Turning **insight** into **action**.

OSIsoft. USERS CONFERENCE 2011

Developing Analytics Over Streaming Data with Microsoft StreamInsight & PI for StreamInsight

Presented by **Erwin Gove, Glenn Moffett and Roman Schindlauer**

Challenges

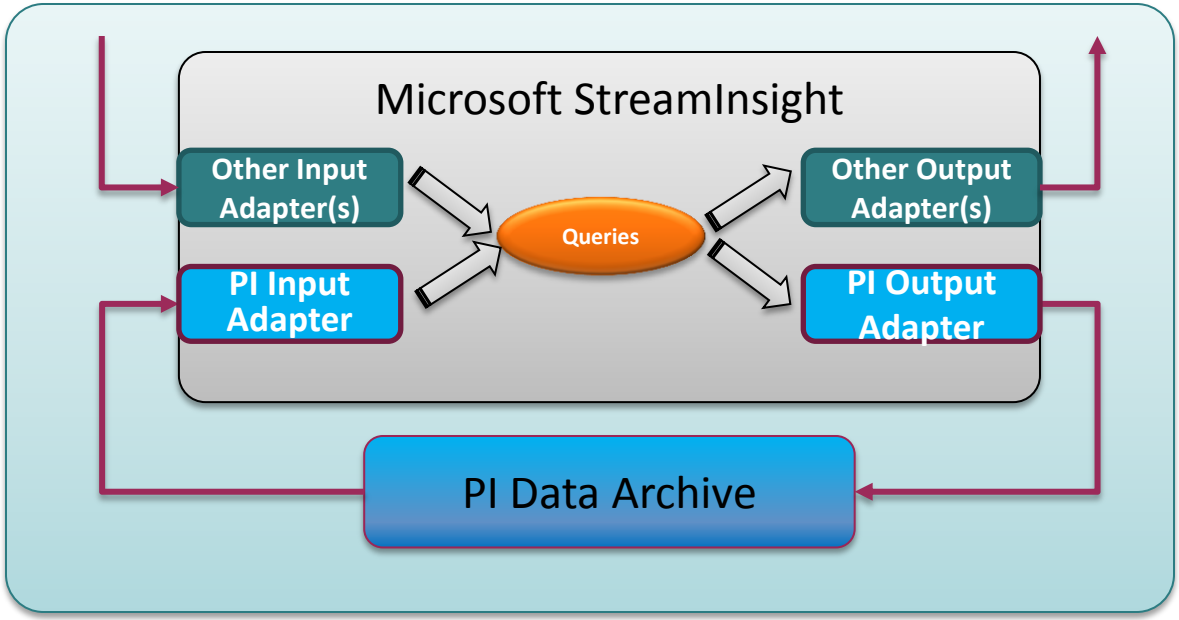
- Increasing amount of data
- Integration of data for analysis
- Data quality management
- Finding better ways to analyze data

Microsoft StreamInsight

- In-memory data processing engine
- Event driven, incremental computations
- .NET and LINQ programming environment
- Correlate multiple, heterogeneous data sources
- Modular design – engine and adapters

PI for StreamInsight

- Input and Output adapters for StreamInsight
- Analytics using Microsoft StreamInsight



Terminology

- Stream is a series of events

- Events contain a payload

Payload
Sensor10
3.14
Good
Gas Turbine 2
Online
…

Typical StreamInsight Queries

- “Give me the Top 3 values every 10 minutes”
- “Filter out sensor readings when the device was in a maintenance period”
- “Tell me when an event A was followed by an event B within 3 minutes”

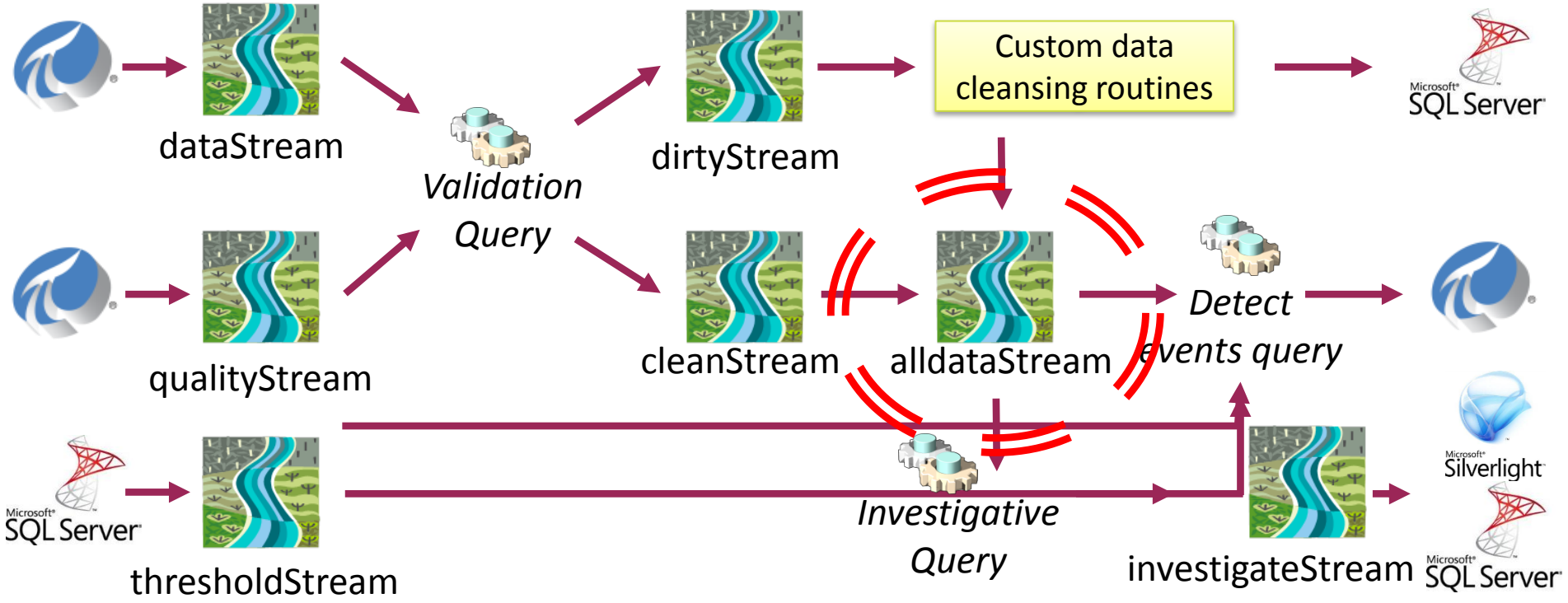
Use Scenarios

- Event detection
  - Vibration analysis & meter data outage management
- Cleanse and monitor sensor data
  - Inform users of data issues
- Other Industry examples
  - Clickstream analysis for user behavior
  - Risk analysis for business activity monitoring

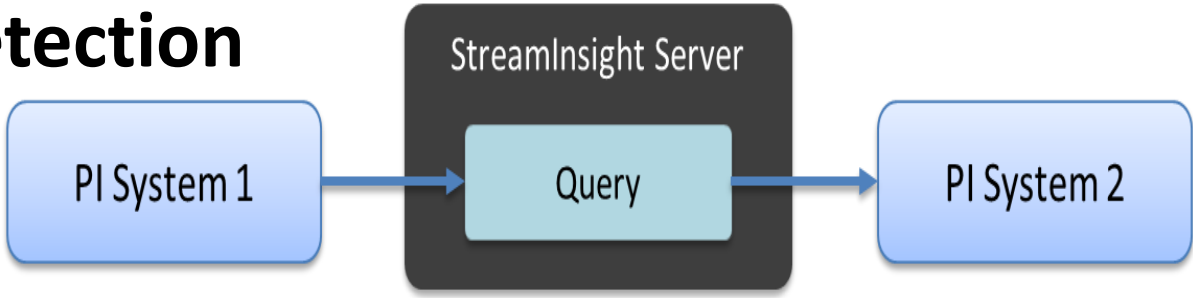
Cleanse and monitor sensor data

- Outside acceptable limits or invalid
- Monitor 1000-10,000s of data points
- Support data bursts upward of 10,000 events/sec
- External meta-data for validation rules
- Call 3rd party or custom library for validation

Cleanse and monitor sensor data

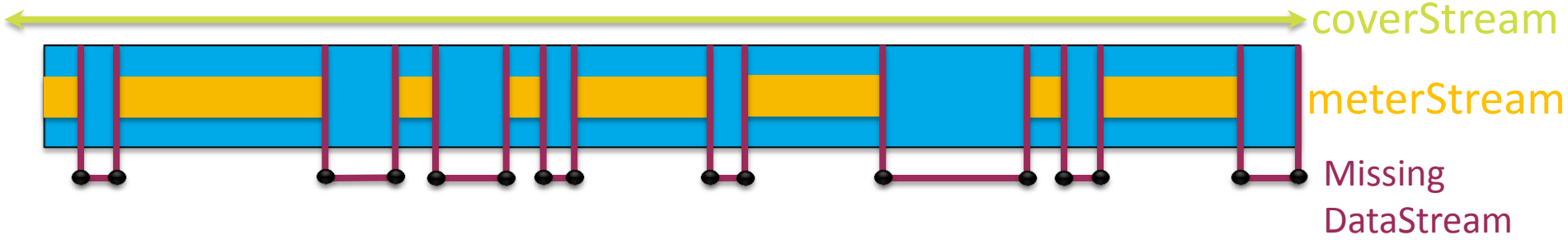


Vibration Event detection & Analysis



- Requirement
  - Detect limit exceedance for 12,000 data streams and events/sec
  - Replicate a data window of 60 minutes to the second PI System on event detection
- Benefit
  - Enable failure analysis with data from each event

Meter reporting validation




```
missingDataStream  = from left in coverStream
                     where (from right in meterStream
                            where left.tag == right.tag
                            select right).IsEmpty()
                            select left;
```

Meter reporting validation – Traditional method

```
public List<DateTime> MissingIntervals(DateTime start, DateTime end, TimeSpan interval, String serverName, String tagName)
{
   List<bool> foundIntervals = new List<bool>();
   for (DateTime intervalStart = start; intervalStart - interval < end; intervalStart += interval)
      foundIntervals.Add(false);
   PISDK.PISDK sdk = new PISDK.PISDK();
   PISDK.Server server = sdk.Servers[serverName];
   PISDK.PIPoint pipoint = server.PIPoints[tagName];
   PISDK.PIValues pivalues = pipoint.Data.RecordedValues(start, end,
      BoundaryTypeConstants.btOutside, string.Empty, FilteredViewConstants.fvShowFilteredState, null);
   foreach (PISDK.PIValue pivalue in pivalues)
   {
      long intervalOffset = ((long)(pivalue.TimeStamp.UTCSeconds * TimeSpan.TicksPerSecond) + interval.Ticks / 2 - start.Ticks) /
interval.Ticks;
      if (intervalOffset >= 0 && intervalOffset < foundIntervals.Count)
         foundIntervals[(int)intervalOffset] = true;
   }
   List<DateTime> missing = new List<DateTime>();
   for (long intervalOffset = 0; intervalOffset < foundIntervals.Count; intervalOffset++)
   {
      if (!foundIntervals[(int)intervalOffset])
         missing.Add(start + TimeSpan.FromTicks(intervalOffset * interval.Ticks));
   }
   return missing;
}
```

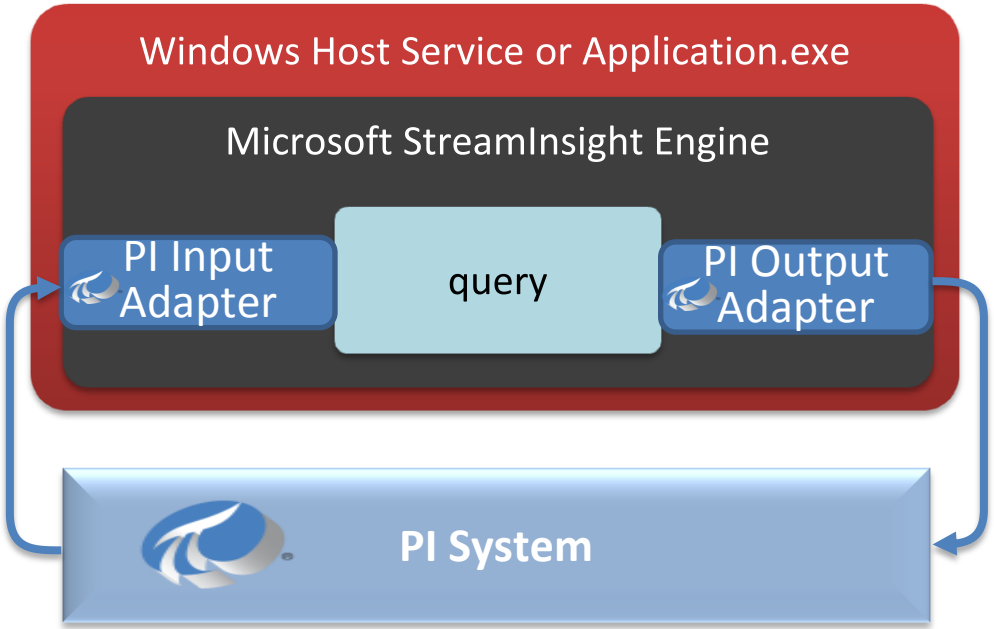
Where does StreamInsight fit with the PI System?

<table>
<tr><th colspan="2">PI Analytics</th></tr>
<tr><th>Configuration</th><th>Programming</th></tr>
<tr><td>Performance Equations</td><td>PI Advanced Computing Engine (PI ACE)</td></tr>
<tr><td>Totalizers</td><td rowspan="2">Microsoft StreamInsight & PI for StreamInsight</td></tr>
<tr><td>Statistical Quality Control</td></tr>
<tr><td>PI AF formula data reference</td><td>PI AF custom data reference</td></tr>
</table>

PI ACE and StreamInsight Features

Features and Capabilities	PI ACE	PI for StreamInsight
PI Server Collective and PI buffering support	✓	✗
Scheduler or can run as a Windows Service	✓	✓ (manual restart of query)
Recalculation	✓	✗
Wizard and Manager	✓	✗
High level query language LINQ (“SQL Like”)	✗	✓
Performance events/sec	1000’s*	> 10,000’s*

How does StreamInsight work?



Microsoft Visual Studio

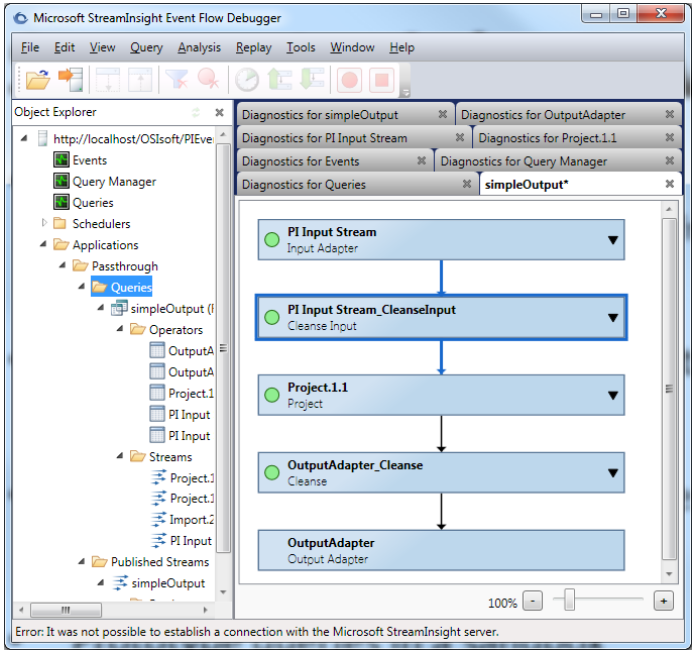
Microsoft .NET

StreamInsight Features

Benefit	StreamInsight Features
Simple configuration for PI Tags	PI Adapters support tag searches
Support complex payloads	Event types are .NET structures
Process high data rates	Low latency and in excess of 10,000 events/sec*
Minimize load on server	In-memory storage and processing of data
Query across different data sources	Normalized events, join streams
Don’t need to know about PI SDKs	.NET, LINQ and PI Configuration for adapters only

StreamInsight Environment

- Modular Queries and collaboration
- Multiple clients sharing a single StreamInsight server
- Trace the event flow of queries
- Prototype queries in a sandbox



StreamInsight Software

- Microsoft SQL Server 2008 R2
  - StreamInsight inside!
  - Standard and Premium versions
- PI for StreamInsight
  - PI Adapters
    - Stream events as they arrive at the PI Data Archive
    - Write events to PI System

More Information

- OSIsoft Virtual Campus
  - OSIsoft vCampus Live 2010 presentation
  - Pre-release (CTP) version of PI for StreamInsight
  - Forums
- Microsoft StreamInsight – blogs, forums
  - Hitchhiker's Guide to StreamInsight

What can StreamInsight do for you?

- Support higher numbers of data and data rates
- Use queries to answer business questions

Thank you

Turning **insight** into **action**.