Home Work 4

Part A

6.5 Ben measures the performance of RPC messages that are small in size but the improvement might not be uniform for all the RPC calls. The effect of faster network might be more for RPCs greater in size. So, he should first monitor which type of RPC calls are most frequent. Also, faster network reduces just the transit time which might not be the only bottleneck of system. Other processes involved in issuing an RPC starting from system call processing to marshalling to undoing everything on the other end adds up to the overall latency. So, he should consider improving these parameters as well.

6.6 Increasing the page size decreases performance due to more memory being wasted by the partially filled pages. Also, since there would be lesser pages stored in the fast memory at a time, number of misses would increases again affecting the performance. But at the same time, as the size of the page increases, it allows much more virtual space for each application. This allows the pages to map to larger blocks in the physical memory and thus reduces the disk access times since the data would be now more contiguous in the physical memory. The smaller the page size, the more fragments will the data be divided in which thus increases the average seek time from the disk. Thus, increasing the page size would hurt temporal locality because of the misses but would benefit spatial locality at the same time.

Question 3.

Average seek latency= 8ms

RPM= 7200 Rev/minute = 120 Rev/sec or 1 revolution every 8.33ms.

Avg. rotational latency = 8.33/2 = 4.17ms

Data in each track = 1.5MB

Data in each sector = 4KB

Data rate at head = 1.5MB*120 Rev/sec= 180Mbps

(i) Average Latency for Read of a sector = Avg seek time + Avg rotational latency + Transmission of 4Kb

= 8ms + 4.17ms + (4KB / (180*1024)Kbps) * 1000ms

= 12.19ms

(ii) Average Throughput for Read of a sector = (1/Avg Latency) * 4KB

= (1/12.19e-3) * 4KB

= 328Kbps

(iii) Average Latency for Read of a track = Avg seek time + Rotational latency for one complete track

= 8ms + 8.33ms

= 16.33ms

(iv) Average Throughput for Read of a track= (1/Avg Latency) * 1.5MB

= (1/16.33e-3) * (1.5MB)

= 91.86 Mbps

(v) Server Utilization = Arrival rate * Service time
0.8 = Arrival rate * 12.19
Arrival rate = 65.63 s^{-1}

(vi) Time of queue= μ (ρ/1- ρ)= 12.19 * (0.8/0.2) = 48.76 ms
Average length of the queue at 80% utilization = Arrival rate * Time of queue
= 65.63 * 48.76e-3
= 3.2 requests