

FIȘA DISCIPLINEI

1. Date despre program

1.1 Instituția de învățământ superior	Universitatea Tehnică din Cluj-Napoca
1.2 Facultatea	Automatică și Calculatoare
1.3 Departamentul	Calculatoare
1.4 Domeniul de studii	Calculatoare și Tehnologia Informației
1.5 Ciclul de studii	Master
1.6 Programul de studii / Calificarea	Inginerie Software/ Master
1.7 Forma de învățământ	IF – învățământ cu frecvență
1.8 Codul disciplinei	10.

2. Date despre disciplină

2.1 Denumirea disciplinei	Procesarea fluxurilor de date în sisteme distribuite pe scară largă				
2.2 Titularii de curs	Conf. dr. ing. Ionut Anghel – ionut.anghel@cs.utcluj.ro				
2.3 Titularul/Titularii activităților de seminar/laborator/proiect	Conf. dr. ing. Ionut Anghel – ionut.anghel@cs.utcluj.ro Prof. dr. ing. Tudor Cioara – tudor.cioara@cs.utcluj.ro				
2.4 Anul de studiu	1	2.5 Semestrul	2	2.6 Tipul de evaluare (E – examen, C – colocviu, V – verificare)	E
2.7 Regimul disciplinei	DA – de aprofundare, DS – de sinteză, DC – complementară				DS
	DI – Impusă, DOp – opțională, DFac – facultativă				DI

3. Timpul total estimat

3.1 Număr de ore pe săptămână	3	din care:	Curs	2	Seminar	-	Laborator	1	Proiect	-
3.2 Număr de ore pe semestru	42	din care:	Curs	28	Seminar	-	Laborator	14	Proiect	-
3.3 Distribuția fondului de timp (ore pe semestru) pentru:										
(a) Studiul după manual, suport de curs, bibliografie și notițe										20
(b) Documentare suplimentară în bibliotecă, pe platforme electronice de specialitate și pe teren										10
(c) Pregătire seminarii / laboratoare, teme, referate, portofolii și eseuri										10
(d) Tutoriat										16
(e) Examinări										2
(f) Alte activități:										-
3.4 Total ore studiu individual (suma (3.3(a)...3.3(f)))										58
3.5 Total ore pe semestru (3.2+3.4)										100
3.6 Numărul de credite										4

4. Precondiții (acolo unde este cazul)

4.1 de curriculum	n/a
4.2 de competențe	Operarea cu fundamente științifice, ingineresti și ale informaticii

5. Condiții (acolo unde este cazul)

5.1. de desfășurare a cursului	Proiector, calculator, tabla, Internet Platforma MS Teams ca și mediul de predare online Site-ul web al disciplinei pentru materiale de curs
5.2. de desfășurare a seminarului / laboratorului / proiectului	Calculatoare, software specific Platforma MS Teams ca și mediul de predare online Site-ul web al disciplinei pentru materiale de laborator

6. Competențele specifice acumulate

6.1 Competențe profesionale	C4 - Integrarea contextuală și integritatea sistemelor software complexe <ul style="list-style-type: none"> • C4.1 - Demonstrarea cunoașterii și înțelegerii elementelor de interoperabilitate și integrare specifice sistemelor software, luate atât în ansamblu cât și pe module • C4.2 - Folosirea unor cunoștințe interdisciplinare pentru adaptarea sistemelor software complexe în raport cu cerințele dinamice ale domeniului de aplicații
-----------------------------	--

	<ul style="list-style-type: none"> • C4.3 - Utilizarea combinată a unor principii și metode clasice și originale pentru integrarea componentelor unor sisteme de calcul complexe • C4.4 - Folosirea standardelor de calitate, siguranță și securitate în prelucrarea informațiilor și în integrarea sistemelor software complexe. • C4.5 - Realizarea de proiecte interdisciplinare, incluzând identificarea și analiza problemei, elaborarea specificațiilor, proiectarea software, implementarea testarea funcțională și evaluarea criteriilor de calitate, securitate și de performanță specifice, precum și validarea sistemului software integrat. <p>C5 - Îmbinarea creativă a cunoștințelor multidisciplinare din domeniul științei calculatoarelor și tehnologiei informației în vederea cercetării, specificării, proiectării, optimizării, implementării, testării și evaluării de teorii, algoritmi, tehnici, metode și metodologii originale specifice sistemelor software complexe.</p> <ul style="list-style-type: none"> • C5.1 - Demonstrarea cunoașterii metodologiei de cercetare, proiectare, implementare, optimizare și testare a sistemelor de calcul complexe • C5.2 - Demonstrarea capacității de a analiza și interpreta situații noi prin prisma cunoștințelor fundamentale din domeniul calculatoarelor și tehnologiei informației • C5.3 - Îmbinarea creativă, bazată pe descoperirea de legături semantice și funcționale noi, a diferite principii de proiectare moderne din domeniul calculatoarelor și tehnologiei informației pentru rezolvarea unor probleme de optimizare • C5.4 - Fundamentarea activității de cercetare și proiectare inovativă din domeniul calculatoarelor pe criterii corecte de evaluare • C5.5 - Realizarea de activități de cercetare cu finalitate practică demonstrată prin prototipuri software funcționale.
6.2 Competențe transversale	N/A

7. Obiectivele disciplinei

7.1 Obiectivul general al disciplinei	Principalul obiectiv al acestei discipline este de introduce concepte de baza si tehnologii specifice integrarii si utilizarii stream-urilor mari de date ce sunt achizitionati din surse de date eterogene (dispozitive IoT, web, social networking, etc.) in aplicatii software. Astfel se urmareste asigurarea bagajului de cunostinte ce poate fi folosit la evaluarea critica a solutiilor de dezvoltare existente, precum si la dezvoltarea si integrare de noi solutii in aplicatii existente sau nou dezvoltate.
7.2 Obiectivele specifice	Pentru atingerea obiectivului general, studentii vor studia arhitectura Lamba si tehnologiile specifice pentru fiecare element al arhitecturii: modele de date, managementul eficient al fluxurilor de date si procesarea lor atat offline cat si in real time pentru a oferi diferite vederi utilizatorilor aplicatiilor web.

8. Conținuturi

8.1 Curs	Nr.ore	Metode de predare	Observații
Introducere, overview curs, motivație, probleme legate procesarea datelor in timp real	2	Expunere la tablă, prezentare cu video-proiectorul si/sau prezentari online folosind platforma MS Teams Site-ul web al cursului Discuții	Nu este cazul.
Arhitecturi, Cerinte nonfunctionale, limbaje de programare	2		
Modele de date scalabile	2		
Stari distribuite, Consens si Coordonare	2		
Managmentul Fluxurilor de Date Distribuite	2		
Cozi si fluxuri de date	2		
Stocarea datelor distribuite	2		
Batch Layer: Map Reduce	2		
Apache Spark Streaming	2		
Apache Spark Machine Learning	2		
Serving Layer	2		
Speed Layer – vederi in timp real	2		

Speed Layer – procesarea streamurilor de micro – batch-urilor	2		
Concluzii, directii de dezvoltare in cercetare	2		
Bibliografie (<i>bibliografia minimală a disciplinei conținând cel puțin o lucrare bibliografică de referință a disciplinei, care există la dispoziția studenților într-un număr de exemplare corespunzător</i>)			
1. Byron Ellis, Real Time Analytics Techniques to Analyze and Visualize Streaming Data-Wiley (2014)			
2. Nathan Marz and James Warren, Big Data, Principles and best practices of scalable realtime data systems, April 2015, ISBN 9781617290343, Manning Publications			
3. Holden Karau, Andy Konwinski, Patrick Wendell, Matei Zaharia - Learning Spark Lightning-Fast Big Data Analysis-O'Reilly Media			
4. Site-ul web al disciplinei			
8.2 Aplicații (seminar/laborator/proiect)*	Nr.ore	Metode de predare	Observații
Mapare teme studenti, Discutarea temelor de prezentare primite de studenți (utilizarea tehnologiilor si uneltelor software existente in proiectarea si implementarea aplicatiilor bazate pe procesarea stream-urilor de date)	2	Expuneri la tablă, referate tematice elaborate ca urmare a cercetarii	Nu este cazul.
Arhitectura Lambda si Apache Trift	2	bibliografiei,	
Apache ZooKeeper, Apache Kafka, Apache Flume	2	prezentare cu video-proiectorul, explicații	
Hadoop file system, MongoDB si Cassandra	2	suplimentare, discutii	
Apache Spark Batch layer si Serving Layer	2	fata in fata si/sau in	
Speed Layer - Micro Batch Stream	2	mediul online folosind	
Finalizare laborator, discutii proiectare aplicatii distribuite complexe	2	platforma MS Teams	
Bibliografie (<i>bibliografia minimală pentru aplicații conținând cel puțin o lucrare bibliografică de referință a disciplinei care există la dispoziția studenților într-un număr de exemplare corespunzător</i>)			
1. Byron Ellis, Real Time Analytics Techniques to Analyze and Visualize Streaming Data-Wiley (2014)			
2. Nathan Marz and James Warren, Big Data, Principles and best practices of scalable realtime data systems, April 2015, ISBN 9781617290343, Manning Publications			
3. Holden Karau, Andy Konwinski, Patrick Wendell, Matei Zaharia - Learning Spark Lightning-Fast Big Data Analysis-O'Reilly Media			
4. Site-ul web al disciplinei			

*Se vor preciza, după caz: tematica seminariilor, lucrările de laborator, tematica și etapele proiectului.

9. Coroborarea conținuturilor disciplinei cu așteptările reprezentanților comunității epistemice, asociațiilor profesionale și angajatorilor reprezentativi din domeniul aferent programului

In cadrul studiilor de licenta, studentii aprofundeaza proiectarea a multor clase de sisteme printre care se numara sistemele distribuite, sistemele paralele etc. Aceasta disciplina isi propune sa completeze varietatea de sisteme acoperite in timpul studiilor de licenta, propunand studierea unei clase de sisteme distribuite la scala mare si aplicatii software complexe care devine din ce in ce mai prezenta atat in aria de cercetare cat si in domeniul comercial.

10. Evaluare

Tip activitate	Criterii de evaluare	Metode de evaluare	Pondere din nota finală
Curs	Abilitatea de a propune solutii unor probleme specifice domeniului. Prezență, (inter)activitate în timpul orelor de curs.	Examen scris fata in fata sau online folosind platforma MS Teams	51%
Seminar			
Laborator	Abilitatea de a identifica probleme și de a utiliza tehnologiile existente în domeniu in dezvoltarea de aplicatii web complexe. Prezență, (inter)activitate în timpul orelor.	Evaluare pe parcursul semestrului fata in fata si/sau online folosind platforma MS Teams	49%
Proiect			

Standard minim de performanta:

Înțelegerea conceptelor de baza ale domeniului și demonstrarea abilității de a utiliza noile tehnologii studiate.

Calcul nota disciplina: 49% (laborator) + 51% (examen final)

Conditii de participare la examenul final: Nota Laborator ≥ 5; Elaborarea unui Referat de Cercetare

Conditii de promovare: Nota Examen final ≥ 5

Data completării:	Titulari	Titlu Prenume NUME	Semnătura
_____	Curs	Conf.dr.ing. Ionut Anghel	
	Aplicații	Conf.dr.ing. Ionut Anghel	
		Prof. dr. ing. Tudor Cioara	

Data avizării în Consiliul Departamentului Calculatoare _____	Director Departament Prof.dr.ing. Rodica Potolea
Data aprobării în Consiliul Facultății de Automatică și Calculatoare _____	Decan Prof.dr.ing. Liviu Miclea