**News Alert RAG System Architecture & Pricing**

**System Components**

**1. Data Collection & Storage**

* **Amazon EventBridge**: Scheduled trigger for news API calls
  + Pricing: Free for first 14M events/month
  + Estimated monthly cost: $0
* **AWS Lambda** (News API Fetcher):
  + 128MB memory, average execution 10s
  + 4 runs per hour = 2,880 executions/month
  + Pricing: $0.20/1M requests + $0.0000166667/GB-second
  + Estimated monthly cost: ~$1-2
* **Amazon S3** (Raw News Storage):
  + Assuming 100KB per article, 2,880 articles/month
  + Standard storage: $0.023/GB
  + Estimated monthly cost: ~$1

**2. RAG Processing Pipeline**

* **Amazon RDS** (PostgreSQL for Vector Store):
  + db.t4g.micro instance (2vCPU, 1GB RAM)
  + Pricing: $0.016/hour
  + Estimated monthly cost: ~$12
* **AWS Lambda** (RAG Processor):
  + 1024MB memory, average execution 30s
  + Processing 2,880 articles/month
  + Pricing: $0.20/1M requests + $0.0000166667/GB-second
  + Estimated monthly cost: ~$3-4

**3. Alert System**

* **AWS Lambda** (Alert Generator):
  + 256MB memory, average execution 5s
  + Pricing: Similar to above Lambda calculations
  + Estimated monthly cost: ~$1
* **Amazon SNS** (Notification Service):
  + $0.50/1M notifications
  + Estimated monthly cost: ~$1
* **Twilio** (for WhatsApp integration):
  + Not an AWS service, but typically:
  + ~$0.0085/message
  + Assuming 1000 alerts/month
  + Estimated monthly cost: ~$8.50

**4. Additional Services**

* **AWS CloudWatch** (Monitoring):
  + Basic monitoring included free
  + 5GB logs/month
  + Estimated monthly cost: ~$5
* **API Gateway**:
  + $1.00/million API calls
  + Estimated monthly cost: ~$2

**Total Estimated Monthly Costs**

**Fixed Costs:**

* RDS: $12
* CloudWatch: $5
* API Gateway: $2

**Variable Costs (based on usage):**

* Lambda functions: ~$6-7
* S3 Storage: ~$1
* Bedrock: ~$15
* SNS: ~$1
* Twilio: ~$8.50

**Total Estimated Monthly Cost: $50-60**

**Cost Optimization Tips**

1. **Lambda Optimization**:
   * Tune memory allocation for best price/performance
   * Implement batch processing where possible
2. **Storage Optimization**:
   * Implement lifecycle policies on S3
   * Clean up old articles regularly
3. **RDS Optimization**:
   * Consider using Aurora Serverless for variable workloads
   * Implement proper indexing for vector searches
4. **Scaling Considerations**:
   * Costs will scale primarily with:
     + Number of articles processed
     + Number of alerts sent
     + Storage duration requirements

**Notes**

* Prices based on US East (N. Virginia) region
* Actual costs may vary based on:
  + Exact implementation details
  + Data volume
  + Processing requirements
  + Region selection
* Free tier benefits not included in estimates
* Recommended to set up AWS Budget Alerts