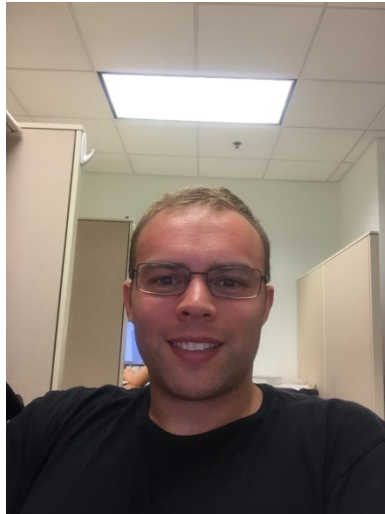


Database Forensic Analysis with DBCarver

James Wagner, **Alexander Rasin**, Tanu Malik,
Karen Heart, Hugo Jehle, Jonathan Grier



Data Systems and Optimization Lab at DePaul



James Wagner



Tanu Malik



Karen Heart



Hugo Jehle



Jonathan Grier



Motivation

- Cyber-crime
 - Detecting (and proving) data theft
 - JP Morgan/Dow Jones
 - Mobile device analysis
 - FBI, 4Discovery
- Involves a database



Motivation

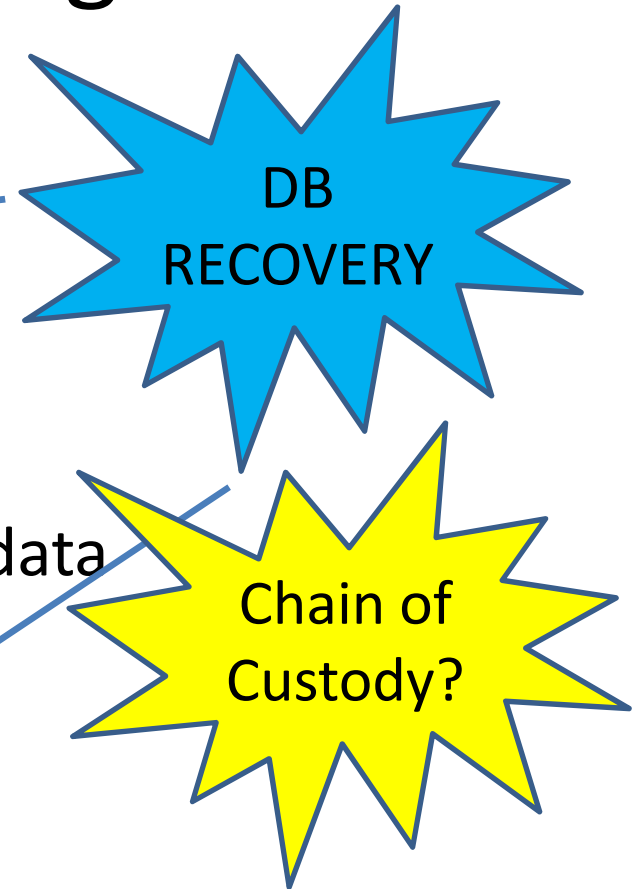
- Example Queries
 - Reconstruct deleted data
 - Identify recent access, modifications
 - Detect catalog/data tampering
- Un-trusted environment

Forensic Analysis Targets

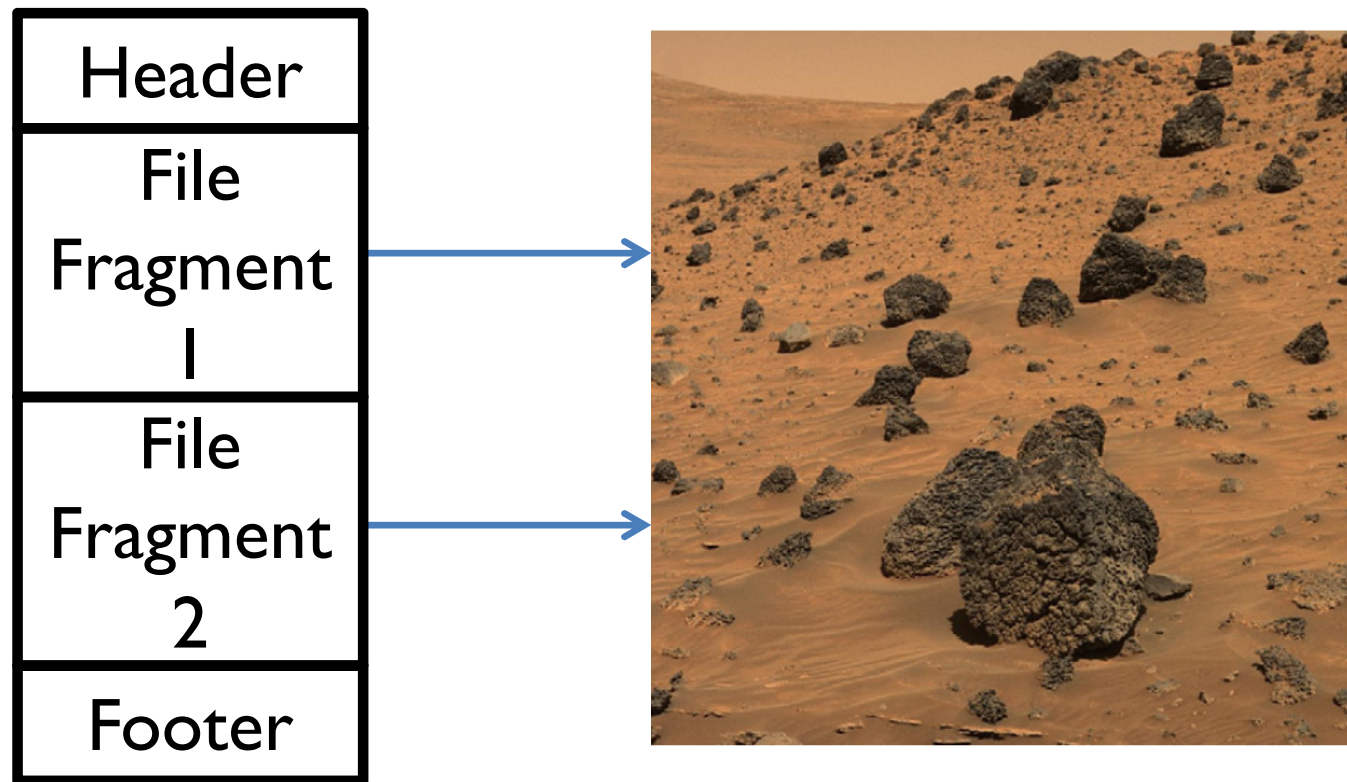
- Logs
 - Audit, Query, WAL
- RAM
 - Buffer cache, intermediate data
- Query-able DB content
 - Tables, MVs, Catalog
- Un-query-able content
 - Indexes, Deleted data, Free-listed data

Forensic Analysis Targets

- Logs
 - Audit, Query, WAL
- RAM
 - Buffer cache, intermediate data
- Query-able DB content
 - Tables, MVs, Catalog
- Un-query-able content
 - Indexes, Deleted data, Free-listed data



File Carving (JPEG)

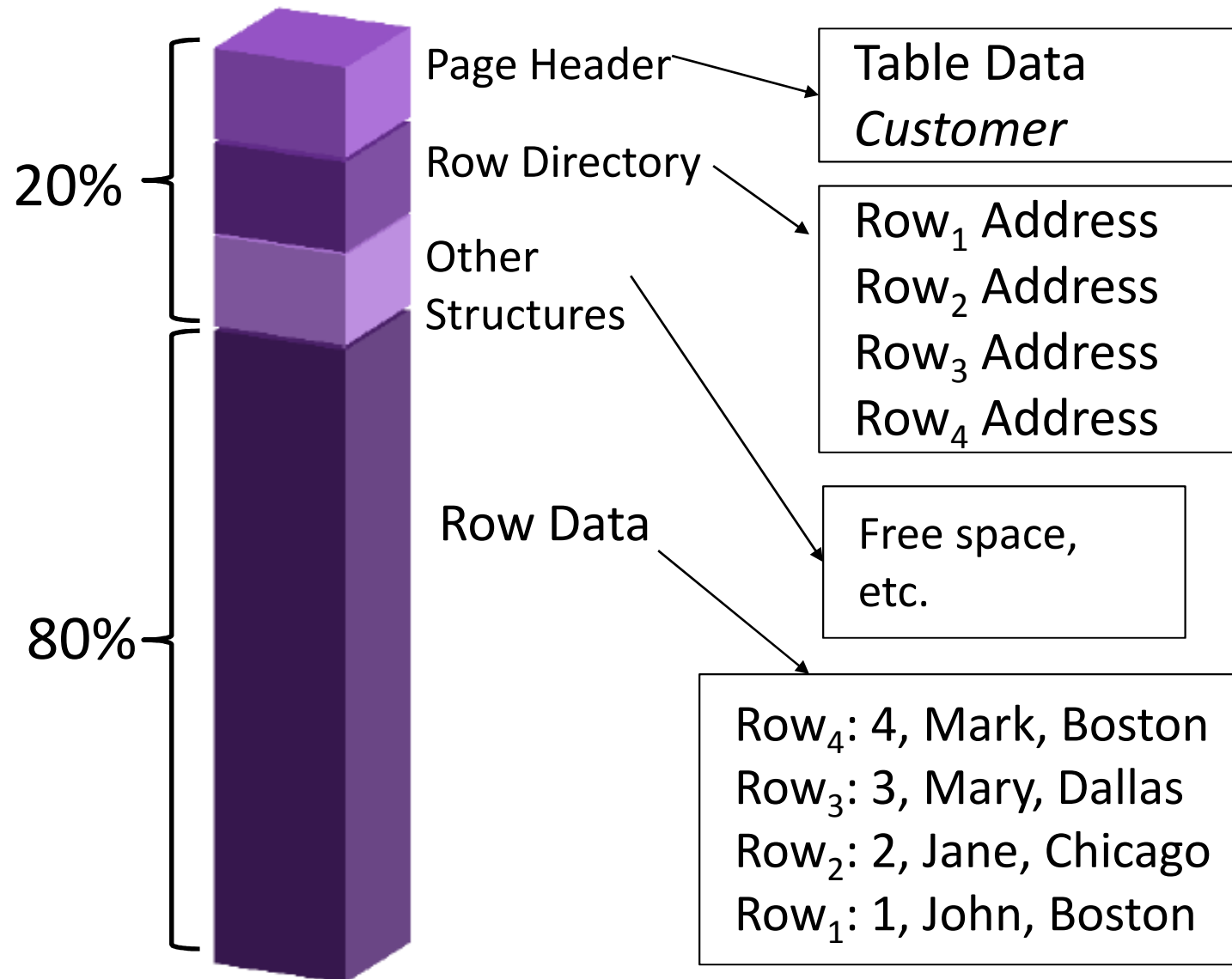


Forensic Analysis Targets

- Logs
 - Audit, Query, WAL
- RAM
 - Buffer cache, intermediate data
- Query-able DB content
 - Tables, MVs, Catalog
- Un-query-able content
 - Indexes, Deleted data, Free-listed data



Generalized Page Carving



Forensic Analysis Targets

- Logs

- Audit, Query, WAL

- RAM

- Buffer cache, int. data

- Query-able DB content

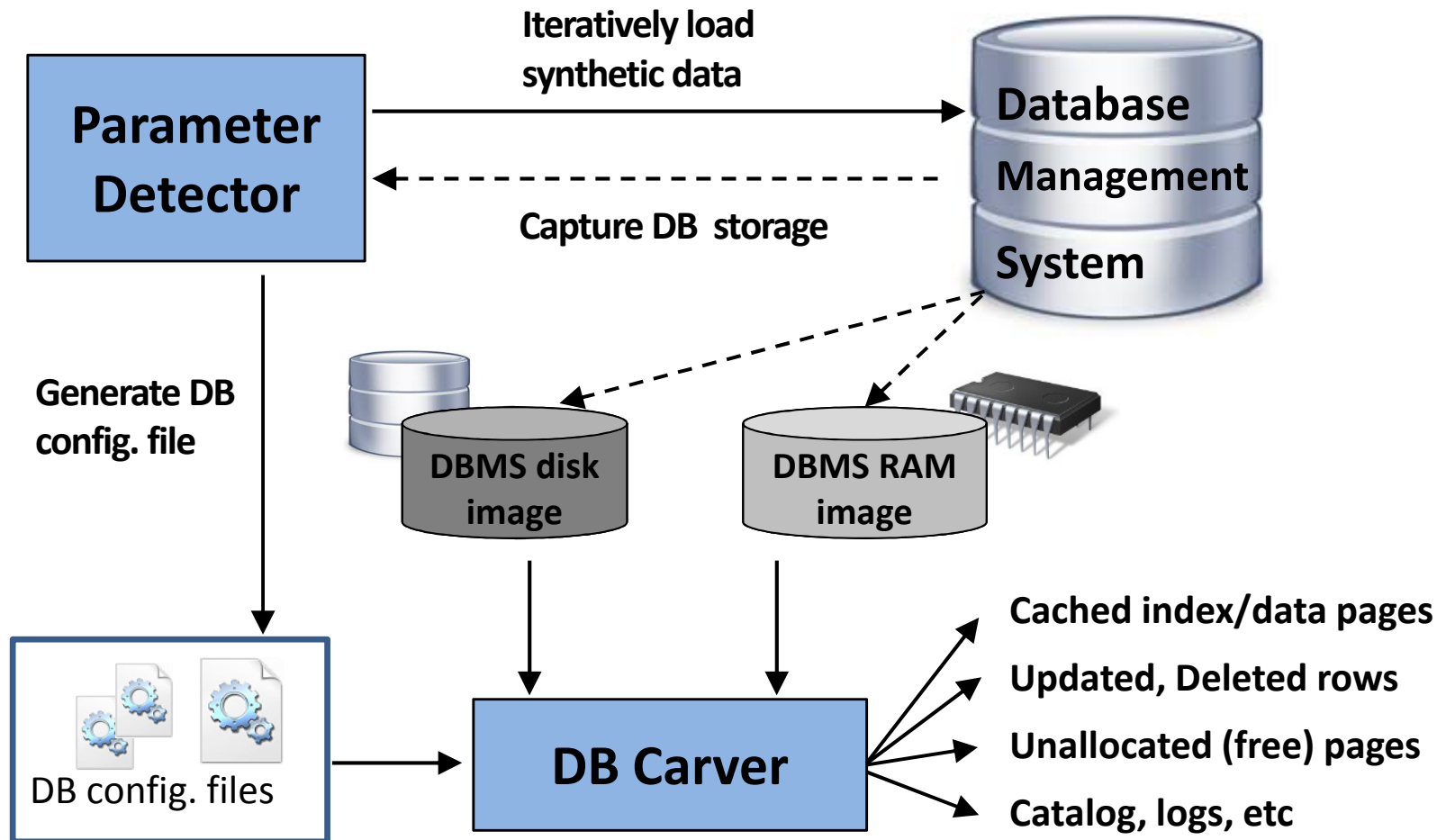
- Tables, MVs, Catalog

- Un-query-able content

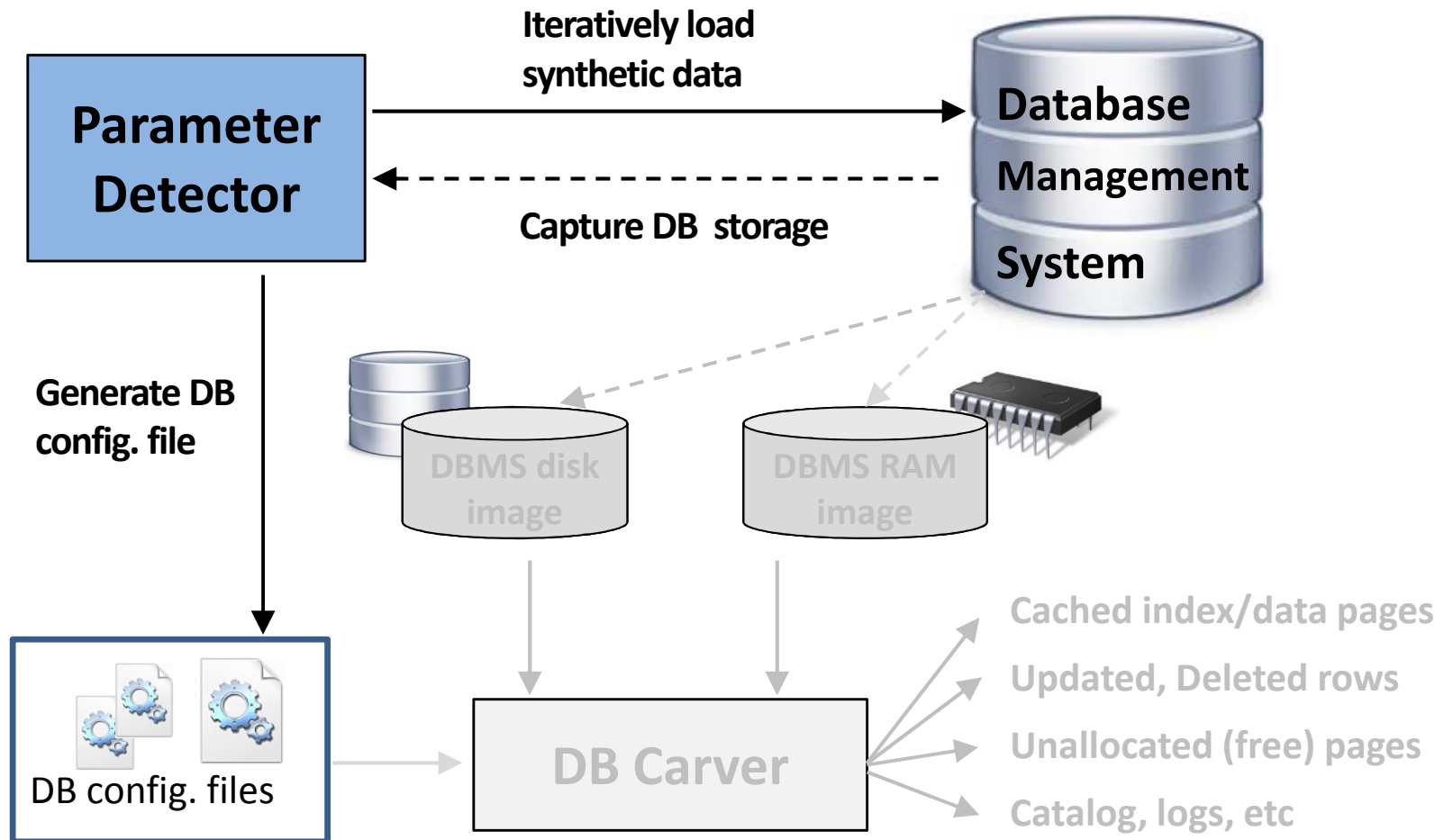
- Indexes, Deleted data, Free-listed data



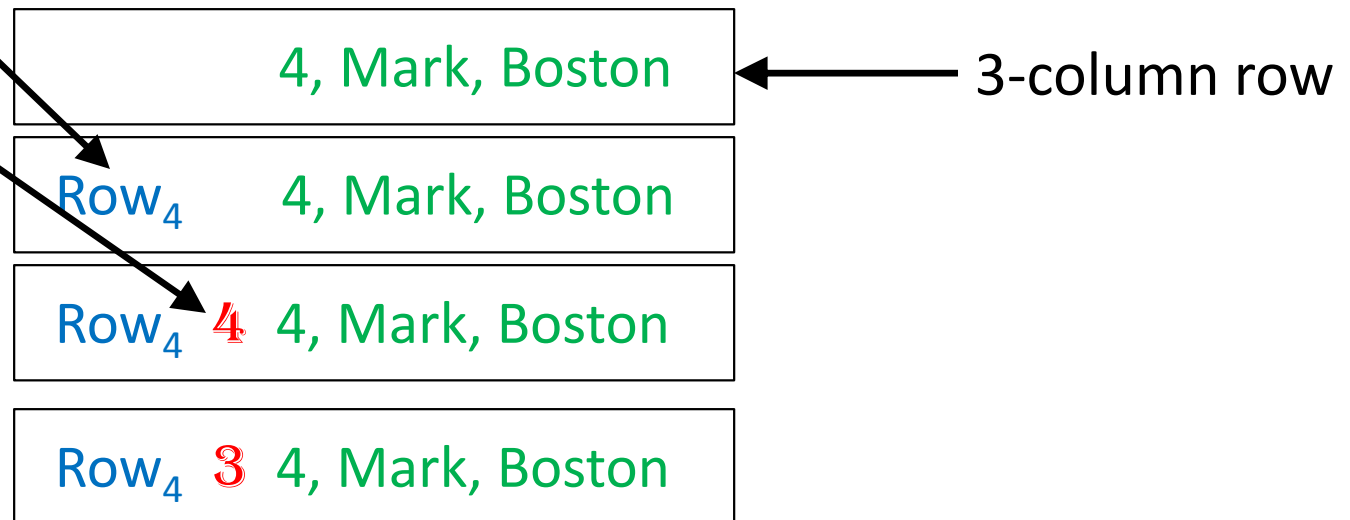
DBCarver Architecture



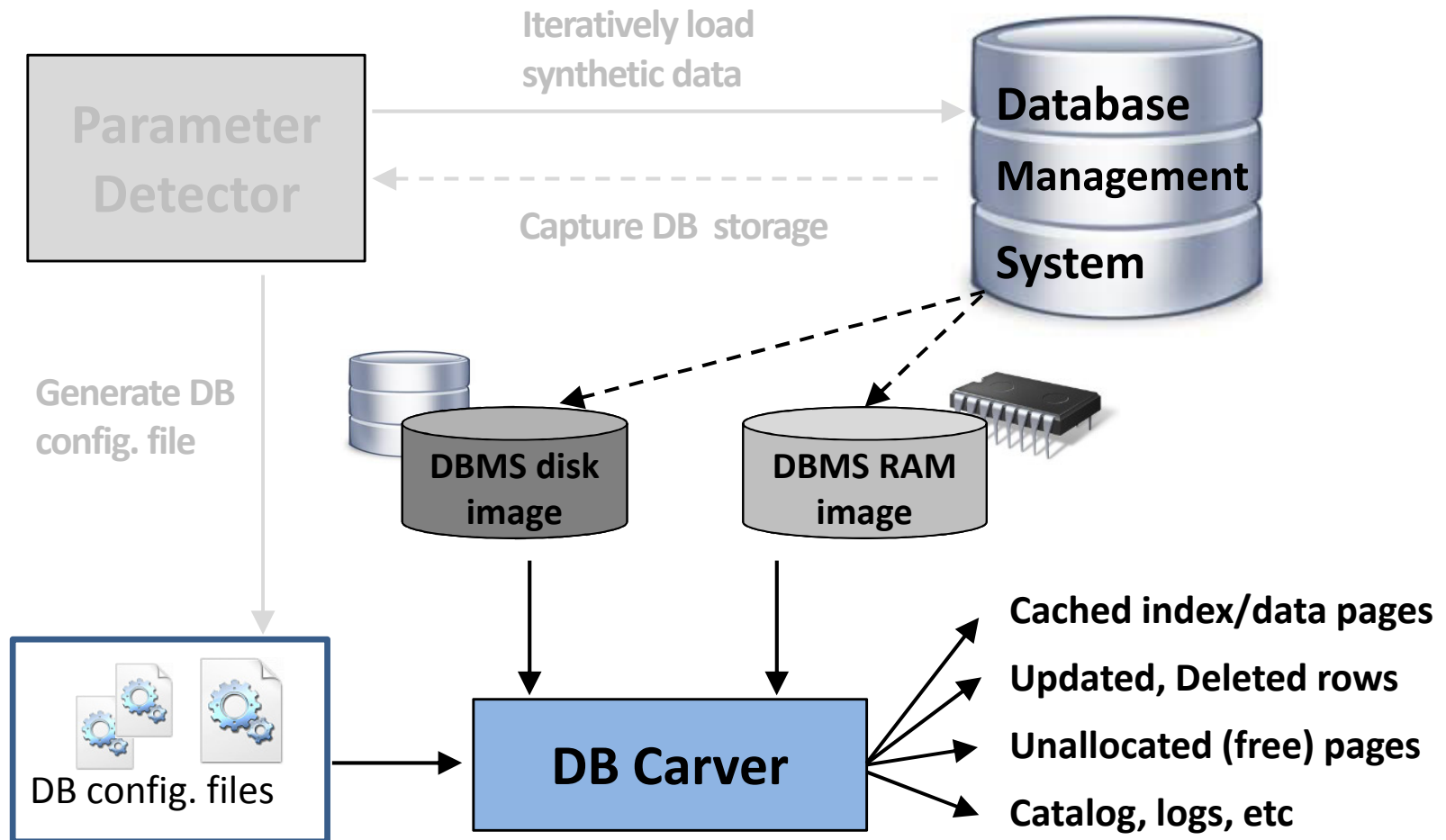
DBCarver Architecture



	Oracle	PostgreSQL	SQLite	Firebird	DB2	SQLServer	MySQL	Apache Derby
Structure Identifier	Yes	No	Yes					No
Unique Page ID	Yes							No
Row Dir. Sequence	Top-to-bottom insertion					Bottom-to-top insertion		
Row Identifier	No	Yes		No			Yes	
Column Count	Yes			No		Yes	No	Yes



DBCarver Architecture



DBCarver Output (SQLite on Android)

Page Address: 2726696960 | Page Type: Table | Record Cnt: 20 | Structure ID:

Status	RowID	Data
+	361	NULL 325 Going to our house today 325 1
+	362	NULL 326 Maybe later why 326 1
+	363	NULL 327 Before 3:30 327 1
+	364	NULL 328 Ya 328 1
+	366	NULL 330 Ok 330 1
+	367	NULL 331 Moms walking him hes cranky 331 1
+	368	NULL 332 Ok 332 1
+	379	NULL 343 Will email you a form to sign 343 1
+	380	NULL 344 When ur free call me plz 344 1
+	381	NULL 345 Cancel that...I talked w Tracey 345 1
	...	
+	389	NULL 353 They said it could take six hours 353 1
+	400	NULL 364 Drop car off tomorrow pm. Work on it we
+	401	NULL 365 Ok 365 1
-		1 NULL NULL ..Just let him out before u leave. N

Number of
Active Rows

Internal
RowID

Deleted
Row

Forensic Value of an Index (Update)

Employee Index
on (LastName)

Doe	
Jack	
Locke	
NotSmith	
Smith	

Employee Table

111	J. Doe	...	Emp.	42K
222	J. Smith	...	Emp.	35K
333	A.Locke	...	Mgr.	65K
444	P. Jack	...	Emp.	37K
222	I. NotSmith	...	Emp.	35K

Forensic Value of Caching (Update)

Memory (RAM)

Disk Storage

Data Page

111	J. Doe	...	Emp.	42K
222	J. Smith	...	Emp.	35K
333	A.Locke	...	Mgr.	65K
444	P. Jack	...	Emp.	37K

Forensic Value of Caching (Update)

**Data Page
(a copy in
RAM)**



111	J. Doe	...	Emp.	42K
222	J. Smith	...	Emp.	35K
333	A.Locke	...	Mgr.	65K
444	P. Jack	...	Emp.	37K

Memory (RAM)

Disk Storage

Data Page



111	J. Doe	...	Emp.	42K
222	J. Smith	...	Emp.	35K
333	A.Locke	...	Mgr.	65K
444	P. Jack	...	Emp.	37K

Forensic Value of Caching (Update)

**Data Page
(a copy in
RAM)**

111	J. Doe	...	Emp.	42K
222	J. Smith	...	Emp.	35K
333	A.Locke	...	Mgr.	65K
444	P. Jack	...	Emp.	37K
222	I. NotSmith	...	Emp.	35K

Memory (RAM)

Disk Storage

Data Page

111	J. Doe	...	Emp.	42K
222	J. Smith	...	Emp.	35K
333	A.Locke	...	Mgr.	65K
444	P. Jack	...	Emp.	37K

Forensic Value of Caching (Update)

**Data Page
(a copy in
RAM)**

111	J. Doe	...	Emp.	42K
222	J. Smith	...	Emp.	35K
333	A.Locke	...	Mgr.	65K
444	P. Jack	...	Emp.	37K
222	I. NotSmith	...	Emp.	35K

Memory (RAM)

Disk Storage

Data Page

111	J. Doe	...	Emp.	42K
222	J. Smith	...	Emp.	35K
333	A.Locke	...	Mgr.	65K
444	P. Jack	...	Emp.	37K
222	I. NotSmith	...	Emp.	35K

Delete Progression

- Storage state:
 - Issue the delete command
 - ??? (Profit?)
 - Value is gone
- Observe disk and RAM state
 - In Table, Index (e.g., *Unique*), MV

Delete Progression

- T_0 : Load the data (Table, Index, MV)
- T_1 : Delete a unique value (222)
- T_2 : Refresh the MV
- T_3 : Flush_buffer_cache()
- T_4 : Overwrite the buffer cache
- T_5 : Vacuum Table, Index and MV

	Table	Index	MV	Table	Index	MV
		Disk			RAM	
T_0	222	222	222			
T_1	222	222	222	222	222	
T_2	222	222	222	222	222	222
T_3	222	222	222	222	222	222
T_4	222	222	222			
T_5						

Recover Corrupted Data

- Load SSBM Scale1 data
- Simulate disk corruption (random writes)

File Percent Damage					
	0%	1%	2%	5%	10%
DWDate	2556 (100%)	2459 (96%)	2384 (93%)	2130 (83%)	2147 (84%)
Supplier	2000 (100%)	1987 (99%)	2000 (100%)	1740 (87%)	1680 (84%)
Customer	120K (100%)	118K (98%)	115K (96%)	108K (90%)	96K (80%)
Part	200K (100%)	195K (97%)	189K (94%)	174K (87%)	146K (73%)
Lineorder	6M (100%)	5.8M (97%)	5.7M (95%)	5.2M (87%)	4.5M (75%)
Full JOIN	6M (100%)	5.3M (88%)	4.9M (81%)	2.9M (49%)	1.9M (31%)

Conclusions/Future Work

- DB Carving
- No apriori assumptions
- Forensic Meta-Queries
 - Reconstruct deleted data
 - Detect recently updated values
 - Identify log tampering

