Warp 10 - A novel approach for time series management and analysis based on

Mathias @Herberts - CTO, Cityzen Data

HBASECON 2017-06-12 Mountain View, CA

What is Warp 10?

Open Source Time Series Platform & Tool Suite

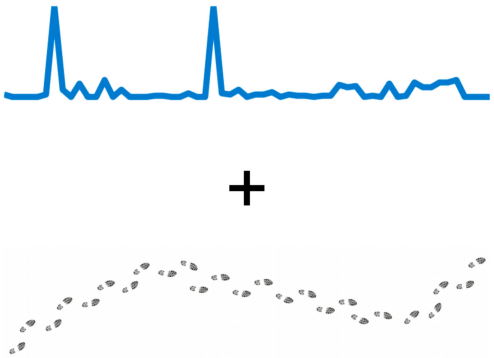
Apache 2.0 licence

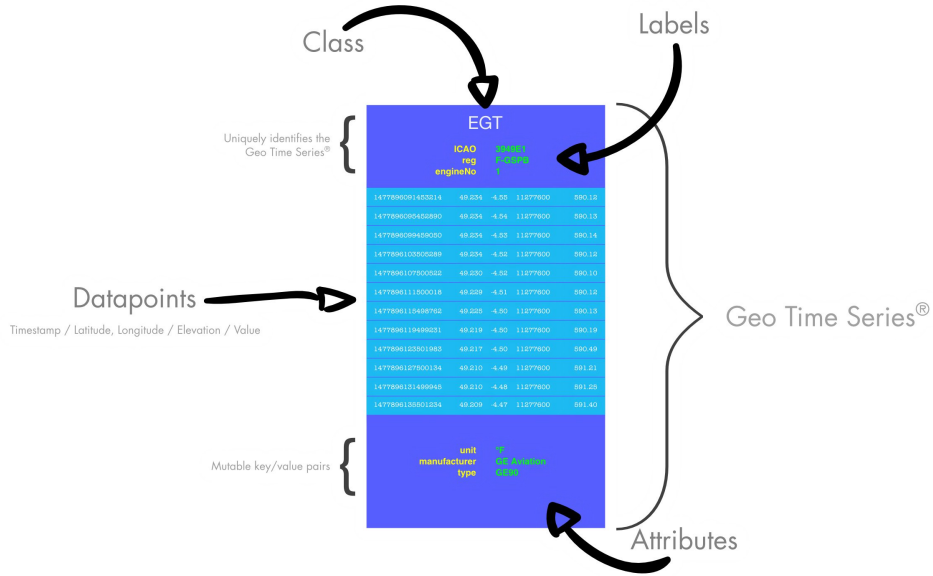
Created from scratch for IoT use cases

Complements the Hadoop Ecosystem

Embedded, standalone and distributed versions

Universal data model - Geo Time Series®

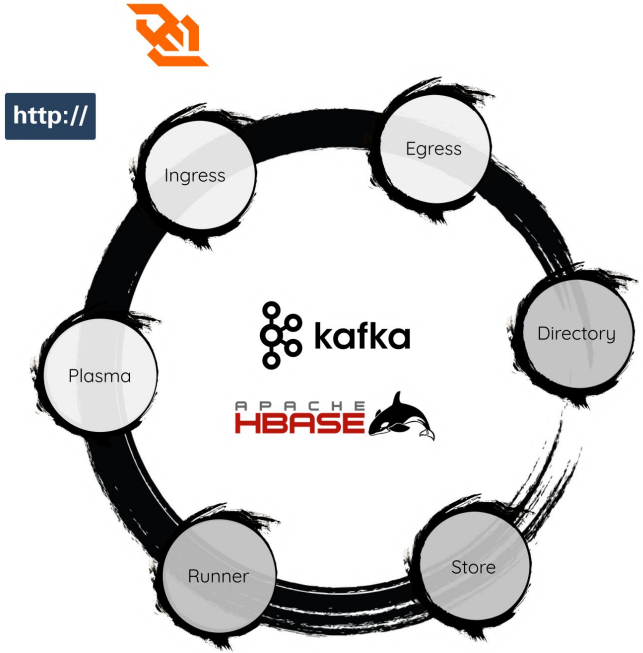




Support for LONG, DOUBLE, BOOLEAN and STRING types

Full UTF-8 support

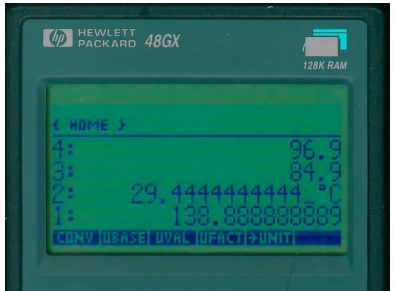
Distributed Architecture



A dedicated language for GTS Analytics

Stack based language with ~800 functions

WarpScript



WarpScript in one slide...

! != % & && * ** + +! - ->B64 ->B64URL ->BIN ->BYTES ->DOUBLEBITS ->FLOATBITS ->GEOHASH ->HEX ->HHCODE ->HHCODELONG ->JSON ->LIST ->MAP ->MAT ->OPB64 ->PICKLE ->Q ->SET ->TSELEMENTS ->V ->VEC ->Z / < << <= == > >= >> >>> ABS ACOS ADDDAYS ADDMONTHS ADDVALUE ADDYEARS AESUNWRAP AESWRAP AGO AND APPEND **APPLY** ASIN ASSERT ATAN ATBUCKET ATINDEX ATTICK ATTRIBUTES AUTHENTICATE B64-> B64TOHEX B64URL-> BBOX BIN-> BINTOHEX BITCOUNT BITGET BITSTOBYTES BOOTSTRAP BREAK BUCKETCOUNT **BUCKETIZE** BUCKETSPAN BYTES-> BYTESTOBITS BYTESTOBITS CALL CBRT CEIL CHUNK CLEAR CLEARDEFS CLEARSYMBOLS CLEARTOMARK CLIP CLONE CLONEEMPTY CLONEREVERSE COMMONTICKS COMPACT CONTAINS CONTAINSKEY CONTAINSVALUE CONTINUE COPYGEO COPYSIGN CORRELATE COS COSH COUNTER COUNTERDELTA COUNTERVALUE COUNTTOMARK CPROB CROP CSTORE CUDF DEBUGOFF DEBUGON DEDUP DEF DEFINED DEFINEDMACRO DELETE DEPTH DET DIFFERENCE DISCORDS DOC DOCMODE DOUBLEBITS-> DOUBLEEXPONENTIALSMOOTHING DROP DROPN DTW DUMP DUP DUPN DURATION DWTSPLIT E ELAPSED ELEVATIONS EMPTY ESDTEST EVAL EVALSECURE EVERY EXP EXPM1 EXPORT FAIL FDWT FETCH FETCHBOOLEAN FETCHDOUBLE FETCHLONG FETCHSTRING FFT FFTAP FILLNEXT FILLPREVIOUS FILLTICKS FILLVALUE **FILTER** FIND FINDSETS FINDSTATS FIRSTTICK FLATTEN FLOATBITS-> FLOOR FOR FOREACH FORGET FORSTEP FROMBIN FROMBITS FROMHEX FUSE GEO.DIFFERENCE GEO.INTERSECTION GEO.INTERSECTS GEO.REGEXP GEO.UNION GEO.WITHIN GEO.WKT GEOHASH-> GEOPACK GEOUNPACK GET GETHOOK GETSECTION GRUBBSTEST GZIP HASH HAVERSINE HEADER HEX-> HEXTOB64 HEXTOBIN HHCODE-> HUMANDURATION HYBRIDTEST HYBRIDTEST2 HYPOT IDENT IDWT IEEEREMAINDER IFFT IFT IFTE IMMUTABLE INTEGRATE INTERPOLATE INTERSECTION INV ISNULL ISNaN ISO8601 ISODURATION ISONORMALIZE JOIN JSON-> JSONLOOSE JSONSTRICT KEYLIST LABELS LASTBUCKET LASTSORT LASTTICK LBOUNDS LFLATMAP LIMIT LIST-> LMAP LOAD LOCATIONOFFSET LOCATIONS LOCSTRINGS LOG LOG10 LOG1P LORAENC LORAMIC LOWESS LR LSORT LTTB MACROBUCKETIZER MACROFILTER MACROMAPPER MACROREDUCER MAKEGTS **MAP** MAP-> MAPID MARK MAT-> MATCH MATCHER MAX MAXBUCKETS MAXDEPTH MAXGTS MAXLONG MAXLOOP MAXOPS MAXPIXELS MAXSYMBOLS MD5 MERGE META METASET METASORT MIN MINLONG MODE MONOTONIC MSGFAIL MSORT MSTU MUSIGMA NAME NBOUNDS NDEBUGON NEWGTS NEXTAFTER NEXTUP NONEMPTY NOOP NORMALIZE NOT NOTAFTER NOTBEFORE NOTIMINGS NOW NPDF NRETURN NSUMSUMSQ NULL NaN ONLYBUCKETS OPB64-> OPB64TOHEX OPS OPTDTW OR PACK PAPPLY PARSE PARSESELECTOR PARTITION PATTERNDETECTION PATTERNS PFILTER PGraphics PI PICK PICKLE-> PIGSCHEMA PREDUCE PROB PROBABILITY PUT Palpha Parc Pbackground PbeginContour PbeginShape Pbezier PbezierDetail PbezierPoint PbezierTangent PbezierVertex Pblend PblendMode Pblue Pbox Pbrightness Pclear Pclip Pcolor PcolorMode Pconstrain Pcopy PcreateFont Pcurve PcurveDetail PcurvePoint PcurveTangent PcurveTightness PcurveVertex Pdecode Pdist Pellipse PellipseMode Pencode PendContour PendShape Pfill Pget Pgreen Phue Pimage PimageMode Plerp PlerpColor Pline Pmag Pmap PnoClip PnoFill PnoStroke PnoTint Pnorm Ppixels Ppoint PpopMatrix PpopStyle PpushMatrix PpushStyle Pquad PquadraticVertex Prect PrectMode Pred PresetMatrix Protate ProtateX ProtateY ProtateZ Psaturation Pscale Pset PshapeMode PshearX PshearY Psphere PsphereDetail Pstroke PstrokeCap PstrokeJoin PstrokeWeight Ptext PtextAlign PtextAscent PtextDescent PtextFont PtextLeading PtextMode PtextSize PtextWidth Ptint Ptranslate Ptriangle PupdatePixels Pvertex Q-> QCONJUGATE QDIVIDE QMULTIPLY QROTATE QROTATION QUANTIZE RAND RANDPDF RANGE RANGECOMPACT REDEFS **REDUCE** RELABEL REMOVE RENAME REPLACE REPLACEALL RESET RESETS RESTORE RETURN REV REVBITS REVERSE REXEC REXECZ RINT RLOWESS ROLL ROLLD ROT ROTATIONQ ROUND RSADECRYPT RSAENCRYPT RSAGEN RSAPRIVATE RSAPUBLIC RSASIGN RSAVERIFY RSORT RTFM RUN RUNNERNONCE RVALUESORT SAVE SECTION SECUREKEY SET SET-> SETATTRIBUTES SETVALUE SHA1 SHA1HMAC SHA256 SHA256HMAC SHRINK SIGNUM SIN SINGLEEXPONENTIALSMOOTHING SINH SIZE SNAPSHOT SNAPSHOTALL SNAPSHOTALLTOMARK SNAPSHOTCOPY SNAPSHOTCOPYALL SNAPSHOTCOPYALLTOMARK SNAPSHOTCOPYTOMARK SNAPSHOTTOMARK SORT SORTBY SPLIT SQRT STACKATTRIBUTE STACKTOLIST STANDARDIZE STL STLESDTEST STOP STORE STRICTMAPPER STRICTPARTITION STRICTREDUCER STU SUBLIST SUBMAP SUBSTRING SWAP SWITCH TAN TANH TEMPLATE TEMPLATE THRESHOLDTEST TICKINDEX TICKLIST TICKS TIMECLIP TIMEMODULO TIMESCALE TIMESHIFT TIMESPLIT TIMINGS TLTTB TOBIN TOBITS TOBOOLEAN TODEGREES TODOUBLE TOHEX TOKENINFO TOLONG TOLOWER TORADIANS TOSELECTOR TOSTRING TOTIMESTAMP TOTIMESTAMP TOUPPER TR TRANSPOSE TRIM TSELEMENTS TSELEMENTS-> TYPEOF UDF ULP UNBUCKETIZE UNGZIP UNION UNIQUE UNLIST UNMAP UNPACK UNSECURE UNTIL UNWRAP UNWRAPEMPTY UNWRAPSIZE UPDATE URLDECODE URLENCODE UUID V-> VALUEDEDUP VALUEHISTOGRAM VALUELIST VALUES VALUESORT VALUESPLIT VEC-> WEBCALL WHILE WRAP WRAPOPT WRAPRAW WRAPRAWOPT Z-> ZDISCORDS ZIP ZPATTERNDETECTION ZPATTERNS ZSCORE ZSCORETEST [[]] ^ bucketizer.and bucketizer.count bucketizer.count.exclude-nulls bucketizer.count.include-nulls bucketizer.count.nonnull bucketizer.first bucketizer.join bucketizer.join.forbid-nulls bucketizer.last bucketizer.mad bucketizer.max bucketizer.max.forbid-nulls bucketizer.mean bucketizer.mean.circular bucketizer.mean.circular.exclude-nulls bucketizer.mean.exclude-nulls bucketizer.median bucketizer.min bucketizer.min.forbid-nulls bucketizer.or bucketizer.percentile bucketizer.sum bucketizer.sum.forbid-nulls d e filter.byattr filter.byclass filter.bylabels filter.bylabelsattr filter.bymetadata filter.last.eq filter.last.ge filter.last.gt filter.last.le filter.last.lt filter.last.ne filter.latencies h m mapper.abs mapper.abscissa mapper.add mapper.and mapper.ceil mapper.count mapper.count.exclude-nulls mapper.count.include-nulls mapper.count.nonnull mapper.day mapper.delta mapper.distinct mapper.dotproduct mapper.dotproduct.positive mapper.dotproduct.sigmoid mapper.dotproduct.tanh mapper.eq mapper.exp mapper.finite mapper.first mapper.floor mapper.ge mapper.geo.approximate mapper.geo.clear mapper.geo.outside mapper.geo.within mapper.gt mapper.hdist mapper.highest mapper.hour mapper.hspeed mapper.join mapper.join.forbid-nulls mapper.kernel.cosine mapper.kernel.epanechnikov mapper.kernel.gaussian mapper.kernel.logistic mapper.kernel.quartic mapper.kernel.silverman mapper.kernel.triangular mapper.kernel.tricube mapper.kernel.triweight mapper.kernel.uniform mapper.last mapper.le mapper.log mapper.lowest mapper.lt mapper.mad mapper.max mapper.max.forbid-nulls mapper.max.x mapper.mean mapper.mean.circular mapper.mean.circular.exclude-nulls mapper.mean.exclude-nulls mapper.median mapper.min mapper.min.forbid-nulls mapper.min.x mapper.minute mapper.mod mapper.month mapper.mul mapper.ne mapper.npdf mapper.or mapper.parsedouble mapper.percentile mapper.pow mapper.product mapper.rate mapper.replace mapper.round mapper.sd mapper.sd.forbid-nulls mapper.second mapper.sigmoid mapper.sum mapper.sum.forbid-nulls mapper.tanh mapper.tick mapper.toboolean mapper.todouble mapper.tolong mapper.tostring mapper.truecourse mapper.var mapper.var.forbid-nulls mapper.vdist mapper.vspeed mapper.weekday mapper.year max.tick.sliding.window max.time.sliding.window ms ns op.add op.add.ignore-nulls op.and op.and.ignore-nulls op.div op.eq op.ge op.gt op.le op.lt op.mask op.mul op.mul.ignore-nulls op.ne op.negmask op.or op.or.ignore-nulls op.sub pi ps reducer.and reducer.and.exclude-nulls reducer.argmax reducer.argmin reducer.count reducer.count.exclude-nulls reducer.count.include-nulls reducer.count.nonnull reducer.join reducer.join.forbid-nulls reducer.join.nonnull reducer.join.urlencoded reducer.mad reducer.max reducer.max.forbid-nulls reducer.max.nonnull reducer.mean reducer.mean.circular reducer.mean.circular.exclude-nulls reducer.mean.exclude-nulls reducer.median reducer.min reducer.min.forbid-nulls reducer.min.nonnull reducer.or reducer.or.exclude-nulls reducer.percentile reducer.product reducer.sd reducer.sd.forbid-nulls reducer.shannonentropy.0 reducer.shannonentropy.1 reducer.sum reducer.sum.forbid-nulls reducer.sum.nonnull reducer.var reducer.var.forbid-nulls s us w { {} | || } ~ ~=

Fully extensible and very flexible

WarpScript extensions

CALL of external programs (such as *TensorFlow*)

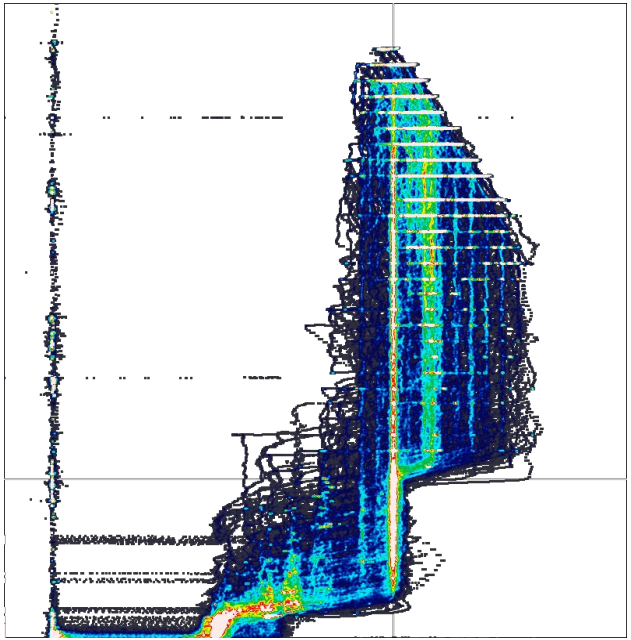
Embeddable in third party applications

Usable with any time series datasource

Integrated with Pig, Flink, Spark and Storm

Integrates visualization features of *Processing*

```
800 'width' STORE 800 'height' STORE
400.0 'maxspeed' STORE 40000.0 'maxalt' STORE
3.0 2.0 2.0 @orbit/heatmap/kernel/triangular 'kernel' STORE
@orbit/heatmap/palette/classic 'palette' STORE
'TOKEN"token' STORE

$width $height '2D' PGraphics
'MULTIPLY' PblendMode 'CENTER' PimageMode
[ $token '~(ALT|CAS)' {} NOW -2000000 ] FETCH
DUP 0 GET LASTTICK 'now' STORE
[ SWAP bucketizer.last $now STU 0 ] BUCKETIZE

// Create heatmap
<%
  7 GET LIST-> DROP 'CAS' STORE 'ALT' STORE
  <% $CAS ISNULL NOT $ALT ISNULL NOT && %>
  <% $kernel $CAS $maxspeed / $width * $ALT $maxalt / 1.0
SWAP - $height * Pimage %>
  IFT
  0 NaN NaN NaN NULL
%> MACROREDUCER 'GRAPHER' STORE
[ SWAP [] $GRAPHER ] REDUCE DROP

// Colorize
Ppixels <% DROP Palpha $palette SWAP GET %> LMAP
PupdatePixels Pencode Pdecode
$width $height '2D' PGraphics

// Do the grid
PnoFill 0 0 $width 1 - $height 1 - Prect
2.0 PstrokeWeight 200.0 Pcolor Pstroke
250.0 $maxspeed / $width * DUP 0 SWAP $height Pline
0 10000 $maxalt / 1.0 SWAP - $height * DUP $width SWAP Pline

SWAP 0 0 Pimage Pencode
```



HBase Schema

No lookup keys

Prefix	Class Id	Labels Id	Reversed Timestamp
	64/128 bits SipHash on Class Name	64/128 bits SipHash on Labels Names & Values	Long.MAX_VALUE - ts

2 column families, *m* and *v*

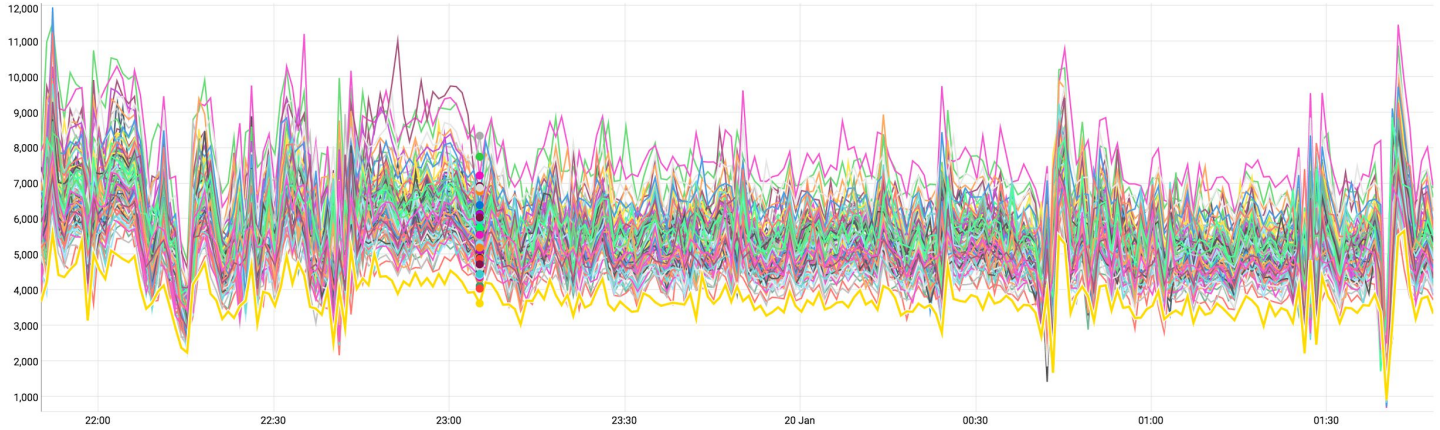
m for metadata, null cq, encrypted serialized thrift structures

v for individual values, null cq, *GTSEncoder* content, possibly encrypted

With FASTDIFF_PREFIX, down to ~10 bytes per cell

No automatic compaction, possibility to *pack* chunks of GTS as STRING values

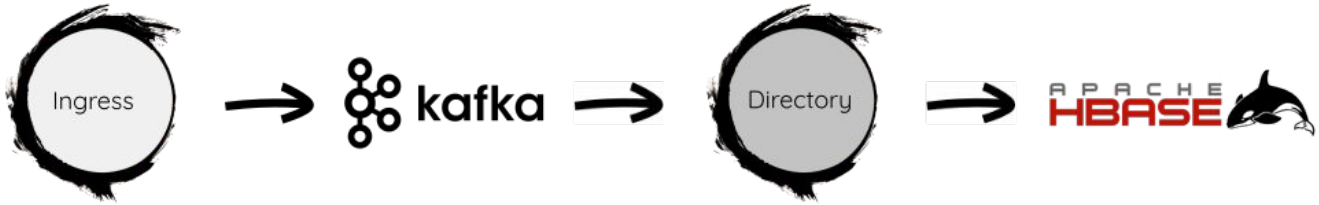
Hash based keys achieve high write distribution



108 Region Servers, typical IT monitoring load (50M active series), ~800k datapoints/s

Write Path

Metadata write path

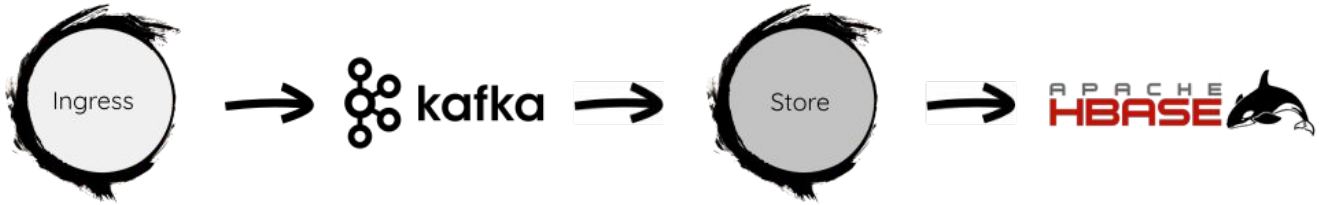


Metadata structures pushed when newly encountered or modified by *Ingress*

Metadata structures saved as they are consumed by *Directory* instances

Content in HBase converges towards latest version

Data write path



Push batches of Put to HBase with time and size thresholds

Reset Kafka consumption when HBase errors are encountered

Sensitive to HBase RS slowdowns due to good key distribution

Write performance mainly driven by Kafka partition count and available CPUs in Store

Deletions

Deletion process

EU Legal requirement to provide deletion capabilities for hosted services

Directory is accessed to retrieve GTS Metadata

Can use the *data* topic or a dedicated one depending on chosen semantics

Delete messages pushed to Kafka, one per GTS to delete (partial or complete)

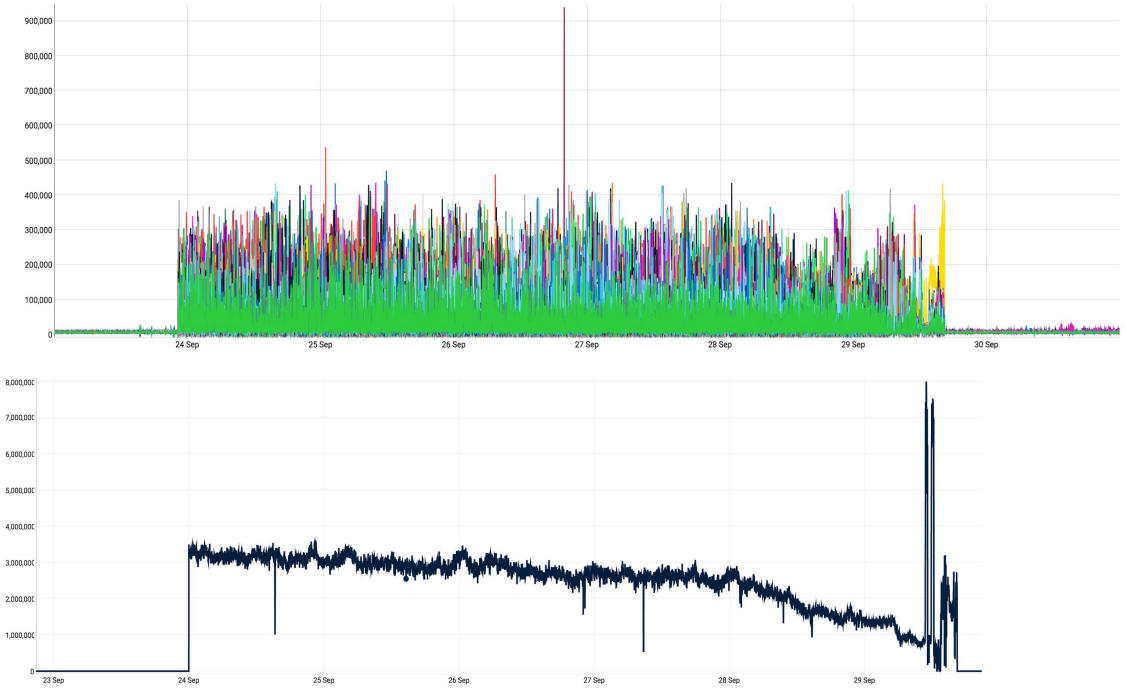
Directory will remove Metadata from its memory and from HBase for complete deletes

Use of the *BulkDelete* coprocessor endpoint from the Store daemons

Deletes and Puts are treated in order of arrival for sequential consistency

Problem of RegionServer becoming *hot* when deleting certain GTS, rely on sorting GTS as first step

Deletion effect on writes/s on RS



1.3T datapoints deleted over the course of 6 days while ingesting 800k/s datapoints

Reads occur simultaneously are roughly 2.75M/s to enable the deletes

Read Path

Life of a fetch query

Retrieve Geo Time Series® metadata from Directory

Fetch cells from HBase

Merge cells into one *GTSEncoder* per GTS

Expand *GTSEncoder* into a data structure WarpScript can understand

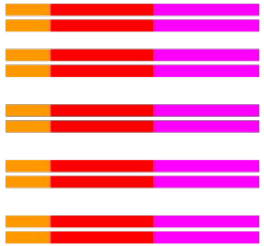
Initial (and fallback) method uses one Scanner per GTS fetched

Only scans data we will retrieve, but using many scanners is a real performance hit

Enters a custom filter

SlicedRowFilter

Concatenation of *slices* of the row key



Ranges of key slices, passed as filter parameter

Filter behavior

Row key is *sliced*

Sliced key is compared with valid ranges, row is accepted at first match

When slices form a prefix of the row key, hinting is possible

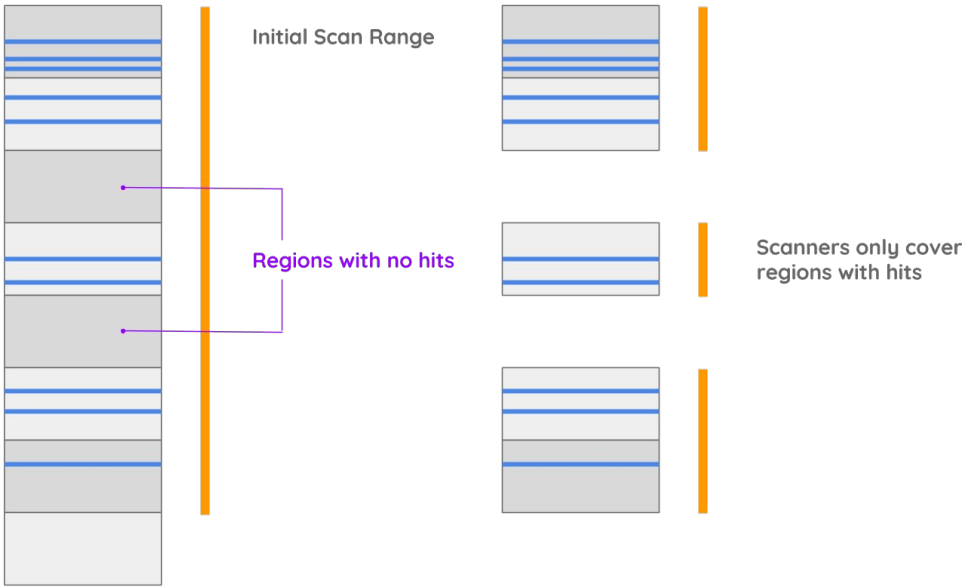
When encountering a row key outside of a valid range, hint to seek to the next range

Only scan retrieved rows, except for the first row of each scanned region which may be skipped

All done in a single Scanner instance

Fetch reloaded

When many regions with no hits in the scanner range, useless open can happen



Only regions with hits are open, all (but 1) rows scanned per region are returned

Fetch revolutions

Split GTS to retrieve in smaller batches which will be treated independently

Use a thread pool to issue multiple Scans in parallel

Reach performance of multiple million datapoints/s

Able to saturate 1Gbps links of RS

Warp10InputFormat

Hadoop InputFormat to retrieve data from Warp 10

One InputSplit per Geo Time Series®

InputSplit combined up to a certain number of GTS, grouped by RS (of most recent datapoint)

Fetcher daemons colocated with Region Servers

Perform fetches as described earlier

With parallel scanners can easily retrieve 5M datapoints/s per fetcher

Conclusion

HBase really rocks for time series data!!!

Test drive Warp 10 in standalone mode (no HBase needed)

```
curl -O -L https://dl.bintray.com/cityzendata/generic/io/warp10/warp10/1.2.7-rc2/warp10-1.2.7-rc2.tar.gz
tar zxpf warp10-1.2.7-rc2.tar.gz
export JAVA_HOME=/path/to/java/home; cd warp10-1.2.7-rc2; ./bin/warp10-standalone.init start
```

Let's talk about your time series projects

@warp10io

http://www.warp10.io/

http://groups.google.com/forum/#!forum/warp10-users

https://github.com/cityzendata