

Specialized File Transfer Service for Large Oil&Gas Datasets

Fabio de Souza^{1,2}, Marcus Salles³, Fabio Campos², Silvio Costa², Myrian Costa¹, Nelson Ebecken¹

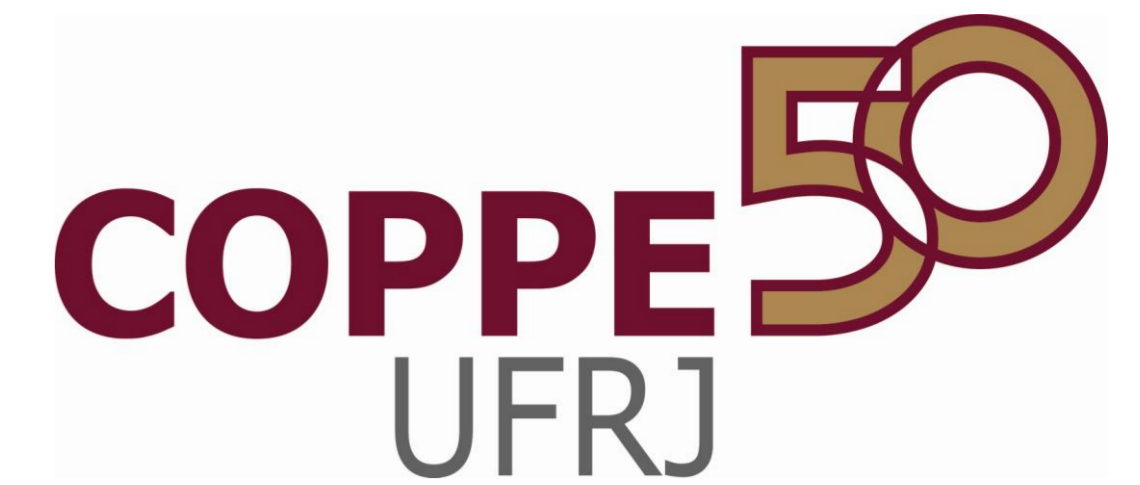
¹COPPE, Federal University of Rio de Janeiro, Brazil

²PETROBRAS – Petróleo Brasileiro S.A., Brazil

³BNDES – Brazilian Development Bank, Brazil



PETROBRAS



Introduction

- SETA is a web application using GridFTP to transfer Oil&Gas Datasets with high performance and security.
- SETA was driven by the challenge of transferring huge data volumes between Petrobras and Brazilian Regulatory Agency of Petroleum (ANP).
- Big challenge in Oil&Gas to handle large Datasets (400GB – TB files).
- Based on previous studies in our environment[1], we built a in house tool with GridFTP.
- Inspired by Globus Online, SETA was developed using the Globus Java API – jGlobus[2].
- Security policies prevent from using Globus Online[3] between Petrobras and ANP.

Features

- 1) File integrity:** comparison between source and destination on a block-by-block basis, in case of mismatch, only the differing blocks will be retransmitted, avoiding a full data retransmission in case of corrupted file transfer, saving time and bandwidth.
- 2) Transfer Resume:** in case of connection interruption or any other failure.
- 3) Graphic Monitoring and Reports:** users can keep up with running and finished transfers.
- 4) Authentication:** this service queries an existing corporative authentication to logon.
- 5) Scheduled Transfer:** able to schedule the transfer to an appropriate time, avoiding link bottlenecks and daytime congestion.
- 6) TCP Parallel Data Connection:** taking advantage of links above 1Gbit/s. Users can choose among 1 (no parallelism), 2, 4 or 8 parallels connections, as an advanced option.

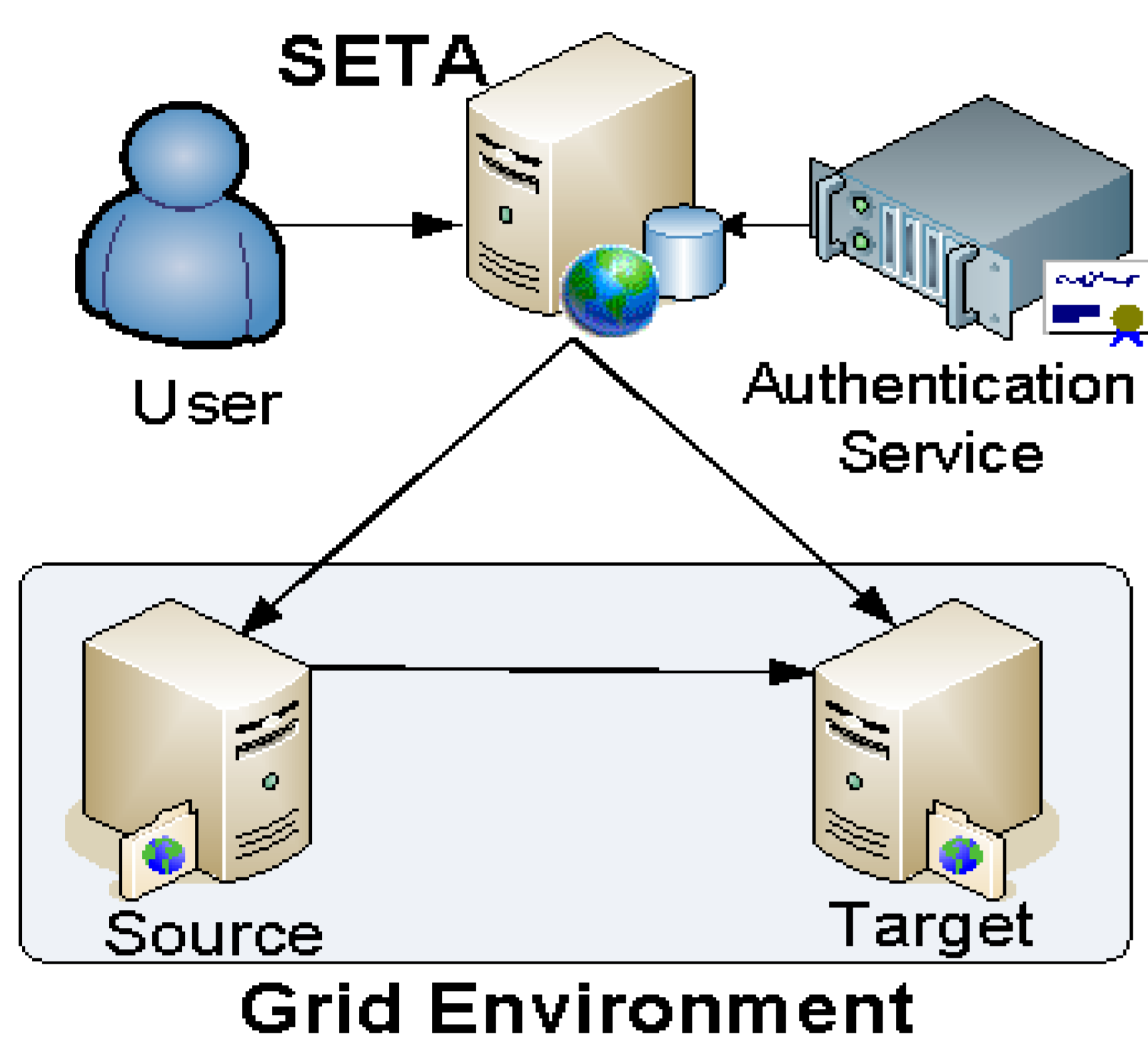


Figure 1 - SETA Design

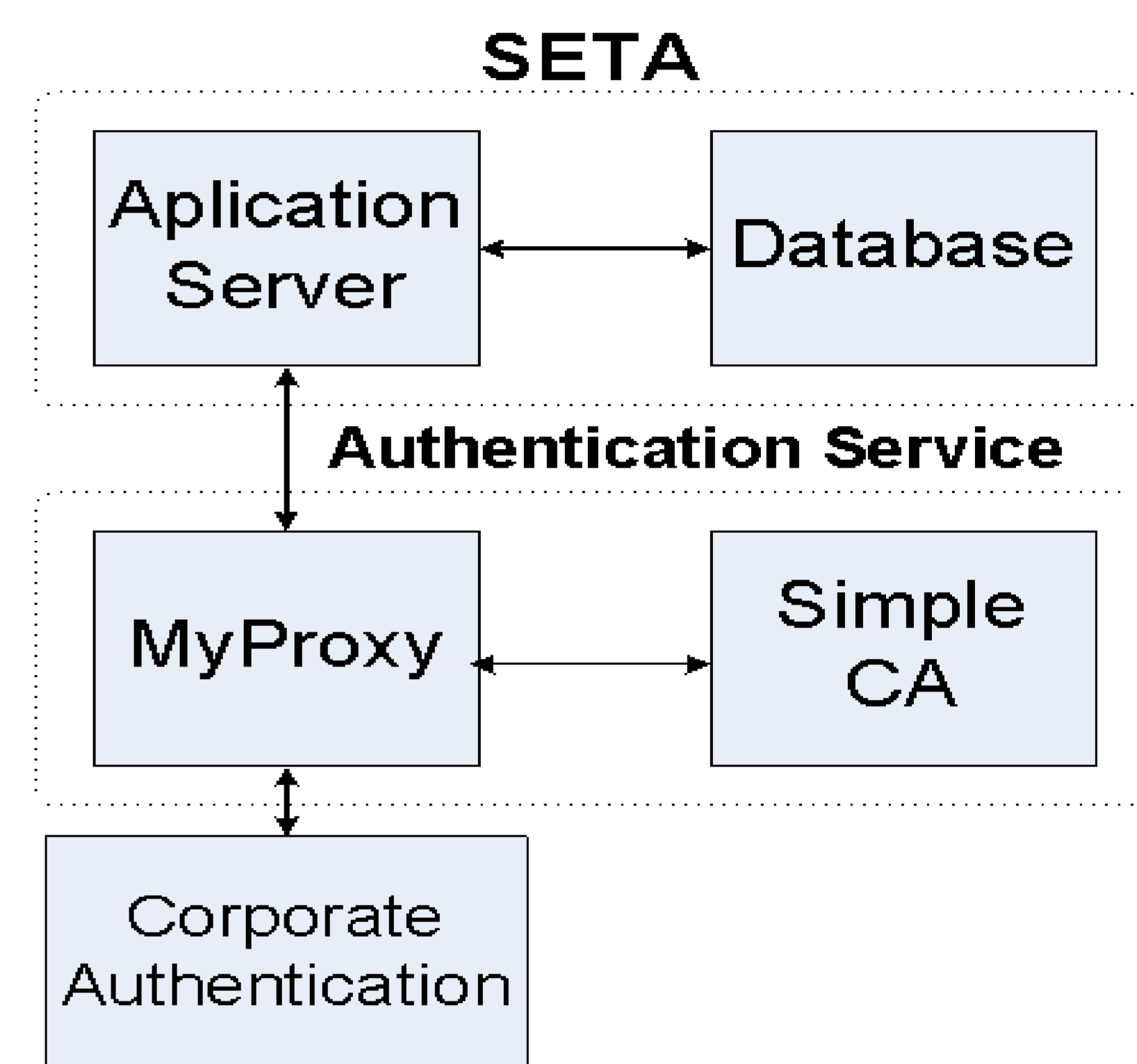


Figure 2 - SETA Infrastructure

User Interface

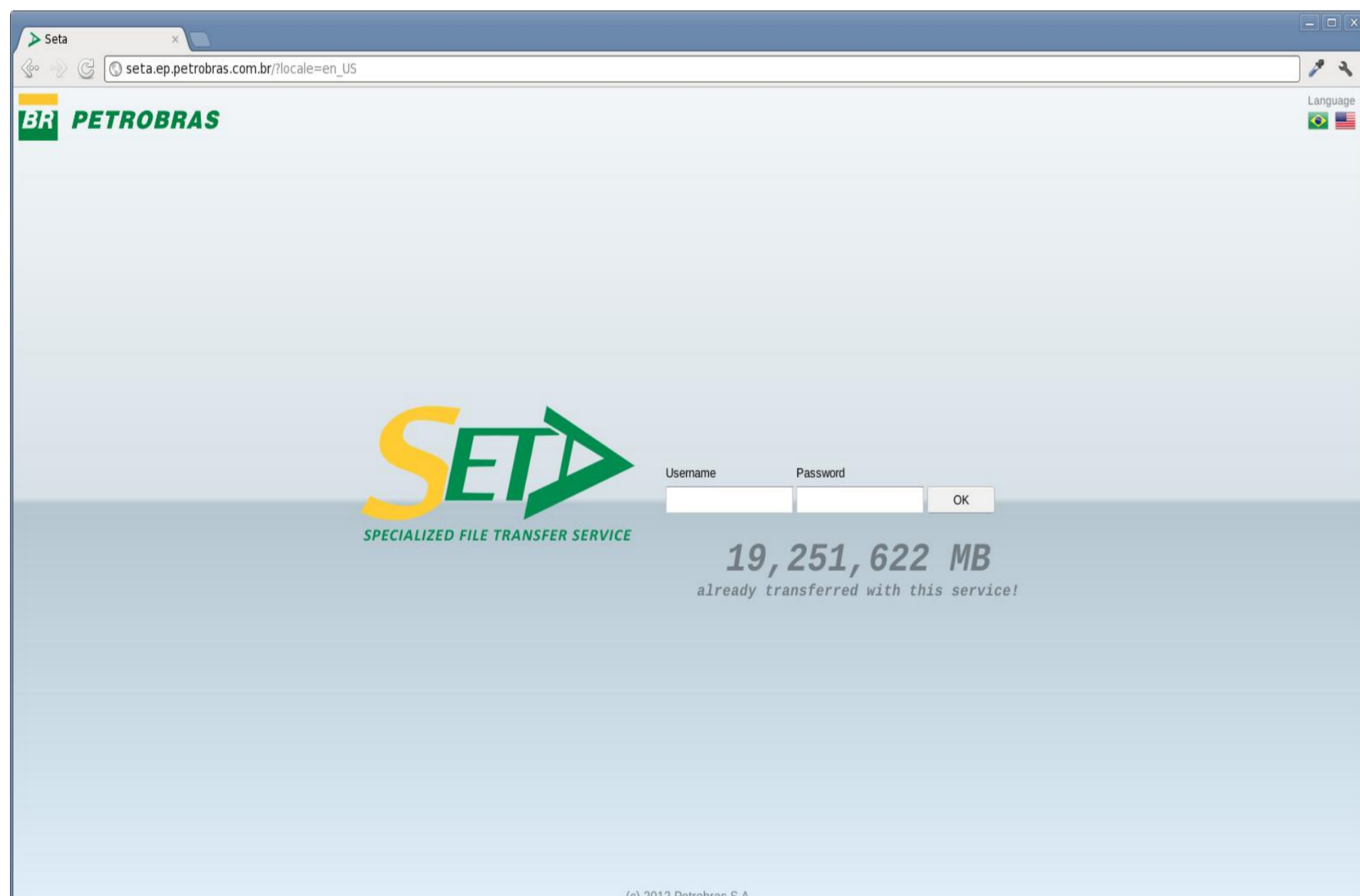


Figure 3 – Login

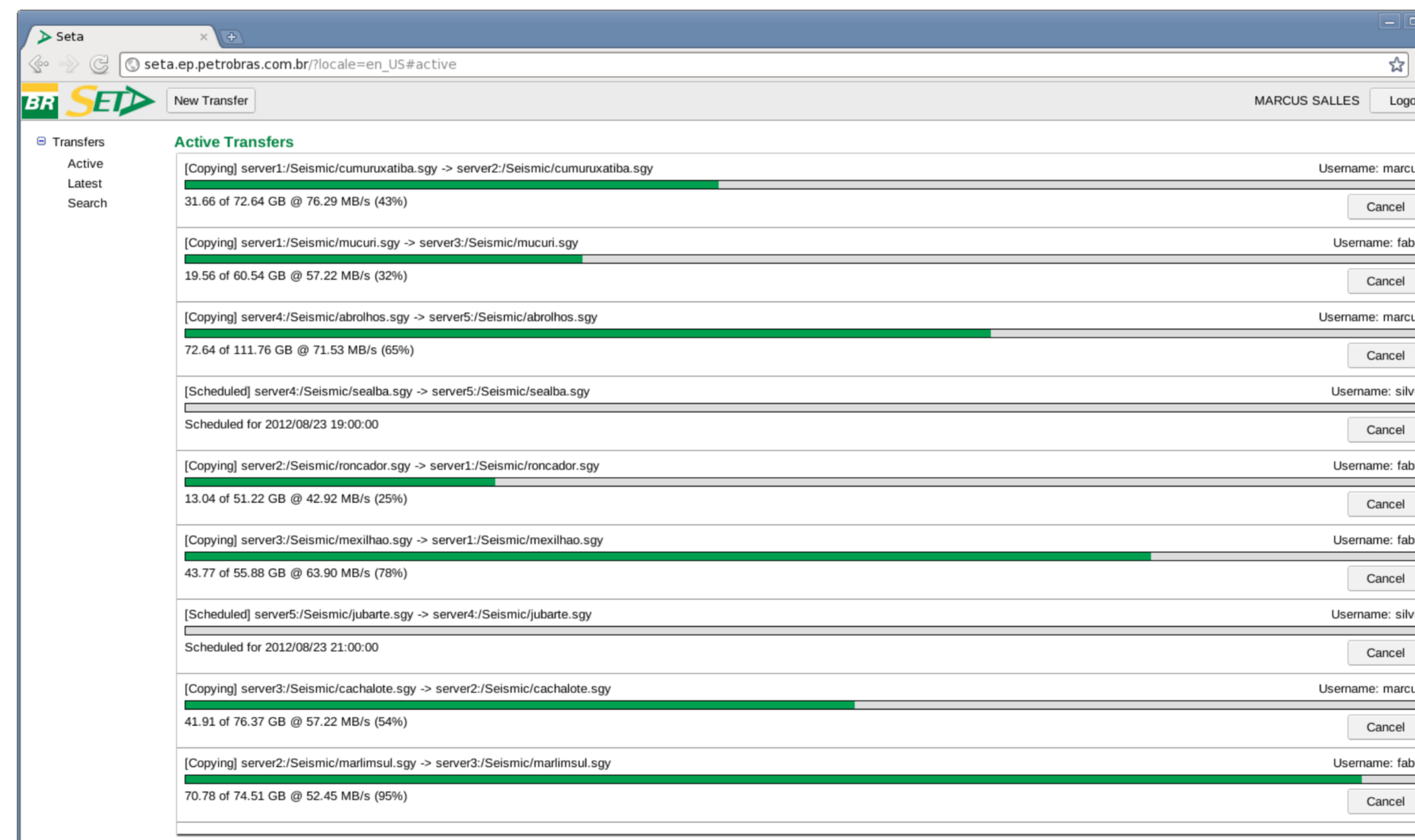


Figure 4 – Scheduled and Active Transfers

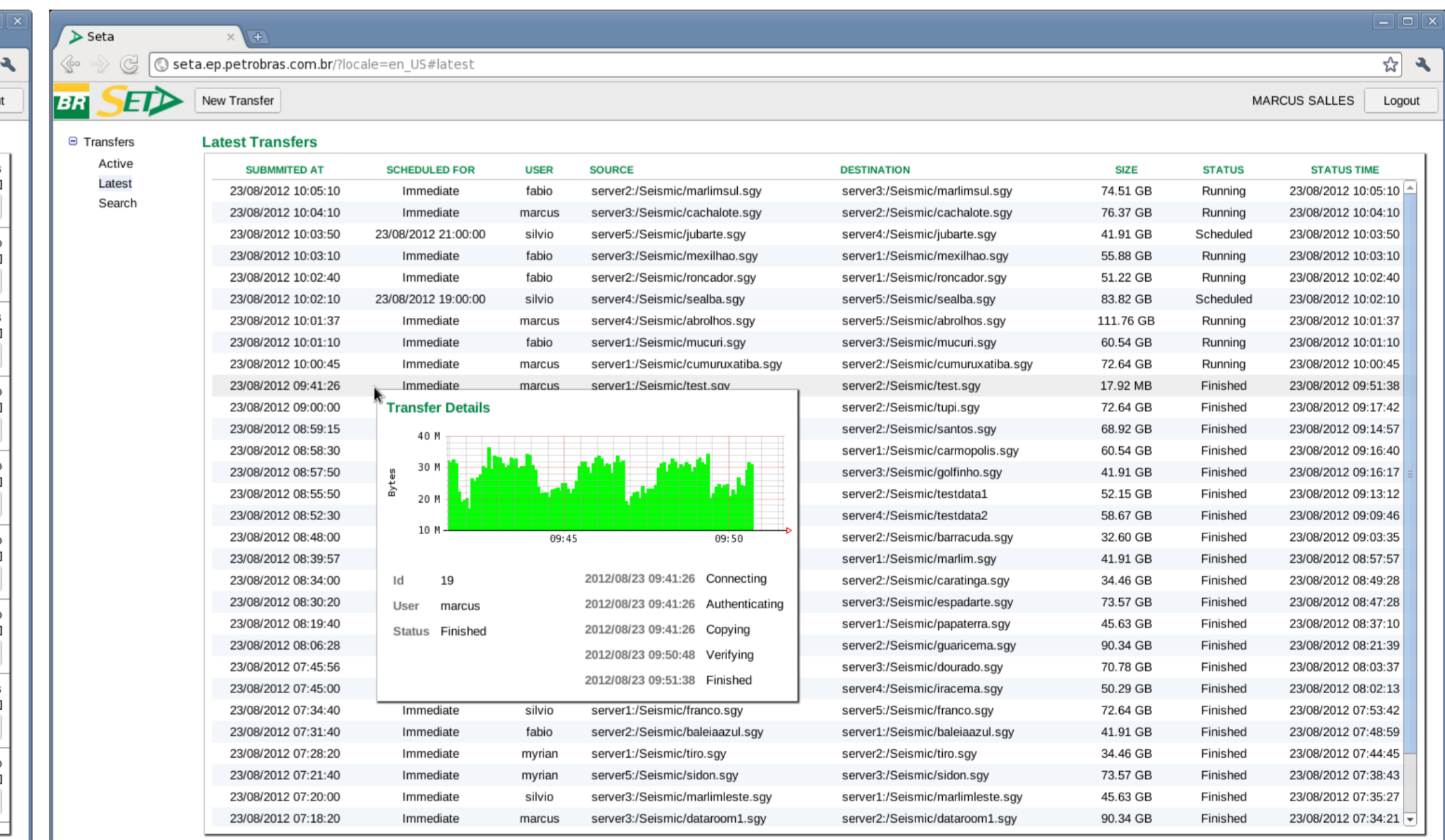


Figure 5 – Latest Transfers

Results

1Gbit/s Link

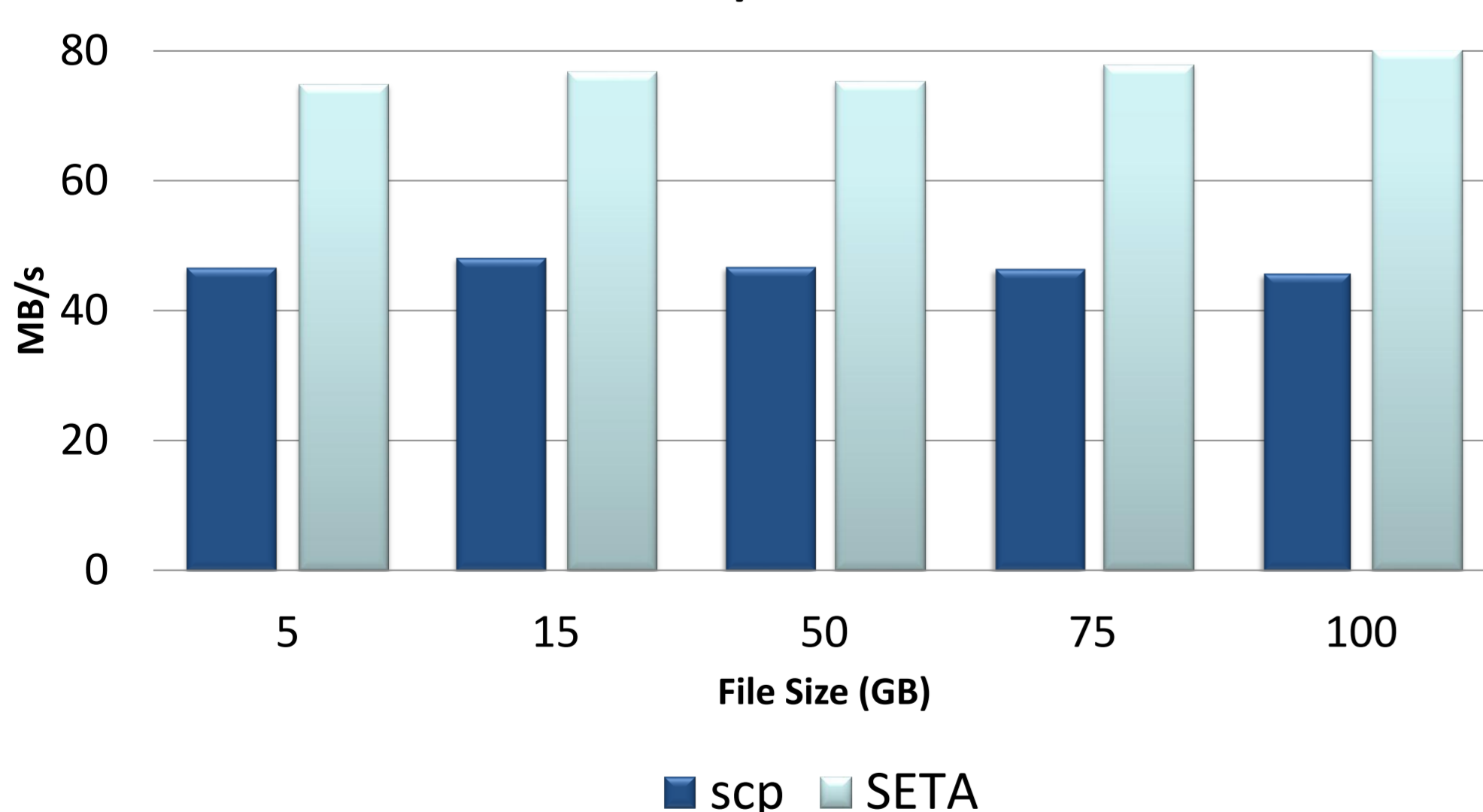


Figure 6 – Case study with 1 Gbit/s link

- Tests with 1Gbit/s links evaluated the performance with Parallel Data Connection.
- Tests with 4Mbit/s links evaluated the integrity, reliability and transmission resume.

4Mbit/s Link

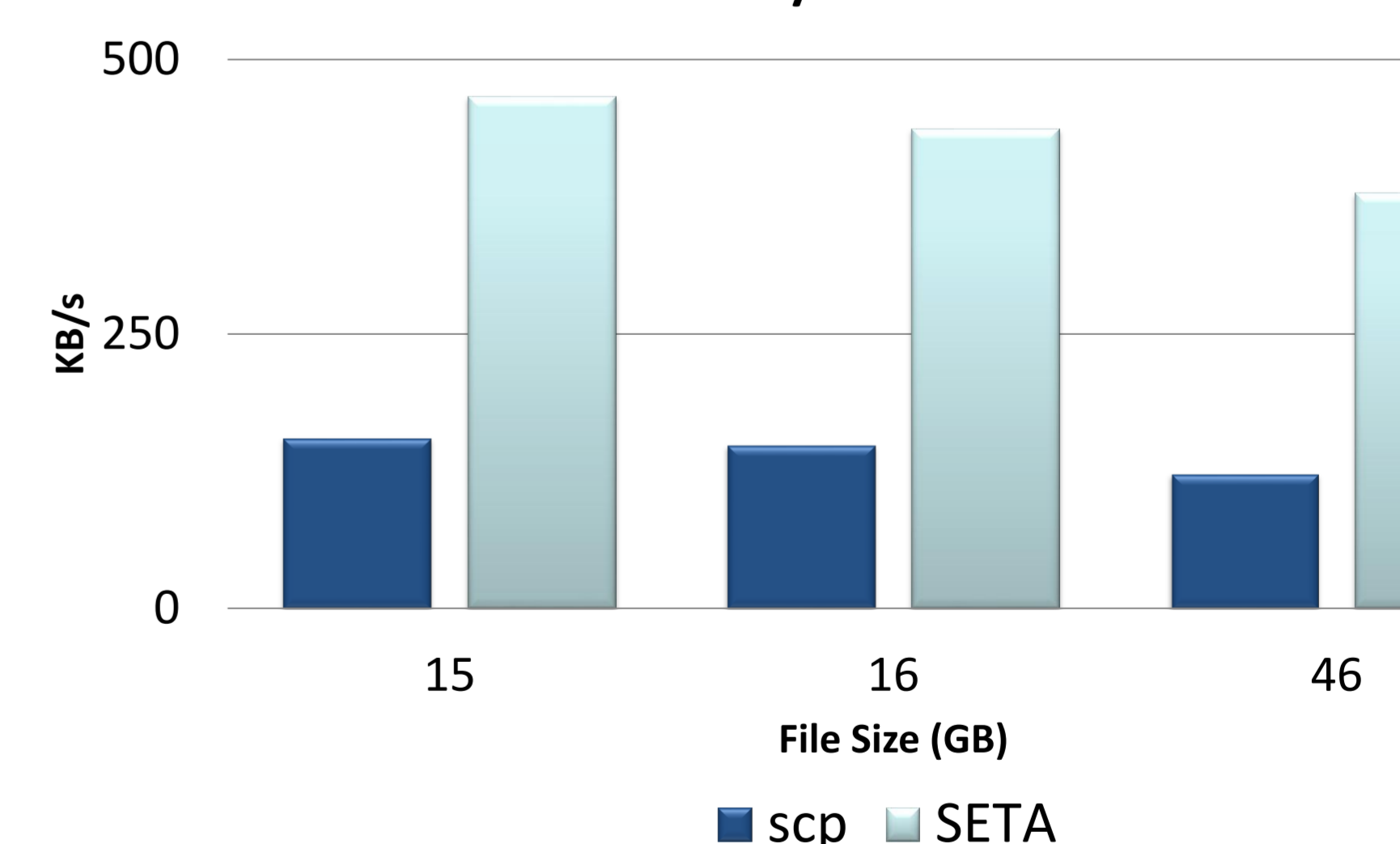


Figure 7 – Case study with 4Mbit/s link

Conclusions

- SETA shows many advantages against current solutions employed by Petrobras.
- SETA ensures fast and safe transmission of large amounts of Oil&Gas datasets.
- Provides a user friendly interface through the web.
- Reduce costs and mitigate the risks of data loss.

Future Work

- Striped data transfers (Striped Mode[4]) with a parallel file systems.
- User notification of finished transfers.
- Synchronization between two remote directories so as to provide mirroring.

References

1. Costa, S. Large Data Volume Transfers on Geographically Distributed Environments, MSc Dissertation – COPPE/UFRJ, 2009 (in Portuguese).
2. jGlobus, <http://www.globus.org/toolkit/jglobus>
3. Allen, B., Bresnahan, J., Childers, L., Foster, I., Kandaswamy, G., Kettimuthu, R., Kordas, J., Link, M., Martin, S., Pickett, K. and Tuecke, S. Globus Online: Radical Simplification of Data Movement via SaaS. Preprint CI-PP-05-0611, Computation Institute, 2011.
4. Allcock, B., Bresnahan, J., Kettimuthu, R., Link, M., Dumitrescu, C., Raicu, I. and Foster, I. The Globus Striped GridFTP Framework and Server. SC'2005, 2005.

Contacts

¹ COPPE, Federal University of Rio de Janeiro, P.O.Box 68516, 21945-970, Rio, RJ, Brazil fabioms@coc.ufrj.br, myrian@nacad.ufrj.br, nelson@ntt.ufrj.br

² PETROBRAS, Technical Support for E&P, Av. República do Chile 65 – Sala 1501, Centro, 20031-912, Rio, RJ, Brazil {fabiomsouza@petrobras.com.br, ffidelis.spassu@petrobras.com.br, scosta@petrobras.com.br}

³ BNDES – Brazilian Development Bank, Av. República do Chile 100 – Centro, 20139-900, Rio, RJ, Brazil marcus.salles@bndes.gov.br