

**Assignment 1** Amdahl's law states the maximum available speedup of a program which has proportion  $P$  of parallel code and  $1 - P$  of sequential code, and  $N$  processors working in parallel. Compute the speedup of code with  $N = 16$ .

- a) In the case  $P = 75\%$ . (1p)
- b) In the case  $P = 85\%$ . (1p)
- c) In the case  $P = 95\%$ . (1p)

**Assignment 2** In this question the task is to do calculations of mean time to failure (MTTF) of hypothetical storage configurations. In the following assume a mean time to failure (MTTF) of a hard disk to be 250 000 hours. Also assume hard drive failures form a Poisson process, i.e., they are independent and the time between failures follows an exponential distribution.

When using 2 TB hard disks in RAID 5 configuration ( $n$  disks for data + 1 disk for parity), compute the mean time to failure in years of a RAID 5 array that consist of the following amounts of storage space. To simplify calculations, we (pessimistically!) assume that failed hard disks are left unrepaired:

- a) 10 TB (1p)
- b) 20 TB (1p)
- c) 40 TB (1p)

Hint: Use the minimum number of disks, e.g., 10 TB will be obtained using 5 data disks + one parity disk. The array fails if two hard disks fail.

**Assignment 3** Briefly (using at most half a page of text maximum) describe what is the statement called as the Brewer's CAP theorem, what implications does it have for tradeoffs in distributed systems design, and give some examples of practical distributed software systems exploiting these tradeoffs to achieve the different goals achievable by a distributed system. (4p)

**Assignment 4** Bloom filters are a probabilistic data structure for storing sets of items. Consider the case of a Bloom filter with 4 megabytes of memory, where we would like to insert at most  $n = 3900000$  items. What is the approximate optimal number of hash functions  $k$  to minimize the number of false positives? What is the false positive probability with using that  $k$  (rounded to the nearest integer) after having inserted  $n$  unique items? (4p)

**Note! More assignments on the other side of the paper.**

---

The name of the course, the course code, the date, your name, your student id, and your signature must appear on every sheet of your answers.

Please give feedback through the course Noppa page link "Course feedback".

**Assignment 5** Please briefly (using maximum of two sentences for each case) define what are the following concepts as used in the course lectures:

- a) BASE (1p)
- b) RAID 10 (1p)
- c) RAID Write hole problem (1p)
- d) Consistent hashing (1p)
- e) Scaling out (1p)
- f) Kryder's law (1p)

**Assignment 6** Briefly (using at most half a page of text maximum) describe the asynchronous consensus problem and the FLP Theorem that is a central result on the asynchronous consensus problem. Also very briefly discuss how does the Paxos algorithm relate to the FLP Theorem.

(4p)

---

The name of the course, the course code, the date, your name, your student id, and your signature must appear on every sheet of your answers.

Please give feedback through the course Noppa page link "Course feedback".