Time, Clocks,

and the Ordering of Events in a Distributed System

a paper by Leslie Lamport

David Murray (@fernomac)

PWL @ Seattle, 2020-10-01

Operating Systems

R. Stockton Gaines
Editor

Time, Clocks, and the Ordering of Events in a Distributed System

Leslie Lamport
Massachusetts Computer Associates, Inc.

Network Working Group
RFC # 677
NIC # 31507

Paul R. Johnson (BBN-TENEX)
Robert H. Thomas (BBN-TENEX)
January 27, 1975

The Maintenance of Duplicate Databases

Preface:

This RFC is a working paper on the problem of maintaining duplicated databases in an ARPA-like network. It briefly discusses the general duplicate database problem, and then outlines in some detail a solution for a particular type of duplicate database. The concepts developed here were used in the design of the User Identification Database for the TIP user authentication and accounting system. We believe that these concepts are generally applicable to distributed database problems.

Operating Systems

R. Stockton Gaines
Editor

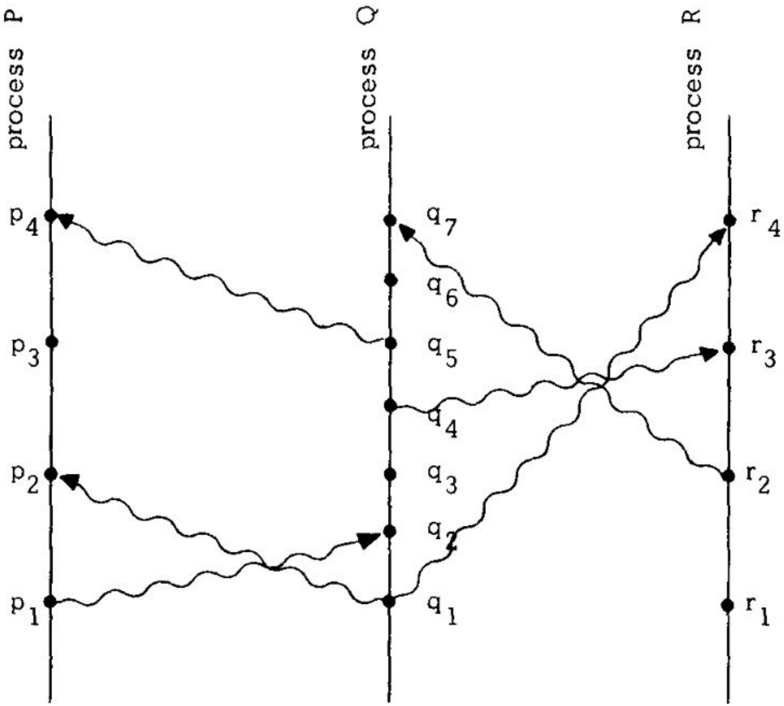
Time, Clocks, and the Ordering of Events in a Distributed System

Leslie Lamport
Massachusetts Computer Associates, Inc.

Partial Ordering

Partial Ordering

Fig. 1.



Partial Ordering

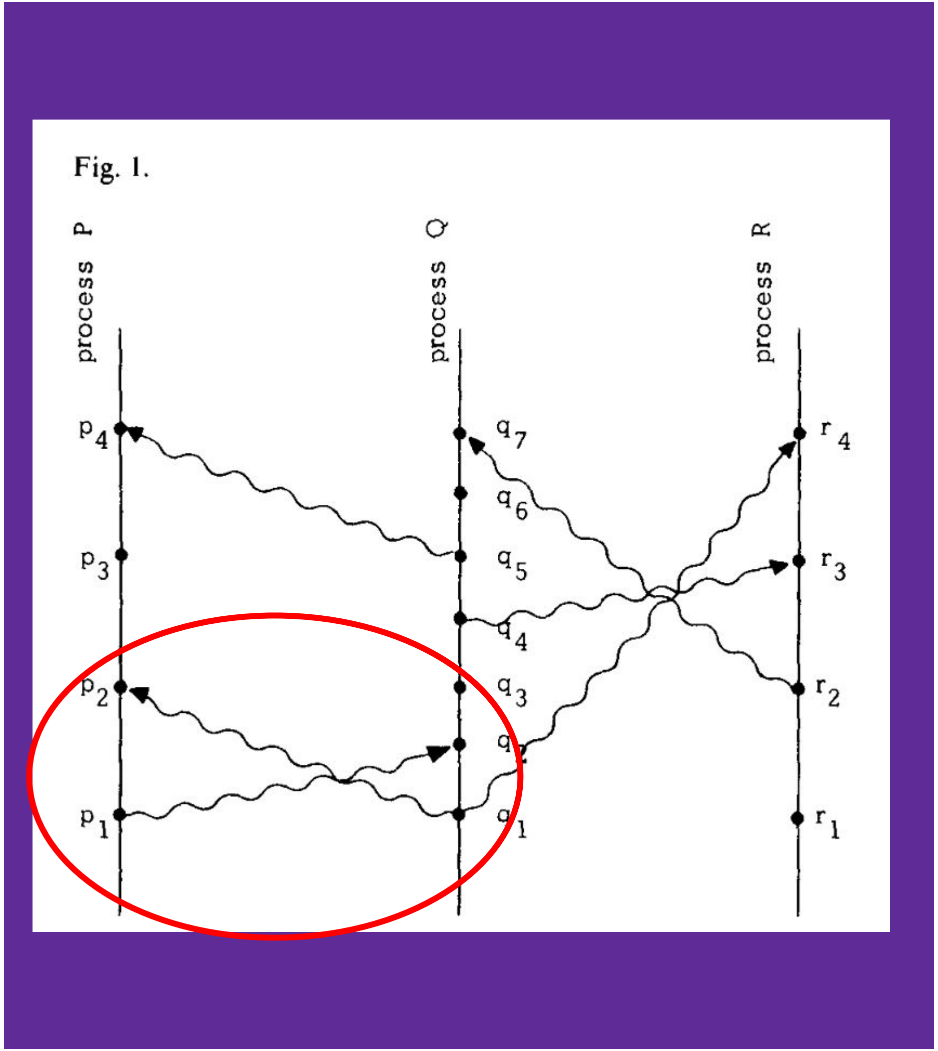
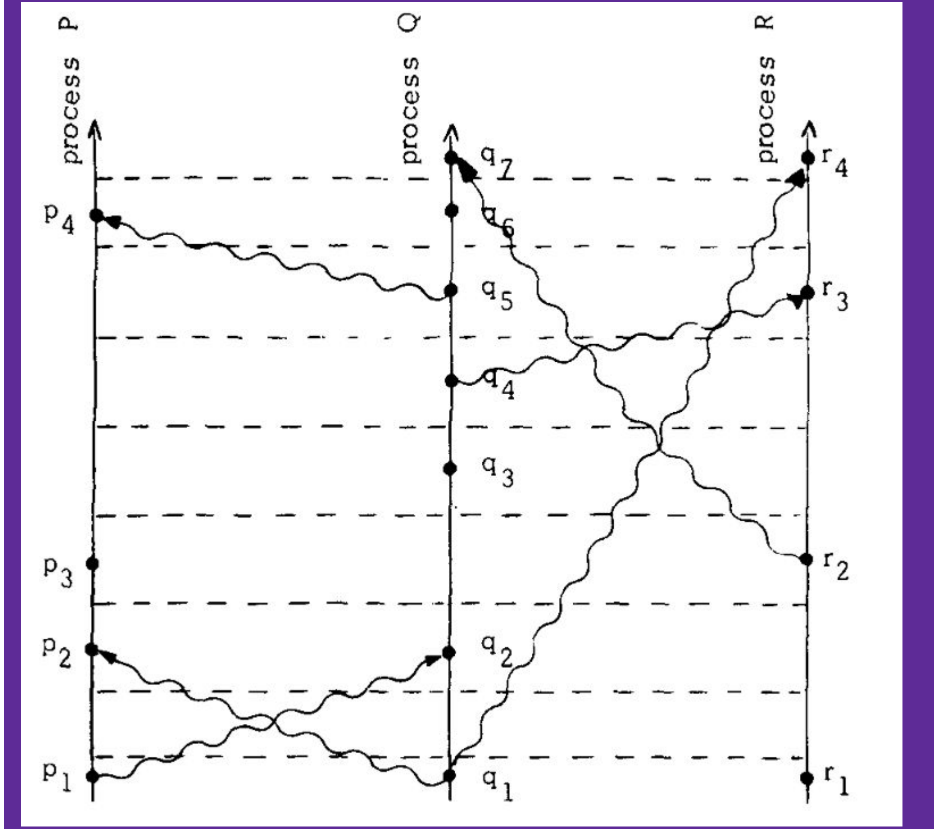


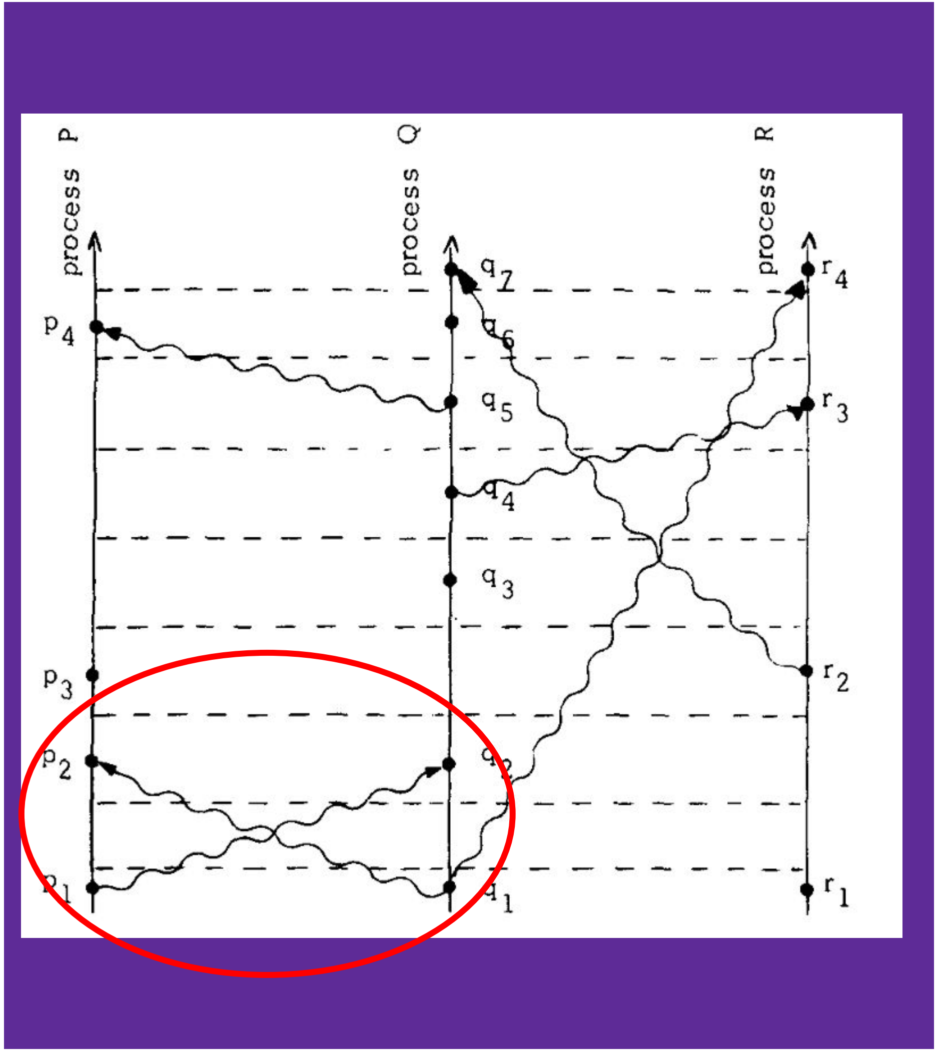
Fig. 1.

Logical Clocks

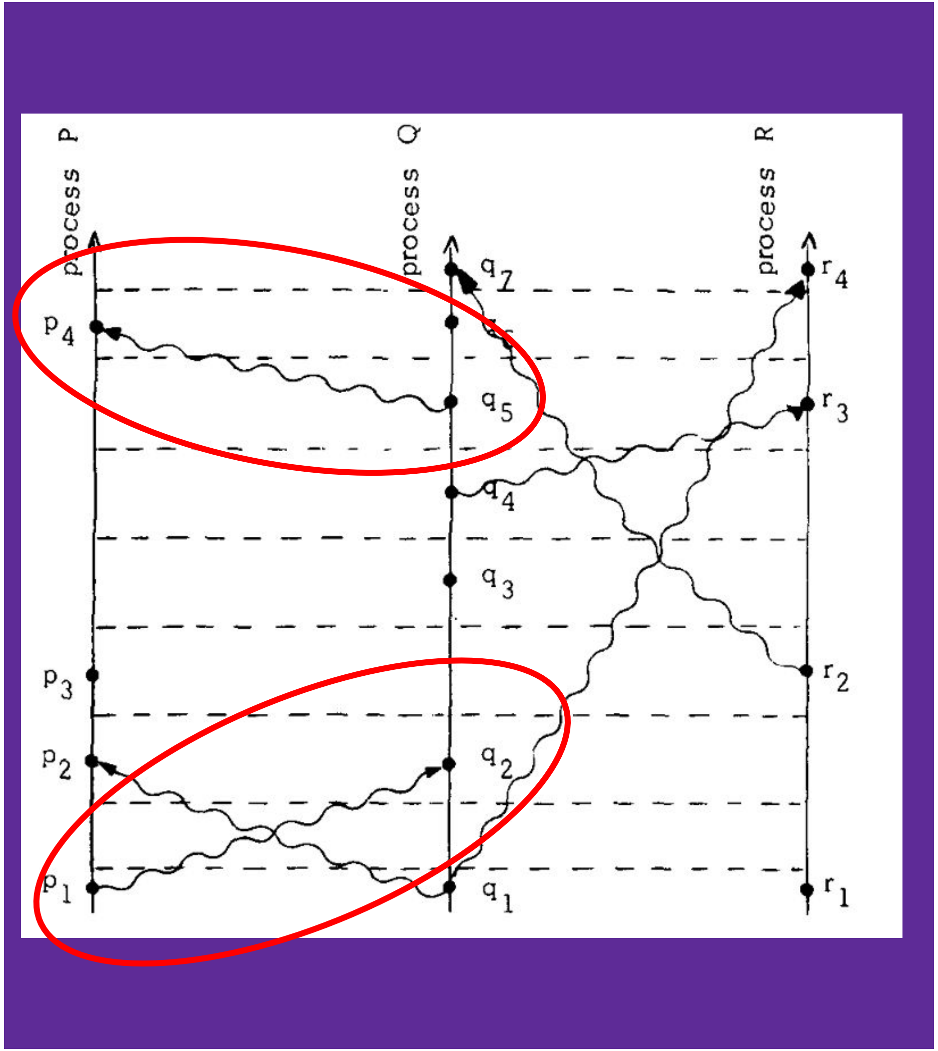
Logical Clocks



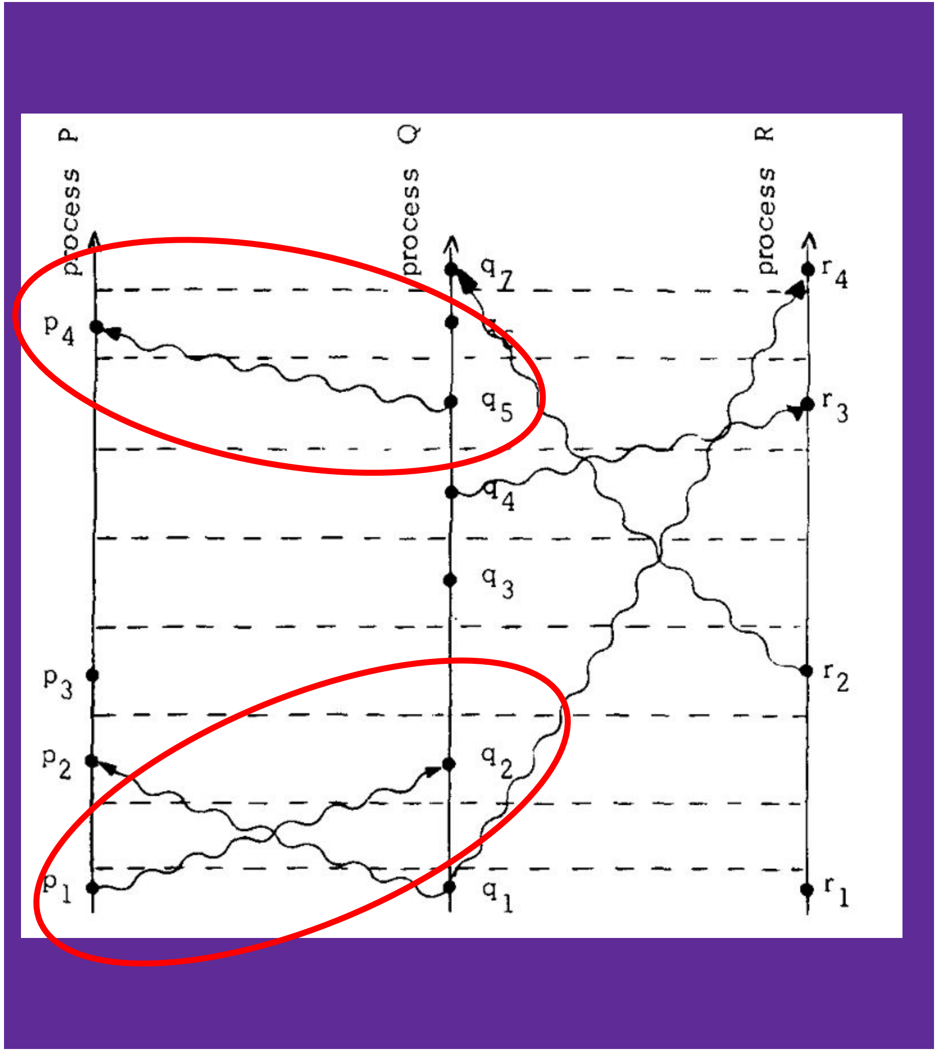
Logical Clocks



Logical Clocks

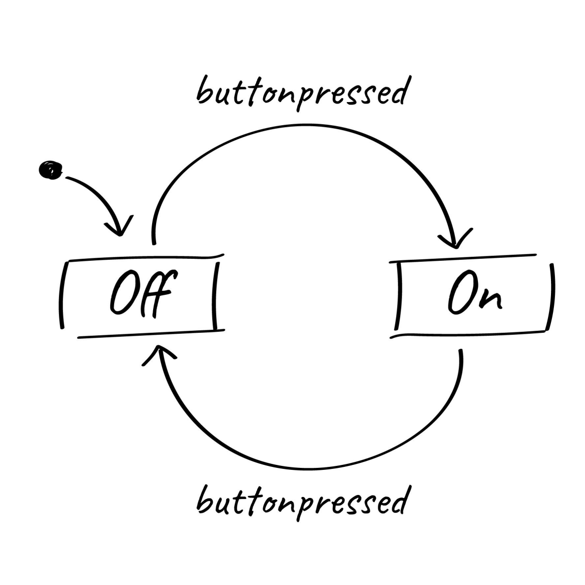


Logical Clocks

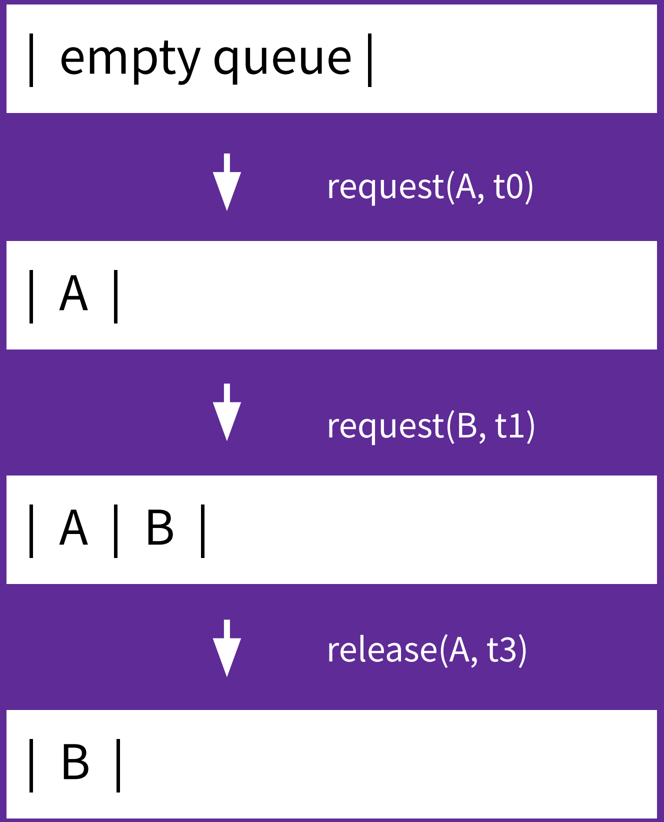


State Machines

State Machines



State Machines



| empty queue |

request(A, t0)

| A |

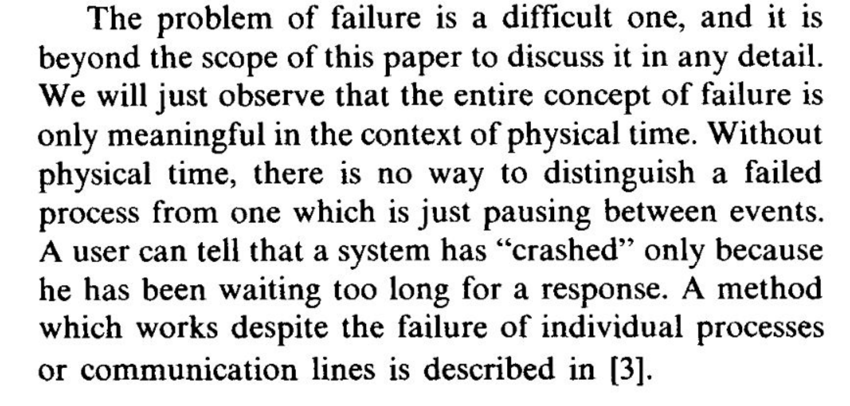
request(B, t1)

| A | B |

release(A, t3)

| B |

State Machines



The problem of failure is a difficult one, and it is beyond the scope of this paper to discuss it in any detail. We will just observe that the entire concept of failure is only meaningful in the context of physical time. Without physical time, there is no way to distinguish a failed process from one which is just pausing between events. A user can tell that a system has “crashed” only because he has been waiting too long for a response. A method which works despite the failure of individual processes or communication lines is described in [3].

Conclusion