

INTRODUCCIÓN

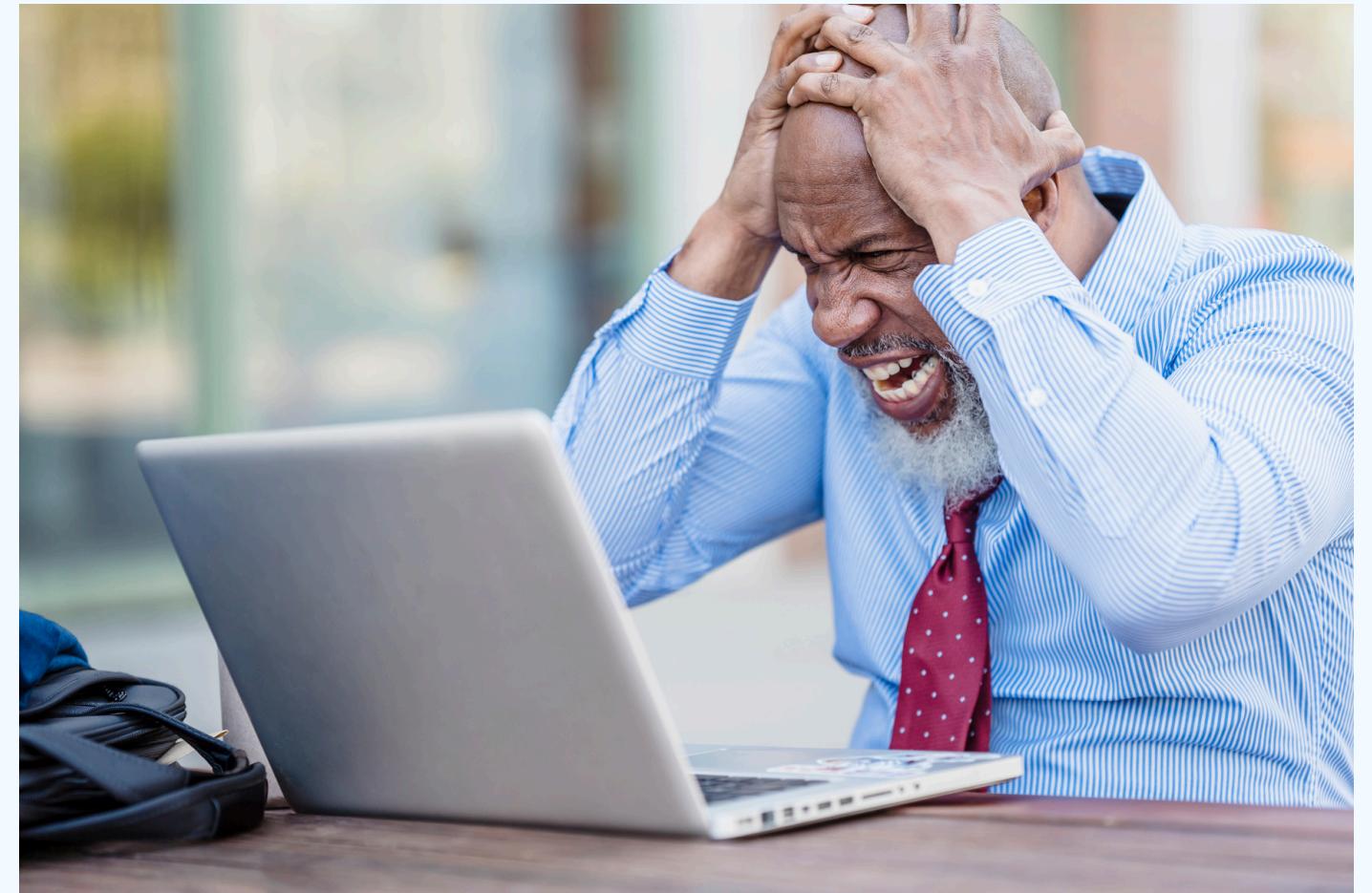
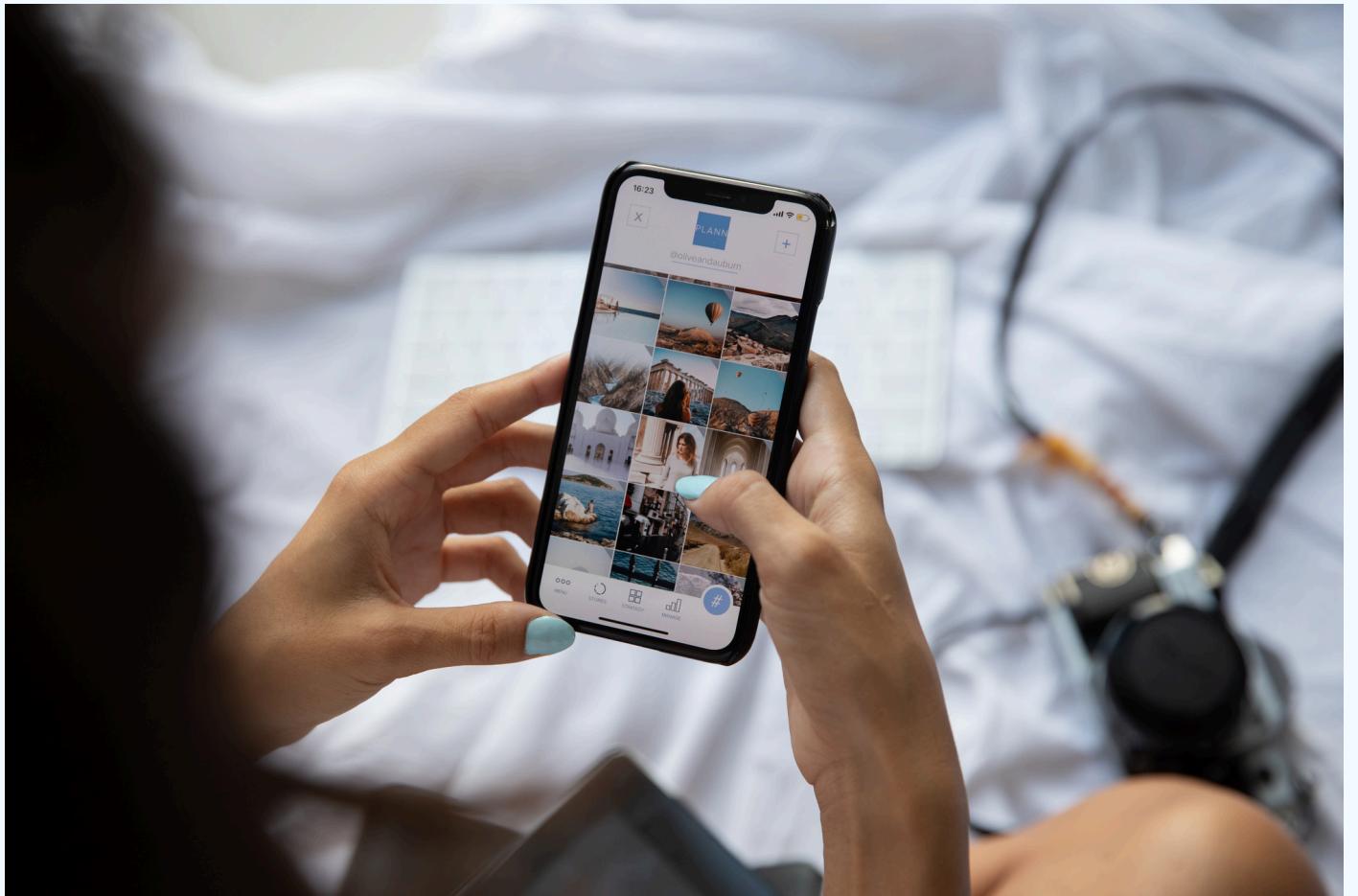
---

# RENDIMIENTO

---

Daniel Blanco Calviño

# INMEDIATEZ



# RENDIMIENTO

- Dos parámetros fundamentales:
  - **Response Time** (tiempo de respuesta)
  - **Throughput** (número de peticiones por unidad de tiempo)

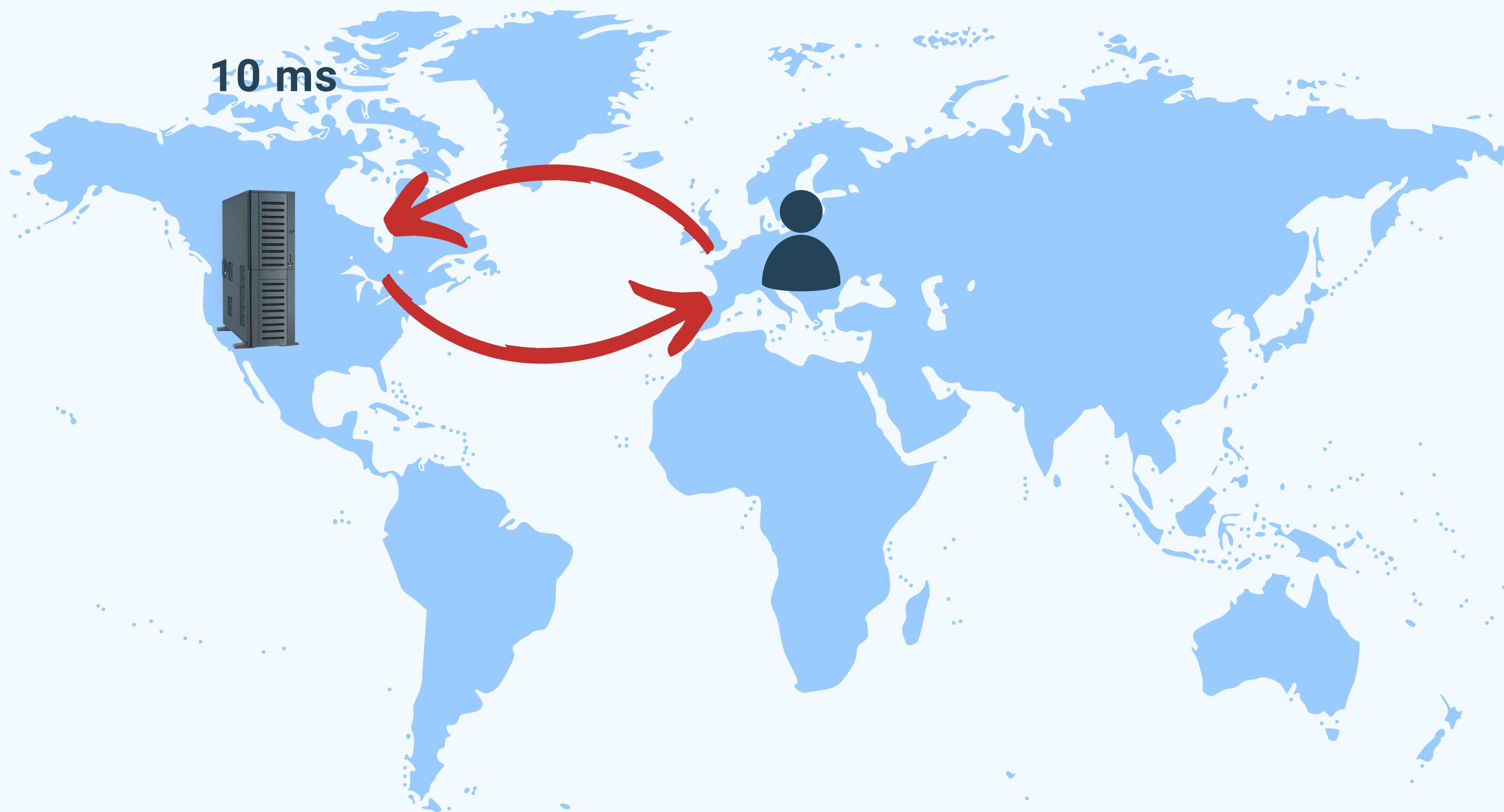
# TIEMPO DE RESPUESTA

- **Tiempo de procesamiento.**
  - Lógica.
  - DB.
- **Tiempo de espera.**
  - Espera por otras peticiones.
  - Cables.
  - Routers.
  - Switches.

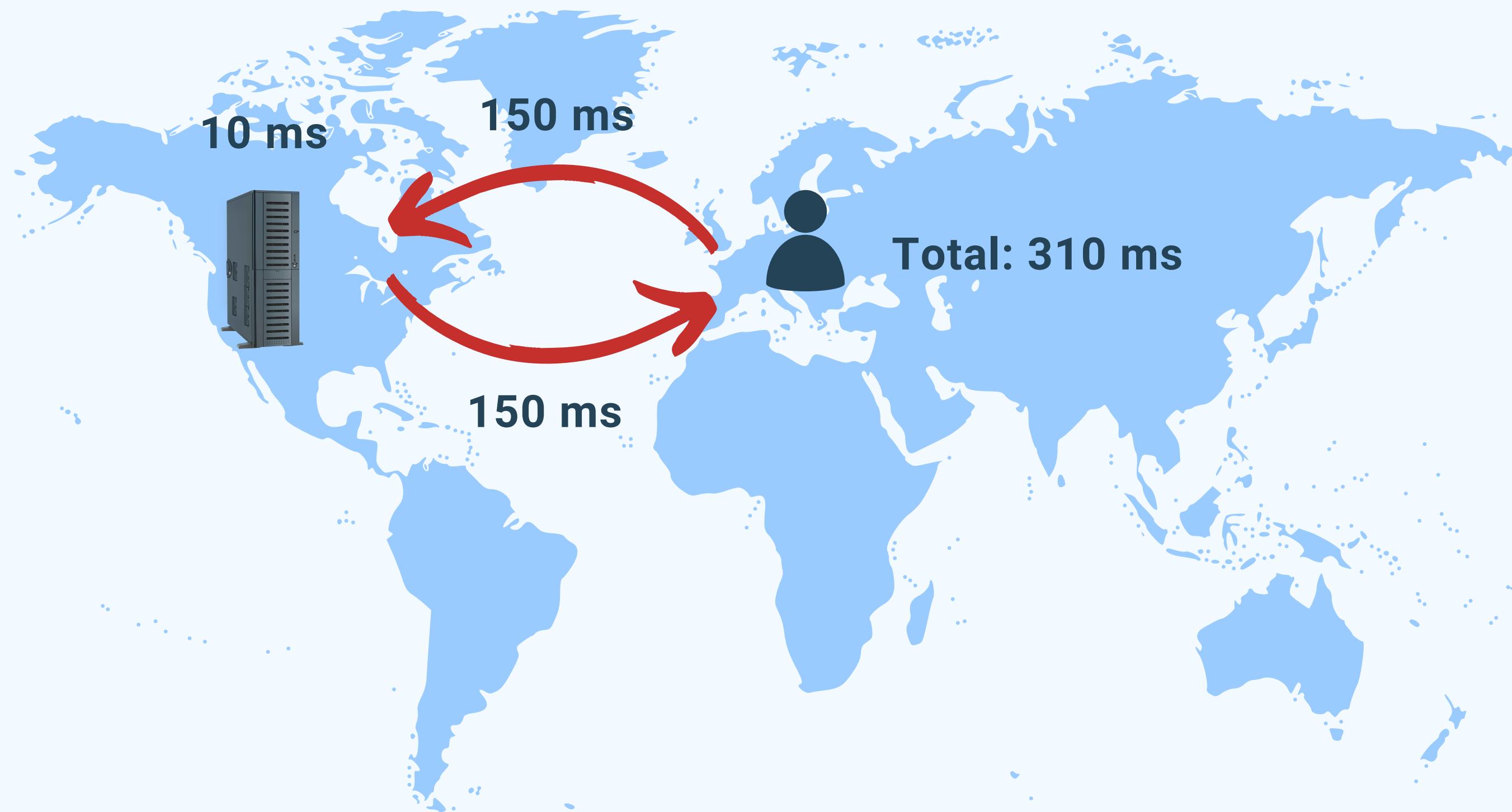
# TIEMPO DE RESPUESTA



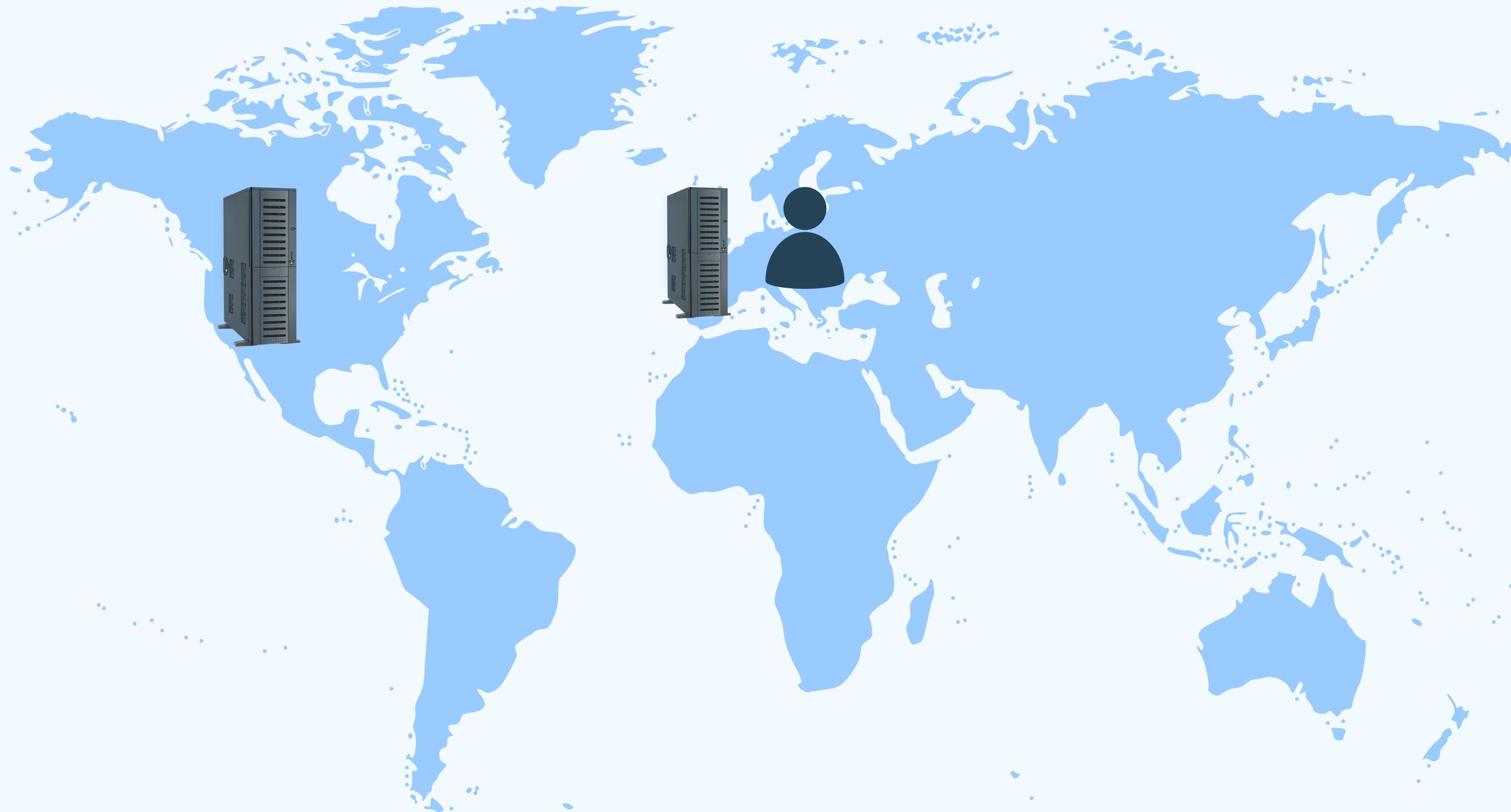
# TIEMPO DE RESPUESTA



# TIEMPO DE RESPUESTA



# TIEMPO DE RESPUESTA



# THROUGHPUT

- **Número de peticiones** que nuestro sistema es capaz de procesar **por unidad de tiempo**.
  - Puede ser incluso más importante que el tiempo de respuesta.

# THROUGHPUT

- **Número de peticiones** que nuestro sistema es capaz de procesar **por unidad de tiempo**.
  - Puede ser incluso más importante que el tiempo de respuesta.
- Sistema que soporta **una petición a la vez**. Tiempo de respuesta 100 ms.
  - Se realizan cinco peticiones. Tiempos de respuesta de cada una:
    - 100 ms
    - 200 ms
    - 300 ms
    - 400 ms
    - 500 ms

# MEDICIONES EN SISTEMAS REALES

- En sistemas reales no es sencillo medir estos parámetros.
  - Muchos usuarios concurrentes.
  - Múltiples servidores.

# MEDICIONES EN SISTEMAS REALES

- En sistemas reales no es sencillo medir estos parámetros.
  - Muchos usuarios concurrentes.
  - Múltiples servidores.



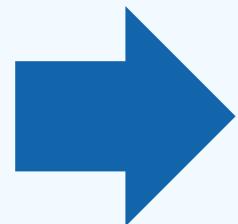
# MEDICIONES EN SISTEMAS REALES

- En sistemas reales no es sencillo medir estos parámetros.
  - Muchos usuarios concurrentes.
  - Múltiples servidores.



# MEDICIONES EN SISTEMAS REALES

- En sistemas reales no es sencillo medir estos parámetros.
  - Muchos usuarios concurrentes.
  - Múltiples servidores.



Distintos tiempos de respuesta y throughput

# MEDICIONES EN SISTEMAS REALES

- En sistemas reales no es sencillo medir estos parámetros.
  - Muchos usuarios concurrentes.
  - Múltiples servidores.
- ¿Cómo analizamos estos datos?
  - ¿Media?
  - ¿Mediana?

# MEDICIONES EN SISTEMAS REALES

- En sistemas reales no es sencillo medir estos parámetros.
  - Muchos usuarios concurrentes.
  - Múltiples servidores.
- ¿Cómo analizamos estos datos?
  - ¿Media?
  - ¿Mediana?



# MEDICIONES EN SISTEMAS REALES

- En sistemas reales no es sencillo medir estos parámetros.
  - Muchos usuarios concurrentes.
  - Múltiples servidores.
- ¿Cómo analizamos estos datos?
  - ¿Media?
  - ¿Mediana?
- Histograma y percentiles



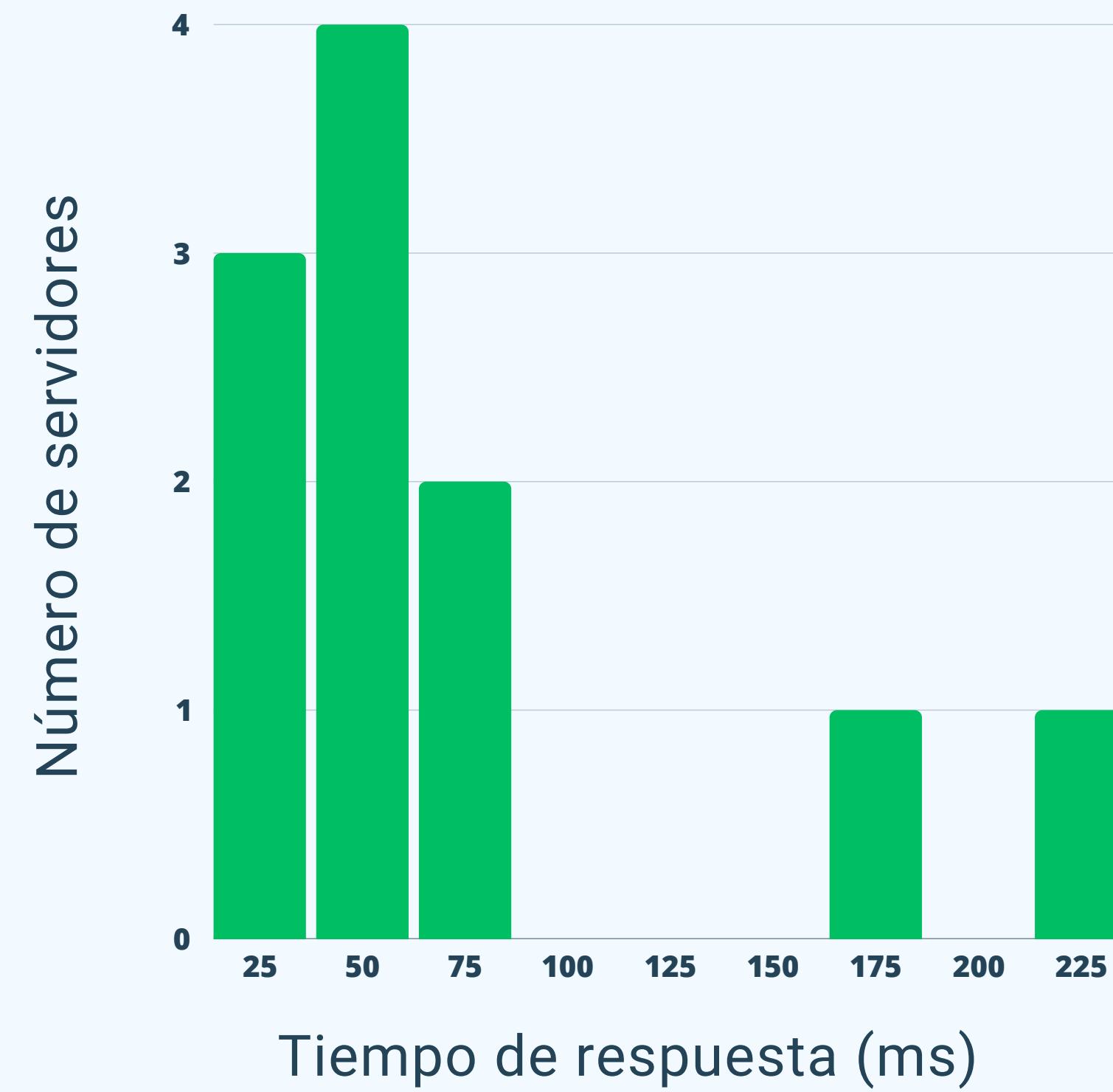
# MEDIA Y MEDIANA

Server	Tiempo de Respuesta (ms)
S1	30
S2	55
S3	20
S4	155
S5	30
S6	205
S7	30
S8	25
S9	15
S10	35
S11	55

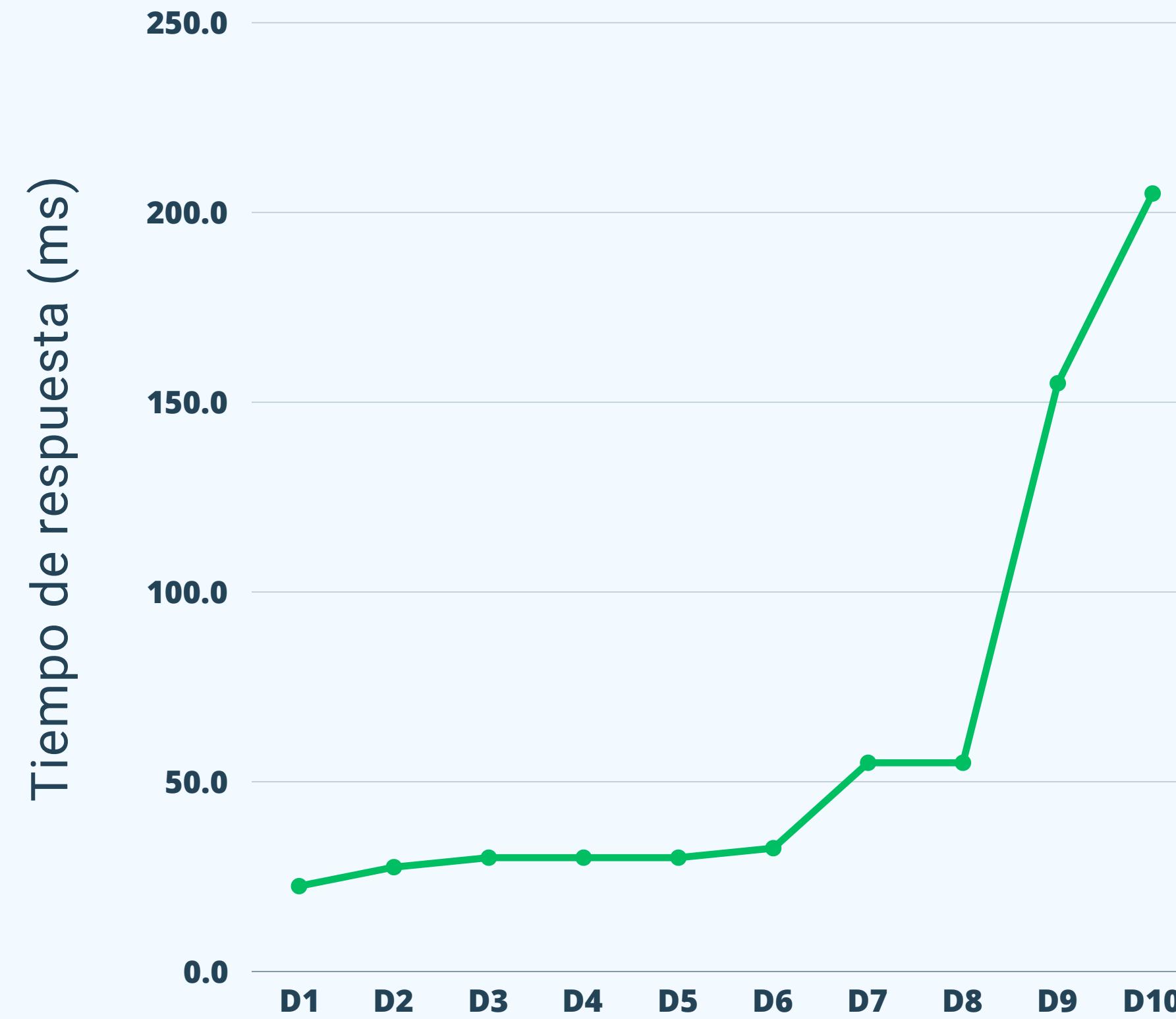
**Media:** 59.5 ms

**Mediana:** 30 ms

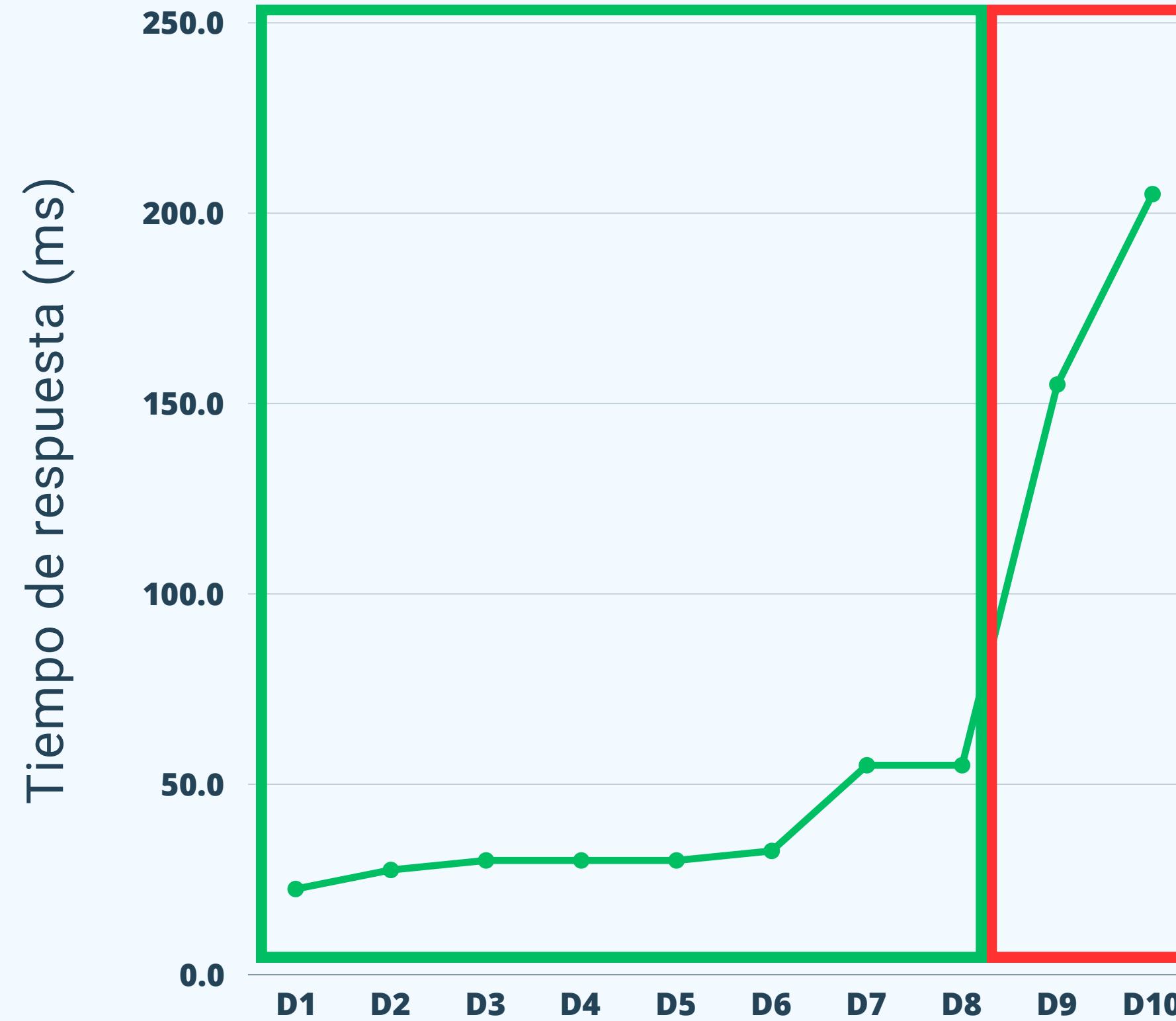
# HISTOGRAMA



# DECILES



# TAIL LATENCY



# REQUISITOS

- A la hora de tomar requisitos sobre tiempo de respuesta:
  - **Evitar media y mediana.**
  - Mejor usar la forma "El 99% de las peticiones deben tener un tiempo respuesta menor a X ms".