

# Análisis de prestaciones de algoritmos de encaminamiento adaptativos para la conmutación de ráfagas ópticas

Elias Hortas, Davide Careglio, Mirosław Klinkowski, Josep Solé-Pareta  
Advanced Broadband Communications Centre (CCABA)  
Universitat Politècnica de Catalunya (UPC)  
Jordi Girona 1-3, Campus Nord, 08034 Barcelona  
E-mail: d8131538@est.fib.upc.edu, {careglio, mlinkow, pareta}@ac.upc.edu

## Resumen

En este artículo se presentan una serie de algoritmos de encaminamiento de tipo adaptativos diseñados para redes ópticas de conmutación de ráfagas orientadas a la conexión. En particular se proponen dos algoritmos del tipo aislado llamados Multi-path y Bypass y se comparan con el "clasico" algoritmo basado en el camino mas corto. Junto a la exposición del funcionamiento de los algoritmos, también se muestran diferentes resultados de la evaluación de las prestaciones de éstos usando dos diferentes algoritmos de reserva de recursos para los nodos. Finalmente se propone una comparativa entre la aplicación del Multi-path y del Bypass en una red de conmutación de ráfagas y en una red de conmutación de paquetes.

## 1. Instrucciones

La conmutación de ráfagas ópticas (Optical Burst Switching, OBS) es un nuevo paradigma surgido en el ámbito de la investigación de las redes completamente ópticas. OBS pretende ser una solución a medio plazo entre la conmutación de circuitos ópticos (OCS) - hoy en día en proceso de estandarización - y la conmutación de paquetes ópticos (OPS) - solución a largo plazo - [1].

Por un lado, OBS usa la multiplexación estadística directamente en el dominio óptico y es por lo tanto mas eficiente que soluciones basadas en OCS. Por el otro, las ráfagas son agregaciones de paquetes clientes construidas en los nodos fronterizos entre el dominio óptico y las redes de acceso (Figura 1). Por lo tanto las ráfagas son generalmente mucho mas largas que los paquetes de OPS. Además, las cabeceras viajan separadas de las ráfagas con un cierto margen de anticipo para dar tiempo a los

nodos intermedios de configurar la conmutación haciéndola, de hecho, mas viables que en el caso de OPS.

Estas ventajas originan los numerosos esfuerzos encaminados a desarrollar la tecnología que haga posible este paradigma. Para su diseño entonces, se tiene que partir de estudios relacionados con elementos discretos de red (los nodos), para posteriormente afrontar el estudio global de la red óptica. En este contexto se han realizado varios modelos de nodos de red, donde se han estudiado aspectos como la arquitectura [2], la señalización [3], la agregación de paquetes para formar las ráfagas [4], la asignación de recursos [5] o el aprovisionamiento de calidad de servicio (QoS) [6], entre otros. Algunos detalles se tratarán en el apartado 2.

La siguiente etapa en el desarrollo de una red OBS es el del estudio de la red propiamente dicha. En este estudio entran conceptos generales de networking,

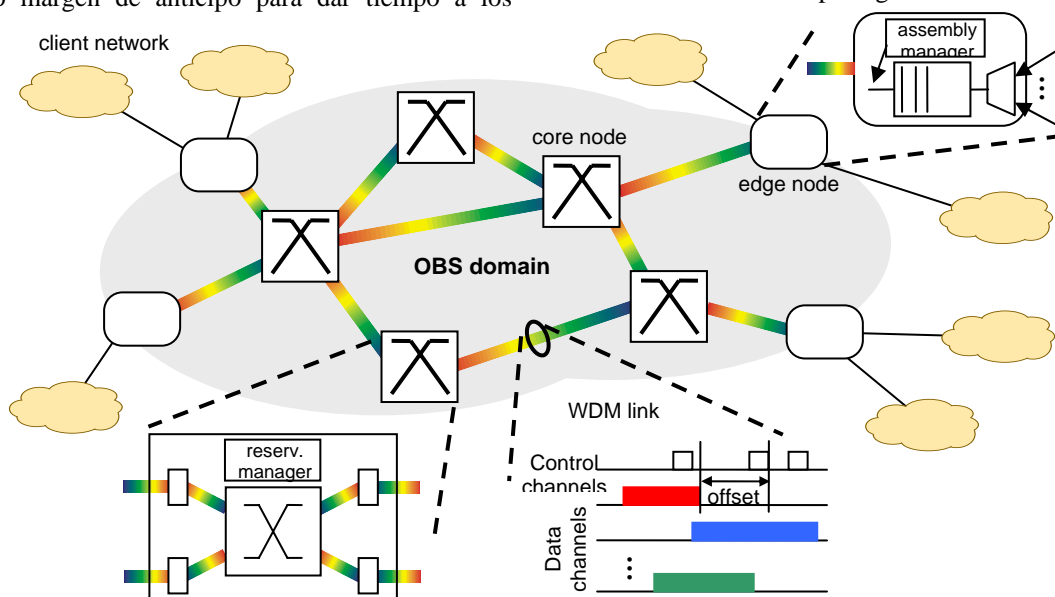


Figura 1. Esquema de una red de conmutación de ráfagas

como topología, congestión, o encaminamiento. Precisamente en este aspecto es donde se centrará este artículo (apartado 2.2). Analizaremos y evaluaremos la aplicabilidad de algoritmos de encaminamiento clásicos al entorno OBS y compararemos estos resultados con los obtenidos para redes OPS.

## 2. Escenario de estudio

Como en toda red, necesitamos un protocolo para gestionar el tráfico. El protocolo que usaremos en nuestra red óptica será el MPLS [7] (*Multiprotocol Label Switching*), que es un protocolo orientado a conexión. En efecto, un sistema orientado a conexión puede ser usado como infraestructura de una red de datagramas, como ocurre en el caso de paquetes IP sobre ATM. Una conexión MPLS se denomina LSP (*Label Switched Path*). Mediante el uso de la tecnología WDM, se pueden establecer varios LSP's en el mismo enlace físico, lo que representa un altísimo volumen de tráfico. Este tráfico viaja por las distintas longitudes de onda de las fibras ópticas encapsulado en ráfagas de datos. En el camino que recorren las ráfagas entre origen y destino, éstos son conmutados en los nodos implementando la técnica denominada OBS.

Como se ha comentado anteriormente, la parte de control (la cabecera) y la parte de datos (la ráfaga) viajan separadas usando diferentes longitudes de onda (Figura 1). Los nodos fronterizos se encargan de agregar paquetes de las redes clientes en ráfagas de longitud variable. Cuando una ráfaga está lista, el nodo genera un paquete de control y lo envía al destino. Pasado un cierto tiempo llamado *offset-time*, el nodo envía la ráfaga. El paquete de control es el que señala a los nodos intermedios la llegada de la ráfaga y es la única información que se convierte a eléctrico para ser procesada. Existen diferentes maneras para efectuar esta señalización. En este trabajo se ha estudiado el tipo *Just-Enough-Time* (JET) [1]. El paquete de control transporta el valor del *offset-time* y por lo tanto cada nodo intermedio reserva los recursos necesarios solo a partir del tiempo de llegada de la ráfaga y los libera una vez transmitida. La ráfaga por lo tanto permanece todo el tiempo en el dominio óptico.

### 2.1. Arquitectura del conmutador

Un nodo OBS tienen una matriz de conmutación  $N \times N$ . Si tenemos en cuenta la multiplexación WDM con  $W$  longitudes de onda, al final tenemos una matriz de entrada salida  $(N \times W)(N \times W)$ . El nodo consta de cuatro partes principales: interfaz de entrada, núcleo de conmutación, interfaz de salida y control lógico de conmutación [1] (Figura 2). En el núcleo de conmutación se encuentran los conversores de longitud de onda y el buffer óptico del conmutador.

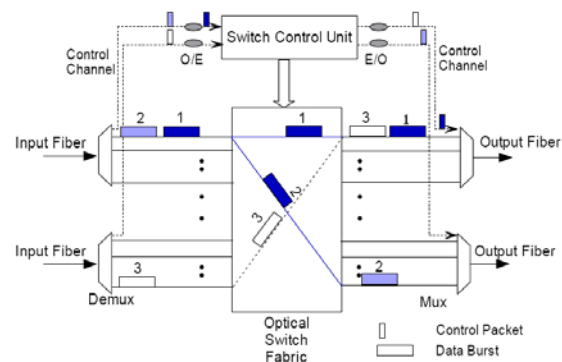


Figura 2. Esquema de un conmutador OBS

### 2.2. Asignación de recursos

La reserva de recurso es el problema principal de cualquier tecnología basada en la conmutación de ráfagas. Para asignar recursos un nodo puede administrar tres dominios distintos:

- La frecuencia usando los conversores de longitud de onda;
- El tiempo mediante el uso de colas ópticas. El problema de la conmutación óptica es que no es posible implementar RAM ópticas, así que se hace uso de múltiples líneas de retardo (*Fibre Delay Lines*, FDLs). Sin embargo, el uso de las FDLs es engorroso y debe limitarse;
- El espacio retransmitiendo la ráfaga por un puerto de salida diferente para seguir un camino alternativo que lleve al mismo destino. Este método necesita la implementación de algoritmos de encaminamiento adaptativos que veremos en el apartado 2.2.

La pérdida será sin embargo inevitable cuando ninguno de estos dominios proporcione una solución viable.

En este trabajo se han implementado 5 mecanismos distintos ya propuestos en literatura que explotan el dominio temporal y la frecuencia. Estos podemos dividirlos en dos grupos:

- el tipo *Horizon* [8] donde los nodos solo consideran libre los recursos a partir del último momento en el cual el recurso está planificado para ser usado. Los dos principales son los algoritmos MINGAP y MINLEN que respectivamente reservan el recurso que proporciona la mínima distancia (gap) entre dos reservas y el que proporciona el mínimo retraso en la cola.
- el tipo *Void Filling* (VF) [5] donde los nodos pueden reservar un recurso (hueco) entre dos reservas ya realizadas. A partir de este concepto hay varias posibles implementaciones que llamamos VF, VF-ST y VF-MM. El primero selecciona el primer hueco libre que encuentra. El segundo selecciona el hueco que proporciona la mínima distancia de la reserva que la precede. El tercero reserva el hueco que más se adapta a la longitud de la ráfaga.

### 2.3. Algoritmos de routing adaptativos

Se han implementado dos algoritmos de encaminamiento adaptativos. Como indica la clasificación en [9], estos algoritmos se llaman de tipo aislado en cuanto los nodos toman decisiones exclusivamente según su propio estado de congestión y por lo tanto no tienen conocimiento del estado de toda la red.

**MultiPath (MP).** En MP los nodos origen realizan una selección del camino a utilizar, es decir, cuentan con un abanico de posibilidades a su elección (p.e. los  $k$  caminos mas cortos). El nodo origen selecciona para cada ráfaga el siguiente nodo en función del estado de sus enlaces (se elige el que está menos congestionado). Por lo tanto, de los  $k$  caminos inicialmente disponibles, solo quedan los que pasan por el nodo elegido. Cada nodo intermedio comprueba cuantos caminos alternativos quedan y selecciona el enlace menos congestionado. Este proceso hace que las alternativas posibles vayan disminuyendo hasta quedarse con un único camino. Así que el primer camino en ser comprobado siempre es el SP. Si pero un enlace está congestionado, el nodo comprueba si tiene una ruta alternativa que tendrá una longitud igual o superior al SP. La figura 3a muestra un ejemplo de funcionamiento en el caso de  $k=3$ . Si el camino mas corto A-D-E fuera por ejemplo congestionado en el enlace D-E, entonces D elegiría el camino D-C-E. Si el enlace C-E resultase congestionado, entonces finalmente la ráfaga se perdería. Para este algoritmo, el nodo origen configura el offset-time entre paquete de control y ráfaga de acuerdo con lo que necesita el camino mas largo entre los  $k$  disponibles.

**Bypass (BP).** Con BP se pretende que una ráfaga pueda evitar un enlace congestionado haciendo un bypass. Es decir, si un nodo no encuentra un recurso disponible en un enlace de salida, se permite dar un salto más a la ráfaga con el objeto de evitar ese enlace. El número de bypass permitido depende de la topología de la red. La figura 3b ilustra un ejemplo. El camino es el A-D-E. Si pero el enlace D-E está congestionado, D puede elegir de bypassar este enlace, pasando por C o por F eligiendo el que tiene menor congestión (es decir menor numero de ráfagas en las colas). Como en el caso anterior, el nodo origen configura el offset-time de acuerdo con lo que necesita el camino mas largo.

Para tener una referencia, tambien se ha implementado el algoritmo **Shortest Path (SP)**. SP asigna a las ráfagas el camino más corto posible entre el nodo origen y el nodo destino. Si hubiera más de uno con el mismo número de saltos se escogería el primero computado. En este caso, solo hay un camino disponible y por lo tanto un nodo no puede resolver la asignación en el tiempo o en frecuencia, la ráfaga se pierde

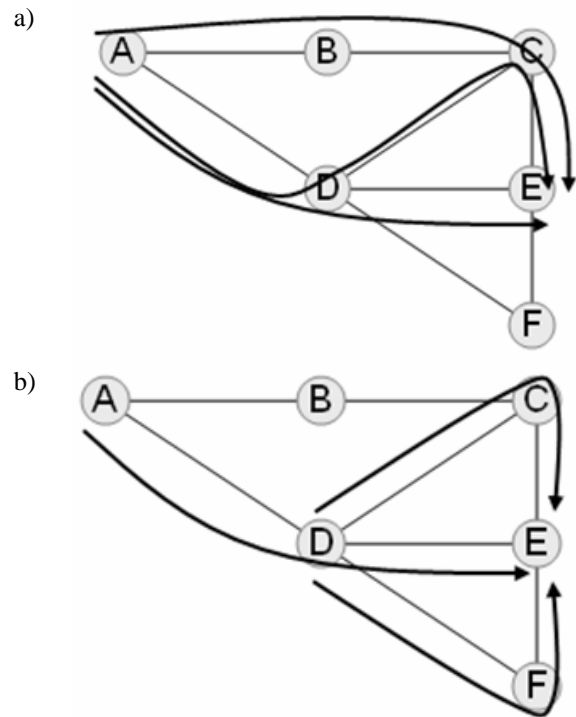


Figura 3. Ejemplo de encaminamiento de los paquetes usando los algoritmos propuestos con  $k=3$ , a) Multi-path, b) Bypass

## 3. Evaluación

### 3.1. Escenario de simulación

A continuación se describen los parámetros de configuración de la red a simular. Las ráfagas de datos son asíncronas, de longitud variable con distribución exponencial y media de 40000 bytes. Este valor es un buen compromiso entre largas ráfagas que necesitan largos tiempos de construcción en los nodos de acceso y cortas ráfagas que necesitan nodos de conmutación. Se considera una distribución de tráfico uniforme. En este estudio se consideran 2 algoritmos de asignación de recursos en los nodos; el MINLEN y VF-MM que respectivamente proporcionan los mejores resultados para el caso Horizon y Void-filling. Entre las varias topologías estudiadas, en este artículo presentamos resultados usando la topología NSFNet (Figura 4), red troncal de los EEUU. Para esta topología asumimos que cada enlace transporta 16 longitudes de onda a 10 Gbps.

### 3.2. Resultados de simulación

En las figuras 5 y 6 se evalúa la probabilidad de pérdidas de ráfagas en función, respectivamente, de la carga total de tráfico (normalizado con la capacidad de la red) y la longitud de las colas. Las diferentes curvas muestran los resultados usando los algoritmos de recursos de reserva MINLEN y VF-MM y los algoritmos de encaminamiento SP, MP y BP descritos en el apartado 2.2 y 2.3.

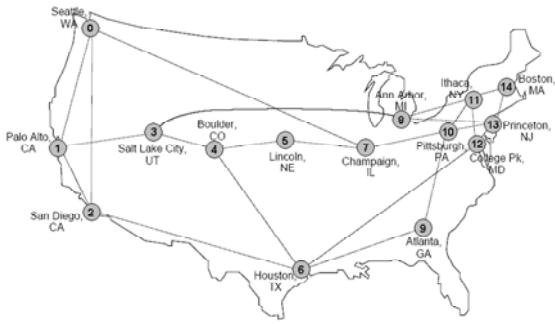


Figura 4. Topología de la red NSFNet

En la figura 5 cada nodo implementa una cola óptica de tamaño  $B = 4$  FDLs. La figura muestra que, en cada escenario, el algoritmo VF-MM presenta mejores resultados que el MINLEN, aunque con alta cargas las prestaciones son parecidas. Este comportamiento puede ser debido a que algunas partes de la red resultan altamente congestionadas y eso perjudica las prestaciones del sistema entero. En estos casos, el grado de optimización del algoritmo de reserva de recurso tiene poca importancia. De la misma manera podemos ver que los algoritmos propuesto Multi-path y Bypass tienen prestaciones parecidas al Shortest-Path en caso de alta cargas.

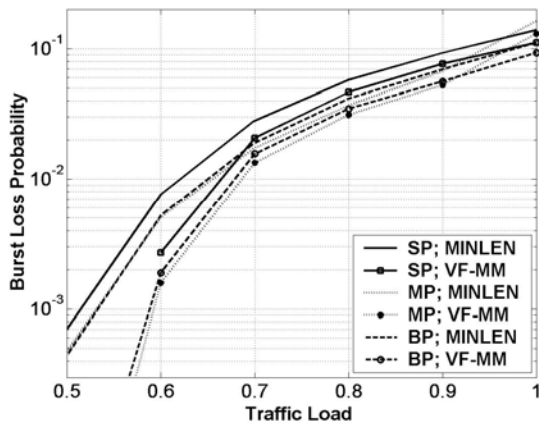


Figure 5. Probabilidad de pérdida de ráfagas en función de la carga de tráfico

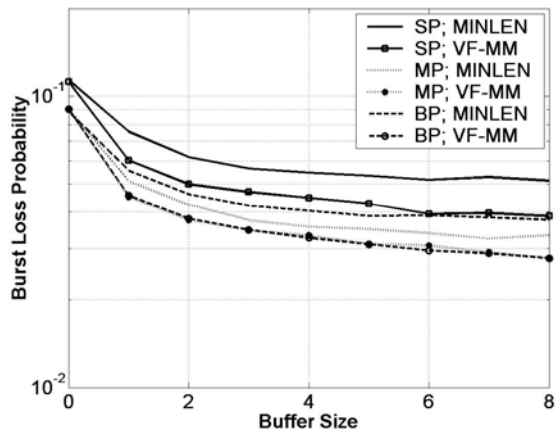


Figure 6. Probabilidad de pérdida de ráfagas en función del tamaño de las colas ópticas

En la figura 6 la carga total de la red normalizada de 0.8. Esta figura muestra que tener FDL aporta ventajas pero su capacidad afecta poco. En efecto los resultados obtenidos con una capacidad  $B > 2$  presentan pérdidas muy parecidas. Como en el caso anterior, los diferentes algoritmos de encaminamiento presentan resultados similares.

Una conclusión de este estudio es que estos algoritmos de tipo aislado no obtienen resultados mejores que el simple SP.

### 3.3. Comparativa con OPS

En esta sección comparamos los resultados obtenidos anteriormente con el caso de una red OPS y aplicando los mismos algoritmos de encaminamiento SP, MP y BP.

La comparativa se ha hecho con escenarios parecidos. Se ha usado la misma topología NSFNet y las mismas características de tráfico descrito anteriormente. Lo que cambia es el tamaño medio de los paquetes en el caso OPS que es de 500 bytes. Para el caso OPS se ha usado el algoritmo MINGAP mientras el VF-MM para OBS.

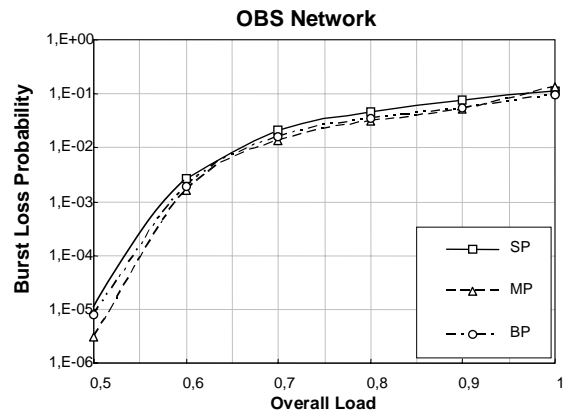


Figure 7. Probabilidad de pérdida de ráfagas en función de la carga de tráfico comparando SP, MP y BP usando VF-MM

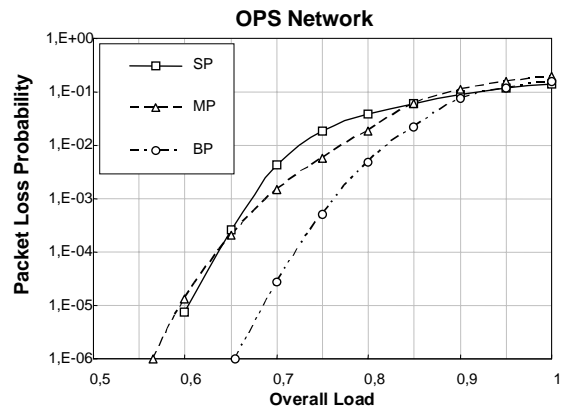


Figure 8. Probabilidad de pérdida de paquetes en función de la carga de tráfico comparando SP, MP y BP usando MINGAP

En las figuras 7 y 8 se comparan las probabilidades de pérdidas de ráfagas y paquetes respectivamente para los dos tipos de redes. Mientras en el caso de OBS los algoritmos de encaminamiento presentan resultados parecidos, en el caso OPS el algoritmo BP obtiene resultados mucho mejores que los otros dos. El porque de este resultado es actualmente bajo estudio pero se puede ya indicar que la razón está relacionada con la longitud de las ráfagas. Sabemos que si desviamos un paquete o una ráfaga hacia un camino más largo del SP, eso implica que el paquete o la ráfaga se quedará en la red más tiempo y eso incrementa la carga de la red. Desviar un paquete pero tiene menos impacto que desviar una ráfaga siendo la primera, en media, dos ordenes de magnitud mas corta.

#### 4. Conclusión

En este artículo hemos presentados dos algoritmos de encaminamiento adaptativos de tipo aislado para entorno OBS orientado a la conexión: el Multi-path y el Bypass. Se han evaluado las prestaciones de estos algoritmos usando dos estrategias de reserva de recurso: el MINLEN y el VF-MM. Los resultados obtenidos indican que algoritmos de tipo aislado no son muy efectivos para este entorno a diferencia de lo que pasaba para redes OPS.

El trabajo futuro será por lo tanto diseñar algoritmos distintos que sean de tipo distribuidos donde los nodos toman decisiones según el estado de toda la red a través de una función de coste.

#### Agradecimientos

Este trabajo ha sido en parte financiado por el Ministerio de Ciencia y Tecnología (MCIT, TIC2002-04334-C02-02) y por la Unión Europea con el proyecto integrado NOBEL (FP6-506760).

#### Referencias

- [1] C. Qiao, M. Yoo, "Optical Burst Switching (OBS) – a new paradigm for an optical Internet", *Journal of High Speed Networks*, vol. 8, no. 1, pp. 69-84, Marzo 1999
- [2] T. Battestilli, H. Perros, "An introduction to Optical Burst Switching", *IEEE Optical Communications*, vol. 41, no. 8, pp. S10-S15, Agosto 2003.
- [3] J.Y. Wei, R.I. McFarland, "Just-in-time signaling for WDM optical burst switching networks", *IEEE/OSA Journal of Lightwave Technology*, vol. 18, no. 12, pp. 2019-2037, Diciembre 2000.
- [4] V.M. Vokkarane, K. Haridoss, J.P. Jue, "Threshold-based assembly policies for QoS support in optical burst-switched networks", *SPIE OptiComm 2002*, Boston, MA, Julio 2002.
- [5] Y. Xiong, M. Vandenhouete, C. Cankaya, "Control architecture in optical burst-switched WDM networks", *IEEE Journal of Selected Areas in Communications*, vol. 18, no. 10, pp. 1838-1851, Octubre 2000.
- [6] M. Yoo, C. Qiao, S. Dixit, "Optical burst switching for service differentiation in the next-generation optical Internet", *IEEE Communications Magazine*, vol. 39, no. 2, pp. 98-104, Febrero 2001.
- [7] E. Rosen, "Multiprotocol label switching", *IETF RFC 3031*, Jan. 2001.
- [8] F. Callegati, W. Cerroni, G. Corazza, "Optimization of wavelength allocation in WDM optical buffers", *Optical Network Magazine*, vol. 2, no. 6, pp. 66-72, Noviembre 2001.
- [9] A. S. Tanenbaum, *Computer Networks*, second edition, Prentice Hall, 1988.
- [10] M. Klinkowski, F. Herrero, D. Careglio and J. Solé-Pareta, "Adaptive routing algorithms for optical packet switching networks", *ONDM 2005*, Milan, Italia, Febrero 2005