**Availability of PRObE: 1000 Nodes for Systems Research Experiments**

**Presentation:**

The purpose of this Birds of Feather session was to raise awareness about the availability of PRObE, a new computing resource for large-scale systems research. Garth Gibson (CMU) introduced the PRObE platform setup and its relationship to Emulab. Currently, PRObE comprises three clusters: Kodiak (1024 nodes), Marmot (128 nodes) and Susitna (34 nodes).

Detailed information can be found here: https://www.nmc-probe.org/wiki/Machines

Currently, the platform hasn’t been fully used, which means not enough resource competition and contention. This BoF was intended to attract new researchers to apply for usage.

**Q&A Session:**

Q: Does PRObE provide network isolation, or a method of knowing current state of network? If a researcher works on a subset of nodes, what control on traffic do they have?

A: Eric Eide (University of Utah) replied that based on Emulab software setup, every experiment owns the entire physical machine and has a dedicated network. PRObE is not intended to be similar to a cloud setup, rather the experiments are intended to run on a dedicated cluster. Therefore, it generically provides isolation per experiment, and has no contention.

Q: Someone in the audience remarked that they had noticed inconsistent network performance when running an experiment on 20 machines at different times.

A: Garth Gibson, after checking out the experimental setup, replied that the experiment has not created any dedicated network topology, so by default it runs on the control network, which is shared by all users for SSH, package downloads, etc. Ideally, each experiment should create a data network for its own. If the users still experience issues, they are encouraged to contact the PRObE support.

Q: Akshat Aranya (NEC Labs) said that he had asked for a 1G ethernet for experiment, but the administrator pointed me to the control network.

A: It was pointed out to him that his experiment was on Kodiak, which does not have 1G ethernet dedicated data network. The ethernet is intended for control and infiniband is intended for data network.

Q: How should we use the network to contact machines outside the PRObE clusters?

A: Ideally, there should be no outside traffic. There is an internet connection, but that exists only to facilitate researchers to install software packages.

Q: What is the storage hardware on PRObE? How does one use it to experiment storage system?

A: It has storage hardware for anything. Each machine in Kodiak has two 1 TB SATA 7200rpm hard disks and three physical data networks to use. Once a user allocate nodes for an experiment, the user owns the entire machine hardware until the experiment is terminated. Among all three networks, the infiniband is perfect to experiment any NAS like system with allocated machines. Each node in Susitna has three 1 TB hard disks and a 64 GB SSD.

Users need to note that the data stored on each machine will be removed at experiment termination. Additionally, there is some shared storage to store experimental output and OS images, which will also be eventually revoked after the end of the experiment.

Q: Is PRObE available for institutions outside US?

A: The hardware is donated by US government, which focuses on research progress in US. The best method to use PRObE resources outside US is to have a US-based partner. Industry can also use this platform but the work has to be published, even if it is on their own website. And, for any publication, authors should cite the PRObE paper.

Q: Are there any heterogeneous clusters in the PRObE setup, comprising nodes with different hardware specifications?

A: No, all clusters are homogeneous. However, users can emulate heterogeneity by using a subset of the available resources on each node.

Q: The PRObE network seems to be flat, is it possible to do routing experiments on it by emulating rich networks?

A: Due to the support of Emulab, users can program any topology and routing policy by themselves.

Q: If someone needs to run an experiment for big data analysis, like Genetics Science research, how can they put a large amount of data onto PRObE?

A: Andree Jacobson (New Mexico Consortium) replied that the PRObE is meant to facilitate systems research. That being said, you can potentially bring your own storage box by approval and directly plug it into the network.

Q: Can I change the low-level routing policy on PRObE?

A: Garth replied that PRObE is not designed to allow users to change the infrastructure environment. Researchers should create experiment on top of the infrastructure by using Emulab. Andree added that if they have an extra switch, they can possibly set up another gigabit ethernet with their choice of network topology.

Q: Can I control the machine placement in an experiment? For example, can I explicitly ask machines belonging to the same rack?

A: Eric Eide replied that there is no direct control for this. But, in fact, Emulab always tries to minimize resource usage, so if you need two machines and a network linking them, it probably end up with two machines connected to the same direct router. Another person in audience confirmed that you can basically get what you want by designing a network topology.

Q: Is it possible to do low-level firmware development?

A: Emulab gives access to serial consoles, which permits logging in to the BIOS and doing whatever you want.

Q: Akshat Aranya commented that Emulab can do a whole lot of things, but PRObE cannot do some of them, and the current documentation is not clear.

A: Eric Eide stated that “Emulab” has different meanings, and is used to refer to both the software, and the network testbed at the University of Utah. Both of these have slightly different functionality, which is often conflated. Andree pointed out the existence of a user forum and a wiki, and encouraged people to add their experiences in wiki and discuss their problems in the forum.

Q: If we need to test kernels that are built up from scratch (not customized from existing image), how can we test it on PRObE?

A: It is operator's job to verify their setup. For Kodiak, it is highly recommended that users start with one of the preconfigured OS images.

Q: Mohammed G. Khatib (HGST Research) asked if there is any policy for experiments that can help discover weaknesses in the hardware?

A: Garth replied that the PRObE group will be happy if there is any research aiming to make the environment last longer. But it might need a special proposal for the approval of the administrator (Andree).

Q: What is the policy for multi-cluster usage if we need to generate traffic among them?

A: Different clusters are currently deployed at different locations, such usage is not encouraged because the cross-cluster traffic costs a lot of money. There are more appropriate platforms for wide-area network traffic such as PlanetLab. If users still want to use PRObE for geo-distributed experiments, they can simulate the topology even within a single cluster, albeit with low latencies.