Replication

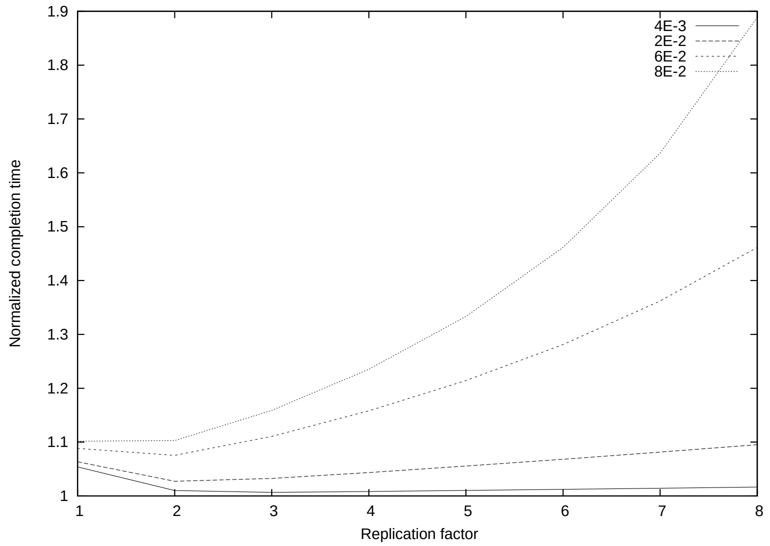
- Experimental study in [Kondo et al. 2002]
- Replication factor n, no keepalive, synchronous

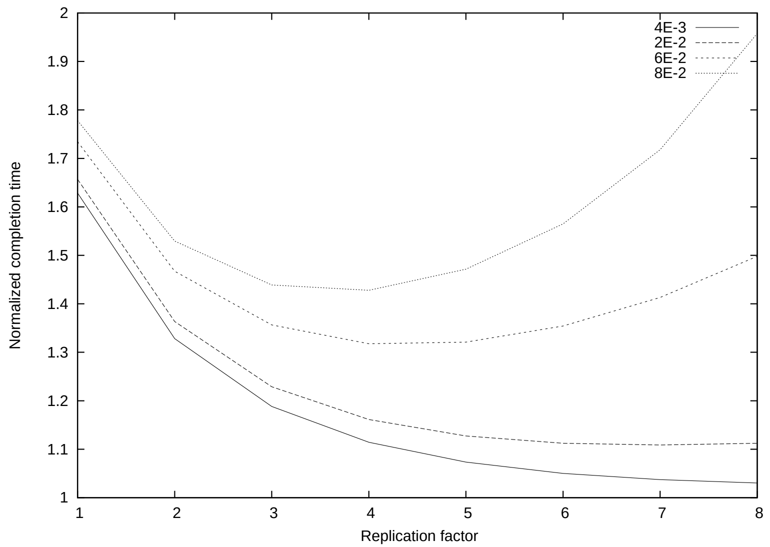
$$\overline{T_n} = \frac{1}{\lambda(1-(1-e^{-\lambda})^n)} \sum_{k=1}^{n} \frac{(1-e^{-\lambda})^k}{k}$$

- Infinite ressource : the optimal latency can be closely approximated

$$\lim_{n \to \infty} T_n = 1$$

- Completion time
  - M/G/1 queue, service time $\overline{T_n}$
  - $n = 2$ or 3 is a reasonable tradeoff





The Cost Model

- A job is falsified with unknown probability p - the defect rate
- p_a the acceptable defect rate
- Non-productive system, $p \geq p_a$: the cost is the number of calls to the oracle $s(p)$
- Productive system, $p < p_a$: the cost is number of calls to the oracle plus erroneous rejections

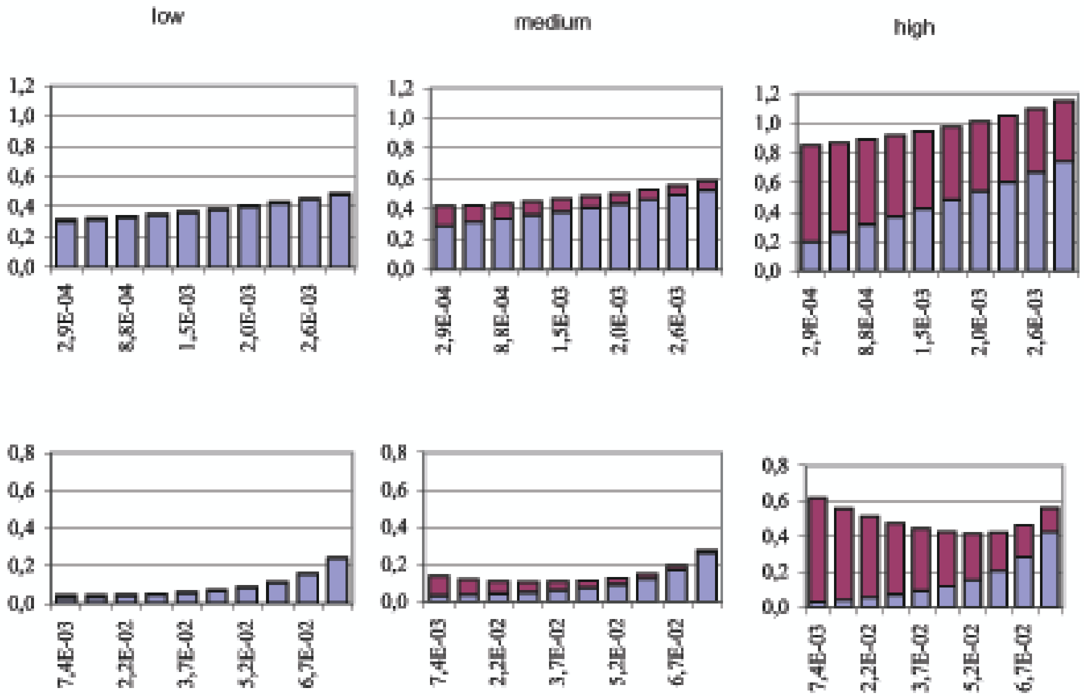
$$C(p) = \frac{1}{B}\left(s(p) + r(p)P_p(\text{reject})\right)$$

- The average cost for a productive system is

$$\mathcal{C}_a = E_Q[C(p)|p \leq p_a]$$

Checking a Batch (1) The simple sequential test

- Acceptable performance for the Normal case
  - The cost is dominated by the number of calls to the oracle

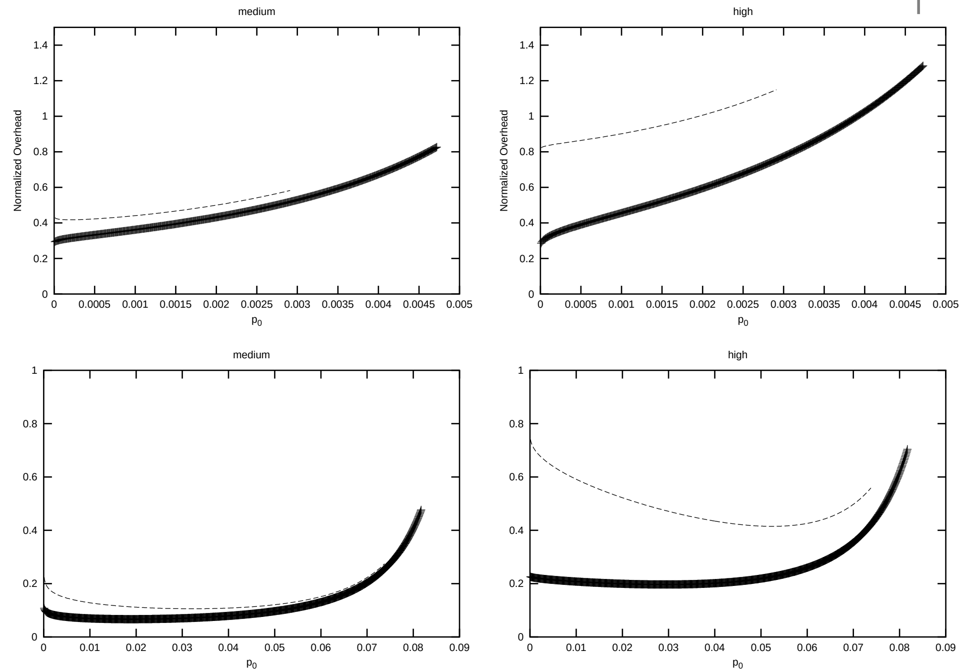


- Vulnerable to denial of service
  - With a uniform distribution of p, the decision of rejection should be reached earlier

Cost breakdown. $\alpha = \beta = 0.05$, $B = 1000$, $p_a = 0.1$ (lower), $p_a = 0.01$ (upper), $\lambda = 0.01, 0.1, 1$

The two-phases test

- Algorithm
  - First step : sense the system - time-sequential test
  - Second step : Test the whole batch
- *Normal and Massive* : Same cost
- *Denial of service* : Much more robust



$\alpha = \beta = 0.05$, $B = 1000$,
$p_a = 0.1$ (lower), $p_a = 0.01$ (upper),
$\lambda = 0.1, 1$