Hrvatski znanstveni i obrazovni oblak (HR-ZOO) : idejno rješenje izvedbe računalnih, spremišnih i mrežnih resursa

Dobrenić, Dobriša; Imamagić, Emir; Klobučar, Mario; Marić, Ivan; Žagar, Damir Danijel

Other document types / Ostale vrste dokumenata

Publication year / Godina izdavanja: **2019**

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:102:784380

Rights / Prava: Attribution-NonCommercial-NoDerivatives 4.0 International/Imenovanje-Nekomercijalno-Bez prerada 4.0 međunarodna

Download date / Datum preuzimanja: **2024-04-25**

srce
Sveučilište u Zagrebu
Sveučilišni računski centar

Repository / Repozitorij:

Digital repository of the University Computing Centre (SRCE)

dabar
DIGITALNI AKADEMSKI ARHIVI I REPOZITORIJI

SVEUČILIŠTE U ZAGREBU

SVEUČILIŠNI RAČUNSKI CENTAR

HRVATSKI ZNANSTVENI I OBRAZOVNI OBLAK (HR-ZOO)

IDEJNO RJEŠENJE IZVEDBE RAČUNALNIH, SPREMIŠNIH I MREŽNIH RESURSA

Zagreb, travanj 2017.

http://www.srce.hr/

Autori dokumenta su djelatnici Srca (abecedni poredak):

- Dobriša Dobrenić
- Emir Imamagić
- Mario Klobučar
- Ivan Marić
- mr. sc. Damir Danijel Žagar.

Verzije dokumenata:

- v2.0, 2017-04-27:
  Idejno rješenje izvedbe računalnih, spremišnih i mrežnih resursa HR-ZOO-a
- v1.0, 2015-03-09:
  Arhitektura i dizajn računalnih, spremišnih i mrežnih resursa HR-ZOO-a

U Zagrebu, 27. travnja 2017.

Ovo djelo je dano na korištenje pod licencom Creative Commons *Imenovanje-Nekomercijalno-Bez prerada* 4.0 međunarodna. Licenca je dostupna na stranici: http://creativecommons.org/licenses/by-nc-nd/4.0/deed.hr

SADRŽAJ

A. OPIS DOKUMENTA 4

B. UVOD 5

C. OBJEDINJENA INFRASTRUKTURA RAČUNALNIH, SPREMIŠNIH I MREŽNIH RESURSA .. 8

D. INFRASTRUKTURA ZA ZAHTJEVNA RAČUNANJA 10

D.1. Računalstvo visokih performansi 10

D.2. Računalstvo s velikom propusnošću 18

E. INFRASTRUKTURA ZA VISOKO SKALABILNO RAČUNALSTVO 22

E.1. Računalna infrastruktura 22

E.2. Spremišna infrastruktura 23

E.3. Mrežna infrastruktura 24

E.4. Posrednički sustavi 25

F. SUSTAVI ZA ODRŽAVANJE INFRASTRUKTURE 28

F.1. Sigurnosna pohrana podataka 28

F.2. Sigurnosna infrastruktura 28

F.3. Nadzorna infrastruktura 29

F.4. Infrastruktura za automatsku uspostavu poslužitelja i upravljanje konfiguracijama 30

G. KORISNIČKI SUSTAVI 31

G.1. Praćenje korištenja i izvještavanje 31

G.2. Autentikacijska i autorizacijska infrastruktura 31

G.3. Služba pomoći i tehnička podrška 31

G.4. Korisnička sučelja 32

H. KATALOG USLUGA 33

I. RASPODJELA RESURSA 35

REFERENCE 37

A. OPIS DOKUMENTA

Ovaj dokument daje pregled dostupnih i nadolazećih tehnologija potrebnih za izvedbu objedinjene infrastrukture računalnih, spremišnih i mrežnih resursa te definira idejno rješenje njihove izvedbe u infrastrukturi HR-ZOO-a, kao i usluge koje će biti na raspolaganju korisnicima te infrastrukture.

Ciljevi ovog dokumenta su:

- omogućiti dionicima sustava znanosti i visokog obrazovanja uvid u planiranu tehnološku izvedbu nacionalne e-infrastrukture te katalog usluga koje će im stajati na raspolaganju po završetku projekta kako bi pojedinci i ustanove mogli bolje planirati svoje aktivnosti
- omogućiti Ministarstvu znanosti i obrazovanja, Hrvatskoj zakladi za znanost, Sveučilištima i znanstvenim institutima, kao i ostalim tijelima i ustanovama, bolje planiranje rješavanja budućih zahtjeva za e-infrastrukturom
- omogućiti krajnjim korisnicima, istraživačima i nastavnicima uvid u katalog usluga HR-ZOO-a
- omogućiti trgovačkim društvima, zainteresiranim integratorima i potencijalnim izvoditeljima uvid u potrebne funkcionalnosti te veličinu pojedinih komponenti e-infrastrukture HR-ZOO-a
- omogućiti općoj javnosti uvid u tehnološku izvedbu te katalog usluga nacionalne e-infrastrukture sustava znanosti i visokog obrazovanja.

Dokument je organiziran na sljedeći način: u poglavlju B (Uvod) dan je kratak osvrt na ulogu e-infrastrukture u modernoj znanosti i obrazovanju, te je opisano postojeće stanje i potrebe za e-infrastrukturom u Hrvatskoj s osvrtom na računalnu i spremišnu komponentu. Poglavlje C (Objedinjena infrastruktura računalnih, spremišnih i mrežnih resursa) sadrži opis općenitih svojstava koje infrastruktura HR-ZOO-a i njene komponente trebaju zadovoljiti. Poglavlja D (Infrastruktura za zahtjevna znanstvena računanja) i E (Infrastruktura za visoko skalabilno računalstvo) daju detaljan pregled dostupnih računalnih, spremišnih i mrežnih komponenata te posredničkih sustava za potrebe infrastruktura za zahtjevna računanja i visoko skalabilno računalstvo. U poglavlju F (Sustavi za održavanje infrastrukture) opisani su sustavi koji omogućavaju učinkovito održavanje cjelokupne infrastrukture. Poglavlje G (Korisnički sustavi) sadrži pregled korisničkih sustava koje je potrebno uspostaviti za jednostavniji pristup infrastrukturi HR ZOO-a te kvalitetniju korisničku podršku. U poglavlju H (Katalog usluga) predstavlja se katalog usluga buduće infrastrukture HR-ZOO-a. Poglavlje I (Raspodjela resursa) daje sažet prijedlog raspodjele resursa pojedinih slojeva infrastruktura unutar HR-ZOO-a koji trebaju zadovoljiti elaborirane zahtjeve kroz poglavlja C – H.

B. UVOD

E-infrastruktura je složeni integrirani sustav koji se sastoji od velikog broja komponenti temeljenih na informacijskoj i komunikacijskoj tehnologiji i predstavlja okruženje u kojem istraživači, nastavnici, studenti i drugi članovi akademske i istraživačke zajednice surađuju i zajednički ostvaruju pristup raspodijeljenim i/ili jedinstvenim elementima istraživačke i obrazovne infrastrukture, bez obzira na vrstu i zemljopisni smještaj tih sredstava.

Kompetitivna i međunarodno relevantna znanost, te u konačnici pretvaranje znanstvenih postignuća u inovativne proizvode i usluge, danas su nezamislivi bez modeliranja i simuliranja (engl. *Simulation Data*), obrade podataka prikupljenih tijekom eksperimenata (engl. *Experimental Data*) ili opažačkih podataka (engl. *Observational Data*) instrumenata, što zahtijeva e-infrastrukturu za:

- brze obrade podataka
- spremanje velikih količina podataka
- prijenos velikih količina podataka
- pristup udaljenoj laboratorijskoj opremi
- suradničke sustave za zajednički rad velikog broja zemljopisno udaljenih istraživača.

Takva zahtjevna računanja i obrade podataka (engl. *Advanced Computing*) omogućuju sljedeće računalne infrastrukture:

- infrastruktura za računalstvo visokih performansi (engl. *High-Performance Computing* ili ***HPC***) – ostvarena npr. u superračunalima (engl. *supercomputers*) ili računalnim klasterima (engl. *computer clusters*)
- infrastruktura za računalstvo s velikom propusnošću (engl. *High-Throughput Computing* ili ***HTC***) – ostvarena npr. kroz računalnu grid-infrastrukturu ili slabo povezane računalne klastere.

Sve je više područja znanosti koja intenzivno trebaju HPC ili HTC e-infrastrukturu; spomenimo samo neka polja: fizika, geofizika, kemija, matematika, biotehnologija, medicina, inženjerstvo.

Međutim, i ostala područja znanosti, kao i segment obrazovanja, koja se možda manje oslanjaju na zahtjevne računalne obrade u svom svakodnevnom radu trebaju različite računalne ili spremišne resurse za unapređenje učinkovitosti istraživanja i edukacije, dakle zahtijevaju e-infrastrukturu koja je:

- lako pristupačna
- visoko prilagodljiva
- skalabilna.

Ovakvo visokoskalabilno računalstvo (engl. *High-Scalability Computing* ili ***HSC***), ovisno o načinu uporabe, naziva se:

- računalstvo u oblacima (engl. *Cloud Computing*) – ostvareno kroz *cloud* usluge kao što su:
    - *Infrastructure as a Service* (IaaS)
    - *Platform as a Service* (PaaS)
    - *Software as a Service* (SaaS)
    - *Data as a Service* (DaaS)
- napredno računalstvo u oblacima (engl. ***HPC in the Cloud***) – ostvareno npr. kao *High Performance Cloud Computing* (***HPC2***).

Nadalje, mala i srednja poduzeća također u svom inovacijskom ciklusu, od faze idejnog koncepta i izvodljivosti preko istraživanja, razvoja i testiranja, pa sve do komercijalizacije i izlaska na tržište, imaju snažnu potrebu za e-infrastrukturom istih svojstava kakve zahtijeva znanost.

E-infrastruktura treba dakle zadovoljiti sve te grupe korisnika, omogućavajući pritom interdisciplinarnost, njihovo umrežavanje i sinergiju u istraživanju, transferu i diseminaciji znanja.

Izgradnja takvih e-infrastruktura predviđena je i u programu Europske unije za istraživanje i inovacije za razdoblje od 2014. do 2020. godine „Obzor 2020.“ (Horizon 2020) u prioritetu "Izvrsna znanost" kao dio aktivnosti „Istraživačke infrastrukture (RI)“. E-Infrastructure Reflection Group ([1]) preporuča nacionalnim vladama provedbu izgradnje nacionalnih e-infrastruktura kroz osiguravanje financijskih sredstava, okupljanje korisničkih zajednica, prilagodbe zakonodavstva radi omogućavanja otvorenog pristupa, kao i kroz poticanje suradnje s europskim e-infrastrukturama.

Osnovni građevni elementi e-infrastrukture za ranije navedene potrebe korisnika o kojima se mora posebno voditi računa su:

- sustavi za autentikaciju (engl. *Authentication*), autorizaciju (engl. *Authorization*) i upravljanje pristupom resursima (engl. *Access control*)
- računalstvo visokih performansi i računalstvo s velikom propusnošću
- računalstvo u oblacima
- lokalno i globalno umrežavanje velikog kapaciteta
- sustavi za spremanje, očuvanje, upravljanje, indeksiranje i pristup podacima
- sustavi za upravljanje i pristup bazama podataka
- podrška razvoja programske podrške za istraživače
- sustavi za upravljanje, nadzor i praćenje korištenja, te pouzdanosti i dostupnosti podatkovnih centara
- sustavi koji osiguravaju energetsku učinkovitost i praćenje utjecaja podatkovnih centara na okoliš.

Teoretski je moguće da svaka grupa korisnika, ili čak svaka ustanova, gradi svoju e-infrastrukturu samo za svoje potrebe. To bi ipak bilo potpuno neučinkovito korištenje

financijskih resursa, uz vjerojatno neadekvatno operativno održavanje. Osim toga tu su i kompleksni i skupi sustavi za fizičko smještanje opreme, energetsko napajanje, klimatsko održavanje, vatrodojavu, nadzor (opisani npr. u dokumentu HR-ZOO „Tehnički opis sustava podatkovnog centra“). Također, nabava potrebne komercijalne programske podrške manje je isplativa u manjim količinama.

Stoga je razumno graditi zajedničke e-infrastrukture, nacionalne i europske, koje će onda korisnici dijeliti i pomoću njih surađivati, uz osigurano središnje stručno upravljanje, te uz osiguran utjecaj samih korisnika na razvoj i održavanje. Zajednička e-infrastruktura poduprijet će također i otvoreni pristup (engl. *Open Access*), otvorene podatke (engl. *Open Data*), otvorena istraživanja (engl. *Open Research*), naposljetku otvorenu znanost (engl. *Open Science*) i stvaranje i podržavanje novih standarda.

Velik je broj ustanova i pojedinaca koji traže skalabilnu *cloud* infrastrukturu za povećanje djelotvornosti svojih standardnih poslovnih procesa (npr. početkom 2017. godine 87 ustanova koristilo je *cloud* infrastrukturu virtualnih privatnih poslužitelja Srca, a 969 pojedinačnih korisnika registrirano je za korištenje *cloud* infrastrukture virtualnih učionica i laboratorija).

Iz navedenog proizlazi da se stanje u Hrvatskoj uvelike poklapa s uobičajenim europskim i svjetskim potrebama, osim što je, nažalost, u korištenju zajedničke e-infrastrukture kod malih i srednjih poduzeća u Hrvatskoj tradicionalno prisutan manji interes. Interes i naposljetku poduzetničko korištenje postojećih infrastruktura u svrhu razvoja, ili kao mjesto dodira znanosti i gospodarstva, iznimka je koju se mora i zajedničkom e-infrastrukturom i suradničkom organizacijom pretvoriti u pravilo.

U Srcu su u posljednjih desetak godina izgrađeni neki dijelovi nacionalne e-infrastrukture koji za sada uglavnom djeluju kao zasebni sustavi:

- AAI@EduHr ([2]) – autentikacijska i autorizacijska infrastruktura sustava znanosti i visokog obrazovanja u RH
- Računalni klaster Isabella ([3]) – HPC sustav
- Hrvatska nacionalna grid-infrastruktura, CRO NGI ([4]) – HTC sustav
- *Cloud* usluga Srca *Virtual Private Server*, VPS ([5]) – IaaS *cloud* sustav
- *Cloud* usluga Srca *Virtual Computing Lab*, VCL ([6]) – IaaS/SaaS *cloud* sustav
- MojOblak ([7]) – spremišni sustav.

Ti sustavi pokazuju da postoje stručnjaci s usvojenim i izgrađenim tehnološkim znanjima, ali i da za potpunu e-infrastrukturu treba uložiti sredstva za širenje sustava, zatim ih integrirati u jedinstvenu cjelinu prilagođenu korisnicima, korisnicima pružati konkretnu pomoć u uporabi e-infrastrukture, omogućiti im djelotvorni utjecaj na njeno upravljanje te naposljetku povezati hrvatsku e-infrastrukturu s europskim sustavima. Jednako je važno tu izgrađenu e-infrastrukturu kasnije financijski, organizacijski i tehnički održavati.

C. OBJEDINJENA INFRASTRUKTURA RAČUNALNIH, SPREMIŠNIH I MREŽNIH RESURSA

Bez napredne i razvijene e-infrastrukture sustav visokog obrazovanja i znanosti ne može uspješno ostvarivati promjene i svoju ulogu u društvu i gospodarstvu. Takva infrastruktura koja će se graditi u okviru projekta Hrvatskog znanstvenog i obrazovnog oblaka (HR-ZOO) omogućit će usluge virtualnih računalnih i spremišnih resursa na načelu računalstva u oblacima, resurse grida, računalne resurse visokih performansi, velike spremišne kapacitete, brzu optičku mrežu i povezivanje prema europskoj akademskoj i istraživačkoj mreži GEANT.

Izgradnja infrastrukture HR-ZOO-a predstavlja izazov izgradnje objedinjene infrastrukture koja će istovremeno osiguravati računalne i spremišne resurse (na prvi pogled) različitim vrstama računalnih infrastruktura – uz zahtjeve za lako dostupnim resursima kojima se zadovoljavaju uobičajene potrebe znanstvene zajednice za računalnim resursima (HSC), takva infrastruktura treba istovremeno zadovoljiti i potrebe za izvođenje zahtjevnih znanstvenih računanja (HPC i HTC). Prilikom izgradnje podatkovnih centara i planiranja računalno-spremišne infrastrukture koja će u njima biti ugrađena nužan je holistički pristup kojim se istovremeno razmatraju različiti aspekti infrastrukture i osigurava infrastruktura koja je:

- Visokoučinkovita – osigurava resurse za:
  - učinkovito izvođenje paralelnih procesa (HPC)
  - podržavanje izvođenja U/I zahtjevnih procesa
  - učinkovito konsolidiranje računalnih poslužitelja (HSC).
- Energetski učinkovita – preduvjet za izgradnju „zelenog“ računalstva:
  - učinkovito hlađena i toplinski upravljana – hlađena tekućinom, uz rad na višim temperaturama ili uz više temperature rashladnog medija)
  - u potpunosti integrirana s energetski učinkovitim podatkovnim centrima (korištenje otpadne toplinske energije)
  - učinkovito izvedenog sustava napajanja (napajana visokim naponom radi smanjenja gubitka distribucije)
  - upravljana u skladu s opterećenjem (neaktivni dijelovi infrastrukture ne troše energiju)
  - učinkovito projektirana (odabrane arhitekture sklopovlja prilagođene su različitim oblicima opterećenja i načina korištenja)
  - energetska učinkovitost jednako je bitna kao i računalna moć.
- Skalabilna – omogućava:
  - dinamičku preraspodjelu resursa među uslugama/korisnicima
  - jednostavno horizontalno skaliranje

  - skaliranje kojim će se zadovoljiti potrebe u planiranom razdoblju eksploatacije.
- Pouzdana
  - za pojedine segmente (HSC) potrebno je osigurati visoku pouzdanost i dostupnost usluge, za što je preduvjet pouzdana i visoko dostupna (HA) infrastruktura.
- Otvorena
  - izvedba zasnovana na korištenju otvorenih tehnologija omogućava implementaciju heterogene infrastrukture čije komponente nisu vezane uz pojedinog dobavljača opreme, već je moguće kombiniranje komponenti različitih dobavljača koje međusobno dobro „surađuju“
  - sistemska programska podrška zasnovana na otvorenom kodu
  - za izvedbu se koriste slobodno dobavljive (COTS, engl. *commercial off the shelf*) komponente.
- Integrirana i modularna
  - učinkovito provedena integracija pojedinačnih komponenti
  - dobro uspostavljene hijerarhijske razine
  - izvedena integracija infrastrukture s elementima podatkovnog centra.
- Dostupna
  - prostorno – bilo gdje
  - vremenski – bilo kad.
- Upravljiva
  - pruža sučelja za učinkovito nadziranje i upravljanje svim komponentama.

D. INFRASTRUKTURA ZA ZAHTJEVNA RAČUNANJA

Infrastruktura za zahtjevna znanstvena računanja omogućava izvođenje programske podrške koja postavlja izuzetno visoke zahtjeve za performansama i dostupnošću računalnih resursa (procesorskim jezgrama, radnom memorijom, diskovnim prostorom itd.).

Jedna od podjela zahtjevnih računanja zasnovana je na međusobnoj interakciji/sprezi poslova koji se u istom trenutku odvijaju na nekoj računalnoj infrastrukturi. Na taj način razlikujemo:

- Računalstvo visokih performansi – HPC (*High Performance Computing*)
- Računalstvo s velikom propusnošću – HTC (*High Throughput Computing*).

Izvođenje takvih poslova postavlja zahtjeve na računalnu infrastrukturu koji uvjetuju i razlike u izvedbama pojedinih segmenata infrastrukture.

D.1. RAČUNALSTVO VISOKIH PERFORMANSI

Računalstvo visokih performansi podrazumijeva istovremeno izvođenje različitih poslova (procesa/dretvi) koji međusobno komuniciraju preko dijeljene (engl. *shared*) ili distribuirane (engl. *distributed*) memorije.

U danas najčešćim izvedbama sustava s dijeljenom memorijom, u kojima pojedini procesori pristupaju udaljenoj lokalnoj memoriji drugih procesora (engl. *Non-Uniform Memory Access*, NUMA), izvedbe se najčešće zasnivaju na mehanizmima koji osiguravaju koherentnost priručne (engl. *cache*) memorije između različitih procesora, koji su izvedeni ili sklopovljem koje je ugrađeno u samim procesorima ili dodatnim sklopovljem koji su razvili pojedini proizvođači sklopovlja. Izvedbe sustava s dijeljenom memorijom ne omogućavaju implementaciju heterogene infrastrukture sklopovljem različitih proizvođača.

Osnova za izvedbu sustava s distribuiranom memorijom je mrežna infrastruktura koja osigurava visoku propusnost i nisku latenciju u prenošenju poruka, pri čemu je prenošenje poruka izvedeno standardnim programskim bibliotekama (engl. *Message Passing Interface*, MPI). Zbog mogućnosti implementacije standardno dostupnim komponentama (poslužiteljima – računalnim čvorovima, mrežnim preklopnicima…) sustavi koji koriste distribuiranu memoriju (računalni klasteri) imaju nižu cijenu i brojniji su u odnosu na sustave s dijeljenom memorijom.

Bitno je naglasiti da izvedbe s dijeljenom memorijom omogućavaju izvedbu jednostavnijih programskih modela, a samim tim i lakšu/jednostavniju izvedbu programske podrške.

Zbog dugotrajnosti i složenosti poslova koji se ponekad izvode na ovakvoj infrastrukturi poželjno je pojedine komponente izvesti kao visoko dostupne ili u višim (upravljačkim) mehanizmima implementirati mehanizme kojima se može povećati pouzdanost izvođenja (mehanizmi za privremeno zaustavljanje i nastavljanje izvršavanja poslova – *checkpoint&restart*).

D.1.1. RAČUNALNA INFRASTRUKTURA

Računalna infrastruktura najbitnija je komponenta u izvedbi HPC infrastrukture. Cilj računalstva visokih performansi je brza obrada velike količine podataka, pa je osnovna metrika koja se koristi u ovom području sposobnost procesora u izvršavanju operacija s pomičnim zarezom u dvostrukoj preciznosti (*double precision floating-point*) i najčešće se iskazuje u obliku FLOPS (engl. *FLoating-point Operations Per Second*).

Računalna infrastrukture namijenjena računalstvu visokih performansi treba osigurati:

- dostupnost računalnih resursa (procesorskih jezgri i radne memorije)
- visoku propusnost i nisku latenciju
    - sučelja za povezivanje procesora
    - priručne i radne memorije
    - sučelja prema perifernim uređajima
    - komunikacijske infrastrukture za povezivanje računalnih čvorova
- optimirano izvođenje aritmetičkih operacija.

Karakteristike danas dostupnih centralnih procesora (CPU), koje omogućavaju ispunjavanje postavljenih zahtjeva, su sljedeće:

- velik broj procesorskih jezgri u pojedinom procesoru
    - 24 (IBM POWER9) ili 32 jezgre (Intel Xeon E5 v5 / AMD Zen)
- visoke frekvencije izvođenja
    - 3 GHz (npr. Intel Xeon) do 4 GHz (npr. IBM POWER9)
- visoka memorijska propusnost
    - 120 GB/s (IBM POWER9 SO), 230 GB/s (IBM POWER9 SU)
    - 6 ili 8 ugrađenih memorijskih upravljačkih jedinica
    - uporaba DDR4 memorije brzine 2,666 GHz
    - više razina priručne memorije (uz postojeće razine L1/L2/L3)
        - L4 eDRAM (IBM POWER9)
- podržana količina radne memorije po pojedinom procesoru
    - 3 TB (Intel Xeon E7), 4TB (IBM POWER9)
- instrukcijski set podržava ubrzavanje izvođenja numeričkih operacija
    - AVX (*Advanced Vector Extensions, Intel*) – inačica AVX-512
    - FMA3 (*Fused Multiply-Add, Intel/AMD*) – inačice FMA3 i FMA4
    - VMX (*Vector Multimedia Extension, IBM*) – inačica VMX3
- visokopropusna sučelja za povezivanje s drugim procesorima i pomoćnim procesorima
    - UPI (*Ultra Path Interconnect*, Intel) – 120 GB/s

    - NVLink 2.0 (Nvidia) – 160 GB/s
    - CAPI (*Coherent Accelerator Processor Interface*, IBM) – 300 GB/s
- visokopropusna sučelja za komunikaciju s perifernim uređajima
    - PCIe Gen 4 – 4 GB/s po stazi
    - podržan veći broj komunikacijskih staza – do 128 staza (npr. AMD Zen).

Očekivanim unapređenjem postojećih tehnologija tijekom izvedbe projekta očekuje se dodatno poboljšanje performansi procesora.

Današnji procesori izvedeni su kao višejezgreni procesori. Ta paradigma još je više istaknuta u procesorima koji su namjenski razvijeni za paralelnu obradu velike količine podataka – grafičkim procesorima (engl. *Graphic Processing Unit,* GPU) i mnogojezgrenim procesorima (engl. *Many Integrated Cores*, MIC). Takvi procesori sposobni su izvršavati velik broj istih (engl. *Single Instruction Multiple Data*, SIMD) ili različitih operacija/instrukcija (engl. *Multiple Instruction Multiple Dana*, MIMD) nad tokom (engl. *stream*) velike količine podataka pa se često nazivaju i procesorima toka (engl. *Stream Processors*). Njihove prednosti u usporedbi s klasičnim procesorima su:

- veće performanse (5-10x, trenutačno ~5 TFLOPs)
- veća energetska učinkovitost (10-15x, trenutačno ~70 GFLOPs/W)
- veća memorijska propusnost (5-10x, trenutačno ~1 TB/s).

Pojavom namjenskih sučelja/sabirnica za povezivanje takvih procesora s centralnim procesorima i za njihovo međusobno povezivanje (NVIDIA NVlink, IBM CAPI) otklonjeni su dosadašnji nedostaci u izvedbi njihove integracije u računalnim poslužiteljima.

Zbog navedenih prednosti, u odnosu na izvedbu koja koristi samo centralne procesore, uporaba GPU/MIC procesora neizostavna je komponenta za izvedbu energetski učinkovitih HPC sustava, pa time i nužna u izvedbi HPC infrastrukture projekta HR-ZOO.

Uz odabrane procesore, performanse računalne infrastrukture ovise i o količini, tipu i načinu pristupa (hijerarhijskoj organizaciji) memoriji iz koje procesor dohvaća podatke. Kako bi povećali memorijsku propusnost, odnosno brzinu dohvaćanja podataka iz glavne memorije, najavljene izvedbe centralnih i pomoćnih procesora, uz uobičajene L1-L3 razine pomoćne memorije, uključit će i veću količinu pomoćne L4 memorije (CPU), odnosno ugrađene glavne memorije (engl. *Embedded DRAM*, eDRAM) na samom procesoru. Planirane izvedbe GPU/MIC procesora podržavat će više od 32 GB ugrađene memorije, uz izvedbe memorijskih sabirnica koje osiguravaju izuzetno brz pristup memoriji i niske latencije pristupa (engl. *High Bandwidth Memory*, HBM).

Izvedba i količina ugrađene memorije utječe i na energetsku potrošnju poslužitelja (energetska potrošnja memorije i procesora iznosi otprilike 1/3 ukupne potrošnje poslužitelja) pa je, uz povećanje količine dostupne radne memorije, nužno osigurati smanjenje potrošnje. Najavljene su i izvedbe postojane (engl. *non-volatile*) glavne memorije koju je moguće priključiti direktno na memorijske upravljačke jedinice procesora, čijom uporabom će se osigurati mogućnost smanjene (pa čak i nulte) potrošnje energije u trenucima kada procesor ne pristupa memoriji.

U nekim područjima znanosti (npr. bioinformatika) količina dostupne memorije značajnija je od dostupnih procesorskih resursa, pa je za potrebe takvih aplikacija potrebno osigurati poslužitelje s velikom količinom memorije. Dostupnošću programske podrške koja omogućava programsko upravljanje/definiranje memorije (engl. *Software Defined Memory*) moguće je udružiti memorijske resurse više poslužitelja unutar jedinstvenog adresnog prostora jednog virtualnog poslužitelja (npr. ScaleMP vSMP [8]). Uz udruživanje na razini radne memorije (RAM), takva programska podrška omogućava i višerazinsko upravljanje memorijom (engl. *tiering*), pri čemu se kao više razine memorije, za pohranu podataka kojima programski kod prilikom izvršavanja pristupa rjeđe, mogu koristiti napredne izvedbe SS (engl. *solid state*) pogona (npr. Intel 3D XPoint, Optane, …) povezane na perifernim (PCIe) ili memorijskim sabirnicama. S obzirom na nižu cijenu takvih medija za pohranu u odnosu na radnu memoriju (DRAM), izvedba višerazinskog upravljanja memorijom omogućava izvedbu sustava s velikom količinom memorije uz nižu cijenu.

D.1.2. SPREMIŠNA INFRASTRUKTURA

Spremišna infrastruktura koja se koristi u području računalstva visokih performansi treba:

- osigurati mogućnost pristupa velikog broja klijentskih poslužitelja
- podržati izvedbu HPC globalnih datotečnih sustava, odnosno:
    - osigurati istovremeni paralelni pristup pojedinim datotekama
    - podržati brz pristup velikim datotekama
    - podržati brz pristup metapodacima (podrška velikog broja datoteka)
    - podržati izvođenje mehanizma za privremeno zaustavljanje/pokretanje poslova (*checkpoint/restart*)
    - osigurati mogućnost pouzdane pohrane i replikacije podataka na istoj ili na udaljenoj lokaciji
    - podržati mogućnost memorijski mapiranog pristupa podacima (npr. DI-MAPP)
- osigurati više razina pohrane podataka:
    - razinu pohrane izuzetno brzog pristupa (3D XPoint pogoni, NVMe sabirnica)
    - razinu pohrane brzog pristupa (SSD pogoni, NVMe sabirnica)
    - razinu pohrane često korištenih podataka (SSD pogoni, SAS/SATA sabirnica)
    - razinu pohrane velikih količina podataka (NL HDD pogoni, SAS/SATA sabirnica)
    - razinu pohrane rijetko korištenih podataka (magnetske trake)
- osigurati mogućnost brzog direktnog pristupa podacima (RDMA)
- podržavati HPC mrežnu infrastrukturu (npr. InfiniBand RDMA ili Ethernet RoCE)
- biti izvedena dostupnim sklopovljem (COTS)

- koristiti programske komponente otvorenog koda
- omogućiti veliku gustoću pohrane podataka (> 5PB spremišnog kapaciteta u ormaru)
- podržati jednostavno horizontalno skaliranje.

D.1.3. MREŽNA INFRASTRUKTURA

Zbog potrebe za brzim prenošenjem velikog broja poruka (MPI) kojima se ostvaruje komunikacija među procesima koji se istovremeno izvode na različitim poslužiteljima, zahtjevi HPC infrastrukture u mrežnom segmentu fokusirani su na ostvarivanje visoke propusnosti i niskih latencija. Uz navedene karakteristike, takve izvedbe zahtijevaju mogućnost jednostavnog i (gotovo) neograničenog skaliranja te imaju niske zahtjeve spram sigurnosti mrežne infrastrukture.

Mehanizam prenošenja poruka zasniva se na brzom prenošenju segmenata memorije u kojima su poruke pohranjene. Kako bi se ovaj mehanizam izvodio što brže, potrebno je osigurati minimalan utrošak vremena na sam proces kopiranja ili u potpunosti isključiti kopiranje iz postupka prenošenja poruka (engl. *zero-copy*), zbog čega se implementiraju mehanizmi koji iz cijelog procesa isključuju uplitanje centralnog procesora (CPU) i prepuštaju izvođenje svih operacija povezanih s prenošenjem poruka sklopovlju ugrađenom u mrežnom sučelju. Naziv kojim su objedinjeni svi takvi mehanizmi je RDMA (engl. *Remote Direct Memory Access*), tj. udaljeni izravni pristup memoriji. InfiniBand je RDMA implementirao od svojih početaka, a Ethernet u svojim posljednjim inačicama uvodi protokol RoCE (engl. RDMA *over Converged Ethernet*).

U okvirima HR-ZOO-a razmatramo/predviđamo u HPC okružju uporabu sljedećih mrežnih tehnologija u inačicama koje će biti dobavljive u planiranom vremenskom periodu:

- Infiniband
  - 100 Gb/s – EDR (*Enhanced Data Rate*)
  - 200 Gb/s – HDR (*High Data Rate*)
- Omni-Path
  - 100 Gb/s
  - 200 Gb/s
- Ethernet
  - 100 Gb/s.

Pojedine izvedbe mrežnih tehnologija, uz RDMA mehanizme, uključuju i druge mehanizme kojima se neki funkcionalni elementi mehanizma prijenosa poruka (MPI) izvode u namjenskom sklopovlju, te time dodatno smanjuju latenciju prijenosa poruka.

Nedostatak mehanizma prenošenja poruka je u tome što ne može u potpunosti zamijeniti izvođenje posla u okružju dijeljene memorije, pa postoje izvedbe posredničke programske infrastrukture (*middleware*) kojima se programski ostvaruje koherencija priručne memorije procesora na poslužiteljima koji su međusobno povezani Ethernet ili InfiniBand mrežom (npr. ScaleMP vSMP). Pri tome takva posrednička infrastruktura često implementira vlastite

mehanizme i protokole prijenosa poruka bez oslanjanja na više (programske) razine koje su ugrađene u mrežnim tehnologijama – koristeći samo osnovne (L1 i L2) razine mrežne infrastrukture.

U vremenskom periodu izvedbe projekta očekuje se i integracija mrežnog sklopovlja za povezivanje udaljenih procesora na samom procesoru (npr. Intel Omni-Path) čime će se omogućiti direktna komunikacija između udaljenih procesora bez posredovanja komunikacijske/periferne sabirnice, odnosno dodatno smanjiti latencija u prijenosu poruka. Ovisno o dostupnosti navedene tehnologije, u okviru projekta HR-ZOO, razmotrit će se njena uporaba u izvedbi HPC dijela infrastrukture.

Mrežna infrastruktura Ethernet uglavnom podržava *closs i mash* topologije. Kako je Ethernet od početka zamišljen kao mrežna infrastruktura u kojoj je uobičajeno da nije potrebno osigurati jednaku propusnost između bilo koja dva povezana poslužitelja, odnosno u kojoj postoji preplaćenost (engl. *oversubscription*), takva infrastruktura uobičajeno je razdijeljena na više razina, koristi se sljedeće nazivlje: *spine* i *leaf*. U izvedbi je uobičajeno koristiti preklopnike različitih performansi, o kojima ovise i cijene, u različitim razinama. Izvedba mrežne topologije kojom bi se osigurala potpuna propusnost između svih poslužitelja u mreži nije uobičajena i iziskivala bi troškove zbog kojih bi uporaba Etherneta za svrhe HPC-a bila potpuno neprihvatljiva.

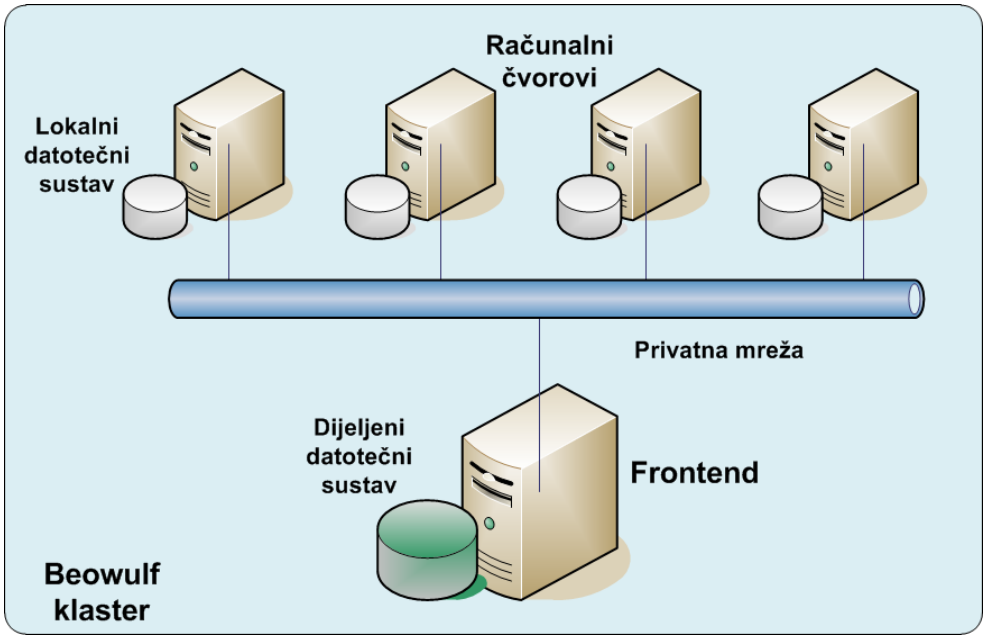
InfiniBand podržava različite mrežne topologije koje se izvode tako da uglavnom osiguravaju punu propusnost između svih povezanih poslužitelja. Uz standardne topologije *tree* ili *fat-tree* (rabi se za *tree* topologiju koja osigurava punu propusnost) i *mesh* (zbog načina izvedbe u kojem je svaki poslužitelj direktno povezan sa svim ostalim poslužiteljima uobičajeno se u InfiniBand terminologiji rabi naziv *all-to-all*), InfiniBand podržava i hibridne topologije koje su specifične za HPC okolinu: *Hypercube, Enhanced Hypercube, 2D/3D Torus, Dragonfly*. Ove topologije osiguravaju mogućnost izvedbe mrežne infrastrukture uporabom manjeg broja jednostavnijih preklopnika i kabliranje uporabom bakrenih kabela na manjim udaljenostima, pa ih je svakako nužno razmotriti u konačnoj izvedbi ove razine HPC infrastrukture.

Kako je mrežna razina zajednička svim poslužiteljima, potrebno je osigurati višu razinu pouzdanosti u njenoj izvedbi. Ethernet i InfiniBand osiguravaju visoku pouzdanost izvođenjem višestrukih putanja između povezanih poslužitelja (Ethernet *Multi-Path* ili InifiniBand *Multi-Rail*). Osim povećane pouzdanosti, izvođenjem višestrukih putanja i agregacijom višestrukih komunikacijskih kanala, omogućeno je i povećanje propusnosti koje je proporcionalno broju korištenih kanala, pa time i povećanje performansi u prijenosu poruka.

Osim za međusobno povezivanje poslužitelja, navedene mrežne tehnologije koriste se u HPC okolini i za izvedbu podatkovnih spremišta i datotečnih sustava te njihovo povezivanje s poslužiteljima. Ranije spomenuta RDMA tehnologija potrebna je i za ostvarivanje ove funkcionalnosti, a podržana je na obje (Ethernet i InfiniBand) mrežne tehnologije. Za povezivanje podatkovnih spremišta i poslužitelja koristi se protokol SRP (*SCSI RDMA Protocol*).

D.1.4. POSREDNIČKI SUSTAVI

Najčešća arhitektura sustava za potrebe računalstva visokih performansi je *Beowulf* klaster. *Beowulf* klaster sastoji se od skupa poslužitelja koji se dijele na *frontend* i radne čvorove. *Frontend* čini središte klastera i na njemu su obično smješteni dijeljeni datotečni sustav te središnje komponente sustava za upravljanje poslovima, sustava za nadzor klastera i sustava za automatsku instalaciju čvorova. Radni čvorovi koriste se isključivo za izvođenje korisničkih poslova. *Frontend* i čvorovi povezani su u jednu ili više privatnih mreža koje su fizički izolirane od javne mreže. Na taj način ostvaruje se učinkovitija komunikacija između čvorova i veća razina sigurnosti za radne čvorove. Korisnici pri korištenju *Beowulf* klastera rade isključivo na *frontendu* koji tako čini jedinstvenu točku pristupa. Arhitektura *Beowulf* klastera prikazana je na slici 1.



Slika 1. Arhitektura *Beowulf* klastera

Za potrebe računalstva visokih performansi potrebno je osigurati sljedeće posredničke sustave:

- sustav za upravljanje poslovima koji korisnicima omogućava učinkovito izvođenje aplikacija na HPC resursima
- raspodijeljeni datotečni sustav za uspostavu spremišta visokih performansi opisanog u poglavlju E1.2
- prevodioce i programske knjižnice za razvoj i izvođenje zahtjevnih aplikacija.

Najznačajnija svojstva sustava za upravljanje poslovima su:

- korisničko komandno-linijsko, *web* i programsko sučelje; poželjna je podrška za standardno programsko sučelje DRMAA (engl. *Distributed Resource Manager Application API*)
- upravljanje poslovima: zaustavljanje, praćenje stanja posla, mijenjanje svojstava poslova

- prikaz povijesti i statistike o izvođenju poslova
- podrška za različite oblike poslova: serijske, paralelne, interaktivne, *workflow*, polja poslova
- napredno opisivanje zahtjeva poslova (npr. količina radne memorije, količina privremenog spremišnog prostora, arhitektura procesora itd.)
- korisničko postavljanje prioriteta poslovima
- napredno opisivanje radnih čvorova klastera (npr. postavljanje vlastitih atributa za potrebe raspoređivanja poslova)
- podrška za *prologue* i *epilogue* skripte – jednostavni programi koji obavljaju pripremu radnih čvorova prije i nakon izvođenja korisničkih poslova (npr. stvaranje privremenih direktorija, postavljanje varijabli okoline itd.)
- podrška za različite politike korištenja resursa (npr. davanje prioriteta određenim grupama korisnika ili poslova, *fair-share* itd.)
- mogućnost integracije posebnih procesorskih jedinica (npr. GPU, DSP, FPGA itd.)
- napredni mehanizmi raspoređivanja i upravljanja poslovima:
    - rezervacija resursa (fiksne i jednokratne) za pojedine korisnike ili vrste poslova
    - *checkpointing* – mogućnost spremanja stanja procesa radi kasnijeg pokretanja iz spremljenog stanja, ostvaruje se pomoću podrške u operacijskom sustavu, programskoj knjižnici ili samoj aplikaciji
    - *preemption* – mogućnost obustavljanja izvođenja jednog posla i pokretanje drugog posla
    - selidba poslova s jednog čvora na drugi radi postizanja ravnomjernog opterećenja klastera (engl. *load balancing*) ili obavljanja administracijskih operacija na čvorovima
    - *backfilling* – mogućnost izvršavanja kratkotrajnih poslova u vremenskom razdoblju između dvaju rezervacija resursa
    - *CPU harvesting* – mogućnost iskorištavanja radnih stanica kada ih njihovi vlasnici ne koriste; ovaj mehanizam nije ključan za Beowulf klastere.

Najpoznatiji sustavi za upravljanje poslovima su:

- HTCondor ([10])
- IBM Platform LSF (engl. *Load Sharing Facility*) ([11])
- PBS Professional (engl. *Portable Batching System*) ([12])
- Slurm (engl. *Simple Linux Utility for Resource Management*) ([13])
- Sun of Grid Engine ([14])
- Univa Grid Engine ([15]).

Najpoznatiji raspodijeljeni datotečni sustavi za uspostavu učinkovite spremišne infrastrukture su:

- BeeGFS ([16])
- GlusterFS ([17])
- IBM GPFS (engl. *General Parallel File System*) ([18])
- Lustre ([19]).

D.2. RAČUNALSTVO S VELIKOM PROPUSNOŠĆU

Kao osnovna razlika između HPC i HTC računalstva navodi se interakcija/sprega poslova koji se u istom trenutku odvijaju na nekoj računalnoj infrastrukturi. Ova razlika uvjetovana je načinom na koji se u ovim paradigmama obrađuju podaci, odnosno o zavisnosti rezultata trenutačne obrade o obradi koja je nedavno završila. U računalstvu visokih performansi trenutačna obrada zavisna je o nedavno obavljenoj obradi koja se možda izvršila na nekom drugom, udaljenom, procesoru pa je za dohvaćanje tog rezultata potrebno izvesti mehanizme za brzo prenošenje tog rezultata, za čiju izvedbu je nužna mrežna infrastruktura koja osigurava veliku propusnost i nisku latenciju. Računalstvo s velikom propusnošću izvodi obrade koje (uglavnom) ne ovise o ranijim rezultatima, pa je vrijeme koje pojedini posao troši na dohvaćanje podataka manje nego kod računalstva visokih performansi. Istovremeno, takav način obrade rezultira obradom u kojoj nije nužno pamćenje/zadržavanje ranijeg stanja (engl. *stateless*), odnosno kao posljedicu ima manje zahtjeve za privremenom pohranom rezultata u lokalnoj memoriji poslužitelja. Metrika koja se zbog toga često koristi u naprednom računalstvu za klasificiranje određenog tipa posla je omjer vremena obrade i vremena potrebnog za komunikaciju/dohvaćanje podataka.

Za razliku od HPC infrastrukture koja je konsolidirana na jednoj lokaciji, zbog nepostojanja zahtjeva za spregom poslova koji se izvršavaju, infrastruktura koja podržava HTC računalstvo često se izvodi raspodijeljeno na različitim fizičkim lokacijama.

D.2.1. RAČUNALNA INFRASTRUKTURA

Računalstvo s velikom propusnošću ne zahtijeva pouzdanu izvedbu računalne infrastrukture. U slučaju kvara pojedine komponente – poslužitelja, posao koji se na njemu izvršavao može se pokrenuti na novom poslužitelju. Zahtjevi koji se postavljaju na računalnu infrastrukturu niži su nego u slučaju infrastrukture za računalstvo visokih performansi, a računalna infrastruktura je cjenovno vrlo osjetljiva, odnosno treba nuditi što veći broj pojedinih (samostalnih) procesorskih jezgri uz što nižu cijenu.

Za izvedbu poslužitelja mogu se koristiti jeftinije inačica procesora (koje sadrže manji broj memorijskih upravljačkih jedinica uz ograničenu skalabilnost na samo 2 procesorska utora) s većim brojem procesorskih jezgri. Količina memorije po pojedinoj jezgri u HTC okolini uglavnom je manja nego u slučaju HPC infrastrukture.

Iznimku u zahtjevima za snagom pojedine procesorske jezgre i količine memorije čini infrastruktura namijenjena paralelnoj obradi velike količine podataka (*bigdata*) i analizom podataka (*data analytics*) uporabom programskog modela MapReduce ([20]) ili baza podataka koje nisu zasnovane na jeziku SQL (NoSQL/NewSQL). Slični zahtjevi postavljaju

se i u slučaju izvedbe simulacija koje se oslanjaju na metodu Monte-Carlo ([21]). Zbog zahtjeva koji se nalaze između HPC i HTC, takav tip obrade podataka često se izdvaja u zasebnu paradigmu računalstva – MTC (engl. *Many Task Computing*).

Zbog niskih zahtjeva za spregom u istovremenom izvođenju i obradi podataka moguća je i konsolidacija HTC računalne infrastrukture s HSC infrastrukturom u segmentu koji ne osigurava visoku pouzdanost.

D.2.2. SPREMIŠNA INFRASTRUKTURA

Ovisno o načinu izvedbe, odnosno u slučaju fizičke distribucije računalnih resursa HTC infrastrukture, na spremišnu infrastrukturu postavljaju se dodatni zahtjevi u slučaju obrade zajedničkog skupa podataka – takva infrastruktura treba podržati/automatizirati distribuciju podataka (bliže mjestu obrade podataka) – bilo mehanizmima migracije/replikacije podataka ili izvedbom distribuiranih datotečnih sustava koji sadrže ove mehanizme.

U slučaju izvedbe spremišne infrastrukture koja podržava MTC obrade, uz izvedbu zajedničke/dijeljene spremišne infrastrukture nužno je osigurati i spremišnu infrastrukturu koja je lokalna za određeni dio računalne infrastrukture (poslužitelj).

Zbog zahtjeva za osiguravanjem visoke propusnosti i podrškom velikog broja ulazno/izlaznih operacija, takva infrastruktura izvodi se uporabom SS (*solid state*) ili NVRAM (*Non-volatile RAM*) pogona za pohranu podataka.

D.2.3. MREŽNA INFRASTRUKTURA

Mrežna infrastruktura koja treba podržavati računalstvo s velikom propusnošću ne treba osigurati visoku propusnost i nisku latenciju u prijenosu podataka, kao što je to slučaj kod HPC mrežne infrastrukture, stoga su prihvatljive izvedbe mrežne infrastrukture zasnovane na uobičajenoj Ethernet tehnologiji. Kao iznimku treba razmatrati slučajeve podatkovno intenzivnih obrada koje zahtijevaju veću propusnost mrežne infrastrukture.

U okvirima HR-ZOO-a razmatramo/predviđamo u HTC okružju uporabu sljedećih tehnologija u inačicama koje će biti dobavljive u planiranom vremenskom periodu:

- Ethernet
    - 25 Gb/s
    - 50 Gb/s
    - 100 Gb/s.

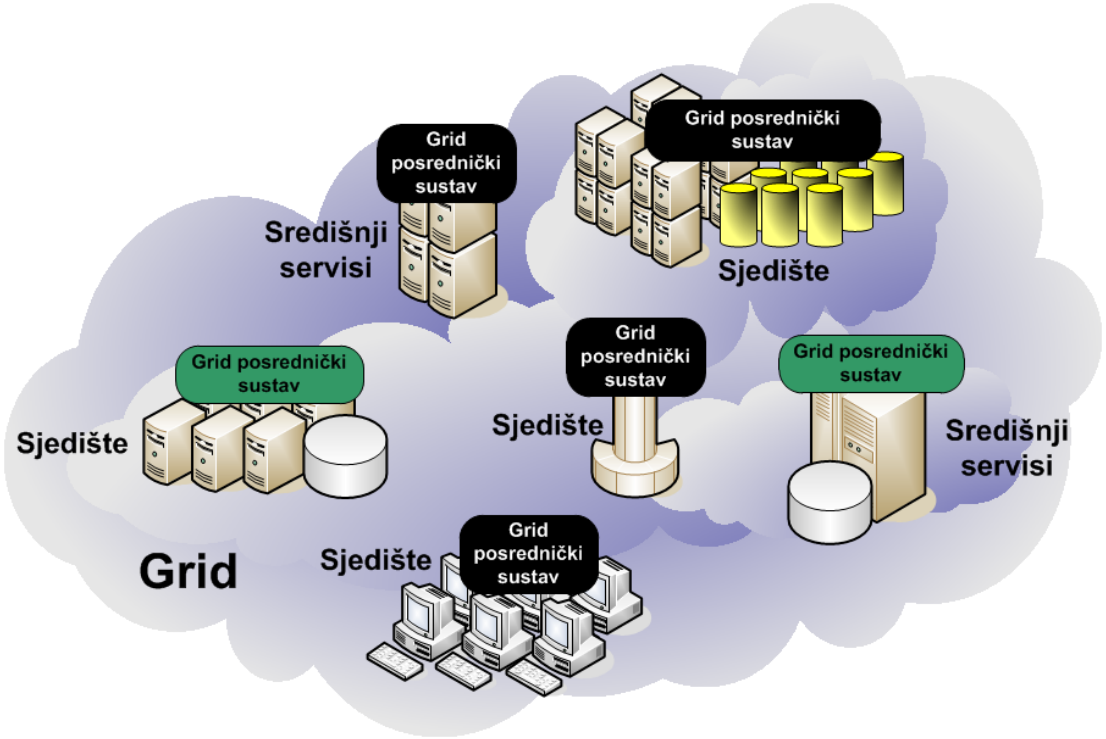
Mrežna infrastruktura Ethernet uglavnom podržava *F* i *mash* topologije. Kako je Ethernet od početka zamišljen kao mrežna infrastruktura u kojoj je uobičajeno da nije potrebno osigurati jednaku propusnost između bilo koja dva povezana poslužitelja, odnosno u kojoj postoji preplaćenost (engl. *oversubscription*) takva infrastruktura uobičajeno je razdijeljena na više razina.

Uz izvedbu protokola SRP (*SCSI RDMA Protocol*), mrežna infrastruktura Ethernet zadovoljava zahtjeve za osiguravanjem niske latencije prilikom pristupa podatkovnim spremištima.

U slučaju potrebe povezivanja resursa koji se nalaze na međusobno udaljenim lokacijama (podatkovnim centrima) nužno je osigurati mogućnost povezivanja udaljenih mrežnih infrastruktura. Povezivanje se izvodi na najmom optičkih kapaciteta i upotrebom WDM-a (engl. *Wavelength Division Multiplexing*) ili odgovarajućih L2 (Ethernet) ili L3 (IP) TK kapaciteta.

D.2.4. POSREDNIČKI SUSTAVI

Za potrebe računalstva s velikom propusnošću, na razini jedne fizičke lokacije, koriste se isti posrednički sustavi kao i za računalstvo visokih performansi. Na razini više fizičkih lokacija najčešće se koriste grid posrednički sustavi koji pružaju standardna sučelja iznad sustava za upravljanje poslovima i pohranu podataka te sigurnosne mehanizme potrebne za rad u raspodijeljenoj okolini. Arhitektura HTC/grid-infrastrukture prikazana je na slici 2.



Slika 2. Arhitekture HTC/grid-infrastrukture

S obzirom da će HTC resursi biti raspodijeljeni na više fizičkih lokacija ključna je i integracija sa autentikacijskim i autorizacijskim infrastrukturama te visoka razina sigurnosti kako bi se osigurao pristup i komunikacija između udaljenih lokacija. Također je ključno da sustav za upravljanje poslovima ima mogućnost raspoređivanja poslova između više računalnih klastera i automatske selidbe poslova u slučaju ispada pojedinog klastera.

Najčešće korišteni grid posrednički sustavi su:

- ARC (engl. *Advanced Resource Connector*) ([22])
- CREAM (engl. *Computing Resource Execution And Management*) ([23])
- Globus Toolkit ([24])
- UNICORE (engl. Uniform Interface to Computing Resources) ([25]).

EGI FedCloud ([26]) je federacija sjedišta koja pristup resursima pružaju koristeći paradigmu oblačnog računalstva. Oblačno računalstvo omogućava fleksibilniji pristup resursima, primjerice korisnici mogu pokretati vlastite virtualne poslužitelje s odgovarajućim operacijskim sustavima i programskim paketima. Sjedišta FedCloud sadrže standardne komponente koje omogućavaju transparentno korištenje različitih implementacija oblačnog računalstva. Najbitnije su:

- sučelje OCCI (engl. *Open Cloud Computing Interface*) za upravljanje virtualnim poslužiteljima na sjedištima
- sučelje CDMI (engl. *Cloud Data Management Interface*) za upravljanje podacima pohranjenim na sjedištima.

U okviru HR-ZOO-a za potrebe HTC infrastrukture moguće je koristiti identičan cloud posrednički sustav kao i za potrebe HSC infrastrukture. Dio virtualizacijskih poslužitelja će tako biti alociran za HTC virtualne poslužitelje, a dio za HSC. Na taj način će se konsolidirati fizički resursi te kvalitetnije iskoristiti cjelokupna infrastruktura.

E. INFRASTRUKTURA ZA VISOKO SKALABILNO RAČUNALSTVO

Infrastruktura za visoko skalabilno računalstvo osigurava mrežno-računalno-spremišne resurse za izvedbu infrastrukture računalstva u oblacima i naprednog računalstva u oblacima.

U slučaju računalstva u oblacima radi se o infrastrukturi namijenjenoj konsolidaciji i osiguravanju dinamičkih računalnih resursa – virtualnih poslužitelja i podatkovnih resursa. Virtualni poslužitelji trebaju podržati izvedbu statičkih ili dinamičkih (skalabilnih) scenarija (slučajeva) raspodjele opterećenja/poslova.

Napredno računalstvo u oblacima najčešće podrazumijeva izvođenje HTC poslova u virtualiziranoj okolini, pri čemu su zahtjevi za infrastrukturom, uz dodatak razine koja sadrži mehanizme za izvedbu virtualizacije i osigurava upravljanje, isti kao i u slučaju izvedbe HTC infrastrukture.

Zajednička karakteristika takve infrastrukture je dinamičko kreiranje/otpuštanje resursa uz istovremenu podršku velikog broja korisnika te osiguravanje pouzdanih resursa na koje se korisnici mogu osloniti te njima u potpunosti supstituirati (postojeće) resurse koji bi inače bili u njihovom vlasništvu. Za dinamičko kreiranje korisničkih resursa nužna je mogućnost programskog upravljanja/definiranja svih sastavnih dijelova infrastrukture.

U nekim segmentima (aplikacijama) zahtijeva se i visoka pouzdanost infrastrukture što se u izvedbi rješava osiguravanjem visoko pouzdanih resursa ili umnogostručivanjem (osiguravanjem redundantnih) resursa i posredničke programske podrške (engl. *middleware*) namijenjene koordinaciji upravljanja redundantnim resursima.

S obzirom na prostornu izvedbu takva infrastruktura može biti izvedena na jednom mjestu (konsolidirana) ili distribuirana na različitim lokacijama. Fizička razdvojenost resursa nužna je za osiguravanje kontinuiteta poslovanja, odnosno neprekidnosti poslovnog procesa (engl. *business continuity*).

E.1. RAČUNALNA INFRASTRUKTURA

HSC računalna infrastruktura treba osigurati prividno neograničenu količinu resursa za izvedbu virtualnih poslužitelja/računala u različitim aplikacijama:

- Poslužitelji u oblacima (IaaS)
- Udomljeni poslužitelji – VPS (engl. *Virtual Private Server)*
- Računala za izvedbu virtualnih radnih površina - VDI (engl. V*irtual Desktop Infrastructure*)
- Računala za izvedbu udomljenih radnih površina – HDI (engl. *Hosted Desktop Infrastructure*)
- Zahtjevna računanja u oblacima.

S obzirom na različite zahtjeve u pojedinim segmentima, takva infrastruktura je heterogena, odnosno treba u pojedinim segmentima biti izvedena kao visoko učinkovita i visoko dostupna (za izvedbu kritičnih udomljenih poslužitelja koji zahtijevaju mogućnost

vertikalnog skaliranja), a u drugim osigurati resurse kod kojih dostupnost nije uvjetovana ali su, zbog potreba za visokom skalabilnošću, bitni čimbenici u izvedbi infrastrukture cijena, gustoća pakiranja i mogućnost jednostavnog horizontalnog skaliranja.

Visoku pouzdanost i mogućnost vertikalnog skaliranja moguće je osigurati uporabom jeftinijih odvojenih (redundantnih) resursa koji su međusobno koordinirani prikladnim programskim komponentama.

E.2. SPREMIŠNA INFRASTRUKTURA

HSC spremišna infrastruktura ovisna je o segmentu izvedbe HSC resursa.

U segmentu koji zahtijeva visoku pouzdanost i replikaciju postoji potreba za izvedbom visokopouzdane i infrastrukture visokih performansi. Takva infrastruktura može biti izvedena zasebnim diskovnim spremištima ili COTS spremišnim poslužiteljima, pri čemu je pohrana podataka programski upravljana/definirana (engl. *Software Defined Storage, SDS*). Kako ova razina visoke dostupnosti zahtijeva geografski raspodijeljena sjedišta – spremišna infrastruktura treba podržavati replikaciju podataka i slika virtualnih poslužitelja među udaljenim lokacijama.

Nužno je razmotriti izvedbu spremišne infrastrukture koja je izvedena postojećom računalnom infrastrukturom (engl. *hyper-converged infrastructure*, HCI). Takva izvedba je programski upravljana/definirana, pri čemu ne koristi zasebne računalne resurse već izvodi spremišne resurse na podatkovnim uređajima ugrađenim u poslužitelje računalne infrastrukture.

Segment koji osigurava visokoskalabilnu infrastrukturu oslanja se na COTS spremišne poslužitelje koji osiguravaju nisku cijenu za ugrađeni diskovni kapacitet, a spremišna infrastruktura izvedena je uporabom odgovarajućih programskih posredničkih razina. Te razine, uz osiguravanje visoke pouzdanosti, često osiguravaju i izvedbu zahtjevnijih funkcionalnosti kao što su *tiering*, deduplikacija, *erasure coding* i replikacija podataka.

U svim navedenim slučajevima spremišna infrastruktura treba podržavati upravljačko sučelje kroz koje su ostvarene funkcije dinamičkog dodjeljivanja/raspodjeljivanja podatkovnih resursa korisnicima. Uz pomoć ovih sučelja pojedine grupe korisnika mogu same upravljati spremišnim resursima koji su im dodijeljeni.

Spremišna infrastruktura treba omogućiti pohranu:

- diskova i slika virtualnih poslužitelja
- dugotrajnih podataka u objektno spremište tipa key-value (npr. Amazon S3, OpenStack Swift)
- dugotrajnih podataka na standardni dijeljeni datotečni sustav (npr. NFS, SMB)
- podataka koji su nastali iz zahtjevnih izračuna ili su dugotrajno potrebne za zahtjevno računanje (HPC ili HTC aplikacije), a za koje nije potreban visoko-učinkoviti raspodijeljeni datotečni sustav.

Kako za gore navedene oblike podataka nije uvijek ključna učinkovitost nego kapacitet spremišta, potrebno je osigurati ekonomične spremišne sustave koji, uz diskovne, koriste i tračne resurse – HSM (engl. *hierarchical storage management*).

E.3. MREŽNA INFRASTRUKTURA

Ethernet je uobičajena mrežna infrastruktura koja se koristi u poslovnoj (*Enterprise*) okolini podatkovnih centara, pa tako i u izvedbi infrastrukture računalstva u oblacima. S obzirom na visoku propusnost, Ethernet zadovoljava sve zahtjeve koji se postavljaju na HSC mrežnu infrastrukturu.

U okvirima HR-ZOO-a razmatramo/predviđamo u HSC okružju uporabu sljedećih tehnologija u inačicama koje će biti dobavljive u planiranom vremenskom periodu:

- Ethernet
    - 25 Gb/s
    - 50 Gb/s
    - 100 Gb/s.

Specifični zahtjevi HSC infrastrukture uvjetovani su mogućnošću dinamičkog osiguravanja (engl. *provisioning*) svih resursa, pa tako i mrežnih, korisnicima infrastrukture. Za izvedbu te funkcionalnosti nužna je izvedba programskog upravljanja mrežnom infrastrukturom – SDN (engl. *Software Defined Networking*) i NFV (engl. *Network Function Virtualisation*). Ova funkcionalnost podržana je od nekih izvedbi Ethernet mrežne infrastrukture kroz programsko upravljanje konfiguracijom preklopnika i integracijom s programskim preklopnicima koji su izvedeni u posredničkoj programskoj podršci za virtualizaciju, odnosno prebacivanjem mrežnih funkcija u virtualizacijsku okolinu. Programsko upravljanje mrežnom infrastrukturom implementirano je na L2/L3 razinama mrežnih protokola (Ethernet i TCP/IP).

Mrežna infrastruktura Ethernet uglavnom podržava *closs* i *mash* topologije. Kako je Ethernet od početka zamišljen kao mrežna infrastruktura u kojoj je uobičajeno da nije potrebno osigurati jednaku propusnost između bilo koja dva povezana poslužitelja, odnosno u kojoj postoji preplaćenost (engl. *oversubscription*), takva infrastruktura uobičajeno je razdijeljena na više razina.

Mrežna infrastruktura Ethernet zadovoljava djelomično i zahtjeve za pristup spremišnoj infrastrukturi uporabom protokola iSCSI ili se na nju u potpunosti oslanja u slučaju izvedbe HCI spremišne infrastrukture.

U slučaju potrebe povezivanja resursa koji se nalaze na međusobno udaljenim lokacijama (podatkovnim centrima) nužno je osigurati mogućnost povezivanja udaljenih mrežnih infrastruktura. Povezivanje se izvodi unajmljenim optičkim kapacitetima i upotrebom WDM-a (engl. *Wavelength Division Multiplexing*) ili odgovarajućim L2 (Ethernet) ili L3 (IP) TK kapacitetima.

E.4. POSREDNIČKI SUSTAVI

Za uspostavu sustava HSC ključni su sljedeći slojevi:

- virtualizacijski sloj koji se sastoji od hipervizora i omogućava upravljanje virtualnom infrastrukturom (poslužiteljima, mrežama, diskovima itd.)
- *cloud* posrednički sustav koji omogućava dinamičko upravljanje cjelokupnom infrastrukturom
- raspodijeljeni datotečni sustav za pohranu virtualnih diskova i datoteka.

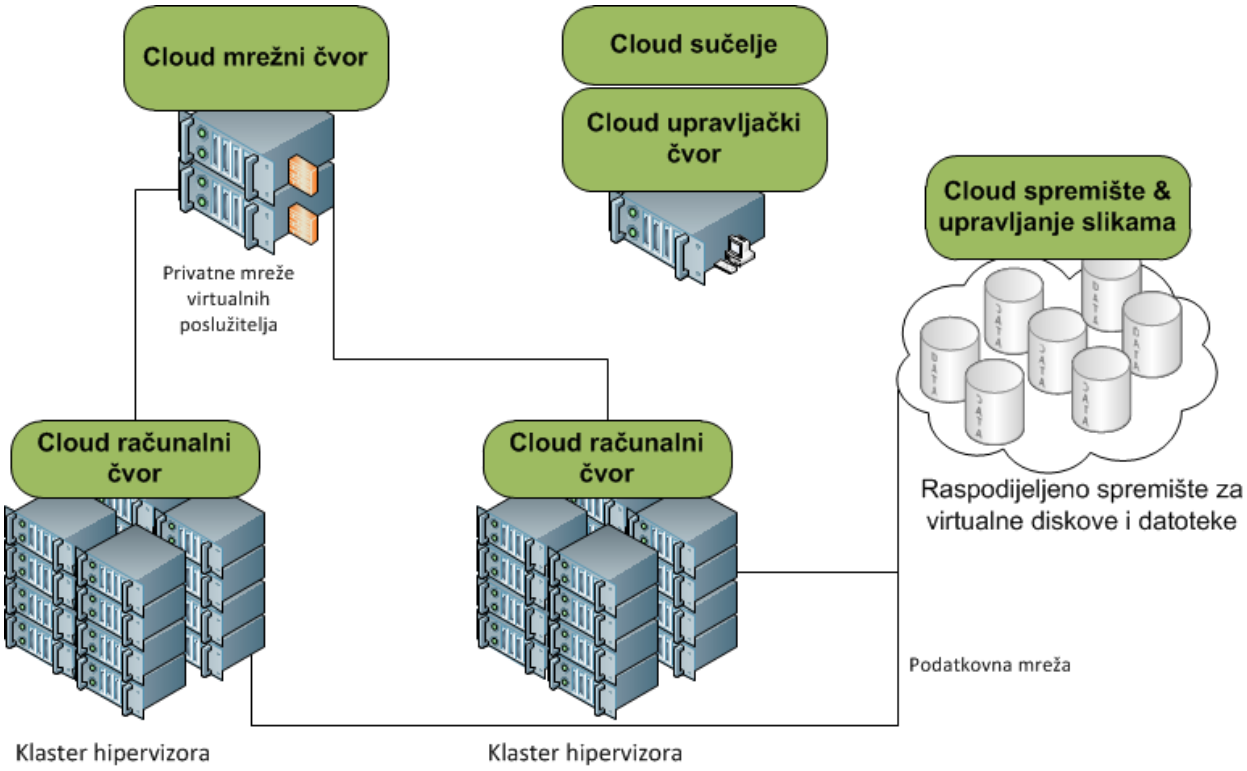
Najčešće korišteni virtualizacijski sustavi su:

- Citrix XenServer ([28])
- KVM ([29])
- VirtualBox ([30])
- VMware ([31])
- Xen ([32]).

Najznačajnija svojstva *cloud* posredničkih sustava su:

- mogućnost integracije s različitim autentikacijskim i autorizacijskim infrastrukturama (npr. digitalni certifikati, LDAP ili AD, integracija s različitim AAI sustavima, virtualne organizacije)
- korisničko i administracijsko komandno-linijsko, *web* i programsko sučelje; poželjna je podrška za standardna i *de facto* standardna programska sučelja (npr. Amazon EC2, S3, OCCI, CDMI itd.)
- korisnički prikaz povijesti i statistike o korištenju resursa
- podrška za različite virtualizacijske platforme (npr. Citix XenServer, KVM, VMware, Xen itd.)
- podrška za različite podatkovne protokole za pristup virtualnim diskovima i slikama (npr. SAN/iSCSI, NFS, SCP)
- upravljanje virtualnim mrežama i ostalim mrežnim komponentama (npr. usmjernik, vatrozid, NAT, load balancer)
- upravljanje slikama – mogućnost jednostavnog stvaranja i izmjene slika za virtualne poslužitelje te dodavanja već postojećih slika ili virtualnih poslužitelja u nekom od standardnih formata
- mogućnost sigurnosne pohrane virtualnih poslužitelja i njihovih slika (engl. *image*)
- mogućnost spremanja stanja virtualnih poslužitelja (engl. *snapshot*)
- mogućnost dugotrajne pohrane podataka u spremište tipa *key-value*
- visoka dostupnost virtualnih poslužitelja – mogućnost automatske selidbe virtualnih poslužitelja na dostupne hipervizore
- podrška za fizički udaljena sjedišta

  - mogućnost definiranja zona i politike dodjele resursa pojedinim grupama korisnika (npr. dodjela resursa po lokaciji korisnika)
  - mogućnost automatske selidbe cjelokupne virtualne infrastrukture na neko od dostupnih sjedišta (npr. VMware *stretched cluster*)
- *cloudbursting* – mogućnost povezivanja *cloud* sustava s drugim javno dostupnim *cloud* sustavima (npr. Amazon)
- *multi-tenancy* – uspostava međusobno izoliranih slika i virtualne infrastrukture za pojedine skupine korisnika; bitna je i mogućnost povezivanja sa sigurnosnom infrastrukturom kako bi se grupe korisnika automatski mapirale u *cloud* sustav.



Slika 3. Arhitektura *cloud* sustava

Tipična arhitektura *cloud* sustava prikazana je na Slici 3. Najčešće korišteni *cloud* sustavi su:

- CloudStack ([33])
- Eucalyptus ([34])
- OpenNebula ([35])
- OpenStack ([36])
- Synnefo ([37])
- VMware vCloud ([38]).

De facto standard za uspostavu *cloud* infrastrukture je sustav OpenStack. OpenStack sastoji se od više komponenata od kojih su najznačajnije:

- Keystone – središnja komponenta za sigurnost (autentikacija i autorizacija)
- Nova – komponenta za upravljanje virtualnim poslužiteljima
- Neutron – komponenta za upravljanje virtualnom mrežnom infrastrukturom
- Swift – komponenta za spremanje podataka tipa *key-value*
- Glance – komponenta za upravljanje slikama virtualnih poslužitelja.

Programska sučelja OpenStack komponenata postala su *de facto* standard pa su implementirana i u drugim *cloud* sustavima (npr. Synnefo). Brojne komercijalne tvrtke pružaju svoje distribucije OpenStacka sa specifičnim nadogradnjama (npr. dodatne mogućnosti integracije, orkestracija itd.). Neke od komercijalnih distribucija OpenStacka su:

- HP Helion ([39])
- IBM Cloud Manager ([40])
- Oracle OpenStack ([41])
- Red Hat Enterprise Linux OpenStack Platform ([42])
- VMware Integrated OpenStack ([43]).

F. SUSTAVI ZA ODRŽAVANJE INFRASTRUKTURE

Infrastruktura HR-ZOO-a sastoji se od velike količine računalnih resursa pa je potrebno osigurati sustave za učinkovito sistemsko održavanje resursa. U nastavku su opisane pojedine infrastrukture te su navedena iskustva koja Srce ima s konkretnim sustavima.

U okviru uspostave sustava za potrebe infrastrukture HR-ZOO-a bit će napravljena evaluacija pojedinih sustava te će biti odlučeno da li će se koristiti identični sustavi ili neki drugi. U slučaju pregovaračkog dijaloga odabir sustava ovisit će prvenstveno o ponuđačima.

F.1. SIGURNOSNA POHRANA PODATAKA

Sustav za sigurnosno pohranjivanje (engl. *backup*) podataka omogućava:

- automatsko periodičko sigurnosno pohranjivanje podataka ili cjelokupnih virtualnih strojeva
- sigurnosno pohranjivanje na zahtjev korisnika
- automatsko periodičko arhiviranje (trajno pohranjivanje) podataka ili cjelokupnih virtualnih strojeva
- vraćanje spremljenih podataka (engl. *restore*) sa sigurnosnih ili arhivskih kopija
- definiranje različitih politika sigurnosnog pohranjivanja za pojedine komponente infrastrukture.

Bitan uvjet za učinkovito i ekonomično skaliranje sigurnosne pohrane podataka je integracija s tračnim sustavima. Srce trenutačno koristi komercijalno rješenje IBM TSM Storage Manager ([44]).

U postupku evaluacije potrebno je odabrati rješenje koje će imati sva gore navedena svojstva te omogućiti integraciju s odabranim sustavima za visokoskalabilno računalstvo.

F.2. SIGURNOSNA INFRASTRUKTURA

Sigurnosna infrastruktura osigurava sljedeće funkcionalnosti:

- sustav za detekciju i prevenciju napada na poslužitelje i servise (engl. *Host Intrusion Detection System*, HIDS)
- sustav za detekciju i prevenciju napada na mrežnoj razini (engl. *Network Intrusion Detection System*, NIDS)
- sustav za središnje upravljanje sistemskim zapisima koji pohranjuje i omogućava pretraživanje sistemskih zapisa svih poslužitelja i servisa infrastrukture
- sustav za provjeru ranjivosti poslužitelja i servisa
- središnje sučelje za prikaz događaja iz sustava NIDS i HIDS.

Srce se u okviru sigurnosne infrastrukture koristi sljedećim sustavima:

- OSSEC ([45]) – sustav HIDS
- Snort ([46]) – sustav NIDS
- Nessus ([47]) – sustav za provjeru ranjivosti
- sustav za središnje pohranjivanje sistemskih zapisa rsyslog
- OSSIM ([48]) – središnje sučelje za prikaz događaja i generiranje periodičkih izvještaja.

S obzirom na vrijednost i količinu resursa u infrastrukturi HR-ZOO-a i sigurnosna infrastruktura mora biti pažljivo dizajnirana, posebno uzimajući u obzir skalabilnost. Uz samu infrastrukturu potrebno je osigurati tim sigurnosnih stručnjaka koji će moći brzo reagirati u slučaju sigurnosnih incidenata.

F.3. NADZORNA INFRASTRUKTURA

Nadzorna infrastruktura ključna je za osiguravanje pouzdanog rada infrastrukture te praćenje iskorištenosti resursa i planiranje budućih proširenja. Nadzorna infrastruktura osigurava sljedeće funkcionalnosti:

- nadzor stanja svih komponenata infrastrukture (servisa, poslužitelja, mrežnih uređaja, spremišta, napajanja, rashladnih uređaja itd.) te dojavljivanje nadležnim osobama u slučaju pojave problema
- nadzor učinkovitosti rada svih komponenata infrastrukture
- inventar svih komponenata infrastrukture s automatskim dohvatom podataka i mogućnošću ručnog unosa podataka
- *web*-sučelje za prikaz svih prikupljenih informacija i grafički prikaz učinkovitosti kroz vrijeme
- programsko sučelje za integraciju nadzorne infrastrukture s drugim sustavima (npr. sustavi za izvještavanje).

Srce se od 2008. godine za nadzor koristi složenim sustavom koji se sastoji od četiri sustava:

- Nagios ([49]) – nadzor stanja
- Zabbix ([50]) – nadzor učinkovitosti
- OCS ([51]) i GLPI ([52]) – inventar računalnih sredstava.

Uz ove sustave razvijen je određen broj lokalnih sustava koji su integrirani s navedena četiri sustava pomoću programskih sučelja. Iskustvo je pokazalo da su sustavi iznimno podesni i omogućavaju učinkovitu integraciju s drugim sustavima.

Za potrebe nadzora mrežne opreme Srca uspostavljen je nadzorni sustav Zenoss ([53]). Zenoss omogućava učinkovitu integraciju s različitom mrežnom opremom protokolom SNMP (engl. *Simple Network Management Protocol*).

Dodatno, Srce od 2006. sudjeluje na sustavu za nadzor dostupnosti i pouzdanosti servisa SAM koji se koristi za nadzor grid-sustava EGI. Sustav omogućava izradu mjesečnih

izvještaja te praćenje zadovoljavanja SLA-ova (engl. *Service Level Agreement*) od strane pojedinačnih grid-sjedišta.

F.4. INFRASTRUKTURA ZA AUTOMATSKU USPOSTAVU POSLUŽITELJA I UPRAVLJANJE KONFIGURACIJAMA

Učinkovito upravljanje velikim brojem poslužitelja i servisa te osiguravanje visoke skalabilnosti nije moguće ostvariti bez infrastrukture za automatsku uspostavu poslužitelja i upravljanje konfiguracijama. Upravljanje konfiguracijama osigurava konzistentnu konfiguraciju poslužitelja i servisa.

Srce za uspostavu i upravljanje računalnim klasterima (Isabella, CRO NGI), ali i standardnim servisima koristi sustave:

- Cobbler ([54]) – sustav za automatsku uspostavu poslužitelja s operacijskim sustavom RHEL Linux
- FAI ([55]) – sustav za automatsku uspostavu poslužitelja s operacijskim sustavom Debian Linux
- Puppet ([56]) – sustav za upravljanje konfiguracijama poslužitelja s operacijskim sustavima Linux.

Pojedine komponente infrastrukture HR-ZOO-a sadržavat će vlastite mehanizme za automatsku uspostavu (npr. sustavi za računalstvo u oblacima, sustavi za uspostavu računalnih klastera). U procesu uspostave cjelokupne infrastrukture napravit će se evaluacija svih sustava. Dodatno, bit će procijenjeno postoji li potreba za jedinstvenom infrastrukturom za automatsku uspostavu poslužitelja i upravljanje konfiguracijama.

G. KORISNIČKI SUSTAVI

G.1. PRAĆENJE KORIŠTENJA I IZVJEŠTAVANJE

Pojedine komponente infrastrukture HR-ZOO-a moraju osigurati praćenje korištenja resursa (engl. *accounting*). Primjerice, u slučaju računalstva visokih performansi praćenje korištenja obavlja sustav za upravljanje poslovima, a kod računalstva u oblacima postoje posebne komponente koje prate iskorištenje resursa (npr. OpenStack Ceilometer). Uz praćenje korištenja resursa na pojedinim komponentama potrebno je uspostaviti jedinstvenu točku gdje će korisnici i udomitelji resursa moći dobiti informaciju o načinu korištenja resursa. Za uspostavu te točke bit će potrebno napraviti integraciju s pojedinim komponentama te nadzornom i sigurnosnom infrastrukturom.

Sustav za praćenje korištenja resursa treba imati mogućnost izračuna troška korištenja (engl. *billing*) primjenom različitih algoritama (npr. potrošnja procesorskog vremena, memorije, spremišnog prostora, električne energije itd.). Ova funkcionalnost je nužna za uspostavu modela naplate po korištenju (engl. *pay-per-use*).

Radi što boljeg informiranja korisnika i partnera potrebno je definirati periodička izvješća (npr. mjesečna, kvartalna i godišnja). Srce već ima dobro definiranu praksu mjesečnih periodičkih izvješća. Kako bi se olakšao postupak izrade, potrebno je razviti sustav za automatsko generiranje izvještaja.

G.2. AUTENTIKACIJSKA I AUTORIZACIJSKA INFRASTRUKTURA

U cilju olakšavanja pristupa resursima bit će potrebno iskoristiti postojeće autentikacijske i autorizacijske infrastrukture, poput sustava AAI@EduHr, sustava eduGAIN te certifikacijskih autoriteta SRCE CA i Terena TCS. Dodatno, korisnicima će se omogućiti korištenje AAI@EduHr virtualnih organizacija ([57]) za potrebe stvaranja skupina korisnika između više različitih ustanova. Za odabir posredničkih sustava nužan uvjet je integracija s gore navedenim autentikacijskim i autorizacijskim sustavima.

Integracija s postojećim sustavima omogućit će veću razinu automatizacije cjelokupne infrastrukture HR-ZOO-a. Tako će, primjerice, prema atributima u sustavu AAI@EduHr određeni korisnici imati ovlasti stvaranja virtualnih strojeva ili infrastrukture na HSC resursima, a pravo korištenja bit će određeno pripadnošću virtualnoj organizaciji.

G.3. SLUŽBA POMOĆI I TEHNIČKA PODRŠKA

Infrastruktura HR-ZOO-a sastojat će se od različitih komponenata te će posluživati različite korisničke skupine. Stoga je potrebno osigurati sustav za službu pomoći koji će omogućiti razdvajanje timova koji će održavati pojedine komponente, odnosno pružati podršku određenim korisničkim skupinama. Također, potrebno je osigurati prvu razinu podrške koja će rješavati jednostavnije probleme, odnosno prosljeđivati upite odgovarajućim timovima.

Ovisno o kritičnosti komponenata infrastrukture bit će potrebno definirati različite procedure reagiranja na upite i probleme u radu komponenata. Primjerice, u slučaju problema u

sustavu za visokoskalabilno računalstvo bit će potrebno reagirati u kraćem roku nego u slučaju problema s čvorovima računalnih klastera.

G.4. KORISNIČKA SUČELJA

Svi odabrani sustavi trebaju imati mogućnost pristupa putem korisničkih *web*-portala. *Web*-portali trebaju omogućiti osnovne funkcionalnosti:

- pokretanje korisničkih poslova i upravljanje korisničkim poslovima na sustavima HPC i HTC
- upravljanje virtualnim poslužiteljima i infrastrukturom na sustavu HSC.

Uz osnovne funkcionalnosti pojedinih sustava korisnicima se trebaju ponuditi prikazi iz sustava opisanih u poglavljima G i H: sigurnosna i nadzorna infrastruktura te praćenje korištenja infrastrukture.

Za potrebe zahtjevnog znanstvenog računanja postoje znanstveni *web*-portali (engl. *science gateway*) koji omogućavaju jednostavno pokretanje poslova na različitim računalnim infrastrukturama. U okviru HR-ZOO-a bit će potrebno uspostaviti ovakav oblik *web*-portala za najzastupljenije znanstvene zajednice (npr. za istraživače iz znanstvenog polja kemije ili istraživače iz znanstvenog polja psihologije) te omogućiti korisnicima da sami pokreću ovakve portale na sustavu HSC. Primjer sustava za uspostavu znanstvenih *web*-portala je gUSE (engl. *grid and cloud user support environment*) ([58]).

H. KATALOG USLUGA

U nastavku je pregled (katalog) usluga koje će biti dostupne korisnicima infrastrukture HR-ZOO-a.

Uslugu naprednog računalstva čine dvije komponente: usluga HPC i usluga HTC. Za odabrane velike zajednice korisnika (npr. kemičari, biolozi, matematičari) bit će pripremljeni znanstveni *web*-portali koji će omogućiti jednostavno izvođenje najčešće korištenih aplikacija.

Usluga HPC pružat će računalne resurse s tehnologijom ubrzivača (GPU/MIC) za korisnike koji imaju potrebu za naprednim računanjem. Korisnici će pokretati aplikacije putem *web*-sučelja ili izravnim pristupom na pristupno računalo protokolom SSH (engl. *Secure Shell*). Resursi usluge datotečnog i objektnog spremišta bit će korišteni za pohranu dugotrajnih podataka. U okviru HPC usluge korisnicima će biti dostupno minimalno:

- ukupno 10.000 procesorskih (CPU) jezgri
- 8 GB radne memorije po procesorskoj jezgri
- mogućnost spajanja memorijskih resursa više poslužitelja do 16 TB radne memorije
- mreža propusnosti 200 Gb/s i kašnjenje <100 ns (*port-to-port*)
- 2 TB višerazinski programski upravljane/definirane memorije po radnom čvoru
- odnos radne memorije (DRAM) i memorije više razine (npr. 3D XPoint) > 1 : 4
- 2 TB spremišnog SSD prostora po radnom čvoru
- 2 PB spremišta visokih performansi.

Usluga HTC omogućit će virtualne računalne resurse s grafičkim procesorima za korisnike koji imaju potrebu za naprednim računanjem, skaliranjem aplikacija (npr. memorijske baze podataka, računalno zahtjevni aplikacijski serveri) ili grafičkim renderiranjem. U okviru usluge HTC korisnicima će biti dostupan skup predložaka virtualnih poslužitelja te će imati mogućnost stvaranja vlastitih predložaka. U okviru HTC usluge korisnicima će biti dostupno minimalno:

- 20.000 procesorskih jezgri
- 4 GB radne memorije po jezgri
- Ukupno 1 TB višerazinski programski upravljane/definirane memorije po radnom čvoru
- Odnos radne memorije (DRAM) i memorije više razine (npr. 3D XPoint) > 1 : 4
- 4 PB spremišta za privremenu pohranu podataka (lokalni *scratch*).

Usluga HSC omogućit će korisnicima pokretanje visokodostupnih virtualnih poslužitelja. Visoka dostupnost bit će ostvarena pomoću *cloud* posredničkog sustava koji će u slučaju ispada jednog HR ZOO sjedišta automatski pokretati virtualne poslužitelje na preostalim sjedištima. Korisnici će uspostavu i upravljanje virtualnim poslužiteljima i pripadnim resursima obavljati putem cloud *web*-sučelja. Usluga će uključiti i spremanje sigurnosnih kopija virtualnih poslužitelja.

U okviru usluge HSC korisnicima će biti dostupno najmanje:

- 300 visokodostupnih virtualnih poslužitelja sljedećih karakteristika:
    - 8 virtualnih procesorskih jezgri po poslužitelju
    - 32 GB radne memorije po poslužitelju.
- 1.500 virtualnih poslužitelja sljedećih karakteristika:
    - 4 virtualne procesorske jezgre po poslužitelju
    - 16 GB radne memorije po poslužitelju
- 100 TB spremišnog prostora za virtualne diskove

te po potrebi:

- mrežni servisi
    - logička segmentacija mreže
    - mrežne usluge DHCP, NAT
    - usmjeravanje prometa
    - upravljanje sigurnošću
    - VPN.

Usluga podatkovnog spremišta pružat će korisnicima mogućnost pohrane podataka u obliku datoteka i objekata. Za ovu uslugu koristit će se raspodijeljeno spremište za datoteke i objekte te tračni resursi. Spremišna infrastruktura osigurava visoku dostupnost pohranjenih podataka.

Karakteristike datotečnog spremišta:

- Podrška za standardni protokole kao što su NFS ili SMB.
- Spremište će biti dostupno prvenstveno u okviru HPC usluge i s virtualnih poslužitelja na *cloud* uslugama. Na zahtjev korisnika bit će moguće ostvariti pristup datotečnom spremištu s resursa korisnika.

Karakteristike objektnog spremišta:

- Pristup će biti moguć raširenim protokolima kao što su Amazon S3 ili OpenStack Swift.
- Pristup podacima bit će moguć iz cijelog svijeta.

U okviru usluge podatkovnog spremišta korisnicima će biti dostupno minimalno:

- 10 PB neto spremišnog prostora
- 20 PB neto spremišnog prostora na tračnim resursima.

Usluga udomljavanja omogućit će korisnicima instalaciju vlastiti računalnih, spremišnih i mrežnih komponenata u podatkovne centre infrastrukture HR ZOO-a. Udomljavanje će biti omogućeno isključivo za resurse koji ne mogu biti ostvareni na drugim uslugama HR-ZOO-a.

I. RASPODJELA RESURSA

Polazeći od analize i rasprave provedene u prethodnim poglavljima ovoga dokumenta u nastavku se daje sažet prijedlog raspodjele resursa pojedinih slojeva infrastruktura unutar HR-ZOO-a koji trebaju zadovoljiti prethodno elaborirane zahtjeve.

Slojevi računalne infrastrukture unutar HR-ZOO-a su sljedeći:

- HPC resursi
- *cloud* resursi za potrebe HTC
- visoko dostupni *cloud* resursi za potrebe HSC.

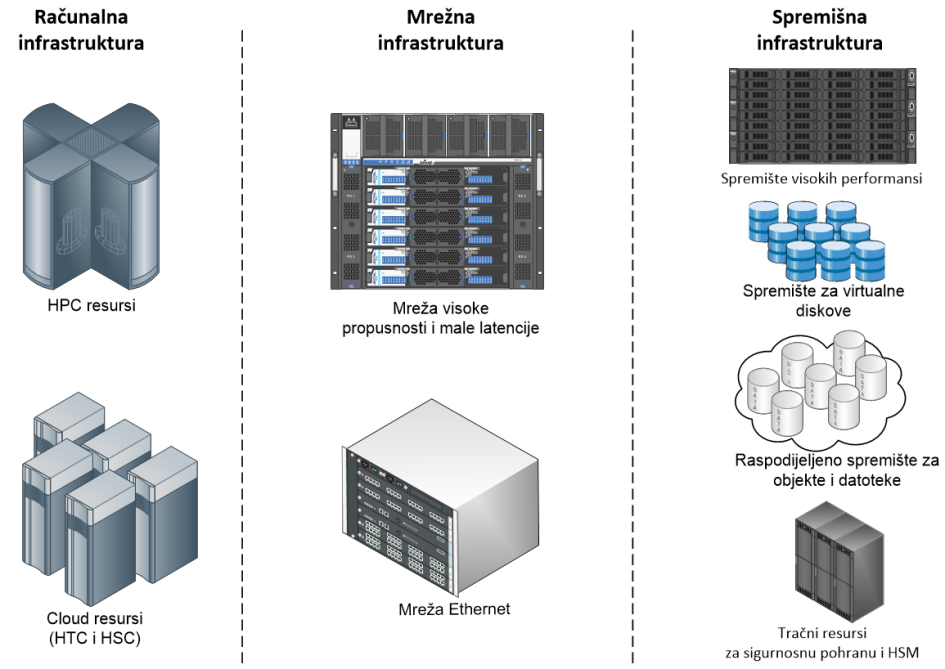
Cloud resursi trebaju koristiti zajedničku posredničku infrastrukturu. Računalni resursi za *cloud* bit će podijeljeni u dvije skupine ovisno o tome da li će se na njima izvoditi HTC ili HSC aplikacije.

Slojevi mrežne infrastrukture podatkovnih sjedišta unutar HR-ZOO-a su sljedeći:

- Pasivna infrastruktura (sastavni dio projektne dokumentacije pojedinih podatkovnih sjedišta)
- Preklopnička infrastruktura:
  - Mreža visoke propusnosti i male latencije za potrebe HPC
  - Mreža Ethernet za potrebe HSC i HTC.
- Usmjerivačka infrastruktura.

Slojevi spremišne infrastrukture unutar HR-ZOO-a su sljedeći:

- Spremište visokih performansi za HPC
- Raspodijeljeno spremište za datoteke i objekte
- Spremište za virtualne diskove
- Tračni resursi za sigurnosnu pohranu i HSM.



Slika 4. Slojevi infrastrukture HR-ZOO-a

Raspodjela resursa pojedinih slojeva po sjedištima prikazana je u tablicama 1., 2. i 3.

Tablica 1. Raspodjela resursa računalne infrastrukture po sjedištima

Sjedište	UNIOS	UNIRI	UNIST	Srce	Borongaj	SK / IRB
HPC resursi					X	
HTC resursi	X	X	X		X	
***Cloud* resursi**	X	X	X		X	

Tablica 2. Raspodjela resursa mrežne infrastrukture po sjedištima

Sjedište	UNIOS	UNIRI	UNIST	Srce	Borongaj	SK / IRB
Mreža visoke propusnosti i male latencije					X	
Mreža Ethernet	X	X	X	X	X	X

Tablica 3. Raspodjela resursa spremišne infrastrukture po sjedištima

Sjedište	UNIOS	UNIRI	UNIST	Srce	Borongaj	SK / IRB
Učinkoviti raspodijeljeni datotečni sustav za HPC					X	
Raspodijeljeno spremište za datoteke i objekte	X	X	X		X	
Spremište za virtualne diskove	X	X	X		X	
Tračni resursi za sigurnosnu pohranu i HSM			X			X

REFERENCE

[1] e-Infrastructure Reflection Group (e-IRG), http://www.e-irg.eu

[2] AAI@EduHr, http://www.aaiedu.hr

[3] Računalni klaster Isabella, http://isabella.srce.hr

[4] Hrvatska nacionalna grid-infrastruktura, CRO NGI, http://www.cro-ngi.hr

[5] *Cloud* usluga Virtual Private Server, VPS, http://www.srce.unizg.hr/vps/

[6] *Cloud* usluga Virtual Computing Lab, VCL, http://www.srce.unizg.hr/vcl/

[7] MojOblak, GSS, http://www.srce.unizg.hr/mojoblak/

[8] ScaleMP, http://www.scalemp.com

[9] SGI UV, http://www.sgi.com/products/servers/uv/

[10] HTCondor, http://research.cs.wisc.edu/htcondor/

[11] IBM Platform LSF, http://www-03.ibm.com/systems/platformcomputing/products/lsf/

[12] PBS Professional, http://www.pbsworks.com/PBSProduct.aspx?n=PBS-Professional&c=Overview-and-Capabilities

[13] Slurm, https://computing.llnl.gov/linux/slurm/

[14] Son of Grid Engine, https://arc.liv.ac.uk/trac/SGE

[15] Univa Grid Engine, http://www.univa.com/products/grid-engine.php

[16] BeeGFS, http://www.beegfs.com

[17] GlusterFS, http://www.gluster.org/

[18] IBM GPFS, http://www-03.ibm.com/software/products/en/software/

[19] Lustre, https://wiki.hpdd.intel.com/display/PUB/HPDD+Wiki+Front+Page

[20] MapReduce, http://en.wikipedia.org/wiki/MapReduce

[21] Monte Carlo, http://en.wikipedia.org/wiki/Monte_Carlo_method

[22] ARC, http://www.nordugrid.org/arc/

[23] CREAM, http://grid.pd.infn.it/cream/

[24] Globus Toolkit, http://toolkit.globus.org/toolkit/

[25] UNICORE, https://www.unicore.eu/

[26] EGI FedCloud, https://wiki.egi.eu/wiki/Fedcloud-tf:Main

[27] Xsigo/Oracle FabricManager, http://www.oracle.com/us/corporate/acquisitions/xsigo/index.html

[28] Citrix XenServer, http://www.citrix.com/products/xenserver/overview.html

[29] KVM, http://www.linux-kvm.org/page/Main_Page

[30] VirtualBox, https://www.virtualbox.org/

[31] VMware, http://www.vmware.com

[32] Xen, http://www.xenproject.org/

[33] CloudStack, http://cloudstack.org

[34] Eucalyptus, http://open.eucalyptus.com

[35] OpenNebula, http://opennebula.org

[36] OpenStack, http://www.openstack.org

[37] Synnefo, https://www.synnefo.org/

[38] VMware vCloud, http://www.vmware.com/products/vcloud/overview.html

[39] HP Helion, http://www8.hp.com/us/en/cloud/hphelion-openstack-overview.html

[40] IBM Cloud Manager, http://www.ibm.com/developerworks/servicemanagement/cvm/sce/

[41] Oracle OpenStack, http://www.oracle.com/us/technologies/linux/openstack/overview/index.html

[42] Red Hat Enterprise Linux OpenStack Platform, http://www.redhat.com/en/technologies/linux-platforms/openstack-platform

[43] VMware Integrated OpenStack, http://www.vmware.com/products/openstack

[44] IBM Tivoli Storage Manager, http://www.ibm.com/software/products/en/tivostormana/

[45] OSSEC, http://www.ossec.net

[46] Snort, http://www.snort.org

[47] Nessus, http://www.tenable.com/products/nessus/

[48] OSSIM, http://www.ossim.org

[49] Nagios, http://www.nagios.org

[50] Zabbix, http://www.zabbix.com

[51] OCS, http://www.ocsinventory-ng.org

[52] GLPI, http://www.glpi-project.org

[53] Zenoss, http://www.zenoss.com

[54] Cobbler, http://www.cobblerd.org

[55] FAI, http://www.fai-project.org

[56] Puppet, http://www.puppetlabs.com

[57] AAI@EduHr virtualne organizacije, http://www.aaiedu.hr/virtualne_organizacije.html

[58] gUSE, http://guse.hu

srce