



Palavras-chave: Cluster de processamento. Equipamento multiusuário

Introdução/Objetivo:

Serviço para processamento de dados de Projetos Científicos, composto por equipamentos de alto desempenho, onde pesquisadores e alunos de pós-graduação do Instituto de Biologia podem executar suas análises computacionais de forma rápida e eficiente. Planejado para atender demandas distintas de Biologia Computacional, teve aporte financeiro inicial via Reserva Técnica Institucional (RTI-FAPESP). Como estratégia de crescimento, qualquer pesquisador pode colaborar. Há colaborações de 8 mil reais, em pentes de memória RAM, e até uma futura aquisição de um novo nó para o cluster.

Metodologia:

Visitas em outras instituições melhorou o entendimento da solução, mas o que mais colaborou com o sucesso do projeto foi o teste piloto realizado com alunos de distintas demandas por análise de dados. Todos os equipamentos são compartilhados entre os pesquisadores do Instituto. A Direção fez um investimento inicial e agora os pesquisadores estão contribuindo para o crescimento da solução, ao invés da aquisição de equipamentos para uso individual

Resultados

Dez projetos científicos foram atendidos em menos de um ano de inauguração do cluster. Todos os Departamentos do Instituto fizeram uso do serviço, mostrando o alto grau de compartilhamento do recurso computacional. Outro resultado positivo é a velocidade de disponibilização do recurso aos pesquisadores, antes apenas o processo de aquisição demorava meses. Agora a máquina virtual é liberada em questão de horas. Integração entre os recursos computacionais dos equipamentos é feita através de um software open-source, chamado XCP-ng, enquanto o gerenciamento é executado via interface gráfica pelo Xen Orchestra. Ambos os softwares permitem o escalonamento eficaz e rápido da solução. A execução foi facilitada pois o instituto possui um Data Center adequado, que conta com: climatização profissional, sistema de combate a incêndio, nobreaks e gerador.

Processamento de Dados Científicos



192 processadores
1,2 TB memória RAM
313 TB espaço em disco

Os projetos científicos fazem uso de máquinas virtuais para análise e processamento de dados.



Conclusão:

Houve dificuldade inicial no convencimento da adoção do Cluster compartilhado no Instituto, porém, vencida esta fase, conseguimos concluir o projeto e atender rapidamente às demandas por recursos computacionais. Análises de dados sobre infecção pelo SARS-CoV-2 foram beneficiadas através do uso deste projeto.