Recitation 7: VM and Malloclab (pt.1)

Friday, February 28th

Introduction

The goal of this activity is to gain a better understanding of TLB benefits, using the lens of locality of access, size of working set, and frequency of context switches! Each group should select one of the two scenarios and make an analysis of the implications that the scenario has on the TLB or page activity.

Scenarios

1. **Large Data Computing**: We are running a large-data computation task, processing data on the magnitude of terabytes. Suppose we have a reasonably good, regular access pattern to data, as well as a reasonable page size (eg. 4KB).
2. **Virtualized Cloud Environment**: We are a virtualized cloud environment like AWS, supporting multiple virtual machines at the same time (think of virtual machines as its own process). Assume all of these processes map onto the same physical machine.

Now let's discuss the features of each workload. We simply present them for the purpose of this recitation, but make sure to briefly discuss in your group about why these features are true!

Scenario 1: Good spatial locality from accessing, but huge working set across the lifetime of the program. This is all running on one process so there are minimal context switches.

Scenario 2: Perhaps good locality within each task (or VM), but no virtual pages are shared across each task - showing low locality. The working set is reasonable within each task, but can get very large across the set of all tasks being served. In order to serve all of these tasks, we will often need to change the process we are currently working on - lots of context switches.

Questions to Answer

Using the categories of locality of access, size of working set, and frequency of context switches, try to see how each scenario does with respect to each category. Think about:

1. Is there good locality in the scenario? What is the total size of the working set (large or small)? How often do we context switch (are there a lot of independent tasks to be processed)?
2. What do the observations in part 1 indicate about TLB behavior? What implications do they have?
3. What changes can we make to improve the benefits we gain from TLB or avoid the pitfalls of TLB? (eg. cache design, page size, ect…)

It may be helpful to organize your thought in the table provided:

Factor	**Does the TLB Help?**	**Why or Why Not?**	**What are some improvements we can make?**
Locality of Access			
Size of Working Set			
Frequency of Context Switches			