

<u>The digital world during and after COVID-19:</u> Why low latency is essential in times of accelerated digitalization

by Dr. Thomas King - CTO DE-CIX

The world has undergone major changes in recent months. The outbreak and rapid spread of COVID-19 has meant the effect that unpredictable events can have on our everyday lives – whether privately or professionally – has become tangible. The result of this situation is that digital transformation must and is being strongly accelerated; under normal circumstances, this process would not have happened as quickly as it has in recent months. Despite all the challenges and limitations, it is reassuring to know that the structure of the Internet functions reliably and, to a large extent, supports both private and professional life – albeit from a distance. On the one hand, people can keep in touch privately with friends and family through video calls and other digital communication tools. Gaming and the streaming of video content also help to keep people entertained. On the other hand, in the world of work and especially business processes, the Internet plays an essential role – equivalent to water from the tap or power from the socket: Millions of employees and employers now use the possibility of working from home, supported by digital collaboration tools. Here too, video calls via Skype, Zoom or Teams are indispensable. Stable VPN connections guarantee access to the work-related materials stored digitally on the file servers in the company network. Internet Exchanges, such as DE-CIX (German Commercial Internet Exchange) in Frankfurt, take on the neutral role of a "data intermediary"; connected networks are securely and reliably interconnected via a direct connection.

In a survey¹ conducted by DE-CIX in Frankfurt – the world's largest Internet Exchange – among its own customers, 81 percent of the participants stated that latency is the most important criterion when concluding new interconnection contracts. What are the drivers behind this strong demand for latency? Here are some examples of technology areas for which low latencies are particularly important:

1. Classic: Web applications

All content that is called up in the web browser or via app should appear as directly and without delay as possible – this is the standard expectation nowadays. If there is a significant delay between user interaction and the presentation of the content, i.e. latency is perceptible, the user expresses impatience and irritation, and the user experience is perceived as poor. A study by Akamai² showed

¹ A Survey on the Current Internet Interconnection Practices (<u>https://www.de-</u>

 <u>cix.net/Files/2e74b254adad13951d5c604928a5d5a202fbdca1/A-Survey-on-the-Current-Internet-Interconnection-Practices.pdf</u>) - For the evaluation, a total of 106 ISPs, enterprises, CDNs and infrastructure providers were surveyed between December 2017 and January 2018.
<u>https://www.akamai.com/uk/en/about/news/press/2017-press/akamai-releases-spring-2017-state-of-online-</u>

retail-performance-report.jsp

long before COVID-19 that even a 2-second delay in the loading time of a website is sufficient to increase the bounce rate (bounce rate of visitors to a website) by more than 100 percent, and the trend towards such impatience is increasing. In addition, a delay of 100 milliseconds (0.1 seconds) in the loading time of websites would reduce the conversion rate (reaching predefined targets on a website, e.g. clicking on a video after reading the teaser text) by 7 percent. This shows that, from the user-experience perspective, latency plays a decisive role.

2. Cloud Gaming

One specific type of web application is cloud gaming, which must be handled separately in this context. Last autumn, Google Stadia was launched, heralding the start of cloud gaming on a larger scale. Until now, hard disk installations remain predominant for computer games. However, against the backdrop of the general trend towards clouds and "as a service" applications, it can be assumed that the gaming sector is also facing similar upheavals. So far, the amount of data transferred during gaming has been limited. Arithmetical operations, which are necessary to create the virtual worlds, still largely take place on the local system. But in cloud gaming, the computer game is run on a server in a data center and the screen content is streamed over the Internet to the user's monitor (e.g. tablet or laptop). This increases the demands on the user's Internet connection enormously. In addition to the additional bandwidth, the requirements for low latency increase dramatically, without which it is not possible to guarantee seamless gaming fun.

3. Virtual Reality

One of the most critical applications when it comes to latency is virtual reality. In order to provide a fluid experience, there must be as little time lag as possible between user actions and the reactions of the virtual environment. Otherwise, virtual reality is quickly perceived as disorientating and therefore annoying. The latency requirements in this range are around 20 milliseconds - to put this in context, a blink of an eye takes 150 milliseconds. Here is a small example of the consequence of this: The speed of light, and therefore the maximum speed at which data packets can be routed over fiber optic lines, is about 300,000,000 meters per second. Multiplying this value by the 20 milliseconds – or 0.02 seconds - longest acceptable time lag, results in a maximum distance of 6,000,000 meters or 6,000 kilometers. The distance from New York to Frankfurt as the crow flies is about 6,200 kilometers. This means that a virtual reality application hosted in New York could not be rendered smoothly in the German city. And this is only a reduced, purely theoretical calculation. In addition, other factors that cause a delay in the transmission play a major role, such as the processing time of the servers in the data center – this can easily take up 15 or even 20ms on its own. Ultimately, this means that users need to be within a few kilometers (<100km) of the hosted virtual reality application to be within the latency tolerance range. The solution to this situation is, firstly, edge computing: decentralized data processing in the immediate vicinity of the user, i.e. at the edge of the network. This could take the form of a mini-data center or exchange within the <100km latency radius of the users. In addition, cloud computing is indispensable: data processing takes place in the cloud and the data is directly available online. These approaches keep the latency as low as possible.

The bottom line: Latencies and the technologies of the future

In addition to the three examples mentioned, there are many other applications and technology areas in which latency plays an important or even decisive role. In the future, autonomous driving will be an integral part of our lives. Cars will sometimes need to make vital "decisions" on the basis of data, so it must be ensured that this data processing is direct and immediate. The theoretical latency requirement in an emergency is 0ms. Industrial robots – catchphrase: Industry 4.0 – must also be able to make fast, data-based decisions. The latency requirements for some applications are at values from 1ms to 10ms. All these examples show how important and necessary the development of latency-saving Internet applications is. We can expect exciting innovations in this area in the coming years. In future, it must be ensured that data is processed as close to the customer as possible, and securely hosted in the cloud.

###

About the author - Dr. Thomas King - Chief Technology Officer

Dr. Thomas King is Chief Technology Officer (CTO) at DE-CIX since 2018. Before, King was Chief Innovation Officer (CIO) at DE-CIX starting in 2016. He started his career as Member of Technical Staff at DE-CIX from 2008 to 2010 being responsible for the BSI IT-Grundschutz / ISO 27001 readiness program which resulted in a successful certification in 2010.

In 2010, Thomas King joined 1&1 Internet AG as a Product Manager responsible for mobile applications and mail. In 2011, he founded a startup company called audriga, in order to provide a service to automatically migrate email accounts from one hosting provider to another.

Dr. King re-joined DE-CIX as Head of Research & Development in 2014 and was promoted in 2016 to the newly-created position of Chief Innovation Officer at DE-CIX.

Thomas King received an M.Sc. degree in Computer Science and Business Administration from the University of Mannheim, Germany in 2004, and a Ph.D. degree at the chair of Computer Networks from the University of Mannheim in 2008.

