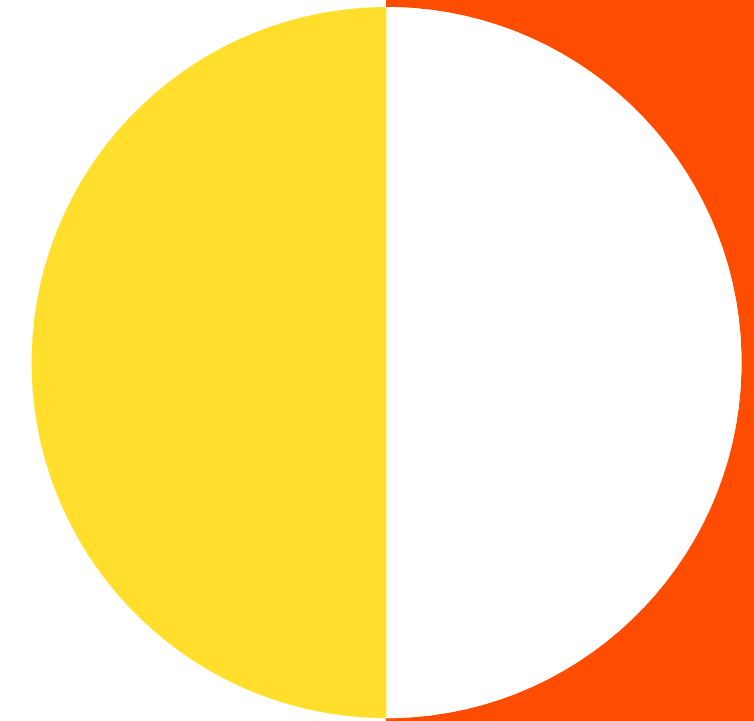


Sistemas Distribuídos:

Buscadores Web



LAVÍNIA BARBOSA

18 DE JUNHO DE 2025

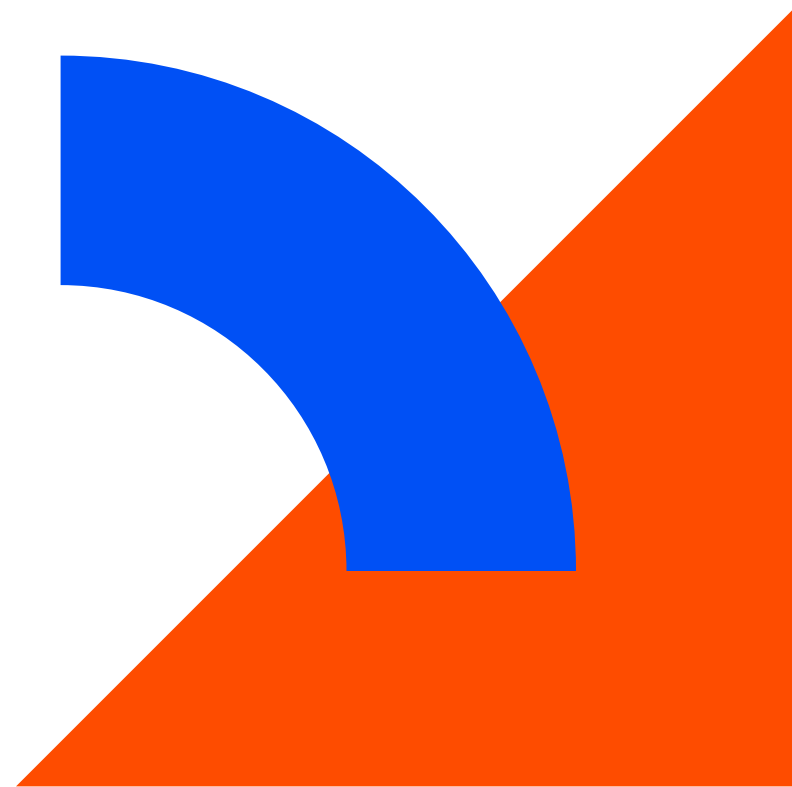


Sumário

VISÃO GERAL

- Introdução: Desvendando os Sistemas Distribuídos
- Aplicação e Importância: Onde os Buscadores Web Se Encaixam
- Desafios
- Soluções
- Tendências

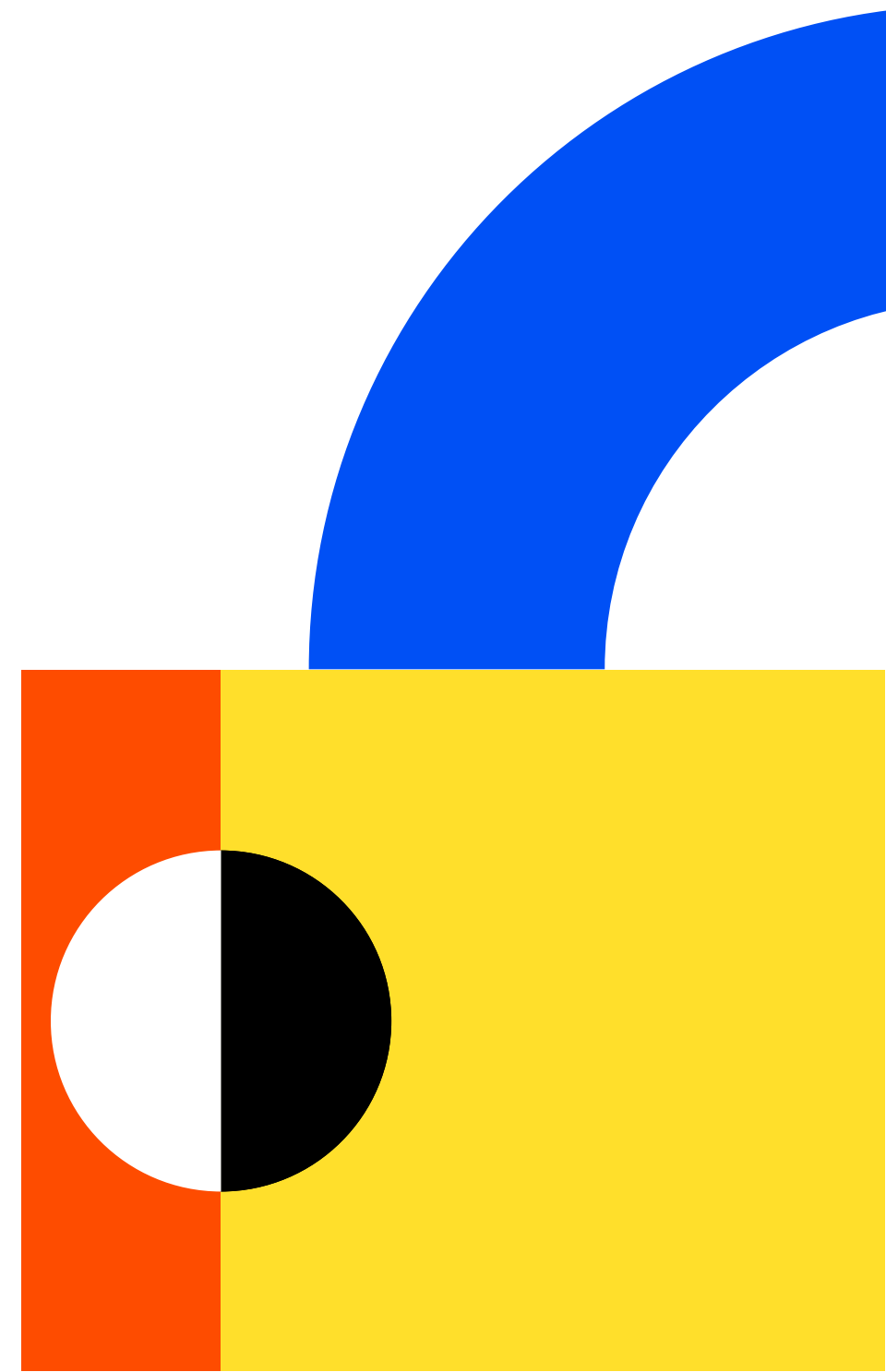
“ Talvez você não perceba, mas toda vez que usa a Pesquisa do Google, verifica seu e-mail ou paga com cartão de crédito, está confiando em um sistema de computação distribuído.”



Introdução

O QUE SÃO SISTEMAS DISTRIBUÍDOS?

Podemos definir esses sistemas como uma **rede de computadores** que, embora separados fisicamente, **trabalham coletivamente** para **realizar tarefas** com **mais eficiência** do que uma **única máquina** poderia.



Aplicação e Importância

- As aplicações reais de sistemas distribuídos incluem diversos **Serviços da web**, como os **Buscadores (Yahoo, Bing, Google)**, websites, comércio eletrônico.
- Imagina ter que indexar bilhões de páginas para processar milhões de consultas e buscas simultâneas.
- O índice da Google possui mais de 100 petabytes de dados, acumulando bilhões de páginas e entregando elas em altíssima velocidade.

Aplicação e Importância

- Nesse contexto, fica clara a importância dos SD's para maior:
 - **Escalabilidade:** Lidar com o crescimento exponencial de dados e usuários (a web não para de crescer!).
 - **Disponibilidade:** Garantir que o buscador esteja sempre online, mesmo se parte do sistema falhar.
 - **Desempenho:** Entregar resultados relevantes em milissegundos.



Desafios

- **Consistência de Dados:**

Manter os índices atualizados e sincronizados em milhares de servidores ao redor do mundo. Imagine um site que acaba de ser lançado ou atualizado!

- **Latência de Rede:**

Atrasos na comunicação entre os diferentes nós do sistema, impactando o tempo de resposta das buscas.

- **Tolerância a Falhas:**

Como garantir que a falha de um servidor (ou data center inteiro!) não derrube o buscador.



Desafios

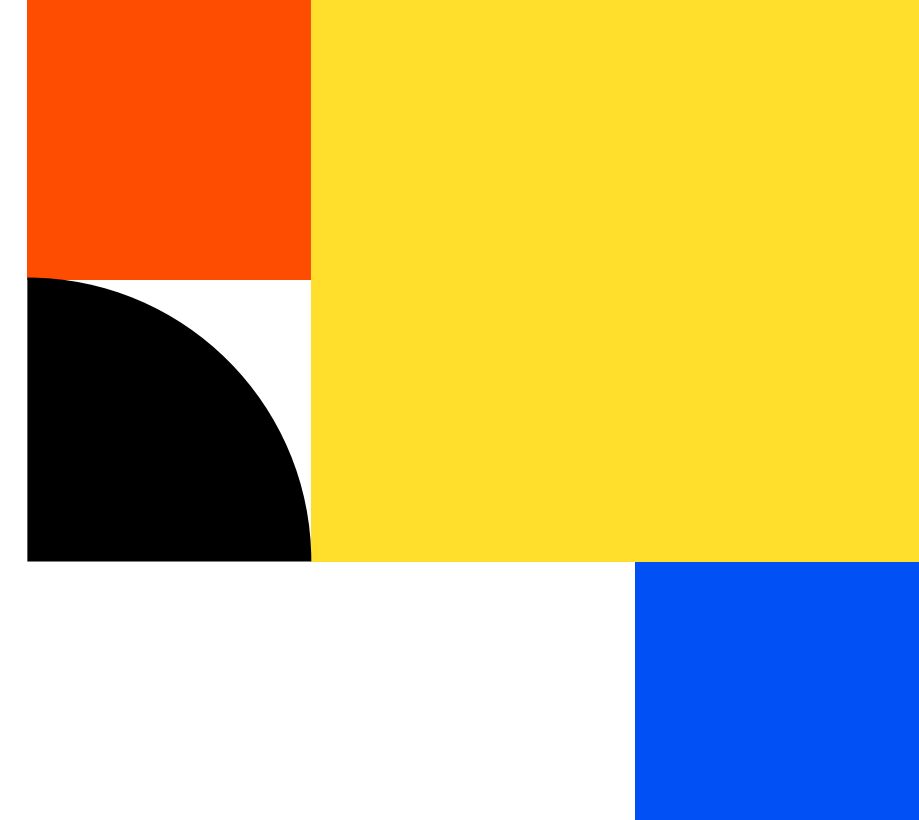
- **Gerenciamento de Recursos:**

Otimizar o uso de poder computacional, memória e armazenamento em uma escala gigantesca.

- **Complexidade Operacional:**

Monitorar, depurar e manter um sistema tão grande e complexo.

Soluções



Particionamento:

Dividir o índice da web e as consultas em partes menores, distribuídas entre vários servidores (Ex: Sharding).

Replicação:

Manter cópias dos dados e serviços em diferentes locais para alta disponibilidade e tolerância a falhas.

Balanceamento de Carga:

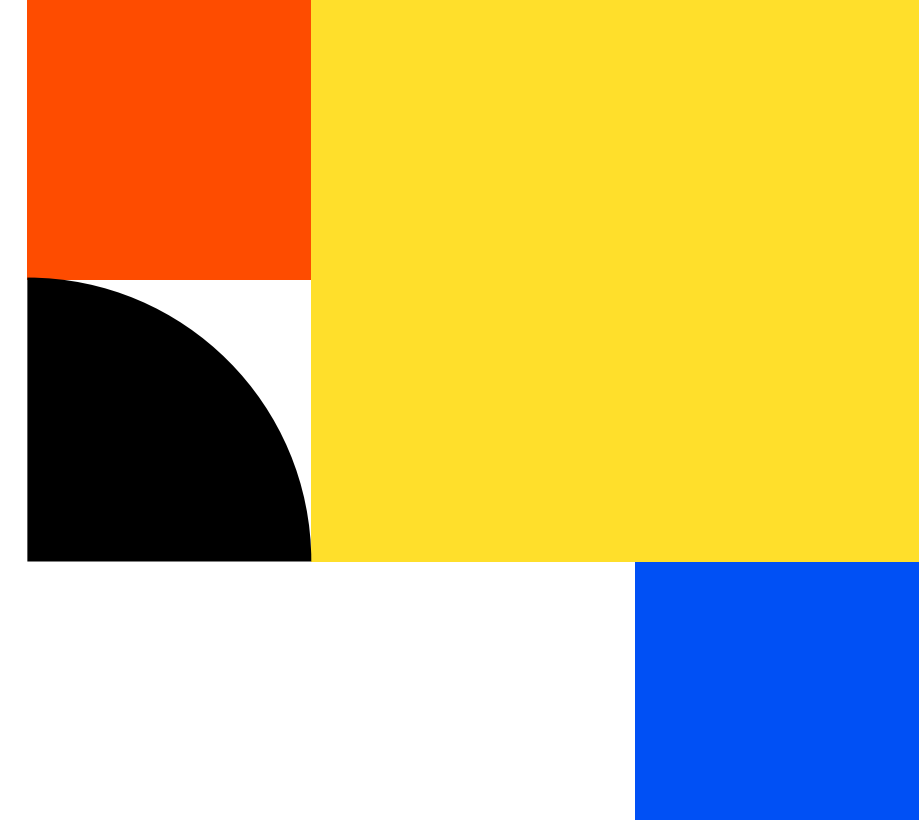
Distribuir as requisições de busca de forma eficiente para evitar sobrecarga em um único servidor.

Soluções

Cache Distribuído:

Armazenar resultados de buscas frequentes mais próximos do usuário para acelerar as respostas.

Aproximadamente 194.000 resultados (0,36s)





Tendências

Inteligência Artificial e Machine Learning:

Aprimorando a relevância dos resultados de busca e a personalização em tempo real.

Computação em Nuvem:

Aumento da utilização de infraestruturas flexíveis e escaláveis para hospedagem de buscadores.



Tendências

Edge Computing:

Processamento de dados mais próximo do usuário final para reduzir a latência, especialmente em buscas por voz e realidade aumentada.

Novas Arquiteturas:

Tornando o desenvolvimento e a manutenção dos buscadores mais ágeis e resilientes.



Referências

Entendendo Sistemas Distribuídos

Disponível em: <[hive](#)>. Acesso em: 12 jun. 2025.

Google Busca: história e como funciona o maior buscador do mundo

Disponível em: <<https://www.mundoconectado.com.br/ciencia/google-busca-historia-e-como-funciona-o-maior-buscador-do-mundo/>>.

Acesso em: 12 jun. 2025.

Como funcionam os mecanismos de busca: Google, Bing e outros

Disponível em: <<https://www.conversion.com.br/blog/mecanismos-de-busca>>. Acesso em: 12 jun. 2025.

Obrigada!

