

Philipps



Universität  
Marburg

# Auf dem Weg zur allwissenden Fabrik

Vertikale Integration auf Basis  
kontinuierlicher Datenverarbeitung

**Christoph Heinz, Jürgen Krämer,  
Tobias Riemenschneider,  
Bernhard Seeger**

**27.09.2007**

# Agenda

## ■ **Production Intelligence**

- Vertikale Integration und deren Nutzen

## ■ **Kontinuierliche Anfragen über Datenströmen**

- Spezifikation und Verarbeitungsprinzipien

## ■ **Unsere Softwareinfrastruktur PIPES**

- Architekturelle Grundlagen und Funktionsumfang

## ■ **Einsatz von PIPES in der Fabrikautomation**

- Kopplung an die Integrationsmiddleware i-Plant
- Synergieeffekte
- Kurze Demo

# Agenda

## ■ **Production Intelligence**

- Vertikale Integration und deren Nutzen

## ■ **Kontinuierliche Anfragen über Datenströmen**

- Spezifikation und Verarbeitungsprinzipien

## ■ **Unsere Softwareinfrastruktur PIPES**

- Architekturelle Grundlagen und Funktionsumfang

## ■ **Einsatz von PIPES in der Fabrikautomation**

- Kopplung an die Integrationsmiddleware i-Plant
- Synergieeffekte
- Kurze Demo

# Motivation

## ■ Anwendungsgebiet: Produktionsindustrie

- Branchen mit hohem Wettbewerbsdruck und hohem Automatisierungsgrad
- Generierung enormer Datenmengen durch Sensoren in modernen Automatisierungssystemen

## ■ Gegenwärtiger Stand

- Daten sind vorhanden, aber Zugriff und Verknüpfung sind schwierig
- Teure Datengewinnung mit großem zeitlichen Verzug
- Schnelle Reaktion auf aktuelle Marktanforderungen oder unvorhersehbare Ereignisse ist schwierig!

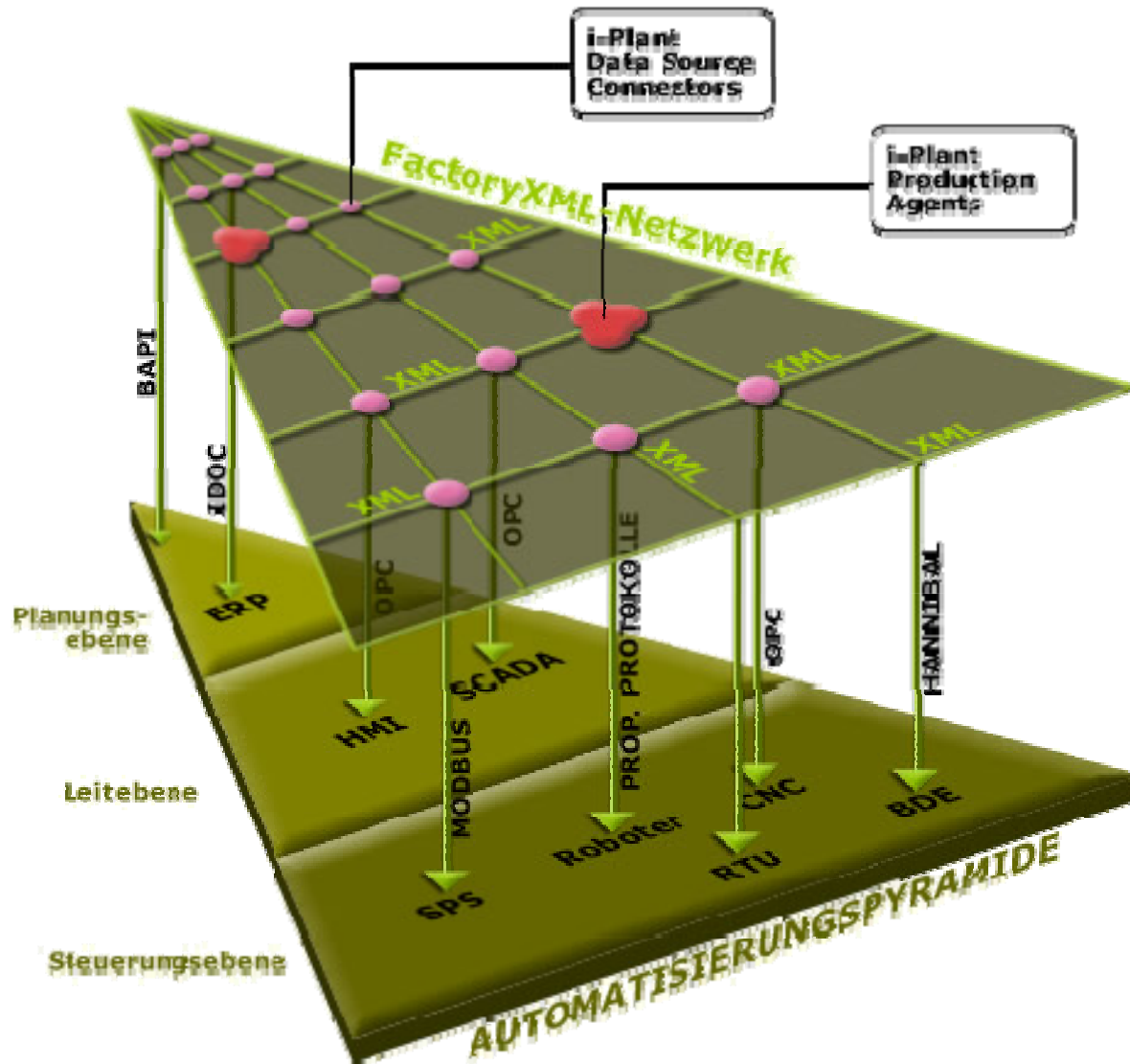
# Production Intelligence

- **Ziel:** Zügige Bereitstellung und Analyse von Live-Daten aus Produktionsprozessen zur Gewinnung betrieblich wichtiger Kennzahlen
  - Ursprung in *Business Intelligence*
    - Messung der Effizienz einzelner Geschäftsprozesse
    - Darstellung betrieblich wichtiger Kenngrößen
      - Zeitnah, verständlich und einfach zugreifbar
  - Wichtige Unterschiede
    - Heterogene Daten und Systeme vs. strukturierte und verdichtete Daten
    - Signifikant höhere Datenaufkommen

# Vertikale Integration

- **Bezug zwischen automatisierungstechnischen und betriebswirtschaftlichen Daten**
- **Leistungsmessung eines Produktionsprozesses**
  - Ausschuss nicht nur in Stück, sondern in €
  - Ausfall einer Maschine in der aktuellen Woche gegenüber dem Vormonat
  - Vergleich von Ist- mit Sollzuständen
- **Durchgängiger Datenfluss zwischen Steuerungs- und Planungsebene**

# Vertikale Integration



# Typische Softwareanwendungen

## ■ Digital Dashboard

- Überblick über wichtige Leistungsindikatoren im Web-Browser mit ständiger Aktualisierung



## ■ Web-Reporting

- Zeitnahe Auskunft über betriebswirtschaftliche Parameter angepasst an Nutzergruppen

## ■ Alarmierung und Nachrichtenversand

- Automatisierte, softwareseitige Alarmierung
- Nachrichten bei signifikanten Ist/Soll-Abweichungen
- Einbeziehung von Web- und Internettechniken sowie mobiler Endgeräte





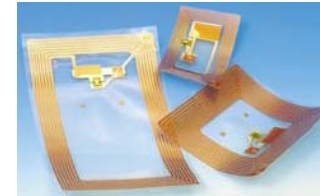
# Agenda

- **Production Intelligence**
  - Vertikale Integration und deren Nutzen
- **Kontinuierliche Anfragen über Datenströmen**
  - Spezifikation und Verarbeitungsprinzipien
- **Unsere Softwareinfrastruktur PIPES**
  - Architekturelle Grundlagen und Funktionsumfang
- **Einsatz von PIPES in der Fabrikautomation**
  - Kopplung an die Integrationsmiddleware i-Plant
  - Synergieeffekte
  - Kurze Demo

# Datenströme

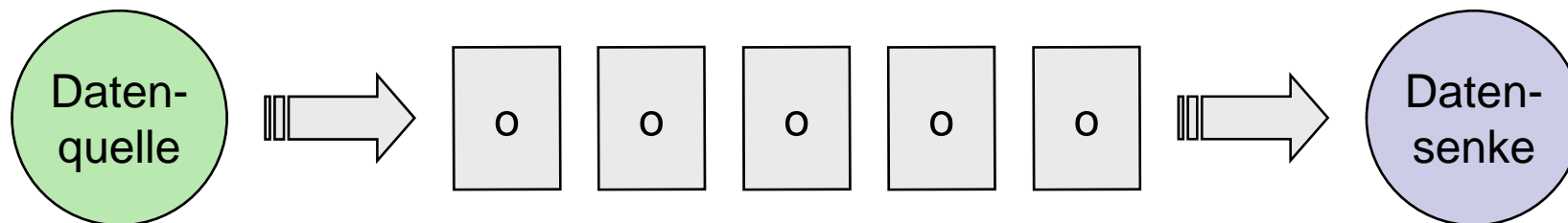
## ■ Aktive Datenquelle

- Datenquelle, die ihre Daten selbständig übermittelt

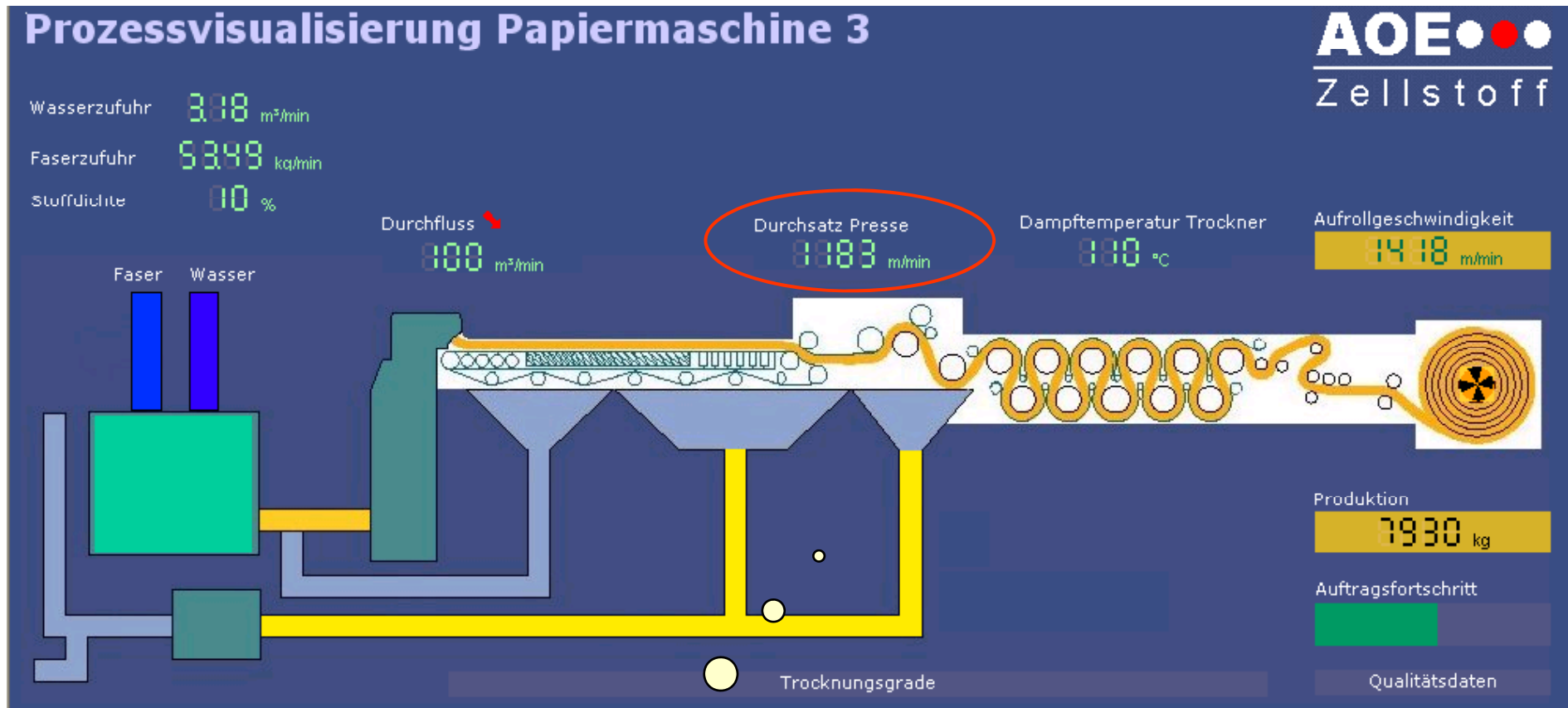


## ■ Datenstrom

- Kontinuierliche Folge von Elementen einer aktiven Datenquelle



# Beispiel Fabrikautomation



Wie sieht ein Datenstrom aus?

# Eingabestrom

## ■ Folge von Paaren (**Tupel**, **Zeitstempel**)

- Alle Tupel sind vom gleichen Typ.
- Jedes Tupel besitzt einen Zeitstempel.
- Datenstrom ist zeitlich geordnet.

## ■ Format

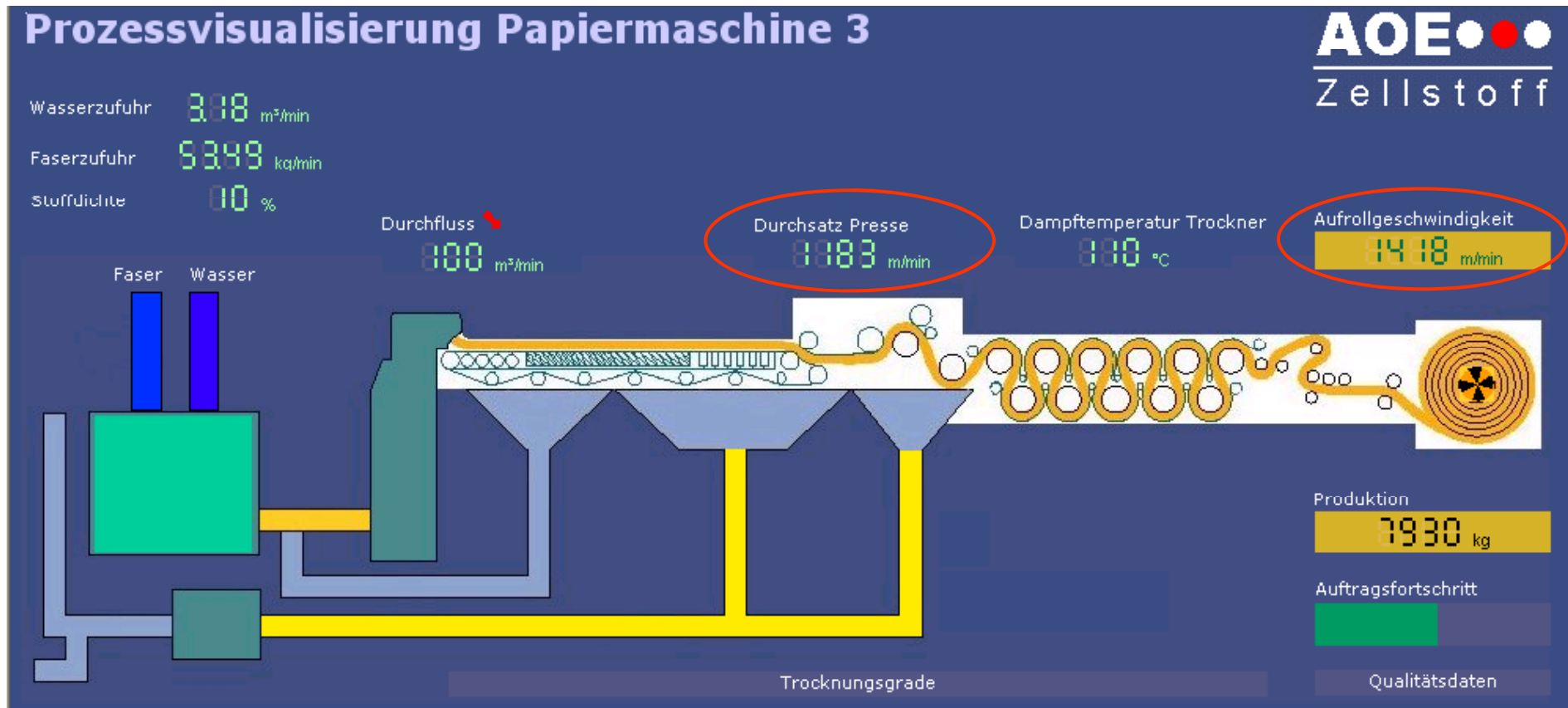
```
PressStream(wert REAL, zeitstempel TIMESTAMP)
```

## ■ Datenstrom

```
(1183, 13:10:02)  
(1120, 13:10:03)  
(1215, 13:10:04)
```

...

# Beispiel für kontinuierliche Anfrage



„Bestimme für den **Pressendurchsatz** und die **Aufrollgeschwindigkeit** alle Wertepaare, deren betragliche Differenz innerhalb der letzten Minute größer als 300 war.“

# Aktuelle Umsetzung

- **Direkte Implementierung**
- **Nachteile bei komplexerer Applikationslogik**
  - Tausende Zeilen an Quellcode
    - Proprietäre Programmiersprachen
  - Hohe Entwicklungs- und Wartungskosten
  - Inflexibilität gegenüber Änderungen/Erweiterungen
    - Extensives Testen und Debuggen

```
STEP 5 - S5KXS01X
FB 201
Netzwerk 1
: B      =DB
:
: L      =PW
: SWW    3
: T      DW 200
: L      KF +314
: >F
: SPB    =M001
: S      =DRHT
: L      KF +0
: T      =XA
: BEA
M001 : RB  =DRHT
:
: L      DW 200
: L      KF +2600
: <F
: SPB    =M002
: S      =BU
: SPA    =M003
FAdressen F Bib.Nr.
1Symb.Anz.2Referenz
```

# Idee

## ■ Verwendung von Datenbanksystemen

- Bewährte Technologie
- Einfache Spezifikation von Applikationslogik
  - SQL als deklarative Anfragesprache
  - Spezifikation der Ergebnismenge statt Berechnungsvorschriften

## ■ Vorteile

- Hohe Akzeptanz bei Anwendern
- Standardisierung und wohldefinierte Semantik
- Geringere Entwicklungs- und Wartungskosten


# Anfrageformulierung

## ■ Textuelle Anfrage

„Bestimme für den **Pressendurchsatz (S1)** und die **Aufrollgeschwindigkeit (S2)** alle **Wertepaare**, deren **betragliche Differenz** innerhalb der letzten Minute **größer als 300** war.“

## ■ SQL-Anfrage

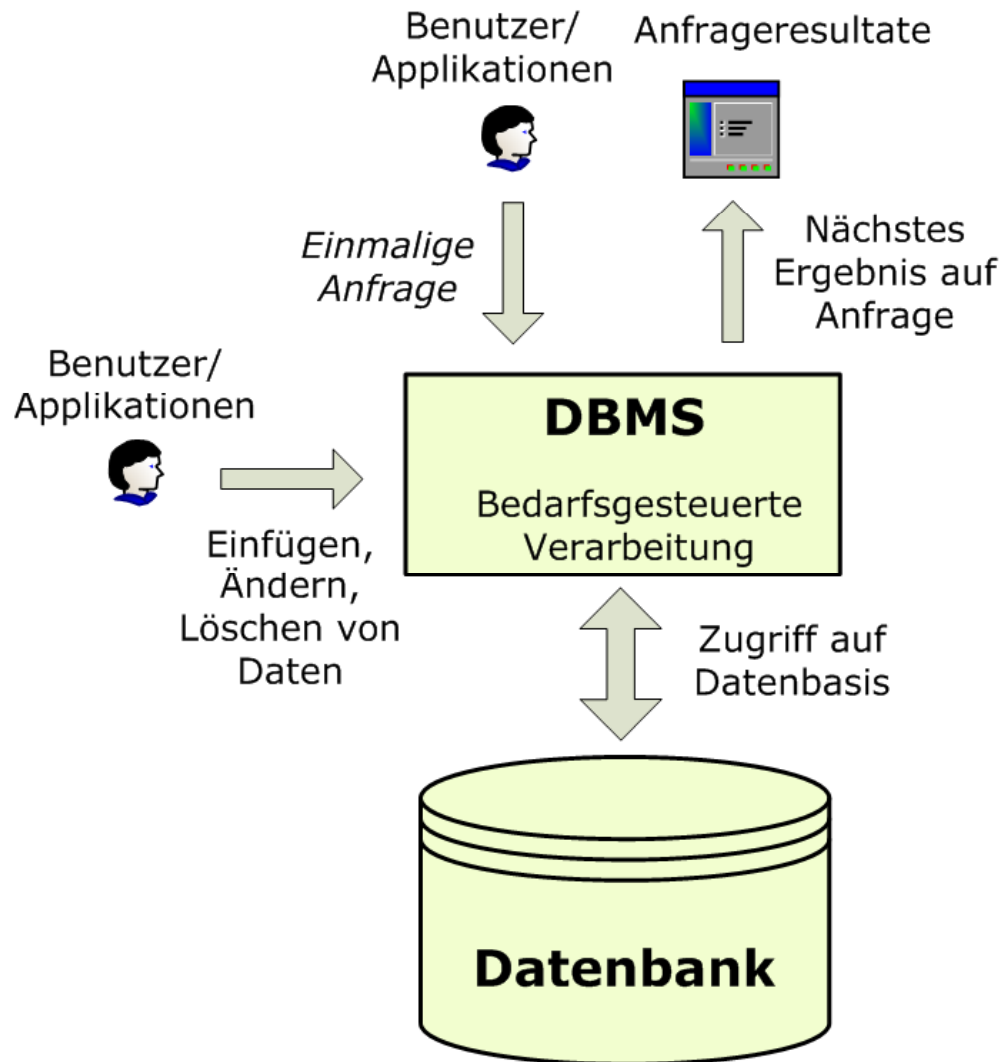
```
SELECT S1.WERT, S2.WERT
FROM   S1 WINDOW(RANGE 1 MINUTE),
       S2 WINDOW(RANGE 1 MINUTE)
WHERE  |S1.WERT - S2.WERT| > 300;
```



Wie wird  
diese  
Anfrage  
verarbeitet?



# Traditionelle Anfrageverarbeitung

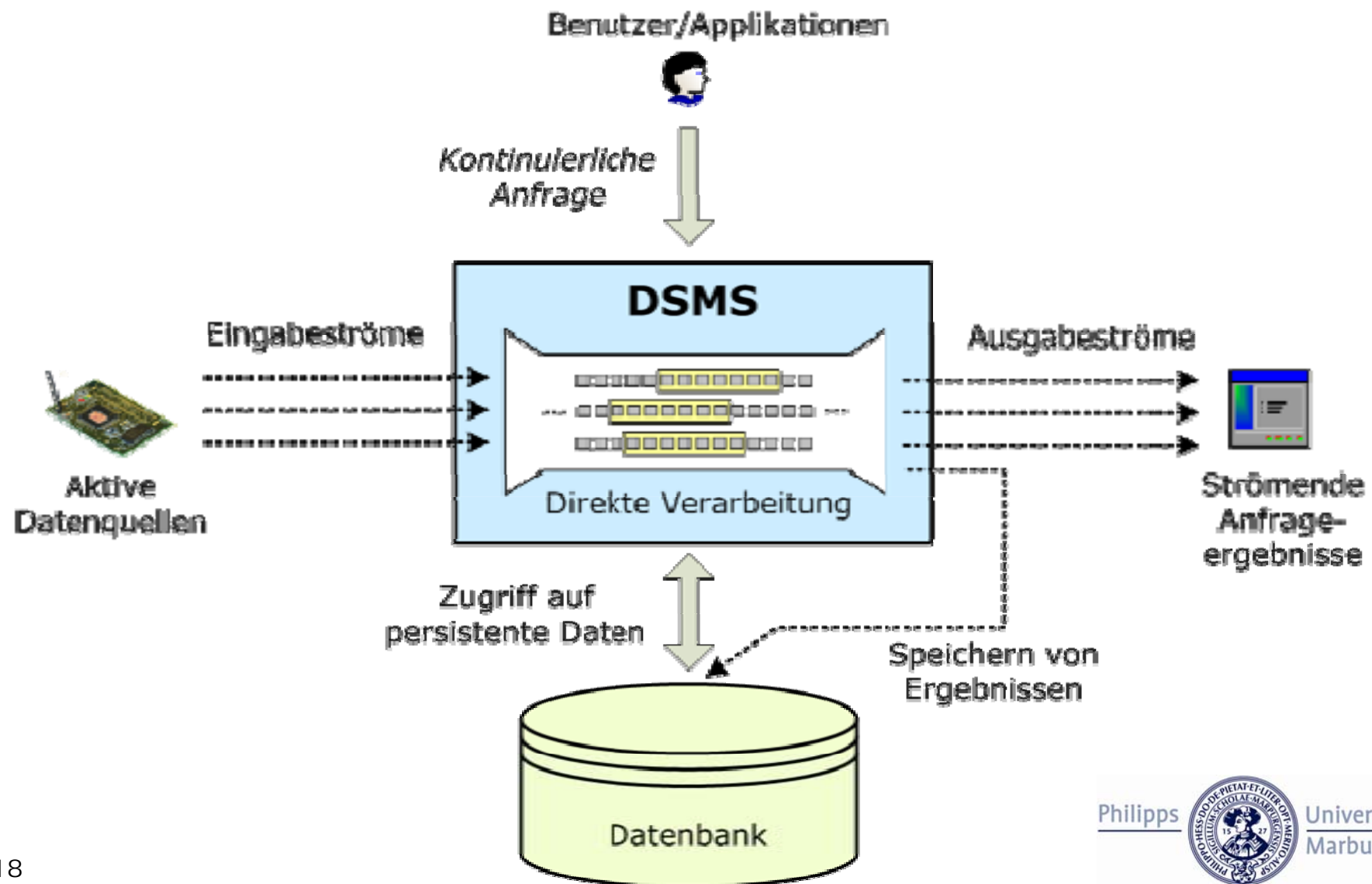


- **Keine kontinuierlichen Anfragen**
- **Vollständige Speicherung**
  - Teure Suche und Analyse
  - *Datenfriedhöfe*



DBMS sind eher ungeeignet für Datenstromanwendungen

# Anfrageverarbeitung auf Datenströmen



# Agenda

- **Production Intelligence**
  - Vertikale Integration und deren Nutzen
- **Kontinuierliche Anfragen über Datenströmen**
  - Spezifikation und Verarbeitungsprinzipien
- **Unsere Softwareinfrastruktur PIPES**
  - Architekturelle Grundlagen und Funktionsumfang
- **Einsatz von PIPES in der Fabrikautomation**
  - Kopplung an die Integrationsmiddleware i-Plant
  - Synergieeffekte
  - Kurze Demo

# PIPES – Public Infrastructure for Processing and Exploring Streams

## ■ Softwareinfrastruktur zur Verarbeitung von Datenströmen

- > 1000 Java-Klassen
- > 150.000 Lines of Code
- Dokumentiertes API
- Zahlreiche Use-Cases

## ■ Gegenwärtiger Stand

- Kommerzielle Weiterentwicklung  
→ Spin-Off
- Gefördert durch BMWi und EU

Deutsche  
Forschungsgemeinschaft

**DFG**



Bundesministerium  
für Wirtschaft  
und Technologie



EUROPÄISCHE UNION  
Europäischer Sozialfonds

Philipps



Universität  
Marburg

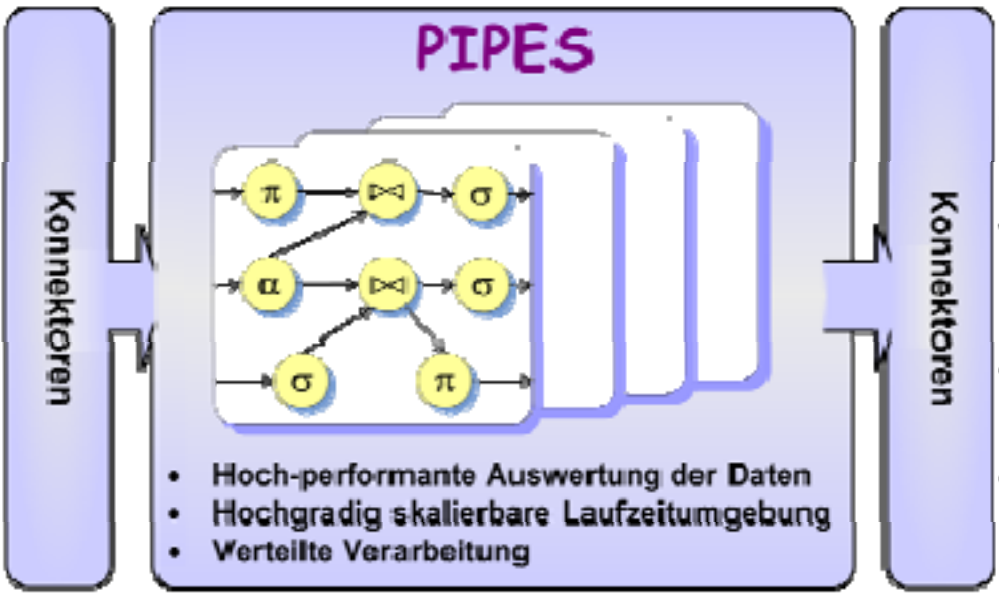


### Entwicklung von Applikationen mit



### Datenquellen

- Markt-daten
- Click-streams
- Netzwerk-pakete
- RFID / Sensoren



### Datensenken

- Alarm-monitor
- Dash-boards
- Data Warehouse
- Web-interfaces



# Anfrageformulierung

## ■ SQL-konforme Anfragesprache

- Standardisierung / weite Verbreitung in der Praxis
- Klare und wohldefinierte Semantik
- Einfache Formulierung komplexer Applikationslogik

## ■ Visuelle Programmierung

- Drag & Drop
- Modellierung als Datenflussdiagramm (Workflows)

## ■ Implementierung mit API

- Wenige zentrale Schnittschnelle
- Zahlreiche Vorimplementierungen

# Anfrageverarbeitung

- Operatoren implementiert mit Online-Algorithmen → direkte Verarbeitung
  - Publish-Subscribe Architektur
    - Zusammenfassen gemeinsamer Teilanfragen
  - Adaptive Laufzeitumgebung
    - Scheduling → Multi-Threading
    - Ressourcenmanagement → Überlastbehandlung
    - Optimierung → statisch und dynamisch
- Hohe Skalierbarkeit und Verfügbarkeit!**

# Weitere Features

## ■ Data Mining auf Datenströmen

- Berechnung einfacher und höherwertigerer Statistiken
- Erkennung von Mustern

## ■ Verteilung auf mehrere Rechner

- Migration von Anfragen zur Laufzeit
- Einsatz in Grid-Umgebungen

## ■ Archivierung von Ausgabeströmen

- Vorverarbeitung und Aufbereitung der Daten
- Unterstützung historischer Anfragen



# Leistungsdaten

## ■ Durchsatz

- >  $10^6$  Elemente pro Sekunde
- Vergleich mit kommerziellen Datenbanksystemen
  - Nachbildung kontinuierlicher Anfragen mittels Triggern
  - Weniger als  $10^2$  Elemente pro Sekunde
  - Keine Anfrageoptimierung möglich

## ■ Skalierbarkeit

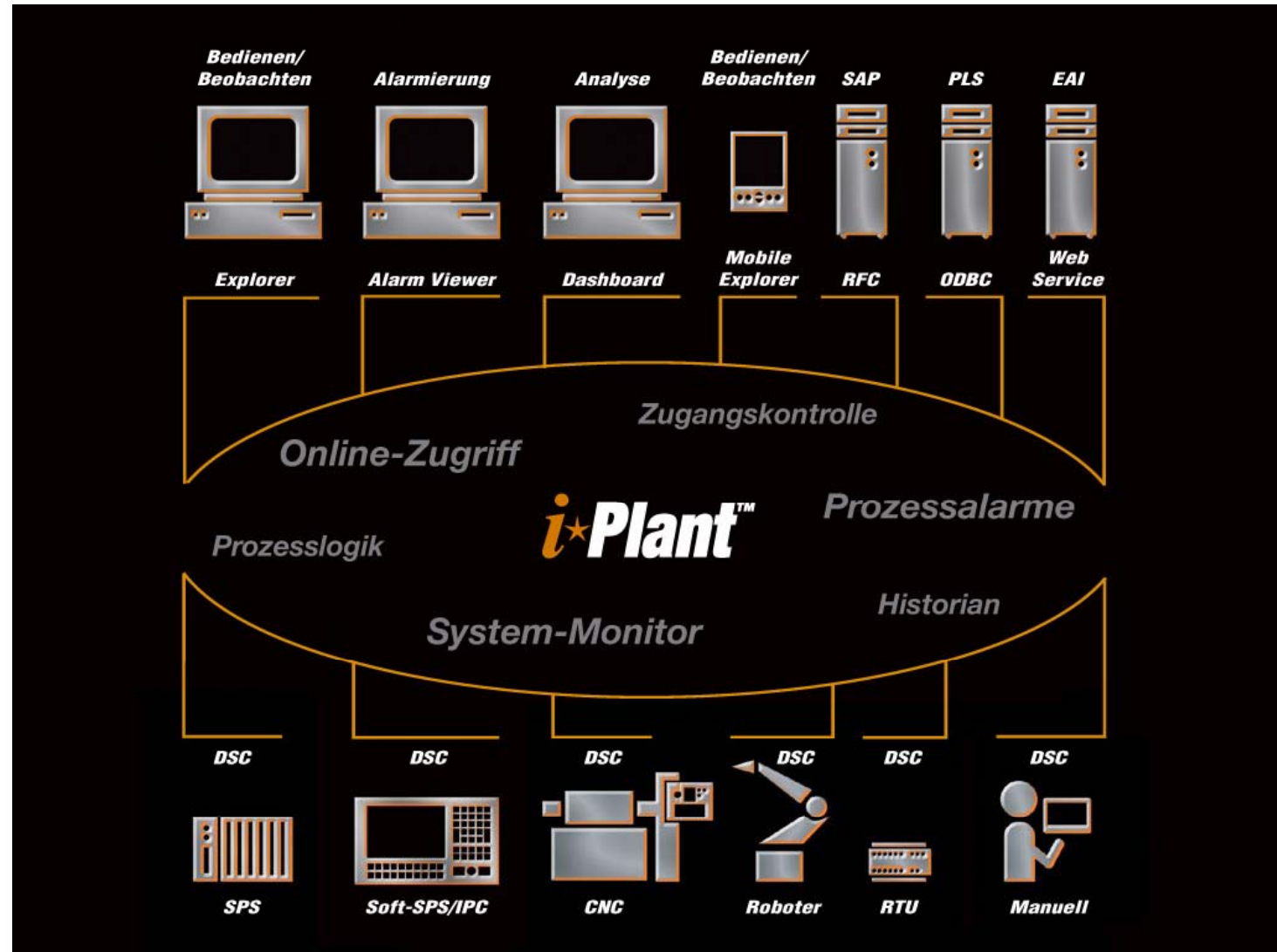
- Langzeitexperimente mit mehr als 5000 kontinuierlichen Anfragen und mehr als 50 Threads
- Deutliche Effizienzsteigerung durch Anfrageoptimierung

# Agenda

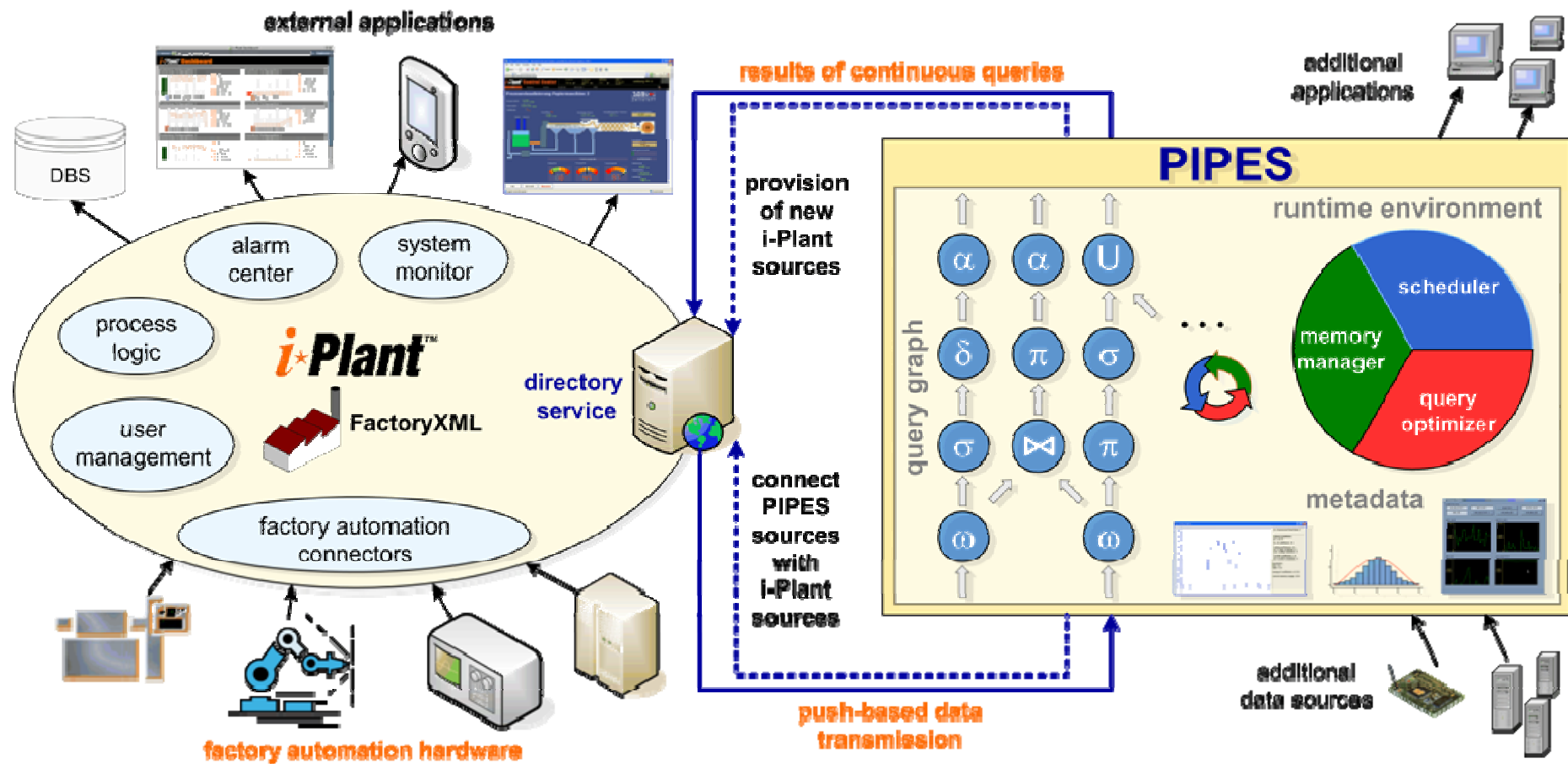
- **Production Intelligence**
  - Vertikale Integration und deren Nutzen
- **Kontinuierliche Anfragen über Datenströmen**
  - Spezifikation und Verarbeitungsprinzipien
- **Unsere Softwareinfrastruktur PIPES**
  - Architekturelle Grundlagen und Funktionsumfang
- **Einsatz von PIPES in der Fabrikautomation**
  - Kopplung an die Integrationsmiddleware i-Plant
  - Synergieeffekte
  - Kurze Demo

# i-Plant als Integrationsplattform

Kooperation  
mit



# Kopplung von i-Plant und PIPES



# Synergieeffekte

## ■ i-Plant

- Unterstützung komplexer kontinuierlicher Anfragen
  - Subquery-Sharing
  - Adaptive Laufzeitumgebung
  - Fortgeschrittene Datenanalysetechniken

## ■ PIPES

- Zugriff auf eine Vielzahl von Konnektoren
- Professionelle Front-Ends sowie Anbindung an diverse externe Applikationen
  - SCM und EAI Systeme, E-Mail und PDA Dienste
- Benutzerverwaltung, Backup-Dienste und Fehlertoleranz beim Zugriff auf i-Plant Quellen

# i-Plant™ Control Center

Produktion 730 kg ↑  
Sollmenge 15000 kg →  
Auftragsfortschritt 5 % ↑  
Auftragsende 27.06.05 19:27

Leistung PM 3

HMI Alarme Archiv Explorer Benutzer Hilfe

## Prozessvisualisierung Papiermaschine 3

AOE Zellstoff

Wasserzufuhr 4.29 m³/min

Faserzufuhr 104.76 kg/min

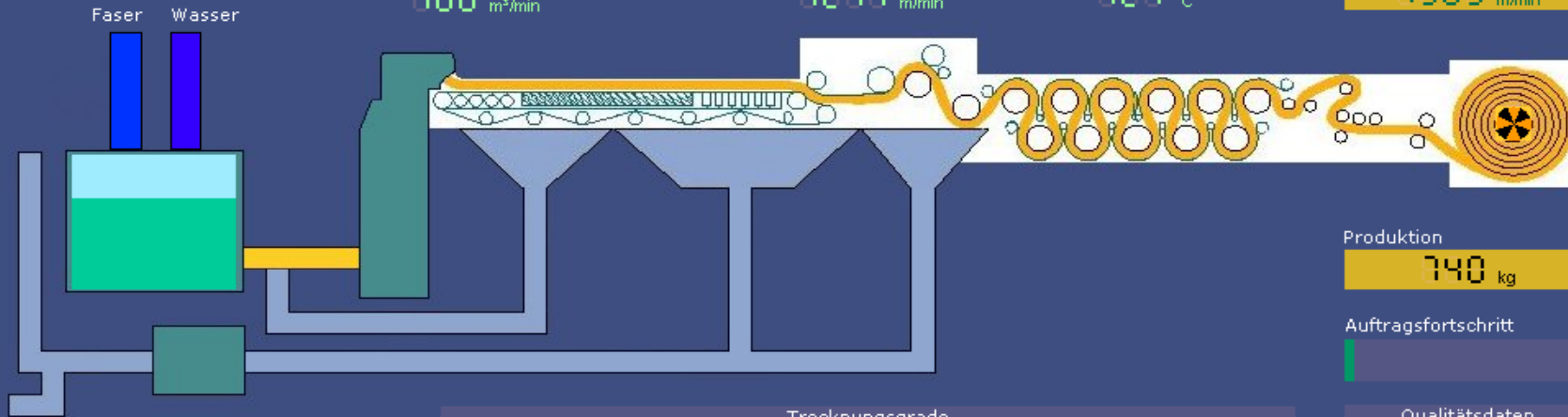
Stoffdichte 8.2

Durchfluss 100 m³/min

Durchsatz Presse 12.14 m/min

Dampftemperatur Trockner 121 °C

Aufrollgeschwindigkeit 1585 m/min



Produktion 740 kg

Auftragsfortschritt

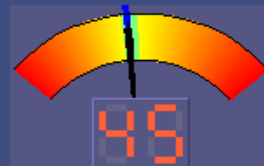
Qualitätsdaten

Trocknungsgrade

Siebpartie

Presspartie

Trockenpartie



Bahnbreite 800 cm

Papierstärke 2000 µm

Reißlänge 482 km

HMI

BDE-MDE

Übersicht

# Zusammenfassung

## ■ Production Intelligence

- Zeitnahe Bereitstellung/Analyse großer Datenmengen
- DBMS und direkte Programmierung ungeeignet
- Direkte Verarbeitung der Datenströme mittels PIPES

## ■ PIPES

- Flexible, modulare und erweiterbare Architektur
- Hohe Effizienz bei der Anfrageauswertung
- Adaptive und skalierbare Laufzeitumgebung
- Einfache Formulierung komplexer Applikationslogik

Philipps



Universität  
Marburg

Vielen Dank für Ihre  
Aufmerksamkeit!

**Haben Sie Fragen?**