Protocol and Network Design for Manageability

Jennifer Rexford, AT&T Labs–Research
http://www.research.att.com/~jrex

Designing an IP network is a challenging task that requires selecting from a wide variety of components, protocols, and mechanisms to meet current and future demands. Equally daunting, *managing* the network requires tuning these protocols and mechanisms over time in response to changing constraints and conditions. However, today's IP protocols were not designed with manageability in mind, and building and operating IP networks is, at best, a black art practiced by an increasingly overwhelmed community of engineers. The poor state-of-the-art suggests three promising avenues for networking research on the Science of Design:

- *Protocols and mechanisms*: Protocols and mechanisms should be designed with their role in a large, complex, tunable system in mind. Good protocols and mechanisms should have several key properties, such as predictable behavior for a given set of inputs and small reactions to small events, to enable the construction of robust, efficient, secure networks. Identifying these properties requires analyzing the limitations of the existing protocols and mechanisms as they are deployed in large network settings.
- *Networks*: Given a collection of existing protocols and mechanisms, networks should be designed to satisfy complex, diverse, and changing goals. Good network design requires a set of principles and best-common practices for driving key design decisions and comparing alternative designs. Identifying these principles and practices requires grappling with nuanced metrics such as robustness, flexibility, and expressiveness, and studying a wide range of existing network designs.
- *Network management systems*: Given an existing network, a network management system should monitor the performance and behavior of the system and tune the configuration as the constraints and conditions change. Good network management systems enable a high-level specification of policies and automate low-level tasks. Identifying the key abstractions requires studying the challenges of managing existing networks to uncover the barriers to flexible operation and automation.

The three design problems are closely related. A good design of the underlying protocols and mechanisms greatly simplifies the process of designing and managing the network. The protocol and network designs determine what kinds of "dials" the network management system can track and what "knobs" can be tuned to affect network behavior. That is, the protocols and the network design determine what "control loop" the network management system needs to model and influence.

To ground these three design problems, we enumerate several desirable properties of a network protocol to simplify network design and management, and illustrate how they are violated by the Border Gateway Protocol (BGP), the de facto interdomain routing protocol for the Internet. The following properties make a network protocol easier to manage:

- *Predictable output for a given input*: To simplify network management, the protocol behavior should be predictable for a given set of inputs; this makes it easier to answer "what if" questions about how the network would respond to a change in the inputs. Yet, BGP is not guaranteed to converge, stable solutions are not necessarily unique, and the protocol includes features that are non-deterministic (such as age-based tie-breaking).

- *Small reactions to small changes*: For a robust and stable network, the protocol should not have a large reaction to a small change in the inputs; this helps avoid over-reacting to minor fluctuations in traffic, routing, and topology. Yet, the common practice of "hot potato" routing makes BGP routing decisions (and, hence, the flow of traffic) inside an Autonomous System extremely sensitive to small changes in the path costs in the intradomain routing protocol.

- *Tractable network-wide models of behavior*: For ease of management, the network-wide behavior of components running the protocol should be easy to model; this reduces the complexity of reasoning about the system and the computational overhead of predicting system behavior. Yet, certain BGP features (such as the multiple-exit discriminator and route reflectors) make it impossible to impose a total ordering on the externally-learned routes, making the modeling of BGP more complicated.

- *Simple tuning of configurable parameters*: For efficient operation of the network, the tunable parameters of the protocol should be easy to configure to satisfy network-wide goals; this makes it possible to select parameter values based on best-common practices or simple optimization techniques. Yet, BGP offers an extremely large set of tunable parameters that have, at best, an indirect affect on the flow of traffic, making it difficult to select good parameter settings.

- *Ease of monitoring and troubleshooting*: To optimize end-to-end performance and diagnose network problems, the protocol should be easy to monitor and troubleshoot; this makes it possible to adapt to changing network conditions and to diagnose and fix problems as they arise. Yet, monitoring a path-vector protocol like BGP requires collecting data from many vantage points, and BGP messages do not provide any explanation for why the router changed from one route to another.

- *Protection from misbehaving participants*: For a secure and reliable network, the protocol should not be vulnerable to misbehaving participants; this helps protect the network and its users from configuration mistakes and malicious attack. Yet, BGP has no mechanisms to prevent a network from announcing routes for destinations it does not own, or tampering with the contents of route advertisements initiated by others; this can lead to blackholed traffic and other routing anomalies.

BGP is not alone in violating these principles. For example, active queue management schemes such as RED (Random Early Detection) are notoriously difficult to model and tune in large network settings, and the Transmission Control Protocol (TCP) has many known vulnerabilities to misbehaving participants. Similarly, the Domain Name System has many persistent configuration problems and is quite difficult to troubleshoot in practice.

In fact, satisfying all of these protocol design goals, along with the more traditional goals of scalability and efficiency, may be extremely difficult if not impossible. As a research community, we need to determine which principles are most important and how to design protocols that satisfy them. Where the protocol design alone cannot solve the problem, we need to determine how to create network designs and operational practices that ensure good system behavior (i.e., how to "engineer" the system to satisfy the remaining properties). Finally, we need to rethink the design of the boundary between the network protocols (the "control plane") and the network management systems (the "management plane"). Ultimately, we may find value in having much simpler network protocols that provide very basic service, while placing increasing responsibility for operating the network in the management systems[1].

[1] Many of the limitations of BGP can be addressed by moving functionality from the routers to a logically-centralized, customizable service for selecting and propagating routes. We describe this idea, and elaborate on many of the BGP weaknesses highlighted above, in "A Case for Separating Routing from Routers" by N. Feamster, H. Balakrishnan, J. Rexford, A. Shaikh, and J. van der Merwe in *Proc. ACM SIGCOMM Workshop on Future Directions in Network Architecture*, August 2004.