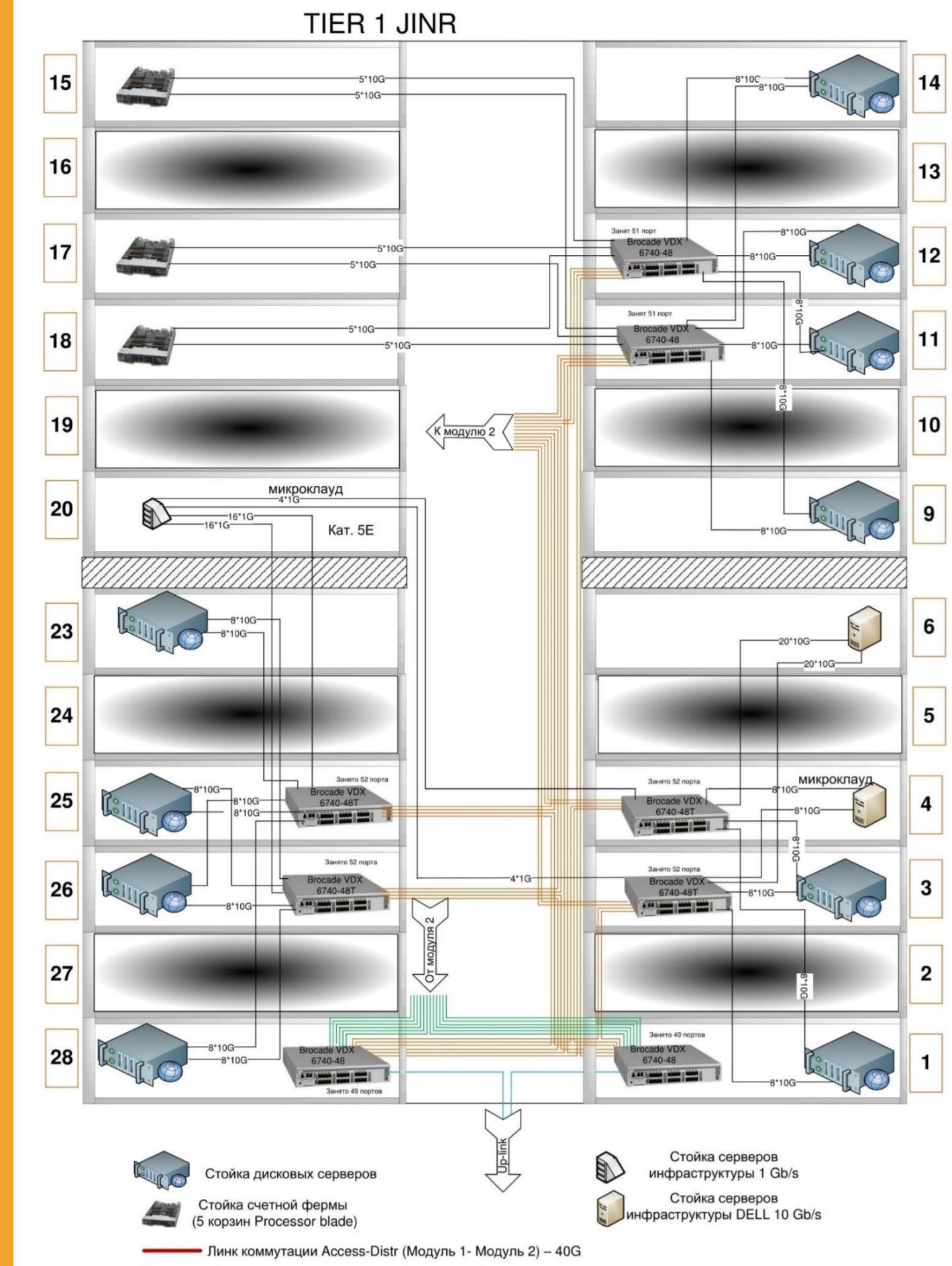




# Equal-cost multi-pathing in high power systems with TRILL

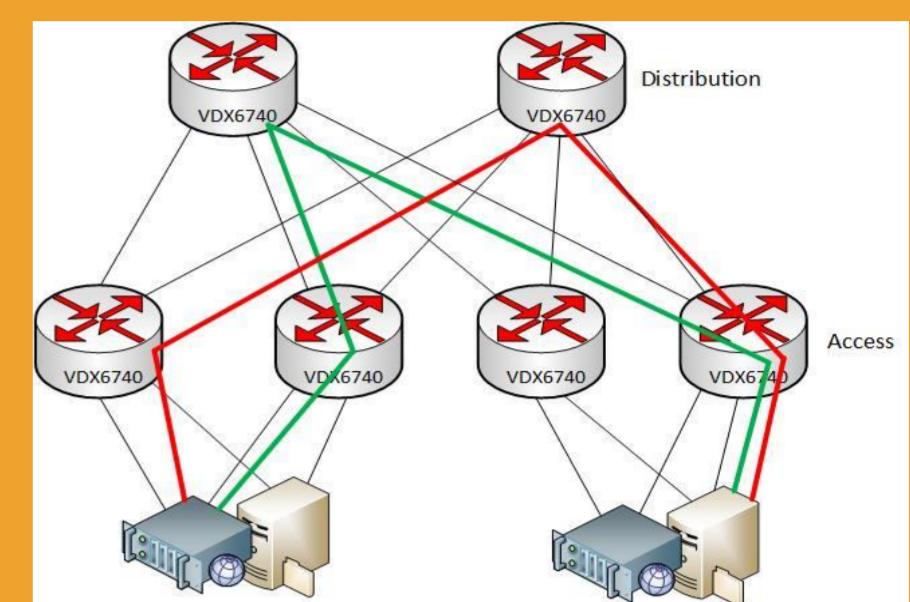
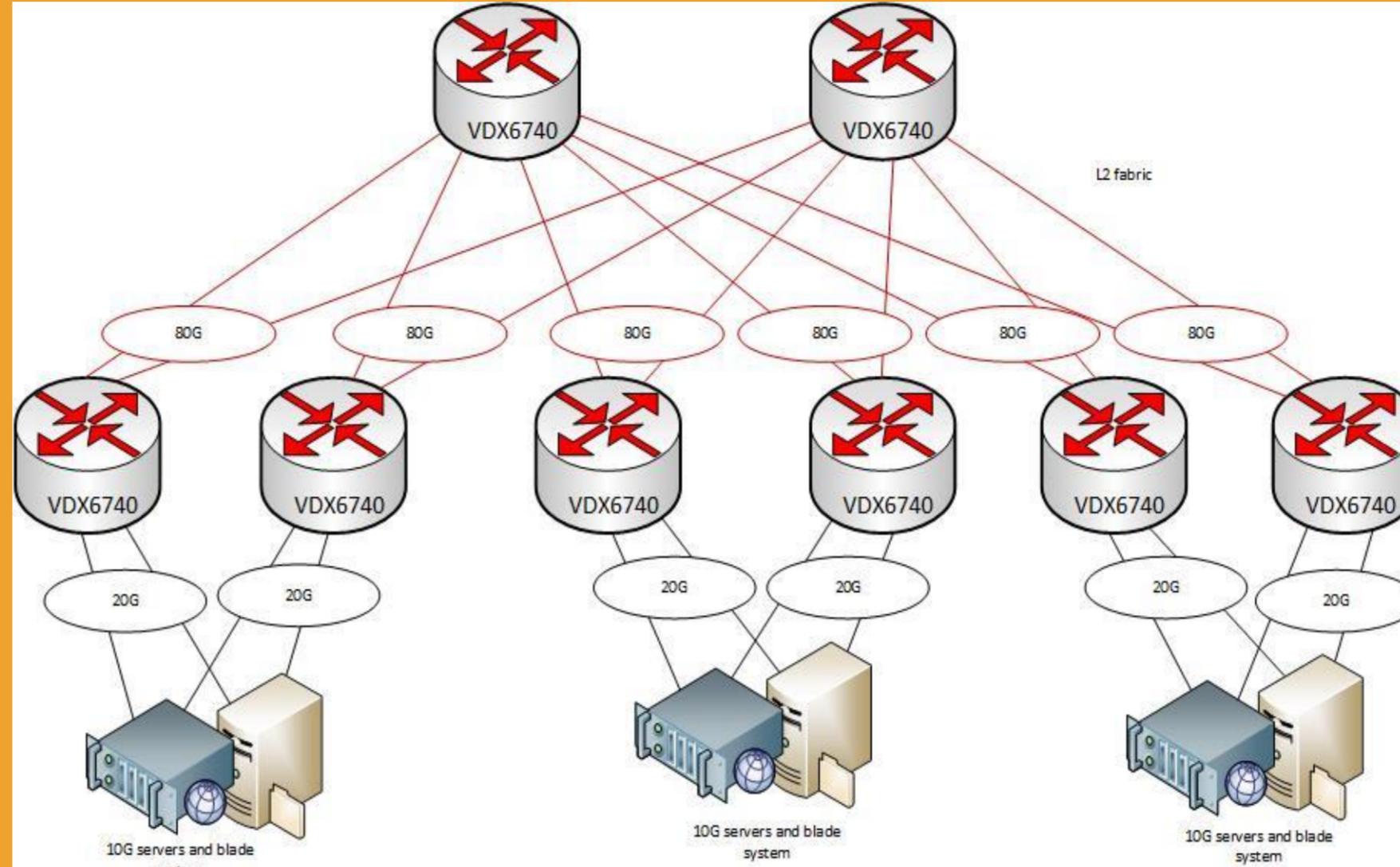
Joint Institute for Nuclear Research  
Laboratory of Information Technologies  
BAGINYAN, Andrey; KORENKOV, Vladimir; DOLBILOV, Andrey; KASHUNIN, Ivan

## Data processing center



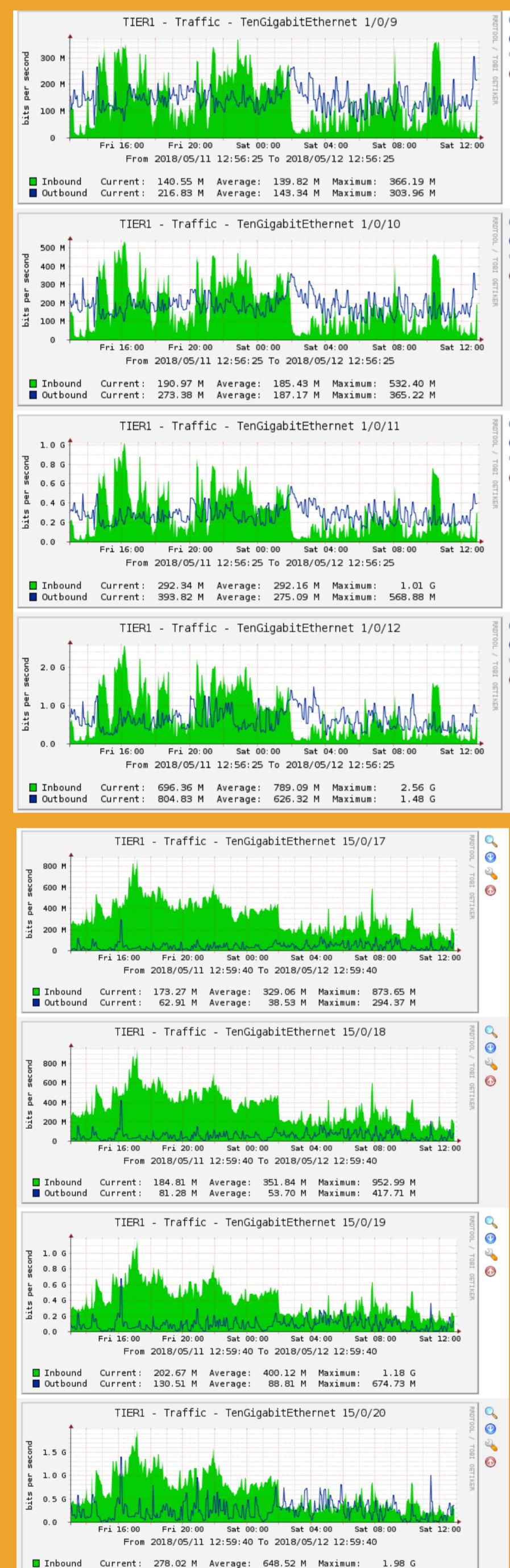
A full redundancy of links is provided at all levels [1]. As a result of such an architecture, a failure of one switch will lead to the reduction in the total traffic capacity of the network segment only by 25% [2]. In this case all servers will have access to the external network.

## SFP protocol and TRILL



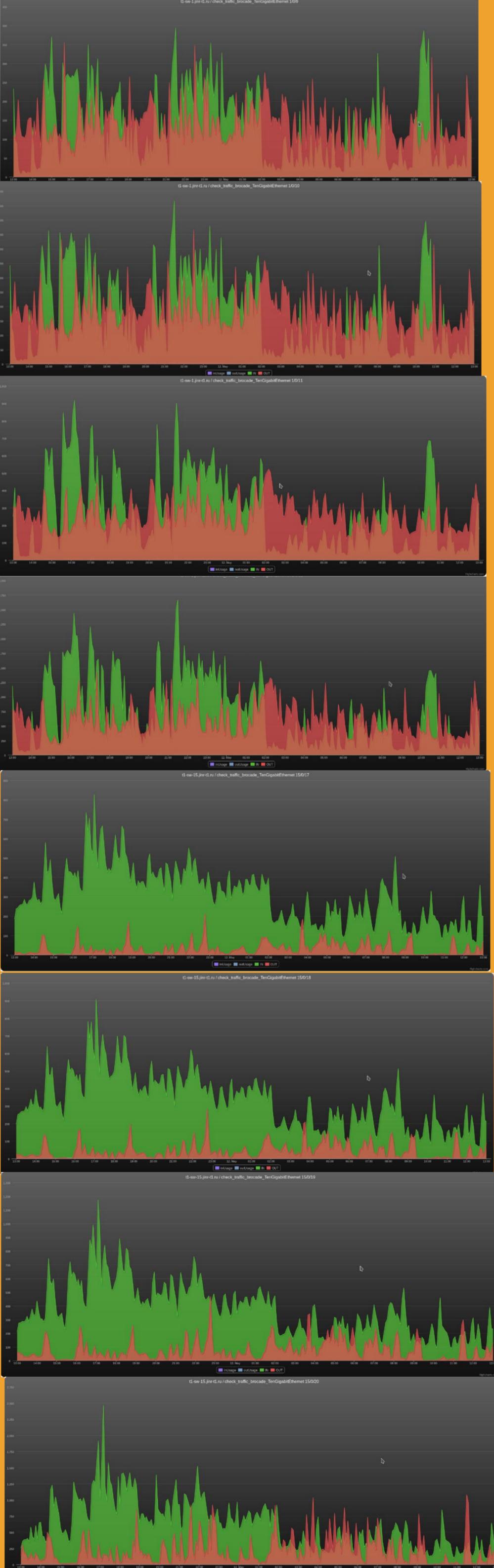
The TRILL protocol allows building an unblockable network architecture which fully provides for the total use of the network imperceptible for users and helps without prejudice to operated devices to involve new servers [3], because all active channels stay redundant.

## Monitoring system Cacti



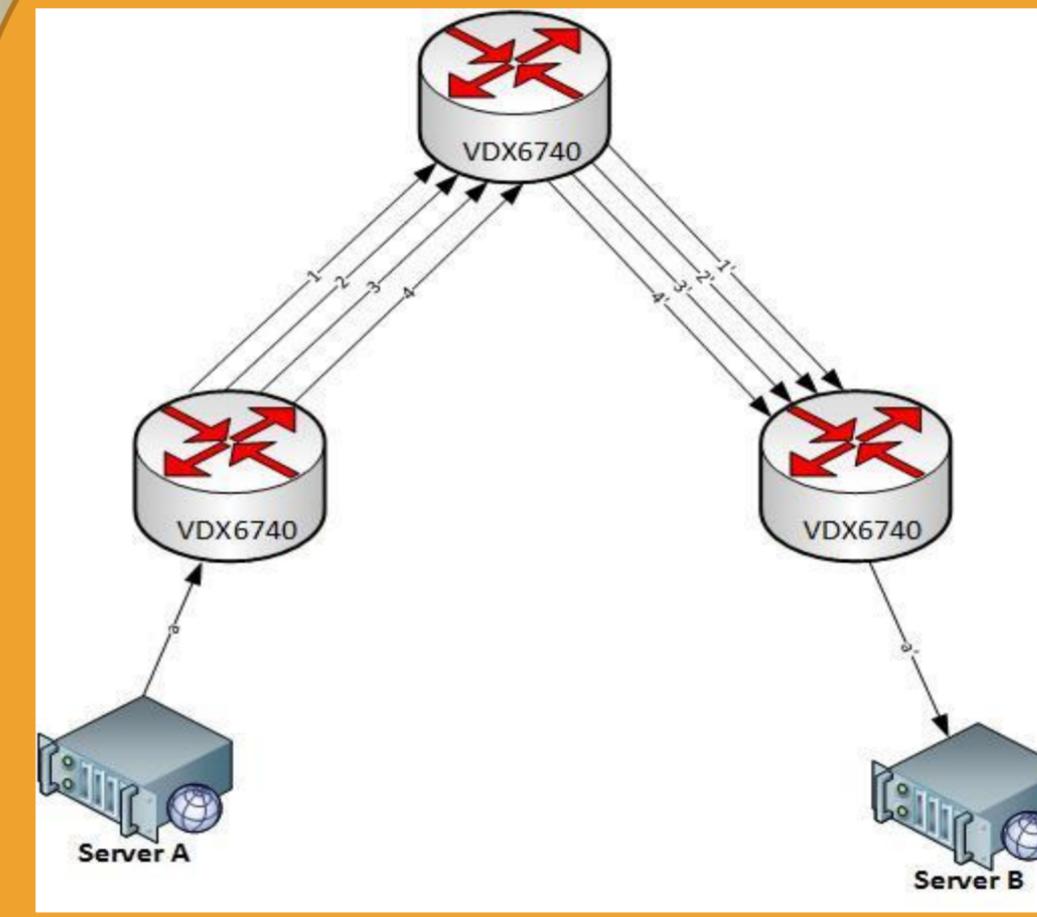
The curves analysis allows one to pay attention to a similar nature of the behavior of sent and received traffic via each communication line. Only vertical axis values differ, and this difference is not systematic. This shows a similar nature of curves' behavior. It turns out that our assumption about equal weight categories for this path is incorrect. To verify this assumption and exclude third party software products, a proprietary monitoring system for communication lines has been developed.

## Monitoring system Litmon



We can observe the similarity of the data acquired by monitoring system Cacti with Litmon. Moreover, the main extrema of the incoming and the outgoing traffic match within intervals. All this proves that the data acquired by monitoring systems are equal, but as we noticed earlier, the vertical axis values are not the same.

## Equal-cost multi-path



The communication between the access and the distribution levels has four 40G paths which will enable 160G data transmission. Each 40G communication line is a sum of four 10 G communication lines [4]. This means that the 40G SFP+ module contains 4 transmitters

optic signal each of which insures a two-way 10G data communication. A scheme of data communication between two data center nodes is shown. We may affirm that the path is a-1-1'-a' ~ a-2-2'-a' ~ 1-3-3'-a' ~ a-4-4'-a'.

Basing on a graph calculation according to Dijkstra's algorithm and on the patent claims for calculation of the shortest path "TRILL optimal forwarding and traffic engineered multi-pathing in cloud switching", we determine the equal cost of the path for all communication lines between the access level and the distribution level. Thus, we can conclude that the load on data channel via path a-1-1'-a' is equal to the traffic via path a-2-2'-a', 1-3-3'-a' and a-4-4'-a'.

## Conclusion

The network architecture of the TIER 1 data center at JINR of the first module is built on hardware supplied by Brocade company using the modern multichannel data transfer protocol TRILL. The so obtained experimental data direct our activities for the further research on the nature of the traffic distribution in redundant topologies. We need to answer several obvious questions. In what way is the distribution done while transmitting packed data by four peer paths provided that the conditions of the patent claims are not met? What will happen if the traffic via one of the communication channels achieves peak values?

We are developing a test bench to carry out a similar experiment using a traffic generator. It is possible that, in order to confirm the collected data, it will be necessary to build, as in all physical experiments, a similar network fabrics running via the TRILL protocol on another manufacturer's equipment.

In case of completing the harmonization of this matter, we will carry out a comparative functional analysis of the TRILL protocol in various data center modules by various manufacturers. This work has immense importance because the collected data will be used to construct the Data Processing Center within the NICA (Nuclotron-based Ion Collider fAcility) megaproject[5].

## REFERENCES

- [1] N.S.Astakhov, A.S.Baginyan, A.G.Dolbilov, N.I.Gromova, I.A.Kashunin, V.V.Korenkov, V.V.Mitsyn, S.V.Shamatov, .A.Strizh,V.V.Trofimov,N.N.Voitishin,V.E.Zhiltssov JINR TIER-1-level computing system for the CMS experiment at LHC: status and perspectives / Computer studies and simulation, vol. 7, No 3, 2015. pp. 455–462.
- [2] Baginyan A.S., Dolbilov A.G., Korenkov V.V. Network for data-center Tier 1 at JINR for experiment CMS (LHC) / T-Comm, Vol. 10, No.1, 2016. pp. 25-29.
- [3] Baginyan A.S., Dolbilov A.G. TCAM - fromIpv4 toIpv6 / T-Comm, No.4, 2013. pp. 24-28.
- [4] N.S.Astakhov, A.S.Baginyan, S.D.Belov, A.G.Dolbilov, A.O.Golunov, N.I.Gromova , I.S.Kadochnikov, I.A.Kashunin, V.V.Korenkov, V.V.Mitsyn, I.S.Pelevanyuk, S.V.Shamatov, T.A.Strizh, E.A.Tikhonenko, V.V.Trofimov, N.N.Voitishin, V.E.Zhiltssov JINR Tier-1 center for the CMS Experiment at LHC / Particles and Nuclei, Letters, vol.13, No 5, 2016. pp. 1103-1107.
- [5] V.V.Korenkov, A.V.Nechaevskiy, G.A.Ososkov, D.I.Pryahina, V.V.Trofimov, A.V.Uzhinskiy Simulation concept of NICA-MPD-SPD Tier0-Tier1 computing facilities / Particles and Nuclei, Letters, vol.13, No 5, 2016, Pp.1074-1083.G. O. Young, "Synthetic structure of industrial plastics (Book style with paper title and editor)," in Plastics, 2nd ed. vol. 3, J. Peters, Ed. New York: McGraw-Hill, 1964, pp. 15–64.