Industrial Automation and Control
Automation College

uniformance PHD Data Processing Overview

Lesson Objective

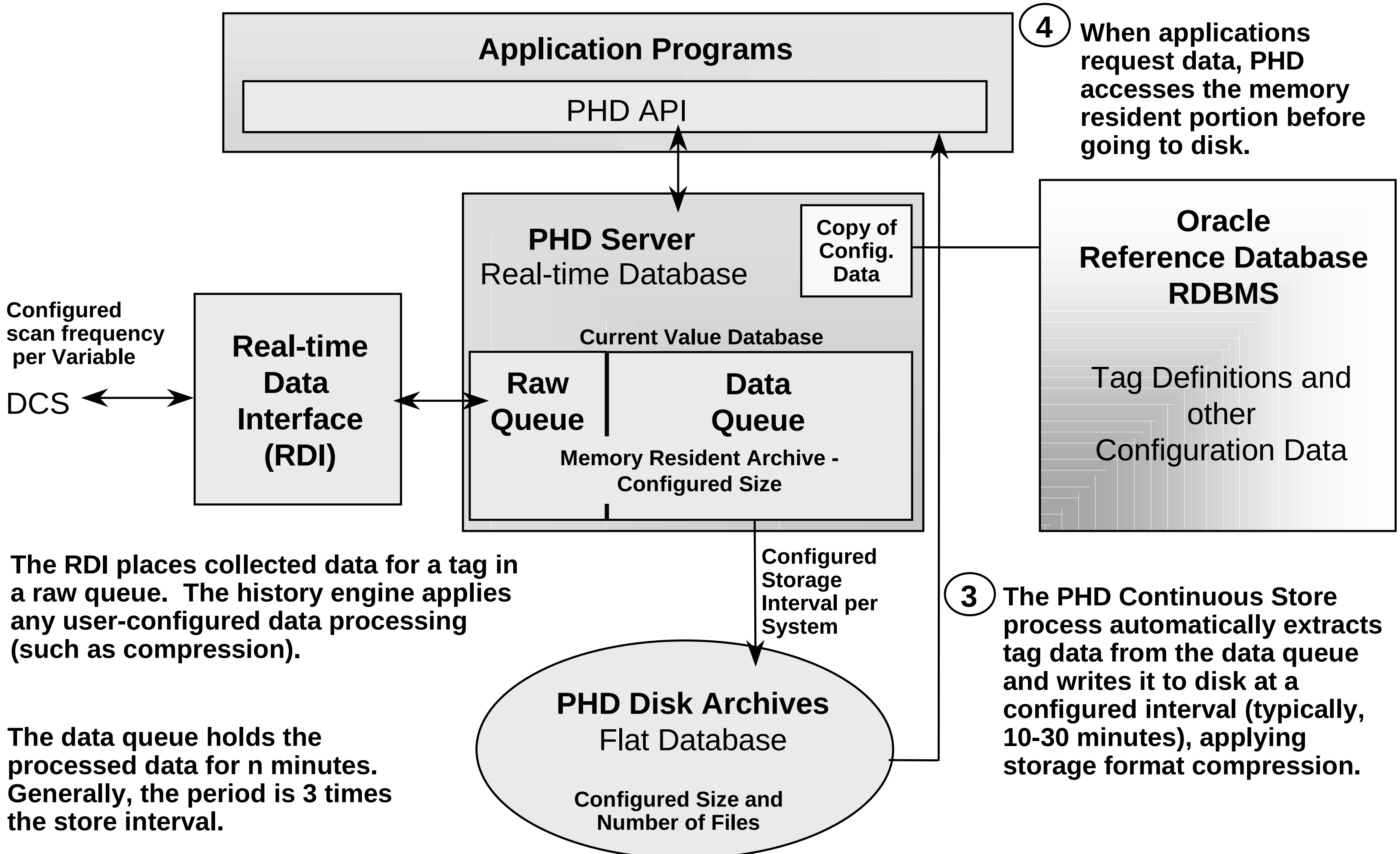
Objective

- Recognize the purpose of the various PHD data processing options

Topics

- System Components and Data Flow
- Gross Error Elimination
- Data Smoothing and Noise Gating
- Data Elimination Compression
- Data Quantizing and Scaling (Storage Format Compression)
- Confidence Factor
- Intelligent Data Retrieval:
    - Time Weighted Data Reductions
    - Interpolation and Extrapolation
    - Engineering Unit Conversion

System Components and Data Flow



1. The RDI places collected data for a tag in a raw queue. The history engine applies any user-configured data processing (such as compression).

2. The data queue holds the processed data for n minutes. Generally, the period is 3 times the store interval.

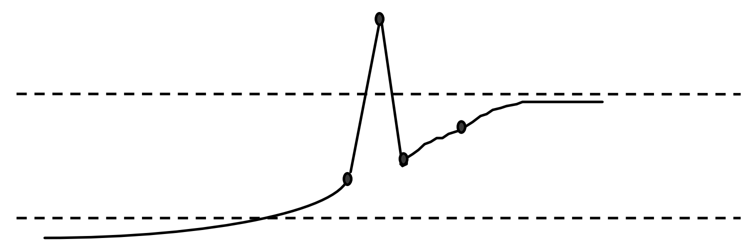
3. The PHD Continuous Store process automatically extracts tag data from the data queue and writes it to disk at a configured interval (typically, 10-30 minutes), applying storage format compression.

4. When applications request data, PHD accesses the memory resident portion before going to disk.

Typically, PHD sites keep high resolution (scan rate) data on-line for 2 to 3 years.

Front End Processing - Statistical Gross Error Elimination

Configuration option - PHD evaluates values to eliminate random instrumentation errors. Performs linear regression on values before and after the value being analyzed. Determines standard deviation of surrounding values. If difference between value and regression line estimate is greater than the threshold, the value is considered a gross error. The system keeps a log of gross errors.



Tag Definitions:

Gr Err Sigma Limit Number of standard deviations that the tag can vary from the regression estimate. Typical Limit = 3 to 5 Default = 0, disables gross error detection for tag.

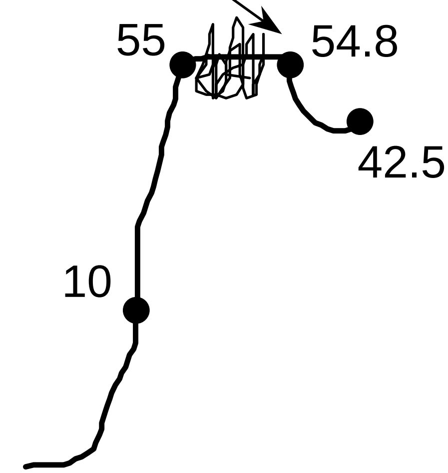
Gr Err Samples Total number of samples used for regression estimate and standard deviation. Typical Samples = 6 or 8

Example: Samples = 6
3 values prior and 3 values after the value being evaluated are used for the regression/deviation estimate.

(see *PHD System Manual,* Data Processing and Formatting and *PHD User Guide*, Forms)

Front End Processing - Data Smoothing

Configuration option used to minimize the effect of random noise on measured values.



Smoothing Constant

Exponential decay filter + feed forward digital filter.

Value of 0.99 indicates maximum smoothing. Typical factor = 0.3 - 0.5

Default = 0, no smoothing.

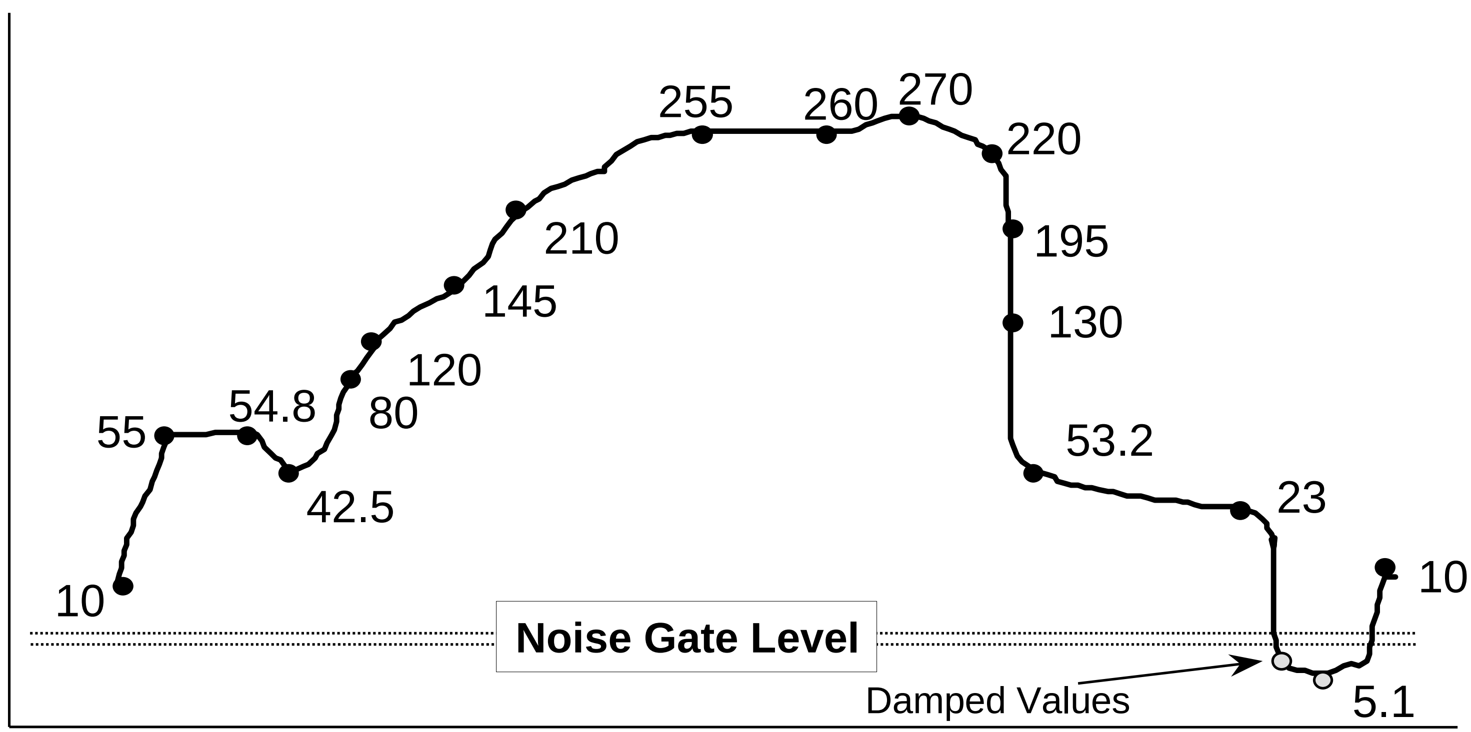
(see *PHD System Manual,* Data Processing and Formatting and *PHD User Guide*, Forms)

Front End Processing - Noise Gating

Configuration option used to minimize the effect of random noise on measured values. Default = no noise gating

Noise Gate Level

Low noise threshold value below which values will be damped to zero.



(see *PHD System Manual,* Data Processing and Formatting and *PHD User Guide*, Forms)

Data Elimination Compression

PHD Data elimination compression reduces the volume of stored data by eliminating redundant values and values that can be interpolated within a specified compression tolerance. The integrity and accuracy of the data are maintained.

In cases where you must preserve the shape of your data curve, such as modeling for advanced process control, it may be necessary for your PHD manager to adjust the tag parameters that control this processing option or disable it.

If compression tolerance = -1, PHD stores ALL values (ex:environmental data). No Compression.

If compression tolerance = 0, PHD stores all but the redundant values. Default = 0

Data Elimination Compression, *continued*

Constrained slope data compression for floating point data

PHD stores floating point data such as inventory measurements and flow rates using Honeywell's constrained slope data compression process. Floating point values that can be successfully interpolated from stored data to within the compression tolerance are eliminated.

PHD calculates the compression tolerance using the equation:

compression tolerance = Tolerance x Compr Toler Factor

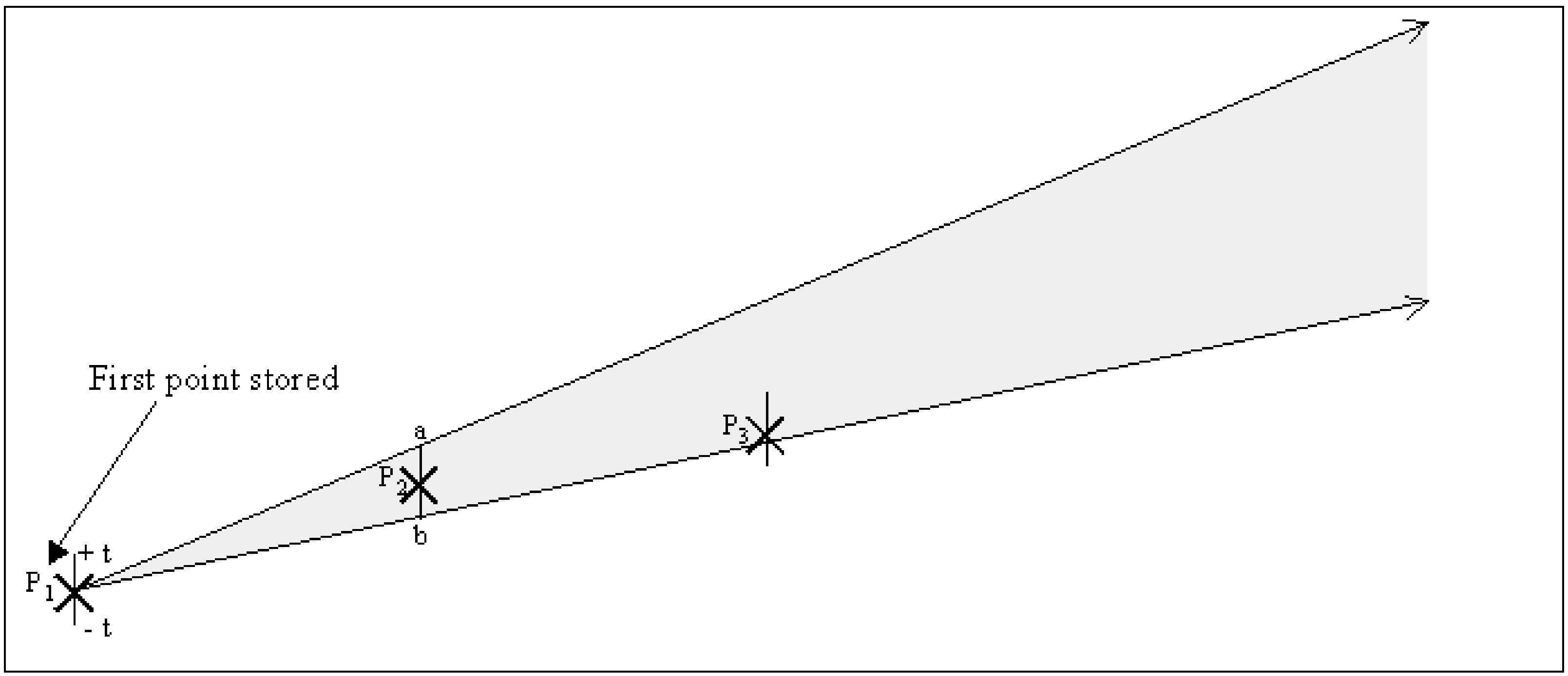
where Tolerance and Compr Toler Factor are values specified in the tag configuration.

Tag Definition	Example Value	Description
Tolerance type	Percent	Absolute value in EUs, or a percentage of the reading, or a percentage of the tag's span.
Tolerance	± .5 (%)	Tolerance Value—Low tolerance values are used for more accurate measurement devices.
Compression Tolerance Factor	.1 (10%)	Tolerance Multiplier—Higher values provide more tolerance and more compression.

(see *PHD System Manual,* Data Processing and Formatting and *PHD User Guide*, Forms)

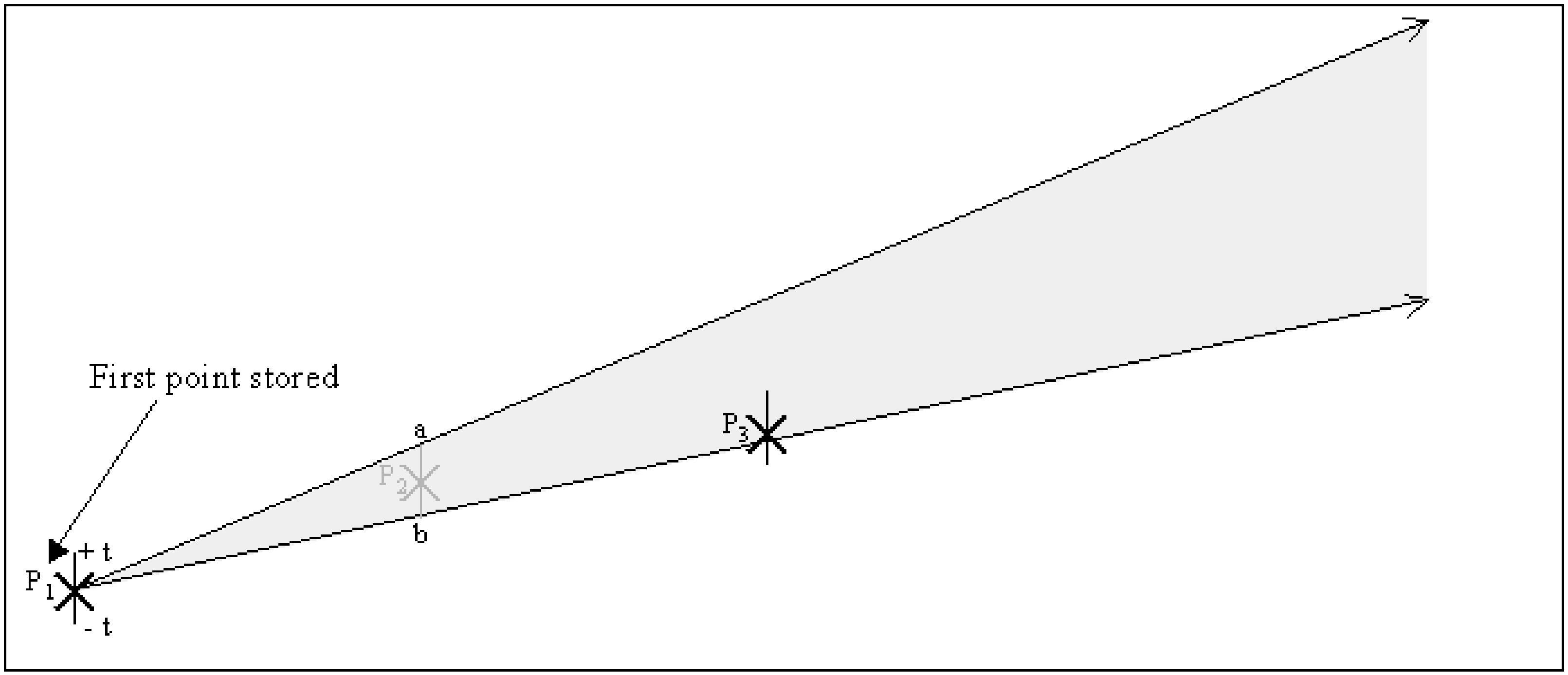
Data Elimination Compression, *continued*

To begin the constrained slope data compression process, PHD stores the first point it collects for the tag, P_1. PHD draws rays connecting P_1 to points a and b. Point a is equal to the next new point, P_2, plus the compression tolerance, t. Point b is equal to P_2 minus the compression tolerance.



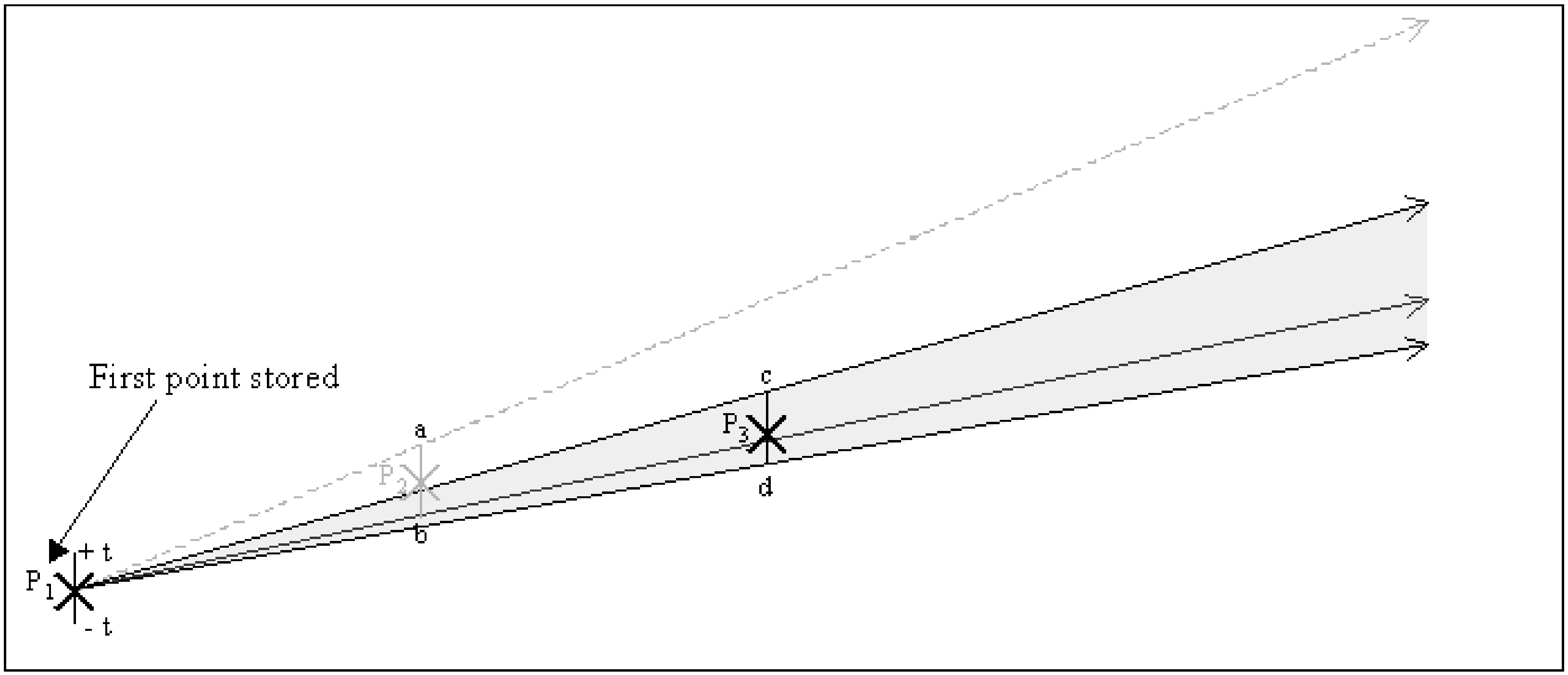
Data Elimination Compression, *continued*

PHD checks the next new point, P_3, to determine whether it lies inside the angle aP_1b. P_3 lies inside the angle so P_2 is not stored. However, the rays drawn to $P_2 \pm t$ continue to play a part in the process.



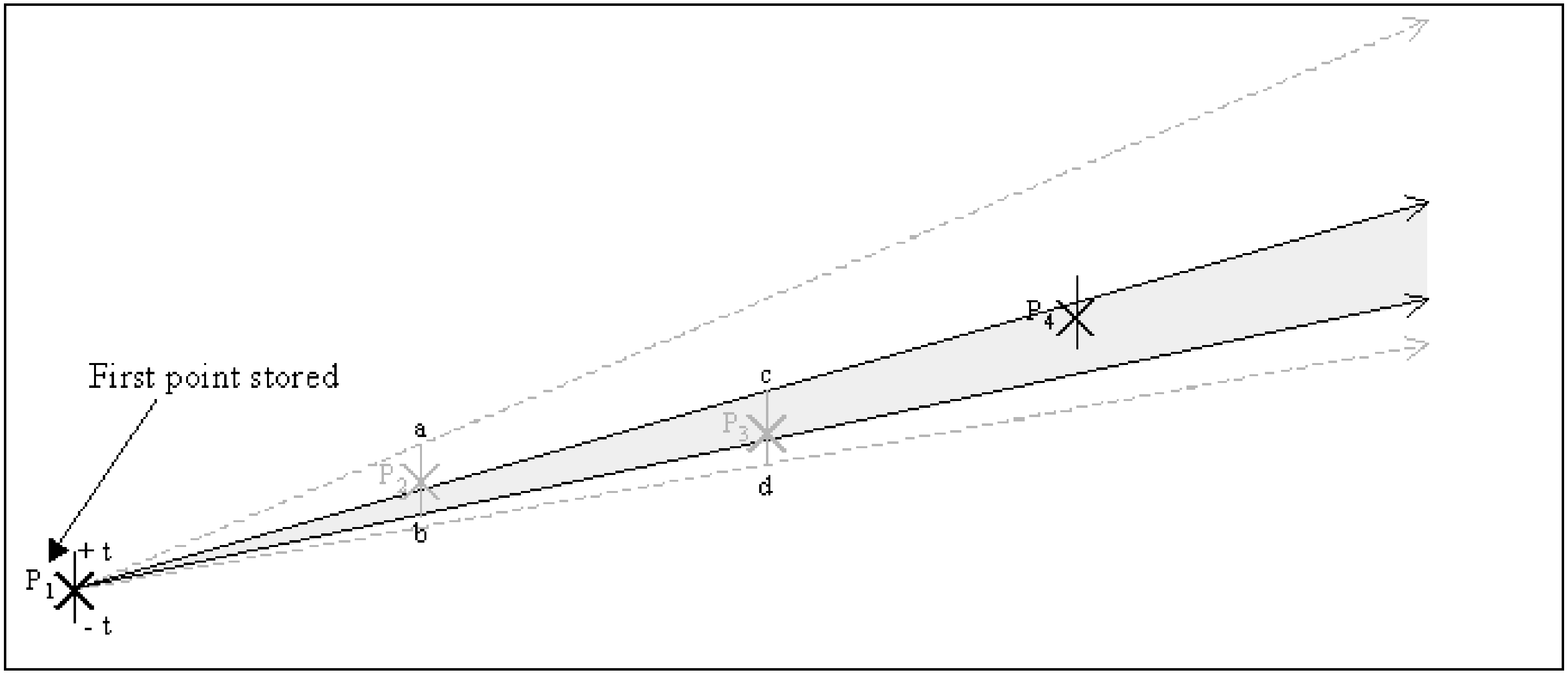
Data Elimination Compression, *continued*

PHD draws rays from P_1 to points c and d, equal to $P_3 \pm t$. P_3 lies inside the angle cP_1d but it also lies inside the smaller angle cP_1b.



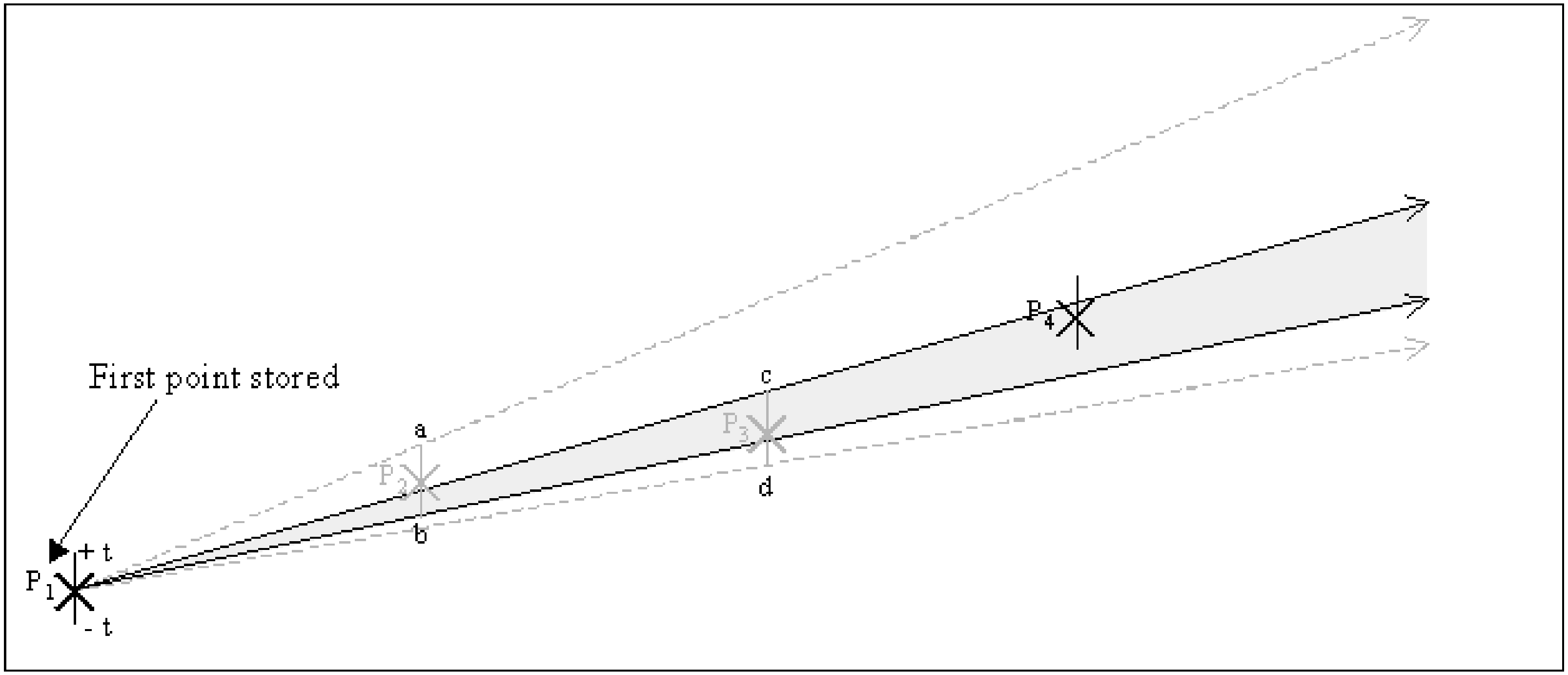
Data Elimination Compression, *continued*

PHD constrains the area in which the next new point must lie to the smaller angle cP_1b.



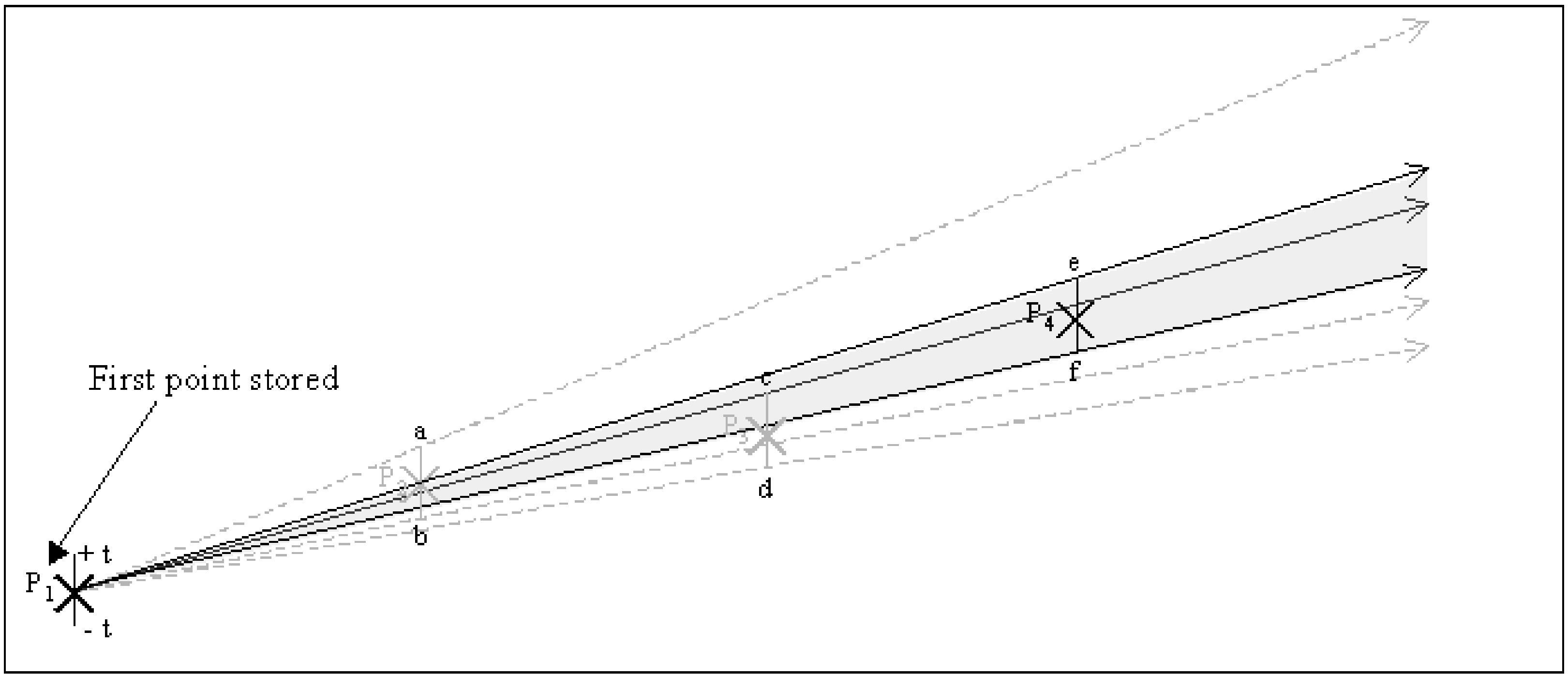
Data Elimination Compression, *continued*

PHD checks the next new point, P_4, to determine whether it lies inside the angle cP_1b. P_4 lies inside the angle so P_3 is not stored. However, the rays drawn from $P_3 \pm t$ continue to play a part in the process.



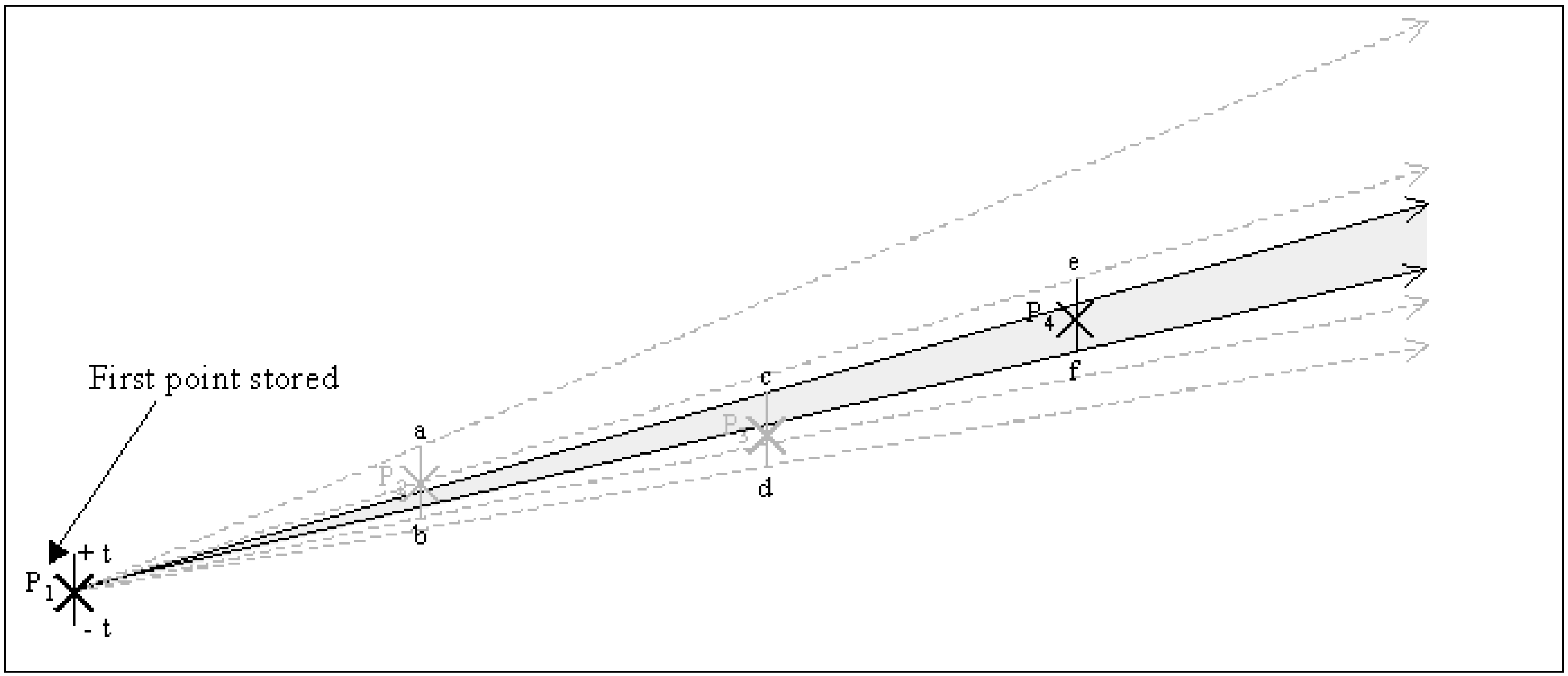
Data Elimination Compression, *continued*

PHD draws rays from P_1 to points e and f, equal to $P_4 \pm t$. P_4 lies inside the angle eP_1f but it also lies inside the smaller angle cP_1f.



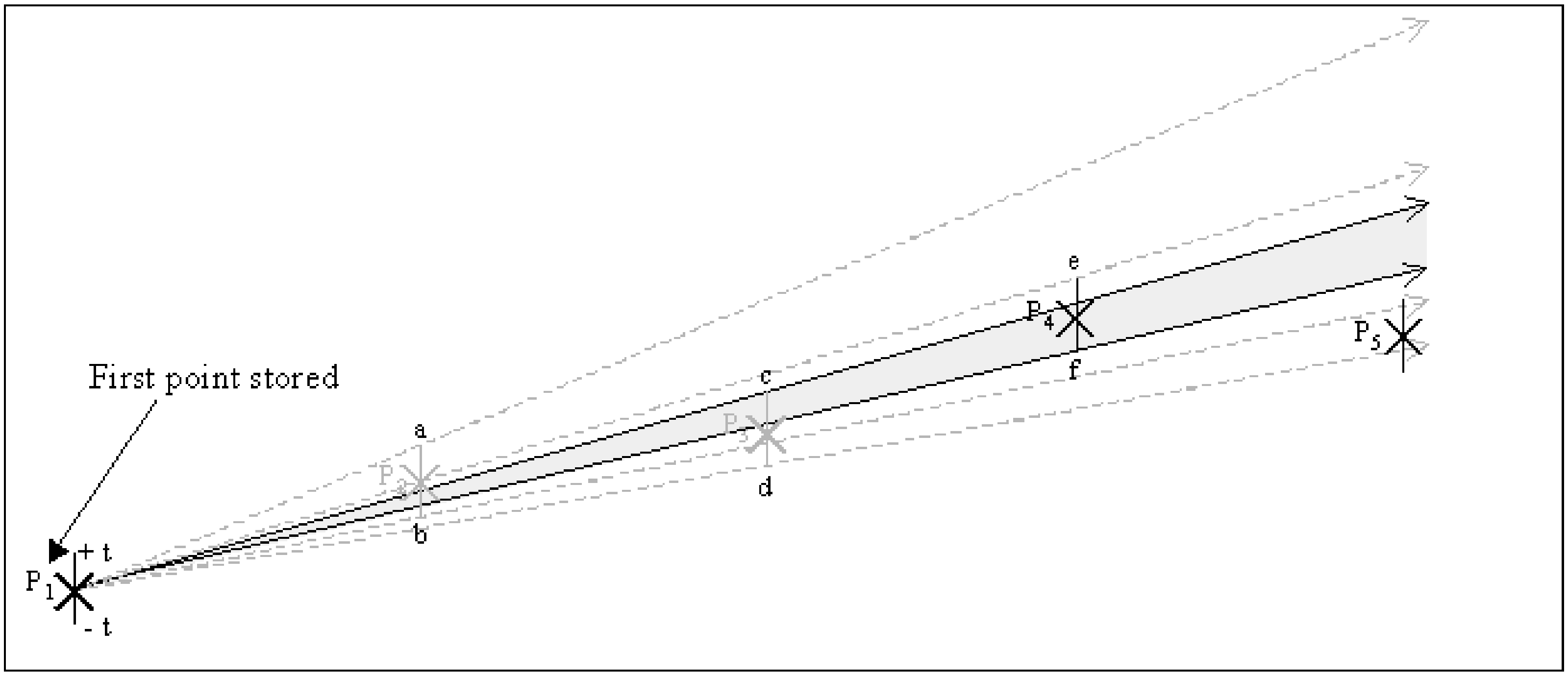
Data Elimination Compression, *continued*

PHD constrains the area in which the next new point must lie to the smaller angle cP_1f.



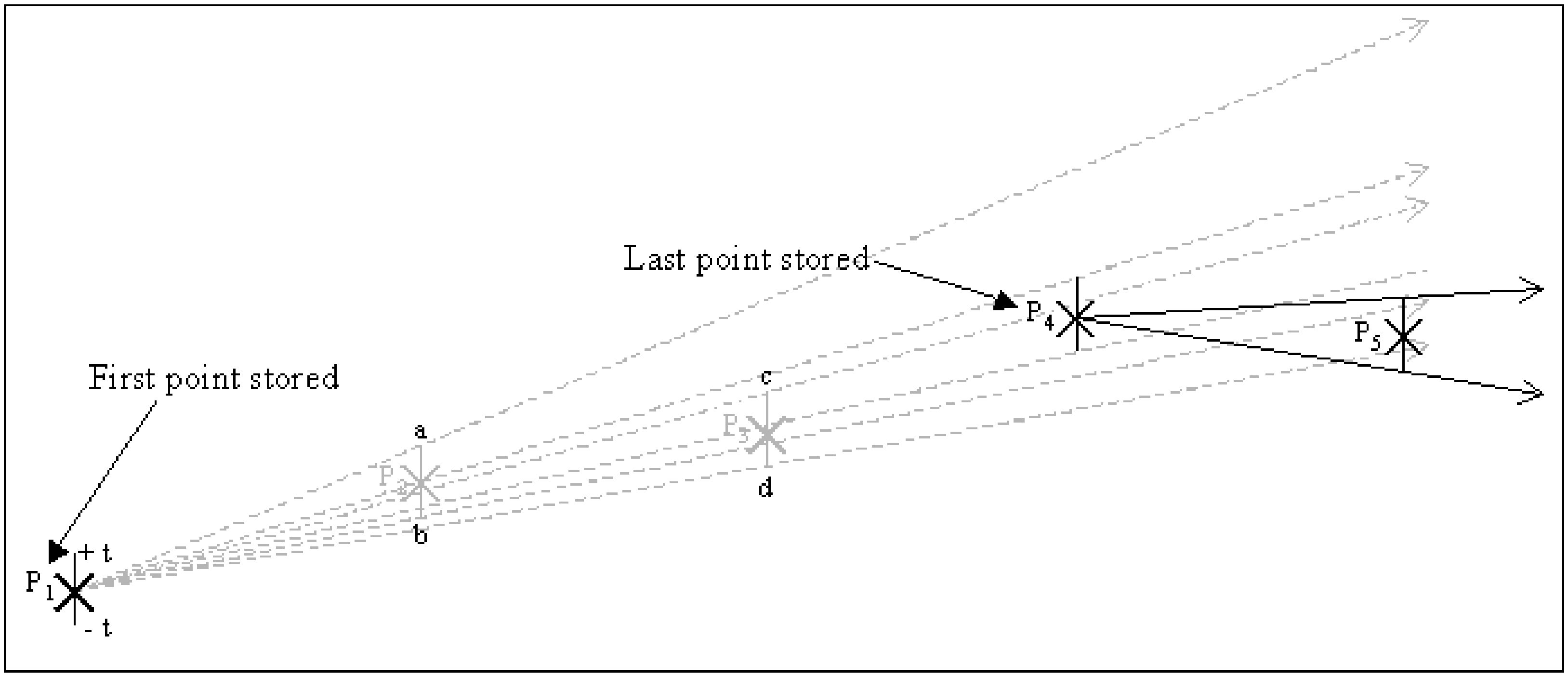
Data Elimination Compression, *continued*

PHD checks the next new point, P_5, to determine whether it lies inside the angle cP_1f. P_5 lies outside the angle so P_4 is stored.



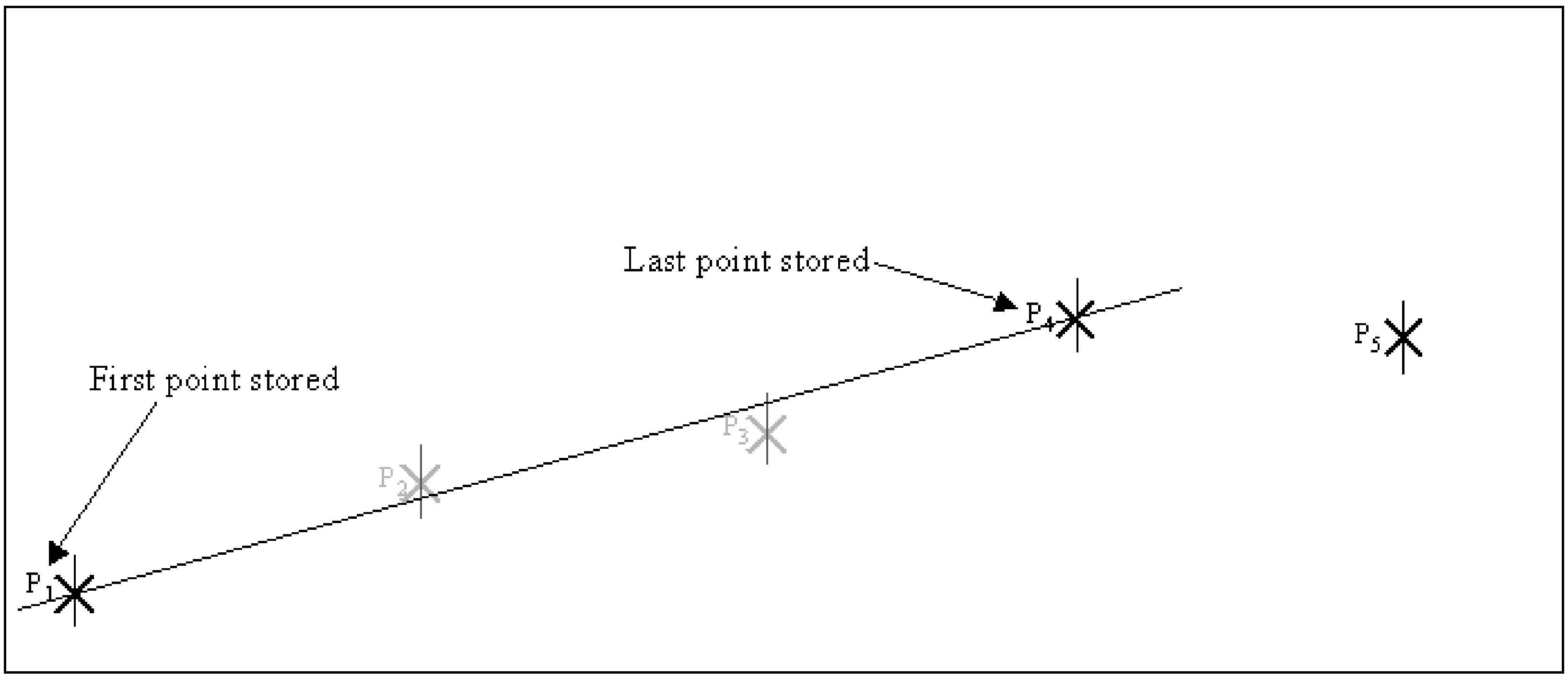
Data Elimination Compression, *continued*

PHD draws rays from the last point stored, P_4, to points equal to the next new point, P_5, plus and minus the compression tolerance. The process continues.



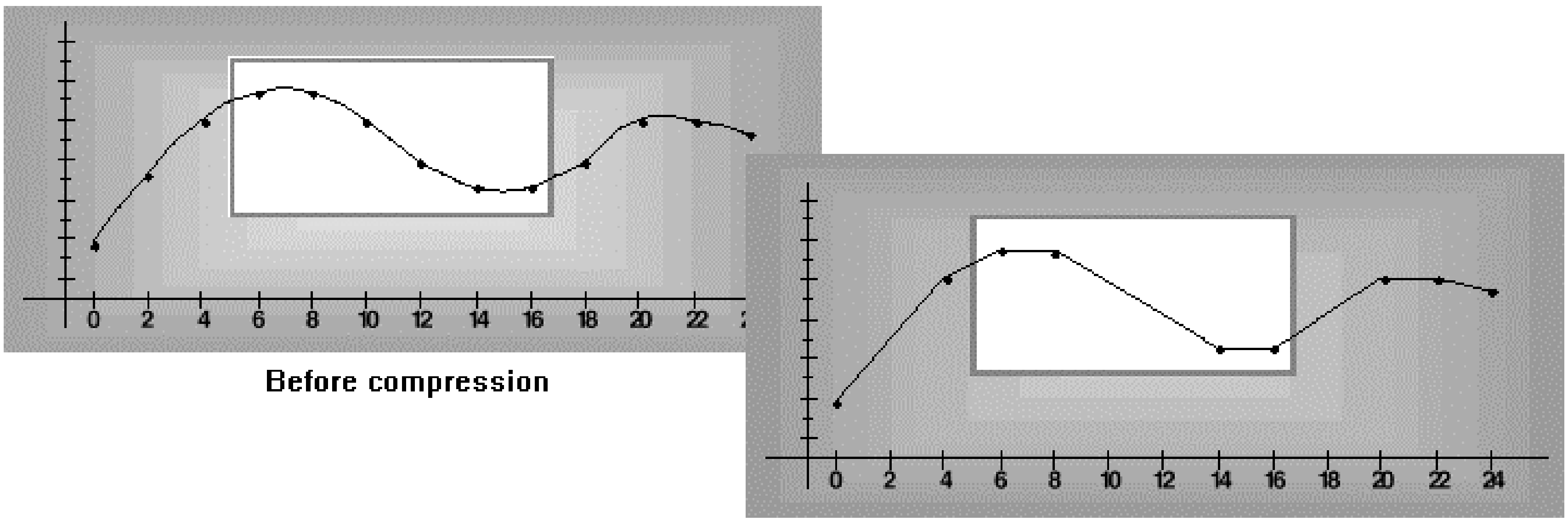
Data Elimination Compression, *continued*

The eliminated points, P_2 and P_3, can be interpolated linearly from the stored points to within the compression tolerance.



Data Elimination Compression, *continued*

The constrained slope data compression process reduces the volume of stored data while it maintains data integrity and accuracy. One effect of the process is the storage sample rate is no longer constant. The sample rate varies with the dynamic range of the data.



Before compression

After compression

Data Elimination Compression, *continued*

Because the storage sample rate varies with the dynamic range of the data, the number of raw values you retrieve in a specified time interval may be different for each tag. If the number of values is different, the timestamps will not have a one-to-one correspondence.

	A	B	C	D	E	F
4	**Date Time**	**Value**	**Confidence**	**Date Time**	**Value**	**Confidence**
5	TIC21941.PV	**From:**	now-1	TIC21941.SP	**From:**	now-1
6		**To:**	now		**To:**	now
7	**Raw**			**Raw**		
8	**Date Time**	**Value**	**Confidence**	**Date Time**	**Value**	**Confidence**
9	08-Oct-98 15:33:05	36.54383087	100	08-Oct-98 15:33:05	36.54383087	100
10	08-Oct-98 15:34:05	32.2522316	100	08-Oct-98 15:34:05	25.00114441	100
11	08-Oct-98 15:35:06	26.43854523	100	08-Oct-98 16:03:06	25.00114441	100
12	08-Oct-98 15:36:05	24.48386383	100	08-Oct-98 16:04:06	89.10505676	100
13	08-Oct-98 15:37:06	24.47241974	100	08-Oct-98 16:18:05	89.10505676	100
14	08-Oct-98 15:39:06	25.00114441	100	08-Oct-98 16:19:05	25.00114441	100
15	08-Oct-98 15:46:05	25.00114441	100	08-Oct-98 16:48:06	25.00114441	100
16	08-Oct-98 15:47:05	23.30052567	100	08-Oct-98 16:49:06	59.96795654	100
17	08-Oct-98 15:48:06	27.52574921	100	08-Oct-98 16:56:05	59.96795654	100
18	08-Oct-98 15:49:05	26.60334206	100	08-Oct-98 16:57:05	25.00114441	100
19	08-Oct-98 15:50:05	25.37880516	100	08-Oct-98 17:19:05	25.00114441	100
20	08-Oct-98 15:51:06	24.90958977	100	08-Oct-98 17:20:05	59.07759094	100
21	08-Oct-98 16:03:06	24.99885559	100	08-Oct-98 17:22:06	59.07759094	100
22	08-Oct-98 16:04:06	48.83955002	100	08-Oct-98 17:23:05	25.00114441	100
23	08-Oct-98 16:05:05	80.42343903	100	08-Oct-98 17:50:05	25.00114441	100
24	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

Redundant values are eliminated.

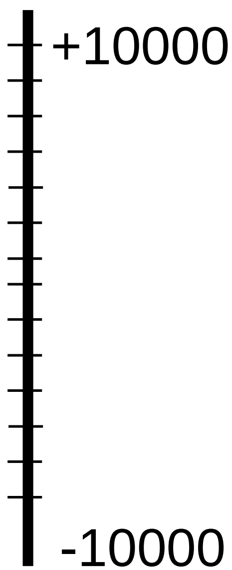
Timestamps for the two tags do not correspond.

Data Quantization and Scaling (Storage Format Compression)

Specifies the precision with which to store a number to make more effective use of disk space for archive files. PHD stores floating point numbers as scaled 16 bit integer values.

Low Extreme/High Extreme = Lowest and Highest possible value for tag.

Quantum = Scaling step size. A larger quantum yields greater compression, but reduces the precision of the stored value.



(High Extreme—Low Extreme) / 65535 = quantum

Quantum = -1 forces PHD to store a full 32 bit floating representation of the number, rather than scaling it.

(see *PHD System Manual,* Data Processing and Formatting and *PHD User Guide*, Forms)

Future Reference - Tag Configuration Data Processing Parameters

PHD Data Processing Option	PHD Tag Configuration Parameter	Definition	Typical Values
Gross Error Elimination	Gr Err Sigma Limit	The number of standard deviations by which the value may deviate from its linear regression. If the tag value deviates greater than this, then it is considered a gross error and will not be entered into PHD. The linear regression is calculated based on the lastest set of tag values as defined in the “Gr Err Samples”.	• 3 to 5 standard deviations • A blank or zero in this or the Gr Err Samples parameter field disables gross error elimination for the tag.
	Gr Err Samples	The number of samples used to perform linear regression and calculate the standard deviation.	• 6 or 8 samples • A blank or zero in this or the Gr Err Sigma Limit parameter field disables gross error elimination for the tag.
Data Quantization and Scaling	Scale High Extreme	The highest possible value for the tag.	Depends on tag.
	Low Extreme	The lowest possible value for the tag.	Depends on tag.
	Quantum	The step size for integer scaling.	• Minimum quantum = (Scale High Extreme - Low Extreme)/65535 where 65535 is a 16 bit quantizing algorithm. • A -1 disables data quantization and scaling.

Future Reference - Tag Configuration Data Processing Parameters, *continued*

PHD Data Processing Option	PHD Tag Configuration Parameter	Definition	Typical Values
Data Elimination Compression	Tolerance	The instrument tolerance value that indicates the accuracy of its measurements. Lower values indicate higher accuracy.	See your instrument documentation for the instrument tolerance.
	Tolerance Type	Whether the Tolerance is a percentage (P) of the reading, or an absolute value (A) in engineering units, or a percentage of span (S) in engineering units.	S See your instrument documentation for the tolerance type.
	Compr Toler Factor	A value multiplied by the Tolerance to give the compression tolerance used to eliminate data.	• A 1 stores all but the values that can be interpolated to within the instrument tolerance * compr toler factor. • A zero stores all but the redundant values. • A -1 disables data elimination compression .
	Min Compr Toler	The minimum absolute compression tolerance in engineering units. Use when the instrument tolerance type is percent (P).	Quantum is used in place of Min Compr Toler when Min Compr Toler is not specified or is specified as zero.

Future Reference - Tag Configuration Data Processing Parameters, *continued*

PHD Data Processing Option	PHD Tag Configuration Parameter	Definition	Typical Values
Data Smoothing	Smoothing Constant	Used with a time based exponential decay to calculate weighting factors for the data smoothing algorithm. It is approximately equal to the portion of the previous smoothed value summed into the smoothed value, given the tag's defined scan interval: Smoothed value = previous smoothed value x smoothing constant + current value x (1 - smoothing constant).	• 0 to 0.99 range • 0.5 is typical • zero or blank disables data smoothing
Noise Gating	Noise Gate	The threshold below which all values are damped to zero.	Blank disables noise gating.

Confidence Factor

- All process data within the PHD system have the following three attributes:
  - Time stamp - Internally represented as the integer number of seconds since midnight, January 1, 1970
  - Value - The process value at the time indicated by the time stamp.
  - Confidence Factor - Indicates the “goodness” of the value. Confidence factors range from -1 to 100.
- For Manual Input Tags, the confidence factor can be input with the value.

Confidence factor	Meaning	Cause
100	Complete confidence in the data	RAW Data Retrieval: • PHD collected the value from the source system and the value was within the PHD low/high extreme values. SNAPSHOT data retrieval: • PHD had a value at the exact timestamp with 100% confidence; or PHD may have interpolated a value between two values that each had a 100% confidence AVERAGE data retrieval: • All of the values used in the average had a 100% confidence; there was 100% data availability for the time span of the average.
-1	No data was available from source system	RAW or SNAPSHOT Data Retrieval: • Inactive tag • Instrument failure • Process variable is not a number
0	No confidence in the data	RAW or SNAPSHOT Data Retrieval: • A value was available from the source system, but it exceeded the PHD low or high extreme value and PHD clamped the value at the high or low extreme.
1-99	Increasing confidence in the data	• Time-weighted average of reduced data in interval where data exceed extreme limit and was clamped. SNAPSHOT Data Retrieval for Collection Tag: • The value was extrapolated beyond one scan period for the tag (50).

Intelligent Data Retrieval

PHD provides application programs with built-in data manipulations and calculations, thus simplifying application development and maintenance.

- **Data Reduction Calculations**
- **Automatic Missing Data Interpolation and Extrapolation**
- **Automatic Engineering Unit Conversions**

(see *PHD System Manual,* Data Manipulation Facilities)

Data Reduction Calculations

When requesting data from PHD, the application may request PHD to perform data reduction calculations. Data reductions, reduce the amount of data to a subset of the total raw samples.

PHD supports data reductions for F (floating point) and I (integer) datatype tags, including the following:

Average The average data over the reduction frequency (time range).

Delta The difference between the values at the end and start of the reduction frequency (time range).

First The first actual (raw) value within the reduction frequency; timestamped with the actual time.

Last The last actual (raw) value within the reduction frequency; timestamped with the actual time.

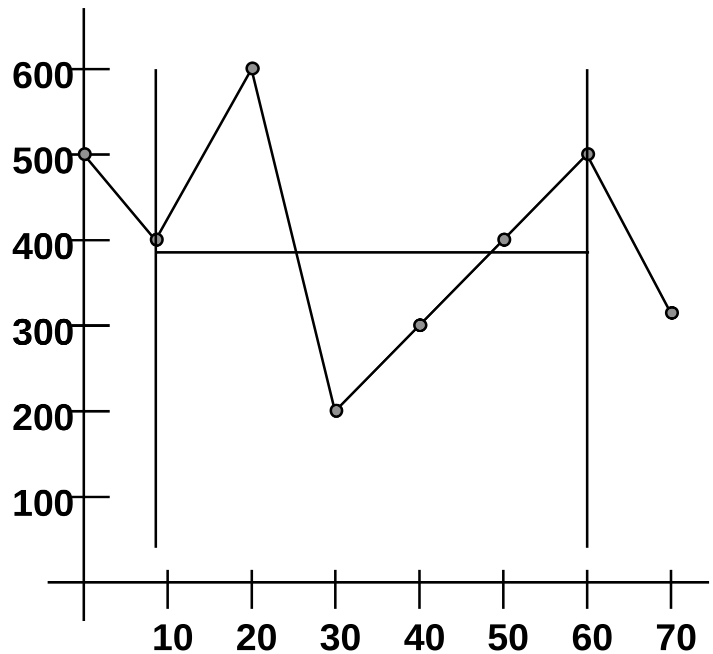
Maximum The maximum value over the reduction frequency; timestamped with the actual time of the maximum value.

Minimum The minimum value over the reduction frequency; timestamped with the actual time of the minimum value.

Time Weighted Data Reductions

PHD calculates a time weighted average value from the raw values in the time interval for these reduction types: average, delta, delta at interval, maximum, maximum at interval, minimum, minimum at interval, regression, and standard deviation.

The reductions return both a reduction value and a composite confidence factor based on the reliability of the source data.



~~**Sample Point Average**~~

~~$$\frac{400+600+200+300+400+500}{6 \text{ samples}} = 400$$~~

~~**Period Total = 400 x 5 = 2000 units**~~

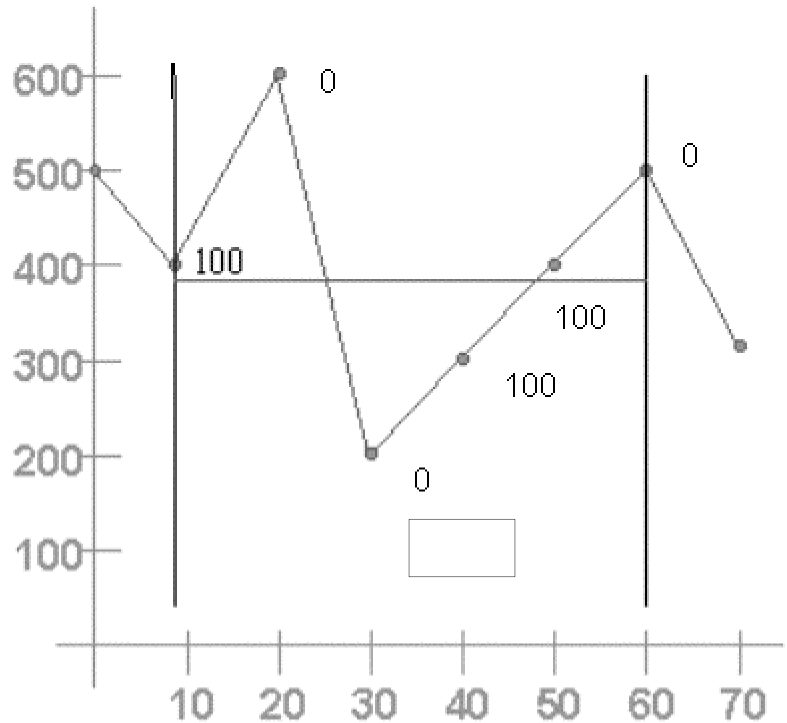
Time Weighted Average

$$\frac{500+400+250+350+450}{5 \text{ intervals}} = 390$$

Period Total = 390 x 5 = 1950 units

Time Weighted Average Confidence Factor

PHD calculates a time weighted average confidence factor for each reduced value calculated using a time weighted average.



Time Weighted Average Confidence Fact

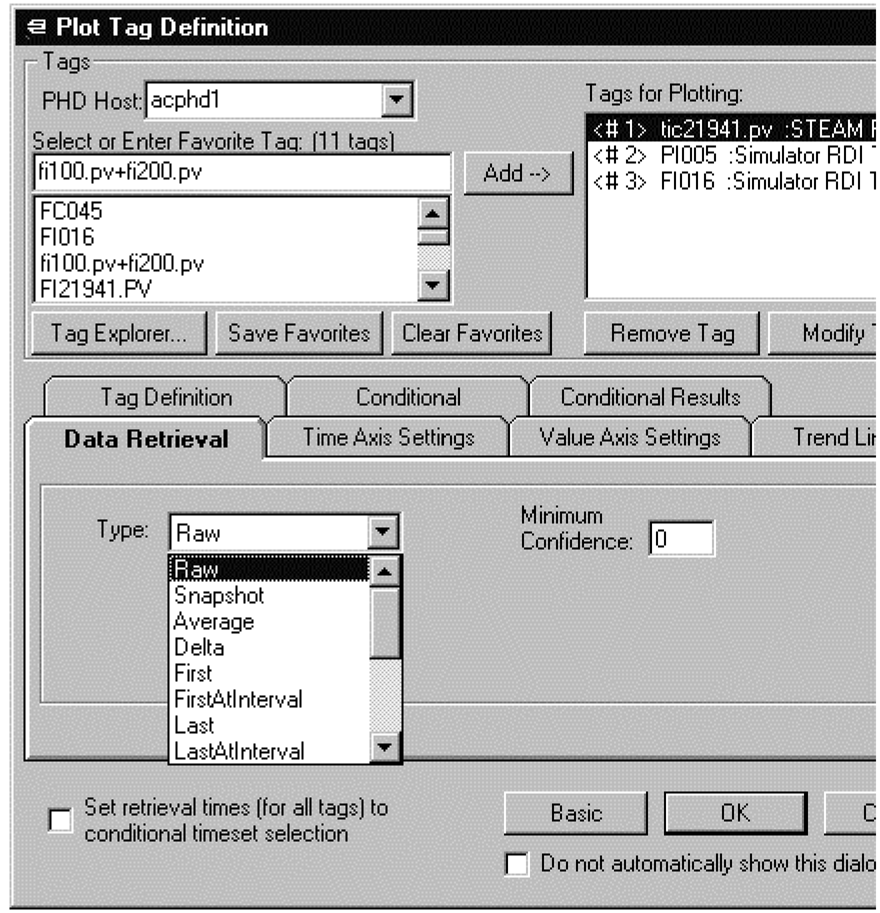
$$\frac{50 + 0 + 50 + 100 + 100}{5 \text{ intervals}} = 60$$

There is an issue regarding retrieval of PHD Confidence values - Refer to knowledge base article 14287.

Retrieval Types

Data Retrieval Types—Sample Method (Raw or Snapshot) or Reduction (such as Average)

Process Trend Example



Snapshot—Interpolates and resamples raw data at the sample frequency you request. Raw data that exist at specific timestamps are retrieved. Other values are interpolated and resampled according to the tag configuration.

Raw—Retrieves the actual sample stored with its timestamp; does not include interpolation. If the actual sample is not there during the time frame, it returns the nearest actual samples.

Reductions (for floating point and integer data)
<u>Time averaged values</u>—PHD calculates a time weighted average value from the raw values in the time interval:

Average, Delta, Delta at Interval, Maximum, Maximum at Interval, Minimum, Minimum at Interval, Regression, and Standard Deviation

<u>Time sampled values</u>—PHD retrieves raw data and confidence factors that exist at specific timestamps:

First, First At Interval, Last, and Last At Interval.

The tag configuration <u>Interpolation Method</u> affects all reduction results and <u>Resample Method</u> affects the results when resampled data are reduced.

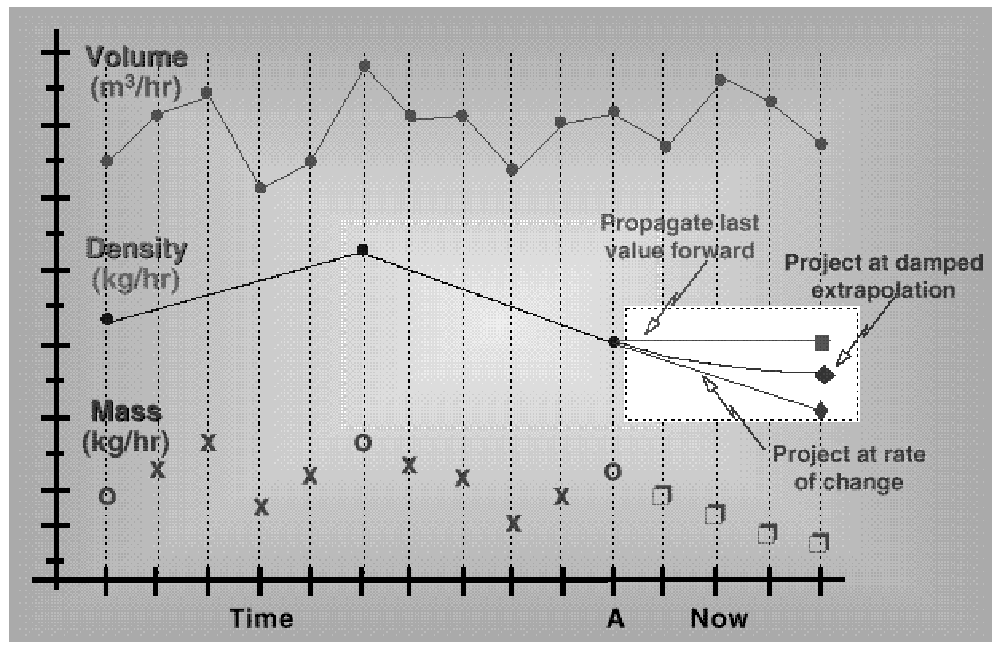
Interpolation and Extrapolation

PHD automatically performs linear interpolation estimates for missing data segments between known (stored) values.

PHD performs an automatic extrapolation for data requests that are beyond the limits of data available for a tag.

Extrapolation for times prior to or after known data use exponential damping between the last known slope of the data and the overall trend of the data. These extrapolation characteristics are configurable on a per-tag (or class-tag) basis:

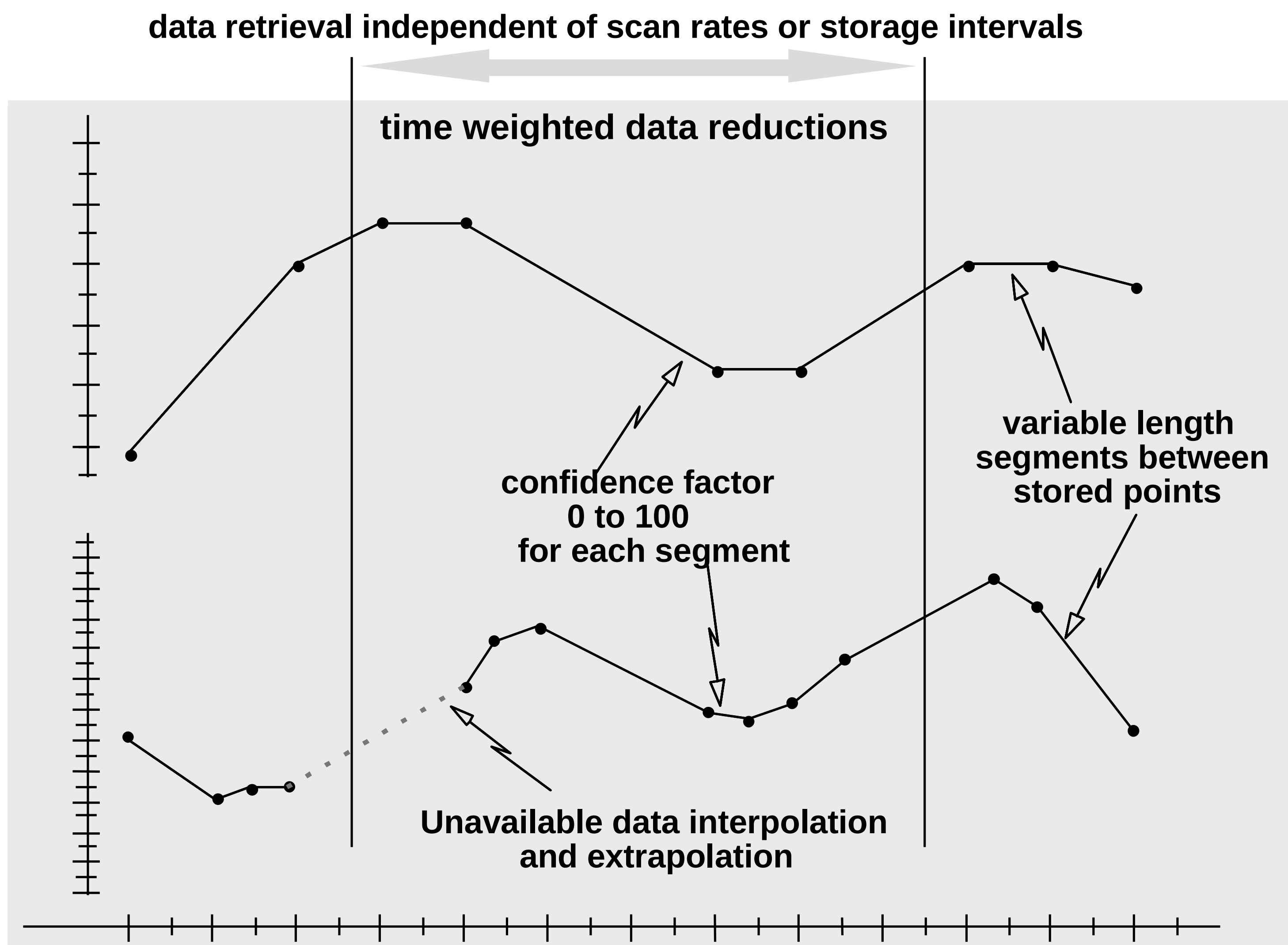
- Last known value extrapolation
- Last known slope extrapolation
- Damped extrapolation



(See *PHD System Manual* and *PHD User Guide)*

PHD Data Representation

Confidence factor for data between stored values - either the confidence of the previous value or the minimum confidence of stored values used in a linear interpolation



Reference - Reduction Types

Data Type	Tag Configuration	Reduction Type	Description	Timestamp (Modify with "Reduction Offset")	Confidence Factor	Combined Tolerance
Floating point Integer	None	Average	Average value over time interval using time weighted averaging	End of time interval	Time weighted average	For floating point values
		Delta	Value at start of time interval subtracted from value at end using time weighted averaging	End of time interval		
		Delta At Interval	Value at start of time interval subtracted from value at end using time weighted averaging	End of time interval		
		First	Value at start of time interval	Start of time interval	Confidence factor of value returned	
		First At Interval	Value at start of time interval	End of time interval		
		Last	Value at end of time interval	End of time interval		
		Last At Interval	Value at end of time interval	End of time interval		
		Maximum	Maximum value over time interval using time weighted averaging	Time of maximum	Time weighted average	
		Maximum At Interval	Maximum value over time interval using time weighted averaging	End of time interval		
		Minimum	Minimum value over time interval using time weighted averaging	Time of minimum		
		Minimum At Interval	Minimum value over time interval using time weighted averaging	End of time interval		

Reference - Reduction Types, *continued*

Data Type	Tag Configuration	Reduction Type	Description	Timestamp (Modify with "Reduction Offset")	Confidence Factor	Combined Tolerance
Floating point	None	Regression Constant	Y-intercept of the regression line using time weighted averaging	End of time interval	Time weighted average	No
		Regression Deviation	Standard deviation of the regression line over the time interval using time weighted averaging			
		Regression Slope	Slope of the regression line over the time interval using time weighted averaging			
		Standard Deviation	Standard deviation of the samples over the time interval using time weighted averaging			

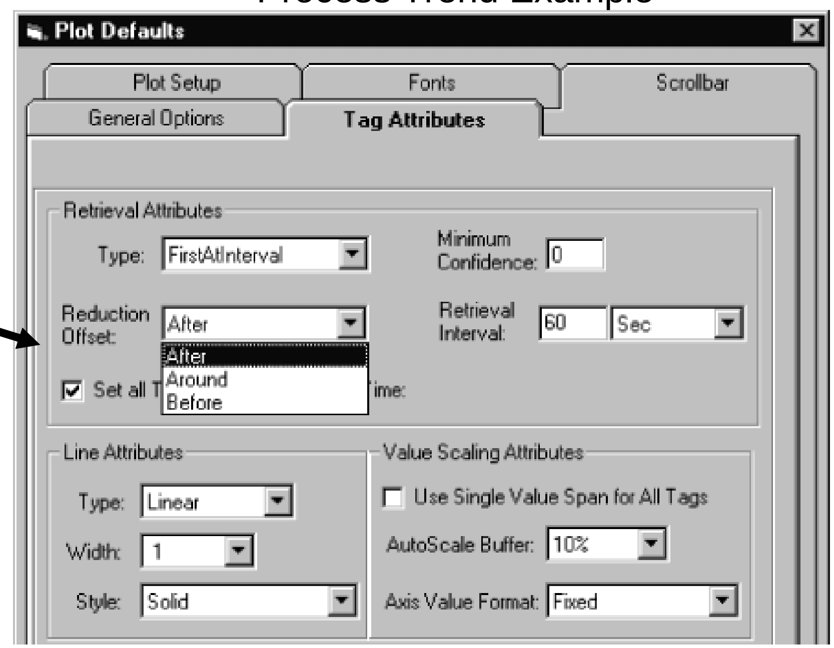
Reduction Offset
Defines how the reduction value timestamp relates to the data in the interval over which the reduction is performed:

After—Refer to “Reduction Type Average” in Customer Priority Notification #9902.

Around—Refer to “Reduction Type Average” in Customer Priority Notification #9902.

Before—Refer to “Reduction Type Average” in Customer Priority Notification #9902.

Process Trend Example



Data Retrieval - Date and Time Examples

Syntax for specifying Relative Time - Uses a keyword and a delta time format

`Keyword` `+n or -n`

Keywords:

NOW	Current time
HOUR	The top of the current hour
TODAY	Today at midnight
TOMORROW	Tomorrow at midnight
YESTERDAY	Yesterday at midnight
YEAR	January 1st at midnight of the current year
MONTH	The 1st day of the current month at midnight

Delta Time Format:

-n = prior times
(n = days:hours:minutes:seconds)

+n = later times
(n = days:hours:minutes:seconds)

Examples:

NOW-::30	(30 minutes ago)
NOW-0:6:30	(now back 6 hours, 30 minutes)
TODAY+365	(365 days from today)
YESTERDAY +0:15:30	(15 hours 30 minutes from midnight Yesterday)

Refer to the *PHD User Guide*, PHD Character Time Formats

Data Retrieval - Date and Time Examples, *continued*

Offset Keywords

An alternative delta time format can use these offset keywords.

`Keyword` `+n or -n`

Keywords:

`NOW`	Current time
`HOUR`	The top of the current hour
`TODAY`	Today at midnight
`TOMORROW`	Tomorrow at midnight
`YESTERDAY`	Yesterday at midnight
`YEAR`	January 1st at midnight of the current year
`MONTH`	The 1st day of the current month at midnight

Delta Time Format - **Offset Keywords**:

Y	Same month, day and time of the year
Mo	Same day and time of the month
W	Same day and time of the week
D	Same time of the day
H	Same minutes, seconds of the hour
M	Same seconds of the minute
S	Same second

Examples:

`NOW -1W`	(one week ago from the current time)
`TODAY-1W4H`	(a week and 4 hours prior to today at midnight)
`TODAY-1MO`	(the same day of the previous month)

Automatic Engineering Unit Conversions

When an application program requests data from PHD, it is possible to specify the units it requires. PHD converts the units from those stored with the tag into the requested units, eliminating having to build conversion functions into the various user tools and applications.

A tag can be defined with any unit code, but is convertible by PHD only if it makes use of the standard PHD base units. A valid convertible unit contains one or two base units combined in one of the following ways:

baseunit
baseunit/constant
baseunit1/baseunit2
baseunit/timeunit

ABSOLUTE	RATE	ACCELERATION
BBL	BBL/D	BBL/Hr^2
M^3	M^3/D	M3/ Hr^2
CF	CFS	CFS^2
Tonne	Tonne/D	Tonne/ Hr^2
LB	LB/Hr	LB/ Hr^2
KG	KG/D	KG/ Hr^2
DEGC	DEGC/Hr	DEGC/ Hr^2
DEGF	DEGF/Hr	DEGF/ Hr^2
KPA	KPA/Hr	KPA/S^2
PSIG	PSIG/Hr	PSIG/S^2

(See *PHD System Manual,* Automatic Unit Conversion)

Honeywell

Helping You Manage Your World

www.iac.honeywell.com

TotalPlant *Support Online:*

www.iac.honeywell.com\pub\mktg\support.stm