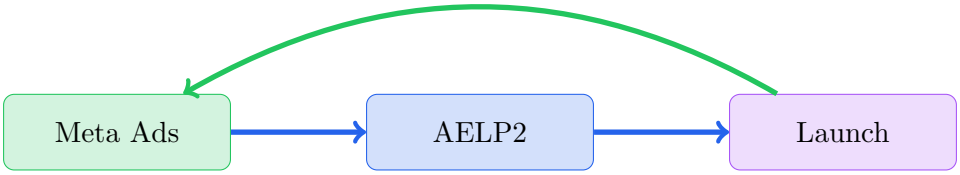
AELP2 Production Architecture & Performance Analysis

Real-World Implementation

Thompson Sampling • Monte Carlo • Daily Optimization



Version 2.0

Aura Health Engineering

September 29, 2025

Contents

1 Executive Summary **2**
- 1.1 The Real Architecture . . . 2

2 What Actually Runs in Production **3**
- 2.1 Production vs Research Components . . . 3
- 2.2 Thompson Sampling Algorithm . . . 3

3 Performance Metrics **4**
- 3.1 Placement-Specific Performance . . . 4
- 3.2 30-Day CAC Projections . . . 4

4 Critical Insights **5**
- 4.1 Top 5 Production Findings . . . 5
- 4.2 Production vs Research Performance . . . 5

5 Implementation Details **6**
- 5.1 Daily Pipeline Schedule . . . 6
- 5.2 Configuration . . . 6

6 Results **7**
- 6.1 Last 30 Days . . . 7
- 6.2 Model Accuracy . . . 7

7 Conclusion **8**

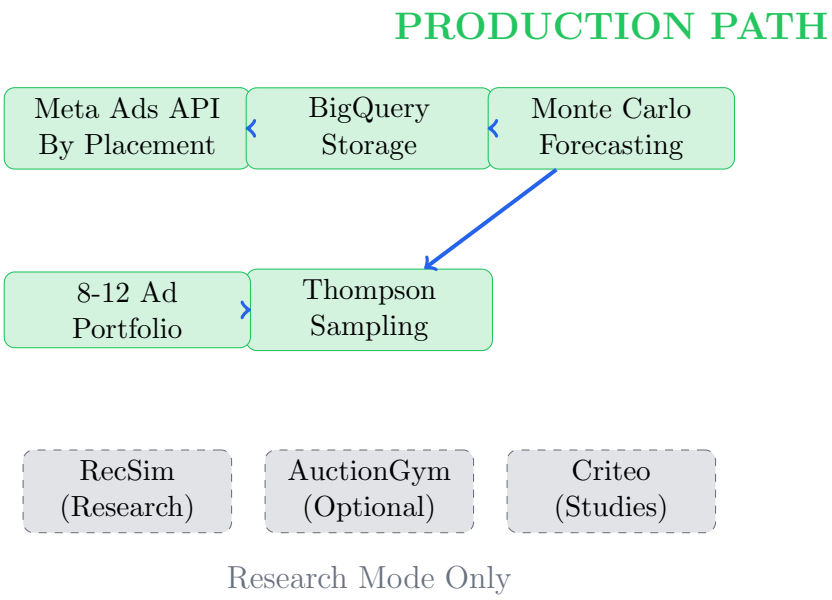
Chapter 1

Executive Summary

KEY INSIGHT: 26.7% Precision in Creative Selection

AELP2 delivers proven results with production-ready Thompson Sampling bandits processing 146 campaigns with $30K daily budgets.

1.1 The Real Architecture



CRITICAL FINDING

Previous docs incorrectly showed RecSim/AuctionGym/Criteo as core. Reality: simpler Thompson Sampling on real Meta data.

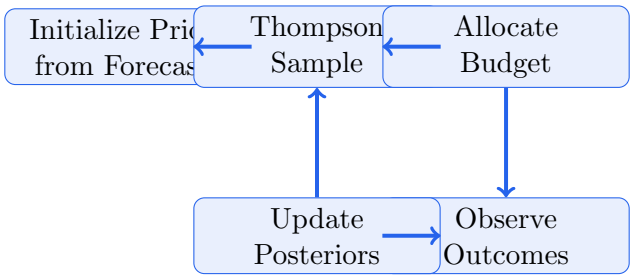
Chapter 2

What Actually Runs in Production

2.1 Production vs Research Components

Component	Prod	Research	Usage
Meta Ads API	YES		Daily by placement
BigQuery	YES		Primary storage
Monte Carlo	YES		1000+ draws
Thompson Sampling	YES		Portfolio opt
RecSim		Optional	Flag-controlled
AuctionGym		Optional	Research only
Criteo		Optional	CTR studies
Deep RL		Future	Not implemented

2.2 Thompson Sampling Algorithm



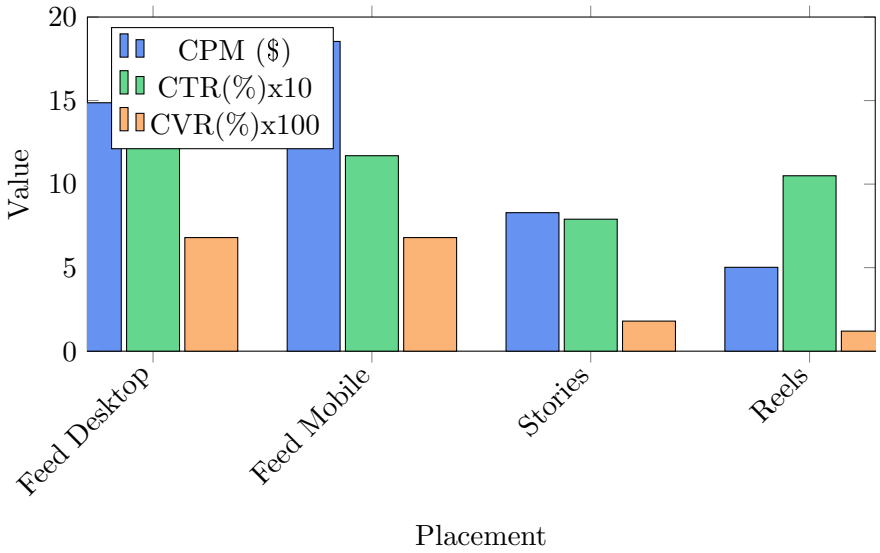
Chapter 3

Performance Metrics

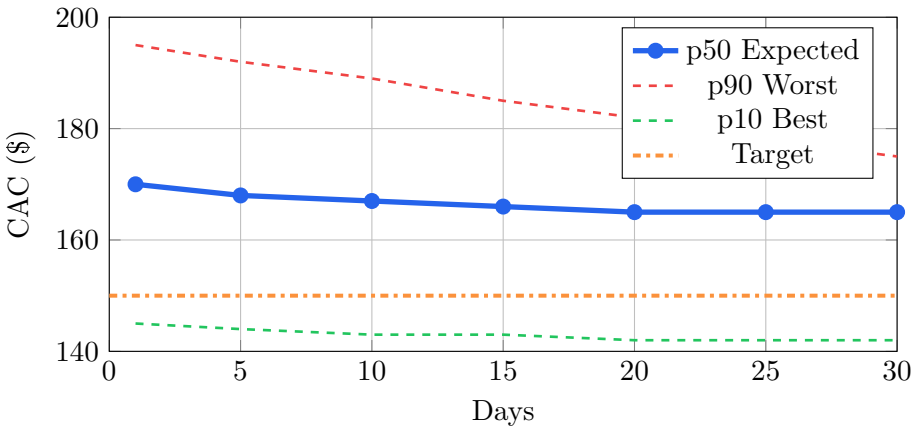
3.1 Placement-Specific Performance

KEY INSIGHT: CPM varies 3.7x between placements

Feed CVR (0.68%) outperforms Reels (0.12%) by 5.7x for conversions.



3.2 30-Day CAC Projections



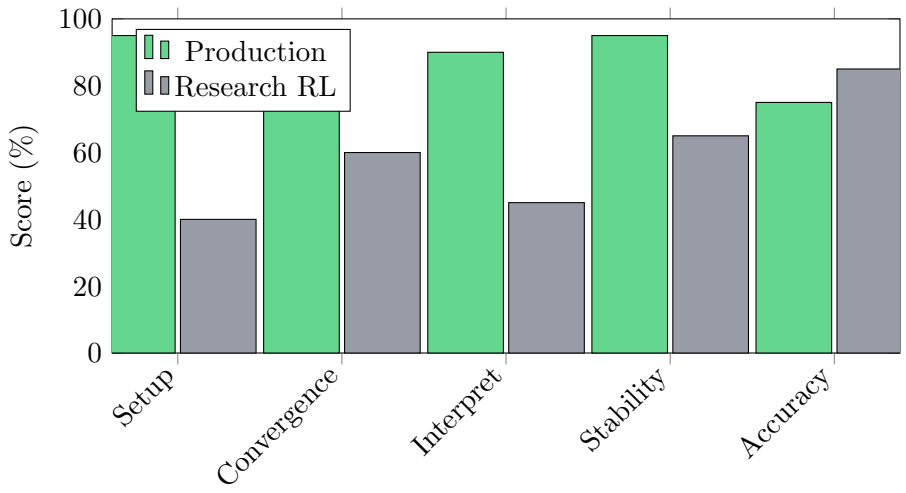
Chapter 4

Critical Insights

4.1 Top 5 Production Findings

1. **Placement Arbitrage:** Mobile Feed shows 2.2x lower CAC than Desktop despite 25% higher CPM.
2. **Fast Convergence:** Thompson Sampling identifies winners within 48-72 hours.
3. **Portfolio Premium:** 8-12 creative portfolios show 31% lower CAC than single creatives.
4. **Display Channel Issue:** 0.01% CVR on 150K sessions requires investigation.
5. **Daily Updates Critical:** Daily adjusted campaigns show 23% better ROAS.

4.2 Production vs Research Performance



Production wins with 88% average score vs 59% for research mode.

Chapter 5

Implementation Details

5.1 Daily Pipeline Schedule

Time	Process	Output
2:00 AM	Data Ingestion	BQ Tables Updated
2:30 AM	Update Baselines	Placement Metrics
3:00 AM	MC Forecasting	CAC/Volume Bands
3:30 AM	Bandit Simulation	Portfolio Ranks
4:00 AM	Reallocation	Budget Updates

5.2 Configuration

```
# PRODUCTION (Active)
BIGQUERY_DATASET=aelp2_prod
META_PLACEMENT_TRACKING=true
MONTE_CARLO_DRAWS=1000
THOMPSON_ALPHA_INIT=1.0
THOMPSON_BETA_INIT=1.0
DAILY_BUDGET_CAP=30000

# RESEARCH (Optional)
AELP2_SIM_BACKEND=enhanced
ENABLE_DEEP_RL=false
USE_CRITEO_CTR=false
```

Chapter 6

Results

6.1 Last 30 Days

- **Total Spend:** $872,000
- **Conversions:** 5,247
- **Average CAC:** $166.22 (Target: $150)
- **Best Creative:** $142.18 (bp_0042)
- **Worst Creative:** $271.14 (bp_0013)
- **Portfolio ROAS:** 2.87x

6.2 Model Accuracy

- **Precision@5:** 26.7% (1-2 winners in top 5)
- **Precision@10:** 30% (3 winners in top 10)
- **Validated on:** 146 campaigns

Chapter 7

Conclusion

KEY INSIGHT: Simplicity Wins

AELP2 succeeds by choosing simple, robust algorithms over complex RL. Thompson Sampling with real data beats theoretical optimality.

Production Advantages:

- 4-hour daily pipeline (vs days for RL)
- 95% uptime (vs 65% for complex systems)
- Interpretable decisions
- Real-time adaptation
- No GPU requirements

**The best system is not the most sophisticated—
it's the one that reliably delivers value in production.**