Sara Salahi
CS 443 – Paper Summary

Title: *BAR Fault Tolerance for Cooperative Services*
Authors: Amitanand S. Aiyer, Lorenzo Alvisi, Allen Clement, Mike Dahlin, Jean-Phillipe Martin and Carl Porth

Summary: In multiple administrative domains, protocols must tolerate two distinct behaviors: 1) Byzantine, when broken, misconfigured or malicious nodes arbitrarily deviate from their specification and 2) Rational, when selfish nodes deviate from their specification to increase their local benefit. The authors 1) introduce the BAR model as a foundation for reasoning about cooperative services, 2) propose a three-level architecture to reduce the complexity of building services under the BAR model, and 3) describe an implementation of BAR-B – the first cooperative backup service to tolerate nodes that exhibit both types of behavior.

Key Ideas

The BAR model: 1) Byzantine – nodes that behave arbitrarily or maliciously, 2) Altruistic – nodes that execute the proposed program, whether it benefits them or not, and 3) Rational – nodes that deviate from the proposed program for the purposes of local benefit. To provide guarantees similar to those from Byzantine fault tolerance to all rational and altruistic nodes, the authors propose the following two protocols: 1) Incentive-Compatible Byzantine Fault Tolerance (IC-BFT), a protocol that guarantees the specified set of safety and liveness properties if it is in the best interest of all rational nodes to follow the protocol exactly, and 2) Byzantine Altruistic Rational Tolerant (BART), a protocol that guarantees the specified set of safety and liveness properties in the presence of all rational deviations from the protocol. In order to extend Byzantine fault tolerance to nodes that are greedy (rational), the authors use game theory tools such as proposing a protocol that provides a Nash Equilibrium, where the nodes have nothing to gain by deviating from the protocol while the other nodes do not (all rational nodes will follow protocol because they have nothing to gain by deviating from it themselves). They also propose a protocol that will punish rational nodes for their greediness by denying them access to a state machine which allows them to complete their objectives. The three-level architecture isolates classes of misbehavior at appropriate levels of abstraction. Level 1 (Basic Primitives) achieves five goals: 1) provide IC-BFT versions of key abstractions using BART-RSM which is based off of PBFT, 2) ensure long-term benefits to participants by rotating the leadership role among the participants, 3) limit non-determinism by communicating proofs of misconduct and the use of Terminating Reliable Broadcast (TRB), 4) mitigate the effects of residual non-determinism by encouraging timeliness by threatening penance, and 5) enforce predictable communication patterns with a message queue. Level 2 (Work Assignment) assigns work of state machine replication to individual nodes using arithmetic codes to provide low overhead fault-tolerant storage. Witness nodes are a "go-between" for communication between two nodes. They provide a proof-of-misconduct if one node fails to perform. A fast-path also exists between two nodes to bypass the witness (like a game of chicken). Level 3 (Application) is the BAR-B cooperative backup system. The system has three operations: to store, retrieve and audit. To store a file, the node breaks it up into pieces, encrypts it and stores the pieces on remote nodes. The remote nodes send back a receipt. To retrieve the file the receipt is sent back to the remote nodes and the remote nodes can respond with the chunk, a demonstration that the lease has expired or a more recent storage receipt. Receipts are the method for auditing – nodes exchange receipts to verify compliance with storage quotas. Overall, the performance is worse than other protocols that do not make the guarantees provided by IC-BFT and BART, but maybe that is acceptable.

Sara Salahi
CS 443 – Paper Summary

Flaws

Because of all the assumptions the authors made, this paper only demonstrates a good start for this project. However, the results are not relevant to true systems and much future work needs to be done in relaxing their assumptions. The results of their tests are not compared against anything. Although there are no comparable protocols available to test against, the authors could have compared against an "optimal" of some sort.

Relevance and Future Work

Research that involves only dealing with Byzantine behaviors is not enough, especially when dealing with Multiple Administrative Domains (MAD) where nodes can deviate from their specification in order to increase their local benefit. This is relevant to Internet routing, file distribution and cooperative backup – for example, 30 co-workers who cooperatively backup their personal home machines.

Future work mainly lies in relaxing the assumptions made in this paper and optimizing the protocol to handle more relaxed assumptions.