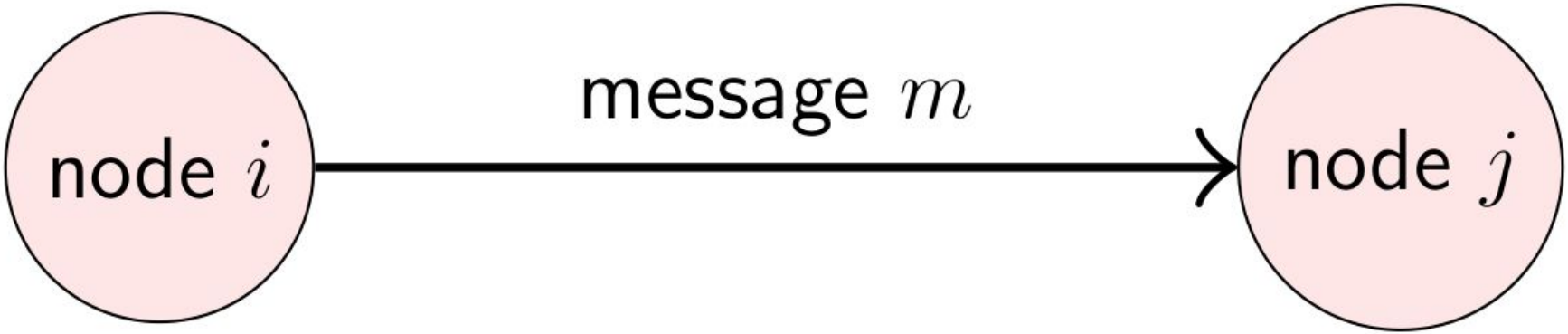
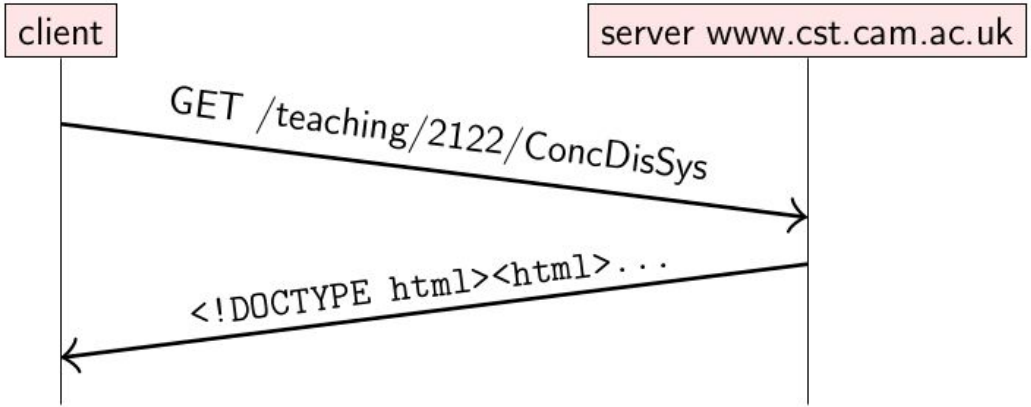
Data Engineering: Protocols

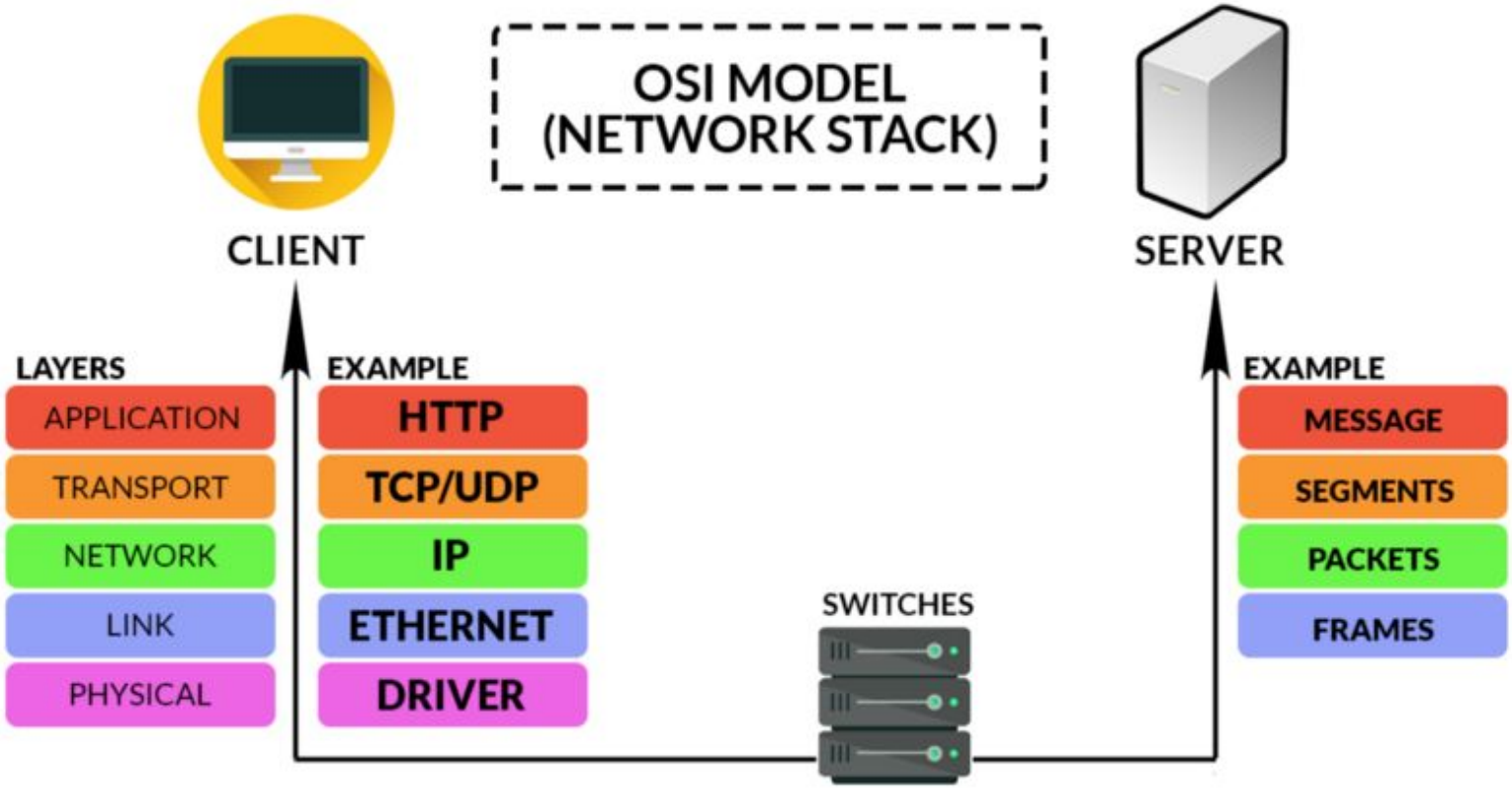
Communication


node i
message m
node j

Client-server example: the web

Time flows from top to bottom.

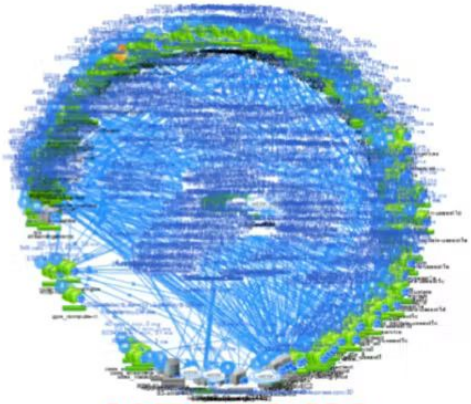



OSI MODEL
(NETWORK STACK)
CLIENT
SERVER
LAYERS
APPLICATION
TRANSPORT
NETWORK
LINK
PHYSICAL
EXAMPLE
HTTP
TCP/UDP
IP
ETHERNET
DRIVER
SWITCHES
EXAMPLE
MESSAGE
SEGMENTS
PACKETS
FRAMES

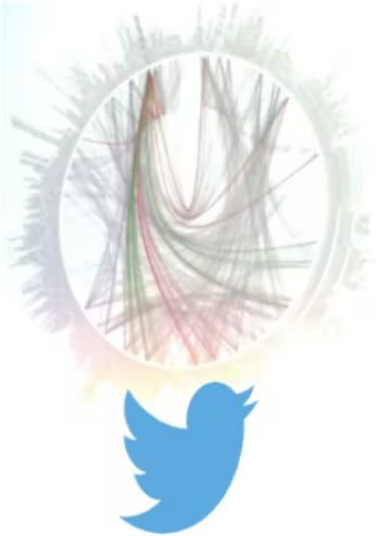
Source

450 microservices

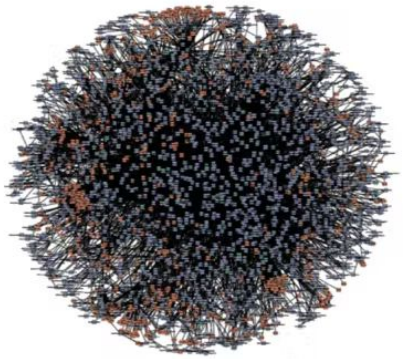
500+ microservices



500+ microservices



500+ microservices



Remote Procedure Calls (RPC)

RPC’s goal is to make distributed programming look like as much as possible like normal programming.

Client-server example: online payments

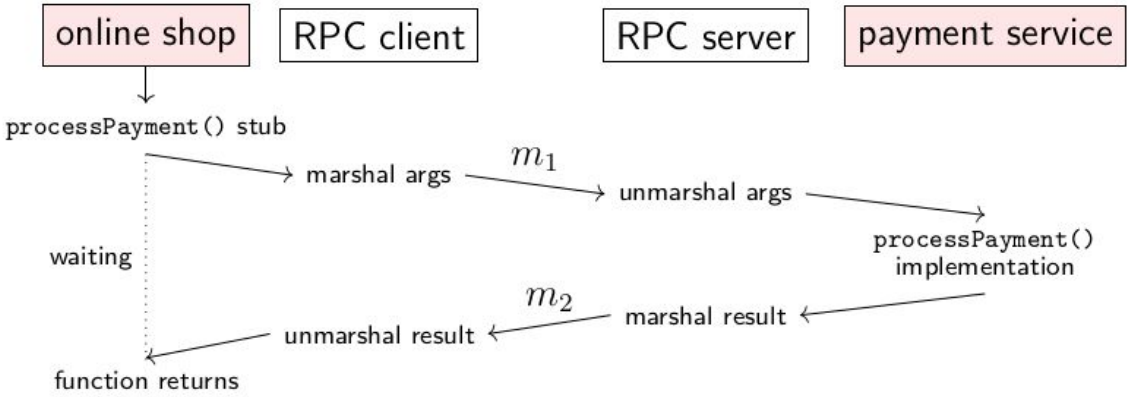
Remote Procedure Call (RPC) example

```
// Online shop handling customer's card details
Card card = new Card();
card.setCardNumber("1234 5678 8765 4321");
card.setExpiryDate("10/2024");
card.setCVC("123");

Result result = paymentsService.processPayment(card,
    3.99, Currency.GBP);

if (result.isSuccess()) {
    fulfilOrder();
}
```

Implementation of this function is on another node!



$m_1 =$

```
{
  "request": "processPayment",
  "card": {
    "number": "1234567887654321",
    "expiryDate": "10/2024",
    "CVC": "123"
  },
  "amount": 3.99,
  "currency": "GBP"
}
```

$m_2 =$

```
{
  "result": "success",
  "id": "XP61hHw2Rvo"
}
```

Ideally, RPC makes a call to a remote function look the same as a local function call.

RPC Failure

- Machine failures at only one end (caller/callee)
- Communication failures

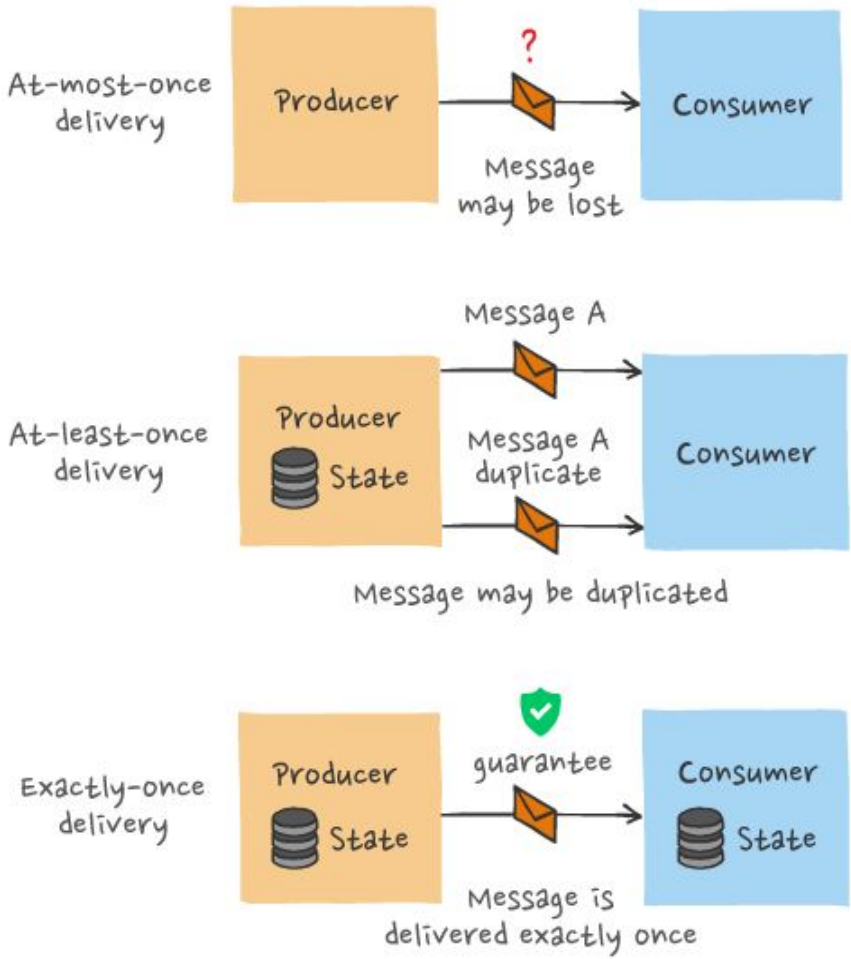
RPCs can return “failure” instead of results.

After Failure!

- Procedure did not execute
- Procedure executed once
- Procedure executed multiple times
- Procedure partially executed

Type of message semantics

- At most once semantic
- At least once semantic
- Exactly once semantic



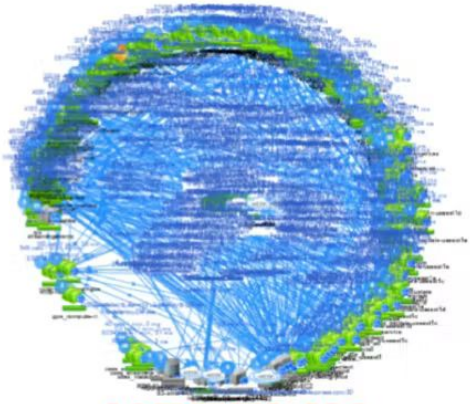
RPC history

- SunRPC/ONC RPC (1980s, basis for NFS)
- CORBA: object-oriented middleware, hot in the 1990s
- Microsoft's DCOM and Java RMI (similar to CORBA)
- SOAP/XML-RPC: RPC using XML and HTTP (1998)
- Thrift (Facebook, 2007)
- gRPC (Google, 2015)
- REST (often with JSON)
- Ajax in web browsers

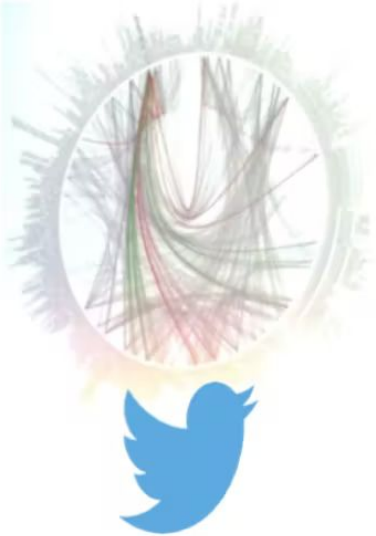
Google Remote Procedure Call (grpc)

450 microservices

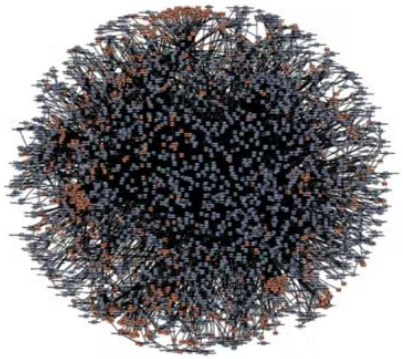
500+ microservices



500+ microservices



500+ microservices



Interface Definition Languages (IDL)

gRPC IDL example

```
message PaymentRequest {
  message Card {
    required string cardNumber  = 1;
    optional int32  expiryMonth = 2;
    optional int32  expiryYear  = 3;
    optional int32  CVC         = 4;
  }
  enum Currency { GBP = 1; USD = 2; }

  required Card     card     = 1;
  required int64    amount   = 2;
  required Currency currency = 3;
}

message PaymentStatus {
  required bool   success      = 1;
  optional string errorMessage = 2;
}

service PaymentService {
  rpc ProcessPayment(PaymentRequest) returns (PaymentStatus) {}
}
```

- Type Consistency
- Compatibility
- Performance

Json vs Protobuf

Reference

- http://www.scs.stanford.edu/20sp-cs244b/
- https://www.cst.cam.ac.uk/teaching/2122/ConcDisSys