Self-Optimized Strategy for IO Accelerator Parametrization

Lionel Vincent, Gaël Goret, Mamady Nabe, Trong Ton Pham

Bull - Atos Technologies

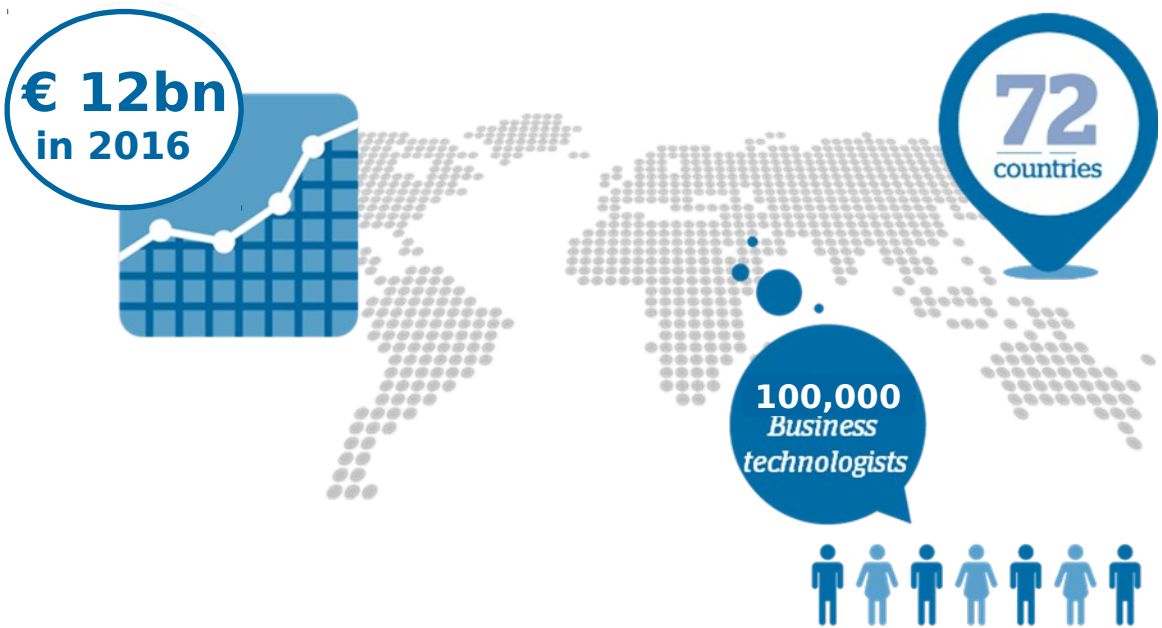
28-06-2018

WOPSSS - 18

Bull atos technologies

We are THE European IT Leader

and a top 5 Digital services player worldwide



Atos

worldline
e-payment services

Bull
atos technologies

UNIFY

Bull
atos technologies

This is our Mission within Atos

WorldLine

Business
& Platforms Solutions

Big Data & Security
700M€, 3500 people

Infrastructure
& Data Management

1 Mission, 3 Go-to-Markets, 3 Business Units

BIG DATA & HPC
High-performance technologies for insight platforms & data management

CYBER SECURITY
Intelligent platforms for digital security, trust & compliance

MISSION-CRITICAL SYSTEMS
Intelligent systems for Defence, Homeland Security, Aerospace & Transport

Intelligence in a data-driven world

Bull
atos technologies

Big Data & HPC Business Unit

End-to End-offering to handle the most complex challenges

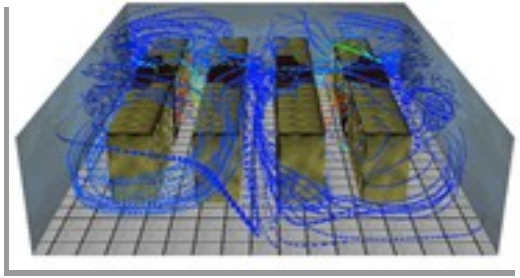
Supercomputers

Data Management

Software

HPCaaS/DLaaS

Data center

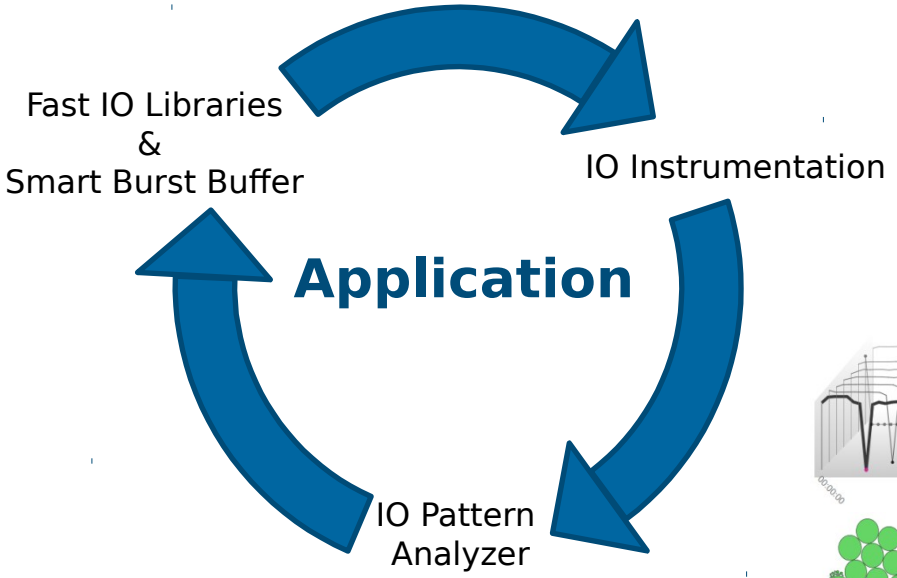


Expertise & services

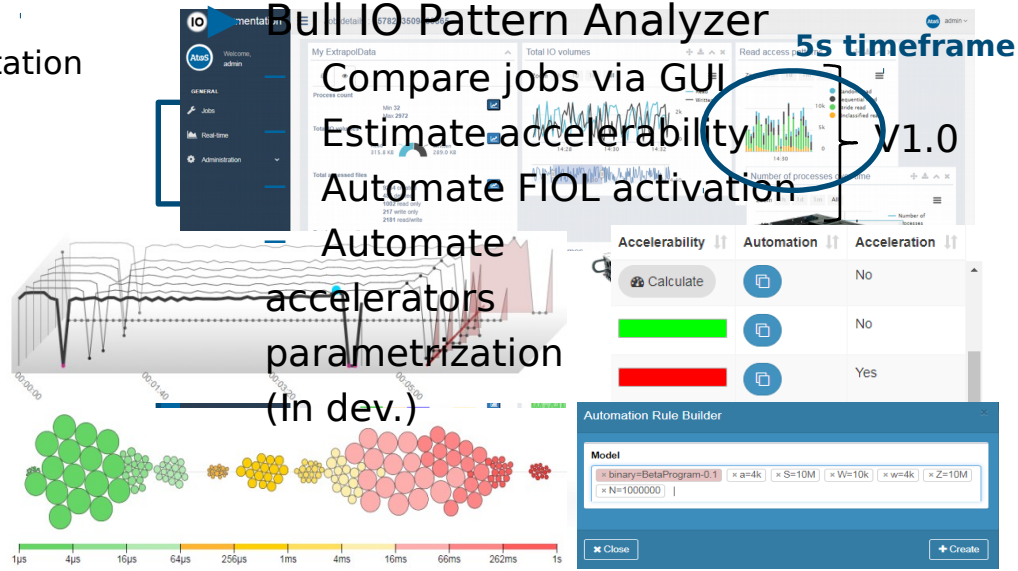
Bull
atos technologies

BDS - HPC R&D – Data Management

Product Overview



- Bull IO Instrumentation V1.1
- Bull Fast IO Libraries V1.0
- Bull Smart Burst Buffer (End 2018)
- Bull IO Pattern Analyzer
  - Compare jobs via GUI
  - Estimate accelerability
  - Automate FIOL activation
  - Automate accelerators parametrization (In dev.)



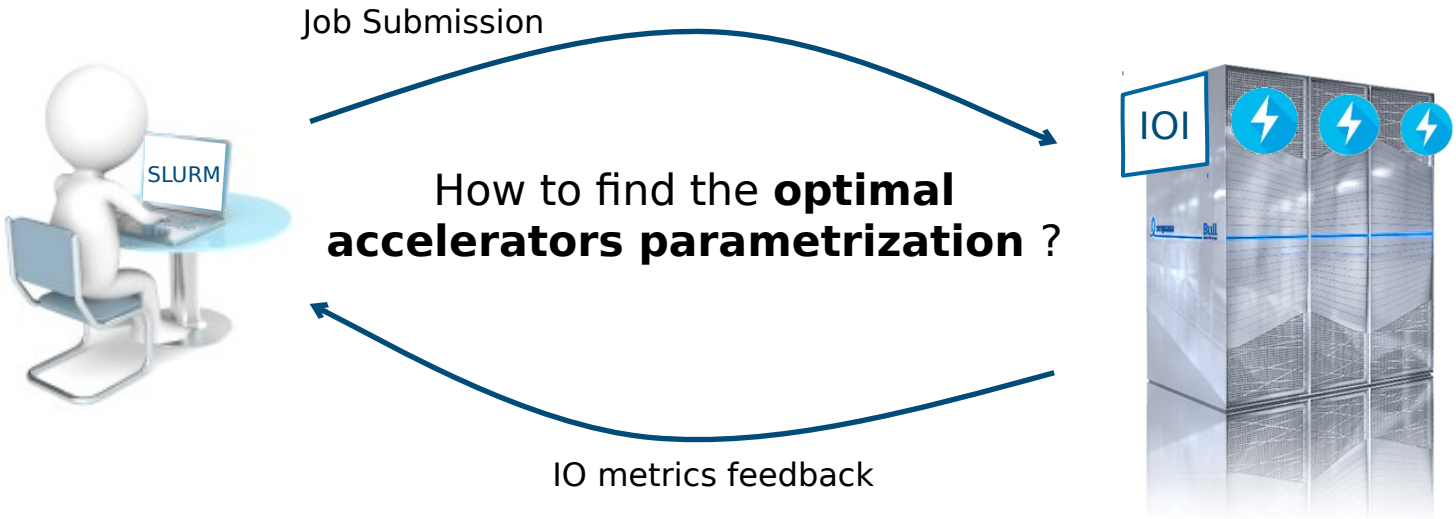
Bull IO Pattern Analyzer

Automate accelerators parametrization

Bull IO Pattern Analyzer

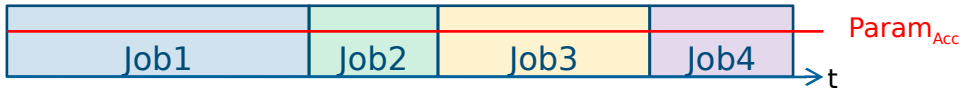
Targeted User Feature

- Run a job with the minimal execution time

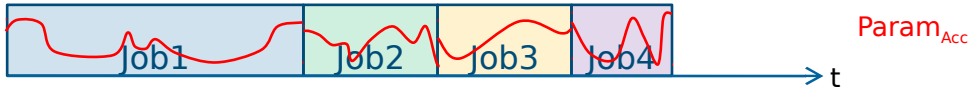


Parametrization approaches

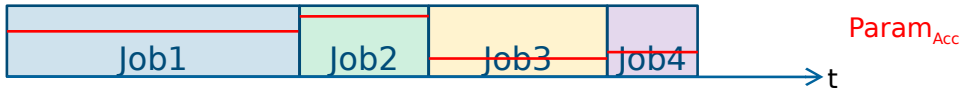
- Static parametrization
  - ⇒ mean performance for a wide range of applications : never optimal



- Dynamic parametrization
  - ⇒ online analysis of application behavior : too costly (high overhead)



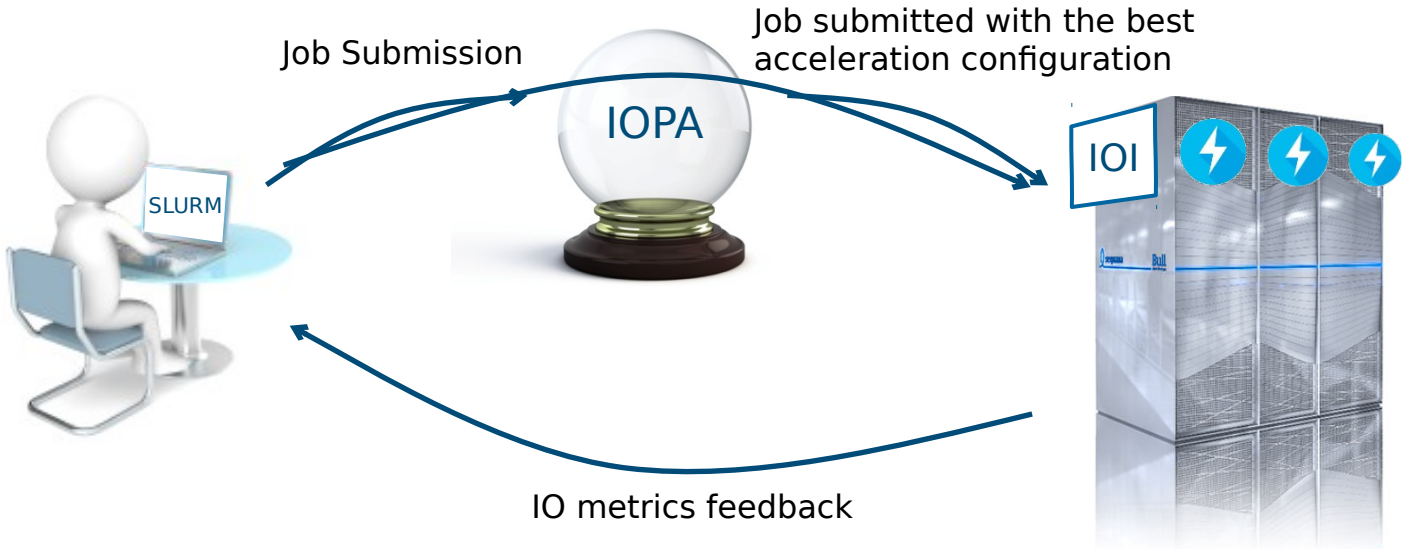
- Self-optimized parametrization
  - ⇒ find the optimal static parametrization for the job : trade-off static/dynamic



Bull IO Pattern Analyzer (IOPA)

Targeted User Feature

- Run a job with the minimal execution time
  - ⇒ **Automatically** find the optimal parametrization of the IO accelerators

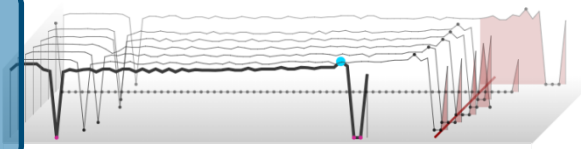


Bull
atos technologies

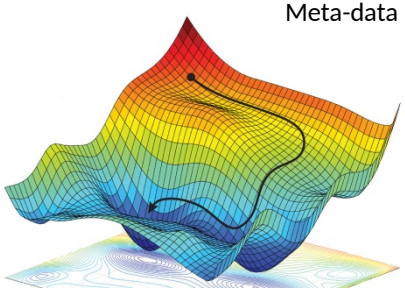
Self-optimized parametrization method

IO Instrumentation
IO Accelerators

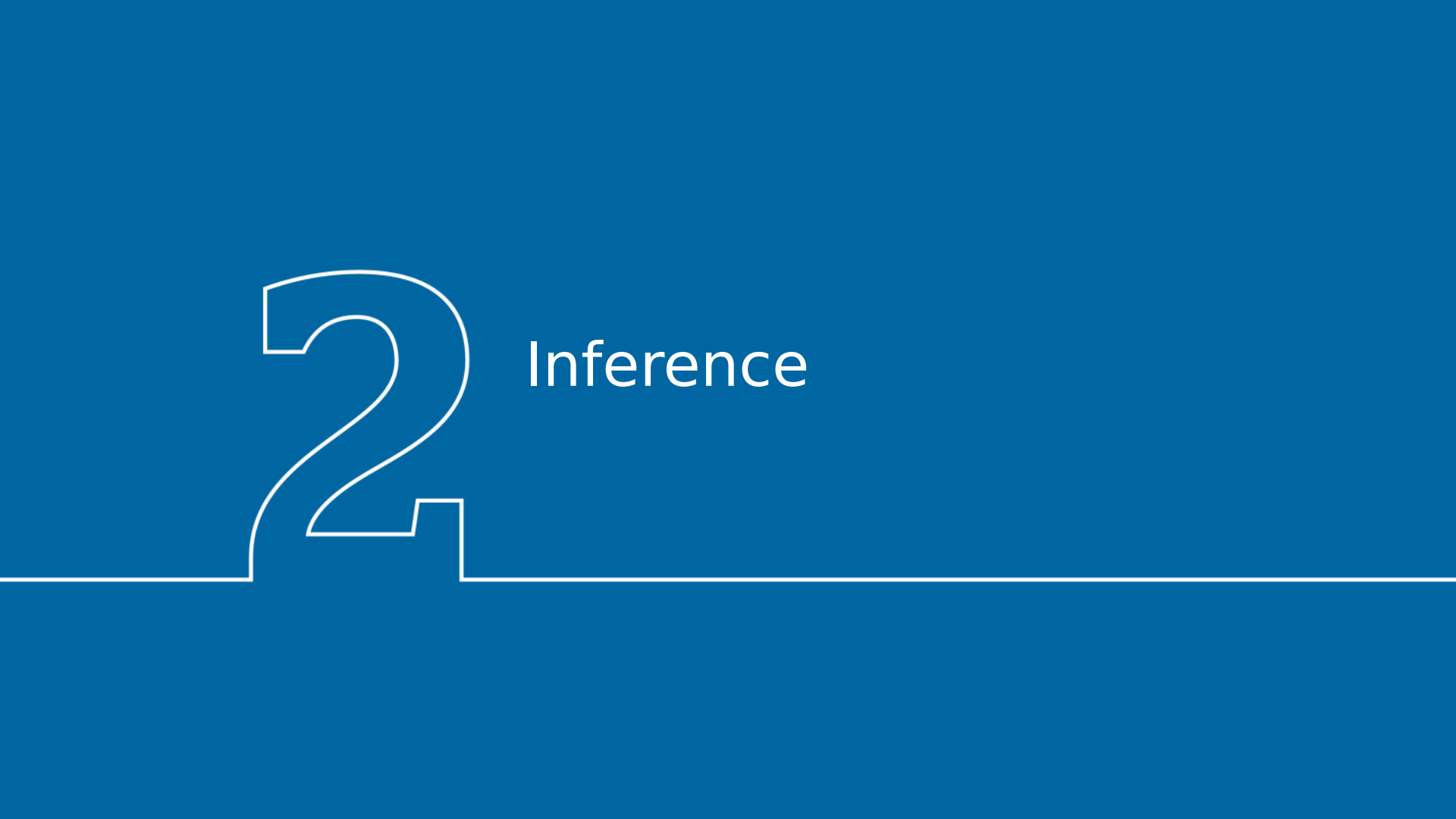
Cluster
IO metrics collection
Runs Database
IO metrics
Parameters
Perf Value



Job
Discriminator
Topology
Parameters
Meta-data
Free Meta-data
(To be optimized)
All runs
Fixed Meta-data
(User defined)
Family
Filter
Sub-set of family related runs
IO Pattern Analyzer
Proposed
Meta-data
Inference
Relevant runs
- Parameters
- Perf Value
Clustering
(Outlier
detection)
Threshold

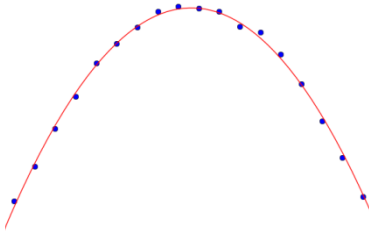


Bull
atos technologies

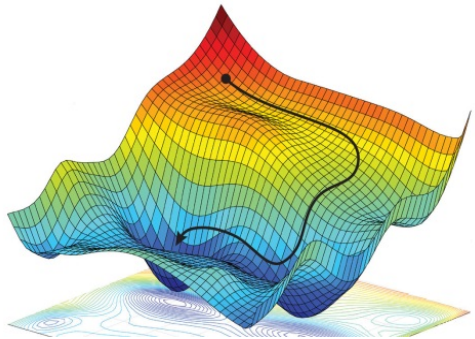

2
Inference

Inference
The « equation »

Inference (of optimal Parameters) = Modeling the objective function + Numerical optimization



- Theoretical ✘
- Interpolation ✘
- Regression ✔



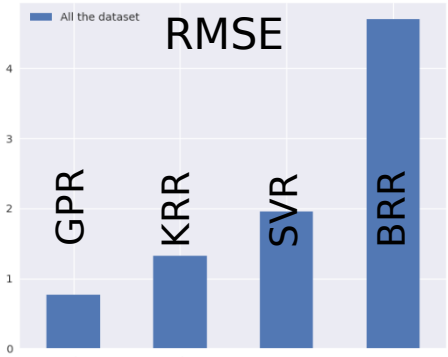
Inference - Regression

From principle to application

- Find a model which estimates the relationship :
  - ⇒ **accelerators parameters vs performance**

$$f(\text{param}) = \text{perf}$$

- Different methods studied
  - Bayesian Ridge Regression (BRR)
  - Kernel Ridge Regression (KRR)
  - Support Vector Machines for Regression (SVR)
  - Gaussian Process Regression (GPR)



Methods	Prediction time*	Train time*
SVR	0,001 s	0,110 s
KRR	0,001 s	0,120 s
GPR	**0,001 s**	**0,069 s**

* Time measured to perform regression/prediction on 168 runs

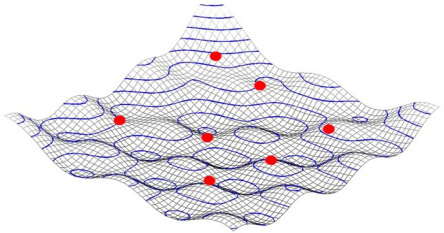
Inference - Optimization

Gradient-free optimization methods

- Find the optimal solutions of the IO accelerator parameters (min execution time)
  - ⇒ Use heuristics to find « good » solutions in a reasonable time
- Gradient-free method
  - ⇒ Less sensitive to local minimum locking

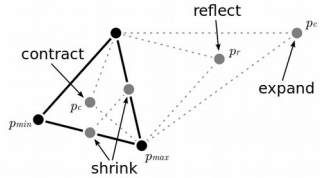
Particle Swarm Optimization (PSO)

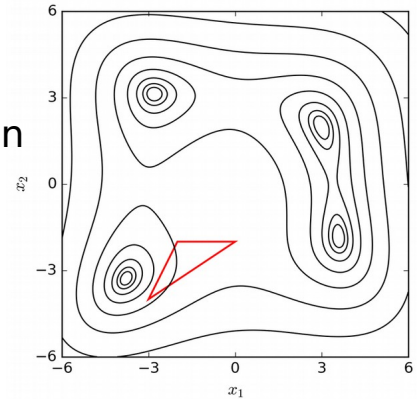
- Use collective behavior to model the problem
- They are described by a position and a velocity



Nelder-Mead (NM)

- A simplex inspired method
- Based on four main transformations

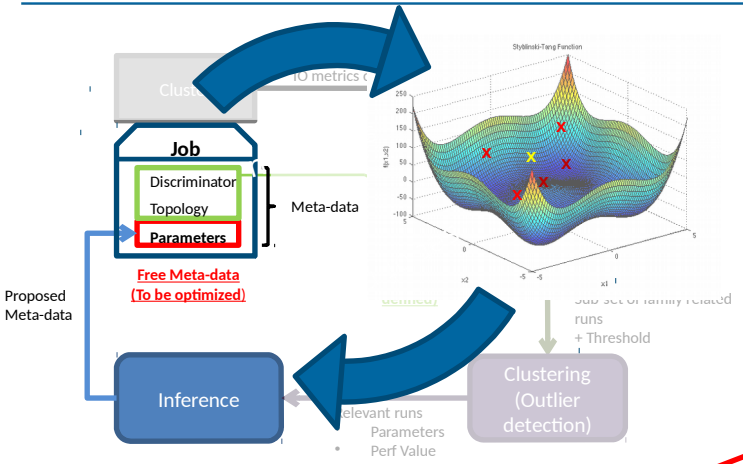




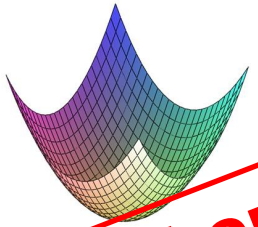
Covariance Matrix Adaptation Evolution Strategy (CMA-ES)

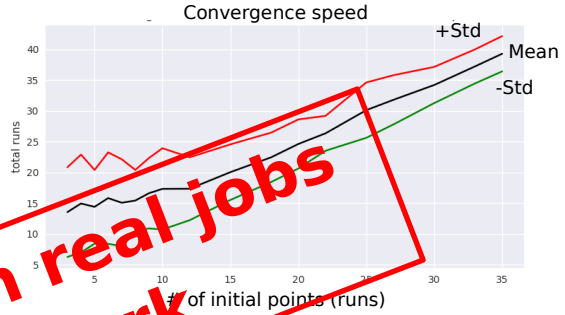
Inference

Convergence validation : simulations



Quadratic

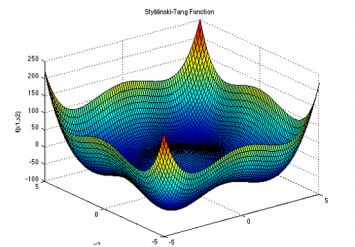


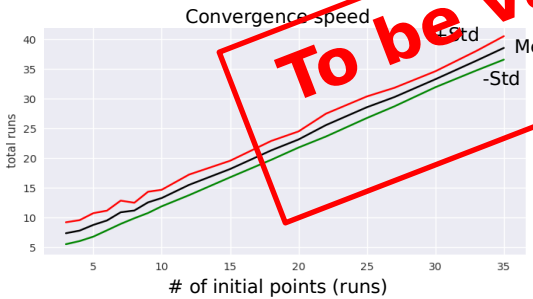


To be validated on real jobs

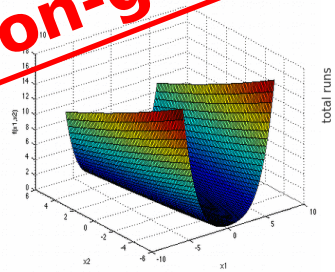
on-going work

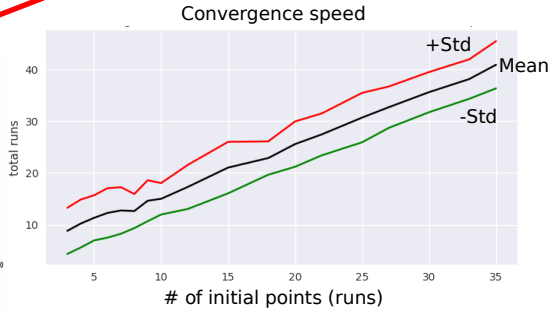
Stiblinsky





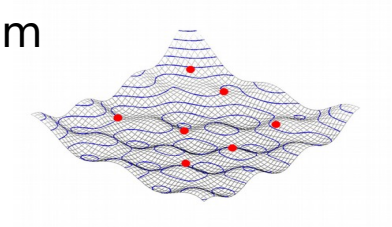
Rosenbrock

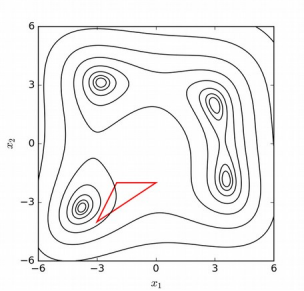




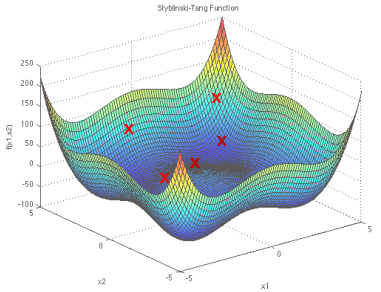
Inference
Perspectives

- Chose the most relevant optimization algorithm





- Setup a parametrization strategy for initialization



- Implement optimization and regression features into our IOPA product

Thank you

Bull
atos technologies