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

- Average image size: 2-5MB with 30-day retention
- Claude 3.5 Haiku model for OCR processing
- Token usage: 2,000 input + 500 output tokens per image
- Lambda functions: 512MB-1024MB memory allocation
- 5-7 API Gateway calls per image processing workflow
- 2-3 DynamoDB write operations and 3-5 read operations per image
- Standard ON DEMAND pricing without reserved capacity

Limitations and Exclusions

- Data transfer costs between AWS services (minimal within same region)
- CloudWatch logging and monitoring costs
- Development and testing environment costs
- Custom domain and SSL certificate costs for API Gateway
- Backup and disaster recovery costs
- Third-party monitoring and alerting tools

Cost Breakdown

Unit Pricing Details

Service	Resource Type	Unit	Price	Free Tier
Amazon Bedrock Foundation Models (Claude 3.5 Haiku)	Input Tokens	1,000,000 tokens	$0.30	No free tier available for Bedrock foundation models
Amazon Bedrock Foundation Models (Claude 3.5 Haiku)	Output Tokens	1,000,000 tokens	$1.50	No free tier available for Bedrock foundation models
AWS Lambda	Requests	1,000,000 requests	$0.20	Always free: 1M requests/month + 400,000 GB-seconds/month
AWS Lambda	Compute Tier1	GB-second (0-6B GB-seconds)	$0.0000166667	Always free: 1M requests/month + 400,000 GB-seconds/month
AWS Lambda	Compute Tier2	GB-second (6B-15B GB-seconds)	$0.0000150000	Always free: 1M requests/month + 400,000 GB-seconds/month
Amazon API Gateway	Rest Api Tier1	1,000,000 requests (first 333M)	$3.50	First 12 months: 1M REST API calls free per month
Amazon API Gateway	Rest Api Tier2	1,000,000 requests (next 667M)	$2.80	First 12 months: 1M REST API calls free per month
Amazon DynamoDB	Read Requests	1,000,000 RRUs	$0.125	Always free: 25GB storage + 25 RCU + 25 WCU per month
Amazon DynamoDB	Write Requests	1,000,000 WRUs	$0.625	Always free: 25GB storage + 25 RCU + 25 WCU per month

Service	Resource Type	Unit	Price	Free Tier
Amazon DynamoDB	Storage	GB-month (after 25GB free)	$0.25	Always free: 25GB storage + 25 RCU + 25 WCU per month
Amazon S3	Standard Storage Tier1	GB-month (first 50TB)	$0.023	First 12 months: 5GB standard storage free
Amazon S3	Standard Storage Tier2	GB-month (next 450TB)	$0.022	First 12 months: 5GB standard storage free

Cost Calculation

Service	Usage	Calculation	Monthly Cost
Amazon Bedrock Foundation Models (Claude 3.5 Haiku)	OCR processing with 2,000 input tokens and 500 output tokens per image (Input Tokens Per Image: 2,000 tokens, Output Tokens Per Image: 500 tokens)	Cost per image: ($0.30/1M × 2K) + ($1.50/1M × 0.5K) = $0.0006 + $0.00075 = $0.00135	$0.30 per 1M input tokens, $1.50 per 1M output tokens

Service	Usage	Calculation	Monthly Cost
AWS Lambda	Three functions: upload handler (512MB, 2s), OCR processor (1024MB, 10s), status handler (256MB, 1s) (Requests Per Image: 3 requests (upload, process, status), Compute Per Image: ~0.5 GB-seconds total)	Cost per 1000 images: ($0.20/1M × 3K requests) + ($0.0000166667 × 500 GB-seconds) = $0.0006 + $0.008 = $0.009	$0.20 per 1M requests + $0.0000166667 per GB-second
Amazon API Gateway	REST API with 5-7 requests per image processing workflow (Requests Per Image: 6 requests average (upload, status polling, results))	Cost per 1000 images: $3.50/1M × 6K requests = $0.021	$3.50 per 1M requests (first 333M)

Service	Usage	Calculation	Monthly Cost
Amazon DynamoDB	On-demand billing for job status and results storage (Writes Per Image: 2.5 WRUs average, Reads Per Image: 4 RRUs average, Storage Per 1000 Images: ~2GB)	Cost per 1000 images: ($0.625/1M × 2.5K WRUs) + ($0.125/1M × 4K RRUs) + (2GB × $0.25) = $0.002 + $0.0005 + $0.50 = $0.503	$0.125 per 1M RRUs, $0.625 per 1M WRUs, $0.25 per GB-month storage
Amazon S3	Standard storage for product images with 30-day retention (Storage Per 1000 Images: ~3GB (assuming 3MB average per image))	Cost per 1000 images: 3GB × $0.023 = $0.069	$0.023 per GB-month (first 50TB)
Total	**All services**	**Sum of all calculations**	**$4.15/month**

Free Tier

Free tier information by service: - **Amazon Bedrock Foundation Models (Claude 3.5 Haiku)**: No free tier available for Bedrock foundation models - **AWS Lambda**: Always free: 1M requests/month + 400,000 GB-seconds/month - **Amazon API Gateway**: First 12 months: 1M REST API calls free per month - **Amazon DynamoDB**: Always free: 25GB storage + 25 RCU + 25 WCU per month - **Amazon S3**: First 12 months: 5GB standard storage free

Cost Scaling with Usage

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

Service	Low Usage	Medium Usage	High Usage
Amazon Bedrock Foundation Models (Claude 3.5 Haiku)	$0/month	$0/month	$0/month
AWS Lambda	$0/month	$0/month	$0/month
Amazon API Gateway	$1/month	$3/month	$7/month
Amazon DynamoDB	$0/month	$0/month	$0/month
Amazon S3	$0/month	$0/month	$0/month

Key Cost Factors

- **Amazon Bedrock Foundation Models (Claude 3.5 Haiku)**: OCR processing with 2,000 input tokens and 500 output tokens per image
- **AWS Lambda**: Three functions: upload handler (512MB, 2s), OCR processor (1024MB, 10s), status handler (256MB, 1s)
- **Amazon API Gateway**: REST API with 5-7 requests per image processing workflow
- **Amazon DynamoDB**: On-demand billing for job status and results storage
- **Amazon S3**: Standard storage for product images with 30-day retention

Projected Costs Over Time

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

Base monthly cost calculation:

Service	Monthly Cost
Amazon Bedrock Foundation Models (Claude 3.5 Haiku)	$0.30
AWS Lambda	$0.20
Amazon API Gateway	$3.50
Amazon DynamoDB	$0.12
Amazon S3	$0.02
Total Monthly Cost	**$4**

Growth Pattern	Month 1	Month 3	Month 6	Month 12
Steady	$4/mo	$4/mo	$4/mo	$4/mo
Moderate	$4/mo	$4/mo	$5/mo	$7/mo
Rapid	$4/mo	$5/mo	$6/mo	$11/mo

- Steady: No monthly growth (1.0x)
- Moderate: 5% monthly growth (1.05x)
- Rapid: 10% monthly growth (1.1x)

Detailed Cost Analysis

Pricing Model

ON DEMAND

Exclusions

- Data transfer costs between AWS services (minimal within same region)
- CloudWatch logging and monitoring costs
- Development and testing environment costs
- Custom domain and SSL certificate costs for API Gateway
- Backup and disaster recovery costs
- Third-party monitoring and alerting tools

Recommendations

Immediate Actions

- Use Claude 3.5 Haiku model for optimal cost-performance ratio in OCR tasks
- Implement S3 lifecycle policies to automatically delete images after 30 days
- Right-size Lambda memory allocation based on actual performance metrics
- Enable DynamoDB TTL for automatic cleanup of old job records #### Best Practices
- Monitor token usage patterns and optimize prompts to reduce input/output tokens
- Consider ARM-based Lambda functions for 20% cost savings on compute
- Implement API Gateway caching for frequently accessed status endpoints
- Use batch processing for high-volume scenarios to reduce per-request overhead
- Set up CloudWatch alarms for cost monitoring and budget alerts

Cost Optimization Recommendations

Immediate Actions

- Use Claude 3.5 Haiku model for optimal cost-performance ratio in OCR tasks
- Implement S3 lifecycle policies to automatically delete images after 30 days
- Right-size Lambda memory allocation based on actual performance metrics

Best Practices

- Monitor token usage patterns and optimize prompts to reduce input/output tokens
- Consider ARM-based Lambda functions for 20% cost savings on compute
- Implement API Gateway caching for frequently accessed status endpoints

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.