The TCO Benefits of Dell's Next-Generation Telco Servers

Peter Fetterolf, Ph.D.

EXECUTIVE SUMMARY

Communication Service Providers (CSPs) worldwide are in the process of deploying 5G networks, many of which use telco cloud architectures. In a telco cloud, virtual RAN functions and virtual packet core functions run on standard x86 servers. This paper focuses on the total cost of ownership (TCO) benefits of moving from the current generation of servers based on Intel's 3rd Gen Intel Xeon Scalable processors to the next generation of telco servers using 4th Gen Intel Xeon Scalable processors with Intel vRAN Boost.

ACG has modeled a hypothetical network that represents a Tier 1 network in a European country or a North American region. We compare previous generation servers with Dell's XR8000 and R760 servers that use the 4th Gen Intel Xeon Scalable processors with Intel vRAN Boost. Our results show a five-year cumulative total cost of ownership savings of 30%, a CapEx savings of 23%, and an OpEx savings of 37%. The TCO savings are a result of several factors:

- Significant performance improvement with the 4th Gen Intel Xeon Scalable processors with Intel vRAN Boost supporting both vDU and core workloads. The estimated improvement on the vDU workload is up to 2X in specific scenarios, and the estimated improvement on the core is approximately 42% in specific scenarios over that of 3rd Gen Intel Xeon Scalable processors, leading to reductions in the number of edge and core servers required
- Innovative designs to simplify operations in telecom environments contribute to the reduction in the number of truck rolls and service disruptions
- Increase in power efficiency with Dell's Smart Cooling design even in harsh environments
- XR8000 servers are designed to leverage existing telcos' cabinets and can use existing power infrastructure, which reduces network installation expenses.

New Generation Telco Servers and Benefits

To help CSPs continue to scale their telco cloud networks, Dell's newer servers for telcos deliver a significant boost in network performance and simplified operations to reduce the total cost of ownership (TCO). Dell's XR8000 and R760 servers have the following benefits that directly reduce network TCO:

- Improved RAN and packet core network performance and throughput
- Dell's Smart Cooling design reduces power consumption, thereby minimizing OpEx and CapEx, and contributes to ESG/green data center goals
- Smart, modular designed for telcos' environments, reducing labor expenses and truck roll expenses
- New edge servers are designed to support deployments in existing telcos' environments (cell site, central office, and data center), which reduce OpEx and CapEx associated with upgrading power and cabinet infrastructure
- Dell has extended support (up to 11 years) to increase longevity of infrastructure

The new telco servers, which have been designed to increase the performance for RAN and packet core applications, use the Intel 4th Gen Intel Xeon Scalable processors with Intel vRAN Boost. The new processors significantly increase the DU capacity (up to 2X in specific scenarios), and the packet core UPF and RAN CU performance increases by 42%. This reduces the TCO for RAN and packet core deployments, and it is a major consideration in our TCO model.

The XR8000 server is designed for next-generation Distributed RAN (D-RAN) and Centralized-RAN (C-RAN) deployments supporting vRAN and ORAN, and deployments supporting packet core distributed UPF functions. The key benefits of the XR8000 are summarized in Figure 1.

1. Flexible form factor with modular SLED design
    - Multiple SLED options enable a common platform for use cases across RAN, edge, and core
    - Independent SLED design supporting different combination of OS, CPUs, DPUs, and GPUs
    - System designed to support up to three generations of CPU

2. Smart cooling design to support harsh edge environments
    - Vapor chamber heat sink options to support higher efficiency cooling, thus drawing less electricity
    - Up to 65C support with certain configurations to support extreme temperature
3. Power and fan design simplify field operations while massively reducing OpEx
    - Dual PSUs capable of supporting up to four SLEDs, massively reducing power infrastructure and consumption
    - Modular and independent PSUs allow for replacement without service disruption
    - SLED integrated fan infrastructure eases operations and serviceability

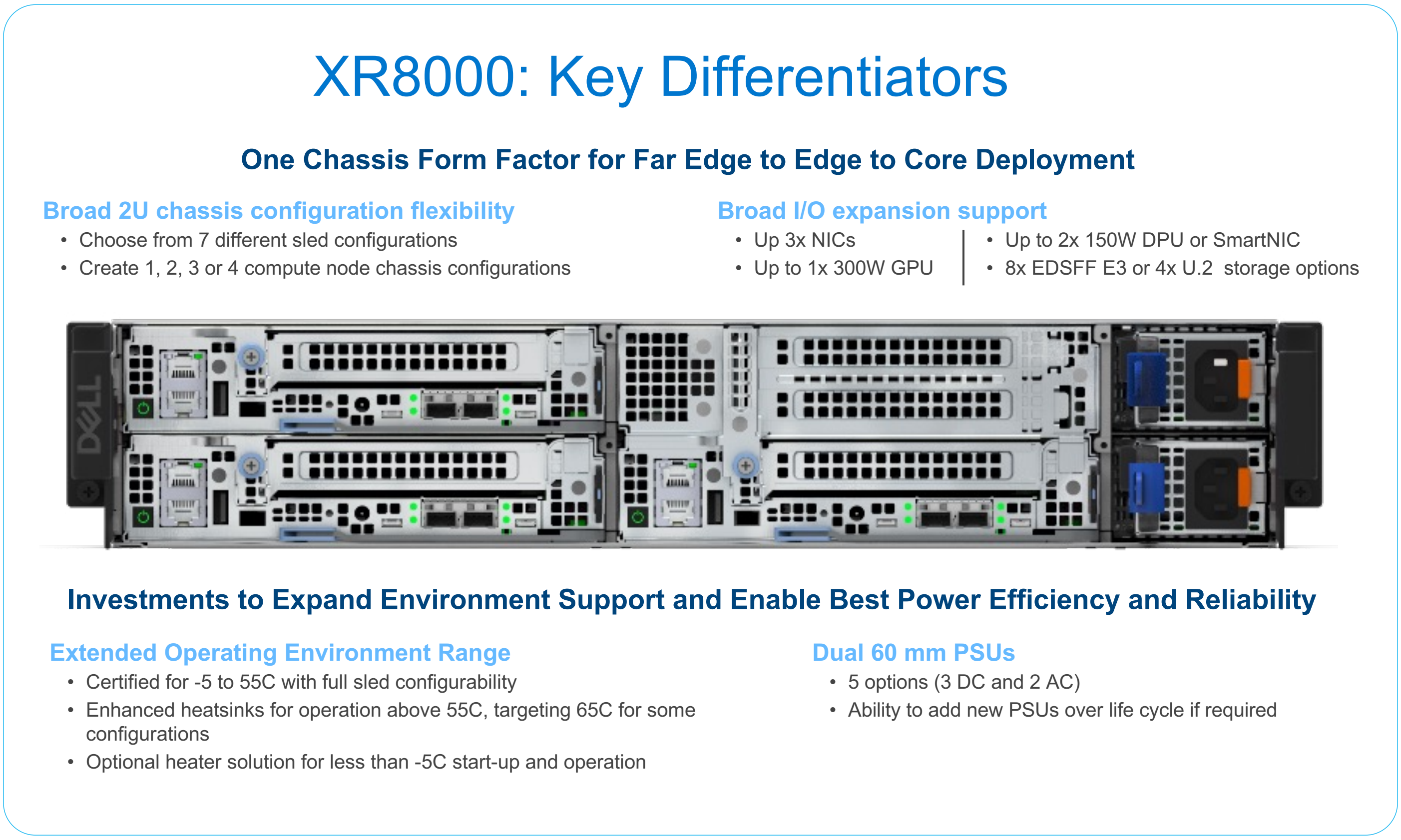

Figure 1. XR8000 Key Differentiators

A summary of XR8000 benefits is presented in Figure 2.

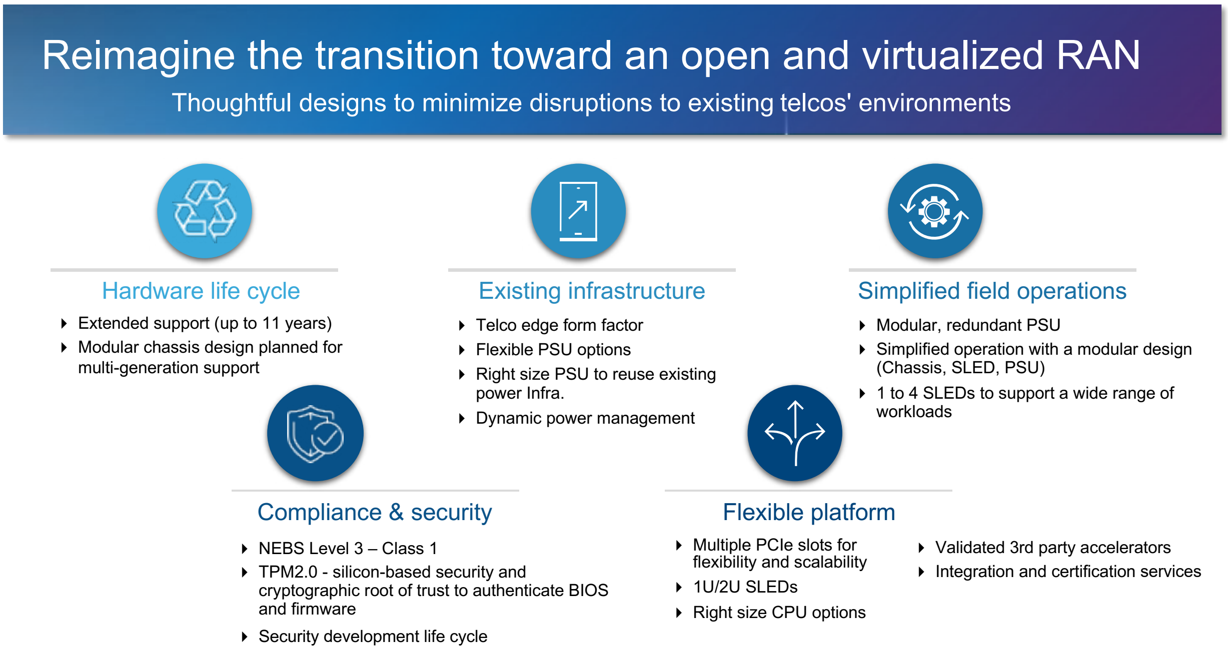

Figure 1. XR8000 Key Differentiators

The PowerEdge R760 series servers are designed to address the fast growing 5G core workloads. The following are the key benefits of the R760 series:

- Max performance: Extract value faster from collected data with high-performance processors, GPUs, next-generation storage for demanding workloads such as automation intelligence/machine learning
- Thoughtfully crafted: To fit your current infrastructure and deliver the right performance to support demanding applications
- Smart cooling options: Flexible options with patented multi-vector air cooling or liquid cooling
- Energy efficient: Adaptive closed loop control optimizes fan and system power consumption

Figure 3. Dell PowerEdge R760

TCO Model Assumptions

The analysis compares the TCO of a mobile network over five years for two scenarios:

1. Present Mode of Operation (PMO): Previous generation servers using 3rd Gen Intel Xeon Scalable processors for the RAN, distributed UPF and the packet core control plane in a regional data center
2. Future Mode of Operation (FMO): XR8000 and R760 servers using 4th Gen Intel Xeon Scalable processors with Intel vRAN Boost for the RAN and distributed UPF and R760 servers for the packet core control plane in a regional data center

We model a network that is typical of a Tier 1 service provider in a European country or a North America region. The key characteristics of this network are:

- 15 million simultaneous active users
- 30,000 cell sites
- 1 regional data center hosting the packet core control plane
- 10 edge data centers hosting the CU and UPF network functions
- 500 far edge data centers providing shared DU pools for C-RAN deployments
- Four types of regions: dense urban, urban, suburban, and rural
- D-RAN, C-RAN, and packet core deployments

We modeled a combination of D-RAN and C-RAN. In D-RAN networks the DU and CU run on one or more servers collocated at the cell site base station. In C-RAN deployments the DU and CU are centralized in DU and CU pooling data centers. The distribution of C-RAN and D-RAN is presented in Table 1. This distribution is representative of many CSPs' deployment scenarios. We assume:

- In urban and dense urban areas, the distance from the cell site to the edge data centers is shorter; therefore, it is possible to build fronthaul networks from the cell sites to the edge data centers without violating latency or jitter requirements.
- In urban and dense urban areas 20% of the cell sites use C-RAN and 80% use D-RAN.
- In suburban areas many cell sites are spread out and not suitable for C-RAN; we assume only 10% of the suburban cell sites use C-RAN.

- In rural sites the distances are generally too great for C-RAN; we assume 100% D-RAN in rural sites.

Table 2 depicts the radio carriers and number of sectors for each carrier in each type of cell site. mmWave is used only in the dense urban area, and M-MIMO 100MHz 64X64 is used in both dense urban and urban areas. Suburban and rural areas use 20MHz 4X4 radio carriers. The number of sectors in each type of area is designed to reflect demand in those areas.

	D-RAN	C-RAN	Total	% C-RAN
Dense Urban	5000	1000	6000	20%
Urban	12000	2400	14400	20%
Suburban	6000	600	6600	10%
Rural	3000	0	3000	0%
Total	26000	4000	30000	15%

Table 1. Distribution of C-RAN and D-RAN in Different Regions

	Sectors			
Radio Carrier	**Dense Urban**	**Urban**	**Suburban**	**Rural**
100MHz 64X64	6	3	0	0
20MHz 4X4	9	9	6	3
800MHz mmWave	3	0	0	0

Table 2. Radio Carriers in Each Type of Cell Site

In D-RAN locations we calculate the servers required based on the radio carrier configuration and the capacity of the Intel processor to host DU/CU functions. The number of servers in each type of location for both the 3rd Gen Intel Xeon Scalable processor-based server and the Dell XR8000 with 4th Gen Intel Xeon Scalable processors with Intel vRAN Boost is presented in Table 3. 4th Gen Intel Xeon Scalable processors with Intel vRAN Boost have been designed to optimize vRAN performance; they require fewer servers and have lower core densities.

	Number of Servers in D-RAN Locations			
Intel Processor	**Dense Urban**	**Urban**	**Suburban**	**Rural**
3rd Gen Intel Xeon Scalable Processor	3	2	1	1
4th Gen Intel Xeon Scalable Processor	2	1	1	1

Table 3. Servers Required in D-RAN Locations for 3rd Gen Intel Xeon Scalable Processors and 4th Gen Intel Xeon Scalable Processors with Intel vRAN Boost

For C-RAN deployments we compare the 3rd Gen Intel Xeon Scalable processors-based server in edge data centers with the new Dell XR8000. The number of servers is calculated based on the pooled demand from all the cell sites and radio carriers that are aggregated at the edge data center.

For both D-RAN and C-RAN the XR8000 results in a lower number of servers required and reduced core counts for servers. This directly decreases both TCO and power consumption. It should also be noted that the XR8000 power usage is more efficient than the 3rd Gen Intel Xeon Scalable processor-based server due to advancements in the Intel 4th Gen Intel Xeon Scalable processors with Intel vRAN Boost process and more efficient server design.

In addition to the server and processor optimization there are additional TCO benefits due to the XR8000 system architecture. Specifically:

- Advanced system power design enables dual PSUs to support up to 4 SLEDs, massively reducing power infrastructure and usage
- Modular architecture of the XR8000 reduces field operations expenses
  - Field operations require union workers, each with specific skill sets for power or compute
  - Truck rolls for maintenance generally require two technicians: one for power and one for compute
  - Modular design allows replacement of SLED or PSU separately; only one technician is needed for a truck roll depending on whether a compute SLED failed or a PSU failed

TCO Modeling Methodology

We use the ACG Business Analytics Engine (BAE), which is a visual, network and cloud economic simulation platform[1], to model the network scenarios. We simulated two scenarios over five years. With the BAE we simulate demand and automatically calculate the resources and cash flows necessary to meet demand. In this model the demand is created by 5G UEs and mobile cell sites. The 5G UEs grow from zero to 15 million over 5 years using an S-Curve or logistics function growth curve[2]. Each UE generates average busy period traffic of 1 Mbps with an annual growth rate of 23%. The combination of 5G UE growth and traffic growth drives the 5G control plane and user plane components. The growth of cell sites and radio carriers primarily drive the RAN DU and the CU. Each radio carrier requires a specific number of cores to support DU functions. CU components are primarily driven by network traffic. We assume cell sites will grow from zero to 30,000 cell sites over five years.

A visual view of the TCO model in BAE model designer is presented in Figure 5. Demand is driven by end points, which are 5G UEs and cell sites. The end points drive services, which are packet core services and radio carrier services. We assume each 5G UE has an average data rate during the busy period of 1 Mbps, and the average data rate grows at 23% annually. The services then drive resources, which include packet core CP and UPF and RAN CU and DU functions. The resources then drive servers in edge and regional data centers. Servers have many CapEx and OpEx cash that drive TCO.

Figure 4. ACG Business Analytics Engine

1 https://www.acgbae.com/

2 https://en.wikipedia.org/wiki/Logistic_function

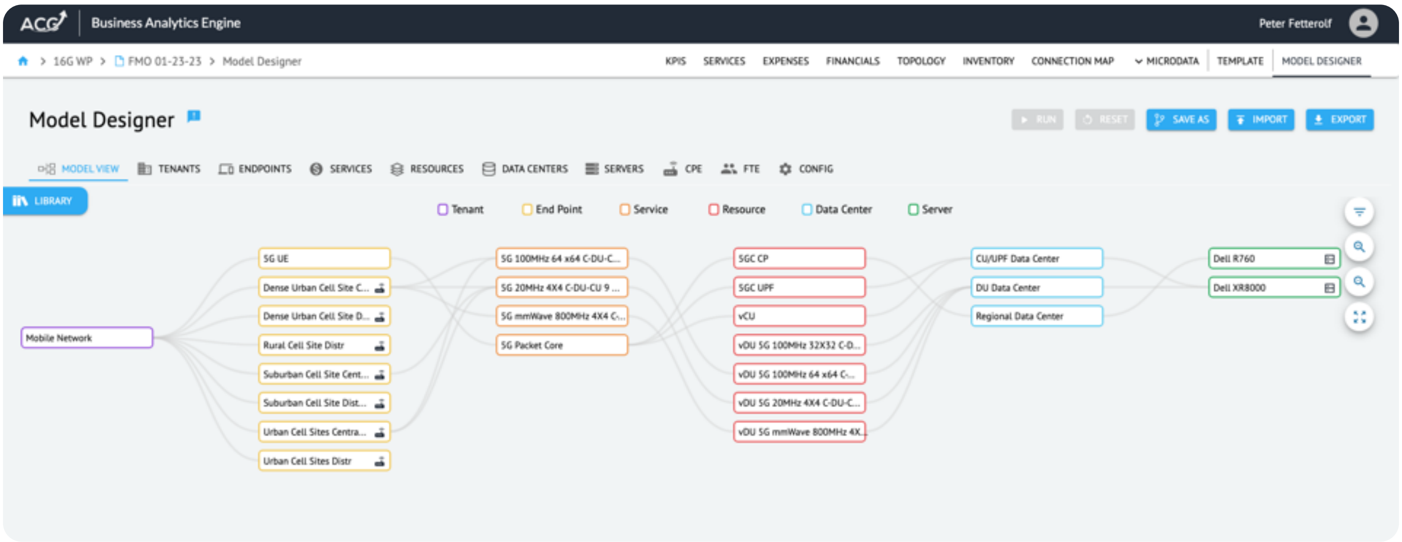

Figure 5. BAE Model Designer Visual View of TCO Model

TCO Model Results

The results of our five-year TCO study show a five-year cumulative TCO savings of 30%, a CapEx savings of 23%, and an OpEx savings of 37%. A breakdown of the five-year CapEx savings are presented in Figure 6, and a breakdown of the five-year OpEx savings are provided in Figure 7. The key drivers for OpEx savings are directly driven by the number of servers, reduced maintenance, server support, and power, cooling, and floorspace. The key CapEx expense reductions are server acquisition, infrastructure power, and installation expenses.

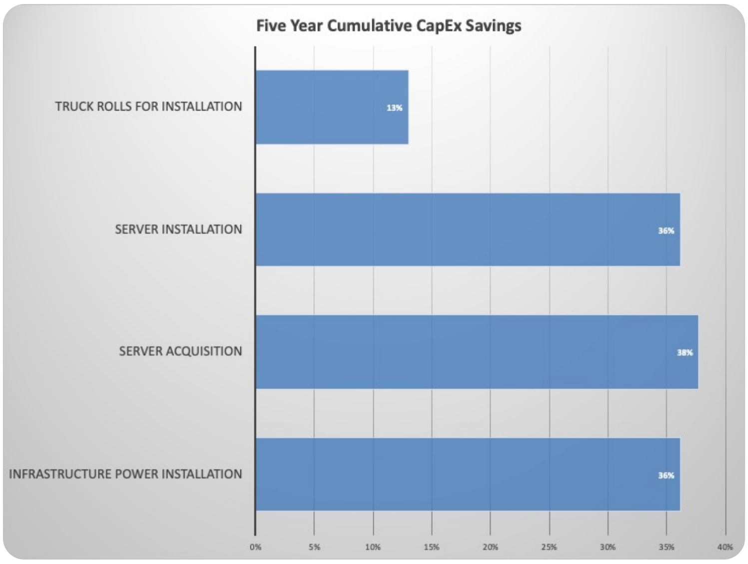

Figure 6. Breakdown of the Five-Year CapEx Savings

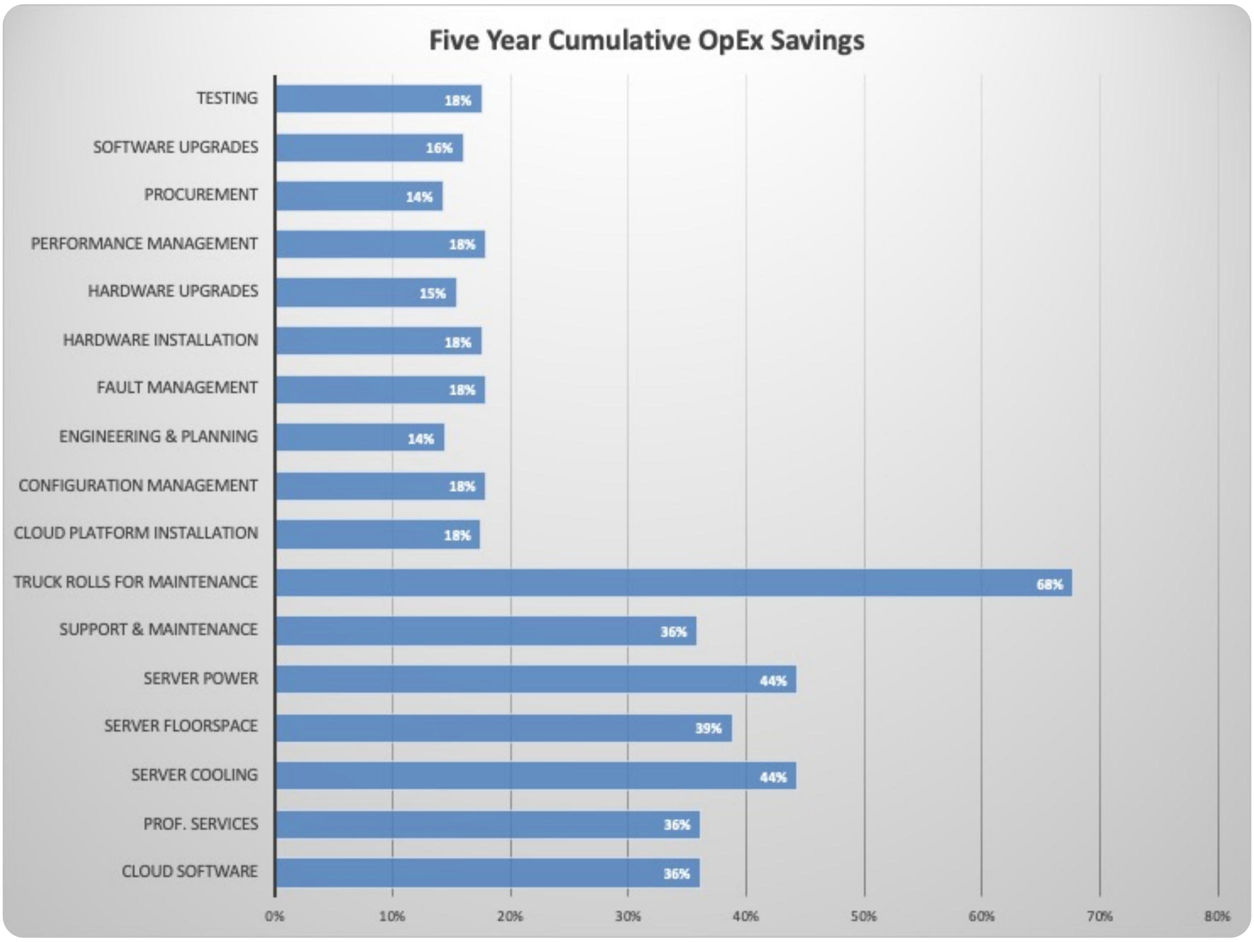

Figure 7. Breakdown of the Five-Year OpEx Savings

Another key area of savings are power and cooling expenses and reduction in CO^2 emissions. In the results we assume average North American power expenses, presented in Table 4. The CO^2 emissions savings translates to driving 55,220 gas powered cars for one year or 49,865 homes electricity use for one year[3].

	Savings
Total Power & Cooling Savings	$54.2 Million
Total kW Hour Savings	362.6 Million
Total CO^2 Emissions Savings	256,276 Metric Tons

Table 4. Five-Year Cumulative Power and Cooling Savings and CO^2 Emissions Savings

1 https://www.epa.gov/energy/greenhouse-gas-equivalencies-calculator

The results are combined results for both the RAN and the 5G core. We also modeled the RAN and 5G core separately; the TCO results are presented in Table 5. In the combined network the RAN has many more servers than the packet core, dominating the overall results. Therefore, the savings for the RAN are essentially the same as the savings for the overall network. The 5G core savings are slightly less because the performance improvements due to the 4th Gen Intel Xeon Scalable processors with Intel vRAN Boost for the 5G core are not as significant as the RAN performance improvements.

	RAN Savings	5G Core Savings
TCO	30%	22%
CapEx	23%	18%
OpEx	37%	25%

Table 5. TCO Savings for RAN and 5G Core

Conclusion

CSPs worldwide are struggling to reduce network CapEx and OpEx while growing service revenues. Mobile networks based on telco cloud require large numbers of servers to support vRAN deployments and packet core deployments. The number of servers drives many elements of CapEx and OpEx, including acquisition, power and cooling, cloud software, and labor expenses. This paper demonstrates that moving to the next generation of Dell servers, including XR8000 and PowerEdge R760 designed for telecom environments, will dramatically reduce the number of servers required in the network, leading to significant decreases in network TCO. Our results show a five-year cumulative TCO savings of 30%, a CapEx savings of 23%, and an OpEx savings of 37%.

Peter Fetterolf

Peter Fetterolf, Ph. D. is an expert in network technology, architecture and economic analysis. He is responsible for financial modeling and whitepapers as well as software development of the ACG Research Business Analytics Engine. Dr. Fetterolf has a multidisciplinary background in the networking industry with over thirty years of experience as a management consultant, entrepreneur, executive manager, and academic. He is experienced in economic modeling, business case analysis, engineering management, product definition, market validation, network design, and enterprise, and service provider network strategy.