

Boosting Software Efficiency:

A Case Study of 100% Performance Improvement in an Embedded C++ System

GILI KAMMA





INTRO

- ☐ The talk today is about software development.
- I worked on a product for two years, and during this time, I drastically improved its capabilities.
- I wanted to share with you the journey I had.
- I hope you will get some useful ideas and inspiration to improve your own product.

HELLO!

I am Gili Kamma

20 years in the industry

B.S.c in Electronics

I love to improve things and solve problems

Team leader @ Priority-software



BACKGROUND

My story begins....



BACKGROUND

The business is meters (water, heat and electricity)



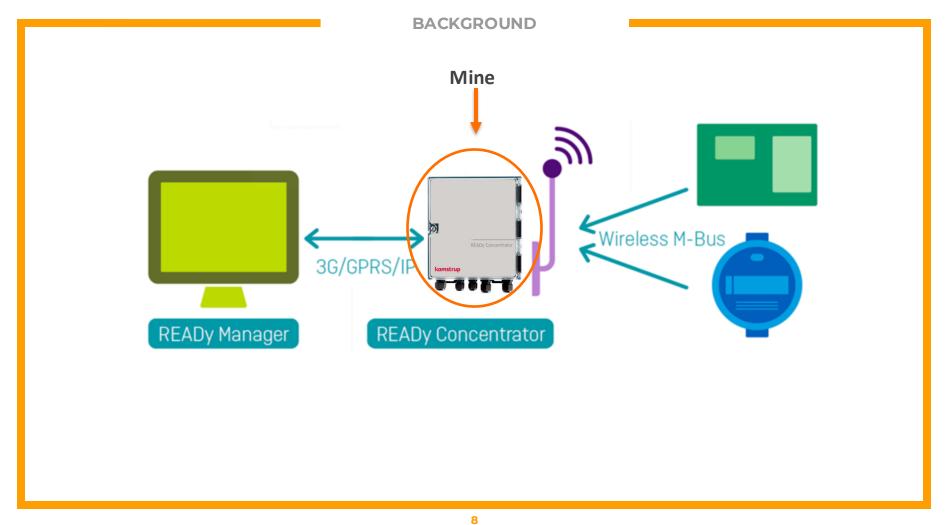
Smart metering

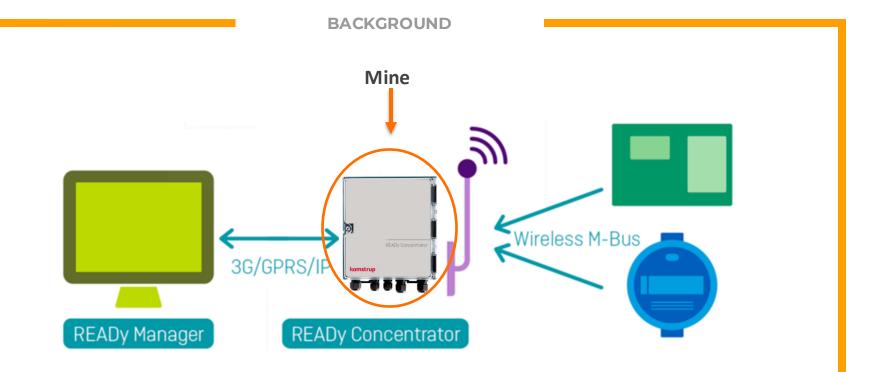


A system that uses digital devices to automatically send your utility data to the provider.

BACKGROUND







Collects meter measurements from water meters (wireless) and sends them to the cloud by cellular modem or ethernet.

BACKGROUND Mine Wireless M-Bus 3G/GPRS/IP **READy Manager READy Concentrator** Loss of data → loss of money.

The Problems



Unstable product.



- Unstable product.
- Unsatisfied customers.



- Unstable product.
- Unsatisfied customers.
- Unhappy developers.

- Unexplained crashes.
- Occasional data loss during network issues.

- Unexplained crashes.
- Occasional data loss during network issues.

- Memory issues we didn't have enough memory.
- Unexplained crashes.
- Occasional data loss during network issues.

- Unexplained crashes.
- Occasional data loss during network issues.

Technology Stack

- Linux kernel, Yocto distribution old image no sources.

- Qt − an application framework − a wrapper to low-level programming.

- Linux kernel, Yocto distribution old image no sources.

- Qt − an application framework − a wrapper to low-level programming.

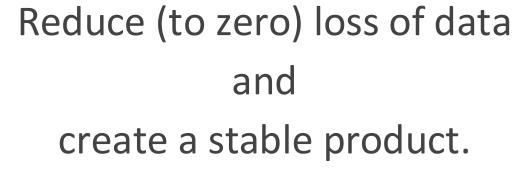
- Linux kernel, Yocto distribution old image no sources.

- Qt − an application framework − a wrapper to low-level programming.

- Linux kernel, Yocto distribution old image no sources.

- Qt − an application framework − a wrapper to low-level programming.





Reduce (to zero) loss of data and create a stable product.

How Am I Going to Do That?

First Impression

A lot of news & deletes

A lot of news & deletes

→ memory leaks & memory fragmentation

Memory fragmentation:

0 100

40 20 40

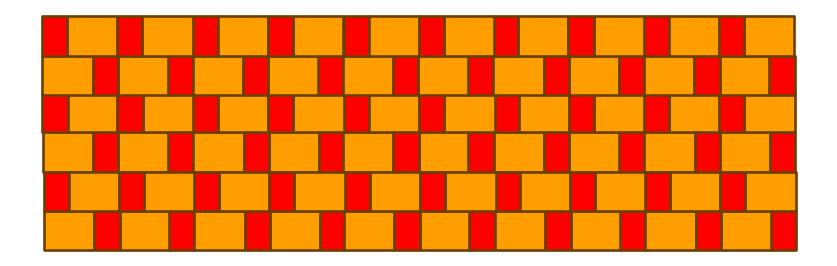
Memory fragmentation:

0 100

40 20 40

20

Memory fragmentation:



A lot of news & deletes memory leaks & memory fragmentation

explains not enough memory

A lot of news & deletes memory leaks & memory fragmentation explains not enough memory

→ allocation fails

A lot of news & deletes memory leaks & memory fragmentation explains not enough memory allocation fails

accessing null

A lot of news & deletes
memory leaks & memory fragmentation
explains not enough memory
allocation fails
accessing null

→ uncontrolled reset

A lot of news & deletes memory leaks & memory fragmentation explains not enough memory allocation fails accessing null uncontrolled reset

Q.E.D



Right? Wrong!

Right? Wrong!

It is not "wrong wrong" but it is wrong in this case

Wrong!

A lot of news & deletes

memory leaks & memory fragmentation

explains not enough memory

allocation fails

accessing null

uncontrolled reset





I needed to test my assumptions...



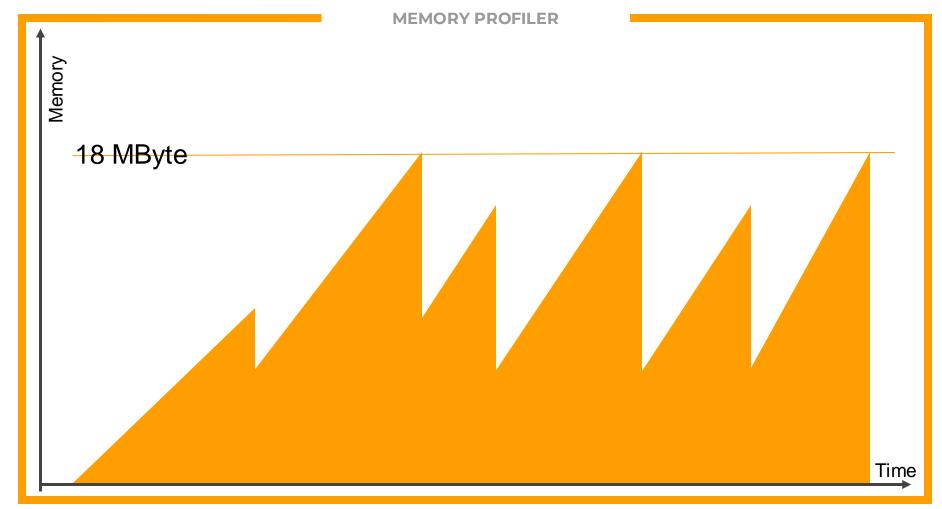


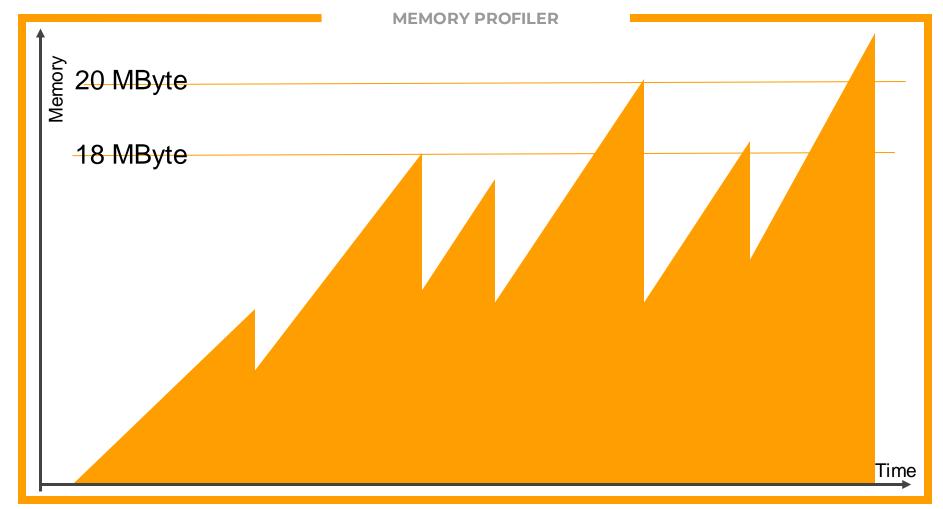
Memory Profiler

Gave me a good understanding of the sizes and the quantities of allocations.



I couldn't use standard memory profilers as the Linux kernel was old.





- Number of allocations per second.
- Current and maximum number of allocations.
- Current and maximum bytes allocated.
- Current and maximum allocations per size value.

- Number of allocations per second.
- Current and maximum number of allocations.
- Current and maximum bytes allocated.
- Current and maximum allocations per size value.

- Number of allocations per second.
- Current and maximum number of allocations.
- Current and maximum bytes allocated.
- Current and maximum allocations per size value.

- Number of allocations per second.
- Current and maximum number of allocations.
- Current and maximum bytes allocated.
- Current and maximum allocations per size value.

Implementation

App's pointer ...

size[bytes]

- array
- Overloaded new and delete operators.
- Added 4 bytes to each allocation.



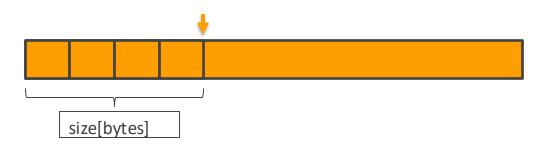
```
void* operator new(size_t size)
    void* ptr = malloc(size+4);
 *reinterpret_cast<size_t*>(ptr) = size;
    profiler.addAllocation(size);
    return (reinterpret_cast<char*>(ptr)+4);
         size[bytes]
```

```
void* operator new(size_t size)
    void* ptr = malloc(size+4);
    *reinterpret_cast<size_t*>(ptr) = size;
 profiler.addAllocation(size);
    return (reinterpret_cast<char*>(ptr)+4);
         size[bytes]
```

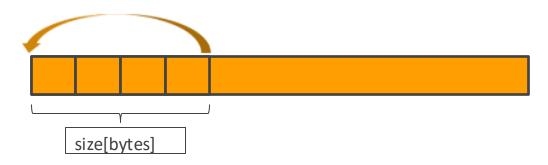
```
void* operator new(size_t size)
    void* ptr = malloc(size+4);
    *reinterpret_cast<size_t*>(ptr) = size;
    profiler.addAllocation(size);
 return (reinterpret_cast<char*>(ptr)+4);
         size[bytes]
```

MEMORY PROFILER



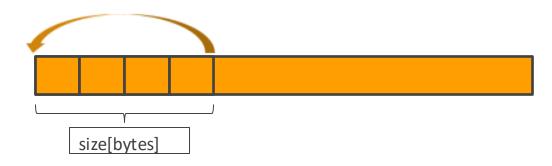






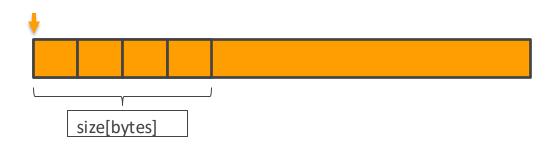
MEMORY PROFILER

```
void operator delete(void* ptr, std::size_t)
{
    size_t size = *reinterpret_cast<size_t*>(reinterpret_cast<char*>(ptr)-4);
    profiler.removeAllocation(size);
    free(reinterpret_cast<char*>(ptr)-4);
}
```



MEMORY PROFILER

```
void operator delete(void* ptr, std::size_t)
{
    size_t size = *reinterpret_cast<size_t*>(reinterpret_cast<char*>(ptr)-4);
    profiler.removeAllocation(size);
    free(reinterpret_cast<char*>(ptr)-4);
}
```



It helped me:

- Find the memory leaks and clean them.
- □ Figure out this is not my main problem.

It helped me:

- Find the memory leaks and clean them.
- Figure out this is not my main problem.

I've come across many problems, some big and some small.



Let's talk more about them

Problems:

- Unexplained crashes.
- Not enough information after crash.
- Occasional data loss during network issues.
- Suddenly, no meters are being detected anymore.

Unexplained crash after crash continuously (with gaps of several minutes) in big sites (>5000 water meters).

How to simulate 5000+ water meters?



- □ 1 data frame to 10-50 fake data frames.
- The unit crashed on my table when I simulated 6k meters and put "aggressive configuration".

```
1 void ReceiveThread(queue<Message>& queue)
1 struct Message
                                     2 {
2 {
                                           Message message = {};
      long opcode;
                                          const bool simulatorEnabled = std::filesystem::is regular file("myfile.txt");
      long id;
                                          while (true)
      long value;
6 };
                                               message = {};
                                               bool isReceived = ReceiveMessage(message);
                                     8
                                               if(isReceived)
                                    10
                                                   queue.push(message);
                                    11
                                                   if(simulatorEnabled)
                                    12
                                    13
                                    14
                                                       Message simulatorMsg = {};
                                                       for(int i=0; i<100; i++)</pre>
                                    15
                                    16
                                                           simulatorMsg = message;
                                    17
                                                           simulatorMsg.id = message.id + rand() % 1000;
                                    18
                                    19
                                                           queue.push(simulatorMsg);
                                    20
                                    21
                                    22
                                    23
                                    24 }
```

```
→ 1 void ReceiveThread(queue<Message>& queue)
    2 {
         Message message = {};
         const bool simulatorEnabled = std::filesystem::is_regular_file("myfile.txt");
         while (true)
    6
             message = {};
    8
             bool isReceived = ReceiveMessage(message);
             if(isReceived)
   10
                 queue.push(message);
   11
                 if(simulatorEnabled)
  12
   13
```

```
1 void ReceiveThread(queue<Message>& queue)
 2 {
      Message message = {};
 4---> const bool simulatorEnabled = std::filesystem::is_regular_file("myfile.txt");
      while (true)
 6
          message = {};
           bool isReceived = ReceiveMessage(message);
           if(isReceived)
10
               queue.push(message);
11
               if(simulatorEnabled)
12
13
```

```
1 void ReceiveThread(queue<Message>& queue)
 2 {
      Message message = {};
      const bool simulatorEnabled = std::filesystem::is_regular_file("myfile.txt");
      while (true)
 6
          message = {};
          bool isReceived = ReceiveMessage(message);
          if(isReceived)
10
              queue.push(message);
11
               if(simulatorEnabled)
12
13
```

```
1 void ReceiveThread(queue<Message>& queue)
 2 {
      Message message = {};
      const bool simulatorEnabled = std::filesystem::is_regular_file("myfile.txt");
      while (true)
 6
          message = {};
 8
          bool isReceived = ReceiveMessage(message);
          if(isReceived)
10
11
          queue.push(message);
              if(simulatorEnabled)
12
13
```

```
1 void ReceiveThread(queue<Message>& queue)
 2 {
      Message message = {};
      const bool simulatorEnabled = std::filesystem::is_regular_file("myfile.txt");
      while (true)
 6
          message = {};
 8
          bool isReceived = ReceiveMessage(message);
          if(isReceived)
10
11
              queue.push(message);
          → if(simulatorEnabled)
12
13
```

```
queue.push(message);
11
12
          → if(simulatorEnabled)
13
14
                   Message simulatorMsg = {};
                   for(int i=0; i<100; i++)</pre>
15
16
17
                       simulatorMsg = message;
18
                       simulatorMsg.id = message.id + rand() % 1000;
                       queue.push(simulatorMsg);
19
20
21
22
23
24 }
```

```
queue.push(message);
11
12
               if(simulatorEnabled)
13
14
                   Message simulatorMsg = {};
                   for(int i=0; i<100; i++)</pre>
15
16
17
                  simulatorMsg = message;
                       simulatorMsg.id = message.id + rand() % 1000;
18
                       queue.push(simulatorMsg);
19
20
21
22
23
24 }
```

```
queue.push(message);
11
12
               if(simulatorEnabled)
13
14
                   Message simulatorMsg = {};
                   for(int i=0; i<100; i++)</pre>
15
16
17
                       simulatorMsg = message;
18
                  simulatorMsg.id = message.id + rand() % 1000;
                       queue.push(simulatorMsg);
19
20
21
22
23
24 }
```

```
queue.push(message);
11
12
               if(simulatorEnabled)
13
14
                   Message simulatorMsg = {};
                   for(int i=0; i<100; i++)</pre>
15
16
17
                       simulatorMsg = message;
18
                       simulatorMsg.id = message.id + rand() % 1000;
                     queue.push(simulatorMsg);
19
20
21
22
23
24 }
```

```
queue.push(message);
11
12
              if(simulatorEnabled)
13
14
                  Message simulatorMsg = {};
              → for(int i=0; i<100; i++)
15
16
17
                       simulatorMsg = message;
18
                       simulatorMsg.id = message.id + rand() % 1000;
                       queue.push(simulatorMsg);
19
20
21
22
23
24 }
```

```
1 void ReceiveThread(queue<Message>& queue)
1 struct Message
                                     2 {
2 {
                                          Message message = {};
      long opcode;
                                          const bool simulatorEnabled = std::filesystem::is_regular_file("myfile.txt");
     long id;
                                          while (true)
      long value;
6 };
                                               message = {};
                                               bool isReceived = ReceiveMessage(message);
                                               if(isReceived)
                                    10
                                                   queue.push(message);
                                    11
                                    12
                                                   if(simulatorEnabled)
                                    13
                                    14
                                                       Message simulatorMsg = {};
                                                       for(int i=0; i<100; i++)</pre>
                                    15
                                    16
                                                           simulatorMsg = message;
                                    17
                                                           simulatorMsg.id = message.id + rand() % 1000;
                                    18
                                    19
                                                           queue.push(simulatorMsg);
                                    20
                                    21
                                    22
                                    23
```

24 }

- □ 1 data frame to 10-50 fake data frames.
- The unit crashed on my table when I simulated 6k meters and put "aggressive configuration".

- No special hardware required.
- Easy access to simulator mode without additional building.
- Consistently crashed the system.

- No special hardware required.
- Simple implementation.
- Easy access to simulator mode without additional building.
- Consistently crashed the system.

- No special hardware required.
- Simple implementation.
- □ Easy access to simulator mode without additional building.
- Consistently crashed the system.

- No special hardware required.
- Simple implementation.
- □ Easy access to simulator mode without additional building.
- Consistently crashed the system.

FIND THE ROOT CAUSE

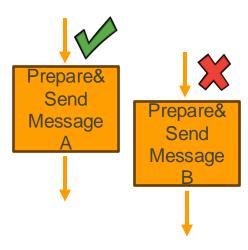
```
void* operator new(size_t size)
    void* ptr = malloc(size+4);
 *reinterpret_cast<size_t*>(ptr) = size;
    profiler.addAllocation(size);
    return (reinterpret_cast<char*>(ptr)+4);
         size[bytes]
```

```
void* operator new(size_t size)
    void* ptr = malloc(size+4);
  *reinterpret_cast<size_t*>(ptr) = size;
    profiler.addAllocation(size);
    return (reinterpret_cast<char*>(ptr)+4);
```

size[bytes]

Two threads tried to create large messages about 5 MB each, at the same time.

The second one always failed.



Imagine this issue on your system.
How would you solve it?



- Changed from asynchronous work to synchronous work.

- Used a static array instead of dynamic allocation.

SOLVE Prepare& Send Message Prepare& Send Prepare& Message Send Prepare& Message Send Message

Asynchronous to synchronous

5MB+5MB → 5MB

So, I am not crashing anymore. Can I stop here?

No!



7500 meters and not just 5000

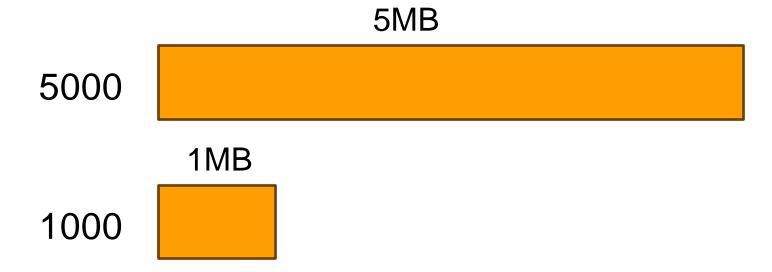
Asynchronous to synchronous

5MB+5MB → 5MB

Why doesn't it happen in all sites?



The message size was proportional to the number of water meters the unit listened to.



- Changed from asynchronous work to synchronous work.
- Split large messages into smaller ones.
- Used a static array instead of dynamic allocation.

SOLVE 5MB 1X 99

SOLVE 5MB 1X 1MB 5X 100

Asynchronous to synchronous

5MB+5MB → 5MB

Split large messages into smaller ones

5MB **→** 1MB

- Changed from asynchronous work to synchronous work.
- Split large messages into smaller ones.
- Used a static array instead of dynamic allocation.

Asynchronous to synchronous

5MB+5MB → 5MB

Split large messages into smaller ones

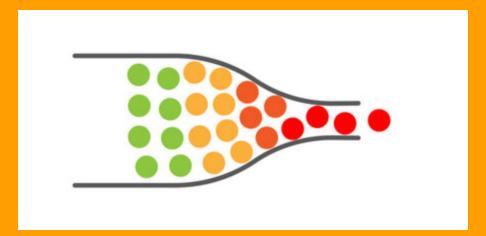
5MB **→** 1MB

Qt to C++

1MB → 0.5MB

- Changed from asynchronous work to synchronous work.
- Split large messages into smaller ones.
- Used a static array instead of dynamic allocation.

Bottle neck of the system



Problems:

- Unexplained resets.
- Not enough information after crash.
- Occasional data loss during network issues.
- Suddenly, no meters are being detected anymore.

- I had standard application logs.
- □ I didn't have a remote access.
- ☐ I had an event mechanism (RAM).

- I had standard application logs.
- ☐ I didn't have a remote access.
- ☐ I had an event mechanism (RAM).

- I had standard application logs.
- ☐ I didn't have a remote access.
- ☐ I had an event mechanism (RAM).

When the unit wakes up, send the last 100 lines from the application log to the backend using the event service.

```
MFO — [Mar 31 23:17]

t/buddy_List-php HIIF/
t/buddy_List-php HIIF/
sence/reconnect-php?
sence/reconnect-php?
n=6&fb_dtsg=AQDJ95ij HIIF/
INFO — [Mar 31 23:17:22]

INFO — [Mar 31 23:17:22]

INFO — [Mar 31 23:17:22]

dPlayer-swf?v=1 HIIF/1

dPlayer-swf?v=1 HIIF/1

INFO — [Mar 31 23:18:32]

INFO — [Mar 31 23:18:32]

INFO — [Mar 31 23:18:32]
```

When the unit wakes up, send the last 100 lines from the application log to the backend using the event service.

```
MFO — [Mar 31 23:17:22]

t/buddy_list.php HIIF/

t/buddy_list.php HIIF/

sence/reconnect.php?

n=6&fb_dtsg=AQDJ95ij HIIF/

INFO — [Mar 31 23:17:22]

dPlayer.swf?v=1 HIIF/1

INFO — [Mar 31 23:18:32]

INFO — [Mar 31 23:18:32]

INFO — [Mar 31 23:18:32]

INFO — [Mar 31 23:18:32]
```

- □ tail -n 100 /var/log/my_log.txt→shortLog
- SendEvent ("reset reason", shortLog)

Good example to good enough solution

80% result / 20% effort

Problems:

- Unexplained resets.
- Not enough information after reset.
- Occasional data loss during network issues.
- Suddenly, no meters are being detected anymore.

Data for transmission remains in RAM, awaiting further processing.

Data for transmission remains in RAM, awaiting further processing.



So, what is the problem with that?

Data for transmission remains in RAM, awaiting further processing.

In case of unstable communication:

- Start to aggregate takes a lot of space.
- → Loss of data in case of reset.

•

Disconnect the Logic from the Network



Disconnect the Logic from the Network

Thread #1 → Logic



Thread #2 → Sending



Disconnect the Logic from the Network

Thread #1 → Logic



Thread #2 → Sending



Always execute the same logic and store the results in nonvolatile memory (regardless of the current network status).

Implementation Achievements:

- Maximum data loss is now limited.

Implementation Achievements:

- Maximum data loss is now limited.
- Not being sensitive any more to network errors.

Implementation Achievements:

- Maximum data loss is now limited.
- Not being sensitive any more to network errors.

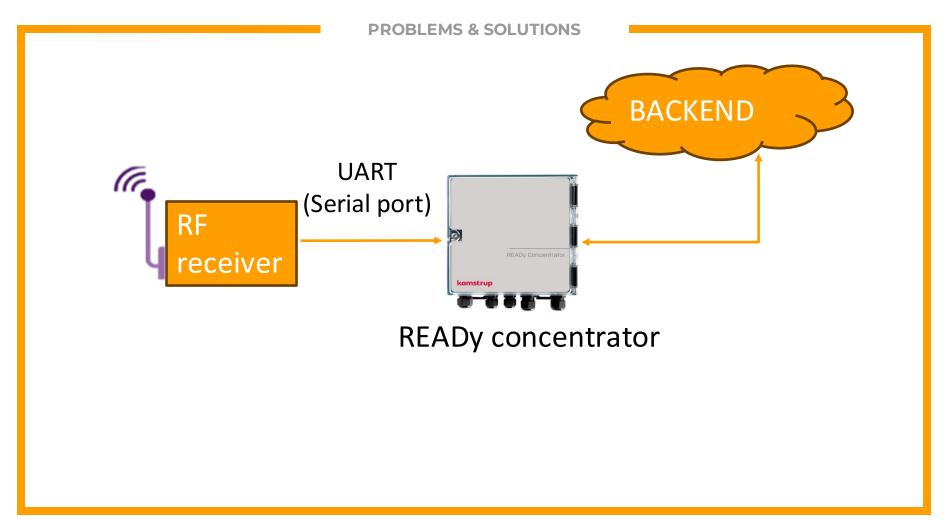


After implementing this change, we no longer experienced any significant data loss.

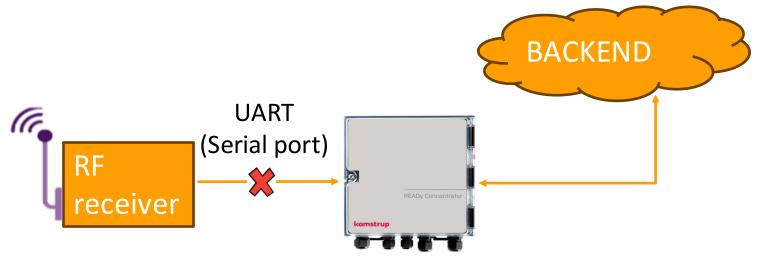
A good example of always using the same scenario to avoid any edge cases

Problems:

- Unexplained resets.
- Not enough information after reset.
- Occasional data loss during network issues.
- Suddenly, no meters are being detected anymore.



PROBLEMS & SOLUTIONS



READy concentrator

Once in a while we stopped receiving frames (The system doesn't recover on its own)

- Customers complained.
- Every time it was a different unit.
- We didn't understand why it happens.
- Controlled Reset solved the problem.

- Customers complained.
- Every time it was a different unit.
- We didn't understand why it happens.
- Controlled Reset solved the problem.

- Customers complained.
- Every time it was a different unit.
- We didn't understand why it happens.
- Controlled Reset solved the problem.

- Customers complained.
- Every time it was a different unit.
- We didn't understand why it happens.
- Controlled Reset solved the problem.

Reset the HW after 1 hour of silence.

Reset the HW after 1 hour of silence.

Second solution:

Reset the HW after 4 minutes of silence (sweet spot).

Reset the HW after 1 hour of silence.

Second solution:

Reset the HW after 4 minutes of silence (sweet spot).

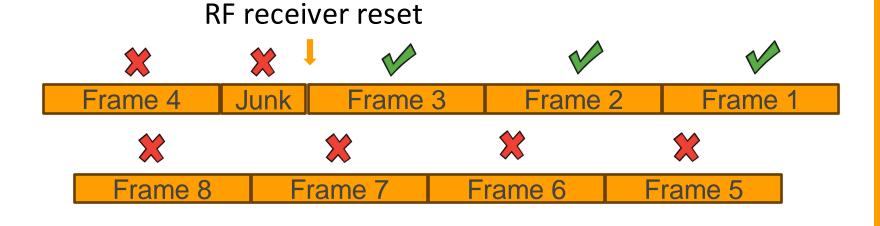
Reset the HW after 1 hour of silence.

Second solution:

Months later, I was looking in logs of a unit that had this silence issue – and I realized the RF receiver reset itself.



As a result, the READY concentrator couldn't sync again on the next message.



A real solution:

Sometimes it takes time to identify the real problem, so we use temporary solutions in the meantime.

Monitoring

How it was when I started to work:

Complaints



Look at the event service

After several weeks:

Look at the event service



! Complaints

How I started:

- Started reading events for specific units, but it was uncomfortable.

How I started:

- Started reading events for specific units, but it was uncomfortable.

How I started:

- Started reading events for specific units, but it was uncomfortable.
- Wrote a Python script to analyze the text and export to an Excel file (sort, filter, plot).

MONITORING

Timestamp	Event		
9/16/23 10:00	data send		
9/16/23 11:00	data send		
9/16/23 12:00	data send		
9/16/23 13:00	data send		
9/16/23 14:00	data send		
9/16/23 15:00	data send		
9/16/23 16:00	data send		
9/16/23 17:00	data send		
9/16/23 18:00	data send		
9/16/23 19:00	data send		
9/16/23 20:00	data send		
9/16/23 21:00	data send		
9/16/23 22:00	data send		
9/16/23 23:00	data send		
9/17/23 0:00	data send		
9/17/23 1:00	data send		
9/17/23 2:00	data send		

How I started:

- □ Ran the scripts every morning, analyzing for 10 minutes daily.
- Added more units and parameters.
- Automated the scripts to run at night and receive results by email in the morning.

How I started:

- □ Ran the scripts every morning, analyzing for 10 minutes daily.
- Added more units and parameters.
- Automated the scripts to run at night and receive results by email in the morning.

How I started:

- Added more units and parameters.
- Automated the scripts to run at night and receive results by email in the morning.

- Max memory usage.
- How many errors.
- Last time to be seen.
- How many resets.
- Current configuration.

- Max memory usage.
- How many errors.
- Last time to be seen.
- How many resets.
- Current configuration.

- Max memory usage.
- How many errors.
- Last time to be seen.
- How many resets.
- Current configuration.

- Max memory usage.
- How many errors.
- □ Last time to be seen.
- How many resets.
- Current configuration.

- Max memory usage.
- How many errors.
- □ Last time to be seen.
- How many resets.
- Current configuration.

- Max memory usage.
- How many errors.
- □ Last time to be seen.
- How many resets.
- How many meters were reported.
- Current configuration.

- Max memory usage.
- How many errors.
- □ Last time to be seen.
- How many resets.
- How many meters were reported.
- Current configuration.

Be proactive!



Use the data you already have



E2E tests:

- Rewrote the test paper from manual tests to automatic tests.
- Implemented the automated tests.
 - □ Two pipelines long and short tests.
 - \square 24/7 to stabilize the tests.

E2E tests:

- Rewrote the test paper from manual tests to automatic tests.
- Implemented the automated tests.
 - □ Two pipelines long and short tests.
 - \square 24/7 to stabilize the tests.

System tests:

System tests:

The question isn't if there's a bug - it's who will find it first

How did it end?

8 Months Later

- Product without any resets.



- Product without any resets.



2 Years Later

- Added features and improvements.

Take Home Messages

Measure Everything – Avoid Assumptions



(Remember – my first impression)

Write Robust Software

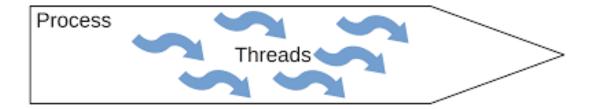


Unexpected inputs are unavoidable, but don't let your system crash – have safeguards in place

- Split large messages into smaller, fixed sizes.
- Decouple logic and network operations.

- Split large messages into smaller, fixed sizes.
- Decouple logic and network operations.

Reduce the Number of Threads



(The most challenging bugs in the system arise from multiple threads running simultaneously)

Monitoring – Be Proactive.



Don't Wait for Customer Complaints

Create Simple Software



Creating simple software is more challenging than making it complex.

Thanks!

Any questions?