

# ANVÄNDNING AV NAND-FLASH SSD I DATORMOLNTJÄNSTER

Gabriel Kivilohkare

Handledare: Jerker Björkqvist

## Referat

Denna avhandling presenterar datormoln, dess teknik och datalagring. Inledningsvis behandlas allmän fakta om datormoln och SSD, varefter det presenteras olika sorters datormoln allmänt samt deras fördelar och nackdelar ur slutanvändarens synvinkel. Därefter beskrivs kort om datormolnets tjänsteleverantörer och datormolnets uppbyggnad. Efter detta kommer det att presenteras detaljerat hårddiskar, hybriddiskar och SSD som datalagrings alternativ i datormoln. Slutligen beskrivs detaljerat om SSD i datormoln ur olika aspekter, samt gör flera jämförelser med olika datalagringsmedium. Tyngdpunkten i detta arbete kommer att vara i hur användningen av SSD påverkar datormoln.

**Nyckelord:** Datormoln, SSD, solid-state-drive, hårddiskar, datalagring, NAND Flash

## Innehållsförteckning

Referat.....	2
<b>1. Inledning.....</b>	<b>4</b>
<b>2. Allmänt om datormoln.....</b>	<b>5</b>
2.1 Infrastruktur som tjänst.....	5
2.2 Plattform som tjänst .....	5
2.3 Mjukvara som tjänst .....	6
2.4 Fördelar med datormoln till slutanvändaren .....	6
2.5 Nackdelar med datormoln.....	7
<b>3. Datormolntjänsteleverantörens grundverktyg.....</b>	<b>8</b>
3.1 Virtualisering .....	8
3.2 Hårdvarans betydelse .....	9
3.3 Servicens tillgänglighet .....	9
<b>4. Datalagring i datormoln .....</b>	<b>10</b>
4.1 Hårddiskar .....	10
4.2 SSD .....	11
4.3 Hybriddiskar .....	12
4.4 Datalagringsflaskhalsar i datormoln .....	12
<b>5. SSD i datormoln .....</b>	<b>13</b>
5.1 Användning av SSD i datormoln .....	13
5.2 SSD:s påverkan i uppstartningen av datormolntjänster .....	13
5.3 Strömförbrukning i SSD.....	14
<b>6. Slutsatser .....</b>	<b>15</b>

# 1. Inledning

Datormoln (eng. ”*Cloud Computing*”) är ett omtalat begrepp inom informationsteknologin. Konceptet av datormoln härstammar från 1950-talet, då omfattande centraldatorer användes av skolor och företag. En av största milstolparna inom datormoln förekom 1999, då Salesforce.com var först med konceptet att leverera system till företag genom en webbläsare. Datormolntjänsternas popularitet har vuxit stort sedan det. För tillfället är 7 % av all data lagrad i datormoln, och detta uppskattas öka till 36 % år 2016 [1]. I början av år 2013 uppskattades det finnas över en exabyte data totalt i datormoln, vilket motsvarar 1 073 741 824 gigabyte. Således är det naturligt datalagringsmediumet spelar stort roll.

NAND flash SSD har blivit mer använt i både konsumentmarknaden och i datacentraler. Tills vidare är traditionella hårddiskar det allmännaste datalagringsmediumet, men populariteten av SSD ökar konstant. Det finns flera sorters SSD, men i denna avhandling koncentrerar jag endast i NAND flash SSD, eftersom den är den allmännaste sorten av SSD. Också finns det skillnader i NAND flash SSD från varandra (se kap. 4.2) vilket lönar sig att beakta vid val av SSD.

## 2. Allmänt om datormoln

Generellt definieras datormoln som tjänster som levereras över Internet. Dessa tjänster befinner sig på datormolnets tjänsteleverantörs fysiska resurser, och körs ofta genom ett lätt användargränssnitt såsom en webbläsare. Således behöver inte slutanvändaren av tjänsten installera något på sin egen dator eller investera i hårdvara. Kunden betalar endast enligt de resurser och tjänster som kunden använder. Underliggande infrastrukturen och dess underhåll sköts av datormolnets tjänsteleverantör. Enligt NIST (National Institute of Standards and Technology) kännetecknas datormoln av följande egenskaper; självbetjäning enligt efterfrågan, bred nätverksåtkomst, resurspoolning, snabb elasticitet och uppmätbar service [2]. Allmänt indelas molntjänster i tre kategorier; infrastruktur som tjänst, mjukvara som tjänst och plattform som tjänst.

### 2.1 Infrastruktur som tjänst

Infrastruktur som tjänst (eng. ”*Infrastructure-as-a-Service (IaaS)*”) är den mest grundläggande formen av molntjänster. Slut användaren har tilldelats grundläggande datorresurser, såsom minne, diskutrymme och processorkraft från datormolnets tjänsteleverantörs delade server. Tjänsten är virtualiserad till användarna. Också kan användaren ofta välja operativsystem och kontrollera diskutrymmet [3]. Ett exempel på IaaS är Amazon EC2.

### 2.2 Plattform som tjänst

Plattform som tjänst (eng. ”*Plattform-as-a-Service (PaaS)*”) är en utvecklingsplattform där utvecklare av program kan arbeta på tjänsteleverantörens fysiska resurser utan att behöva fokusera på underliggande infrastrukturen. Programmen körs på ovannämnda infrastruktur, och levereras till utvecklarna via

Internet. Slut användaren ser endast färdiga programmet [4]. Exempel på dessa tjänster är Google App Engine och Microsoft Azure. Också en nämnvärd tjänsteleverantör är [www.elastx.com](http://www.elastx.com), vars plattform som tjänst är byggd på SSD.

### **2.3 Mjukvara som tjänst**

Mjukvara som tjänst (eng. ”*Software-as-a-service (SaaS)*”) är tjänster där användaren köper mjukvara av tjänsteleverantören och behöver inte själv investera i hårdvara eller licenser. Mjukvaran används typiskt genom en webbläsare. Datormoln tjänsteleverantören använder en s.k. multitenant arkitektur, där användarna har sin egen miljö, och således är infrastrukturen delad till användarna. Exempel på mjukvara som tjänst är [Salesforce.com](http://Salesforce.com).

### **2.4 Fördelar med datormoln till slut användaren**

Datormoln tjänster är väldigt kostnadseffektiva för användning, upprätthållning och uppdatering. Köparen av datormoln behöver inte investera i hårdvara, mjukvara eller licenser, utan man betalar enligt den mängd service man har behov till. Allmännaste faktureringsätt är betalning enligt användning och betalning av ett fast pris. Betalning enligt användning medför att tjänsten är elastisk och skalbar, vilket är speciellt viktigt för växande företag och kunder med ändrande behov. För fast prissatta tjänster bör man känna till sina krav för att hålla tjänsten kostnadseffektivt.

En annan viktig fördel med datormoln är dess tillgänglighet. Man kan åtkomma tjänsten var som helst i världen, bara man har tillgång till Internet [5]. Också är i ibruktagningen av datormoln oerhört snabbt.

## 2.5 Nackdelar med datormoln

Valet av datormoln tjänsteleverantör är ytterst viktigt med tanke på säkerheten av speciellt sensitivt data, eftersom all data befinner sig på tjänsteleverantörens fysiska resurser. Administratoreernas tillträde till data och säkerställandet av datans raderande [6] är fortfarande utmaningar i datormoln. Utöver detta kan tjänsteleverantörerna ha tekniska problem. Dessa kan bero bland annat på eventuella intrångsförsök till datormolnet, överbelastning hos servern och hög latens.

Som lösning till en del nackdelar har vissa datormoln tjänsteleverantörer tagit i bruk servicenivåavtal (eng. ”*Service-level-agreement*”), som är ett kontrakt mellan kunden och tjänsteleverantören. Denna bestämmer specifika krav för servicen, och då denna servicenivå inte uppfylls, följer överenskomna sanktioner. Till problemet med latens har SSD en positiv inverkan.

### **3. Datormolntjänsteleverantörens grundverktyg**

Tjänsteleverantörerna av datormolnen (eng. ”*Internet Service Provider*”) underhåller den infrastruktur där datormoln tjänsterna befinner sig. De hyr delade resurser av server maskiner, där kunderna betalar enligt den mängd de använder av servicen. Tekniken som möjliggör partitioneringen av resurserna kallas *virtualisering*.

#### **3.1 Virtualisering**

Virtualiseringen av fysiska resurser är en av huvudpunkterna i datormoln. Med hjälp av virtualisering kan en fysisk resurs tilldelas till flera logiska resurser. Tyngdpunkten med virtualisering är att dela hårdvaru resurser osynligt till slutanvändaren, applikationer, plattformar, datalagringsystem, datorer och infrastrukturella komponenter [7]. Med virtualisering och *hypervisors* (se kapitel 3.3) kan en fysisk resurs bilda flera virtuella maskiner (eng. ”*Virtual machine*”) med egna operativsystem.

Virtualisering möjliggör att man kan använda mera resurser än vad som är fysiskt möjligt med underliggande fysiska resurserna. Detta sker genom att använda inaktiva resurser till flera virtuella maskiner, då servern har låg trafik [8]. Vanliga virtualiseringsteknik är VMware och Xen.

##### **3.1.1 Hypervisor**

Hypervisors (eng. ”*Hypervisor, också Virtual Machine Manager (VMM)*”) är en mjukvara eller fast programvara som möjliggör att flera gästmaskiner kan köras från en värdmaskin (eng. ”*Host*”). Virtuella maskinerna är gästmaskiner med egna operativsystem, medan värdmaskinen är en fysisk dator [9] som delar på de



fysiska resurserna. Hypervisorns huvuduppgift är att ge gästmaskinerna sina egna isolerade omgivningar. SSD i värdmaskinen påverkar stort på gästmaskinernas funktionalitet.

### **3.2 Hårdvarans betydelse**

Datormolnets tjänsteleverantör ansvarar för den underliggande infrastrukturen hos datormolnet. Detta omfattar bland annat fysiska resurserna, strömförsörjningen och internetanslutningen.

Som i personliga datorer spelar hårdvaran också stor roll hos server datorer. Användarna av datormoln använder resurser ur den underliggande infrastrukturen. Hårdvaran har stor betydelse med tanke på servicens funktionalitet, tillgänglighet och smidighet. Eftersom SSD är snabbt jämfört med hårddiskar, förbättrar de datormolntjänsternas prestanda. Också påverkar andra komponenter såsom CPU och minne på prestandan. Mängden resurser fördelas till användarna av datormolntjänsterna, och datormolnets köpare betalar beroende på mängden resurser som behövs.

### **3.3 Servicens tillgänglighet**

Tjänsteleverantörerna av datormoln lovar typiskt en tillgänglighet på över 99 %. Till exempel har Amazon EC2, en IaaS tjänst, 99.95% servicenivåavtal [10], vilket innebär att servicen får maximalt ha en tid ur funktion på dryga fyra timmar per år. Majoriteten av användarna av datormoln kommer antagligen inte ha någon tid ur funktion, men de som har, kommer antagligen att ha betydligt längre tid ur funktion [11]. SSD kan förbättra en tjänsts drifttid eftersom de är mer tillförlitliga än hårddiskar.

## **4. Datalagring i datormoln**

Molntjänsternas data lagras i tjänsteleverantörens fysiska resurser. För att nå dessa data, måste man ansluta sig till ifrågavarande fysiska resurser. Vanligtvis använder sig datormolnets tjänsteleverantör av hårddiskar, SSD eller hybrida skivor som lagringsmedium.

Val av lagringsmedium till en fysisk resurs är ytterst viktigt, eftersom denna påverkar direkt på hur snabbt datum kan läsas ur servern och skrivas till servern. Också har olika datalagrings alternativ olika levnadslängder, vilket lönar sig att beakta vid val av lagringsmedium. SSD är känt för lång livslängd och god hållbarhet. SSD är betydligt dyrare än hårddiskar, då SSD:s priser börjar från 0,59\$/GB [12] och hårddiskarnas priser börjar från 0,05\$/GB. Trots stora prisskillnaden, håller SSD på att bli allt populärare. Hybriddiskar, som är en blandning av SSD och hårddiskar, kostar typiskt en aning mer än hårddiskar.

### **4.1 Hårddiskar**

Hårddiskar är för tillfället det mest använda datalagringsmediet i överlag. Detta mest på grund av hårddiskens låga pris jämfört med SSD. Fördelar med hårddiskar, förutom låga priset, är att de har en hög lagrings kapacitet.

Hårddisken sparar data magnetiskt på roterande skivor, där en läsarm läser av binära enheterna som befinner sig på skivan. Det faktum att hårddisken har flera rörliga delar orsakar att den kan upphettas, hålla högt ljud och har en rätt begränsad livslängd.

## 4.2 SSD

SSD (eng. ”*solid-state-drive*”) har inga rörliga mekaniska delar. Nuförtiden baserar sig SSD oftast på NAND flash teknik i datalagring. SSD:s marknadsdel har vuxit stort under de senaste åren. Den använder flera NAND flash minneskretsar parallellt, och är således snabbare än en ända NAND flash minneskrets [13].

SSD:s mest omtalade egenskap är utan tvekan dess datalagrings hastighet, som beror delvis av att den sparar och åtkommer data rakt på minneskretsar [14]. Eftersom SSD inte har rörliga delar, är dess försöndring osannolikt. Dess livslängd beror till det mesta på hur mycket data det skrivs till denna. Enligt Intel skall dess 525 seriens SSD klara av minst 20GB skrivning i dagen under fem år [15].

Tre celltyper som SSD har är SLC (single-level cell), MLC (multi-level cell) och TLC (triple-level cell). SLC sparar endast en bit per cell, medan MLC sparar två bitar per cell och TLC sparar tre bitar per cell. TLC har kortaste livslängden och sämsta prestandan av de tre celltyperna. Den klarar av omkring 1000 skrivcykler per cell, emedan MLC klarar av omkring 10 000 skrivcykler per cell och SLC till och med 100 000 skrivcykler per cell [16]. Därmed kan man räkna med att SSD med en bestämd lagringskapacitet med celltypen SLC har tio gånger längre livstid än MLC med samma lagringskapacitet. SLC är dyraste att tillverka per megabyte, medan TLC är billigast. Vanligen kostar SLC ungefär dubbelt mer än MLC. Också har SLC högsta överföringshastigheten och minsta strömförbrukningen. Till företag som kräver prestanda och tillförlitlighet, är SLC bästa alternativet. MLC är däremot vanligare inom kundmarknaden.

### **4.3 Hybriddiskar**

En hybriddisk (eng. ”*Solid State Hybrid Drive (SSHD)*”) är en blandning av hårddiskar och SSD. Dessa sparar prestanda krävande funktioner, såsom ofta använt data, på NAND flash minnet, medan mindre krävande funktioner i hårddisken. Hybriddiskar blir således mycket billigare än SSD, och har samma lagringskapacitet som hårddiskar. Hybriddiskarna består av en liten SSD del och en större hårddisk del. SSD delen är ofta omkring 8GB och hårddisk delen flera hundra gigabyte. Största nackdelen med hybriddiskar är att SSD delen är liten, och ifall prestanda krävande data ändrar konstant, förminskar nyttan av SSD delen.

### **4.4 Datalagringsflaskhalsar i datormoln**

Datalagringsflaskhalsar kan uppstå i datormoln då det inte finns tillräckligt mycket kapacitet för datahanteringen. Detta kan leda till dataflödet försämras eller stannar fullständigt. En av allmännaste flaskhalsarna beror på datalagrings I/O, till vilket SSD ger en lösning eftersom SSD har snabb datahämtning, snabb dataskrivning och mindre latens i jämförelse med hårddiskar [17].

IOPS är en standard mätenhet för mängden input/output operationer en datalagrings enhet kan göra i sekunden. SSD möjliggör väldigt mycket mera IOPS jämfört med hårddiskar. Detta försnabbar flera datormolntjänster, exempelvis webbutiker. Webbutiker använder ofta databaser aktivt, och SSD skivorna försnabbar webbutiken signifikant, eftersom de möjliggör väldigt snabba databas förfrågningar. Krävande webbapplikationer har också nytta av SSD. Kortfattat gör SSD webbutiker och andra webbapplikationer snabbare på grund av väldigt snabb datahämtning och skrivning av data.

## **5. SSD i datormoln**

Andelen av SSD i datormolntjänster håller på att stiga och priset på dem sjunker konstant. Flera datormoln tjänsteleverantörer marknadsför kraftigt att de använder SSD i sina tjänster, vilket ännu idag är rätt sällsynt. Trots att priset på SSD sjunker, är de fortfarande mångfalt dyrare än traditionella hårddiskar. Därför har majoriteten av datormoln tjänsteleverantörer hårddiskar i sina fysiska resurser.

### **5.1 Användning av SSD i datormoln**

Flera datormolntjänster försnabbas drastiskt av SSD skivor. Också påverkar dessa på hastigheten av ibrukstagningen av tjänsten och vissa administrativa funktioner, såsom användningen av olika loggar. Hastigheten av uppstartningen av tjänsterna är relevant för tjänstens kvalitet samt kundnöjdheten. Också påverkar detta på hela upplevelsen av tjänsten.

SSD:s latens är ytterst liten i jämförelse med hårddiskar, vilket medför att data kan åtkommas med mindre dröjning. Vanligen har SSD fem gånger mindre latens än hårddiskar [18].

### **5.2 SSD:s påverkan i uppstartningen av datormolntjänster**

Uppstartningen av alla datormolntjänster blir snabbare med SSD än hårddiskar, speciellt uppstartningen av krävande tjänster. Som exempel nämner jag uppstartningen av molnbaserade virtuella maskiner, som är utan tvekan den vanligaste tjänsten under kategorin infrastruktur som tjänst. Uppstartningen av

dessa med SSD är betydligt snabbare än med hårddiskar. Enligt Altech:s undersökning tar uppstartningen av tre virtuella maskiner med en 7200RPM HDD 62 sekunder, medan uppstartningen av samma mängd virtuella maskiner med Corsair Neutron GTX –SSD tar endast 16 sekunder. Då de gjorde undersökningen om med sex virtuella maskiner, blev uppstartningen med SSD fortfarande 16 sekunder, medan med 7200RPM hårddisken dröjdes uppstartningen till 134 sekunder [19].

Utöver den snabba uppstartningstiden, blir även ibruktagningstiden av datormolntjänsterna snabbare med SSD. Exempelvis erbjuder Digital Ocean SSD infrastruktur tjänster med endast 55 sekunders ibruktagningstid efter gjord beställning.

### **5.3 Strömförbrukning i SSD**

SSD har betydligt mindre strömförbrukning än hårddiskar eller hybriddiskar. Detta mest på grund av att SSD har inga rörliga delar. Av samma orsak uppvärms inte SSD lika mycket som hårddiskar, vilket leder till att datacentralers behov av förkylning blir mindre. Detta sparar märkvärdiga mängder ström.

Då datalagringsmediumet är i viloläge, använder den en viss mängd ström. Enligt Hewlett-Packards undersökning använder SSD endast omkring en fjärdedel av den ström som hårddiskar använder i viloläge [20]. Också är strömförbrukningen under belastningstoppar och i överlag mindre i SSD än i hårddiskar. Detta faktum medför att SSD ger signifikanta strömbesparningar till datacentraler.

SSD använder mindre än 80 % av den ström som hårddiskar använder. Utöver detta, skulle det behövas hundrafalt av hårddiskar till att uppnå samma IOPS prestanda. Detta leder till att för varje I/O genomförande använder SSD bara 0,51 % av den ström som hårddiskar använder [21]. Därmed blir förhållandet mellan strömanvändning och prestanda utmärkt i SSD.

## 6. Slutsatser

Detta arbete har diskuterat om att SSD är ett verkligen anmärkningsvärt alternativ till datalagringsmedium i datormoln. Andelen av SSD i datormoln kommer antagligen att öka stort på kommande åren. Populäriteten av SSD i persondatorer har redan vuxit ytterst mycket, och samma sak kommer antagligen att förstärkas inom datacentraler. Detta mest på grund av att SSD blir allt billigare. Dessutom är hårddiskar rätt gammalmodiga, medan SSD är väldigt trendigt. Berömda företag, såsom Netflix, har övergått till SSD. Eftersom övergången har fått mycket publicitet, har SSD:s ställning blivit avsevärt bättre.

I jämförelserna med SSD mot hårddiskar, var SSD bättre på alla tekniska perspektiv. I flera datormolntjänster behövs den prestanda som SSD ger, och därför är det rekommendabelt för tjänsteleverantörer att övergå till SSD.

## 7. Litteraturförteckning

- [1] Dixie, Alex <http://www.olswang.com/articles/2013/05/content-meets-the-cloud-infographic/> hämtad 13.9.2013
- [2] Mell, Peter – Grance, Timothy, The NIST Definition of Cloud Computing, <http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf> 17.6.2013
- [3] Fu, Lixin – Gondi, Chandana, Cloud Computing Hosting, Department of Computer Science University of North Carolina, Greensboro, 978-1-4244-5540-9/10 ©2010 IEEE
- [4] Ahson, Syed A. – Ilyas Mohammad, Cloud Computing and Software Services Theory and Techniques, 2011
- [5] ] Kumar, Arjun – Lee, Byung – Lee, Hoon – Kumari, Anu, Secure storage and access of data in cloud computing 978-1-4673-4828-7/12 ©2012 IEEE
- [6] Badger, Lee – Grance, Tim – Patt-Comer, Robert, Cloud Computing Synopsis and Recommendations, National Institution of Standards and Technology 29.05.2012
- [7] Chorafas, Dimitris, Cloud Computing Strategies, Taylor and Francis Group LCC 2011
- [8] Siddhisena, Buddhika – Warusawithana, Lakhmal – Mendis, Mithila, Next Generation Multi-Tenant Virtualization Cloud Computing Platform ISBN 978-89-5519-155-4



[9] Jones, Ken – Musgrave, Devon, Woolley, Valerie, Understanding Microsoft Virtualization Solutions, From the Desktop to the Datacenter, 2nd Edition, Microsoft Corporation

[10] Amazon Web Services, <http://aws.amazon.com/ec2/> hämtad 17.08.2013

[11] Chorafas, Dimitris, Cloud Computing Strategies, Taylor and Francis Group LCC 2011

[12] <http://techreport.com/news/24636/crucial-m500-ssd-now-selling-for-59-cents-per-gig> hämtad 13.9.2013

[13] Park, Kwanghee – Lee, Dong-Hwan – Woo, Youngjoo – Lee, Geunhyong – Lee, Ju-Hong – Kim Deok-Hwan, Reliability and Performance Enhancement Technique for SSD array storage system using RAID mechanism, 978-1-4244-4522-6/09 ©2009 IEEE

[14] Ko, Dae-Sik – Cheong, Seung-Kook, Web Performance Enhancement of E-business System using the SSD, 978-0-7695-3546-3/08 © 2008 IEEE

[15] Intel,  
<http://www.intel.com/content/dam/www/public/us/en/documents/product-specifications/ssd-525-specification.pdf> 328354-002US, hämtad 05.08.2013

[16] EMC Corporation, Considerations for Choosing SLC versus MLC Flash, <http://www.emc.com/collateral/hardware/technical-documentation/300-013-740-choosing-slc-versus-mlc-flash-technical-notes.pdf>  
Hämtad 14.09.2013

[17] Mellanox Technologies, [http://www.mellanox.com/related-docs/whitepapers/WP\\_Solving\\_IO\\_Bottlenecks.pdf](http://www.mellanox.com/related-docs/whitepapers/WP_Solving_IO_Bottlenecks.pdf), Solving I/O Bottlenecks to Enable Superior Cloud Efficiency, hämtad 18.08.2013

[18] Huang, Howie – Li, Shan, Performance Modeling and Analysis of Flash-based Storage Devices, 978-1-4577-0428-4/11/\$26.00 c 2011 IEEE

[19] Altech Computers,

[http://www.altech.com.au/Launch/Article/1045/SSD\\_Benefits\\_For\\_Windows\\_Server\\_2012\\_Part\\_2](http://www.altech.com.au/Launch/Article/1045/SSD_Benefits_For_Windows_Server_2012_Part_2), hämtad 18.08.2013.

[20] Hewlett-Packard Development Company,

<http://h20195.www2.hp.com/v2/GetPDF.aspx/4AA0-0248ENW.pdf>

Hämtad 12.9.2013

[21] Stenc Inc, Solid State Drive - The Performance and Power Advantages

[http://www.stec-inc.com/downloads/whitepapers/Performance\\_Power\\_Advantages.pdf](http://www.stec-inc.com/downloads/whitepapers/Performance_Power_Advantages.pdf)

Hämtad 13.9.2013