

How to estimate the peak QPS?

Estimating peak Queries Per Second (QPS) is crucial for designing systems that are robust enough to handle maximum traffic loads without degrading performance. Here's a step-by-step guide to help you estimate peak QPS based on known metrics:

1. Understand Your Traffic Pattern

Identify traffic patterns from historical data. Look for patterns of daily, weekly, and seasonal variations. Analyze logs and metrics to understand when the traffic peaks typically occur.

2. Gather Metrics

Collect data regarding:

- **Total Daily Requests:** The total number of requests your system handles in a typical day.
- **Total Monthly Requests** (if available): This can help identify monthly variations and highest traffic days.
- **Hourly Requests during Peak Days:** Data on how requests are distributed over peak days.

3. Calculate Average QPS

Start by calculating the average QPS. This gives a baseline from which to start:

$$\text{Average QPS} = \frac{\text{Total Daily Requests}}{86400 \text{ seconds/day}}$$

4. Estimate Peak QPS Using Multipliers

- **Daily Multiplier:** If you know your traffic increases during certain hours of the day, apply a multiplier to the average QPS. For instance, if traffic doubles during peak hours:

$$\text{Peak QPS} = \text{Average QPS} \times 2$$

- **Event-driven Peaks:** If specific events (e.g., product launches, marketing campaigns) cause spikes, factor these into your estimates. Use past event data to predict increases.

5. Consider Maximum Observed QPS

Look at monitoring tools or logs to find the highest recorded QPS. This empirical method provides a concrete number but relies on having previously encountered high traffic.

6. Adjust for Growth and Buffer

Adjust the estimation to accommodate future growth and to include a buffer to handle unexpected surges:

- **Growth Rate:** If expecting growth, increase the QPS estimate by the growth rate. For example, a 10% expected increase in traffic might mean a 10% higher peak QPS.
- **Buffer:** Adding a buffer (like 20-50% more capacity) ensures the system remains stable during unforeseen spikes.

Example Calculation

Suppose your website receives 50 million requests daily and experiences double the traffic during peak hours (6 PM - 9 PM).

1. Calculate Average QPS:

$$\text{Average QPS} = \frac{50,000,000}{86400} \approx 579$$

2. Estimate Peak QPS:

$$\text{Peak QPS} = 579 \text{ QPS} \times 2 = 1158$$

3. Adjust for Growth and Buffer (let's say a 10% growth and 30% buffer):

$$\text{Adjusted Peak QPS} = 1158 \times 1.1 \times 1.3 \approx 1656$$