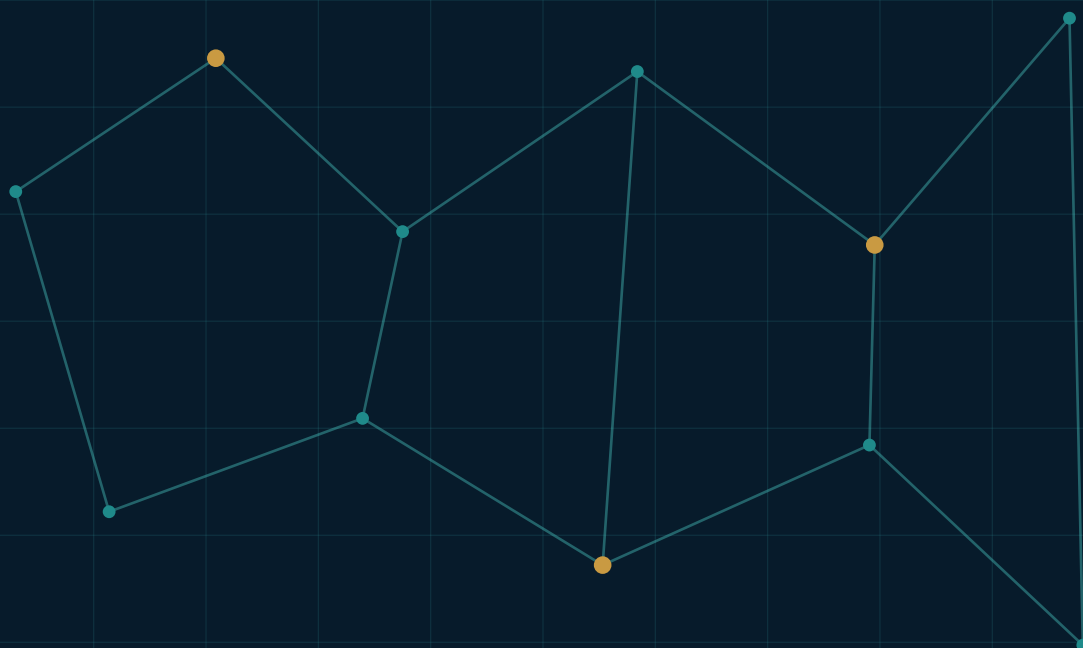




AI INFRASTRUCTURE & DIGITAL ECONOMY IN ASIA

Strategic Outlook 2026-2030

HONGKONG YIHE

Independent market research | 9 August 2026

How to read this report

Scope, evidence hierarchy and decision use

KEY TAKEAWAY

Asia should be analysed as a portfolio of specialised infrastructure and demand hubs, not as one homogeneous AI market.

Decision supported. This report helps corporate strategists, infrastructure investors, technology vendors and market-entry teams decide where to build, partner, sell or allocate capital across the Asian AI stack from 2026 to 2030.

Geographic scope. East Asia, South Asia and Southeast Asia, with focused analysis of China, Japan, South Korea, India, Singapore, Malaysia, Indonesia, Vietnam, Thailand, the Philippines, Taiwan and Hong Kong. Australia is used selectively in regional infrastructure benchmarks but is not treated as a core country chapter.

Evidence hierarchy. Tier 1 evidence includes government publications, multilateral institutions and company filings. Tier 2 includes established market research and industry surveys. Tier 3 includes HongKong YIHE planning models that convert evidence into prioritisation scores and scenarios.

Important comparability rule. Data centre capacity can refer to operational IT load, colocation supply, hyperscale self-build, on-premise capacity or a development pipeline. Digital economy figures can refer to GDP contribution, platform GMV, sector revenue or digitally delivered services. The report never adds these definitions together as if they were one market total.

As-of date

Public evidence reviewed through 9 August 2026. Fast-moving announcements after that date are outside the report snapshot.

This publication is strategic research, not investment, legal, tax or engineering advice. Any site selection requires utility confirmation, land and water diligence, grid studies, export-control screening and country-specific legal review.

Contents

Thirty pages organised from thesis to market action

Pages	Section	Decision question
04-06	Executive view	What is the central thesis and where are the structural bottlenecks?
07-10	Demand and semiconductor stack	Which parts of the AI value chain are most exposed to Asian capacity and know-how?
11-14	Data centres, power and networks	Where can high-density compute be delivered with reliable power and connectivity?
15-23	Market deep dives	Which economies fit compute, supply-chain, sovereign-AI or application roles?
24-25	Enterprise and digital monetisation	Where will infrastructure translate into durable revenue pools?
26-29	Capital, risk and action	How should companies sequence investment, partnerships and compliance?
30	Sources and producer	What evidence supports the analysis and who prepared it?

Core visual evidence

- APAC data centre capacity path from 32 GW in 2025 to 57 GW in 2030.
- Global and six-market APAC data centre power-demand comparisons.
- Country-role heatmap across demand, supply chain, power, interconnection, governance and deployability.
- Enterprise AI adoption comparison and 2030 infrastructure scenarios.

Notation	All dollar figures are US dollars unless a local currency is stated. GW means gigawatts of data centre capacity or IT load as defined by the cited source; TWh means annual electricity consumption.
----------	--

Executive Summary

The bottleneck is shifting from chips alone to coordinated infrastructure delivery

KEY TAKEAWAY

Asia is the indispensable production base for AI hardware and the fastest-scaling demand region, but value creation will concentrate where compute, power, interconnection, capital and governance can be assembled as one system.

- Infrastructure is entering a second constraint phase. Global data centre electricity demand rises from about 415 TWh in 2024 to 945 TWh in 2030 in the IEA base case. In Asia, power connection lead times, renewable availability and grid quality increasingly determine which announced projects become operational.
- Asia remains the critical hardware control point. Taiwan leads advanced foundry and packaging; Korea leads high-bandwidth memory; Japan supplies materials, equipment and precision manufacturing; China is building an increasingly independent compute stack; Malaysia and Vietnam are gaining assembly, test and electronics roles.
- The deployment map is multi-centred. Japan and Korea combine industrial demand with deep supply chains; India combines scale, software talent and public compute; Singapore is the governance and interconnection hub; Johor and other Southeast Asian markets provide land and expansion capacity; China operates as a large but increasingly self-contained ecosystem.
- Digital demand is broad enough to support local inference. Southeast Asia's digital economy reached roughly \$305 billion of GMV and \$135 billion of revenue in 2025. India expects its digital economy to approach 20% of GVA by 2029-30. Enterprise adoption is high, but production ROI, data quality and workflow redesign lag employee tool use.
- The best risk-adjusted exposure may sit beside the GPU. Grid upgrades, transformers, cooling, network fabrics, fibre, subsea cables, advanced packaging, memory, security and AI operations can capture growth without depending on one model vendor.

Five strategic conclusions

What management teams should act on now

57 GW JLL APAC capacity outlook for 2030	945 TWh IEA global data centre electricity demand in 2030	\$305B Southeast Asia digital GMV in 2025	78% APAC employees reporting weekly AI use
--	---	---	--

1. Power-ready land is more valuable than cheap land. A nominally attractive site without a contracted energisation path can destroy project economics. Treat utility capacity, curtailment risk, generation mix and substation delivery as acquisition conditions.
2. Training and inference will separate geographically. Frontier training concentrates in a small number of power-rich, capital-intensive clusters. Inference expands near users, regulated data and industrial workflows, creating a wider set of regional nodes.
3. Supply-chain resilience requires dual architectures. Export controls and national security policy are producing a US-aligned stack and a China-centred stack. Companies need product, software and compliance roadmaps for both where commercially necessary.
4. The enterprise opportunity is workflow redesign, not licence resale. High reported usage does not automatically produce productivity. Durable budgets move to vendors that connect models to proprietary data, measurable process outcomes and governance controls.
5. Country selection should follow role, not rank. A market may be excellent for financing and regional headquarters but poor for hyperscale power. Another may be suitable for data centres but weaker for high-value enterprise demand. Portfolio design should combine complementary locations.

Management implication	Use a two-stage gate: first establish infrastructure feasibility and compliance, then test unit economics and customer demand. Do not let announced investment substitute for delivered capacity or contracted revenue.
------------------------	---

The AI infrastructure bottleneck stack

Capacity is usable only when six layers are delivered together

Layer	Primary constraint	Asian control points	What to verify
Compute silicon	Accelerator availability, export controls, software compatibility	Taiwan foundry and packaging; Korea memory; China domestic accelerators	Allocation, permitted end use, software ecosystem
System integration	HBM, substrates, optics, network fabrics, rack power density	Korea, Taiwan, Japan, Malaysia	Bill of materials, thermal design, qualification
Facility	Power density, cooling, construction and commissioning	Japan, Korea, Singapore, Johor, Mumbai, Batam	Delivered MW, not announced campus size
Energy	Grid connection, generation mix, price and curtailment	Market-specific; often the binding layer	Firm power, substation schedule, PPA quality
Network	Latency, fibre diversity, subsea routes and cloud on-ramps	Singapore, Hong Kong, Tokyo, Mumbai; emerging ASEAN corridors	Route diversity, landing-station concentration
Data and governance	Data rights, localisation, model risk and cybersecurity	China-specific stack; India DPDP; ASEAN soft-law; Korea act	Workload classification and cross-border path

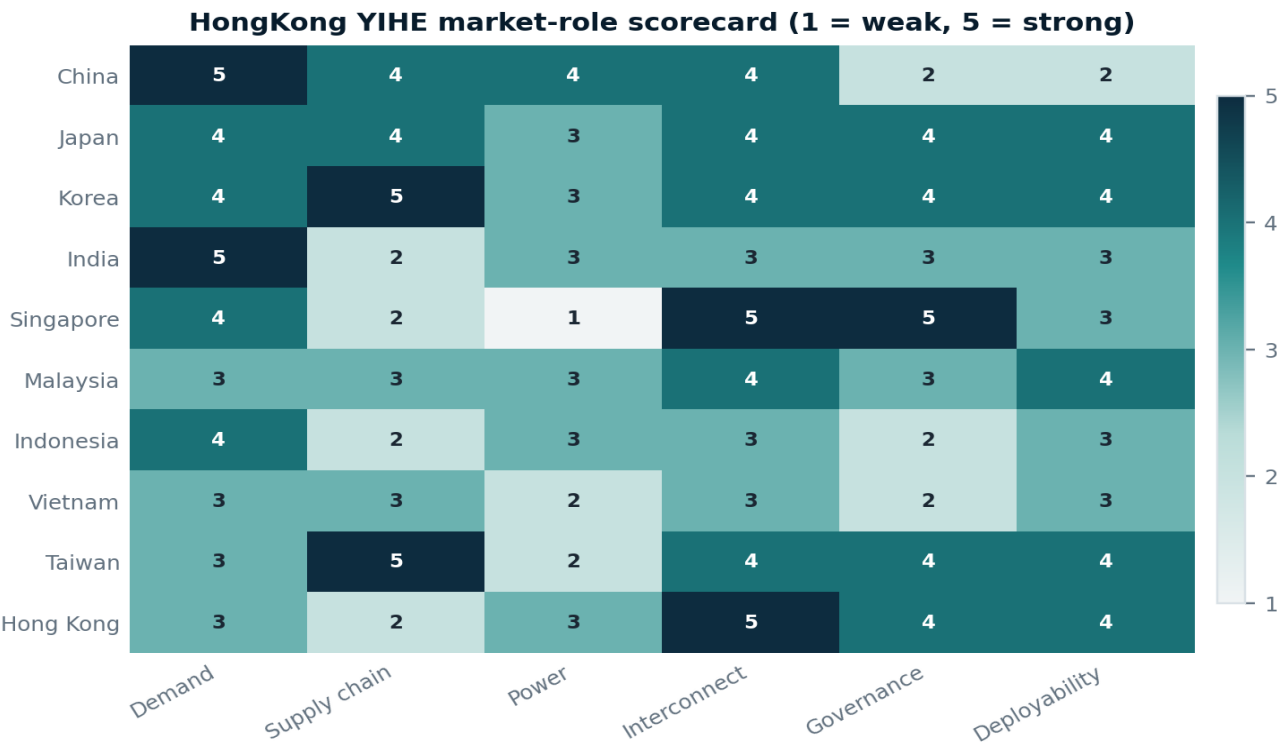
KEY TAKEAWAY

The commercial unit is not the chip or the building. It is a commissioned, networked and compliant megawatt that can support a defined workload at an acceptable total cost.

This changes diligence. GPU delivery does not prove service capacity; a land bank does not prove power; a national AI strategy does not prove usable data; and employee adoption does not prove enterprise ROI. Each layer needs an owner, evidence and a delivery date.

Asia is a portfolio of specialised hubs

Market roles are complementary rather than interchangeable



How to use the scorecard. Scores indicate suitability for a role, not an investability ranking. China scores highly on demand and system depth but lower on cross-border deployability. Singapore leads on interconnection and governance but is intentionally constrained by power. Taiwan and Korea dominate strategic hardware capabilities. India offers scale and talent while still building infrastructure depth.

Portfolio design	A common regional architecture pairs a control hub such as Singapore or Hong Kong, a scalable compute location such as Malaysia or India, and specialised supply-chain exposure in Taiwan, Korea or Japan.
------------------	--

Digital demand is becoming compute demand

Cloud migration, digital services and AI workloads reinforce one another

66% Asia-Pacific population online in 2024	\$305B SEA digital GMV in 2025	\$135B SEA digital-economy revenue in 2025	20% India digital-economy GVA target for 2029-30
--	--	--	--

Asia's AI infrastructure cycle is not supported by model training alone. It rests on mobile commerce, digital payments, streaming, gaming, cloud migration, industrial automation and public digital services. Training clusters centralise around high-density power and system performance; inference distributes toward users, data rights and latency. The demand base creates recurring revenue pools for local inference and model customisation.

The region still contains a major demand gap. ITU estimated that 66% of the Asia-Pacific population was online in 2024, close to the global average but with deep differences by income, geography, age and service quality. Future value therefore comes from both higher AI intensity among connected users and first-time or higher-quality connectivity for the rest.

Southeast Asia offers the clearest platform-revenue signal. e-Conomy SEA 2025 estimated approximately \$305 billion of GMV and \$135 billion of revenue across the digital economy. The distinction matters: GMV shows transaction scale, while revenue is closer to the pool available to platforms and service providers.

India's official measurement provides a wider GDP lens. The digital economy represented 11.74% of national income in 2022-23 and is projected to reach about 20% of GVA by 2029-30. These figures are not comparable with Southeast Asian GMV, but together they show the depth of the regional demand base.

Sources / note: Sources: ITU Asia-Pacific digital development; e-Conomy SEA 2025; Government of India digital-economy measurement.

Asia controls the critical semiconductor path

The AI system bill of materials runs through multiple Asian economies

Control point	Regional strength	Why it matters	Primary risks
Advanced foundry	Taiwan leads leading-edge contract manufacturing; Korea has advanced logic capability	Determines accelerator performance, yield and supply	Geopolitics, fab concentration, tool access
High-bandwidth memory	Korea is the strongest HBM production centre; Japan and Taiwan support equipment and packaging	HBM bandwidth is integral to AI system throughput	Qualification cycles, yields, allocation
Advanced packaging	Taiwan leads CoWoS-class integration; Korea and Japan invest in next-generation packaging	Combines compute dies, HBM and interconnect	Capacity, substrate, thermal and testing bottlenecks
Materials and equipment	Japan retains deep positions in specialty chemicals, wafers, components and precision equipment	Upstream quality and process control are hard to substitute	Export controls, single-source dependencies
Assembly, test and electronics	Malaysia, Vietnam, the Philippines and China provide scale	Enables diversification and lower-cost manufacturing	Talent depth, infrastructure and product mix

Taiwan's semiconductor industry data illustrate the concentration of value. The Taiwan Semiconductor Industry Association reported IC manufacturing revenue of NT\$3.42 trillion in 2024, up 28.4% year on year. TSMC's 2025 annual report continued to position high-performance computing and advanced packaging as central growth engines.

Investment implication

The most persistent bottlenecks often sit at the interfaces between technologies: HBM qualification, packaging yield, substrates, optics, power delivery and cooling. These interfaces deserve separate diligence from headline accelerator supply.

Sources / note: Sources: TSIA 2025 overview; TSMC 2025 Annual Report; SIA industry research; company filings.

HBM and packaging capture disproportionate value

AI memory economics reveal the scarcity premium in the system

KRW 97.1T SK hynix FY2025 revenue	KRW 47.2T SK hynix FY2025 operating profit	49% SK hynix FY2025 operating margin
---	--	--

SK hynix's record 2025 performance is a useful indicator of the economic value accruing to AI memory. The company reported KRW 97.1 trillion of revenue and KRW 47.2 trillion of operating profit, citing HBM and other high-value AI memory products as the central driver.

The lesson is structural rather than company-specific. AI accelerators require a tightly qualified combination of compute die, HBM stacks, interposer or bridge technology, substrates and high-speed interconnect. Performance depends on the package as a system, not on the processor alone.

Three factors support continued value capture. First, qualification cycles slow substitution. Second, yields are sensitive across multiple process steps. Third, the transition from HBM3E to HBM4 and larger packages increases integration complexity even as suppliers expand capacity.

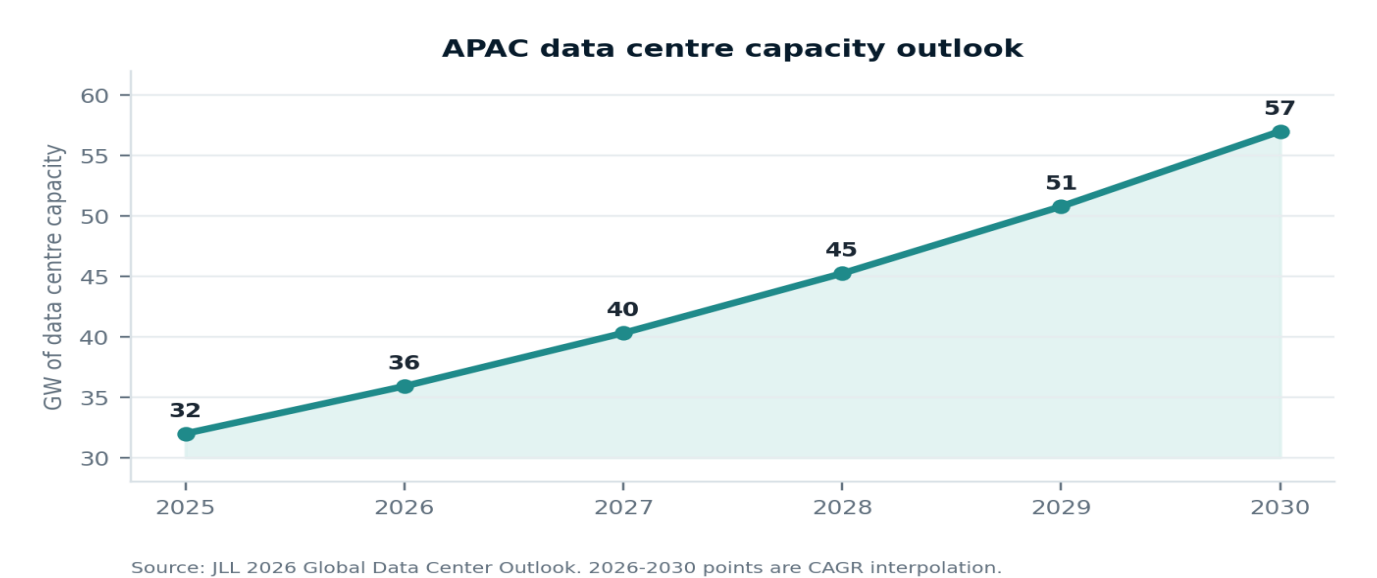
The counter-risk is cyclical. Memory pricing, customer concentration, yield ramps and aggressive capacity investment can compress margins. Investors and procurement teams should monitor bit growth, HBM mix, conventional DRAM pricing, customer qualification, advanced-packaging output and capex discipline together.

KEY TAKEAWAY

AI hardware scarcity is not one bottleneck. It is a chain of linked yields. The weakest qualified component can limit the saleable output of the entire rack.

APAC capacity is scaling toward 57 GW

Colocation growth and cloud migration broaden the investment cycle



JLL projects APAC data centre capacity to expand from 32 GW in 2025 to 57 GW by 2030, a roughly 12% compound annual growth rate. The same outlook expects colocation to lead growth while enterprise on-premise capacity declines as workloads migrate to cloud and shared facilities.

A second dataset provides a narrower but useful operational view. Cushman & Wakefield reported 13.8 GW of operational capacity in the markets it tracked at the end of 2025, after 1.56 GW of additions during the year. Its H1 2026 pipeline reached 26.5 GW, including 4.8 GW under construction and 21.7 GW planned.

Do not combine the totals

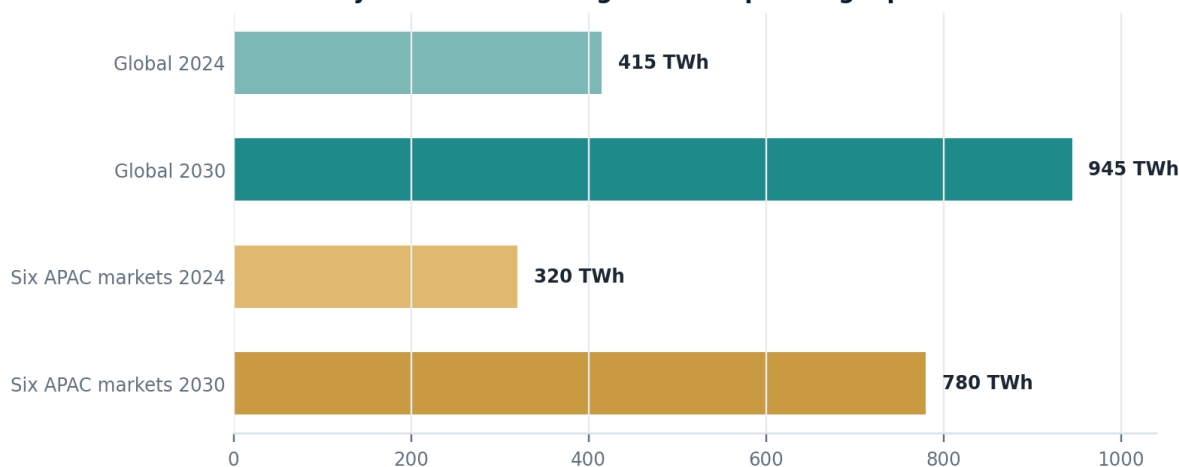
JLL's 32 GW regional base and Cushman & Wakefield's 13.8 GW tracked operational stock use different scopes. The strategic message is consistent, but the figures should not be added or used as substitutes.

Growth is moving beyond the traditional Singapore, Hong Kong, Tokyo and Sydney hubs toward Johor, Mumbai, Melbourne, Batam and other power-accessible corridors. However, secondary-market expansion raises execution risk in transmission, water, skills, permitting and network diversity.

Electricity is the binding constraint

Demand growth is faster than the surrounding power system

Data centre electricity demand is moving from an operating input to a site-selection constraint



Sources: IEA Energy and AI; PwC six-market APAC study. Definitions are not identical.

The IEA base case projects global data centre electricity use to rise from about 415 TWh in 2024 to 945 TWh by 2030. China accounted for approximately 25% of global data centre electricity use in 2024, making Asian power-system decisions central to the global AI build-out.

PwC's six-market Asia-Pacific study estimates data centre electricity consumption rising from 320 TWh in 2024 to 780 TWh in 2030. The study expects only 32% of the 2030 demand to be met by renewable energy under the assessed trajectory. This is a scenario for selected large markets, not a directly comparable subset of the IEA global total.

Power risk has four components: connection timing, firmness, price and carbon content. Facility readiness adds rack density, liquid-cooling capability, water constraints and measured efficiency at real load. A signed allocation can still face substation or transmission delays; a powered site can still fail customer thermal, reliability or decarbonisation requirements.

KEY TAKEAWAY

The most bankable AI data centre is not the one with the lowest headline tariff. It is the one with a credible energisation schedule, firm capacity, manageable price risk and a pathway to lower-carbon supply.

AI clusters depend on network geography

Interconnection and subsea diversity determine usable regional reach

78 systems New subsea systems expected online in 2023-2025 map cycle	>300,000 km Cable length in that build cycle	~50% Major technology-company role in trans-Pacific projects
--	---	--

Training clusters require east-west bandwidth inside and between facilities. Inference requires reliable north-south reach to users, data sources and cloud regions. Subsea cables, metro fibre, internet exchanges and cloud on-ramps therefore form a single commercial layer.

TeleGeography estimated that the 2023-2025 cable boom would bring 78 systems and more than 300,000 kilometres online. OECD analysis notes that major technology companies back nearly half of trans-Pacific cable investments scheduled to enter service between 2023 and 2025. This vertical integration can improve capacity but also concentrates route ownership and strategic influence.

Singapore, Hong Kong and Tokyo remain major interconnection hubs. Mumbai is strengthening its position on South Asia routes. Johor benefits from proximity to Singapore but must prove route and facility independence. Indonesia's archipelagic geography increases the value of domestic fibre and resilient international landings.

KEY TAKEAWAY

A new data centre market is investable only when power growth is matched by diverse network growth. A single landing corridor or cross-border dependency can turn cheap capacity into fragile capacity.

Cloud investment is localising AI capacity

Announced commitments show where providers expect demand and policy support

Provider / market	Announced commitment	Period / purpose
AWS - Japan	JPY 2.26 trillion	Existing cloud infrastructure through 2027
AWS - Singapore	Additional SGD 12 billion	Cloud infrastructure through 2028; total planned investment exceeds SGD 23 billion
AWS - India	US\$12.7 billion	Local cloud infrastructure through 2030
AWS - Thailand	>US\$5 billion	New region and long-term infrastructure plan
Microsoft - Japan	US\$2.9 billion	AI and cloud investment announced for a two-year period
Microsoft - Indonesia	US\$1.7 billion	Cloud and AI infrastructure plus skills over four years
Microsoft - Malaysia	US\$2.2 billion	Cloud and AI infrastructure and ecosystem investment

The figures are commitments in different currencies, periods and cost categories; they are signals of direction, not a comparable ranking. Local investment is driven by data residency, latency, enterprise demand and public procurement. Sovereign-AI programmes add demand for assured compute, governed datasets, local models and jurisdictional cloud control, but countries do not need to own every layer.

Cloud financials also show the cost of scaling. Microsoft's 2025 annual report noted pressure on Azure gross margin from AI infrastructure expansion. Amazon reported \$128.3 billion of cash capital expenditure in 2025, primarily reflecting technology infrastructure and fulfilment capacity, and guided to higher spending in 2026.

Commercial test	An investment announcement is not contracted local AI capacity. Verify the availability zone, accelerator type, service launch date, quota, pricing, data-residency terms and disaster-recovery path.
-----------------	---

Sources / note: Sources: AWS investment announcements; Microsoft 2025 Annual Report and country announcements; Amazon 2025 Annual Report.

China: scale with a separate technology stack

A vast demand pool, strong infrastructure and rising policy separation

25% China share of global data centre electricity use in 2024	50.2T yuan China digital economy estimate in 2022	1B+ Internet-user scale
---	---	---

China combines one of the world's largest digital economies with broad domestic cloud, model, device and industrial ecosystems. The IEA estimated that China accounted for about one quarter of global data centre electricity consumption in 2024. This scale supports training, consumer inference, industrial AI and government applications.

The strategic constraint is technology bifurcation. Export controls limit access to advanced foreign accelerators, while policy increasingly supports domestic chips, software and state-funded infrastructure. The result is not simply a slower copy of the global stack; it is a separate optimisation path with different hardware, frameworks, procurement rules and economics.

China remains attractive for companies with local demand, approved products, domestic partners and the ability to operate a localised data and technology architecture. Cross-border services face more complex cybersecurity, algorithm, generative-AI and data-export requirements.

Recommended role	Treat China as a dedicated market architecture with its own product roadmap, supply chain, cloud choices, compliance model and unit economics. Do not assume a regional platform built elsewhere can be extended unchanged.
------------------	---

Watchpoints: domestic accelerator performance and software compatibility; state procurement; data-centre power utilisation; model price competition; AI-generated content labelling; and any change in export-control enforcement.

Japan: industrial AI and trusted infrastructure

A mature economy with manufacturing depth and a digital-diffusion gap

Japan offers a distinctive combination: advanced materials and equipment, sophisticated manufacturing customers, large financial and industrial groups, reliable institutions and substantial cloud investment. Its opportunity is less about consumer-platform hypergrowth and more about industrial, enterprise and physical AI.

Strength	Strategic relevance
Semiconductor materials and equipment	Provides upstream control points and supports domestic capacity revival
Industrial base	Creates high-value use cases in robotics, automotive, machinery, energy and healthcare
Cloud investment	AWS plans JPY 2.26 trillion through 2027; Microsoft announced US\$2.9 billion over two years
Governance	METI and MIC updated AI Guidelines for Business to Version 1.2 in 2026
Regional infrastructure	Tokyo and Osaka are major cloud and interconnection markets; new supply faces power and construction constraints

Japan's adoption paradox is commercially useful. Excellent infrastructure and engineering coexist with slower digital diffusion in parts of the service sector and SME base. Vendors that integrate AI into established workflows, deliver Japanese-language quality and support change management can unlock value that generic tools miss.

Recommended role	Use Japan for trusted enterprise inference, industrial AI, semiconductor and materials partnerships, and high-value sector solutions. Price for integration, governance and service quality rather than commodity model access.
------------------	---

South Korea: memory leadership meets AI ambition

A compact market with global hardware champions and demanding digital users

South Korea is a strategic hardware and application market. It combines leadership in HBM and memory, advanced electronics, dense connectivity, sophisticated consumer platforms and strong manufacturing use cases. The country's national AI agenda seeks to add domestic compute and model capability to that hardware base.

KRW 97.1T SK hynix 2025 revenue	0.95 Korea 2025 OECD Digital Government Index	2026 AI Framework Act effective year
---	---	--

The AI Framework Act took effect in January 2026, establishing a national governance structure and obligations around transparency and high-impact AI. Korea's policy direction is supportive of innovation but more formal than a purely voluntary regime. International suppliers need Korean-language compliance, local representation where required and strong documentation.

For infrastructure, the key questions are grid capacity, local permitting, compute-centre delivery and the balance between national capability and global cloud providers. For enterprise applications, chaebol supply chains can provide scale but require long sales cycles, qualification and local partnerships.

Recommended role	Prioritise Korea for HBM and advanced-system partnerships, manufacturing AI, telecom and consumer inference, and high-value enterprise deployments. Build around local champions rather than treating the market as a simple cloud-resale opportunity.
------------------	--

India: scale, talent and public compute

The region's largest demand-and-talent expansion case

34,000+ Common-compute GPUs by May 2025	20% Projected digital-economy share of GVA by 2029-30	1,800+ Global capability centres reported in 2025
---	---	---

India combines a vast user base, software and services talent, global capability centres, digital public infrastructure and fast-growing cloud investment. Its official digital-economy report projects the sector to reach about one fifth of GVA by 2029-30, from 11.74% in 2022-23.

The IndiaAI Mission uses public-private capacity to lower compute access barriers. Common compute crossed 34,000 GPUs in May 2025, and the government later reported 38,000 deployed GPUs. The opportunity is not just national models: startups, researchers, MSMEs and public agencies can use subsidised capacity for local applications.

Execution risk remains uneven power quality, land and grid bottlenecks, state-level variation, data-centre concentration, water stress and enterprise procurement. India's data-protection regime is also moving into implementation, requiring disciplined consent, security and breach processes.

Recommended role	Use India for AI development, global capability centres, cost-sensitive inference, local-language applications and selected hyperscale deployment. Pair national strategy with state-level diligence on power, fibre, land and incentives.
------------------	--

Sources / note: Sources: Digital India; Government of India digital-economy measurement; IndiaAI Mission; DPDP Rules 2025.

Singapore: control tower, not unlimited compute basin

Governance, capital and interconnection offset physical constraints

Singapore remains Southeast Asia's premium control hub for cloud services, regional headquarters, finance, cybersecurity, interconnection and AI governance. Its constraint is intentional scarcity: land, power and carbon considerations limit unbounded data-centre growth.

Role	Why Singapore leads	Constraint
Regional headquarters	Stable legal system, talent, capital and multinational density	High operating cost
Interconnection	Dense carrier, cloud and subsea ecosystem	Concentration risk and dependence on regional routes
Trusted AI	Testing, governance and public-private coordination	Voluntary frameworks still require company implementation
Enterprise demand	Finance, logistics, healthcare and government use cases	Small domestic population
Data centres	Existing hyperscale and colocation base	Power and sustainability allocation

The Green Data Centre Roadmap aims to unlock at least 300 MW of additional capacity through energy efficiency and clean energy. AWS plans an additional SGD 12 billion of infrastructure investment through 2028. These commitments reinforce Singapore's hub position without removing the physical constraint.

Recommended role	Place governance, regional sales, cloud control planes, interconnection and regulated enterprise workloads in Singapore. Use Johor and other markets for scalable capacity where latency, resilience and regulatory design permit.
------------------	--

Malaysia and Johor: scale next to the hub

The leading Southeast Asian capacity expansion story now faces a quality test

RM 144.4B Approved data-centre and cloud investment, 2021 to mid-2025	400 MW Approximate Johor live capacity cited in early 2025 research	3.3 GW Johor planned capacity cited in early 2025 research
---	---	--

Malaysia has emerged as the main expansion corridor for Singapore-linked data centre demand. Johor offers land, proximity and a developing operator ecosystem, while Kuala Lumpur provides enterprise demand and connectivity. MIDA reported RM 144.4 billion of approved data centre and cloud-computing investment between 2021 and mid-2025.

The next phase is about delivery quality rather than announcement volume. Power-system impact, water use, local content, chip-export controls and project sustainability are receiving more scrutiny. Projects that cannot demonstrate resource efficiency and credible end users may face slower approvals.

The Johor-Singapore pairing is commercially powerful but requires independent resilience. A facility in Johor should not rely on one cross-border network path, one power allocation or assumptions that all Singapore-based workloads can be moved across the border.

Recommended role	Use Malaysia for scalable regional capacity, cloud regions, AI inference and selected high-density clusters. Make power, water, route diversity, end-user compliance and accelerator export controls explicit investment conditions.
------------------	--

Indonesia: demand scale with infrastructure dispersion

A large digital economy needs more domestic compute and fibre depth

Indonesia has Southeast Asia's largest population and one of its largest digital-consumer markets. This creates strong demand for local cloud, payments, commerce, content, gaming, government and AI inference. Data localisation and latency add to the rationale for domestic capacity.

Opportunity	Evidence	Execution question
Consumer inference	Large mobile-first population and digital-platform use	Can unit costs support mass-market local-language services?
Cloud localisation	Global-provider investment and local data-centre expansion	Are zones, accelerators and network routes live at required scale?
Batam corridor	Proximity to Singapore and new campus development	Is power and cable diversity independent enough?
Sovereign and public-sector AI	National AI roadmap and public-service needs	Can procurement, skills and governance support production systems?
Industrial AI	Resources, logistics, manufacturing and agriculture	Can solutions operate across dispersed geography and variable connectivity?

A DayOne-INA campus in Batam was financed for approximately 72 MW, while CoreWeave announced three Indonesian sites with a combined 360 MW planned for 2028. Announcements indicate momentum, but delivered capacity, customer commitments and network integration remain the real milestones.

Recommended role	Prioritise local inference, consumer platforms, public services and industrial applications. Use Jakarta for demand and cloud access, and assess Batam as a regional capacity corridor with separate power and connectivity diligence.
------------------	--

Sources / note: Sources: e-Economy SEA 2025; Indonesian AI roadmap reporting; DayOne-INA and CoreWeave announcements.

Vietnam, Thailand and the Philippines

Three demand markets with different infrastructure and industrial roles

Market	Primary AI role	Infrastructure edge	Binding constraints
Vietnam	Manufacturing AI, software services, consumer and local-language applications	Electronics supply chain, young technical talent, rapid digital growth	Power reliability, data regulation, advanced compute access
Thailand	Cloud region, automotive and industrial AI, digital commerce	AWS launched a region and plans more than US\$5B of investment	Skill depth, policy execution, power and regional competition
Philippines	BPO transformation, customer-service AI, fintech and content	English-language talent, large services sector, subsea connectivity	Power cost, latency, infrastructure concentration and reskilling

Vietnam targets a digital economy equal to 20% of GDP and 30% by 2030 under its strategy. The country's strongest differentiated opportunity is the intersection of electronics manufacturing, software talent and digital-consumer growth. However, high-density infrastructure must be matched by reliable power and a clear compliance route.

Thailand has moved from cloud demand to infrastructure delivery with the AWS Asia Pacific (Thailand) Region launched in January 2025. Its automotive, electronics, tourism and retail sectors create sector-specific AI use cases.

The Philippines has a large services workforce exposed to both augmentation and automation. The strategic opportunity is to move from labour-only BPO toward AI-enabled operations, workflow design, quality assurance and regulated data services.

Portfolio implication	Treat these markets as application and industry nodes first. Scale infrastructure only where power, network and anchor customers are evidenced.
-----------------------	---

Sources / note: Sources: Vietnam digital-economy strategy; AWS Thailand; e-Economy SEA; ITU connectivity research.

Taiwan and Hong Kong: supply-chain and capital gateways

Two compact markets with outsized regional functions

Dimension	Taiwan	Hong Kong
Core role	Advanced foundry, packaging, electronics and AI hardware ecosystem	Capital, professional services, regional connectivity and Greater Bay Area gateway
Demand	Semiconductor, manufacturing, devices and cloud	Finance, trade, logistics, professional services and cross-border business
Infrastructure	Strong technical ecosystem; power and geopolitical concentration risk	Mature connectivity and data centres; land and cost constraints
Strategic value	Irreplaceable near-term hardware control points	Financing, treasury, deal-making and governance functions
Risk	Cross-strait tension, power resilience, concentration	China-West policy exposure, regional competition, market scale

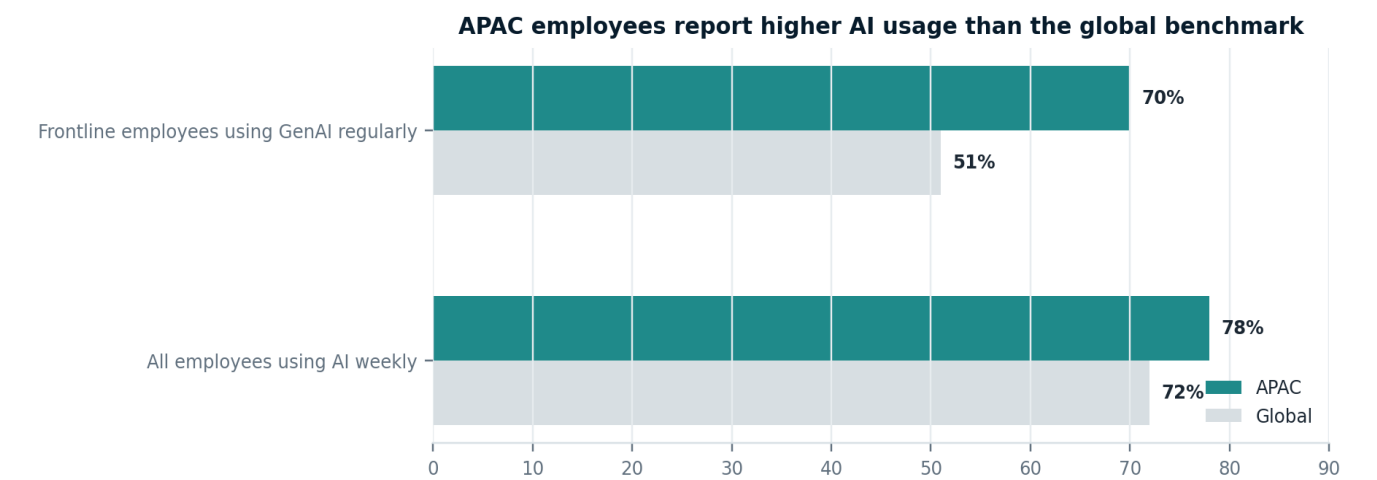
Taiwan's strategic value is physical and technical. Its advanced manufacturing and packaging base is central to accelerator supply, while the surrounding design, substrate, testing and equipment ecosystem creates dense know-how. Diversification outside Taiwan is increasing, but leading-edge substitution takes years.

Hong Kong's value is connective. It remains a major financial centre, cloud and network location, professional-services hub and route into the Greater Bay Area. The Shenzhen-Hong Kong-Guangzhou innovation cluster ranked first in WIPO's 2025 cluster ranking, illustrating the broader regional ecosystem.

Recommended role	Use Taiwan for hardware and supply-chain partnerships. Use Hong Kong for regional finance, corporate structuring, client coverage, professional services and selected regulated workloads. Avoid treating either as a low-cost hyperscale deployment market.
------------------	--

Adoption is high; operating maturity is uneven

Reported usage is ahead of workflow redesign and measured ROI



Source: BCG, October 2025. Survey results measure reported use, not realized enterprise ROI.

BCG reported that 78% of APAC respondents used AI at least weekly in 2025, compared with 72% globally. Frontline adoption was 70% in APAC versus 51% globally. Separate Microsoft research found extensive employee-led usage and a large share of users bringing their own AI tools into work.

This is a strong demand signal but a weak ROI metric. Employee use measures behaviour, not whether the enterprise has redesigned processes, improved cycle time, reduced error, increased revenue or managed risk. It can also create data leakage and shadow-AI exposure.

Maturity stage	Primary management question	Infrastructure implication
Tool access	Who can use which model safely?	Identity, secure endpoints and model gateway
Copilot	Which tasks become faster or higher quality?	Data retrieval, logging and cost controls
Workflow	Can a full process be redesigned?	Integration, orchestration, evaluation and human review
Agentic operations	Can systems act with bounded autonomy?	Permissions, observability, rollback and liability

Sources / note: Sources: BCG APAC AI Adoption 2025; Microsoft Work Trend research. Survey results are self-reported.

Where AI infrastructure monetises

Revenue pools differ in latency, regulation, data intensity and willingness to pay

Sector	High-value AI workloads	Infrastructure requirement	Commercial test
Financial services	Fraud, risk, service, research, compliance	Low latency, secure data, auditability, local regulation	Loss reduction, revenue uplift, control evidence
Commerce and media	Search, recommendation, advertising, content, agents	High concurrency, low unit cost, multilingual inference	Gross profit per query or transaction
Manufacturing	Vision, predictive maintenance, engineering, robotics	Edge compute, private network, deterministic response	Yield, downtime, quality and safety
Healthcare	Imaging, documentation, triage, discovery	Sensitive data, validation, clinical governance	Outcome, throughput and liability
Government	Citizen service, language, fraud, planning	Sovereign control, accessibility, procurement	Service quality, inclusion and trust
Software and BPO	Code, testing, service operations, workflow agents	Secure model access, integrations and observability	Revenue per employee and quality
Mobility and logistics	Routing, autonomy, maintenance, supply chains	Edge-cloud coordination and real-time data	Utilisation, time and loss reduction

KEY TAKEAWAY

The highest-value applications combine proprietary data, high-frequency decisions and measurable economic outcomes. The most scalable applications also minimise inference cost and human exception handling.

Market-entry teams should size the revenue pool by workflow and customer segment, not by multiplying a national AI-market forecast. Infrastructure demand follows real workloads, and workloads follow budgets controlled by specific business owners.

Invest across the stack, not only in compute

Risk-adjusted value is strongest where scarcity meets repeat demand

Layer	2026-2030 thesis	What could go right	Primary downside
Accelerators	Demand remains strong but vendor and export-control concentration is high	Performance leadership and scarcity pricing	Custom silicon, policy, utilisation and price compression
HBM and packaging	Complexity and qualification support value	Higher content per system and linked yields	Capacity overshoot and memory cyclical
Power and grid	Every delivered MW needs generation and connection	Long-duration contracted demand	Regulation, project delays and capital intensity
Cooling and electrical	Rack density drives redesign	Content growth per MW and retrofit demand	Standards change and customer concentration
Fibre and subsea	Regional inference and cloud require route diversity	Recurring capacity demand	Overbuild, route concentration and geopolitics
Data centres	Capacity expands with cloud and AI	Lease-up and real-asset value	Power delays, stranded sites, financing cost
AI platforms and software	Value shifts to workflow integration and data	High gross margins and vertical scale	Model commoditisation and weak ROI

A balanced capital strategy separates volume exposure from scarcity exposure. Accelerators and data-centre shells depend heavily on utilisation. Transformers, switchgear, cooling, interconnect, memory and packaging may benefit from system complexity but can also face capacity cycles.

Investment discipline	Track delivered capacity, contracted customers, backlog quality, capitalised interest, financing structure, power milestones and return on invested capital. Avoid using addressable-market growth as a substitute for project economics.
-----------------------	---

Regulation is becoming infrastructure design

Data, model and hardware policy now shape architecture before launch

Jurisdiction	Direction as of August 2026	Architecture implication
China	Generative-AI, algorithm, cybersecurity, data-export and content-labelling controls	Dedicated local stack, filings, data mapping and content controls
South Korea	AI Framework Act effective January 2026 with transparency and high-impact provisions	Governance documentation, local obligations and model-risk review
Japan	Soft-law-led AI Guidelines for Business Version 1.2	Lifecycle governance, voluntary controls and sector regulation
India	DPDP Rules 2025 with phased implementation	Consent, security, breach response and children's-data controls
ASEAN	Expanded GenAI Guide, Responsible AI Roadmap and AI Safety Network	Interoperable governance baseline, but country laws still control
Singapore	Pro-innovation governance, testing and voluntary assurance mechanisms	Evaluation, documentation and sector-specific compliance

Regulation changes where data can move, which model can be used, how outputs are labelled, who is responsible and whether a foreign provider needs local presence. These are system requirements, not a legal appendix after deployment.

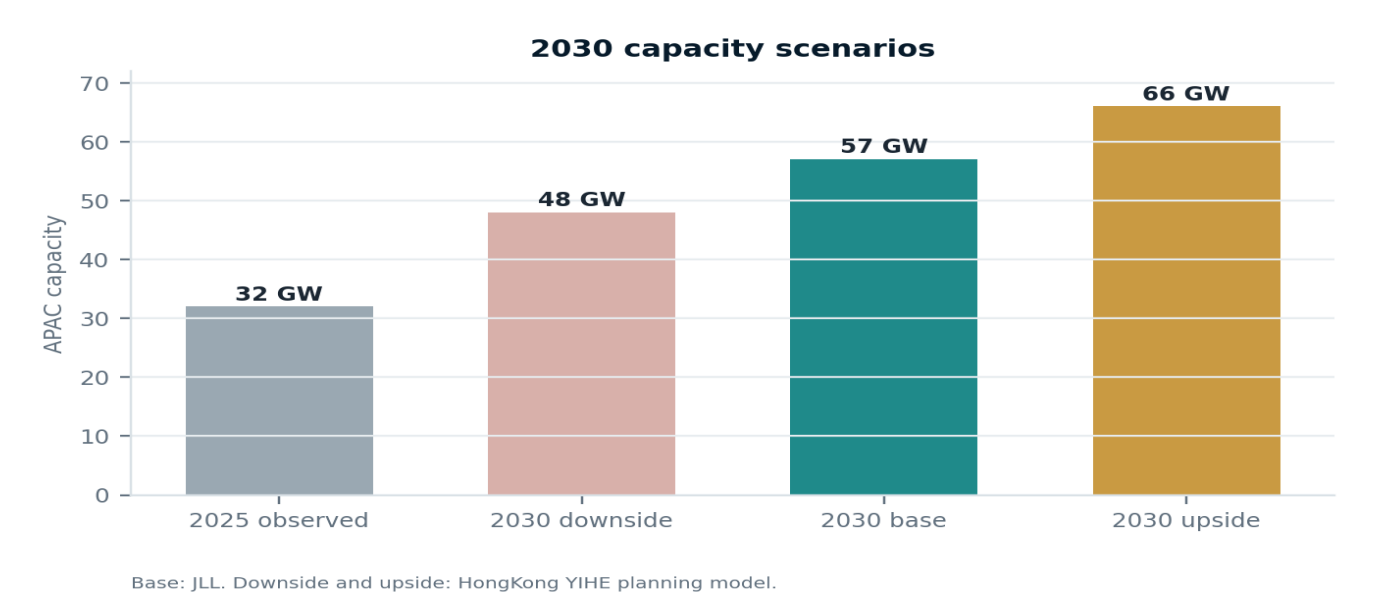
Export controls create an additional hardware-governance layer. End-user identity, workload, chip performance, destination and ownership structure can affect accelerator access. Regional intermediaries and data-centre operators need stronger customer and end-use diligence.

Minimum control set	Maintain a workload register, data lineage, model inventory, evaluation protocol, human-oversight rules, incident plan, vendor and end-user screening, and jurisdiction-specific deployment matrix.
---------------------	---

Sources / note: Sources: ASEAN AI governance materials; Japan METI; India MeitY; Korea AI Framework Act; China AI governance measures.

2030 scenarios: power and ROI define the range

The base case remains strong, but announced capacity is not guaranteed delivery



Scenario	Capacity	Conditions	Strategic response
Downside	48 GW	AI ROI disappoints; export controls tighten; grid and financing delays cancel projects	Prioritise contracted infrastructure, retrofits and efficiency
Base	57 GW	JLL outlook; cloud migration and inference scale; power constraints slow but do not stop growth	Build a multi-market portfolio with power and customer gates
Upside	66 GW	Inference demand broadens; sovereign projects accelerate; power and cooling delivery improves	Secure land, power and network options before scarcity reprices

The downside and upside values are HongKong YIHE planning scenarios around JLL's 57 GW base, not third-party forecasts. They are intended to test capital resilience. The downside still implies growth from the 2025 base; the question is how much of the pipeline earns its cost of capital.

Leading indicators	Monitor utility connection queues, construction starts, accelerator delivery, cloud-service quotas, HBM and packaging output, colocation pre-leasing, enterprise AI budgets, model inference prices and regulatory restrictions.
--------------------	--

A 100-day Asia AI infrastructure plan

Move from regional narrative to investable and executable decisions

Timing	Workstream	Required output	Go / no-go gate
Days 1-20	Demand and workload	Priority workloads, customer segments, latency, data and density requirements	Named budget owner and measurable outcome
Days 1-30	Market-role screen	Country shortlist by control hub, compute, supply chain and demand role	No single composite score without role definition
Days 15-45	Infrastructure diligence	Power letter, substation path, cooling design, water, fibre and land review	Credible energisation and route diversity
Days 25-55	Technology and supply chain	Accelerator, memory, network, software and export-control matrix	Qualified and permitted bill of materials
Days 35-65	Governance	Data map, model inventory, risk tier, localisation and incident plan	Workload approved for target jurisdiction
Days 50-80	Commercial model	TCO, utilisation, pricing, customer commitments and downside case	Risk-adjusted return exceeds hurdle
Days 70-100	Partnership and launch	Utility, operator, cloud, telecom, government and channel workplan	Contracted milestones, owners and stop-loss points

Board decision template. Approve the role, market, workload, capacity, customer, power path, compliance architecture, downside liquidity and milestones together. Defer any project where one of these remains a narrative rather than evidence.

KEY TAKEAWAY

The winning Asia strategy is a portfolio: specialist supply-chain exposure in Northeast Asia, scalable infrastructure in selected South and Southeast Asian corridors, and commercial control in trusted regional hubs.

HongKong YIHE can extend this report into a country-entry study, site and partner shortlist, competitive map, investment thesis, customer demand assessment or project-level diligence framework.

Sources and producer

Selected evidence base; full URLs are embedded as clickable links

1. International Energy Agency
Energy and AI: Energy demand from AI
2. International Energy Agency
Energy and AI: Executive summary
3. JLL
2026 Global Data Center Outlook
4. Cushman & Wakefield
APAC Data Centre Update H1 2026
5. PwC
Closing the clean energy gap for Asia Pacific data centres
6. CBRE
Global Data Center Trends 2025
7. Google, Temasek and Bain
e-Conomy SEA 2025
8. World Bank
Digital Progress and Trends Report 2025: AI Foundations
9. Asian Development Bank
Asian Economic Integration Report 2025
10. International Telecommunication Union
State of digital development and trends in Asia and the Pacific 2025
11. TSMC
2025 Annual Report
12. Taiwan Semiconductor Industry Association
Overview on Taiwan Semiconductor Industry 2025
13. SK hynix
FY2025 Financial Results
14. Singapore IMDA
Green Data Centre Roadmap
15. Malaysia Investment Development Authority
Building Resilience Through Localisation
16. Digital India
India common compute capacity crosses 34,000 GPUs
17. Government of India
India digital economy to contribute one-fifth of national income
18. AWS
Japan cloud infrastructure investment through 2027
19. AWS
Additional S\$12 billion Singapore investment through 2028
20. AWS
US\$12.7 billion India cloud infrastructure plan through 2030
21. AWS
Thailand Region and more than US\$5 billion planned investment
22. Microsoft
2025 Annual Report
23. Amazon
2025 Annual Report
24. BCG
Asia Pacific leads the world in AI adoption
25. Stanford HAI
2025 AI Index Report
26. ASEAN
Expanded ASEAN Guide on AI Governance and Ethics
27. ASEAN
Responsible AI Roadmap 2025-2030
28. Japan METI
AI Guidelines for Business Version 1.2
29. India MeitY
Digital Personal Data Protection Rules 2025
30. OECD
Competition in artificial intelligence infrastructure
31. TeleGeography
Submarine Cable Map 2024
32. Hong Kong Digital Policy Office
Digital Economy Development in Hong Kong
33. Vietnam Ministry of Information and Communications
Digital economy development strategy
34. OECD
Digital Government Outlook 2026: Korea

Produced by HongKong YIHE

Asia market intelligence, digital commerce, growth strategy and implementation research