Innovation
Asset
Collective

IAC Landscape Report
Energy Efficiency in Data Centre

Adoption of environmental, social and corporate governance (ESG) practices

Evaluation of commercial trends and patenting activity

www.ipcollective.ca

About Innovation Asset Collective

Innovation Asset Collective (IAC) is an independent membership based not-for-profit selected by the Canadian Government's Department of Innovation, Science and Economic Development (ISED) to assist Canadian small and medium-sized enterprises (SMEs) in the data-driven clean technology sector with their IP needs.

Led by experts in IP education, strategic counsel, IP generation and patent access, IAC helps Canadian SMEs understand and harness the value and power of their IP so that their innovations can be commercialized and protected for the benefit of the Canadian economy.

With the help of the IAC team, member companies will maximize the value of their intangible assets, while benefiting from the services of the collective and setting the stage for international growth. By realizing the inherent value of IP, IAC will foster Canadian innovation, which will see more Canadian companies succeed globally.

To learn more about IAC, visit www.ipcollective.ca

About the landscape report

Data centres are central to the critical digital infrastructure that supports advanced knowledge-based economies. Exponential growth in computing workloads driven by a massive rise in global data consumption has led to a significant increase in the number of data centres worldwide. However, the data centres have also become heavy consumers of electricity and, thus, indirect emitters of greenhouse gases (GHG). A recent study reports that a large data centre with a capacity of 100MW can utilize power, equivalent to 80,000 US households[1]. Stricter environmental policies and ESG commitments have driven innovation in data centres and are expediting the adoption of energy-efficient technologies and techniques. IAC's team examines some of these technologies and techniques to introduce energy efficiencies in data centres' operation and facility management.

In this digital age, organizations face a growing challenge to service legacy workloads to maintain business continuity alongside next-generation applications reliant on high performance compute and network resources. To accommodate new requirements, organizations are moving away from the traditional on-premise architecture to adaptive ownership and management models allowing flexibility in scaling infrastructure (performance and capacity) through third-party managed physical facilities and multi-cloud environments. Increased capital investments by incumbents, new innovations in data storage, computing and networking, and consolidation of assets through mergers and acquisitions (M&A) point towards high commercial activity and growth in the near future.

This landscape study illustrates the data centre value chain by examining the trends and technological shifts that reflect on the adoption of energy-efficient means in the three essential data centre components - namely, storage, compute and networking. The landscape study also identifies and examines technological and patenting trends in power management techniques to monitor and manage data centre resources (such as servers, virtual machines, storage systems, processors, networking components) and minimize energy losses. The value chain identifies different groups of organizations involved in delivering products and services needed to operationalize data centres. These groups of organizations range from incumbents with strong controlling positions mediating the markets and technological standards to newer organizations building atop a large knowledge base offering unique value propositions.

Patent search, which forms the primary source of data for this landscape, is bounded by the explicit use of “energy efficiency”

1 Energy Innovation, Mar 2020, How Much Energy Do Data Centers Really Use?, https://energyinnovation.org/2020/03/17/how-much-energy-do-data-centers-really-use/

in “data centres” in the patents (variations of words included while collecting data). This data is supplemented with industry-related information, including public policy changes, market activities such as mergers and acquisitions (M&A), licensing deals, alliances, and key projects undertaken. At a granular level, this report provides a lay of the competitive space by examining organizations’ patenting strategies worldwide and, further, identifies any risks and opportunities through market insights and white-spaces in technology segments related to the energy efficiency in data centres at large.

How can landscape reports help businesses?

Knowledge-based assets are increasingly becoming a source of competitive advantage and profits for businesses. Managing such intellectual assets requires complex processes integrated with the other business growth and internal processes. Foremost of these processes is to identify such knowledge-based assets that can differentiate a business in the market from others, support expansion of market share and generate profits. In other words, to extract value from knowledge-based assets, businesses should have a deeper awareness of the competitive landscape and be able to distinctly identify their own core knowledge-based assets that should be strategically appropriated to confer competitive advantage and maximize returns.

Awareness of the competitive landscape can impart the much-needed intelligence to identify strengths and weaknesses of a firm, opportunities that lie ahead and threats present in the market. Patent landscape studies offer a systematic way to sketch the competitive landscape by observing patterns in patenting strategies of rival firms. These trends can be examined at a higher or collective level as well as allow disaggregating data at a granular level to analyze individual patterns. The insights drawn from the landscape studies can enable a firm to develop strategies to compete in the market.

IAC’s landscape studies sketch sector-level value chains, along with supplementing the patent research with market research. Such a study could be a valuable source for understanding the market structure, the interactions between different sectoral divisions based on application areas or technologies, organizations' pervasiveness in these segments and the power to mediate the operation across the stack.

It should be, however, noted that the competitive landscape undergoes continuous change due to industrial innovations, market consolidations, emerging innovative markets and economy-related developments. Therefore, a single landscape study may become obsolete soon. A process is needed to continuously monitor and analyze the patterns of strategies and activities undertaken by rival firms. Only after embedding such processes with other business growth processes, a firm can have strategic information to stay ahead in the market.

Increasing Dependence on Digital Infrastructure

The data centre sector has been relatively untouched by the economic slowdown due to COVID-19. Instead, an even higher dependence is noticeable on the digital infrastructure due to higher workload demands and increased needs to digitize services and products and make these available online. This increase in demand is reflected in the rapid rise in infrastructural investments and mergers and acquisitions by hyperscalers, colocation providers and other data centre operators.

A recent Synergy Research[2] reports a rise of 18% from 2019 to 2020 in data centres, with the majority of investments coming from hyperscalers driven by Amazon, Google and Microsoft, followed by Facebook (Meta), Apple, Alibaba, Tencent and IBM. Additionally, Cisco Global Cloud Index[3] forecasts that "hyperscale data centers will represent more than half (53%) of all installed data center servers by 2021, and hyperscale data centers will account for 55% of total data center traffic in 2021, up from 39% last year."

Newer ownership and management models also suggest that though hyperscalers may own and manage their data centres, a large number of hyperscale providers (including SaaS providers, Salesforce, SAP, Workday, PayPal) and platform operators (Uber, Lyft) are leasing wholesale facilities from colocation providers, who also in some cases act as data centre operators and are responsible for long-term growth. Equinix and Digital Reality are two such colocation giants, actively growing their base across the world.

A recent article[4] reports 279 data centres hosted in Canada, placing it fifth in data centre density internationally. With the

2 Synergy Research Group, Sep 2021, Hyperscale Data Center Count Grows to 659 – ByteDance Joins the Leading Group, https://www.srgresearch.com/articles/hyperscale-data-center-count-grows-to-659-bytedance-joins-the-leading-group

3 Cisco Global Cloud Index, Feb 2018, Global Cloud Index Projects Cloud Traffic to Represent 95 Percent of Total Data Center Traffic by 2021, https://newsroom.cisco.com/press-release-content?articleId=1908858

4 The Globe and Mail, May 25 2021, Demand for storage space in the cloud creates a land rush, https://www.theglobeandmail.com/business/industry-news/property-report/article-demand-for-storage-space-in-the-cloud-creates-a-land-rush/

growth forecasted to double each year, much growth can be expected in this sector in Canada. Canadian cities with high data centre density include Toronto (56 sites, 174.03 MW), Montreal (33 sites, 384.85.00 MW), Vancouver (16 sites, 13.6 MW) and Calgary (13 sites, 22.48 MW)[5].

Data centres have a growing share in power consumption, which has become one of the central issues when it comes to energy efficiency. Power draw ranges from a few kW for a rack of servers in a closet to several tens of MW for large facilities. Some facilities have power densities of more than 100 times that of a typical office building. The latest report by AFCOM and Data Center World indicates, "For higher power density facilities, electricity costs are a dominant operating expense and account for over 10% of the total cost of ownership (TCO) of a data center." Reported figures indicate that global data centres consumed 416 TWh in 2016 (nearly 40% more than the consumption of the entire United Kingdom), while the US data centres' total annual consumption was more than 90 billion kWh[6].

International Data Corporation (IDC) notes the age of data centres to be close to 10 years[7], indicating renewal and modernization of infrastructure to follow. ESG practices further pressure data centre owners to adopt measures to lower their environmental footprint and replace ageing equipment with newer technologies and techniques.

5 Canada Data Center Market, https://baxtel.com/data-center/canada

6 Forbes, Dec 15 2017, Why Energy Is A Big And Rapidly Growing Problem For Data Centers, https://www.forbes.com/sites/forbestechcouncil/2017/12/15/why-energy-is-a-big-and-rapidly-growing-problem-for-data-centers/?sh=6387b6ab5a30

7 AFCOM and Data Center World, May 13 2021, Fifth Annual State of the Data Center Industry Report Reveals Shifting Trends, https://www.datacenterworld.com/sites/default/files/AFCOM_State%20of%20the%20Data%20Center_FINAL_2021_5-10-21.pdf

Lowering Environmental Footprint Due to Expanding Digital Footprint

Domo, a data visualization company, estimates that by 2020, "1.7 MB of data will be created every second for every person on earth,"[8] and IDC (a study sponsored by Seagate) projects a data growth of 30% a year, reaching 163 ZB by 2025[9]. The skyrocketing digital data creation and consumption creates the urgency to reconsider now on how we are expanding our needs and reliance on data centres.

Acting as a central host to an organization's critical and, in some cases, proprietary applications and data, the data centre becomes vital for the operation and continuance of the service for any modern-day organization. A data centre centralizes many functions for an organization around its IT operations and facilitates storing, processing, and disseminating data and applications.

Cisco defines a data centre as:
"At its simplest, a data center is a physical facility that organizations use to house their critical applications and data. A data center's design is based on a network of computing and storage resources that enable the delivery of shared applications and data. The key components of a data center design include routers, switches, firewalls, storage systems, servers, and application-delivery controllers."

The modernization of data centres is focused on limiting their impact on the environment. This can be done in many ways – namely, optimizing resource utilization, maximizing the utility of storage systems, upgrading to deploy energy-efficient equipment(s), opting for cleaner energy resources, and strategically locating facilities at destinations where environmental conditions can be utilized for maintaining ambient server room temperatures.

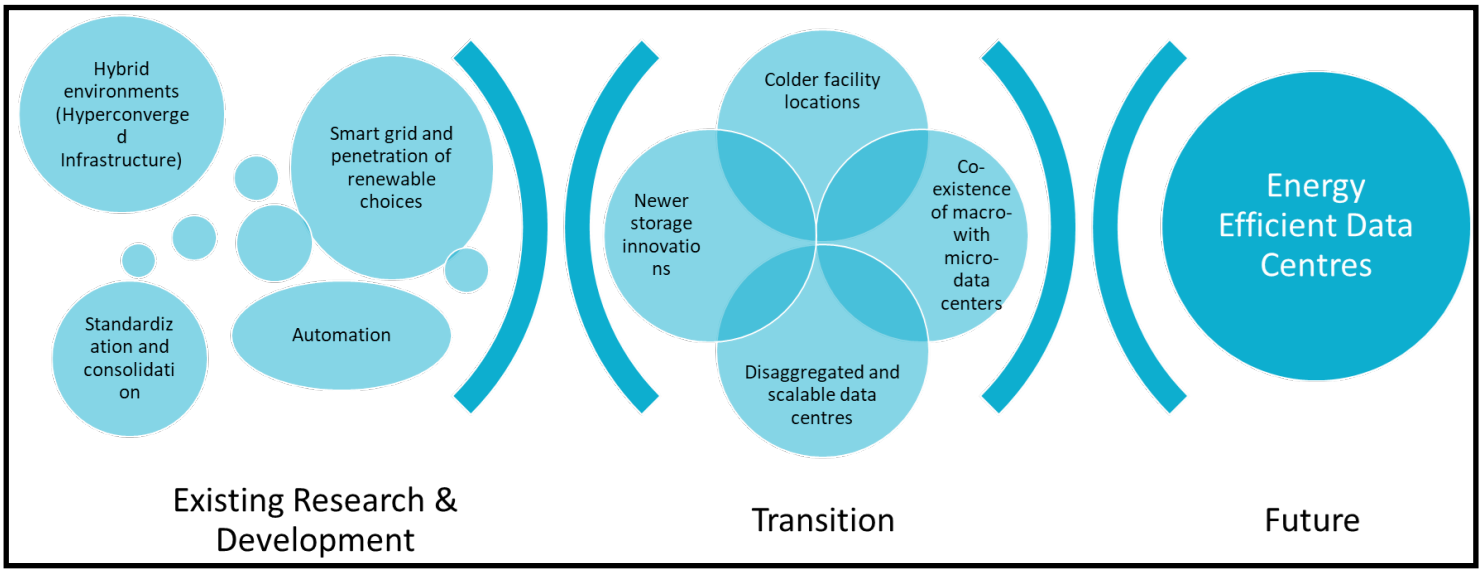


Figure I: Innovation in data centre towards attaining energy efficiency

8 DOMO, 2018, Data Never Sleeps 6.0, https://www.domo.com/assets/downloads/18_domo_data-never-sleeps-6+verticals.pdf

9 IDC Seagate sponsored study, 2018, https://blog.seagate.com/business/enormous-growth-in-data-is-coming-how-to-prepare-for-it-and-prosper-from-it/

Innovations Related to Energy Efficiency in Data Centres

1) Infrastructure and resource utilization:

Standardization and consolidation

Enterprises rely on data centres for continuity in operation, involving reliance on all fronts - data processing, storage and back-up, and adaptability to scale quickly to meet demands. From a data centre standardization and consolidation perspective, a lot of planning needs to be undertaken to identify functions, processes, services, products and people impacted by it. Consolidation efforts are aided by standardization, often involving replacing ageing data centre equipment, and focused on reducing the number of data centers and avoiding server sprawl (both physical and virtual).

Shift to composable infrastructure

Different ownership and management models for data centres have appeared over time, including large facilities owned and operated by hyperscalers, in some cases, hyperscalers leasing facilities from colocation owners, multi-tenant spaces leased and managed by colocation owners, as well as enterprise-owned facilities that range from micro- to mid- and large- sized data centres. Cloud infrastructure has introduced new possibilities for organizations to make a shift from managing on-premises physical servers to utilizing virtualized environments such as multi-cloud workspaces.

On the architecture front, innovation in hardware-based (converged) and software-based (hyperconverged) infrastructure solutions have paved the way for improving resource utilization, lowering energy consumption, and minimizing Capital Expenditure (CapEx) and Operational Expenditure (OpEx). Hyperconverged infrastructure (HCI) can be seen as preconfigured, software-based infrastructure that abstracts compute, storage and network components, and consolidates them into logical resource pools. Recent developments in architecture indicate that it is possible to compose data centres of merely software-defined infrastructure that is made up of disaggregated, commodity components such that compute, storage and network resources are abstracted and presented as unified services that can be allocated on-demand. Such an architecture would allow unlimited scalability as resources can be provisioned when required. Progress in this area can be seen through the Facebook (Meta) Open Compute project started in 2009.

2) Energy-efficient means:

Maximizing data density on existing facilities

Reported investments in data centre building and renovation have seen a slower increase compared to spending on enhancing the data density of the existing ones. The increased investments have led to a 4x increase in rack density and increased energy efficiency, from 2.4 kW/rack in 2011 to 7-10 kW/rack[6].

While Hard Disk Drives (HDDs) are the dominant storage systems at present, some other solutions include Solid State Devices (SSDs), NVM-NAND, NVM-Others, Optical and Tapes. The use of tape is expected to grow from 14% (2018) to 18% (2024). Where smaller data centres utilize flash drives more, cloud environments will see a continued increase in the use of hard drives and tapes.

Drives supporting interface protocols such as Serial Advanced Technology Attachment (SATA), Serial Attached SCSI (SAS), Non-Volatile Memory Express (NVMe) enable low latency and internal parallelism for solid-state storage devices. A growing field also seems to be towards extending the NVMe protocol to HDDs. The drive interface is typically used locally in a JBOD (Just a Bunch Of Disks), or a JBOF (Just a Bunch Of Flash) box and interconnected to the storage system/ compute system using NVMe-oF (NVMe over Fabric), where the fabric is typically Fiber channel, Infiniband, or Ethernet. The new variant NVMe-oF draws advantage by allowing workloads to benefit from disaggregated storage.

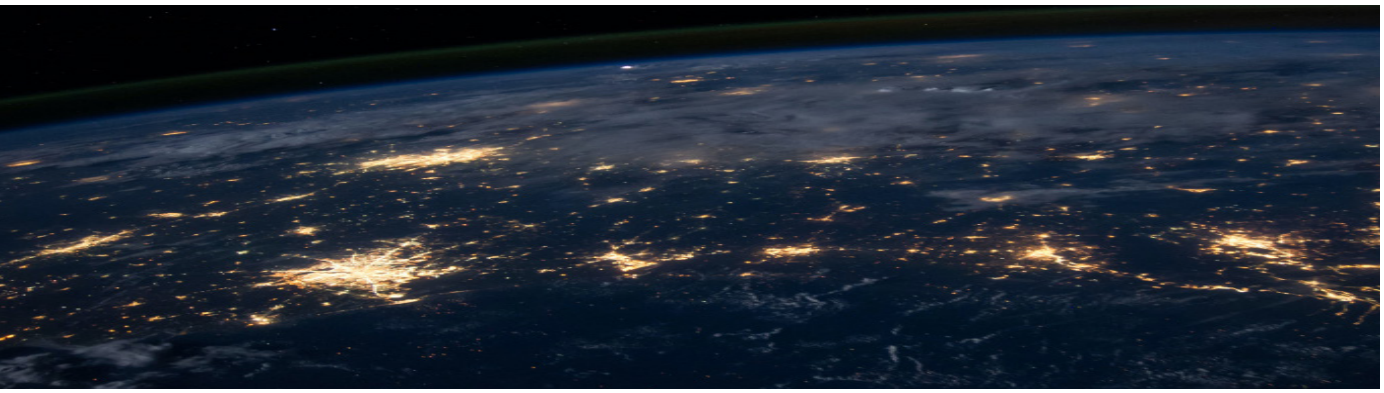
Some new storage solutions are also on the horizon; these include shingled magnetic recording (SMR), heat-assisted magnetic recording (HAMR) and liquid-state storage, which promise greater capacity and reduced costs.

A software-driven approach to storage systems may also be a widely adopted future trend. Existing trends in larger data centres provide evidence of form factors and interface protocols converging to create standards that are device agnostic. Such standards would eventually facilitate greater application of software across the data centre stack - bringing significant benefits of systems optimization.

Innovation in high-performance interconnects

It should also be noted that the greatest impact on power reduction exists in the aggregation layer of a data centre, which includes intra-data centre storage and compute interconnects and network interconnect. The interconnect technology determines the performance capabilities and scalability of a storage system and computing environment.

As understood, NVMe-oF is promising storage interconnect technology for high-performance computing environments.



Switchless interconnect, on the other hand, enables its own routing by replacing centralized switch fabrics with distributed and high-performance interconnects. It can be applied to 6D Torus, Dragonfly or Slimfly topologies, simplifying the large configurations that high-performance computing environments demand.

Innovation in power supplies

Lowering power usage by making use of energy-efficient chipsets and reducing CPUs power consumption, and introducing efficiencies in other server components such as power supply equipment (i.e., Uninterruptible Power Supply, UPS and Power Distribution Units, PDUs) are some of the other ways data centres become energy efficient. Adaptive power supply management and distributing power at higher voltages, while making use of improved cooling systems has also helped data centres manage and lower their power consumption.

3) Renewable energy resources:

Smart grids and power consumption

Smart grid has made it possible for deeper penetration of renewable resources over an electrical grid, enabling switching to cleaner sources of power easier and convenient.

4) Location:

Cooling techniques

A large percentage of data centre costs can be attributed to the cooling systems, such as maintaining the ambient temperature in server rooms, racks, cooling systems, etc. Canada appears as a leading country to build low-energy data centres in an index released by the New Statesman Media Group[10]. The index is calculated based on the total energy supply per 100,000 people and average annual temperature for maintaining reliably cool temperatures. Cold countries offer benefits such as leveraging cool air and water as primary cooling methods.

10 Energy Monitor, Dec 2020, Canada: The best country for energy-efficient data centres, https://www.energymonitor.ai/tech/energy-efficiency/canada-the-best-country-for-energy-efficient-data-centres

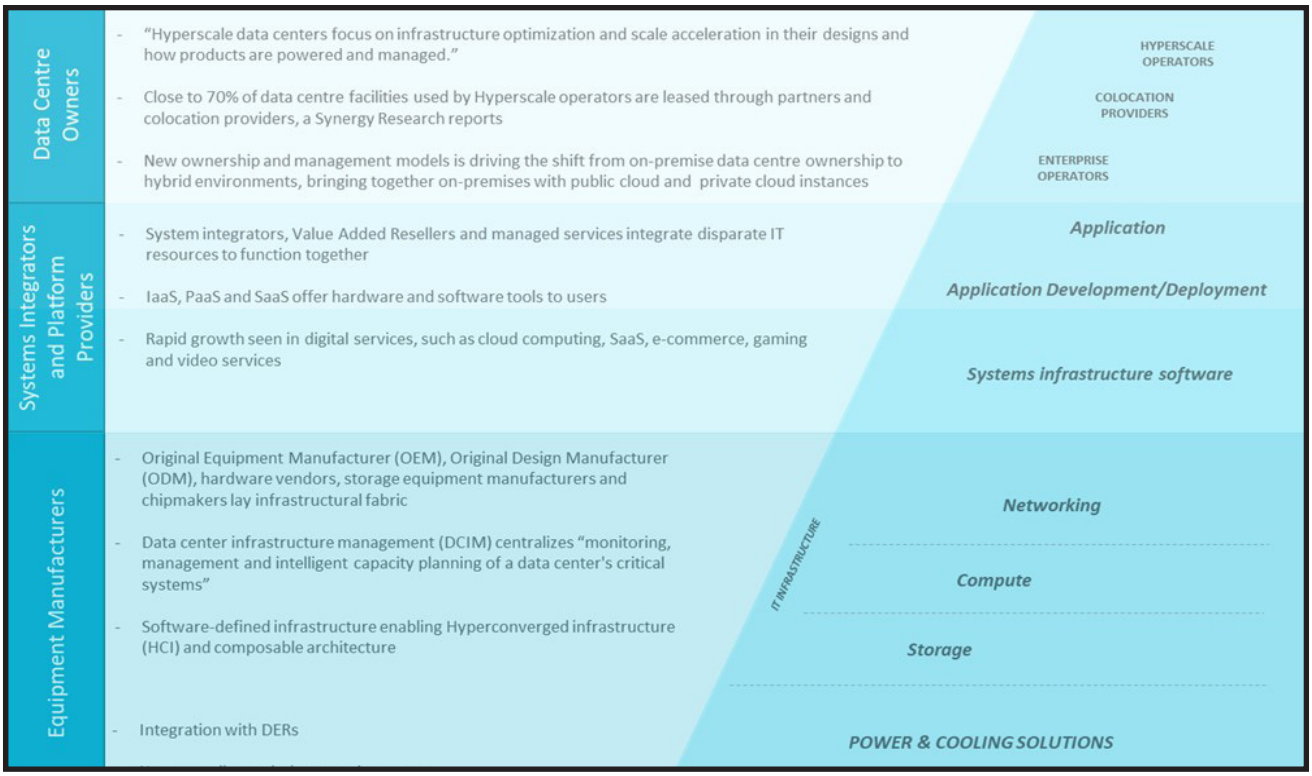


Figure II: Data Centre Value Chain

The detailed landscape report contains the following information:

- Data centre value chain with key organizations placed in the three layers (as shown in Fig. II).
- Risks and opportunities in the data centre sector.
- Multiple time-series illustrations to analyze patenting trends [including by technology segments, geographic location, top filers, technology cycle time (TCT)] and summaries of the innovation activities based on these illustrations.
- Time-series competitive landscape illustrations (by benchmarking patent portfolios using various quality indices) to elucidate changing competitive positioning of the companies in each technology segment over time.

Summary of findings

As part of critical infrastructure, data centres seem to be unimpacted by the economic slowdown and investments holdoff due to COVID-19. The past few years have seen an even higher demand for data centre services and rapid growth and investments in infrastructure and innovation.

The patenting activity in the technological areas relevant for introducing energy efficiency in data centres suggests a positive growth. Chip maker Intel seems to be aggressively filing in all segments of data centres, indicating its intent to protect its energy-efficient ways across the data centre technology stack. On the other hand, hyperscalers like Amazon and Google (Alphabet), which are also making significant

investments in new physical facilities, or leasing from colocation providers, are among the active filers in this sector. Active filers also include IBM, Dell Technology, Samsung Group, Huawei, Microsoft, Oracle, Cisco System, Hitachi, Furukawa Co Group Qualcomm and HP Enterprise.

Some players have a more focused approach and seem to be protecting more in one of the three segments. For example, names like Commvault, Western Digital Corp, Pure Storage, Rubrik, Seagate Technology, Micron Technology appear to aggregate patents in Storage, while Oracle, Broadcom, Marvell Technology Group, Juniper Networks, Nippon Telegraph & Telephone (NTT) and Corning appear to be filing in Networking. Some other names appearing in Compute are Advanced Micro Devices, Citrix System, Nvidia and Apple.

China state-owned entities and Chinese universities are also noticeable in this sector, with State Grid Corp of China and ZTE filing actively in Compute and Networking, with a few patents in Storage and other energy-efficient techniques, and Nanjing University, Chongqing University and Beijing University of Posts and Telecom appearing among the top filers in a few segments.

Other Chinese players include Suzhou Inspur Intelligent Technology and Zhengzhou Yunhai Information Technology.

The detailed landscape report also summarizes some of the recent consolidation activities and examines the patents identified in the dataset that have been used in litigation.

Authors

Arushi Sharma
Senior Patent Analyst, IAC

Parteek Saxena
Junior Patent Analyst, IAC

The detailed version of this, as well as future Patent Landscape Reports will be availableto all IAC members. If you are a member and have any questions about these reports,please feel free to get in touch with Melissa Bouffard, Relationship Manager at **mbourffard@ipcollective.ca**.

If you are a Canadian data-driven cleantech SME and are interested in joining IAC, for more information, please connect with Rasha Shamat, Business Development Manager at **rshamat@ipcollective.ca**.

We would also invite private or public organizations to connect with us at **partner@ipcollective.ca**.

If this is a topic of interest and to explore how we might be able to partner together to further the discussion of IP in the cleantech sector.