Trusted Friend Computing: Securely share OMICS data

B. Wolf, P. Kuonen, J. Stoppani, *University of Applied Sciences Western Switzerland*
D. Atlan, *Phenosystems, Belgium*

Introduction

Sharing OMICS data is becoming increasingly essential when analysing patient data. Yet many legal and ethical constraints slow down or make it impossible to share the data. This leads to a situation where many laboratories store their valuable patient data only in internal databases, unable to share them even with others in the same consortium. In an effort to solve this problem, we developed TFC (Trusted Friend Computing), an extension to the POP-Java programming language, enabling developers to create applications that can share data and computing power in a secure way. We integrated this technology into GensearchNGS, an existing NGS data analysis software developed by Phenosystems SA, to enable users to easily and securely share variant data.

POP-Java

We extended POP-Java, a Java language extension for distributed computing, by introducing the concept of Trusted Friend Computing (TFC), which allows to programmer to create secure networks. Inside those networks objects can be distributed, with all connections being secured by SSL/TLS 1.2. To join a network, the public keys of one member and the new member need to be exchanged. The base concept is that on every node the same code (application) is running and thus no code is transiting over the network. The programmer can bring the calculations to where the data is and thus remove the security risks of transfering patient data between laboratories. If enabled, statistics of the resource usage can be gathered and used for billing.

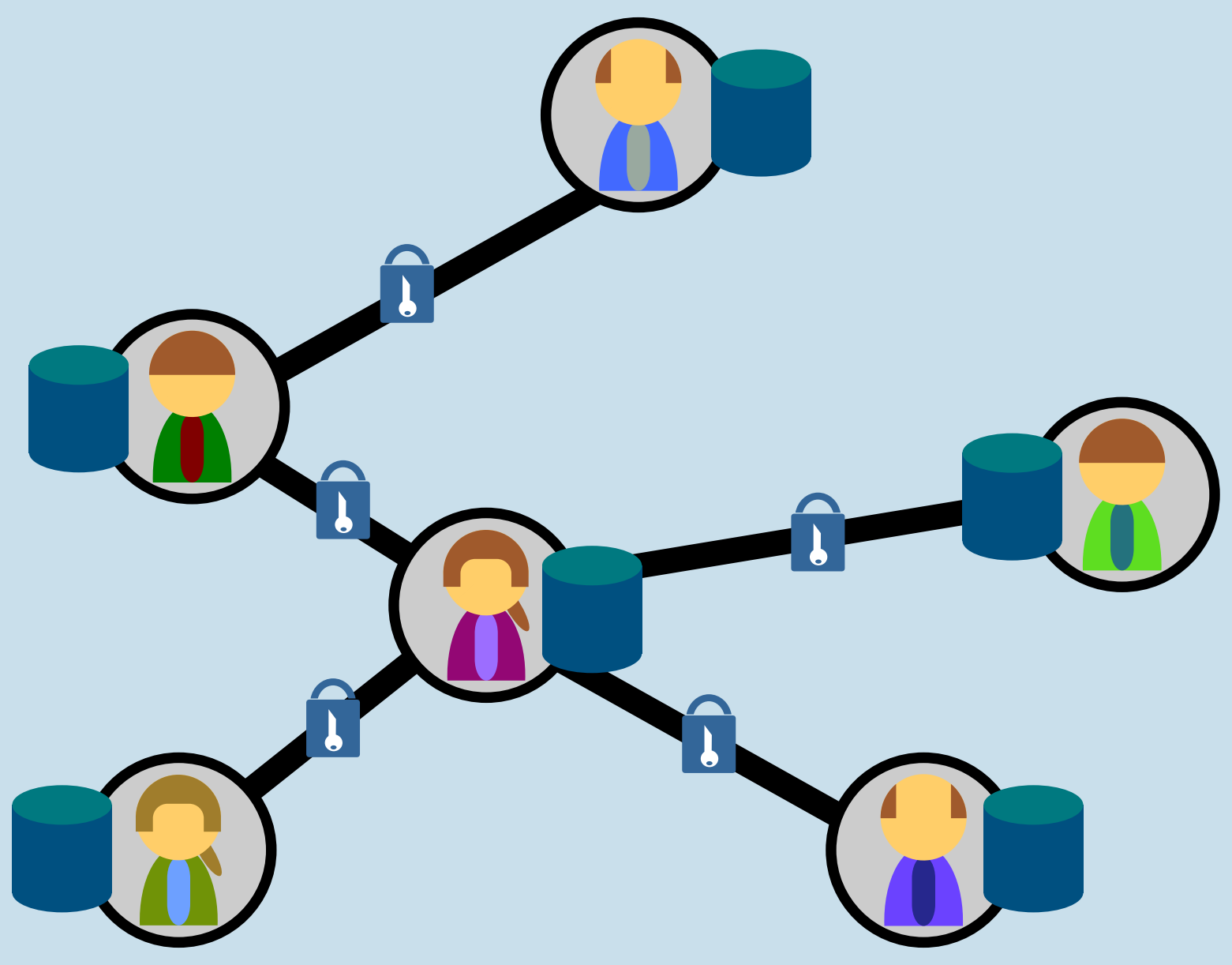


Figure 1: Example layout of a TFC network using secured conections

The developed technology is independent of bio-informatics, but fits the requirements of many projects in the field.
POP-Java with TFC can be integrated in any existing Java project through a maven package.
POP-Java is released under the LGPL licence.

GensearchNGS

We implemented several features in GensearchNGS that make use of the possibilities provided by POP-Java TFC. One of those features is the ability to share variants throughout multiple laboratories. No information about the patients or the variants themselves leave the respective laboratories, but if desired, a laboratory can allow its friend laboratories to ask if it has seen a certain variant before. Figure 1 and 2 show the implementation of this feature.

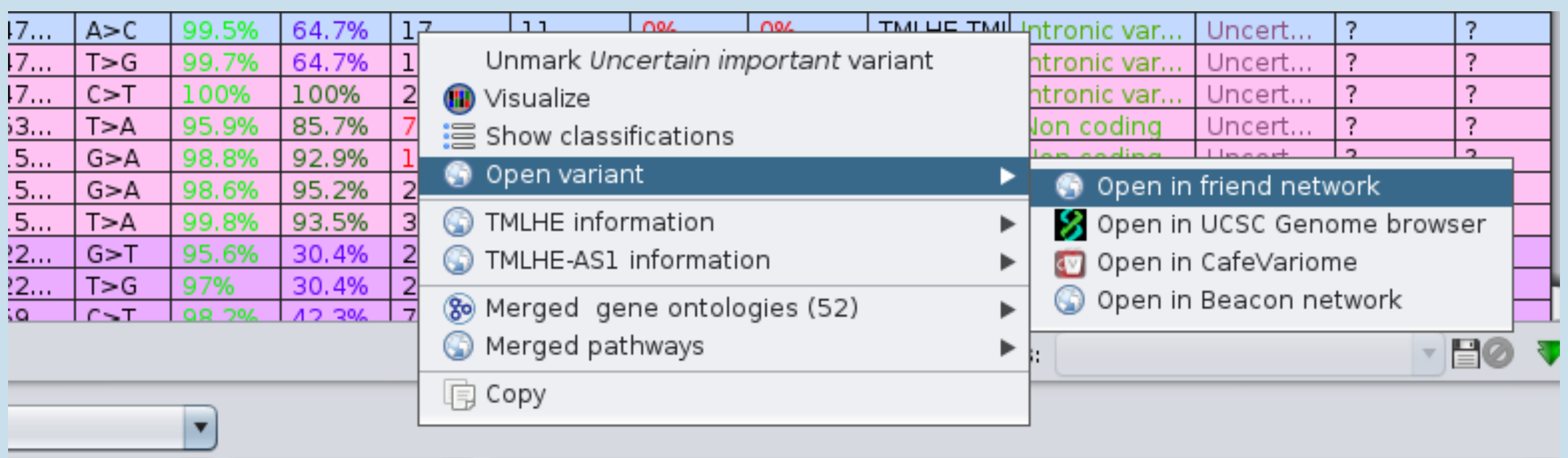
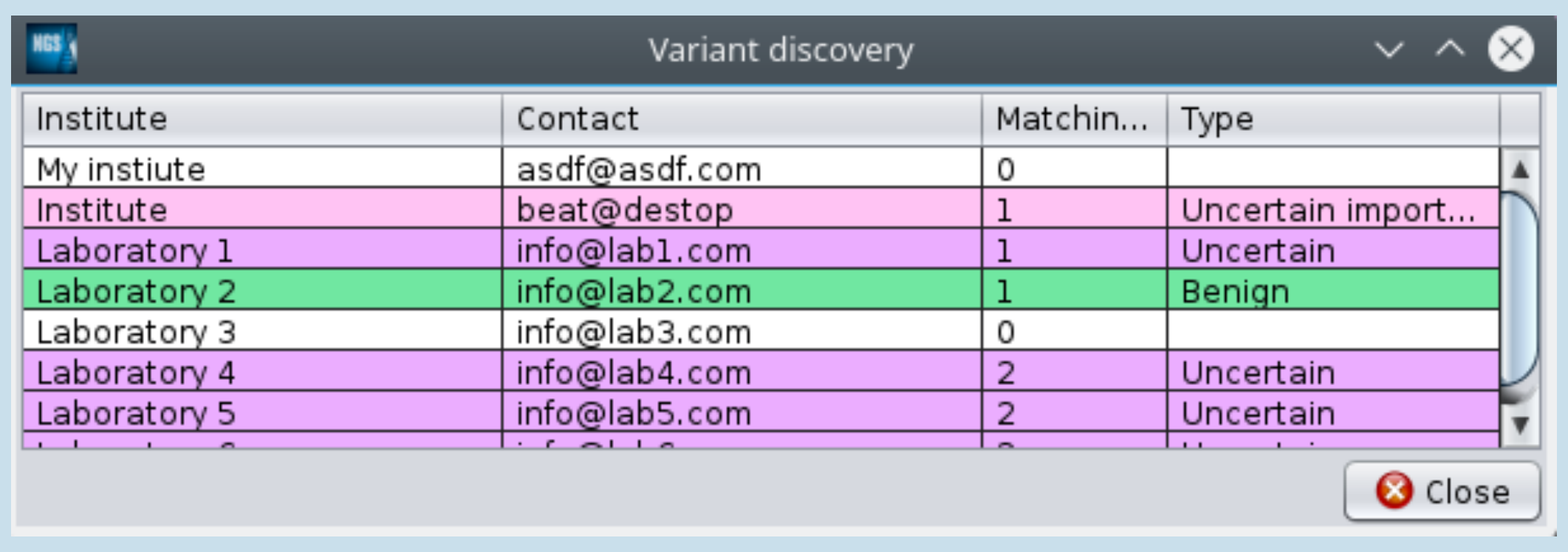


Figure 2: Transparent integration of the prototype into the variant list



Variant discovery

Institute	Contact	Matchin...	Type
My instiute	asdf@asdf.com	0	
Institute	beat@destop	1	Uncertain import...
Laboratory 1	info@lab1.com	1	Uncertain
Laboratory 2	info@lab2.com	1	Benign
Laboratory 3	info@lab3.com	0	
Laboratory 4	info@lab4.com	2	Uncertain
Laboratory 5	info@lab5.com	2	Uncertain

Figure 3: List of friendly laboratories and their classification of a specific variant

We also adapted the GensearchNGS aligner to be able to use a TFC network to distribute the alignment process over multiple computers. Especially smaller laboratories can profit from this feature, as it allows to speed up the analysis without requiring new infrastructure.

Conclusion

We propose a new framework to develop secure distributed systems, like those used in bio-informatics, called POP-Java. This open-source Java library allows programmers to easily distribute objects over a network of trusted computers, allowing them to use their resources. We applied this technology to GensearchNGS to more easily share OMICS data. Those features include:

-Distributed variant database
-Distributed alignment

The underlying technologies, POP-Java and TFC, are freely available, with the commercial NGS data analysis software GensearchNGS implementing those features using that technology.

Grants

CTI no. 18781.1 PFES-ES

Contact: Beat Wolf, beat.wolf@hefr.ch

Website POP-Java (TFC):
https://github.com/pop-team/pop-java
Website GensearchNGS:
http://www.phenosystems.com/

PhenoSystems
GensearchNGS
MOLECULAR GENETICS MADE EASY

Hes·so FRIBOURG FREIBURG
Haute Ecole Spécialisée de Suisse occidentale
Fachhochschule Westschweiz
University of Applied Sciences Western Switzerland

iCoSys
Institute of Complex Systems