Reasoning with MAD Distributed Systems

Lorenzo Alvisi and Edmund L. Wong

Laboratory for Advanced Systems Research (LASR)
Department of Computer Science, The University of Texas at Austin
2317 Speedway, 2.302
Austin, TX 78712 USA
{lorenzo,elwong}@cs.utexas.edu

> *How does one reason about and build dependable distributed systems in which no component is guaranteed to follow the specified protocol?*

While the setting of this question may appear implausible, this is precisely the environment in which services that span multiple administrative domains (MAD) must function. In such services—which include applications such as content dissemination (e.g.., [2]),, file backup (e.g., [6]), volunteer computing (e.g., [5]), multi-hop wireless networking (e.g., [4]), and Internet routing—resources are not under the control of a single administrative domain, so the necessary cooperation cannot simply be achieved by fiat. Instead, it is imperative that the service be structured so that nodes—which are administered by different, potentially selfish entities—have an incentive to help sustain it. Indeed, such issues are not imaginary: ample evidence suggests that a large number of peers will free-ride or deviate from the assigned protocol if it is in their interest to do so (e.g., [3,9,16,21]).

The presence of rational nodes challenges all approaches to dependability that rely on a clean separation between correct and faulty nodes. The standard approach to fault tolerance that relies on correct nodes to take appropriate action to mask or tolerate faulty nodes no longer applies when nodes that are not faulty may nonetheless selfishly deviate from their correct specification. Even the basic question of deciding on a failure model appropriate for reasoning about dependable MAD systems does not offer an obvious answer. Of course one could model all deviations, whether due to faults or to selfishness, as Byzantine faults [17], but many interesting problems in distributed computing become unsolvable once the number of Byzantine nodes exceeds a third of the total [17]—and there is no inherent reason why the combined number of faulty and selfish node should conveniently stay below that threshold. Alternatively, one could apply classic notions from game theory to model selfish behavior, but these typically only account for rational behavior, and become brittle if some (faulty) nodes behave in a seemingly irrational fashion. This is particularly the case in cooperative services, where the nodes themselves are often unreliable personal machines riddled with malware and other exploits [1,11]. The natural way forward, then, is to somehow combine insight from game theory and fault-tolerant distributed computing.

One strategy is to specify a notion of equilibrium that draws inspiration from traditional Byzantine fault tolerance and to aim for a solution concept in which rational nodes prefer the specified strategy despite the presence of a threshold of arbitrary failures [7,8,15]. This is the approach adopted by the elegant

P.R. D'Argenio and H. Melgratti (Eds.): CONCUR 2013, LNCS 8052, pp. 1–4, 2013.

(k, t)-robustness solution concept [7]. Rational nodes have no incentive to deviate from a (k, t)-robust equilibrium despite up to t Byzantine nodes; further, unlike Nash equilibria, which are robust only against unilateral deviations, (k, t)-robust equilibria can tolerate collusions of up to k rational nodes. Unfortunately, the elegance of (k, t)-robustness comes at the cost of strong assumptions, which in principle can limit the practical applicability of this solution concept in realistic scenarios (more on this later).

An alternative strategy, explored in the BAR approach [10,14] is to classify nodes as belonging to one of three classes (Byzantine, Acquiescent, and Rational) and to model explicitly the expectations held by rational nodes concerning the behavior of Byzantine nodes. On the positive side, the BAR model has been successfully applied to build several real systems [10,18,19] that tolerate both malicious and rational deviations; however, these systems, as well as other work that has relied on models similar to BAR [20], suffer from several limitations of their own. First, they assume that rational nodes always model that Byzantine behavior as malicious, with Byzantine nodes hell-bent on producing the worst possible outcome for every other node; second, while they do not *rely* on acquiescent nodes to provide their guarantees, they also do not take advantage of their presence; finally, they do not explicitly handle collusion among rational nodes: at best, colluding rational nodes are modeled as Byzantine [10,19].

What should then be the basis for a rigorous treatment of cooperative services? How should participating nodes be modeled and what guarantees should we aim for? And can these models and guarantees be applied to real systems? This talk reviews our recent progress in trying to answer these questions.

Which solution concept can offer rigorous and practical basis for dependable cooperative services? To answer this question, we introduce a *communication game* that captures the key characteristics of most distributed systems that tolerate arbitrary faults. Specifically, our game models systems in which (a) some node-to-node communication is necessary to achieve some desired functionality (b) bandwidth is not free; and (c) the desired functionality is achievable despite t Byzantine failures. We find that notions of equilibrium inspired by traditional Byzantine fault-tolerant techniques, such as (k, t)-robustness, are capable of achieving equilibrium in communication games only under very limited circumstances, severely limiting their practical usefulness [22,24]. Our findings suggest that practical solution concepts must explicitly model the *beliefs* of rational nodes when it comes to Byzantine behavior.

What is the role of acquiescent nodes in cooperative services? Although real MAD systems include a sizable fraction of acquiescent (correct and unselfish) nodes, their impact on the incentive structure of MAD services is not well understood. In particular, systems built under the BAR model have sidestepped the challenge by designing protocols that neither depend on nor leverage the presence of acquiescent nodes—indeed, it seems possible that the very presence of acquiescent nodes may demotivate selfish rational nodes from contributing their share of resources, in the hope of free-riding off the acquiescent nodes' good will.

Can that good will be leveraged without imperiling rational participation? By distilling this question to a rational peer's *last* opportunity to cooperate, we find that not only is the good will of acquiescent nodes not antithetical to rational cooperation, but that, in a fundamental way, rational cooperation can only be achieved in the presence of the scintilla of altruism that acquiescent nodes bring to the system [23].

How should collusion be managed? The literature offers two approaches to guarantee that deviations resulting from collusion do not affect the incentives provided to rational nodes. The first is to model collusion as a fault and colluding nodes as Byzantine—which, similar to modeling rational deviations as Byzantine, forces an artificially low cap on the number of colluders. The second approach—taken by *strong Nash* [12], *k-resilient* equilibria [7,8], and *coalition-proof* Nash equilibria [13], to name a few—is to deny any benefit to colluders: if the equilibrium is a best response not just to every individual, but also to every possible coalition, then collusion poses no harm to the equilibrium's stability, since nodes gain no benefit by colluding. However, nodes that collude are likely to trust each other more and, more generally, be able to hold stronger assumptions about one another. Since stronger assumptions typically lead to more efficient protocols, identifying a single strategy that is a best response both inside and outside of every possible coalition is in practice very hard.

To overcome this challenge, we propose a fundamentally different approach to dealing with coalitions, based on the observation that while finding a single best response between all nodes is sufficient to prevent nodes from deviating, it is not *necessary* to achieve such stability. We introduce two new notions of equilibrium that leverage the observation that coalitions (including the trivial singleton coalition of one non-colluding node) will not deviate from an equilibrium as long as the equilibrium specifies a best-response strategy for every *coalition*. We thus allow the strategy a node follows to depend on whom the node is colluding with, thereby enabling the equilibrium to explicitly account for the advantages of coalition members while guaranteeing that nodes have no incentive to deviate from the specified equilibrium [22].

∼ ∗ ∼

We are working on the design and implementation of a new hybrid (in that it relies on both servers and peer-to-peer cooperation) content distribution system that aims to apply these insights towards building a scalable, robust, and dependable method for distributing content.

References

1. 1 in 5 Macs has malware on it. Does yours?, `http://nakedsecurity.sophos.com/2012/04/24/mac-malware-study`
2. BitTorrent, `http://bittorrent.com`
3. Kazaa Lite, `http://en.wikipedia.org/wiki/Kazaa_Lite`
4. OpenGarden, `http://opengarden.com`
5. SETI@home, `http://setiathome.ssl.berkeley.edu`

6. SpaceMonkey, `http://spacemonkey.com`
7. Abraham, I., Dolev, D., Gonen, R., Halpern, J.: Distributed computing meets game theory: Robust mechanisms for rational secret sharing and multiparty computation. In: Proceedings of the 25th Annual ACM SIGACT-SIGOPS Symposium on Principles of Distributed Computing (July 2006)
8. Abraham, I., Dolev, D., Halpern, J.Y.: Lower bounds on implementing robust and resilient mediators. In: Proceedings of the 5th IACR Theory of Cryptography Conference (March 2008)
9. Adar, E., Huberman, B.A.: Free riding on Gnutella. First Monday 5(10), 2–13 (2000), `http://www.firstmonday.org/issues/issue5_10/adar/index.html`
10. Aiyer, A.S., Alvisi, L., Clement, A., Dahlin, M., Martin, J.P., Porth, C.: BAR fault tolerance for cooperative services. In: Proceedings of the 20th ACM Symposium on Operating Systems Principles (October 2005)
11. Alyias, D., Batchelder, D., Blackbird, J., Faulhaber, J., Felstead, D., Henry, P., Jones, J., Kuo, J., Lauricella, M., Li, L., Ng, N., Rains, T., Sekhar, V., Stewart, H., Thomlinson, M., Zink, T.: Microsoft Security Intelligence Report, July through December, vol. 14 (2012)
12. Aumann, R.J.: Acceptable points in general cooperative n-person games. Annals of Mathematics Study 40(4), 287–324 (1959)
13. Bernheim, B.D., Peleg, B., Whinston, M.D.: Coalition-proof Nash equilibria, I. Concepts. Journal of Economic Theory 42(1), 1–12 (1987)
14. Clement, A., Li, H.C., Napper, J., Martin, J.P., Alvisi, L., Dahlin, M.: Bar primer. In: DSN, pp. 287–296 (2008)
15. Eliaz, K.: Fault tolerant implementation. Review of Economic Studies 69, 589–610 (2002)
16. Hughes, D., Coulson, G., Walkerdine, J.: Free riding on Gnutella revisited: The bell tolls? IEEE Distributed Systems Online 6(6) (June 2005)
17. Lamport, L., Shostak, R., Pease, M.: The Byzantine generals problem. ACM Transactions on Programming Languages and Systems 4(3), 382–401 (1982)
18. Li, H.C., Clement, A., Marchetti, M., Kapritsos, M., Robison, L., Alvisi, L., Dahlin, M.: FlightPath: Obedience vs. choice in cooperative services. In: Proceedings of the 8th USENIX Symposium on Operating Systems Design and Implementation (December 2008)
19. Li, H.C., Clement, A., Wong, E.L., Napper, J., Roy, I., Alvisi, L., Dahlin, M.: BAR Gossip. In: Proceedings of the 7th USENIX Symposium on Operating Systems Design and Implementation (November 2006)
20. Moscibroda, T., Schmid, S., Wattenhofer, R.: When selfish meets evil: Byzantine players in a virus inoculation game. In: Proceedings of the 25th Annual ACM SIGACT-SIGOPS Symposium on Principles of Distributed Computing (July 2006)
21. Saroiu, S., Gummadi, P.K., Gribble, S.D.: A measurement study of peer-to-peer file sharing systems. In: Proceedings of the 9th Annual ACM/SPIE Multimedia Computing and Networking (January 2002)
22. Wong, E.L., Alvisi, L.: What's a little collusion between friends? In: Proceedings of the 32nd Annual ACM Symposium on Principles of Distributed Computing (July 2013)
23. Wong, E.L., Leners, J.B., Alvisi, L.: It's on me! The benefit of altruism in BAR environments. In: Proceedings of the 24th International Symposium on Distributed Computing (September 2010)
24. Wong, E.L., Levy, I., Alvisi, L., Clement, A., Dahlin, M.: Regret freedom isn't free. In: Proceedings of the 15th International Conference on Principles of Distributed Systems (December 2011)