User Uptime in Practice

Anika Mukherji
Pinterest • SRE
About Me

Anika Mukherji
SRE at Pinterest
Background in iOS and Performance
Responsible for the reliability of pinner facing systems (Core API, Core Web, Traffic Infrastructure)
Importance of availability metrics

- Site **Reliability** Engineering
- Reliability of what? How do we measure it?
- Who is the audience?
- How do we use it?
- Success Rate
Success Rate

- Percentage of successful requests
- Health from the perspective of the server
- Overactive users
- All requests treated equally
- User behavior
  - Different types of errors
  - Retries
  - Switch surfaces
What is User Uptime?

Meaningful Availability - USENIX

User events are probes into the system that assess its health from that user’s perspective.

\[
\text{user-uptime} = \frac{\sum_{u \in \text{users}} \text{uptime}(u)}{\sum_{u \in \text{users}} \text{uptime}(u) + \text{downtime}(u)}.
\]

- **User 1:**
  - T0: Success
  - T1: Failure
  - T2: Success
  - T3: Success
  - Uptime = 2
  - Downtime = 1
  - Uptime %: 2/4 = 50%
  - SR = 2/3 = 66.7%

- **User 2:**
  - T0: Failure
  - T1: Success
  - T2: Success
  - Uptime = 1
  - Downtime = 3
  - Uptime %: 1/4 = 25%
  - SR = 1/3 = 33.3%

- **User 3:**
  - T0: Success
  - T1: Success
  - T2: Success
  - T3: Success
  - Uptime = 4
  - Uptime %: 3/4 = 75%
  - SR = 3/4 = 75%

**General Calculation:**

- Uptime %: 8/12 = 66.7%
- SR = 7/9 = 77.8%
User Uptime MVP - Server Only

- Core Pinner REST API
- What counts as a success and what counts as a failure?
- Breaking down the product (and our API) by criticality

```python
endpoint_classifications=(EndpointCriticality.CRITICAL,)
```

- Degraded product quality
- Imperfect WIP
User Uptime MVP - Technologies

- API -> Kafka -> Xenon Platform/Flink (Stream Processing Platform)
Beyond MVP - Integrating Client Side Metrics

- Server is just one piece of the puzzle
- Mobile crashes
  - App crash logs are delayed until after app restart
- Web rendering - error boundaries
  - Page level error boundaries
- Traffic infrastructure (DNS, CDN, GSLB)
  - Distinguishing between Pinterest and internet infra
  - Third party infrastructure
- Current implementation: offline report available after one day delay
  - Requires manual manipulation when reporting to larger audience to account for:
    - Loss in uptime
    - Error types that are not yet integrated
    - Outages related to misconfigured third party infra
What is next?

- Incident/bug classification guidelines
- Continue client work to better understand UX
- Product criticality contract
- Identifying high ROI projects (actionability)
- Extending user uptime to non-organic (ads/partners) products and unauth experience
- Performance - how slow is too slow?
Summary

- It is difficult to classify what “failed” means for a user
- User uptime has become a critical tool for understanding how outages impact to pinners, *in conjunction with success rate*
- Not yet possible to take manual intervention completely out of calculations
- New class of client side failure types that contribute the majority of downtime
- User Uptime makes it simpler to communicate impact
Thank You

amukherji@pinterest.com