Δ-Reliable Broadcast: A Probabilistic Measure of Broadcast Reliability

Patrick Th. Eugster Rachid Guerraoui Petr Kouznetsov

Sun Microsystems
Switzerland

Distributed Programming Lab, EPFL
Switzerland

Outline

1. The problem: scalable "reliable" broadcast
2. Reliable Broadcast specification [HT94].
3. Δ-Reliable Broadcast.
4. Reliability distribution function.
5. Examples: reliability analysis of Bimodal Multicast [BHO^{+}99] and IP Multicast [DC90].

Broadcast protocols

- Best-effort: Multicast Usenet (MUSE), IP Multicast, Reliable Multicast Transfer Protocol (RMTP), etc.
- Probabilistic: Bimodal Multicast, lpbcast, etc.

What is the problem?

Traditional specification: Reliable Broadcast[HT94]

Integrity. For any message m, every process delivers m at most once, and only if m was previously broadcast by sender(m).

Validity. If a correct process p broadcasts a message m, then p eventually delivers m.

Agreement. If a correct process delivers a message m, then every correct process eventually delivers m.

Informal specification: Atomicity [BHO$^+$99]

A broadcast protocol provides a bimodal delivery guarantee if there is

- a high probability that a broadcast message will reach almost all processes,
- a low probability that a broadcast message will reach just a very small set of processes, and
- a vanishingly small probability that a broadcast message will reach some intermediate number of processes.

Bridging the gap: Δ-Reliable Broadcast

Let $\Delta = (\psi, \rho) \in [0,1] \times [0,1]$. A broadcast protocol is Δ-*Reliable* iff the following properties are *simultaneously satisfied with probability* ψ:

Integrity. For any message m, every process delivers m at most once, and only if m was previously broadcast by $sender(m)$.

Validity. If a correct process p broadcasts a message m then p eventually delivers m.

Δ-Agreement. If a correct process delivers a message m, then eventually at least a fraction ρ of correct processes deliver m.

Δ-Reliable Broadcast: ρ and ψ

$\Delta = (\psi, \rho)$ is a "reliability measure" of a given protocol.

Reliability degree ρ: the fraction of correct processes that eventually deliver a broadcast message.

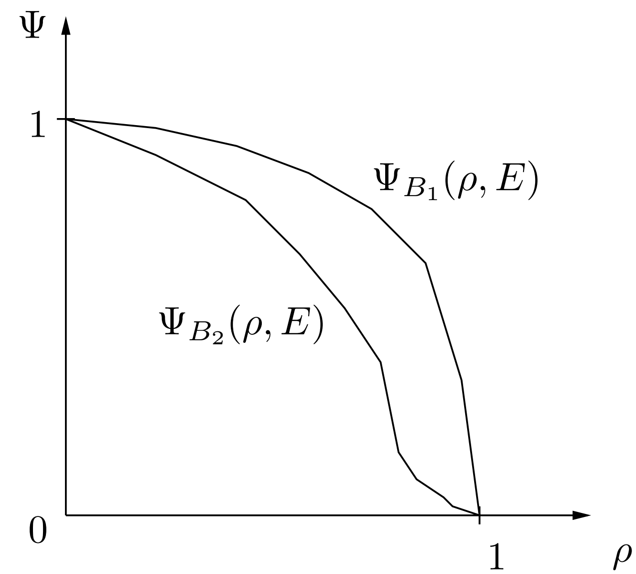
Reliability probability ψ: - the probability that "enough" (correct) processes deliver a broadcast message and no fake or duplicate messages are delivered.

Reliability distribution function

Let $\mathbb{E}$ be an *environment space* and B be a broadcast protocol.

A function $\psi_B : [0,1] \times \mathbb{E} \mapsto [0,1]$ is the *reliability distribution function* of B iff

$\forall \rho \in [0,1]\ \forall \mathcal{E} \in \mathbb{E}$:
B is Δ-Reliable with $\Delta = (\psi_B(\rho, \mathcal{E}), \rho)$.



B_1 is more reliable than B_2 in E.

Reliability distribution function: examples

- *Dreamcast* (Reliable Broadcast [HT94]) in a given $\mathcal{E} \in \mathbb{E}$:

$$\forall \rho \in [0,1] : \psi(\rho, \mathcal{E}) = 1.$$

- *Spellcast* (does nothing):

$$\forall \rho \in]0,1], \forall \mathcal{E} \in \mathbb{E} : \psi(\rho, \mathcal{E}) = 0.$$

Atomicity

Atomicity predicate of Bimodal Multicast: given a protocol B, $\sigma \in [0, 0.5]$ and an environment $\mathcal{E} \in \mathbb{E}$, a broadcast message reaches more than a fraction σ, but less than a fraction $1 - \sigma$ of correct processes with probability:

$$P(\sigma \leq \rho < 1 - \sigma) = \psi_B(\sigma, \mathcal{E}) - \psi_B(1 - \sigma, \mathcal{E})$$

Bimodal Multicast [BHO^{+}99]

Environment:
- n processes
- Fanout β
- Message loss probability ε
- Process crash probability τ
- Number of gossip rounds T

deliver_and_gossip(m, round) {* Auxiliary function *}
    if received_already(m) then return
    bmdeliver(m)
    if round=T then return
    choose $S \subset \Pi$, such that $|S| = n\beta$
    for each p in S send to p *gossip*(m,round+1)

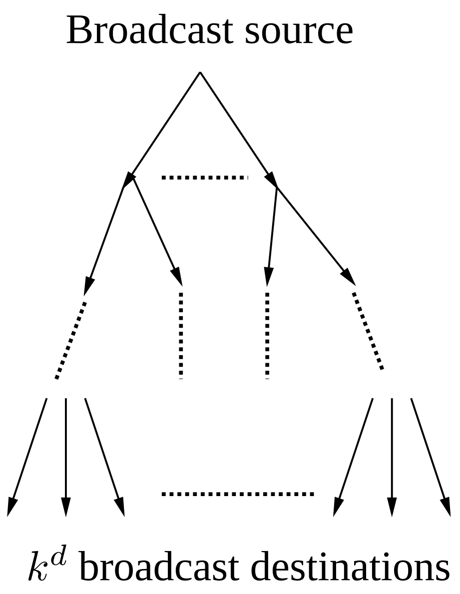
On bmcast(m):
    deliver_and_gossip(m,0)

On receive gossip(m,round):
    deliver_and_gossip(m,round)

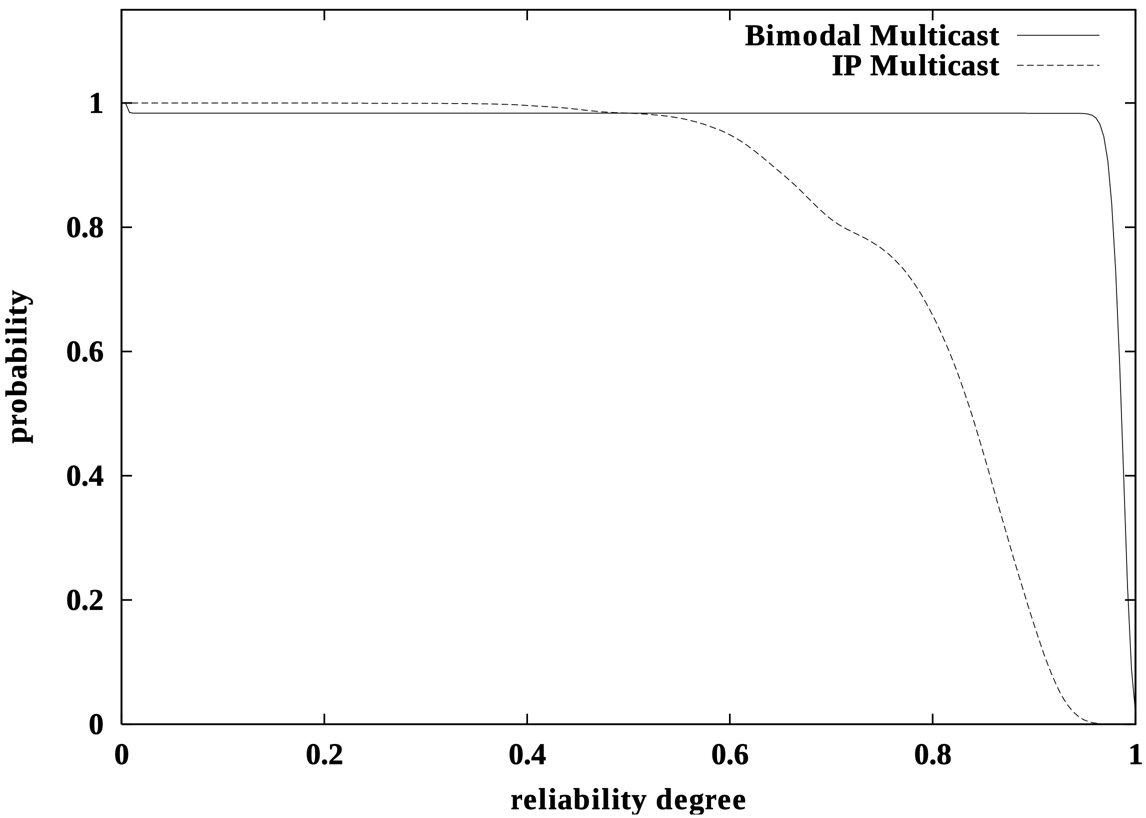
IP Multicast (PIM-SM) [DC90, FHHK00]

Environment:

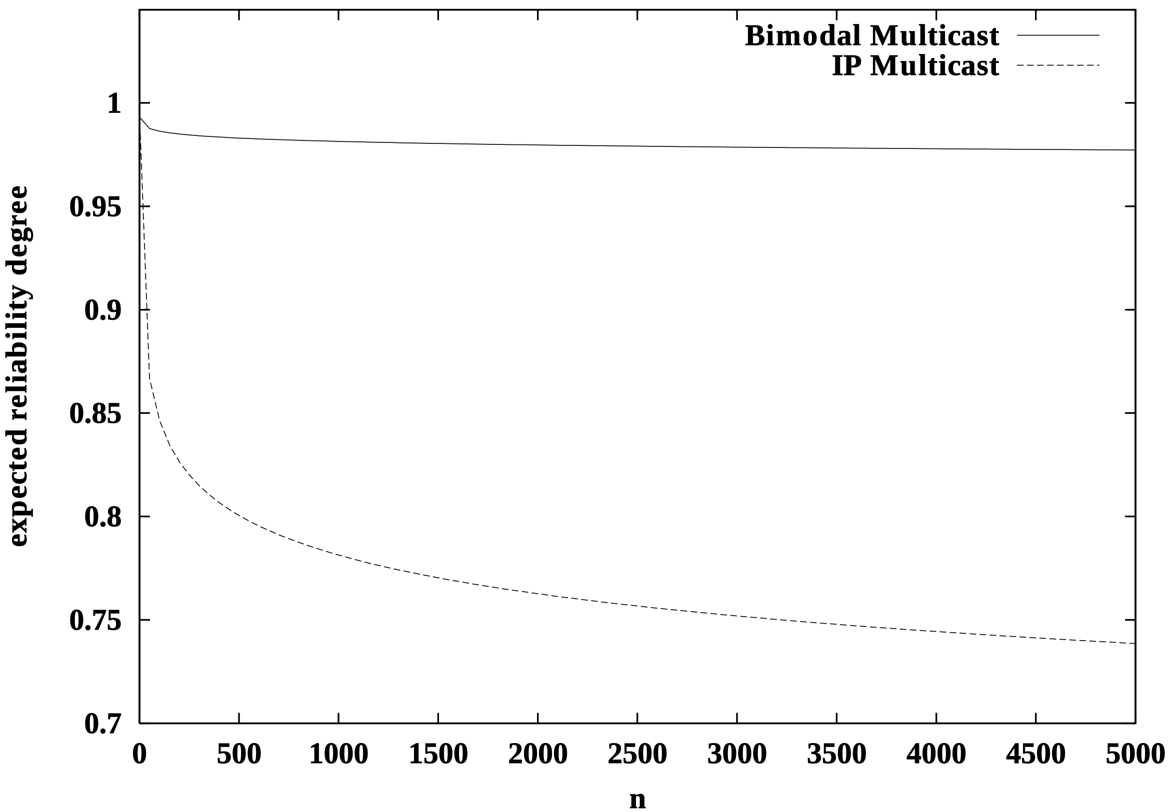
k-ary spanning tree of depth d
k^d processes
Message loss probability ε_l
Process crash probability τ
Router crash probability γ



Reliability distribution functions



Average reliability degrees



Outline

1. The problem: scalable "reliable" broadcast
2. Reliable Broadcast specification [HT94].
3. Δ-Reliable Broadcast.
4. Reliability distribution function.
5. Examples: reliability analysis of Bimodal Multicast [BHO$^+$99] and IP Multicast [DC90].

References

[BHO+99] Kenneth P. Birman, Mark Hayden, Oznur Ozkasap, Zhen Xiao, Mihai Budiu, and Yaron Minsky. Bimodal multicast. *ACM Transactions on Computer Systems*, 17(2):41–88, 1999.

[DC90] S. Deering and D. Cheriton. Multicast Routing in Datagram Internetworks and Extendexd LANs. *ACM Transactions on Computer Systems*, 8(2):85–110, May 1990.

[FHHK00] B. Fenner, M. Handley, H. Holbrook, and I. Kouvelas. Protocol Independent Multicast-Sparse Mode (PIM-SM):

Protocol Specification (Revised). *Internet Engineering Task Force (IETF)*, November 2000.

[HT94] Vassos Hadzilacos and Sam Toueg. A modular approach to fault-tolerant broadcast and related problems. Technical report, Cornell University, Computer Science, May 1994.