPro schedules agents as a first-class workload

LAYER	US-CENTRIC STACK TODAY	HUAWEI'S STACK
Model access	Closed APIs, per-token billing	Hosted plus open-weight models
Agent runtime	Application-level frameworks	ThinkPro inside the OS itself
Silicon	CUDA and Nvidia accelerators	Ascend NPUs, open PTO ISA
Tool catalogue	MCP servers, fragmented per vendor	5,000+ general and 1,000+ industry assets, catalogued
Compute entry	Pay per token or per hour	Shared clusters plus a free baseline allocation
Governance	US and EU regulation	Domestic regime and export controls

THE ROW THAT MATTERS

Everyone else ships agents as applications. Putting the scheduler inside the operating system bets that agents become the normal way software runs.

WHY THE CATALOGUE IS THE HEADLINE

An agent with twenty tools is a demo; one with five thousand is infrastructure. MCP stays a shared, vendor-neutral protocol either way.

IF YOU ARE NOT IN THAT ECOSYSTEM

Nothing changes tomorrow, and something useful changes about your options: protocol-level portability becomes a real strategy, and the local, self-hosted agent stops being the exotic choice. Be sceptical of catalogue sizes — a tool your agent can call is also a tool that can call your agent.

SIGNAL TO WATCH	WHY IT MATTERS
Is the catalogue actually maintained?	Catalogue size is a press number. Maintenance is the only thing that makes a tool safe to call.
Does MCP portability survive?	If the protocol stays neutral, the work of building tools transfers between ecosystems. If it forks, it does not.
Do other operating systems copy ThinkPro?	An OS-level agent scheduler is the strongest signal yet that agents are becoming the default way software runs.

Sources: Huawei (September 2026), Power Technology, Huawei Connect 2026. Full analysis at wolffi.sh/blog/huawei-open-agentic-cloud