Exascale systems, such as CERN's EOS, may suffer from poor data layout. They contain many disks that potentially span over multiple generations and locations. Large distributed systems serve constant traffic from multiple users that do not have similar hardware or data demands. Different nodes in a distributed system may offer varying levels of performance. In the case workloads shift rapidly towards a piece of data located on a slower node, that node will not respond to all the request as fast as the previous node thus causing the sustained IO throughput of the system to decrease unless the data can be moved to a faster node. Exascale systems also see frequent hardware updates and failures which will modify system configuration. Heuristics rely on the system's configurations to place data depending on the usage. To update a system's configuration, a system administrator will need to bring down the system so that they can reconfigure it.

Geomancy modifies the layout of the data to increase the throughput of a workload previously observed on the system without human intervention. It uses a deep neural network to find trends in data access patterns and predict future throughput values of future accesses. The neural network has one LSTM layer and seven convolutional layers. The LSTM layer relates newer features to past values and the convolutional layer builds a sparse representation of the data, which allows the neural network to disentangle the information present and build a clearer representation of which action caused a certain state. Detecting the access patterns allows Geomancy to calculate how accesses may change in the future and identify which node might get overwhelmed by requests from users. This allows Geomancy to preemptively move data before any bottlenecks happens or when hardware gets added/removed from the system.

When hardware changes happen, Geomancy takes care of integrating the new hardware or removing the broken hardware in its data placement layouts with no human intervention. As long as it is given access to the new hardware it is be able to integrate in on the fly with little to no system down time. If a hardware fails, static algorithms might not know that it failed and requires system down time to reconfigure the data layout. Geomancy ignores any failed drives and routes data to active drives.

Data accesses can be represented a large number of features, such as the name of the file or the amount of bytes being accessed by the system, and not all of them affect the performance of the access. The CERN EOS access logs uses 32 features to describe each access. Selecting the wrong features to use to represent data accesses will have detrimental effects on the performance of the system accesses, such as lower IO throughput, and will confuse any understanding of how the throughput varies at each access.

Using CERN EOS performance logs, we found five features that were highly correlated with the throughput of the system. One of those features is the timestamp of when a file is opened or closed. These features affect the throughput since if there is a drop in the throughput when a file is opened, it indicates that other files in the same location are being accessed at the same time. We used those features in our experiment on the PNNL servers.

Pacific Northwest National Lab (PNNL) provided us a scientific simulation that accessed 24 files. The node that this simulation runs on has a home NFS mount, a temporary RAID 1 mount, a RAID 5 mount, and an externally mounted Dell st2000dl HDD. The RAID5 mount has the highest performance in I/O throughput while the externally mounted HDD has the lowest.

Using the PNNL system, we compared Geomancy's placement policy against having all the data on each mount and to heuristics. The tested heuristics are having the most accessed or larger data on the faster nodes, or spread the files randomly across the available mounts. Geomancy's data placement policy outperformed placing all files on one mount, and did as well as the other heuristics without having to move as many files. Unnecessary movement of files, as experienced with heuristic data layouts, added unneeded congestion to the system. Geomancy was able to get a 49% speedup over the original PNNL data placement, which had all the data on the home mount.

Geomancy can achieve the highest throughput of 5.65 GB/s, compared to the three other heuristics and placing all the files on one mount, with reduced data movement. In the future, when Geomancy will move the data live, it will have a reduced overhead caused by moving files from one node to another compared to a placement that mapped the most accessed files over all the available nodes. By moving less files, Geomancy needs to keep track of a lower number of files, thus reducing overhead when improving performance.