Tech Stack & Justification

The technology stack for the full IoT-based mining site monitoring architecture is designed to support the layered system from sensors (child nodes) to cloud applications, using AWS-native services for seamless integration, edge-to-cloud data flow, and extensibility to Digital Twins. It handles multiple sites with parent nodes (data acquisition devices), gateways (LoRaWAN, WiFi, Cellular), edge compute (Greengrass on Raspberry Pi/Jetson), and cloud components (IoT Core, Kinesis, Timestream). The stack prioritizes Python for its use in IoT scripting and ML and availability of stable and feature rich device and cloud SDks from AWS.

The choices are justified based on performance (efficiency in harsh environments), maintainability (managed services, community support), scalability (auto-scaling for thousands of sensors), and cost (pay-per-use with free tiers). Costs are based on 2025 AWS pricing for US East (N. Virginia), with assumptions for moderate scale (1,000 devices, 10 messages/min/device, $500-1,500/month total).

Tech Stack used:

Languages

- **Python 3.x**: Core language for sensor scripts (e.g. synthetic_generator.py), edge Lambda functions, data processing, and ML models.

Frameworks & Libraries

- **Paho MQTT (Python Client)**: For MQTT communication from sensors/gateways to edge/cloud.
- **Boto3 (AWS SDK for Python)**: For AWS service interactions (e.g. IoT Core, S3).

Cloud Services

- **AWS IoT Core**: Central MQTT broker, device management, rules engine.
- **AWS IoT Greengrass**: Edge runtime on Raspberry Pi/Jetson for local processing, buffering, and gateways (LoRaWAN/WiFi/Cellular integration).
- **AWS Lambda**: Serverless compute for data enrichment, rule actions, and ML inference.
- **Amazon Kinesis Data Streams**: Real-time streaming for high-volume telemetry.
- **Amazon SNS**: Alerting via email/SMS/push.
- **Amazon CloudWatch**: Metrics, logs, alarms.
- **AWS KMS**: Key/cert management.
- **Amazon SageMaker**: ML for anomaly detection (e.g. gas spikes).
- **AWS IoT Device Shadow**: Device state synchronization.
- **AWS IoT TwinMaker (Optional)**: Digital Twin modeling.

Databases & Storage

- **Amazon Timestream**: Time-series DB for hot telemetry.
- **Amazon S3**: Object storage for cold data, logs.
- **Amazon Athena**: SQL querying on S3 data.

Visualization Tools

- **Amazon QuickSight**: BI dashboards for analytics.
- **Grafana**: Custom dashboards integrated with CloudWatch/Timestream.

Component	Justification
Python 3.x	**Performance**: Efficient for real-time IoT scripting and ML (e.g. edge anomaly detection). **Maintainability**: Vast libraries (e.g. NumPy for analytics), community support. **Scalability**: Runs on Lambda/Greengrass for auto-scaling. **Cost**: Free/open-source.
Paho MQTT	**Performance**: Lightweight for low-bandwidth mining comms (LoRa/BLE). **Maintainability**: Simple API, integrates with Python. **Scalability**: Handles thousands of connections with QoS 1. **Cost**: Free.
Boto3	**Performance**: Low-latency AWS API calls. **Maintainability**: Official SDK with auto-retries. **Scalability**: Parallel ops for high throughput. **Cost**: Free.
AWS IoT Core	**Performance**: Low-latency MQTT for real-time data. **Maintainability**: Managed rules/shadows. **Scalability**: Millions of devices/messages. **Cost**: $1/1M messages, $0.08/1M connect-minutes; free tier (500K messages, 2.25M connect-minutes).
AWS IoT Greengrass	**Performance**: Local ML/Lambda reduces latency. **Maintainability**: OTA updates. **Scalability**: Groups for site sharding. **Cost**: $0.16/device/month; first 3 devices free/year. Ideal for offline scenarios.
AWS Lambda	**Performance**: Millisecond execution. **Maintainability**: Serverless, versioned. **Scalability**: Auto-scales to millions. **Cost**: $0.20/1M requests, $0.0000166667/GB-second; free tier (1M requests, 400K GB-seconds).
Amazon Kinesis Data Streams	**Performance**: Sub-second streaming. **Maintainability**: Managed shards. **Scalability**: Elastic for TBs/day. **Cost**: $0.015/shard-hour, $0.014/1M PUT units; no free tier mentioned.
Amazon SNS	**Performance**: Instant notifications. **Maintainability**: Topic-based. **Scalability**: Billions of messages. **Cost**: Not detailed; general $0.50/1M publishes, $0.06/100K deliveries; 1M free publishes/month.
Amazon CloudWatch	**Performance**: Real-time ingestion. **Maintainability**: Built-in dashboards. **Scalability**: Petabytes. **Cost**: $0.30/metric (first 10K), $0.50/GB ingested logs; free tier (10 metrics, 5GB logs).
AWS KMS	**Performance**: Hardware-accelerated. **Maintainability**: Auto-rotation.

	Scalability: Unlimited keys. **Cost**: $1/key/month; free for AWS-managed keys.
Amazon SageMaker	**Performance**: GPU training/inference. **Maintainability**: Notebooks/managed endpoints. **Scalability**: Distributed. **Cost**: Not detailed for ml.t3.medium; general $0.0464/hour (ml.t3.medium), free tier (250 hours/month for notebooks).
AWS IoT Device Shadow	**Performance**: Instant state queries. **Maintainability**: JSON-sync. **Scalability**: Per-device. **Cost**: $1.25/1M operations; free tier (225K operations).
AWS IoT TwinMaker	**Performance**: Real-time 3D rendering. **Maintainability**: Scene composer. **Scalability**: Complex sites. **Cost**: $0.05/entity/month, $0.50/10K queries; free tier (50M API calls/12 months). Optional for Twins.
Amazon Timestream	**Performance**: 1,000x faster queries. **Maintainability**: SQL, auto-tiering. **Scalability**: 1M+ writes/sec. **Cost**: Not detailed; free tier (50GB ingest, 100GB magnetic, 750GB-hour memory, 24 TCU-hours queries/1 month).
Amazon S3	**Performance**: High durability. **Maintainability**: Lifecycles. **Scalability**: Unlimited. **Cost**: $0.023/GB-month (Standard), $0.005/1K GET requests; free tier (5GB, 20K GET, 2K PUT). Glacier for savings.
Amazon Athena	**Performance**: Serverless queries. **Maintainability**: SQL on S3. **Scalability**: TBs scanned. **Cost**: $5/TB scanned
Amazon QuickSight	**Performance**: In-memory engine. **Maintainability**: Drag-drop UI. **Scalability**: Thousands users. **Cost**: Author $24/user/month, Reader $3/user/month; free trial (4 users/30 days). Over Grafana for AWS integration.
Grafana	**Performance**: Custom visualizations. **Maintainability**: Plugins. **Scalability**: Integrates with CloudWatch. **Cost**: Open-source free; managed $0.30/metric. Complements QuickSight for ops dashboards.

This stack is cost-effective system ($500-1,500/month for scale), with AWS services reducing latency in mining environments. Equivalents (e.g. Azure IoT) were considered but AWS chosen for unified ecosystem.