

Infraestrutura Tecnológica da PCDaS

Lucas Zinato Carraro
Tecnologista em Saúde
Pública

Evolução da Plataforma

A versão inicial foi montada inicialmente em uma máquina doada pelo Cemaden:

- ✓ 12 Cores
- ✓ 32Gb de memória RAM
- ✓ 2 Tb de Disco
- ✓ Todos serviços rodando no mesmo servidor

O desafio:

- ✓ Migrar toda arquitetura de forma transparente para os usuários para rodar em um Cluster de Servidores.
- ✓ Continuar utilizando apenas softwares livres para garantir as atualizações e evolução da plataforma.

Plataforma 1.5:

- ✓ 6 servidores fisicamente localizados no LNCC
- ✓ 192 Cores
- ✓ 512Gb de memória RAM
- ✓ 28,2 Tb de Disco úteis
- ✓ Arquitetura de Cluster com Serviços distribuídos entre os Servidores



Principais Softwares da Plataforma:



Serviços Oferecidos:

- ✓ Cluster ElasticSearch
- ✓ Dashboards Customizados pelas necessidades dos parceiros
- ✓ Dashboard de análise Geral ao usuário
- ✓ JupyterHub com áreas e recursos específicos para parceiros
- ✓ Rstudio Server com conexão aos dados para todos os usuários
- ✓ Dashboards R Shinny específicos para parceiros.

Tempo de resposta da Plataforma:

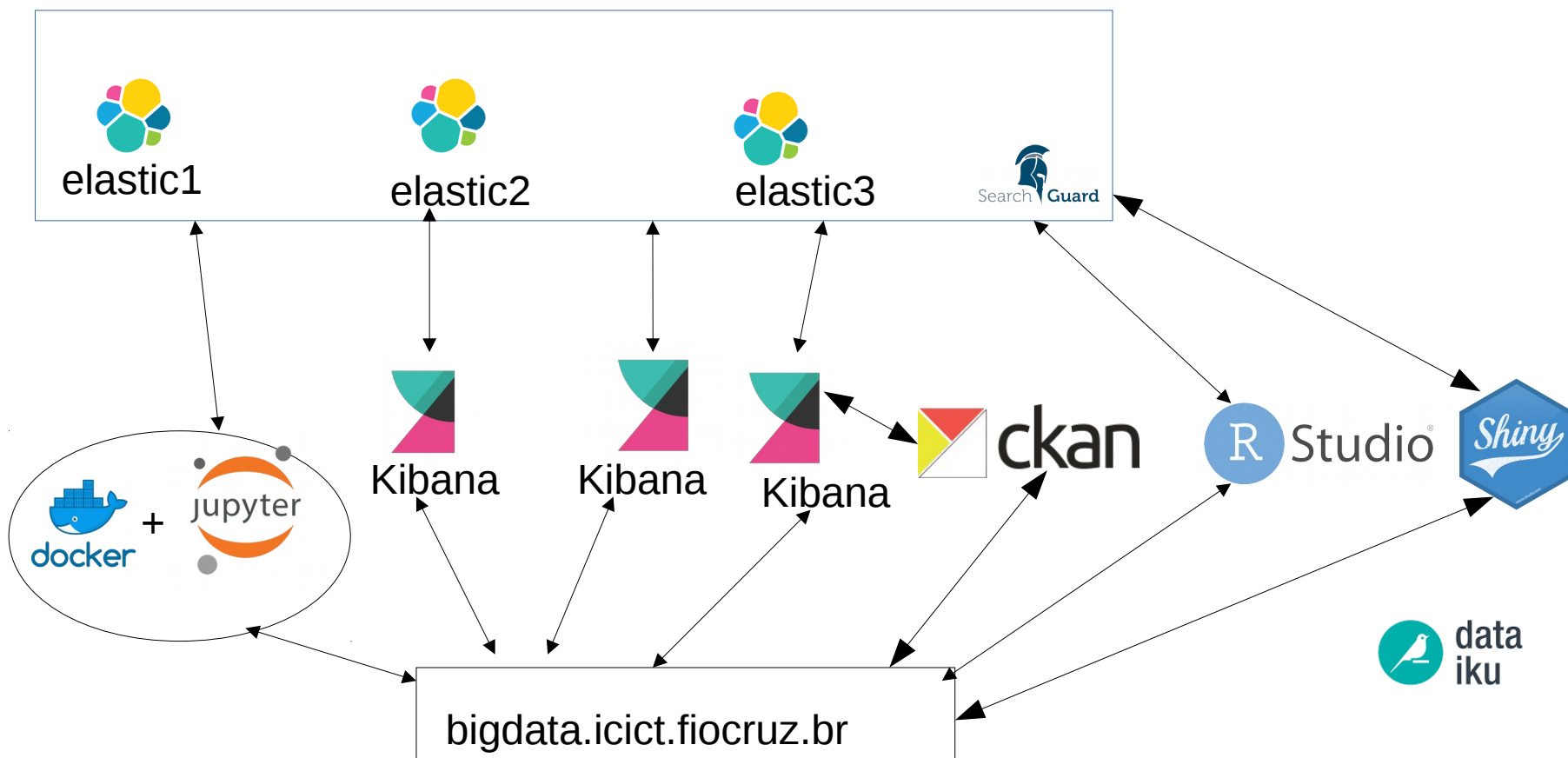
Para uma consulta com retorno de 10.000 agregações:

- ✓ - 0.53s consulta efetuada pela Internet
- ✓ - 0.256s consulta efetuada usando o Cluster Local
- ✓ - 1.2s servidor Google Colab*

* apenas testes não disponível aos usuários ainda

Desenho Geral da Solução:

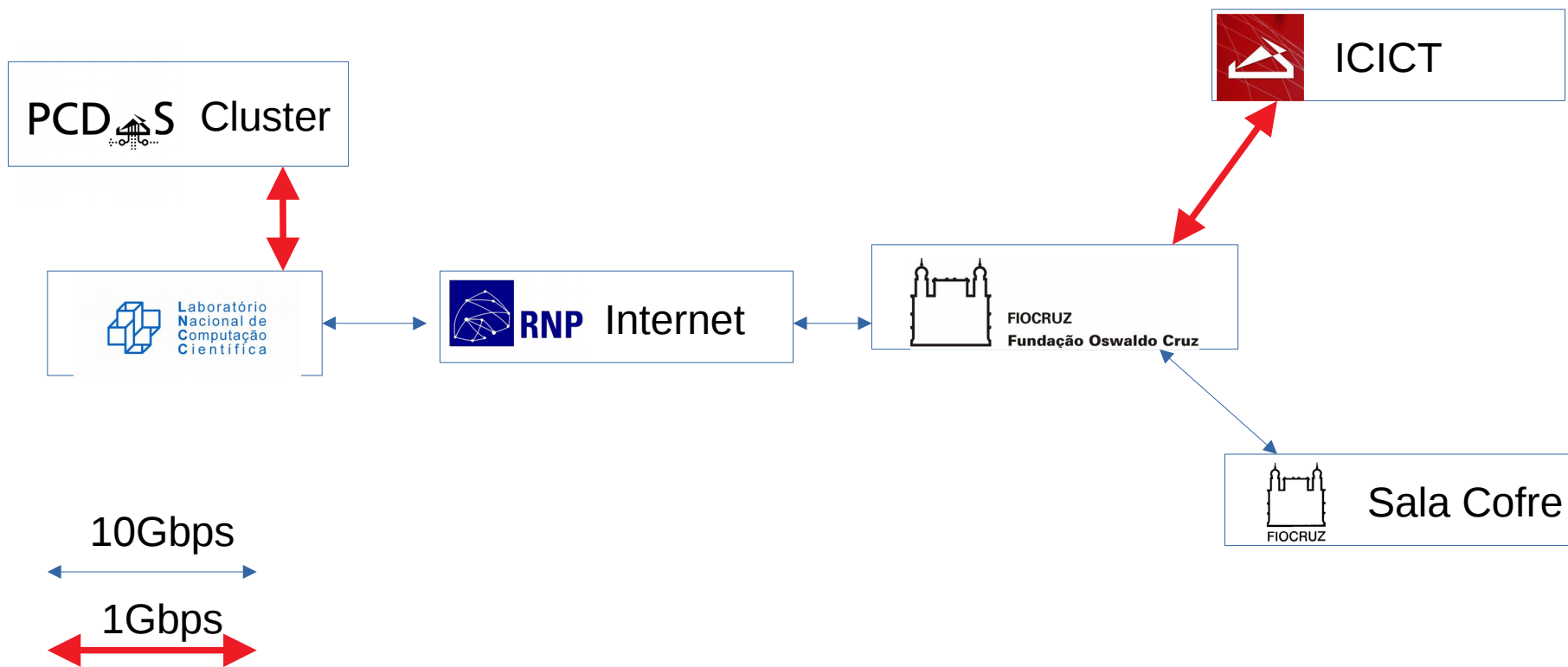
Cluster



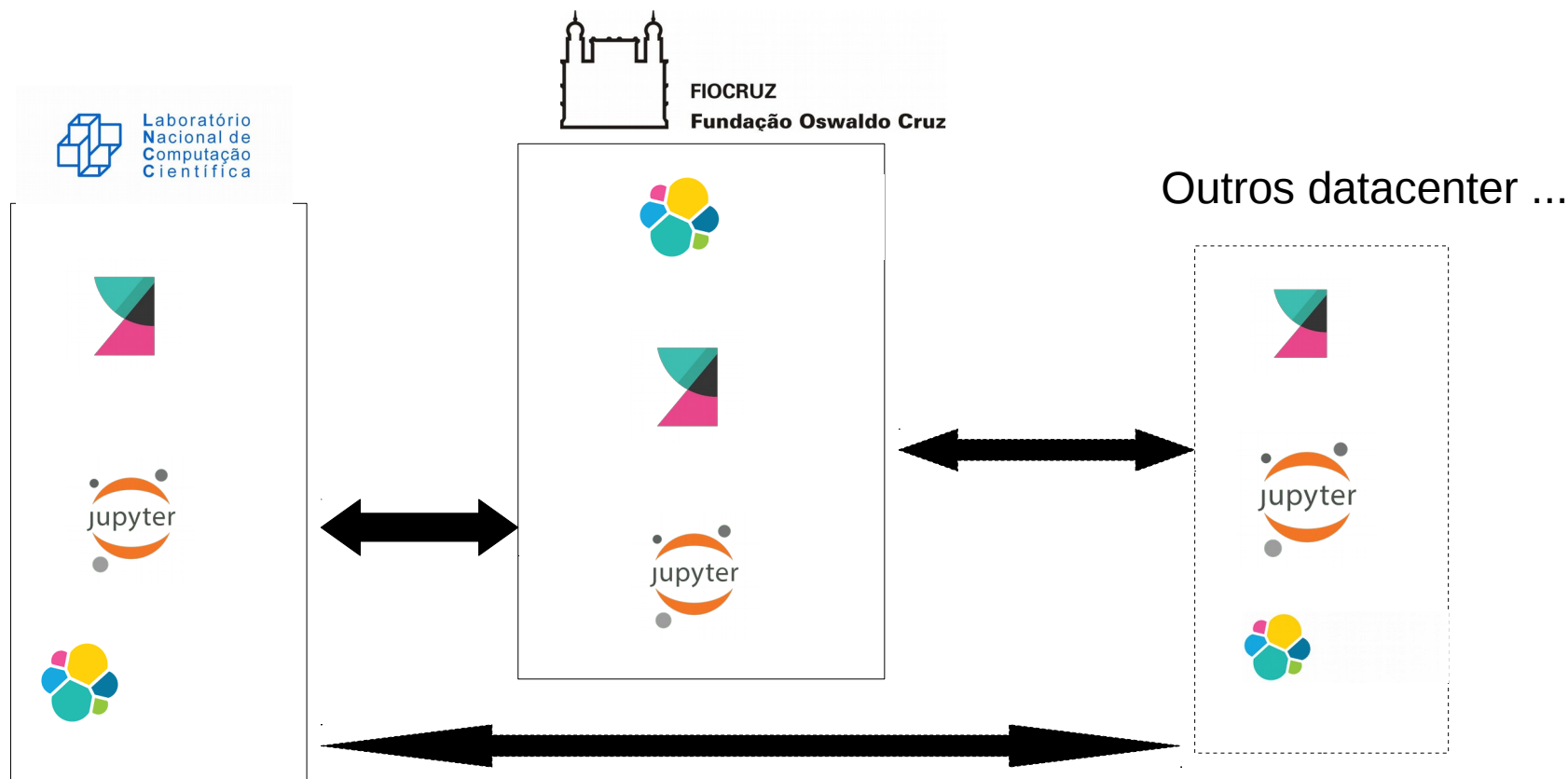
Futuro da plataforma:

- ✓ Rede de 10Gbps entre todos os pontos
- ✓ Expansão de memória e storage dos servidores
- ✓ Servidores distribuídos entre vários serviços de Cloud
- ✓ Cluster Elastic para os Parceiros e para a Internet
- ✓ Processamento por meio de GPU (Graphics Processing Unit)
- ✓ Gráficos e monitoramento histórico de toda a performance das máquinas do cluster

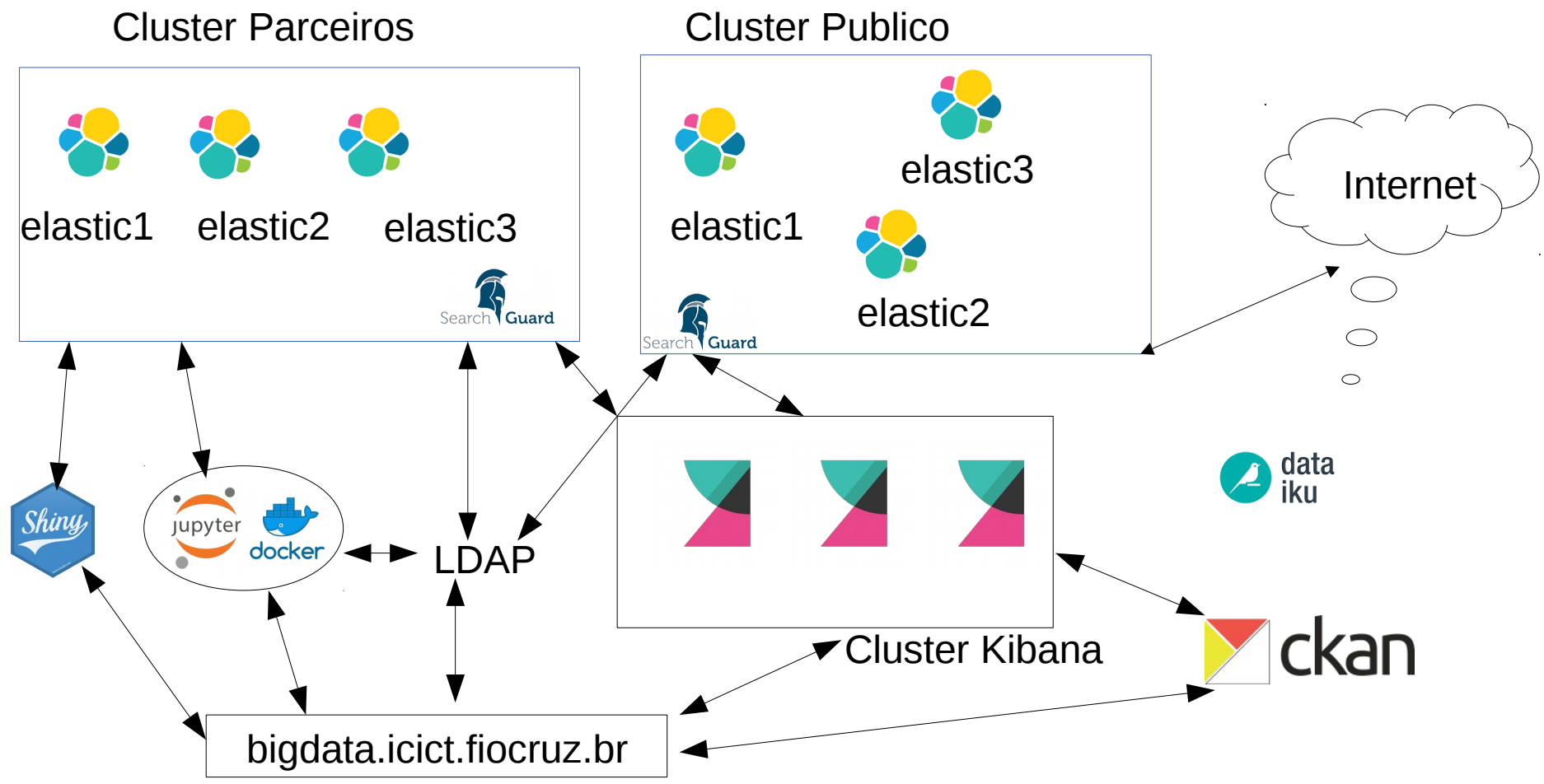
Rede de 10Gbps entre todos os pontos:



Servidores distribuídos entre vários provedores de Cloud:



Cluster Elastic para os Parceiros e para a Internet:



Processamento por meio de GPU (Graphics Processing Unit):

- ✓ GPUs são chips de computador que realizam cálculos matemáticos de forma paralelizada e veloz.
- ✓ Memória mais rápida que a utilizada pelas CPUs
- ✓ Existem várias bibliotecas de data science prontas para uso de GPU.
- ✓ Custo de uma placa gráfica GPU: Aproximadamente R\$ 50.000,00

bigdata.icict.fiocruz.br



www.facebook.com/fiocruz.icict

[twitter.com/@Icict_fiocruz](https://twitter.com/Icict_fiocruz)

www.youtube.com/videosaudefio

www.icict.fiocruz.br