On the Quality of Service of Failure Detectors

Present by :Lihua Ran

(Some of the slides are made by the authors of the paper:

Sam Toueg, Wei Chen, M.K. Aguilera)

April 18, 2002

Presentation Outline

- Introduction of QoS
- On the QoS Specification of Failure Detectors
- The Design and Analysis of a New Failure Detector Algorithm
- Configuring the Failure Detector to Satisfy QoS Requirements
- Conclusion Remarks

What is the QoS of Failure Detectors

- QoS is a specification that quantifies

  a) speed: how fast the failure detector detects actual failures.

  b) accuracy: how well it avoids false detections.

Why do we need to study the QoS of Failure Detectors

- Roughly speaking, a failure detector provides some information on which processes have crashed.
- The information, typically given in the form of a list of *suspects*, is not always up-to-date or correct.
  - A failure detector may take a long time to start suspecting a process that has crashed;
  - It may erroneously suspect a process that has not crashed (in practice this can be due to message losses and delays).

Why do we need to study the QoS of Failure Detectors (cont.)

- For asynchronous systems, failure detectors specified in terms of their eventual behavior (e.g., a process that crashed is eventually suspected.) are appropriate.
- But applications that have timing constraints require failure detectors that provide a quality of service (QoS) with some quantitative timeliness guarantees.

Presentation Outline

- Introduction of QoS
- On the QoS Specification of Failure Detectors
- The Design and Analysis of a New Failure Detector Algorithm
- Configuring the Failure Detector to Satisfy QoS Requirements
- Conclusion Remarks

The Failure Detector Model

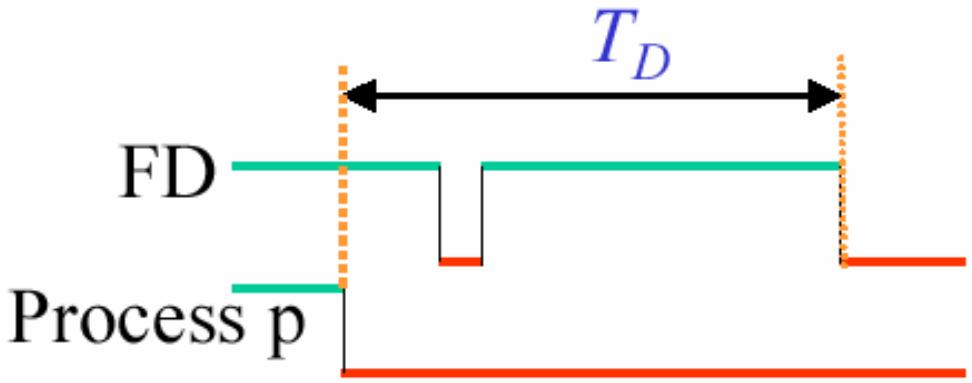
- We consider a system of two processes *p* and *q*., the failure detector at *q* monitors *p* , and *q* does not crash. Failure detector at *P* is unreliable, it can erroneously suspect a process that has not crashed.
- The output of the failure detector at *q* at time t is either *S* or *T*, which means that *q* suspects or trusts *p* at time *t*.
- Transition:

  S-transition: the output at q changes from T to S.

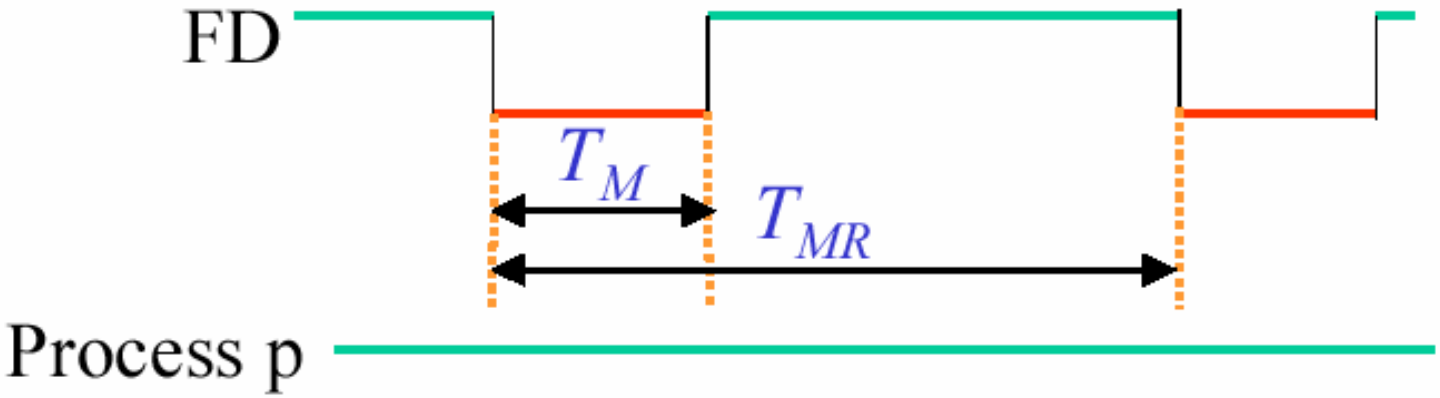
  T-transition: the output at q changes from S to T.

Primary Metrics

- Detection time T_D



- Mistake recurrence time T_{MR}
- Mistake duration T_M



Derived Metrics

- Average mistake rate(λ_M): this measures the rate at which a failure detector make mistakes.
- Query accuracy probability (P_A): the probability that the failure detector's output is correct at a random time.
- Good period duration (T_G): the length of a good period (the time that elapses from a T-transition to the next S-transition).
- Forward good period duration (T_{FG}): a random variable representing the time that elapses from a random time at which q trusts p, to the time of the next S-transition.

Relations Among Accuracy Metrics

$$T_G + T_M = T_{MR}$$

$$\lambda_M = 1 / E(T_{MR})$$

$$P_A = E(T_G) / E(T_{MR})$$

$$\Pr(T_{FG} \le x) = \frac{1}{E(T_G)} \int_0^x \Pr(T_G > y)\, dy$$

$$E(T_{FG}^k) = \frac{E(T_G^{k+1})}{(k+1)E(T_G^{k+1})}$$

$$E(T_{FG}) = \frac{E(T_G^2)}{2E(T_G^2)} = \frac{E(T_G)}{2}\left(1 + \frac{V(T_G)}{E(T_G)^2}\right)$$

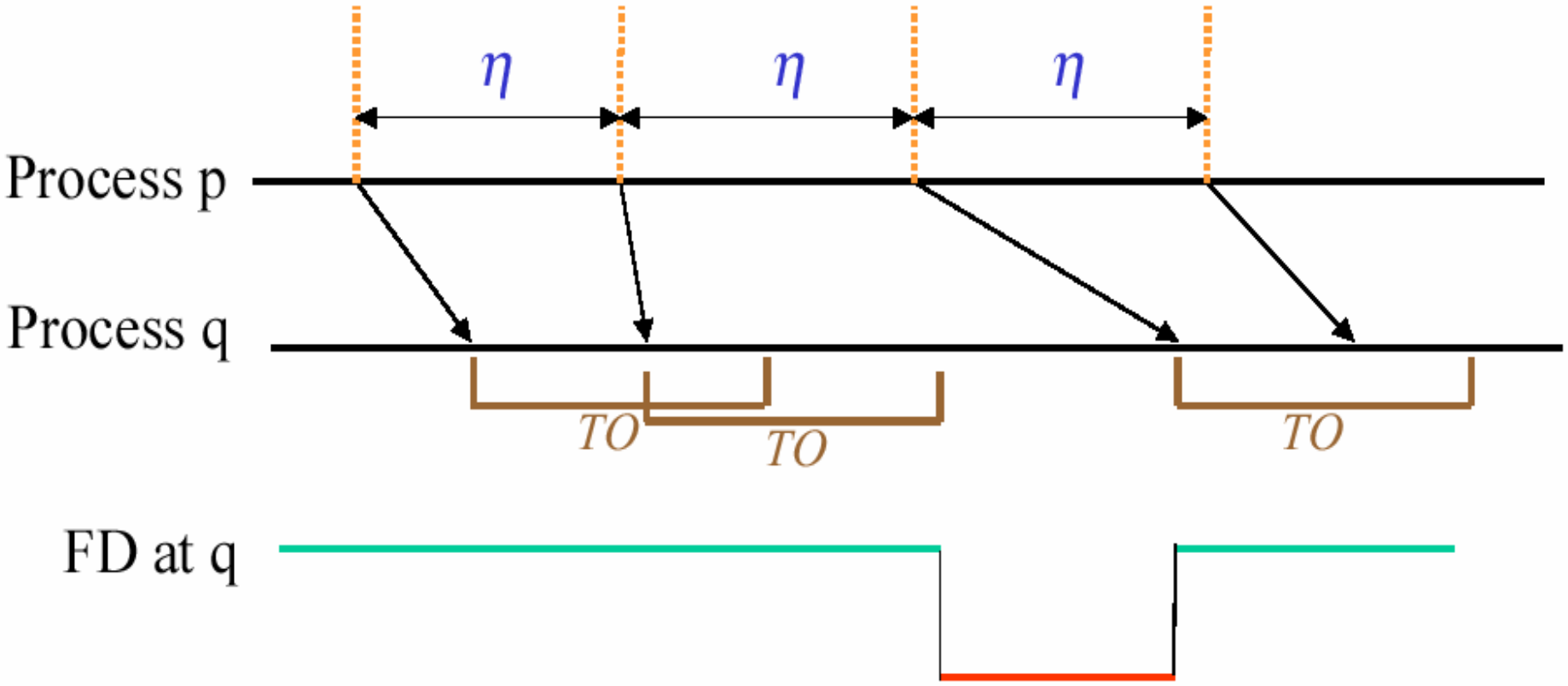
Presentation Outline

- Introduction of QoS
- On the QoS Specification of Failure Detectors
- The Design and Analysis of a New Failure Detector Algorithm
- Configuring the Failure Detector to Satisfy QoS Requirements

The Probabilistic Network Model

- Process p and q are connected by a link that does not create or duplicate messages, but may delay or drop messages.
- Message loss probability p_L
- Message delay time D
- Process p and q have access to synchronized clocks.

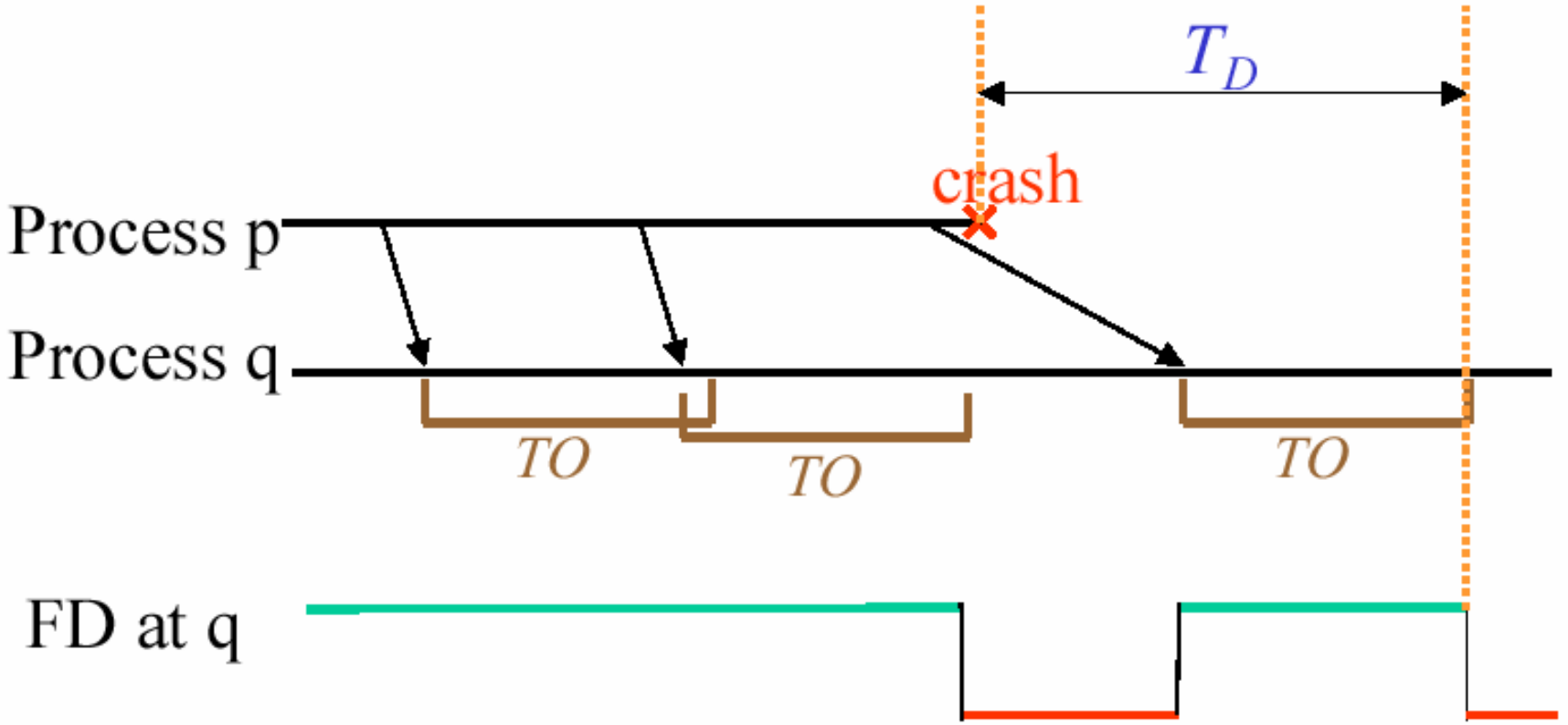
A Simple FD Algorithm



- Timing out depends on two consecutive messages

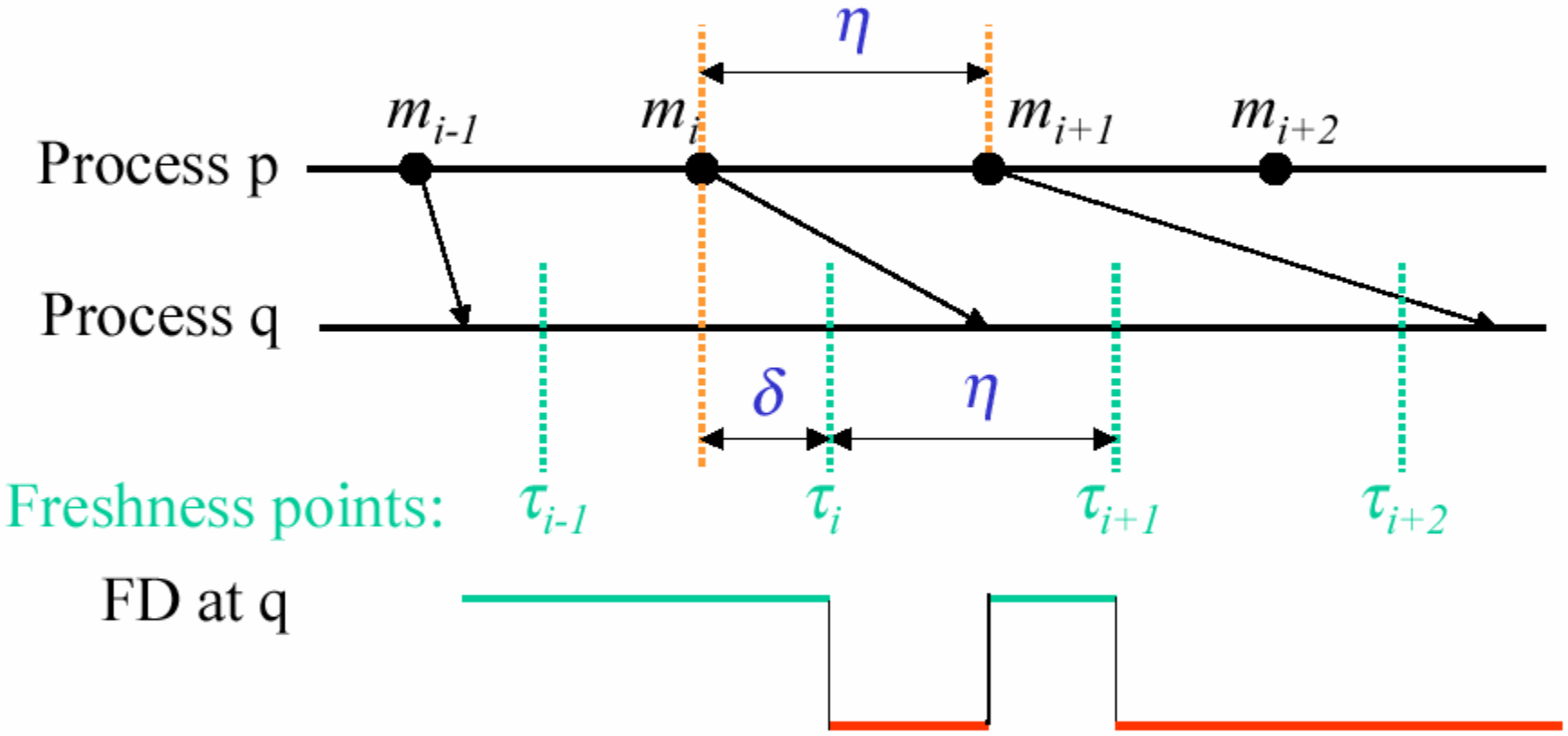
Large Detection Time

- Depends on the delay of the last message sent by p



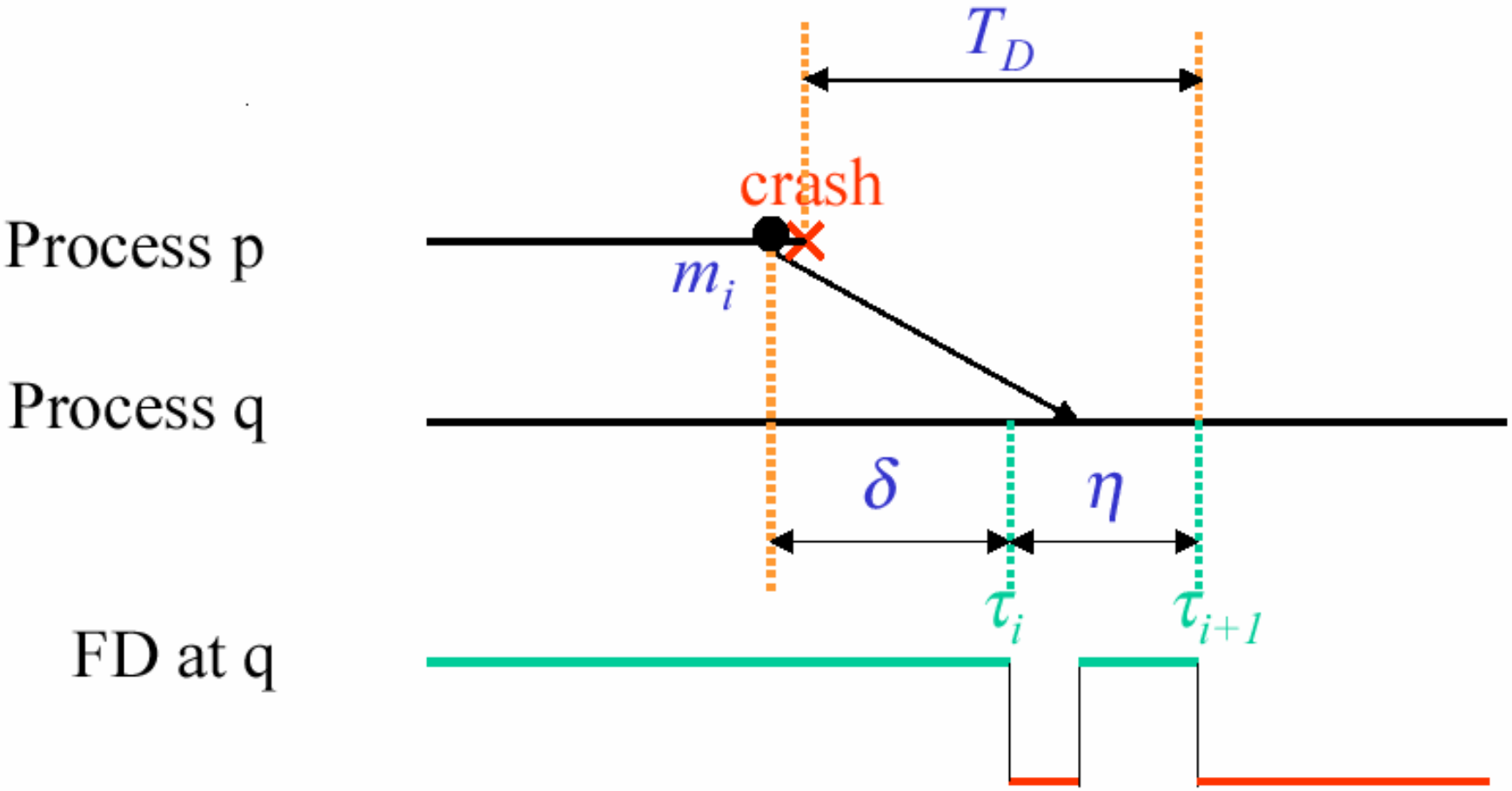
$$T_D \leq \max(D) + TO$$

New Algorithm (w/ synchronized clocks)



- At any time $t \in [\tau_i, \tau_{i+1})$, FD trusts p iff q has received heartbeat message m_i or higher.

Detection Time



$$T_D \leq \delta + \eta$$

An Optimality Result

Among all FD algorithms such that

- the monitored process p sends a message every η,
- the detection time is always less than a given bound,

our new algorithm provides the best query accuracy probability.

Presentation Outline

- Introduction of QoS
- On the QoS Specification of Failure Detectors
- The Design and Analysis of a New Failure Detector Algorithm
- Configuring the Failure Detector to Satisfy QoS Requirements
- Conclusion Remarks

Satisfying QoS Requirements

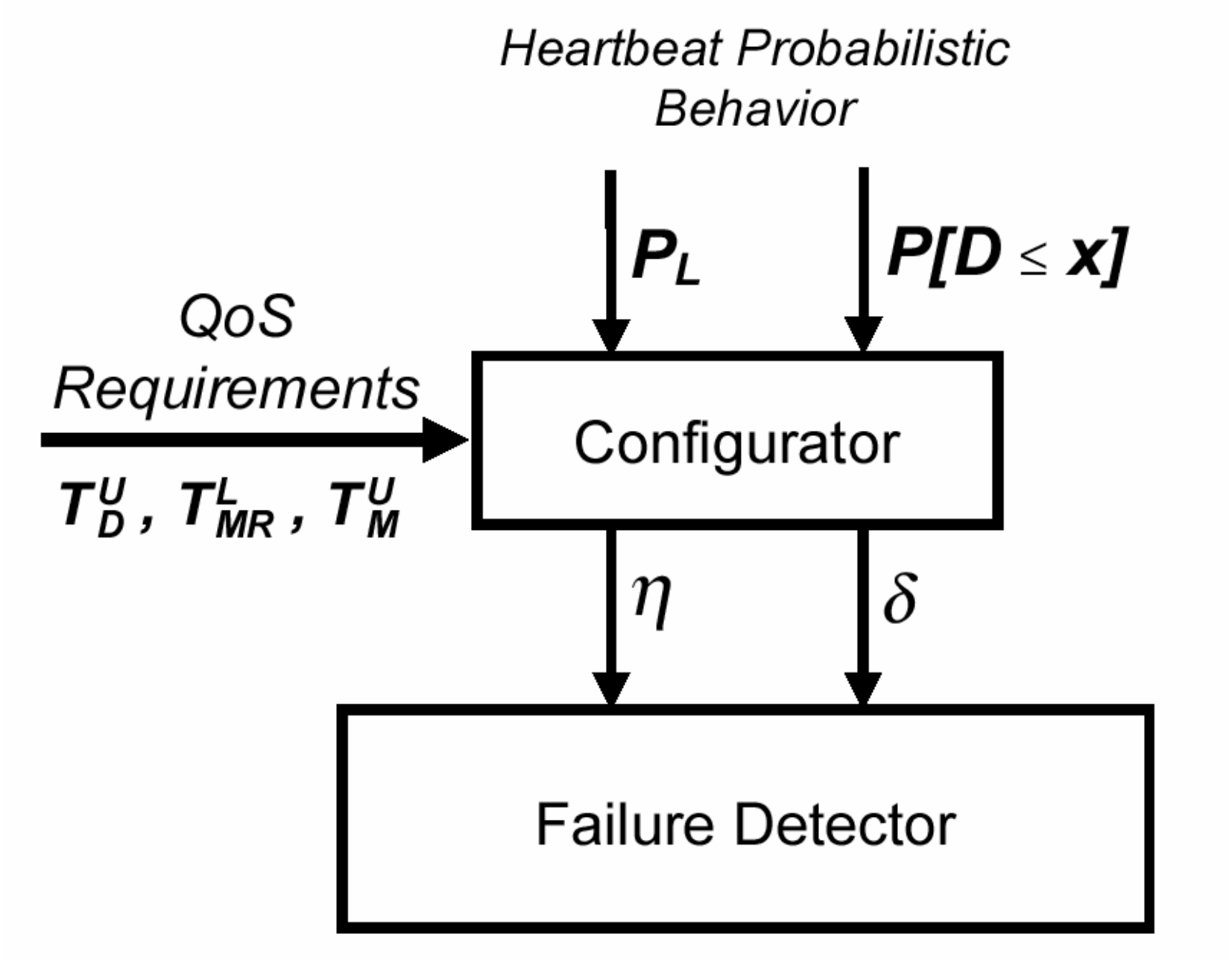
- Given a set of QoS requirements as a tuple (T_D^U, T_{MR}^L, T_M^U) such that

$$T_D \leq T_D^U$$

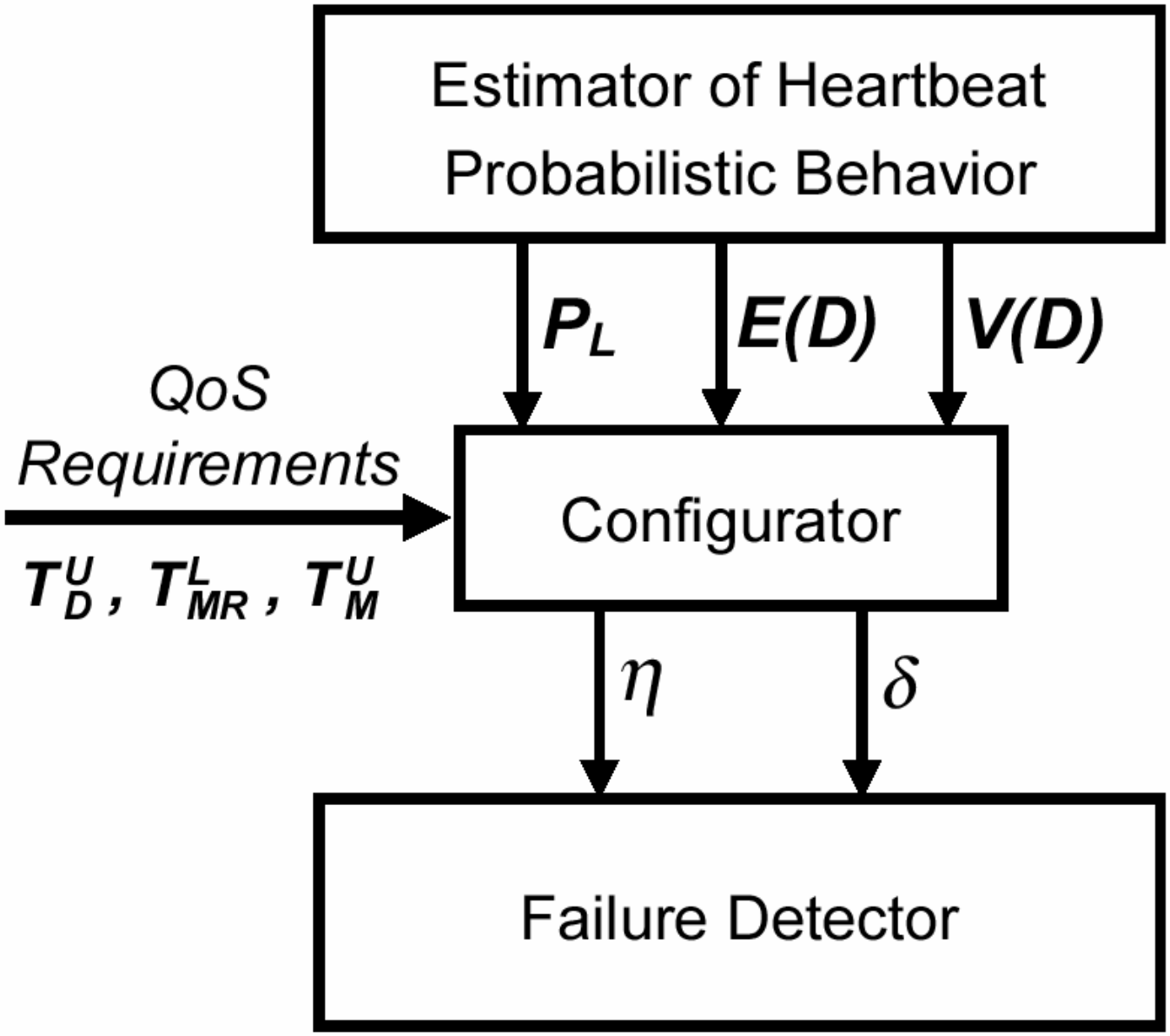
$$E(T_{MR}) \geq T_{MR}^L$$

$$E(T_M) \leq T_M^U$$

- Find η and δ to achieve these requirements



The probabilistic behavior of heartbeats is **given**



The probabilistic behavior of heartbeats is **unknown**

Main Idea

- Bound $\Pr(D \leq x)$ using $E(D)$ and $V(D)$
- Modify configuration procedure to use $E(D)$ and $V(D)$ instead of $\Pr(D \leq x)$
- Estimate $E(D)$, $V(D)$ and p_L using heartbeats
- Use estimates to run configuration procedure

An adaptive Failure Detector

- In some networks, the probabilistic behavior of heartbeat message changes along the time.
- The procedure used for unknown message behavior can be used to make failure detector adaptive.
- Main idea: to periodically estimate the current values of $E(D)$, $V(D)$ and p_L using the n most recent heartbeats.

Presentation Outline

- Introduction of QoS
- On the QoS Specification of Failure Detectors
- The Design and Analysis of a New Failure Detector Algorithm
- Configuring the Failure Detector to Satisfy QoS Requirements
- Conclusion Remarks

Concluding Remarks

- This work is the first comprehensive and systematic study of the QoS of failure detectors using probability theory.
- The new algorithm presented provides the best query accuracy probability.
- It shows how to compute the failure detector parameters to satisfy the given QoS requirements with or without knowing the probabilistic behavior of heartbeat messages.
- Adaptive failure detector forms the core of a failure detection service that is currently being implemented and evaluated.