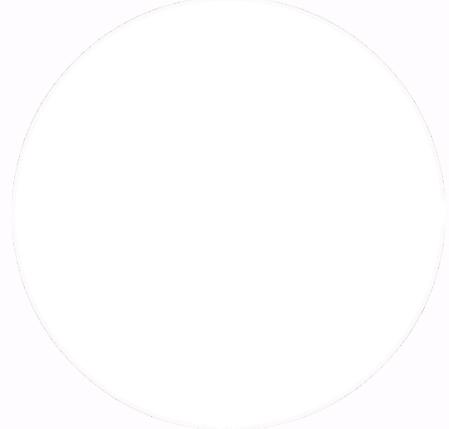


Grow in Software Engineering

Key Performance Metrics



@tauseeffayyaz 



System & Application Performance

Monitors speed, load, and system efficiency.

Reliability & Resilience

Tracks system stability and failure recovery.

Network & Infrastructure

Evaluates connectivity, speed, and reliability.

Business-Centric Metrics

Connects technical metrics to user outcomes.

Development & CI/CD

Optimizes code quality and delivery speed.

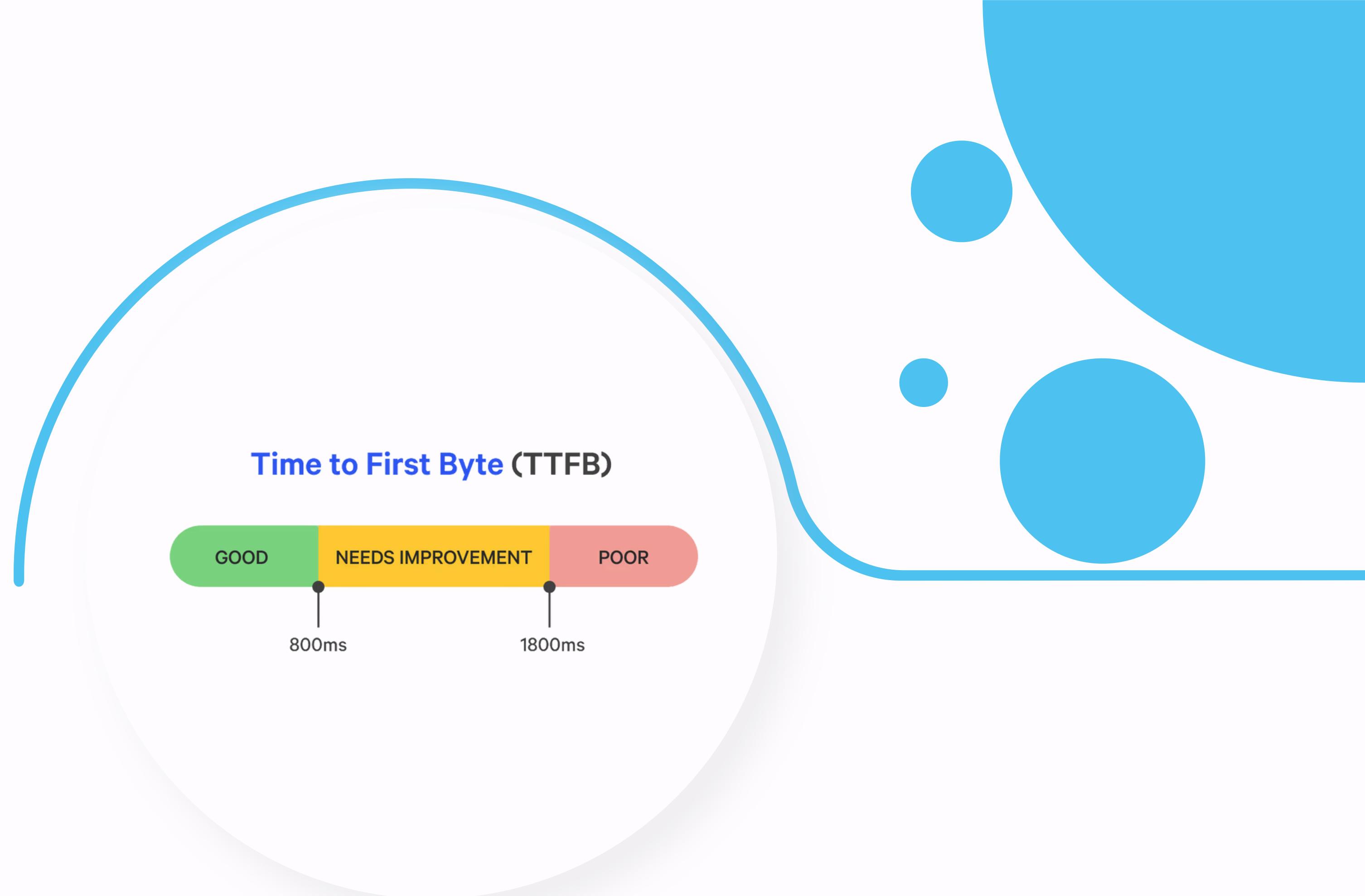
User Experience

Improves speed, interactivity, and usability.



@tauseeffayyaz 





Time to First Byte (TTFB)

Measures how quickly the server responds with the first byte after a request is sent. It reflects server responsiveness and network latency. A low TTFB is crucial for fast page load times.

Measures initial server responsiveness speed.



Throughput

Indicates the number of transactions or requests a system can process per second. It measures overall system capacity and is vital for scaling and performance benchmarking.

Evaluates system's handling capacity/load.



Latency

The delay between sending a request and receiving the first response byte. Lower latency means faster interactions, important for real-time systems and user experience.

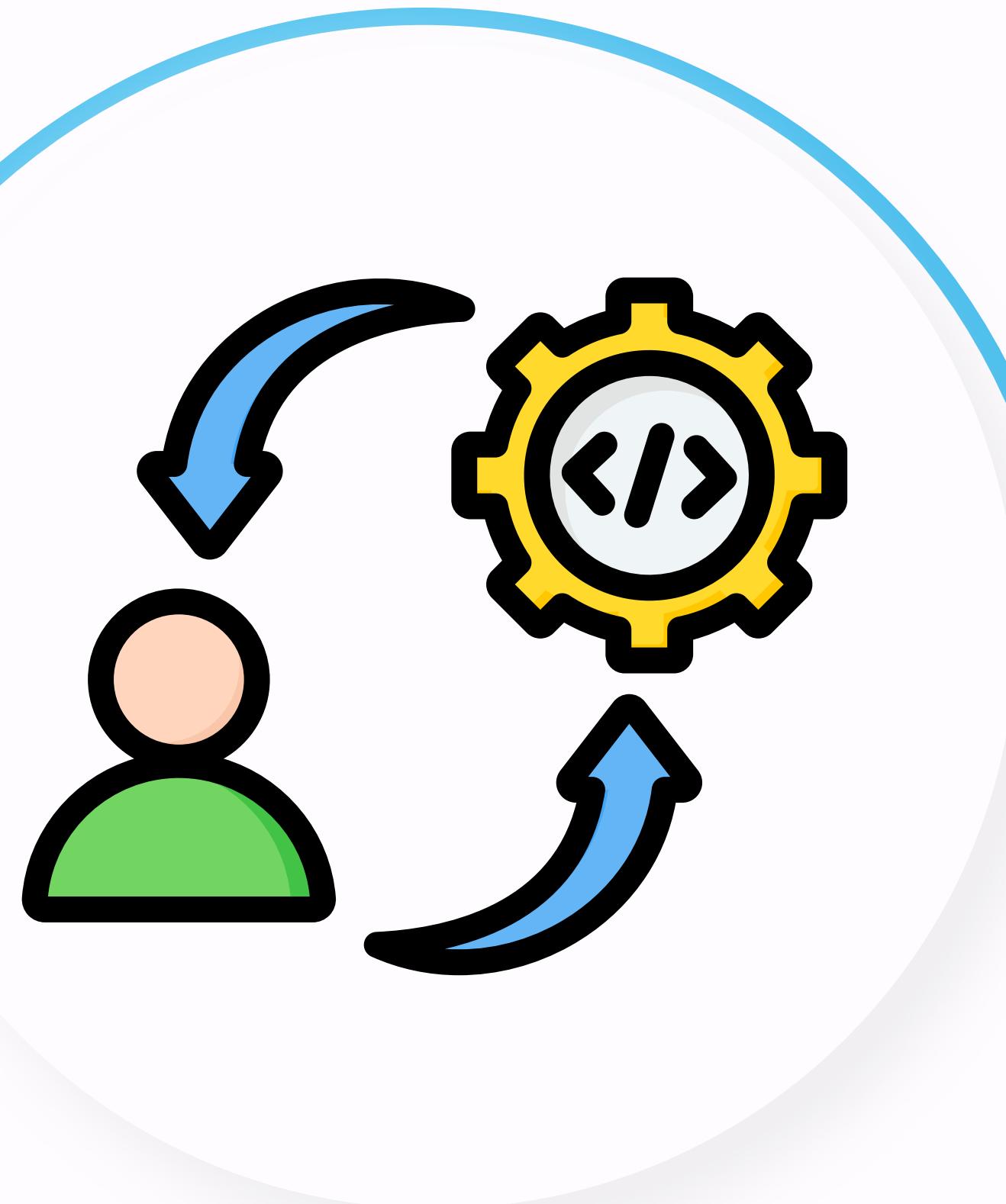
Assesses delay in communication response.



Response Time

The total time taken to receive the full response after a request is made. It includes latency, processing time, and data transfer.

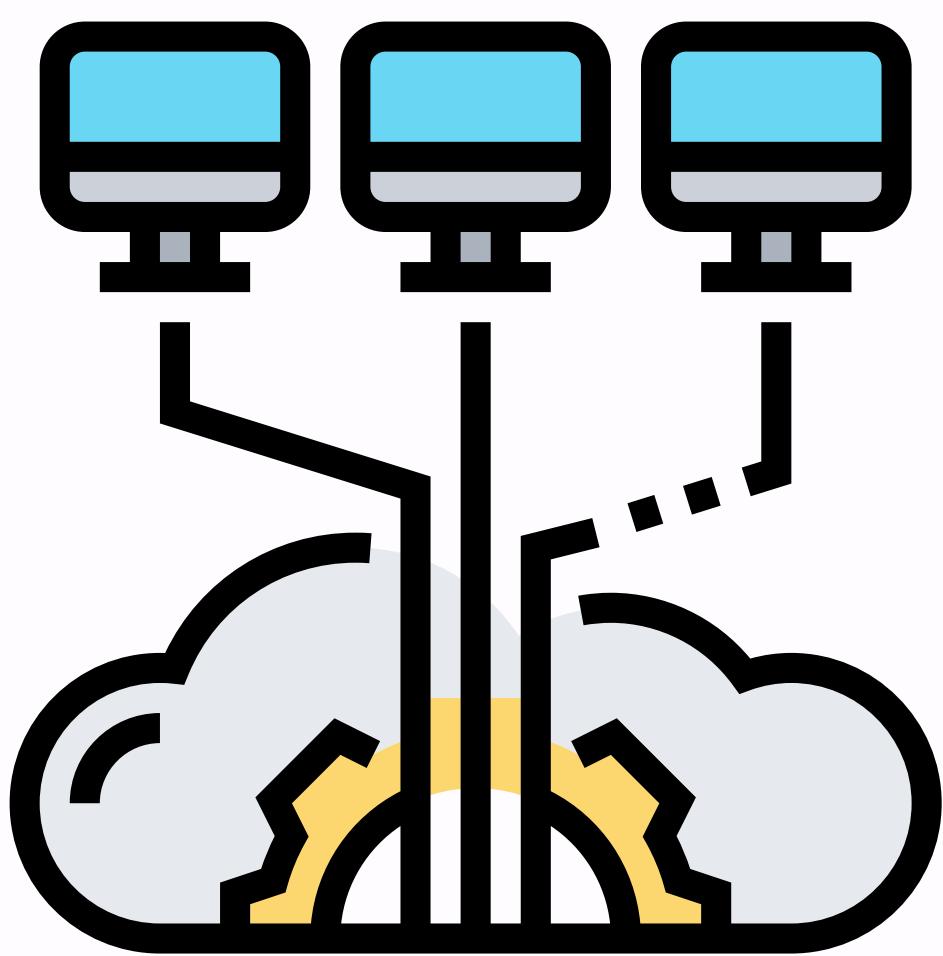
Tracks total request completion duration.



Request Rate (RPS/QPS)

Reflects the number of requests or queries handled per second. Useful for measuring server load and performance under peak traffic.

Monitors server's load handling rate.



Concurrent Connections/Users

Number of users or sessions active simultaneously.
Determines system capacity and is crucial for load testing and scaling.

Measures multi-user handling capability.



System Uptime

Percentage of time the system is operational and available. Higher uptime indicates better availability and reliability.

Indicates system reliability and availability.



Resource Utilization

Tracks usage of CPU, memory, disk, and GPU. Helps identify bottlenecks and optimize resource allocation for better performance.

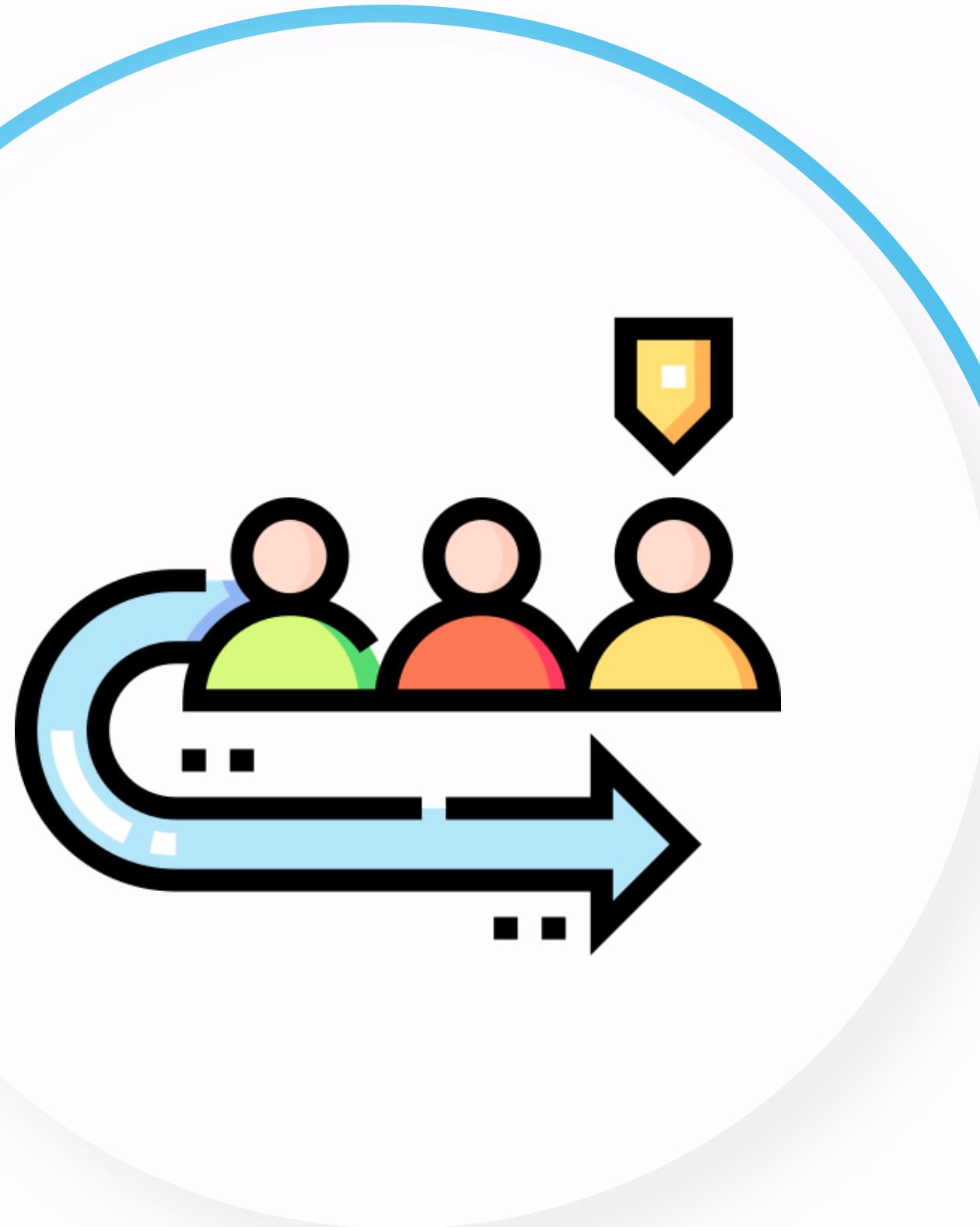
Optimizes CPU, memory, and disk.



Thread/Process Count

The number of active threads or processes in the system. Useful for diagnosing resource leaks, concurrency issues, or thread pool tuning.

Detects process overload or leakage.



Queue Length

Number of requests or tasks waiting in a queue. Long queues can signal performance bottlenecks or under-provisioned services.

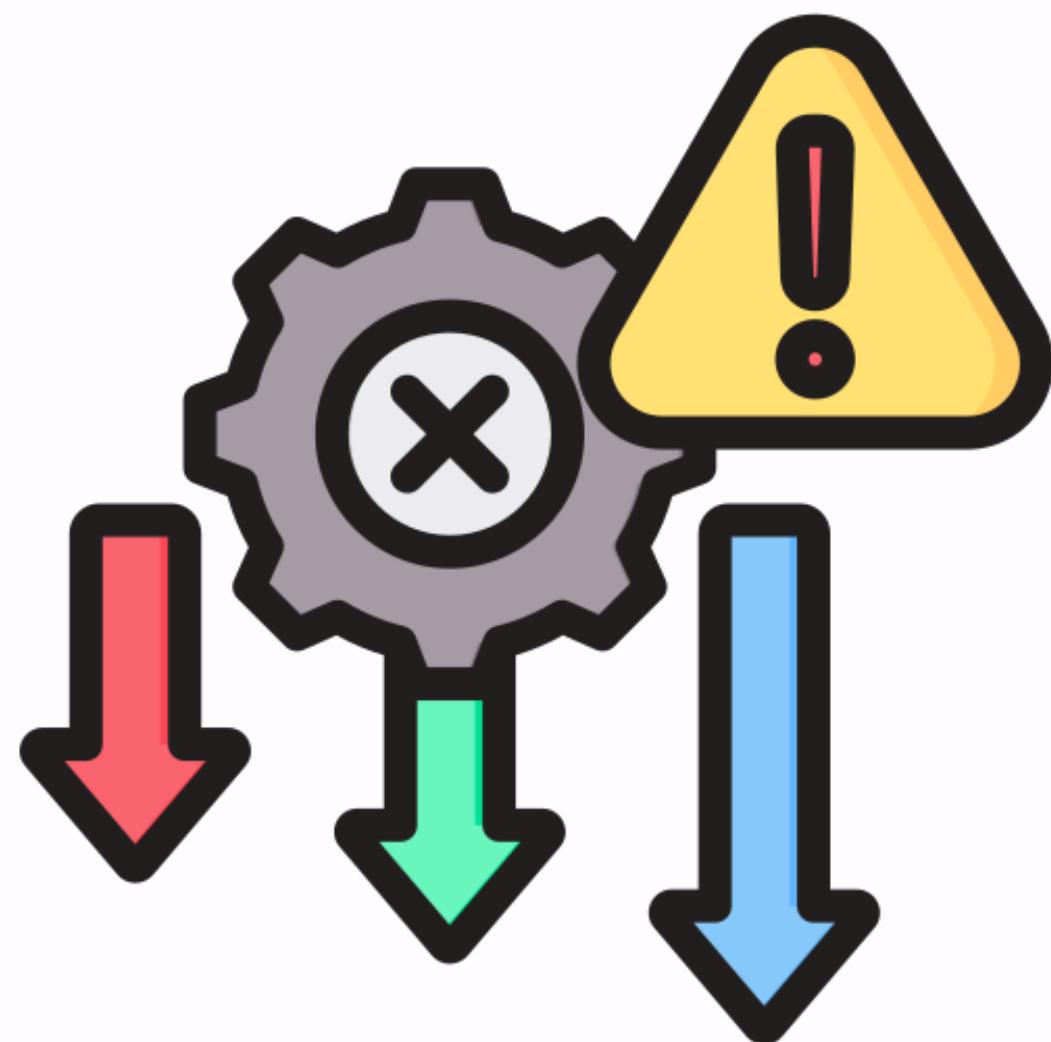
Reveals processing backlog or bottlenecks.



Error Rate

Percentage of failed operations or requests. A high error rate indicates systemic issues affecting reliability.

Tracks frequency of failed operations.



Failure Rate

Measures how often the system crashes or throws unhandled exceptions. Frequent failures imply poor resilience or bugs in code.

Identifies crash or fault occurrences.



Mean Time Between Failures (MTBF)

Average time between two consecutive failures. A higher MTBF signals a more stable and reliable system.

Measures system stability over time.



Mean Time to Repair (MTTR)

Time taken to fix a failure and restore functionality.
Lower MTTR means quicker recovery and minimal downtime.

Evaluates recovery time after failure.



Availability

Ratio of system uptime to total time, often presented in “nines” (e.g., 99.99%). It reflects how dependable the system is.

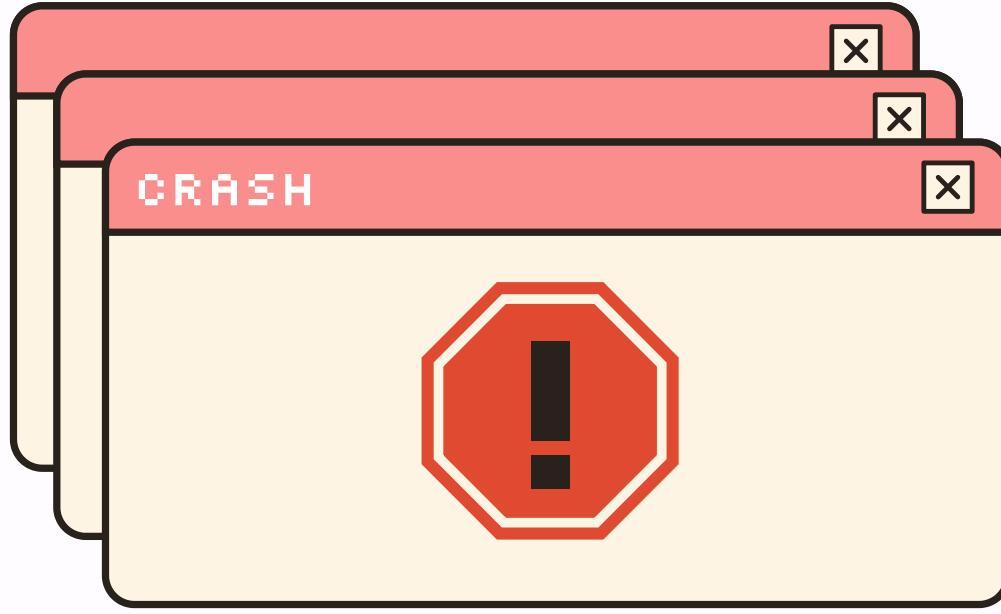
Shows service uptime reliability percentage.



Retry Rate

Indicates how often operations are retried due to temporary failures. High retry rates may mask underlying instability.

Identifies instability or transient failures.



Crash Rate

Number of crashes per session, user, or time period.
Useful for mobile or client-side application monitoring.

Measures crash frequency and severity.



Network Bandwidth

Measures the maximum data transfer rate of a network. Higher bandwidth enables faster data communication.

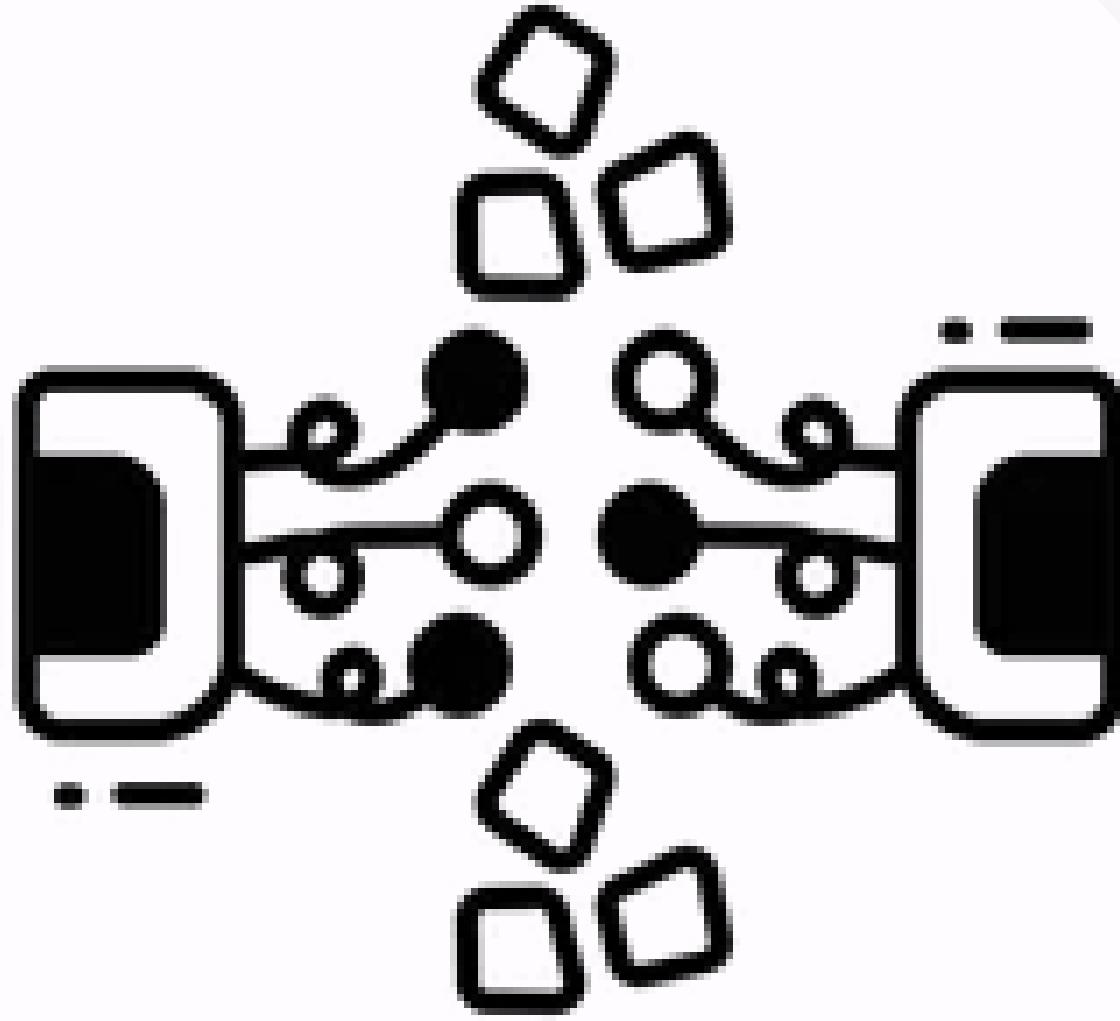
Assesses data transfer speed capacity.



Network Latency

Time taken for data to travel across the network.
Critical for real-time applications like video conferencing or gaming.

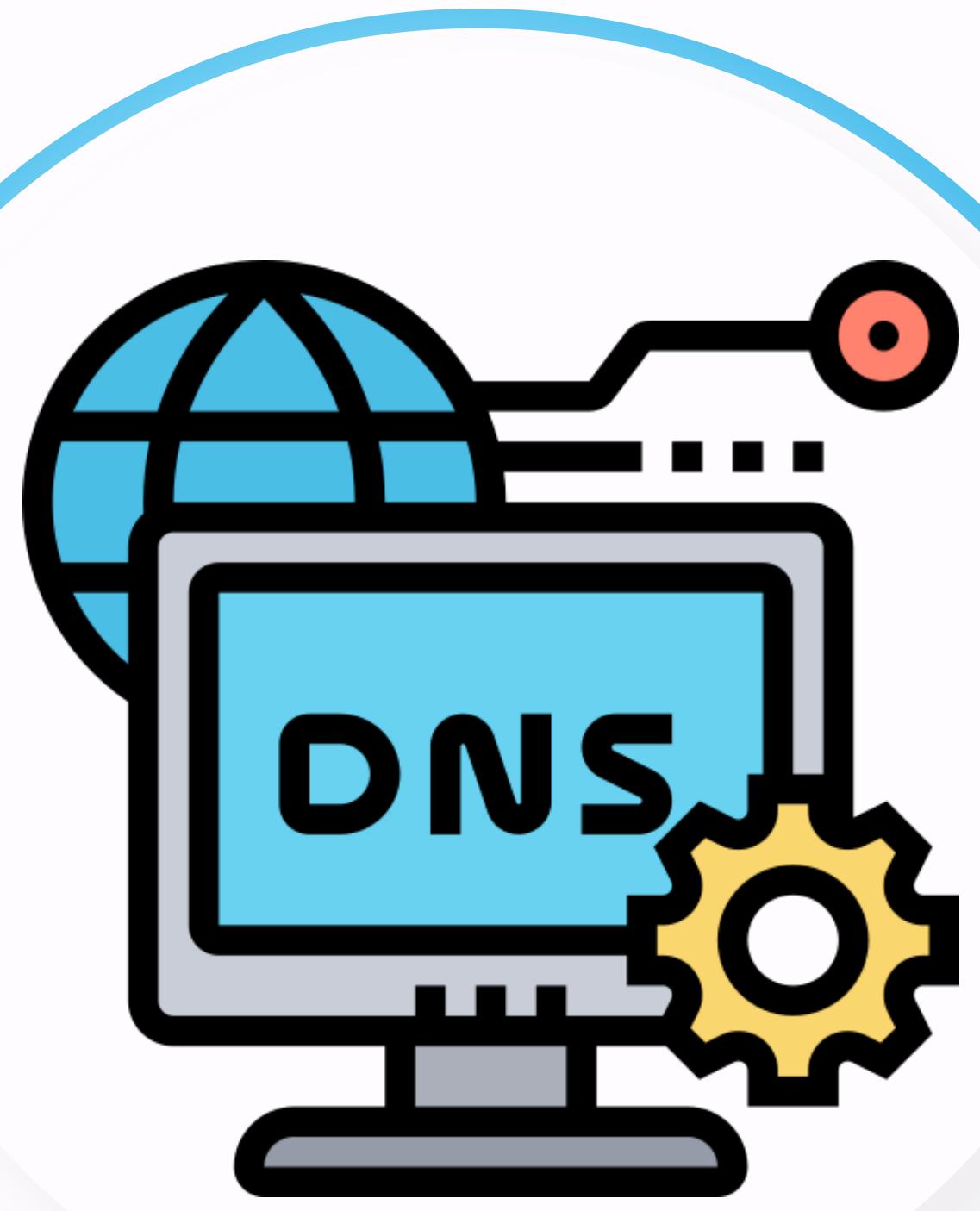
Measures delay in network response.



Packet Loss

Percentage of data packets lost during transmission.
Even small losses can degrade performance or cause data corruption.

Identifies data transmission reliability issues.



DNS Resolution Time

Time taken to resolve a domain name to an IP address. Slow DNS can delay page loads or cause timeouts.

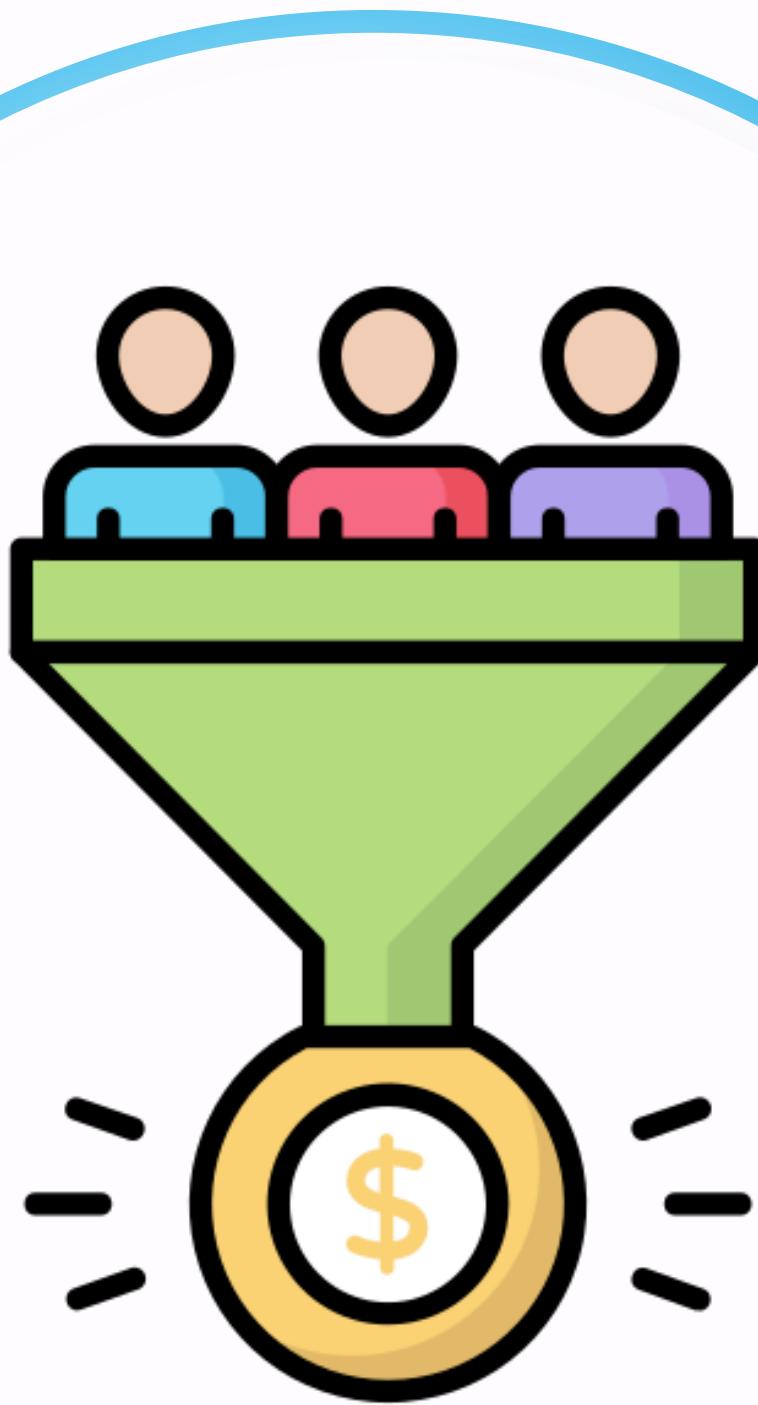
Tracks speed of domain resolution.



Load Average

The average system load over 1, 5, and 15 minutes.
Helps assess whether a system is under stress or performing normally.

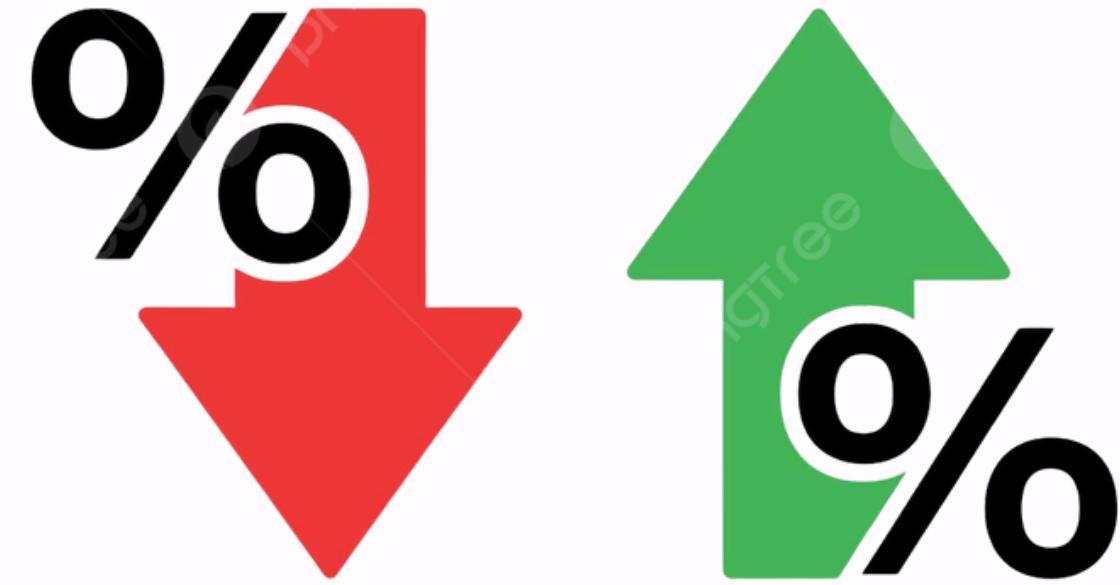
Monitors system workload over time.



Conversion Rate

Percentage of users who complete a goal action like signing up or making a purchase. A key metric for business success.

Measures effectiveness of user actions.



Drop-Off Rate

Points where users abandon a process or funnel. Helps identify UX or functional issues causing user loss.

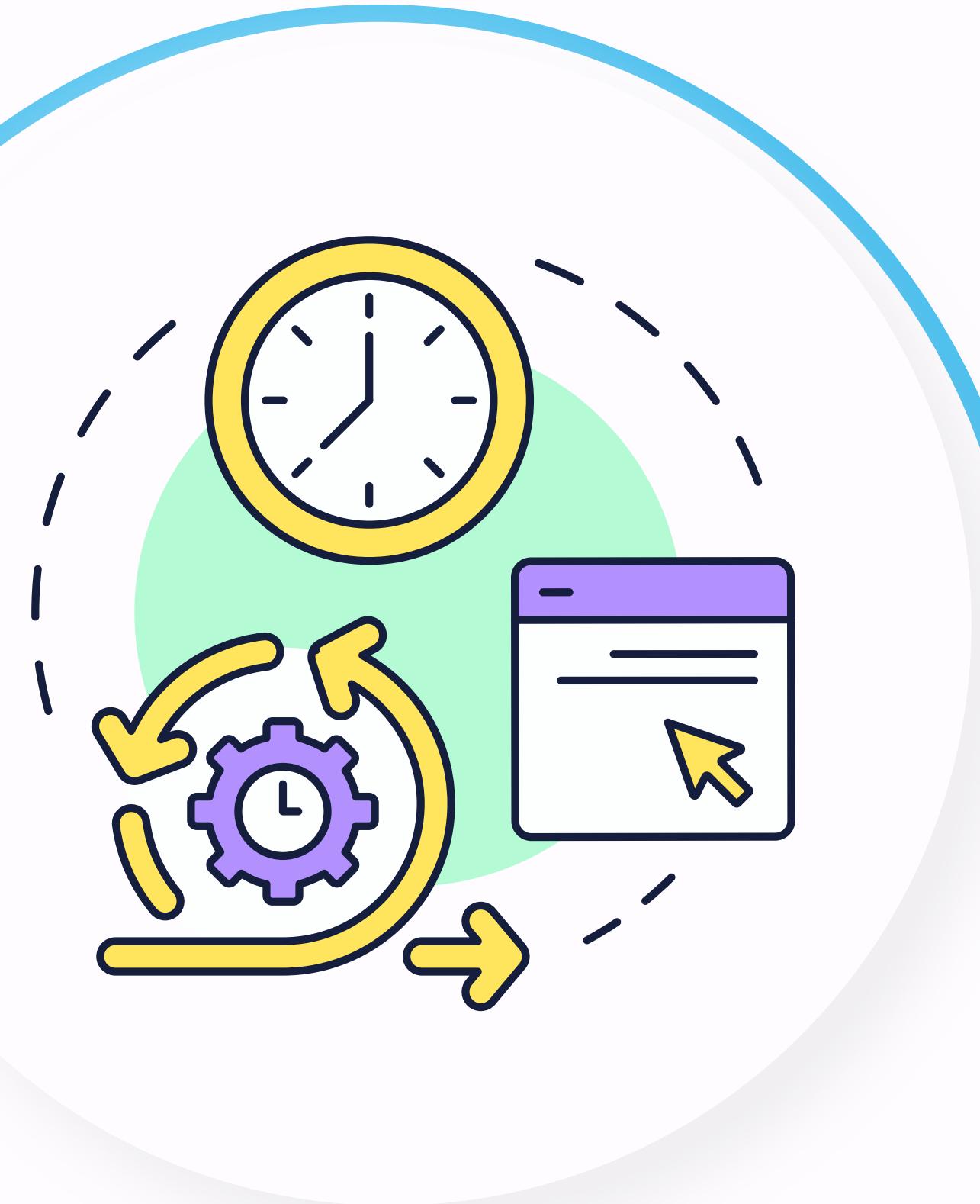
Pinpoints user disengagement or abandonment.



Adoption Rate

Speed at which users begin using a new feature or product. Indicates success of rollouts and user engagement.

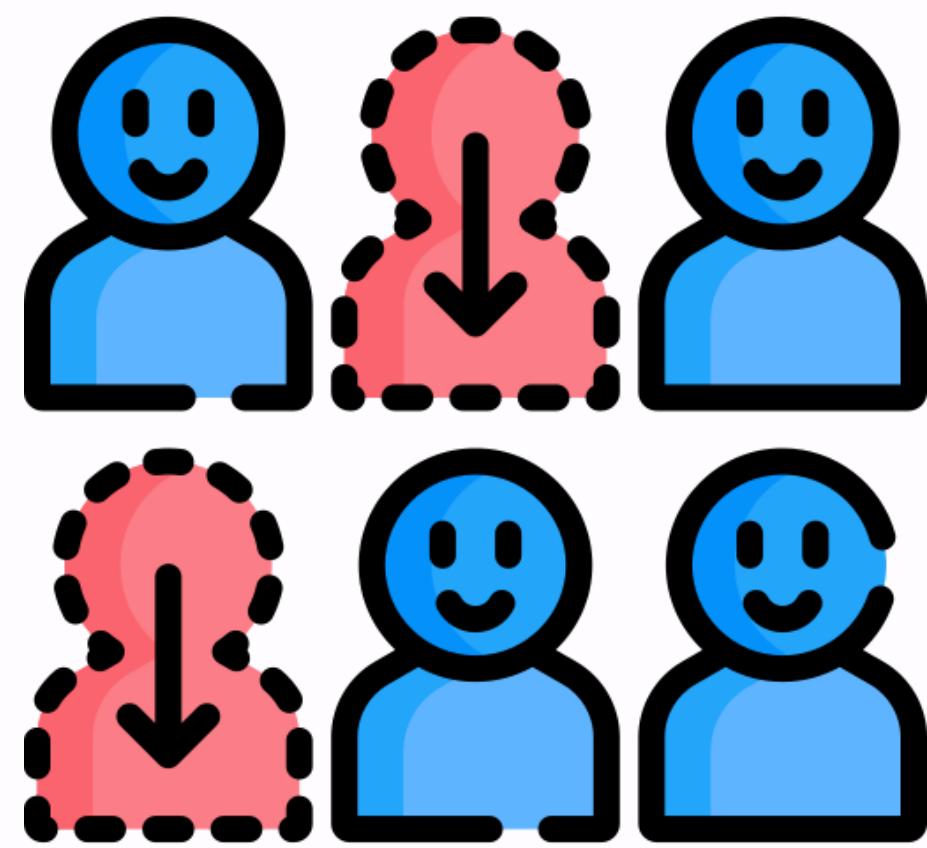
Tracks new feature user uptake.



Session Duration

Average time a user spends during one session. Longer durations can signal engagement, while short ones may hint at issues.

Gauges user engagement per visit.



Churn Rate

Percentage of users who stop using the system over time. Lower churn suggests better retention and satisfaction.

Indicates user retention performance.



Customer Satisfaction (CSAT)

Measures user happiness via surveys or ratings. Higher scores reflect better product/service experience.

Reflects user happiness and feedback.



Net Promoter Score (NPS)

Indicates how likely users are to recommend your product to others. A high NPS reflects strong brand loyalty.

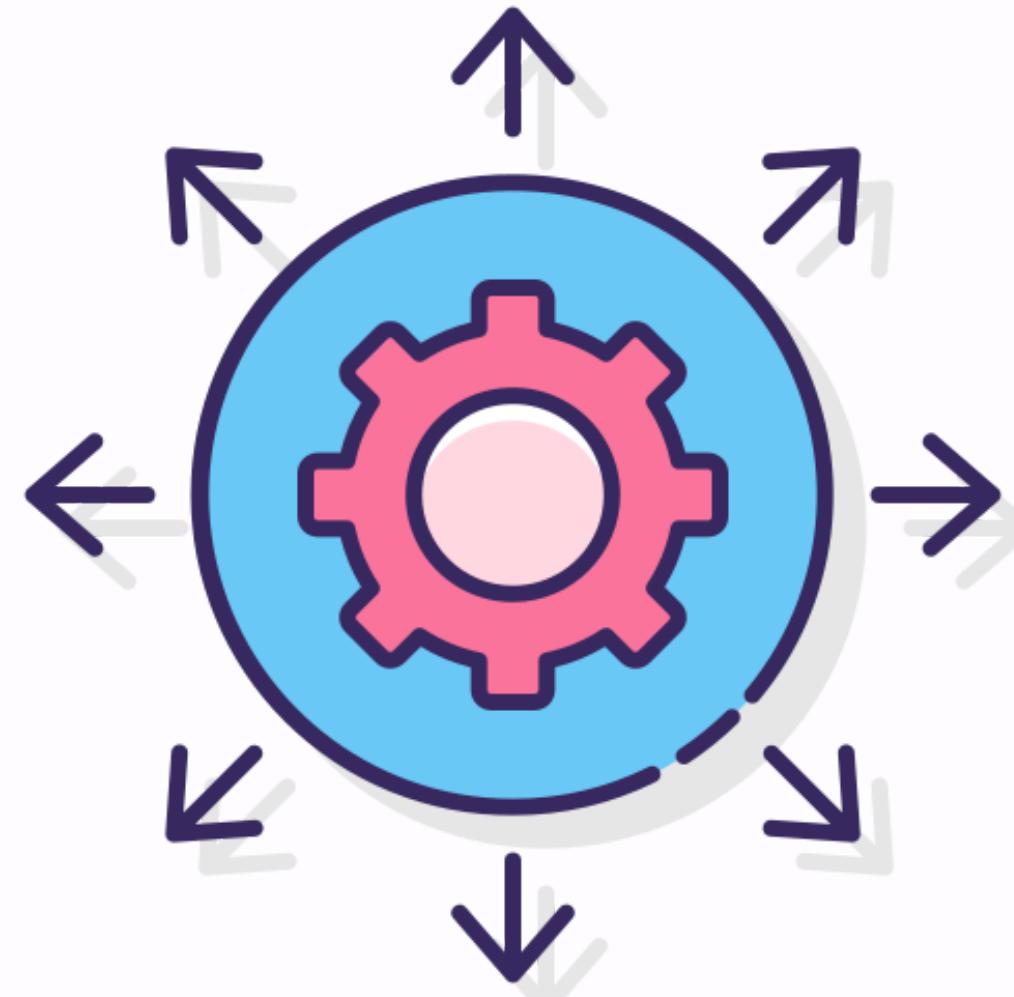
Shows likelihood of user referrals.



Build Time

Time required to compile, test, and package code.
Shorter build times speed up the development cycle.

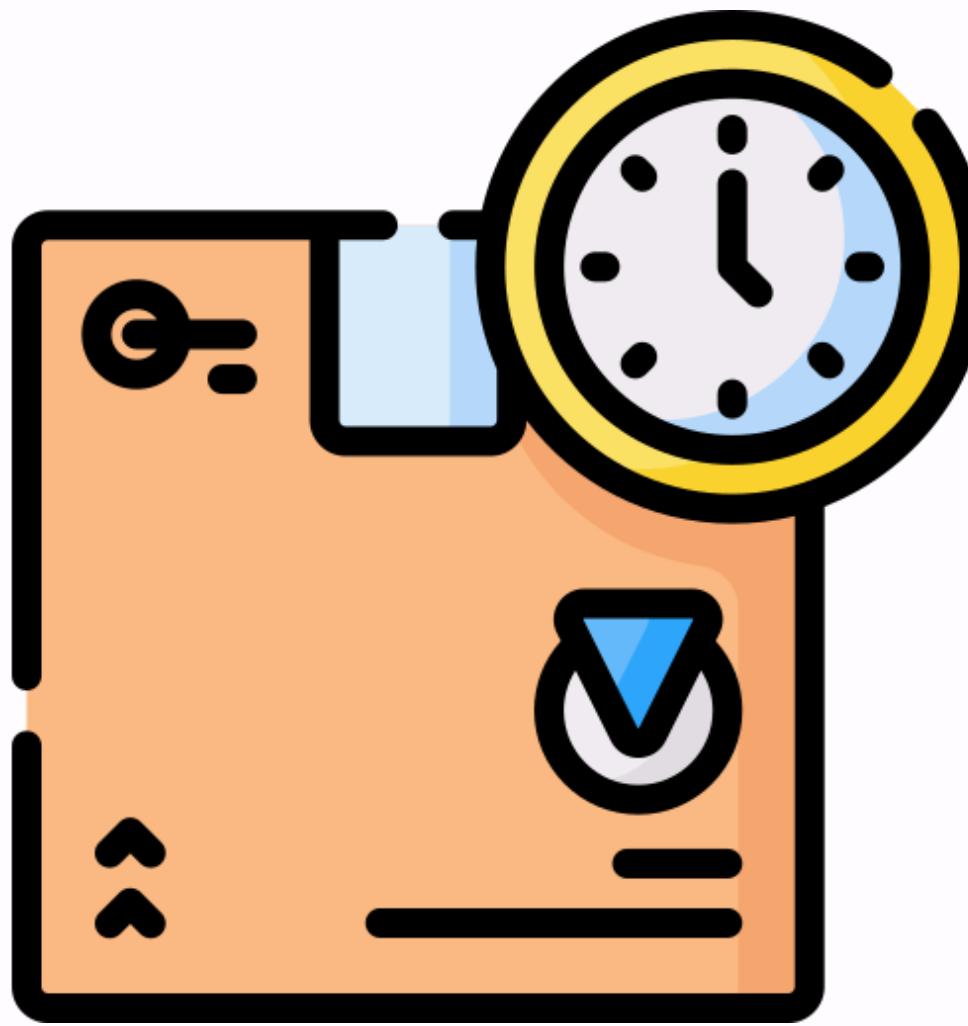
Monitors code compilation efficiency.



Deployment Frequency

Number of times code is deployed to production.
Frequent deployments indicate agile and mature
DevOps practices.

Measures release agility and speed.



Lead Time for Changes

Time from code commit to production release. A lower lead time means quicker delivery of new features or fixes.

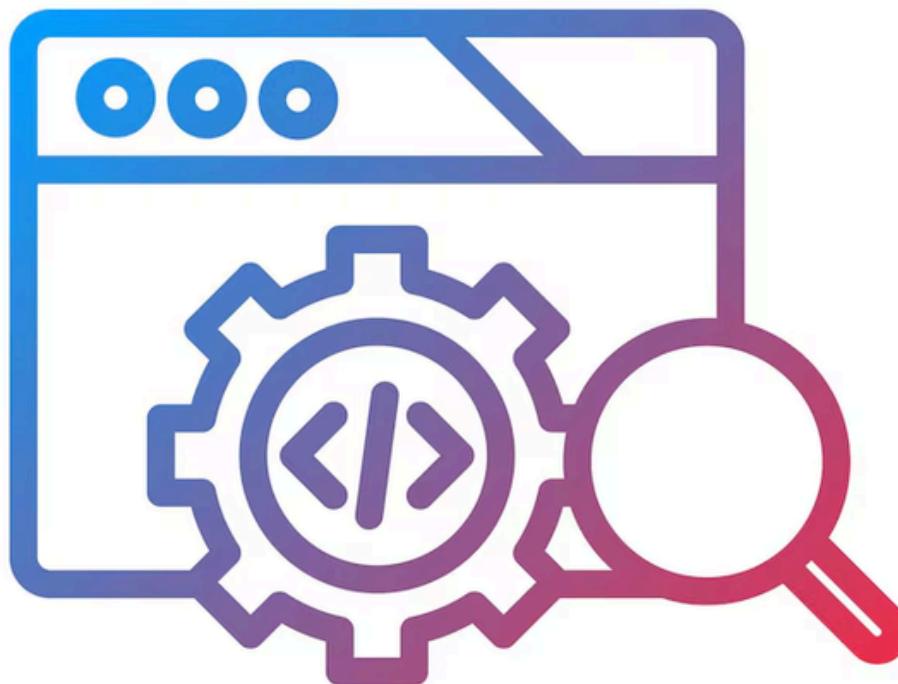
Tracks time to production readiness.



Change Failure Rate

Percentage of releases that cause failures or issues in production. Lower rates suggest a reliable release process.

Evaluates deployment stability risk.



Code Coverage

Measures how much of the codebase is tested by automated tests. High coverage helps reduce the risk of undetected bugs.

Ensures testing of codebase logic.



Test Pass Rate

Percentage of test cases passing in CI pipelines. Low pass rates may reflect unstable builds or regressions.

Validates build quality and consistency.



Code Quality Metrics

Includes metrics like cyclomatic complexity, duplication, and lint errors. Helps maintain clean, maintainable code.

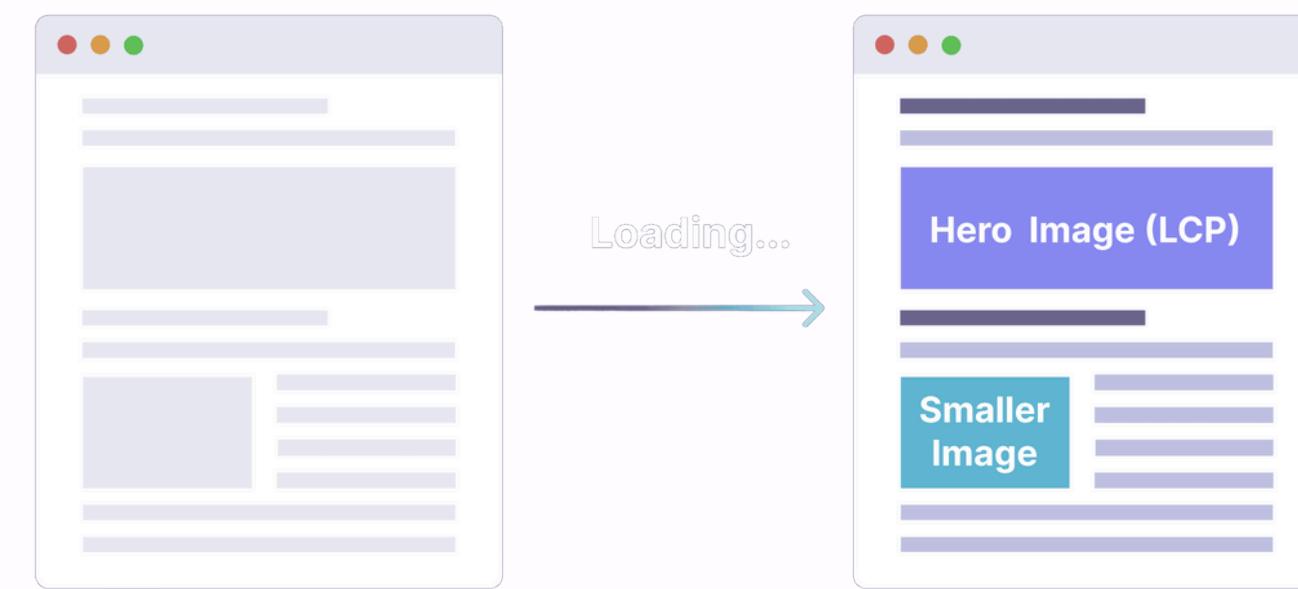
Tracks maintainability and technical debt.



First Contentful Paint (FCP)

Time until the first piece of content appears on screen.
Indicates how quickly users perceive progress after a request.

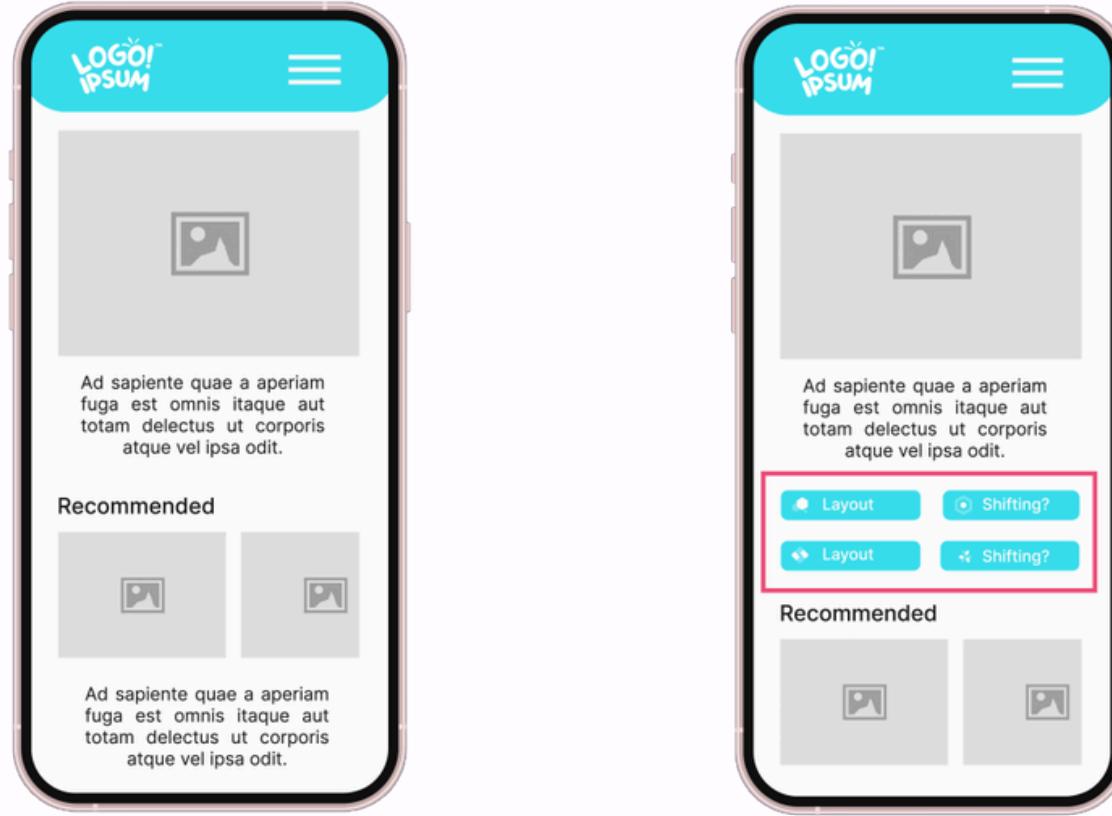
Measures perceived load start time.



Largest Contentful Paint (LCP)

Time it takes for the main visible element to render. A key metric for user-perceived load speed.

Tracks main content render speed.



Cumulative Layout Shift (CLS)

Measures unexpected layout changes during page load.
High CLS disrupts usability and visual stability.

Identifies unexpected visual shifts.

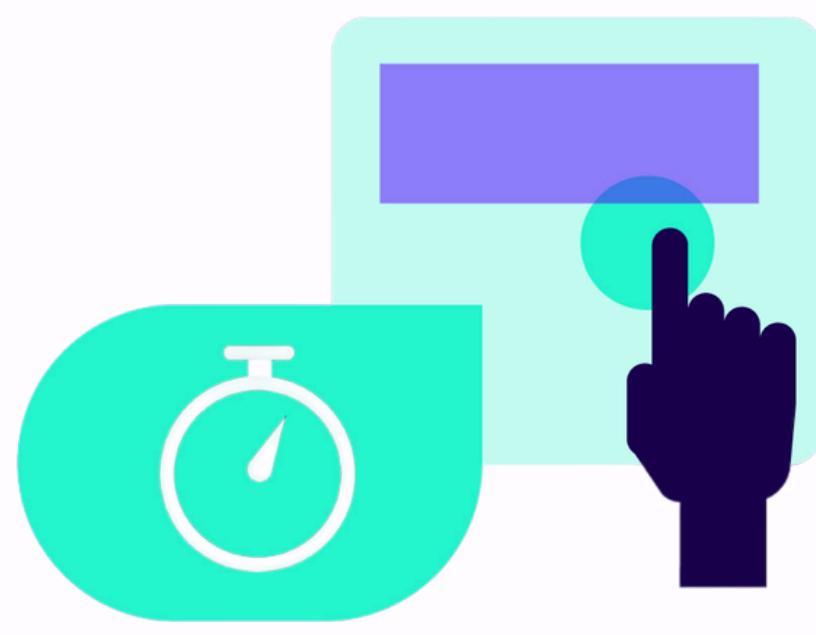
@tauseeffayyaz



Interaction to Next Paint (INP)

Measures responsiveness of UI to user actions. Lower INP means snappier, more responsive interfaces.

Evaluates real-time user interaction delay.



Time to Interactive (TTI)

Time taken for a page to become fully usable and responsive to user inputs. A key performance metric for SPAs and web apps.

Measures readiness for full user input.



About Tauseef Fayyaz

Full-Stack Engineer, Ranked #1 in Computer Engineering (Pakistan), Career Mentor for Aspiring Engineers, Tech Content Creator

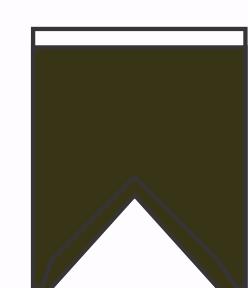
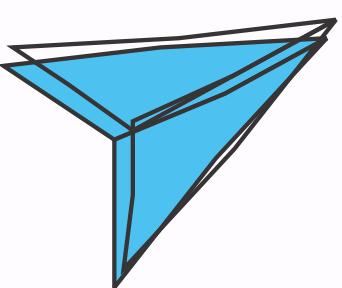
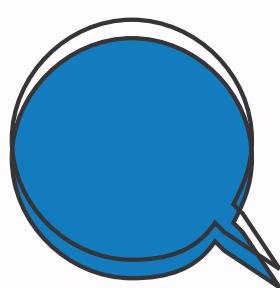
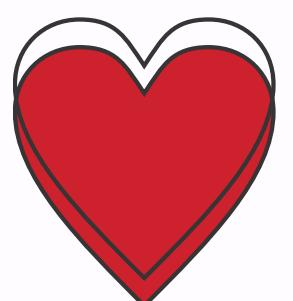
You can book a dedicated mentorship session with me for guidance, support, or just honest career advice.

→ www.topmate.io/tauseeffayyaz

@tauseeffayyaz



Your support keeps
me motivated.



@tauseeffayyaz