Lyddon
CONSULTING

+44 (0) 7979 132 341
enquiries @lyddonconsulting.com
www.lyddonconsulting.com

13th March 2025

T2 Disaster Recovery arrangements are synthetic, befitting a synthetic currency

Introduction

T2 (the renamed TARGET2) is a systemically-important payment system and the Real-Time Gross Settlement System (RTGS) for the euro. As such it should be a '5-nines' system, enjoying 99.999% availability during its scheduled operating hours.

On 27th February it went down from 10:15 until 18:00 CET, according to the operational update message sent by the European Central Bank (ECB) on 28th February at 17:10 CET.[1]

The explanation of the outage, when set against how a top New York commercial bank had configured its Disaster Recovery and how that performed on 9/11, demonstrates that the Disaster Recovery arrangements behind T2 are synthetic to the point of non-existence.

The proof of that is that they failed when they were called upon to kick in.

The euro as a synthetic currency

The euro is a synthetic currency in that it enjoys no 'full faith and credit' backing from any single country.[2] This backing cannot exist when there are separate Eurozone member states. No Eurozone member state fulfils the tests of being a 'sovereign' towards the euro: a country that is the sole user of the currency and which controls all of the monetary policy tools regarding the currency. The Eurozone member states share usage of the currency and have surrendered control of the monetary policy toolkit to the ECB.

There is therefore no 'sovereign borrower' amongst the Eurozone member states. Nor do any of the EU supranationals qualify as 'sovereign borrowers': the European Union itself, the European Investment Bank, the European Stability Mechanism and the European Financial Stability Facility.

The member states are 'sub-sovereigns' due to their deviations from the definition of 'sovereign'. The EU supranationals sit above the member states in the topology, but they are no more than porous membranes which allow their creditors to look through them to the creditworthiness of the member states. Creditors lend their money to the supranational but their credit risk is on the member states, although not always on a joint-and-several-liability basis.

[1] https://www.ecb.europa.eu/paym/target/html/t2s_history.en.html

[2] This means that every tax-liable legal person – both natural and non-natural legal persons – in the country is responsible for the country's debts and on a joint-and-several-liability basis: every entity is liable to pay, whether or not the others do, up to a point where the 'last entity standing' is liable for the entire debt

Only a debt of the EU itself delivers joint-and-several-liability. Debts of the other supranationals deliver different versions of the lesser 'several-but-not-joint liability': each member state is only liable up to a fixed figure, and a creditor cannot cause the others to pay more if one member state fails to meet its liability in full.

The versions of 'several-but-not-joint liability' are based on member state guarantees (the European Financial Stability Facility) or on capital that has been subscribed by member states but not yet called up (the European Investment Bank and the European Stability Mechanism).

Because there is no 'sovereign' in the Eurozone, there cannot be a Eurozone 'safe asset'.

Terminology

Despite the reality of the situation, the EU and Eurozone authorities use the same terminology as US, Japanese, UK and other authorities that are a 'sovereign', do have a 'sovereign borrower', and 'safe assets'.

The same problem spills over into T2. All the right words are there on the surface, but the substance underneath diverges.

Sharing is politically-driven and results in weakness

T2 is a partially-centralised system, unlike the Bank of England's RTGS/CHAPS.

All of the Eurozone national central banks, and the ECB, and several non-Eurozone EU national central banks share usage of T2 via access to the Single Shared Platform (SSP).

'SSP' is not a synonym for 'the IT platform on which T2 runs'. The SSP does not house the totality of functions that add up to 'T2'. Some functions – like running the accounts of banks that a national central bank has sponsored into T2 – can be done outside the SSP.[3]

The term 'SSP' is an oxymoron as its operation is shared around four of the Eurozone national central banks: the original triumvirate of Germany, France, and Italy, and now with Spain as well. The sharing, or distribution, of T2 across four member states undermines the contention that is a 'single' platform in the way the casual observer would understand the term.

The sharing is politically-driven: no one member state would want to have a single, other member state control the system entirely. A 'system' distributed like this across four member states cannot be stronger than a fully-centralised one. Indeed it is liable to be weaker.

Weakess evidence in topology and performance of T2 Disaster Recovery arrangements

The topology of the T2 Disaster Recovery arrangements has the potential for weakness, despite a service level being published that implies robustness. The way T2 is being run in practice is evidently not in line with the published service level.

This deficit is both the result of the structural weakness and a contributor to it: the two compound one another.

[3] This would occur on the HAM, or 'Home Accounts Module' outside the SSP

The 'Information Guide for TARGET participants Part 2 – CLM & RTGS Version R2024.NOV' is the document that gives the most detail about the TARGET2 Disaster Recovery arrangements, and specifically para 4.1.2.5 starting on p. 44 under the heading 'Service continuity (failover to the second site or second region)'.[4]

Different elements in the arrangements are spread around 'regions' and 'sites' in different member states. The arrangements are written as if work will be handed off in-country and/or cross-border, phrasing which only makes sense if the SSP is loaded in multiple IT complexes in parallel, each one of which is independent of the others but kept in complete or near-complete synchronisation.

It remains opaque whether 'sites' in the same member state are fully independent of one another.

If they are not independent, then they do not fulfil a basic qualification for acting as standby sites for one another.

'Sites', if they located in different member states, are perforce independent of one another, but can only act as standby sites for one another if they are kept in complete (for hot standby) or near-complete (for warm standby) synchronisation with the 'sites' they are acting as standby for.

To compound that further, the roles of the different 'sites' are rotated.

This introduces an enormous level of complexity, whose customary impacts are:

- Difficulties in fault-finding when something does go wrong;
- More difficult to repair the fault;
- No time to test the repair adequately;
- Longer outage;
- Higher risk of repeat failure.

These issues could be mitigated by:

1. Going to single centralised IT platform, probably housed in the most neutral and unobjectionable EU venue – Luxembourg – and having two standby sites also in Luxembourg; or
2. Large extra monetary investment in what exists now.

'Regions'

Despite the inclusion of Banco de Espana into the inner circle of central banks running the operation of T2, there is no indication in the papers that the basics of how the system is run in normal production have changed, nor of change to the Disaster Recovery arrangements.

They appear to be as they were in 2020, the previous time there was a major TARGET2 outage. If Banco de Espana has taken on an operational role, though, that would further increase the complexity.

[4] CLM stands for Centralised Liquidity Management

The page references in this section are to the User Book valid at the time of the 2020 outage:

- The SSP has three 'regions'. Regions 1 and 2 (payments and accounting) are run in data centres of the Bundesbank and Banca d'Italia; Region 3 (Customer related service system) is run in data centres of the Banque de France;
- The Regions are hundreds of kilometres apart, although the 'sites' within the 'regions' may well be closer together, and indeed may even be in the same IT complex;
- Region 1 is the live system; Region 2 is the disaster recovery and the testing system. They are supposedly replicas of one another, kept 'in synch' via the 4CB communications network;
- The role of being Region 1 is rotated between the Bundesbank and Banca d'Italia: when the Bundesbank has Region 1, Banca d'Italia has Region 2 and vice versa.

'Sites'

Each Region contains two sites (User Book 1 for v14.0 p367): a primary site and a recovery site. It is not specifically stated that these 'sites' are in separate data centres of the central bank concerned.

So the first failover for live operation is the recovery site within Region 1; the second failover is in Region 2, and presumably to the primary site in Region 2.

What happens if all 'regions' and 'sites' go down

Para 4.1.2.4 of the T2 2024 Information Guide for CLM and RTGS, starting on p. 36, describes the ECONS II procedure that 'supports the settlement of very critical and critical cash transfer orders in the event that CLM and/or RTGS are unavailable'.

These are called 'Contingency arrangements', whereas the casual observer might understand that phrase as being the main Disaster Recovery procedure.

ECONS II was invoked on 27th February.

What happened on 27th February

These are the key extracts from the ECB communication:

- Due to a defective hardware component, both services [meaning T2 and T2S - TARGET2 Securities] suffered from a slowdown and subsequent unavailability from respectively 08:07 and 10:15 CET;
- Contingency arrangements were however activated in T2 at around 11:30 [presumably there are no such arrangements in T2S];
- First analysis had indicated an issue at the level of the system's database. Owing to the real-time synchronisation between the databases of the two operational sites, the possibility of executing a failover to the secondary site was initially excluded, as the issue would have likely been replicated in the other database;
- At around 15:45, once it was confirmed that the problem was actually due to an infrastructure component, the decision was taken to activate the failover of T2S and T2 to the secondary site. The failover was technically completed around 17:15;
- Owing to the nature of the incident, further precautionary checks had to be performed before both systems could resume operations, shortly after 18:00;
- The closing of T2S and T2 was postponed by 6 hours. The systems eventually closed orderly at 00:00 (instead of 18:00 on a normal day).

Examination of the phrase 'synchronisation between the databases of the two operational sites'

A tenet of Disaster Recovery is that three environments are needed and that they all be replicas of one another: production, hot standby and warm standby.

The production and hot standby environments receive and process all items as they come in: the cut-over from one to the other is instantaneous.

That being the case, no fault can occur in the one without it also occurring in the other. In this case a fault occurred in the production environment without its occurring in the hot standby.

In this case it is implied that that the slowdown was not visible in the hot standby environment: the hot standby was working more quickly than the production environment.

The ECB feared that if they switched to the hot standby the fault would appear there: 'the possibility of executing a failover to the secondary site was initially excluded, as the issue would have likely been replicated in the other database'.

All this can mean is that the 'synchronisation between the databases of the two operational sites' was partial rather than being complete. The two environments were not being run as exact counterparts of one another.

It is a matter of speculation whether the fault might have been isolated more easily if the two environments had been counterparts and there had been an attempt at switchover. As it is the T2 Disaster Recovery failed four key tenets:

1. Not running the production and hot standby environments as complete replicas of one another, including processing all the traffic;
2. Allowing a hardware component to fail;
3. Allowing it to fail in one environment and not the other;
4. Taking hours to isolate the fault.

Hardware discrepancy between the two 'sites'

The explanation of a hardware fault is unacceptable. RTGS systems run by definition on 'always-on', fault-tolerant computers. The entire system must run on that computer, and not on several interconnected ones, which may not all be of the same standard.

The application environment must be simple: handling both high volume and high value demands a simple application architecture.

The production, hot standby and warm standby environments must be separate and independent environments but all under the one hand of control.

There must be a control room overseeing the performance of the entire system. Where is it? Or is the control function distributed?

The occurrence of a hardware fault found in one 'site' but not the other should be impossible, and should not cause downtime even if it did occur:

- The computers are fault-tolerant;
- The hot standby should kick in at once if the fault is in the production site;
- If the fault is in the hot standby site, it needs to be fixed as soon as possible but does not interrupt processing at the production site.

Measured against that, the ECB explanation leaves more questions open than it answers, such as:

- Why did it take so long to isolate the fault at the production site?
- What work needed to be undertaken at the production site to remedy the fault?
- By when was it remedied?
- Were checks made on the hardware at the second 'site' between 11:30 and 15:45 to make sure that either the error was either not present there, or could be repaired before the 'intra-region failover' was attempted?
- Why was the 'intra-region failover' required, if it was the case that the hardware fault could be checked for and, if existing, fixed on the 'second' site during that time window: why could it not also be fixed on the first one?

Key provisions of para 4.1.2.5 'Service continuity (failover to the second site or second region)'

The ECB operational bulletin can be compared with the provisions of para 4.1.2.5 'Service continuity (failover to the second site or second region)':

- An intra-region failover means the failing over from site A to site B within the same region. As a synchronous copy is applied, the databases at both sites are exactly the same, and no reconciliation is required after the failover;
- An intra-region failover ensures the continuation of normal business within a maximum of one hour after the decision to recover from the other site is taken;
- A wide-scale regional disruption that causes a severe interruption of transportation, telecommunication, power or other critical infrastructure across a metropolitan or a geographical area will require the failing over to the second region;
- As an asynchronous copy is applied, the databases in the two regions show the processing status with a time discrepancy of no longer than two minutes;
- Loss of data following an inter-region failover when both sites within Region 1 become unavailable is possible. In this situation, there is no alternative but to fail over to Region 2 and reconcile the missing traffic. Still, the resumption of operations in Region 2 should be enabled within two hours of the decision-making process.

Key interim conclusions

The key interim conclusions to be drawn are:

- No attempt was made to do an inter-region failover to 'Region 2';
- The incident was managed in between the two 'sites' within the then-current 'Region 1';
- The databases at the two sites in 'Region 1' were not 'exactly the same';

Continuation of Key interim conclusions:

- This explains why, when the incident was initially diagnosed as a database issue, it was decided not to attempt the 'intra-region failover' to the second 'site' on the grounds that it would introduce a fault into the second 'site'. If the two sites had been 'in synch', the fault would already have been there;
- After it had been discovered that the fault was a hardware one, and presumably after it had been checked that the same fault was not present in the second 'site', the 'intra-region failover' to the second 'site' was attempted and was successful;
- It took 75 minutes after T2 failed – which was 203 minutes after T2S failed – to 'activate' the T2 so-called 'Contingency arrangements', which is only a portion of the Disaster Recovery;
- It took a further 255 minutes to complete the diagnosis of the problem, work out how to proceed and then to decide on an 'intra-region failover';
- It took 90 minutes to carry out the 'intra-region failover' to the second 'site';
- 'Further precautionary checks' were carried out lasting 45 minutes 'before both systems could resume operations, shortly after 18:00';
- By that stage T2S had been down for 9 hours and 53 minutes, and T2 for 7 hours and 45 minutes.

Comparison with service level

- The above timeline compares to the service level for an 'intra-region failover' of 'a maximum of one hour after the decision to recover from the other site is taken';
- This phrasing demonstrates the loopholes in the service level, which does not mention a delay between the service going down and the decision being taken to cut over to the second 'site': the casual observer would assume that a decision would be taken after a few minutes. The professional observer might be a surprised that the failover did not happen automatically and that a decision was required at all;
- The 'intra-region failover' firstly took 30 minutes longer than is stated in the service level as a 'maximum of one hour';
- Then the service level does not mention a need for 'further precautionary checks' lasting another 45 minutes.

Inter-region failover not attempted

The conditions written into the service level were not met for an 'inter-region failover': there was no 'wide-scale regional disruption that causes a severe interruption of transportation, telecommunication, power or other critical infrastructure across a metropolitan or a geographical area'. This a high bar and an irrelevant one: why does the cause of an outage matter? The more relevant condition would be that the production site had gone down and that the hot standby site could not be brought up to replace it.

The conditions serve only to speculate what might be happening in the outside world to cause a lengthy outage. They appear to be a deterrent to invoking the 'inter-region failover' and one that introduces considerations extraneous to the task in hand. The length of the outage should determine the response not the presence or absence of certain indicators outside the building. The interruption was such that an 'inter-region failover' would have been appropriate, unless for some reason it could not be invoked.

The service level maintains that the T2 traffic is going into the database in 'Region 2' albeit at a short delay: 'As an asynchronous copy is applied, the databases in the two regions show the processing status with a time discrepancy of no longer than two minutes'.

It then admits that 'Loss of data following an inter-region failover' is possible but that 'the resumption of operations in Region 2 should be enabled within two hours of the decision-making process'.

This means that an 'inter-region' failover is a valid response for a quite short outage: it ought to be possible to cutover within some minutes if the need arose. It certainly ought to be available for a lengthy intraday outage: it does not have to be reserved for a full-day or multi-day outage.

Yet it was not attempted on 27th February and the question remains open as to why not. The answer to that can only be induced from extrapolating what we suspect applies to the extent of the synchronisation between the two 'sites' in the same 'region':

- The synchronisation between the two 'sites' in the same 'region' is not complete, as the ECB's explanation of the outage confirms;
- The 'synchronisation' between the two 'regions' is 'asynchronous' to the point of non-existence;
- The 'inter-region failover' procedure is not readily available in reality;
- The production 'site' in Region 2 is cold, not warm.

Instead, 'Contingency arrangements were…activated in T2 at around 11:30 in order to settle the most critical payments': this is the ECONS II procedure described in para 4.1.2.4. There was no such contingency behind T2S. ECONS II is applied to a very small subset of the volume of T2 payments, but a much higher portion of the traffic's total value. Its naming, though, would be understood by the casual observer as meaning it can be applied to the entirety of the traffic.

Disaster Recovery arrangements in a New York commercial bank

These were the arrangements put in place by a New York commercial bank and leading clearer of Fedwire, CHIPS and SWIFT payments around the year 2000:

- Production site near the Twin Towers
- Hot standby site across the Hudson River on separate power supply
- Warm standby site in a data centre in upstate New York
- These roles were not rotated
- The day's traffic is processed across the hot standby environment in parallel to its being processed across the production environment
- The day's traffic is processed across the warm standby environment in batches every fifteen minutes
- Assumption that traffic is 95%+ Straight-Through Processed (STP) overall, with even higher STP rates for the most time-critical payments
- Operators are needed only to work on queued non-STP messages
- There is no 'standby' team in upstate New York to take over the queue management if the production and hot standby sites go down
- Operators should be able to work remotely (i) from the production site on the hot standby machine; (ii) from the hot standby site on the hot standby machine; (iii) from the production site or the hot standby site on the warm standby machine

Continuation of Disaster Recovery arrangements in a New York commercial bank

- The service level for non-STP traffic is that the bank would complete it by close-of-business as long as it was not blocked for credit or Money Laundering/Terrorist Financing issues
- The sites are all exact replicas of one another, in separate buildings at least a few kilometres apart, and on separate power networks;
- The failover from production to hot is instantaneous;
- The failover from production or hot to warm is up to thirty minutes – ten minutes to bring it to online processing from batch, and twenty minutes to process (i) all the traffic that has collected in the warehouse since the last batch was retrieved and processed; and (ii) all the traffic that comes in after the switchover from batch to online;
- No checks are needed before the hot or warm site takes over.

How these failed on 9/11

The bank's Disaster Recovery plans were 5-Star, and duly certified as such by financial authorities, but they failed because the scale of the disaster was beyond what could be planned for.

It affected multiple elements in the plan, individually and badly :

- The building housing the production site was badly damaged and later had to be demolished;
- The power to both the production site and hot standby was cut off: the failover procedure could not be initiated either between the two or to the warm standby;
- Staff could not work in either the production site building nor travel across the Hudson River to the hot standby site building.

Why that was acceptable and the T2 incident was not

9/11 caused a situation in the outside world even more severe than the T2 conditions for an 'inter-region failover', which are given as 'wide-scale regional disruption that causes a severe interruption of transportation, telecommunication, power or other critical infrastructure across a metropolitan or a geographical area'.

Nothing like that happened in Region 1 on 27th February, and yet the length of the outage matched what could have been expected had those conditions prevailed around Region 1.

The fact that no 'inter-region failover' was attempted is a strong indicator that it is a procedure that only exists on paper and not in reality.

9/11 caused unprecedented difficulties, of the type that no-one could have conceived of. The T2 outage was a bread-and-butter event for Disaster Recovery professionals.

Given that the ECB is the supreme monetary authority in the Eurozone, which even more supreme authority has certified its Disaster Recovery plans for T2?

Summary and conclusions

27th February proved that the Disaster Recovery arrangements behind T2 are synthetic.

The 'inter-region failover' procedure is synthetic to the point of not existing.

The occurrence of the outage, its duration, the reason for its occurrence, and the ECB's explanation of its handling of the outage are all unacceptable. They point to Disaster Recovery arrangements that would not be accepted by the financial authorities overseeing a major New York clearing bank and direct member of the CHIPS and Fedwire systems.

What is expected of a member of a Financial Market Infrastructure (the official umbrella term applied inter alia to payment and securities clearing and settlement systems) should be all the more expected of the Financial Market Infrastructure itself.

T2 is a priori weak because of its concept of 'sharing' and the compromises that derive from that.

The T2 concept needs to be completely revised, and the hardware and application architectures and the Disaster Recovery arrangements brought up to a 5-Star standard.

BL/13.3.25