

Adaptive Network: Ett ramverk för att förstå nätverkskraven i Edge Cloud

Sammanfattning

Molntjänster tar över allt mer. Molntjänster finns oftast i bakgrunden när människor och verksamheter konsumerar innehåll och data. Det kan gälla allt från individuella användare som sträckttittar på Over-the-Top-video (OTT) till storföretag som installerar programvara som en tjänst (SaaS). Sedan många år tillbaka har dessa tjänster mest handlat om stora, centraliserade datacenter och molnarkitektur.

Nu kommer en hel generation med rena molnapplikationer inom exempelvis underhållning, detaljhandel, tillverkning och bilindustrin. Dessa kommer i många fall att kräva mer datorkraft och vara latenskänsligare. De traditionella centraliserade molnarkitekturerna kommer inte att klara kraven på Quality of Experience (QoE) för dessa applikationer. Istället krävs en mer dynamisk och distribuerad molnmodell. Det kommer att leda till att datorkraften och lagringen behöver flytta närmare nätverkskanten, där ju innehållet genereras och konsumeras. Det är detta som kallas Edge Cloud.

Övergången till en distribuerad Edge Cloud-modell kommer uppskattningsvis att kräva tre gånger fler datacenter vid nätverkskanten än idag. Alla som är en del av ekosystemet måste tänka om kring saker som rör nätverkets kopplingar¹.

Den här artikeln handlar om vad som driver utvecklingen mot Edge Computing och vad den kommer att innebära. Vi visar också hur Cienas Adaptive Network™ kan utgöra ett effektivt ramverk när man rör sig mot en distribuerad Edge Cloud-arkitektur.

Världen förändras och data flyttar närmare nätverkskanten.

Föreställ dig någon som ska åka på en affärsresa och ställer följande enkla fråga: "Hej Siri, visa hur jag tar mig till flygplatsen". På bara några sekunder tar Apple Kartor fram den kortaste vägen och erbjuder en vägbeskrivning med alla delsteg.

Under färden visar popupnotiser uppdaterade vägval för att undvika nyligen uppkomna köer, detta utan att ankomsten försenas med mer än några minuter. Under processens alla steg använde sannolikt Apples karttjänst ett centraliserat datacenter som kan ha legat tusentals kilometer bort, vilket gav en extra tidsfördröjning när förfrågningarna skulle bearbetas. För en icke-kritisk applikation, som Apples navigeringshjälp, är en viss fördröjning acceptabel. I de allra flesta fall kan användaren navigera till sin destination på ett korrekt sätt.

Låt oss titta på ett annat exempel. En kund går in i en livsmedelsbutik och checkar in med sin mobilapp. Rörelserna – när kunden plockar upp och tar med sig varor från hyllorna – fångas av kameror som byggts in i taket. AI analyserar rörelserna för att avgöra vad kunden har köpt och drar pengarna från kreditkortet direkt. På så vis slipper kunden alla de vanliga momenten i kassan på väg ut. För att åstadkomma detta behövs betydande databearbetningsresurser, antingen i butiken eller i nätverkskanten, för att utföra denna omedelbara bildbearbetning så att allt fungerar smidigt för kunden.

En av drivkrafterna bakom Edge Computing är just värdet i att kunna bearbeta stora datamängder som genererats lokalt och därmed minska backhaul-trafiken tillbaka till det centrala molnet. Målet är att minska latensen och mängden backhaul-trafik till det centrala molnet. Dessutom vill man effektivisera den storskaliga analys som krävs för att skicka inferenser och prediktioner till enheter i kanten, vilket förbättrar prestandan i applikationerna.

Många kommande applikationer kräver lägre latens än vad som är möjlig via centraliserade datacenter, så som visas i bild 1. Intäkterna från dessa applikationer väntas få en 42-procentig årlig tillväxttakt (Compound Annual Growth Rate, CAGR), från 1,2 miljarder USD 2020 till över 5 miljarder USD 2024. Video- och innehållsnätverk, molndataspel och fordonsapplikationer blir allra mest lönsamma¹.

¹ Mobile Experts: "Edge Computing for Enterprises 2019", juli 2019.

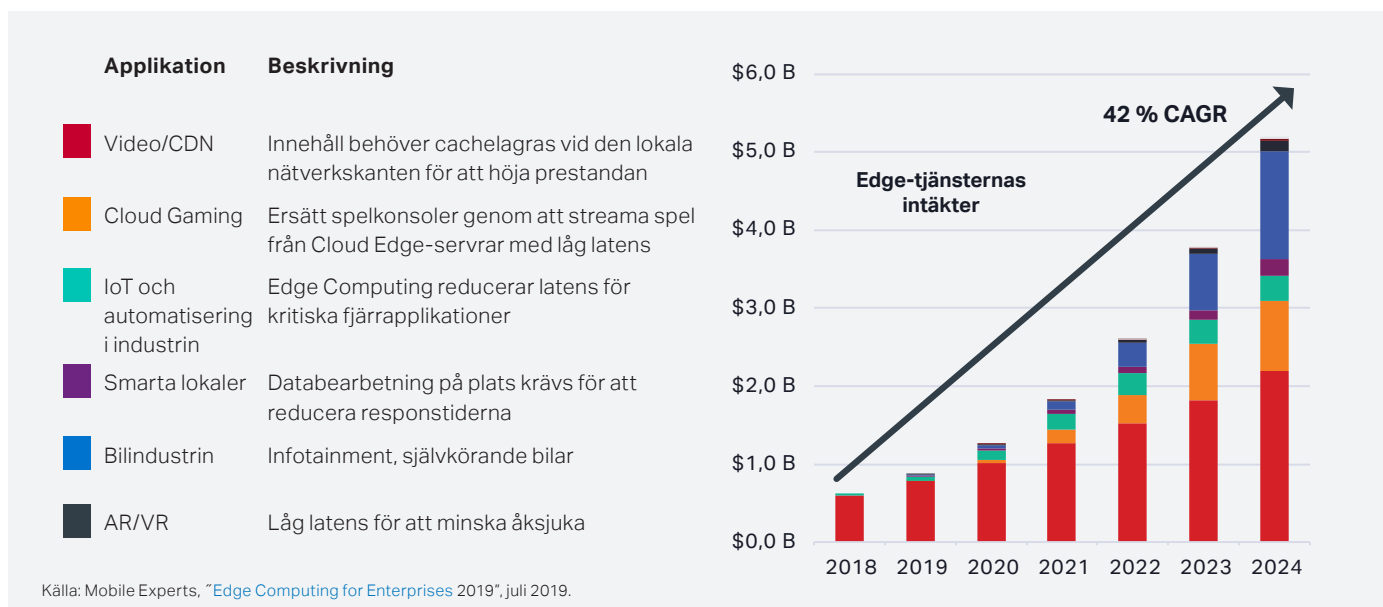


Bild 1. De största applikationsområdena inom Edge Computing sorterade efter intäkt

Förutom behovet av låg applikationslatens och kraven på QoE, tillkommer dessutom ytterligare krav från mobildelen av nätverket. Mobilappar inom molnspel och AR/VR kräver allt mer datakraft, vilket förstärker påverkan på mobilens batteriprestanda. Så i de här fallen kan datakraft och lagring med fördel flyttas till ett molndatacenter nära användaren med förbättrad batteritid som följd.

Dessutom ålägger flera regeringar molntjänstföretagen att lagra kundernas data inom respektive land, för att möta kraven på ökad integritet. Tidigare har många molnföretag drivit sina molntjänster från ett enda land i en världsdel eller stor region. Exempelvis har centraler i Storbritannien servat hela den europeiska marknaden. I och med att regeringar reser digitala gränser, blir molnföretagen ålagda att placera sina molnresurser i Edge-datacenter i de berörda länderna. Då minskar naturligtvis avståndet till användarna i de allra flesta fall.

Branschen bemöter dessa utmaningar genom att arbeta mot en distribuerad och dynamisk molnmodell. Här flyttas molnresurser från centraliserade datacenter till center i nätverkskanten, närmare användaren.

I den här artikeln tittar vi närmare på vad en distribuerad molnlösning, det vill säga Edge Cloud, innebär. Ciena definierar Edge Cloud som ett flexibelt molnekosystem som omfattar Edge Computing-komponenter för bearbetning och lagring från olika leverantörer, och ett skalbart nätverk. I detta skickas trafik fram och tillbaka mellan datacenter, som i realtid

kan känna av och anpassa sig till applikationernas behov på ett säkert sätt.

Var finns då nätverkskanten?

Många i branschen vill gärna slå fast exakt var nätverkskanten finns. Men sanningen är att den kommer att finnas på vitt skilda platser, helt beroende på QoE-förväntningar och varje enskild applikations resurskrav och -tillgänglighet. Var kommer Edge Cloud att finnas? Svaret beror på vilket perspektiv användarna, nätverksoperatörerna eller applikationsleverantörerna har.

I denna artikel beskrivs följande grupper av platser där en applikation kan finnas rent fysiskt (som visas i bild 2):

- 1. Metro Edge:** En blandning av stora datacenter med multitenans – Global Content Networks (GCN) och Data Center Operators (DCO) – och dessutom centrala nav för Communications Service Providers (CSP), som tillgodoser marknadens behov i regioner eller städer
- 2. Far Edge:** En mix av centrala CSP-nav och Cable Multi-Service Operator (MSO). För det sistnämnda avses huvudände eller platser för mobila distribuerade enheter (DU) för 5G, belägna nära användaren
- 3. User / On-premises Edge:** En mix av små och stora företags lokaler, inklusive deras datacenter och lokalkontor. I vissa fall kan nätverket även innefatta exempelvis transportnav, gruvor och fabriker

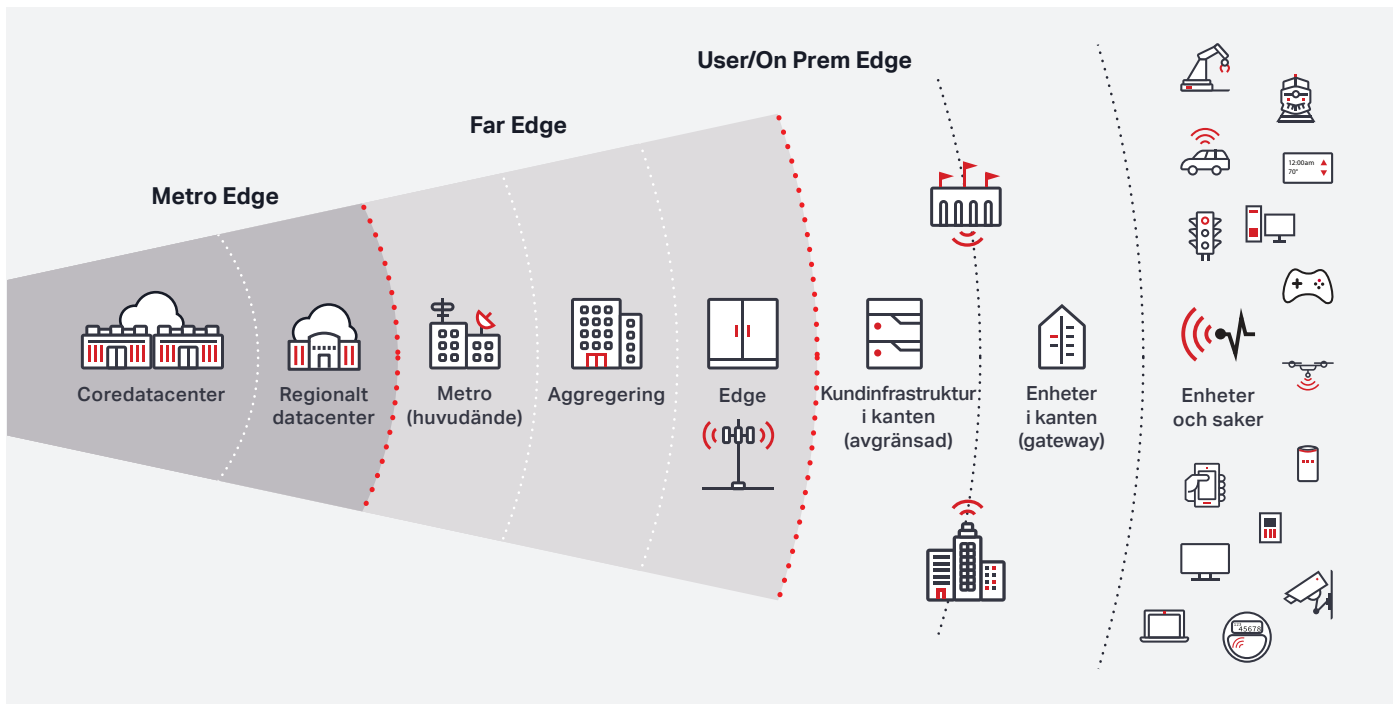


Bild 2. Här finns nätverkskanten

Edge Cloud suddar ut gränserna allt mer. Olika företag börjar samarbeta kring molnresurser som innefattar flera moln och tjänsteleverantörer. Dessutom tillkommer lösningar som utvecklats av CGN eller till och med slutanvändarna själva.

Idag finns det ungefär 10 000 datacenter i världen. Prognoser (se bild 3) över den allt snabbare utvecklingen mot Edge Cloud visar att det kommer att finnas upp till tre gånger fler nya datacenter i Metro/Far Edge och User/On-premises Edge inom fyra år¹.

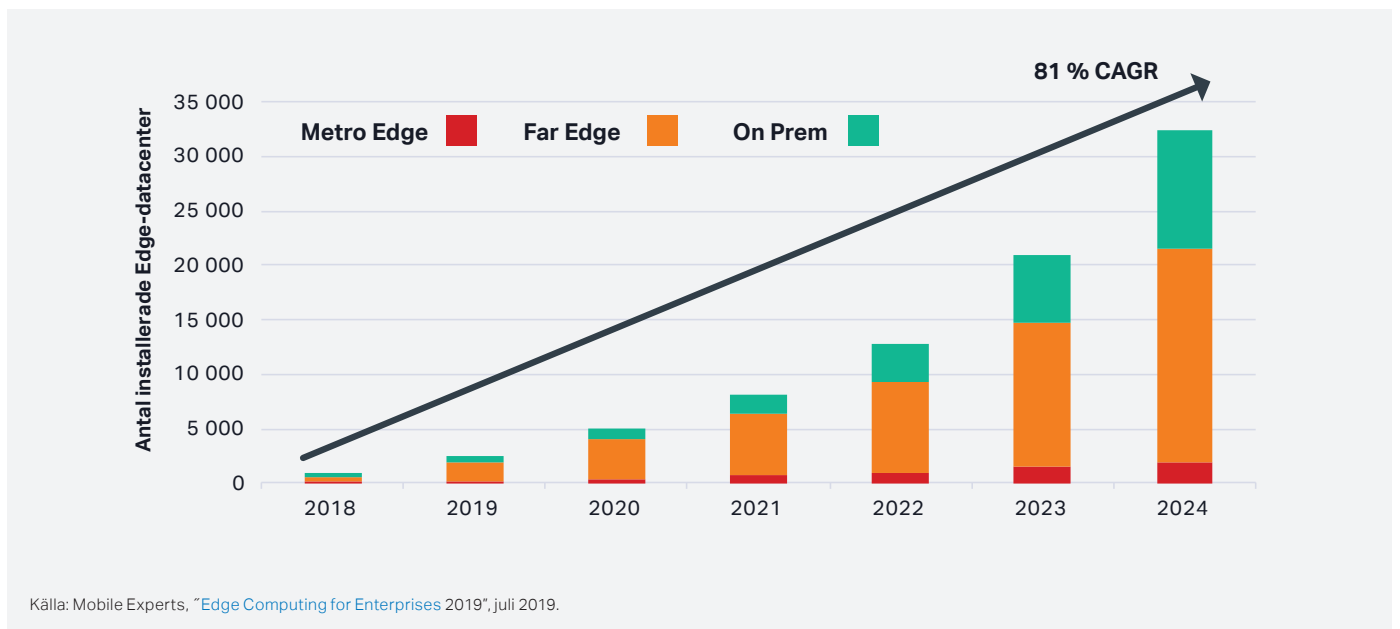


Bild 3. Antalet datacenter för Edge Computing blir allt fler

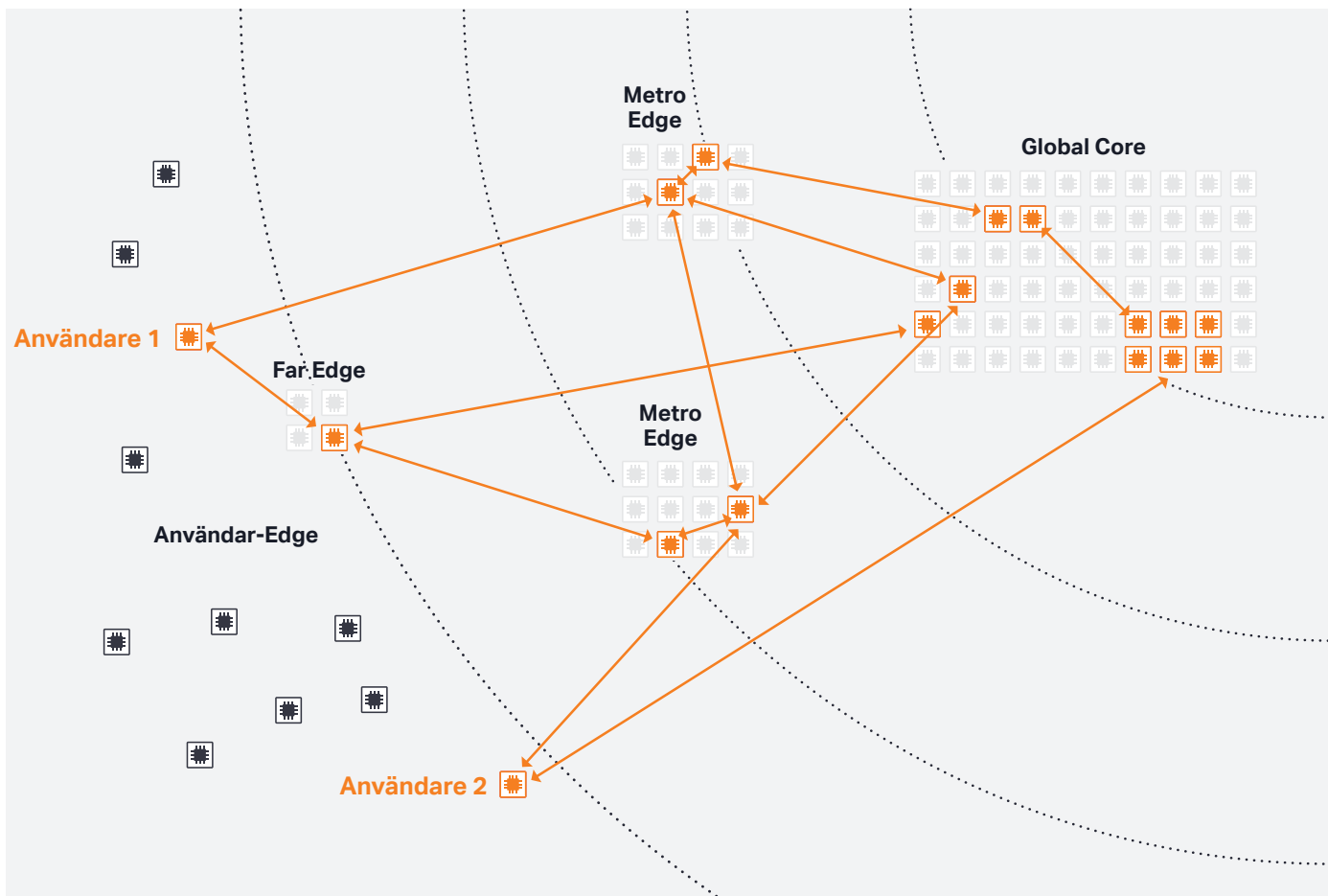


Bild 4. Den dynamiska miljön för applikationer i Edge Cloud

Det är inte bara var nätverkskanten finns som är flytande; det gäller även karaktären av applikationerna vid kanten på nätverken, som kan vara antingen statiska eller dynamiska. Varje ruta i bild 4 representerar en Edge Computing-plats för en applikation, mikrotjänst eller nätverksfunktion som tjänar en slutanvändare (konsument eller företag). Under den tid som en användarapplikation körs kan den komma att utnyttja bearbetningsresurser på flera olika Edge Cloud-platser. Olika användare kan komma att använda molnresurser på olika platser (Far Edge, Metro Edge, Global Core), beroende på vad det är för applikation och vilka molnresurser som finns tillgängliga och kan tillgodose QoE-kraven under applikationens session. Applikationerna kommer alltså att röra sig mellan olika Edge Cloud-platser och det är denna dynamik som kräver nya nätverksresurser i Edge Cloud.

Vilka är Edge Cloud-leverantörerna?

Många aktörer behöver bygga nya affärsrelationer för att kunna dra nytta av Edge Cloud, inte bara applikationsutvecklarna. Här finns också hyperscalers, GCN-operatörer, DCOs och CSPs. När kapplöpningen mot Edge Cloud snabbas upp, kommer leverantörerna i ekosystemet att behöva samarbeta. Partnerskap växer redan fram och vi kan vänta oss en dramatisk ökning under kommande år.

I följande avsnitt tittar vi närmare på hur olika Edge Cloud-leverantörer förväntas utveckla sina datacenterstrategier för att lägga grunden till Edge Cloud (bild 5).

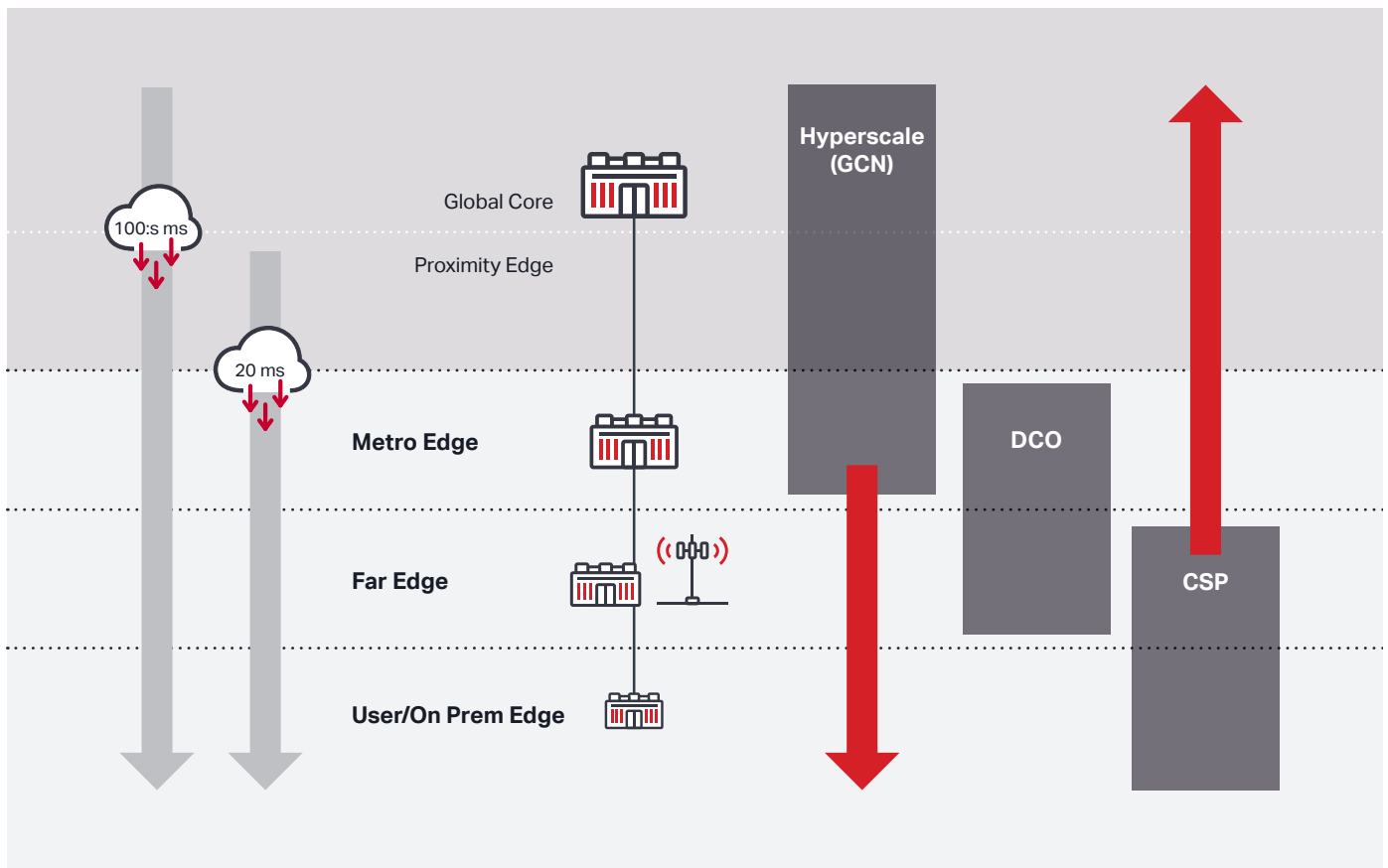


Bild 5. Edge Cloud-leverantörer jämfört med Edge-datacenter

GCNs: GCNs har byggt ut en omfattande uppsättning hyperscalecenter (global core) som en del av den decentraliserade molnarkitekturen. De har också inlett samarbeten med DCOs så att deras metronätverk kan expandera mot nätverksskanten. För att kunna gå mot distribuerade molntjänster med lägre latens behöver de expandera sitt Edge-fotavtryck till Far Edge och User / On-premises Edge-platser. Det kan de göra, antingen på egen hand eller tillsammans med de CSPs som dominerar detta utrymme.

Exempel på framväxande strategiska partnerskap mellan GCNs och CSPs är AT&Ts samarbete med Google Cloud och Azure, och Verizons samarbete med AWS. Edge Cloud-tjänsterna kommer att konsumeras dynamiskt. Därför förväntar sig GCNs att DCO- och CSP-nätverken ska tillhandahålla högre Edge Cloud-medvetenhet, när GCN-applikationerna körs på dem. För att ta ett exempel: När GCNs streamar OTT-videor i sina moln måste de försäkra sig om att CSPs har tillräckligt med nätverksresurser tillgängliga, så att konsumenterna får stabila videostreamar.

DCOs: DCOs har byggt ut ett stort antal datacenter i metronätverken, så att deras företagskunder kan outsourca sina egna datacenter. De erbjuder också utrymme och datorkraft till molnleverantörer och har utbyten med innehålls- och molnleverantörer. Deras primära intäktbas består av fastigheter, datorkraft och anslutningar, men de inser att de även behöver verka högre upp i molnstacken för att öka intäkterna och marginalerna. De kommer även fortsättningsvis att spela en nyckelroll i utbyggnaden av Edge Cloud.

CSPs: CSPs är de som huvudsakligen tillhandahåller anslutningar och infrastruktur till slutanvändare – konsumenter och företag – idag. För att åstadkomma detta har de byggt tusentals centrala anläggningar/huvudändrar genom åren. De virtualiserar också sina interna nätverk; de får en Edge-arkitektur som ligger nära slutanvändaren. Som vi nämnde ovan, satsar CSPs på nya samarbeten med GCNs och skapar medvetenhet om nätverksresurserna som kopplar deras Edge Cloud-platser till applikationerna som tillhandahålls av GCNs och vice versa.

Krav för att leverera Edge Cloud-tjänster

Edge Cloud är dynamiskt och kräver att de olika nätverksaktörerna i ekosystemet ändrar sitt sätt att tänka. För att lyckas med en Edge Cloud-installation, måste GCNs, DCOs och CSPs förstå vad deras nya nätverk måste klara och hur de måste agera. Här följer de viktigaste kraven på Edge Cloud-nätverk:

1. Applikationsmedvetenhet: Det är applikationsnätverken som kommer definiera hur molntjänster och applikationer hanteras i nästa generationens nätverk. Idag körs applikationer som drivs av GCNs över en virtualiserad infrastruktur långt från den fysiska infrastrukturen. Man strävar mot en effektiv drift av applikationer som körs över geografiskt spridda Edge Computing-resurser. För att åstadkomma detta krävs det att applikationsnätverken och de fysiska infrastrukturnätverken är medvetna om varandras karaktäristik och krav (overlay och underlay, se bild 6).

- 2. Visibilitet över nätverkets och applikationernas arbetsbelastning och placering:** För att möta de skiftande kraven vid nätverkskanten krävs att både applikations- och nätverksleverantörerna får bättre visibilitet över infrastruktur- och applikationslagren, så att de kan se var trafikstockningar uppstår och var man kan förvänta sig problem. Denna nivå av visibilitet måste stödjas av alla lager och fungera i en miljö med produkter från flera tillverkare.
- 3. Säkerhet:** Det är mer komplicerat att upprätthålla en stabil säkerhetsnivå när applikationerna blir mer utspridda och dynamiska. Det är enklare för företag att upprätthålla enhetlig teknisk och fysisk säkerhet när molnresurser och applikationer huserar i ett enda datacenter. Edge Cloud är betydligt mer komplex. Här måste ekosystemet hantera både framväxande säkerhetsmodeller och fysiska säkerhetsparametrar med nolltillit, och detta för tre gånger fler, utspridda Edge-dataplatser.

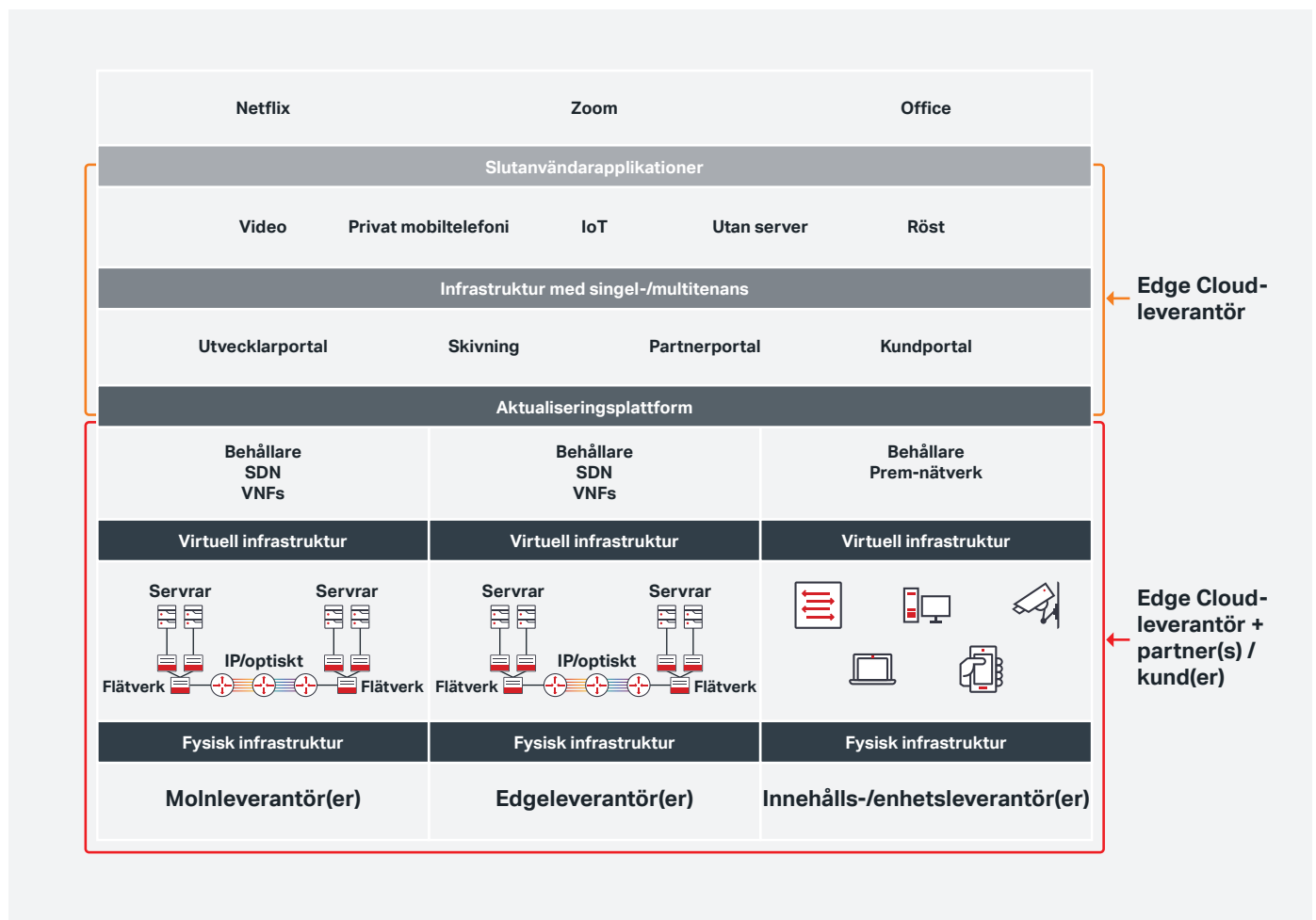


Bild 6. Komponenterna i Edge Cloud-ekosystemet

4. Analys: Realtidsdata är en viktig affärstillgång och datamängden kommer att fortsätta att växa när mer data samlas in från slutpunkter, särskilt när det gäller den väntade explosiva tillväxten av IoT-enheter. Man behöver betrakta dataanalys från två perspektiv:

- Analys av data i rörelse – hämtning och analys av data i realtid eller nästan realtid; viktig metadata skickas norrut för lagring nära källan, ofta vid nätverkskanten
- Data i vila – data aggregeras/slås samman med andra datamängder för olika former av affärsanalys i ett centraliserat datacenter

Datan är dock bara meningsfull om analysen leder till insikter som man kan agera utifrån – resultaten måste kunna leda till åtgärder. AI, maskininlärning och analysverktyg är centrala för att kunna förstå de skiftande kraven från Edge-applikationerna och för att kunna förbättra deras och nätverkets prestanda.

5. Skalbar kapacitet för Edge-datacenter (inter/intra)

nätverk: Edge Computing skapar nya förväntningar på högpresterande, ständigt aktiva applikationer i och mellan Edge-datacentren och det centrala molnet. Trafiken i metrodelen av nätverken förväntas öka i högre takt än i core delen. Med Edge Computing för höga bandbredder spås detta på ytterligare. Denna kapacitetsinversion för core-till-metro kommer att kräva en skalbar paket-optisk infrastruktur, både i och mellan Edge-datacentren, och med upprätthållen konnektivitet till det centrala molnet.

6. Intelligent Edge Cloud-orkestrering och

automatisering: Målet är att optimera användandet av Edge Cloud-resurser och på samma gång tillgodose efterfrågan på och kraven från dynamiska applikationer. Då krävs intelligent automatisering, med ett fågelperspektiv över nätverket och Edge Cloud. Separata orkestreringsinstanser kommer att hantera moln-/Edge-virtualisering, plattform, infrastruktur och applikation för att placera och koppla samman applikationskomponenter till lämpliga Edge Cloud-vårdar. Dessa baseras på slutanvändarens plats; applikationsresursen, QoS och tjänstspecifikationerna; värdens förmåga, kapacitet, kostnad och tillgänglighet; nätverkets kapacitet och prestanda; samt begränsningar när det gäller exempelvis operatör, regelverk och klient. Det centraliserade molnet levererar till en stor mängd användare, men varje Edge-applikation knyts till en mindre mängd. Då måste den kunna svara omedelbart, dynamiskt och automatiskt på varje enskild kunds krav och detta i en miljö med begränsade resurser.

7. Edge Cloud-skivning för multitenans: Med Edge Cloud har nätverksleverantörer möjlighet att dynamiskt allokera olika moln- och nätverksresurser för varje enskild nätverkskund över sina Edge-datacenter. Man kan kalla detta Edge Cloud-skivning. Det handlar om att leverera databearbetning, lagring och nätverksresurser till nätverkskanten från slutpunkt till slutpunkt, baserat på nätverkskundens applikations- och servicenivåkrav.

Vad är Edge Cloud?



Adaptive Network™ banar väg för Edge Cloud

En av de största utmaningarna för Edge-leverantörer är effektiv och intelligent hantering av nätverks- och applikationsresurserna för datacenter i Edge Cloud under perioder med högt användande. Med Cienas Adaptive Network får leverantörerna ett ramverk för Edge Cloud så att de kan förverkliga ett end-to-end-nätverk med kollektiva krafter. Detta blir smartare och smidigare för varje dag och har tillräckligt stor skala för att kunna svara dynamiskt på trycket utifrån.

Adaptive Network gör att Edge Cloud-leverantörer kan optimera sina befintliga system samtidigt som de inför ny teknik och nya arbetsmönster, för att möta kraven från Edge Cloud. Adaptive Network bygger på fyra grundstenar: programmerbar infrastruktur, analys och intelligens, programvarukontroll och automatisering, samt tjänster. De bidrar var och en för sig till bättre resultat i nätverket och ekonomin och när de samverkar blir resultaten än bättre.

Programmerbar infrastruktur: En programmerbar paket-optisk infrastruktur för nätverkskanten ska kunna nås och konfigureras via öppna gränssnitt. Den ska också vara skalbar och kunna exportera nätverkets prestandadata i realtid till applikationslagret i Edge Cloud och kunna justera sina resurser efter vad applikationslagret behöver. Uppnås dessa viktiga faktorer får man ett applikationsmedvetet nätverk och den skalbarhet som behövs för att koppla samman Edge Cloud-flätverk i och mellan Edge Cloud-datacentren. Nätverksskivning i infrastruktur-lagret kommer också att vara centralt för leverantörer, när man vill kunna erbjuda Edge Cloud-tjänster till flera olika molnleverantörer och applikationsöverlägg.

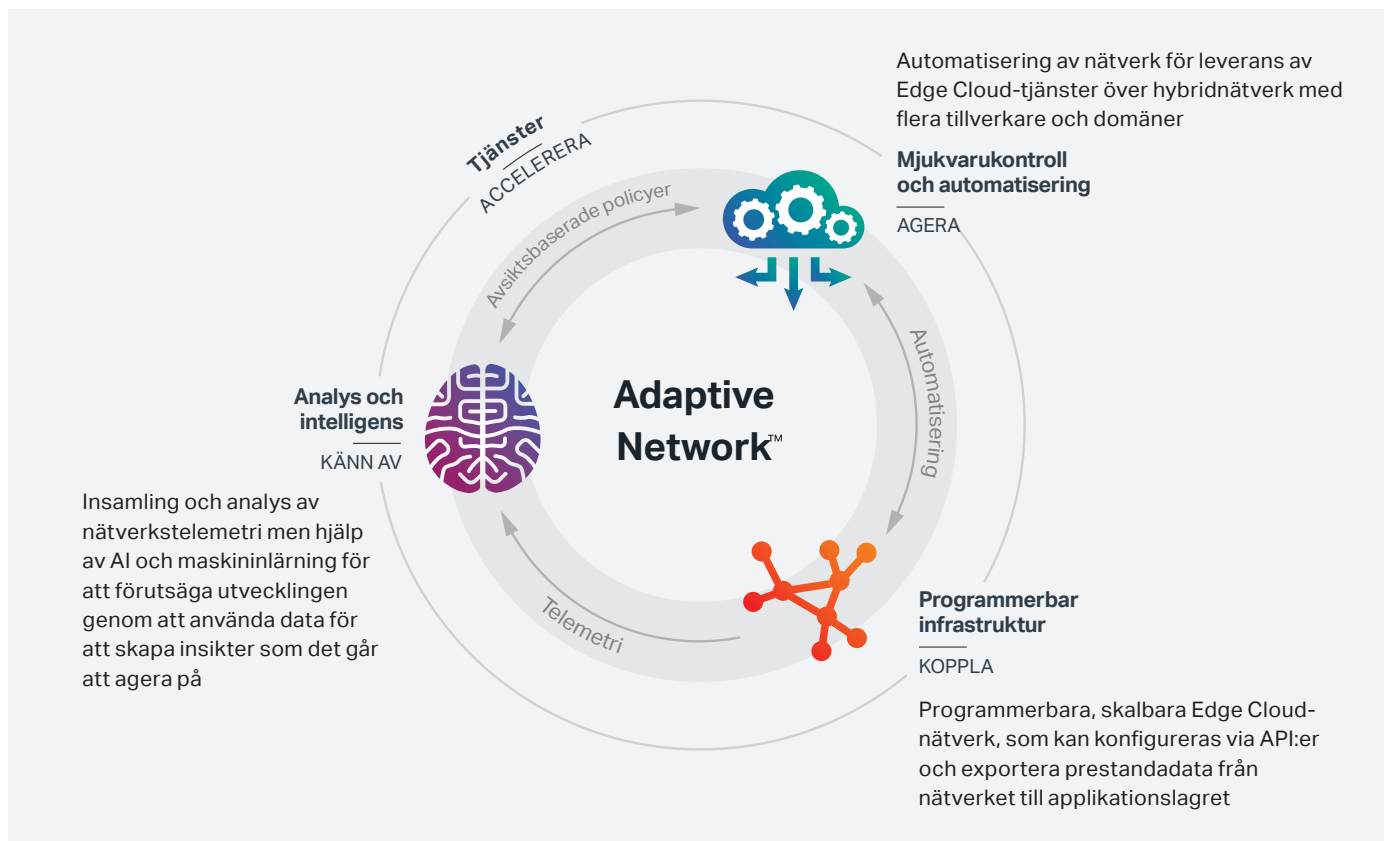


Bild 7. Adaptive Network: ett ramverk som möjliggör Edge Cloud

Analys och intelligens: Edge Cloud är en utveckling av Cloud Computing som till stor utsträckning bygger på automatisering. Denna bygger i sin tur på tolkning av enorma strömmar med telemetri/nyckeltal (Key Performance Indicators, KPI) från underliggande resurser. Applikationsmedvetenhet och automatisering kommer att vara beroende av insamling och analys av nätverks-, server- och virtualiseringsresurser med hjälp av AI. Dessutom krävs förmågan att förutsäga mönster genom att utvinna användbara insikter ur gigantiska datamängder. Dessa insikter kommer att ge ett applikationsmedvetet nätverk som kan känna av och anpassa sig till Edge-applikationernas behov, i realtid och på ett säkert sätt.

Programvarukontroll och automatisering: De viktigaste begränsningarna och kraven när det gäller Edge Cloud innebär att man måste kunna automatisera och placera ut arbetsbelastningar efter vad applikationerna kräver nästan

i realtid. Genom att implementera DN, NFV och öppna API:er kan leverantörer förenkla styrningen, säkringen och automatiseringen av sina nätverk från slutpunkt till slutpunkt, så att de kan leverera Edge Cloud-tjänster över hybridnätverk med flera tillverkare och domäner.

Konsulttjänster: Tekniska och professionella konsulttjänster krävs för att kunna hjälpa leverantörerna att ta fram bästa möjliga strategi och arkitektur för Edge Cloud och för att bygga, driva och förbättra sina nätverk – och samtidigt skynda på vägen mot Adaptive Network.

Adaptive Network-visionen
Läs mer



Sammanfattning

Vi pratar om de första stegen mot Edge Cloud och en distribuerad molnarkitektur. Man ska inte betrakta nätverkskanten som ett specifikt datacenter, utan den kommer att finnas på vitt skilda platser, helt beroende på QoE-förväntningar och varje enskild applikations resurskrav och -tillgänglighet. Var en applikation befinner sig kan variera mellan olika Edge-datacenter under dess livscykel. Det driver på behovet av att kunna skala infrastrukturen på ett intelligent sätt både i och mellan Edge-datacenter och det centrala molnet. Detta samtidigt som arbetsbelastningar automatiseras mellan platser i nätverkskanten.

GCNs har redan visat att de framgångsrikt kan skala i en centraliserad publik hybridmodell. Men när de går mot en distribuerad Edge Cloud-modell krävs samarbeten med DCOs och CSPs, så att de kan använda deras omfattande infrastruktur närmare slutanvändarna. För att den distribuerade Edge Cloud-modellen ska kunna nå sin fulla potential måste man ta hänsyn till nya nätverkskrav. Först och främst måste applikationslagren i molnstacken ha dynamisk medvetenhet om resurserna i nätverkslagren och vice versa.

Redan nu spelar Cienas Adaptive Network-lösningar en central roll i några av världens största interdatacenter och molnarkitekturer. Ciena är marknadsledande både när det gäller Global DCI och Metro DCI. Därför är företaget perfekt rustat för att ta sina djupa kunskaper om molnet och DCI-marknader till nätverkskanten. Adaptive Network tillhandahåller ett ramverk för samtliga leverantörer i Edge-ekosystemet. Här tacklas olika utmaningar i Edge Cloud-modellen genom programmerbar och skalbar infrastruktur, analys och automatisering. På så vis kan både nätverks- och applikationsresurser skalas upp och ned dynamiskt, för att leva upp till slutanvändarnas förväntningar. En anpassning till Adaptive Network-ramverket bidrar till att en Edge Cloud-modell blir tillräckligt skalbar och anpassningsbar, så att den kan möta de ständigt föränderliga kraven i nätverkskanten.



Var innehållet användbart?

Ja

Nej