

Cloud Architect Exclusive job

Job Information

Recruiter
NEXUS Corporation

Job ID
1617022

Industry
IT Consulting

Company Type
Large Company (more than 300 employees) - International Company

Non-Japanese Ratio
About half Japanese

Job Type
Permanent Full-time

Location
Tokyo - 23 Wards

Salary
Negotiable, based on experience

Refreshed
November 25th, 2026 00:00

General Requirements

Minimum Experience Level
Over 6 years

Career Level
Mid Career

Minimum English Level
Business Level

Minimum Japanese Level
Business Level

Minimum Education Level
Bachelor's Degree

Visa Status
Permission to work in Japan required

Job Description

The Cloud Architect is responsible for designing and delivering scalable, secure, and production-ready cloud solutions that support enterprise applications and AI-driven business initiatives. This role combines cloud architecture, application engineering, platform integration, and technical solution delivery to translate complex business requirements into robust, scalable technical implementations.

The role requires strong technical depth in modern application development, cloud-native systems, and Generative AI implementation, along with the ability to operate effectively in ambiguous and fast-moving delivery environments.

Key Responsibilities:

- Engage with customers and stakeholders to understand business challenges, technical requirements, and operational constraints, translating them into effective cloud-based solution designs.

- Define application and cloud solution architectures that support scalability, reliability, maintainability, and production readiness.
- Design and implement cloud-native applications and backend services using Python, JavaScript, or related technologies.
- Develop and integrate AI-powered solutions, including LLM-based applications, conversational systems, automation workflows, and knowledge-based applications.
- Design and implement Retrieval-Augmented Generation (RAG) pipelines, including document ingestion, embedding strategies, retrieval optimization, and response orchestration.
- Support the design of agent-based architectures, tool integrations, task orchestration, and execution workflows.
- Guide application deployment across cloud platforms such as AWS, Microsoft Azure, and Google Cloud Platform (GCP).
- Implement containerized deployments and support cloud-native development practices, CI/CD pipelines, and infrastructure provisioning.
- Ensure that cloud-based applications meet requirements for scalability, resilience, maintainability, and operational efficiency.
- Identify technical risks, integration challenges, and implementation bottlenecks, developing mitigation strategies to support successful delivery.
- Collaborate with data engineers, architects, product teams, and technical leads to integrate applications and AI workflows into enterprise systems.
- Balance rapid prototyping with production-grade engineering standards, ensuring quality, security, and long-term sustainability.
- Advise customers and internal stakeholders on technical trade-offs, implementation approaches, and solution delivery decisions.

Required Skills

Experience:

- Minimum 8 years of experience in software engineering, technical implementation, or related technical delivery roles
- Proven experience driving projects with direct client engagement and stakeholder management
- Experience working in fast-paced, ambiguous delivery environments with strong ownership and execution capability

Application Engineering:

- Strong hands-on development experience using Python and/or JavaScript
- Experience building production-grade backend services, APIs, and application workflows
- Strong understanding of software engineering fundamentals including modular design, testing, and maintainability
- Experience integrating frontend and backend components for end-to-end solution delivery

Cloud & Platform Engineering:

- Experience building and deploying systems in cloud environments (AWS, GCP, Azure)
- Experience with containerized deployment and cloud-native architecture patterns
- Understanding of deployment automation, CI/CD pipelines, and infrastructure provisioning
- Familiarity with application scalability, resilience, and cloud cost optimization

Customer Delivery & Stakeholder Engagement:

- Strong ability to engage with customers to clarify requirements, align expectations, and drive delivery decisions
- Ability to explain complex technical concepts to business and non-technical stakeholders
- Experience balancing customer priorities with technical feasibility and delivery constraints

Soft Skills:

- Strong problem-solving orientation with a customer-first mindset
- Ability to balance hands-on coding with customer-facing engagement
- Strong decision-making capability under ambiguity
- Ability to maintain delivery speed without compromising quality
- Strong ownership and accountability for delivery outcomes
- Ability to remain composed and effective in high-pressure delivery environments

Nice to Have:

- Business-level or higher Japanese language proficiency
- Experience implementing large-scale enterprise systems
- Experience in data security, governance, and access control design
- Experience working in startup or new business environments
- Experience collaborating with global or distributed teams
- Strong English communication skills
- Experience with observability tools and operational monitoring for AI systems

Language Requirements:

- English: Business proficiency required
- Japanese: Business-level proficiency N1 & N2 preferred

Company Description