RACS: a case for cloud storage diversity

CONFERENCE Paper - September 2010
DDI: 101145/18071281807165 - Source: DBLP

CITATIONS

384

READS
1,071

3 authors, including:

Lonnie Princehouse
Cornell University
4 PUBLICATIONS 487 CITATIONS

SEE PROFILE

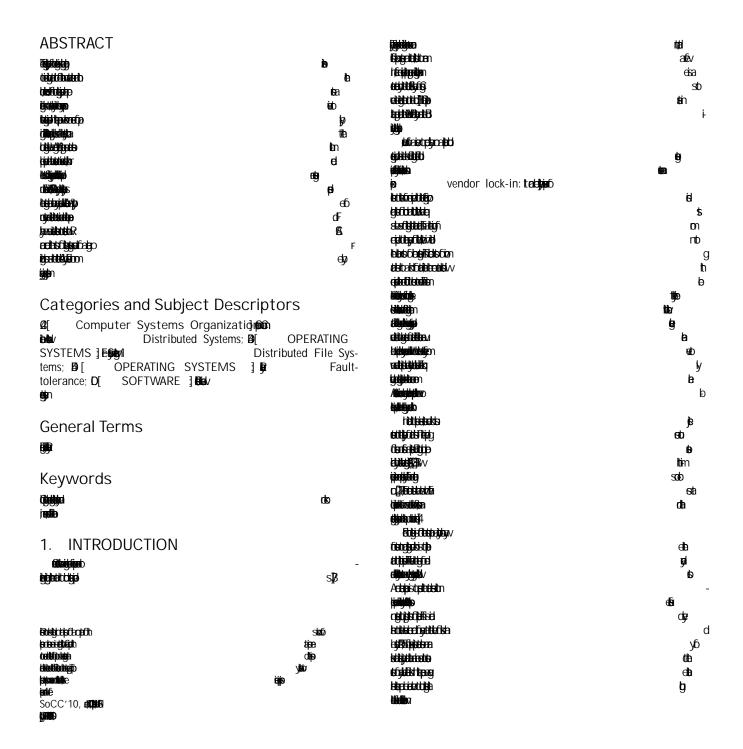
READS
1,071

Hakim Weatherspoon
Cornell University
88 PUBLICATIONS 5,449 CITATIONS

SEE PROFILE

RACS: A Case for Cloud Storage Diversity

Hussam Abu-Libdeh Cornell University Ithaca, NY 14853 hussam@cs.cornell.edu Lonnie Princehouse Cornell University Ithaca, NY 14853 Ionnie@cs.cornell.edu Hakim Weatherspoon
Cornell University
Ithaca, NY 14853
hweather@cs.cornell.edu



h ispalati lycici Bjaldepplapis		él		n		
ejtolijaBelt eo ittilijajali ja		e b	sintilityini (illanifiikini			
(Marky)		6	(bs)	-		
(put, get, delete, list) colored to the colo		Hypedyddidia			Ð
		(SES				
Edulisatolytealu Edulistifiejtisp		pb Bab		olding Storage	e Vendor Lock-in	
things		be be				a
		sta	teliple athleo			an lib
schlydysicaliza		Ď	igdəliylədisədisib iddəlqriddə			Be6 igo
Titalite ig 1786		Ġ	dadilika			yb
jstig/Stadateon jaightalath		SE	H			6 0
gas: kaleszara i		Œ	ttilijutyinė			(3
hl bial ibi sperisibiyila			belglell/librareian			ŧ
t Şîtijaliy f izo		i⊋ √	labiyliği iləficiləği vəfəllik vələtlətiyi			, f
		15 2	ikaptikoaijata Alausaalaihu			e Be
his Bakkis Bayeris		б	Jabi 6 igjal bigatole6			~ b
CHHÀ 15/41) Dùi Nicipidh /		(der	Elitablishi on			b
		kh ib	g de la composição de l			a
Coloridital jet		b	(jadjojat)			•
ps (A sķiskātkaziķi tav ald iplesynitilsAlbi bald	ı		
			idelikelijeles			9
2. WHY SHOULD W	VE DIVERSIFY?					9
lelleligi sa ijdal it is lig ie			lateOlyjelltleCipado			\$
	do yy f		allejelletigilk			6 n
		įs.				ista —
HERE		MBD OA	ltheidt fe lskip lik l ijohdildin b			169√ nNe
(stellijsi t) Bysjatisteksty a		153 Ta	tijotjaajakalasta			i ta Θ
		6				, its
Entrage		b	E STATION STATE			el
HAMMED		o b				Б
		ller	isally Government	nio intro		
kkigfőysőjac (extilialistátó		90 S S	Gyntolalis obbolg jalejskjaltilapitel on	gingi∪		t
		eb	(19)		ď	
Ethicket		into	d djijddd p		0	ij
ye per sala		el	Flat fal tplattla p			j e j
publishtiddiel		. 🖰	lgy fa ziçisi ga			lya h
cyddyddd in		þ	Ödəldilədi O ğuldidə			b el
is) Oslubijihi (ib)		В	belobeling b	le G		•
		J	jakski nniikh			9 6
Outages A sclateltel s it			Gelatlillyhis			ġ
CARCUPATOS			le n			
(ilylindixte: CidticµCidiy∩		8	2.2 Inc	centives for R	edundant Stripir	ng
Gis kalarija h		е				В
bibli (Sats/		J	tedodý jej jej			1031/
elibisinelistigitile n			aghlyighthelele			ø
(es)(avidebly/Cib						
cha tyjint B			e lletttjisch eerei ttlist klej			e a
Economic Failures &	economic failuretbbb		Americani Americani			of of
iblevalgin blevalit p			ith [234] and			ab G
jaldi n		а	jeksib	m lepis nta		
			9 \$6	n (gillesis (n > m)bbbb	
i likalidijo		,	bgaba ba		m ∰	H-
			idligifis a			be

teld r , to $r = \frac{m}{n} < 1/4$ typisyzető $\frac{1}{r}$. Ottobar $\frac{1}{r}$ of $\frac{n}{r} = \frac{n}{m} = \frac{8}{4} = 28$ of $m = 16$	figure $r = \frac{1}{2}$ $m = 40 \text{ bases}$ 466	ei
Reference tertifiel terripiel	m = A n = Mgs 1: [arbs]	s ta v
is nis y∜hDoddp		Jagev .
ca itiúifie (n n m ìghalla i		ebh
Balander College		B
Tolerating Outages hapby		
lytaisijalbesthaath istilijalab	f	ebs pa
k6kkg@dakinea	•	-
Estimistos Estimistos		ie Ib
dagan dagan		
Tolerating Data Loss		9
		6
Djûrk Clektv K iphtwiteile E i	!	nten eta
₩		.
/ Egathidistalia et tinalici ta		ig yis
		Θ
l igistletide l igistletidde		el el
the public	ef	ÎĐ
Adapting to Price Changes	iid ele n	
idala dainida ab		I£V
Managa p Edugato E	1	b) tep
statement.		ø
iginaliki (kudi Hilb		g i
Adapting to New Providers (and the state of t	
Control Monetary Spending		_
i kkitijachu ek titijajali ka	•	b b
attipisellus:		fib
g Choice in Data Recoveryhted	a lific o	
dajlajla/		je
dugligigato Galilladdien		stal gel
Soliphi str		Θ
that the state of		lte
1 Declar Billion		S
e	m (fjala	j an
schij(ing)(in) bv m istratustrati		L- :
c Deficition	e)m (s)b	
ethilitestsel[<i>e aledallat</i> s6	2 2
ENZERENER E Grafan io		m
3·19·		

put	bucket,	key, object
get	bucket,	key
delete	bucket,	key
create	bucket	•
delete	bucket	
list	k ∮n	bucket
list	Ädas	

Table 1: Amazon S3 operations



3. DESIGN

Schließelste n		
d£Static h	buckets tia	
(a)	keys ib v objects ka	
ti thigh atathis		pat5
lyfibilybeitldy m		tt b
(dit))(peld) (V		€ 5
		Ē
ķisstkikalstija m		8
Boddeteidjili tv		B
diligibili pa		ts
add Capital		
Spis eapp istalte		t
jacht 6	n repositorie; sid o	
		9 8
jje stickikic		jga
put ##	m data	shares
	=m (Maja)ka √	
m < n deglación		
t iezib i	(n m) redundant share, £££ 6	n
		уbа
ető m t ádádálájá tá		h
Hitizija Hitizija		
Jantis a	get #	m
er kladd silor i		b≩

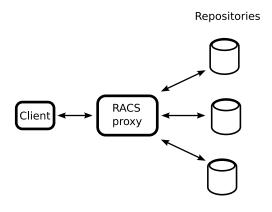


Figure 1: RACS single-proxy architecture

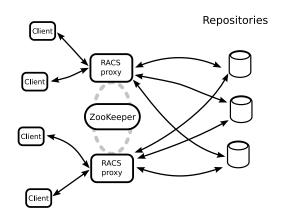
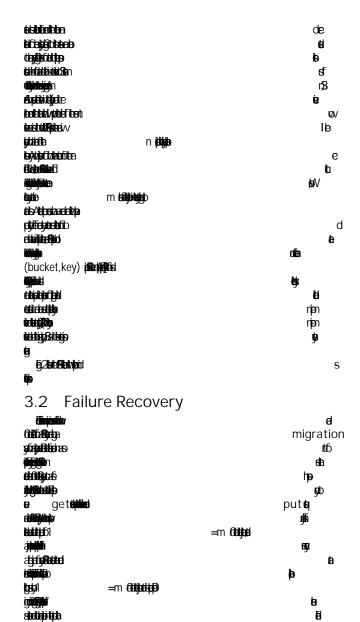


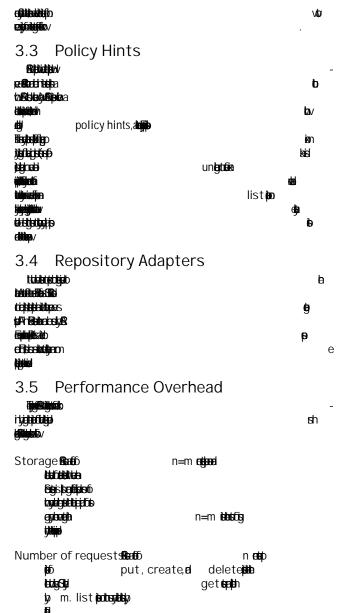
Figure 2: Multiple RACS proxies coordinate their actionsi**ug** ZooKeeper



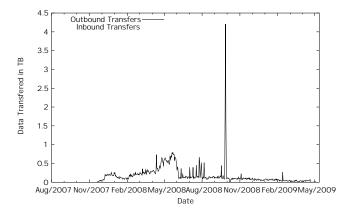
put d deleteblijh

ф

etderej vettijat



Bandwidth	n	put je	HAMP O	tet _
lyza tá tó ca táratáka	$\frac{n}{m}$, ethers	get t list u	ettiritēja Heitiritēja	ntā D
		n n	things	288
ebzalgiei tatatobeta			digitaligh.	tia
		_		_
Codbyldyldyn Ye n		е	45ilidde 4llylliddo	i5 889
•			E NGCHU E	SE E
Latency Bythyth	ut jasinõleisõl e	put e	Bio\$BilitateEi √	į t
jajatajaja Pr	и с развишения			ė,
ibbs	list bielgena p		At Mandate Man	eis Bg
catal city		0	tylint (this step)	b g
sl upidhi rda d dhhib		get ja	Constitution	-
(statutaga		get p a mi b	铀	
ERFJANDING ZA		· I	Millip	У
je Egelijje t		þ	yhathigilla	
yatiksiteteteli ithysteriya				
iji je da kalija i		m	Mulgjidgo	
de propie			ghidhalis Gal a	
3.6 RACS P	rototype Impleme	ntation	e n	
618 pistint 1	rototypo impiomo	Traction .	Water College B.	
b6BobbasBb			Abbahalis R	
(the Cartestisten		d	4.1 Trace Characteristics	
chalatelia@Miles/		h	Hilladifilib	f
i gsRedeglidi ∧ St p:Njdsigedige		h -	CONCENTRALIE A	b.
layag Recyclicated		g	tiller it form	e6
6 9bid is ys tpela		-	diskilyaligy HasialialikA	48a√ 1 2 a
is ttilotyiiti lo Glischyliy s		€r En		.1b
		þn	etital dansp	tte
ingeriyitte n		['] e8	chytrixishebyel	Ð
		ė –		
isteriteijate Vijetskijate/		lga Itada	4.2 Cost of Moving to The Cloud	
Agenting		ap ap	A RIGITY H A	е
****			elibedisperiositele	el e
El-Machaly io		- del	i juljulus N atu s	n6 beal
		etsl ,apo	i i i julijas Lijuu i i i i i i i i i i i i i i i i i i	n
686) (10))]	MEASTACH	eĐ
		ė		nded
şilişiri S atiy çizli Siyili Sabaki ya		þ	indigity seeding the seeding s	pentangan pentangan Sentangan pentangan
tilephiliphici		ġ	tial light sallad in	l i n
Table 1		3	den	
		el No.	hēj5e Hbis6ēgble BigigijiMā s	b n
ks tillingijs HillusgilhjeR		e E	Citibilit	ee
manufacture (5	si delitatisky t	İs
4. AN INTE	ERNET ARCHIVE	IN THE		Macou
CLOUD			ahvebbgbbintelc takjabb@mbn	d h
etri jih ishtis		n	SEELEND	þ ''
k iliyasilabil o		b	l tlykai stalligrai d	g
laterbiglatikt ja		Αjp	CONTRACTOR	op 4
t eritiki (dilit a) Big <u>i</u> tijaki		p g	Operation in the contract of t	d
		eta g	titutiti	9



	SB	36	R	6	ial	EB Ah
86iB	•	\$	₿		₿	\$
爾 多	€	₩	₽	9	₿	⊕
9	€	6	€	69	6 9	€
putel liste(D	•	\$	*			
putal listando getallando	•	6				
(det*	put 🛍	•	•	deletetaje		•

Table 2: Simpli ed pricing schemes of different cloud stgeaproviders.

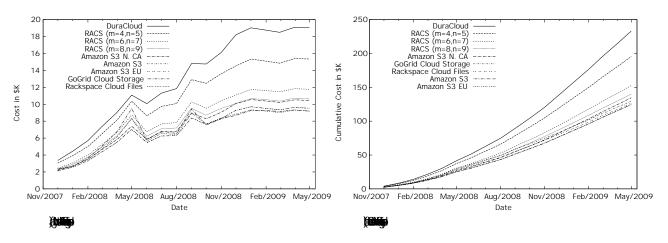


Figure 5: Estimated monthly and cumulative costs of hosbinghe cloud

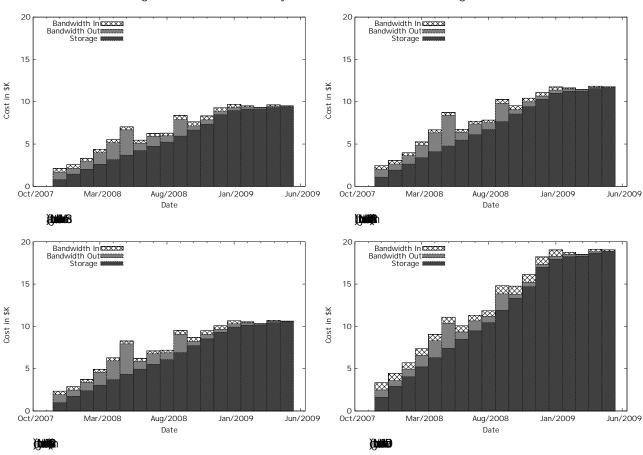


Figure 6: Breakdown of cloud storage costs

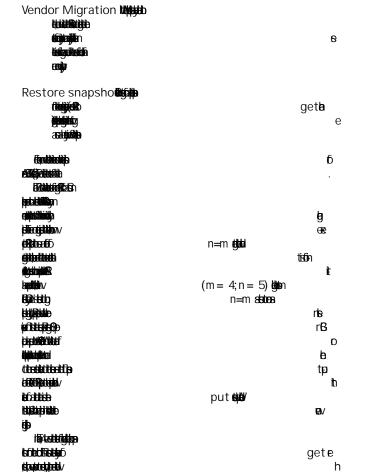
-	-
5 -	-
0 -	-
-5 -	-
-10 -	-
-15 -	-
-20 -	-
-25 -	-

						ď	R 18	ß	ß
put d	lis	t G			22		B	9	74
get q					Ø		8	8	2
##	(₽	8	8	Б				
	2	9		3	3		8		
360 0	Ф	₽	Ф	₿					
BBB (5)	Φ	₿	₽	(4)					
		₿	Ø	Ø	@				
Œ ∰e		1			5	3	5		
W			1		5	2	8		

Table 3: Upload Snapshot Benchmark. Amazon S3 and estimated Rackspace (RS) costs. Monthly cost is the same for both

						NR.	B	₿	B	
pute	list	9			4		4	1	4	. 4
get q				3		8	}	2	28	
##	8		G)	9		3			
30(5) 80(5) 30(6) 30(6)	φ	28	æ	Φ						
389		₽)	Ф	Ф	Ф				
BB B		Ø		Ø	Ø	Ø				
3			1			(1)	6	(20	

Table 4: Restore Snapshot Benchmark. Amazon S3 and estimated Rackspace (RS) costs.



attesteette

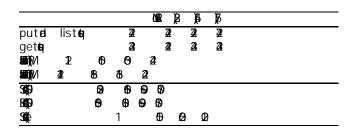


Table 5: Vendor Migration Benchmark. Amazon S3 and estimated Rackspace (RS) costs.

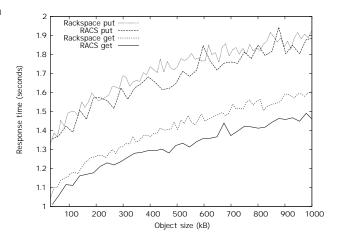
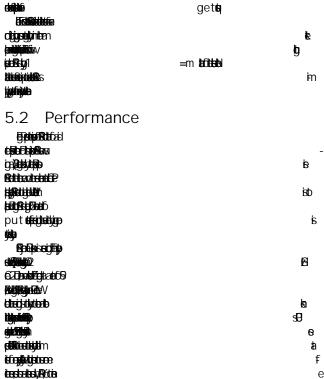


Figure 9: How long it takes RACS and Rackspace to handle object put and get requests, as a function of object sizeeAll sponse times averaged over four runs.

b

ertlefühlis**eleß**en

h



ei jb rakfig	(m;n)	sleshelgalhista v	
e n	(,)	B/#J/# 5	Ð
(IB:LEI)vid:R		tfaki Mathrius	d
	p∂ Sobble	(ji:Chape)	en d
elatrialitation	e zme	lytig/Raddid alih/dighitishas	db
djenishst o	n P	i tijikit kija n	d
te(takking	ia de la companya de	d anikabij an	Þ
	ide .	child differ	Sed
Giffajo	ig √	illibidise.	þ
udyndyddig Urbfighaddfil	igh ef	(1.1.1.1.1.1.1.1.1.1.1.1.1.1.1.1.1.1.1.	ilea/
in the second se	G	(makanila)	
		8. FUTURE WORK	
6. RELATED WORK		MAKUB RA KU N	
Adjuttato	ď		o n _
kinji i jilikin dajajaj Gajakin	dde it	CiploCelagCipia Isajpubbigo	6 id
(alia) Britata Calair Acada	t6	elpistratshupers	m
(iii)	to	(the feater	•••
#######		itato accigitalgetata l/	
	8	grab ik fak	inta .
griffes	Ð	istelletoffstylden	_ yn
guluggigel in bisligig ge lagan	add el	klydyritkilien Giffigjerithi	ba eMaz∨
Retelled	by6	din	e e
ETTO PARTIES AND P	ds ds	dipity	,
g ulli dib i	в	AN B hat be the second of the	S
ebutat histori	t	Absolution	_
Injustic	n B	işləşstələtəsə2000;Fixta ellətininktişlərin	e
klyn Tilphtapistath	k	is/2011tites	eCite
Strouvisite	e	lightly in the second	(P)
this in the second seco	8£ n	ja@djatijsts a	6
EXPLICITION		(Hiptifet)	s tb
GIRICIDA	e	dealtg	
Egilly light Efficiency	b e		
Basisish isipis erimentatio	d d	9. AVAILABILITY	
Historia	te e	Brack Late Conc.	
Epitotic plans	Ġ Đ	is is the second of the second	É
	n g n_	4.0 4.0((4),0)4/(,50,0),451/(70	
chaidtealaightach Alabathleaidte/l	B	10. ACKNOWLEDGMENTS	
Office fights	gel Ka		
plist Michigan	t a	Jettiljet in rijivi etitijaeth	- h
digjaddt A	t a	istAlianturRi	l a
stirite (til)	g	13	
	Œ	3	
continuenta p		11. REFERENCES	
Amilitiks	Đ	∄ http://aws.amazon.con	1/63
aniuj	95	Page	1/30 .
	i s k:	http://www.networkworld.com/new	s/2008/
		072108-amazon-outages.html	
		http://aws.amazor	n.com/s3-sla .
7. CONCLUSIONS		140000	
III. Caleiri istjička:/	t	http://bit.ly/cloud_outage	
skillifiyetin	economic	http://www.duraspace.org/duraclo	ıd nhn
entity istoididis SUNNECARD E	d	http://www.duraspace.org/duracion	ia.prip .
ayuwaacayya Ahbigii isigaliya	is	http://www.atmosonline.com/?pag	e_id=7 .
D - 3-3-	-	.,	-

700				
	http://www.gogrid.com/cloud-hosting			
	http://www.gogrid.com/legal/sla.php		BINNAND	
PMA	http://www.archive.org/		Single Si	000010
	http://www.mim.comin.com/h.com/h.com/		樹 h Proceedings of the 2009 9th IEE	CCGRID '09:
	http://www.nirvanix.com/how-to-buy/ self-service-pricing.aspx		Symposium on Cluster Computin	
] (806)	con corrido prioriigidopx			
	http://www.rackspacecloud.com/cloud_		5 2XXXXXX	w
	hosting_products/files .		http://allmydata.org/trac/zfec	.y , Q
Pigag	http://www.bbc.co.uk/blogs/technology/		Biographics	,-
	2009/10/the_sidekick_cloud_disaster.html			Proc. of ACM
BAR	http://www.rackspacecloud.		SIGMOD Conf., (9) B. And the	
	com/legal/cloudfilessla .		(Piggs)	Software Practice and
	http://hadoop.apache.org/zookeeper		Experience,	
J EH N ON				d Drag of
		brony of	JAMAN USENIX FAST, 2	Proc. of
	文章	brary of	B. B	
	//www.loc.gov/today/pr/2009/09-140.html) (IEEE I I
	I S		ggith n F Computing. ∏®2	Proc. of IEEE Internet
	ELEMAND)		Biribiti	
			i i i i i i i i i i i i i i i i i i i	Proc. of
	A M G U		HPCS, (2007) BY MARING MARY	
	Proc. of ACM Conference on Computer and	n	igi liyajigi u Tanakar fanasar	
	Communications Security (CCS),		HOTOS'07: Proceedings of the 1	
BIODE		d	Hot topics in operating systems,	
	Magneni Magnetic Proc. of IEEE ICDE,∰and		B/MMR	
	9			
PERM	• •		g Bi sapa	
	NSMARKET RATINGET	Proc.	NEAPPEL To management	Gartner Newsroom,
	of USENIX NSDI,		Online Ed., de http:	
DIDIDE		5	//www.gartner.com/it/page.jsp)?id=707508 .
	of ACM SOSP, 669	Proc.		Trans. Storage,
			Q	5 • A
				Drog of IDTDC
ТОТИМЕНТ	h Proc. of USENIX NSDI, MA	n	xiiip h 8 2 1	Proc. of IPTPS,
		n Proc. of		
	the Workshop on Design Issues in Anonymity and			
	Unobservability, y 2	I III		
Biggg	ns No PCWorld,n©o	kel		
1211/ME In	Proc. of FOCS Symp.,			
-	tip)			
		Proc. of		
	ACM STOC,	I IUG. UI		
	_			
	ट्याप्रकृतिका İlba n Proc. of ACM S	TOC 6 4		
	99	100,gvi		