

**eScience = infrastruktur + metodik + tillämpningar**

**Sw-eScience – Förslag till ett nationellt forskningsprogram  
för metodologiska problem och frågeställningar  
inom eScience i ett SNIC-perspektiv**

Rapporten är utarbetad av Erik Elmroth och Bo Kågström  
vid HPC2N på uppdrag av SNIC.

Umeå 2007-04-20

Professor Bo Kågström  
HPC2Ns föreståndare

## Sammanfattning

Denna rapport och de förslag som läggs fram är en del i en sammanhållen forskningssatsning på nationell nivå, vilken ska möjliggöra en internationell förstärkning av vår kompetens och konkurrensförmåga inom eScience - forskningens tredje stöttepelare och nödvändiga komplementet till vetenskapens klassiska par teori och experiment. Rapporten fokuserar på metod- och infrastrukturforskning inom de beräkningsinriktade delarna av eScience.

I Sektion 1 görs en kortfattad beskrivning av eScience stora betydelse idag och i framtiden inom i princip alla forskningsdiscipliner samt inom näringslivet och samhället i stort. Bakgrund ges till behovet av ett nationellt forskningsprogram gällande metodologiska problem- och frågeställningar om eller med stora behov av infrastruktur i form av metoder, algoritmer och programvara för storskalig, effektiv användning av nätverk, hårdvara för beräkning, visualisering och forskningsdatabaser. Behovet av denna forskning drivs av de vetenskapliga frågeställningarna inom de beräkningsvetenskapliga disciplinerna och tillämpningarna, tillsammans med den kraftfulla informationsteknologiska utvecklingen.

Sektion 2 ger en kortfattad internationell utblick av pågående och planerade satsningar inom eScience med fokus på de Nordiska grannländerna, Europa och USA. I Sverige har satsningar inom eScience huvudsakligen avsett en påbörjad uppbyggnad av en långsiktig, uthållig och ändamålsenlig infrastruktur. Det är nu hög tid - ja nödvändigt - att denna infrastruktursatsning kompletteras med en motsvarande sammanhållen forskningssatsning inom eScience.

Denna utrednings övergripande målsättning (se bilaga 1) är att identifiera om och hur en nationell satsning på eScience bör utformas för att på bästa sätt bidra till att infrastrukturen inom SNIC kan nyttjas så effektivt som möjligt av SNICs nuvarande och framtida användare. I Sektion 3 klargörs utredningens avgränsningar och övergripande forskningsbehov beskrivs. Viktiga byggstenar för eScience inom SNICs verksamhetsområde är metodologisk forskning om SNIC-relaterad infrastruktur samt om numeriska och icke-numeriska algoritmer och programvara för modellering, simulering och visualisering, utvecklad för att lösa vetenskapliga problem inom en rad olika discipliner med hjälp av beräkningsvetenskapliga tekniker.

Sektion 4 beskriver drivkrafterna bakom nya forskningsbehov inom eScience och kan sammanfattas i följande tre övergripande punkter: (i) Många etablerade, nya och framtida beräkningsvetenskapliga tillämpningar har stora och växande behov av resurser både för storskaliga beräkningar och datahantering som leder till nya behov av metoder och algoritmer; (ii) Den snabba datorutvecklingen leder ständigt till nya beräkningstekniska problem för att på ett effektivt och robust sätt nyttja datorernas ökade kapacitet och potential; (iii) Nya samarbetsmönster och behov i nya storskaliga projekt leder till forskningsbehov för distribuerade infrastrukturlösningar.

I Sektion 5 föreslås en långsiktig satsning i form av ett nationellt forskningsprogram, **Sw-eScience**, med fokus på metodologiska problem och frågeställningar inom eScience som ett nödvändigt komplement för att få maximal effekt av infrastruktursatsningar inom SNIC. Förslag till utformning, organisation och finansiering av forskningsprogrammet beskrivs. Omslutningen av programmet föreslås minst omfatta 180 MSEK fördelade över minst sex år.

Slutligen, i Sektion 6 beskrivs värdet av en samlad forskningssatsning inom området samt dess betydelse för SNIC och för VRs övriga infrastruktursatsningar inom eScience.

## 1 eScience i ett SNIC-perspektiv – infrastruktur och metodik för beräkningsvetenskaper

Användningen av storskaliga datorresurser för simulering och modellering är idag ofta lika självklar som nödvändig inom många forskningsområden. Inom naturvetenskap beskrivs begreppet *eScience* idag ofta som forskningens tredje stöttepelare - det idag nödvändiga komplementet till vetenskapens klassiska par teori och experiment. I andra sammanhang har användningen av eScience en ännu tydligare särställning, som en direkt ersättning för verkliga experiment. Anledningarna till att ersätta verkliga experiment med modellering och datorsimuleringar är många. Det kan röra sig om experiment som är praktiskt omöjliga att utföra (t.ex. att förutsäga globala klimatförändringar, studier av Universum) eller experiment som är för farliga att utföra (t.ex. att studera effekter av giftutsläpp i en stadsmiljö eller radioaktiva utsläpp). Inom livsvetenskaperna kan användningen av eScience erbjuda möjligheter att ersätta experiment som är oetiska eller kontroversiella av andra anledningar. Andra exempel är fullskaleexperiment som är alltför dyra att utföra och där storskaliga beräkningar och simuleringar med fördel används (t.ex. vindtunneltester av nya flygplansmodeller). I de flesta av dessa exempel finns en stark koppling mellan teori, experiment och eScience som i samverkan är drivkraften till nya stora genombrott inom framförallt medicinsk, natur- och teknikvetenskaplig forskning. Användningen av eScience börjar också alltmer bli ett viktigt komplement inom humanistisk, beteendevetenskaplig och samhällsvetenskaplig forskning.

Under senare tid har *eScience* rönt oerhörd uppmärksamhet som absolut centralt för såväl framgångsrik frontlinjeforskning inom i stort sett alla ämnesområden som för näringslivets och samhällets utveckling i stort. Den SNIC-relaterade beräkningsinriktade forskningen som är föremål för denna utredning utgör den kanske mest centrala kärnan inom eScience. Kärt barn har många namn - förutom begreppet eScience så används ofta de engelska benämningarna *scientific computing*, *computational science*, eller det längre *computational science and engineering*, samt *cyberinfrastructure* för relevanta delar av detta område. Exempel på motsvarande svenska begrepp är *beräkningsvetenskap(er)*, *beräkningsteknik* och *teknisk-vetenskapliga beräkningar*. Beräkningsvetenskaperna utvecklas i tillämpningarna och nära sitt disciplinära sammanhang vilket definierar delområden som t.ex. beräkningsfysik, beräkningskemi, beräkningsmekanik, vilka alla är drivande i den pågående expansiva utvecklingen av området.

Det vi beskrivit ovan gäller framförallt användningen av eScience för att skapa ny kunskap inom olika vetenskapliga domäner. Tillsammans med *infrastruktur* och *metodik* för dess effektiva användning utgör dessa *tillämpningar* de tre huvudkomponenterna inom eScience<sup>1</sup>. *De forskningssatsningar som föreslås i denna utredning avser huvudsakligen viktiga metodologiska problem- och frågeställningar inom de SNIC-relaterade delarna av eScience. Dessa definieras dels av de vetenskapliga frågeställningarna och tillämpningarna, dels av den kraftfulla informationsteknologiska utvecklingen.*

Genom tillgången till kraftfulla parallella datorsystem och effektiv, lättanvänd programvara byggd på moderna högpresterande och robusta algoritmer kan simuleringar och beräkningar utföras för att studera nya klasser av problem. På samma sätt ger tillgång till storskaliga

---

<sup>1</sup> eScience = infrastruktur + metodik + tillämpningar

databaser och effektiva analysverktyg möjligheter att lösa nya forskningsproblem, t.ex. inom medicin, humaniora och samhällsvetenskap. För att uppnå detta och samtidigt stärka Sveriges internationella ställning inom eScience så krävs en kraftfull nationell satsning på forskning kring *metodologiska problem- och frågeställningar, inklusive metoder, algoritmer och programvara för storskalig, effektiv användning av infrastruktur i form av högpresterande datorer, nätverk, visualiseringssystem och forskningsdatabaser.*

## 2 eScience – en internationell utblick

Nedan illustrerar vi, utan ambition att vara heltäckande, hur program och förslag till satsningar inom eScience tar sig uttryck i våra nordiska grannländer, Europa och USA. För mer detaljerade beskrivningar, inklusive olika satsningars faktiska utformning och budget hänvisas till respektive källa. För ett svenskt perspektiv på internationell infrastruktur inom SNICs verksamhetsområde se även Appendix i SNICs landskapsdokument [16].

Ett internationellt tongivande dokument som beskriver behovet av beräkningsinriktad forskning är rapporten *Computational Science: Ensuring America's Competitiveness* utgiven av US President's Information Technology Advisory Committee (PITAC) [18], från vilken vi inleder med två citat:

*“While it is itself a discipline, computational science serves to advance all of science. The most scientifically important and economically promising research frontiers in the 21st century will be conquered by those most skilled with advanced computing technologies and computational science applications.”*

*“Together with theory and experimentation, computational science now constitutes the “third pillar” of scientific inquiry, enabling researchers to build and test models of complex phenomena – such as multi-century climate shifts, multidimensional flight stresses on aircraft, and stellar explosions – that cannot be replicated in the laboratory, and to manage huge volumes of data rapidly and economically. Computational science’s models and visualizations – of, for example, the microbiological basis of disease or the dynamics of a hurricane – are generating fresh knowledge that crosses traditional disciplinary boundaries. In industry, computational science provides a competitive edge by transforming business and engineering practices.”*

PITAC-rapporten är en av flera tunga amerikanska rapporter som entydigt pekar på behovet av ytterligare satsningar på beräkningsteknik. Andra exempel kommer från The Executive Office of the President [5], The Office of Science [11] och The National Science Foundation (NSF) [13], The NSF Cyberinfrastructure Council [10] och The Towards 2020 Science Group [17]. Dessa rapporter uppvisar en närmast slående samstämmighet angående det samlade vetenskapssamhällets stora behov av starka satsningar på beräkningsvetenskaplig forskning.

På den europeiska arenan har Storbritanniens mycket omfattande satsning inom eScience varit vägledande sedan den inleddes 2001. Satsningen är mycket bred och griper över forskning med hemvist på samtliga brittiska forskningsråd. Illustrativa exempel på denna satsnings utformning finns i [1, 6, 7], och en utgångspunkt för ytterligare information är programmets webbsida på Research Councils UK [12]. Programmet har influerat satsningar inom eScience över hela Europa. Vad gäller Europeisk forskningsinfrastruktur, inklusive eScience, så har

ESFRI utfört en omfattande kartläggning av behov och resurser och utformat en Europeisk roadmap [3], liksom the e-Infrastructures Reflection Group (e-IRG) [2].

I Norden är Norge föregångare i utformningen av ett genomgripande program för eScience, eVITA, som inkluderar både infrastruktur och beräkningsvetenskaplig forskning [4, 8]. Behovet av liknande satsningar har identifierats också på andra håll i Norden, och bland annat manifesterats i en utredning om ett Finländskt program för eScience [15] och i Nordiska Ministerrådets tillsättande av en arbetsgrupp för utformning av en samnordisk strategi för satsningar inom eScience.

Inom Sverige organiseras infrastrukturen för eScience sedan 2006 av Vetenskapsrådet (VR) via Kommittén för Forskningens Infrastrukturer (KFI) och dess beredningsgrupp för eScience. Under KFI/eScience är infrastrukturen för den beräkningsinriktade delen av eScience sedan tidigare organiserad inom det nationella metacentret Swedish National Infrastructure for Computing (SNIC) [14]. Såväl SNIC som KFI har nyligen dokumenterat ambitiösa och långsiktiga framtidsplaner [16, 19]. Hittills har den svenska satsningen inom eScience huvudsakligen inriktats mot en påbörjad uppbyggnad av en långsiktig och uthållig och ändamålsenlig infrastruktur.

Det är nu hög tid att denna infrastruktursatsning kompletteras med en motsvarande och sammanhållen forskningssatsning inom eScience. Denna utredning och de förslag som läggs fram utgör en viktig del i en sådan sammanhållen forskningssatsning på nationell nivå, vilken ska möjliggöra en internationell förstärkning av kompetens och konkurrensförmåga inom eScience.

### **3 eScience – avgränsningar och målsättningar**

Utredningens övergripande målsättning är att identifiera om och hur en nationell satsning på eScience bör utformas för att på bästa sätt bidra till att infrastrukturen inom SNIC kan användas så effektivt som möjligt av SNICs nuvarande och framtida användare.

För att kunna nyttja SNICs infrastruktur så effektivt som möjligt bör satsningen omfatta såväl forskning gällande infrastrukturen som grundläggande metoder och tekniker inom eScience, samt metodologisk forskning inom specifika forskningsområden och tillämpningar som har stora behov av förstklassig infrastruktur för eScience. I enlighet med uppdraget så är utredningen avgränsad så att den t.ex. inte behandlar frågor av ren infrastrukturkaraktär (annat än forskning om infrastruktur) eller disciplinär forskning som faller utanför SNICs verksamhetsområde. Forskning som använder infrastruktur eller verktyg för eScience men som inte bedrivs i avsikt att utveckla eller förstå eScience inkluderas inte heller i utredningen.

Viktiga byggstenar för eScience inom SNICs verksamhetsområde är forskning om SNIC-relaterad infrastruktur samt om numeriska och icke-numeriska algoritmer och programvara för modellering och simulering, utvecklad för att lösa vetenskapliga problem inom en rad olika discipliner med hjälp av beräkningsvetenskapliga tekniker.

Denna målsättning och avgränsning ligger helt i linje med den forskningssatsning som SNIC efterfrågade i sitt landskapsdokument [16], genom att då citera PITAC-rapporten [18]:

*“Important building blocks for computational science are:*

- Algorithms (numerical and non-numerical) and modeling and simulation software developed to solve science (such as biological, physical, and social), engineering, and humanities problems,*
- Computer and information science that develops and optimizes the advanced system hardware, software, networking, and data management components needed to solve computationally demanding problems,*
- The computing infrastructure that supports both the science and engineering problem solving and the developmental computer and information science.”*

## **4 eScience – exempel på nya behov av forskning**

Vetenskapliga behov stimulerar forskning om och utveckling av nya beräkningsvetenskapliga tekniker, vilka i sin tur möjliggör forskning inom nya vetenskapliga domäner. För att denna symbios ska leda till stark och uthållig forskning inom såväl det beräkningsvetenskapliga området som inom de vetenskapliga domäner som nyttjar beräkningsvetenskap så krävs långsiktig planering av dessa forskningsinsatser. I ett SNIC-perspektiv inbegriper detta att insatser på HPC-infrastruktur måste koordineras med forskning som långsiktigt leder till att största möjliga vetenskapliga utbyte av denna infrastruktur.

Behovet av forskning inom SNICs verksamhetsområde drivs åtminstone av följande tre faktorer: tillämpningarnas växande behov, den snabba datorutvecklingen och utbredningen av ny infrastruktur för storskaliga samarbeten och komplex problemlösning. Nedan illustrerar vi kortfattat forskningsbehov som uppkommer av var och en av dessa tre drivkrafter, samt ger också exempel på forskningsbehov som är gemensamma för dessa.

Många *etablerade, nya och framtida beräkningsvetenskapliga tillämpningar har stora och växande behov av beräkningskraft* beroende på problemets komplexitet, t.ex., som en följd av geometri, bivillkor, variabla tidsskalor, ökade antal dimensioner, ökade krav på upplösning och noggrannhet, vilket i sin tur ställer motsvarande krav på nyutveckling av algoritmer och programvaror. Viktiga tillämpningar med stora behov av datorkapacitet har nyligen presenterats vid Nordic Computational Grand Challenge Workshop [9], och inkluderar bl.a. svensk forskning inom molekyl-modellering, bioinformatik och biologiska simuleringar. Andra stora behov, inom flödes- och klimatberäkningar, materialmodellering, kvantkemi m.fl. har lyfts fram i och med SNIC och KAWs gemensamma utlysning om medel för stora ämnesspecifika datorsystem. Behovet av stor datorkapacitet går ofta hand i hand med behov av metodutveckling, såväl inom dessa områden som generellt inom t.ex. beräkningskemi, beräkningsfysik, partikelfysik, astronomi, rymdfysik, rymdstrålning, och proteinkristallografi. Mycket aktuella exempel på sådana forskningproblem drivna av tillämpningarnas behov är t.ex. adaptiva metoder, felbegränsningar, generering av beräkningsnät, hierarkiska metoder, multigrad-tekniker, wavelet-baserade metoder, optimeringsmetoder, samt diskretiserings- och tidsstegsmetoder som bibehåller specifika egenskaper hos problemet. Allt detta är starkt knutet till kopplade fysikaliska modeller med olika tidsskalor. Hantering av stora datamängder och visualisering är andra viktiga områden med stark tillämpningsanknytning. Den snabba utvecklingen av forskningsområden med behov av data mining, inom främst livsvetenskaperna, men också t.ex. ekonomi, samhällsvetenskap och humaniora, ställer också nya krav på utveckling av relevant tillämpningsnära programvara för databasrelaterade tillämpningar med behov av högpresterande datorer.

Den snabba datorutvecklingen leder ständigt till nya beräkningstekniska problem för att på ett effektivt och robust sätt nyttja datorernas ökade kapacitet och potential. Exempel på sådana forskningsproblem inkluderar utveckling av parallella algoritmer för såväl generella som tillämpningsnära problem, effektiv hantering av minneshierarkier, effektivt nyttjande av datorernas allt fler och mer komplexa enheter, inklusive s.k. special-purpose-maskinvara för olika typer av beräkningar. Exempel på forskningsbehov inkluderar algoritmer och metoder för att hantera olika former av datalokalitet, inklusive minneshierarkier, parallellisering av adaptiva algoritmer, inklusive datapartitionering och kommunikation. Ytterligare behov innefattar algoritmer och verktyg som dynamiskt kan hantera de svårigheter som uppstår i storskaliga och heterogena miljöer, inklusive t.ex. lastbalansering, hybrida parallelliseringsmetoder, samt återvinning av data och omstart av beräkningar och simuleringar som av någon anledning ofrivilligt avbrutits.

Nya samarbetsmönster och behov i nya storskaliga projekt leder till forskningsbehov för distribuerade infrastrukturlösningar. Den växande utbredningen av Grid-teknologi ställer nya krav på forskning inom i stort sett hela Grid-området, inklusive forskning om grundläggande Grid-infrastruktur för t.ex. resurshantering, säkerhetslösningar, informationssystem, resursallokeringslösningar, betalningssystem, datahanteringslösningar, distribuerade databaser och mer tillämpningsnära forskning för användargränssnitt, portaler, workflow-system, problemlösningssmiljöer och rena Grid-tillämpningar. Infrastrukturforskning av hög relevans inkluderar utveckling av verktyg för ett effektivt nyttjande av högpresterande datorsystem.

Gemensamt för hela det beräkningsvetenskapliga området, inklusive de ovan nämnda områdena, är ett behov av stark forskning om effektivt nyttjande av programspråk samt utveckling av användarvänlig och effektiv matematisk programvara, programbibliotek och miljöer för beräkning, simulering och visualisering [13]. I detta sammanhang är det viktigt att identifiera gemensamma problem inom olika tillämpningsområden.

## 5 Sw-eScience – ett nationellt forskningsprogram

Sammanfattningsvis, kan de övergripande forskningsbehov, som presenterats i sektion 4, för att säkerställa ett effektivt nyttjande av SNICs nuvarande och framtida infrastruktur samlas under följande tre punkter som harmoniserar med PITAC-rapportens ”building blocks” och det norska eVITA-programmets struktur:

- SNIC-relaterad metodologisk forskning för specifika tillämpningar och tillämpningsområden inom eScience.
- Generisk beräkningsteknisk forskning inom eScience-metodologiska discipliner (framförallt datavetenskap, parallella och teknisk-vetenskapliga beräkningar, tillämpad matematik).
- Forskning som syftar till utvecklande av eller effektivt nyttjande av ny infrastruktur, central för SNICs verksamhetsområde.

Vi föreslår en långsiktig riktad forskningsinsats inom dessa tvärvetenskapliga och mångdisciplinära områden som ett nödvändigt komplement för att få maximal effekt av infrastruktursatsningar inom SNIC. En viktig målsättning är att stimulera interdisciplinära samarbeten med ämnesöverskridande kompetens som gemensamt kan ta sig an nya vetenskapliga utmaningar. Inte minst avses här synergier mellan de tillämpningar som driver området och nödvändiga metodologiska discipliner.

Vi föreslår att satsningen görs i form av ett nationellt forskningsprogram med riktade projektanslag och ramanslag på minst 3+3 år. Satsningen måste göras långsiktig och ha en omfattning så att resultaten kan bli de som förväntas. En årlig omslutning på 30 MSEK skulle kunna ge stöd till 10-20 forskargrupper. Räknat över sex år innebär det att den totala omfattningen av forskningsprogrammet bör vara minst 180 MSEK. Alternativt kan de internationellt mest framstående forskargrupperna ges större ramanslag över en längre tidsperiod.

Som lämpliga finansiärer för satsningen ser vi i första hand (i bokstavsordning) Knut och Alice Wallenbergs Stiftelse (KAW), Stiftelsen för Strategisk Forskning (SSF) och Vetenskapsrådet. Satsningen bör koordineras med relaterade Nordiska initiativ, som t.ex. eventuella initiativ av Nordiska Ministerrådet och Nordforsk, där potentiell medfinansiering är tänkbar.

## **6 Sw-eScience – programmets betydelse för SNIC och för VRs övriga infrastruktursatsningar**

Det är uppenbart att den metodologiska forskning som beskrivs i sektionerna 4 och 5 är helt central för SNICs verksamhetsområde och att den på såväl kort som lång sikt borgar för ett effektivare nyttjande av SNICs infrastruktur och att SNICs användare ges möjlighet att ta sig an nya utmaningar och utföra mer avancerad forskning genom effektivare infrastruktur-utnyttjande och med hjälp av mer avancerade och fullständiga modeller och programvaror.

Det finns också stora fördelar med att samla satsningen inom ett program. Till detta hör de förbättrade möjligheterna att mellan forskningsgrupper inom olika tillämpningsområden sprida mer generellt användbara resultat för generiska metodologiska problem. Identifieringen av och fokuseringen på viktig forskning inom eScience kan också i kombination med ett samlat program höja områdets status och därmed attrahera nya kompetenta forskare, och exempelvis ge ytterligare incitament för SNICs användare att ägna större uppmärksamhet åt den metodutveckling som långsiktigt förbättrar forskningens kvalitet. Vidare förbättrar en samlad satsning möjligheterna för deltagande i stora internationella samarbeten, dels genom bättre synbarhet externt, dels genom bättre samordning internt. En samlad satsning skulle också motverka problem med att multidisciplinär beräkningsvetenskaplig forskning hamnar mellan stolarna vid t.ex. VRs ordinarie ansökningsförfarande.

Satsningens organisation bestäms delvis av vem som ansvarar för dess finansiering. Vid finansiering via VR bör satsningen placeras under ämnesrådet för Naturvetenskap och Teknik. Oavsett var satsningen placeras så bör den koordineras med såväl SNICs verksamhet som med KFI:s övriga i sammanhanget relevanta infrastruktursatsningar. Exempel på de sistnämnda är DISC, Nordic DataGrid Facility (NDGF) och en planerad satsning på bioinformatikinфраstruktur. Samordningen ska i första hand vara avsedd för att effektivt dra nytta av uppkomna forskningsresultat, men också i någon mån styra satsningens inriktning. I SNICs fall finns t.ex. ett antal tillämpningsexperter som naturligt kan kanalisera forskningsresultat till SNICs användare. Liknande arrangemang bör också kunna realiseras bland de andra, nyare infrastrukturerna. Sist men inte minst bör poängteras att satsningens koppling till SNIC också bidrar till att förstärka den metodologiska forskningen i sig, genom större närhet till tillämpningar, nödvändig infrastruktur och kompetensnätverk.

## Referenser

- [1] *Developing the UK's e-infrastructure for science and innovation*. Report from the OSI e-Infrastructure Working Group.
- [2] *e-Infrastructures Roadmap*. e-Infrastructures Reflection Group (e-IRG), December, 2005. <http://www.e-irg.eu>.
- [3] *European Roadmap for Research Infrastructures*, Report 2006. European Strategy Forum on Research Infrastructures (ESFRI), 2006.
- [4] *eVITA. eVITenskap og Anvendelser. Forskning i en ny epoke*. Maj 2004.
- [5] Executive Office of the President of the United States. *Federal Plan for High-End Computing*, 2004.
- [6] G. Fox and D. Walker. *e-Science Gap Analysis*, The UK e-Science Technical Report Series, 2003.
- [7] T. Hey and A.E. Trefethen. The UK e-Science Core Programme and the Grid, *Future Generation Computing Systems*, Vol. 18, page 1017- (2002)
- [8] *Infrastructure, Theory and Application (eVITA). Work Programme*. Division for Science eScience, The Research Council of Norway, September 2005.
- [9] Nordic Computational Grand Challenge Survey, March 2007. <http://www.csc.fi/english/csc/courses/archive/FSD2007>
- [10] *NSF's Cyberinfrastructure Vision for the 21<sup>st</sup> Century Discovery*. NSF Cyberinfrastructure Council, NSF, July 2006 (draft).
- [11] Office of Science. *Strategic Plan February 2004*. US Department of Energy, 2004.
- [12] Research Councils UK. [www.rcuk.ac.uk/escience](http://www.rcuk.ac.uk/escience)
- [13] *Simulation-Based Engineering Science. Revolutionizing Engineering Science through Simulation*. Report from the National Science Foundation Blue Ribbon Panel on Simulation-Based Engineering Science, February 2006.
- [14] *SNIC. Swedish National Infrastructure for Computing. Ett Svenskt Metacenter för Högpresterande Datorsystem*. Strategidokument, 2003.
- [15] *Suomen eScience-ohjelma (Finlands eScience-program)*. Finska undervisningsministeriet, arbetsgruppen för gridstrategin, January 2007. [http://www.minedu.fi/OPM/Julkaisut/2007/Suomen\\_eScience\\_ohjelma.html](http://www.minedu.fi/OPM/Julkaisut/2007/Suomen_eScience_ohjelma.html)
- [16] *The Swedish HPC Landscape 2006–2009. Visions and Roadmap*. The Swedish National Infrastructure for Computing (SNIC), 2006.
- [17] *Towards 2020 Science*. Report by the Towards 2020 Science Group for the Microsoft Corporation, 2006.
- [18] US President's Information Technology Advisory Committee (PITAC). *Computational Science: Ensuring America's Competitiveness*. Report to the President, June 2005.
- [19] *Vetenskapsrådets guide till infrastrukturen*, Vetenskapsrådet, Juni, 2006.

## Bilaga 1: Beskrivning av utredningsuppdrag om ett svenskt eScience-program

**Bakgrund:** Den snabba utvecklingen inom informationsteknologi har haft genomslag på nästan alla vetenskapliga områden. Genom tillgången till kraftfulla parallella datorsystem och effektiv, lättanvänd programvara byggd på moderna högpresterande algoritmer kan simuleringar och beräkningar användas för att studera nya klasser av problem. På samma sätt ger tillgång till storskaliga databaser och effektiva analysverktyg nya verktyg för att lösa andra nya forskningsproblem. Begreppet eScience innefattar *forskning om eller med stora behov av infrastruktur i form av programvara och algoritmer för storskalig, effektiv användning av nätverk, hårdvara och forskningsdatabaser.*

SNICs huvuduppdrag är att tillhandahålla medel för och koordinera storskalig dator- och lagringskapacitet för svenska forskare. Detta uppdrag har stora beröringspunkter med övriga eScienceområdet, och SNIC utför också vissa utvecklingsprojekt inom t ex Grid-computing. SNICs planerade utvecklingsprojekt och kopplingarna till forskning inom olika grenar av eScienceområdet beskrivs på olika ställen i SNICs landskapsdokument, t ex handlar kapitel 5.2.1 om ”Resource and Grid Environments”, kapitel 6 behandlar ”Needed SNIC development efforts”, och Kapitel 7 ”Interface to Research in Scientific Computing Techniques and Software Tools”.

Utvecklingsprojekt av den typen som beskrivs i landskapsdokumentet kan, och bör, finansieras inom SNICs ram. Viktiga delar av eScience-forskningen faller dock utanför SNICs uppdrag. Om alternativa huvudmän för denna typ av forskning inte kan identifieras kan det finnas en risk att framsteg inom dessa områden som skulle vara av direkt nytta för SNICs användare inte kommer till stånd. Exempel på fält där denna risk kan vara uppenbar är utveckling av ny teori, nya modeller, metoder, algoritmer och programvaror riktade mot viktiga beräkningsintensiva och/eller dataintensiva grundläggande frågeställningar och tillämpningar inom akademi eller näringsliv.

**Uppdraget:** SNICs styrelse beslutade vid sitt möte 2006-10-25 att ge HPC2N i uppdrag att, i samarbete med övriga SNIC-centra, *utreda hur ett svenskt program för eScience-forskning av direkt nytta för SNICs övriga verksamhet skulle kunna se ut och implementeras.* Utredningen skall presenteras i en sammanhållen rapport på ca 10 sidor som produceras av HPC2N. Rapporten skall behandlas i STAC-kommittén, och en preliminär version skall vara STAC tillhanda senast 2007-03-15, och en slutlig version skall finnas 2007-03-31. Efter att rapporten godtagits av STAC kommer den att behandlas i SNICs styrelse. Utredningen skall belysa åtminstone följande frågeställningar:

- Vilka eScience-områden är mest betydelsefulla för att infrastrukturen inom SNIC skall kunna användas så effektivt som möjligt av SNICs nuvarande och framtida användare?
- Hur omfattande forskningsverksamhet inom dessa områden behövs i Sverige? Vilka effekter för användningen av SNICs infrastruktur kan denna typ av forskning antas få, och vilka blir följderna om denna typ av verksamhet inte kommer till stånd.
- Vad är fördelarna respektive nackdelarna med att samla satsningar enligt ovan inom ett svenskt eScience-program?
- Vilka kopplingar och relationer till SNIC skulle ett svenskt eScience-program kunna ha?
- Vilka tänkbara finansiärer finns för ett svenskt eScience-program?