Generate a secure policy for API rate limiting. Ensure security best practices.

Here's a sample **API Rate Limiting Policy** that incorporates security best practices. This policy can be tailored to fit the needs of your organization or API platform:

**API Rate Limiting Policy**

**1. Purpose**  
This policy establishes guidelines for controlling the rate at which clients can access the API, ensuring optimal performance, availability, and security of services.

**2. Scope**  
This policy applies to all APIs exposed to external clients, including third-party developers, partners, and internal services interacting through public endpoints.

**3. Policy Statement**

**3.1. Rate Limits**  
Clients are subject to the following rate limits based on their classification:

* **Public Users (Unauthenticated):** 60 requests per minute
* **Registered Users (Authenticated):** 1,000 requests per minute
* **Premium/Partner Clients:** Custom limits based on service agreement (up to 10,000 requests per minute)

**3.2. Burst Control**  
A burst allowance of up to 2x the regular rate may be allowed for short durations (e.g., 1–2 seconds) to accommodate traffic spikes, using token bucket or leaky bucket algorithms.

**3.3. Throttling & Quotas**

* Throttling will return HTTP 429 (Too Many Requests) status code.
* Clients exceeding sustained quota thresholds may be rate limited, temporarily blocked, or have their API key suspended.
* Daily, weekly, or monthly quotas may be applied in addition to per-minute limits.

**3.4. Authentication Requirements**  
Rate limits are tied to:

* API keys
* OAuth tokens
* IP addresses (fallback when authentication is unavailable)

**3.5. Abuse Detection and Mitigation**

* Automated detection of excessive request patterns, scanning, or DDoS behavior.
* Integration with a WAF (Web Application Firewall) and anomaly detection tools.
* Blacklisting and throttling of suspicious IP addresses or tokens.

**3.6. Developer Notification**

* Headers will communicate remaining rate limits:
* X-RateLimit-Limit: 1000
* X-RateLimit-Remaining: 250
* X-RateLimit-Reset: 1631120400
* Clients approaching limits will receive warnings.

**3.7. Exemptions**  
Internal services and trusted automation may request exemptions via written approval and must undergo security evaluation.

**3.8. Logging and Monitoring**

* All requests are logged with metadata for auditing.
* Alerts are triggered when thresholds are frequently exceeded.

**3.9. Review and Revision**  
This policy is reviewed annually or after major API architecture changes.

**4. Enforcement**  
Violations of this policy may result in:

* API key revocation
* IP bans
* Legal or contractual penalties