Providing Superpowers to Bioinformaticians

Time is everything for a bioinformatician. With an active R&D group constantly creating new datasets that need to be analyzed, keeping up with their requests is an uphill battle. Sombeet Sahu, Associate Director of Bioinformatics understood this predicament all too well. He joined Nkarta Therapeutics after years of experience building custom analysis pipelines for companies like Mission Bio and Thermo Fisher Scientific.

His vision was to pioneer a scalable bioinformatics strategy that could match the increasing analysis needs of this growing cell therapy company.

basepair

nkarta

SUMMARY

BEFORE

- Increasing volume of NGS data created a never-ending backlog of work for bioinformaticians
- Bench scientists relied on bioinformatics for primary analysis of data sets
- Bottlenecked process resulted in stalled R&D projects and overworked teams
- Needed a solution that was user-friendly (point-and-click GUI) for bench scientists
- Needed a way to leverage the existing compute and storage resources in their AWS account

AFTER

- Bench scientists were empowered to analyze their NGS data without needing to code or wait for bioinformatics
- Bioinformaticians were free to devote their time to designing custom pipelines and helping with advanced analysis questions
- Research data was safe & secure in their own AWS account
- Benefitted from economies of scale with AWS by using their own computing and storage resources
- Bioinformaticians use the CLI and APIs to automate pipelines and automate data ingestion from third party sequencers

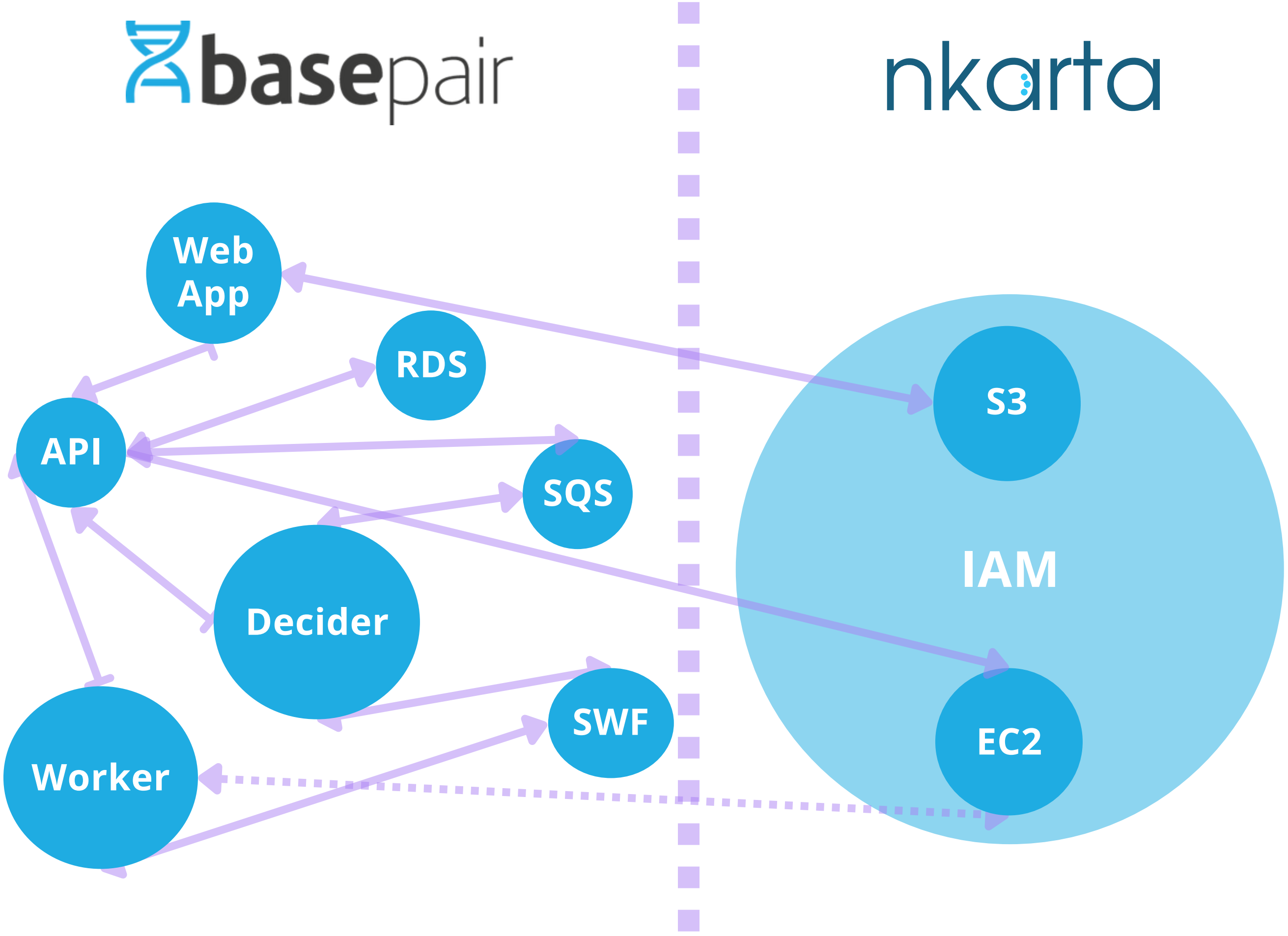
basepair

nkarta

Nkarta scientists work around the clock to discover and develop natural killer (NK) cell therapies to treat cancer. Throughout their development process, they create increasing amounts of NGS data across multiple application areas that require analysis. But without a way for scientists to perform routine analysis independently, they were experiencing a significant bottleneck in analysis that delayed scientific insights. **“We wanted to enable every non-bioinformatician to become some sort of data scientist by giving them tools and visualizations to be able to analyze their data. With that vision, we started evaluating possible solutions including building our own in-house software and commercially available off-the-shelf solutions.”** Traditionally, a combination of automation and adding headcount as analysis needs increase would have been the standard approach to take. Sombeet knew from experience that this approach was not scalable and set out to find a better way to help bioinformaticians and bench scientists work more efficiently together. He had experience using first-generation bioinformatics platforms that offered analysis capabilities but he was disappointed with their scalability and ease of use. He had a vision for a user-friendly, point-and-click interface for bench scientists with out-of-the-box core visualization and reporting capabilities that could integrate multiple data sets.

“We needed a solution in which initial analysis could be done by a bench scientist before a bioinformatician would jump in to assist with more advanced methodologies involving machine learning and AI. The objective is to help them (bench scientists) gain deeper insights from their data before needing our help. Ideally, this is what a bioinformatician should be doing as opposed to crunching primary or routine analysis data that can otherwise be accomplished by bench scientists running automated pipelines.” One approach would have been to run bioinformatics pipelines directly on AWS. However, over the years he had found this approach to be neither user-friendly for bench scientists, nor the most cost-effective from a cloud resourcing perspective. As a result, their scientists were still fully reliant on bioinformatics personnel to run routine analyses. In parallel, the bioinformatics team was drowning in analysis requests from R&D and unable to allocate time to solving complex problems and designing new custom pipelines. They wanted the benefits of a SaaS platform without losing control of the cloud resources being used to store and analyze their data. Eventually, he was recommended by a colleague to try Basepair, a next-generation cloud-based NGS analysis and visualization platform designed with both the bench scientist and the bioinformatician in mind.

During the evaluation period with Basepair, Sombeet was pleased that instead of having to sit through a 3 or 4-day training course, the bench scientists only needed an hour or two of training to learn how to run the approved pipelines that had been built and deployed by the bioinformatics team. After training, they were able to interpret results using the built-in interactive visualizations and reports. Being able to simply plug their own cloud account into Basepair allowed Nkarta to leverage the existing compute and storage resources in their own cloud account to analyze and store their data. This meant they could benefit from the economies of scale with their cloud provider. Finally, Basepair's command line interface (CLI) meant the bioinformaticians could interact with the platform in a way that was more conducive to the way they worked, while the APIs greatly reduced the level of effort needed to get data in and out of the platform. This would enable them to focus their time on more valuable tasks such as the use of data lakes to integrate genomic data sets and subsequently, leverage machine learning (ML) and artificial intelligence (AI) approaches to deliver scientific insights.



What is the mission of a bioinformatician?

Associate Director of Bioinformatics, Sombeet Sahu, believes that by providing pipeline automation through Basepair he can free his team from doing the repetitive analysis work on which most bioinformatics spend most of their time. Instead, they're able to work on building more advanced pipelines which accelerates time to scientific insight.

Basepair is not meant to replace bioinformaticians but to remove the burden of "busy work" often left for them. While growing a team and adding resources will help, it is not enough nor is it scalable. ***By implementing Basepair as an NGS analysis platform they give their bioinformaticians and bench scientists superpowers and allow them to use their exceptional capabilities on projects that are worthy of their time.***

To learn more,
visit our website.

Sombeet Sahu
Associate Director of Bioinformatics