



GÖTEBORGS UNIVERSITET

Val av plattform för IBM WebSphere Application Server och IBM WebSphere Portal

**Microsoft Windows Server 2003 jämfört med Linux med
avseende på prestanda ur ett tekniskt och administrativt
perspektiv**

Choice of platform for IBM WebSphere Application Server and IBM WebSphere Portal

**Microsoft Windows Server 2003 compared to Linux in terms of performance from a
technical and administrative perspective**

Martin Samuelsson

Kandidatuppsats i Informatik

**Rapport nr. 2010:023
ISSN: 1651-4769**

Choice of platform for IBM WebSphere Application Server and IBM WebSphere Portal

Microsoft Windows Server 2003 compared to Linux in terms of performance from a technical and administrative perspective

Martin Samuelsson

Department of Applied Information Technology

Abstract

Today there are two dominant technologies among operating systems on servers of companies and organizations. The two are Windows Server and Unix / Unix-like operating systems where Linux is the most common. Choice of operating system can be based on various parameters such as performance, purchase or licensing costs, operating costs and expertise within the organization. IBM WebSphere Portal is a product that can be installed on various platforms including Windows Server and Linux. This thesis intended to explore the differences between Linux and Windows Server 2003 as a platform for WebSphere Portal and thus for WebSphere Application Server in terms of technical performance and maintainability. The pre-study identified the main technical performance factors to be CPU usage, RAM usage and disk activity. From a user perspective response times from WebSphere Portal was considered to be the main performance factor. The pre-study showed that maintainability was an too broad and subjective area to explore within the scope of the study. The factors identified in the pre-study served as a framework for the benchmarks conducted in the main study. The main study showed that Windows Server 2003 consistently outperformed the tested Linux distribution on the same virtual hardware in all tests conducted.

This thesis is written in swedish.

Keywords: WebSphere, Performance, Maintainability, Benchmark, Windows, Linux

Abstrakt

Idag finns två dominerande teknologier bland operativsystemen på servrar hos företag och organisationer. Dessa båda är Windows Server och Unix/unixliknande operativsystem där Linux är det vanligast förekommande. Att välja operativsystem kan göras utifrån olika parametrar som prestanda, inköps- eller licenskostnad, driftskostnad och kompetens inom organisationen. IBM WebSphere Portal är en produkt som kan installeras på bland annat Windows Server och Linux. Uppsatsen ämnade att undersöka vilka skillnader som kunde påvisas mellan Linux och Windows Server 2003 som plattform för WebSphere Portal och därmed också för WebSphere Application Server med avseende på teknisk prestanda och administrerbarhet. Genom en förstudie identifierades de viktigaste tekniska prestandafaktorerna till processoranvändning, RAM-minnesanvändning och hårddiskaktivitet. Sett ur ett användarperspektiv ansågs svarstider från WebSphere Portal som den viktigaste prestandafaktorn. Förstudien visade att administrerbarhet var ett allt för brett och subjektivt område att undersöka inom ramen för studien. Faktorerna som identifierades i förstudien tjänade som ett ramverk för den benchmark-undersökning som genomfördes och som visade att Windows Server 2003 konsekvent presterade bättre än den testade linuxdistributionen på samma virtuella hårdvara i samtliga tester som genomfördes.

Nyckelord: WebSphere, Prestanda, Administrerbarhet, Benchmark, Windows, Linux

Tack

Jag vill särskilt tacka Daniel Gedgudas för ovärderlig hjälp med WebSphere och annat under tiden på Atea IM samt Arne Billestedt för att ha möjliggjort hela studien och varit mån om att alla nödvändiga resurser funnits till mitt förfogande. Tack också till intervjupersoner och övriga medarbetare på Atea IM som hjälpt mig och varit goda medmänniskor. Vidare vill jag tacka min uppsatshandledare Lennart Petersson för goda diskussioner, viktiga synpunkter och all övrig hjälp. Till sist vill jag tacka Malin Sun för att ha hållit mig vid gott mod och varit förstående under långa och svåra skrivstunder.

Innehållsförteckning

1. Bakgrund.....	3
1.1. Problemområde	3
1.2. Syfte och frågeställning	4
1.3. Avgränsningar.....	4
1.4. Disposition	5
2. Teori.....	6
2.1 Plattformar / Operativsystem.....	6
2.1.1 Windows Server (2003).....	6
2.1.2 Linux.....	6
2.2 Applikationsservrar	7
2.2.1 IBM Websphere Application Server.....	7
2.3 Portaler.....	7
2.3.1 IBM Websphere Portal	8
2.4 Java Platform, Enterprise Edition (Java EE)	8
2.5 Benchmarking - Att mäta prestanda	8
3. Arbetsprocess och metod - förstudie	10
3.1 Val av metod	10
3.2 Förstudie genom kvalitativ metod	10
3.2.1 Försthandserfarenhet och utbildning.....	10
3.2.2 Intervjuer	11
3.2.2.1 Beskrivning av intervjupersoner	12
3.2.3 Litteraturstudier	12
4. Resultat - förstudie	13
4.1 Val av kvalitetsattribut	13
4.2 Prestanda - tekniska prestandafaktorer	14
4.3 Administrerbarhet - administrativa prestandafaktorer.....	17
4.4 Sammanfattning av förstudie.....	19
5. Arbetsprocess och metod - huvudsaklig studie	20
5.1 Metod.....	20
5.2 Beskrivning av testfall.....	20

5.2.1 Maskinvaruspecifikation	20
5.2.2 Operativsystem	21
5.2.3 Mjukvaruprodukter	21
5.2.4 Lasttestserver och mjukvara.....	21
5.3 Mätning av tekniska prestandafaktorer.....	22
6. Resultat.....	23
6.1 Test 1 - Stegring 1-60 klienter under 25 minuter.....	23
6.1.1 Processoranvändning	23
6.1.2 Svarstider	24
6.2 Test 2 - Stegring 1-100 klienter under 5 minuter.....	25
6.2.1 Processoranvändning	25
6.2.2 Svarstider	26
6.3 Test 3 - 30 samtidiga klienter under 5 minuter	26
6.3.1 Processoranvändning	27
6.3.2 Svarstider	27
6.4 Minnes- och hårddiskaktivitet	28
6.4.1 Minnesanvändning.....	28
6.4.2 Hårddiskaktivitet	29
7. Diskussion.....	30
7.1 Reflektion kring resultatet	30
7.2 Förslag till fortsatta studier.....	31
8. Slutsats.....	32
9. Referenser	33
Bilagor	35

1. Bakgrund

Idag finns det två dominerande och konkurrerande teknologier bland operativsystemen på servrar hos företag och organisationer. På ena sidan finns Microsoft med sina versioner av Windows Server, på andra sidan finns en uppsjö olika unixbaserade eller unixliknande (*nix) operativsystem. Linux är det vanligaste unixliknande operativsystemet.

Unix, Linux eller Windows Server – det är frågan. Att som organisation välja operativsystem för en eller flera servrar är ett stort ämne. Man kan välja utifrån olika parametrar som prestanda, inköps- eller licenskostnader, driftskostnader, kompetens inom organisationen och även välja utifrån olika perspektiv på framtiden.

Viss mjukvara är utvecklad för att exekveras på ett specifikt operativsystem medan annan mjukvara i större grad är utvecklad för att kunna användas på ett större urval av operativsystem och möjliggör därmed att beslut om plattform fattas på andra grunder.

IBM Websphere Application Server är en applikationsserver för Java EE-arkitekturen. Den kan installeras och exekveras på en rad mer eller mindre kända plattformar där Linux och Windows Server 2003 är två av dessa[7].

Konsultbolaget Atea IM sysslar bland mycket annat med att installera, implementera och administrera produkten IBM Websphere Portal åt sina kunder. IBM Websphere Portal är en mjukvara som installeras på IBM Websphere Application Server för att integrera företagets information och applikationer i en portal.

1.1. Problemområde

Medarbetare på Atea IM som arbetar med IBM Websphere Portal och IBM Websphere Application Server upplever Windows Server 2003 som "inlåst" och menar att "mycket [system]resurser går åt till att driva alla extraprocesser som kommer med Windows"[5]. För att kunna administrera systemet i windowsmiljön tillfredsställande måste även ytterligare applikationer installeras på serverna. Samma medarbetare är av uppfattningen att mycket skulle vara vunnet, både vad gäller utvecklings- och administrationstid samt prestanda om man valde just Linux framför Windows Server 2003 som underliggande plattform för sina installationer av WebSphere Portal. En expert från Tyskland som ingick i IBM:s europeiska performance-team för WebSphere uttryckte under ett besök hos Atea IM när en kundimplementation av systemet lasttestades en förvåning över att man hade valt Windows Server 2003-plattformen framför Linux. Experten från IBM anses vara mycket kompetent och hans långa erfarenhet har visat att Linux ger en generellt bättre prestanda på samma hårdvara[5].

Inledande samtal^[4] med Atea IM pekar på att företagets säljare säljer Windows Server-lösningar av "gammal vana" och/eller att de upplever linuxlösningar som "svårare att sälja in". Det i kombination med att deras kunder "bara har windowsfolk på sina IT-avdelningar" leder till att de uteslutande säljer lösningar där Windows Server 2003 utgör grundplattformen för WebSphere Portal.

Det har också framkommit att det finns en kunskapsbrist kring vilka skillnader som finns mellan de båda plattformarna vilket också leder till att man som expert på området inte har möjlighet att upplysa sina kunder om detta. Det finns dock ett intresse av fördjupade kunskaper och den kommande uppsatsen syftar till att bidra med detta.

1.2. Syfte och frågeställning

Uppsatsen syftar till att genom en fallstudie undersöka eventuella skillnader vad gäller prestanda och administrerbarhet mellan att installera, implementera och administrera IBM Websphere Portal 6.1 på IBM Websphere Application Server 6.1 på Windows Server-plattformen och Linux-plattformen. Idag gör Atea IM detta uteslutande på Windows Server-plattformen men är intresserade av underlag för att kunna hjälpa sina kunder fatta bättre beslut kring val av underliggande plattform. Det finns idag en brist på kunskaper i det specifika fallet och uppsatsen kommer därför bidra med upplysande forskning.

Framförallt kommer rent tekniska konsekvenser av att välja Windows Server-plattformen eller Linux-plattformen studeras. Detta kommer genomföras genom en så kallad benchmark-undersökning. Utöver de tekniska konsekvenserna kommer skillnader i administrerbarhet mellan de olika plattformarna att jämföras.

Den huvudsakliga frågeställningen blir därmed:

Vilka skillnader kan påvisas mellan Linux och Windows Server 2003 som plattform för IBM WebSphere Portal och därmed också för IBM WebSphere Application Server med avseende på teknisk prestanda och administrerbarhet?

1.3. Avgränsningar

Studien avgränsades till att studera hur WebSphere Portal och nödvändig kring-mjukvara fungerar på en standardinstallation av respektive operativsystem i frågeställningen. De enda ytterligare inställningar för respektive operativsystem som gjordes var de som inkluderades i de officiella installationsmanualerna tillhandahållna av IBM. Utöver dessa inställningar finns det ytterligare en mängd optimeringar som kan göras både av de testade operativsystemen och WebSphere Portal samt WebSphere Application Server. Dessa optimeringar kräver noggranna analyser av WebSphere

Portals användningsområde i det specifika fallet. Det har av tidsmässiga skäl inte funnits några möjligheter att göra den typen av analyser och det studien tar upp är endast ett testfall.

1.4. Disposition

Kapitel 2 (Teori) syftar till att ge läsaren upplysning kring i första hand tekniska begrepp som rör studien. Detta kapitel tar upp de operativsystem, tekniker och mjukvaror som förekommer i studien. Här behandlas också begreppet *benchmarking*. I Kapitel 3 (Arbetsprocess och metod - förstudie) beskrivs arbetsprocessens gång, valet av forskningsansats samt hur datainsamlingen gått till i förstudien. Studien är upp delad i en förstudie och en huvudsaklig studie. Kapitel 4 (Resultat - förstudie) presenterar resultatet av den förstudie som gjordes. Detta görs i form av ett ramverk av faktorer som sedan användes i den huvudsakliga studien. I kapitel 5 (Arbetsprocess och metod - huvudsaklig studie) presenteras val av metod för den huvudsakliga studien samt hur arbetet praktiskt gick till med mjukvara för lasttest och prestandamätning. I detta kapitel beskrivs också de båda testade systemen. I Kapitel 6 (Resultat) beskrivs resultatet av den huvudsakliga studien. Kapitel 7 (Diskussion) diskuterar resultatet från den huvudsakliga studien samt presenterar förslag på fortsatta studier. Kapitel 8 (Slutsats) presenterar slutligen en slutsats för studien som helhet.

2. Teori

För att läsaren av uppsatsen bättre skall kunna förstå problemområdet som uppsatsen rör sig inom kräver det att vissa nyckelbegrepp förklaras.

2.1 Plattformer / Operativsystem

Här beskrivs de två huvudsakliga plattformar eller operativsystem som förekommer i uppsatsen och som kommer jämföras inom studiens kontext. Är man redan bekant med dessa kan avsnittet hoppas över.

2.1.1 Windows Server (2003)

Windows Server 2003 är när uppsatsen skrivs det näst senaste operativsystemet från Microsoft avsett för drift av servrar främst avsett för företag och organisationer. Windows Server är en proprietär programvara, det vill säga att licensen man förbinder sig att följa genom att använda produkten sätter restriktioner vad gäller att använda, modifiera och kopiera den.

En standardinstallation av Windows Server innehåller ett stort antal tjänster och processer som automatiskt startar när operativsystemet startas upp. Utöver dessa tjänster och processer startas även ett grafiskt gränssnitt som delvis är integrerat med den så kallade kärnan, det vill säga att man inte kan välja bort det.

Windows Server 2003 finns i ett flertal varianter^[3, 17], däribland Web Edition som är avsett för webbapplikationer, Standard Edition som är avsett för små och medelstora organisationer samt Enterprise Edition som är avsett för medelstora till stora organisationer.

2.1.2 Linux

Linux är i dagligt tal ett operativsystem som är baserat på den så kallade Linuxkärnan och som helt eller till största delen består av fri programvara. Utöver kärnan och de fria programvarorna behövs ett antal programvarubibliotek för att styra maskinvaran. Linux paketeras vanligen i så kallade distributioner, vilket innebär att man utöver kärnan och de fria programvarorna, som i regel alltid finns med, även inkluderar programvarubibliotek och annan programvara som är anpassade för specifika uppgifter som att styra en specifik hårdvara och ge en komplett användarupplevelse. Linux är alltså, i motsats till Windows Server, fri programvara med öppen källkod^[10]. Licensen man förbinder sig att följa genom att använda operativsystemet lägger inga begränsningar på användaren vilket innebär att man fritt kan använda, modifiera och kopiera det.

Man kan egentligen inte prata om en standardinstallation av Linux utan istället får man utgå från någon distribution. Det finns en mängd linuxdistributioner för olika ändamål, på ett vis liknande modellen hos Windows Server med olika varianter. En ändamålsenlig installation av godtycklig linuxdistribution innehåller endast de komponenter som är nödvändiga för ändamålet. Linux i sig självt är inte specialanpassat för serverdrift eller för någonting annat heller. Det är upp till varje distribution att anpassa Linux med de komponenter som passar ändamålet.

2.2 Applikationsservrar

En applikationsserver kan ses som kärnan i en 3- eller flerskiktad klient-server-arkitektur. Anrop tas emot av exempelvis en webbserver eller annan tjänst, behandlas av mjukvara på servern och returneras tillbaka. För att kunna behandla anrop och skicka tillbaka relevant information anropar applikationsservern i sin tur vanligen en eller flera databaser och/eller ett eller flera så kallade legacy-system, till exempel ekonomisystem och lönesystem, och behandlar därefter den inkommande datan.

Syftet med applikationsservern är att centralisera organisationens applikationer. Centraliseringen innebär många fördelar som att exempelvis förenkla underhåll och utveckling genom att man inte riskerar att gamla versioner av applikationer existerar som kan modifiera exempelvis data i en databas i ett icke-kompatibelt mönster. Datasäkerhet är också en viktig faktor som säkerställs genom att flytta ansvaret för autentisering från det potentiellt osäkra klientskiktet till applikationsservern som därmed utgör ett lager mellan klienten och databasen[1].

2.2.1 IBM Websphere Application Server

Websphere Application Server är en applikationsserver från IBM avsedd som värdsystem för applikationer skrivna för Java EE-standarderna.

2.3 Portaler

Portaler är i sammanhanget webbaserade system som abstraherar användaren från underliggande applikationer samt den komplexitet som ett stort antal datakällor inom en organisation innebär. Genom att knyta ihop information management, knowledge management samt applikationer kan portaler erbjuda en samlad och överskådlig vy för anställda, kunder och affärspartners[20].

Av definitionen kan två typer av portaler identifieras; externa och interna. Externa portaler riktar sig till kunder och partners medan interna används inom organisationen, ofta som bas för typiska intranät där de underlättar informationssökning och stödjer beslutsfattande samt underlättar samarbete.

2.3.1 IBM Websphere Portal

Websphere Portal är IBM:s portalprodukt avsedd för att exekveras på Websphere Application Server. Portalen kan skräddarsys efter olika roller, säkerhetsbehov och genom personliga inställningar anpassas för den enskilde medarbetaren, kunden eller partnern. I portalen kan arbetsflöden definieras för att stödja olika typer av processer. Websphere Portal kan också genom en single sign-on tjänst fungera som utgångspunkt för de underliggande applikationerna och skapar ett gemensamt gränssnitt.

2.4 Java Platform, Enterprise Edition (Java EE)

Java Platform, Enterprise Edition (Java EE) är en utvecklingsplattform från Sun Microsystems. Java EE är baserat på Java Platform, Standard Edition (Java SE) som används främst för skrivbordsapplikationer på persondatorer. Java EE innehåller ett utökat komponentbibliotek jämfört med Java SE och används för utveckling av webbaserade enterprise-applikationer och tjänster som exekveras på applikationsservrar.

Java EE erbjuder funktionalitet och tekniker för att utveckla flerskiktade, modulbaserade applikationer. Java EE-applikationer består typiskt av fyra skikt; klientskikt, webbskikt och affärsskikt som körs på applikationsservern samt dataskikt som körs på databasservern.

Klientskiktet består av klienten som anropar applikationsservern. Klienten är ofta en webbläsare men kan också vara en skrivbordsapplikation eller ett annat system.

Webbskiktet innehåller komponenter som ansvarar för kontakten mellan klientskikt och affärsskikt. Webbskiktet hanterar klientens tillstånd (state) och status, innehåll genereras dynamiskt till klienten och förfrågningar från klienten behandlas och skickas till affärsskiktet.

Affärsskiktet innehåller komponenter för affärslogik och ansvarar också för kommunikationen med dataskiktet. Affärslogiken kapslas in i Enterprise Java Beans (EJB). Kommunikationen mellan EJB:s hanteras av en EJB container som hanterar kommunikationen mellan EJB:s vilket innebär att utvecklaren kan koncentrera sig på att implementera affärslogiken.

Dataskiktet består vanligen av en extern databasserver.

2.5 Benchmarking - Att mäta prestanda

För att kunna utvärdera prestandan hos en mjukvara måste man utföra en mätning av dess prestanda. Denna studie kommer göra detta genom Benchmarkingundersökningar eller Benchmarks:

“In computing, a benchmark is the act of running a computer program, a set of programs, or other operations, in order to assess the relative performance of an object, normally by running a number of standard tests and trials against it.”[2]

Eller mera tydligt förklarat:

“Simply stated, benchmarking refers to running a set of programs on various computer and network configurations and measuring the results [...]. Specifically, a computer benchmark is a computer program that performs a strictly defined workload (i.e., a set of operations) and returns some form of measurable result (i.e., metric), describing how the system under test (SUT) performed.”[23]

Benchmarks kan delas in i två grupper: standardbenchmarks och applikationsspecifika[29]. Standardbenchmarks kan skilja två eller flera datorsystem åt genom att belasta dessa på samma sätt och avläsa hur de olika komponenterna i systemet som exempelvis processor och RAM-minne presterar. Den här typen av benchmarks kan vara missvisande om man använder resultat för att bedöma vilket system som lämpar sig bäst för en applikation som inte följer precis samma mönster som benchmarkapplikationen. Den andra benchmarktypen, applikationsspecifik, löser det problemet genom att mäta prestandan hos den specifika produkten. En mjukvara belastar inte varje hårdvarukomponent i systemet för sig utan är beroende av att en kombination av hårdvara fungerar optimalt. Denna benchmark kan göras på två eller flera datorsystem och man kan utifrån den insamlade datan avgöra hur produkten påverkas annorlunda av att exekveras på de olika systemen.

3. Arbetsprocess och metod - förstudie

Detta kapitel diskuterar och motiverar val av forskningsmetod för förstudien samt beskriver det praktiska genomförandet.

3.1 Val av metod

Studien har indelats i en förstudie och en huvudsaklig studie. Förstudien använder sig av en kvalitativ metod för att upprätta ett ramverk för den huvudsakliga studien. Till den huvudsakliga studien har en kvantitativ metod valts. Kombinationen av kvalitativ och kvantitativ metod har valts för att göra studien så heltäckande som möjligt.

3.2 Förstudie genom kvalitativ metod

Förstudien syftade till att klarlägga de faktorer som ansågs vara av intresse för den huvudsakliga studien. Dessa faktorer, tekniska och ur administrativ synvinkel, rör alla prestanda i någon form. Författaren var själv ingen expert på varken WebSphere Application Server, WebSphere Portal eller hur arbetet med dessa båda produkter på Windows Server 2003 och Linux fungerade i praktiken. För att stärka validiteten i förstudien så mycket som möjligt inom den begränsade tidsramen för dess genomförande användes triangulering. Triangulering kombinerar flera olika datainsamlingsmetoder. Informationen från dessa vägdes sedan samman i analysen för att ge en så komplett bild som möjligt utifrån dess förutsättningar[27].

Förstudien kombinerade författarens förstahandserfarenhet med produkterna och operativsystemen, intervjuer med två personer på Atea IM som arbetar med produkterna samt en litteraturstudie.

3.2.1 Förstahandserfarenhet och utbildning

Arbetet med förstudien inleddes med att författaren bekantade sig med WebSphere Application Server och WebSphere Portal på plats hos Atea IM. Med hjälp av IBM:s anvisningar och dokumentation samt instruktioner av handledare på Atea IM utfördes installationer, uppgraderingar, och konfigurationer av WebSphere Portal inklusive WebSphere Application Server. Tillhörande mjukvara som databasmotorn DB2 installerades och konfigurerades också. Detta arbete utfördes både på Windows Server 2003 samt Ubuntu Linux Server Edition för att få en grundläggande känsla för arbetsprocessens gång på de två skilda operativsystem som kom att ingå i studien och för att få en grundläggande produktkännedom.

3.2.2 Intervjuer

Kvalitativa intervjuer genomfördes med två personer på Atea IM som arbetar med WebSphere-plattformen. Anledningen till att intervjuerna endast blev två till antalet hade främst att göra med att det endast är tre personer på avdelningen som arbetar med WebSphere och att det endast var två av dessa som fanns tillgängliga under tiden för förstudien. Förstudien hade troligen kunnat berikats med intervjuer med kunder till Atea IM som köpt produkterna i fråga, men tyvärr gavs inte möjlighet till sådana intervjuer.

Intervjuerna genomfördes i mitten av April 2010 i Ateas lokaler i Sisjön i Göteborg. Under planeringen av intervjuerna var dessa beräknade att ta maximalt 30 minuter var i anspråk. Intervjupersonerna visade sig dock välvilliga att utveckla sina svar vilket resulterade i att intervjuerna istället tog mellan 45 och 60 minuter att genomföra.

Intervjupersonerna upplystes inför intervjuerna om syftet med den huvudsakliga studien och att deras svar skulle komma att behandlas anonymt och vägas in i en förstudie som utöver intervjuerna skulle innefatta litteraturstudier och förstahandserfarenhet. Intervjupersonerna garanterades anonymitet i studien. Intervjuerna spelades med intervjupersonernas medgivande in med hjälp av ett ljudinspelningsprogram på intervjuarens dator. Intervjuerna transskriberades så snabbt som möjligt efter intervjutillfällena.

Inför intervjuerna upprättades en intervjumall med fasta frågor (se Bilaga A). Intervjumallen tjänade under intervjuerna som en utgångspunkt för diskussion och frångicks vid flera tillfällen då följdfrågor ställdes. På förhand var det bestämt att intervjuerna skulle ha en låg grad av standardisering. Låg grad av standardisering innebär att det inte finns några fasta svarsalternativ utan intervjupersonen ges gott om utrymme att själv formulera svaren med egna ord. "Syftet med en kvalitativ intervju är att upptäcka och identifiera egenskaper och beskaffenheten hos något, t.ex. den intervjuades /./ uppfattningar om någon fenomen"[27]. Intervjuerna följde i stort den upprättade mallen och kan därmed anses ha haft en relativt hög grad av strukturering.

Det är en fördel att som intervjuare i förväg skaffa sig kunskaper om området intervjun rör[27]. Detta gjordes inför intervjuerna främst genom förstahandserfarenhet av de produkter och system som frågorna rörde samt läsning av dokumentation. Det är inte nödvändigt att som intervjuare ha en teoretisk utgångspunkt inför intervjuerna då syftet med dessa är induktivt[27]. Ett induktivt förhållningssätt syftar till att upptäcka någonting nytt[27].

3.2.2.1 Beskrivning av intervjupersoner

PERSON 1

Intervjuperson 1 var man och IT-arkitekt på Atea IM. För att förtydliga beskrev han sig som integrationsarkitekt vilket innebar att han både tar fram och designar lösningar åt kunder inom WebSphere. Han menade också att eftersom de inte var så många som jobbade med WebSphere hos Atea IM i Göteborg så var han delaktig i hela kedjan från att ta fram och designa nya lösningar till att installera och administrera. Intervjuperson 1 har jobbat på Atea IM sedan 2007 dit han kom efter avslutat 4-årigt Systemvetenskapligt program på Göteborgs Universitet. Han har inte haft några tidigare IT-relaterade anställningar.

PERSON 2

Intervjuperson 2 var man och WebSphere-administratör på Atea IM. Utöver WebSphere arbetar han också med administration av Lotus Domino och även en del utveckling för den plattformen. Intervjuperson 2 har jobbat på Atea IM sedan 2008. Han har jobbat i IT-branschen sen 2000, främst med Domino-utveckling.

3.2.3 Litteraturstudier

För att öka validiteten i det ramverk som förstudien resulterade i genomfördes också litteraturstudier. Artiklar om hur man mäter mjukvaruprestanda, så kallad benchmarking, har varit särskilt intressanta, liksom dokumentation om hur man justerar WebSphere Portal och WebSphere Application Server. Den senare dokumentation har varit nyttig inte så mycket för kunskaperna om hur man justerar WebSphere för bättre prestanda som för att den klargör vilka komponenter hos värdsystemet och hos operativsystemet som är kritiska för en god prestanda. Vidare har artiklar om prestandamätning av WebSphere Portal och WebSphere Application Server varit till god hjälp samt även artiklar som jämför Linux och Windows ur olika perspektiv.

4. Resultat - förstudie

Detta kapitel presenterar resultatet av den förstudie som genomfördes i syfte att upprätta ett ramverk med faktorer till den huvudsakliga undersökningen. Faktorerna har delats in i två huvudgrupper - Tekniska prestandafaktorer och Administrativa prestandafaktorer. Kapitlet presenterar en redogörelse över dessa samt beskriver hur och varför de valts ut.

4.1 Val av kvalitetsattribut

Mjukvarukvalitet kan definieras som till vilken grad en mjukvara följer ett antal önskade kvalitetskrav eller attribut. Kvalitetskrav som en mjukvara skall uppfylla kan delas in i två huvudgrupper baserat på vilken typ av kvalitet de efterfrågar[26]:

1. Kvalitetskrav baserade på en utvecklarens perspektiv (development quality requirement) som exempelvis administrerbarhet (maintainability) och begriplighet (understandability).
2. Kvalitetskrav baserade på mjukvarans funktionsduglighet eller användarens perspektiv (operational quality requirements) som exempelvis prestanda (performance) och användbarhet (usability).

Kvalitetskrav sätts upp av beställaren medan kvalitetsattribut kan definieras som det som mjukvaran faktiskt levererar.

Två olika kvalitetsattribut har valts ut för studien. Prestanda och Administrerbarhet. IEEE standard 610.12-1990[18] definierar dessa båda som:

Prestanda: "The degree to which a system or component accomplishes its designated functions within given constraints, such as speed, accuracy, or memory usage."

Administrerbarhet: "The ease with which a software system or component can be modified to correct faults, improve performance or other attributes, or adapt to a changed environment."

Dessa båda har valts ut därför att de speglar en stor del av det perspektiv en arkitekt, utvecklare och administratör på ett företag som Atea IM har på exempelvis WebSphere Portal samt det speglar användarens och kundens upplevelse av den levererade produkten.

4.2 Prestanda - tekniska prestandafaktorer

Alla datorsystem har en begränsad prestanda. Prestandan begränsas exempelvis av hastigheten med vilken processorn kan utföra beräkningar, storleken samt hastigheten hos RAM-minnet där den data som behövs för stunden lagras och söktiden samt hastigheten på hårddiskarna där data långtidslagras.

Med tekniska prestandafaktorer avses faktorer som direkt påverkar prestandan hos WebSphere Portal, eller mera korrekt, den prestanda slutanvändaren av WebSphere Portal upplever. Båda respondenterna framhöll att det är slutanvändarens upplevelse av prestanda, i slutändan att tidsåtgången för att utföra en åtgärd, som är den viktigaste faktorn:

“Ja, det viktigaste är väl naturligtvis hur snabbt användaren får svar i sin applikation. Eftersom det är webbaserade applikationer så är det ju hur snabbt applikationen svarar.”

“Naturligtvis tycker jag väl att om användaren upplever att det är långsamt så är det långsamt och då är det mycket som kan ligga bakom, men användarupplevelsen är det viktiga.”

“Att det funkar, att man inte behöver en lång väntetid för att komma in i systemet. Om det är ett webbsystem, skriv in URL:en och ska inte behöva vänta mer än en sekund eller nåt innan man får upp en förstasida.”

Bakom att en applikation ger användaren ett snabbt svar ligger flera faktorer. WebSphere Portal kan i det här fallet förenklat ses som en webbaserad databasdriven applikation som exekveras på WebSphere Application Server som i sin tur exekveras i en JVM (Java Virtual Machine) på operativsystemet. Mellan klienten (slutanvändarens webbläsare) och WebSphere Portal finns en webbserver som kan ses som en del av systemet. Webbservern agerar mellanhand och levererar exempelvis statisk information som bilder, javascript-filer och stilmallar direkt ifrån hårddisken och skickar övriga förfrågningar till WebSphere Portal. Till detta kommer den data som applikationen behandlar och som i regel finns i en databas.

Det finns med andra ord en rad komponenter som är inblandade när en slutanvändare anropar WebSphere Portal.

En av respondenterna är av den uppfattningen att databasens prestanda påverkar slutanvändarens upplevelse:

“[...] naturligtvis hur snabbt databasen kan svara påverkar hur snabbt användaren får sitt svar.”

Att databasen är en viktig del av WebSphere Portal bekräftas av IBM:s *WebSphere Portal 6.1.X Performance Tuning Guide*[19]:

“[...] WebSphere Portal V6.1 uses database servers for core functionality.”

Ett snabbt svar från databasen är enligt samma respondent beroende av mängden tillgängligt RAM-minne:

“Databasen är väl en av de största felkällorna för det sker så mycket databasanrop hela tiden. Det här är väl också en stor grej då. Vi lever ju i 32-bitarsmiljöer så 4 GB minne är ju max. Det är nästan alltid det som gör att saker bli slöa, att vi inte kan allokeras mer minne. På databassidan då i alla fall.”

Att tillgängligt RAM-minne är viktigt i databassammanhang bekräftas i en artikel ur IBM Database Magazine[25]:

“Understanding DB2 [standarddatabas för WebSphere] memory usage is even more critical in a 32-bit environment, in which processes are restricted to 4GB of virtual address space.”

Det står klart att databasens prestanda påverkas av RAM-minnet. Båda respondenterna ser i slutändan databasen och därmed RAM-minne som den allra viktigaste begränsande faktorn. Även den andra respondenten uppfattar på en direkt fråga RAM-minne som den viktigaste faktorn bakom god responstid:

“Ja det är ju RAM-minne i slutet egentligen. Att du har tillräckligt med RAM-minne för att utföra, att du har ledigt minne för att processa frågor till systemet.”

RAM-minne är därför en teknisk prestandafaktor som skall mätas i den huvudsakliga undersökningen.

Utöver RAM-minnet är det endast två ytterligare faktorer som nämnts av båda respondenterna och som kan härledas till de tekniska prestandafaktorer. Dessa båda är processorn och hårddisken.

När intervjuaren pratar utifrån egna erfarenheter av applikationer som använder processorn intensivt tillerkänner en av respondenterna processorn en viss betydelse, framförallt hos webbservern som ju är inblandad mellan klienten och WebSphere Portal, när intervjuaren pratar utifrån egna tidigare erfarenheter:

“[...] det skulle nog kunna vara i webbservern som sådana problem skulle kunna uppstå. Där kan den [processorn] jobba på rätt rejält om man har mycket last, men Apache är ganska bra. Det är framför allt där jag har märkt av det - det är i webbserversammanhang.”

Den andra respondenten menar också att processorn kan vara en begränsande faktor:

“Det är väl ett samspel mellan processor och minne.”

Denna respondent menar dock att processorn är av underordnad betydelse jämfört med RAM-minnet:

“Säg att du har ett system med 2 GB RAM och en 2.0 GHz processor, det första steget är inte att byta processor utan det är att öka RAM-minnet om man har dålig prestanda.”

De båda respondenterna menar alltså att processorn inte har samma begränsande betydelse som RAM-minnet, men litteraturen verkar säga annorlunda. Här citeras IBM WebSphere Portal Performance Troubleshooting Guide[28]:

“In most cases, the processor load at the node(s) running WebSphere Portal will be the limiting factor in the system.”

Eftersom det finns motsägelsefulla uppgifter anser jag att det finns grund att i den huvudsakliga undersökningen ta med processoranvändning som en teknisk prestandafaktor.

När det kommer till hårddisken så är de båda respondenterna av lite olika uppfattning. Den ena respondenten svarar så här på intervjuarens fråga om operativsystemets hantering av förfrågningar till hårddisken kan ha någon betydelse för den övergripande prestandan:

“Ja det tror jag definitivt. Både på webbserversidan och på databassidan tror jag det kan vara jättestor skillnad [mellan olika operativsystem].”

Den andra respondenten verkar mena att hårddiskar generellt hamnar långt ner på skalan över begränsande faktorer och att de faller in under “generell översikt” av servern:

“Har man fortfarande dålig prestanda så kanske man får göra en generell översikt; vad kan vi göra för att förbättra servern totalt? Då kanske man behöver ha snabbare [hård] diskar om det sker kommunikation med hårddiskar. Då kanske man behöver snabbare sådana.”

Denna respondent menar dock att det kan finnas fall där hårddisken är en begränsande faktor och därför anses att det finns grund att i den huvudsakliga undersökningen ta med hårddiskaktivitet som en teknisk prestandafaktor.

4.3 Administrerbarhet - administrativa prestandafaktorer

WebSphere Portal administreras huvudsakligen på två sätt, dels genom ett webbgränssnitt och dels via en stor uppsättning konfigurationsfiler på serverns hårddisk. Webbgränssnittet är med tanke på WebSphere Portals natur som webbapplikation likadant oberoende av plattform vilket gör det ointressant att ta upp i studien.

Att konfigurera, administrera och installera WebSphere Portal kan skilja sig åt beroende på vilket operativsystem man väljer. Operativsystemet självt kräver givetvis också ett visst underhåll som skiljer sig åt. Båda respondenterna menar med anledning av det att val av operativsystem bör göras utifrån kompetens hos kunden:

“Först och främst får man kolla vem som skall administrera systemet. Är dom pro det ena eller pro det andra för man väga in det väldigt starkt i bedömningen.”

“Ja hos kunden. Vi är så illa tvugna att göra vad de vill i slutändan. Men, framförallt ska man ju kolla vad dom kan och vad dom känner.”

“Driftpersonal är väl den viktigaste frågan egentligen tycker jag, om man skall se ur kundens synvinkel.”

En respondent menar också att kompetensen hos Atea IM kan vara en avgörande faktor:

“Sen också kolla kompetensen hos oss för att installera det. Har vi den kompetensen som krävs för de olika så är det inte en faktor man behöver väga in men har vi inte det får man kolla vad vi klarar av.”

På frågan om skillnader i administrerbarhet mellan Windows Server 2003 och Linux har de båda respondenterna olika utgångspunkter. En av respondenterna kan beskrivas som mycket linuxkompetent medan den andra endast har ringa erfarenheter av Linux. Båda respondenterna har god kompetens när det kommer till Windows Server-administration. En respondent svarar på frågan om allmänna skillnader i prestanda mellan installationer av Windows Server 2003 och Linux:

“Jag tycker ju att, ur administrationssynvinkel, är det smidigare med Linux för att dom har helt enkelt bättre applikationer. Det är ett system som är mer utformat för att kunna

köras som server tycker jag. Det är ju mycket smidigare att kunna ssh:a och använda touch och vim och tail.”

Samma respondent identifierar en av de huvudsakliga skillnaderna en administratör upplever mellan Windows Server 2003 och Linux. Windows Server 2003 är starkt knutet till det grafiska användargränssnittet och kräver att man använder sig av ett fjärrskrivbord för administration. Linux har också den möjligheten, men inget krav finns att använda sig av ett grafiskt användargränssnitt. Istället kan man ansluta till servern genom ett textbaserat gränssnitt:

*“Ja det är ju en j*vla skillnad. Man måste ju använda Remote Desktop [i Windows Server 2003] och hålla på och klicka runt. För att göra de enklaste grejerna så måste man ändå... ja det är ju väldigt omständigt och bökigt. Det är i och för sig förmodligen enklare och mindre risk att göra fel i Windows. Det finns en viss sak för allting. Men det är ju å andra sidan j*vligt krångligt och tidsödande att göra det.”*

Respondenten säger också:

“Kopiera filer, starta applikationer, ja allting är ju enklare i Linux, men det är kanske en vanesak också men det tar ju längre tid att gå igenom det grafiska gränssnittet.”

Respondenten med ringa erfarenhet av Linux säger:

“Ja det är Remote Desktop som gäller, fast det är lite skakigt ibland.”

“[...] jag har inte så mycket erfarenhet av att administrera Linux tyvärr. Windows är väl inte så jättelätt att administrera alltid heller. Den gömmer grejer lite här och där och man får googla mycket för att hitta grejerna på Windows.”

Av svaren ovan kan konstateras att det finns stora skillnader i administrationen av de båda operativsystemen. Under intervjuerna upplevdes det också också att respondenterna hade subjektiva åsikter om de båda operativsystemen. Den ena respondenten var pro Linux medan den andra inte hade någon erfarenhet av detta operativsystem och därför föredrog att administrera ett windowssystem.

Litteraturen påvisar att ämnet administrerbarhet är viktigt, inte minst när det kommer till kostnader. Kostnaden som läggs på att administrera ett datorsystem är flera gånger högre än den sammanlagda kostnaden för hårdvara och mjukvara[22]. Det sägs att det är mänskligt att fela och handhavandefel orsakar hälften av alla avbrott i IT-system[21]. Dessa avbrott innebär sedan stora kostnader för uteblivna intäkter och att återställa[24].

Områdets subjektiva natur, den begränsade mängden respondenter och den begränsade tid som finns för undersökningens genomförande möjliggör tyvärr inte att gå vidare med en undersökning kring skillnader i administrerbarhet. Utöver det som presenterats i förstudien kommer detta område inte ytterligare behandlas i uppsatsen mer än i diskussionen.

4.4 Sammanfattning av förstudie

Förstudien har resulterat i ett ramverk av faktorer som kommer bedömas i den huvudsakliga studien. Eftersom förstudien visat att administrativa prestandafaktorer av olika anledningar inte kan bedömas inom ramen för denna undersökning kommer den huvudsakliga studien att helt rikta in sig på tekniska prestandafaktorer. De tre tekniska prestandafaktorer som kommer utvärderas på respektive Windows Server 2003 och Linux är utnyttjande av RAM-minne och processor samt hårddiskaktivitet.

Utöver de tre identifierade tekniska prestandafaktorerna kommer även en helhetsbedömning göras utifrån de mätdata som levereras av lasttestapplikationen som kommer nyttjas för att lastsätta systemen från en klientdator. Responstider från systemet är ett viktigt mått för att bedöma användarens upplevelse av systemet och kommer därför mätas.

5. Arbetsprocess och metod - huvudsaklig studie

Detta kapitel beskriver det praktiska genomförandet av den huvudsakliga studien och redogör för val och beslut som fattades samt metoden som användes för att analysera den insamlade datan.

5.1 Metod

Frågeställningen handlar i huvudsak om faktorer som är mätbara. Tidsåtgång och systemresurser är de viktigaste faktorerna. Att samla in kvantitativ information handlar om att göra en mätning vilket i det här fallet innebär att numeriska värden på ett entydigt sätt tilldelas det som studeras. En egenskap som studieobjektet har betraktas som en mängd varpå man applicerar en skala och läser av mängden som ett siffervärde. Den studerade egenskapens värde betraktas som en sanning som inte varierar slumpmässigt[27]. Viktigt i en kvantitativ undersökning är att veta att man undersöker det man avser att undersöka, det vill säga att undersökning har god validitet. Förutom god validitet måste undersökningen ske på ett tillförlitligt sätt, den måste alltså ha god reliabilitet[27].

Den huvudsakliga undersökningen bestod av en så kallad benchmark-undersökning. Benchmarking-undersökningen som gjordes resulterade i en mängd data i sifferform och är därmed en bra grund för en kvantitativ undersökning.

5.2 Beskrivning av testfall

I detta avsnitt kommer det specifika testfallet att beskrivas med en genomgång av relevanta parametrar.

5.2.1 Maskinvaruspecifikation

Windows Server 2003 och Linux installerades på identisk maskinvara. Maskinvaran var virtuella servrar på en VMware ESX 4-lösning som hade 40 GB RAM-minne och 4 processorer om vardera 3,66GHz till sitt förfogande. Varje virtuell server som användes som testmiljö (2 stycken) tilldelades följande resurser:

- 1 x processor á 3,6 GHz
- 4 GB RAM-minne
- 15 GB hårddiskutrymme

Noterbart är att de två testsystemen inte delades upp på flera servrar vilket är brukligt i produktion. Anledningen till detta var att det skulle bli enklare att få en helhetsbild av den tekniska prestandan med hela systemet på samma virtuella hårdvara.

5.2.2 Operativsystem

Windows Server 2003 valdes på ett tidigt stadium i undersökningen då det är den senaste versionen av Windows Server som stöds av IBM för WebSphere Portal[8]. Enterprise Edition av Windows Server 2003 användes. Som linuxdistribution valdes Ubuntu 8.04 Server LTS. Ubuntu är inte en linuxdistribution som stöds av IBM för WebSphere Portal. Anledningen till att Ubuntu valdes var en kombination av undersökarens tidigare erfarenheter samt tidsbrist då det uppstod problem vid installationerna av både RedHat och SUSE i den virtuella miljön. Att en icke-stödd linuxdistribution valdes för studien är inte optimalt då många faktorer hos distributionen utöver kärnan, som är densamma oavsett distribution[9], kan spela en roll i prestandasammanhang[12].

En standardinstallation utfördes av respektive operativsystem. De ytterligare inställningar för respektive operativsystem som gjordes var de som inkluderades i de officiella installationsmanualerna tillhandahållna av IBM. Utöver dessa inställningar finns det ytterligare en mängd optimeringar som kan göras både av de testade operativsystemen och WebSphere Portal samt WebSphere Application Server. Dessa optimeringar kräver noggranna analyser av WebSphere Portals användningsområde i det specifika fallet. Det har av tidsmässiga skäl inte funnits några möjligheter att göra den typen av analyser och det studien tar upp är endast ett begränsat testfall.

5.2.3 Mjukvaruprodukter

Identiska versioner av WebSphere Application Server (v.6.1.0.19), WebSphere Portal (v.6.1.0.1) och DB2 (v.9.1.0.4) har installerats på de båda operativsystemen. IBM:s installationsanvisningar *Setting up a stand-alone production server*[13, 14] har följts för respektive operativsystem. I installationen av WebSphere Portal valdes även att installera IBM Lotus Web Content Management (v6.1.0.1). Web Content Management installerar ett exempel på en extern webbtjänst samt ett exempel på en intranättjänst. Dessa användes som mål för testerna som senare utfördes. Ingenting ytterligare har installerats.

5.2.4 Lasttestserver och mjukvara

Lasttestservern installerades på samma VMWare ESX-server som de två testserverna. Operativsystem för lasttestservern valdes utifrån mjukvaran som skulle användas för testerna. Mjukvaran var avsett för Windows och därför valdes Windows Server 2003 som operativsystem för servern. Efter utprovning av virtuell maskinvara fastställdes att 1 processor á 3,6 Ghz, 4 GB RAM-minne samt 10 GB hårddisk var tillräckligt för att kunna belasta båda testsystemen samtidigt utan

att det egna systemet skulle sätta begränsningar. Lasttestservern befann sig på samma lokala nätverk vilket minimerade risken att flaskhalsar i nätverket kunde påverka resultatet.

Mjukvaran som användes var OpenSTA. OpenSTA används för att utföra lasttester av webbapplikationer och andra applikationer som utnyttjar HTTP- och HTTPS-protokollen. Denna mjukvara valdes därför att Atea IM själva använder den och har goda erfarenheter. OpenSTA kan genom en proxyserver "spela in" hur användaren navigerar genom en webbapplikation. Utifrån de inspelade kommandona genererar programmet kod som sedan kan modifieras av användaren. Den genererade koden är ett manus som används för att testa webbapplikationen i fråga. Under ett test väljer man parametrar som antal simulerade användare, hur lång tid de spenderar på varje enskild sida och hur många gånger manuset skall upprepas för respektive simulerad användare. Med korrekt konfigurerad hårdvara och med tillräcklig bandbredd kan OpenSTA simulera ett tusental samtidiga användare från samma maskin.[11]

5.3 Mätning av tekniska prestandafaktorer

Mätning av prestanda skedde dels med hjälp av OpenSTA på lasttestservern och dels på ESX-servern. OpenSTA levererar efter avslutat test en mängd rapporter som kan användas för analys. I denna undersökning användes OpenSTA:s rapportfunktion för att mäta responstider under last från de testade systemen. På ESX-servern användes kommandot *esxtop* i batch-läge under testen. *Esxtop* kan användas för att samla in data för alla typer av systemresurser som processor, minne, hårddisk och nätverk.[16] Under testen genererades loggfiler av all aktivitet hos de två testsystemen. Ur dessa filer extraherades sedan datan som specifikt rörde de tre tekniska prestandafaktorerna som identifierades i förstudien: processoranvändning, minnesanvändning samt hårddiskaktivitet. Metoden att mäta den tekniska prestandan direkt på ESX-servern istället för på respektive operativsystem valdes därför att det eliminerar eventuella skillnader i mätresultatet som kan uppstå eftersom de två testsystemen bygger på skilda teknologier. Nackdelen med att inte mäta direkt på testsystemen var att ingen enskild data för olika komponenter som JVM, DB2 och systemet självt kunde mätas. Detta ansågs godtagbart då studiens syfte var att utvärdera de två systemens prestanda i sin helhet.

6. Resultat

Detta kapitel beskriver de tester som genomfördes mot de två testsystemen samt redogör för resultatet av dessa.

Tre olika test genomfördes. Inför dessa test genererades i OpenSTA identiska manus för de båda testsystemen där en simulerad användare navigerar runt under ca 3 minuter på demotjänsterna som installerades. Samma manus användes i alla samtliga test som beskrivs nedan.

6.1 Test 1 - Stegring 1-60 klienter under 25 minuter

I detta test konfigurerades OpenSTA att under 25 minuter simulera 1 till 60 samtidiga klienter på de båda testsystemen. Testet startade med en simulerad klient och var 20:e sekund lades en ny klient till. De sista fem minuterna belastades systemet således av samtliga 60 klienter. Varje simulerad klient utförde sitt manus och väntade sedan 20 sekunder innan manuset påbörjades igen.

6.1.1 Processoranvändning

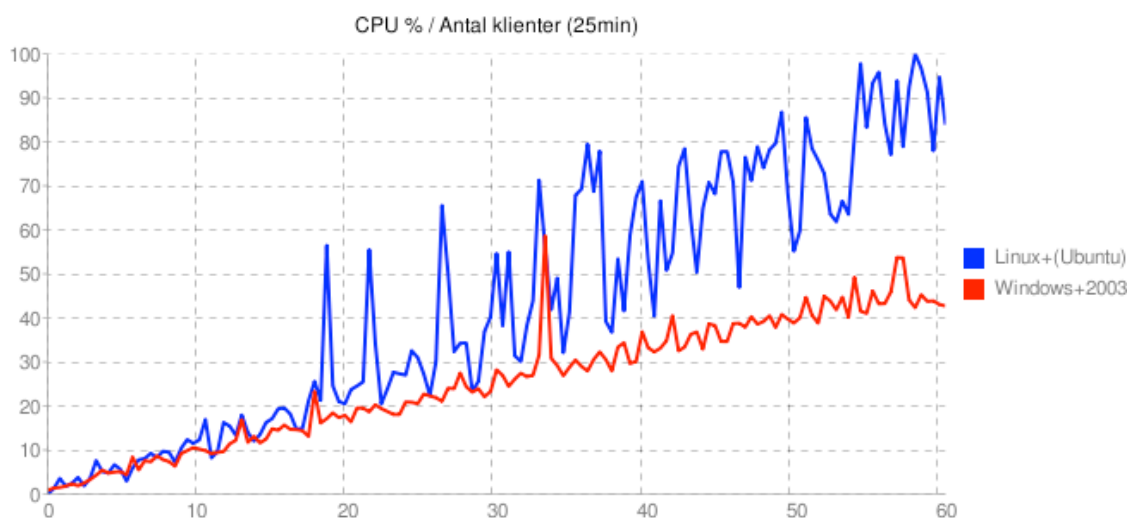


Fig 1. Processoranvändning - 1-60 klienter under 25 minuter

Fig. 1 visar den uppmätta processoranvändningen hos vardera testsystem i procent över tiden 25 minuter. Esxtop har här konfigurerats att läsa av processorutnyttjandet var 15:e sekund under testet. Med ett fåtal klienter (upp till ca 20 stycken) kunde det konstateras att de båda testsystemen uppvisade likvärdiga siffror. Mellan 20 och 60 klienter använde linuxsystemet generellt betydligt mer av de tillgängliga processorresurserna, många gånger mer än dubbelt så mycket som windowssystemet. Mot slutet av testet använde linuxsystemet större delen av de tillgängliga resurserna vid flera av avläsningarna medan windowssystemet stabiliserade sig mellan 40 och 50%.

Noterbart är också att processorutnyttjandet hos linuxsystemet varierade mycket kraftigt mellan många av avläsningarna, något som endast kunde påvisas i enstaka fall hos windowssystemet.

6.1.2 Svarstider

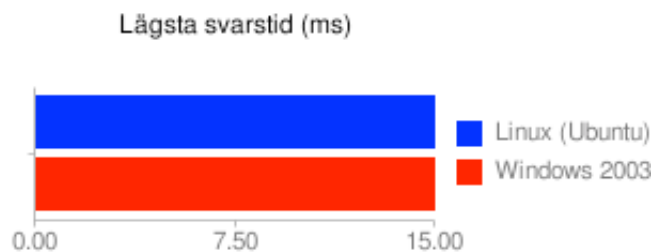


Fig 2. Lägsta svarstid - 1-60 klienter under 25 minuter

Båda systemen visade samma uppmätta lägsta svarstid – 15 ms. Att denna siffra var exakt 15 ms för båda systemen är troligen en begränsning hos OpenSTA. OpenSTA visade 0 ms som svarstid vid ett flertal tillfällen men inga svarstider mellan 0 och 15 ms. Svarstider på 0 ms är inte rimligt utan troligare är att båda systemen vid flera tillfällen uppvisade svarstider under 15 ms men att dessa var för låga för att OpenSTA kunde hantera detta. Samma mönster uppvisades i samtliga tester. Ingen dokumentation kring detta fenomen kunde hittas men eftersom OpenSTA även rapporterar att dessa anrop lyckades (statuskod 200[6]) utgår undersökningen ifrån att 0 ms inte innebär timeout. Vad gäller genomsnittliga svarstider så har dessa beräknats genom att ersätta 0 ms med 7,5 ms i rapporten från OpenSTA.

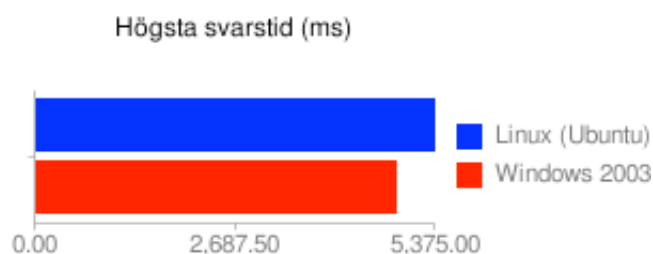


Fig 3. Högsta svarstid - 1-60 klienter under 25 minuter

Båda systemen uppvisade mycket höga högsta svarstider. Linuxsystemets högsta uppmätta svarstid var 5375 ms medan windowssystemet låg omkring en halv sekund lägre på 4859 ms.

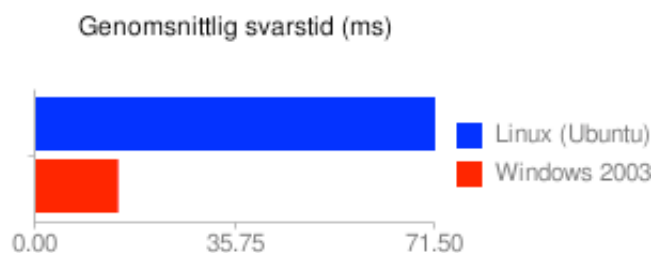


Fig 4. Genomsnittlig svarstid - 1-60 klienter under 25 minuter

Linuxsystemet levererade i genomsnitt betydligt högre svarstider än vad windowssystemet gjorde. Den genomsnittliga svarstiden för linuxsystemet var 71,5 ms vilket är knappt 5 gånger högre än windowssystemets 14.8 ms.

6.2 Test 2 - Stegring 1-100 klienter under 5 minuter

I detta test konfigurerades OpenSTA att under 5 minuter simulera 1 till 100 samtidiga klienter på de båda testsystemen. Testet startade med en simulerad klient och varannan sekund startade en ny klient. Detta innebär att efter 200 sekunder belastades testsystemen med vardera 100 simulerade klienter. Varje klient utförde sitt manus och väntade sedan 20 sekunder innan det på nytt påbörjades.

6.2.1 Processoranvändning

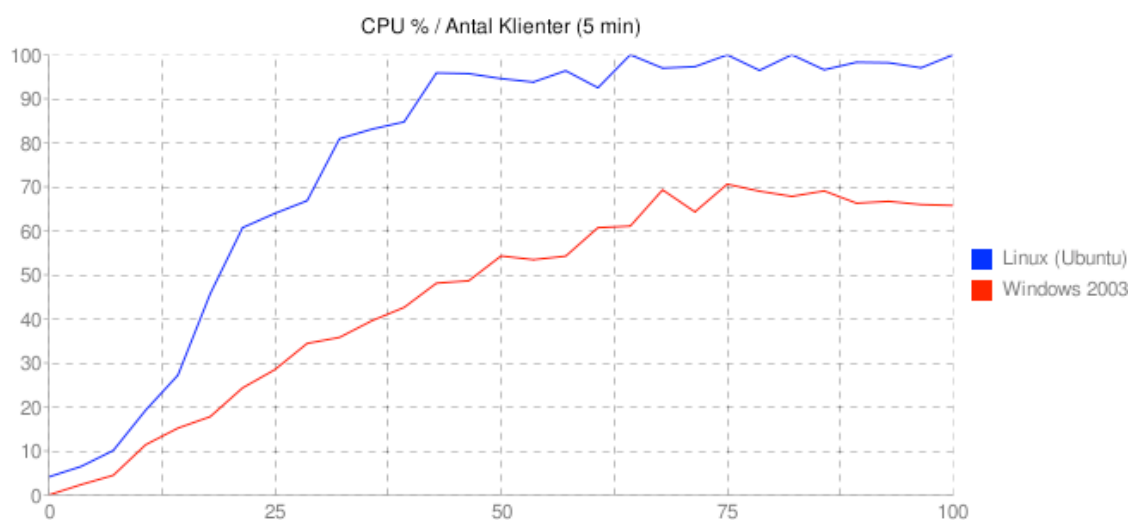


Fig 5. Processoranvändning - 1-100 under 5 minuter

Fig. 5 visar det uppmätta processorutnyttjandet hos de båda testsystemen över 5 minuter. Linuxsystemet ligger generellt på högre värden än windowssystemet. Vid ungefär 60 samtidiga klienter når linuxsystemet 100% processoranvändning samtidigt som windowssystemet når 70%. Windowssystemet stabiliserar sig sedan runt 70% resten av testet. Linuxsystemet låg samtidigt stadigt runt 100%.

6.2.2 Svarstider

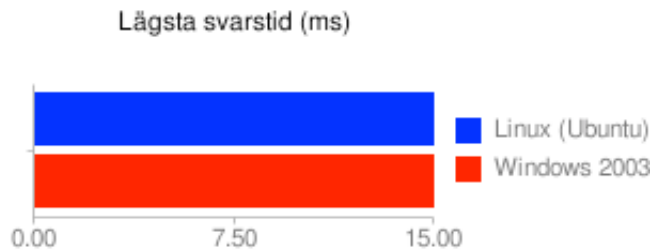


Fig 6. Lägsta svarstid - 1-100 klienter under 5 minuter

Lägsta svarstid som uppmättes var 15 ms hos båda systemen. Anledningen till detta resultat förklaras i test 1.

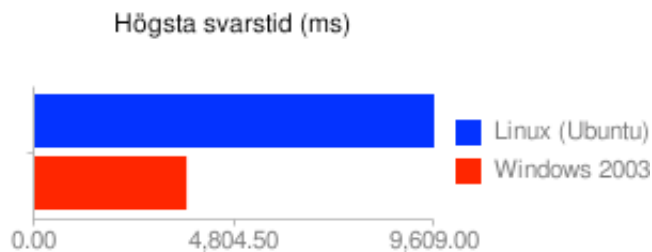


Fig 7. Högsta svarstid - 1-60 klienter under 25 minuter

Linuxsystemet visade på den högsta svarstiden – 9609 ms, alltså nästan 10 sekunder. Windowssystemets högsta svarstid var under mätningen 3656 ms.

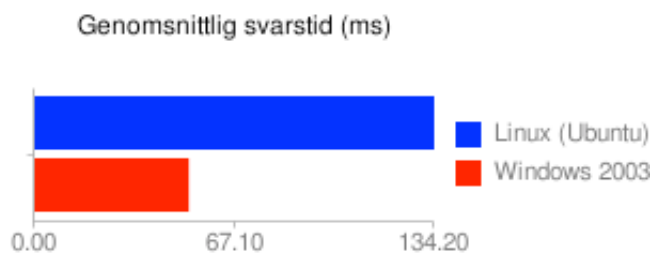


Fig 8. Genomsnittlig svarstid - 1-100 klienter under 5 minuter

Den genomsnittliga svarstiden i test 2 var hos linuxsystemet 134.2 ms och hos windowssystemet 51.8 ms. Linuxsystemet uppvisar i genomsnitt alltså 2,5 gånger högre svarstider än windowssystemet i detta test.

6.3 Test 3 - 30 samtidiga klienter under 5 minuter

I detta test konfigurerades OpenSTA att belasta de båda testsystemen med 30 samtidiga klienter vardera. 30 klienter valdes utifrån resultatet av test 1 där en generell ökning i processorutnyttjandet kunde visas efter 20 samtidiga klienter. 30 klienter valdes för att kontrollera om

processorutnyttjandet följde mönstret i de tidigare testen att linuxsystemet utnyttjar processorna i högre grad än windowssystemet vid ett högre antal simulerade klienter samt för att kontrollera om processorutnyttjandet ökade eller minskade över tid vid ett konstant antal klienter.

6.3.1 Processoranvändning

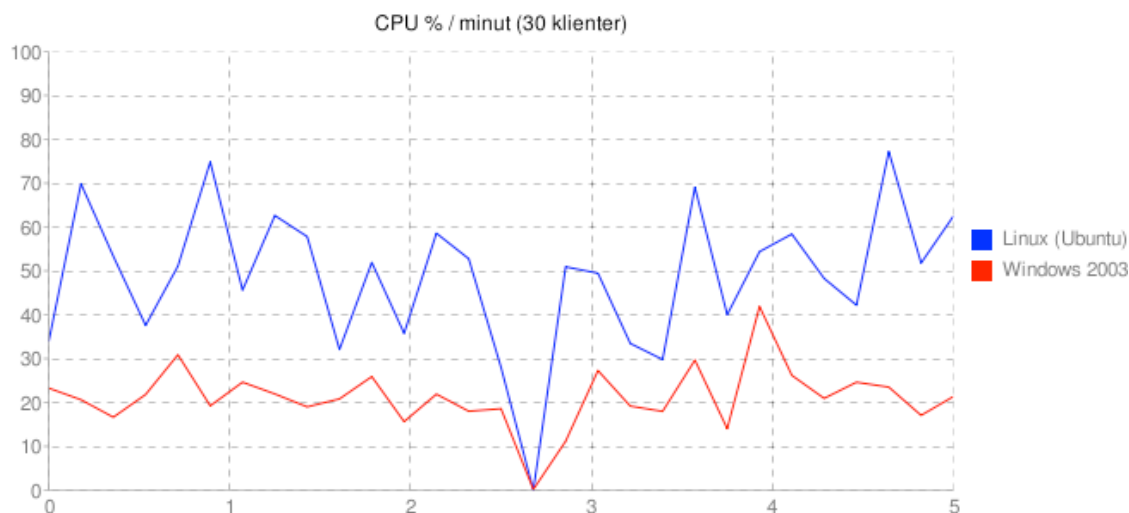


Fig 9. Processoranvändning - 30 klienter under 5 minuter

Fig. 9 visar att processoranvändningen hos de båda testsystemen höll sig relativt konstant över testets 5 minuter. Linuxsystemet låg liksom i de övriga testerna generellt högre än windowssystemet. Snittanvändningen hos linuxsystemet låg på 50% och hos windowssystemet på 24%, alltså knappt hälften.

6.3.2 Svarstider

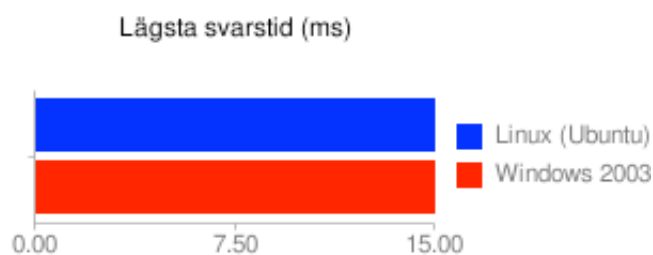


Fig 10. Lägsta svarstid - 30 klienter under 5 minuter

Liksom i övriga test rapporterade OpenSTA 15 ms som lägsta svarstid för de båda testade systemen.

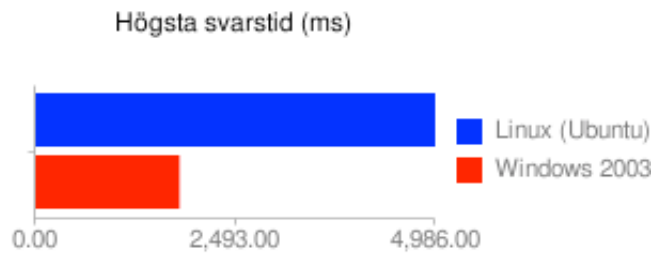


Fig 11. Högsta svarstid - 30 klienter under 5 minuter

Linuxsystemet visade på den högsta svarstiden av de två – 4986 ms. Windowssystemets långsammaste inladdning skedde på 1796 ms.

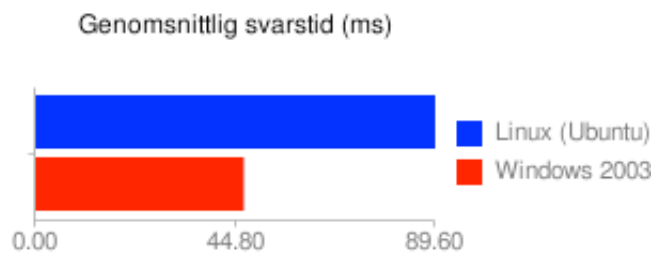


Fig 12. Genomsnittlig svarstid - 30 klienter under 5

Den genomsnittliga svarstiden för linuxsystemet var i detta test 89,6 ms vilket är nästan dubbelt så högt som windowssystemet som i snitt hade 46,7 ms.

6.4 Minnes- och hårddiskaktivitet

Uppgifter om minnes- och hårddiskaktivitet har utelämnats från redovisningarna av de tre utförda lasttesterna. Inga skillnader hos de enskilda testsystemen kunde påvisas mellan de olika lasttesterna och därför redovisas dessa data separat.

6.4.1 Minnesanvändning

Linuxsystemet allokerade totalt i genomsnitt ca 2 550 MB minne (se bilaga B). Den totala minnesallokeringen varierade mellan 2470 MB och 2600 MB både under last och när systemet var obelastat. Delen aktivt minne, alltså den del av det allokerade minnet som utnyttjades av systemet under last varierade mellan 630 MB och 1250 MB med en ungefärlig snittförbrukning på 900 MB (se appendix).

Windowssystemet allokerade totalt i genomsnitt ca 2040 MB minne (se bilaga B) med mycket små variationer mellan obelastat och belastat läge. Delen aktivt minne låg på en snittförbrukning runt 650 MB och inte heller här kunde någon ökning under last påvisas (se appendix).

6.4.2 Hårddiskaktivitet

Endast mycket små skillnader kunde påvisas mellan de två testsystemen vad gäller hårdiskaktivitet (se bilaga C). Esxtop rapporterade ingen läsning alls från hårddisken under något av testen vilket sannolikt berodde på att data cachades av WebSphere Portal när manusen för OpenSTA skapades. Vad gäller skrivning av data till hårddisken förekom detta på en mycket låg nivå - 6 kB/sek i genomsnitt hos linuxsystemet och 21 kB/sek i genomsnitt hos windowssystemet.

7. Diskussion

Detta kapitel reflekterar över resultatet som presenterades i föregående kapitel. Vidare kommer också förslag på fortsatta studier presenteras.

7.1 Reflektion kring resultatet

Det är entydigt att windowssystemet presterade bättre på samma virtuella hårdvara än linuxmotsvarigheten i samtliga tester som genomfördes. Den kanske viktigaste indikationen på hur en slutanvändare upplever systemets prestanda – svarstider – var i genomsnitt mellan dubbelt och fem gånger så höga på linuxsystemet som på windowssystemet. En webbsida är uppbyggd av ett flertal komponenter och är beroende av att samtliga dessa snabbt kan laddas in för att fungera i sin helhet. Linuxsystemets svarstider uppgick som mest till knappt 10 sekunder för en simulerad klient medan den högsta uppmätta svarstiden för en simulerad klient på windowssystemet var hälften eller knappt 5 sekunder. Detta innebär i värsta fall att en användare skulle behöva vänta motsvarande 10 och 5 sekunder på en sidladdning hos respektive system. Viktigt att notera är att ingen cachning av hämtade filer gjordes av OpenSTA, något som är brukligt hos en konventionell webbläsare. Detta medförde en högre belastning på de testade systemen då filer som vanligen bara hämtas en gång som javascript-, css- och bildfiler hämtades för varje sidladdning. De båda testsystemen skulle därför rimligen kunna hantera många fler verkliga användare än vad resultatet antyder.

Det är också entydigt att linuxsystemet konsumerade mer vad gäller tillgängliga processorresurser än vad windowssystemet gjorde. Vid två av lasttesterna nådde linuxsystemet systemets maxgräns för tillgänglig processorkapacitet vilket troligen ledde till ökade svarstider.

De testade systemen baserades på standardinstallationer både vad gäller operativsystem och övriga mjukvaror. Inga ytterligare justeringar gjordes för att resultatet skulle visa på en så rättvis bild som möjligt därför att eventuella optimeringar som görs är specifika för respektive operativsystem. Undersökningens resultat speglar därför inte resultatet som skulle uppkomma i en produktionsmiljö. Att hela systemet inklusive databas lades på samma server som övriga systemet är också ovanligt i en produktionsmiljö där man vanligen separerar applikationsserver, webbserver och databas.

Det föll också utanför studiens omfattning att undersöka hur de testade operativsystemen presterade i en virtuell miljö jämfört med på traditionell maskinvara, men det faktum att åtminstone Ubuntu levererar en version som är optimerad för virtuell maskinvara^[15] antyder att det kan finnas skillnader som kan avspeglas i resultatet.

7.2 Förslag till fortsatta studier

Fortsatta komparativa studier i ämnet skulle kunna fokusera mer på ett utvecklarperspektiv och då lägga mer tid på administrerbarhet. Förstudien visade att det ämnet utan tvekan hade kunnat ta en hel uppsats i anspråk. God administrerbarhet kan leda till sänkta kostnader då utvecklingstid förkortas och öka driftsäkerhet när färre fel begås i administrativa processer. I fortsatta studier bör även försäkra sig om att på ett tidigt stadium ha tillgång till en linuxdistribution som IBM stöder för WebSphere Portal. Framöver kommer det också vara intressant hur Windows Server 2008 fungerar med WebSphere Portal vilket är en naturlig utveckling av att Microsoft så småningom själva slutar ge support på Windows Server 2003. Framtida studier bör också fokusera på att jämföra miljöer som mer liknar produktionsmiljöer med optimeringar av WebSphere Portal, WebSphere Applikation Server, DB2 och de underliggande operativsystemen samt undersöka vilka skillnader som kan påvisas genom att separera de inblandade komponenterna till egna servrar.

8. Slutsats

Klara prestandamässiga skillnader kunde påvisas mellan de båda testsystemen. Linuxsystemet krävde konsekvent mer processorresurser samtidigt som svarstiderna var över de dubbla jämfört med windowsmotsvarigheten. Även vad gäller minnesanvändning krävde linuxsystemet mer av maskinvaran. Vad gäller hårddiskaktivitet så utnyttjade linuxsystemet hårddisken marginellt mindre än windowssystemet.

Resultatet av studien skall ses mot bakgrund av att inget av de testade systemen hade anpassats utöver vad de officiella installationshänvisningarna krävde samt att den testade linuxdistributionen inte stöds av IBM varför resultatet inte kan appliceras på ett optimerat system i produktion. Vidare hade hela systemet installerats på samma server vilket är ovanligt i produktion. Undersökningen tog inte heller upp hur de båda operativsystemen påverkades av den virtuella miljön.

9. Referenser

Referenser utan angiven författare listas först, därefter listan referenser i bokstavsordning efter författare.

1. *Application server - Wikipedia, the free encyclopedia*. URL: http://en.wikipedia.org/wiki/Application_Server (2010-05-20)
2. *Benchmark (computing) - Wikipedia, the free encyclopedia*. URL: [http://en.wikipedia.org/wiki/Benchmark_\(computing\)](http://en.wikipedia.org/wiki/Benchmark_(computing)) (2010-04-30)
3. *Comparison of Windows Server 2003 Editions: General*. URL: <http://technet.microsoft.com/en-us/library/cc758523%28WS.10%29.aspx> (2010-05-18)
4. *Förberedande gruppsamtal på Atea IM* (2010-03-10).
5. *Förberedande samtal på Atea IM* (2010-03-09).
6. *HTTP/1.1: Status Code Definitions*. URL: <http://www.w3.org/Protocols/rfc2616/rfc2616-sec10.html> (2010-05-17)
7. *IBM - WebSphere Application Server - System requirements*. URL: <http://www-01.ibm.com/software/webservers/appserv/was/requirements/> (2010-03-17)
8. *IBM - WebSphere Portal V6.1 hardware and software requirements - Microsoft Windows*. URL: <http://www-01.ibm.com/support/docview.wss?rs=688&uid=swg27011314> (2010-05-07)
9. *Linux kernel - Wikipedia, the free encyclopedia*. URL: http://en.wikipedia.org/wiki/Linux_kernel (2010-05-20)
10. *Linux Online - GNU General Public License*. URL: <http://www.linux.org/info/gnu.html> (2010-03-16)
11. *Open Testware Reviews - OpenSTA (Open System Testing Architecture)*. URL: <http://tejasconsulting.com/open-testware/feature/opensta.html> (2010-05-13)
12. *[Phoronix] Linux Reviews & Articles on Distributions*. URL: <http://www.phoronix.com/scan.php?page=category&item=Distributions> (2010-05-20)
13. *Setting up a stand-alone [server] on Linux*. URL: *Setting up a stand-alone [server] on Linux* (2010-05-07)
14. *Setting up a stand-alone server on Windows*. URL: http://publib.boulder.ibm.com/infocenter/wpdoc/v6r1/topic/com.ibm.wp.ent.doc_v6101/install/win_standalone.html (2010-05-07)

15. *Ubuntu Server Edition JeOS | Ubuntu*. URL: <http://www.ubuntu.com/products/whatisubuntu/serveredition/jeos> (2010-05-18)
16. *VMware Communities: Understanding esxtop*. URL: <http://communities.vmware.com/docs/DOC-11495> (2010-05-14)
17. *Windows Server 2003 - Wikipedia, the free encyclopedia*. URL: http://en.wikipedia.org/wiki/Windows_Server_2003 (2010-05-18)
18. *IEEE Std 610.12, Glossary of Software Engineering Terminology*. 1990.
19. *IBM WebSphere Portal 6.1.X Performance Tuning Guide v 2.1*. 2009, URL: [http://www-10.lotus.com/ldd/portalwiki.nsf/0/2B32B87526796329852576780069E846/\\$FILE/wp61x.tuning_guide_FinalDraft2.pdf](http://www-10.lotus.com/ldd/portalwiki.nsf/0/2B32B87526796329852576780069E846/$FILE/wp61x.tuning_guide_FinalDraft2.pdf) (2010-04-23)
20. Benbya, H., et al., *Corporate portal: a tool for knowledge management synchronization*. *International Journal of Information Management*, 2004(24): p. 201-220.
21. Brown, A. & Patterson, D.A. *To Err is Human*. in *Workshop on Evaluating and Architecting System Dependability*. 2001. Göteborg, Sweden.
22. Brown, A., *Towards Availability and Maintainability Benchmarks: A Case Study of Software RAID Systems*. 2001.
23. Farwell, T. *Measuring Up | AIX | IBM Systems Magazine*. URL: <http://www.ibmssystemsmag.com/aix/februarymarch05/features/6963pl.aspx> (2010-04-30)
24. Ganek, A.G. & Korbi, T.A., *The dawning of the autonomic computing era*. *IBM Systems Journal*, 2003. 42(1): p. 5-18.
25. Kumar, R. *Monitoring and Tuning Memory in DB2 9 for Linux, Unix, and Windows*. 2007, URL: <http://www.dbmag.intelligententerprise.com/story/showArticle.jhtml?articleID=202805692> (2010-04-23)
26. Mattson, M., et al., *Software Architecture Evaluation Methods for Performance, Maintainability, Testability, and Portability*, in *Second International Conference on the Quality of Software Architectures*. 2006.
27. Patel, R. & Davidson, B., *Forskningsmetodikens grunder*. 2003, Lund: Studentlitteratur.
28. Presler-Marshall, M. *IBM WebSphere Portal Performance Troubleshooting Guide v 2.0, s. 9*. 2008, URL: <http://www-01.ibm.com/support/docview.wss?uid=swg27007059>
29. Seltzer, M., et al., *The Case for Application-Specific Benchmarking*. *Proceedings of the Seventh Workshop on Hot Topics in Operating Systems*, 1999: p. 102-107.

Bilagor

A. Intervjufrågor

Hur länge har du jobbat här på Atea?

Beskriv din roll på företaget

Beskriv på vilket sätt du tycker ditt arbete är intressant

Berätta om din bakgrund vad gäller utbildning och eventuella tidigare jobb

Utifrån vilka faktorer anser du att man skall bedöma ett datorsystems interna prestanda, med vilket jag menar icke nätverksrelaterad prestanda?

Oberoende av operativsystem, vilka faktorer anser du är viktiga för att bedöma en mjukvaras prestanda?

Har du erfarenheter av mjukvara som kan installeras på mer än ett operativsystem?

Tänk dig att man installerar samma mjukvaruprodukt på två olika operativsystem. Vilka faktorer hos systemet utöver mjukvaruprodukten är viktiga att titta på när det kommer till att bedöma prestanda?

Tänk dig en mjukvaruprodukt som kan installeras på flera operativsystem. Vilka faktorer anser du är viktiga att bedöma hos mjukvaruprodukten för att finna vilken installation som kan anses vara den optimala i sammanhanget?

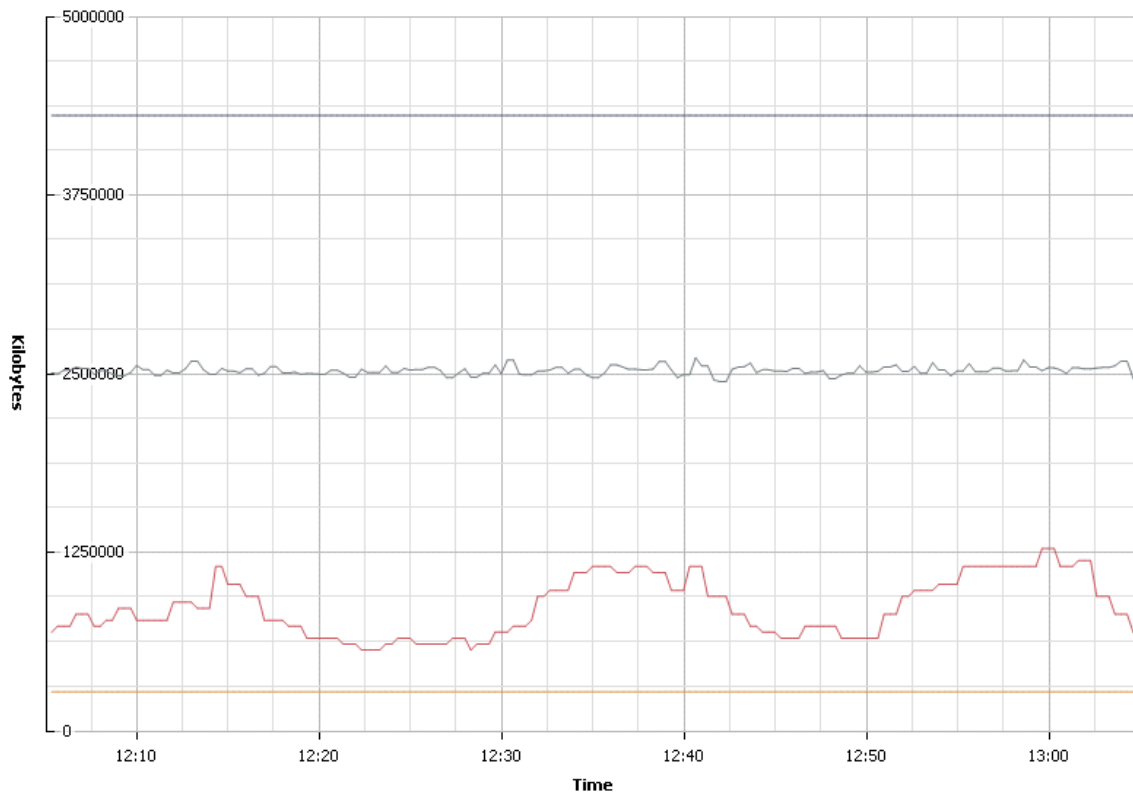
Om du har jobbat med WebSphere Portal, beskriv dina erfarenheter.

Hur tycker du man skall gå tillväga för att bedöma vilken plattform, Linux eller Windows Server 2003, som är bäst lämpad för att använda som grund för WebSphere Portal?





Utöver de faktorer du identifierat tidigare under intervjun, har du några tillägg som kan vara viktiga i det specifika fallet med WebSphere Portal när det kommer till att jämföra prestanda mellan installationer på olika operativsystem förutsatt en identisk hårdvara?

B. Minnesanvändningsexempel från lasttest

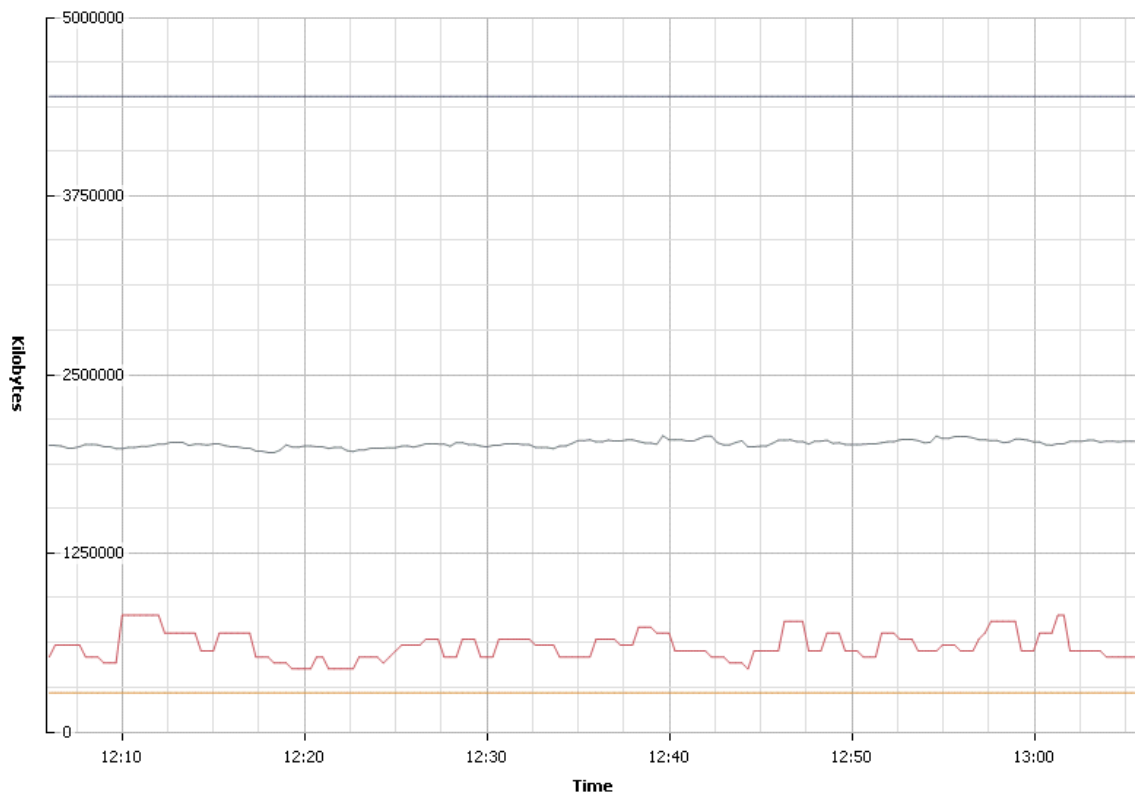
Memory/Real-time, 2010-05-11 12:05:03 - 2010-05-11 13:05:03 - Martin Exjobb Ubuntu 2



Performance Chart Legend

Key	Object	Measurement	Rollup	Units
	Martin Exjobb Ubuntu 2	Granted	average	Kilobytes
	Martin Exjobb Ubuntu 2	Active	average	Kilobytes
	Martin Exjobb Ubuntu 2	Balloon	average	Kilobytes
	Martin Exjobb Ubuntu 2	Consumed	average	Kilobytes

Memory/Real-time, 2010-05-11 12:05:51 - 2010-05-11 13:05:51 - Martin Exjobb Win2003

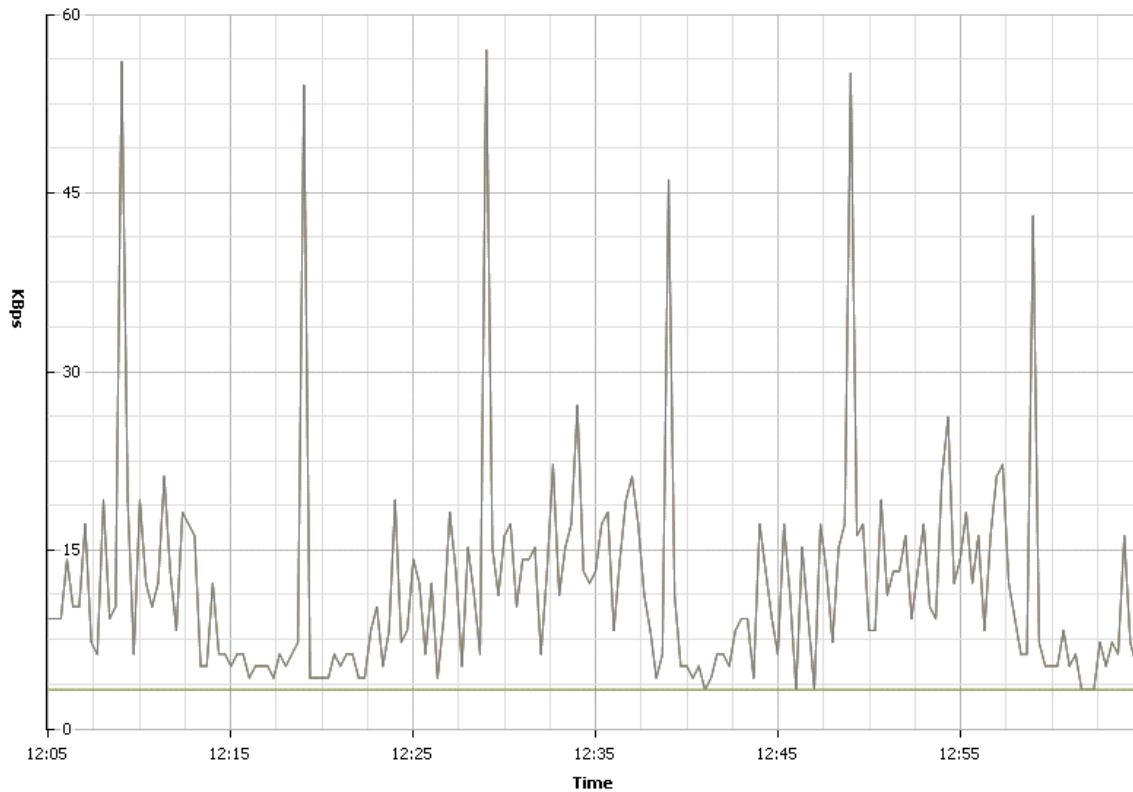


Performance Chart Legend

Key	Object	Measurement	Rollup	Units
■	Martin Exjobb Win2003	Granted	average	Kilobytes
■	Martin Exjobb Win2003	Active	average	Kilobytes
■	Martin Exjobb Win2003	Balloon	average	Kilobytes
■	Martin Exjobb Win2003	Consumed	average	Kilobytes

C. Hårddisk användningsexempel från lasttest

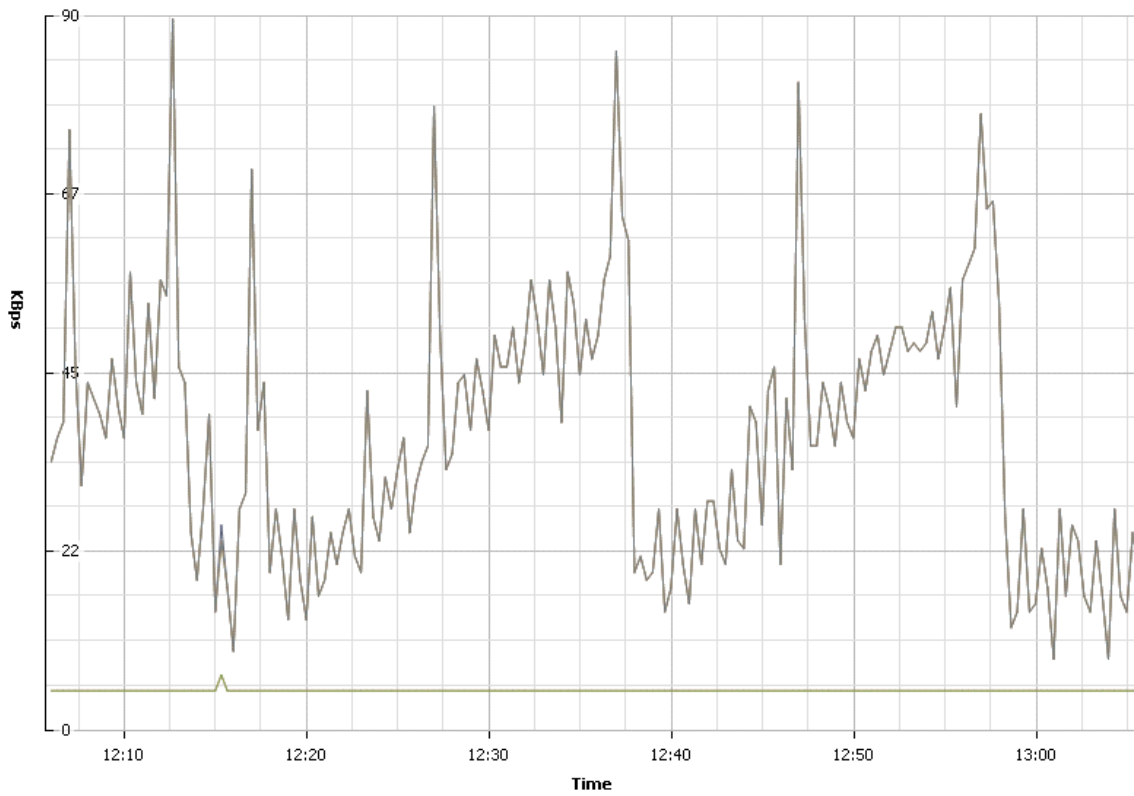
Disk/Real-time, 2010-05-11 12:04:53 - 2010-05-11 13:04:53 - Martin Exjobb Ubuntu 2



Performance Chart Legend

Key	Object	Measurement	Rollup	Units
	Martin Exjobb Ubuntu 2	Usage	average	KBps
	Martin Exjobb Ubuntu 2	Read rate	average	KBps
	mpx.vmhba0:C0:T0:L0	Write rate	average	KBps
	Martin Exjobb Ubuntu 2	Write rate	average	KBps
	mpx.vmhba0:C0:T0:L0	Read rate	average	KBps

Disk/Real-time, 2010-05-11 12:05:41 - 2010-05-11 13:05:41 - Martin Exjobb Win2003



Performance Chart Legend

Key	Object	Measurement	Rollup	Units
	Martin Exjobb Win2003	Usage	average	KBps
	Martin Exjobb Win2003	Read rate	average	KBps
	mpx.vmhba0:C0:T0:L0	Write rate	average	KBps
	Martin Exjobb Win2003	Write rate	average	KBps
	mpx.vmhba0:C0:T0:L0	Read rate	average	KBps