

IoTからデータ解析の道筋での PostgreSQLの使われ方

PostgreSQL Conference Japan 2019

(2019/11/15 r1)

 **住友金属鉱山株式会社**

工務本部 設備技術開発部(青梅)
グループ長 佐藤 健司

目 次

1. 住友金属鉱山の紹介
2. IoT事例
3. IoTからデータ解析に至る道筋
4. IoT領域のデータ処理に必要な技術
5. データ解析の為に必要なDB技術
6. 付加価値の高いエッジでのデータ処理
7. まとめ

1-1. 住友金属鉱山 (Sumitomo Metal Mining) の紹介

創業 429年 (1590～)

海外拠点

国内拠点



・鉱山

USA
ペルー
チリ

モレンシー
セロ・ベルデ
ケブラダ・ブランカ
シエラ・ゴルダ, カンデラリア,
オホス・デル・サラド
ノースパークス

オーストラリア

・製錬所

中国
フィリピン

金隆
コーラルベイ, タガニート,
ニッケルアジア

インドネシア
ニューカレドニア

ソロワコ
フィゲスバル

・工場

中国
台湾

上海, 広東省東莞,
高雄

1-2. 住友金属鉱山の主な事業

資源 金鉱石、銅鉱石などの採掘



銅

ニッケル

金

製錬 鉱石から有価金属(金、銀、銅、ニッケル、コバルトなど)を取り出す



金



ニッケル



コバルト



銅

材料 金属を加工した素材を製造



電池材料



ペースト



銅ポリイミド基板

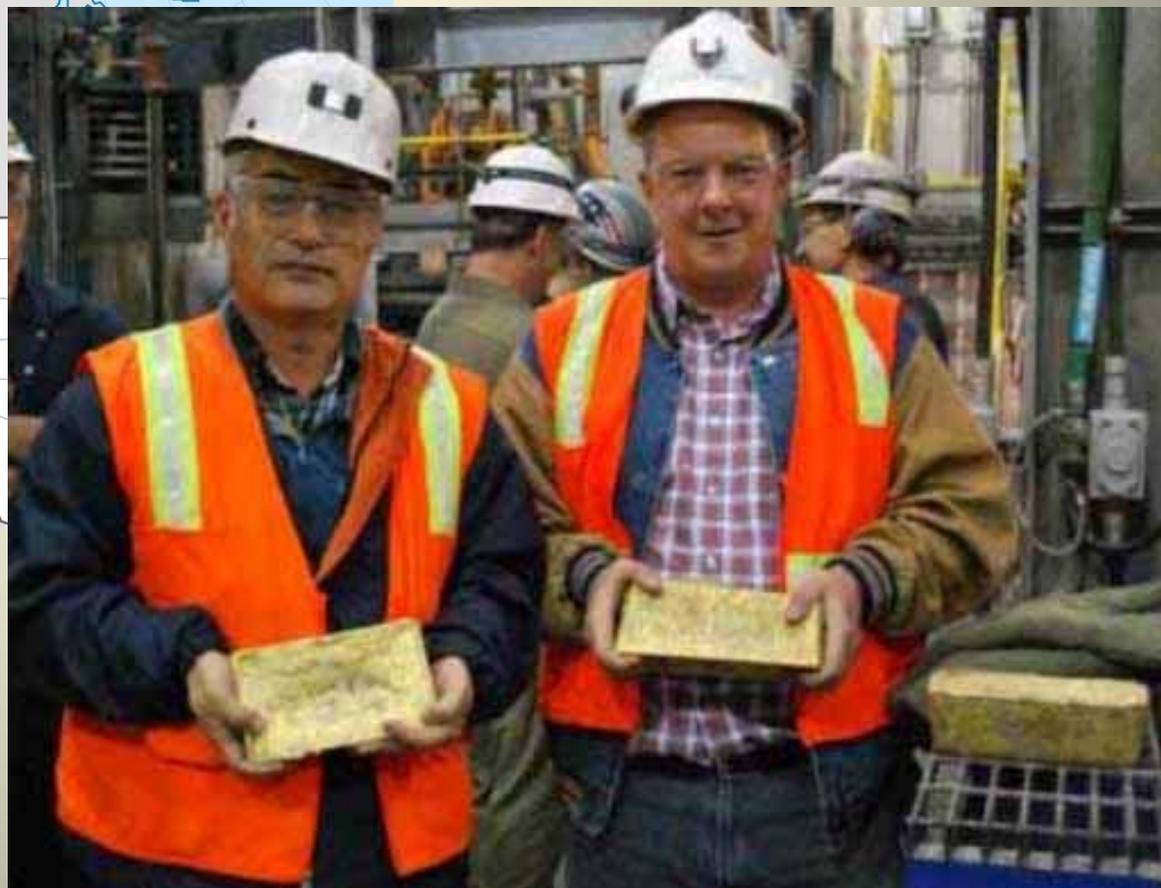
目 次

1. 住友金属鉱山の紹介
- 2. IoT事例**
3. IoTからデータ解析に至る道筋
4. IoT領域のデータ処理に必要な技術
5. データ解析の為に必要なDB技術
6. 付加価値の高いエッジでのデータ処
7. まとめ

1-5. 海外鉱山の一例



米国のアラスカ州にある
金鉱山

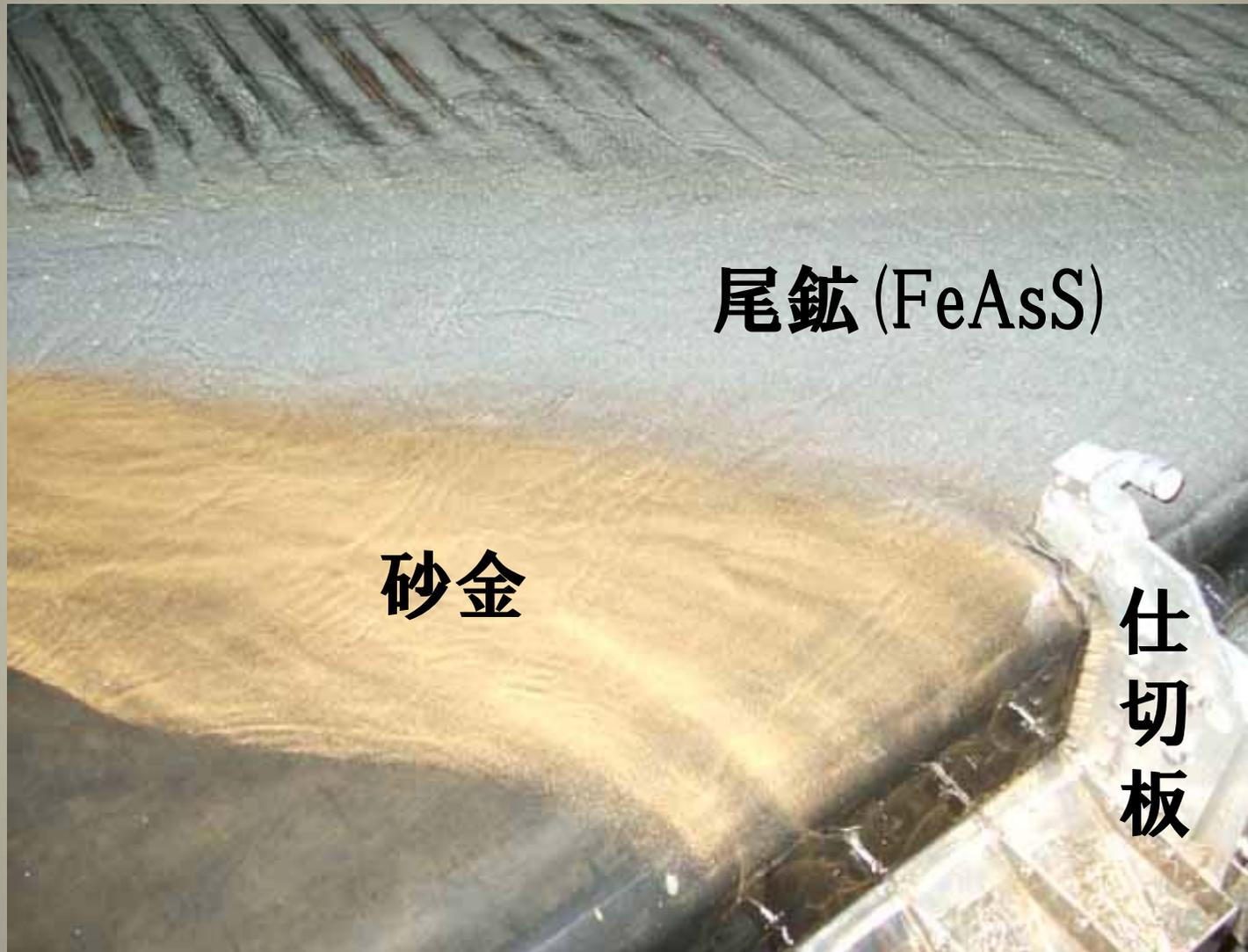


Pogo鉱山は、2018/10に
Northern Star Resources Ltd.へ売却

Sumitomo Metal Mining Co.LTD.

17.42kg 9100万円

2-2. IoT事例 [テーブル選鉱機上の仕切板制御]



日本鉱業協会主催 全国鉱山・製錬所現場担当者会議で発表 (2016/6)

Sumitomo Metal Mining Co.LTD.

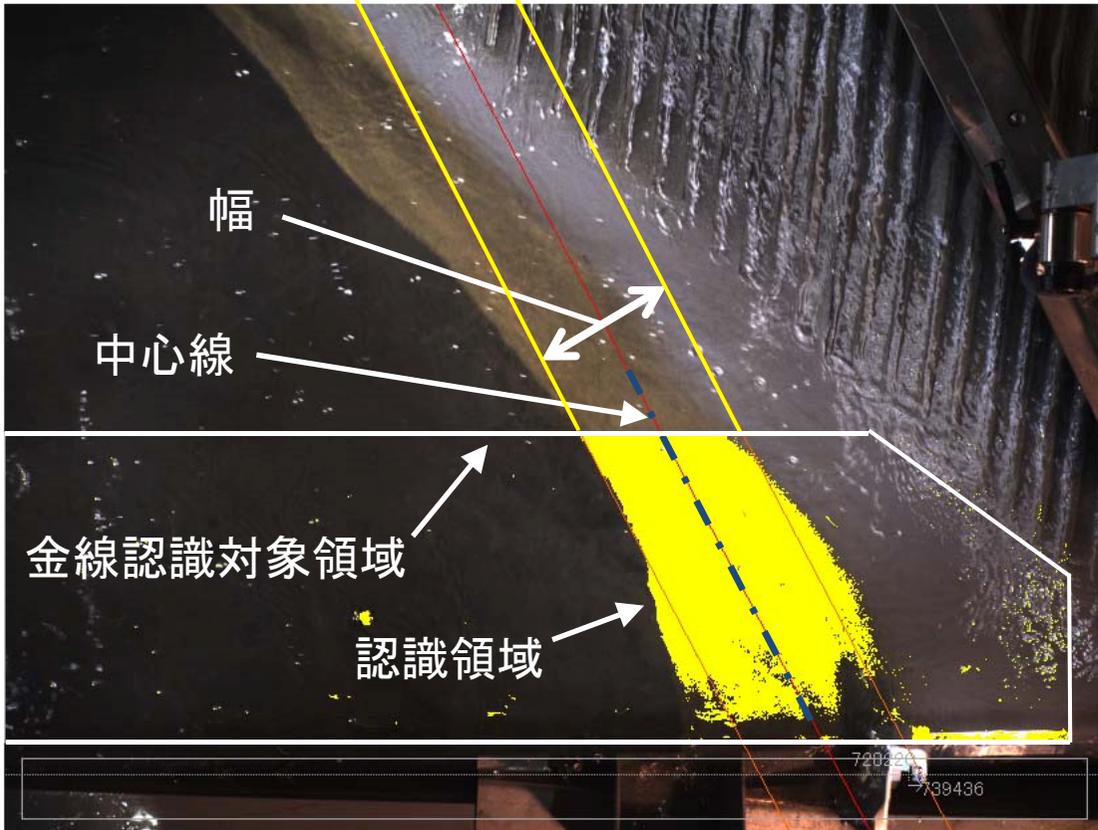
2-3. 金の認識と仕切板の位置制御

10.20.254.156/pogo_tc/html/job110.php?n=117372&t=2014-01-28 19:40:28

Monitor Real time monitor pogo : sato Table concentrator

Close

File name img_2014012819402800.png Create time 2014/01/28 19:40:28.00 Show line Original size



幅

中心線

金線認識対象領域

認識領域

Show binarize image

Show marker detection

Binarize threshold (lower) 40

Binarize threshold (upper) 255

Show line detection

Binarize threshold (lower) 30

Binarize threshold (upper) 250

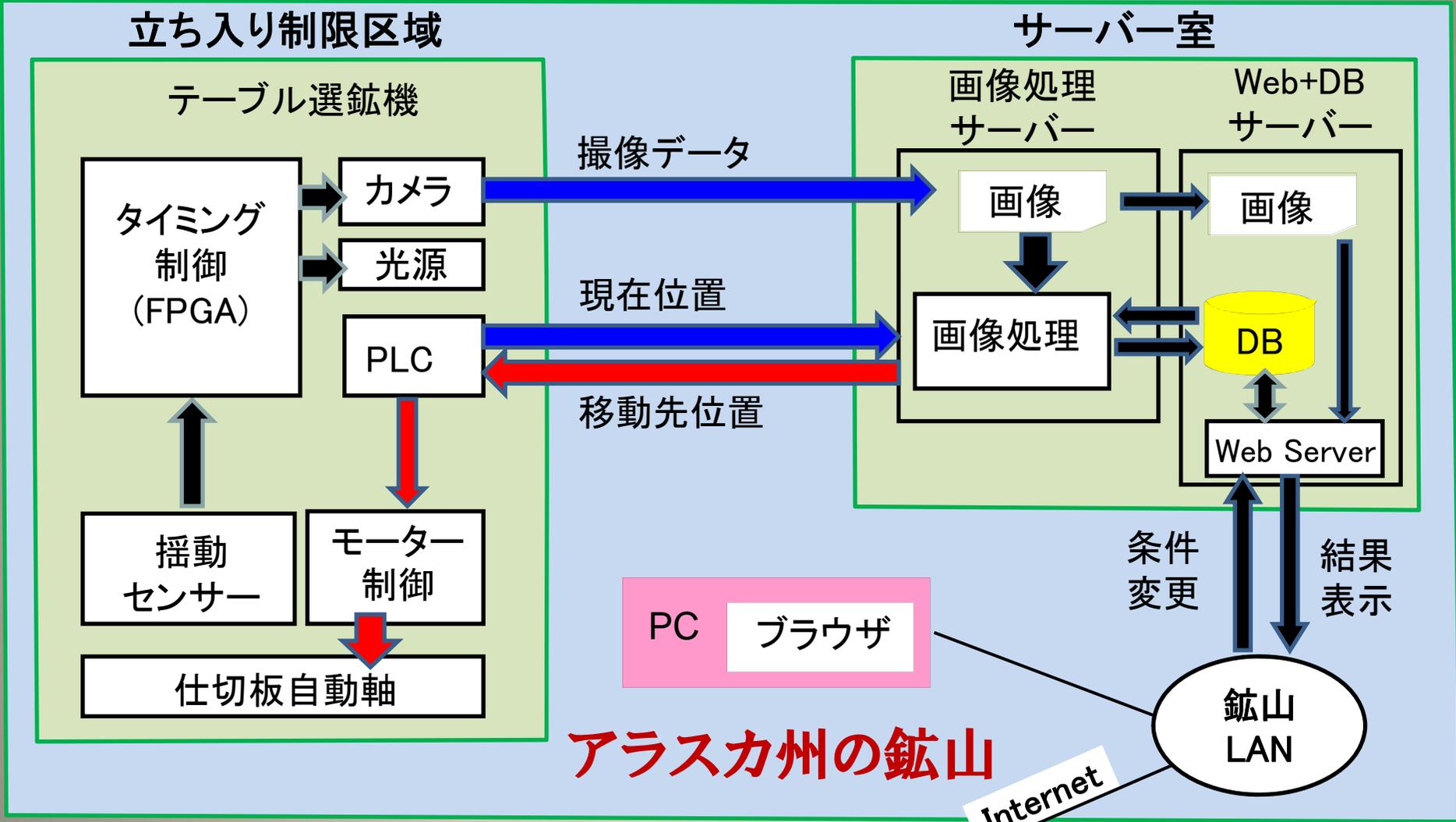
Show overlay image

Reload

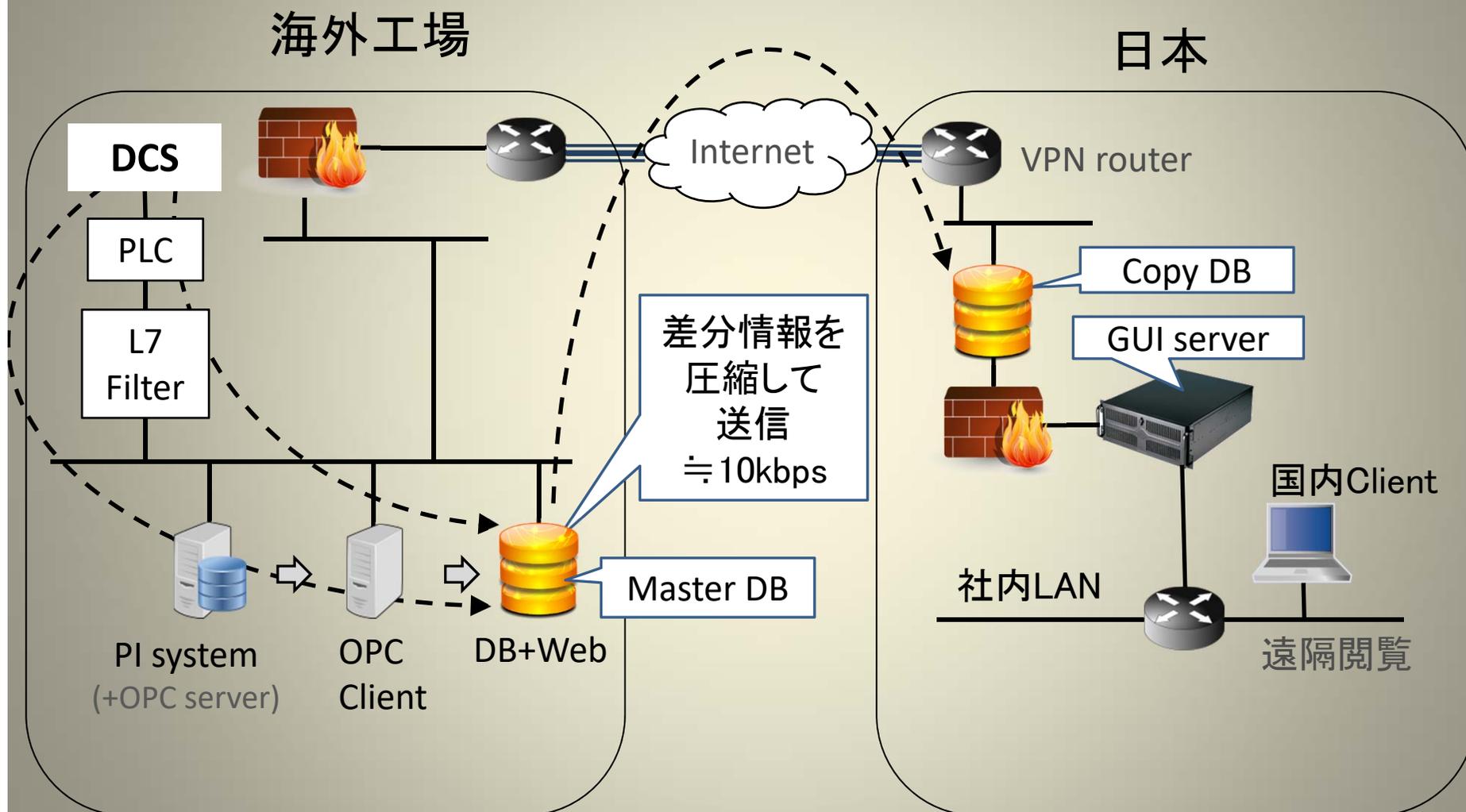
739436

739436

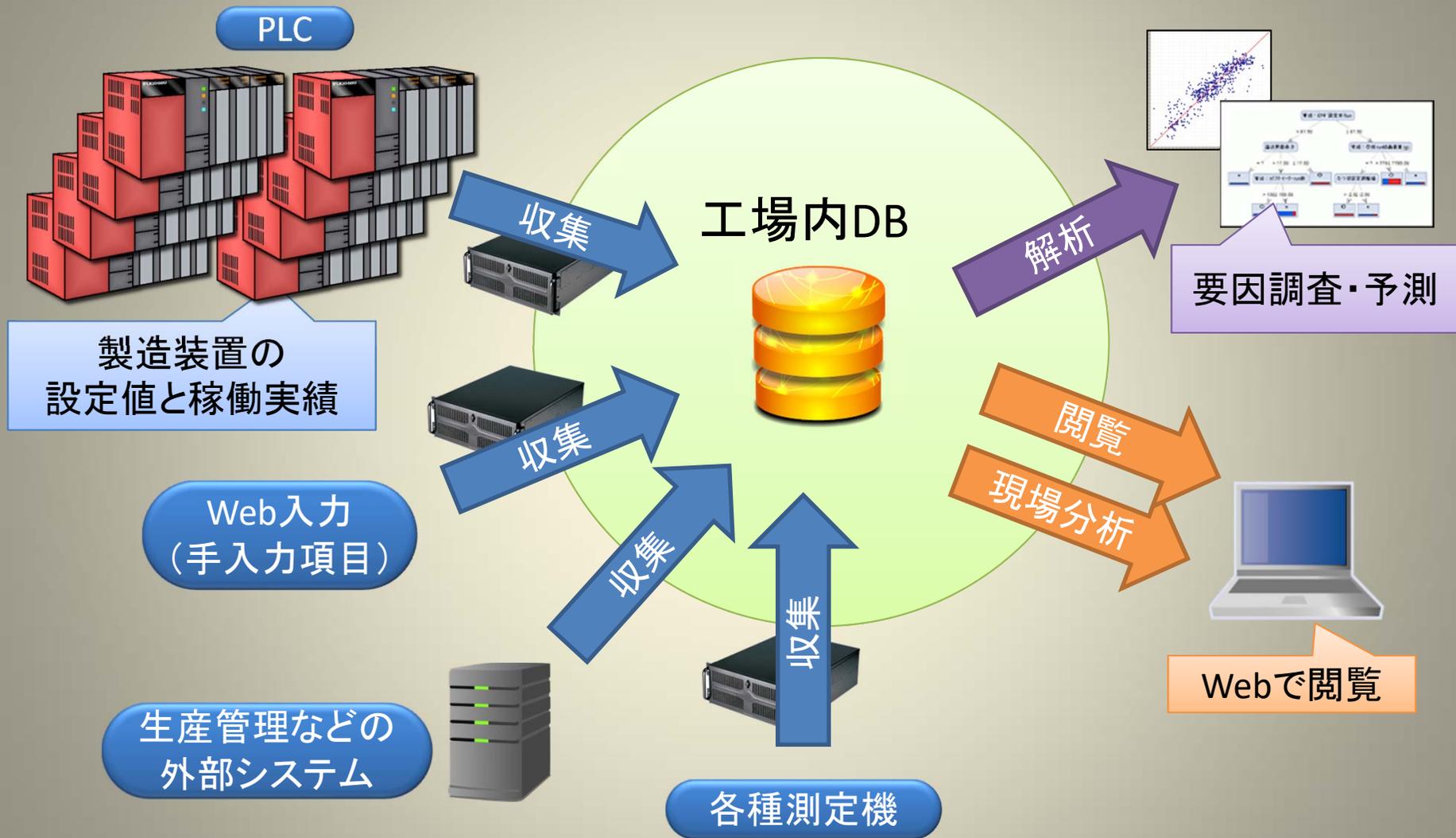
2-4. システム構成



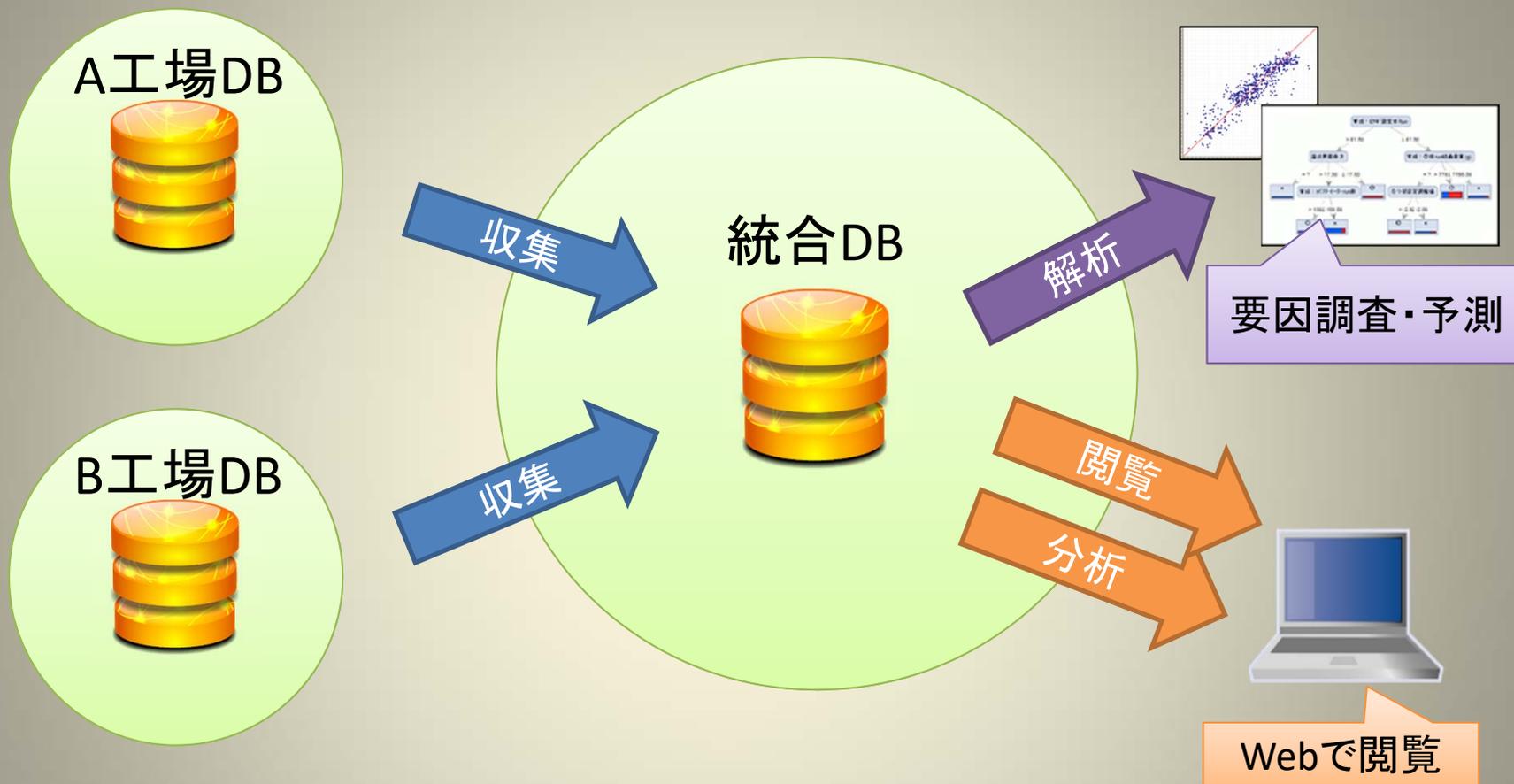
2-5. 遠隔閲覧/制御



2-6. 大規模データ収集システム



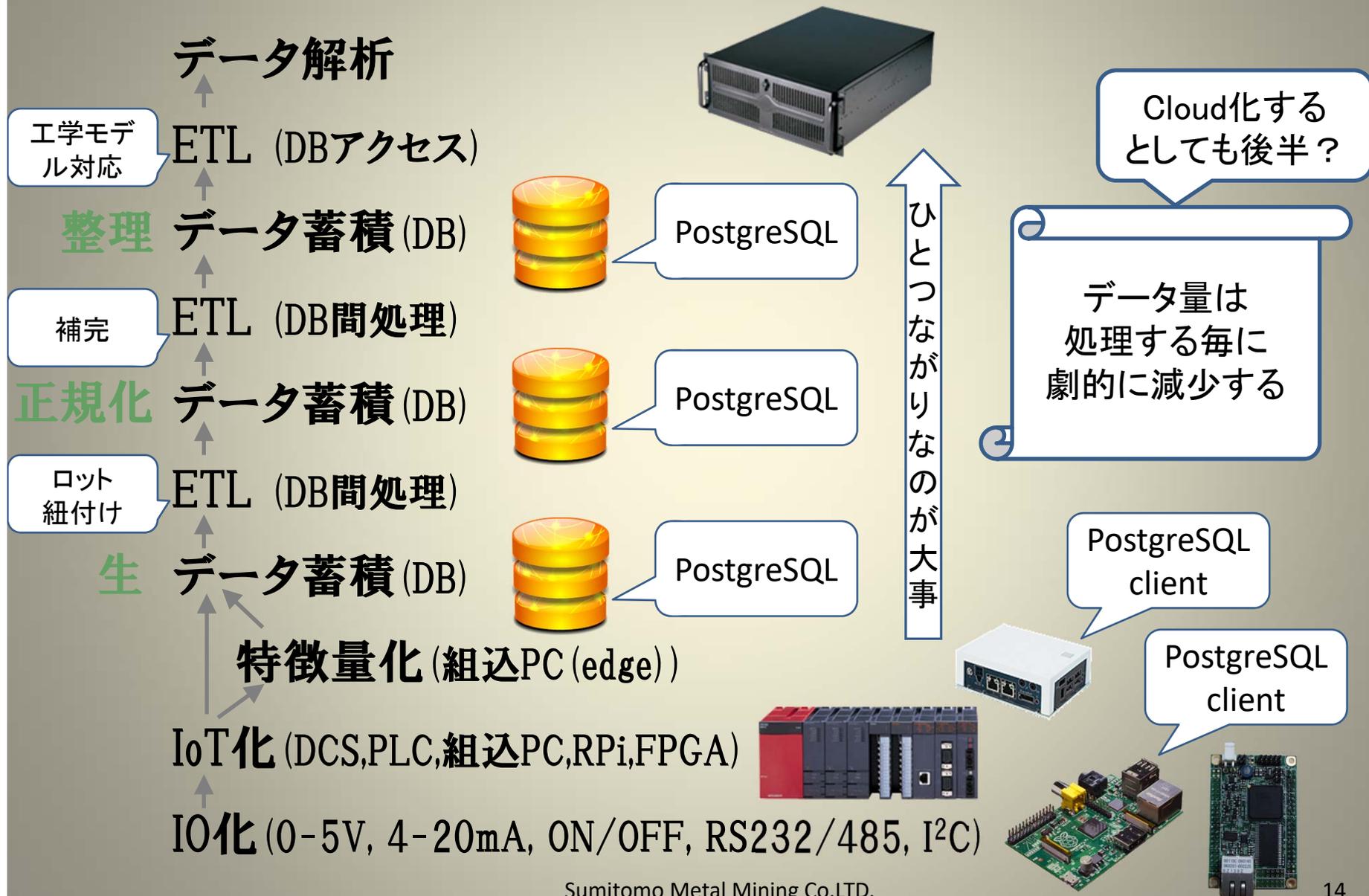
2-7. 統合データ管理システム



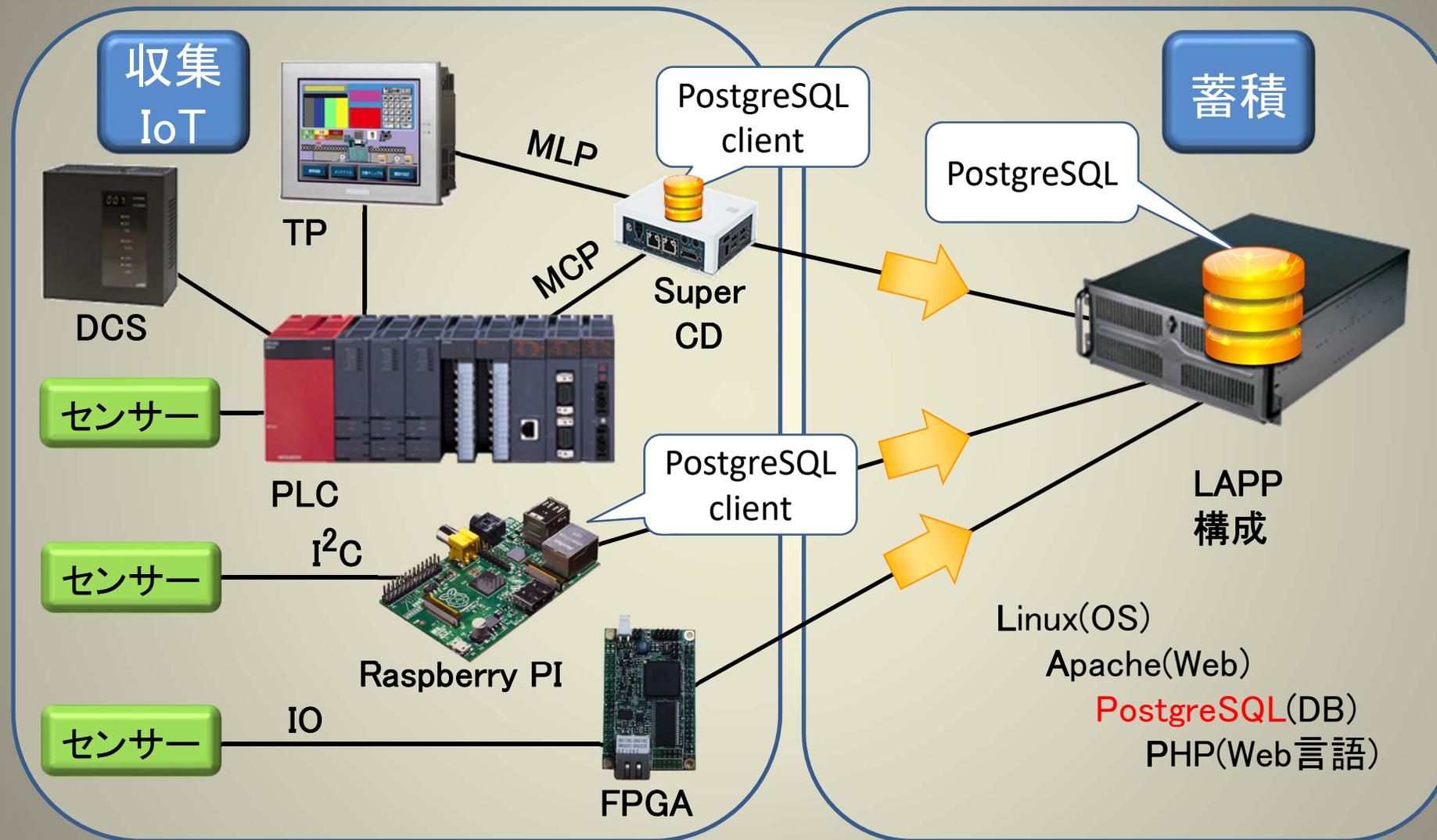
目 次

1. 住友金属鉱山の紹介
2. IoT事例
- 3. IoTからデータ解析に至る道筋**
4. IoT領域のデータ処理に必要な技術
5. データ解析の為に必要なDB技術
6. 付加価値の高いエッジでのデータ処理
7. まとめ

3-1. IoTからデータ解析に至る道筋



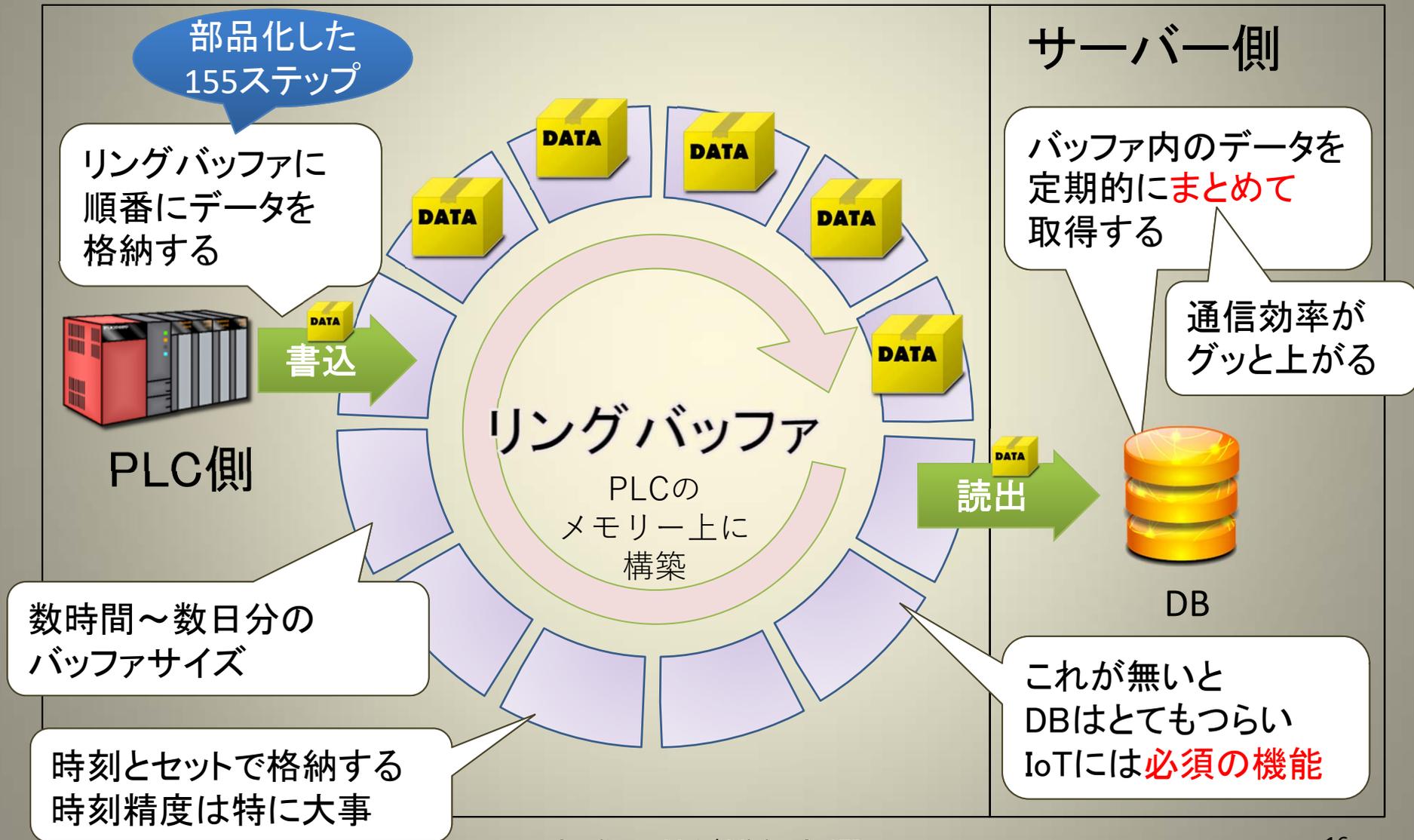
3-2. IoTの実装イメージ



※1. MLP : Memory Link Protocol

※2. MCP : MELSEC Communication Protocol

3-3. PLCに追加したリングバッファ機能



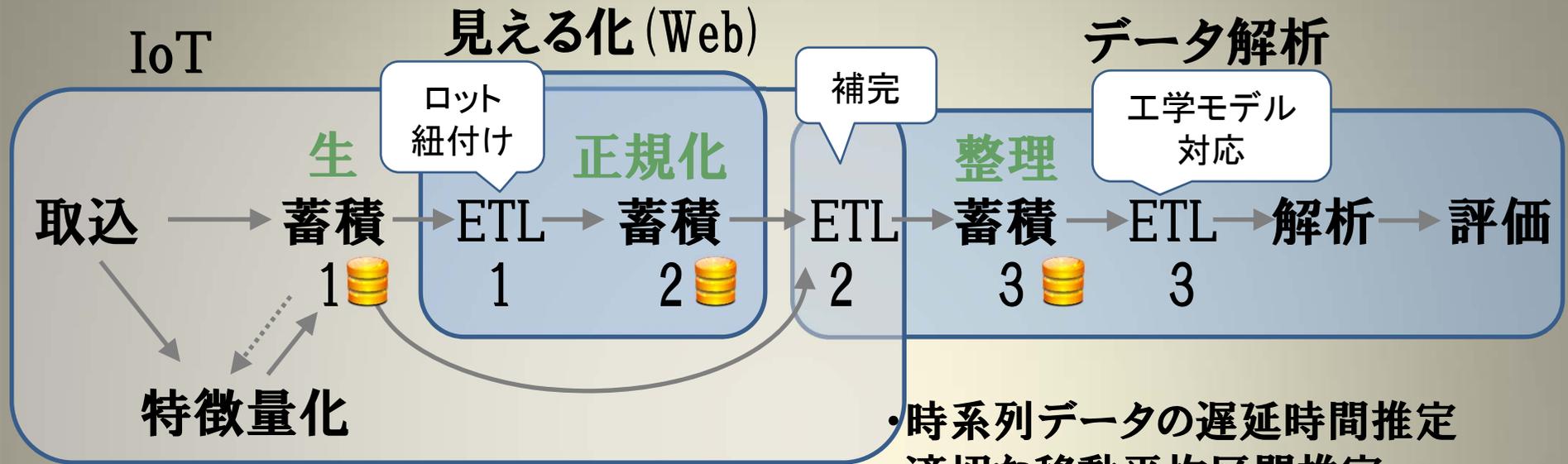
3-4. PLCの時刻同期

- PLCの時計はズれる（±8秒/日くらいの覚悟が必要）
- 原因は水晶振動子の周波数精度が±50～100ppmくらいだから
- ntp同期の機能が無いものがある
- ntpサーバー設定を1つしかできないものがある
- ntp同期に1度失敗すると二度と同期しない仕様のものがある
- 外部でPLCの時刻ズレ量を監視して閾値を超えたら自動設定

目 次

1. 住友金属鉱山の紹介
2. IoT事例
3. IoTからデータ解析に至る道筋
- 4. IoT領域のデータ処理に必要な技術**
5. データ解析の為に必要なDB技術
6. 付加価値の高いエッジでのデータ処理
7. まとめ

4-1. IoT領域のデータ処理に必要な技術



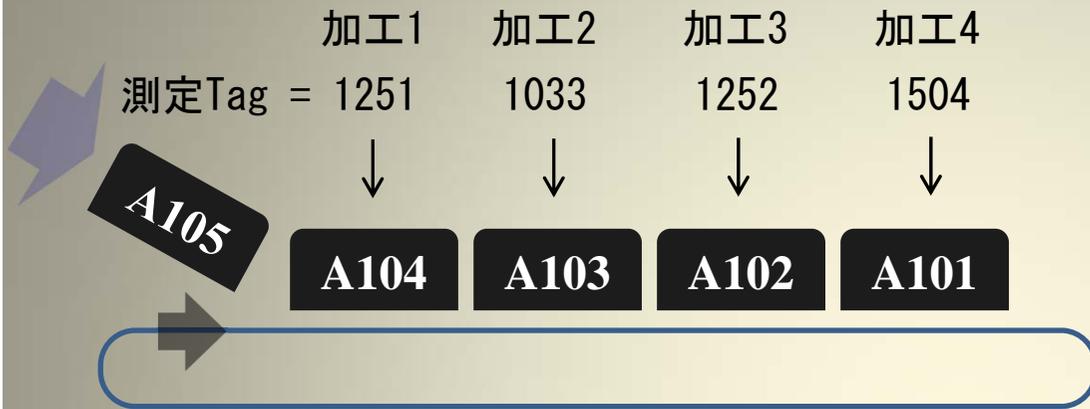
- 通信プロトコル
MCP, MLP, I²C, http
- 通信制御
PHP, C
- IO制御
ラダー, C, PHP, VHDL

- 蓄積DB制御
SQL
- ETL
SQL, C, PHP, Python

- 時系列データの遅延時間推定
- 適切な移動平均区間推定
- 補完
- 説明変数と目的変数の関係性調査
- 異常値の排除
- 重回帰
- Lasso
- 交差検定
- 製造プロセスにおける
物理化学モデルの理解



4-2. 時系列生データとロット別の整理データ



Tag No	A101	A102	A103
1251	0.80	0.81	0.80
1033	1076	1077	1077
1252	0.82	0.68	0.78
1504	1.20	1.25	1.14

Tag No	10:00	10:01	10:02	10:03	10:04
1251	0.80	0.81	0.80	0.79	0.78
1033	1075	1076	1077	1077	1078
1252	0.78	0.79	0.82	0.68	0.78
1504	1.04	1.10	1.13	1.20	1.25

整理

生

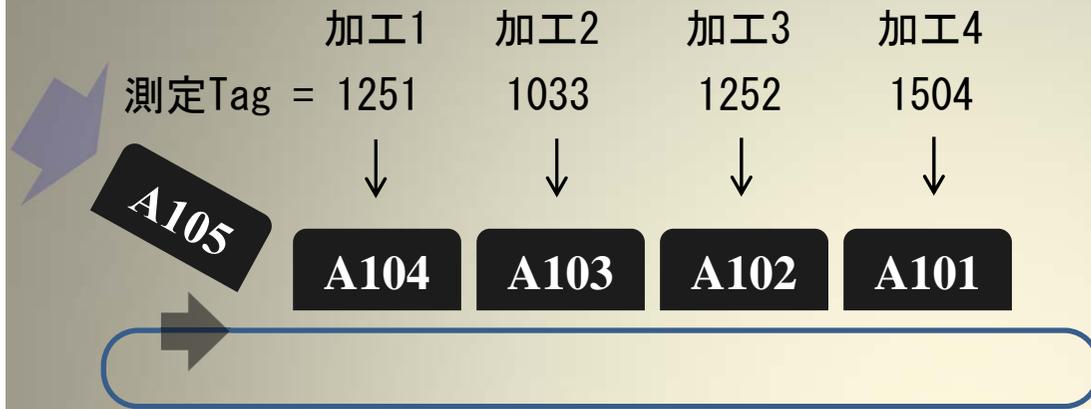
品質バラツキ原因をデータ解析で調査するにはID別の整列データへ変換しなければならない
(変換できるようにIDやTagを割り振らなければならない)

同一ロット
ロットID=A101

ID=A102

ID=A103

4-3. 補完や値の推定処理

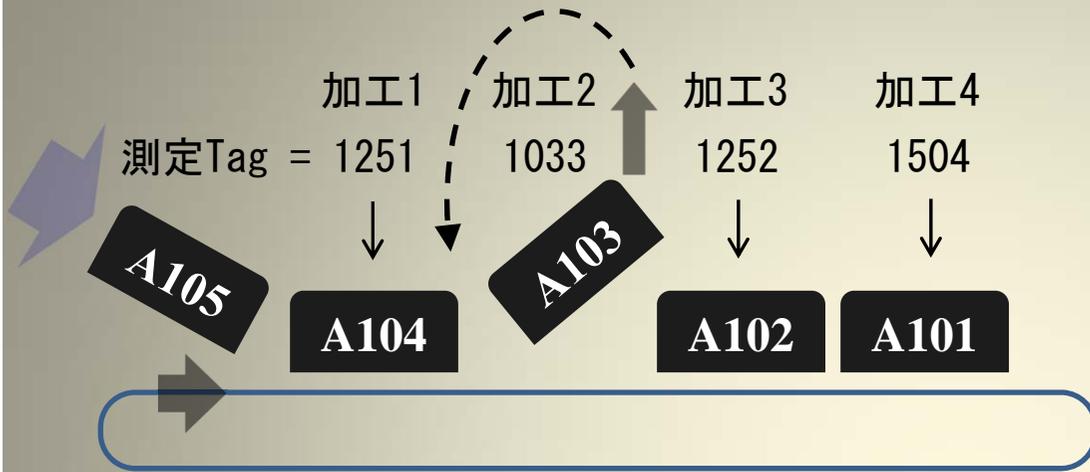


Tag No	10:00	10:01	10:02	10:03	10:04
1251	0.80	0.81	欠測	0.79	0.78
1033	1075		1077		1078
1252	0.78			0.68	
1504	1.04	1.10	1.13	1.20	1.25

同一ロット

- ・現実には、こんな状況になっている
- ・欠測値を補完し、測定と測定間の推定値を求める必要がある
- ・イベントにより大きく変動する値は補完や推定が困難なので、イベント駆動形の測定が大事

4-4. イベント駆動式記録の重要性



工程異常の懸念でA103ロットを取り出され、A104の進み具合が早まったら？
 (ログデータとロットの時間関係がずれる)



A103の調査の結果、問題が無かったのでA105が通り過ぎた後にA103を流したら!?
 (物の順番が入れ替わったら、ログデータとロットを紐付ける整理作業はどうなる?)



Tag No	10:00	10:01	10:02	10:03	10:04
1251	0.80	0.81	0.80	0.79	0.78
1033	1075	1076	1077	1077	1078
1252	0.78	0.79	0.82	0.68	0.78
1504	1.04	1.10	1.13	1.20	1.25

根本的な解決には、ロットの切り替わりの度にデータを取り、ロットと紐付けてデータを保存すべき(=イベント駆動式記録)

ログデータ

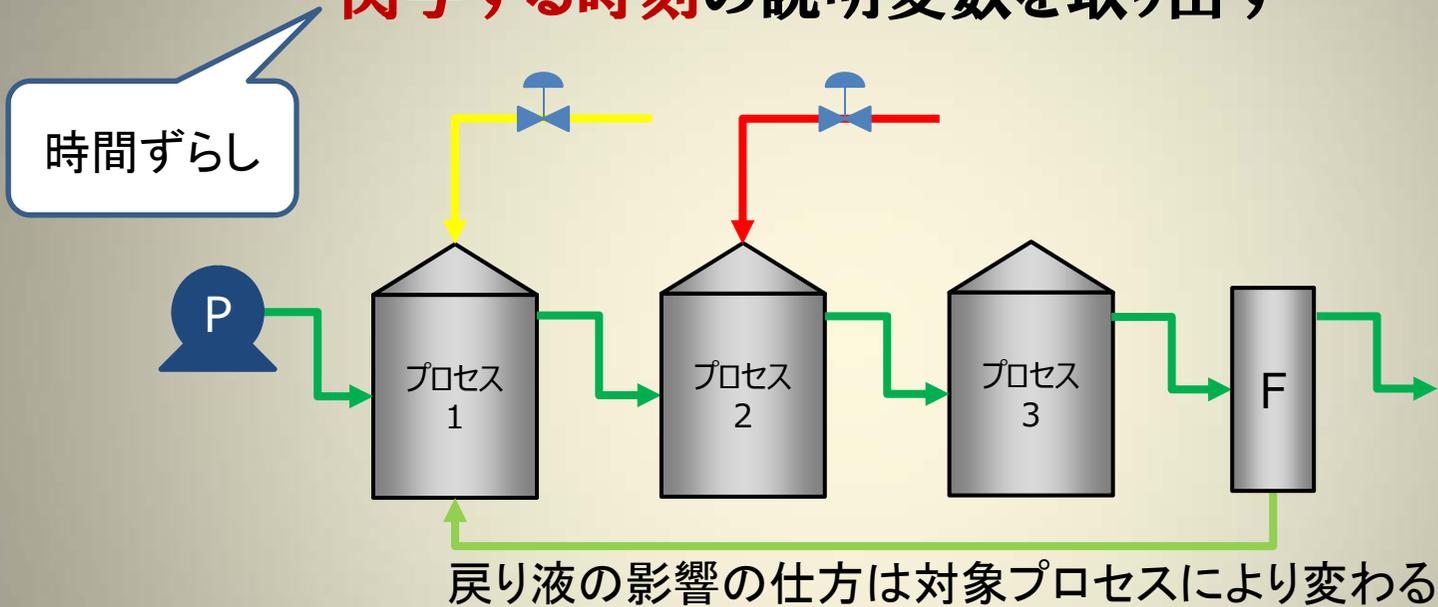
同一ロット
ロットID=A101

ID=A102

ID=A103 ?

4-5. 解析対象別の整理

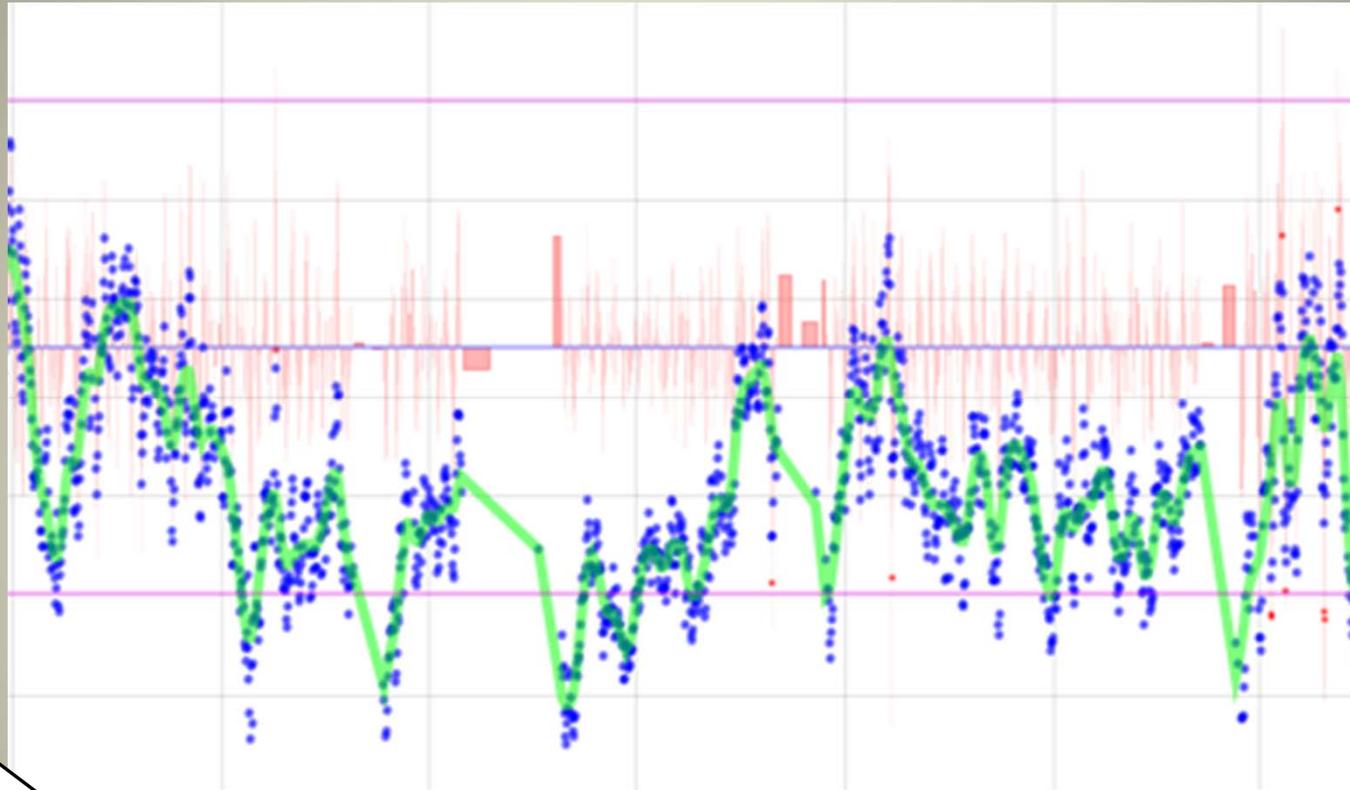
- ・連続操作データから目的変数に
関与する時刻の説明変数を取り出す



解析対象

プロセス1	-0.1h	±0.0h	+0.5h	+1.0h	+1.3h	+1.5h
プロセス2	-0.6h	-0.5h	±0.0h	+0.5h	+0.8h	+1.0h
					-0.1h	-0.6h

4-6. 時系列データ処理 (移動平均)



これが
曲者

- 適切な移動平均区間による
欠損値補間と
任意時刻の値取り出し

補完

適切な移動平均区間と
解析に都合の良い
データ間隔と
適切な時間ずらし量を
決めたら
SQLで一気にデータ変換

目 次

1. 住友金属鉱山の紹介
2. IoT事例
3. IoTからデータ解析に至る道筋
4. IoT領域のデータ処理に必要な技術
- 5. データ解析の為に必要なDB技術**
6. 付加価値の高いエッジでのデータ処理
7. まとめ

5-1. データ解析の為に必要なDB技術

- ETL (3)

解析用の抽出



- 蓄積 (整理)

解析目的別の巨大な単一テーブル

- ETL (2)

解析用の抽出、異常値処理、欠損値補間、移動平均



- 蓄積 (正規化)

通常のWebシステムと同様、見える化Web

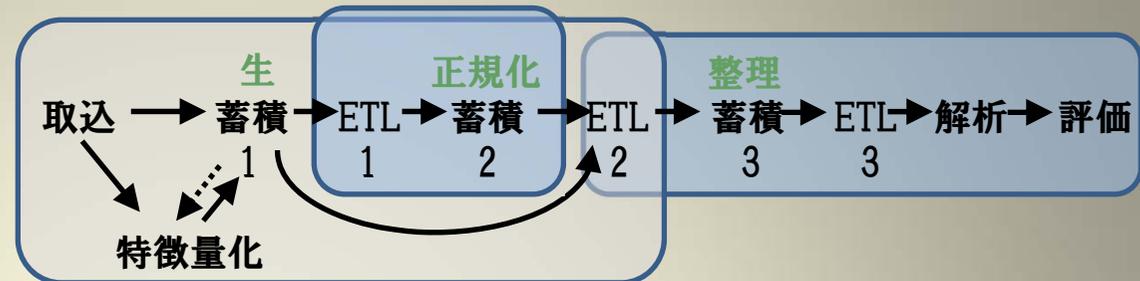
- ETL (1)

ロット紐付け



- 蓄積 (生)

月や年単位でのテーブル・パーティショニング



データ前処理の
殆どをDB内演算で
済ませている

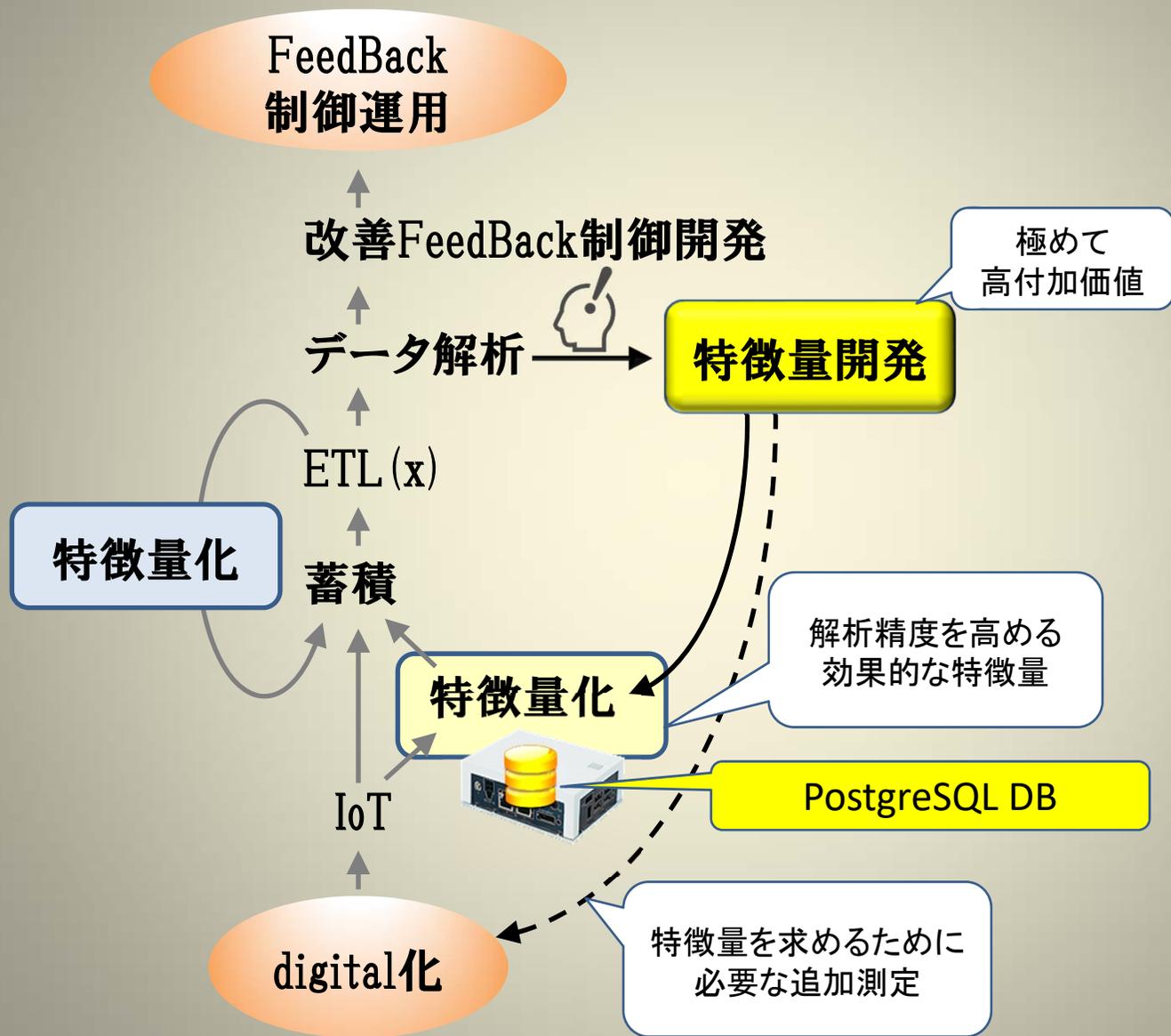


解析ソフトの
負荷が低くなる

目 次

1. 住友金属鉱山の紹介
2. IoT事例
3. IoTからデータ解析に至る道筋
4. IoT領域のデータ処理に必要な技術
5. データ解析の為に必要なDB技術
- 6. 付加価値の高いエッジでのデータ処理**
7. まとめ

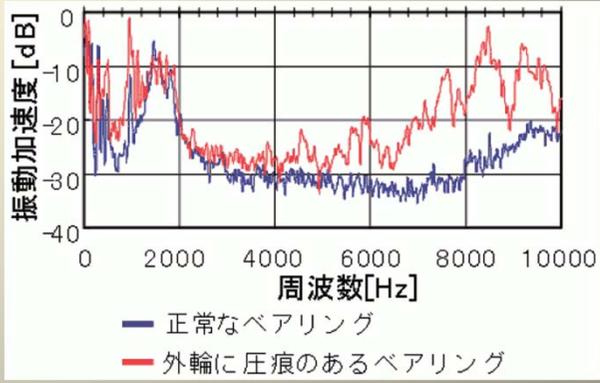
6-1. エッジでのデータ処理



6-2. エッジ・データ処理方式のデータ量推移

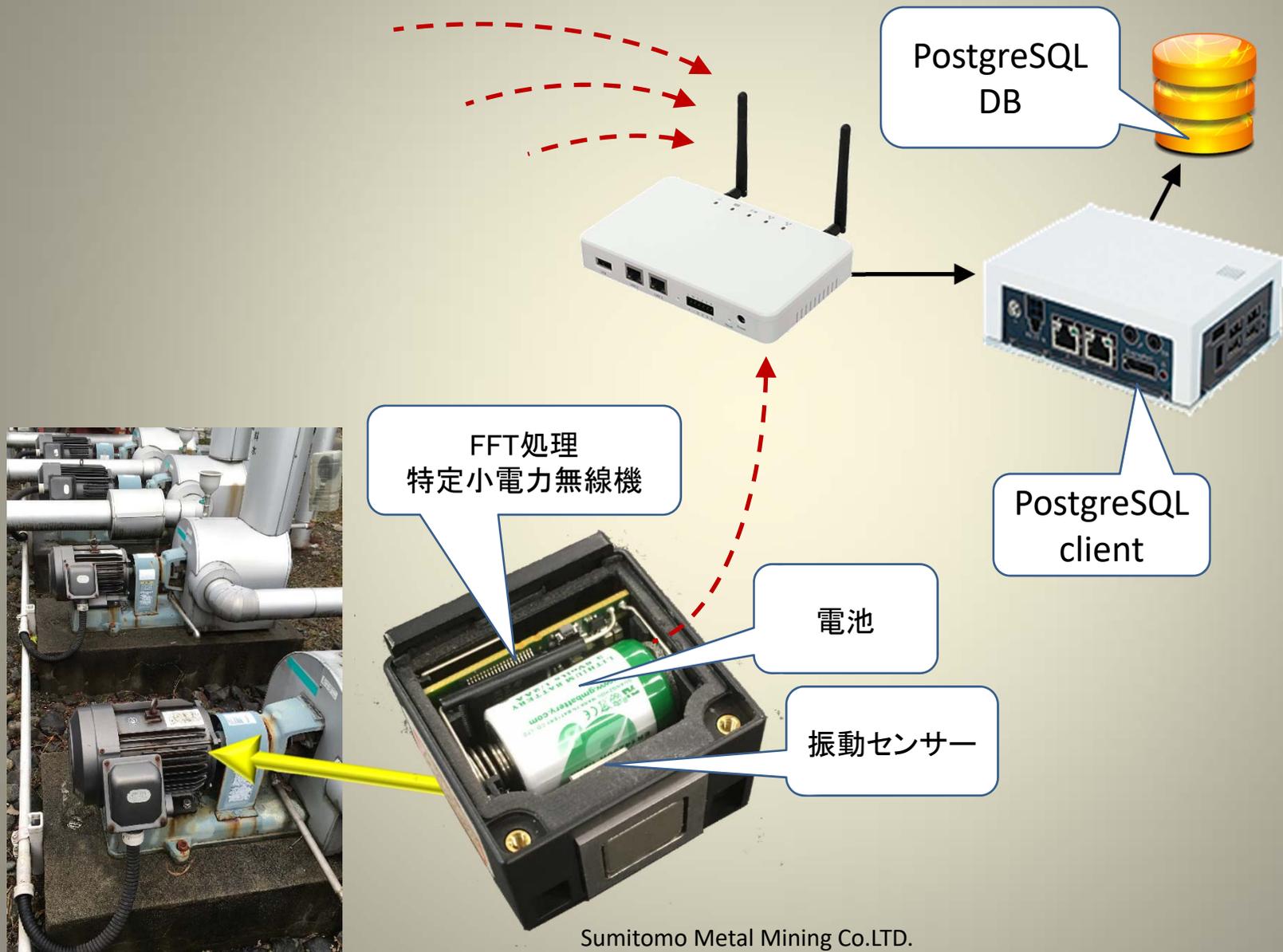


多点化 300箇所分	72M bps	1.6k bps	2.4k bps
エッジ (特徴量化)	なし 240k bps	FFT化 上位5ヶ/分 (f,a)x5 5.3 bps	Wavelet化 上位5ヶ/分 (f,a,ts,te)x5 8 bps
IoT	10kHz sampling x 3軸 x 8bit分解能 = 240k bps		



NEDO インフラモニタリング技術 p35 より引用

6-3. エッジ・データ処理方式



目 次

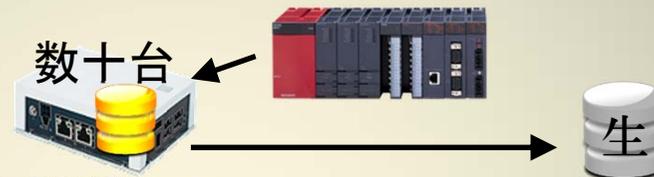
1. 住友金属鉱山の紹介
2. IoT事例
3. IoTからデータ解析に至る道筋
4. IoT領域のデータ処理に必要な技術
5. データ解析の為に必要なDB技術
6. 付加価値の高いエッジでのデータ処理
- 7. まとめ**

まとめ

PostgreSQLは工場の中で沢山使われている

- IoT

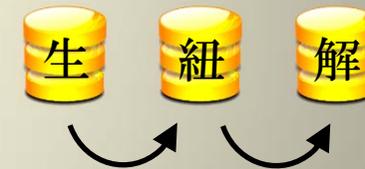
パッシブなPLC等の制御機器とDBサーバーを繋ぐedge



- 所謂DBらしい使い方のDB

生データDB、製品紐付DB、データ解析用DB

nTB 0.0nTB 0.00nTB



- 見える化用

Webシステムや品質管理システムとしてのDB

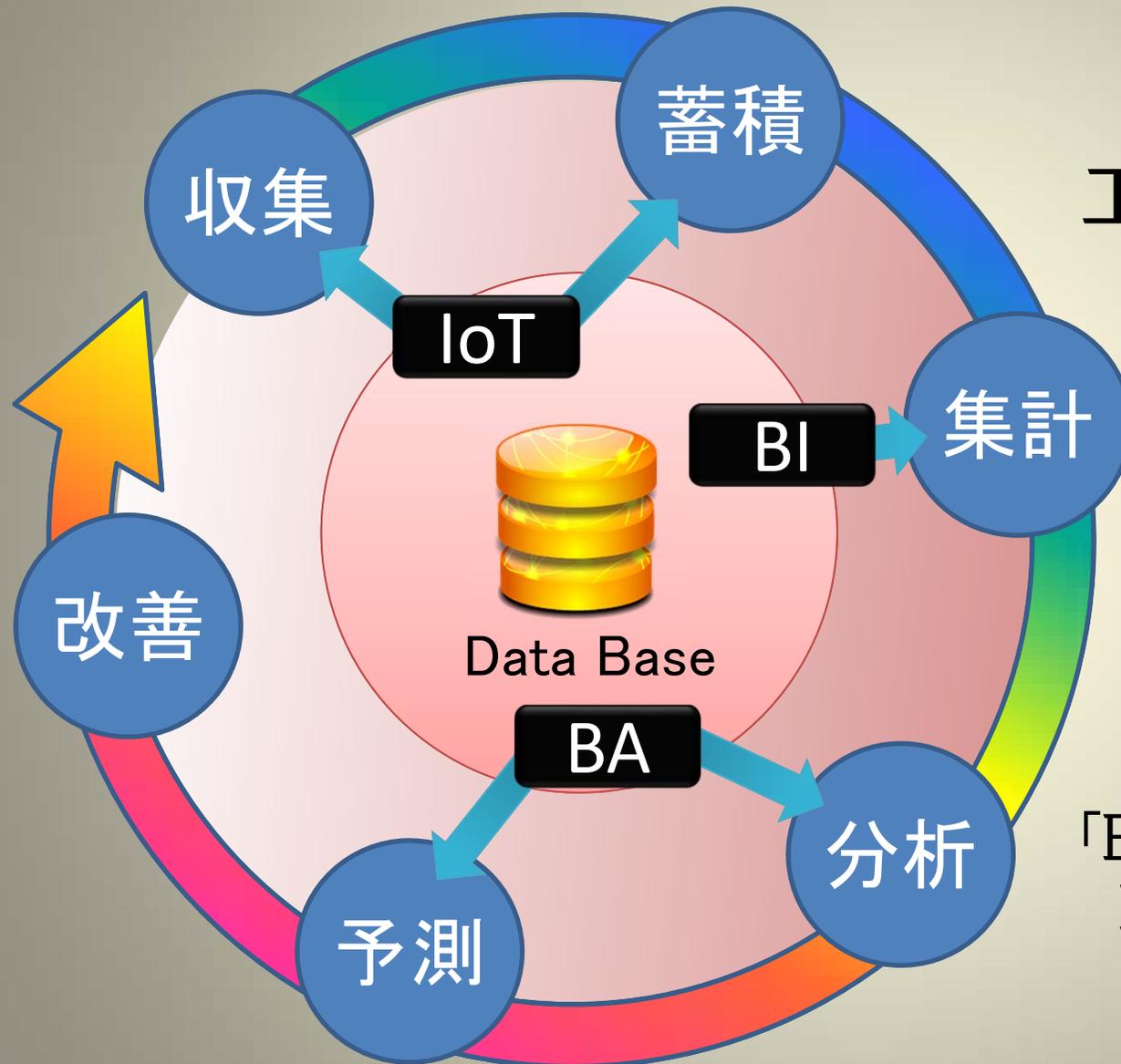


- 特徴量化

edgeコンピューティング用DB



現場の熱意が大事



工場が本気になると
ループが自然に
回りだす

「Excelの限界を超えるには
DataBaseの活用が必須」
という啓蒙活動が大事