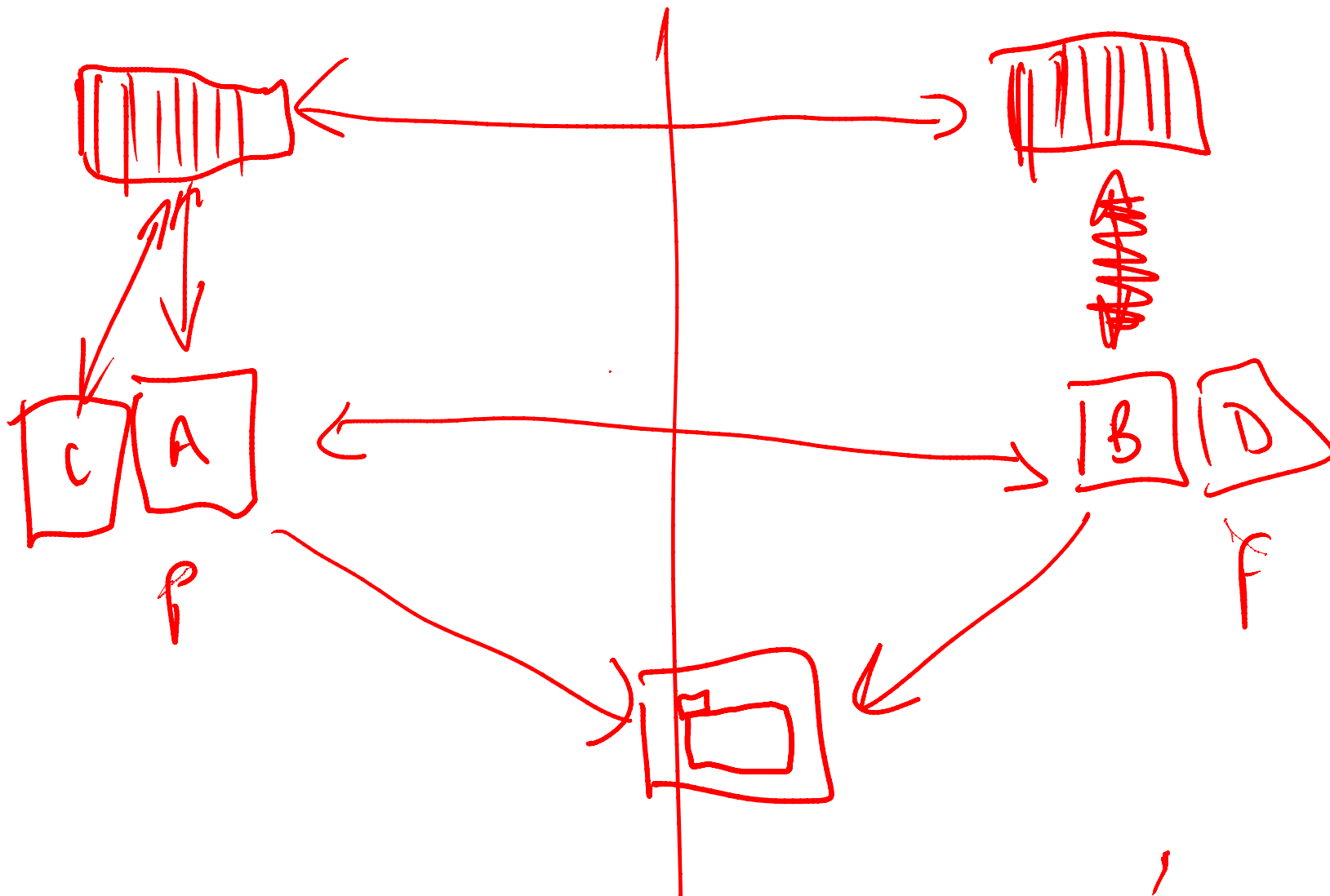
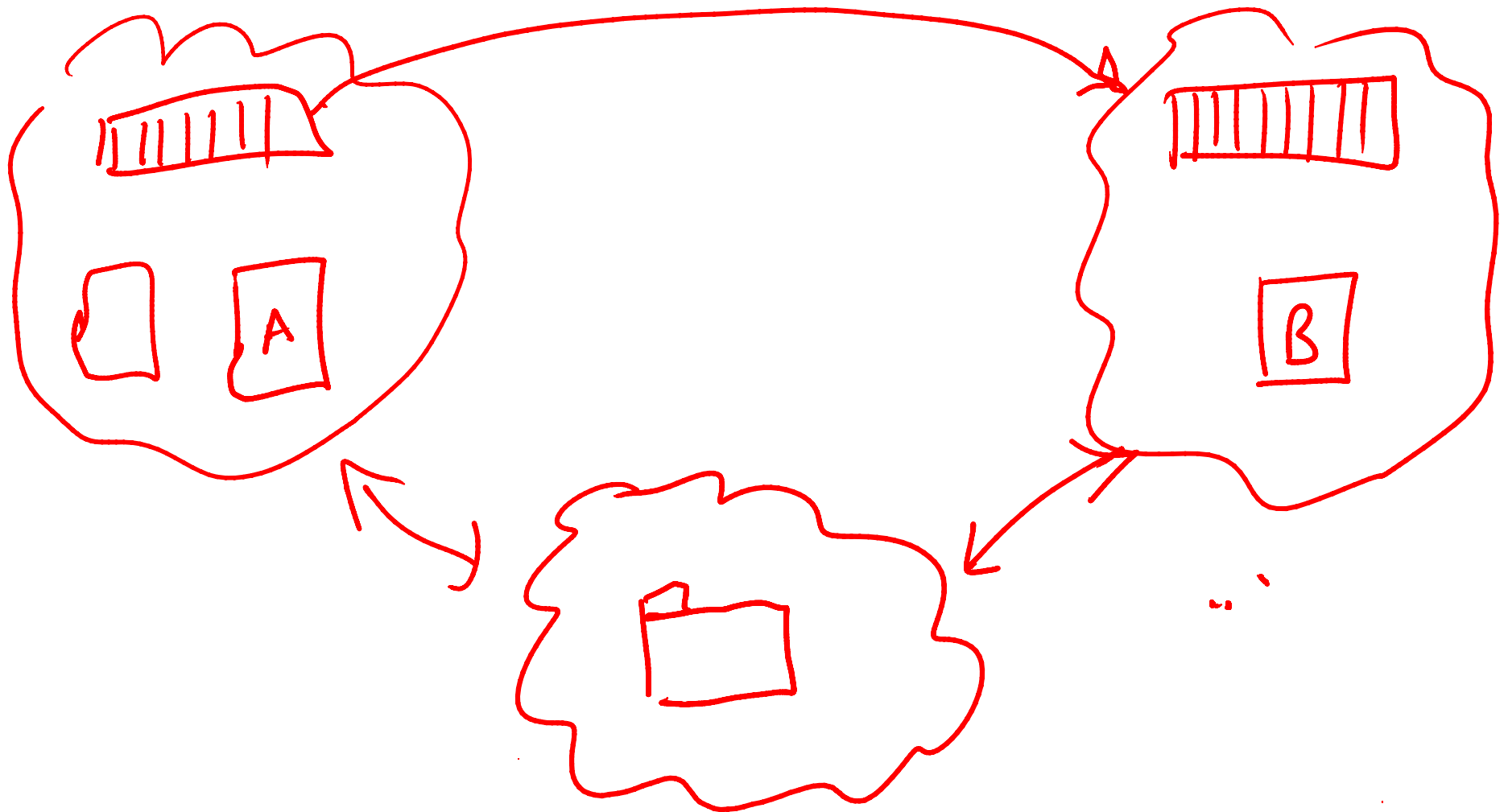


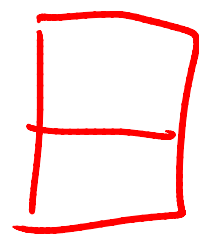
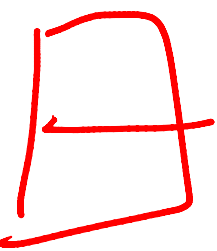


Geo-Clustering requirements. This slide shows the required components to setup a geo-cluster for SQL Server. The use of multiple nodes per site is possible, but the file share for quorum needs to be centrally located where all nodes have access to support a majority of voters in the event of a site failure.



Geo-Clustering requirements. This slide shows the problem of having an odd number of nodes in a geo-cluster configuration. If the site on the left goes down, the cluster will not have a majority of voters in the configuration and will not failover in the remote location.

128GB

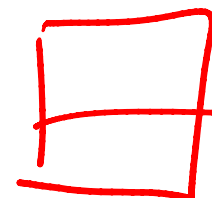
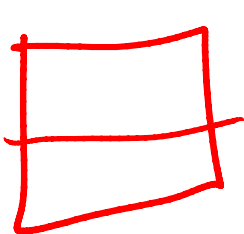


128GB 48 cores

512GB RAM

4-12 core
MagnyCours

128GB



128GB

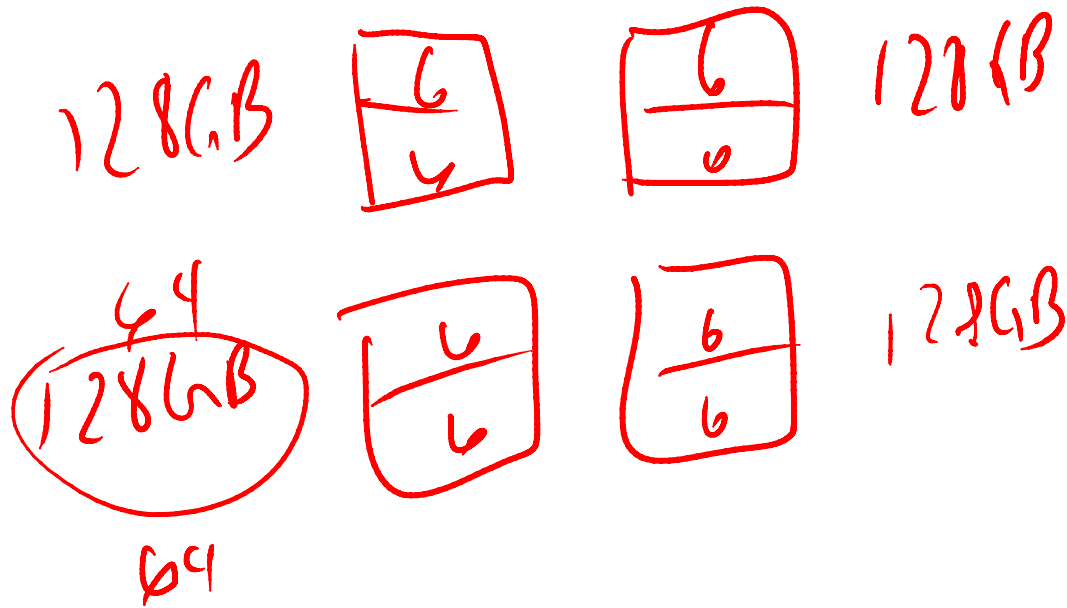
8 NUMA Nodes

480GB Buffer Cache

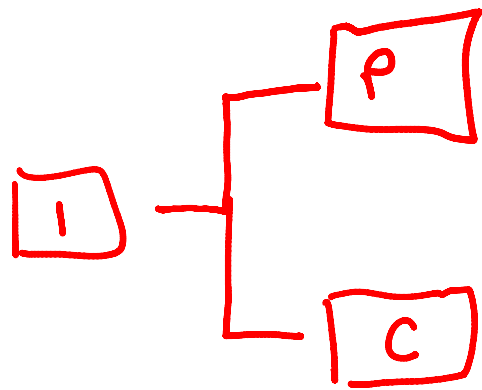
60GB

Explanation of NUMA using an example system of 4 AMD Magny-Cours 12 core processors and 512GB RAM. The Magny-Cours processors have dual memory controllers and present 2 NUMA nodes per socket. The SQL Buffer pool is therefore divided amongst 8 LazyWriters, and with a 'max server memory' configuration of 480GB, each would have 60GB to manage.

VMware NUMA



Vmware NUMA optimizations are only enabled if the VM is smaller than a single NUMA node in ESX 4.1 and lower. If the host has 4 AMD 12 core processors and 512GB of RAM, the largest VM size that would have the NUMA optimizations enabled is 6 vCPUs and 64GB RAM. If a VM exceeds this size, it would not have the NUMA optimizations in the hypervisor enabled.



DOP 8

17 workers

8 Producer
8 Consumer
1 Control

17

Parallelism

48 Producers 6
48 Consumers 6
1 Control 1

97 threads 13 threads

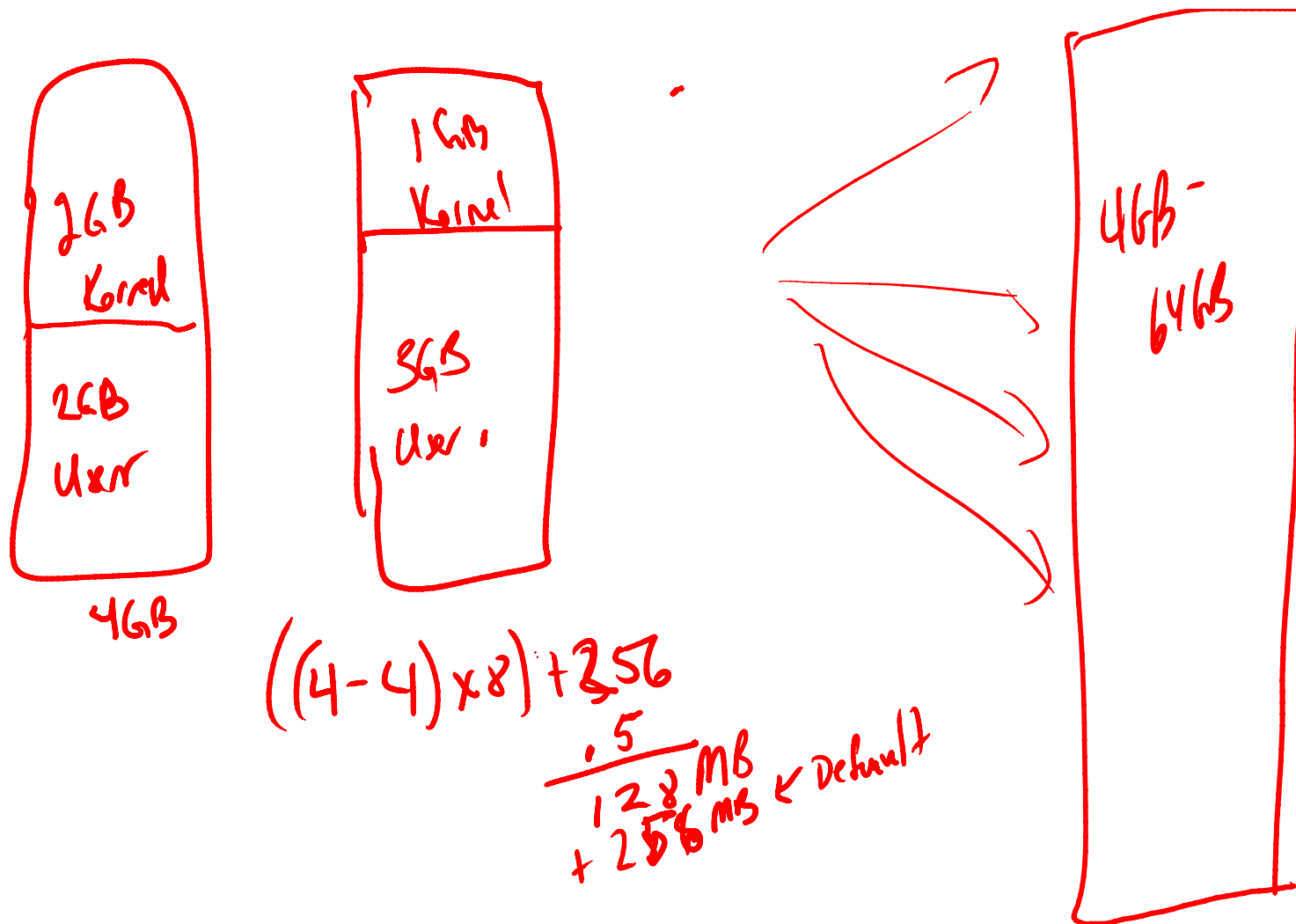
$$\left((48 - 4) \times 16 \right) + 512$$

264

44

$$\frac{44}{104} \times 12 = 1216$$

Worker thread usage under parallelism is cumulative and can lead to Threadpool starvation issues if the default configuration is used for servers with a large number of CPUs. Each parallel task will use 2 threads (1 producer, 1 consumer) per level of DOP and 1 control thread for the parent task. At DOP 8 this equates to 17 workers per execution. At DOP 0 on a 48 core server this equates to 97 threads per execution. A server with 48 cores would have a total of 1216 worker threads by default allowing just under 13 concurrent executions.



32-bit systems have 4GB of total VAS that is divided into 2GB kernel mode and 2GB user mode by default. If the /3GB switch is used in the boot.ini, this shifts to 1GB kernel mode and 3GB user mode, but limits the maximum memory available on the server to 16GB due to reduced page table entries (PTEs) in the kernel. PTEs are used to map the AWE allocated memory above 4GB. The VAS Reservation (aka MemToLeave) is calculated as the worker thread stack (max worker threads * .5MB) + startup reservation amount (256MB by default or -g startup parameter value if used)

128 MB Thread Slack

256 MB Default Reservation

384 MB VAS Reservation

⁹
2048 MB

- 384 MB

1664 MB

3072 MB

384

2688 MB

Sizing the buffer pool on 32-bit instances that are not using AWE. Without /3GB the max buffer pool size is 1664MB. Using /3GB it is 2688MB, which is why /3GB was so popular on 32-bit instances.

$$(\text{Installed RAM} / 4) * 300$$

•

How to calculate a more effective Page Life Expectancy value instead of using the traditional 300 value that is dated and based on significantly different hardware than currently available. As pointed out in class it should be $(\text{BufferPoolSize}/4)*300$, but in reality, the appropriate value should be based on the baseline of the server, rather than a fixed value.