ContentCraft AI

Infrastructure Cost Estimation

AWS-Powered Architecture with Third-Party Alternatives

Prepared by: - Himanshu Tiwari

Document Version: 1.0

February 2026

Table of Contents

Right-click the TOC and select "Update Field" to refresh page numbers

Executive Summary

This document provides a comprehensive cost estimation for deploying and operating the ContentCraft AI platform, an AI-powered content creation platform with intelligent automation, trend analysis, and multi-platform distribution. The analysis covers both AWS-native services and third-party alternatives to help make informed architectural decisions.

Key Findings

- Monthly infrastructure costs range from $850 (Startup) to $8,500+ (Enterprise) depending on scale
- AWS Bedrock offers significant cost savings (60-80%) compared to OpenAI API for AI workloads
- Serverless architecture (Lambda + Fargate) reduces fixed costs by 40% for variable workloads
- Reserved Instances can reduce compute costs by up to 50% for predictable workloads
- Free tier services (CloudFront, Lambda, S3) provide $200+ monthly value for new deployments

Architecture Overview

System Components

ContentCraft AI is built on a modern microservices architecture with the following key components:

Component	Technology	Function
Backend API	Node.js + Express	REST API, business logic
Frontend	React + TypeScript	User interface
Database	MongoDB	Document storage
Cache	Redis	Session, queue, caching
AI Engine	Gemini/DALL-E	Content generation
Queue	Bull/Redis	Background jobs

Table 1: Core System Components

AWS Service Mapping

The following table maps each system component to its recommended AWS service equivalent:

Component	AWS Service	Notes
Backend API	ECS Fargate / EC2	Auto-scaling containers
Frontend	S3 + CloudFront	Static site hosting
Database	DocumentDB	MongoDB compatible
Cache	ElastiCache	Redis cluster mode
AI Engine	Bedrock / SageMaker	Managed AI services
Queue	SQS / ElastiCache	Message queuing

Table 2: AWS Service Mapping

Detailed Cost Breakdown

This section provides detailed pricing for each AWS service category. All prices are based on the US East (N. Virginia) region and are subject to change. Prices shown are for On-Demand usage unless otherwise specified.

Compute Services

Amazon EC2

Instance Type	vCPU/RAM	On-Demand	Reserved (1yr)
t3.medium	2 vCPU / 4 GB	$0.0416/hr	$0.0243/hr
t3.large	2 vCPU / 8 GB	$0.0832/hr	$0.0486/hr
m6g.large	2 vCPU / 8 GB	$0.077/hr	$0.046/hr
c6g.xlarge	4 vCPU / 8 GB	$0.136/hr	$0.081/hr

AWS Fargate (ECS)

Resource	Linux/x86	Linux/ARM (Graviton)
vCPU per hour	$0.04048	$0.03238
GB per hour	$0.004445	$0.00356
Spot vCPU	$0.01244 (70% off)	$0.00995 (70% off)

AWS Lambda

Metric	Pricing
Requests	1M free/month, then $0.20 per 1M
Duration (128MB)	400K GB-sec free, then $0.0000167/GB-sec
Duration (512MB)	400K GB-sec free, then $0.0000667/GB-sec
Duration (1GB)	400K GB-sec free, then $0.0000167/GB-sec

Database Services

Amazon DocumentDB (MongoDB Compatible)

Instance Type	Specs	On-Demand	Reserved
db.t4g.medium	2 vCPU / 4 GB	$0.075/hr	$0.045/hr
db.r6g.large	2 vCPU / 16 GB	$0.219/hr	$0.131/hr
db.r6g.xlarge	4 vCPU / 32 GB	$0.437/hr	$0.262/hr
Storagez	Per GB	$0.10/GB-month	-
I/O	Per million	$0.20/million	-

Amazon ElastiCache (Redis)

Node Type	Specs	On-Demand	Reserved
cache.t4g.micro	2 vCPU / 0.5 GB	$0.012/hr	$0.007/hr
cache.r6g.large	2 vCPU / 13 GB	$0.219/hr	$0.131/hr
cache.r6g.xlarge	4 vCPU / 26 GB	$0.437/hr	$0.262/hr
cache.r7g.large	2 vCPU / 16 GB	$0.233/hr	$0.140/hr

AI/ML Services

Amazon Bedrock

Amazon Bedrock provides access to foundation models including Anthropic Claude, Meta Llama, and Amazon Titan. Pricing is per-token based with separate rates for input and output.

Model	Input (per 1K)	Output (per 1K)	Use Case
Claude 3.5 Sonnet	$0.003	$0.015	General purpose
Claude 3 Haiku	$0.00025	$0.00125	Fast, cost-effective
Llama 3.2 90B	$0.002	$0.002	Open source
Titan Text G1	$0.0008	$0.0016	AWS native
Titan Image Gen	$0.08/image	-	Image generation

Amazon SageMaker

For custom model training and hosting, Amazon SageMaker offers flexible pricing based on instance usage.

Instance Type	Specs	Price/Hour
ml.t3.medium	2 vCPU / 4 GB	$0.058/hr
ml.m5.large	2 vCPU / 8 GB	$0.115/hr
ml.g4dn.xlarge	4 vCPU / 16 GB + T4	$0.736/hr
ml.p3.2xlarge	8 vCPU / 61 GB + V100	$3.825/hr

Storage & CDN

Amazon S3

Storage Class	Storage	Requests
Standard	$0.023/GB-month	$0.005 per 1K PUT
Intelligent-Tiering	$0.023/GB-month	$0.005 per 1K PUT
Standard-IA	$0.0125/GB-month	$0.01 per 1K PUT
Glacier	$0.004/GB-month	$0.03 per 1K PUT
Data Transfer Out	First 100GB free, then $0.09/GB	-

Amazon CloudFront

Metric	Price	Free Tier
Data Transfer Out	$0.085/GB (US/EU)	1 TB/month
HTTP Requests	$0.0075 per 10K	10M/month
HTTPS Requests	$0.01 per 10K	10M/month
CloudFront Functions	$0.10 per 1M	2M/month

Networking & Security

Service	Component	Price
Application Load Balancer	Per hour	$0.0225/hr
ALB LCU	Per LCU-hour	$0.008/LCU-hr
Route 53	Hosted Zone	$0.50/zone-month
Route 53	Queries	$0.40 per 1M
Secrets Manager	Per secret	$0.40/secret-month
CloudWatch Logs	Ingestion	$0.50/GB
API Gateway HTTP	Per 1M requests	$1.00
API Gateway REST	Per 1M requests	$3.50

Cost Comparison Analysis

AWS vs Third-Party Services

The following comparison shows the cost differences between AWS-native services and popular third-party alternatives. AWS services often provide better integration, security, and cost predictability.

Service	AWS	Third-Party	Savings
AI Text Gen	Bedrock: $0.003/1K	OpenAI: $0.01/1K	70%
AI Image Gen	Titan: $0.08/img	DALL-E: $0.04-0.12/img	33%
Database	DocumentDB: $0.075/hr	MongoDB Atlas: $0.12/hr	37%
Cache	ElastiCache: $0.012/hr	Redis Cloud: $0.018/hr	33%
Email	SES: $0.10/1K	SendGrid: $0.50/1K	80%

Table 3: AWS vs Third-Party Cost Comparison

AI Service Cost Analysis

For AI content generation, AWS Bedrock offers significant cost advantages over OpenAI and Google Gemini:

Provider	Model	Input	Output
AWS Bedrock	Claude 3.5 Sonnet	$3.00	$15.00
AWS Bedrock	Claude 3 Haiku	$0.25	$1.25
AWS Bedrock	Llama 3.2 90B	$2.00	$2.00
OpenAI	GPT-4	$30.00	$60.00
OpenAI	GPT-3.5 Turbo	$0.50	$1.50
Google	Gemini 2.5 Pro	$1.25-2.50	$10.00-15.00
Google	Gemini 2.5 Flash	$0.15	$0.60-3.50

Table 4: AI Service Pricing Comparison (per 1M tokens)

Three-Tier Cost Estimates

Based on different usage patterns and business stages, we have prepared three cost estimation scenarios:

Tier 1: Startup (Up to 1,000 Users)

Service	Configuration	Monthly Cost
ECS Fargate	1 vCPU, 2GB (2 tasks)	$60
DocumentDB	db.t4g.medium	$55
ElastiCache	cache.t4g.micro	$9
S3	100GB Standard	$2.30
CloudFront	500GB transfer	$35
ALB	1 load balancer	$16
Bedrock	500K input, 200K output	$200
Lambda	2M requests	$0 (free tier)
CloudWatch	10GB logs	$5
Route 53	1 hosted zone	$0.50
Secrets Manager	10 secrets	$4
API Gateway	HTTP API	$5
SES	10K emails	$1

Total Monthly Cost: $850 - $1,200

Tier 2: Growth (1,000 - 10,000 Users)

Service	Configuration	Monthly Cost
ECS Fargate	2 vCPU, 4GB (4 tasks)	$240

Service	Configuration	Monthly Cost
DocumentDB	db.r6g.large	$160
ElastiCache	cache.r6g.large	$160
S3	500GB Standard	$11.50
CloudFront	2TB transfer	$160
ALB	1 load balancer + LCU	$35
Bedrock	5M input, 2M output	$1,500
Lambda	10M requests	$1.80
CloudWatch	50GB logs	$25
Route 53	1 hosted zone + queries	$5
Secrets Manager	25 secrets	$10
API Gateway	HTTP API	$25
SES	100K emails	$10
OpenSearch	t3.small (dev)	$25

Total Monthly Cost: $2,800 - $4,500

Tier 3: Enterprise (10,000+ Users)

Service	Configuration	Monthly Cost
ECS Fargate	4 vCPU, 8GB (6 tasks)	$720
DocumentDB	db.r6g.xlarge cluster	$650
ElastiCache	cache.r6g.xlarge cluster	$650
S3	2TB Standard + IA	$35
CloudFront	10TB transfer	$800
ALB	2 load balancers	$50

Service	Configuration	Monthly Cost
Bedrock	20M input, 8M output	$4,000
Lambda	50M requests	$9
CloudWatch	200GB logs	$100
Route 53	1 hosted zone + queries	$15
Secrets Manager	50 secrets	$20
API Gateway	HTTP API	$100
SES	500K emails	$50
OpenSearch	m6g.large cluster	$350
WAF	1 Web ACL	$25

Total Monthly Cost: $6,500 - $8,500+

Cost Optimization Recommendations

Immediate Actions (0-30 Days)

0. **Leverage AWS Free Tier**: Use 750 hours/month of EC2 t2/t3.micro, 1M Lambda requests, 5GB S3 storage
1. **Enable Cost Explorer**: Set up billing alerts at 80% of expected monthly spend
2. **Use Spot Instances**: For non-critical workloads, Spot instances offer up to 90% savings

Short-Term Optimizations (1-3 Months)

3. **Purchase Reserved Instances**: 1-year upfront payment yields 30-40% savings on baseline compute
4. **Implement Auto Scaling**: Scale compute resources based on actual demand
5. **Optimize S3 Storage Classes**: Move infrequently accessed data to S3 Infrequent Access

Long-Term Strategies (3+ Months)

6. **Savings Plans**: Compute Savings Plans offer up to 72% discount with flexible commitment
7. **Graviton2/3 Processors**: ARM-based instances provide 20-40% better price-performance
8. **Serverless-First Architecture**: Migrate to Lambda and Fargate for variable workloads

Conclusion

The ContentCraft AI platform can be deployed on AWS with predictable monthly costs ranging from $850 for a startup configuration to $8,500+ for enterprise-scale deployments. Key recommendations include:

- Start with AWS Bedrock for AI workloads to achieve 60-80% cost savings over OpenAI
- Use serverless services (Lambda, Fargate) to minimize fixed costs during growth phases
- Implement Reserved Instances once baseline usage is established
- Monitor costs using AWS Cost Explorer and set up billing alerts
- Consider multi-year Savings Plans for predictable, long-term workloads

The AWS ecosystem provides a robust, scalable foundation for ContentCraft AI with extensive security, compliance, and integration capabilities that third-party services cannot match at comparable price points.

ContentCraft AI

AI-Powered Content Creation Platform

Built by: Himanshu Tiwari, Sonu Jhajharia, Vrinda Gupta, Harsh Kumar