# Requirements for Open Communication Standards for Efficient Data Exchange

## A Data-Driven Guide to Selecting the Right Standard for Efficient Data Exchange


Graduate

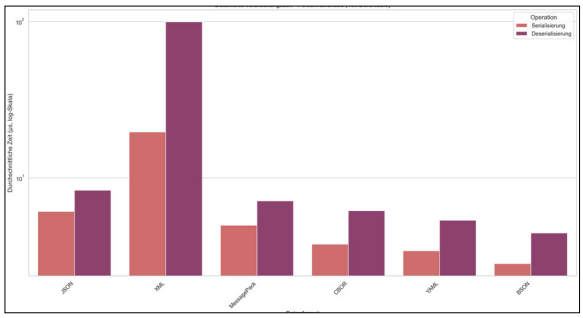
**Alessandro Shun Bühler**


Initial Situation: The digital transformation is leading to a flood of data in networked systems like the Internet of Things (IoT) and microservice architectures, making efficient data exchange a critical success factor. However, developers face the challenge of selecting the optimal choice from a vast array of open standards such as HTTP, MQTT, JSON, or Protobuf. As well-founded empirical data is often lacking, architectural decisions are frequently based on general assumptions. This thesis addresses this gap through a systematic, quantitative analysis of the fundamental trade-offs between network overhead, processing speed, data size, and robustness to create a data-driven foundation for practical application.

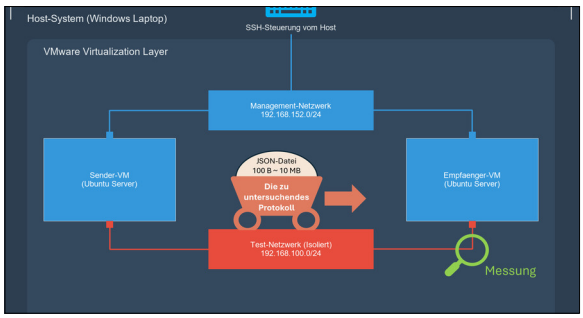
Approach / Technology: To enable a comprehensive evaluation, a dual-part research design was implemented in a fully isolated, virtual laboratory environment, separating the analysis of format properties from protocol performance on the network. Serialization Quality Test (SQ-Test): This protocol-independent test analyzed 13 data formats (including JSON, XML, Protobuf) directly in memory to assess their intrinsic performance characteristics: data size (compactness), processing speed (CPU time), and robustness in handling faulty data.
Systemic Protocol Efficiency Analysis (SPEA-Test): This test measured the end-to-end performance of 5 protocols (HTTP, MQTT, CoAP, TCP, UDP) in a completely isolated network between two virtual machines (VMs). To ensure a fair comparison, a uniform JSON payload was used in sizes ranging from 100 bytes to 10 MB. Systemic metrics such as network overhead, packet count ("chattiness"), and application latency (RTT) were recorded and their scaling behavior analyzed.

Result: The central finding is that the optimal technology choice is context-dependent and primarily dominated by message size. The results can be summarized in three "golden rules": Rule 1: For small messages ( < 10 KB), the protocol dominates. Choosing a lean protocol like MQTT instead of HTTP can reduce network overhead by over 70% and is the decisive efficiency factor. Rule 2: For large messages ( > 1 MB), the data format dominates. The protocol overhead becomes negligible ( < 1%). The compactness and processing speed of a binary format like Protobuf now determine the data throughput. Rule 3: Simplicity and readability are an expensive but often justified luxury. Text formats like JSON are measurably less efficient, but their robustness and vast ecosystem often justify the performance trade-off through faster development and easier debugging.

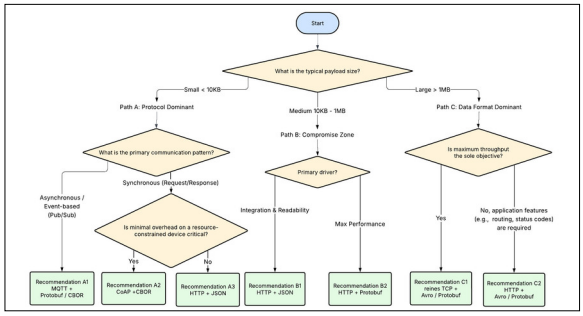
**Detailed processing time for the complex Java object "PersonAdvanced" (containing lists, maps, and nested objects)**
Own presentment




**Architecture of the virtualized test environment**
Own presentment




**Decision-making model for the systematic selection of a communication standard, based on the experimental results**
Own presentment




Advisor
**Prof. René Pawlitzek**

Co-Examiner
**Prof. Guido Piai**

Subject Area
**Computer Science**