World's fastest supercomputer speeds up genomic analysis

It has been somewhat difficult to ascertain what problems the fastest supercomputer on the planet has been chewing on since it was announced in 2013, but there are some signs that China is now pinning the machine's mission on the future of genomics, among other areas.

A team working on the Chinese supercomputer was able to achieve a 45x speedup on a single node of the system without a loss in precision by refining their approach to parallelization of a critical part of the genomic analysis pipeline. By revamping how a commonly used SNP detection framework shares the load via the team's mSNP framework, they could take this single node performance and scale it to just over 4,000 nodes of the Xeon Phi-boosted super.

The existing tool is called SOAPsnp, which the team says took more than one full week to analyze for one human genome with 20-fold coverage. To put the critical nature of this step in larger genomics workflow into context, consider the role of SNP detection in the future of medicine. The single nucleotide polymorphism (SNP) is the genetic equivalent of a bit flip, a spot in the DNA sequence where variation can be spotted. These are useful to identify in sequences since they can pinpoint vulnerabilities to certain diseases, map more targeted pharmaceutical routes, and highlight other genetic markers of importance. And these are not few and far between—there are several million SNPs that have been identified in the human genome alone.

The GLP aggregated and excerpted this blog/article to reflect the diversity of news, opinion and analysis. Read full, original post: China's top supercomputer dramatically speeds genomics pipeline