Indirection	2VCC		References
00000000	0000000	00000	0

Overview of Concurrency in L-Store: 2VCC - Two-version Concurrency Control

Mohammad Sadoghi

Exploratory Systems Lab University of California, Davis

ECS165a - Winter 2021







Indirection	2VCC	References
••••••		

1 Data Velocity: Index Maintenance

2 Data Volume: MVCC Concurrency

3 Decentralized & Democratic Data Platform

4 References







Reducing	Index maintenance	Velocity Dimension	
00000000	0000000	00000	
Indirection	2VCC	Vision	References

In the absence of in-place updates in operational multi-version databases, the cost of index maintenance becomes a major obstacle to cope with data velocity.

Reducing Index ma	aintenance [.] Ve	locity Dimension	
0000000	0000000	00000	0
Indirection	2VCC	Vision	References

In the absence of in-place updates in operational multi-version databases, the cost of index maintenance becomes a major obstacle to cope with data velocity.

Extending storage hierarchy (using fast non-volatile memory) with *an extra level of indirection* in order to

Reducing Index	maintenance.	Velocity Dimension	
0000000	0000000	00000	
Indirection	2VCC	Vision	References

In the absence of in-place updates in operational multi-version databases, the cost of index maintenance becomes a major obstacle to cope with data velocity.

Extending storage hierarchy (using fast non-volatile memory) with an extra level of indirection in order to Decouple Logical and Physical Locations of Records to Reduce Index Maintenance





Updating random leaf pages

Mohammad Sadoghi (UC Davis)

ECS165a - 2021 5 / 24

э

The distance MALLET	and the standard strength		
000•00000	0000000	00000	0
Indirection	2VCC	Vision	References

Traditional Multi-version Indexing: Updating Records





	and the second second second second			
000•••0000	0000000	00000		
Indirection	2VCC	Vision	References	







Indirection		2VCC	Vision	References
000000000				
	1 8 4 1.1	• • •	• • • • •	







Indirection Indexir	g: Updating Records		
	0000000	Vision 00000	O O



3

イロト イヨト イヨト イヨト

Indirection	2VCC	Vision	References
00000000			

Indirection Indexing: Updating Records



3

< □ > < 同 > < 回 > < 回 > < 回 >



э

< □ > < 同 > < 回 > < 回 > < 回 >



Eliminating random leaf-page updates

Mohammad Sadoghi (UC Davis)

< □ > < □ > < □ > < □ >

э



Eliminating random leaf-page updates

Mohammad Sadoghi (UC Davis)

ECS165a - 2021 6 / 24



Eliminating random leaf-page updates

Mohammad Sadoghi (UC Davis)

ECS165a - 2021 6 / 24

< □ > < □ > < □ > < □ >

э

Indirection	2VCC	Vision	References
000000000			

Analytical & Experimental Evaluations

Indirection	2VCC			Vision	References
0000000000	00000				
1.12.12	 ~	1.1.1.1	A 1		

Indirection Time Complexity Analysis

	Legend
Κ	Number of indexes
LB	LIDBlock size
М	Number of matching records

Method	Туре	Imm. SSD	Def. SSD	Imm. HDD	Def. HDD
Base	Deletion	0	0	2 + K	$\leq 1 + K$
	Single-attr. update	0	0	3 + K	\leq 2 + K
	Insertion	0	0	1 + K	$\leq 1 + K$
	Search Uniq.	0	0	2	0
	Search Mult.	0	0	1 + M	0
Indirection	Deletion	2	0	2	≤ 3
	Single-attr. update	2	0	4	\leq 3
	Insertion	2 + 2K	2K/LB	1	$\leq 1+2K/LB$
	Search Uniq.	2	0	2	0
	Search Mult.	1 + M	0	1 + M	0

2

イロト イヨト イヨト イヨト

Indirection	2VCC	Vision	References
000000000			
Experimental Setti	ng		

Hardware:

- (2 × 8-core) Intel(R) Xeon(R) CPU E7-4820 @ 2.00GHz, 32GB, 2 × HDD, SSD Fusion-io
- Software:
 - Database: IBM DB2 9.7
 - Prototyped in a commercial proprietary database
 - Prototyped in Apache Spark by UC Berkeley
 - LIBGist v.1.0: Generalized Search Tree C++ Library by UC Berkeley (5K LOC) (Predecessor of Generalized Search Tree (GiST) access method for PostgreSQL)
 - LIBGist^{mv} Prototype: Multi-version Generalized Search Tree C++ Library over LIBGist supporting Indirection/LIDBlock/DeltaBlock (3K LOC)

Data:

- TPC-H benchmark
- Microsoft Hekaton micro benchmark





Substantially improving the update time ...

Mohammad Sadoghi (UC Davis)

ECS165a - 2021 10/24

Update (Base)



... Consequently affording more indexes and significantly reducing the query time

Number of Indexes

0.2

0 7 (PKs) 8 9 10 11 12 13 14 15 16 17

Indirection	2VCC	References
	0000000	

1 Data Velocity: Index Maintenance

2 Data Volume: MVCC Concurrency

3 Decentralized & Democratic Data Platform

4 References

Intua du ain a N	Audation Comp		
00000000	0000000	00000	
Indirection	2VCC	Vision	References

Introducing Multi-version Concurrency Control



Generalized	Concurrency Contro	I. Volume Dimens	sion
	0000000		
Indirection	2VCC	Vision	References

In operational multi-version databases, there is a tremendous opportunity to avoid clashes between readers (scanning a large volume of data) and writers (frequent updates).

Generalized	Concurrency Contro	Volume Dimens	sion
00000000	0000000	00000	
Indirection	2VCC	Vision	References

In operational multi-version databases, there is a tremendous opportunity to avoid clashes between readers (scanning a large volume of data) and writers (frequent updates).

Introducing a (latch-free) *two-version concurrency control (2VCC)* by extending indirection mapping (i.e., central coordination mechanism) and exploiting existing two-phase locking (2PL) in order to

Generalized	Concurrency Contro	d. Volume Dimens	sion
00000000	0000000	00000	
Indirection	2VCC	Vision	References

In operational multi-version databases, there is a tremendous opportunity to avoid clashes between readers (scanning a large volume of data) and writers (frequent updates).

Introducing a (latch-free) *two-version concurrency control (2VCC)* by extending indirection mapping (i.e., central coordination mechanism) and exploiting existing two-phase locking (2PL) in order to Decouple Readers/Writers to Reduce Contention (Pessimistic and Optimistic Concurrency Control Coexistence)





Extending the indirection to committed/uncommitted records

э

< □ > < 同 > < 回 > < 回 > < 回 >



Mohammad Sadoghi (UC Davis)

ECS165a - 2021 14 / 24

э

< □ > < 同 > < 回 > < 回 > < 回 >



Decoupling readers/writers to eliminate contention

э



Decoupling readers/writers to eliminate contention

Mohammad Sadoghi (UC Davis)

ECS165a - 2021 14 / 24

э



Decoupling readers/writers to eliminate contention

э





Two-phase locking (2PL) consisting of growing and shrinking phases





Two-phase locking (2PL) consisting of growing and shrinking phases

< ロト < 同ト < ヨト < ヨ



Overview of Two-version Concurrency Control Protocol



Two-phase locking (2PL) consisting of growing and shrinking phases

Mohammad Sadoghi (UC Davis)

ECS165a - 2021 15 / 24





Extending 2PL with certify phase

Mohammad Sadoghi ((UC Davis)
--------------------	------------

э







Exclusive locks held for shorter period (inherently optimistic)

Mohammad Sadoghi (UC Davis)

ECS165a - 2021 15 / 24

(I) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1))







Exclusive locks held for shorter period (inherently optimistic)

(I) < (II) <





Relaxed exclusive locks to allow speculative reads (increased optimism)

★ ∃ ► ★





Trade-offs between blocking (i.e., locks) vs. non-blocking (i.e., read counters)

(I) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1)) < ((1))

Indirection	2VCC	Vision	Referen
	00000000		

Experimental Analysis

2\/CC · Effect of	of Porallel Undate Tra	insoctions	
00000000	00000000	00000	
Indirection	2VCC	Vision	References



Update Only Workload; High Contention

Substantial gain by reducing the read/write contention & using non-blocking operations



2VCC: Effect of Parallel Update Transactions



Lock Statistics Comparison; High Contention

Substantial gain by reducing the read/write contention & using non-blocking operations

< □ > < 凸

Indirection	2VCC	Vision	References
		0000	

1 Data Velocity: Index Maintenance

2 Data Volume: MVCC Concurrency

3 Decentralized & Democratic Data Platform

4 References

Indirection	2VCC	Vision	References
		0000	

Recap: Data Management Challenges at Microscale



OLTP and OLAP data are isolated at microscale

Mohammad Sadoghi (UC Davis)

æ

イロト 不得 トイヨト イヨト

Indirection	2VCC	Vision	References
		0000	

Recap: Data Management Challenges at Microscale



First step is to unify OLTP and OLAP

Mohammad Sadoghi (UC Davis)

æ

< □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ >





Moving towards distributed environment

Mohammad S	Sadoghi ((UC Davis
------------	-----------	-----------

æ



Platform Scaling: Non-blocking Agreement Protocols



Message redundancy vs. latency trade-offs [EasyCommit, EDBT'18]

Mohammad Sadoghi (UC Davis)

ECS165a - 2021 20 / 24

< □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ >

Central Control	l· Data Gate Kee	ner	
		00000	
Indirection	2VCC	Vision	References



Conform to trusting the central authority and governance

Mohammad	Sadoghi	(UC Davis
----------	---------	-----------

イロト イヨト イヨト イ

æ





Seek trust in decentralized and democratic governance [PoE (EDBT'21), RCC (ICDE'21)]

(日)



Seek trust in decentralized and democratic governance [PoE (EDBT'21), RCC (ICDE'21)]

(I) < (II) <





Self-managed infrastructure

ECS165a - 2021 22 / 24

э

< □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ >

Indirection 00000000	2VCC 00000000	Vision ○○○●○	References O
Global-scale	Reliable Platform	over Unreliable H	ardware



Cloud-managed infrastructure (trust the provider)

Nohammad S	Sadoghi ((UC Davis)
------------	-----------	------------

э

Image: A mathematical states and a mathem

Global-scale Relia	ble Platform over l	Jnreliable Hardwa	are
		00000	
Indirection	2VCC	Vision	References



Cloud-managed infrastructure (trust the provider)

Mohammad S	adoghi ((UC	Davis
------------	----------	-----	-------

(日)

э

22 / 24

ECS165a - 2021

Global-scale R	eliable Platform	over Unreliable H	ardware
		00000	
Indirection	2VCC	Vision	References



Light-weight, fault-tolerant, trusted middleware [Blockplane, (ICDE'18)]

э

< □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ >



Global Scale fault-tolerant protocols [GeoBFT (VLDB'20), Delayed Replication (ICDT'20)]

< ロト < 同ト < ヨト < ヨ

Indirection	

2VCC 00000000 Vision

References 0

Questions? Thank you!

Exploratory Systems Lab (ExpoLab) Website: https://expolab.org/





Related Publications (Patents Omitted)

- 1. Abdelante, A. Folson, O. Hanarmadoh, P. Zhang, M. Sarlight Large-scale viscotraal and tentical visibility based mixing of boundings graph to predict drug drug interaction
- In J. Web Son, pages 100-117, 2017. 5. Februari G. Hanaravaldi, M. Sadashi and F. Zhana
- Q. A. Francis, G. Franziscov, in: January, and F. Atang. Predicting drug drug infranzismi. Grouph similarity hand Tah pendiction over ande data. In Proceedings of the 21th International Conference on World 1999 1996, WWW 2016, Manimel, Canada, April 11 38, 2018, Companier Volume, pages 179–178, 2018.
- A. Folume, M. Bankephi, O. Homorenshoh, and P. Zhang,
 Predicting desg desg interactions through large-scale similarity hand link prediction.
- In The Imputing Windowski and Bara Dimonsional 20th Immunities (2010) 2011 (2
- Accelerating dealase methanis by selecter barbare system on design.
 In 2014 IEEE International Conference on Data Engineering. ICER 2019, Holdest, Fisland, Bday 38 20, 2018, pages 1428–0413, 2018.
 D. A. Chardto, C. Hourswalto, M. Konder, M. Sanderi, and D. Nivartano.
- A. Chambi, D. Homawalith, K. Koulas, M. Endight, and D. Schuntava. Enchanalising designation approximate sciencian predicates.
- A. Farvahli, M. Badaghi, and H.-J. Jamines. Transfer underskillig based intraction with newst presenting. In Proceedings of the 10th 2400 international conference on Databased week based entern. 2010. 2011. name 107–202. New York, Ne
- D. Hanaraucht, M. Sadighi, and R. J. Miles. Accuracy of approximate string joins using gram.
- b. V. Garit and J. Harmann, editors, Proceedings of the Fifth International Workshop on Quality in Educations, QCW 2007 at MEDI 2007 conference, Vienna, Anniea, Styteminer 20, 2007, pages 11–50, 200
 W. Hormanipung, R. Monimules, and M. Sadagpit. Earch A distribution in sommer how the simulation.
- In Presentings of the 17th Annual Ministerary Conference, Trente, Bulg. December 12–18, 2018, 2018.
 M. Jargies, M. Sarlaghi, and H.-R. Jacoban.
- [10] M. Joppi, M. Kadagki, and K. A. Jaminon. COMPUTE: A surgery and industriant to furthering data contrict and human and human products. *Molecular, Mathematica, Mary 21 - June 4, 2023, pages 1027–1029, 2023.* In *Workshop of the 2020 IEEE INFORMATION Computing Conference on Management of Data, Molecular, Marked, Mary 21 - June 4, 2023, pages 1027–1029, 2023.*
- M. M. Jorgiev, M. Kanlaghi, and H.-J. Jaminov. Gen-distribution of Brailer basisma processon user publish/solvariler paradigm. In Proceedings of the 17th Annual Mildingure Conference, Trents, Edg. Disconders 12–18, 2028, 2028.
- P. Menor, T. Rald, M. Basleghi, and H. Jamines. Califordia: An 35D Journal bryoalar stars.
- In IEEE 20th International Conference on Data Engineering, Oringe, ICDE 2014, R. 1558, March H. April 4, 2014, pages 1110–1117, 2014.
- Optimizing lay-value stores for hybrid storage architectum. In 1972 Distances and the store of the stores of the store of the store
- The FQP value. Findle party promiting as a monitorial surgesting fairs. NUMOD Record - Special Junct on University Inter in Data Management , 40(2), pages 1–30, 2019.
- R. V. Mohammy, Y. Yaon, M. Bahnghi and K.A. Javoham, "oblythm: suppring field initiation downstantinum and Mann. In: Proceedings of the 18th ACM international conference on Distributed receiption, IEEE 2013, pages 381–382, New York, New Distributed and Academic and Academic Conference on Distributed receiption, IEEE 2013, pages 381–382, New York, New Distributed and Academic and Academic Conference on Distributed receiption, IEEE 2013, pages 381–382, New York, New Distributed and Academic Conference on Distributed receiption, IEEE 2013, pages 381–382, New York, New Y
- Finalis pary promote as FPEA. PALOR 2012 6(12) 1110-1113, 2013
- M. Nijah, M. Salaghi, and H. Jawitson. Configurable holesane based intensing andimisture using online programmable blocks. In the 2018 International Conference on Data Engineering. ICOI 2025, Sourd. South Kerne, April 1217, 2021, pages 129–10.
- M. Nicht, M. Sadaght, and H. Janstein. The FQP vision: Finalist query promoting on a meanfigurable computing faders. IEEE/DDF Reveal, 54(1):1-10, 2015. M. Nicola, M. Sadarbit, and M. Janstein.
- M. Kojek, M. Balaghi, and H. Janolson. *Billion: A weak of the starting of a scheduler with adjustative setting prevalues*. *In 2018 IEEE/CVF Annual Technical Conference*, *IEEE/CVFC 2018*, Lenser, CO, UMA, Amer 27.212, 2028, pages 839–939, 2 *7*. Reid, M. Barghell, S. Gareev, V. Manthal, Marken, H.A. Janness, and S. Meshanki.
- [2] T. Kald, M. Balaghi, S. Ginner, V. Martin, Marin, H.-J. Javelsen, and S. Martanshi. Bolog hig data shallness for minipule application preformance management. *Proceedings of the VLDB Endocument, PVEDI 2023*, 5(12):1724–1778, Aug. 2013.
- T. Rali, M. Sashighi, K. Zhang, and H. Janaham. Partie: MM205 - a multi-layered, adaptive, distributed event stare.
- 7 Ruli, K. Zhang, M. Sadoghi, N. K. Paviny, A. Ngan, C. Wang, and H.-J. Javahon. Boling manufactoring equipment manifolds through villationic samples near interacting. Exits grant hadrogs. In Proceedings of the 80 ACM International Conference on Datibuted Ensemble Reveal Parallel SULT, pages 310–383. Berlin, Germany, 2022. ACM.
- In Proceedings of the Bit ACM International Conference on Databased Event Reard Systems, DERS 2012, pages 131–300, Review, 2012. ACI M. Kanlaghi
- ¹⁰ Towards an restandale officient rewet presenting locad. In Proceedings of the MIDACO/PICES 2012 FED Symposium, PMD 2012, pages 3–8, Sontitudale, Arisona, USA, 2013. ACM.
- 8. Salight 1. Borns, and H.-A. Jaminon. IFX Mathew - growing Backup predicts hand 374th segmetics matches in Promotion of the 10th biotestimol Conference on Entending Ladawa Technology. IEEE//CET 2011, pages 49–98, Uppeds, Sond
- M. Sashighi, M. Cavin, B. Duritariuriye, F. Nagel, and K. A. Ross. Enducing database looking contentian through multi-arreian concernment.
- Philips 2014 7(13) 1331-1342, 2014
- Maptine parallel compressed neuron matching.
 In IEEE 20th International Conference on Data Engineering, Orisogo, NZW 2014, IL, USA, March IV April 4, 2014, pages 201–1275, 2014.
- M. Sadaghi and H. A. Jandows.
 H. Starlaghi and H. A. Jandows.
 H. Ster, as index situative in efficiently match. Random supervision, new high dimensional discrete spars.
 In Proceedings of the 2023 ACM INTERNATION International Conference on Public, INSERCO 2011, pages 827–468, Advers, 2011. ACM.
- B M. Sanlaght and H. A. Junchum. Location-hand matching in public/valuenche metalled.
- M. Redight and H. A. Jacobson.
 Kényeser mattery: Capitalizing in Intelligent matching in publish/subsector).
- In HEE 20th International Conference on Data Engineering, ICDE 2012, pages 708-717, Adeques, Veginia, USA, 2012. IEEE Computer Society.
- M. Sadaghi and H. A. Jaminos. Analysis and optimization for Basinan represent indusing.
- ACM Trans. Database Syst., TODS 2013, 38(2) 81–847, July 2013

- M. Sadaghi, H.-A. Jarviner, M. Lalerenpe, W. Show, and H. Singh. Efficient neuroperating through maniforcable hardware for algorithmic totaling
- Pennedige of the SLDB Endowneet, PELDB 2010, 3(2):1120–1120, 2010.
 M. Sadaghi, K. Javel, N. Tandilar, H. Singh, R. Palariagnan, and H.-A. Janzinee.
- Mukit party virtuar promoting on figure In Promoting of the 2022 EEEE 2016 International Conference on Data Engineering, ICDE 2023, pages 1220–1230, Washington, DC, USA, 2023. IEEE Computer Society M. Sashinghi, M. Jorgier, H.-H. Jamison, R. Holl, and R. Yanath.
- [2] M. Eadaghi, M. Jorgier, H. A. Janobow, R. Hull, and R. Vasudo. Safe distribution and parallel researching of data sensitiv sarellouss new the publich/subsection abstraction. IEEE Trans. on Forum 2 and Data Data. TXXII 2015. 2015.
- B. M. Sarlaghi, M. Jorgier, H.-S. Janvison, N. Holl, and R. Varulin. Safe distribution and parallel researcing of data sensitic workflows new the publich/subservice adstruction.
- M. Satisfiel and N. Soudan
 Willing Data Data Chapter Protein
 - In the 15th International Conference of Center for Advanced Studies on Collaborative Research, IBM CASCON 2028. Online 17:32, 2028. 2028.
- M. Sadaghi, K. A. Ross, M. Canin, and R. Ebattanharjon. Making spikens disk1/D fermily using SSDs.
- M. Sadaghi, K. A. Ross, M. Casto, and R. Bhattacharjos. Exploiting such in speculianal multicension databases.
- M. Sadaghi, S. Ehstacherger, E. Ehstacherger, and M. Carter, Liferer: A realities GLTP and GLAP system.
- M. Sadaghi, H. Singh, and H.-A. Janshon. Spac TuPUS: Incompress room processing on FPGAs.
- figs TaPSS: line speed next proceeding on FPGAs.
 In Proceedings of the Std ACM International conference on Distributed scene kaund system, DBIS 2013, pages 3779–374, New York, Mar. York, USA, 2013. ACM
- M. Malaghi, H. Singh, and H. A. Annione. Transch highly goodfer small presenting through menofiqueditis hardware. In Proceedings of the Denseth Honoralized Distribution on Earlier Manhaume, Caddinit 2021 at 30201020, pages 27–52, Advense, Graves, 2021. ACM.
- A standards
 Manufacture Data Submort Platham
 Incomparison of the Displander of the Displander Standards
 Incomparison of the Displander of the Displa
- 2 In Analysis A. 2000a, D. Patashalari, Y.L. Charg, B. Labir, A. Panini, and P.A. Panini. Influencing databases. In Proceedings of the 18th International Conference on Estimating Database Technology, EDIT'14, Rookaan, Proceed March 19, 2020, pages 033.
- 1. Nypey, M. Steligere Mars, O. Hawarashi, A. Materichan Ghow, M. Jalagki, Jake Learning of Load and Ghade Partners for Entity Editing of Neural Networks. In Proceedings of the 20th International Conference on Computational Linguistics, COS INT' 30, Outra, Japan. Describer 11 38, 2019, pages 2135–2120, 2028.
- K. Zhang, M. Xashighi, V. Matheamy and H.-B. Januham. Distributed sedimi faits distribution in smill retroots.
- In Presendings of E2-d International Conference on Distributed Computing Systems, IEDCS 2023, Washington, DC, USA, 2023. IEEE Computer Society K. Zhang, M. Sashinghi, and H.-A. Janshew.
- [8] K. Dong, M. Badaphi, ed K. M. Jamhen. Ex share: A deviational lybrid CET and CLAP data proceeding regime. In 2018 IEEE International Conference on Distributed Computing Journe, ICDCS 2018, Nava, Jopen, Jour 27-20, 2018, pages 709–770, 2016.
- M. Nicjól, K. Zhang, H. A. Janshon, M. Ballaghi Hatharan Antomization Londways for Direktowine Paul fore Analysis. Votans and Linkastere. In UEE 17th International Conferences on Distribution Computing Systems (EDCS 2027), Michael, GA, 556, 52 Aur. 2027, pages 2010–2048, 22
- Bhair Rainbarni, Maratu Carin, Mihammad Latighi, Eshawarajin Bhatsaharjin, Yuan Chi Dang, and Panis Kales. Isommatid Response subgraph miting on long multiling paphs. IEEE Tennes Hourd: Data for UNIX 21154-2112 2021.
- Bhát Rédobarolt, Manada Carrin, Mohammad Tarlighi, Rohammyin Bhattanharjin, Yuan Chi Chang, and Panns Kalnis. Incommuted Import subgraph mining an large realizing graphs.
- Genals 1 Dias, Arbite Folices, and Mohammad Sadoghi
 - Indexis, Staldah, uninkey awar papit selectings. In Proceedings of the 23th International Conference on Extending Database Technology, EDET 2028, Vienne, Austria, March 20-20, 2028, pages 433–430, 2028.
- Anna Genera and Mohammad Baloght. Hybrid GLTP and GLAP.
- In 1994 Take and Albert Zemaps, white, Encyclopedia of Eig East Technologies, pages 1–13. Springer International Publishing, Cham, 2020 Speech Gapta and Materia Zemapia. Biological Technologies
- Hadahain Tuanatine Processing. In Stored Sales and Alloret Zamaya, million, Encyclopedia of Rig Easte Technologies, pages 1–11. Springer International Publishing, Cham, 2028.
- n r memory or ne concentrational Conference on Extending Distance Probability. EEET 2028, Verna, Austria, March 20-20, 2028, pages 307–308, 201 30 Marcend S. Haman, Talana Kaumitana, Han Chai Joney, Malif G. Arel, and Mahammad Sadaghi. 31 Second Science Science Andreas Content and Science Andreas Andreas Science Science Science Andreas An
- In Proceedings of the 25th International Conference on Estimating Candonse Technology, ECRT 2023, Vienue, Austria, March 26 29, 2028, pages 29–88, 2028.
 Michaeved S. Honson, Factore Konewhenes, Myon Chel. Joney, Billel G. Aref, and Michaeved Stadight.
 Orbitation: Conduct and Endon March and Area Candidates relations.
- Environmenter Bigle, Martinerer and Berner and Ber
- A walable streadar pipeline design for multi-may stream joins in hardware. In 31th IEEE International Conference on Data Engineering, ICEE 2018, Paris, Paris, Paris, April 10-10, 2018, 20
- Shharmad Sadagh, Sandi Bhatacheyn, Bolszwarjen Blattanhayn, and Matala Carin.
 Leiser: A malitime OCFF and OLFF spinn.
 In Promotion of the 20th Broading Conference on Entended Database Technology EDET 2022. Versa, Anaton. Mach 20:29, 2022. versa Mil
- Ration Zhong, Mahammad Satinghi, Vinnel Michoareng and Harn-Amer Janakom. Efficient moving for top & Ultraing in motionic hand publick/valuer-for spatime.
- 17. Quicks and M. Sadoghi. QuCC: A quasa standard, marked free concurrency architecture. In Proceedings of the 18th ACM/1997 SIGENCE Middleware Conference, Researc, France, December 23 - 14, 2018, 2018.

▲ロト ▲圖 ▶ ▲ 臣 ▶ ▲ 臣 ▶ ─ 臣 ─ のへで