

Some Comments on Hourglasses

Rui L Aguiar
Universidade de Aveiro
Instituto de Telecomunicações
3810-193 Aveiro Portugal
ruilaa@ua.pt

This article is an editorial note submitted to CCR. It has NOT been peer reviewed. Authors take full responsibility for this article's technical content. Comments can be posted through CCR Online.

ABSTRACT

The hourglass concept has been undisputable ruler of networking visions on the last years. As network evolution is now a hot topic, this article aims to reflect on this concept, highlighting its current shortcomings, and suggesting its evolution.

Categories and Subject Descriptors

C.2.0 [Computer Systems Organization]: Computer-Communications networks: General

General Terms

Design, Theory

Keywords

Network design, Network architecture

1. INTRODUCTION¹

Currently IP networks have taken the center stage on all communication networks. From a past where IP was a niche technology for specific environments, with islands communicating through links relying on completely different technologies, IP-based technologies have gradually imposed themselves on all types of networks. This has led to many novel challenges for data networks, but often to a misuse of (all) concepts associated to networks. The “all-IP” network concept, appearing associated with novel wireless networks, or the “future-Internet” research terms, are a good example of how established IP-network concepts have become both for the researchers on the area, and to all communication professionals. Unfortunately, with this recognition, many concepts became widespread without any critical review on why they appeared, and how (or if) they were still valid. Examples of these are the discussions on the end-to-end principle [3-4], or the discussions on the different usage of IP addresses [2], or finally, the so much taunted hourglass architecture concept [1].

The “hourglass” concept basically claims that all networks should be based on IP-technologies, since it is “obviously the single layer where all communication converges to a single protocol” (See Fig. 1, with some art and technical liberties).

¹ This discussion document is based on a talk delivered to NeXtworking'2007.

Unfortunately, this type of comment suits itself to misuse – which we should be careful to avoid when considering any evolution of the network. The following tries to present some potentially ideas on the reality of this hourglass argument, and its usefulness for the ongoing development of the network.

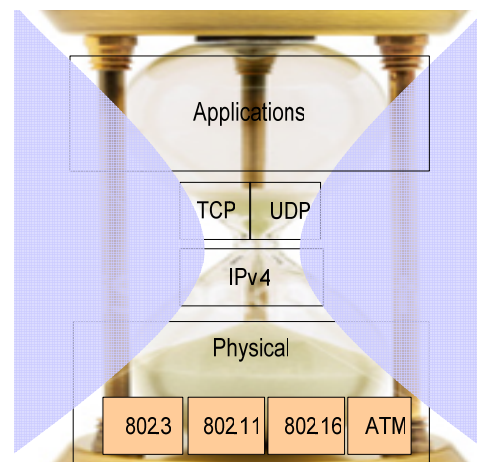


Figure 1. The hourglass vision of the IP-stack.

2. CONTROL AND DATA

Albeit there are several ways of looking into a communication system, one of the most convenient is to consider what happens in terms of the data and the control paths. The data path refers to the set of units/protocols that actually are related with the flow of information, while the control path refers to the set of units/protocols which may impact the behavior of the elements associated to the data path. The hourglass architecture refers in practice to the data path, and avoids control path issues altogether.

Control path was not really an issue for IP networks. The ideas behind the inception of IP networks were basically of *best-effort* coupled with in-band signaling. The idea of packet based communication implicitly relied on in-band control, with the control path becoming (apparently) equal to the data path: the packets transported all the information required for its processing at all network units (routers). So, gradually, the notion of control path, essential to other communication technologies, apparently became redundant. Naturally, this was not completely correct - after all, there were internal and external routing protocols already

defined, which provide out-of-band control information to the routers – but was a convenient way of portraying the advantages of the novel communication concept at the time: packet base communications.

This portrait was valid under a simple set of features: i) all (most) of the information required for the control of the data packet was provided inside the packet itself; and ii) the control actions to be performed on the packet were independent from packet to packet (mostly, these actions were the decision on which interface to forward the packet to). So the hourglass description was an adequate approximation on the early days of IP-networks. No separation between the control and data paths were necessary to consider, below we could have whatever technology desired, and above we could have whatever application needed.

The IP network layer, responsible for transferring information hop-by-hop across all the networks from a source to a destination, and the TCP/UDP transport layer, responsible for performing an end-to-end service according to some expected behavior, became gradually to be seen as the only relevant layers for communications. IP networks were an efficient, lean, vehicle for the movement of packets from one point to another a bicycle.

By nature, the network layer is the one responsible for the transferral of information across all the networks. Although several protocols could co-exist (and did for some years), the simplicity of having a single protocol at that layer overcomes other considerations that may exist for different protocols, and led to the gradual universality of the IP protocol. We want a network layer in order to be able to communicate across the whole world, with a consistent addressing, and we want a single network layer to maximize interoperability, reduce interworking complexity and minimize failure possibilities. Thus IP networks become so popular that the prevalent vision of communications stack became the so-called TCP-IP stack, which condensed everything above the transport layer simply as the “application” layer, instead of the some-times controversial functional breakdown of the the OSI stack, with its presentation and session layers [5]. Discussions on control versus data plane simply faded away.

3. CONTROL STRIKES BACK

Currently, data networks are no longer the simpler best-effort packet networks that were the reference for early IP development – and current trends indicate their complexity will continue to increase [6]. IP networks have been evolving during these years trying to cope with new requirements, which continuously appeared as IP-infrastructure gradually began to replace all other communication infrastructures. Society fought back technology concepts, and the simple efficient bicycle was pulled in multiple directions – most of them at the control plane.

3.1 Multimedia

Multimedia communications was one of the key drivers leading to new developments on IP-networks. Multimedia requirements led to changes on the transport and application layer, with the introduction of new protocols (like RTP and SIP [7]) to overcome some of the deficiencies of the original TCP/IP stack. The challenges were on mechanisms able to provide a network layer adequate to transport different multimedia codecs. In fact, the IP bicycle had to be developed into a container truck, able to carry the adequate load for its end-points.

Interestingly, some of these protocols led to some sort of session concept, as an overlay, albeit such notion was not recognized in the traditional IP stack represented in Figure 1. Although parts of these changes were carried as well inside the packets (the RTP header, e.g.), added control information (packets) were required for the adequate operation of this multimedia environments. In SIP-based, e.g., most data packets will not transverse the same set of links than the control packets, due to the introduction of new functional entities in the network. Here, new entities (SIP proxy, e.g.) appeared dedicated to the control of the network *per se*. The best-effort network became a multimedia service network, and the simple all-purpose communication infrastructure developed specific entities for specific types of communication.

3.2 Multicast

A simpler aspect, often mistook as a simple evolution of the already existent (and neglected) routing control path, was the introduction of multicast routing, with protocols like PIM and IGMP. While multimedia mostly increased the complexity of the end points, and new functional entities in the network, multicast routing led to an increased complexity on the routers (the data path units) – its control flow, originally based only on the final destination, now became dependent on processing two different tables, handling an indirection for each packet. The router no longer understands the final destination, but only the interfaces (which may be any number) where to process that packet. This is not surprising: after all, our IP bicycle now had to behave as a bus, delivering packets to multiple destinations. The simpler routing action is no more.

3.3 Quality of Service (QoS)

QoS became an issue in IP networks mostly by the increased usage of real-time communications in many production networks. The data plane had now to comply with multiple restrictions, and a different set of protocols were developed to inform routers of how to change the actions to perform to each data packet. Following this trend of increased intelligence on the routers, QoS came also into play in routers, processing end-user signaling protocols like RSVP (or those associated with the NSIS framework), and control protocols like COPS.

Routing actions were no longer simple forwarding, but became prioritizing, dropping, marking, and so on, in able to overcome all contention aspects arising from infrastructure sharing by many users. The issues of multicast router complexity paled compared with these new evolutions – with increasingly complex control planes. In fact, new functional entities were considered in such networks (such as bandwidth brokers [8], or policy managers).

The bicycle was now supposed to morph into a luxury vehicle for some packets, and to keep its unstable balance behavior for many others.

3.4 Security

The introduction of security, in multiple aspects, was yet another major change to the original concept of a simple network. The network had to provide (just to name a few) information confidentiality, integrity and access control, all of which led to large changes in all the communication entities, from end-points to routers - and even with the introduction of new functional entities (like authentication servers or firewalls).

Protocols like IPSEC (and the IKE framework), or Radius and Diameter, were deployed to provide a new set of security functionalities on the network. Once again, with these protocols packets became interrelated, either explicitly from the source point of view, or implicitly by the network operation - control actions were again applied to large sets of packets. Overall, the simple bicycle had to provide the transport assurances of a tank.

3.5 Mobility

Last, but not least, mobility changed the face of IP networks. Current society is mobile. And as such, this was a requirement posed to IP networks –albeit not deployed in most networks, mobility protocols has already entered the standardization track of major fora. Mobility changed both the data and control planes, defining new usages of the IP protocol – the identifier/locator dichotomy [9] is a good example of this.

Besides the basic Mobile IP protocol, now on deployment plan for future mobile networks, fast mobility protocols appeared (most specially for IPv6, such as [10]). At this level, all the discussion on data or control plane becomes difficult to separate. For optimum performance, most of these protocols operate at the data plane, but with temporary changes on the control plane, in order to handle the basically slow operation of Mobile IP. Our bicycle has to perform as a sports car.

This mobility problem became compounded by the development of multihoming concepts [11], leading to a much higher complexity of the interconnection of a device than it was ever considered twenty years ago.

4. HOURGLASS TIMING IS OVER?

When discussing the evolution of future networks, accepting the existence of multiple technologies is obvious. Having a single, consistent, network and transport layers is of paramount concern, as discussed, and IP networks seem to be the only technology that is in position to deliver that universality. But with the current large requirements we have for data networks, the assumptions

that led to a neglect of the control path in the early IP networks are no longer sustainable: we have sessions (security, multimedia, QoS, etc..) of interrelated packets on the network, and each packet does not contain in itself all the information required for its processing.

So a concept like the hourglass should be reconsidered. While in 2001, Steve Deering was worried that the hourglass was getting fat [1], now we should consider in what sense its existence can be sustained.

4.1 Do we have a hourglass architecture?

The advantage of having a simple common layer understandable across the whole network is undeniable. But this simple network layer does not exist for a long time. It is true that we have a single simple layer for interoperation of networks – in fact this layer simplicity is the major hurdle to the global development of new features on the network.

As soon as we consider all the aspects mentioned in the previous section, then the network can hardly be claimed to have anything near an hourglass. The analysis of the control plane sheds a completely different light on this “hourglass figure” (see Figure 2, where a large number of approximations were taken by graphic representation limitations).

The control plane is actually thicker on the middle layers – because for each novel requirement posed to the network, the solution has been the development of new protocols, aiming to handle optimally each one of those aspects. The result is that our simple “bicycle” is no longer simpler, but it is also not a truck, or a bus, or a tank, or a sports or luxury car. The original bicycle that moved packets between any two end-points, became, according to the circumstances, quite different transport vehicle: our network is no longer global– the amount of control protocols required for all these features are so large that no widespread deployment across the network can be found.

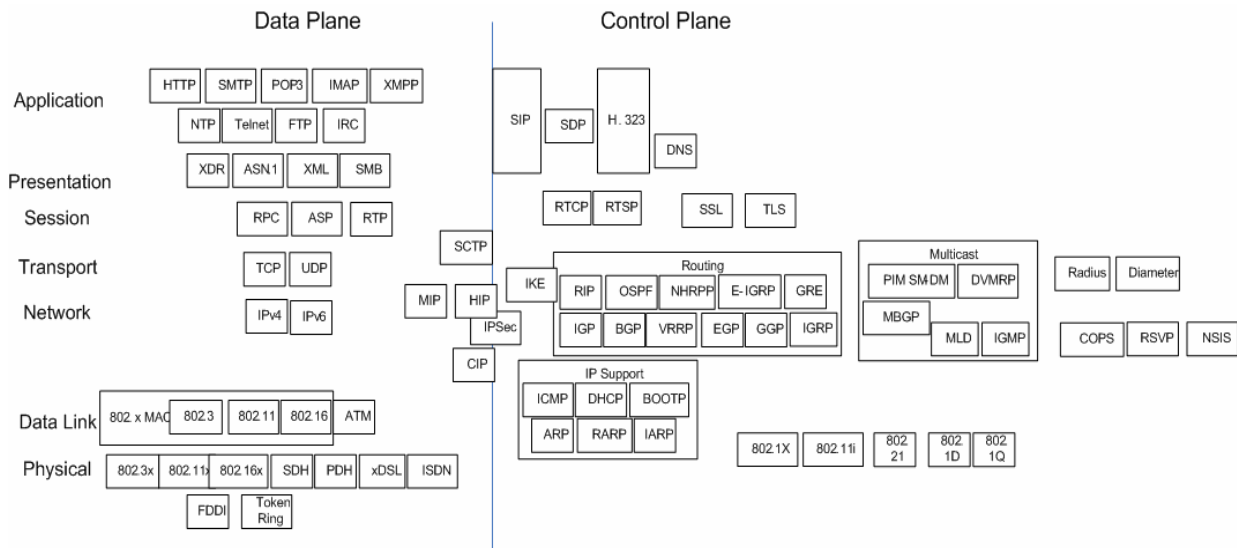


Figure 2. A liberal view on IP data and control protocols.

We have now a segmented network, with different features in different parts of the network. Minimal communication is the only global assurance – much like the early days of IP deployment.

Curios enough, the wide deployment corresponds to the “hourglass” side: mostly the data plane, with minimal control features.

4.2 Do we need two hourglasses?

The original hourglass concept, of developing a common layer able to interact across the whole network, is still interesting, and its simplicity is still needed for widespread development.

However, the hourglass argument relied on a joint vision of the control and data plane. This is not possible anymore – in fact the control plane is not simple. We do not have a common functional vision on the network, with different entities in the control plane according to the requirements to fulfill.

The evolution of IP networks should again strive towards simplicity of the control plane – although not necessarily jointly handled with the data plane. Note that as bitrates increase, the overall processing per packet becomes prohibitive, specially when packets are structured in flows (or flow aggregates), related with (user or application) sessions. Separate control concepts are then needed, with a control plane configuring actions to perform on each session. Although packet processing at the routers will be more than simple forwarding, no hard processing should be required to understand the actions to perform.

Even in this condition, we need to retake the original objectives of IP: a simple, common layer able to maximize interoperability, reduce interworking complexity and minimize failure possibilities. However, this cannot be achieved by neglecting the large amount of diversified usage seen in the network today.

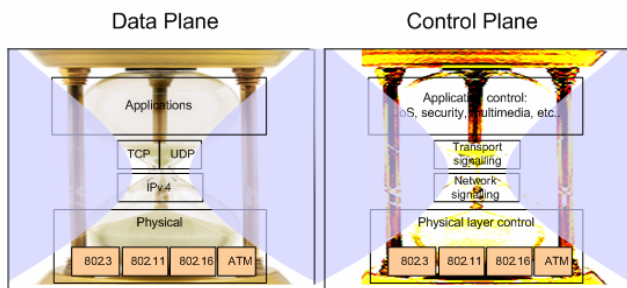


Figure 3. Desired IP-stack: two hourglasses.

A novel control plane, able to cope with the different requirements we are currently witnessing needs to be in place, and able to operate inter-network(s). Part of these objectives was taken in the NSIS framework, but this was too much centered in QoS-related issues. The development of this control plane should lead to clear, simplified, interactions across networks, covering the multiplicity of tasks being requested today – developing in effect the same hourglass concept into the control plane as well

(see Figure 3). The best architecture to handle the complexity associated with this approach is an interesting challenge.

5. CONCLUSIONS

Discussions on the future of IP networks reuse multiple arguments from its inception. This discussion aims to reflect on the current reality of the so-called hourglass concept, and on its evolution for the future. Overall, the hourglass concept assumed simple control planes, and no major deviation between the control and data plane. Current requirements on networks are not compatible with these simple control assumptions. The simplicity of a common protocol for the network should be kept, but this needs to incorporate control plane aspects as well. Further discussion on the issue is welcomed!

6. REFERENCES

- [1] DEERING, S., Watching the waist of the .protocol hourglass, IETF 51, London August 2001.
- [2] CARPENTER, B., CROWCROFT, J., AND REKHTER, Y., RFC2101 - IPv4 Address Behaviour Today, February 1997
- [3] SALTZER, J.H. , REED, D. P. , AND CLARK, D. D. , End-To-End Arguments in System Design, *ACM TOCS*, Vol 2, Number 4, November 1984, pp 277-288.
- [4] CARPENTER, B., AND BRIM, S. , RFC 3234 - Middleboxes: Taxonomy and Issues, February 2002.
- [5] ISO/TC97/SC16, “Reference model of open systems interconnection”, Doc N227, June 1979.
- [6] PATEL, G. AND DENNETT, S., The 3GPP and 3GPP2 Movements Toward an All-IP Mobile Network, *IEEE Personal Communications*, Aug 2000, Volume: 7, Issue: 4, page(s): 62-64
- [7] ROSENBERG, J., SCHULZRINNE, H. *et all*, RFC 3261, SIP: Session Initiation Protocol, June 2002
- [8] NICHOLS, K., JACOBSON, V., AND ZHANG, L. RFC2638 - A Two-bit Differentiated Services Architecture for the Internet, July 1999
- [9] FALK, A., The IETF, the IRTF, and the networking research community, *ACM SIGCOMM Computer Communication Review* Volume 35 , Issue 5, October 2005, pp: 69 - 70
- [10] KOODLI, R., (Ed)., RFC 5268 - Mobile IPv6 Fast Handovers, June 2008.
- [11] ATKINSON, R, BHATTI, S., HAILES, S., A proposal for unifying mobility with multi-homing, NAT, & security, *5th ACM international workshop on Mobility management and wireless access table of contents*, 2007 Chania, Crete Island, Greece, Pages: 74 - 83