



# Production Patterns



## Error Handling

Wrap API errors with context. Return meaningful messages Claude can relay.



## Caching

LRU cache with TTL. Reduces API calls, improves response time.



## Bulk Operations

Batch requests when possible. 77 tools in Miro, batch updates save time.



## Authentication

Support env vars for API keys. Never hardcode. Validate on startup.



## Rate Limiting

Token bucket or sliding window. Respect API limits, add backoff.



## Retries

Exponential backoff for transient failures. Don't retry 4xx errors.