Self-Organizing Raspberry Pi Clusters

Tanaka Khondowe ‘22, George Kokalas ‘22, Khang Huynh ‘23
Mentor: Richard A. Brown

Abstract

- The Raspberry Pi single-board hardware architecture provides a low-cost, robust platform for learning high-performance computing. We have improved and extended the Self Organizing Cluster (SOC) system we demonstrated in 2020, in order to make learning about cluster computing more convenient and accessible at St. Olaf and elsewhere.
- The SOC consists of Raspberry Pi units and other commodity hardware components and standard Linux-based open-source software (i.e., a Beowulf cluster), and is designed for automated cluster configuration, making it possible for beginners to conveniently setup, learn from, and tear down a cluster within a single class period.

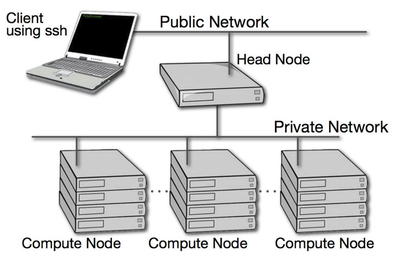
The Raspberry Pi

- The Raspberry Pi is a single-board computer that is a powerful-enough platform useful for understanding complex computer science concepts, despite its small size.

- Despite the Pi motherboard being about the size of a credit card the Pi’s hardware mirrors standard computer infrastructure in functionality. It’s size makes it convenient for building a portable computer cluster.

Beowulf Clusters

- A cluster is a group of computer networked together in order to work on shared tasks.
- A Beowulf cluster is a type of computer cluster built using off-the-shelf computers(nodes) and components. These computers are networked together into a local network. Libraries and software are installed on the cluster in order for facilitating process sharing among the cluster nodes.



A Beowulf cluster connected to a laptop

Building The Self-Organizing Cluster

- The Self-organizing cluster(SOC) was built on a layer of various networking technologies and protocols. In order to create the local Local Area Network of Raspberry Pi cluster nodes, we took advantage of the following:
- Dynamic Host Configuration Protocol (DHCP). We set up the “head” node as the DHCP server which *dynamically* assigns Internet Protocol (IP) addresses to the rest of the nodes on the cluster, the clients. This dynamic allocation of IP addresses is one point of differentiation between the SOC and other Beowulf Clusters in which each node has a fixed IP address.
- Network File System (NFS) mounting where the “head” node shares its files/directories with other nodes such that changes on one node are reflected on all the other nodes.
- Passwordless Secure Shell (SSH), a more secure form of SSH was setup between the nodes in order to facilitate the running of jobs.

The Cluster

- The SOC cluster was built to be *“plug and play”*, that is, cluster users can easily reduce or scale up the size of their clusters by simply plugging or unplugging their Pis from the Ethernet switch.

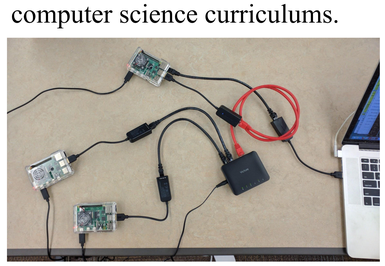
- The SOC was built from commodity-grade components: RJ45 Ethernet cables, USB-Ethernet Dongles, an Ethernet switch, power supply cables and at least two Raspberry Pi computers. The Ethernet cables and the switch establish the LAN on the most basic, physical level. Other protocol configurations on the LAN such as SSH, NFS and DHCP depend on this physical layer connection.
- The USB-to-Ethernet dongle allows for the participants to connect their Raspberry Pis to laptops and access their Pi [“head”] node via a tool called VNC. VNC shows the node’s desktop environment and users can use their laptop’s keyboard and mouse on the Pi.

Application

- The SOC’s portability and “plug-and-play” ability extends its usefulness to a variety of groups. So far, the SOC has been used for learning Parallel and Distributed Computing by participants of the 2021 CSinParallel workshop and students in the Hardware Design Course at Saint Olaf College.
- The cluster has been configured to suit the aforementioned groups’ different needs such as the need to use alternative Network interfaces for the cluster “worker nodes” for convenience and cost saving.
- Flexibility in the cluster’s infrastructure allows us to ship the technology to other institutions beyond the Saint Olaf College campus - institutions whose students could benefit immensely from learning high-performance computing concepts through sight, sound and touch.

Impact

- This Summer (2021), 37 educators from across the United States attended the CSinParallel virtual workshop. The workshop was a 4-day event which was hosted by Professor Richard Brown (Saint Olaf College), Professor Libby Shoop (Macalester College), Professor Joel Adams (Calvin University) and Professor Suzanne Matthews (United States Military Academy (Westpoint)).
- Our team shipped SOC kits containing complete cluster components which they successfully set up and used for learning, sharing and discussing modular approaches and materials for incorporating Parallel computing into their computer science curriculums.



A SOC with head node connected to laptop

A Raspberry Pi Kit

Future Plans

- Currently, the cluster uses passwordless secure shell access over RJ45, Ethernet cables which is fairly robust.
- At the same time, each node can connect to the Wide Area Network(WAN) via wifi which, in a way, undermines the security benefits of using passwordless SSH since attackers can attempt brute force attacks on the WAN especially if the wifi is open, that is, requires no password.
- In the future, we will incrementally improve the security on the cluster as well as well as the SOC’s portability.
- One way we improved the cluster was using a networking tool called iptables which enables the “worker nodes” to use the WAN connection on the “head node” such that only the “head node” is potential visible to the outside world.

Acknowledgements

- Special thanks to the CSinParallel Project who, through the NSF grant (NSF DUE-1822480) made this research possible, and St Olaf College for supporting our efforts. We would also like to thank Max Narvaez ‘21 and Anway De ‘21 for assisting with building the SOC and the cluster management training at Saint Olaf College.