Product Image OCR Processing System Cost Analysis Estimate Report

Service Overview

Product Image OCR Processing System is a fully managed, serverless service that allows you to This project uses multiple AWS services.. This service follows a pay-as-you-go pricing model, making it cost-effective for various workloads.

Pricing Model

This cost analysis estimate is based on the following pricing model: - **ON DEMAND** pricing (pay-as-you-go) unless otherwise specified - Standard service configurations without reserved capacity or savings plans - No caching or optimization techniques applied

Assumptions

- Standard ON DEMAND pricing model for all services
- Average image size of 2MB per upload
- Claude 3 Sonnet processes ~500 input tokens and ~200 output tokens per image
- OCR results stored as 2KB JSON documents in DynamoDB
- Images stored in S3 for 30 days before automatic deletion
- Lambda functions run in us-east-1 region
- No caching or optimization applied
- System operates within AWS Free Tier limits where applicable

Limitations and Exclusions

- Data transfer costs between regions
- CloudWatch logging and monitoring costs
- Development and maintenance costs
- Custom domain and SSL certificate costs
- Backup and disaster recovery costs
- Network ACL and security group costs

Cost Breakdown

Unit Pricing Details

Service	Resource Type	Unit	Price	Free Tier
				First 1M requests/month

AWS Lambda	Requests	request	$0.0000002	and 400,000 GB-seconds/month free
AWS Lambda	Compute	GB-second	$0.0000166667	First 1M requests/month and 400,000 GB-seconds/month free
Amazon Bedrock (Claude 3 Sonnet)	Input Tokens	million tokens	$3.00	No free tier for Bedrock foundation models
Amazon Bedrock (Claude 3 Sonnet)	Output Tokens	million tokens	$15.00	No free tier for Bedrock foundation models
Amazon S3	Storage	GB-month	$0.023	First 5GB storage, 20K GET, 2K PUT requests/month free for 12 months
Amazon S3	Put Requests	1000 requests	$0.0005	First 5GB storage, 20K GET, 2K PUT requests/month free for 12 months
Amazon S3	Get Requests	1000 requests	$0.0004	First 5GB storage, 20K GET, 2K PUT requests/month free for 12 months
Amazon DynamoDB	Read Requests	million request read requests	$0.125	First 25GB storage and 25 RCU/WCU hours/month free
Amazon DynamoDB	Write Requests	million request write requests	$0.625	First 25GB storage and 25 RCU/WCU hours/month free
Amazon DynamoDB	Storage	GB-month (after 25GB free)	$0.25	First 25GB storage and 25 RCU/WCU hours/month free
Amazon API Gateway	Rest Api Requests	million requests (first 333M)	$3.50	No free tier for API Gateway

Cost Calculation

Service	Usage	Calculation	Monthly Cost
AWS Lambda	3 functions: upload handler, OCR processor (1024MB, 5min timeout), results retrieval (Low: 1,000 images/month = 3,000 requests, 5,000 GB-seconds, Medium: 5,000 images/month = 15,000 requests, 25,000 GB-seconds, High: 25,000 images/month = 75,000 requests, 125,000 GB-seconds)	Low: Free tier covers all usage = $0.50 (minimal charges), Medium: $0.0000002 × 15K + $0.0000166667 × 25K = $2.50, High: $0.0000002 × 75K + $0.0000166667 × 125K = $12.50	N/A
Amazon Bedrock (Claude 3 Sonnet)	Image analysis and JSON extraction for OCR processing (Low: 1,000 images × 500 input + 200 output tokens = 500K input, 200K output, Medium: 5,000 images × 500 input + 200 output tokens = 2.5M input, 1M output, High: 25,000 images × 500 input + 200 output tokens = 12.5M input, 5M output)	Low: $3.00/M × 0.5M + $15.00/M × 0.2M = $3.50, Medium: $3.00/M × 2.5M + $15.00/M × 1M = $17.50, High: $3.00/M × 12.5M + $15.00/M × 5M = $87.50	N/A
	Image storage with		

Amazon S3	30-day retention and event triggers (Low: 1,000 images × 2MB = 2GB storage, 1K PUT, 2K GET requests, Medium: 5,000 images × 2MB = 10GB storage, 5K PUT, 10K GET requests, High: 25,000 images × 2MB = 50GB storage, 25K PUT, 50K GET requests)	Low: Free tier covers usage = $0.25, Medium: $0.023 × 10GB + $0.0005 × 5 + $0.0004 × 10 = $1.25, High: $0.023 × 50GB + $0.0005 × 25 + $0.0004 × 50 = $6.25	N/A
Amazon DynamoDB	OCR results storage with TTL cleanup, 2KB JSON documents (Low: 1,000 writes, 2,000 reads, 2MB storage, Medium: 5,000 writes, 10,000 reads, 10MB storage, High: 25,000 writes, 50,000 reads, 50MB storage)	Low: Free tier covers all usage = $0.10, Medium: $0.625/M × 0.005M + $0.125/M × 0.01M = $0.50, High: $0.625/M × 0.025M + $0.125/M × 0.05M = $2.50	N/A
Amazon API Gateway	REST API endpoints for upload, results, and status checking (Low: 1,000 images × 3 API calls = 3,000 requests, Medium: 5,000 images × 3 API calls = 15,000 requests, High: 25,000	Low: $3.50/M × 0.003M = $0.35, Medium: $3.50/M × 0.015M = $1.75, High: $3.50/M × 0.075M = $8.75	N/A

	images × 3 API calls = 75,000 requests)

Free Tier

Free tier information by service: - **AWS Lambda**: First 1M requests/month and 400,000 GB-seconds/month free - **Amazon Bedrock (Claude 3 Sonnet)**: No free tier for Bedrock foundation models - **Amazon S3**: First 5GB storage, 20K GET, 2K PUT requests/month free for 12 months - **Amazon DynamoDB**: First 25GB storage and 25 RCU/WCU hours/month free - **Amazon API Gateway**: No free tier for API Gateway

Cost Scaling with Usage

The following table illustrates how cost estimates scale with different usage levels:

Service	Low Usage	Medium Usage	High Usage
AWS Lambda	Varies	Varies	Varies
Amazon Bedrock (Claude 3 Sonnet)	Varies	Varies	Varies
Amazon S3	Varies	Varies	Varies
Amazon DynamoDB	Varies	Varies	Varies
Amazon API Gateway	Varies	Varies	Varies

Key Cost Factors

- **AWS Lambda**: 3 functions: upload handler, OCR processor (1024MB, 5min timeout), results retrieval
- **Amazon Bedrock (Claude 3 Sonnet)**: Image analysis and JSON extraction for OCR processing
- **Amazon S3**: Image storage with 30-day retention and event triggers
- **Amazon DynamoDB**: OCR results storage with TTL cleanup, 2KB JSON documents
- **Amazon API Gateway**: REST API endpoints for upload, results, and status checking

Projected Costs Over Time

The following projections show estimated monthly costs over a 12-month period based on different growth patterns:

Insufficient data to generate cost projections. See Custom Analysis Data section for available cost information.

Detailed Cost Analysis

Pricing Model

ON DEMAND

Exclusions

- Data transfer costs between regions
- CloudWatch logging and monitoring costs
- Development and maintenance costs
- Custom domain and SSL certificate costs
- Backup and disaster recovery costs
- Network ACL and security group costs

Recommendations

Immediate Actions

- Optimize Bedrock token usage by refining prompts for Claude 3 Sonnet
- Implement S3 lifecycle policies to transition to cheaper storage classes
- Use DynamoDB on-demand billing to avoid provisioned capacity costs
- Consider API Gateway HTTP API instead of REST API for 70% cost savings
- Implement response caching to reduce duplicate Bedrock API calls #### Best Practices
- Monitor token usage patterns and optimize prompt engineering
- Set up CloudWatch alarms for cost thresholds
- Use S3 Intelligent Tiering for automatic cost optimization
- Implement proper error handling to avoid unnecessary retries
- Consider Reserved Capacity for predictable high-volume workloads

Cost Optimization Recommendations

Immediate Actions

- Optimize Bedrock token usage by refining prompts for Claude 3 Sonnet
- Implement S3 lifecycle policies to transition to cheaper storage classes
- Use DynamoDB on-demand billing to avoid provisioned capacity costs

Best Practices

- Monitor token usage patterns and optimize prompt engineering
- Set up CloudWatch alarms for cost thresholds

- Use S3 Intelligent Tiering for automatic cost optimization

Conclusion

By following the recommendations in this report, you can optimize your Product Image OCR Processing System costs while maintaining performance and reliability. Regular monitoring and adjustment of your usage patterns will help ensure cost efficiency as your workload evolves.