INTERNET POLICY REVIEW
Journal on internet regulation

Volume 3 | Issue 1

Decentralised internet governance: the case of a 'peer-to-peer cloud'

Francesca Musiani
MINES ParisTech, France, francesca.musiani@mines-paristech.fr

Published on 13 Feb 2014 | DOI: 10.14763/2014.1.234

Abstract: This article retraces the early stages of development of the 'peer-to-peer cloud' storage service Drizzle, with the aim of providing an example of decentralised network architecture as internet governance 'in practice'. More specifically, this paper sheds light on how changes in the architectural design of networked services affect the circulation, storage and privacy of data, as well as the rights and responsibilities exerted by different actors on them. This article does not mean to be a compendium of the implications of the decentralisation option in building a cloud platform, which entails a number of technical complications as well as advantages, including how to ensure the reliability and redundancy of data, and the soundness of the encryption mechanism. However, the privacy-related design choices described here are some of the many possible ways to illustrate the extent to which changes in network architecture are, indeed, changes in network governance.

Keywords: Internet governance, Science and technology studies (STS), Peer-to-peer (P2P), Decentralisation, Privacy by design, Internet architecture, Cloud

Article information

Received: 30 Nov 2013 **Reviewed:** 10 Feb 2014 **Published:** 13 Feb 2014

Funding: This work is supported by the ANR project ADAM - Architectures distribuées et applications multimédia and by the FP7 STREP project P2Pvalue - Techno-social platform for sustainable models and value generation in commons-based peer production in the Future Internet
Competing interests: The author has declared that no competing interests exist that have influenced the text.

URL:
http://policyreview.info/articles/analysis/decentralised-internet-governance-case-peer-peer-cloud

Citation: Musiani, F. (2014). Decentralised internet governance: the case of a 'peer-to-peer cloud'. *Internet Policy Review*, *3*(1). DOI: 10.14763/2014.1.234

The architecture of a networked system is its underlying technical structure - its logical and structural layout. In my last article for the *Internet Policy Review* (Musiani, 2013a), I have built upon the work of several authors in science and technology studies, economics, law and computer science (e.g. Star, 1999; van Schewick, 2010; Elkin-Koren, 2006; Agre, 2003) to discuss the idea of network architecture as internet governance. I have suggested that, by changing the design of the networks subtending internet-based services and the global internet itself, the politics of the *network of networks* are affected – the balance of rights between users

and providers, the capacity of online communities to engage in open and direct interaction, the fair competition between actors of the internet market.

This article retraces the early stages of development of a 'peer-to-peer cloud' storage service, Drizzle, with the aim of providing an example of decentralised network architecture as internet governance 'in practice'. More specifically, the paper sheds light on how changes in the architectural design of networked services affect the circulation, storage and privacy of data, as well as the rights and responsibilities exerted by different actors on them. This article does not mean to be a compendium of the implications of the decentralisation option in building a cloud platform, which entails a number of technical complications as well as advantages, including how to ensure the reliability and redundancy of data, and the soundness of the encryption mechanism. However, the privacy-related design choices described here are some of the many possible ways to illustrate the extent to which changes in network architecture are, indeed, changes to network governance.

DECENTRALISING THE CLOUD

In early 2007, when Drizzle first sees the light, the industry of online data storage - a service allowing users to store, save and share data on one or several terminals connected to the internet - has "never felt better" (Guerrini, 2010). Google, Amazon, Microsoft and Oracle, to name but a few, propose their storage platforms, each with its specificities and one common denominator: the 'cloud'. According to this model, the service provider is in charge of both the physical infrastructure and the software. Thus, the service provider hosts applications and data at once - in a location, and according to modalities, unknown or at best ambiguous to the user (Mowbray, 2009). The so-called 'server farms' proliferate, to support and manage this increasing *remoteness* of data from users and users' terminals.

In this context, Drizzle[1], a small start-up founded by two developers and computer programmers who we will call Dietrich and Kurt, makes an unusual foundational decision: its cloud storage platform will mainly be composed – alongside more 'classical' data centres – of portions of the users' hard disks, directly linked in a peer-to-peer, decentralised network architecture (Schollmeier, 2001; Taylor & Harrison, 2009). This choice entails a number of peculiar features. On the one hand, the implementation of a technical process defined as "encrypted fragmentation"[2], which consists in encrypting locally – on the user's computer, and by means of a previously installed Drizzle P2P client – the content that will be stored. The content is then divided into fragments, duplicated to ensure redundancy, and spread out to the network. In return, users need to accept to 'pool' - put at the disposal of other users and their computers - the computational and material resources necessary for the operations related to the storage of content. As the service's terms of use point out:

> "The user acknowledges that Drizzle may use processor, bandwidth and hard disk (or other storage media) of his computer for the purpose of storing, encrypting, caching and serving data that has been stored in Drizzle by the user or any other users. The user can specify the extent to which local resources are used in the settings of the Drizzle client software. The amount of resources the user is allowed to use in Drizzle depends on the amount of local resources the user is contributing to Drizzle."

The interdependent and egalitarian model subtending the platform will allow its users to barter their local disk space with an equivalent space in the decentralised cloud, thereby improving the quality of this storage space, which will become permanently available and accessible. By shaping their decentralised storage service, the developers of Drizzle carry on a double experimentation: with the frontier between centralisation and decentralisation, and with sharing modalities that blend peer-to-peer, social networking and the cloud.

PEER-TO-PEER STORAGE: THE CLOUD MEETS *PRIVACY BY DESIGN*

"In 2007, it was all starting to get social," Dietrich recalls three years later. Indeed, social media, Facebook and Twitter in particular, were at that moment entering the daily life of millions of internet users in an increasingly pervasive way. Drizzle's first steps are taken in a community of research and development that tries to counter the social media "explosion" by developing P2P systems as an alternative to a variety of internet-based services, including social networks, structured in a centralised manner (Le Fessant, 2009; Musiani, 2010a; Musiani, 2010b).

In 2007, Facebook had been in existence for three years. Millions of users had taken part in it, thereby contributing to the massive success of these Web-based services that allow individuals to build a public or semi-public profile within a system, define a list of other users with whom to interact, and see/browse the list of their and others' connections made in 'public mode' within the system (Boyd & Ellison, 2007). In parallel to their spectacular growth, social networks raise vibrant discussions and controversies, both within the expert community and among the general public. The ways in which social networking service providers leverage personal information and user data remains controversial, since they sometimes mean allowing external applications to access them, while on other occasions they pursue direct commercial purposes (Boyd, 2008). The rise of the so-called *cloud* does nothing to mitigate the impression of risk for informed users, as applications and data are increasingly hosted in locations and ways unknown or at best ambiguous. User exposure on social networking sites and on cloud-based services positions privacy, more than ever, at the foreground of discussions.

In this context, several developers – including Drizzle's – identify in a peer-to-peer type of network architecture a possible way of approaching the protection of personal data privacy with a different angle: through the relocation and "re-appropriation" of data within the terminals of users, who would be able to host their own profiles and the information they contain (see also Moglen, 2010; Aigrain, 2010, 2011).

As in the development of Drizzle, a conception of privacy and confidentiality of personal data, which is conceived of and enforced via technical means – called *privacy by design* (Cavoukian, 2010; Schaar, 2010), is at work. This conceptualisation of privacy is defined by means of the constraints and the opportunities linked to the treatment and the location of data, according to the different moments and the variety of operations taking place within the system. In particular, the confidentiality of data (personal data as well as the content stored in the P2P cloud) is defined by a peculiar role and enhanced features attributed to the password that identifies the user *vis-à-vis* the network, and by the implementation of the resource allocation system on which Drizzle is based.

PASSWORD AND USER RESPONSIBILITY

In Dietrich's intentions, the role of the user-selected and user–generated password for the Drizzle system should have "stri[cken] the user as soon as he had access to the system for the very first time." Indeed, the virtual form that is served to users upon subscription may come as a surprise: it informs that

> "We do not know your password as it never leaves your computer. Please, do not forget your password and use, if needed, your password hint."

The status of the password is thus negotiated, beyond its usual meaning of unique identifier *vis-à-vis* the system, to define, detail and legitimise the process of local encryption and decryption of data within the Drizzle system. This feature comes to symbolise the specificity of Drizzle's promise of security and privacy as well as users' trust, as it becomes the symbol and the graphical representation of the 'local' dimension of the encryption process – as it never leaves the computer of the user who created it. The operations, for the most part automatically managed, that are linked to the protection of personal data are thus hosted on the terminals of users. Indeed, this entails a modification of the user's role within the service's architecture: node among equal nodes, it becomes a server itself, instead of a starting point and a final point for operations that are otherwise conducted on another machine or group of machines.

Through the attribution of this status to the password, the developers of Drizzle are also proposing an alternative to the balance between the rights exerted by users on their own data and the rights acquired by the service provider on these same data – a balance that is usually heavily bent on the provider's side. However, this reconfiguration in the balance of rights comes with a trade-off. As the password stays with the user and is not sent to the servers controlled by the firm, the latter cannot retrieve the password if needed. Thus, users do not only see their privacy reinforced, but at the same time and for the same reasons, the responsibility for their actions is augmented – while the service provider renounces to some of its control on the content that circulates thanks to the service it manages. The meaning of this 'renunciation', Dietrich explains, is double: on the one hand, the Drizzle team wishes to make it evident, almost *translate* into a specific object the user can easily relate to, the 'obscure' and unfamiliar process of client-side encryption, which is an ongoing source of controversies and perplexities. On the other hand, it is also a matter of Drizzle's business model: the more the firm knows about its users, the more it is mandatory for it to submit the users to regular surveillance and control – and this requires an investment of material resources and time that, in its first phases of existence, the firm does not have:

> "If we can know what is in your account, starting with your password, we have heightened obligations to police the content and to make sure nobody can eavesdrop on the traffic."

DATA PRIVACY AND RESOURCE ALLOCATION

Another aspect that contributes to define rights and responsibilities is the detailing of the conditions for allocation and management of the computational resources provided by the different computers participating in the system.

As briefly described above, the choice to decentralise the platform makes it necessary, due to the very particular status of the resources used by the system, to detail several aspects in the terms of use: the role of computers belonging to users, the types of resources that Drizzle is able to use, their purpose. It also becomes necessary to detail the extent to which users are able to decide – and communicate to their P2P client, thus to the system – the maximum quantity of local resources that the rest of the network/storage system can use. However, it is also necessary to define the articulation between the availability of resources and the different operations to which these resources will be destined to within the system.

The articulation of these two aspects has important implications for the confidentiality of data circulating in the system (both personal information and content stored by users). Several users, giving feedback to the developers in the early stages of the system, warn that the resource allocation process could be framed as a possible 'surveillance' or 'monitoring' of these resources, in a way that can potentially be highly automatised, invasive, privacy-threatening.

After a discussion between these concerned users and the developers, *via* the Drizzle forum, two modifications were applied to the terms of use: while the general terms now state that "resources are allocated and monitored in accordance with the Privacy Policy," the privacy policy itself details the extent of automation and pervasiveness of the system that allocates and monitors resources:

> "In order to ensure a fair allocation of resources within Drizzle, various data about the computers participating in the Drizzle network is collected. This data includes their IP addresses, disposability and the amount of resources they are contributing (e.g. bandwidth, memory). [...] Drizzle keeps track of how much storage space you have used and earned [...] Drizzle collects statistical information for the purposes of monitoring, debugging and improving the system. This includes automatically generated problem, performance, network analysis and general usage reports, as well as logs of the connections and queries made to Drizzle's servers (including the involved IP addresses), as well as analytical data about the usage of the Drizzle website. However, none of this data contains information from your private or shared files."

Thus, the correct functioning of the allocation system indeed implies the gathering of several pieces of information concerning the material, computational and memory resources pooled by each participating computer. The pooling of the storage equipment (i.e., users' local resources, made available by each of them) is necessary for the system to work; however, it is not meant to imply an intrusion in the stored content itself, which remains protected by the local safeguard of the password and the encryption of content. The collection of information, the developers of Drizzle affirm, has the purpose of automatically computing the storage space made available by each user – and, as we have analysed elsewhere (Musiani, 2013b), of establishing the extent to

which each user can reclaim her place in the 'P2P cloud', an equivalent storage space in the network of participating users.

CONCLUSIONS

The development of Drizzle's 'peer-to-peer cloud' allows to observe how changes in the architectural design of networked services affect data circulation, storage and privacy - and in doing so, reconfigure the articulation of the 'locality' and the 'centrality' in the network (Akrich, 1989: 39), suggesting a model of decentralised governance "by architectural design" for the service.

Ultimately, decentralising the cloud leads to a reformulation and 're-balancing' of the relationship between the user and the service provider. The local, client-side encryption of data first, and its fragmentation afterwards - both operations conducted within the P2P client installed by the user, and entirely taking place on his terminal - are proposed by Drizzle as evidence that the firm, in its own words, "does not even have the technical means" to betray the trust of users.

In particular, this conception of *privacy by design* takes shape around the password, that remains locally stored in the user's P2P client and unknown to the service provider. In doing so, it becomes a form of disengagement of the service provider with respect to security issues, its 'auto-release' from responsibility: a detail whose importance may seem small at first, but eventually leads to changes in the forms of technical solidarity (Dodier, 1995) established between users and service provider.

For the purpose of this article, I have focused in particular on aspects such as the strengthening of *privacy by design* and the increase in responsibility attributed to the user, arguably among the "positive" aspects of a peer-to-peer cloud. However, it should be pointed out that an important part of the decentralisation choice made by the Drizzle team has involved assessing its possible downsides: reliability and redundancy of data, slow downloading performances, soundness of the encryption mechanism, and - no less important - the perception of these issues by users. A heated discussion among developers, and between developers and some pioneer users, also occurred on the topic of the 'legality' of the system, especially in jurisdictions such as that of the United States. All of these are complex issues and most of them could not be accounted for here - it has been done in a much more detailed manner elsewhere (Musiani, 2013: 123-173), by analysing, with tools derived from the field of science and technology studies (STS), a number of socio-technical controversies related to the development of the platform. However, the privacy-related dynamics provided here are a few of the several possible ways to flesh out the extent to which changes in network architecture are, indeed, changes in network governance.

The example of Drizzle has illustrated in practice the implications of 'architectures as governance' we had introduced in the previous article: the repartition of competences and responsibilities between service providers, content producers, users and network operators; the articulation between the individual and the collective; the shaping of user rights and 'community' norms; the definition of 'contributor' in internet-based services. In light of Edward Snowden's leaks about certain surveillance practices by the US National Security Agency, the potential of architectural choices - choices that would make the internet less centralised and more distributed - as a means of *de facto* privacy advocacy and promotion of decentralised

governance has never been more evident. The goal, as *The New Yorker* recently reported, “isn’t to end surveillance, but to make it harder to do en masse” (Kopstein, 2013).

FOOTNOTES

1. The name is fictitious ('light rain') and recalls the fragmentation and the distribution of data in the system’s storage mechanism. The names of the developers are pseudonyms, as well. I have no direct interest in Drizzle - I use it as a case study of a possible 'decentralisation of the cloud'.

2. Unless otherwise noted, citations are derived from in-depth interviews with the developers of Drizzle, conducted within a period of online and 'live' ethnography of Drizzle’s development, design and innovation process (see Vinck, 2003) between 2010 and 2011.

REFERENCES

Agre, P. (2003). "Peer-to-Peer and the Promise of Internet Equality." *Communications of the ACM*, 46 (2): 39-42.

Aigrain, P. (2010). "Declouding Freedom: Reclaiming Servers, Services and Data." In 2020 FLOSS Roadmap (2010 Version/3rd Edition), https://flossroadmap.co-ment.com/text/NUFVxf6wwK2/view/

Aigrain, P. (2011). "Another Narrative. Addressing Research Challenges and Other Open Issues session." *PARADISO Conference*, Brussels, 7–9 Sept. 2011.

Akrich, M. (1989). "De la position relative des localités. Systèmes électriques et réseaux socio-politiques." *Cahiers du Centre d'Études pour l'Emploi*, 32 : 117-166.

Boyd, D. (2008). "Facebook's Privacy Trainwreck: Exposure, Invasion, and Social Convergence." *Convergence*, 14 (1).

Boyd, D. & Ellison, N. (2007). "Social Network Sites: Definition, History, and Scholarship." *Journal of Computer-Mediated Communication*, 13 (1).

Callon, M., Lascoumes, P. & Barthe, Y. (2001). *Agir dans un monde incertain. Essai sur la démocratie technique*, Paris: Seuil.

Cavoukian, A. (eds., 2010). Special Issue: Privacy by Design: The Next Generation in the Evolution of Privacy. Identity in the Information Society, 3(2).

Dodier, N. (1995). *Les Hommes et les Machines. La conscience collective dans les sociétés technicisées*. Paris: Métailié.

Elkin-Koren, N. (2006). "Making Technology Visible: Liability of Internet Service Providers for Peer-to-Peer Traffic." *New York University Journal of Legislation & Public Policy*, 9 (15), 15-76.

Guerrini, Y. (2010). "Wuala : le P2P comme solution de stockage." http://www.presence-pc.com/actualite/Wuala-stockage-cloud-P2P-39035/#xtor=RSS-11

Kopstein, J. (2013). "The mission to de-centralize the Internet." The New Yorker, 13 December 2013, http://www.newyorker.com/online/blogs/elements/2013/12/the-mission-to-decentralize-the-internet.html

Le Fessant, F. (2009). "Les réseaux sociaux au secours des réseaux pair-à-pair." *Défense nationale et sécurité collective*, 3 : 29-35.

Moglen, E. (2010). "Freedom in the Cloud: Software Freedom, Privacy and Security for Web 2.0 and Cloud Computing." Keynote, ISOC Meeting, New York Branch, 5 February 2010.

Mowbray, M. (2009). "The Fog over the Grimpen Mire: Cloud Computing and the Law." *SCRIPTed*, 6(1): 132-146.

Musiani, F. (2013a). "Network architecture as internet governance." *Internet Policy Review*, 24 October 2013, http://policyreview.info/articles/analysis/network-architecture-internet-governance

Musiani, F. (2013b). *Nains sans géants. Architecture décentralisée et services Internet*. Paris : Presses des Mines.

Musiani, F. (2012). "Caring About the Plumbing: On the Importance of Architectures in Social Studies of (Peer-to-Peer) Technology." *Journal of Peer Production*, 1.

Musiani, F. (2010). "Ménager le droit à la vie privée, entre anonymat et connaissance de l'identité: les débuts des réseaux sociaux en pair-à-pair." *Terminal*, 105: 107-116.

Musiani, F. (2010b). "When Social Links Are Network Links: the Dawn of Peer-to-Peer Social Networks and Its Implications for Privacy." *Observatorio*, 4(3), 185-207.

Schaar, P. (2010). "Privacy by Design." Identity in the Information Society, 3(2): 267-274.

Schollmeier, R. (2001). "A definition of peer-to-peer networking for the classification of peer-to-peer architectures and applications." *Proceedings of the First International Conference on Peer-to-Peer Computing*, 27–29.

Star, S. L. (1999). "The Ethnography of Infrastructure." American Behavioral Scientist, 43 (3): 377-391.

Taylor, I. & Harrison, A. (2009). *From P2P to Web Services and Grids: Evolving Distributed Communities. Second and Expanded Edition*. London: Springer-Verlag.

van Schewick, B. (2010). *Internet Architecture and Innovation*. Cambridge, MA: The MIT Press

Vinck, D. (Ed., 2003). *Everyday Engineering. An Ethnography of Design and Innovation*. Cambridge, MA: The MIT Press.