Aggregation and Degradation in JetStream: Streaming Analytics in the Wide Area

Ariel Rabkin Princeton University asrabkin@cs.princeton.edu

Work done with Matvey Arye, Siddhartha Sen, Vivek S. Pai, and Michael J. Freedman

Today's Analytics Architectures



 Backhaul is inefficient and inflexit
--

9 도표표표 6 1 % () 실 사 M %										
(time	C-10	co-usern	ane s-sterane	e-computername	1-10	1-cot	cs-method	ct-uri-stem	ce-uri-query	
22225	127.0.0.1		W3SVC1	60GBOA!	127.0.0.1	00	621	/32xcopyoct20068		
22:25:54	127.0.0.1		W39VC1	8058CA7	127.0.0.1	00	GET	(32):00pyoct20009		
22:29:54	127.0.0.1		W39VC1	80GBOAT	127.0.0.1	90	GET	(32x0pyoct2000)		
22:25:54	127.0.0.1		W39VC1	80GBOAT	\$27.0.0.1	00	GET	(32x0gyoct2006#		
22:25:54	127.0.0.1		w39VC1	8058OA7	127.0.0.1	80	681	(32x00y0x1200t)		
22:25:54	127.0.0.1		w39vC1	8058CA7	127.0.0.1	80	GET	(32x00pyoct2000)		
22:25:54	127.0.0.1		w35VC1	80GBOA7	127.0.0.1	80	Q(T	/32x0gypct2000/		
22,25,54	127.0.0.1		w39VC1	6058CA7	127.0.0.1	80	GET	(32×mpyort2000#		
22:25:54	127.0.0.1		W39VC1	6558OA7	127.0.0.3	80	CET	(32×00gypd2000)		
22:25:54	127.0.0.1		W3SVCI	60580A7	127.0.0.1	80	CET	/32x00pyoct2000#		
22:25:54	127.0.0.1		W39VCI	8058OA7	127.0.0.1	00	OCT .	(32vccevoct2000)		
22:25:54	127.0.0.1		W35VC1	6058OA7	127.0.0.1	00	OCT	(32x00yoct2000#		
22:25:55	127.0.0.1		W35VCI	60580A7	127.0.0.1	80	GET	(32x00yoct2000)		
22:29:55	127.0.0.1		W35VC1	60580A7	127.0.0.1	00	GET	(32):00pyoct20008		
22:29:55	127.0.0.1		W22VC1	6058CA7	127.0.0.1	00	GET	(32):00pyoct2000/		
22:25:55	127.0.0.1		W35VC1	60G8OA7	127.0.0.1	80	GET	(32xxxpyoct2000)		
22:25:55	127.0.0.1		W35VC1	6058OA7	127.0.0.1	00	GET	(32):00pyoct2000/		
22:20:55	127.0.0.1		W22VC5	BOGBOAT	127.0.0.1	00	CET 120	(32):copyoct2606/		
22:25:55	127.0.0.1		W39VC1	60GBOA7	\$27.0.0.1	90	GET	(32xxxpyort2006)		
22:25:55	127.0.0.1		W39YC1	80580A7	127.0.0.1	80	681	(32x00pyoct2000#		
22:29:55	127.0.0.1		W39VC1	8058OA7	127.0.0.1	80	GET	(32xxxpyoct2606)		
22:25:55	127.0.0.1		W39VC1	8058OA7	127.0.0.3	90	GET	(32x0gyoct2000#		
22:25:55	127.0.0.1		w39VC1	8058CA2	127.0.0.1	80	021	(32×mpyort2000)		
22:25:55	127.0.0.1		W39VC3	8058OA7	127.0.0.1	80	OUT	(32×000yoct2000)		
22:05:56	127.0.0.1		w39VC1	8558047	127.0.0.3	80	CET	(32x00pyort2000)		
22/25/56	127.0.0.1		w39VC1	8058OA7	127.0.0.1	80	001	(32×mpyor52000)		
22 (22.5%)	127.0.0.1		W35VC1	8058OA7	127.0.0.1	80	021	(32×00/v0/2000/		
22:25:58	127.0.0.1		W35VC1	8558047	127.0.0.1	80	CET	(32vmpyor)2000#		
22:25:58	127.0.0.1		w39VC8	6058OA7	127.0.0.1	80	QUT	[32xxxpyoct20007		
22:22:56	127.0.0.1		WOOVCE	6052OA7	127.0.0.1	00	OCT	(32xxxxyxxt2000)		
22:25:56	127.0.0.1		W35VC1	6058OA7	127.0.0.1	60	GET	(32xxxevoct2000)		
22:25:56	127.0.0.1		W35VC1	60G8OA7	127.0.0.1	80	GET	(32xxxpyoct2000)		
22:29:56	127.0.0.1		W39VC1	EDGECAT	127.0.0.1	00	CET	(32)mpyoct2000/		
22:25:56	127.0.0.1		WOSVCS	60GBOAT	127.0.0.1	90	621	(32x00pyoct2000)		
22:25:58	127.0.0.1		W39VC1	8058CAT	127.0.0.1	00	621	[32x00pyoct2006]		
22:29:58	127.0.0.1		WOSVCI	8058CA7	127.0.0.1	90	68.1	(32) copyoct2000/		
22:25:58	127.0.0.1		W39VC1	\$5GBCAT	127.0.0.1	90	GET	(32x00pyoct2000)		
22:25:58	127.0.0.1		w39VC1	EDGBCA/	127.0.0.1	00	GET	(32x00yort2000)		

Tomorrow's Architecture: JetStream



alla .									
					% 🛤 🐂	★♥ 24 34	四月11	2 1 2	9
ct-uri-stem ct-uri-query	cs-method	teot	t-ip	e-computername	s-skerone	co-usemane	100	time	
(32)xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx	681	00	127.0.0.1	polipica!	W30VC1		127.0.0.1	220350	
(32)copyoct20064 -	GET	00	127.0.0.1	#DSBCA7	W294C1		127.0.0.1	22:29:54	
(32xcoevoct2000/) -	GET	90	127.0.0.1	\$258CAT	W29YC1		127.0.0.1	22:29:54	
(32x00y0d20000 -	901	00	127.0.0.1	EDGECAT	W200C1		122.0.0.1	22:25:54	
(32x00y0d20000 -	68.1	80	127.0.0.1	foldoat	WINKI		122.0.0.1	22:25:54	
(32x00y0ct20009 -	661	80	127.0.0.1	R0GRCA7	W39/C1		127.0.0.1	22:25:54	
(32x00yart20007 -	001	80	127.0.0.1	R0GBCA7	1010001		122.0.0.1	22:05:54	
/32×mpyor52000/s -	021	80	127.0.0.1	8058CA7	w39VCI		127.0.0.1	22.25.54	
(32x00gypet2006/) -	081	80	127.0.0.1	6058047	W39VCI		127.0.0.1	22:25:54	
232xmpyort2006/1 -	001	80	127.0.0.1	60GBOAT	W39VCI	(a.	127.0.0.1	22:25:54	1
(32)reserver(2000) -	021	00	127.0.0.1	80580A7	W39VCI		127.0.0.1	22-25-54	-1
/32x00yoct2006/1 -	021	00	127.0.0.1	6058OA7	W39VCI		127.0.0.1	22:25:54	-1
(32x00pvoct2000.0 -	021	80	127.0.0.1	60580A7	W39VCI		127.0.0.1	22:25:55	-
(32xt00voct20008 -	GET	60	127.0.0.1	60580A7	W39VC1		127.0.0.1	22:29:55	
/32xxxxvoct2000/0 +	GET	00	127.0.0.1	6058CA7	W22VC1		127.0.0.1	22:29:55	-
(32x00yoct2000) -	CET .	80	127.0.0.1	60580A7	W39VC1		127.0.0.1	22:25:55	-
(32x00x0ct2000C -	GET	00	127.0.0.1	6052OA7	W39YC1		127.0.0.1	22:25:55	-
(32):000voct20006 -	GET	00	127.0.0.1	805BOAT	W22VC1		127.0.0.1	22:20:55	
(32)x00y0ct20007 -	GET	90	127.0.0.1	60GBOAT	W220C3		127.0.0.1	22-25.55	
(32x00y0ct200EA -	68.1	00	127.0.0.1	#DSECAT	W39YC1		127.0.0.1	22:25:55	
(32)(00/0012000/1 -	GET	80	127.0.0.1	sostcar	W39/C1		127.0.0.1	22:29:55	
(32x00y0d20008 -	001	90	127.0.0.1	ROGBCAT	W290C1		127.0.0.1	22:25:55	
02x00x0020000 -	011	80	127.0.0.1	6050CA7	W320C1		122.0.0.1	22.05.55	
02x00x0020008 -	(11)	80	127.0.0.1	R01ROAT	10/320/03		127.0.0.1	22 (29 55	
02x00x0020000 +	01	80	127.0.0.1	8558047	1039003		122.0.0.1	22:05:56	=
/32xeepyoet2000/1 -	001	80	127.0.0.1	8058CA7	w39VC8		127.0.0.1	22/25/56	
/32x00yoct20000 -	021	00	127.0.0.1	8058OA7	W35VCI		127.0.0.1	22 (29.56	
(32x00yod2000h -	021	80	127.0.0.1	8058047	W38VCI	in .	127.0.0.1	22:25:55	
/32xcopyoct20000 -	921	80	127.0.0.1	60580A7	W35VC8		127.0.0.1	22:25:56	-1
(32x00ypet2000/1 -	001	00	127.0.0.1	8058OA7	WORKE		127.0.0.1	22.25.56	
/32x00pvort2000/1 -	GET	60	127.0.0.1	60580A7	W39VC1		127.0.0.1	22-25-56	
(32x00yoct2000) -	921	80	127.0.0.1	60GBOAT	W39/CI		127.0.0.1	22-25-56	=1
(32)rcspvoct2000/1 -	CET .	00	127.0.0.1	605DOAT	W22VC1	1.0	127.0.0.1	22-29-56	=1
/32x00voct20000 -	68.1	90	127.0.0.1	\$05BOAT	W32VC1		127.0.0.1	22:25:56	=1
(32x00yoct2000/1 -	921	00	127.0.0.1	6050CAT	W22VC1		127.0.0.1	22:25:58	
(32)(00)voct20004 -	64.1	90	127.0.0.1	8058CA7	W29/C1		127.0.0.1	22:29:58	=1
(32):000voct2000/1 -	GAT	90	127.0.0.1	\$5GECAT	W29VC1		127.0.0.1	22:25:58	=
(32x00y0d20009 -	art	00	127.0.0.1	EDGECA!	W200C1		122.0.0.1	22:25:58	=
,									e fi

Backhaul is inefficient and inflexible
Goal: optimize use of WAN links by exposing them to streaming system.

Backhaul is Intrinsically Inefficient



Stream Processing Basics



Some Operators in JetStream:

Filtering (count > 100) Sampling (drop 90% of data) Image Compression Quantiles (95th percentile) Query stored data

The JetStream System

What: Streaming with aggregation and degradation as first-class primitives

Where: Storage and processing at edge

Why: Maximize goodput using aggregation and degradation

How: Data cubes and feedback control

An Example Query



Mechanism 1: Storage with Aggregation



Mechanism 2: Adaptive Degradation



Requirements for Storage Abstraction

• Update-able (locally and incrementally)



Data size is reducible (with predictable accuracy cost)



Merge-able (without accuracy penalty)



The Data Cube Model

Cube: A multidimensional array, indexed by a set of *dimensions*, whose cells hold *aggregates*.

Counts by URL	12:00	12:01	12:02
www.mysite.com/a	3	5	0
www.mysite.com/b	0	2	0
www.yoursite.com	5	4	
www.her-site.com	8	12	

Aggregation used for:

- Updates
- Roll-ups
- Merging cubes
- Summarizing cubes

Cubes have aggregation function: $Agg(= , =) \rightarrow =$

Cubes can be "Rolled Up"

Cube: A multidimensional array, indexed by a set of *dimensions*, whose cells hold *aggregates*.

Counts by URL	12:00	12:01	12:02
www.mysite.com/a	3	5	0
www.mysite.com/b	0	2	0
www.yoursite.com	5	4	
www.her-site.com	8	12	

Counts by URL	12:00	12:01	12:02
*	16	23	•••

Counts by URL	*
www.mysite.com/a	8
www.mysite.com/b	2
www.yoursite.com	9
www.her-site.com	20

Cubes Unify Storage and Aggregation



Degradation: The Big Picture



- Level of degradation auto-tuned to match bandwidth.
- Challenge: Supporting mergeability and flexible policies

Mergeability Imposes Constraints



Insight: Degradation may be discontinuous

There Are Many Ways to Degrade Data

• Can coarsen a dimension



Can drop low-rank values



Coarsening Does Not Always Help



Degradations Have Trade-offs

Name	Fixed BW Savings	Fixed Accuracy cost	Parameter
Dim. Coarsening	Usually no	Yes	Dimension Scale
Drop values (locally)	Yes	Νο	Cut-off
Drop values (globally)	No, multi-round protocol	Yes	Cut-off
Audiovisual downsampling	Yes	Yes	Sample rate
Histogram Coarsening	Yes	Yes	Number of Buckets

A Simple Idea that Does Not Work



- We have sensors that report congestion....
- Have operators read sensor and adjust themselves?

A Simple Idea that Does Not Work



- We have sensors that report congestion....
- Have operators read sensor and adjust themselves?

Challenge: Composite Policies



Chaos if two operators are simultaneously responding to the same sensor

Interfacing with Operators



Experimental Setup



80 nodes on VICCI testbed at three sites (Seattle, Atlanta, and Germany)





Princeton



Policy: Drop data if insufficient BW

Without Degradation



Degradation Keeps Latency Bounded



Showing maximum latencies



Programming Ease

Scenario	Lines of code
Slow requests	5
Requests by URL	5
Bandwidth by node	15
Bad referrers	16
Latency and size quantiles	25
Success by domain	30
Top 10 domains by period	40
Big Requests	97

Conclusions and Future Work

- Useful to embed aggregation and degradation abstractions in streaming systems.
- Aggregation can be unified with storage.
- System must accommodate degradation semantics.
- Open questions:
 - How to guide users to the right degradation policy?
 - How to embed abstractions in higher-level language?